

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

**INFERÊNCIA E DIAGNÓSTICO NO MODELO ESPACIAL LINEAR T-STUDENT
REPARAMETRIZADO: APLICAÇÕES A DADOS AGRÍCOLAS**

ROSANGELA CARLINE SCHEMMER

CASCADEL - PR
FEVEREIRO - 2021

ROSANGELA CARLINE SCHEMMER

**INFERÊNCIA E DIAGNÓSTICO NO MODELO ESPACIAL LINEAR T-STUDENT
REPARAMETRIZADO: APLICAÇÕES 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: Profa. Dra. Fernanda De
Bastiani

Prof. Dr. Manuel Galea
Rojas

CASCADEL - PR
FEVEREIRO - 2021

Ficha de identificação da obra elaborada através do Formulário de Geração Automática do Sistema de Bibliotecas da Unioeste.

Schemmer, Rosangela Carline

INFERÊNCIA E DIAGNÓSTICO NO MODELO ESPACIAL LINEAR T-STUDENT REPARAMETRIZADO: APLICAÇÕES A DADOS AGRÍCOLAS / Rosangela Carline Schemmer; orientador(a), Miguel Angel Uribe-Opazo; coorientador(a), Fernanda De Bastiani, coorientador(a)II, Manuel Galea Rojas, 2021.

150 f.

Tese (doutorado), Universidade Estadual do Oeste do Paraná, Campus de Cascavel, Centro de Ciências Exatas e Tecnológicas, Programa de Pós-Graduação em Engenharia Agrícola, 2021.

1. Distribuição t-Student. 2. Geoestatística. 3. Análise de influência. 4. Modelo espacial linear. I. Uribe-Opazo, Miguel Angel . II. De Bastiani, Fernanda. III. Galea Rojas, Manuel . IV. Título.

BIOGRAFIA

Nome: Rosangela Carline Schemmer.

Ano de nascimento: 1987

Naturalidade: Medianeira-PR

Graduação: Licenciatura em Matemática pela Universidade Estadual do Oeste do Paraná (UNIOESTE) Campus Foz do Iguaçu no ano de 2009.

Pós-Graduação Lato Sensu: Gestão Bancária e Negócios pela Uniguaçu Faculdade de Ensino Superior de São Miguel do Iguaçu (FAESI) 2011, Metodologia do Ensino da Matemática pela Faculdade de Administração, Ciências, Educação e Letras (FACEL) Campus Curitiba 2012, Matemática pela Universidade Tecnológica Federal do Paraná (UTFPR) Campus Campo Mourão 2013.

Mestrado: Engenharia Agrícola pela Universidade Estadual do Oeste do Paraná (UNIOESTE) Campus Cascavel no ano de 2016.

Professora da Secretaria de Estado da Educação (SEED) de 2010 até 2013, Professora no Centro Universitário da Fundação Assis Gurgacz (FAG) ano de 2013, Professora categoria Substituto na Universidade Tecnológica Federal do Paraná (UTFPR) Campus Toledo de 2013 até 2015.

Atualmente: Bolsista Capes

*Aos meus pais, Romeu e Bernardete, com todo meu
amor.*

AGRADECIMENTOS

A Deus, por me fortalecer, permitir que meus desejos se tornem realidade e principalmente por me proporcionar tantos amigos, por tantas amizades as quais obtive durante este doutorado e que irei levar para a vida.

Ao Programa de Pós-Graduação em Engenharia Agrícola (PGEAGRI) da UNIOESTE, pela oportunidade da formação acadêmica e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio financeiro.

Aos docentes do programa de pós-graduação do PGEAGRI, pelos ensinamentos durante o doutorado. À secretaria acadêmica, pelos atendimentos prestados.

Ao professor Dr. Miguel Angel Uribe-Opazo, pela grande amizade, pela excelente orientação, paciência e disponibilidade ao me acompanhar na realização deste trabalho.

Ao professor Dr. Manuel Galea Rojas, minha gratidão pelos ensinamentos, pela paciência, carinho e pelos cafés durante as conversas de orientação, as quais eu não tinha nenhum papel para anotar, em que pude aprender a escutar e que tudo com calma se resolve.

À professora Dra. Fernanda De Bastiani, pelo carinho, atenção, por sua disponibilidade de tempo em me ensinar e passar seus conhecimentos.

À professora de inglês, Angélica Esmael, e a seu esposo, Robson, os quais me fortaleceram em meus estudos, imensamente grata.

À professora Dra. Rosangela Aparecida Botinha Assumpção, orientadora de mestrado, e que me acompanhou durante o doutorado, meu muito obrigada por disponibilizar seu tempo para me ensinar.

Aos meus amigos do Laboratório de Estatística Espacial (LEE) e Laboratório de Estatística Aplicada (LEA), pelas conversas, pelas mensagens, pelas opiniões e principalmente por me lembrarem das datas importantes dos relatórios.

Aos amigos da Facultad De Matemáticas Pontificia Universidad Católica de Chile, pelo apoio, carinho e atenção durante a estadia de estudos.

A todos meus amigos e colegas que, de alguma forma, contribuíram na realização deste trabalho.

RESUMO

INFERÊNCIA E DIAGNÓSTICO NO MODELO ESPACIAL LINEAR T-STUDENT REPARAMETRIZADO: APLICAÇÕES A DADOS AGRÍCOLAS

Este trabalho visou estudar e desenvolver técnicas de inferência e diagnósticos para o modelo espacial linear t-Student reparametrizado sem e com repetição, aplicado para dados agrícolas. Inicialmente, realizou-se uma reparametrização da distribuição t-Student supondo a existência do segundo momento finito, sendo apresentadas algumas propriedades recorrentes. Provaram-se expressões analíticas para a função score e matriz de Fisher da distribuição reparametrizada. Abordou-se um enfoque para a estimação dos parâmetros desenvolvendo um algoritmo iterativo. Apresentaram-se critérios de escolha para o melhor modelo através do parâmetro de forma η . Desenvolveu-se a análise de diagnóstico para detectar a presença de observações influentes e possíveis *outliers*. Estes procedimentos foram desenvolvidos sem e com repetição. Três artigos foram construídos a partir do desenvolvimento deste trabalho: diagnóstico de influência em um modelo espacial linear t-Student reparametrizado, inferência estatística no modelo espacial linear t-Student reparametrizado, modelo espacial linear t-Student reparametrizado com repetição. Conclui-se que a modelagem espacial linear t-Student reparametrizada permite uma modelagem mais robusta na presença de observações influentes.

Palavras—chave: Algoritmo iterativo; análise de influência; geoestatística; modelagem robusta.

ABSTRACT

INFERENCE AND DIAGNOSTICS IN THE REPARAMETERIZED T-STUDENT SPATIAL LINEAR MODEL: APPLICATIONS TO AGRICULTURAL DATA

This research aims at studying and developing the inference and diagnostic techniques for reparameterized t-Student linear spatial model, with and without replication, applied to data agricultural. Initially, a t-Student distribution reparameterization was carried out, assuming the existence of the second finite moment, with some recurring properties. Analytical expressions were tested for the score function and Fisher matrix of reparameterized distribution. The approach occurred to estimate some parameters, based on the development of an Iterative algorithm. Some criteria were shown to choose the best model by the shape parameter η . A diagnostic analysis was developed to detect the presence of influential observations and possible outliers. These procedures were developed without and with replication. Three papers were written due to the development of this research: diagnosis of influence in a reparameterized t-Student spatial linear model, statistical inference in reparameterized t-Student spatial linear model, reparameterized t-Student spatial linear model with replication. Finally, it can be concluded that reparameterized t-Student linear spatial modeling allows a more robust modeling in the presence of influential observations.

Keywords: Iterative algorithm; influence analysis; geostatistics; robust modeling.

SUMÁRIO

LISTA DE TABELAS	x
LISTA DE FIGURAS	xi
1 INTRODUÇÃO	1
1.1 Motivação	2
2 OBJETIVOS	5
2.1 Objetivo Geral	5
2.2 Objetivos específicos	5
3 REVISÃO BIBLIOGRÁFICA	6
3.1 Distribuição t-Student multivariada reparametrizada	6
3.2 A distribuição t-Student multivariada e a Geoestatística	9
3.2.1 Geoestatística	9
3.2.2 Modelagem Geoestatística	10
3.2.3 Modelo da família Matérn	15
3.2.4 Krigagem	17
3.2.5 Krigagem com <i>drift</i> externo	18
3.3 Caracterização das variáveis	20
3.3.1 Produtividade da soja	20
3.3.2 Dados químicos	21
3.3.3 Matéria orgânica	22
3.3.4 pH	23
3.4 Referências	23
4 ARTIGOS	30
4.1 ARTIGO 1: Diagnóstico de influência em um modelo espacial linear t-Student reparametrizado	30
4.2 Introdução	30
4.3 Modelo espacial linear t-Student	31
4.3.1 Distribuição t-Student multivariada reparametrizada	31
4.3.2 Modelo espacial linear t-Student	32
4.3.3 Identificabilidade do modelo	32
4.3.4 Estimativa de máxima verossimilhança	34
4.3.5 Função de influência	35

4.4	Diagnóstico e análise de influência	35
4.4.1	Estatística escore para <i>outlier</i>	36
4.4.2	Influência local	37
4.5	Aplicação	38
4.6	Estudo de simulação	43
4.6.1	Avaliação das estimativas dos parâmetros para o modelo t-Student reparametrizado	43
4.6.2	Medidas de influência	43
4.7	Conclusões	44
4.8	Apêndice: a matriz de informação observada para o modelo linear espacial t-Student	44
4.9	Apêndice: matriz para o esquema de perturbação Δ da matriz de covariância ..	45
4.10	Apêndice: a matriz de informação para o esquema de perturbação da matriz de covariância	47
4.11	Referências	48
4.12	ARTIGO 2: Inferência estatística no modelo espacial linear t-Student reparametrizado	52
4.13	Introdução	52
4.14	Modelo espacial linear t-Student	53
4.14.1	Estimador de máxima verossimilhança	55
4.14.2	Algoritmo iterativo	56
4.14.3	Estimativa do erro padrão assintótico	57
4.14.4	Critério de seleção do parâmetro de forma (η)	58
4.14.5	Gráfico QQ plot	58
4.15	Diagnóstico de influência	59
4.15.1	Influência global	59
4.15.1.1	Influência global baseada na verossimilhança	59
4.15.1.2	Influência global baseada na Q-function	59
4.15.2	Influência local	60
4.15.2.1	Influência local baseada no afastamento da verossimilhança	61
4.15.2.2	Influência local baseada na Q-function	62
4.15.3	Alavanca Generalizada	63
4.16	Aplicação ao conjunto de dados reais	64
4.16.1	Dados experimentais	64
4.16.2	Aplicação a dados de Aquífero	70
4.17	Conclusões	75
4.18	Referências	75
4.19	ARTIGO 3: Modelo espacial linear t-Student reparametrizado com repetição	79
4.20	Introdução	79
4.21	Modelo espacial linear com repetição	80

4.22 Função escore	82
4.23 Matriz de informação esperada de Fisher	82
4.24 Estimador de máxima verossimilhança	83
4.25 Algoritmo iterativo parcial linear	88
4.26 Critério de escolha para o parâmetro de forma η	90
4.27 Teste de hipóteses	91
4.27.1 Teste de hipótese para o modelo t-Student reparametrizado com r repetições independentes	93
4.28 Análise de diagnósticos	94
4.28.1 Influência global	95
4.28.2 Influência global baseada na verossimilhança	95
4.28.3 Influência global baseada na Q-function	97
4.28.4 Influência local	99
4.28.5 Influência local baseada na Q-function	100
4.28.6 Seleção de esquemas de perturbação	101
4.28.7 Esquemas de perturbação	102
4.28.8 Ponderação de casos	102
4.28.9 Perturbação na variável resposta	103
4.28.10 Perturbação na matriz escala	104
4.29 Material e Métodos	105
4.29.1 Área de estudo e dados	105
4.29.2 Coleta de dados	105
4.29.3 Resultados	106
4.30 Conclusões	112
4.31 Apêndice	113
4.32 Função escore	113
4.33 Derivada da matriz de covariância	115
4.34 Matriz de informação observada	116
4.35 Matriz de informação esperada de Fisher	123
4.36 Q-function	130
4.37 Derivada Q-function	131
4.37.1 Derivada de primeira ordem da Q-function	131
4.37.2 Derivada de segunda ordem da Q-function	133
4.37.3 Esperança da Q-function	136
4.38 Esquemas de perturbação para a influência local	137
4.38.1 Ponderação de casos	137
4.38.2 Função escore perturbada para encontrar a perturbação apropriada para a variável resposta	139
4.38.3 Perturbação na variável resposta	140
4.38.4 Derivada de $\Sigma^{1/2}$	142

4.38.5 Derivada da matriz de covariância	143
4.38.6 Perturbação na matriz escala	144
4.38.7 Representação $\ddot{Q}(\hat{\theta})$ para os esquemas de perturbação	144
4.39 Referências	144
5 CONSIDERAÇÕES FINAIS	150

LISTA DE TABELAS

Tabela 1	Função de covariância da família Matérn	16
Tabela 2	Análise descritiva para o conjunto de dados reais	39
Tabela 3	Estimativa dos parâmetros do modelo escolhido para a produtividade da soja considerando a função de covariância Matérn com $\kappa = 0,5$ e os respectivos erros padrão estão entre parênteses	40
Tabela 4	Estimativas dos parâmetros para o modelo escolhido para dados de produtividade da soja definidas sem as observações potencialmente influentes (OPI), considerando a função de covariância Matérn com $\kappa = 1,5$. Os respectivos erros padrão estão entre parênteses	41
Tabela 5	Porcentagens de vezes que o modelo t-Student reparametrizado com $\eta = 0,45$, $\eta = 0,05$ e $\eta = 0,01$ foram escolhidos usando os critérios de validação cruzada (VC) e traço da matriz de covariância assintótica (Tr) em 1000 dados simulados	43
Tabela 6	Porcentagem de vezes que as medidas de influência detectaram a observação perturbada em 1000 dados simulados. O modelo simulado é o t-Student reparametrizado com $\eta = 0,45$ e os modelos ajustados são o t-Student reparametrizado com $\eta = 0,45$, $\eta = 0,05$ e $\eta = 0,01$	44
Tabela 7	Estatística descritiva para as variáveis produtividade (Prod), Fósforo (P), Potássio (K), Potencial hidrogeniônico (pH) e Matéria orgânica (OM)	65
Tabela 8	Estimativas de parâmetros usando o algoritmo iterativo e erros padrão assintóticos entre parênteses	68
Tabela 9	Estimativas dos parâmetros e erros padrão entre parênteses	70
Tabela 10	Valores do traço da matriz de covariância assintótica ($T_r(\eta)$) e validação cruzada ($VC(\eta)$), para o aquífero Vale de Saratoga	71
Tabela 11	Potenciais observações influentes identificadas por diferentes autores usando diagnósticos globais e locais influentes sob diferentes abordagens.	73
Tabela 12	Análise descritiva da produtividade da soja referente a cada repetição	107
Tabela 13	Estimativa dos parâmetros do modelo escolhido para a produtividade da soja considerando a função de covariância da família Matérn com parâmetro de forma $\kappa = 2,9315$ com os respectivos erros padrões entre parênteses	109
Tabela 14	Teste da razão de verossimilhanças (RV) sob o vetor de parâmetros β do modelo espacial linear t-Student reparametrizado com repetição	110

LISTA DE FIGURAS

Figura 1	Área georreferenciada e Mapa da produtividade de soja em função das covariáveis químicas	3
Figura 2	Gráficos referentes à produtividade de soja	3
Figura 3	Função semivariância e seus respectivos parâmetros	13
Figura 4	Relação entre a função de covariância $C(h)$ e função semivariância $\gamma(h)$	15
Figura 5	Função de correlação da família Matérn para diferentes valores de κ em que (a) e (b) apresentam distância de 0 a 1,5 e 0 a 3,5, respectivamente	17
Figura 6	Funções escores para uma variável aleatória $T(0; 1; 0, 25)$ versus ϵ	36
Figura 7	(à esquerda) <i>Postplot</i> dos dados coletados na safra 2016/2017 e (à direita) boxplot da produtividade da soja	39
Figura 8	Modelo de covariância Matérn com ajuste de parâmetros $\kappa = 0,5$, $\hat{\phi}_1 = 0,0184$, $\hat{\phi}_2 = 0,5016$ e $\phi_3 = 0,4$	40
Figura 9	Gráfico do índice da estatística escore para a hipótese nula (à esquerda) e o gráfico de influência local para a matriz de covariância perturbada, com $\kappa = 0,5$ e $\eta = 0,05$ (à direita).	41
Figura 10	Mapas de produtividade da soja, considerando Matérn $\kappa = 0,5$ e $\eta = 0,05$ com todas as observações(a), Matérn $\kappa = 1,5$ e $\eta = 0,05$ sem a observação #34 (b), e Matérn $\kappa = 1,5$ e $\eta = 0,05$ sem as observações #34, #46 e #65 (c)	42
Figura 11	Desvio normal esperado versus distância transformada considerando (a) todas as observações, (b) sem observação #34, e (c) sem observações #34, #46 e #65.	42
Figura 12	(a) <i>Boxplot</i> , (b) <i>Post-plot</i> e (c) <i>semivariogramas direcionais</i> para o conjunto de dados de produtividade da soja ($t\ ha^{-1}$) no ano safra 2014/2015.	66
Figura 13	Gráficos de diagnóstico de influência global $D_{i\beta}$ e $D_{i\phi}$	66
Figura 14	Gráficos de diagnóstico de influência global $QD_{i\beta}$ e $QD_{i\phi}$	67
Figura 15	Gráficos de diagnóstico de influência local C_i e $ L_{max} $	67
Figura 16	Gráficos de diagnóstico de influência local CQ_i e $ Q_{max} $	67
Figura 17	Gráficos da alavanca generalizada considerando (a) a função logarítmica de verossimilhança e (b) a Q-function	69
Figura 18	QQ-plot dos resíduos (a) todos os pontos e (b) sem o ponto #33	69
Figura 19	Mapas da produtividade da soja ($t\ ha^{-1}$) no ano safra 2014/2015, para o conjunto de dados (a) com todas as observações e (b) exclusão da observação #33 .	69
Figura 20	Gráficos de diagnóstico de influência global $D_{i\beta}$ e $D_{i\phi}$ para os dados do aquífero Vale de Saratoga	71
Figura 21	Gráficos de diagnóstico de influência global $QD_{i\beta}$ e $QD_{i\phi}$ para os dados do aquífero Vale de Saratoga	72
Figura 22	Gráficos de diagnóstico de influência local C_i e $ L_{max} $ para os dados do aquífero Vale de Saratoga	72

Figura 23	Gráficos de diagnóstico de influência local CQ_i e $ Q_{max} $ para os dados do aquífero Vale de Saratoga	72
Figura 24	Gráfico de alavanca generalizada para os dados do aquífero Vale de Saratoga	73
Figura 25	QQ-plots(a) com todas as observações (b) sem as observações #49 e #50 e (c) sem as observações #01 e #61	74
Figura 26	Localização dos pontos considerados influentes para os dados do aquífero Vale de Saratoga	74
Figura 27	Mapas para os dados do aquífero Vale de Saratoga (a) com todas as observações (b) sem observações #49 e #50 e (c) sem observações #01 e #61	74
Figura 28	Localização da área em estudo.	106
Figura 29	Scatter plot da produtividade da soja referente a cada repetição	107
Figura 30	<i>Boxplot</i> da variável explicativa produtividade da soja e das covariáveis Fósforo, Potássio, Cálcio e Magnésio nos anos agrícolas 2012/2013, 2013/2014, 2014/2015, 2015/2016 e 2016/2017	108
Figura 31	Gráficos de diagnósticos de influência local perturbação na variável resposta	110
Figura 32	Gráficos de diagnósticos de influência local e ponderação de casos	110
Figura 33	Gráficos de diagnósticos de influência global baseada na verossimilhança ..	111
Figura 34	Gráficos de diagnósticos de influência global baseada na Q-function	112

1 INTRODUÇÃO

A Geoestatística tem um papel fundamental na Agricultura de Precisão, pois suas técnicas possibilitam a construção de mapas temáticos que determinam a estrutura da dependência espacial dos atributos físicos e químicos do solo associados à produtividade da cultura. Assim, é possível detalhar os fatores de produção e definir tanto a aplicação localizada de insumos como o manejo diferenciado das culturas. O conhecimento da estrutura de dependência espacial da produtividade da soja e sua relação com os atributos químicos e físicos do solo são essenciais para o manejo adequado da cultura. Nora et al. (2017) constatam que um sistema de manejo dos cultivares de milho, trigo e soja requer informações precisas sobre a dependência espacial, e a interação entre alta qualidade química é uma estratégia eficiente para o potencial de rendimento dos cultivares. Porém, observações atípicas podem afetar a variabilidade espacial, pois levam a resultados menos próximos da realidade, especialmente quando a distribuição normal de probabilidade está sob suposição. Assumpção et al. (2014) e Schemmer et al. (2017) relatam sobre a sensibilidade da distribuição normal para valores atípicos. Os autores ainda destacam que tal modelagem com distribuição normal na presença de valores atípicos pode gerar mapas imprecisos.

Uma solução para este problema é a utilização de modelos robustos. Segundo Cysneiros et al. (2005), os modelos da classe das distribuições simétricas permitem reduzir a influência dos valores atípicos, incorporando parâmetros adicionais que ajustam a curtose da distribuição dos dados. A distribuição t-Student pertence à classe das distribuições simétricas, apresentam propriedades de maior flexibilidade quanto ao grau de curtose, tem como parâmetro adicional o grau de liberdade e apresentam caudas mais pesadas que a distribuição normal. Sob estas considerações, vários trabalhos foram realizados. Kotz e Nadarajah (2004) apresentam quase todos os resultados disponíveis na literatura sobre distribuições multivariadas t-Student publicados nos últimos 50 anos. Kibria e Joarder (2006) descrevem uma breve revisão sobre a distribuição t-Student multivariada. Nadarajah e Kotz (2008) mostram métodos de estimação para a distribuição t-Student multivariada e Grigelionis (2013) apresenta como construir a distribuição t-Student multivariada a partir de uma mistura da distribuição normal e a distribuição gaussiana inversa.

Lange et al. (1989) propuseram duas reparametrizações no parâmetro de forma para a distribuição t-Student, em que permite a existência do segundo momento finito e uma comparação direta entre a matriz de covariância das distribuições t-Student e normal. Sutradhar (1993) desenvolveu uma versão reparametrizada da distribuição t-Student e teste de escore para a matriz de covariância. Fiorentini et al. (2003) apresentam estimativas e inferências de máxima verossimilhança com base em uma distribuição t-Student reparametrizada. A reparametrização auxilia o estimador de máxima verossimilhança e na construção do algoritmo

EM (DEMPSTER et al., 1977). Galea e Castro (2017) apresentam a inferência robusta em um modelo funcional linear com replicações usando a distribuição t-Student.

Na Geoestatística, muitos trabalhos foram desenvolvidos considerando a matriz de covariância a partir da distribuição normal. Logo, é importante que, quando outros modelos sejam considerados, uma reparametrização na matriz de covariância seja feita para torná-la comparável. E, além de poucos trabalhos já terem sido apresentados considerando uma distribuição t-Student reparametrizada, nenhum deles considera a forma paramétrica da matriz de covariância para estudar a dependência espacial entre observações. Portanto, a proposta desta tese é a extensão do trabalho apresentado por Assumpção et al. (2014), a qual trabalharam com modelos espaciais t-Student e De Bastiani et al. (2017), a qual trabalharam com modelos lineares gaussianos com repetições e, para este estudo, apresentam-se os modelos lineares t-Student reparametrizada sem e com repetições.

1.1 Motivação

O Brasil é o quarto maior consumidor de fertilizante do mundo. Nas últimas décadas, o consumo tem crescido em média 6% ao ano (ANDA, 2014). As culturas da soja e do milho são responsáveis pelo consumo de 50% dos fertilizantes no Brasil (CUNHA et al., 2010). A Agricultura de precisão (AP) surge como uma importante ferramenta para evitar a utilização desnecessária de fertilizantes. Segundo Richart et al. (2016), a AP é capaz de maximizar a produção de forma racional e evitar a superdosagem ou subdosagem de insumos. Magalhães et al. (2012) ressaltam que a AP busca aplicar os insumos de forma localizada, geralmente em taxas variáveis, levando em consideração a variabilidade das áreas produtivas. Molin (2001) relata que a partir da variabilidade espacial, por meio da amostragem de solo georreferenciadas, tomam-se as decisões necessárias para que se possam prescrever operações e aplicações dos insumos em dosagem prescritas e de forma localizadas. Isto se torna possível, pois na AP, são aplicados os princípios da geoestatística para caracterizar a variabilidade espacial dos atributos que fazem parte dos fatores de produção agrícola (AMADO et al., 2007). Com isso, essa ferramenta permite identificar zonas com restrições químicas que possam limitar o rendimento e, posteriormente, realizar a correção de maneira adequada (SANCHEZ et al., 2009).

A utilização da geoestatística na Agricultura de Precisão possibilita a construção de mapas que determinam a variabilidade espacial dos atributos físicos-químicos do solo, bem como a aplicação de insumos agrícolas em locais corretos e quantidades adequadas para a obtenção de áreas homogêneas. Como exemplo, verifica-se na Figura 1 que a área em estudo apresenta regiões de menor produtividade da soja ($t\ ha^{-1}$). As duas menores classes são 2, 29 – 2, 33 e 2, 33 – 2, 36 ($t\ ha^{-1}$) e em sua maior totalidade estão localizadas na região central do mapa. A partir desta análise, define-se um plano estratégico para o melhoramento na fertilidade do solo desta região.

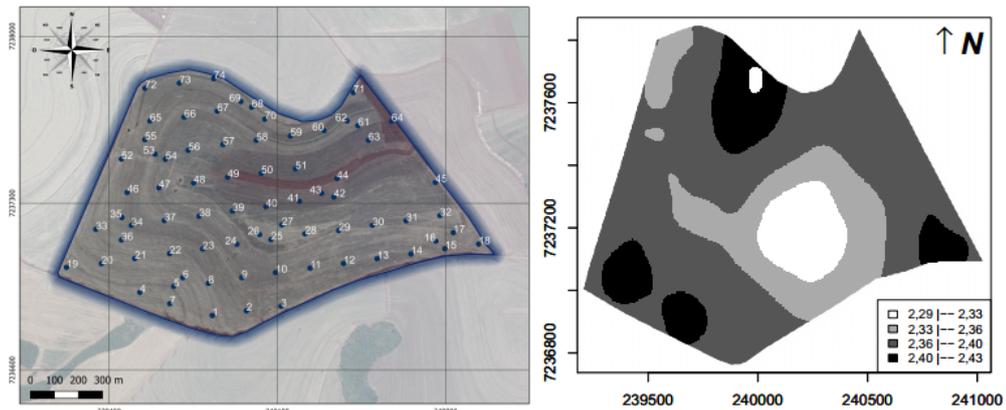
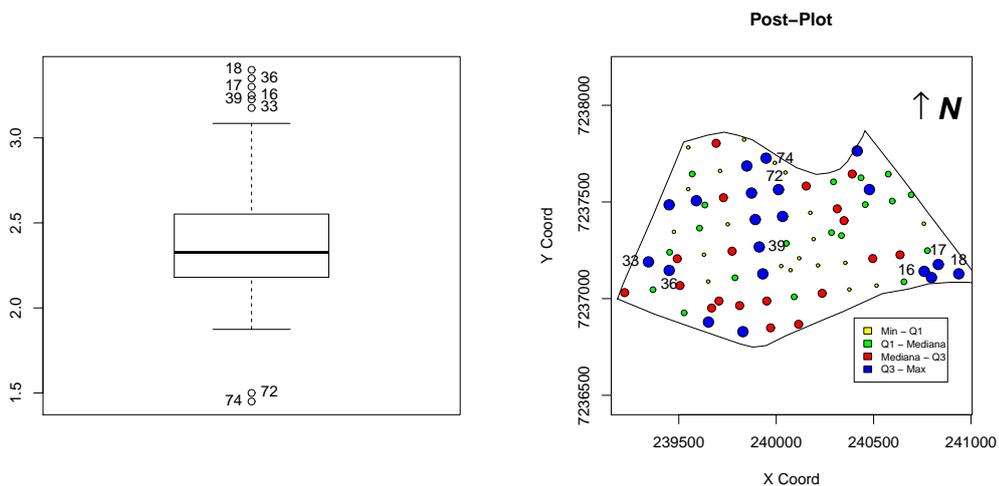


Figura 1 Área georreferenciada e Mapa da produtividade de soja em função das covariáveis químicas

Além disso, a partir do modelo espacial linear, podem-se obter informações valiosas de cada localização, assim como o valor específico de cada atributo para determinar a dose adequada para a correção do solo.



(a) Box plot

(b) Post plot

Figura 2 Gráficos referentes à produtividade de soja

A Figura 2 (a) apresenta um exemplo da presença de pontos discrepantes, os quais estão identificados espacialmente na Figura 2 (b). Estes, por sua vez podem alterar consideravelmente os resultados das análises de dependência espacial e conseqüentemente influenciar na construção dos mapas temáticos (RICHETTI et al., 2016). Em grande parte dos estudos, a modelagem estatística é baseada na distribuição normal, sendo essa sensível a valores atípicos. Assim, a distribuição t-Student multivariada torna-se particularmente atrativa, pois possui caudas mais pesadas que a normal, por isso, torna-se essencial quando se analisam e inferem resultados de dados provenientes de estudos agrícolas. Os mapas devem representar uma estimativa eficiente das condições reais do campo, já que são importantes informações utilizadas para a manutenção de um sistema agrícola de manejo localizado, visando à maior

produtividade e otimização da aplicação de insumos agrícolas.

2 OBJETIVOS

2.1 Objetivo Geral

Desenvolver inferência estatística e estudar a análise de diagnóstico de pontos influentes no modelo espacial linear t-Student reparametrizado com e sem repetições.

2.2 Objetivos específicos

- (i) Desenvolver diagnóstico de influência em um modelo espacial linear t-Student reparametrizado;
- (ii) Aplicar inferência estatística no modelos espaciais lineares t-Student reparametrizado;
- (iii) Propor um modelo espacial linear t-Student reparametrizado com repetições.

3 REVISÃO BIBLIOGRÁFICA

3.1 Distribuição t-Student multivariada reparametrizada

A necessidade de analisar observações contínuas multivariadas surge com frequência em diversas áreas do conhecimento. Em particular, a maioria das inferências estatísticas associadas às estimativas e dos testes de hipóteses, neste contexto, baseia-se no pressuposto da normalidade (MARDIA et al., 1992; ANDERSON, 2003; HÄRDLE; SIMAR, 2012). Embora a maioria destes estudos esteja relacionada à simplicidade e à utilidade da distribuição normal, sabe-se que a inferência estatística baseada nessa distribuição é suscetível à presença de dados atípicos, pois os momentos de quarta ordem da distribuição normal são determinados pelo primeiro e segundo momento da distribuição, impossibilitando o ajuste da curtose das observações.

Especificamente, é necessário considerar distribuições alternativas que nos permitam contornar essa limitação da distribuição normal multivariada. Nessa direção, vários autores apresentam o uso da distribuição t-Student multivariada como alternativa à distribuição normal. Esta distribuição tem sido amplamente utilizada no estudo de dados reais por possuir caudas mais pesadas e permitir que os pontos discrepantes, presentes no conjunto de dados, sejam englobados (LANGE et al., 1989; OSORIO et al., 2007). Além disso, é uma distribuição simétrica com parâmetro de forma extra, denominado grau de liberdade, representado por v ($v > 0$), que permite modelar a curtose dos dados.

Existem casos especiais da distribuição t-Student, tais como a distribuição denominada Central se $\mu = 0$. Note que se $n = 1$, $\mu = 0$ e $\Sigma = 1$, então a função densidade de probabilidade (fdp) reduz para a distribuição t-Student univariada. Se $n = 2$, então, a equação fdp é uma ligeira modificação de superfície bivariada de Pearson. Se o grau de liberdade $v = 1$, a fdp é a distribuição n-variada de Cauchy. Se $(v + n)/2 = m$, sendo m um inteiro, a fdp é a distribuição n-variada de Pearson tipo VII. A forma limitante quando $v \rightarrow \infty$ é a função de densidade de probabilidade conjunta da distribuição normal n-variada com vetor de média μ e matriz de covariância Σ (NADARAJAH; KOTZ, 2008).

Dentre os diversos trabalhos com a distribuição t-Student multivariada, tem-se: a estimação de máxima verossimilhança para a distribuição t-Student multivariada usando o algoritmo EM e certas variações, discutidos por (LITTLE, 1988; LIU; RUBIN, 1995; LIU, 1997). Lange et al. (1989) e Lange e Sinsheimer (1993) demonstraram o uso da distribuição t-Student em regressão e análise multivariada. Sutradhar (1993) propôs um teste $C(\alpha)$ para provar a hipótese que afirma que a matriz de covariância é igual a uma determinada matriz considerando a distribuição t-Student. Por outro lado, Bolferine e Arellano-Valle (1994) e Bolferine e Rojas (1996) aplicaram a distribuição t-Student em modelos de erro de medição e calibração comparativa estrutural, respectivamente. Xie et al. (2007) apresentaram medidas de

diagnóstico para avaliar as observações influentes. Enquanto Finegold e Drton (2011) usaram duas abordagens para a modelagem robusta de dados de expressão gênica considerando a distribuição t-Student multivariada.

Porém, para obter uma comparação direta entre a matriz de covariância da distribuição t-Student n-variada e a normal n-variada, Lange et al. (1989) sugerem a reparametrização do grau de liberdade da distribuição, que permite a existência do segundo momento finito. Além disso, citam que há melhora da inferência quando o grau de liberdade apresenta uma das seguintes transformações: $v = 1/\eta$ ou $v = \log(\eta)$. Sutradhar (1993) desenvolveu uma versão reparametrizada da distribuição t-Student e do teste de escore para a matriz de covariância. Bolfarine e Rojas (1996) relatam a substituição da distribuição normal pela distribuição t-Student. Fiorentini et al. (2003) apresentam uma das primeiras aplicações da reparameterização do grau de liberdade sugerido por Lange et al. (1989), assim como a estimativa e a inferência da máxima verossimilhança com base na distribuição t-Student reparametrizada. Esta reparametrização auxiliou no procedimento de estimativa da máxima verossimilhança e na construção do algoritmo EM (DEMPSTER et al., 1977). Schemmer et al. (2017) apresentam a distribuição t-Student reparametrizada referente à transformação $v = 1/\eta$, ao aplicarem este estudo a dados de produtividade da soja. Galea et al. (2020) estenderam os trabalhos de Lange et al. (1989) e Sutradhar (1993), apresentando a estimação de parâmetros, utilizando o algoritmo EM e discutiram os testes assintóticos sendo: a razão de verossimilhança, Wald, escore e estatística de gradiente.

Para maior clareza, apresentam-se algumas propriedades básicas da distribuição t-Student multivariada baseada na reparametrização introduzida por Sutradhar (1993) e Osorio e Galea (2015). O resultado a seguir apresenta um resumo das propriedades da distribuição $\tau_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$.

Definição 3.1. *Seja Y um vetor aleatório n-dimensional com função de densidade definida por:*

$$f_Y(y) = K_n(\eta) |\boldsymbol{\Sigma}|^{-\frac{1}{2}} [1 + c(\eta)\delta]^{-\frac{1}{2\eta}(1+n\eta)}, \quad (3.1)$$

em que:

$$K_n(\eta) = \left(\frac{c(\eta)}{\pi}\right)^{\frac{n}{2}} \frac{\Gamma\left(\frac{1+n\eta}{2\eta}\right)}{\Gamma\left(\frac{1}{2\eta}\right)}, 0 < \eta < \frac{1}{2},$$

com $c(\eta) = \eta/(1 - 2\eta)$, $\delta = (\mathbf{Y} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \boldsymbol{\mu})$ e $\Gamma(\cdot)$ é a função gama. Como Y tem função densidade de probabilidade dada por (3.1), denota-se por $Y \sim \tau_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$, portanto, o vetor Y tem distribuição t-Student n-variada reparametrizada com vetor de média $\boldsymbol{\mu}$, matriz de covariância $\boldsymbol{\Sigma}$ e parâmetro de forma η .

Logo, o vetor aleatório Y apresenta as seguintes propriedades:

Proposição 3.1. *O vetor aleatório Y pode ser escrito usando a seguinte representação:*

$$\mathbf{Y}|u \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}u^{-1}) \text{ e } u \sim \text{Gama}\left(\frac{1}{2\eta}, \frac{1}{2c(\eta)}\right),$$

com $c(\eta) = \eta/(1 - 2\eta)$.

Proposição 3.2. Segundo Pinheiro et al. (2001), tem-se que $Gama(\alpha, \beta)$ é a distribuição Gama com parâmetros α e β , definidos pela função de densidade de probabilidade dada por:

$$h(u) = \frac{\beta^\alpha u^{\alpha-1} \exp(-u\beta)}{\Gamma(\alpha)},$$

em que $u > 0$, $\alpha > 0$ e $\beta > 0$. Há também as seguintes definições:

$$E(u) = \frac{\alpha}{\beta},$$

$$E(u^{-a}) = \frac{\beta^a \Gamma(\alpha - a)}{\Gamma(\alpha)}, \quad a < \frac{1}{\eta},$$

$$\text{Var}(u) = \frac{\alpha}{\beta^2},$$

$$E(\log(u)) = \psi(\alpha) - \log(\beta),$$

com $\psi(\alpha) = \partial \log(\Gamma(u))/\partial u$ é denominada a função Digama. Em nosso estudo, os parâmetros α e β são definidos, sendo $u \sim Gama\left(\frac{1}{2\eta}, \frac{1}{2c(\eta)}\right)$, logo:

$$E(u) = \frac{c(\eta)}{\eta},$$

$$E(u^{-a}) = \left(\frac{1}{2c(\eta)}\right)^a \frac{\Gamma\left(\frac{1}{2\eta} - a\right)}{\Gamma\left(\frac{1}{2\eta}\right)},$$

$$\text{Var}(u) = \frac{2c^2(\eta)}{\eta},$$

$$E(\log(u)) = \psi\left(\frac{1}{2\eta}\right) - \log\left(\frac{1}{2c(\eta)}\right),$$

Proposição 3.3. O vetor de médias e a matriz de covariância de Y são dados por $E(Y) = \mu$ e $\text{Cov}(Y) = \Sigma$, respectivamente. Além disso, considerando-se a esperança condicional, obtém-se:

$$E(Y) = E(E(Y|u)) = \mu,$$

$$\text{Cov}(Y) = E(\text{Cov}(Y|u)) + \text{Cov}(E(Y|u)) = \Sigma.$$

Proposição 3.4. Seja $\delta = (Y - \mu)^\top \Sigma^{-1} (Y - \mu)$ a distância de Mahalanobis. Então, a variável aleatória :

$$F = \left(\frac{1}{1 - 2\eta}\right) \frac{\delta}{n} \sim F\left(n, \frac{1}{\eta}\right).$$

Note que $E(\delta) = n$.

Proposição 3.5. O coeficiente de curtose multivariado (MARDIA, 1970). Definido por:

$$\beta_{2,n} = E\{[(Y - \mu)^\top \Sigma^{-1} (Y - \mu)]^2\} = n(n + 2)(k_* + 1),$$

em que $k_* = 2\eta/(1 - 4\eta)$, representando o excesso de curtose.

Especificamente para nosso caso, tem-se $\varepsilon = \mathbf{Y} - \boldsymbol{\mu} \stackrel{d}{=} u^{-1/2}\mathbf{Y}$, com $\mathbf{Y} \sim N_n(\mathbf{0}, \boldsymbol{\Sigma})$. Então, o coeficiente de curtose é definido por:

$$\begin{aligned}\beta_{2,n} &= \mathbf{E} \left(\boldsymbol{\varepsilon}^\top \mathbf{E}^{-1} (\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^\top) \boldsymbol{\varepsilon} \right)^2 \\ &= \mathbf{E} \left(u^{-1} \mathbf{Y}^\top \mathbf{E}^{-1} (u^{-1} \mathbf{Y} \mathbf{Y}^\top) \mathbf{Y} \right)^2 \\ &= \frac{\mathbf{E}(u^{-2})}{\mathbf{E}^2(u^{-1})} \mathbf{E} \left(\mathbf{Y}^\top \mathbf{E}^{-1} (\mathbf{Y} \mathbf{Y}^\top) \mathbf{Y} \right)^2 \\ &= \left(\frac{1 - 2\eta}{1 - 4\eta} \right) n(n + 2),\end{aligned}$$

sendo que $\eta < 1/4$. Além disso, por conta de $(1 - 2\eta)/(1 - 4\eta) > 1$ para $0 \leq \eta < 1/4$, tem-se que a curtose da distribuição t-Student multivariada é sempre maior do que a distribuição normal (YAMAGUCHI, 2003).

Proposição 3.6. Utilizando os resultados de Abramowitz e Stegun (1970), segue-se que

$$\lim_{\eta \rightarrow 0^+} K_n(\eta) = 1/(2\pi)^{n/2} \quad \text{e} \quad \lim_{\eta \rightarrow 0^+} (1 + c(\eta)\delta)^{-1/2\eta} = \exp(-\delta/2),$$

portanto, quando $\eta \rightarrow 0^+$, obtém-se a distribuição normal multivariada, cuja a função densidade é dada por:

$$f_N(\mathbf{y}) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1/2} \exp(-\delta/2), \quad \text{em que } \mathbf{y} \in \mathbb{R}^n.$$

3.2 A distribuição t-Student multivariada e a Geoestatística

3.2.1 Geoestatística

A geoestatística difere da estatística clássica, pois os modelos da estatística clássica estão geralmente voltados para a verificação da distribuição de frequência dos dados, enquanto a geoestatística incorpora a interpretação da correlação espacial das amostras. Este aspecto da geoestatística está intimamente associado com a distribuição estatística dos dados no espaço. Os primeiros trabalhos foram publicados por Matheron (1962, 1963). O autor formalizou a metodologia aplicada por Krige, o qual trabalhou com dados de concentração de ouro na África do Sul (KRIGE, 1951). E denominou seu estudo de “Teoria das variáveis regionalizadas”, a qual define uma variável regionalizada com uma função espacial numérica, que varia de um local para outro, com uma continuidade. São variáveis cujos valores estão relacionados de algum modo com a posição espacial que ocupam. Segundo Landim (2006), a geoestatística assume que a diferença entre dois pontos amostrados depende da distância entre eles e da orientação dos pontos, ou seja, pares de observações mais próximas são mais semelhantes do que pares de observações mais distantes.

Desde as publicações de Matheron, a geoestatística passou a ter ampla aplicação na agricultura e, entre os diversos trabalhos desenvolvidos, podem-se citar Mercante et al. (2003), que estudaram a variabilidade espacial de atributos do solo utilizando métodos estatísticos

com semivariogramas escalonados. Assumpção et al. (2007) fizeram uso da krigagem indicatriz para avaliação da probabilidade da produtividade de soja usando como fatores de comparação os padrões regional, estadual e nacional. Guedes et al. (2008) realizaram a análise da anisotropia no estudo da variabilidade espacial de alguns atributos químicos do solo espacialmente referenciados. Dalposso et al. (2012) utilizam diversas medidas derivadas da matriz dos erros para comparar mapas temáticos gerados a partir de métodos geoestatísticos. Uribe-Opazo et al. (2012) desenvolveram métodos de diagnósticos de influência local no caso do estudo de modelos espaciais lineares com distribuição gaussiana. De Bastiani et al. (2012) compararam mapas de variabilidade espacial da resistência do solo à penetração, construídos com e sem covariáveis a partir do modelo espacial linear. Assumpção et al. (2014) realizaram a análise de influência local em geoestatística utilizando a distribuição t-Student para dados da produtividade da soja. De Bastiani et al. (2015) realizaram a análise da influência para a distribuição elíptica. Garcia-Papani et al. (2017) realizaram estudos com modelo espacial Birnbaum-Saunders e diagnósticos aplicados a dados agrícolas, especificamente para a variável fósforo, a qual apresenta efeito positivo ao crescimento e nutrição da soja, além do fato que a variabilidade espacial deste macronutriente ser de extrema importância para a gestão agrícola. Schemmer et al. (2017) analisaram a variabilidade espacial da produtividade de soja aplicado ao modelo espacial linear com distribuição t-Student reparametrizada. Fagundes et al. (2018) investigaram a variabilidade da produtividade da soja em função das propriedades químicas do solo aplicado ao modelo espacial linear slash.

Assim, a geoestatística é utilizada quando se verifica a existência de dependência espacial entre os elementos amostrais, de forma a descrever a existência a partir da função de correlação espacial, a qual representa o comportamento da dependência espacial dos dados a partir de um modelo.

3.2.2 Modelagem Geoestatística

Para o estudo da dependência espacial, considera-se um processo estocástico¹ $\{Y(s_i), s_i \in S\}$, estacionário de segunda ordem em que $S \subset \mathbb{R}^2$ e \mathbb{R}^2 é um espaço euclidiano bi-dimensional, as localizações espaciais podem ser alinhadas em uma grade regular ou distribuída irregularmente no domínio (YAMAMOTO; LANDIM, 2015). Aditem-se as seguintes características estocásticas do processo aleatório:

Definição 3.2. *Um processo estocástico $Y(s)$ em um espaço bi-dimensional S é estacionário se $E[Y(s)] = \mu(s)$ e não depende de s .*

Definição 3.3. *Um processo estocástico $Y(s)$ em um espaço bi-dimensional S é estacionário de segunda ordem se $\text{Var}[Y(s)]$ existe para todo $s \in S$, tais que:*

- (i) $Y(s)$ é estacionário;
- (ii) $C(s_i, s_j) = \text{Cov}[Y(s_i), Y(s_j)]$ depende de s_i e s_j somente pela diferença dada por $h =$

¹Em aplicações geoestatísticas, os dados são considerados uma realização parcial de um processo aleatório, assim denominamos de processo estocástico ou campo aleatório.

$s_j - s_j$.

Aplicações para espaço bi-dimensional são apresentadas em Piegorsch e Bailer (2005), Waller e Gotway (2004), que investigam solos contaminados com Dioxina e estudam a saúde do ecossistema através da biomassa de camarão capim, respectivamente.

A partir das definições 3.2 e 3.3, tem-se $\mathbf{Y} = (Y(s_1), \dots, Y(s_n))^T$ um vetor $n \times 1$ da variável resposta que corresponde às localizações espaciais conhecidas em s_i com $i = 1, \dots, n$, sendo $\mathbf{Y} \sim \tau_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$. Pode-se escrever $Y(s_i)$ como:

$$Y(s_i) = \mu(s_i) + e(s_i), \quad (3.2)$$

sendo $\mu(s_i)$ o termo determinístico de locação e $e(s_i)$ o termo estocástico, ambos dependentes da localização em que $Y(s_i)$ atua. Admite-se que o erro estocástico e tem média zero, ou seja, $E[e(s_i)] = 0$. E a variação entre os pontos no espaço é determinada por alguma função de covariância $C(s_i, s_j) = Cov[e(s_i), e(s_j)]$. A média do processo estocástico é definida pelo termo determinístico $\mu(s_i)$ dado por:

$$\mu(s_i) = \sum_{j=1}^n \beta_j \mathbf{X}_j(s_i),$$

em que $\mathbf{X}_1(s_i)=1$, para todo $s_i = 1, \dots, n$, (X_2, \dots, X_n) são as covariáveis fixas tomadas na posição (s_i) e $(\beta_1, \dots, \beta_n)^T$ são parâmetros desconhecidos a serem estimados. A Equação (3.2) pode ser reescrita na forma matricial:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e},$$

em que, $E(\mathbf{e}) = \mathbf{0}$ e a matriz de covariância $\boldsymbol{\Sigma} = [(\sigma_{ij})]$ sendo $(\sigma_{ij}) = C(s_i, s_j)$. Assume-se então que $\boldsymbol{\Sigma}$ é não singular, simétrica e definida positiva, enquanto \mathbf{Y} segue uma distribuição t-Student reparametrizada com vetor de médias $\mathbf{X}\boldsymbol{\beta}$ e matriz de covariância $\boldsymbol{\Sigma}$ definida na forma paramétrica por:

$$\boldsymbol{\Sigma} = [(\sigma_{ij})] = \phi_1 \mathbf{I}_n + \phi_2 \mathbf{R}, \quad (3.3)$$

sendo :

ϕ_1 o parâmetro que define o efeito pepita;

ϕ_2 o parâmetro que define a contribuição;

\mathbf{I}_n a matriz identidade $n \times n$;

$\mathbf{R} = \mathbf{R}(\phi_3)$ é uma matriz simétrica $n \times n$, em que o alcance (a) é uma função de ϕ_3 , ou seja, $a = g(\phi_3)$.

Os elementos σ_{ij} da matriz $\boldsymbol{\Sigma}$ são $C(s_i, s_j) = C(h_{ij}) = \phi_2 r_{ij}$ em que $h_{ij} = \|s_i - s_j\|$ é a distância euclidiana entre os pontos s_i e s_j , para $i \neq j = 1, \dots, n$, sendo os elementos r_{ij} de \mathbf{R} definido por:

$$r_{ij} = \begin{cases} 1, & \text{se } i = j \\ \frac{1}{\phi_2} C(h_{ij}), & \text{se } i \neq j \end{cases}.$$

Os parâmetros que determinam a estrutura de dependência espacial apresentada na Equação (3.3) são efeito pepita (ϕ_1), contribuição (ϕ_2), alcance ($a = g(\phi_3)$) e patamar

$$(C = \phi_1 + \phi_2).$$

Estes parâmetros, por sua vez, são utilizados na função semivariância para medir a dependência espacial. Na Figura 3, apresenta-se o semivariograma típico com características próximas à ideal, as quais são discutidas a seguir. Seu comportamento representa o que, intuitivamente, se deve esperar de dados no campo. Espera-se que as diferenças $[Y(s_i) - Y(s_i + h)]$ decresçam assim que h , a distância que os separa, diminua. Conseqüentemente, as medições localizadas próximas tornem-se mais parecidas entre si do que aquelas separadas por grandes distâncias. Dessa maneira, é de se esperar que $\gamma(h)$ aumente com a distância h . Por definição, $\gamma(0) = 0$, como pode ser visto pela equação (3.6), quando $h = 0$. Entretanto, na prática, à medida que h tende para 0, $\gamma(h)$ aproxima-se de um valor positivo denominado de efeito pepita e que recebe o símbolo (ϕ_1) . Resultados com valores de efeito pepita maiores que zero são frequentemente encontrados em trabalhos com dados agrícolas (GARCIA-PAPANI et al., 2018; FAGUNDES et al., 2018; DE BASTIANI et al., 2017).

O efeito pepita representa o erro ou a variabilidade em pequena escala, isto é, revela a descontinuidade do semivariograma para distâncias menores do que a menor distância entre os pontos amostrados. Segundo Isaaks e Srivastava (1989), a descontinuidade deve-se à variabilidade de pequena escala não captada pela malha de amostragem ou possível erro de medição. A Contribuição (ϕ_2) é denominada pela variância de dispersão e representa as diferenças espaciais entre os valores de uma variável tomada em dois pontos separados por distâncias cada vez maiores, definida pela diferença entre o patamar $(C = \phi_1 + \phi_2)$ e o efeito pepita (ϕ_1) . O Patamar (C) é o valor no qual o semivariograma se estabiliza, ou seja, à medida que h aumenta, $\gamma(h)$ também aumenta até um valor máximo no qual se permanece. Este por sua vez correspondente à soma do efeito pepita (ϕ_1) e à contribuição (ϕ_2) , sendo aproximadamente igual à variância dos dados $Var(Y)$. A distância na qual $\gamma(h)$ atinge o patamar é chamada de alcance, sendo denominado por a , representada pela função de parâmetro ϕ_3 ($a = g(\phi_3)$). Medições localizadas a distâncias maiores que a têm distribuição espacial aleatória e por isto são independentes entre si. Para estas amostras, a Estatística Clássica pode ser aplicada sem restrições. Por outro lado, amostras separadas por distâncias menores que a são correlacionadas umas às outras, o que permite que se façam interpolações para espaçamentos menores do que os amostrados. Dessa maneira, o alcance a é a linha divisória para a aplicação da geoestatística ou Estatística Clássica.

Segundo Vieira et al. (2000), os dados que apresentam semivariogramas semelhantes ao da Figura 3, provavelmente, serão estacionários de segunda ordem, pois apresentam o patamar claro, definido e, certamente, estarão sob a hipótese intrínseca. Se o semivariograma, ao invés de ser crescente e dependente de h como o mostrado na Figura 3, for constante e igual ao patamar para qualquer valor de h então, tem-se um efeito pepita puro ou ausência total de dependência espacial. Isto significa que o alcance a , para os dados em questão, é menor do que o menor espaçamento entre amostras. Para esses dados, a única estatística aplicável é a Estatística Clássica, pois apresenta-se uma distribuição espacial completamente aleatória (EGUCHI et al., 2002).

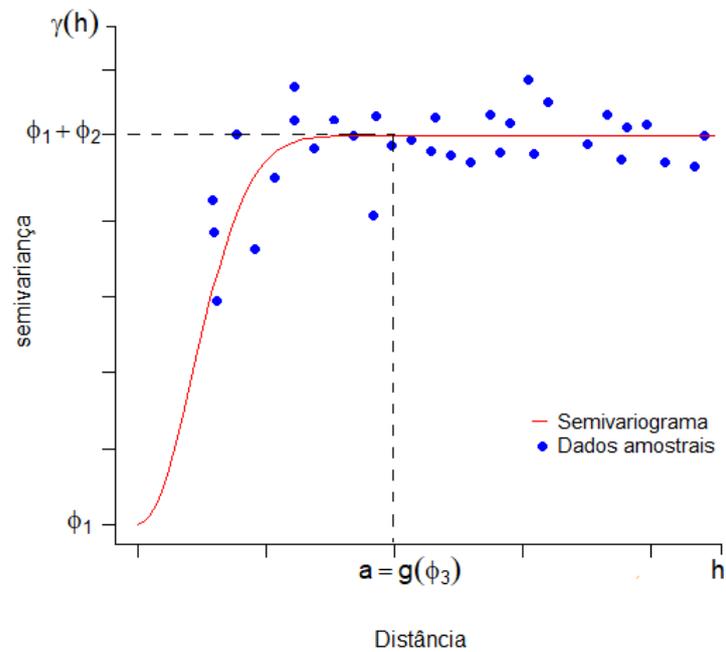


Figura 3 Função semivariância e seus respectivos parâmetros

Outro tipo de semivariograma que pode ocorrer é aquele que cresce, sem limites, para todos os valores de h calculados. Esse semivariograma indica a presença de fenômeno com capacidade infinita de dispersão, o qual não tem variância finita e para o qual a covariância não pode ser definida. Ele indica, também, que o tamanho do campo amostrado não foi suficiente para exibir toda a variância e é provável que exista uma grande tendência nos dados em determinada direção. A seguir, definimos a função semivariância e covariância.

Formalmente, uma variável regionalizada é estacionária se os momentos estatísticos da variável aleatória $Y(s_i + h)$ são os mesmos para qualquer vetor h . De acordo com o número k_m de momentos estatísticos que são constantes, a variável é chamada de estacionária de ordem k_m . Segundo Olea (1975), na geoestatística é requerida a estacionariedade de ordem $k_m = 2$, ou seja, estacionária de segunda ordem.

Por consequência das definições 3.2 e 3.3, supõe-se que o processo estocástico $Y(s_i)$ tem valor esperado:

$$E[Y(s_i)] = \mu(s_i) \text{ e } E[Y(s_i + h)] = \mu(s_i + h),$$

e variâncias

$$\text{Var}[Y(s_i)] \text{ e } \text{Var}[Y(s_i + h)],$$

correspondentes às localizações s_i , $s_i + h$ e qualquer vetor de distância h . Então, a função de covariância espacial é apresentada em (3.4):

$$C(s_i, s_i + h) = C(h) = E[Y(s_i)Y(s_i + h)] - E[Y(s_i)]E[Y(s_i + h)], \quad (3.4)$$

e a função de semivariância é definida por

$$\gamma(h) = \gamma(s_i, s_i + h) = \frac{1}{2}E[(Y(s_i) - Y(s_i + h))^2], \quad (3.5)$$

o estimador clássico da função de semivariância, proposto por Matheron (1963), baseado no método dos momentos (Cressie, 1993), é dado em (3.6):

$$\hat{\gamma}(h) = \hat{\gamma}(s_i, s_k) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (Y(s_i) - Y(s_k))^2, \quad (3.6)$$

em que, $N(h)$ é o número de pares com distância h de $\{(s_i, s_k) : \|s_i - s_k\| = h, \forall s_i \neq s_k\}$.

Com um único elemento amostral em cada ponto, o que se sabe de um processo estocástico $Y(s_i)$ é uma única realização. Então, quando se deseja estimar valores para locais não amostrados, é necessária a suposição de que a variável regionalizada seja estacionária de segunda ordem. Sob a hipótese de estacionariedade, Isaaks e Srivastava (1989) e Wackernagel (2003) demonstraram, a partir das equações apresentadas em (3.4) e (3.5), a função semivariância dada por

$$\gamma(h) = C(0) - C(h). \quad (3.7)$$

A partir da equação (3.7), tem-se a relação $C(h) = C(0) - \gamma(h)$, denominada função de covariância. E ao se dividir esta relação por $C(0)$, em que $C(0) \neq 0$, tem-se:

$$\frac{C(h)}{C(0)} = \frac{C(0)}{C(0)} - \frac{\gamma(h)}{C(0)},$$

$$\frac{C(h)}{C(0)} = 1 - \frac{\gamma(h)}{C(0)},$$

obtem-se assim a função de correlação espacial dada por

$$\rho(h) = 1 - \frac{\gamma(h)}{C(0)}.$$

Na Figura 4, apresenta-se a relação da função de covariância espacial $C(h)$ e a função semivariância $\gamma(h)$ com seus respectivos parâmetros ϕ_1 , ϕ_2 e ϕ_3 , pelos quais se busca quantificar a dependência espacial das variáveis regionalizadas em estudo. Observa-se que a forma do gráfico é a mesma, porém, para a função covariância é invertida. Enquanto o gráfico da função de covariância inicia a partir de um valor de máximo $C = \phi_1 + \phi_2$ em $h = 0$ e decresce para zero, o semivariograma inicia em zero e cresce até um máximo $C = \phi_1 + \phi_2$

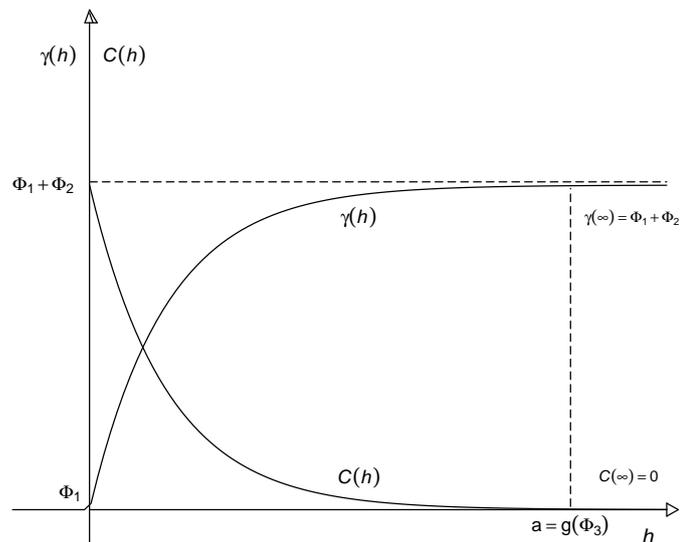


Figura 4 Relação entre a função de covariância $C(h)$ e função semivariância $\gamma(h)$

Vale destacar ainda que uma análise detalhada do semivariograma fornece informações importantes, pois, quando o semivariograma é aproximadamente idêntico para todas as direções convencionais: 0° , 45° , 90° e 135° , sugerindo o mesmo padrão de continuidade espacial, o estudo da dependência espacial pode ser feito através de um único semivariograma, chamado de semivariograma omnidirecional. Obtendo-se assim um comportamento isotrópico, sendo de suma importância para a estimação de locais não amostrados. Caso sejam detectados diferentes padrões de dependência espacial, isto é, apresenta uma variabilidade que não é a mesma em todas as direções, o fenômeno em estudo é chamado de anisotrópico. Portanto, uma correção deve ser aplicada, obtendo-se assim um comportamento isotrópico. Tem-se a seguinte definição de isotropia, dada por Piegorsch e Bailer (2005):

Definição 3.4. Um processo estocástico $Y(s)$ em um espaço bi-dimensional S é isotrópico se $Var[Y(s)]$ existe para todo $s \in S$, tais que:

- (i) $Y(s)$ é estacionário;
 - (ii) Para $h = \|s_i - s_k\|$, $C(s_i, s_k) = Cov[Y(s_i), Y(s_k)]$ depende somente s_i e s_k através de h .
- Um processo anisotrópico é um processo $Y(s)$ que não satisfaz os critérios (i) e (ii).

Para mais detalhes, veja (ISAAKS; SRIVASTAVA, 1989; DIGGLE, 2007; GUEDES et al., 2008).

3.2.3 Modelo da família Matérn

A partir do estudo da dependência espacial, definindo-se $\phi = (\phi_1, \phi_2, \phi_3)^\top$, é possível modelar essa estrutura com o auxílio dos modelos teóricos. Segundo DIGGLE (2007), esses modelos são divididos em não transitivos, os quais não possuem patamar, enquanto os modelos transitivos possuem patamar. Dentre os modelos transitivos destaca-se o modelo da família Matérn, cuja a expressão é definida em função dos termos do modelo teórico do semivariograma

dado por:

$$\gamma(h) = \begin{cases} \phi_1, & \text{se } h = 0, \\ \phi_1 + \phi_2 \left[1 - (2^{\kappa-1}\Gamma(\kappa))^{-1} \left(\frac{h}{\phi_3} \right)^\kappa K_\kappa \left(\frac{h}{\phi_3} \right) \right], & \text{se } h > 0, \end{cases}$$

sendo $K_\kappa(u) = \frac{1}{2} \int_0^\infty x^{k-1} e^{\frac{1}{2}u\left(\frac{x^2+1}{x}\right)} dx$, em que $K_\kappa(\cdot)$ é a função de Bessel de terceiro tipo de ordem k e κ é o parâmetro forma ($\kappa \in \mathbb{R}, \kappa > 0$) (GRADSHTEYN; RYZHIK, 2014). A função covariância da família Matérn é particularmente atraente porque seu comportamento pode mudar dependendo do valor de κ , ou seja, quando $\kappa = 0,5$, corresponde à estrutura de covariância do modelo Exponencial, para $\kappa = 1$ refere-se ao modelo Whittle e quando $\kappa \rightarrow \infty$, a estrutura de covariância se aproxima do modelo Gaussiano, suas equações estão descritas na Tabela 1. Neste modelo, a relação entre o alcance (a) e o parâmetro ϕ_3 depende do valor do parâmetro de forma κ , ou seja, quanto maior o valor de κ , maior o valor do alcance. Assim, define-se a função de covariância por:

$$C(h) = \begin{cases} \phi_1 + \phi_2, & \text{se } h = 0, \\ \phi_2 \left[(2^{\kappa-1}\Gamma(\kappa))^{-1} \left(\frac{h}{\phi_3} \right)^\kappa K_\kappa \left(\frac{h}{\phi_3} \right) \right], & \text{se } h > 0. \end{cases}$$

Na Tabela 1, apresentam-se os casos especiais da função de covariância da família Matérn:

Tabela 1 Função de covariância da família Matérn

Parâmetros	Covariância	Modelos
$\kappa = 1/2$	$C(h) = \varphi_2 \exp(-h/\varphi_3)$	Exponencial
$\kappa = 1$	$C(h) = \varphi_2 (h/\varphi_3) K_\kappa(h/\varphi_3)$	Whittle
$\kappa \rightarrow \infty$	$C(h) = \varphi_2 \exp\left(- (h/\varphi_3)^2\right)$	Gaussiano

$K_\kappa(\cdot)$: é a função de modificação Bessel de ordem k

Fonte: De Bastiani et al. (2016)

A função de correlação espacial é expressa como:

$$r_{ij}(h) = \begin{cases} 1, & \text{se } h = 0, \\ (2^{\kappa-1}\Gamma(\kappa))^{-1} \left(\frac{h}{\phi_3} \right)^\kappa K_\kappa \left(\frac{h}{\phi_3} \right), & \text{se } h > 0. \end{cases}$$

Na Figura 5, verifica-se o comportamento da função de correlação para a família Matérn considerando diferentes valores do parâmetro de forma κ . Observa-se que, ao se fixar o valor de ϕ_3 e se aumentar o valor de κ , sabe-se que o valor de a/ϕ_3 também aumenta. Conforme a distância aumenta, a função de correlação se aproxima para os diferentes valores κ .

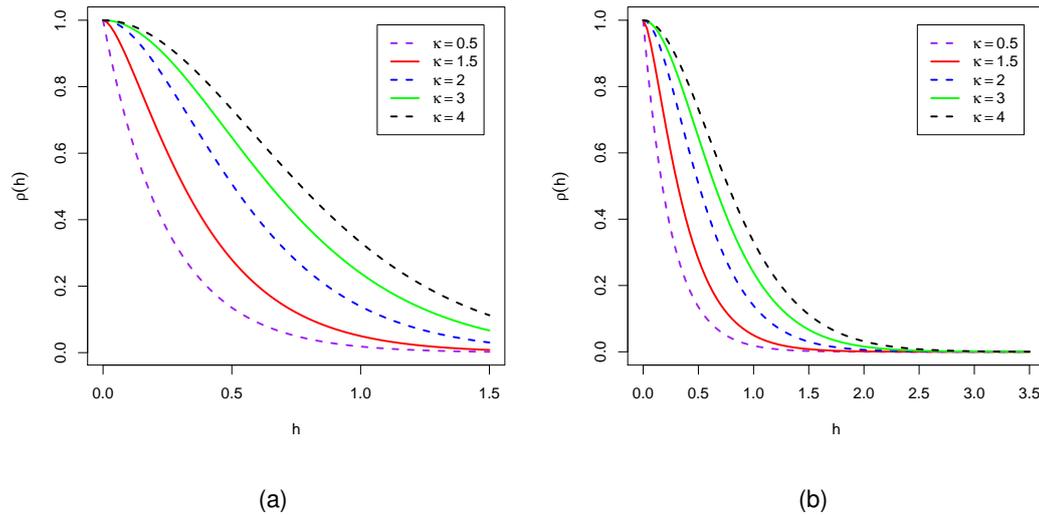


Figura 5 Função de correlação da família Matérn para diferentes valores de κ em que (a) e (b) apresentam distância de 0 a 1,5 e 0 a 3,5, respectivamente

Matheron (1963) mostrou mais precisamente que se um campo aleatório permite um variograma que está em toda parte contínua, exceto na origem, então este campo aleatório é a soma de dois campos aleatórios não correlacionados, um associado ao efeito pepita puro e outro com um variograma contínuo em toda parte. Os resultados de Zhang (2004) demonstraram que no modelo da família Matérn com os parâmetros ϕ_2 e ϕ_3 , sendo $\kappa = 0,5$, a razão ϕ_2/ϕ_3 é muito mais estimada de forma estável que ϕ_2 e ϕ_3 . Stein (1999) recomendou fortemente utilizar o modelo da família Matérn para calcular e representar graficamente a função verossimilhança para os parâmetros desconhecidos dos modelos com estruturas de covariância.

Para o caso da função de covariância da família Matérn (ou outra com mais de três parâmetros), sugerimos utilizar a ideia apresentada em Minasny e McBratney (2005) para uma única realização, que assume κ como fixo para descobrir qual κ fornece o valor máximo para o perfil de log-verossimilhança e, em seguida, usá-lo como um valor inicial e estimá-lo.

3.2.4 Krigagem

Após obter o modelo com os parâmetros estimados, por meio da modelagem do semivariograma e critérios de validação, aplica-se a Krigagem, que consiste em um método de interpolação, o qual utiliza a estrutura de dependência espacial para a estimação do valor de um atributo em locais não amostrados com base em pontos amostrados na mesma área. Baseia-se no princípio da Primeira Lei da Geografia definida por Waldo Tobler, o qual afirma que “todas as coisas estão relacionadas, mas coisas mais próximas são mais parecidas do que coisas mais distantes” (TOBLER, 1970). Por sua vez, a krigagem utiliza-se de funções matemáticas para acrescentar pesos maiores nas posições mais próximas aos pontos amostrais e pesos menores nas posições mais distantes, a fim de obter novos pontos interpolados com base nessas combinações lineares de dados.

Este procedimento é semelhante ao de interpolação por médias móveis ponderadas, exceto porque, aqui, os pesos são determinados a partir de uma análise espacial baseada na estimação de parâmetros obtidos pela modelagem da dependência espacial. Segundo Assumpção et al. (2011), os estimadores fornecidos pela Krigagem não são tendenciosos, pois, em média, a diferença entre o valor predito e o valor amostrado na mesma localização deve ser nula, e possuir variância mínima, chamados de BLUE (Best Linear Unbiased Estimator) ou BLUP (Best Linear Unbiased Predictor), considerados por LANDIM (2000) como os melhores estimadores lineares não viesados.

Portanto, a krigagem é uma técnica especificamente desenvolvida para a geoestatística, pois leva em consideração os parâmetros obtidos pela modelagem da dependência espacial. E também possibilita a construção de mapas temáticos com alta precisão, que apresentam a real variabilidade local.

3.2.5 Krigagem com *drift* externo

Na krigagem ordinária assume-se a existência de variância finita e a hipótese de estacionariedade em relação à média, isto é, $E(Z(s)) = \mu$, com μ constante para todo s do espaço amostral. No entanto, se uma informação auxiliar, correlacionada com a variável resposta, está disponível em todos os pontos da grade, distintos espacialmente, 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 previsões e diferem apenas em passos metodológicos (HENGL et al., 2003; HENGL, 2009).

O objetivo da krigagem com drift externo é encontrar λ_i , $i = 1, \dots, n$, tal que $\hat{Y}(s_0) = \sum_{i=1}^n \lambda_i Y(s_i)$, com erro de predição não enviesado e variância mínima, em que $s_0 \in S \subset \mathbb{R}^n$, é um ponto não amostrado.

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

$$X_{0j} = \sum_{i=1}^n \lambda_i X_{ij}, \quad j = 0, \dots, p, \quad (3.8)$$

em que

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

Para o estabelecimento das condições que minimizem a variância dos erros, é conveniente definir que $Y(s) = \mu(s) + \epsilon(s)$, sendo $E(\epsilon(s)) = 0$, a covariância entre $Y(s_i)$ e $Y(s_j)$, depende apenas da distância entre s_i e s_j , isto é, $\text{Cov}(s_i, s_j) = C(h_{ij})$, em que h_{ij} é a distância euclidiana entre s_i e s_j (DAVIS, 2002; GARCIA-PAPANI et al., 2018). Desta forma,

a variância da estimação pode ser escrita em função dos resíduos $\epsilon(s)$, em que:

$$\begin{aligned}
 \widehat{Y}(\mathbf{s}_0) - Y(\mathbf{s}_0) &= \left(\sum_{i=1}^n \lambda_i Y(\mathbf{s}_i) \right) - Y(\mathbf{s}_0) \\
 &= \left(\sum_{i=1}^n \lambda_i (\mu(\mathbf{s}_i) + \epsilon(\mathbf{s}_i)) \right) - (\mu(\mathbf{s}_0) + \epsilon(\mathbf{s}_0)) \\
 &= \left(\sum_{i=1}^n \lambda_i (\mu(\mathbf{s}_i)) - \mu(\mathbf{s}_0) \right) + \left(\sum_{i=1}^n \lambda_i (\epsilon(\mathbf{s}_i)) - \epsilon(\mathbf{s}_0) \right) \\
 &= \left(\sum_{i=1}^n \lambda_i \epsilon(\mathbf{s}_i) \right) - \epsilon(\mathbf{s}_0) \\
 &= \widehat{\epsilon}(\mathbf{s}_0) - \epsilon(\mathbf{s}_0).
 \end{aligned}$$

Portanto, a variância estimada é dada por:

$$\begin{aligned}
 \text{Var} \left(\widehat{Y}(\mathbf{s}_0) - Y(\mathbf{s}_0) \right) &= \text{E} \left((\widehat{\epsilon}(\mathbf{s}_0) - \epsilon(\mathbf{s}_0))^2 \right) \\
 &= C(0) + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(\mathbf{s}_i, \mathbf{s}_j) - 2 \sum_{i=1}^n C(\mathbf{s}_0, \mathbf{s}_i).
 \end{aligned} \tag{3.9}$$

O problema de minimização é idêntico à krigagem ordinária. Portanto, deriva-se a equação 3.9 em relação à λ_i , $i = 1, \dots, n$ e igualar as derivadas obtidas a zero. Contudo, pretende-se ainda que sejam cumpridas as condições estabelecidas pelas $p + 1$ equações definidas em (3.8). Assim, acrescentam-se $p + 1$ termos nulos na equação 3.9, correspondentes a estas condições. Então, tem-se que

$$\text{Var} \left(\widehat{Y}(\mathbf{s}_0) - Y(\mathbf{s}_0) \right) = C(0) + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(\mathbf{s}_i, \mathbf{s}_j) - 2 \sum_{i=1}^n C(\mathbf{s}_0, \mathbf{s}_i) + 2 \sum_{j=0}^p \eta_j \sum_{i=1}^n \lambda_i X_{ij} - X_{0j},$$

pois η_j são parâmetros de Lagrange. Igualando a zero as $n + p + 1$ derivadas em relação à λ_i , $i = 1, \dots, n$ e η_j , $j = 0, \dots, p$, obtém-se o seguinte sistema de equações lineares:

$$S = \begin{cases} \sum_{k=1}^n \lambda_k C(\mathbf{s}_i, \mathbf{s}_k) + \sum_{k=0}^p \eta_k X_{ik} = C(\mathbf{s}_i, \mathbf{s}_0), & i = 1, \dots, n. \\ \sum_{k=1}^n \lambda_k X_{kj} = X_{0j}, & j = 0, \dots, p. \end{cases}$$

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

$(n + p + 1)$, dada por,

$$C = \begin{pmatrix} C(\mathbf{s}_1, \mathbf{s}_1) & \cdots & C(\mathbf{s}_1, \mathbf{s}_n) & 1 & X_{11} & \cdots & X_{1p} \\ C(\mathbf{s}_2, \mathbf{s}_1) & \cdots & C(\mathbf{s}_2, \mathbf{s}_n) & 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ C(\mathbf{s}_n, \mathbf{s}_1) & \cdots & C(\mathbf{s}_n, \mathbf{s}_n) & 1 & X_{n1} & \cdots & X_{np} \\ 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ X_{11} & \cdots & X_{n1} & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ X_{p1} & \cdots & X_{np} & 0 & 0 & \cdots & 0 \end{pmatrix};$$

λ é o vetor, $(n+p+1) \times 1$, dado por $\lambda = (\lambda_1, \dots, \lambda_n, \eta_0, \dots, \eta_p)^\top$ e A é o vetor, $(n+p+1) \times 1$, dado por $A = (C(\mathbf{s}_1, \mathbf{s}_0), \dots, C(\mathbf{s}_n, \mathbf{s}_0), 1, X_{01}, \dots, X_{0p})$. Desta forma, os pesos $\lambda_i, i = 1, \dots, n$ são obtidos por $\lambda = C^{-1}A$.

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 et al., 2003). Um exemplo de krigagem universal é apresentado em (ASSUMPÇÃO et al., 2014; SCHEMMER et al., 2017; GARCIA-PAPANI et al., 2018).

3.3 Caracterização das variáveis

O objetivo desta seção é apresentar uma breve revisão sobre as variáveis que serão utilizadas neste trabalho. O modelo proposto considera como variável resposta a produtividade da soja e com variáveis explicativas os atributos físicos e químicos do solo.

3.3.1 Produtividade da soja

A soja (*Glycine max*) é um dos principais produtos de exportação e uma das principais *commodities* do Planeta. É cultivada praticamente em todo o território nacional, sendo o principal produto agrícola do País. Com o desenvolvimento de cultivares adaptadas às diferentes regiões agroclimáticas, o Brasil tornou-se o segundo maior produtor mundial de soja, predominantemente utilizada para o processamento do grão em óleo e proteína. A proteína processada denominada de torta ou farelo é utilizada como suplemento protéico na ração animal. Na safra 2016/2017, a cultura ocupou uma área de 33,89 milhões de hectares, o que totalizou uma produção de 113,92 milhões de toneladas e, a produtividade média da soja brasileira foi de 3,362 kg ha⁻¹ (EMBRAPA, 2018b). De modo geral, as cultivares brasileiras de soja são classificadas em grupos de maturação (GM), diferentes em cada região, no estado do Paraná são definidas como: precoce (até 115 dias); semiprecoce (116 – 125 dias); médio (126 – 137 dias) e semitardio (138 – 145 dias) (EMBRAPA, 2018a).

São expressivos os estudos no melhoramento genético da soja, apresentando cultivares com alto potencial produtivo, com boa sanidade e precocidade, os quais estão associados ao equilíbrio nutricional da planta (EMBRAPA, 2018b). A alta produtividade da lavoura, está

relacionada com a análise de solo, o histórico da área, o manejo do solo, entre outras tecnologias para melhorar as recomendações e aplicações de adubação. O custo crescente dos insumos agrícolas exige, cada vez mais, a adoção de técnicas de cultivos adequadas na produção das culturas anuais. Segundo Dalchiavon et al. (2011), a integração de diversas fontes de dados, dentre elas os dados físicos e químicos do solo, proporciona o melhor gerenciamento da produção agrícola, de maneira a reconhecer que as culturas e os solos não devem ser manejados individualmente e de forma homogênea. Montanari et al, (2010) ressaltam que a variabilidade espacial dos atributos do solo deve ser bem conhecida, visando minimizar os erros de amostragem e de manejo, pois o solo apresenta heterogeneidade, tanto vertical quanto horizontalmente. Neste sentido, Basso et al. (2011) afirmaram que os mapeamentos destes atributos em áreas agrícolas são fundamentais para as recomendações de práticas de manejo. Portanto, a aplicação de tecnologia associada à variabilidade espacial se faz necessária, sobretudo na pesquisa agrícola, que estuda o solo e sua capacidade produtiva (GREGO;VIEIRA, 2005).

3.3.2 Dados químicos

Os dados químicos do solo influenciam diretamente o crescimento e o desenvolvimento da cultura da soja. Conhecer a disponibilidade de macro, micronutrientes, a relação entre os nutrientes e as condições de acidez do solo é necessário para determinar a fertilidade do solo (LOPES, 1998). Dezesesseis 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: macronutrientes primários são nitrogênio, fósforo e potássio, os quais geralmente tornam-se deficientes no solo antes dos demais, pois as plantas utilizam em quantidades relativamente grandes; os macronutrientes secundários são cálcio, magnésio e enxofre, essenciais para o crescimento das plantas (ABREU et al., 2007); e os micronutrientes definidos por boro, cloro, cobre, ferro, manganês, molibdênio e zinco, responsáveis pelo metabolismo vegetal, seja como constituintes de compostos ou como reguladores do funcionamento enzimático (FERREIRA et al., 2001). Os macronutrientes secundários e os micronutrientes são geralmente menos deficientes e usados pelas plantas em quantidades menores, mas são tão importantes quanto os macronutrientes primários.

O objetivo desta Seção é apresentar um breve relato sobre os atributos químicos do solo, em especial sobre os nutrientes minerais que serão utilizados no estudo: fósforo (P), potássio (K), magnésio (Mg) e cálcio (Ca).

O potássio (K) é absorvido pelas plantas na forma de íon K^+ . Na solução do solo, o potássio é móvel e, também, sujeito às perdas por lixiviação, ocorrendo, principalmente, em solos ácidos e com baixa CTC, erosão e remoção por colheitas. Um dos principais problemas causados pela deficiência de K nos solos é a ocorrência de acamamento das

plantas, enfraquecimento de caules e atrasos na floração, conseqüentemente prejudicando a produtividade da cultura.

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. Segundo Malavolta (1997), a maior parte do P absorvida pela planta é transferida e armazenada para o fruto ou grão e concentra-se nas áreas mais ativas de crescimento. As formas mais solúveis do P no solo ocorrem na amplitude de pH entre 5,5 a 7,0.

O cálcio (Ca) promove a redução da acidez do solo, melhora o crescimento das raízes, o aumento da atividade microbiana, o aumento da disponibilidade de molibdênio e de outros nutrientes (LOPES, 1998). O cálcio reduz a acidez do solo e diminui a toxidez do alumínio (Al), cobre (Cu) e manganês (Mn). Plantas que apresentam altos teores de cálcio resistem melhor à toxidez destes elementos. As vagens chochas na soja e as folhas enroladas no milho são sintomas de deficiência de cálcio.

O magnésio (Mg) 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 (OLIVEIRA et al., 1991). A deficiência de magnésio resulta em dificuldade na translocação de carboidratos para a raiz e interfere no desenvolvimento do sistema radicular que, por sua vez, reduzirá a absorção de outros nutrientes. Assim como a presença do magnésio, em dose inadequada no sistema solo-planta prejudica o crescimento e a vitalidade da planta (LOPES, 1998).

3.3.3 Matéria orgânica

Segundo Conceição et al. (2005), a matéria orgânica (MO) é considerada indicadora de qualidade do solo, por sua influência nos demais atributos essenciais para que o solo desempenhe suas funções. Lal (2004) afirma a importância da MO, que serve de base para a produtividade agrícola sustentável, pois através dos seus efeitos diretos, é capaz de modular as condições químicas, físicas e biológicas do solo.

Unger et al. (1991) relatam que a cobertura do solo, a partir da palhada das lavouras, auxilia no controle da erosão hídrica, e que os resíduos da superfície do solo o protegem do impacto direto da chuva, pois eles dissipam os pingos de chuva, bem como minimizam o desprendimento das partículas de solo, posto que auxilia na infiltração de água e reduzem o transporte das partículas superficiais. Boulal et al. (2011) e Conceição et al. (2005) confirmam os benefícios da MO para o solo no sistema agrícola, com destaque para o controle térmico, a melhor retenção de água, estimulação da microbiota do solo, condicionamento físico do

solo, efeito tampão biológico e químico, fonte de nutrientes às plantas e a ciclagem de nutrientes. Segundo Blanco-Canqui e Lal (2009) quanto maior a quantidade de resíduos orgânicos retornados para o solo, maior a cobertura da superfície do solo e maior a proteção da estrutura do solo contra perturbações naturais e antropogênicas.

3.3.4 pH

Segundo Malavolta (1997), entre os atributos do solo, o pH, índice que indica o grau de acidez ativa do solo, talvez seja, isoladamente, o mais relevante, no que se refere à utilização de fertilizantes. Pois, este atributo auxilia na disponibilidade de nutrientes contidos no solo, ou adicionada por meio de adubação para a planta, ou seja, a interação entre a planta e o solo. Segundo Da Silva et al. (2011), níveis de pH muito baixos podem interferir negativamente na sobrevivência de bactérias fixadoras de nitrogênio. A acidez do solo traz inúmeros prejuízos à produção das culturas anuais, dentre eles destaca-se a presença do Al trocável. A presença deste elemento ocasiona o engrossamento das raízes, desorganizando a estrutura das membranas e compromete a absorção de nutrientes (TAIZ et al., 2017). Assim, a calagem prévia dos solos ácidos, além de proporcionar aumento do pH e da saturação por base, promove a neutralização do alumínio e de grande parte do ferro e do manganês; aumenta a atividade biológica e a eficiência dos fertilizantes, resultando ainda em diminuição na capacidade de fixação via precipitação do P, favorece conseqüentemente, o desenvolvimento vegetal (ERNANI et al., 1996). Segundo Rutkowska et al. (2015), a utilização da calagem às plantas enaltece a importância na correção da acidez do solo, ou seja, de um pH adequado para o cultivar, para que a cultura possa expressar seu máximo potencial produtivo.

3.4 Referências

- ABRAMOWITZ, M., STEGUN, I. A. (Ed.). **Handbook of mathematical functions with formulas, graphs, and mathematical tables**. US Government printing office, 1970.
- ABREU, C. D., LOPES, A. S., SANTOS, G. D.. Micronutrientes. **Fertilidade do solo**, v. 1, p. 645-736, 2007.
- AMADO, T. J. C., PONTELLI, C. B., SANTI, A. L., VIANA, J. H. M., SULZBACH, L. A. D. S. Variabilidade espacial e temporal da produtividade de culturas sob sistema plantio direto. **Pesquisa Agropecuária Brasileira**, v. 42, n. 8, p. 1101-1110, 2007.
- ANDA. **Anuário estatístico do setor de fertilizantes**. [S.l.], 2014.
- ANDERSON, T. **An Introduction to Multivariate Statistical Analysis**. 3. ed. New York: Wiley, 2003.
- ASSUMPÇÃO, R. A. B.; URIBE-OPAZO, M. A.; GALEA, M. Analysis of local influence in geostatistics using Student's t-distribution. **Journal of Applied Statistics**, v. 41, n. 11, p. 2323-2341, 2014.

ASSUMPÇÃO, R. A. B., OPAZO, M. A. U., GALEA, M. Local influence for spatial analysis of soil physical properties and soybean yield using student's t-distribution. **Revista Brasileira de Ciência do Solo**, v. 35, n. 6, p. 1917-1926, 2011.

ASSUMPÇÃO, R. A. B., URIBE-OPAZO, M. A., DE SOUZA, E. G., JOHANN, J. A. Uso da krigagem indicatriz na avaliação da probabilidade da produtividade de soja segundo os padrões regional, estadual e nacional. **Acta Scientiarum. Agronomy**, v. 29, n. 2, p. 165-171, 2007.

BASSO, F. C., ANDREOTTI, M., CARVALHO, M. D. P., LODO, B. N., MONTANARI, R. Correlação linear e espacial entre a produtividade e o teor de proteína bruta do guandu anão e os atributos de um Latossolo. **Revista Brasileira de Ciências Agrárias**, p. 521-530, 2011.

BLANCO-CANQUI, H., LAL, R. Crop residue removal impacts on soil productivity and environmental quality. **Critical reviews in plant science**, v. 28, n. 3, p. 139-163, 2009.

BOLFARINE, H., ROJAS, M. G. On structural comparative calibration under a t-model. **Computational Statistics**, v. 11, n. 1, p. 63-85, 1996.

BOLFERINE, H., ARELLANO-VALLE, R. B. Robust modelling in measurement error models using the t distribution. **Brazilian Journal of Probability and Statistics**, p. 67-84, 1994.

BOULAL, H., GÓMEZ-MACPHERSON, H., GÓMEZ, J. A., MATEOS, L. Effect of soil management and traffic on soil erosion in irrigated annual crops. **Soil and Tillage Research**, v. 115, p. 62-70, 2011.

CONCEIÇÃO, P. C., AMADO, T. J. C., MIELNICZUK, J., SPAGNOLLO, E. Qualidade do solo em sistemas de manejo avaliada pela dinâmica da matéria orgânica e atributos relacionados. **Revista Brasileira de Ciência do Solo**, v. 29, n. 5, p. 777-788, 2005.

CRESSIE, N. **Statistics for Spatial Data**. [S.l.]: John Wiley, New York, 1993.

CUNHA, J. D., CASARIN, V., PROCHNOW, L. I. Balanço de nutrientes na agricultura brasileira. **Informações agronômicas**, v. 130, p. 1-11, 2010.

CYSNEIROS, F. J. A., PAULA, G. A., GALEA, M. Modelos simétricos aplicados. **Escola de Modelos de Regressao**, v. 9, 2005.

DALPOSSO, G. H., URIBE-OPAZO, M. A., MERCANTE, E., JOHANN, J. A., BORSSOI, J. A. Comparison measures of maps generated by geostatistical methods. **Engenharia Agrícola**, v. 32, n. 1, p. 174-183, 2012.

DALCHIAVON, F. C., CARVALHO, M. P., FREDDI, O. S., ANDREOTTI, M., MONTANARI, R. Variabilidade espacial da produtividade do feijoeiro correlacionada com atributos químicos de um Latossolo Vermelho Distroférrico sob sistema de semeadura direta. **Bragantia**, v. 70, n. 4, p. 908-916, 2011.

DA SILVA, A. F., SCHONINGER, E. L., MONTEIRO, S., CAIONE, G., DE CARVALHO, M. A. C., DALCHIAVON, F. C., NOETZOLD, R. Inoculação com bradyrhizobium e formas de aplicação de cobalto e molibdênio na cultura da soja. **Agrarian**, v. 4, n. 12, p. 98-104, 2011.

DAVIS, J. C. **Statistics and Data Analysis in Geology**. 3rd. ed. New York, NY, USA: John Wiley & Sons, Inc., 2002.

DE BASTIANI, F., GALEA, M., CYSNEIROS, A. H. M. A., URIBE-OPAZO, M. A. Gaussian spatial linear models with repetitions: An application to soybean productivity. **Spatial Statistics**, v. 21, p. 319-335, 2017.

DE BASTIANI, F., DE AQUINO CYSNEIROS, A. H. M., URIBE-OPAZO, M. A., GALEA, M. Influence diagnostics in elliptical spatial linear models. **Test**, v. 24, n. 2, p. 322-340, 2015.

DE BASTIANI, F., URIBE-OPAZO, M. A., DALPOSSO, G. H. Comparison of maps of spatial variability of soil resistance to penetration constructed with and without covariables using a spatial linear model. **Engenharia agrícola**, v. 32, n. 2, p. 393-404, 2012.

DEMPSTER, A. P., LAIRD, N. M., RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 39, n. 1, p. 1-22, 1977.

DIGGLE, P., RIBEIRO, J. Pj. Model based geostatistics. **SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA**, v. 14, p. 192, 2007.

EGUCHI, E. S., SILVA, E. L. D., OLIVEIRA, M. S. D. Variabilidade espacial da textura e da densidade de partículas em um solo aluvial no Município de Lavras, MG. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 6, n. 2, p. 242-246, 2002.

EMBRAPA. **Características da soja. Brasil**, 2018a. Disponível em: <http://www.agencia.cnptia.embrapa.br/gestor/soja/arvore/CONTAG01_24_271020069131.html>.

EMBRAPA. **Produção de soja no Brasil cresce mais de 13% ao ano**. Brasil, 2018b. Disponível em: <<https://www.embrapa.br/soja/cultivos/soja1>>.

ERNANI, P. R., FIGUEIREDO, O. A., BECEGATO, V., ALMEIDA, J. A. Decréscimo da retenção de fósforo no solo pelo aumento do pH. **Revista brasileira de ciência do solo**, v. 20, n. 1, p. 159-162, 1996.

FAGUNDES, R. S., URIBE-OPAZO, M. A., GUEDES, L. P. C., GALEA, M. Slash spatial linear modeling: soybean yield variability as a function of soil chemical properties. **Revista Brasileira de Ciência do Solo**, v. 42, 2018.

FERREIRA, M. E., CRUZ, M. D., RAIJ, B. V., ABREU, C. D. **Micronutrientes e elementos tóxicos na agricultura**. Jaboticabal: CNPq/FAPESP/POTAFOS, 2001.

FINEGOLD, M., DRTON, M. Robust graphical modeling of gene networks using classical and alternative t-distributions. **The Annals of Applied Statistics**, p. 1057-1080, 2011.

FIORENTINI, G., SENTANA, E., CALZOLARI, G. Maximum likelihood estimation and inference in multivariate conditionally heteroscedastic dynamic regression models with Student t innovations. **Journal of Business & Economic Statistics**, v. 21, n. 4, p. 532-546, 2003.

GALEA, M., CADEMARTORI, D., CURCI, R., MOLINA, A. Robust Inference in the Capital

- Asset Pricing Model Using the Multivariate t-Distribution. **Journal of Risk and Financial Management**, v. 13, n. 6, p. 123, 2020.
- GALEA, M., DE CASTRO, M. Robust inference in a linear functional model with replications using the t distribution. **Journal of Multivariate Analysis**, v. 160, p. 134-145, 2017.
- GARCIA-PAPANI, F., 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**, v. 31, n. 1, p. 105-124, 2017.
- GARCIA-PAPANI, F., LEIVA, V., URIBE-OPAZO, M. A., AYKROYD, R. G. Birnbaum-Saunders spatial regression models: Diagnostics and application to chemical data. **Chemometrics and Intelligent Laboratory Systems**, v. 177, p. 114-128, 2018.
- GUEDES, L. P. C., URIBE-OPAZO, M. A., JOHANN, J. A., SOUZA, E. G. D. Anisotropia no estudo da variabilidade espacial de algumas variáveis químicas do solo. **Revista Brasileira de Ciência do Solo**, v. 32, n. 6, p. 2217-2226, 2008.
- GRADSHTEYN, I. S., RYZHIK, I. M. **Table of integrals, series, and products**. Academic press, 2014.
- GREGO, C. R., VIEIRA, S. R. Variabilidade espacial de propriedades físicas do solo em uma parcela experimental. **Revista Brasileira de Ciência do Solo**, v. 29, n. 2, p. 169-177, 2005.
- GRIGELIONIS, B. **Student's t-distribution and related stochastic processes**. Berlin: Springer, 2013.
- HÄRDLE, W.; SIMAR, L. **Applied multivariate statistical analysis**. [S.l.]: Springer, New York, 2012.
- HENGL, T., HEUVELINK, G. B., STEIN, A. **Comparison of kriging with external drift and regression kriging**, Enschede, Netherlands: ITC, 2003.
- HENGL, T. **A practical guide to geostatistical mapping**. [S.l.]: Hengl, v. 52, 2009.
- ISAAKS, E. H., SRIVASTAVA, M. R. **Applied geostatistics**. 1989.
- KIBRIA, B. G., JOARDER, A. H. A short review of multivariate t-distribution. **Journal of Statistical research**, v. 40, n. 1, p. 59-72, 2006.
- KRIGE, D.G. A statistical approach to some basic mine evaluation problems on the witwatersrand. **Journal of South African Institution of Mining and Metallurgy**, Johannesburg, p.119-139, 1951.
- KOTZ, S., NADARAJAH, S. **Multivariate t-distributions and their applications**. Cambridge University Press, 2004.
- LAL, R. Soil carbon sequestration impacts on global climate change and food security. **Science**, v. 304, n. 5677, p. 1623-1627, 2004.

- LANDIM, P. M. B. Sobre geoestatística e mapas. **Terra e Didática**, v. 2, n. 1, p. 19-33, 2006.
- LANDIM, P. M. B. Introdução aos métodos de estimação espacial para confecção de mapas. **Rio Claro: UNESP**, p. 20, 2000.
- LANGE, K. L., LITTLE, R. J., TAYLOR, J. M. Robust statistical modeling using the t distribution. **Journal of the American Statistical Association**, v. 84, n. 408, p. 881-896, 1989.
- LANGE, K., SINSHEIMER, J. S. Normal/independent distributions and their applications in robust regression. **Journal of Computational and Graphical Statistics**, v. 2, n. 2, p. 175-198, 1993.
- LITTLE, R. J. Robust estimation of the mean and covariance matrix from data with missing values. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**, v. 37, n. 1, p. 23-38, 1988.
- LIU, C., RUBIN, D. B. ML estimation of the t distribution using EM and its extensions, ECM and ECME. **Statistica Sinica**, p. 19-39, 1995.
- LIU, C. ML Estimation of the Multivariate Distribution and the EM Algorithm. **Journal of Multivariate Analysis**, v. 63, n. 2, p. 296-312, 1997.
- LOPES, A. S. Manual internacional de fertilidade do solo. **Piracicaba: Associação Brasileira para a Pesquisa da Potassa e do Fosfato**, 1998.
- MAGALHÃES, O., FABRI, A., CALARGA, H. **Agricultura de precisão: gestão da lavoura**. [S.l.], 2012.
- MALAVOLTA, E. Avaliação do estado nutricional das plantas: princípios e aplicações/Eurípedes Malavolta, Godofredo Cesar Vitt, Sebastião Alberto de Oliveira. **Piracicaba: Potafos**, 1997.
- MARDIA, K., KENT, J., BIBBY, J. **Multivariate analysis**. Academic Press, London, 1992.
- MARDIA, K. V. Measures of multivariate skewness and kurtosis with applications. **Biometrika**, v. 57, n. 3, p. 519-530, 1970.
- MATHERON, G. Principles of geostatistics. **Economic geology**, v. 58, n. 8, p. 1246-1266, 1963.
- MERCANTE, E., URIBE-OPAZO, M. A., SOUZA, E. G. Variabilidade espacial e temporal da resistência mecânica do solo à penetração em áreas com e sem manejo químico localizado. **Revista Brasileira de Ciência do Solo**, v. 27, n. 6, p. 1149-1159, 2003.
- MINASNY, B., MCBRATNEY, A. B. The Matérn function as a general model for soil variograms. **Geoderma**, v. 128, n. 3-4, p. 192-207, 2005.
- MOLIN, J. P. **Agricultura de precisão: o gerenciamento da variabilidade**. 2001.
- MONTANARI, R., DE PASSOS, M., ANDREOTTI, M., DALCHIAVON, F. C., LOVERA, L. H.,

DE OLIVEIRA HONORATO, M. A. Aspectos da produtividade do feijão correlacionados com atributos físicos do solo sob elevado nível tecnológico de manejo. **Revista Brasileira de Ciência do Solo**, v. 34, n. 6, p. 1811-1822, 2010.

NADARAJAH, S., KOTZ, S. Estimation methods for the multivariate t distribution. **Acta Applicandae Mathematicae**, v. 102, n. 1, p. 99-118, 2008.

NORA, D. D., AMADO, T. J. C., NICOLOSO, R. D. S., GRUHN, E. M. Modern high-yielding maize, wheat and soybean cultivars in response to gypsum and lime application on no-till Oxisol. **Revista Brasileira de Ciência do Solo**, v. 41, 2017.

OLEA, R. A. **Optimum mapping techniques using regionalized variable theory**. Kansas Geological Survey, 1975.

OLIVEIRA, A. D., GARRIDO, W. E., ARAÚJO, J. D., LOURENÇO, S. **Métodos de pesquisa em fertilidade do solo**. Brasília: Embrapa-SEA, v. 19791, 1991.

OSORIO, F., GALEA, M. Statistical inference in multivariate analysis using the t-distribution. **Unpublished manuscript**, 2015.

OSORIO, F., PAULA, G. A., GALEA, M. Assessment of local influence in elliptical linear models with longitudinal structure. **Computational Statistics & Data Analysis**, v. 51, n. 9, p. 4354-4368, 2007.

PIEGORSCH, W. W., BAILER, A. J. **Analyzing environmental data**. John Wiley & Sons, 2005.

PINHEIRO, J. C., LIU, C., WU, Y. N. Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution. **Journal of Computational and Graphical Statistics**, v. 10, n. 2, p. 249-276, 2001.

RICHART, A., PICCIN, A. L., KAEFER, K. A. C., MORATELLI, G., KONOPATZKI, M. R. S., KAEFER, J. E., ECCO, M. Análise espaço-temporal de atributos químicos do solo influenciados pela aplicação de calcário de cloreto de potássio em taxa variável. **Scientia Agraria Paranaensis**, v. 15, n. 4, p. 391-400, 2016.

RICHETTI, J., URIBE-OPAZO, M. A., BASTIANI, F. D., JOHANN, J. A. Técnicas para detecção de pontos influentes em variáveis contínuas regionalizadas. **Engenharia Agrícola**, v. 36, n. 1, p. 152-165, 2016.

RUTKOWSKA, B., SZULC, W., HOCH, M., SPYCHAJ & FABISIĄK, E. Forms of AI in soil and soil solution in a long-term fertilizer application experiment. **Soil Use and Management**, v. 31, n. 1, p. 114-120, 2015.

SANCHEZ, R. B., MARQUES JÚNIOR, J., SOUZA, Z. M. D., PEREIRA, G. T., MARTINS FILHO, M. V. Variabilidade espacial de atributos do solo e de fatores de erosão em diferentes pedoformas. **Bragantia**, v. 68, n. 4, p. 1095-1103, 2009.

SCHEMMER, R. C., URIBE-OPAZO, M. A., GALEA, M., ASSUMPÇÃO, R. A. Spatial variability of soybean yield through a reparameterized t-Student model. **Engenharia Agrícola**, v. 37, n. 4, p. 760-770, 2017.

- STEIN, M. L. **Interpolation of Spatial Data**: Some Theory for Kriging. New York: Springer, 1999.
- SUTRADHAR, B. C. Score test for the covariance matrix of the elliptic t-distribution. **Journal of Multivariate Analysis**, v. 46, n. 1, p. 1-12, 1993.
- TAIZ, L., ZEIGER, E., MØLLER, I. M., MURPHY, A. **Fisiologia e desenvolvimento vegetal**. Artmed Editora, 2017.
- TOBLER, W. R. A computer movie simulating urban growth in the Detroit region. **Economic geography**, v. 46, n. sup1, p. 234-240, 1970.
- UNGER, P. W., STEWART, B. A., PARR, J. F., SINGH, R. P. Crop residue management and tillage methods for conserving soil and water in semi-arid regions. **Soil and Tillage Research**, v. 20, n. 2-4, p. 219-240, 1991.
- URIBE-OPAZO, M. A., BORSSOI, J. A., GALEA, M. Influence diagnostics in Gaussian spatial linear models. **Journal of Applied Statistics**, v. 39, n. 3, p. 615-630, 2012.
- VIEIRA, SIDNEY ROSE et al. Geoestatística em estudos de variabilidade espacial do solo. **Tópicos em ciência do solo. Viçosa**: Sociedade Brasileira de Ciência do Solo, v. 1, p. 1-53, 2000.
- WACKERNAGEL, H. **Multivariate Geostatistics**: an introduction with applications. Germany: Springer, 388p, 2003.
- WALLER, L. A., GOTWAY, C A. **Applied spatial statistics for public health data**. John Wiley & Sons, 2004.
- XIE, F. C., WEI, B. C., LIN, J. G. Case-deletion influence measures for the data from multivariate t distributions. **Journal of Applied Statistics**, v. 34, n. 8, p. 907-921, 2007.
- YAMAGUCHI, K. Robust Model and the EM Algorithm: ERROR DISTRIBUTIONS. In: **The EM Algorithm and Related Statistical Models**. CRC Press, p. 49-76, 2003.
- YAMAMOTO, J. K., LANDIM, P. M. B. **Geoestatística**: conceitos e aplicações. [S.l.]: Oficina de textos, 2015.
- ZHANG, H. Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. **Journal of the American Statistical Association**, v. 99, n. 465, p. 250-261, 2004.

4 ARTIGOS

4.1 ARTIGO 1: Diagnóstico de influência em um modelo espacial linear t-Student reparametrizado

Resumo:

Neste artigo ¹, considera-se o modelo espacial linear sob a distribuição t-Student multivariada com o segundo momento finito. Esta distribuição, que contém a distribuição normal, oferece uma estrutura mais flexível para a modelagem de dados espaciais. Utiliza-se a versão reparametrizada da distribuição t-Student multivariada, de forma que a matriz escala corresponda à matriz de covariância dos dados espaciais. O objetivo principal deste trabalho foi desenvolver medidas de influência para detectar a presença de observações influentes e de possíveis *outliers*, com base no deslocamento da verossimilhança e na estatística escore. Um modelo heteroscedástico foi considerado como um esquema de perturbação para a matriz de covariância do modelo espacial linear t-Student. Problemas de identificabilidade e aspectos de robustez dos estimadores de máxima verossimilhança também foram discutidos. Os resultados são ilustrados a partir de dados reais de produção de soja e de um estudo de simulação.

Palavras-chave: Estatística escore para *outlier*; Geoestatística; Identificabilidade.

4.2 Introdução

Este artigo foca nos modelos espaciais lineares no campo da geoestatística. Esses modelos são utilizados em uma gama de disciplinas, tais como agricultura, geologia, ciência do solo, hidrologia, climatologia, epidemiologia e planejamento espacial, entre muitas outras. Na literatura, os modelos espaciais lineares baseados na distribuição normal são extensos, como podem ser vistos nos livros por (CRESSIE, 1993; WALLER; GOTWAY, 2004; SCHABENBERGER; GOTWAY, 2017). No entanto, está bem sabido que a distribuição normal sofre de falta de robustez e as inferências são bastante sensíveis a *outliers*. Alternativas robustas para a distribuição normal foram consideradas na literatura. Por exemplo, inferência estatística e diagnósticos de influência para o modelo espacial linear, sob distribuições elípticas, foram discutidos por Assumpção et al. (2014), De Bastiani et al. (2015) e Fagundes et al. (2018). Desde o trabalho pioneiro de Lange et al. (1989), a distribuição t-Student tem provado ser muito útil para a modelagem robusta em muitos modelos de regressão.

A inferência estatística baseada na distribuição t-Student combina a simplicidade metodológica e computacional. A distribuição t-Student incorpora um parâmetro adicional, que permite modelar dados espaciais com alta curtose. Consideramos a reparametrização da distribuição t-Student com o segundo momento finito. Isso permite uma comparação mais

¹Este artigo já foi publicado na revista *Spatial Statistics*, cuja referência é Uribe-Opazo et al. (2021).

direta com a distribuição normal. Baseado em Fiorentini et al. (2003) e Galea et al. (2020), os quais utilizaram a reparametrização dos graus de liberdade sugeridos por Lange et al. (1989) para os dados do modelo financeiro, essa versão da distribuição t-Student é adotada aqui para modelagem de dados espaciais além de Schemmer et al (2017). Veja também (SCHEMMER et al., 2017). A também alguns dos trabalhos como os de Assumpção et al. (2014), De Bastiani et al. (2015) e nesse artigo serão discutimos o problema de identificabilidade do modelo espacial linear sob a distribuição t-Student, a robustez dos estimadores de máxima verossimilhança (EMV) e as medidas de influência.

Portanto, as contribuições principais deste artigo correspondem à especificação do modelo espacial linear t-Student utilizando uma reparametrização na distribuição t-Student multivariada, permitindo que a matriz escala corresponda efetivamente à matriz de covariância. Uma discussão sobre o modelo de identificabilidade através da função de influência mostra que os estimadores de máxima verossimilhança são estimadores robustos. Assim, foi proposto um modelo heteroscedástico como esquema de perturbação e discutiu-se a estatística score e a estatística do gradiente para detecção de *outlier* no modelo espacial linear t-Student.

O artigo é estruturado como segue. Seção 4.3 apresenta a distribuição t-Student reparametrizada multivariada, o modelo espacial linear t-Student, alguns problemas de identificabilidade e mostramos que os estimadores de máxima verossimilhança são estimadores robustos. A seção 4.4 apresenta diagnósticos de influência baseados no método de influência local e a estatística score para *outlier*. Um conjunto de dados reais da produtividade da soja coletados na safra 2016/2017 é analisado na Seção 4.5. Na Seção 4.6 é apresentado um estudo de simulação. Algumas conclusões observadas estão na seção 4.7. Mais detalhes dos resultados são apresentados no Apêndice 4.8.

4.3 Modelo espacial linear t-Student

4.3.1 Distribuição t-Student multivariada reparametrizada

Um vetor aleatório $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$ que segue uma distribuição t-Student multivariada reparametrizada com parâmetro de forma fixo η , em que $0 < \eta < \frac{1}{2}$, com matriz de covariância Σ de ordem $n \times n$, o vetor de média $E(\mathbf{Y}) = \boldsymbol{\mu}$ de ordem $n \times 1$, o qual é denotado como $\mathbf{Y} \sim T_n(\boldsymbol{\mu}, \Sigma, \eta)$, tem função densidade de probabilidade dada por

$$f(\mathbf{y}, \boldsymbol{\theta}) = K_n(\eta) |\Sigma|^{-\frac{1}{2}} (1 + c(\eta)\delta)^{-\frac{1}{2\eta}(1+n\eta)}, \quad (4.1)$$

em que $K_n(\eta) = \left(\frac{c(\eta)}{\pi}\right)^{\frac{n}{2}} \left[\Gamma\left(\frac{1+n\eta}{2\eta}\right)/\Gamma\left(\frac{1}{2\eta}\right)\right]$, com $\delta = (\mathbf{Y} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{Y} - \boldsymbol{\mu})$ sendo a distância de Mahalanobis, $c(\eta) = \eta/(1 - 2\eta)$ e $0 < \eta < 1/2$. Para algumas propriedades dessa versão de distribuição t-Student, veja Galea et al. (2020).

4.3.2 Modelo espacial linear t-Student

Para estudar a dependência espacial, considere um processo estocástico isotrópico de segunda ordem $\{\mathbf{Y}(\mathbf{s}_i), \mathbf{s}_i \in \mathcal{S}\}$, sendo $\mathcal{S} \subset \mathbb{R}^2$ e \mathbb{R}^2 um espaço Euclidiano bidimensional. Seja $\mathbf{Y} = (Y(\mathbf{s}_1), \dots, Y(\mathbf{s}_n))^\top$ um vetor de resposta correspondente às localizações \mathbf{s}_i com $i = 1, \dots, n$, com ordem $n \times 1$, seguindo uma distribuição t-Student n-multivariada reparametrizada, cada elemento $Y(\mathbf{s}_i)$ pode ser escrito como $Y(\mathbf{s}_i) = \mu(\mathbf{s}_i) + \epsilon(\mathbf{s}_i)$, $i = 1, \dots, n$. Ambos os termos determinístico $\mu(\mathbf{s}_i)$ e estocástico $\epsilon(\mathbf{s}_i)$ dependem da localização espacial em que $Y(\mathbf{s}_i)$ é observado. Assumi-se que os erros aleatórios $\epsilon(\mathbf{s}_i)$ têm esperança igual a zero, ou seja, $\mathbb{E}[\epsilon(\mathbf{s}_i)] = 0$, e a variação entre os pontos é determinada por alguma função de covariância $\text{Cov}[\epsilon(\mathbf{s}_i), \epsilon(\mathbf{s}_j)] = \text{Cov}[Y(\mathbf{s}_i), Y(\mathbf{s}_j)] = C(\mathbf{s}_i, \mathbf{s}_j) = \sigma_{ij}$, para $i, j = 1, \dots, n$. Suponha que, para alguma função conhecida de \mathbf{s}_i , $x_1(\mathbf{s}_i), \dots, x_p(\mathbf{s}_i)$ a média do processo estocástico é

$$\mu(\mathbf{s}_i) = \sum_{j=1}^p x_j(\mathbf{s}_i) \beta_j,$$

e β_1, \dots, β_p são parâmetros desconhecidos a serem estimados. Na notação de matriz, o modelo espacial linear t-Student é dado por

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (4.2)$$

com $\mathbf{Y} \sim T_n(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}, \eta)$, \mathbf{X} é uma matriz de posto completo de ordem $n \times p$, com a i -ésima linha um vetor $n \times p$, definido por $\mathbf{x}_i^\top = (x_{i1}, \dots, x_{ip})$ com as variáveis explicativas na localização \mathbf{s}_i , $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ é o vetor de parâmetros desconhecidos a serem estimados de ordem $p \times 1$ e $\boldsymbol{\epsilon} = (\epsilon(\mathbf{s}_1), \dots, \epsilon(\mathbf{s}_n))^\top$ são componentes correlacionados da variância, $\boldsymbol{\epsilon} \sim T_n(\mathbf{0}, \boldsymbol{\Sigma}, \eta)$.

Para modelagem espacial da família Matérn, Matheron (1963), Stein (1999), Minasny e McBratney (2005) apresentam opções das funções de covariância disponível na literatura. Para esta família, os elementos da matriz de covariância, $\boldsymbol{\Sigma} = [(\sigma_{ij})]$, são dados por

$$\sigma_{ij} = C(h_{ij}) = \tau_1 \xi_{ij} + \frac{\tau_2 (h_{ij}/\tau_3)^\kappa}{2^{\kappa-1} \Gamma(\kappa)} K_\kappa(h_{ij}/\tau_3), \quad h_{ij} \geq 0, \quad (4.3)$$

em que ξ_{ij} , $i, j = 1, \dots, n$, é o delta de Kronecker, $\tau_1, \tau_2 \geq 0$, $\tau_3 > 0$, e $\kappa > 0$ são parâmetros e K_κ é a função de Bessel modificada de terceiro tipo de ordem κ .

4.3.3 Identificabilidade do modelo

A identificabilidade de um modelo estatístico é essencial, pois se o modelo não for identificável, não existem estimadores consistentes para o vetor de parâmetros. A identificabilidade garante a unicidade da distribuição em função do parâmetro. De acordo com Demidenko (2013), na regressão do modelo uma condição necessária para a identificabilidade é a unicidade dos primeiros dois momentos como funções dos parâmetros do modelo. Nesta seção, alguns detalhes sobre a identificabilidade do modelo espacial linear t-Student são fornecidos, considerando as questões relacionadas aos vetores $\boldsymbol{\beta}$ e $\boldsymbol{\tau}$. A distribuição multivariada t-Student, definida na Equação (4.1), tem momentos finitos de primeira e segunda ordem, que não dependem do parâmetro de forma, η , e para qualquer valor de $\eta \in (0, 1/2)$,

a distribuição é definida por μ e Σ . Assim, levando-se em consideração que a distribuição t-Student multivariada é totalmente especificada pelo vetor de média e da matriz de covariância, as condições de identificabilidade para o modelo espacial linear t-Student, é dada na Equação (4.2), sendo uma extensão natural do modelo espacial linear gaussiano, quando η é conhecido. Então, seguindo os resultados apresentados em Demidenko (2013), na Seção 3.2, sobre identificabilidade no modelo de regressão normal, segue-se a identificabilidade do modelo espacial linear t-Student, o qual pode ser apresentado.

Seja o modelo de regressão definido por $Y \sim N_n(\mathbf{X}\beta, \Sigma)$, com β um vetor de parâmetros de ordem $p \times 1$ e $\Sigma = \Sigma(\tau)$ uma matriz de covariância de ordem $n \times n$, dependendo do parâmetro vetorial τ com ordem $q \times 1$. Demidenko (2013) mostrou que o modelo de regressão é identificável se e somente se $\mathbf{X}\beta_1 = \mathbf{X}\beta_2$ e $\Sigma(\tau_a) = \Sigma(\tau_b)$ implica que $\beta_1 = \beta_2$, $\tau_a = \tau_b$. Em outras palavras, o modelo de regressão é identificável se e somente se a matriz \mathbf{X} tem posto completo e se Σ é especificado exclusivamente por τ . Assim, com essas restrições nos parâmetros, o modelo espacial linear t-Student dado na Equação (4.2) é identificável.

Sabe-se que nem todos os parâmetros da função de covariância, dados na Equação (4.3), são consistentemente estimáveis sob a estrutura assintótica de domínio fixo. Na verdade, os parâmetros τ_2 e τ_3 não podem ser estimados de forma consistente quando o processo subjacente é observado em uma região limitada de \mathbb{R}^d por $d \leq 3$ (ZHANG, 2004; GENTON; ZHANG, 2012; KAUFMAN; SHABY, 2013). Contudo, Zhang (2004) também mostrou que se τ_3 é fixado em um valor arbitrário, então o estimador de máximo verossimilhança para $\tau^* = \tau_2/\tau_3^{2\kappa}$ é consistente. O parâmetro τ^* é chamado de parâmetro microergódico (STEIN, 1999).

Kaufman e Shaby (2013) estenderam os resultados de Zhang (2004) e mostraram que o estimador de máximo verossimilhança para τ^* também é consistente quando os parâmetros τ_2 e τ_3 são estimados em conjunto. Genton e Zhang (2012) mostraram que esses mesmos problemas de estimativa permanecem para campos aleatórios com contornos elípticos.

Para remover os problemas de identificabilidade, assume-se que o parâmetro τ_3 é fixado a priori Zhang (2004), Zhang e Zimmerman (2007), Zhang e El-Shaarawi (2010), e uma reparameterização alternativa da função de covariância Matérn sugerida por (STEIN, 1999) é usada,

$$C(h_{ij}) = \phi_1 \xi_{ij} + \frac{\phi_2}{2^{\kappa-1} \Gamma(\kappa)} (h_{ij} \tau_3)^\kappa K_\kappa(h_{ij}/\tau_3), \quad h_{ij} \geq 0,$$

$i, j = 1, \dots, n$, com $\phi_1 = \tau_1$ e $\phi_2 = \tau_2/\tau_3^{2\kappa}$. Se τ_3 é fixo a priori, a matriz de covariância tem uma estrutura linear, Anderson (1973), $\Sigma = \phi_1 \mathbf{I}_n + \phi_2 \mathbf{R}$, em que agora a matriz \mathbf{R} com ordem $n \times n$ tem elementos r_{ij} dados por:

$$r_{ij} = \begin{cases} 1, & \text{se } i = j, \\ \frac{1}{2^{\kappa-1} \Gamma(\kappa)} (h_{ij} \tau_3)^\kappa K_\kappa(h_{ij}/\tau_3), & h_{ij} > 0, \end{cases}$$

para $i, j = 1, \dots, n$, e para estimar ϕ_1 e ϕ_2 , o algoritmo proposto por Anderson (1973) pode ser usado. O vetor de parâmetro $(p+2)$ -dimensional agora é $\theta = (\beta^\top, \phi^\top)^\top$, com $\phi^\top = (\phi_1, \phi_2)$.

4.3.4 Estimativa de máxima verossimilhança

Para estimar os parâmetros do modelo, adaptamos um algoritmo proposto por Anderson (1973). A log-verossimilhança para o modelo espacial linear t-Student é dado por

$$\mathcal{L}(\boldsymbol{\theta}) = \log(K_n(\eta)) - \frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2\eta} (1 + n\eta) \log(1 + c(\eta)\delta), \quad (4.4)$$

com $\log(K_n(\eta)) = \frac{n}{2} \log\left(\frac{c(\eta)}{\pi}\right) + \log \Gamma\left(\frac{1+n\eta}{2\eta}\right) - \log \Gamma\left(\frac{1}{2\eta}\right)$, $\delta = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$ e $c(\eta) = \eta/(1 - 2\eta)$, $0 < \eta < \frac{1}{2}$. Conforme observado por Zellner (1976), a função log-verossimilhança (4.4) é uma função decrescente de η , e então não pode ser estimado por máximo verossimilhança. Veja também De Bastiani et al. (2015).

As funções escores para o modelo espacial linear t-Student são fornecidas por $U_\theta(\boldsymbol{\theta}) = (U_\beta^\top, U_\phi^\top)^\top$, em que

$$U_\beta = \partial \mathcal{L}(\boldsymbol{\theta}) / \partial \boldsymbol{\beta} = w(\delta) \mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \quad \text{and} \quad U_\phi = \partial \mathcal{L}(\boldsymbol{\theta}) / \partial \phi,$$

com o j -ésimo elemento de U_ϕ , dado por $U_{\phi_j} = \partial \mathcal{L}(\boldsymbol{\theta}) / \partial \phi_j = -\frac{1}{2} \text{tr}(\boldsymbol{\Sigma}^{-1} (\partial \boldsymbol{\Sigma} / \partial \phi_j)) + \frac{1}{2} w(\delta) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} (\partial \boldsymbol{\Sigma} / \partial \phi_j) \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}$, para $j = 1, 2$, $w(\delta) = \left(\frac{1+n\eta}{\eta}\right) \left(\frac{c(\eta)}{1+c(\eta)\delta}\right)$ e $\psi(x)$ é a função Digama. Nesse artigo, $\boldsymbol{\Sigma} = \phi_1 \mathbf{I}_n + \phi_2 \mathbf{R}$, então $\partial \boldsymbol{\Sigma} / \partial \phi_1 = \mathbf{I}_n$ e $\partial \boldsymbol{\Sigma} / \partial \phi_2 = \mathbf{R}$. Dado $\boldsymbol{\Sigma}$, a função log-verossimilhança (4.4) é maximizada em

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Y}, \quad (4.5)$$

e de $U_\phi = \mathbf{0}$ tem-se que,

$$i) \phi_1 \text{tr}(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1}) + \phi_2 \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1}) = w(\delta) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon},$$

$$ii) \phi_1 \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1}) + \phi_2 \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \mathbf{R}) = w(\delta) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}.$$

As equações $i)$ e $ii)$ podem ser escritas como $\mathbf{A}\boldsymbol{\phi} = \mathbf{b}$, veja Anderson (1973), sendo

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad \text{e} \quad \boldsymbol{\phi} = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \quad (4.6)$$

com $a_{11} = \text{tr}(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1})$, $a_{12} = \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1})$, $a_{22} = \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \mathbf{R})$, $b_1 = w(\delta) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}$ e $b_2 = w(\delta) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}$. Isso sugere o seguinte processo iterativo para estimar os vetores de parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\phi}$.

Etapa 1: Dado um estimador na k -iteração $\hat{\boldsymbol{\theta}}^{(k)} = (\hat{\boldsymbol{\beta}}^{(k)}, \hat{\boldsymbol{\phi}}^{(k)})$, calcule $w(\delta^{(k)})$.

Etapa 2: Usando as equações (4.5) e (4.6), atualizar as estimativas para obter $\hat{\boldsymbol{\theta}}^{(k+1)} = (\hat{\boldsymbol{\beta}}^{(k+1)}, \hat{\boldsymbol{\phi}}^{(k+1)})$ para $k = 0, 1, \dots$

Se $\|\hat{\boldsymbol{\theta}}^{(k+1)} - \hat{\boldsymbol{\theta}}^{(k)}\| < \varepsilon$ parar e definir $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}^{(k+1)}$ e $\hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\phi}}^{(k+1)}$. Caso contrário, $k = k + 1$ e volte para a Etapa 1. O valor de ε é escolhido para alcançar a convergência.

O erro padrão assintótico pode ser calculado invertendo a matriz de informação esperada, $\mathbf{F}_\theta = \text{diag}(F_\beta, F_\phi)$. As expressões explícitas para essas matrizes podem ser encontradas em De Bastiani et al. (2015).

4.3.5 Função de influência

Na estatística robusta, a função de influência mede a sensibilidade de um estimador $\hat{\theta}$ a pequenas quantidades de contaminação na distribuição. A função de influência de $\hat{\theta}(F)$, sendo que F é uma função de distribuição, definida por Hampel et al. (1986) com

$$\text{IF}(\mathbf{z}, \hat{\theta}) = \lim_{\xi \rightarrow 0} \frac{\hat{\theta}(F_\xi) - \hat{\theta}(F)}{\xi} = \left. \frac{\partial \hat{\theta}(F_\xi)}{\partial \xi} \right|_{\xi=0},$$

em que $F_\xi = (1 - \xi)F + \xi\zeta_{\mathbf{z}}$ e $\zeta_{\mathbf{z}}$ tem sua variação em \mathbf{z} e $0 \leq \xi \leq 1$. Em que $\text{IF}(\mathbf{z}, \hat{\theta})$ reflete a influência em $\hat{\theta}$ por adicionar uma observação em \mathbf{z} . Em outras palavras, é uma medida de influência de uma contaminação infinitesimal no ponto \mathbf{z} no estimador $\hat{\theta}$, é uma ferramenta útil para avaliar as propriedades de robustez de um estimador $\hat{\theta}$. Se a função de influência for limitada, $\hat{\theta}$ é dito B-robusto. Para mais detalhes veja Hampel et al.(1986).

A função de influência do estimador máximo verossimilhança Hampel et al. (1986) é dado por,

$$\text{IF}(\mathbf{Y}, \hat{\theta}) = \mathbf{F}_\theta^{-1}(\hat{\theta})U_\theta(\hat{\theta}),$$

com $U_\theta(\hat{\theta}) = (U_\beta^\top, U_\phi^\top)^\top$, com

$$\begin{aligned} U_\beta &= w(\delta)\mathbf{X}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}, \\ U_\phi &= \frac{1}{2} \frac{\partial \text{vec}^\top(\boldsymbol{\Sigma})}{\partial \phi} \{w(\delta) \text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1}) - \text{vec}(\boldsymbol{\Sigma}^{-1})\}, \end{aligned}$$

e $\boldsymbol{\epsilon} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}$. Este é um resultado importante, portanto, basta que a função escore U_θ seja limitada para que $\hat{\theta}$ seja um estimador robusto. Pode-se observar que, quando $\delta \rightarrow \infty$, a função escore para β tende a zero e a função escore para ϕ torna-se limitada. A Figura 6 mostra o comportamento das funções escores para uma variável aleatória $T(0; 1; 0, 25)$. Neste caso, a estimativa de máximo verossimilhança dos parâmetros do modelo espacial linear t-Student, dado na Equação (4.2) têm função de influência limitada, logo, os estimadores são robustos. Como é bem sabido, a função de influência do estimador de máximo verossimilhança no modelo espacial normal é ilimitada e, portanto, os estimadores de máximo verossimilhança não são robustos.

Dois critérios foram aplicados para determinar o valor de η : validação cruzada (VC) e o traço da covariância assintótica da média $Tr = Tr(\eta)$, e um QQ -plot das distâncias transformadas pode ser usado para avaliar o ajuste do modelo espacial linear t-Student. Para mais detalhes veja De Bastiani et al.(2015).

4.4 Diagnóstico e análise de influência

Existem várias técnicas para avaliar a influência das perturbações em um conjunto de dados e no modelo dos parâmetros estimados. O estudo de *outliers* e detecção de observações influentes é uma etapa importante na análise de dados espaciais.

Para a obtenção de referências sobre os artigos na literatura no que tange ao diagnóstico

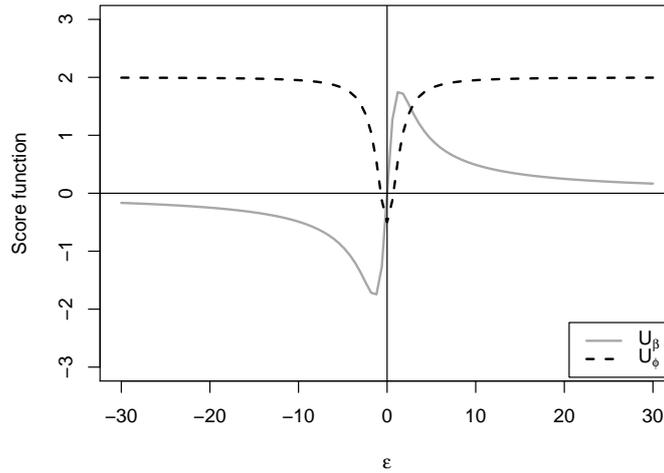


Figura 6 Funções escores para uma variável aleatória $T(0; 1; 0, 25)$ versus ϵ
 Fonte: Uribe-Opazo et al. (2021)

de influência em modelos espaciais lineares, consulte De Bastiani et al. (2018).

4.4.1 Estatística escore para *outlier*

Esta seção apresenta a estatística escore e a estatística de gradiente para detectar *outlier* sob o modelo espacial linear t-Student. Considera-se o modelo de desvio médio dado por

$$\mathbf{Y} = \mathbf{d}\alpha + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} = \mathbf{W}\boldsymbol{\gamma} + \boldsymbol{\epsilon}, \quad (4.7)$$

com $\mathbf{W} = (\mathbf{d}, \mathbf{X})$, $\boldsymbol{\gamma} = (\alpha, \boldsymbol{\beta}^\top)^\top$, \mathbf{d} um vetor unitário tendo 1 (um) na i -ésima posição e zero nas demais posições e $\mathbf{Y} \sim T_n(\mathbf{W}\boldsymbol{\beta}, \boldsymbol{\Sigma}, \eta)$. Nesse caso, o interesse é testar a hipótese

$$H_0 : \alpha = 0 \text{ versus } H_1 : \alpha \neq 0.$$

Se H_0 for rejeitado, o i -ésimo caso pode ser um *outlier*. A respectiva log-verossimilhança é dada por,

$$\mathcal{L}(\boldsymbol{\theta}) = \log K_n(\eta) - \frac{1}{2\eta}(1 + n\eta) \log(1 + c(\eta)\delta),$$

com $\delta = (\mathbf{Y} - \mathbf{W}\boldsymbol{\gamma})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{W}\boldsymbol{\gamma})$ e agora $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top, \phi^\top)^\top$. A função escore para α e $\boldsymbol{\beta}$ são fornecidas por,

$$\begin{aligned} U_\alpha &= w(\delta)\mathbf{d}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon} = w(\delta)\mathbf{d}^\top \boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \alpha\mathbf{d} - \mathbf{X}\boldsymbol{\beta}) \mathbf{e} \\ U_\beta &= w(\delta)\mathbf{W}^\top \boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \alpha\mathbf{d} - \mathbf{X}\boldsymbol{\beta}), \end{aligned}$$

respectivamente, sendo $w(\delta) = \left(\frac{1+n\eta}{\eta}\right) \left(\frac{c(\eta)}{q}\right) = \frac{1+n\eta}{1-2\eta+\eta\delta}$ e $q = 1 + c(\eta)\delta$.

Os estimadores de máxima verossimilhança de α e β são dados por

$$\begin{aligned}\hat{\alpha} &= \mathbf{d}^\top \Sigma^{-1}(\mathbf{I}_n - \mathbf{G})\mathbf{Y} / \mathbf{d}^\top \Sigma^{-1}(\mathbf{I}_n - \mathbf{G})\mathbf{d} \text{ e} \\ \hat{\beta} &= (\mathbf{X}^\top \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \Sigma^{-1}(\mathbf{Y} - \hat{\alpha}\mathbf{d}),\end{aligned}$$

respectivamente, em que $\mathbf{G} = \mathbf{X}(\mathbf{X}^\top \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \Sigma^{-1}$.

A matriz de informação de Fisher para o modelo de desvio médio (4.7) é dada por

$$\mathbf{F}_M = \begin{pmatrix} F_\gamma & \mathbf{0} \\ \mathbf{0} & F_\phi \end{pmatrix},$$

sendo

$$F_\gamma = c_1(\eta)(1 - 2\eta)^{-1} \begin{pmatrix} \mathbf{d}^\top \Sigma^{-1} \mathbf{d} & \mathbf{d}^\top \Sigma^{-1} \mathbf{X} \\ \mathbf{X}^\top \Sigma^{-1} \mathbf{d} & \mathbf{X}^\top \Sigma^{-1} \mathbf{X} \end{pmatrix}.$$

Seja $\hat{\theta}$ e $\tilde{\theta}$ o estimador de máxima verossimilhança de θ no modelo de desvio médio (4.7) e no modelo espacial linear t-Student, (4.2), respectivamente. De acordo com Boos e Stefanski (2013), a estatística escore para a hipótese H_0 é dada por

$$Sc_i = U_\theta^\top(\tilde{\theta}) \tilde{\mathbf{F}}_M^{-1} U_\theta(\tilde{\theta}) = \frac{w^2(\delta)}{c_1(\eta)(1 - 2\eta)^{-1}} \frac{b_i^2}{q_{ii}},$$

$i = 1, \dots, n$ sendo $\mathbf{Q} = \Sigma^{-1}(\mathbf{I}_n - \mathbf{G})$, $\mathbf{e} = \mathbf{Y} - \mathbf{X}\tilde{\beta}$ e $\mathbf{b} = \Sigma^{-1}\mathbf{e}$, com $\tilde{\beta} = (\mathbf{X}^\top \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \Sigma^{-1} \mathbf{Y}$, com q_{ij} o elemento (i, j) da matriz \mathbf{Q} e b_i elemento i -ésimo de \mathbf{b} , para $i, j = 1, \dots, n$.

De acordo com Terrell (2002) e Lemonte (2016), a estatística do gradiente para a hipótese H_0 é dado por

$$Gr_i = U_\theta^\top(\tilde{\theta})(\hat{\theta} - \tilde{\theta}) = w(\delta) \frac{b_i^2}{q_{ii}},$$

$i = 1, \dots, n$. Observa-se que $Sc_i = (w(\delta)(1 - 2\eta)/c_1(\eta))Gr_i$, $i = 1, \dots, n$. Para o caso normal ($c_1(\eta) = w(\delta) = 1$), $Sc_i = Gr_i$ para $i = 1, \dots, n$. Além disso, se $\Sigma = \sigma^2 \mathbf{I}_n$, então $\mathbf{G} = \mathbf{P} = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$, a bem conhecida matriz de Hat, e

$$Sc_i = Gr_i = \frac{e_i^2}{\hat{\sigma}^2(1 - p_{ii})} = r_i^2,$$

com r_i os resíduos t-Student, para $i = 1, \dots, n$. Assim, os gráficos de índice de Sc_i ou Gr_i , $i = 1, \dots, n$ podem revelar possíveis *outliers* nas observações.

4.4.2 Influência local

O método de influência local sugerido por Cook (1986) avalia o efeito simultâneo das observações no estimador de máxima verossimilhança sem removê-las do conjunto de dados. Seja ω um vetor de perturbação $r \times 1 \in \Omega$ subconjunto de \mathbb{R}^r , e o próximo modelo estatístico perturbado $\mathcal{M} = \{f(\mathbf{y}, \theta, \omega) : \omega \in \Omega\}$, sendo $f(\mathbf{y}, \theta, \omega)$ a função de densidade de \mathbf{Y} , (4.1), perturbado por ω e $\mathcal{L}(\theta, \omega) = \log f(\mathbf{y}, \theta, \omega)$ sua função de log-verossimilhança correspondente. Denota-se o vetor não perturbado, o vetor nulo, por ω_0 , suponha que $\mathcal{L}(\theta, \omega_0) = \mathcal{L}(\theta)$ dado

em (4.4). A influência da perturbação ω no estimador de máximo verossimilhança pode ser avaliada pelo deslocamento da verossimilhança dado por $LD(\omega) = 2\{\mathcal{L}(\hat{\theta}) - \mathcal{L}(\hat{\theta}_\omega)\}$, em que $\hat{\theta}$ é o estimador de máximo verossimilhança de $\theta = (\beta^\top, \phi^\top)^\top$ no modelo postulado, com $\beta = (\beta_1, \dots, \beta_p)^\top$, $\phi = (\phi_1, \phi_2)^\top$ e $\hat{\theta}_\omega$ é o estimador de máximo verossimilhança de θ no modelo perturbado por ω . Cook (1986) propôs estudar o comportamento local de $LD(\omega)$ em torno de ω_0 e mostrou que a curvatura normal C_l de $LD(\omega)$ em ω_0 na direção de algum vetor unitário l , é dado por $C_l = C_l(\theta) = 2|l^\top \Delta^\top \mathbf{L}^{-1} \Delta l|$, com $\|l\| = 1$, sendo \mathbf{L} a matriz de informação observada (dada no Apêndice 4.8), avaliada em $\theta = \hat{\theta}$ e Δ é uma matriz $(p+2) \times r$ dada por $\Delta = (\Delta_\beta^\top, \Delta_\phi^\top)^\top$, em que $\Delta_\beta = \partial^2 \mathcal{L}(\theta, \omega) / \partial \beta \partial \omega^\top$ e $\Delta_\phi = \partial^2 \mathcal{L}(\theta, \omega) / \partial \phi \partial \omega^\top$, avaliada em $\theta = \hat{\theta}$ e em $\omega = \omega_0$.

O gráfico dos elementos $|l_{max}|$ versus i (ordem dos dados) pode revelar que tipo de perturbação tem mais influência sobre $LD(\omega)$, na vizinhança de ω_0 , (COOK, 1986). Considerando-se $C_i = 2|f_{ii}|$, em que f_{ii} forma a diagonal principal da matriz $\mathbf{B} = \Delta^\top \mathbf{L}^{-1} \Delta$, pode ser usado o gráfico do índice de C_i para avaliar a presença de observações influentes.

Uma vez que C_l não é invariante sob mudança uniforme de escala, Poon e Poon (1999) propuseram a curvatura normal conformal $B_l = C_l / \text{tr}(2\mathbf{B})$, veja (ZHU; LEE, 2001). Uma propriedade interessante da curvatura conformal é que para qualquer direção unitária l , segue que $0 \leq B_l \leq 1$. Isso permite, por exemplo, a comparação de curvaturas entre diferentes modelos elípticos. Denotado por $B_i = 2|f_{ii}| / \text{tr}(2\mathbf{B})$ a curvatura conformal na direção da unitária com a i -ésima entrada 1 e todas as outras entradas 0. De acordo com Zhu e Lee (2001), a i -ésima observação é potencialmente influente se $B_i > \bar{B} + 2\text{sd}(B)$, em que $\bar{B} = \sum_{i=1}^r B_i / r$ e $\text{sd}(B)$ é o desvio padrão de B_1, \dots, B_r .

Neste artigo, um modelo heteroscedástico é considerado como um esquema de perturbação, ou seja, $\Sigma_\omega = \Sigma \odot \mathbf{J}_\omega$, em que \odot denota o produto Hadamard de duas matrizes, e \mathbf{J}_ω é uma matriz $n \times n$ com elementos diagonais, definido pelo vetor de perturbação $\omega = (\omega_1, \omega_2, \dots, \omega_n)^\top$ e um caso contrário. Nesse caso $\omega_0 = \mathbf{1}$. No Apêndice 4.10, usando a proposta de Zhu et al. (2007), é proposto um esquema de perturbação apropriado. Este resultado é novo para o modelo considerado neste artigo. As matrizes Δ_β e Δ_ϕ para este esquema de perturbação são apresentadas no Apêndice 4.9.

4.5 Aplicação

A aplicação em estudo refere-se a produtividade da soja associada às características do solo, importantes para a alimentação mundial. Os resultados apresentados neste trabalho podem ajudar a aumentar a produtividade da soja e reduzir os impactos ambientais por meio da correção do solo com a quantidade certa de insumos químicos. O conjunto de dados de produtividade da soja foi coletado em uma área agrícola de 127,16 ha, na cidade de Cascavel no estado do Paraná, Brasil, com solo classificado como Latossolo. Foram coletados a produtividade de soja (Prod) em $t \text{ ha}^{-1}$ e quatro teores químicos do solo durante a safra 2016/2017 com 74 observações cada. As variáveis explicativas foram: fósforo (P) medido em

$mg\ dm^{-3}$, potássio (K), cálcio (Ca) e magnésio (Mg) medido em $cmolc\ dm^{-3}$. Uma breve análise descritiva é apresentada na Tabela 2. O valor máximo da produtividade é $4,03\ t\ ha^{-1}$ que corresponde à observação #46 e é detectada como potencialmente influente pela estatística de escore (Figura 9). Observa-se que as informações fornecidas na Tabela 2 são uma primeira análise para compreender o conjunto dos dados, no entanto, este artigo considera um modelo linear t-Student com tendência espacial subjacente.

Tabela 2 Análise descritiva para o conjunto de dados reais

	Mín	Q_1	Mediana	Média	Q_3	Max
Prod	1,518	2,80	3,15	3,06	3,44	4,03
P	4,620	12,74	19,88	20,57	26,01	56,48
K	0,100	0,20	0,250	0,27	0,33	0,58
Ca	7,330	28,27	32,31	31,51	35,95	45,91
Mg	0,600	1,10	1,50	1,69	2,17	4,20

Prod: produtividade da soja, P: fósforo, K: potássio, Ca: cálcio,
Mg: magnésio, Mín: mínimo, Q_1 : primeiro quartil, Q_3 : terceiro quartil,
Max: máximo

Fonte: Uribe-Opazo et al. (2021)

A Figura 7 (à esquerda) mostra como os dados são distribuídos e a Figura 7 (à direita) mostra o boxplot da variável produtividade de soja e identifica os *outliers*. Os *outliers* da produtividade de soja correspondem às observações #68, #67, #65 e #69 (do menor para o maior), com valores $1,518, 1,661, 1,677$ e $1,797\ t\ ha^{-1}$, respectivamente. A observação #34 com valor $2,05$ é destacada na Figura 7 (à esquerda) porque aparece mais tarde na Figura 9 como potencialmente influente.

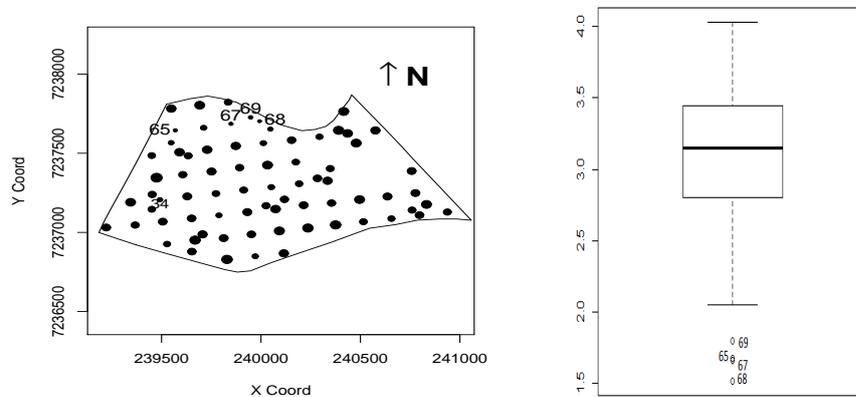


Figura 7 (à esquerda) *Postplot* dos dados coletados na safra 2016/2017 e (à direita) boxplot da produtividade da soja

Fonte: Uribe-Opazo et al. (2021)

As estimativas dos parâmetros para o modelo escolhido, com função de covariância Matérn com $\kappa = 0,5$, para a produtividade de soja considerando a função de covariância Matérn, e o respectivo erro padrão, entre parênteses, é apresentado na Tabela 3. O parâmetro τ_3 foi fixado em $0,4$, obtido em análises anteriores e considerando o método dos mínimos quadrados, bem conhecido em Geoestatística. Os parâmetros $\eta = 0,05$ e $\kappa = 0,5$ foram selecionados de acordo com os critérios de validação cruzada e traço da matriz de covariância assintótica.

A Figura 8 mostra o modelo de covariância ajustado com estimativas de parâmetros

Tabela 3 Estimativa dos parâmetros do modelo escolhido para a produtividade da soja considerando a função de covariância Matérn com $\kappa = 0,5$ e os respectivos erros padrão estão entre parênteses

Intercepto	P	K	Ca	Mg	efeito pepita	f(contribuição)
$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\phi}_1$	$\hat{\phi}_2$
2.8295	0.0039	0.2969	-0.0021	0.1017	0.0184	0.5016
(0.4673)	(0.0073)	(0.6988)	(0.0122)	(0.1235)	(0.2516)	(0.3709)

Fonte: Uribe-Opazo et al. (2021)

fornecidas na Tabela 3, em que podemos ver como a função de covariância diminui à medida que a distância entre as observações aumenta.

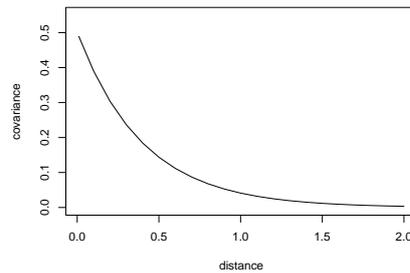


Figura 8 Modelo de covariância Matérn com ajuste de parâmetros $\kappa = 0.5$, $\hat{\phi}_1 = 0.0184$, $\hat{\phi}_2 = 0.5016$ e $\phi_3 = 0.4$

Fonte: Uribe-Opazo et al. (2021)

A Figura 9 (à esquerda) apresenta os gráficos do índice da estatística escore S_{c_i} para a hipótese nula. Os resultados para as estatísticas de gradiente são omitidos aqui, pois são semelhantes à estatística escore. Como esperado, as duas parcelas apresentaram o mesmo comportamento. As observações #34, #46, e #65 são destacadas como *outliers*, que correspondem aos valores de 2,05, 4,03 e 1,667 $t\ ha^{-1}$ de produtividade da soja, respectivamente. A partir do *postplot* dado na Figura 9, pode-se ver que a observação #34 tem um valor muito pequeno quando comparado aos valores de seus vizinhos. A Figura 9 (à direita) mostra o gráfico da influência local para a matriz de covariância perturbada com uma linha de corte, para a função de covariância Matérn com $\kappa = 0,5$ e considerando-se os graus de liberdade, para a distribuição t-Student, igual a $\eta = 0,05$. Os mesmos pontos destacados na Figura 9 (à esquerda) são indicados neste gráfico como observações potencialmente influentes.

A Tabela 4 mostra as estimativas de parâmetros para o modelo escolhido para dados de produtividade de soja, definidos sem as observações potencialmente influentes, #34, e sem #34, #46, e #65, considerando-se a função de covariância Matérn com $\kappa = 1,5$. Comparando-se as Tabelas 3 e 4, pode-se notar que as estimativas dos parâmetros para os β são muito semelhantes, e mantêm o mesmo sinal, o que significa a influência de cada variável explicativa na variável resposta ainda está na mesma direção. Porém, é evidente que os valores dos erros padrão assintóticos aumentam. Por outro lado, os valores dos parâmetros relacionados à função de covariância aumentam quando as observações são retiradas (e o respectivo erro padrão),

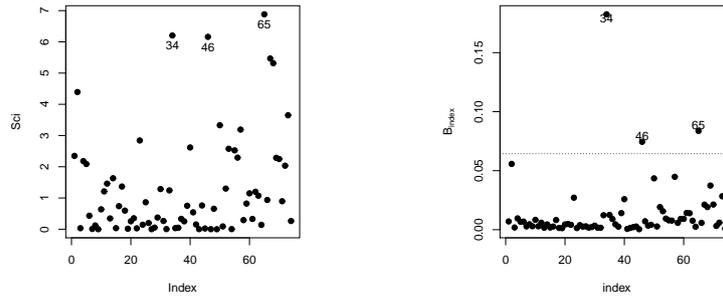


Figura 9 Gráfico do índice da estatística escore para a hipótese nula (à esquerda) e o gráfico de influência local para a matriz de covariância perturbada, com $\kappa = 0,5$ e $\eta = 0,05$ (à direita).

Fonte: Uribe-Opazo et al. (2021)

e isso mostra uma mudança na estrutura de dependência espacial entre as observações. Comparando-se as estimativas dos parâmetros entre as Tabelas 3 e 4, o valor do efeito pepita estimado ($\hat{\phi}_1$) aumenta em relação à ($\hat{\phi}_1 + \hat{\phi}_2$) quando a observação #34 é removida, e também quando as observações #34, #46 e #65 são removidas, indicando que a dependência espacial entre as observações diminui.

Tabela 4 Estimativas dos parâmetros para o modelo escolhido para dados de produtividade da soja definidas sem as observações potencialmente influentes (OPI), considerando a função de covariância Matérn com $\kappa = 1,5$. Os respectivos erros padrão estão entre parênteses

	Intercepto	P	K	Ca	Mg	efeito pepita	f(contribuição)
OPI	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\phi}_1$	$\hat{\phi}_2$
-#34	2,7881 (0,7299)	0,0020 (0,0120)	0,3686 (1,1937)	-0,0021 (0,0207)	0,1293 (0,1979)	0,2415 (1,6055)	1,0073 (1,6954)
-#(34, 46, 65)	2,8752 (0,7110)	0,0018 (0,0116)	0,1227 (1,1589)	-0,0008 (0,0199)	0,0987 (0,1920)	0,2293 (1,4933)	0,9204 (1,5751)

Fonte: Uribe-Opazo et al. (2021)

A Figura 10 apresenta os mapas da variável resposta produtividade de soja com todas as observações, sem a observação #34, e sem as observações #34, #46 e #65. Em todos os casos, as mesmas variáveis explicativas foram consideradas, ou seja, P , K , Ca e Mg , mas a função de covariância mudou de função Matérn com $\kappa = 0,5$ para a função Matérn $\kappa = 1,5$, quando uma observação foi removida do conjunto de dados. Para a construção dos mapas, os valores das variáveis explicativas foram preditos, e então esses valores também foram considerados para a obtenção dos valores preditos para a variável resposta.

O conjunto de dados sem a observação #34 e sem as observações #34, #46 e #65 foram analisados. O modelo escolhido para a estrutura de covariância é a função de covariância Matérn, mas com $\kappa = 1,5$ em vez de $\kappa = 0,5$. Para ambos os cenários, fica evidente a mudança nos valores previstos quando o conjunto de dados é analisado sem as observações potencialmente influentes.

A Figura 11 apresenta o desvio normal esperado versus a distância transformada considerando (a) todas as observações, (b) sem a observação #34, e (c) sem as observações

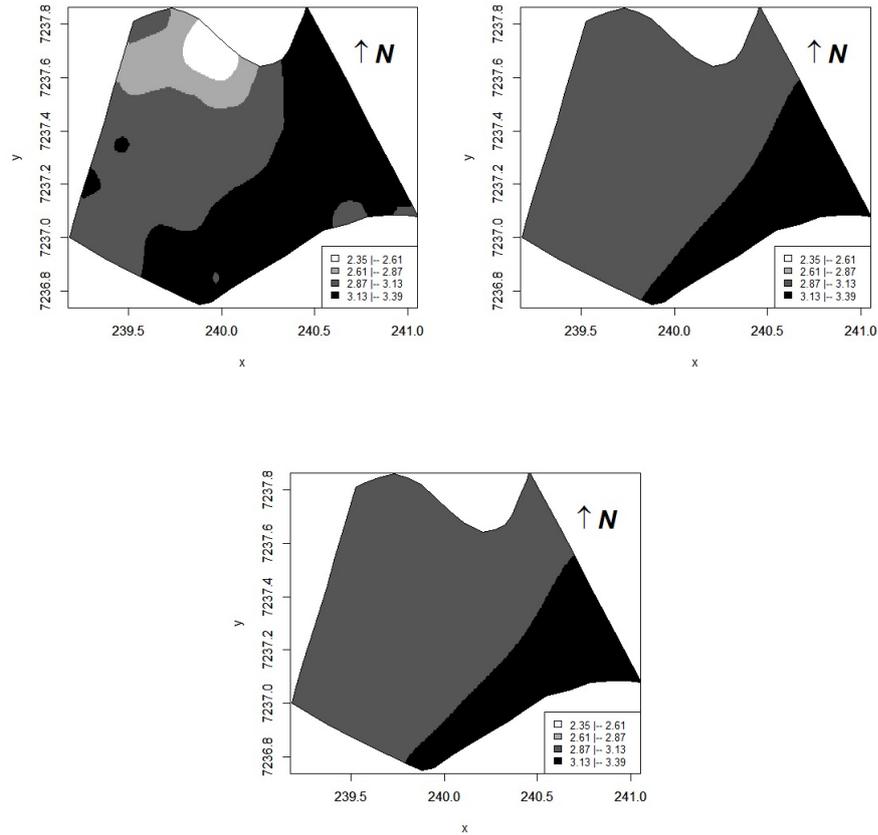


Figura 10 Mapas de produtividade da soja, considerando Matérn $\kappa = 0,5$ e $\eta = 0,05$ com todas as observações(a), Matérn $\kappa = 1,5$ e $\eta = 0,05$ sem a observação #34 (b), e Matérn $\kappa = 1,5$ e $\eta = 0,05$ sem as observações #34, #46 e #65 (c)

Fonte: Uribe-Opazo et al. (2021)

#34, #46 e #65. Pode-se observar que os ajustes do modelo espacial linear t-Student reparametrizado são razoáveis, e o ajuste com todas as observações apresenta a maioria dos pontos ao longo da reta.

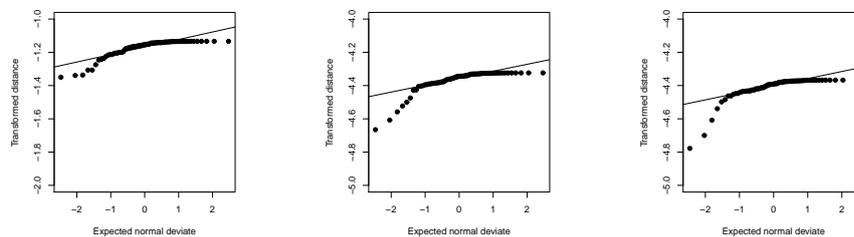


Figura 11 Desvio normal esperado versus distância transformada considerando (a) todas as observações, (b) sem observação #34, e (c) sem observações #34, #46 e #65.

Fonte: Uribe-Opazo et al. (2021)

4.6 Estudo de simulação

4.6.1 Avaliação das estimativas dos parâmetros para o modelo t-Student reparametrizado

O objetivo principal é avaliar as estimativas dos parâmetros do modelo t-Student reparametrizado multivariado. Além disso, esta simulação visa comparar a t-Student reparametrizada multivariada com a distribuição normal sob uma especificação incorreta do modelo. Os parâmetros considerados para a simulação foram baseados no conjunto de dados reais apresentado na Seção 4.16. Com o total de $n = 74$ observações com a localização sendo a mesma apresentada na Figura 7 (à esquerda). Os parâmetros considerados para o modelo simulado foram $\phi_1 = 0,1$, $\phi_2 = 0,5$, $\phi_3 = 0,4$, a função de covariância Matérn com $\kappa = 0,5$, $\mu = 0$ e $\eta = 0,45$.

De acordo com a Tabela 5, em 91,3% dos casos, t-Student multivariado reparametrizado com $\eta = 0,45$ foi selecionado. O caso em que $\eta = 0,01$ pode ser visto como aproximadamente a distribuição normal multivariada, como foi selecionado apenas 8,6%. Tanto o critério da validação cruzada (VC) quanto do traço da matriz de covariância assintótica (Tr) apresentam a mesma porcentagem de vezes para cada modelo.

Tabela 5 Porcentagens de vezes que o modelo t-Student reparametrizado com $\eta = 0,45$, $\eta = 0,05$ e $\eta = 0,01$ foram escolhidos usando os critérios de validação cruzada (VC) e traço da matriz de covariância assintótica (Tr) em 1000 dados simulados

Modelo ajustado	VC	Tr
t-Student $\eta = 0,01$	8,6	8,6
t-Student $\eta = 0,05$	0,1	0,1
t-Student $\eta = 0,45$	91,3	91,3

Fonte: Uribe-Opazo et al. (2021)

4.6.2 Medidas de influência

Os parâmetros considerados para o modelo simulado foram $\phi_1 = 0,1$, $\phi_2 = 0,5$, $\phi_3 = 0,4$, a função de covariância Matérn com $\kappa = 0,5$, $\mu = 0$ e $\eta = 0,45$. Em cada um dos 1000 dados simulados, o valor máximo foi perturbado, passando a ser o valor máximo mais o valor de 5. O modelo t-Student reparametrizado com $\eta = 0,45$, $\eta = 0,05$ e $\eta = 0,01$ foi ajustado.

A Tabela 6 mostra as porcentagens de vezes que as medidas de influência detectaram a observação perturbada em 1000 dados simulados, para cada um dos modelos t-Student reparameterizado multivariado ajustado, com $\eta = 0,45$, $\eta = 0,05$ e $\eta = 0,01$. Os resultados são muito semelhantes, com maior percentual para t-Student reparametrizado multivariado com $\eta = 0,01$.

Tabela 6 Porcentagem de vezes que as medidas de influência detectaram a observação perturbada em 1000 dados simulados. O modelo simulado é o t-Student reparametrizado com $\eta = 0,45$ e os modelos ajustados são o t-Student reparametrizado com $\eta = 0,45$, $\eta = 0,05$ e $\eta = 0,01$

Modelos	Estatística escore	influência local para a matriz de covariância
t-Student $\eta = 0,01$	87,9	85,5
t-Student $\eta = 0,05$	87,9	84,3
t-Student $\eta = 0,45$	87,9	84,0

Fonte: Uribe-Opazo et al. (2021)

4.7 Conclusões

Este trabalho se propôs a modelar a média da produtividade da soja coletada na região sul do Brasil considerando um modelo espacial linear t-Student reparametrizado, e também para verificar a presença (ou não) de observações influentes que podem alterar os resultados. Um esquema de perturbação apropriado para a matriz de covariância de um modelo espacial linear t-Student reparametrizado foi proposto, baseado no deslocamento de verossimilhança.

Foram obtidas expressões explícitas para as matrizes necessárias para implementar essas técnicas de diagnóstico usando a família Matérn de funções de covariância. Esse tipo de resultado é raro em estudos anteriores sobre o assunto. Além disso, foi abordado um estudo sobre problemas de identificabilidade.

A aplicação ao conjunto de dados reais mostrou que as observações detectadas pelo estudo de influência local afetaram a seleção da função de covariância, para compreender a variabilidade espacial e obter os valores previstos apresentados no mapa.

4.8 Apêndice: a matriz de informação observada para o modelo linear espacial t-Student

A função log-verossimilhança para o modelo t-Student reparameterizado é dada por:

$$\mathcal{L}(\boldsymbol{\theta}) = \log K_n(\eta) - \frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2\eta} (1 + n\eta) \log(1 + c(\eta)\delta),$$

sendo $K_n(\eta) = \left(\frac{c(\eta)}{\pi}\right)^{\frac{n}{2}} \left[\Gamma\left(\frac{1+n\eta}{2\eta}\right) / \Gamma\left(\frac{1}{2\eta}\right)\right]$, com $\delta = (\mathbf{Y} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \boldsymbol{\mu})$, e $c(\eta) = \eta / (1 - 2\eta)$ para $0 < \eta < \frac{1}{2}$. Portanto,

$$\begin{aligned} \frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} &= w(\delta) \mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}, \\ \frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \phi_1} &= -\frac{1}{2} \{tr(\boldsymbol{\Sigma}^{-1}) - w(\delta) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}\}, \\ \frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \phi_2} &= -\frac{1}{2} \{tr(\boldsymbol{\Sigma}^{-1} \mathbf{R}) - w(\delta) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}\}, \end{aligned}$$

com $w(\delta) = \frac{1+n\eta}{\eta} \frac{c(\eta)}{1+c(\eta)\delta}$. As derivadas de $w(\delta)$ são

$$\begin{aligned}\frac{\partial w(\delta)}{\partial \beta^\top} &= w_\beta(\delta) = 2w(\delta) \left(\frac{c(\eta)}{1+c(\eta)\delta} \right) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}, \\ \frac{\partial w(\delta)}{\partial \phi_1} &= w_{\phi_1}(\delta) = 2w(\delta) \left(\frac{c(\eta)}{1+c(\eta)\delta} \right) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}, \\ \frac{\partial w(\delta)}{\partial \phi_2} &= w_{\phi_2}(\delta) = 2w(\delta) \left(\frac{c(\eta)}{1+c(\eta)\delta} \right) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon},\end{aligned}$$

A matriz de informação observada é obtida a partir da segunda derivada da função log-verossimilhança, em relação à $\boldsymbol{\theta}$, é dada por

$$\mathbf{L} = - \begin{pmatrix} \mathbf{L}_{\beta\beta} & \mathbf{L}_{\beta\phi} \\ \mathbf{L}_{\beta\phi}^\top & \mathbf{L}_{\phi\phi} \end{pmatrix},$$

sendo

$$\begin{aligned}\mathbf{L}_{\beta\beta} &= w(\delta) \mathbf{X}^\top \left\{ 2 \left(\frac{c(\eta)}{1+c(\eta)\delta} \right) \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon} \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1} \right\} \mathbf{X}, \\ \mathbf{L}_{\beta\phi} &= \begin{pmatrix} \mathbf{L}_{\beta\phi_1} & \mathbf{L}_{\beta\phi_2} \end{pmatrix}, \\ \mathbf{L}_{\phi\phi} &= \begin{pmatrix} l_{11} & l_{12} \\ l_{12} & l_{22} \end{pmatrix},\end{aligned}$$

em que

$$\begin{aligned}\mathbf{L}_{\beta\phi_1} &= w(\delta) \mathbf{X}^\top \left\{ 2 \left(\frac{c(\eta)}{1+c(\eta)\delta} \right) (\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}) \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \right\} \boldsymbol{\epsilon}, \\ \mathbf{L}_{\beta\phi_2} &= w(\delta) \mathbf{X}^\top \left\{ 2 \left(\frac{c(\eta)}{1+c(\eta)\delta} \right) (\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}) \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \right\} \boldsymbol{\epsilon}, \\ l_{11} &= \frac{1}{2} \{ \text{tr}(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1}) - 2w(\delta) (\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}) + w_{\phi_1}(\delta) (\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}) \}, \\ l_{12} &= \frac{1}{2} \{ \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1}) - w(\delta) (\boldsymbol{\epsilon}^\top \{ \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \} \boldsymbol{\epsilon}) \\ &\quad + w_{\phi_2}(\delta) (\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}) \}, \\ l_{22} &= \frac{1}{2} \{ \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \mathbf{R}) - 2w(\delta) (\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}) + w_{\phi_2}(\delta) (\boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}) \}.\end{aligned}$$

4.9 Apêndice: matriz para o esquema de perturbação Δ da matriz de covariância

A função log-verossimilhança perturbada, considerando-se o esquema de perturbação da matriz de covariância, é dada por

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega}) = \log K_n(\eta) - \frac{1}{2} \log |\boldsymbol{\Sigma}_\omega| - \frac{1}{2\eta} (1+n\eta) \log(1+c(\eta)\delta_\omega), \quad (4.8)$$

sendo $\delta_\omega = \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}_\omega^{-1} \boldsymbol{\epsilon}$, $\boldsymbol{\Sigma}_\omega = \boldsymbol{\Sigma} \odot \mathbf{J}_\omega = \phi_1 \mathbf{D}(p(\boldsymbol{\omega})) + \phi_2 \mathbf{R}_\omega$, com $\mathbf{D}(p(\boldsymbol{\omega})) = \text{diag}(p_1(\boldsymbol{\omega}), p_2(\boldsymbol{\omega}), \dots, p_n(\boldsymbol{\omega}))$ e \mathbf{R}_ω é a matriz \mathbf{R} com a diagonal substituída por $p(\boldsymbol{\omega}) = (p_1(\boldsymbol{\omega}), \dots, p_n(\boldsymbol{\omega}))^\top = \mathbf{1}_n + \mathbf{G}_0^{-1/2} (\boldsymbol{\omega} - \mathbf{1}_n)$, em que $\mathbf{G}_0^{1/2}$ é definido no Apêndice D em (DE

BASTIANI et al., 2015). Os elementos são dados por

$$r_{\omega ij} = \begin{cases} p_i(\boldsymbol{\omega}), & \text{se } i = j \\ r_{ij}, & \text{se } i \neq j, \end{cases}$$

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

Lema 1. Note que $\boldsymbol{\Sigma}_\omega = \boldsymbol{\Sigma} \odot \mathbf{J}_\omega = \phi_1 \mathbf{D}(p(\boldsymbol{\omega})) + \phi_2 \mathbf{R}_\omega$, então

- (i) $\frac{\partial p(\boldsymbol{\omega})}{\partial \omega_i} = \mathbf{g}_i, i = 1, \dots, n,$
- (ii) $\frac{\partial \mathbf{D}(p(\boldsymbol{\omega}))}{\partial \omega_i} = \mathbf{D}(\mathbf{g}_i), i = 1, \dots, n,$
- (iii) $\frac{\partial \mathbf{R}_\omega}{\partial \omega_i} = \mathbf{D}(\mathbf{g}_i), i = 1, \dots, n,$
- (iv) $\frac{\partial \boldsymbol{\Sigma}_\omega}{\partial \omega_i} = (\phi_1 + \phi_2) \mathbf{D}(\mathbf{g}_i), i = 1, \dots, n,$
- (v) $\frac{\partial \boldsymbol{\Sigma}_\omega^{-1}}{\partial \omega_i} = -(\phi_1 + \phi_2) \boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(\mathbf{g}_i) \boldsymbol{\Sigma}_\omega^{-1}, i = 1, \dots, n,$
- (vi) $\frac{\partial \boldsymbol{\Sigma}_\omega^{-1}}{\partial \phi_1} = -\boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(p(\boldsymbol{\omega})) \boldsymbol{\Sigma}_\omega^{-1},$
- (vii) $\frac{\partial \boldsymbol{\Sigma}_\omega^{-1}}{\partial \phi_2} = -\boldsymbol{\Sigma}_\omega^{-1} \mathbf{R}_\omega \boldsymbol{\Sigma}_\omega^{-1},$

sendo $\mathbf{g}_i = \mathbf{G}_0^{-1/2} \mathbf{e}_{in}$, com $\mathbf{e}_{in}^\top = (0, \dots, 1, \dots, 0)$, vetor unitário de \mathbb{R}^n , para $i = 1, \dots, n$.

Da Equação (4.8),

$$\begin{aligned} \frac{\partial \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial \boldsymbol{\beta}} &= w(\delta_\omega) \mathbf{X}^\top \boldsymbol{\Sigma}_\omega^{-1} \boldsymbol{\epsilon}, \\ \frac{\partial \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial \phi_1} &= -\frac{1}{2} \left\{ \text{tr}(\boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(p(\boldsymbol{\omega}))) - w(\delta_\omega) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(p(\boldsymbol{\omega})) \boldsymbol{\Sigma}_\omega^{-1} \boldsymbol{\epsilon} \right\}, \\ \frac{\partial \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial \phi_2} &= -\frac{1}{2} \left\{ \text{tr}(\boldsymbol{\Sigma}_\omega^{-1} \mathbf{R}_\omega) - w(\delta_\omega) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}_\omega^{-1} \mathbf{R}_\omega \boldsymbol{\Sigma}_\omega^{-1} \boldsymbol{\epsilon} \right\}, \end{aligned} \quad (4.9)$$

sendo $w(\delta_\omega) = \frac{1 + n\eta}{\eta} \frac{c(\eta)}{1 + c(\eta)\delta_\omega}$. Então, $w_{id}(\delta_\omega) = \frac{\partial w(\delta_\omega)}{\partial \omega_i} = \left(\frac{c(\eta)}{1 + c(\eta)\delta_\omega} \right) w(\delta_\omega) (\boldsymbol{\epsilon}_i^\top \mathbf{C}_i \boldsymbol{\epsilon})$, com $\mathbf{C}_i = (\phi_1 + \phi_2) \boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(\mathbf{g}_i) \boldsymbol{\Sigma}_\omega^{-1}$, para $i = 1, \dots, n$.

As segundas derivadas de (4.9) em relação à ω_i são dadas por:

$$\begin{aligned} \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial \boldsymbol{\beta} \partial \omega_i} &= \mathbf{X}^\top \left\{ w_{id}(\delta_\omega) \boldsymbol{\Sigma}_\omega^{-1} - w(\delta_\omega) \mathbf{C}_i \right\} \boldsymbol{\epsilon}, \\ \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial \phi_1 \partial \omega_i} &= -\frac{1}{2} \left\{ \text{tr}(\boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(\mathbf{g}_i)) - (\phi_1 + \phi_2) \text{tr}(\boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(\mathbf{g}_i) \boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(p(\boldsymbol{\omega}))) \right. \\ &\quad \left. - w(\delta_\omega) \{ \boldsymbol{\epsilon}^\top (\boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(\mathbf{g}_i) \boldsymbol{\Sigma}_\omega^{-1}) \boldsymbol{\epsilon} - 2 \boldsymbol{\epsilon}^\top (\mathbf{C}_i \mathbf{D}(p(\boldsymbol{\omega})) \boldsymbol{\Sigma}_\omega^{-1}) \boldsymbol{\epsilon} \} \right. \\ &\quad \left. - w_{id}(\delta_\omega) \boldsymbol{\epsilon}^\top (\boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(p(\boldsymbol{\omega})) \boldsymbol{\Sigma}_\omega^{-1}) \boldsymbol{\epsilon} \right\}, \\ \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial \phi_2 \partial \omega_i} &= -\frac{1}{2} \left\{ \text{tr}(\boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(\mathbf{g}_i)) - (\phi_1 + \phi_2) \text{tr}(\boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(\mathbf{g}_i) \boldsymbol{\Sigma}_\omega^{-1} \mathbf{R}_\omega) \right. \\ &\quad \left. - w(\delta_\omega) \{ \boldsymbol{\epsilon}^\top (\boldsymbol{\Sigma}_\omega^{-1} \mathbf{D}(\mathbf{g}_i) \boldsymbol{\Sigma}_\omega^{-1}) \boldsymbol{\epsilon} - 2 \boldsymbol{\epsilon}^\top (\mathbf{C}_i \mathbf{R}_\omega \boldsymbol{\Sigma}_\omega^{-1}) \boldsymbol{\epsilon} \} - w_{id}(\delta_\omega) \boldsymbol{\epsilon}^\top \boldsymbol{\Sigma}_\omega^{-1} \mathbf{R}_\omega \boldsymbol{\Sigma}_\omega^{-1} \boldsymbol{\epsilon} \right\}, \end{aligned} \quad (4.10)$$

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

Avaliando as derivadas, a partir da Equação (4.10), em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e em $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, a matriz Δ assume a forma:

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_{\phi_1} \\ \Delta_{\phi_2} \end{pmatrix},$$

com $\Delta_\beta = (\Delta_{1\beta}, \dots, \Delta_{n\beta})$, $\Delta_{\phi_1} = (\Delta_{1\phi_1}, \dots, \Delta_{n\phi_1})$ e $\Delta_{\phi_2} = (\Delta_{1\phi_2}, \dots, \Delta_{n\phi_2})$, com

$$\begin{aligned} \Delta_{i\beta} &= \left. \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \boldsymbol{\beta} \partial \omega_i} \right|_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}, \boldsymbol{\omega} = \boldsymbol{\omega}_0} \\ \Delta_{i\phi_j} &= \left. \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \phi_j \partial \omega_i} \right|_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}, \boldsymbol{\omega} = \boldsymbol{\omega}_0}, \end{aligned}$$

para $i = 1, \dots, n$ e $j = 1, 2$.

4.10 Apêndice: a matriz de informação para o esquema de perturbação da matriz de covariância

Zhu et al. (2007) propuseram um método para selecionar os esquemas de perturbação em modelos estatísticos baseados na matriz de informações de Fisher com relação à $\boldsymbol{\omega}$, qual é $\mathbf{G}(\boldsymbol{\omega}) = (g_{kl}(\boldsymbol{\omega})) = \mathbb{E}_\omega(U_\omega U_\omega^\top)$, com

$$g_{kl}(\boldsymbol{\omega}) = \mathbb{E}_\omega \left\{ \frac{\partial \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial \omega_k} \frac{\partial \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial \omega_l} \right\}, \quad k, l = 1, \dots, r,$$

em que \mathbb{E}_ω denota a esperança com relação à densidade do modelo perturbado $f(\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\omega})$, e $U_\omega = \partial \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega}) / \partial \boldsymbol{\omega}$. Os elementos da diagonal $\{g_{kk}(\boldsymbol{\omega}), k = 1, \dots, r\}$ de $\mathbf{G}(\boldsymbol{\omega})$ são as variâncias das funções escores em relação aos elementos de $\boldsymbol{\omega}$, que indica a quantidade de perturbação do k -ésimo componente de $\boldsymbol{\omega}$. Os elementos fora da diagonal principal $\mathbf{G}(\boldsymbol{\omega})$ são as covariâncias das funções escores com respeito aos componentes de $\boldsymbol{\omega}$ e representa a associação linear entre diferentes componentes de $\boldsymbol{\omega}$. Observe que se $\mathbf{G}(\boldsymbol{\omega})$ é uma matriz diagonal, os componentes de $\boldsymbol{\omega}$ são ortogonais (COX; REID, 1987). A perturbação $\boldsymbol{\omega}$ é considerada apropriada se satisfaz

a condição de $\mathbf{G}(\omega_0) = \mathbf{G}_0 = c\mathbf{I}_r$, com $c > 0$.

Se $\mathbf{G}(\omega_0) \neq c\mathbf{I}_r$, podemos reparameterizar o modelo perturbado, \mathcal{M} , considerando-se a transformação de um para um $\omega = p(\tilde{\omega}) = \omega_0 + \mathbf{G}_0^{-1/2}(\tilde{\omega} - \omega_0)$, em que $\tilde{\omega}$ também tem dimensão r . Com efeito, a matriz de informações de Fisher com relação à $\tilde{\omega}$, é dada por $\mathbf{G}(\tilde{\omega}) = \mathbf{G}_0^{-1/2}\mathbf{G}(p(\tilde{\omega}))\mathbf{G}_0^{-1/2}$, de onde segue que $\mathbf{G}(\tilde{\omega})$ avaliado em ω_0 é igual a \mathbf{I}_r (BOOS; STEFANSKI, 2013). Sob esta nova parametrização, o modelo perturbado pode ser escrito como $\mathcal{M} = \{f(\mathbf{y}, \boldsymbol{\theta}, p(\tilde{\omega})) : \tilde{\omega} \in \tilde{\Omega}\}$, com $\tilde{\Omega} = \{\omega_0 + \mathbf{G}_0^{1/2}(\omega - \omega_0) : \omega \in \Omega\}$.

Para este artigo, inicialmente o seguinte esquema de perturbação foi considerado $\Sigma_\omega = \Sigma \odot \mathbf{J}_\omega$, em que \odot denota o produto Hadamard de duas matrizes, e \mathbf{J}_ω é uma matriz de ordem $n \times n$, cuja a diagonal é dada pelo vetor de perturbação $\omega = (\omega_1, \omega_2, \dots, \omega_n)$ e um caso contrário. Nesse caso, $\omega_0 = \mathbf{1}$.

O seguinte lema, adaptado de Sutradhar (1993) e Bolfarine e Galea (1996).

Lema 2. Se $\epsilon \sim \mathcal{T}_n(\mathbf{0}, \Sigma_\omega, \eta)$ e se $w(\delta_\omega) = \frac{(1+n\eta)}{\eta} \frac{c(\eta)}{1+c(\eta)\delta_\omega}$ com $\delta_\omega = \epsilon^\top \Sigma_\omega^{-1} \epsilon$, então

- (i) $\mathbb{E}(w(\delta_\omega)\epsilon^\top \mathbf{A}\epsilon) = \text{tr}(\Sigma_\omega \mathbf{A})$,
- (ii) $\mathbb{E}(w^2(\delta_\omega)(\epsilon^\top \mathbf{A}\epsilon)(\epsilon^\top \mathbf{B}\epsilon)) = \frac{(1+n\eta)}{(1+(n+2)\eta)} \{2 \text{tr}(\Sigma_\omega \mathbf{A} \Sigma_\omega \mathbf{B}) + \text{tr}(\Sigma_\omega \mathbf{A}) \text{tr}(\Sigma_\omega \mathbf{B})\}$.

Considerando-se $U_{i\omega} = \partial \mathcal{L}(\boldsymbol{\theta}, \omega) / \partial \omega_i$, $i = 1, \dots, n$. Para este esquema de perturbação inicial,

$$\begin{aligned} U_{i\omega} &= -\frac{1}{2}(\phi_1 + \phi_2) \{ \mathbf{e}_{in}^\top \Sigma_\omega^{-1} \mathbf{e}_{in} - w(\delta_\omega) \epsilon^\top \mathbf{A}_i \epsilon \}, \mathbf{e}, \\ U_{i\omega} U_{j\omega} &= \frac{1}{4}(\phi_1 + \phi_2)^2 \left\{ (\mathbf{e}_{in}^\top \Sigma_\omega^{-1} \mathbf{e}_{in})(\mathbf{e}_{jn}^\top \Sigma_\omega^{-1} \mathbf{e}_{jn}) - (\mathbf{e}_{in}^\top \Sigma_\omega^{-1} \mathbf{e}_{in}) w(\delta_\omega) \epsilon^\top \mathbf{A}_j \epsilon \right. \\ &\quad \left. - (\mathbf{e}_{jn}^\top \Sigma_\omega^{-1} \mathbf{e}_{jn}) w(\delta_\omega) \epsilon^\top \mathbf{A}_i \epsilon + w^2(\delta_\omega) (\epsilon^\top \mathbf{A}_i \epsilon) (\epsilon^\top \mathbf{A}_j \epsilon) \right\}, \end{aligned}$$

com $\mathbf{A}_i = \Sigma_\omega^{-1} \mathbf{e}_{in} \mathbf{e}_{in}^\top \Sigma_\omega^{-1}$ para $i, j = 1, \dots, n$. Usando o Lema (2),

$$g_{ij}(\omega) = \mathbb{E}(U_{i\omega} U_{j\omega}) = \frac{1+n\eta}{1+(n+2)\eta} (2\gamma_{ij}^2 + \gamma_{ii}\gamma_{jj}) - \gamma_{ii}\gamma_{jj},$$

sendo γ_{ij} é o elemento i, j da matriz Σ_ω^{-1} , $i, j = 1, \dots, n$. Para n grande o suficiente, $\mathbf{G}_0 = \mathbf{G}(\omega_0) = 2\Sigma^{-1} \odot \Sigma^{-1}$. Então, este esquema de perturbação não é apropriado (ZHU et al., 2007). No entanto, de acordo com o que foi discutido acima, $\omega = p(\tilde{\omega}) = \mathbf{1}_n + \mathbf{G}_0^{-1/2}(\tilde{\omega} - \mathbf{1}_n)$ é um esquema de perturbação apropriado.

4.11 Referências

ANDERSON, T. W. Asymptotically efficient estimation of covariance matrices with linear structure. **The Annals of Statistics**, v. 1, n. 1, p. 135-141, 1973.

ASSUMPÇÃO, R. A. B.; URIBE-OPAZO, M. A.; GALEA, M. Analysis of local influence in geostatistics using Student's t-distribution. **Journal of Applied Statistics**, v. 41, n. 11, p. 2323-2341, 2014.

BOLFARINE, H.; GALEA, M. On structural comparative calibration under a t-model. **Computational Statistics**, PHYSICA VERLAG GMBH TIERGARTENSTRASSE 17, W-6900 HEIDELBERG, GERMANY, v. 11, n. 1, p. 63-85, 1996.

BOOS, D. D., STEFANSKI, L. **Essential Statistical Inference**, Theory and Methods. [S.I.]: Springer, New York, 2013.

COOK, R. D. Assessment of local influence. **Journal of the Royal Statistical Society. Series B (Methodological)**, [Royal Statistical Society, Wiley], v. 48, n. 2, p. 133-169, 1986. ISSN 00359246. Disponível em: <<http://www.jstor.org/stable/2345711>>.

COX, D. R., REID, N. Parameter orthogonality and approximate conditional inference. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 1-39, 1987.

CRESSIE, N. **Statistics for Spatial Data**. [S.I.]: John Wiley, New York, 1993.

DE BASTIANI, F.; URIBE-OPAZO, M., GALEA, M., CYSNEIROS, A. Case-deletion diagnostics for spatial linear mixed models. **Spatial statistics**, Elsevier, v. 28, p. 284-303, 2018.

DE BASTIANI, F., DE AQUINO CYSNEIROS, A. H. M., URIBE-OPAZO, M. A., GALEA, M. Influence diagnostics in elliptical spatial linear models. **Test**, v. 24, n. 2, p. 322-340, 2015.

DEMIDENKO, E. **Mixed Models Theory and Applications with R**. Second Edition. [S.I.]: John Wiley & Sons, New Jersey, 2013.

FAGUNDES, R. S., URIBE-OPAZO, M. A., GUEDES, L. P. C., GALEA, M. Slash spatial linear modeling: soybean yield variability as a function of soil chemical properties. **Revista Brasileira de Ciência do Solo**, v. 42, 2018.

FIORENTINI, G., SENTANA, E., CALZOLARI, G. Maximum likelihood estimation and inference in multivariate conditionally heteroscedastic dynamic regression models with Student t innovations. **Journal of Business & Economic Statistics**, v. 21, n. 4, p. 532-546, 2003.

GALEA, M., CADEMARTORI, D., CURCI, R., MOLINA, A. Robust Inference in the Capital Asset Pricing Model Using the Multivariate t-Distribution. **Journal of Risk and Financial Management**, v. 13, n. 6, p. 123, 2020.

GENTON, M. G., ZHANG, H. Identifiability problems in some non-gaussian spatial random fields. **Chilean Journal of Statistics**, v. 3, n. 2, p. 171-179, 2012.

HAMPEL, F. R., RONCHETTI, E. M., ROUSSEUW, P. J., STAHEL, W. A. **Robust statistics The approach based on influence functions**. [S.I.]: John Wiley, New York, 1986.

KAUFMAN, C. G., SHABY, B. A. The role of the range parameter for estimation and prediction in geostatistics. **Biometrika**, v. 100, n. 2, p. 473-484, 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, n. 408, p. 881-896, 1989.

LEMONTE, A. **The Gradient Test: Another Likelihood-Based Test**. [S.I.]: Academic Press,

London, 2016.

MATHERON, G. Principles of geostatistics. **Economic geology**, v. 58, n. 8, p. 1246-1266, 1963.

MINASNY, B., MCBRATNEY, A. B. The Matérn function as a general model for soil variograms. **Geoderma**, v. 128, n. 3-4, p. 192-207, 2005.

POON, W. Y., POON, Y. S. Conformal normal curvature and assessment of local influence. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 61, n. 1, p. 51-61, 1999.

SCHABENBERGER, O., GOTWAY, C. A. **Statistical methods for spatial data analysis**. CRC press, 2017.

SCHEMMER, R. C., URIBE-OPAZO, M. A., GALEA, M., ASSUMPÇÃO, R. A. Spatial variability of soybean yield through a reparameterized t-Student model. **Engenharia Agrícola**, v. 37, n. 4, p. 760-770, 2017.

STEIN, M. L. **Interpolation of Spatial Data: Some Theory for Kriging**. New York: Springer, 1999.

SUTRADHAR, B. Score test for the covariance matrix of the elliptic t-distribution. **Journal of Multivariate Analysis**, v. 46, n. 1, p. 1-12, 1993. ISSN 0047-259X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0047259X83710432>>.

TERRELL, G. The gradient statistic. **Computing Science and Statistics**, v. 34, p. 206-215, 2002.

URIBE-OPAZO, M. A., DE BASTIANI, F., GALEA, M., SCHEMMER, R. C., ASSUMPÇÃO, R. A. B. Influence diagnostics on a reparameterized t-Student spatial linear model. **Spatial Statistics**, p. 100481, 2021.

WALLER, L. A., GOTWAY, C. A. **Applied spatial statistics for public health data**. John Wiley & Sons, 2004.

ZELLNER, A. Bayesian and non-bayesian analysis of the regression model with multivariate student-t error terms. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 71, n. 354, p. 400-405, 1976.

ZHANG, H. Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. **Journal of the American Statistical Association**, v. 99, n. 465, p. 250-261, 2004.

ZHANG, H.; ZIMMERMAN, D. L. Hybrid estimation of semivariogram parameters. **Mathematical Geology**, v. 39, n. 2, p. 247-260, 2007.

ZHANG, H.; EL-SHAARAWI, A. On spatial skew-gaussian processes and applications. **Environmetrics**, v. 21, p. 33-47, 2010.

ZHU, H.-T.; LEE, S.-Y. Local influence for incomplete data models. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 63, n. 1, p.111-126, 2001.

ZHU, H.; IBRAHIM, J. G.; LEE, S.; ZHANG, H. et al. Perturbation selection and influence measures in local influence analysis. **The Annals of Statistics**, Institute of Mathematical Statistics, v. 35, n. 6, p. 2565-2588, 2007.

4.12 ARTIGO 2: Inferência estatística no modelo espacial linear t-Student reparametrizado

Resumo: A principal contribuição neste trabalho é o desenvolvimento dos métodos de influência global, local e alavanca generalizada para estudar dados espaciais correlacionados usando a distribuição t-Student reparametrizada. Esta distribuição tem sido utilizada como alternativa à distribuição normal, quando os dados apresentam valores discrepantes e uma distribuição com caudas pesadas se ajusta melhor. Além disso, o parâmetro de forma é limitado, supondo a existência do segundo momento finito, o qual permite a viabilidade da estimação dos parâmetros por máxima verossimilhança e a implementação do algoritmo iterativo. Técnicas de diagnóstico de influência global e local são utilizadas para identificar a presença de observações influentes que podem interferir na estimativa de parâmetros, seleção de modelos e construção de mapas. A metodologia é ilustrada com dois conjuntos de dados reais. Os resultados permitiram concluir que a presença de observações consideradas influentes nos dados exerce uma forte influência alterando a estrutura de dependência espacial, o qual também depende da sua localização e os valores de sua vizinhança.

Palavras-chave: Dependência espacial, modelagem robusta, pontos influentes.

4.13 Introdução

A geoestatística difere da estatística clássica, pois os modelos da estatística clássica estão geralmente voltados para verificar a distribuição de frequência dos dados, enquanto a geoestatística incorpora a interpretação da correlação espacial das amostras. Segundo Landim (2006), a geoestatística assume que a diferença entre dois pontos amostrais depende da distância entre eles e da orientação dos pontos, ou seja, pares de observações mais próximos são mais semelhantes do que pares de observações mais distantes. Porém, na prática existe a ocorrência de observações atípicas que podem afetar o comportamento da dependência espacial, especialmente quando a distribuição normal está sob suposição. Vários autores como Osorio et al. (2007), Prado et al. (2013), Assumpção et al. (2014), De Bastiani et al. (2015), Garcia-Papani et al. (2017), Galea e Castro (2017), Schemmer et al. (2017), Garcia-Papani et al. (2018), Fagundes et al. (2018a), Fagundes et al. (2018b), Martínez et al. (2019), Ordoñez et al. (2020) e Sánchez et al. (2020) apresentam abordagem com distribuições mais robustas, cujas estimativas dos modelos sejam menos sensíveis a estas observações. Uma alternativa proposta por Lange et al. (1989) e posteriormente estudada por Schemmer et al. (2017) e Uribe-Opazo et al. (2021) é a utilização da distribuição t-Student reparametrizada, a qual pertence à classe das distribuições simétricas. Ela permite reduzir a influência dos *outliers*, bem como a existência do segundo momento finito, e uma comparação direta entre a matriz da distribuição t-Student e da normal. Segundo Richetti et al. (2016), é possível que uma única observação possua grande influência nos resultados de uma análise de dados espaciais e possa alterar consideravelmente tanto os resultados que definem a estrutura de dependência espacial, como a construção dos mapas.

Seguindo este pensamento, é de suma importância a análise de diagnósticos para a detecção destas observações. Diversos trabalhos de análise de diagnósticos são apresentados, Osorio et al. (2007) avaliaram a influência local em modelos lineares elípticos com estrutura longitudinal, Uribe-Opazo et al. (2012) utilizaram técnicas de diagnóstico para avaliar a sensibilidade dos estimadores de máxima verossimilhança, funções de covariância e preditor linear a pequenas perturbações nos dados e nos pressupostos do modelo linear espacial. Assumpção et al. (2014) apresentaram o estudo da influência local por meio de dois métodos conhecidos como deslocamento de verossimilhança e Q-deslocamento de verossimilhança ambos utilizando a distribuição t-Student. De Bastiani et al. (2017) trabalharam com técnicas de diagnóstico para modelos espaciais gaussianos com repetição. Lachos et al. (2017) trabalharam com diagnósticos de influência em modelos espaciais com respostas censuradas. Garcia-Papani et al. (2018) apresentaram técnicas de diagnóstico para avaliar a influência de observações atípicas nas estimativas de máxima verossimilhança de seus parâmetros com modelo de regressão espacial Birnbaum-Saunders. Leiva et al. (2020) trabalharam com diagnósticos global e local para um modelo geoestatísticos com base em uma nova abordagem para regressão de quantis.

Assim, a proposta deste artigo é uma extensão dos artigos de Schemmer et al. (2017) e Uribe-Opazo et al. (2021) para a aplicação de análise de diagnósticos na distribuição t-Student reparametrizada com base nas técnicas de influência global e local desenvolvidas por máximo verossimilhança e a Q-function. O artigo está organizado da seguinte forma: a Seção 4.14 apresenta o modelo linear espacial t-Student reparametrizado, na Seção 4.14.1 apresenta a estimação de máximo verossimilhança e o desenvolvimento do processo de estimação dos parâmetros, na Seção 4.14.2 a descrição do algoritmo iterativo, a estimativa do erro padrão assintótico definido na Seção 4.14.3, na Seção 4.14.4 descreve-se os critérios de seleção do parâmetro de forma η , na Seção 4.14.5 descreve-se a metodologia para obter o gráfico QQ-plot, na Seção 4.15, foi desenvolvida a análise de diagnósticos para detectar pontos influentes com base nas técnicas de influência global, local e alavanca generalizada. Nas Seções 4.16 e 4.16.2, aplicou-se a metodologia a dados reais para o modelo t-Student reparametrizado. E na Seção 4.17, apresentam-se as conclusões obtidas a partir da metodologia aplicada.

4.14 Modelo espacial linear t-Student

Uma alternativa para a distribuição normal multivariada é a distribuição t-Student multivariada. Esta distribuição tem sido amplamente utilizada no estudo de dados reais por possuir caudas mais pesadas e permitir que os pontos atípicos presentes no conjunto de dados sejam englobados (LANGE et al., 1989; OSORIO et al., 2007). Além disso, é uma distribuição simétrica com parâmetro de forma extra, denominado grau de liberdade, representado por v ($v > 0$), que permite modelar a curtose dos dados.

Seguindo a metodologia de Uribe-Opazo et al. (2021), os quais apresentam a distribuição t-Student multivariada reparametrizada, aplicou-se a transformação $v = 1/\eta$.

Considera-se $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$ um vetor aleatório segue uma distribuição t-Student reparametrizada com parâmetro de forma fixo η , em que $0 < \eta < \frac{1}{2}$, com matriz de covariância Σ , $n \times n$, vetor de média $E(\mathbf{Y}) = \boldsymbol{\mu}$. Sua função de densidade de probabilidade é dada na Equação (4.11) por:

$$f_{\mathbf{Y}}(\mathbf{y}) = K_n(\eta) |\Sigma|^{-\frac{1}{2}} [1 + c(\eta)\delta]^{-\frac{1}{2\eta}(1+n\eta)}, \quad (4.11)$$

em que

$$K_n(\eta) = \left(\frac{c(\eta)}{\pi} \right)^{\frac{n}{2}} \frac{\Gamma\left(\frac{1+n\eta}{2\eta}\right)}{\Gamma\left(\frac{1}{2\eta}\right)},$$

com $\delta = (\mathbf{Y} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{Y} - \boldsymbol{\mu})$ é a distância de Mahalanobis, $c(\eta) = \eta/(1 - 2\eta)$ para $0 < \eta < \frac{1}{2}$. Seja $\mathbf{Y} \sim T_n(\boldsymbol{\mu}, \Sigma, \eta)$ denota um vetor aleatório seguindo uma distribuição t-Student reparametrizada n-multivariada.

A representação estocástica de \mathbf{Y} é dada por

$$\mathbf{Y} \stackrel{d}{=} \boldsymbol{\mu} + V^{-1/2} \mathbf{Z},$$

em que $\mathbf{Z} \sim N_n(\mathbf{0}, \Sigma)$, $V \sim Gama\left(\frac{1}{2\eta}, \frac{1}{2c(\eta)}\right)$, $c(\eta) = \eta/(1 - 2\eta)$, $0 < \eta < \frac{1}{2}$; V e \mathbf{Z} são independentes.

Para estudar a dependência espacial, considere um processo estocástico isotrópico de segunda ordem $\{\mathbf{Y}(\mathbf{s}_i), \mathbf{s}_i \in \mathcal{S}\}$, em que $\mathcal{S} \subset \mathbb{R}^2$ e \mathbb{R}^2 é um espaço Euclidiano bidimensional. Seja $\mathbf{Y} = (Y(\mathbf{s}_1), \dots, Y(\mathbf{s}_n))^\top$ um vetor de resposta com ordem $n \times 1$, correspondente à localização \mathbf{s}_i com $i = 1, \dots, n$. Segue uma distribuição t-Student n-multivariada reparametrizada, denota por $\mathbf{Y} \sim T_n(\boldsymbol{\mu}, \Sigma, \eta)$, em que cada elemento $Y(\mathbf{s}_i)$ pode ser escrito como $Y(\mathbf{s}_i) = \mu(\mathbf{s}_i) + e(\mathbf{s}_i)$, $i = 1, \dots, n$. Em que ambos o termo determinístico $\mu(\mathbf{s}_i)$ e estocástico $e(\mathbf{s}_i)$ podem depender da localização espacial em que $Y(\mathbf{s}_i)$ é observado. Assume-se que os erros aleatórios $e(\mathbf{s}_i)$ têm esperança igual a zero, ou seja, $E[e(\mathbf{s}_i)] = 0$, e a variação entre os pontos é determinada por alguma função de covariância $Cov[e(\mathbf{s}_i), e(\mathbf{s}_u)] = Cov[Y(\mathbf{s}_i), Y(\mathbf{s}_u)] = C(\mathbf{s}_i, \mathbf{s}_u) = \sigma_{iu}$, para $i, u = 1, \dots, n$. Suponha que para alguma função conhecida de \mathbf{s}_i , $x_1(\mathbf{s}_i), \dots, x_p(\mathbf{s}_i)$, a média do processo estocástico é dada por (4.12)

$$\boldsymbol{\mu}(\mathbf{s}_i) = \sum_{j=1}^p x_j(\mathbf{s}_i) \beta_j. \quad (4.12)$$

em que β_1, \dots, β_p são parâmetros desconhecidos a serem estimados.

Na notação matricial, o modelo linear espacial é dado por

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \quad (4.13)$$

em que $\mathbf{Y} \sim T_n(\mathbf{X}\boldsymbol{\beta}, \Sigma, \eta)$, \mathbf{X} é uma matriz de posto completo, com ordem $n \times p$, com a i -ésima linha de um e com o vetor das variáveis explicativas $\mathbf{x}_i^\top = (x_{i1}, \dots, x_{ip})$ referente à localização \mathbf{s}_i , $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ é um vetor $p \times 1$ de parâmetros desconhecidos a serem estimados e $\boldsymbol{\varepsilon} = (e(\mathbf{s}_1), \dots, e(\mathbf{s}_n))^\top$ são componentes correlacionados da variância com $\boldsymbol{\varepsilon} \sim T_n(\mathbf{0}, \Sigma, \eta)$.

A modelagem espacial dada na equação (4.13) depende da estrutura da matriz de

covariância $\Sigma = [(\sigma_{iu})]$, em que $\sigma_{iu} = C(\mathbf{s}_i, \mathbf{s}_u)$ para $i, u = 1, \dots, n$ do processo estocástico Y . A função de covariância $C(\mathbf{s}_i, \mathbf{s}_u)$ é usada no estudo da dependência espacial do processo estacionário e é especificada por um vetor tridimensional $\tau = (\tau_1, \tau_2, \tau_3)^\top$. Conforme apresentado por Uribe-Opazo et al. (2021), a forma paramétrica é dada por

$$\Sigma = \tau_1 \mathbf{I}_n + \tau_2 \mathbf{R},$$

em que τ_1 é o efeito pepita; $\tau_1 \geq 0$, τ_2 é conhecido como contribuição, $\tau_2 \geq 0$; \mathbf{I}_n é uma matriz de Identidade de ordem n ; \mathbf{R} é a matriz simétrica de ordem $n \times n$, em que os elementos estão em função de $\tau_3 > 0$, $\mathbf{R} = \mathbf{R}(\tau_3) = [(r_{ij})]$, sendo $r_{ii} = 1$, e $r_{ij} = \tau_2^{-1} C(\mathbf{s}_i, \mathbf{s}_j)$ para $\tau_2 \neq 0$, e $r_{ij} = 0$ para $\tau_2 = 0$, $i \neq j = 1, \dots, n$, em que r_{ij} depende da distância euclidiana, $h_{ij} = \|\mathbf{s}_i - \mathbf{s}_j\|$, entre pontos \mathbf{s}_i e \mathbf{s}_j . Uma reparametrização alternativa na função de covariância $C(\mathbf{s}_i, \mathbf{s}_u) = C(h_{iu})$ é sugerida por Stein (2012), sendo adaptado por Uribe-Opazo et al. (2021) auxilia na identificabilidade do modelo. Neste artigo considerou-se os parâmetros $\phi_1 = \tau_1$, $\phi_2 = \tau_2/\tau_3^{2\kappa}$ e $\phi_3 = \tau_3$ este último parâmetro considerado fixo e a reparametrização em $C(h_{iu})$ nos mesmos critérios estabelecidos por Uribe-Opazo et al. (2021).

4.14.1 Estimador de máxima verossimilhança

Supondo que $Y \sim T_n(\mathbf{X}\beta, \Sigma, \eta)$, em que η é o parâmetro de forma considerado fixo e $\theta = (\beta^\top, \phi^\top)^\top$, com $\beta = (\beta_1, \dots, \beta_p)^\top$ e $\phi = (\phi_1, \phi_2, \phi_3)^\top$ sendo ϕ_1 e ϕ_2 parâmetros desconhecidos e ϕ_3 parâmetro a ser definido e fixado em função do semivariograma, o logaritmo da função verossimilhança para a distribuição t-Student reparametrizada é dada na equação (4.14)

$$\mathcal{L}(\theta) = \log K_n(\eta) - \frac{1}{2} \log |\Sigma| - \frac{1}{2\eta} (1 + n\eta) \log(1 + c(\eta)\delta), \quad (4.14)$$

em que

$$\log K_n(\eta) = \frac{n}{2} \log\left(\frac{c(\eta)}{\pi}\right) + \log \Gamma\left(\frac{1+n\eta}{2\eta}\right) - \log \Gamma\left(\frac{1}{2\eta}\right), \quad \delta = (\mathbf{Y} - \mathbf{X}\beta)^\top \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\beta), \quad c(\eta) = \eta / (1 - 2\eta), \quad 0 < \eta < \frac{1}{2}.$$

A função escore pode ser escrita como:

$$U(\beta) = \frac{\partial \mathcal{L}(\theta)}{\partial \beta} = w(\delta) \mathbf{X}^\top \Sigma^{-1} \varepsilon, \\ U(\phi) = \frac{\partial \mathcal{L}(\theta)}{\partial \phi_j} = -\frac{1}{2} [tr(\Sigma^{-1} \dot{\Sigma}_j) - w(\delta) \varepsilon^\top \Sigma^{-1} \dot{\Sigma}_j \Sigma^{-1} \varepsilon],$$

em que $\dot{\Sigma}_j = \partial \Sigma / \partial \phi_j$ em que $j = 1, 2$, $w(\delta) = ((1 + n\eta)/\eta)c(\eta)q^{-1}$, $q = (1 + c(\eta)\delta)$ e $\varepsilon = (\mathbf{Y} - \mathbf{X}\beta)$. O parâmetro $\theta = (\beta^\top, \phi^\top)^\top$, pode ser estimado a partir da solução das funções escores $U(\beta) = \mathbf{0}$ e $U(\phi) = \mathbf{0}$, em que :

Para o parâmetro $\hat{\beta}$, a solução da função é obtida de forma imediata:

$$\hat{\beta} = (\mathbf{X}^\top (\Sigma^{(k)})^{-1} \mathbf{X})^{-1} (\mathbf{X}^\top (\Sigma^{(k)})^{-1} \mathbf{Y}),$$

Todavia, para se obter a estimativa do parâmetro $\hat{\phi}$, estimou-se ϕ_j com $j = 1, 2$ por meio da resolução do sistema (4.19). A partir de $U(\phi) = \mathbf{0}$ temos:

$$tr(\Sigma^{-1} \dot{\Sigma}_j) - w(\delta) \varepsilon^\top \Sigma^{-1} \dot{\Sigma}_j \Sigma^{-1} \varepsilon = 0 \quad (4.15)$$

A equação (4.15) pode ser reescrita como:

$$\text{tr}(\Sigma^{-1}\dot{\Sigma}_j) = \text{tr}(\Sigma^{-1}\dot{\Sigma}_j\Sigma^{-1}\mathbf{E}), \quad (4.16)$$

com $\mathbf{E} = w(\delta)\varepsilon\varepsilon^\top$, equivalente à equação (4.16),

$$\text{tr}(\Sigma^{-1}\dot{\Sigma}_j\Sigma^{-1}\Sigma) = \text{tr}(\Sigma^{-1}\dot{\Sigma}_j\Sigma^{-1}\mathbf{E}). \quad (4.17)$$

Considerando-se $\Sigma = \sum_{l=1}^2 \phi_l \dot{\Sigma}_l$, em que $\dot{\Sigma}_1 = \mathbf{I}_n$ e $\dot{\Sigma}_2 = \mathbf{R}(\phi_3)$, aplicando em (4.17)

$$\sum_{l=1}^2 \text{tr}(\Sigma^{-1}\dot{\Sigma}_j\Sigma^{-1}\dot{\Sigma}_l)\phi_l = \text{tr}(\Sigma^{-1}\dot{\Sigma}_j\Sigma^{-1}\mathbf{E}). \quad (4.18)$$

Note que, em notação matricial o sistema (4.18), pode ser escrito como:

$$\mathbf{A} \cdot \boldsymbol{\phi} = \mathbf{b},$$

sendo a matriz \mathbf{A} de ordem 2×2 , com elementos $a_{jl} = \text{tr}(\Sigma^{-1}\dot{\Sigma}_j\Sigma^{-1}\dot{\Sigma}_l)$ e \mathbf{b} um vetor (2×1) com elementos $b_j = \text{tr}(\Sigma^{-1}\dot{\Sigma}_j\Sigma^{-1}\mathbf{E})$, sendo correspondentes a $a_{11} = \text{tr}(\Sigma^{-1}\Sigma^{-1})$, $a_{12} = a_{21} = \text{tr}(\Sigma^{-1}\Sigma^{-1}\mathbf{R}(\phi_3))$, $a_{22} = \text{tr}(\Sigma^{-1}\mathbf{R}(\phi_3)\Sigma^{-1}\mathbf{R}(\phi_3))$, $b_1 = \text{tr}(\Sigma^{-1}\Sigma^{-1}\mathbf{E})$ e $b_2 = \text{tr}(\Sigma^{-1}\mathbf{R}(\phi_3)\Sigma^{-1}\mathbf{E})$. O processo de estimação é dado pela resolução do sistema (4.19):

$$\hat{\boldsymbol{\phi}} = \mathbf{A}^{-1} \cdot \mathbf{b}, \quad (4.19)$$

Na Seção 4.14.2, segue a descrição do algoritmo, o qual irá obter a estimativa dos parâmetros de $\boldsymbol{\theta}$ através de iterações, ou seja, irá obter $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(k)}$ quando atingir o critério de convergência na k -ésima iteração.

4.14.2 Algoritmo iterativo

Baseados em Acosta et al. (2016), para se obter a estimação dos parâmetros de $\boldsymbol{\theta}$, devemos utilizar um algoritmo iterativo, que segue os seguintes procedimentos:

1º Iteração: Seja $k = 0$, em que k representa a fase de iteração:

1º passo: Neste passo, define-se um chute inicial para o parâmetro a ser estimado $\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^0$, em que $\boldsymbol{\theta}^0 = (\boldsymbol{\beta}^0, \phi^0, \eta^0)^\top$. Neste estudo, optou-se em definir os parâmetros iniciais $\boldsymbol{\beta}^0$ e ϕ^0 , obtendo-os em um modelo de regressão com distribuição normal e η^0 foi definido fixo para todas as iterações, o qual foi posteriormente escolhido pelos critérios CV e Tr apresentados na Seção 4.14.4.

2º passo: Calculam-se a partir dos parâmetros iniciais obtidos no 1º passo, as seguintes equações:

$$\Sigma^{(k)} = \phi_1^{(k)} \mathbf{I} + \phi_2^{(k)} \mathbf{R}^{(k)}(\phi_3), \text{ em que } \phi_3 \text{ é fixo no chute inicial.}$$

$$c(\eta) = \frac{\eta}{1-2\eta},$$

$$q^{(k)} = (1 + c(\eta)\delta^{(k)}),$$

$$\delta^{(k)} = (\boldsymbol{\varepsilon}^{(k)})^\top (\Sigma^{(k)})^{-1} \boldsymbol{\varepsilon}^{(k)},$$

$$\boldsymbol{\varepsilon}^{(k)} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}^{(k)}),$$

$$\begin{aligned}
w(\delta)^{(k)} &= \left(\frac{1+n\eta}{\eta} \right) c(\eta)(q^{(k)})^{-1}, \\
\beta^{(k)} &= (\mathbf{X}^\top (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{X})^{-1} (\mathbf{X}^\top (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{Y}), \\
\mathbf{E}^{(k)} &= w(\delta)^{(k)} \boldsymbol{\varepsilon}^{(k)} (\boldsymbol{\varepsilon}^{(k)})^\top,
\end{aligned}$$

em que \mathbf{R} é a função de covariância que depende do modelo exponencial, gaussiano ou família Matérn, em que \mathbf{R} é apresentada por Uribe-Opazo et al. (2012). Verifique que estas equações calculadas no 2º passo são consideradas iniciais, ou seja, com $k = 0$.

3º passo: A partir deste momento, atualizam-se os parâmetros para $(k + 1)$ em que $\boldsymbol{\theta}^{(k+1)} = (\beta^{(k+1)}, \phi^{(k+1)}, \eta^0)^\top$. Considerando-se os seguintes procedimentos:

3.1 passo: Atualizam-se os parâmetros lineares ϕ_1 e ϕ_2 de $\boldsymbol{\Sigma}$, através do sistema linear:

$$\mathbf{A}^{(k)} \cdot \phi^{(k+1)} = \mathbf{b}^{(k)},$$

com a matriz $\mathbf{A}^{(k)}$ de ordem 2×2 , com elementos $a_{jl}^{(k)} = tr((\boldsymbol{\Sigma}^{(k)})^{-1} \dot{\boldsymbol{\Sigma}}_j (\boldsymbol{\Sigma}^{(k)})^{-1} \dot{\boldsymbol{\Sigma}}_l)$ e $\mathbf{b}^{(k)}$ um vetor (2×1) com elementos $b_j^{(k)} = tr((\boldsymbol{\Sigma}^{(k)})^{-1} \dot{\boldsymbol{\Sigma}}_j (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{E}^{(k)})$, em que $j, l = 1, 2$ sendo correspondente a $a_{11}^{(k)} = tr((\boldsymbol{\Sigma}^{(k)})^{-1} (\boldsymbol{\Sigma}^{(k)})^{-1})$, $a_{12}^{(k)} = a_{21}^{(k)} = tr((\boldsymbol{\Sigma}^{(k)})^{-1} (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{R}^{(k)}(\phi_3))$, $a_{22}^{(k)} = tr((\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{R}^{(k)}(\phi_3) (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{R}^{(k)}(\phi_3))$, para o vetor $b_1^{(k)} = tr((\boldsymbol{\Sigma}^{(k)})^{-1} (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{E}^{(k)})$ e $b_2^{(k)} = tr((\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{R}^{(k)}(\phi_3) (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{E}^{(k)})$.

3.2 passo: Obtendo-se $\phi^{(k+1)}$ dos passos (3.1), atualiza-se $\boldsymbol{\Sigma}^{(k+1)}$, a qual será utilizada para atualizar o parâmetro $\beta^{(k+1)}$, pela expressão dada por:

$$\beta^{(k+1)} = (\mathbf{X}^\top (\boldsymbol{\Sigma}^{(k+1)})^{-1} \mathbf{X})^{-1} (\mathbf{X}^\top (\boldsymbol{\Sigma}^{(k+1)})^{-1} \mathbf{Y}),$$

4º passo: Finaliza-se a iteração, definindo-se $\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(1)}$, obtida pelas atualizações (3.1) e (3.2). A partir de $\boldsymbol{\theta}^{(1)}$, aplica-se o critério de convergência, que para este algoritmo é definido: se $\|\boldsymbol{\theta}^{(k+1)} - \boldsymbol{\theta}^{(k)}\| < e_1$ como η e ϕ_3 são fixos, somente verifica-se a convergência para $\|\beta^{(k+1)} - \beta^{(k)}\| < e_1$ e $\|\phi_j^{(k+1)} - \phi_j^{(k)}\| < e_1$ em que $j = 1, 2$ ou $|\mathcal{L}_c(\boldsymbol{\theta}^{(k+1)}) - \mathcal{L}_c(\boldsymbol{\theta}^{(k)})| < e_2$ pare e defina $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(k+1)}$, caso contrário $k = k + 1$ retorne para a 2º passo. Sendo que e_1 e e_2 são tolerâncias constantes. Em geral os valores típicos para as tolerâncias são 10^{-3} e 10^{-6} , respectivamente.

4.14.3 Estimativa do erro padrão assintótico

O erro padrão assintótico pode ser calculado ao se inverter a matriz de informação esperada, $\mathbf{F}(\boldsymbol{\theta})$, em que $\mathbf{F}(\boldsymbol{\theta}) = E[-\mathbf{L}(\boldsymbol{\theta})]$, com $\mathbf{L}(\boldsymbol{\theta}) = \partial^2 \mathcal{L}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top$. A t-Student reparametrizada $\mathbf{F}(\boldsymbol{\theta})$ é dada por (WALLER; GOTWAY, 2004),

$$\mathbf{F} = \mathbf{F}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{F}_{\beta\beta} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{\phi\phi} \end{pmatrix},$$

sendo

$$\mathbf{F}_{\beta\beta} = \mathbf{F}_{\beta\beta}(\beta) = \frac{(1 + n\eta)}{(1 - 2\eta)(1 + (n + 2)\eta)} \mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X},$$

e

$$\mathbf{F}_{\phi\phi} = \mathbf{F}_{\phi\phi}(\phi) = \frac{1}{4} \left(\frac{\partial \text{vec}^T(\boldsymbol{\Sigma})}{\partial \phi} \right) \left[2 \frac{(1+n\eta)}{(1+(n+2)\eta)} (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{N}_n + \left(\frac{(1+n\eta)}{(1+(n+2)\eta)} - 1 \right) \text{vec}(\boldsymbol{\Sigma}^{-1}) \text{vec}^\top(\boldsymbol{\Sigma}^{-1}) \right] \left(\frac{\partial \text{vec}(\boldsymbol{\Sigma})}{\partial \phi^\top} \right),$$

com, $\mathbf{N}_n = \frac{1}{2} (\mathbf{I}_{n^2} + \mathbf{K}_n)$, \mathbf{K}_n é a matriz de comutação de ordem $n^2 \times n^2$ e \mathbf{I}_{n^2} a matriz de Identidade de ordem $n^2 \times n^2$ (MAGNUS; NEUDECKER, 1999).

4.14.4 Critério de seleção do parâmetro de forma (η)

Segundo Zellner (1976) e De Bastiani et al. (2015), a função log-verossimilhança dada em (4.14) é decrescente em η , então a estimação deste parâmetro não pode ser obtida através da maximização da log-verossimilhança ($\mathcal{L}(\theta)$). Como alternativa, Kano et al. (1993) propuseram usar o traço da matriz de covariância assintótica de uma média estimada como critério na seleção de um modelo melhor para a classe das distribuições elípticas. Seguindo este pensamento, De Bastiani et al. (2015) apresentaram o critério traço e a validação cruzada para selecionar o grau de liberdade ν do modelo espacial linear t-Student. A partir destas propostas, Schemmer et al. (2017) apresentaram o critério traço (T_r) e validação cruzada (CV) para selecionar o parâmetro de forma η para o modelo espacial linear t-Student reparametrizado, sendo este aplicado no presente artigo, e proposto por Uribe-Opazo et al. (2021). Para os dois métodos, o melhor parâmetro de forma (η) é determinado pelos menores valores da validação cruzada e traço. Depois de escolher η , o parâmetro do modelo Matérn (κ) é selecionado de acordo com o menor erro padrão assintótico.

4.14.5 Gráfico QQ plot

Seguindo a metodologia apresentada por Dalposso et al. (2018), considera-se o conjunto de dados $\{\mathbf{Y}(s_1), \dots, \mathbf{Y}(s_n)\}$, determina-se o vetor de resíduo $\hat{\boldsymbol{\varepsilon}} = (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})$, em que $\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{X})^{-1} (\mathbf{X}^\top \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{Y})$, sendo o vetor de parâmetros $\hat{\boldsymbol{\beta}}$ e a matriz de covariância $\hat{\boldsymbol{\Sigma}}$, obtidas através das seções 4.14.2 e 4.14.4. A partir da estimativa da matriz de covariância $\hat{\boldsymbol{\Sigma}}$, utiliza-se a decomposição de Cholesky obtida por $\hat{\boldsymbol{\Sigma}} = \hat{\mathbf{L}}\hat{\mathbf{L}}^\top$, em que $\hat{\mathbf{L}}$ é matriz triângular inferior de ordem n . Utilizando a matriz inversa $\hat{\mathbf{L}}^{-1}$, determina-se $\hat{\boldsymbol{\varepsilon}}_{nc} = \hat{\mathbf{L}}^{-1}\hat{\boldsymbol{\varepsilon}}$, definido como o vetor de resíduo não correlacionado. Em seguida, utilizou-se a metodologia de Almeida et al. (2018) com os pacotes qqplotr e ggplot2 no programa R para a construção dos QQ-plots, considerando como dados amostrais o vetor de resíduo não correlacionado $\hat{\boldsymbol{\varepsilon}}_{nc}$ e dados teóricos o vetor aleatório com distribuição t-Student ambos de ordem n . Para a obtenção dos intervalos de confiança, utilizamos o pacote desenvolvido por Wickham (2011) com a aplicação do método de Boot que cria bandas de confiança com base em um Bootstrap paramétrico.

4.15 Diagnóstico de influência

Um estudo de grande importância na análise de diagnóstico é a detecção de observações influentes, isto é, pontos que exercem um peso desproporcional nas estimativas dos parâmetros do modelo ou até mesmo na significância dos parâmetros. A detecção de pontos talvez seja a técnica mais conhecida para avaliar o impacto da retirada de uma observação particular nas estimativas da regressão. Neste artigo, consideram-se dois tipos de diagnóstico de influência: influência global e influência local.

4.15.1 Influência global

A exclusão de caso é uma maneira comum de avaliar o efeito de uma observação na estimativa do processo. Esta é uma análise de influência global uma vez que o efeito da observação é avaliado pela eliminação no conjunto de dados.

4.15.1.1 Influência global baseada na verossimilhança

A análise de influência global fundamenta-se na eliminação de uma ou mais observações consideradas influentes no conjunto de dados e, assim avalia o impacto nas estimativas dos parâmetros. Este tipo de técnicas de diagnóstico é discutido por Cook; Weisberg (1982), Chatterjee; Hadi (2009) e Pan et al. (2014). Uma das medidas mais utilizadas para mensurar as mudanças dos parâmetros estimados, após excluídas as observações, é denominada de distância de Cook (COOK, 1977). Esta medida foi proposta inicialmente para modelos normais e rapidamente expandida para diversas classes de modelos. Seguindo a proposta de Pan et al. (2014), a distância de Cook baseada no estimador ML $\hat{\theta}$, de $\theta = (\beta^\top, \phi^\top)^\top$, é dada por

$$D_{i\theta}^1 = U_{(i)}^\top \left(\hat{\theta} \right) \left[F \left(\hat{\theta} \right) \right]^{-1} U_{(i)} \left(\hat{\theta} \right), \text{ para } i = 1, \dots, n,$$

o qual pode ser decomposto em $D_{i\theta}^1 = D_{i\beta} + D_{i\phi}$, para $i = 1, \dots, n$, sendo $D_{i\beta} = U_{(i)}^\top \left(\hat{\beta} \right) F_{\beta\beta}^{-1} \left(\hat{\beta} \right) U_{(i)} \left(\hat{\beta} \right)$ e $D_{i\phi} = U_{(i)}^\top \left(\hat{\phi} \right) F_{\phi\phi}^{-1} \left(\hat{\phi} \right) U_{(i)} \left(\hat{\phi} \right)$, em que $U_{(i)} \left(\hat{\beta} \right)$ e $U_{(i)} \left(\hat{\phi} \right)$ são as funções escores dos estimadores β e ϕ , respectivamente, sem a i -ésima observação.

4.15.1.2 Influência global baseada na Q-function

Segundo Pan et al. (2014), é difícil estender o método de exclusão de casos para outros modelos se a função de verossimilhança não tiver forma analítica. Assim, Zhu e Lee (2001), Pan et al. (2014) e De Bastiani et al. (2018) apresentaram a distância de Cook a partir do uso da esperança condicional do logaritmo da função verossimilhança completa, denominada Q-function. Seguindo este raciocínio, considere a Q-function dada por

$$\begin{aligned} Q(\theta|\theta^*) &= E[l_c(\theta|Y_c), Y, \theta^*] = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log|\Sigma| - \frac{1}{2} \log\left(\Gamma\left(\frac{1}{2\eta}\right)\right) \\ &\quad + \frac{1}{2\eta} \log\left(\frac{1}{2c(\eta)}\right) + \frac{n}{2} a^* + \frac{1}{2\eta} [a^* - w(\delta^*)] - \frac{1}{2} w(\delta^*) \delta^*, \end{aligned} \quad (4.20)$$

em que $a^* = -\log(w(\delta^*)) - \psi\left(n + \frac{1}{\eta}\right) + \log\left(\frac{1+n\eta}{\eta}\right)$, com $w(\delta^*) = \left(\frac{1+n\eta}{\eta}\right) c(\eta) q^{*-1}$, $q^* = (1 + c(\eta)\delta^*)$, $\delta^* = (\mathbf{Y} - \mathbf{X}\beta^*)^\top \Sigma^{*-1} (\mathbf{Y} - \mathbf{X}\beta^*)$, $\psi\left(n + \frac{1}{\eta}\right)$ é a função Digama, para $0 < \eta < \frac{1}{2}$, Zhu e Lee (2001) propuseram uma medida de influência global como alternativa para obter $\hat{\theta}_{(i)}$, com a seguinte aproximação

$$\hat{\theta}_{(i)} = \hat{\theta} + \left[-\ddot{Q}_{(i)}(\hat{\theta}|\hat{\theta})\right]^{-1} \dot{Q}_{(i)}(\hat{\theta}|\hat{\theta}), \text{ para } i = 1, \dots, n, \text{ em que } \dot{Q}_{(i)}(\hat{\theta}|\hat{\theta}) = \frac{\partial Q_{(i)}(\theta|\hat{\theta})}{\partial \theta} \Big|_{\theta=\hat{\theta}} \text{ e } \ddot{Q}_{(i)}(\hat{\theta}|\hat{\theta}) = \frac{\partial^2 Q_{(i)}(\theta|\hat{\theta})}{\partial \theta \partial \theta^\top} \Big|_{\theta=\hat{\theta}}.$$

Zhu e Lee (2001) propuseram substituir $\left[-\ddot{Q}_{(i)}(\hat{\theta}|\hat{\theta})\right]^{-1}$ por $\left[-\ddot{Q}(\hat{\theta}|\hat{\theta})\right]^{-1}$, em que $\ddot{Q}(\hat{\theta}|\hat{\theta}) = \frac{\partial^2 Q(\theta|\hat{\theta})}{\partial \theta \partial \theta^\top} \Big|_{\theta=\hat{\theta}}$. Então a distância aproximada de Cook, com base na Q-function, é dada por:

$$QD_{i\theta}^* = \left[\dot{Q}_{(i)}(\hat{\theta}|\hat{\theta})\right]^\top \left[-\ddot{Q}(\hat{\theta}|\hat{\theta})\right]^{-1} \left[-\dot{Q}_{(i)}(\hat{\theta}|\hat{\theta})\right], \text{ para } i = 1, \dots, n.$$

Pan et al. (2014) apresentaram um teorema mostrando que é possível substituir $\left[-\ddot{Q}(\hat{\theta}|\hat{\theta})\right]^{-1}$ por $E\left[-\ddot{Q}(\hat{\theta}|\hat{\theta})\right]^{-1}$, então a nova modificação da distância de Cook com base na Q-function é dada na equação (4.21) por

$$QD_{i\theta}^1 = \left[\dot{Q}_{(i)}(\hat{\theta}|\hat{\theta})\right]^\top E\left[-\ddot{Q}(\hat{\theta}|\hat{\theta})\right]^{-1} \left[-\dot{Q}_{(i)}(\hat{\theta}|\hat{\theta})\right] \text{ para } i = 1, \dots, n. \quad (4.21)$$

com $E\left[-\ddot{Q}_{(i)}(\hat{\theta}|\hat{\theta})\right]^{-1}$ é uma matriz diagonal em bloco em relação à β e ϕ . Assim, a estatística modificada de Cook $QD_{i\theta}^1$ de (4.21) pode ser escrita como :

$$QD_{i\theta}^1 = QD_{i\beta}^1 + QD_{i\phi}^1, \text{ para } i = 1, \dots, n,$$

em que,

$$QD_{i\beta}^1 = \left[\dot{Q}_{\beta(i)}(\hat{\theta}|\hat{\theta})\right]^\top E\left[-\ddot{Q}_{\beta\beta}(\hat{\theta}|\hat{\theta})\right]^{-1} \left[-\dot{Q}_{\beta(i)}(\hat{\theta}|\hat{\theta})\right], \text{ e}$$

$$QD_{i\phi}^1 = \left[\dot{Q}_{\phi(i)}(\hat{\theta}|\hat{\theta})\right]^\top E\left[-\ddot{Q}_{\phi\phi}(\hat{\theta}|\hat{\theta})\right]^{-1} \left[-\dot{Q}_{\phi(i)}(\hat{\theta}|\hat{\theta})\right].$$

Portanto,

$$\dot{Q}_{\beta(i)}(\hat{\theta}|\hat{\theta}) = w_{(i)}(\hat{\delta}) \mathbf{X}_{(i)}^\top \hat{\Sigma}_{(i)}^{-1} \hat{\epsilon}_{(i)},$$

$$\dot{Q}_{\phi(i)}(\hat{\theta}|\hat{\theta}) = \frac{1}{2} \left(\frac{\partial \text{vec}^\top(\hat{\Sigma}_{(i)})}{\partial \phi} \right) \left[w_{(i)}(\hat{\delta}) \text{vec} \left(\hat{\Sigma}_{(i)}^{-1} \hat{\epsilon}_{(i)} \hat{\epsilon}_{(i)}^\top \hat{\Sigma}_{(i)}^{-1} \right) - \text{vec} \left(\hat{\Sigma}_{(i)}^{-1} \right) \right],$$

$$E\left[-\ddot{Q}_{\beta\beta}(\hat{\theta}|\hat{\theta})\right] = E\left[-\frac{\partial^2 Q(\theta|\hat{\theta})}{\partial \beta \partial \beta^\top}\right] \Big|_{\theta=\hat{\theta}} = w(\hat{\delta}) \mathbf{X}^\top \hat{\Sigma}^{-1} \mathbf{X},$$

$$E\left[-\ddot{Q}_{\phi\phi}(\hat{\theta}|\hat{\theta})\right] = E\left[-\frac{\partial^2 Q(\theta|\hat{\theta})}{\partial \phi \partial \phi^\top}\right] \Big|_{\theta=\hat{\theta}} = \frac{1}{2} \left(\frac{\partial \text{vec}(\hat{\Sigma})}{\partial \phi} \right)^\top \left(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1} \right) \left(\frac{\partial \text{vec}(\hat{\Sigma})}{\partial \phi} \right),$$

com $w(\hat{\delta}) = \left(\frac{1+n\eta}{\eta}\right) c(\eta) \hat{q}^{-1}$, $w_{(i)}(\hat{\delta}) = \left(\frac{1+n\eta}{\eta}\right) c(\eta) \hat{q}_{(i)}^{-1}$, $\hat{q} = (1 + c(\eta)\hat{\delta})$ e $\hat{q}_{(i)} = (1 + c(\eta)\hat{\delta}_{(i)})$.

4.15.2 Influência local

O método de influência local, proposto por Cook (1986), consiste em avaliar a robustez das estimativas obtidas considerando uma medida de influência e sob pequenas perturbações aplicadas ao modelo ou aos dados, ou seja, para verificar a presença de observações que podem causar distorções nos resultados, sob pequenas perturbações. Este método não requer a eliminação de observações.

Considere o modelo espacial linear dado na equação (4.13). Supondo que o vetor de

erro aleatório ε segue uma distribuição t-Student, com média igual a um vetor de zeros e matriz de covariância Σ , ou seja, $\varepsilon \sim T_n(\mathbf{0}, \Sigma, \eta)$, é possível obter o modelo de regressão $\hat{Y} = X\hat{\beta}$. A observação é chamada de influente se causar alterações relevantes nos resultados. O esquema apropriado da perturbação para a variável resposta de acordo com De Bastiani et al. (2015) é dado por

$$Y_\omega = Y + A\omega, \quad (4.22)$$

em que ω é um vetor $n \times 1$, pertencente ao espaço de perturbação Ω .

4.15.2.1 Influência local baseada no afastamento da verossimilhança

Seja $\mathcal{L}(\theta|\omega)$ definida como a log verossimilhança perturbada. A influência da perturbação, causada pelo vetor ω , nas estimativas dos parâmetros ML de θ , pode ser avaliada pelo afastamento da verossimilhança, definido como

$$LD(\omega) = 2 \left(\mathcal{L}(\hat{\theta}) - \mathcal{L}(\hat{\theta}_\omega) \right),$$

em que $\hat{\theta}$ é o estimador de ML de θ do modelo postulado e $\hat{\theta}_\omega$ é o estimador de ML de θ do modelo perturbado por $\omega \in \Omega$. Cook (1986) propôs estudar o comportamento local de $LD(\omega)$ sobre $\omega_0 \in \Omega$, de tal modo que $\mathcal{L}(\theta|\omega_0) = \mathcal{L}(\theta)$. Assim, é utilizada a curvatura normal C_l de $LD(\omega)$ em ω_0 na direção de um vetor unitário l :

$$C_l = 2 \left| l^\top \Delta^\top L \Delta l \right|.$$

em que $\|l\| = 1$; L : é a matriz Hessiana avaliada em $\theta = \hat{\theta}$; Δ : é uma matriz de ordem $(p+q) \times n$, definida por $\Delta = (\Delta_\beta^\top, \Delta_\phi^\top)^\top$, avaliada em $\theta = \hat{\theta}$ e em $\omega = \omega_0$, em que $\Delta_\beta = \frac{\partial^2 \mathcal{L}(\hat{\theta})}{\partial \beta \partial \omega^\top}$ e $\Delta_\phi = \frac{\partial^2 \mathcal{L}(\hat{\theta})}{\partial \phi \partial \omega^\top}$.

De Bastiani et al. (2015) consideraram a perturbação apropriada generalizada de Zhu et al. (2007), como sendo $Y_\omega = Y + A\omega$, na variável resposta e utilizando a matriz A , $n \times n$, que não depende de β e nem de ω , de modo que a matriz de informações de Fisher para Y_ω com relação ao vetor perturbado ω é $G(\omega_0) = cA^\top \Sigma^{-1} A$, em que c é uma constante positiva. Em geral, $A^\top \Sigma^{-1} A \neq I_n$, entretanto, se $A = \Sigma^{1/2}$, então $G(\omega_0) = cI_n$ em que a perturbação é apropriada, considerando uma distribuição t-Student reparametrizado, $Y_\omega = Y + \Sigma^{1/2}\omega$, é dada na equação (4.22). Neste estudo, a matriz $\Delta = (\Delta_\beta^\top, \Delta_\phi^\top)^\top$ avaliada em $\theta = \hat{\theta}$ e em $\omega = \omega_0$, é dada por

$$\Delta_\beta = w(\hat{\delta}) \left[X^\top \hat{\Sigma}^{-1/2} - 2c(\eta)\hat{q}^{-1} X^\top \hat{\Sigma}^{-1} \hat{\varepsilon} \hat{\varepsilon}^\top \hat{\Sigma}^{-1/2} \right],$$

e Δ_ϕ com elementos

$$\Delta_{\phi_j} = w(\hat{\delta}) \left[\hat{\Sigma}^{-1/2} \frac{\partial \hat{\Sigma}^{1/2}}{\partial \phi_j} \hat{\Sigma}^{-1/2} \hat{\varepsilon} - c(\eta)\hat{q}^{-1} \hat{\Sigma}^{-1/2} \hat{\varepsilon} \hat{\varepsilon}^\top \hat{\Sigma}^{-1} \frac{\partial \hat{\Sigma}}{\partial \phi_j} \hat{\Sigma}^{-1} \hat{\varepsilon} \right], j = 1, 2,$$

com $\hat{\varepsilon} = \hat{\varepsilon}(\omega_0) = (Y - X\hat{\beta})$, $w(\hat{\delta}) = \left(\frac{1+n\eta}{\eta} \right) c(\eta)\hat{q}^{-1}$, $\hat{q} = (1 + c(\eta)\hat{\delta})$, $\hat{\delta} = (Y - X\hat{\beta})^\top \hat{\Sigma}^{-1} (Y - X\hat{\beta})$, $c(\eta) = \eta/(1 - 2\eta)$ e $0 < \eta < \frac{1}{2}$.

Considere a matriz $B = \Delta^\top L^{-1} \Delta$ e $C_i = 2 |b_{ii}|$, para $i = 1, \dots, n$, em que b_{ii} são

os elementos da diagonal principal da matriz B . A representação gráfica de C_i versus i (ordem dos dados) pode ser usada para detectar possíveis observações influentes. Verbeke e Molenberghs (2009) propuseram considerar como observação potencialmente influente a i -ésima observação com $C_i > 2\bar{C}$, em que $\bar{C} = \frac{1}{n} \sum_{i=1}^n C_i$. Outra proposta apresentada por Cook (1986) define L_{max} como o primeiro autovetor normalizado e associado ao maior autovalor da matriz $B = \Delta^\top L^{-1} \Delta$. Assim, com os elementos de $|L_{max}|$ versus i (ordem dos dados), obtém-se um gráfico que pode revelar qual tipo de perturbação tem a maior influência em $LD(\omega)$ em ω_0 .

4.15.2.2 Influência local baseada na Q-function

O objetivo principal é comparar $\hat{\theta}$, a estimativa de ML de θ do modelo postulado, e $\hat{\theta}_\omega$, a estimativa de ML de θ do modelo perturbado, quando $\omega \in \Omega$. Valores próximos indicam que a perturbação tem pequeno efeito no procedimento de estimação. Por outro lado, se diferirem consideravelmente, então é possível que o procedimento de estimação seja sensível à presença de algumas observações. Para medir esta distância, Zhu e Lee (2001) e Zhu et al. (2007) propuseram o deslocamento da Q-function, calculando a diferença entre $\hat{\theta}$ e $\hat{\theta}_\omega$, definido por:

$$f_Q(\omega) = 2 \left(Q(\theta|\hat{\theta}) - Q(\theta, \omega|\hat{\theta}) \right),$$

sendo que $Q(\theta|\hat{\theta}) = E[l_c(\theta|Y_c), Y, \hat{\theta}]$ e $Q(\theta, \omega|\hat{\theta}) = E[l_c(\theta, \omega|Y_c), Y, \hat{\theta}]$ definida na equação (4.20), com $f_Q(\omega) \geq 0, \forall \omega \in \Omega$ e $f_Q(\omega_0) = 0$. Similarmente a Cook (1986), Zhu e Lee (2001) estudaram o comportamento da superfície $r(\omega) = (\omega^\top, f_Q(\omega)^\top)$ e calcularam a curvatura normal para a direção unitária $l \in \mathbb{R}^p$, dada por

$$C_Q = 2 \left| l^\top \Delta_\omega^\top \ddot{Q}(\hat{\theta}|\hat{\theta})^{-1} \Delta_\omega l \right|,$$

em que, $\|l\| = 1$; $\ddot{Q}(\hat{\theta}|\hat{\theta}) = \frac{\partial^2 Q(\theta|\hat{\theta})}{\partial \theta \partial \theta^\top} \Big|_{\theta=\hat{\theta}, \omega=\omega_0}$, avaliado em $\theta = \hat{\theta}, \omega = \omega_0$; Δ_ω é uma matriz $(p+3) \times n$, fornecida por $\Delta_\omega = (\Delta_{\omega\beta}^\top, \Delta_{\omega\phi}^\top)^\top$, avaliado em $\theta = \hat{\theta}, \omega = \omega_0$. Portanto:

$$\ddot{Q}(\hat{\theta}|\hat{\theta}) = \begin{pmatrix} \ddot{Q}(\hat{\beta}\hat{\beta}) & \ddot{Q}(\hat{\beta}\hat{\phi}) \\ \ddot{Q}(\hat{\phi}\hat{\beta}) & \ddot{Q}(\hat{\phi}\hat{\phi}) \end{pmatrix},$$

em que

$$\begin{aligned} \ddot{Q}(\hat{\beta}\hat{\beta}) &= -w(\hat{\delta}) \mathbf{X}^\top \hat{\Sigma}^{-1} \mathbf{X}, \\ \ddot{Q}(\hat{\beta}\hat{\phi}) &= -w(\hat{\delta}) \mathbf{X}^\top \hat{\Sigma}^{-1} \frac{\partial \hat{\Sigma}}{\partial \phi} \left(\hat{\Sigma}^{-1} \otimes \mathbf{I}_3 \right) \left(\mathbf{I}_3 \otimes \hat{\varepsilon} \right) = \ddot{Q}^\top(\hat{\phi}\hat{\beta}), \\ \ddot{Q}(\hat{\phi}\hat{\phi}) &= \frac{1}{2} \frac{\partial \text{vec}^\top(\hat{\Sigma})}{\partial \phi} \left(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1} \right) \frac{\partial \text{vec}(\hat{\Sigma})}{\partial \phi} \\ &\quad - \frac{1}{2} \left(\text{vec}^\top(\hat{\Sigma}^{-1}) \otimes \mathbf{I}_3 \right) \frac{\partial \text{vec}}{\partial \phi} \left(\frac{\partial \text{vec}^\top(\hat{\Sigma})}{\partial \phi} \right) \\ &\quad + w(\hat{\delta}) \frac{\text{vec}^\top(\hat{\Sigma})}{\partial \phi} \left(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1} \hat{\varepsilon} \hat{\varepsilon}^\top \hat{\Sigma}^{-1} \right) \frac{\partial \text{vec}^\top(\hat{\Sigma})}{\partial \phi} \\ &\quad + w(\hat{\delta}) \left(\text{vec}^\top \left(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1} \hat{\varepsilon} \hat{\varepsilon}^\top \hat{\Sigma}^{-1} \right) \otimes \mathbf{I}_3 \right) \frac{\partial \text{vec}}{\partial \phi} \left(\frac{\partial \text{vec}^\top(\hat{\Sigma})}{\partial \phi} \right), \end{aligned}$$

e a matriz Δ_ω tem elementos

$$\Delta_{\omega\beta} = \frac{\partial^2 Q(\theta, \omega | \hat{\theta})}{\partial \beta \partial \omega^\top} = w(\hat{\delta}) \mathbf{X}^\top \hat{\Sigma}^{-1/2},$$

$$\Delta_{\omega\phi} = \frac{\partial^2 Q(\theta, \omega | \hat{\theta})}{\partial \phi \partial \omega^\top}, \text{ com elementos}$$

$$\Delta_{\omega\phi_j} = w(\hat{\delta}) \hat{\Sigma}^{-1/2} \frac{\partial \hat{\Sigma}^{1/2}}{\partial \phi_j} \hat{\Sigma}^{-1/2} \hat{\epsilon}, \quad j = 1, 2.$$

sendo $w(\hat{\delta}) = \left(\frac{1+n\eta}{\eta} \right) c(\eta) \hat{q}^{-1}$, com $\hat{q} = (1 + c(\eta)\hat{\delta})$ e $\hat{\delta} = (\mathbf{Y} - \mathbf{X}\hat{\beta})^\top \hat{\Sigma}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\beta})$.

4.15.3 Alavanca Generalizada

O conceito da alavanca generalizada é medir a influência que o valor observado na variável de resposta y_i sobre o próprio valor ajustado \hat{y} (HOAGLIN; WELSCH, 1978; LAURENT; COOK, 1992; WEI et al., 1998).

Seja $\mu = \mathbf{X}\beta$ o valor esperado de \mathbf{Y} . Então, com base na alavanca generalizada para Wei et al. (1998), a matriz de alavanca generalizada, $GL(\hat{\theta}) = \partial \mathbf{Y} / \partial \hat{\mathbf{Y}}$, em que $\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta}$ e $\hat{\theta}$ estimador ML de θ . A alavanca generalizada é dada por

$$GL(\theta) = \mathbf{D}_\theta (-\mathbf{L}(\theta))^{-1} \mathbf{L}_{\theta\mathbf{Y}},$$

sendo $\mathbf{D}_\theta = \partial \mu / \partial \theta^\top = (\mathbf{X}, \mathbf{0})$, e $\mathbf{L}_{\theta\mathbf{Y}} = \partial^2 \mathcal{L}(\theta) / \partial \theta \partial \mathbf{Y}^\top = (L_{\beta\mathbf{Y}}^\top, L_{\phi\mathbf{Y}}^\top)^\top$ com

$$L_{\beta\mathbf{Y}} = \partial^2 \mathcal{L}(\theta) / \partial \beta \partial \mathbf{Y}^\top = w(\delta) \mathbf{X} \Sigma^{-1} - 2 \left(\frac{\eta}{1+n\eta} \right) w(\delta)^2 \mathbf{X}^\top \Sigma^{-1} \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^\top \Sigma^{-1},$$

$L_{\phi\mathbf{Y}} = \partial^2 \mathcal{L}(\theta) / \partial \phi \partial \mathbf{Y}^\top$ com elementos

$$L_{\phi_j\mathbf{Y}} = \partial^2 \mathcal{L}(\theta) / \partial \phi_j \partial \mathbf{Y}^\top = w(\delta) \boldsymbol{\varepsilon}^\top \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} - \left(\frac{\eta}{1+n\eta} \right) w(\delta)^2 \boldsymbol{\varepsilon}^\top \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^\top \Sigma^{-1},$$

para $j = 1, 2$, com $w(\delta) = \left(\frac{1+n\eta}{\eta} \right) c(\eta) q^{-1}$, $c(\eta) = \frac{\eta}{1-2\eta}$ e $q = (1 + c(\eta)\delta)$.

Depois de alguma álgebra, tem-se $LG(\theta) = \mathbf{GL}_1 + \mathbf{GL}_2$, em que

$$\mathbf{GL}_1 = \mathbf{X} (\mathbf{L}_{\beta\beta} - \mathbf{L}_{\beta\phi} \mathbf{L}_{\phi\phi}^{-1} \mathbf{L}_{\phi\beta})^{-1} (-\mathbf{L}_{\beta\mathbf{Y}}) \text{ e}$$

$$\mathbf{GL}_2 = \mathbf{X} (\mathbf{L}_{\beta\beta} - \mathbf{L}_{\beta\phi} \mathbf{L}_{\phi\phi}^{-1} \mathbf{L}_{\phi\beta})^{-1} (\mathbf{L}_{\beta\phi} \mathbf{L}_{\phi\phi}^{-1} \mathbf{L}_{\phi\mathbf{Y}}),$$

de tal modo que,

$$\mathbf{L}_{\beta\beta} = \frac{\partial^2 \mathcal{L}(\theta)}{\partial \beta \partial \beta^\top} = 2w(\delta)^2 \left(\frac{\eta}{1+n\eta} \right) (\mathbf{X}^\top \Sigma^{-1} \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^\top \Sigma^{-1} \mathbf{X}) - w(\delta) \mathbf{X}^\top \Sigma^{-1} \mathbf{X},$$

$$\mathbf{L}_{\beta\phi} = \frac{\partial^2 \mathcal{L}(\theta)}{\partial \beta \partial \phi^\top}, \text{ com elementos}$$

$$\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta})}{\partial \beta \partial \phi_l} = w(\delta)^2 \left(\frac{\eta}{1+n\eta} \right) \left(\boldsymbol{\varepsilon}^T \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon} \right) \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon} - w(\delta) \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}, \quad l = 1, 2,$$

$\mathbf{L}_{\phi\phi} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta})}{\partial \phi \partial \phi^T}$, com elementos

$$\begin{aligned} \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta})}{\partial \phi_l \partial \phi_j} &= -\frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}^{-1} \left(\frac{\partial^2 \boldsymbol{\Sigma}}{\partial \phi_l \partial \phi_j} - \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \right) \right] \\ &- \frac{1}{2} \left[w(\delta)^2 \left(\frac{\eta}{1+n\eta} \right) \left(\boldsymbol{\varepsilon}^T \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon} \right) \left(\boldsymbol{\varepsilon}^T \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon} \right) \right] \\ &+ \frac{1}{2} w(\delta) \boldsymbol{\varepsilon}^T \boldsymbol{\Sigma}^{-1} \left(\frac{\partial^2 \boldsymbol{\Sigma}}{\partial \phi_l \partial \phi_j} - 2 \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \right) \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}, \text{ para } j, l = 1, 2. \end{aligned}$$

Os elementos diagonais GL_{ii} para $i = 1, \dots, n$, da matriz $\mathbf{LG}(\hat{\boldsymbol{\theta}})$ são usados como ferramenta de diagnóstico de influência no vetor $\hat{\mathbf{Y}}$. A i -ésima resposta é potencialmente influente se $GL_{ii} > \overline{GL} + 2sd(GL)$, em que $\overline{GL} = \frac{1}{n} \sum_{i=1}^n GL_{ii}$ e $sd(GL)$ é o desvio padrão de GL_{11}, \dots, GL_{nn} (DE BASTIANI et al., 2015).

Com base na proposta de Wei et al. (1998), bem como em Zhu e Lee (2001), a matriz de alavanca generalizada para modelos com dados completos é definida pela equação (4.23):

$$\mathbf{LG}_{\mathbf{Q}}(\boldsymbol{\theta}) = \mathbf{D}_{\boldsymbol{\theta}}(-\ddot{\mathbf{Q}})^{-1} \mathbf{Q}_{\boldsymbol{\theta}\mathbf{Y}}, \quad (4.23)$$

sendo,

$$\ddot{\mathbf{Q}} = \ddot{\mathbf{Q}}(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}) = \partial^2 \mathbf{Q}(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T,$$

$$\mathbf{Q}_{\boldsymbol{\theta}\mathbf{Y}} = \partial^2 \mathbf{Q}(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}) / \partial \boldsymbol{\theta} \partial \mathbf{Y}^T |_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = (\mathbf{Q}_{\beta\mathbf{Y}}^T, \mathbf{Q}_{\phi\mathbf{Y}}^T)^T, \text{ com}$$

$$\mathbf{Q}_{\beta\mathbf{Y}} = \partial^2 \mathbf{Q}(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}) / \partial \beta \partial \mathbf{Y}^T |_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = w(\hat{\delta}) \mathbf{X}^T \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{e}$$

$$\mathbf{Q}_{\phi\mathbf{Y}} = \partial^2 \mathbf{Q}(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}) / \partial \phi \partial \mathbf{Y}^T, \text{ com elementos,}$$

$$\mathbf{Q}_{\phi_j \mathbf{Y}} = \partial^2 \mathbf{Q}(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}) / \partial \phi_j \partial \mathbf{Y}^T |_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = w(\hat{\delta}) \boldsymbol{\varepsilon}^T \hat{\boldsymbol{\Sigma}}^{-1} \frac{\partial \hat{\boldsymbol{\Sigma}}}{\partial \phi_j} \hat{\boldsymbol{\Sigma}}^{-1},$$

para $j = 1, 2$, com $w(\hat{\delta}) = \left(\frac{1+n\eta}{\eta} \right) c(\eta) \hat{q}^{-1}$, $c(\eta) = \frac{\eta}{1-2\eta}$ e $\hat{q} = (1 + c(\eta) \hat{\delta})$.

4.16 Aplicação ao conjunto de dados reais

4.16.1 Dados experimentais

O conjunto de dados foi coletado no ano-safra de 2014/2015, em uma área agrícola comercial de 127,18 ha, situada nas proximidades da cidade de Cascavel, região oeste do Paraná, Brasil. As coordenadas de latitude e longitude da área são aproximadamente 24,95° Sul e 53,57° Oeste, com altitude média de 650 m. O solo da área é classificado como Latossolo Vermelho Distroférico, com textura argilosa (EMBRAPA, 2011).

A variável resposta é a produtividade da soja (Prod) [t ha⁻¹] com teores químicos do

solo considerados como variáveis explicativas, fósforo [P] (mg dm^{-3}), potássio [K] (cmolc dm^{-3}), potencial hidrogeniônico [pH] e matéria orgânica [OM] (g dm^{-3}). O modelo espacial linear para a produtividade da soja na localização s_i é dado por $\text{Prod}(s_i) = \beta_1 + \beta_2 P(s_i) + \beta_3 K(s_i) + \beta_4 \text{pH}(s_i) + \beta_5 \text{OM}(s_i) + e(s_i)$, $i = 1, \dots, 78$.

A Tabela 7 mostra uma análise descritiva das variáveis em estudo. Pode-se perceber que a produção mínima foi $1,874 \text{ t ha}^{-1}$, a máxima $3,176 \text{ t ha}^{-1}$ e a produtividade média da soja foi $2,374 \text{ t ha}^{-1}$, um valor inferior à média da produtividade nacional da soja de $3,03 \text{ t ha}^{-1}$, para o ano da colheita 2014/2015 (IBGE, 2015). Estas informações são as primeiras análises preliminares que servem para identificar e compreender o conjunto de dados.

Tabela 7 Estatística descritiva para as variáveis produtividade (Prod), Fósforo (P), Potássio (K), Potencial hidrogeniônico (pH) e Matéria orgânica (OM)

Variáveis	Prod	P	K	pH	OM
n	78	78	78	78	78
Média	2,374	19,19	0,313	4,819	50,63
Mínimo	1,874	3,400	0,100	4,200	38,62
Máximo	3,176	58,60	0,670	6,100	66,37
Mediana	2,326	16,90	0,285	4,750	50,81

n : número de observações

O boxplot da produtividade da soja (Figura 12(a)) apresenta um *outlier*, que corresponde à observação #33, dado como o valor máximo dos dados com produtividade igual a $3,176 \text{ t ha}^{-1}$ (Tabela 7). De acordo com o postplot da produtividade da soja dado na Figura 12(b), a observação #33 é cercada por observações com produtividade de soja inferior a $2,55 \text{ t ha}^{-1}$.

Nos boxplots das variáveis explicativas (não mostradas), foram observados três *outliers* para a variável P, as observações #62, #68 e #69 com respectivos valores de 50,3, 52,5 e 58,6 mg dm^{-3} . Para a variável pH, o gráfico destaca as observações #56, #55, #14 e #4, com os respectivos valores de 5,8, 6,0, 6,0, e 6,1 e para a variável OM apenas um *outlier* foi detectado a observação #71 com $63,37 \text{ g dm}^{-3}$. Na análise dos semivariogramas direcionais dados na Figura 12(c), para direções 0° , 45° , 90° e 135° , é indicado que é razoável assumir isotropia.

A Tabela 8 apresenta as estimativas dos parâmetros para o modelo linear t-Student reparametrizado considerando diferentes valores para o parâmetro η (considerado como fixo) e diferentes valores de κ (considerado como fixo) da família Matérn dos modelos geoestatísticos e erros padrão assintóticos (entre parênteses). O parâmetro ϕ_3 , foi fixado em 0,3 obtido a partir da análise prévia e considerando-se o método dos mínimos quadrados ordinários.

O modelo escolhido é $\hat{P}_{rod}(s_i) = 1,9712 + 0,0030P(s_i) - 0,0997K(s_i) + 0,0379\text{pH}(s_i) + 0,0040\text{OM}(s_i)$, com matriz de covariância estimada dada por $\hat{\Sigma} = \hat{\phi}_1 \mathbf{I}_{78} + \hat{\phi}_2 \mathbf{R}(0,3)$, sendo $\hat{\phi}_1 = 0,0585$ e $\hat{\phi}_2 = 0,0857$, em que os elementos da matriz de correlação \mathbf{R} determinado a partir da família de modelos Matérn com $\kappa = 0,5$ e $\eta = 0,25$ foram selecionados de acordo com os critérios de Validação Cruzada e Traço da matriz de covariância assintótica. Fagundes et al. (2018b) apresentaram estimativas de parâmetros semelhantes aos valores obtidos e analisaram

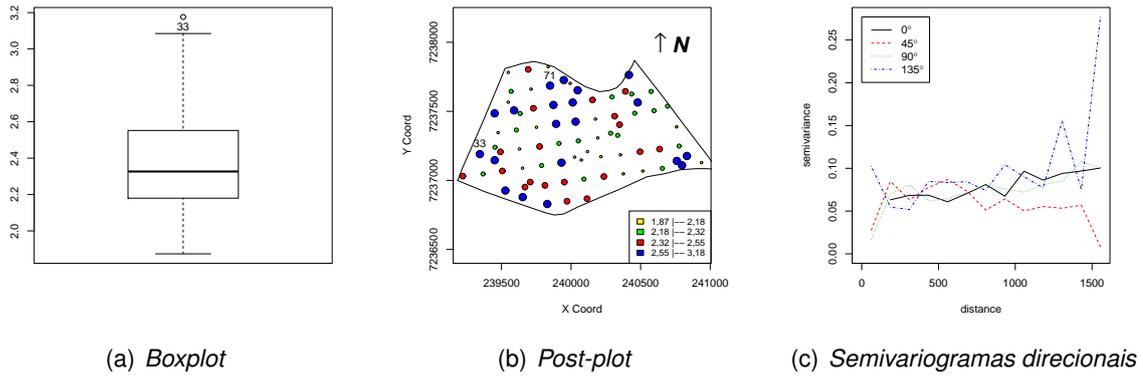


Figura 12 (a) *Boxplot*, (b) *Post-plot* e (c) *semivariogramas direcionais* para o conjunto de dados de produtividade da soja ($t\ ha^{-1}$) no ano safra 2014/2015.

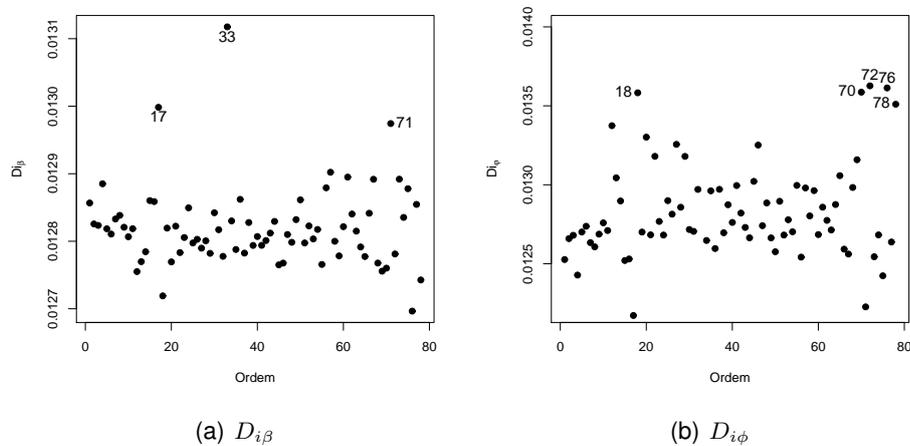


Figura 13 Gráficos de diagnóstico de influência global $D_{i\beta}$ e $D_{i\phi}$

o mesmo conjunto de dados considerando um modelo espacial linear com distribuição Slash reparametrizada.

A análise de diagnóstico para a variável resposta, dada nas Figuras 13 e 14, as observações #18, #33 e #76 estão sempre em destaque para a influência global.

As Figuras 15 e 16 mostram os gráficos de influência local, as observações #33 e #71 foram consideradas pontos influentes. A observação #33 é a mesma detectada na Figura 12 (a) pelo boxplot.

Alavanca generalizada detectou as observações #62, #68 e #69 para P, #4, #14, #55 e #56 para pH e a observação #71 para OM. As mesmas dos boxplots das variáveis explicativas.

Os QQ-plots dos resíduos estão apresentados na Figura 18. Nota-se que quando a observação #33 é retirada, um melhor ajuste foi detectado na parte superior do QQ-plot (Figura 18 (b)), com a maioria dos pontos mais próximos da reta $x = y$. Todos os pontos pertencem ao intervalo de confiança em ambos os gráficos, evidenciando que os dados seguem a distribuição assumida. Segundo Lange et al. (1989), valores que se afastam consideravelmente da reta indicam casos discrepantes, logo podemos observar que quando retirou-se a observação #33,

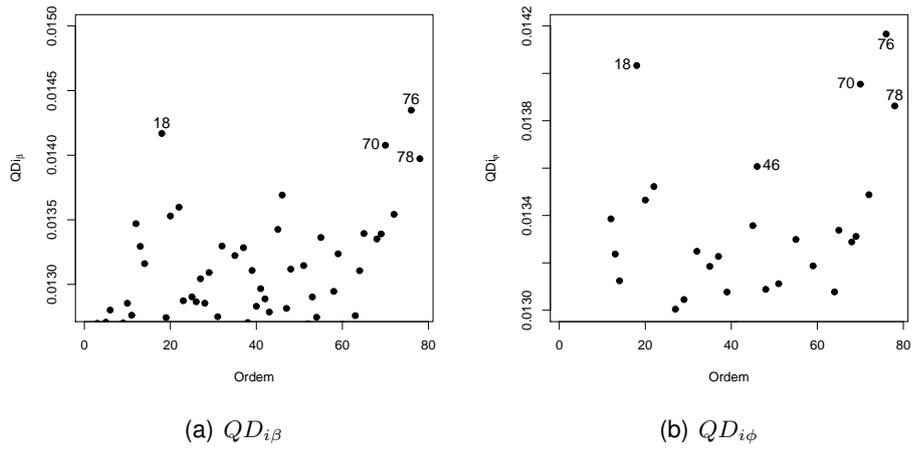


Figura 14 Gráficos de diagnóstico de influência global $QD_{i\beta}$ e $QD_{i\phi}$

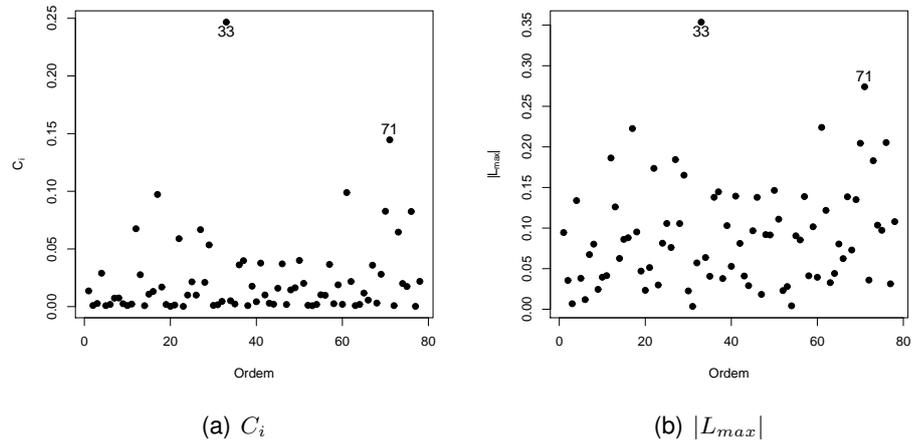


Figura 15 Gráficos de diagnóstico de influência local C_i e $|L_{max}|$

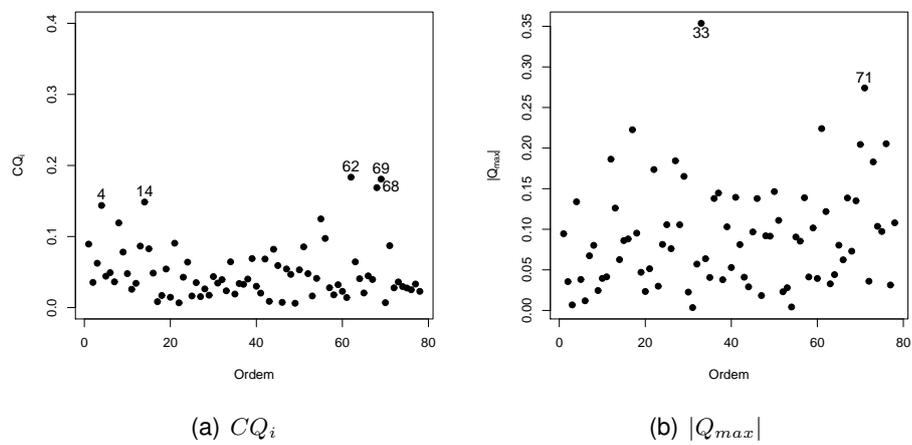


Figura 16 Gráficos de diagnóstico de influência local CQ_i e $|Q_{max}|$

Tabela 8 Estimativas de parâmetros usando o algoritmo iterativo e erros padrão assintóticos entre parênteses

κ	η	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\phi}_1$	$\hat{\phi}_2$
0,5	0,05	1,9712 (0,4227)	0,0030 (0,0028)	-0,0997 (0,2269)	0,0379 (0,0752)	0,0040 (0,0055)	0,0330 (0,0428)	0,0482 (0,0491)
	0,10	1,9712 (0,4217)	0,0030 (0,0028)	-0,0997 (0,2263)	0,0379 (0,0750)	0,0040 (0,0054)	0,0368 (0,0493)	0,0540 (0,0576)
	0,25	1,9712 (0,4205)	0,0030 (0,0028)	-0,0997 (0,2257)	0,0379 (0,0748)	0,0040 (0,0054)	0,0585 (0,0849)	0,0857 (0,1032)
	0,45	1,9712 (0,4231)	0,0030 (0,0029)	-0,0997 (0,2271)	0,0379 (0,0753)	0,0040 (0,0055)	0,2960 (0,4693)	0,4336 (0,5915)
	1,0	0,05	2,0125 (0,5684)	0,0037 (0,0039)	-0,1263 (0,3156)	0,0425 (0,1043)	0,0025 (0,0073)	0,1362 (0,0971)
1,0	0,10	2,0125 (0,6664)	0,0037 (0,0046)	-0,1263 (0,3700)	0,0425 (0,1223)	0,0025 (0,0085)	0,2102 (0,1644)	0,0174 (0,1309)
	0,25	2,0125 (0,8984)	0,0037 (0,0062)	-0,1263 (0,4988)	0,0425 (0,1649)	0,0025 (0,0115)	0,6101 (0,5852)	0,0504 (0,3813)
	0,45	2,0125 (1,1362)	0,0037 (0,0078)	-0,1263 (0,6308)	0,0425 (0,2085)	0,0025 (0,0146)	4,8771 (5,6256)	0,4032 (3,0600)
	1,5	0,05	2,0118 (1,0406)	0,0037 (0,0072)	-0,1258 (0,5774)	0,0425 (0,1909)	0,0025 (0,0133)	0,3658 (0,8091)
1,5	0,10	2,0118 (1,3459)	0,0037 (0,0093)	-0,1258 (0,7468)	0,0425 (0,2469)	0,0025 (0,0173)	0,6869 (1,5368)	0,2411 (1,5100)
	0,25	2,0118 (1,9992)	0,0037 (0,0137)	-0,1258 (1,1093)	0,0425 (0,3667)	0,0025 (0,0256)	2,4208 (5,5842)	0,8496 (5,3471)
	0,45	2,0118 (2,6287)	0,0037 (0,0181)	-0,1258 (1,4586)	0,0425 (0,4822)	0,0025 (0,0337)	20,9157 (50,0842)	7,3408 (46,4495)

κ : valor do parâmetro de forma do modelo da família Matérn ; η : parâmetro de forma da distribuição t-Student reparametrizada; $\hat{\beta}_i$: parâmetros estimados, $\hat{\phi}_i$: estimativas de parâmetros da matriz de covariância, $i = 1, 2$ em que $\phi_1 = \tau_1$; $\phi_2 = \tau_2/\tau_3^{2\kappa}$ sendo τ_1 o efeito pepita e τ_2 a contribuição (URIBE-OPAZO et al., 2021)

os demais pontos se ajustaram melhor à reta. Assim, há indícios de que esta observação é um ponto influente.

A Figura 19 apresenta os mapas considerando todas as observações e excluindo a observação #33, assim, fica evidente a mudança substancial no mapa em que a observação #33 está localizada, veja também a Figura 12(b). Esta observação é um *outlier* e também uma observação influente para a obtenção dos valores previstos.

Para mensurar a similaridade entre o mapa de referência (todos os pontos) e o mapa modelo (#33), determinou-se a exatidão global sendo $EG = 0,73$, segundo Anderson (1976), indicando que os mapas não são semelhantes. Segundo a classificação de Krippendorff (2018) os índices Tau ($T = 0,64$) e Kapa ($K = 0,59$) são considerados com baixa exatidão. Dessa forma, é possível verificar que existe diferença na classificação dos mapas temáticos em estudo, com e sem o ponto #33 detectado como influente. As fórmulas dos índices EG , T e K são apresentadas por De Bastiani et al. (2012).

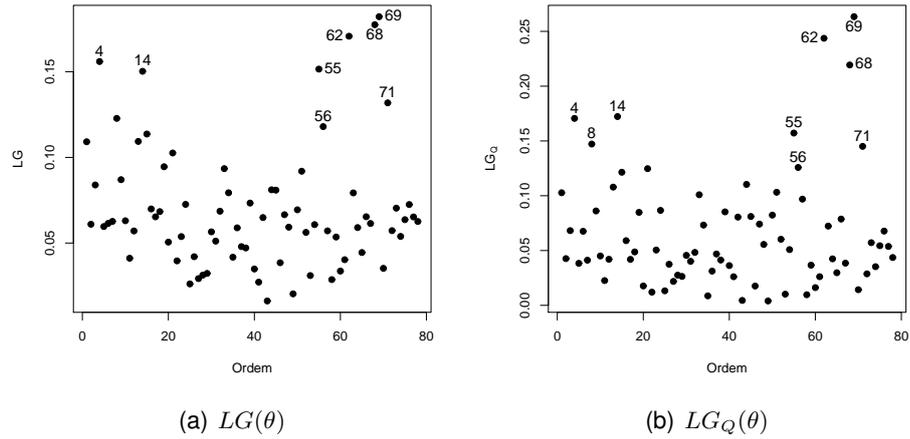


Figura 17 Gráficos da alavanca generalizada considerando (a) a função logarítmica de verossimilhança e (b) a Q-function

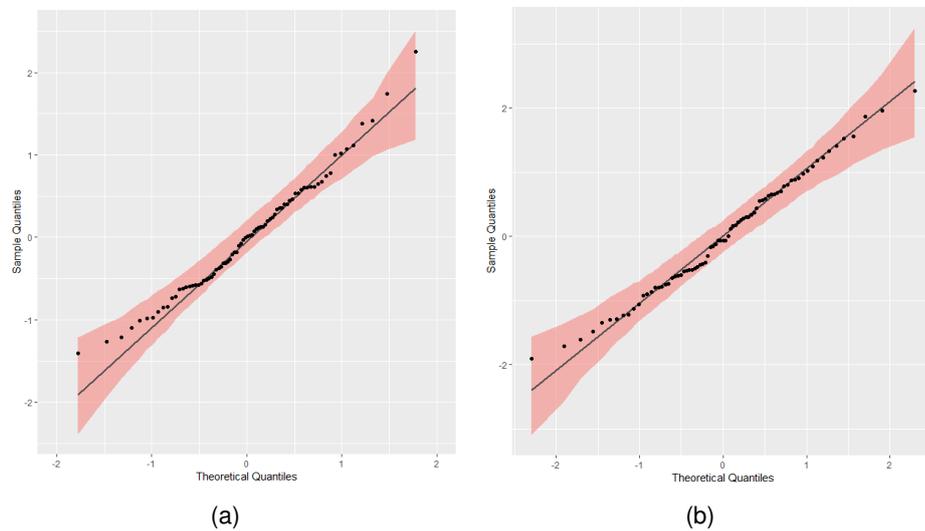


Figura 18 QQ-plot dos resíduos (a) todos os pontos e (b) sem o ponto #33

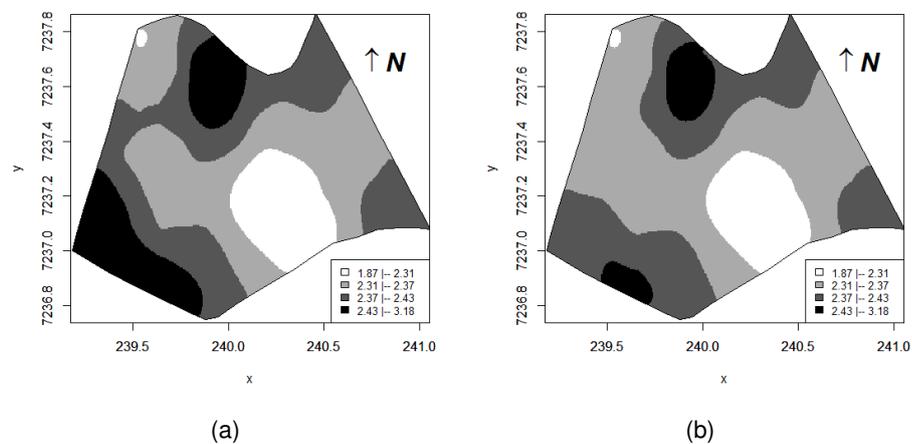


Figura 19 Mapas da produtividade da soja ($t\ ha^{-1}$) no ano safra 2014/2015, para o conjunto de dados (a) com todas as observações e (b) exclusão da observação #33

4.16.2 Aplicação a dados de Aquífero

A segunda aplicação apresenta os dados de 93 poços em um único aquífero próximo ao Vale Saratoga, Wyoming, fornecidos em (JONES, 1989). Cada valor observado da variável resposta $Y(s_i)$ corresponde à altura da água (em metros) acima do nível do mar de cada poço de água. As coordenadas $s_i = (Coordx_i, Coordy_i)$ também foram consideradas como variáveis explicativas, veja De Bastiani et al. (2015) para detalhes. O modelo espacial linear t-Student reparametrizado considerado para a altura do poço de água é dado por $Y(s_i) = \beta_0 + \beta_1 x_i + \beta_2 y_i + \varepsilon(s_i)$, em que $\varepsilon(s_i) \sim \tau_n(\mathbf{0}, \Sigma, \eta)$. Jones (1989) assume isotropia.

Na Tabela 9, as estimativas dos parâmetros são apresentadas, para diferentes valores de η e κ . De acordo com os critérios $(T_r(\eta))$ e $(VC(\eta))$ (Tabela 10), os valores selecionados foram $\eta = 0, 10$ e $\kappa = 1, 5$. Assim, o modelo espacial linear reparametrizado t-Student ajustado para a variável resposta é dado por $\hat{Y}(s_i) = 2274,33 - 0,723781x_i - 3,372901y_i$, com matriz de covariância $\hat{\Sigma} = 2379,264 + 5,832065R(7, 399)$. O parâmetro ϕ_3 foi fixado em 7,399, obtido em análises anteriores e considerando o método dos mínimos quadrados.

Tabela 9 Estimativas dos parâmetros e erros padrão entre parênteses

κ	η	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\phi}_1$	$\hat{\phi}_2$
1, 5	0, 10	2274, 33429	-0, 723781	-3, 372901	2379, 264	5, 832065
		(52, 45971)	(1, 722392)	(0, 909642)	(1213, 042)	(2, 98202)
	0, 25	2274, 33429	-0, 723781	-3, 3729015	3861, 232	9, 464674
		(52, 86439)	(1, 735679)	(0, 9166591)	(2906, 017)	(7, 132724)
		2274, 33429	-0, 723781	-3, 3729015	19600, 64	48, 04519
		(53, 27608)	(1, 749196)	(0, 9237977)	(19354, 5)	(47, 47856)
2, 0	0, 10	2290, 43565	-0, 9780093	-3, 401286	5061, 917	1, 6798836
		(75, 40705)	(2, 3549273)	(1, 325684)	(2634, 368)	(0, 8757784)
	0, 25	2290, 43565	-0, 9780093	-3, 401286	8208, 882	2, 724257
		(75, 96125)	(2, 3722348)	(1, 335427)	(6237, 405)	(2, 071678)
		2290, 43565	-0, 9780093	-3, 401286	41767, 86	13, 86138
		(76, 64223)	(2, 3935014)	(1, 347399)	(41473, 83)	(13, 77038)
2, 5	0, 10	2307, 834	-1, 361453	-3, 370272	10972, 913	0, 4933241
		(108, 329)	(3, 237912)	(1, 918037)	(5793, 428)	(0, 2607125)
	0, 25	2307, 8337	-1, 361453	-3, 370272	17800, 32	0, 800273
		(109, 1424)	(3, 262224)	(1, 932438)	(13617, 79)	(0, 6125132)
		2307, 8337	-1, 361453	-3, 370272	90567, 91	4, 071784
		(110, 1193)	(3, 291424)	(1, 949736)	(90291)	(4, 060427)

κ : valor do parâmetro de forma do modelo da família Matérn; η : parâmetro de forma da distribuição t-Student reparametrizada; $\hat{\beta}_i$: parâmetros estimados, $\hat{\phi}_i$: estimativas de parâmetros da matriz de covariância, $i = 1, 2$ em que $\phi_1 = \tau_1$; $\phi_2 = \tau_2/\tau_3^{2\kappa}$ sendo τ_1 o efeito pepita e τ_2 a contribuição (URIBE-OPAZO et al., 2021)

As Figuras 20 e 21 apresentam os gráficos de diagnósticos de influência global com base na verossimilhança ($D_{i\beta}$ e $D_{i\phi}$) e com base na Q-função ($QD_{i\beta}$ and $QD_{i\phi}$), destacando as observações #49, #50, #69, #80 e #81. Por outro lado, os gráficos de diagnósticos de influência local, apresentados nas Figuras 22 e 23, são baseados na verossimilhança e Q-função, respectivamente, e as observações destacadas são #1, #13, #15, #38, #40, #48, #61, #92 e

Tabela 10 Valores do traço da matriz de covariância assintótica ($T_r(\eta)$) e validação cruzada ($VC(\eta)$), para o aquífero Vale de Saratoga

κ	η	$T_r(\eta)$	$VC(\eta)$
	0, 10	77439,31	2519,313
1, 5	0, 25	78638, 67	2520, 751
	0, 45	79868, 26	2522, 271
2, 0	0, 10	174345, 6	13917, 8
	0, 25	176917, 8	15987, 03
	0, 45	180104	19899, 28
2, 5	0, 10	387215, 9	16435, 64
	0, 25	393052, 4	16457, 47
	0, 45	400120, 5	16484, 14

κ : valor do parâmetro de forma do modelo da família Matérn; η : parâmetro de forma da distribuição t-Student reparametrizada; destacaram-se os valores mais baixos de $T_r(\eta)$ e $VC(\eta)$ em negrito.

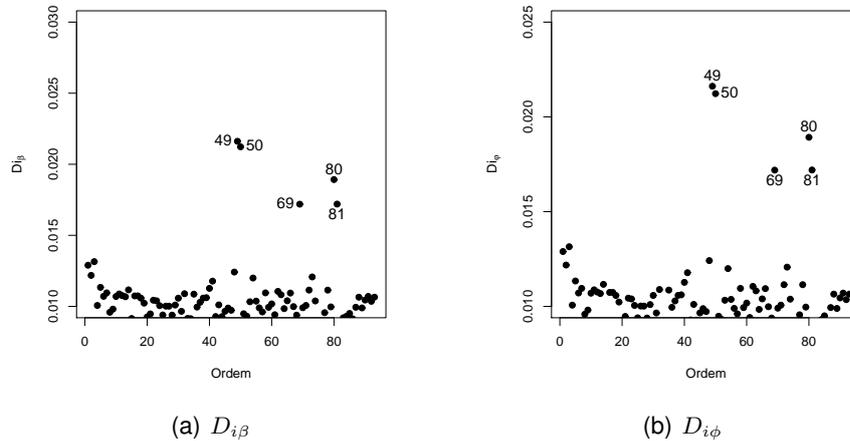


Figura 20 Gráficos de diagnóstico de influência global $D_{i\beta}$ e $D_{i\phi}$ para os dados do aquífero Vale de Saratoga

#93. A Figura 24 apresenta o gráfico de alavanca generalizada.

Alguns outros autores já analisaram este conjunto de dados, e a Tabela 11 dá um resumo dos resultados. Observa-se que a análise de influência global apresentada por Christensen et al. (1992) e Fagundes et al. (2018a) também identificou as observações #49, #50 e #69 como potencial influente, enquanto Uribe-Opazo et al. (2012) identificaram a observação #49 entre outros. De Bastiani et al. (2015) identificaram a observação #1 como potencial influente na análise de influência local, e Fagundes et al. (2018a) também identificaram observações #80 e #81.

Na Figura 25, apresentam-se os gráficos QQ-plots, em que a Figura 25 (a) contém todos os pontos observados. Verifica-se que, em sua maioria, os pontos estão próximos à reta xy , e que somente na parte superior, os pontos estão afastados da reta e também fora do intervalo de confiança, observa-se na Figura 25 (b) e (c), quando retiram-se os pontos considerados influentes, os pontos superiores nestes gráficos, há uma leve aproximação do intervalo de

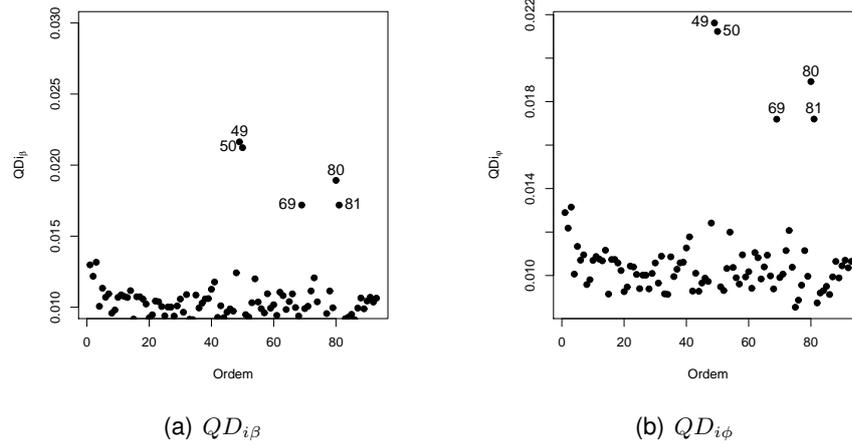


Figura 21 Gráficos de diagnóstico de influência global $QD_{i\beta}$ e $QD_{i\phi}$ para os dados do aquífero Vale de Saratoga

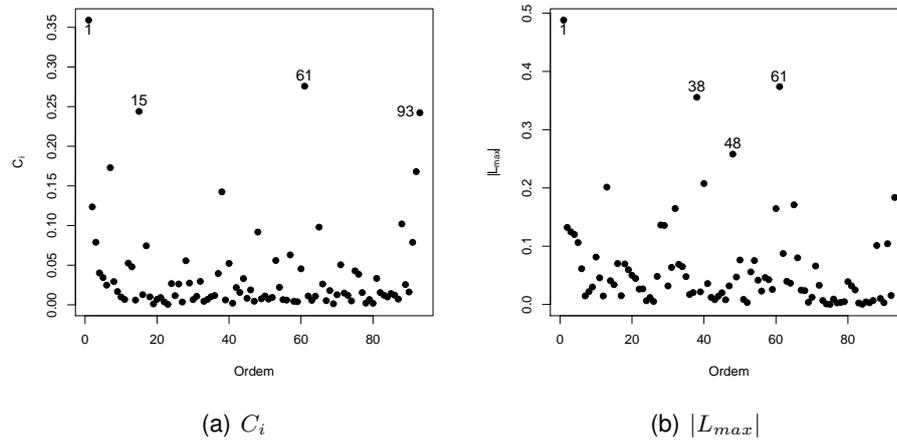


Figura 22 Gráficos de diagnóstico de influência local C_i e $|L_{max}|$ para os dados do aquífero Vale de Saratoga

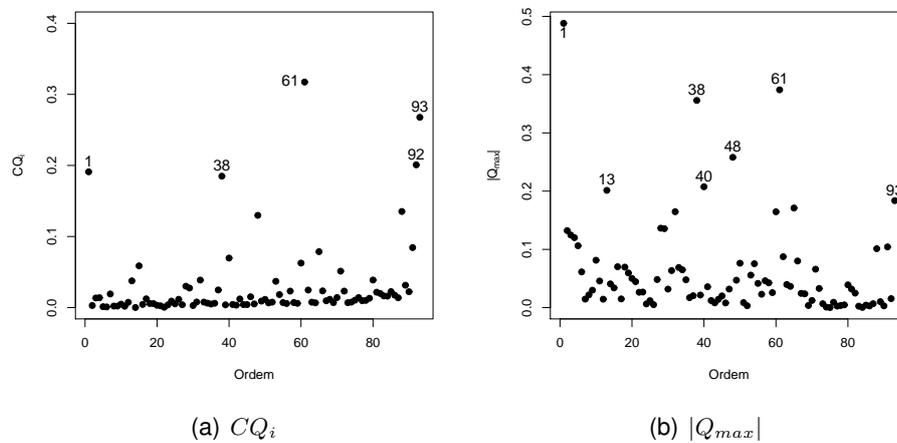


Figura 23 Gráficos de diagnóstico de influência local CQ_i e $|Q_{max}|$ para os dados do aquífero Vale de Saratoga

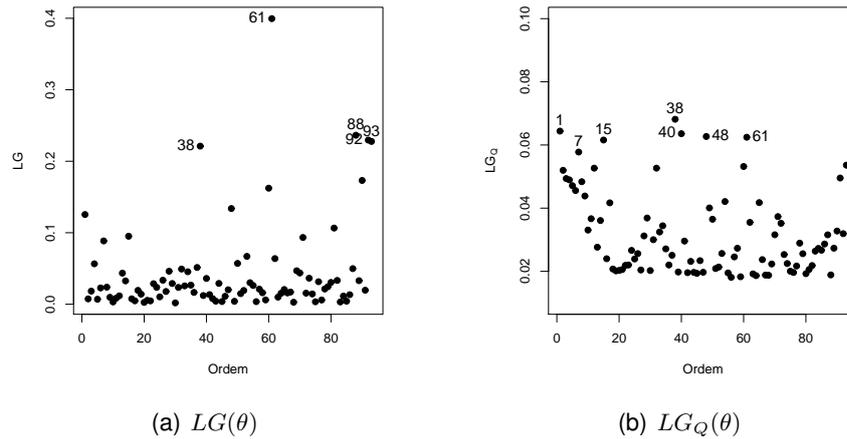


Figura 24 Gráfico de alavanca generalizada para os dados do aquífero Vale de Saratoga

Tabela 11 Potenciais observações influentes identificadas por diferentes autores usando diagnósticos globais e locais influentes sob diferentes abordagens.

	Autores	Observações
Global	Christensen et al. (1992) (Normal)	#1, #2, #49, #50, #69
	Fagundes et al. (2018) (Slash)	#49, #50, #69
	Uribe-Opazo et al. (2012) (Normal)	#1, #49, #64
	Reparameterized t-Student (loglikelihood)	#49, #50, #69, #80, #81
	Reparameterized t-Student (Q-function)	#49, #50, #69, #80, #81
Local	De Bastiani et al. (2015) (t-Student)	#1, #22, #78, #79, #81, #85
	Fagundes et al. (2018) (Slash)	#79, #80, #81
	Reparameterized t-Student (loglikelihood)	#1, #15, #38, #48, #61, #93
	Reparameterized t-Student (Q-function)	#1, #13, #38, #40, #48, #61, #92, #93,

confiança, quando são comparados com a Figura 25 (a).

Os dados foram analisados novamente após a remoção das observações #49 e #50, relacionados ao diagnóstico de influência global. Os parâmetros fixos escolhidos são $\kappa = 1,5$ e $\eta = 0,10$, com estimativas de parâmetros $\beta_0 = 2272,43$, $\beta_1 = -0,67369$, $\beta_2 = -3,35056$, com matriz de covariância $\hat{\Sigma} = 2332,23 + 2320,477R(7,399)$. Além disso, os dados foram analisados após a remoção das observações #01, e #61, relacionados à análise de diagnóstico de influência local. O modelo escolhido foi $\kappa = 1,5$ e $\eta = 0,10$, $\beta_0 = 2261,853$, $\beta_1 = 0,11794$, $\beta_2 = -3,69652$, com matriz de covariância $\hat{\Sigma} = 2321,09 + 2304,34R(7,399)$. Sendo que os modelos escolhidos, quanto aos valores fixados de κ e η foram obtidos através dos critérios Traço e Validação Cruzada. A Figura 27 apresenta os mapas obtidos a partir dos modelos ajustados.

Verificou-se a similaridade entre o mapa referência (a) contendo todos os pontos e os mapas modelos (b) sem as observações #49 e #50 e (c) sem as observações #01 e #61. Calcularam-se os índices de comparação Exatidão Global e Kapa, obtendo-se $EG = 0,99$ e $K = 0,98$ para a comparação entre os mapas (a) e (b). Quando comparados os mapas (a) e (c) foram obtidos $EG = 0,98$ e $K = 0,97$. Dessa forma, verifica-se a alta exatidão em ambas as comparações (ANDERSON, 1976).

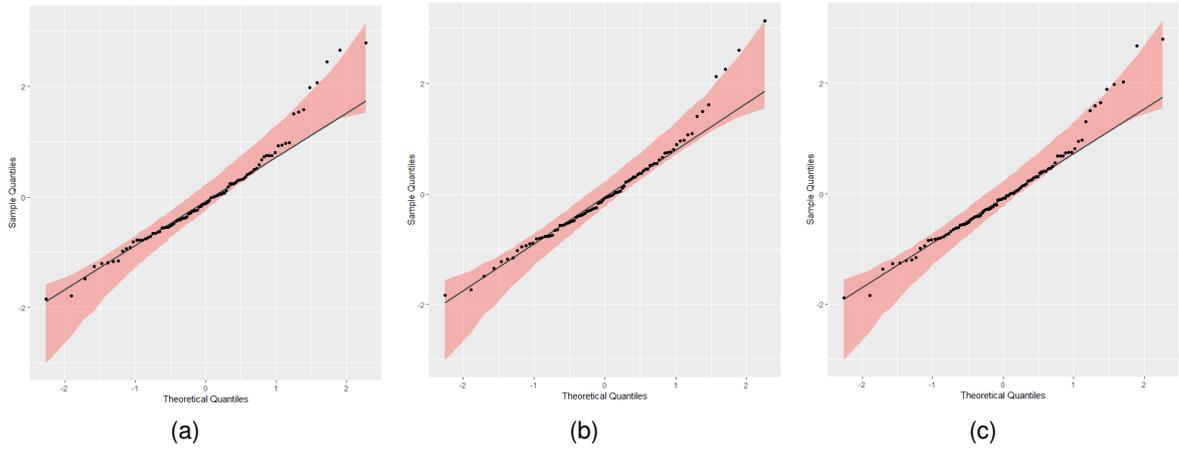


Figura 25 QQ-plots(a) com todas as observações (b) sem as observações #49 e #50 e (c) sem as observações #01 e #61

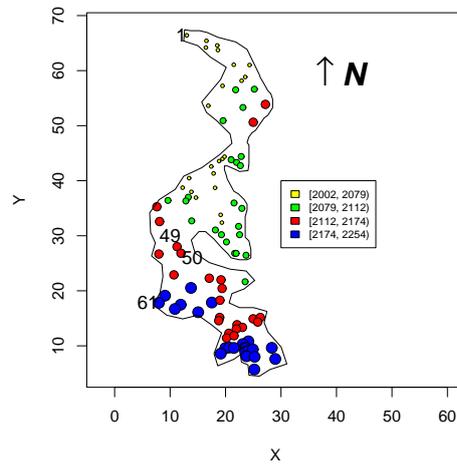


Figura 26 Localização dos pontos considerados influentes para os dados do aquífero Vale de Saratoga

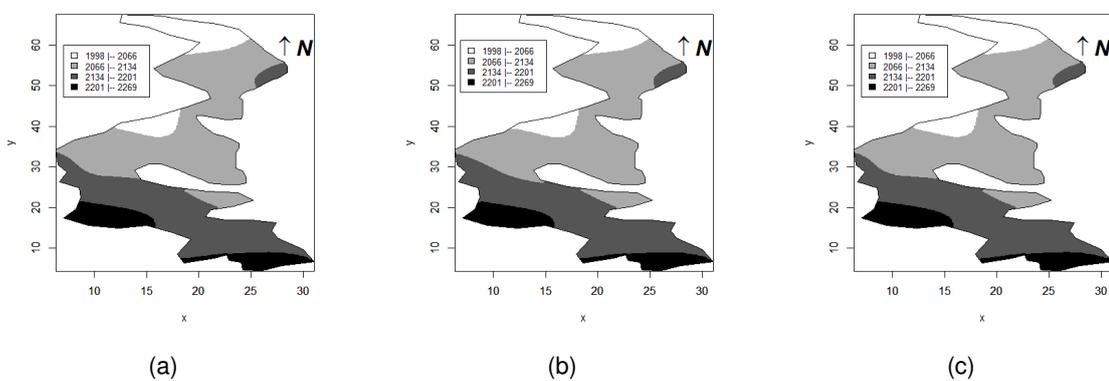


Figura 27 Mapas para os dados do aquífero Vale de Saratoga (a) com todas as observações (b) sem observações #49 e #50 e (c) sem observações #01 e #61

Logo, pode-se notar que não houve mudança relevante nos mapas na retirada dos pontos considerados influentes. Isto se deve ao fato de que estes pontos estão localizados em vizinhanças com valores similares, como se pode notar na Figura 26. Consequentemente, a retirada destes pontos não influenciou fortemente na estimação do modelo, tão pouco na construção dos mapas da Figura 27.

4.17 Conclusões

Foi apresentada a distribuição t-Student reparametrizada aplicada à geoestatística. A reparametrização permite uma comparação direta entre a matriz de covariância de distribuições multivariadas, especialmente para comparar com as distribuições normais multivariadas. O artigo também apresentou o uso em diagnósticos de influência global e local para escolher um modelo mais apropriado para ajustar os dados.

A modelagem espacial linear t-Student reparametrizada permite uma modelagem mais robusta na presença de observações influentes. É de suma importância a detecção dos pontos influentes e relacionar a sua localização espacial. Assim, pode-se chegar a conclusões de grande valia para a retirada ou não destes pontos, definidos na análise de diagnósticos. A investigação destes pontos é imprescindível na tomada de decisão, garantindo assim que as informações contidas nos mapas gerados tenham maior precisão com a realidade.

4.18 Referências

ACOSTA, J., OSORIO, F., VALLEJOS, R. Effective sample size for line transect sampling models with an application to marine macroalgae. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 21, p. 407-425, 2016.

ALMEIDA, A., LOY, A., HOFMANN, H. ggplot2 compatible quantile-quantile plots in r. **The R Journal**, v. 10, n. 2, p. 248-261, 2018.

ANDERSON, J. R. **A land use and land cover classification system for use with remote sensor data**. [S.l.]: US Government Printing Office, v. 964, 1976.

ASSUMPÇÃO, R. A. B.; URIBE-OPAZO, M. A.; GALEA, M. Analysis of local influence in geostatistics using Student's t-distribution. **Journal of Applied Statistics**, v. 41, n. 11, p. 2323-2341, 2014.

CHATTERJEE, S., HADI, A. S. **Sensitivity analysis in linear regression**. [S.l.]: John Wiley & Sons, v. 327, 2009.

COOK, R. D. Detection of influential observation in linear regression. **Technometrics**, [Taylor & Francis, Ltd., American Statistical Association, American Society for Quality], v. 19, n. 1, p. 15-18, 1977. ISSN 00401706. Disponível em: <<http://www.jstor.org/stable/1268249>>.

COOK, R. D.; WEISBERG, S. **Residuals and influence in regression**. [S.l.]: New York:

Chapman and Hall, 1982.

COOK, R. D. Assessment of local influence. **Journal of the Royal Statistical Society**. Series B (Methodological), [Royal Statistical Society, Wiley], v. 48, n. 2, p. 133-169, 1986. ISSN 00359246. Disponível em: <<http://www.jstor.org/stable/2345711>>.

CHRISTENSEN, R.; JOHNSON, W.; PEARSON, L. M. Prediction diagnostics for spatial linear models. **Biometrika**, JSTOR, p. 583-591, 1992.

DALPOSSO, G. H., URIBE-OPAZO, M. A., JOHANN, J. A., GALEA, M., BASTIANI, F. D. Gaussian spatial linear model of soybean yield using bootstrap methods. **Engenharia Agrícola**, SciELO Brasil, v. 38, n. 1, p. 110-116, 2018.

DE BASTIANI, F., URIBE-OPAZO, M., GALEA, M., CYSNEIROS, A. Case-deletion diagnostics for spatial linear mixed models. **Spatial statistics**, Elsevier, v. 28, p. 284-303, 2018.

DE BASTIANI, F., GALEA, M., CYSNEIROS, A. H. M. A., URIBE-OPAZO, M. A. Gaussian spatial linear models with repetitions: An application to soybean productivity. **Spatial Statistics**, v. 21, p. 319-335, 2017.

DE BASTIANI, F., DE AQUINO CYSNEIROS, A. H. M., URIBE-OPAZO, M. A., GALEA, M. Influence diagnostics in elliptical spatial linear models. **Test**, v. 24, n. 2, p. 322-340, 2015.

DE BASTIANI, F., URIBE-OPAZO, M. A., DALPOSSO, G. H. Comparison of maps of spatial variability of soil resistance to penetration constructed with and without covariables using a spatial linear model. **Engenharia agrícola**, v. 32, n. 2, p. 393-404, 2012.

EMBRAPA. Manual de métodos de análise de solo. **Centro Nacional de Pesquisa de Solos**, 2011.

FAGUNDES, R.; URIBE-OPAZO, M.; GALEA, M.; GUEDES, L. Spatial variability in slash linear modeling with finite second moment. *Journal of Agricultural, Biological and Environmental Statistics*, Springer, v. 23, n. 2, p. 276-296, 2018a.

FAGUNDES, R. S., URIBE-OPAZO, M. A., GUEDES, L. P. C., GALEA, M. Slash spatial linear modeling: soybean yield variability as a function of soil chemical properties. **Revista Brasileira de Ciência do Solo**, v. 42, 2018b.

GALEA, M., DE CASTRO, M. Robust inference in a linear functional model with replications using the t distribution. **Journal of Multivariate Analysis**, v. 160, p. 134-145, 2017.

GARCIA-PAPANI, F., LEIVA, V., URIBE-OPAZO, M. A., AYKROYD, R. G. Birnbaum-Saunders spatial regression models: Diagnostics and application to chemical data. **Chemometrics and Intelligent Laboratory Systems**, v. 177, p. 114-128, 2018.

GARCIA-PAPANI, F., 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**, v. 31, n. 1, p. 105-124, 2017.

HOAGLIN, D. C., WELSCH, R. E. The hat matrix in regression and anova. **The**

American Statistician, Taylor & Francis, v. 32, n. 1, p. 17-22, 1978. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/00031305.1978.10479237>>.

IBGE. **Instituto Brasileiro de Geografia e Estatística-IBGE**. 2015.

JONES, R. H. Fitting a stochastic partial differential equation to aquifer data. **Stochastic Hydrology and Hydraulics**, Springer, v. 3, n. 2, p. 85-96, 1989.

KANO, Y., BERKANE, M., BENTLER, P. M. Statistical inference based on pseudo-maximum likelihood estimators in elliptical populations. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 88, n. 421, p. 135-143, 1993.

KRIPPENDORFF, K. **Content analysis: An introduction to its methodology**. [S.l.]: Sage publications, 2018.

LACHOS, V. H.; MATOS, L. A.; BARBOSA, T. S.; GARAY, A. M.; DEY, D. K. Influence diagnostics in spatial models with censored response. **Environmetrics**, Wiley Online Library, v. 28, n. 7, p. e2464, 2017.

LANDIM, P. M. B. Sobre geoestatística e mapas. **Terra e Didática**, v. 2, n. 1, p. 19-33, 2006.

LANGE, K. L., LITTLE, R. J., TAYLOR, J. M. Robust statistical modeling using the t distribution. **Journal of the American Statistical Association**, v. 84, n. 408, p. 881-896, 1989.

LAURENT, R. T. S.; COOK, R. D. Leverage and superleverage in nonlinear regression. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 87, n. 420, p. 985-990, 1992.

LEIVA, V., SÁNCHEZ, L., GALEA, M., SAULO, H. Global and local diagnostic analytics for a geostatistical model based on a new approach to quantile regression. **Stochastic Environmental Research and Risk Assessment**, Springer, v. 34, n. 10, p. 1457-1471, 2020.

MAGNUS, J. R.; NEUDECKER, H. Matrix differential calculus with applications in statistics and econometrics. **Google Scholar**, p. 501-502, 1999.

MARTÍNEZ, S., GIRALDO, R., LEIVA, V. Birnbaumsaunders functional regression models for spatial data. **Stochastic Environmental Research and Risk Assessment**, Springer, v. 33, n. 10, p.1765-1780, 2019.

ORDOÑEZ, J. A., PRATES, M. O., MATOS, L. A., LACHOS, V. H. Objective bayesian analysis for spatial student-t regression models. **arXiv preprint arXiv:2004.04341**, 2020.

OSORIO, F., PAULA, G. A., GALEA, M. Assessment of local influence in elliptical linear models with longitudinal structure. **Computational Statistics & Data Analysis**, v. 51, n. 9, p. 4354-4368, 2007.

PAN, J.; FEI, Y.; FOSTER, P. Case-deletion diagnostics for linear mixed models. **Technometrics**, Taylor & Francis, v. 56, n. 3, p. 269-281, 2014. Disponível em: <<https://doi.org/10.1080/00401706.2013.810173>>.

PRADO, N. V. d.; URIBE-OPAZO, M. A.; GALEA, M.; ASSUMPÇÃO, R. A. Influência local em um modelo espacial linear da produtividade da soja utilizando distribuição t-student. **Engenharia Agrícola**, SciELO Brasil, v. 33, n. 5, p. 1003-1016, 2013.

RICHETTI, J., URIBE-OPAZO, M. A., BASTIANI, F. D., JOHANN, J. A. Técnicas para detecção de pontos influentes em variáveis contínuas regionalizadas. **Engenharia Agrícola**, v. 36, n. 1, p. 152-165, 2016.

SÁNCHEZ, L.; LEIVA, V.; GALEA, M.; SAULO, H. Birnbaum-saunders quantile regression models with application to spatial data. **Mathematics**, Multidisciplinary Digital Publishing Institute, v. 8, n. 6, p. 1000, 2020.

SCHEMMER, R. C., URIBE-OPAZO, M. A., GALEA, M., ASSUMPÇÃO, R. A. Spatial variability of soybean yield through a reparameterized t-Student model. **Engenharia Agrícola**, v. 37, n. 4, p. 760-770, 2017.

STEIN, M. L. **Interpolation of spatial data: some theory for kriging**. [S.l.]: Springer Science & Business Media, 2012.

URIBE-OPAZO, M. A., BORSSOI, J. A., GALEA, M. Influence diagnostics in Gaussian spatial linear models. **Journal of Applied Statistics**, v. 39, n. 3, p. 615-630, 2012.

URIBE-OPAZO, M. A., DE BASTIANI, F., GALEA, M., SCHEMMER, R. C., ASSUMPÇÃO, R. A. B. Influence diagnostics on a reparameterized t-Student spatial linear model. **Spatial Statistics**, p. 100481, 2021.

VERBEKE, G.; MOLENBERGHS, G. **Linear mixed models for longitudinal data**. [S.l.]: Springer Science & Business Media, 2009.

WALLER, L. A., GOTWAY, C. A. **Applied spatial statistics for public health data**. John Wiley & Sons, 2004.

WEI, X., SAMARABANDU, J., DEVDHAR, R. S., SIEGEL, A. J., ACHARYA, R., BEREZNEY, R. Segregation of transcription and replication sites into higher order domains. **Science**, American Association for the Advancement of Science, v. 281, n. 5382, p. 1502-1505, 1998.

WICKHAM, H. ggplot2. **Wiley Interdisciplinary Reviews: Computational Statistics**, Wiley Online Library, v. 3, n. 2, p. 180-185, 2011.

ZELLNER, A. Bayesian and non-bayesian analysis of the regression model with multivariate student-t error terms. **Journal of the American Statistical Association**, Taylor Francis Group, v. 71, n. 354, p. 400-405, 1976.

ZHU, H.-T., LEE, S.-Y. Local influence for incomplete data models. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 63, n. 1, p.111-126, 2001.

ZHU, H., IBRAHIM, J. G., LEE, S., ZHANG, H. et al. Perturbation selection and influence measures in local influence analysis. **The Annals of Statistics**, Institute of Mathematical Statistics, v. 35, n. 6, p. 2565-2588, 2007.

4.19 ARTIGO 3: Modelo espacial linear t-Student reparametrizado com repetição

Resumo: Este trabalho tem como objetivo desenvolver técnicas de diagnósticos para dados espaciais correlacionados usando a distribuição multivariada t-student reparametrizada com repetições. Essa distribuição tem sido usada como alternativa à distribuição normal na presença de *outliers* no conjunto de dados, uma vez que tem cauda mais pesada e geralmente se ajusta melhor que a distribuição normal. A vantagem da distribuição t-student reparametrizada é o fato da forma da matriz de covariância ser a mesma da distribuição normal multivariada, quando o segundo momento é finito, o que permite uma comparação direta entre elas. As estimativas dos parâmetros são obtidas pelo algoritmo iterativo parcial linear. Técnicas de diagnóstico de influência são utilizadas para identificar a presença de observações influentes que podem interferir na estimativa dos parâmetros, seleção do modelo e valores previstos. Uma aplicação para dados agrícolas é apresentada. O conjunto de dados da produtividade da soja foi coletado ao longo de cinco anos com quatro variáveis químicas explicativas. Cada ano foi considerado uma realização do processo. Verificou-se a qualidade do ajuste do modelo e realizaram-se as análises de diagnóstico de influência. Os resultados mostraram que os métodos foram eficazes para detectar observações influentes. Extensivos cálculos são apresentados no Apêndice.

Palavras-chave: Algoritmo iterativo, análise de diagnósticos, razão de verossimilhança.

4.20 Introdução

Vários autores apresentam estudos com distribuição t-Student tais como Lange et al. (1989), Sutradhar (1993), Fiorentini et al. (2003), Kotz e Nadarajah (2004), Assumpção et al. (2014), Schemmer et al. (2017), Galea e Castro (2017) e Uribe-Opazo et al. (2021). Uma razão para se utilizar o modelo espacial linear t-Student consiste por se tratar de um modelo robusto, o qual reduz a influência de valores atípicos, também incorpora um parâmetro adicional em relação à distribuição normal, o que permite modelar os dados espaciais com alta curtose. E considerando-se a reparametrização apresentada por Lange et al. (1989), tem-se o segundo momento finito, o qual permite uma comparação direta com a distribuição normal.

A contribuição principal deste artigo é apresentar um estudo do modelo espacial linear t-Student reparametrizado com repetição, por ser um diferencial aos demais artigos citados e estar baseado no artigo de De Bastiani et al. (2017), o qual apresentaram o modelo espacial linear para distribuição gaussiana. O artigo está organizado da seguinte forma: na seção 4.21, apresenta-se o modelo t-Student reparametrizado com repetição, descreve-se a matriz de covariância, função densidade de probabilidade e a representação estocástica, na seção 4.22 as expressões da função escore definidas para cada parâmetro em estudo; na seção 4.23, a matriz informação esperada de Fisher; na seção 4.24, apresenta-se o método de estimação de máxima verossimilhança de forma que se obtém uma equação definida para cada parâmetro estimado, que auxilia na implementação do algoritmo parcial linear apresentado na seção 4.25. Na seção 4.26 descreve-se o critério de validação cruzada e o traço para a escolha do parâmetro de forma. Na seção 4.27, apresentam-se os testes de hipóteses, em que

optou-se por trabalhar com o teste da razão de verossimilhança, pois é invariante perante a reparametrização. Na seção 4.28, descreve-se a análise de diagnósticos de influência global e local para a verificação de pontos influentes. Na seção 4.29, aplicou-se a metodologia para os dados reais. Na seção 4.30, apresentam-se as conclusões obtidas a partir da metodologia aplicada. Mais detalhes dos resultados são apresentados no Apêndice 4.31.

4.21 Modelo espacial linear com repetição

Um estudo de dependência espacial considera um processo estocástico $\{Y_i(s), s \in S \subset \mathfrak{R}^2\}$ em um espaço euclidiano bidimensional \mathfrak{R}^2 e $i = 1$ representa uma única repetição do processo. Denota-se o valor médio na localização s por:

$$\mu(s) = E[Y_1(s)],$$

Assumindo-se que a variância de $Y_1(s)$ existe para todo $s \in S$. O processo Y_1 é dito t-Student reparametrizado se, para qualquer $n \geq 1$ e localizações s_1, \dots, s_n , o vetor $(Y_1(s_1), \dots, Y_1(s_n))$ segue uma distribuição t-Student reparametrizada multivariada.

O processo de dependência espacial até o momento tem sido baseado no pressuposto de que a inferência deve ser baseada em uma única repetição no campo aleatório $Y_i(s)$ em que $i = 1$. Smith (2001) relata que podem-se obter estimativas melhores se houver conjuntos de várias repetições do processo. Smith (2001) e De Bastiani et al. (2017) estudaram os dados de cada ano como independentes, o que equivale a assumir que existem múltiplas repetições independentes.

Seja $Y = Y(s) = \text{vec}(Y_1(s), \dots, Y_r(s))$ um vetor aleatório $nr \times 1$ de r processos estocásticos independente de n elementos cada, que pertencem à família de distribuições t-Student e dependem da posição $s \in S \subset \mathfrak{R}^2$. O operador “vec” transforma uma matriz em um vetor, empilhando as colunas da matriz uma em abaixo da outra. Assume-se o i -ésimo processo estocástico sendo $Y_i = Y_i(s) = \text{vec}(Y_i(s_1), \dots, Y_i(s_n))$ com $i = 1, \dots, r$, em que $s = (s_1, \dots, s_n)^\top$ representa o vetor $n \times 1$ em que $s \in S \subset \mathfrak{R}^2$. Considera-se a notação matricial $\mu_i(s) = X_i(s)\beta$, o modelo é expresso por:

$$Y_i(s) = X_i(s)\beta + \epsilon_i(s),$$

sendo $i = 1, \dots, r$, com $\beta = (\beta_1, \dots, \beta_p)^\top$ é o vetor de parâmetros desconhecidos $p \times 1$, $X_i(s) = X_i$ é a matriz $n \times p$ com v -ésima linha x_v^\top , em que $x_{vj} = x_j(s_v)$ com $x_i^\top = (x_{i1}, \dots, x_{ip})$ para $v = 1, \dots, n$ e $j = 1, \dots, p$ as variáveis explicativas. A matriz de projeção X_i é diferente para todas as r repetições em que $s \in S \subset \mathfrak{R}^2$ e ϵ_i é o erro estocástico, caso particular quando X_i é igual para todas as repetições é apresentado por De Bastiani et al. (2017) para o modelo espacial linear com distribuição gaussiana.

Seja $\Sigma_i = [C_i(s_u, s_v)]$ a matriz de covariância $n \times n$ de $Y_i(s)$ para a i -ésima repetição, $i = 1, \dots, r$. A matriz Σ_i é não-singular, simétrica e positiva definida, associada ao vetor $Y_i(s)$, em que, para o processo estacionário e isotrópico, os elementos $C_i(s_u, s_v)$ depende da

distância euclidiana entre os pontos s_u e s_v . Se processo estacionário espacial aleatório tem a propriedade de que a dependência entre quaisquer duas observações depende apenas da distância entre suas localizações, $h_{uv} = \|s_u - s_v\|$, independentemente da direção, então o processo é considerado isotrópico. Caso contrário, é dito anisotrópico. Os processos isotrópicos são convenientes para lidar porque existem formas paramétricas amplamente utilizadas para $C_i(h_{uv}) = C_i(s_u, s_v)$. Na mesma abordagem de Smith (2001) e De Bastiani et al. (2017), assume-se um processo homogêneo. Considera-se \mathbf{Y}_i um vetor aleatório que tem distribuição t-Student reparametrizada com parâmetro de forma η fixo, em que $0 < \eta < \frac{1}{2}$, com matriz de covariância Σ_i , $n \times n$, vetor de média $E(\mathbf{Y}_i) = \boldsymbol{\mu}_i$, $n \times 1$, se sua função densidade de probabilidade (fdp) é dada em (4.24) por:

$$f(\mathbf{Y}, \boldsymbol{\theta}) = \prod_{i=1}^r f(\mathbf{Y}_i, \boldsymbol{\theta}) = \prod_{i=1}^r K_n(\eta) |\Sigma_i|^{-\frac{1}{2}} [1 + c(\eta)\delta_i^2]^{-\frac{1}{2\eta}(1+n\eta)}, \quad (4.24)$$

em que,

$$K_n(\eta) = \left(\frac{c(\eta)}{\pi}\right)^{\frac{n}{2}} \frac{\Gamma\left(\frac{1+n\eta}{2\eta}\right)}{\Gamma\left(\frac{1}{2\eta}\right)},$$

sendo, $\delta_i^2 = (\mathbf{Y}_i - \boldsymbol{\mu}_i)^\top \Sigma_i^{-1} (\mathbf{Y}_i - \boldsymbol{\mu}_i)$ é a distância de Mahalanobis, $c(\eta) = \eta/(1 - 2\eta)$ e $0 < \eta < \frac{1}{2}$. Denota-se por $\mathbf{Y} \sim \mathcal{T}_{nr}(1_r \otimes \boldsymbol{\mu}_i, \mathbf{I}_r \otimes \Sigma_i, \eta)$, que o vetor \mathbf{Y} tem distribuição t-Student multivariada reparametrizada. A representação estocástica de \mathbf{Y}_i , uma variável aleatória t-Student reparametrizada, é expressa por:

$$\mathbf{Y}_i \stackrel{d}{=} \boldsymbol{\mu}_i + V^{-1/2} \mathbf{Z}_i,$$

em que $\mathbf{Z}_i \sim N_n(\mathbf{0}, \Sigma_i)$, $V \sim \text{Gama}\left(\frac{1}{2\eta}, \frac{1}{2c(\eta)}\right)$, $c(\eta) = \eta/(1 - 2\eta)$, $0 < \eta < \frac{1}{2}$; V e \mathbf{Z}_i são independentes. Considera-se a mesma estrutura de covariância para cada repetição, ou seja, a matriz de covariância $\Sigma_i = \Sigma = [C_i(s_u, s_v)]$ tem uma estrutura que depende do vetor de parâmetros $\boldsymbol{\phi} = (\phi_1, \dots, \phi_q)$. A matriz de covariância é dada por:

$$\Sigma_i = \Sigma = \phi_1 \mathbf{I}_n + \phi_2 \mathbf{R}, \quad (4.25)$$

em que $i = 1, \dots, r$, em que $\phi_1 \geq 0$ é o parâmetro denominado como efeito de pepita (nugget); \mathbf{I}_n é uma matriz de identidade $n \times n$; $\phi_2 \geq 0$ é parâmetro contribuição (sill); \mathbf{R} é uma matriz simétrica $n \times n$, cujos elementos estão em função dos parâmetros $\phi_3 > 0$, $\mathbf{R} = \mathbf{R}(\phi_3) = [(r_{uv})]$ com elementos na diagonal sendo $r_{uu} = 1$, $u = 1, \dots, n$; $r_{uv} = \phi_2^{-1} C(s_u, s_v)$ para $\phi_2 \neq 0$ e $r_{uv} = 0$ para $\phi_2 = 0$ quando $u \neq v = 1, \dots, n$, em que r_{uv} depende da distância euclidiana $h_{uv} = \|s_u - s_v\|$ entre os pontos s_u e s_v ; ϕ_3 é função do alcance do modelo (a).

Assume-se que $\boldsymbol{\mu}_i(s)$ é constante e a variância é definida por:

$$\text{var}[\mathbf{Y}_i(s_u) - \mathbf{Y}_i(s_v)] = 2\gamma(\|s_u - s_v\|),$$

A função $\gamma(\cdot)$ é denominada de semivariância e só faz sentido se a variância entre as observações depender de s_u e s_v somente através da sua diferença ($s_u - s_v$). Se o processo é

estacionário então:

$$\gamma_i(h_{uv}) = C_i(0) - C_i(h_{uv}).$$

Métodos estatísticos para processos estacionários de segunda ordem podem ser considerados em termos de covariâncias ou em termos de variogramas.

4.22 Função escore

Sendo $\mathbf{Y}_i(s) = \text{vec}(\mathbf{Y}_1(s_1), \dots, \mathbf{Y}_r(s_n))$, $nr \times 1$ o vetor aleatório de r vetores distribuídos independentes e identicamente distribuídos. O logaritmo da função verossimilhança para r repetições independentes é definido por:

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{i=1}^r \mathcal{L}_i(\boldsymbol{\theta}), \quad (4.26)$$

em que \mathcal{L}_i representa para cada repetição $i = 1, \dots, r$ (4.27):

$$\mathcal{L}_i(\boldsymbol{\theta}) = \log K_n(\eta) - \frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2\eta} (1 + n\eta) \log(1 + c(\eta) \delta_i^2). \quad (4.27)$$

A função escore definida por $\mathbf{U}(\boldsymbol{\theta}) = \sum_{i=1}^r \mathbf{U}_i(\boldsymbol{\theta})$ em que cada elemento pode ser escrito como $\mathbf{U}_i(\boldsymbol{\theta}) = (\mathbf{U}_i^\top(\boldsymbol{\beta}), \mathbf{U}_i^\top(\boldsymbol{\phi}), \mathbf{U}_i(\eta))^\top$ com:

$$\mathbf{U}_i(\boldsymbol{\beta}) = w_i \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i, \quad (4.28)$$

$$\mathbf{U}_i(\boldsymbol{\phi}) = \frac{1}{2} D_\phi^\top \text{vec} \left[\left(w_i \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \right) - \boldsymbol{\Sigma}^{-1} \right], \quad (4.29)$$

$$\mathbf{U}_i(\eta) = \frac{1}{2\eta^2} \left[\psi \left(\frac{1}{2\eta} \right) - \psi \left(\frac{1+n\eta}{2\eta} \right) + c(\eta) (n - w_i \delta_i^2) + \log(q_i) \right], \quad (4.30)$$

em que $w_i = ((1+n\eta)/\eta) q_i^{-1} c(\eta)$, $q_i = (1 + c(\eta) \delta_i^2)$, $\boldsymbol{\varepsilon}_i = (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta})$, detalhes das derivadas de $D_\phi^\top = (\partial \text{vec}^\top(\boldsymbol{\Sigma}) / \partial \boldsymbol{\phi})$ com respeito a $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3)$ são apresentados no Apêndice 4.33 e $\psi(z) = (\partial \log \Gamma(z) / \partial z)$ denominada função Digama. As resoluções das expressões $\mathbf{U}_i(\boldsymbol{\beta})$, $\mathbf{U}_i(\boldsymbol{\phi})$ e $\mathbf{U}_i(\eta)$ estão apresentadas no Apêndice 4.32.

4.23 Matriz de informação esperada de Fisher

A matriz informação esperada de Fisher para $\boldsymbol{\theta}$ com base no logaritmo da função verossimilhança definida em (4.27) é denotada por $\mathcal{I}(\boldsymbol{\theta})$ e pode ser escrita como:

$$\mathcal{I}(\boldsymbol{\theta}) = \frac{1}{r} \sum_{i=1}^r E \left\{ \mathbf{U}_i(\boldsymbol{\theta}) \mathbf{U}_i^\top(\boldsymbol{\theta}) \right\},$$

sendo a função escore $\mathbf{U}_i(\boldsymbol{\theta}) = (\mathbf{U}_i(\boldsymbol{\beta})^\top, \mathbf{U}_i(\boldsymbol{\phi})^\top, \mathbf{U}_i(\eta))^\top$, associada à i -ésima componente da função log-verossimilhança com $i = 1, \dots, r$ é definida nas equações (4.28), (4.29) e (4.30) e o valor esperado $E(\cdot)$ é obtido em relação à função densidade de probabilidade (fdp). Em

seguida, obtém-se cada elemento da matriz informação de Fisher dada por:

$$\mathcal{I}(\boldsymbol{\theta}) = \begin{pmatrix} \mathcal{I}_{\beta\beta}(\boldsymbol{\theta}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathcal{I}_{\phi\phi}(\boldsymbol{\theta}) & \mathcal{I}_{\phi\eta}(\boldsymbol{\theta}) \\ \mathbf{0} & \mathcal{I}_{\eta\phi}(\boldsymbol{\theta}) & \mathcal{I}_{\eta\eta}(\boldsymbol{\theta}) \end{pmatrix}, \quad (\text{M1})$$

com elementos em forma matricial:

$$\begin{aligned} \mathcal{I}_{\beta\beta}(\boldsymbol{\theta}) &= \frac{c_{\beta}(\eta)}{r} \sum_{i=1}^r \mathbf{X}_i^{\top} \boldsymbol{\Sigma}^{-1} \mathbf{X}_i, \\ \mathcal{I}_{\phi\phi}(\boldsymbol{\theta}) &= \frac{1}{4} \mathbf{D}_{\phi}^{\top} \left[2c_{\phi}(\eta) (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{N}_n + (c_{\phi}(\eta) - 1) \text{vec}(\boldsymbol{\Sigma}^{-1}) \text{vec}^{\top}(\boldsymbol{\Sigma}^{-1}) \right] \mathbf{D}_{\phi}, \\ \mathcal{I}_{\phi\eta}(\boldsymbol{\theta}) &= - \left[\frac{c(\eta)(n+2)}{(1+(n+2)\eta)(1+n\eta)} \right] \mathbf{D}_{\phi}^{\top} \text{vec}(\boldsymbol{\Sigma}^{-1}), \\ \mathcal{I}_{\eta\eta}(\boldsymbol{\theta}) &= - \frac{1}{2\eta^2} \left[\frac{n}{(1-2\eta)^2} \left(\frac{1+n\eta(1-4\eta)-8\eta^2}{(1+n\eta)(1+(n+2)\eta)} \right) - \mathcal{B}'(\eta) \right] \end{aligned}$$

em que, $\mathbf{D}_{\phi} = \frac{\partial \text{vec}(\boldsymbol{\Sigma})}{\partial \phi^{\top}}$, $\mathbf{N}_n = \frac{1}{2} (\mathbf{I}_{n^2} + \mathbf{K}_n)$, \mathbf{K}_n é a matriz de comutação de ordem $n^2 \times n^2$, \mathbf{I}_{n^2} a matriz de identidade $n^2 \times n^2$ Magnus e Neudecker (1999), $c_{\beta}(\eta) = c_{\phi}(\eta) / (1 - 2\eta)$, $c_{\phi}(\eta) = (1 + n\eta) / (1 + (n + 2)\eta)$ e

$$\mathcal{B}'(\eta) = - \frac{1}{2\eta^2} \left\{ \psi' \left(\frac{1+n\eta}{2\eta} \right) - \psi' \left(\frac{1}{2\eta} \right) \right\}$$

em que ψ' é denominada a função Trigama. Pode-se observar que a matriz de informação de Fisher (M1) é bloco diagonal, formada pelos blocos $\mathcal{I}_{11}(\boldsymbol{\theta})$ e $\mathcal{I}_{22}(\boldsymbol{\theta})$, com o parâmetro β no primeiro bloco e os parâmetros ϕ e η no segundo bloco, respectivamente, dada por:

$$\mathcal{I}_{11}(\boldsymbol{\theta}) = \mathcal{I}_{\beta\beta}(\boldsymbol{\theta}) \text{ e } \mathcal{I}_{22}(\boldsymbol{\theta}) = \begin{pmatrix} \mathcal{I}_{\phi\phi}(\boldsymbol{\theta}) & \mathcal{I}_{\phi\eta}(\boldsymbol{\theta}) \\ \mathcal{I}_{\eta\phi}(\boldsymbol{\theta}) & \mathcal{I}_{\eta\eta}(\boldsymbol{\theta}) \end{pmatrix}. \quad (\text{M2})$$

Os quais colaboram com resultados encontrados para distribuições pertencentes à classe de distribuições mistura escala normal (GALEA et al., 2020). As resoluções das expressões $\mathcal{I}_{\beta\beta}(\boldsymbol{\theta})$, $\mathcal{I}_{\phi\phi}(\boldsymbol{\theta})$, $\mathcal{I}_{\phi\eta}(\boldsymbol{\theta})$, $\mathcal{I}_{\eta\phi}(\boldsymbol{\theta})$ e $\mathcal{I}_{\eta\eta}(\boldsymbol{\theta})$ estão apresentadas no Apêndice 4.34.

4.24 Estimador de máxima verossimilhança

O método de estimação de máxima verossimilhança (ML) é considerado um estimador de parâmetros assintoticamente normal, não viciado e eficiente (MARDIA; MARSHALL, 1984). Ele consiste em maximizar o logaritmo da função verossimilhança em relação aos efeitos fixos e aos componentes dos efeitos aleatórios.

A função de verossimilhança pode ser entendida como a intensidade de contribuições dos parâmetros na produção de uma amostra, ou seja, o quão fortemente os dados suportam os parâmetros desconhecidos. A escolha do melhor modelo para o processo $\mathbf{Y}_i(s)$ corresponde ao valor máximo do logaritmo da função verossimilhança completa.

A função de verossimilhança de um vetor de dados completos $\mathbf{Y}_c = (\mathbf{Y}_i^{\top}, V)^{\top}$, tal

que \mathbf{Y}_i^\top é o vetor de dados observados e V é uma variável adicional, definida como a função densidade conjunta dada por $f_{\mathbf{Y}_i|V}(y_i, \nu)$ e admite-se o seguinte modelo hierárquico:

$$\mathbf{Y}_i|V = \nu \sim N_n(\mathbf{X}_i\boldsymbol{\beta}, \nu^{-1}\boldsymbol{\Sigma}),$$

sendo

$$V \sim \text{Gama}\left(\frac{1}{2\eta}, \frac{1}{2c(\eta)}\right),$$

com $c(\eta) = \eta/(1 - 2\eta)$. Em que *Gama* (α, β) é a distribuição gama de probabilidade com os parâmetros α e β , definidos pela função densidade de probabilidade $h(\nu) = \beta^\alpha \nu^{\alpha-1} \exp(-\nu\beta)/\Gamma(\alpha)$, $\nu > 0$, $\alpha > 0$, $\beta > 0$. Pinheiro et al. (2001) apresentaram também as seguintes definições: $E(V) = \alpha/\beta$, $Var(V) = \alpha/\beta^2$ e $E(\log(\nu)) = \psi(\alpha) - \log(\beta)$, em que $\psi(\alpha) = \partial \log(\Gamma(\alpha))/\partial \alpha$ é denominada a função Digama. A distribuição condicional $V|Y_i$ definida por

$$V|Y_i \sim \text{Gama}\left(\frac{1+n\eta}{2\eta}, \frac{1+c(\eta)\delta_i^2}{2c(\eta)}\right),$$

A função densidade conjunta dos dados completos é da forma

$$f_{\mathbf{Y}_i,V}(y_i, \nu) = f_{\mathbf{Y}_i|V}(y_i|\nu) \cdot h(\nu),$$

sendo

$$f_{\mathbf{Y}_i|V}(y_i|\nu) = \frac{1}{(2\pi)^{\frac{n}{2}} |\nu^{-1}\boldsymbol{\Sigma}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{Y}_i - \mathbf{X}_i\boldsymbol{\beta})^\top \nu \boldsymbol{\Sigma}^{-1}(\mathbf{Y}_i - \mathbf{X}_i\boldsymbol{\beta})},$$

e

$$h(\nu) = \left(\frac{1}{2c(\eta)}\right)^{\frac{1}{2\eta}} \frac{1}{\Gamma\left(\frac{1}{2\eta}\right)} \nu^{\frac{1}{2\eta}-1} e^{-\frac{\nu}{2c(\eta)}}.$$

Assim, obtêm-se a distribuição t-Student n-variada reparametrizada $\mathbf{Y}_i \sim \mathcal{T}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$, com parâmetros desconhecidos do modelo, $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top, \eta)^\top$, em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$, $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3)^\top$ e $0 < \eta < 1/2$ com função densidade conjunta dos dados completos apresentados em 4.31.

O logaritmo da função verossimilhança completa é dado por

$$\mathcal{L}_c(\boldsymbol{\theta}, \mathbf{Y}_c) = \sum_{i=1}^r \mathcal{L}_{ci}(\boldsymbol{\theta}, \mathbf{Y}_c) \quad (4.31)$$

em que

$$\mathcal{L}_{ci}(\boldsymbol{\theta}, \mathbf{Y}_c) = \log(f_{\mathbf{Y}_i,V}(y_i, \nu)) = \log(f_{\mathbf{Y}_i|V}(y_i|\nu)) + \log(h(\nu)),$$

sendo

$$\log(f_{\mathbf{Y}_i|V}(y_i|\nu)) = -\frac{n}{2} \log(2\pi) + \frac{n}{2} \log(\nu) - \frac{1}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2} \nu \delta_i^2,$$

em que $\delta_i^2 = (\mathbf{Y}_i - \mathbf{X}\boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{Y}_i - \mathbf{X}\boldsymbol{\beta})$ é a distância de Mahalanobis e $\log(h(\nu))$ é dado por

$$\log(h(\nu)) = +\frac{1}{2\eta} \log\left(\frac{1}{2c(\eta)}\right) - \log\left(\Gamma\left(\frac{1}{2\eta}\right)\right) + \frac{1}{2c(\eta)} \log(\nu) - \frac{1}{2c(\eta)} \nu,$$

como $((1/2\eta) - 1) = (1/2c(\eta))$. Assim,

$$\begin{aligned} \mathcal{L}_{ci}(\boldsymbol{\theta}, \mathbf{Y}_c) &= -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2} \delta_i^2 \nu + \frac{1}{2\eta} \log\left(\frac{1}{2c(\eta)}\right) \\ &\quad - \log\left(\Gamma\left(\frac{1}{2\eta}\right)\right) + \frac{n}{2} \log(\nu) + \frac{1}{2c(\eta)} [\log(\nu) - \nu]. \end{aligned}$$

Um procedimento muito utilizado para a estimação de parâmetros é a utilização de algoritmos iterativos, que consiste em maximizar o logaritmo da função de verossimilhança completa (4.31), definido como:

$$\mathcal{L}_c(\hat{\boldsymbol{\theta}}, \mathbf{Y}_c) = \max |\mathcal{L}_c(\boldsymbol{\theta}, \mathbf{Y}_c)|. \quad (4.32)$$

A maximização de (4.32) é obtida utilizando-se um processo iterativo. Neste caso, desenvolveu-se o algoritmo iterativo parcial linear, baseado em Acosta et al. (2016) que geram uma aproximação por ML de forma iterativa. Calcula-se a esperança condicional do logaritmo da função verossimilhança completa definida como a Q-fuction:

$$Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) = \sum_{i=1}^r Q_i(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)})$$

sendo $Q_i(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) = E[\mathcal{L}_{ci}(\boldsymbol{\theta}, \mathbf{Y}_c) | V, \boldsymbol{\theta}^{(k)}]$. Cálculos são apresentados no Apêndice 4.36. Obtém-se:

$$Q_i(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) = -\frac{1}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2} \delta_i^2 w_i^{(k)} + \frac{1}{2\eta} \log\left(\frac{1}{2c(\eta)}\right) - \log\left(\Gamma\left(\frac{1}{2\eta}\right)\right) + \frac{1}{2c(\eta)} c_i^{(k)}, \quad (4.33)$$

em que $c_i^{(k)} = \psi\left(\frac{1+n\eta^{(k)}}{2\eta^{(k)}}\right) - \log\left(\frac{1+n\eta^{(k)}}{2\eta^{(k)}}\right) + \log w_i^{(k)} - (w_i^{(k)})$, $w_i^{(k)} = \left(\frac{1+n\eta^{(k)}}{\eta^{(k)}}\right) c(\eta^{(k)}) (q_i^{(k)})^{-1}$ e $q_i^{(k)} = (1 + c(\eta^{(k)}) \delta_i^2)$. Portanto, a esperança condicional do logaritmo da função verossimilhança completa para cada i , em que $i = 1, \dots, r$, representada em (4.33), pode ser expressa como ver:

$$Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) = Q_1(\boldsymbol{\tau} | \boldsymbol{\theta}^{(k)}) + Q_2(\eta | \boldsymbol{\theta}^{(k)}) \quad (4.34)$$

sendo $\boldsymbol{\tau} = (\boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top)^\top$, tem-se que:

$$Q_1(\boldsymbol{\tau} | \boldsymbol{\theta}^{(k)}) = -\frac{r}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{i=1}^r \delta_i^2 w_i^{(k)},$$

e

$$Q_2(\eta|\boldsymbol{\theta}^{(k)}) = r \left\{ \frac{1}{2\eta} \log\left(\frac{1}{2c(\eta)}\right) - \log\left(\Gamma\left(\frac{1}{2\eta}\right)\right) + \frac{1}{2c(\eta)} \left[\psi\left(\frac{1+n\eta^{(k)}}{2\eta^{(k)}}\right) - \log\left(\frac{1+n\eta^{(k)}}{2\eta^{(k)}}\right) + \frac{1}{r} \sum_{i=1}^r (\log w_i^{(k)} - w_i^{(k)}) \right] \right\}.$$

Para mais detalhes ver Apêndice 4.36.

Maximiza-se a função resultante $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$ dada em (4.34), em relação à $\boldsymbol{\theta} = (\boldsymbol{\tau}, \eta)^\top$, assim obtém-se a estimativa para $\boldsymbol{\theta}$. Inicialmente, maximiza-se o primeiro termo $Q_1(\boldsymbol{\tau}|\boldsymbol{\theta}^{(k)})$ em relação à $\boldsymbol{\beta}$ e ϕ , para obter $\hat{\boldsymbol{\beta}}$ e $\hat{\phi}$. Após, maximiza-se o segundo termo $Q_2(\eta|\boldsymbol{\theta}^{(k)})$ em relação à η , para obter $\hat{\eta}$. Assim, segue a resolução das estimativas:

A derivada em relação a $\boldsymbol{\beta}$ é dada por:

$$\dot{Q}_1(\boldsymbol{\tau}|\boldsymbol{\theta}^{(k)}) = \frac{\partial Q_1(\boldsymbol{\tau}|\boldsymbol{\theta}^{(k)})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^r w_i^{(k)} \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta}),$$

igualando a zero, tem-se que;

$$\boldsymbol{\beta} \left(\sum_{i=1}^r w_i^{(k)} \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right) = \sum_{i=1}^r w_i^{(k)} \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{Y}_i,$$

portanto, obtém-se a estimativa do parâmetro $\boldsymbol{\beta}$;

$$\hat{\boldsymbol{\beta}} = \left(\sum_{i=1}^r w_i^{(k)} \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right)^{-1} \left(\sum_{i=1}^r w_i^{(k)} \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{Y}_i \right).$$

Para obter a estimativa do parâmetro ϕ , determinou-se em duas partes, estimou-se ϕ_j com $j = 1, 2$ por meio da resolução do sistema (4.39) e a estimativa para ϕ_j em que $j = 3$, pela resolução da equação (4.40). Inicialmente, deriva-se $Q_1(\boldsymbol{\tau}|\boldsymbol{\theta}^{(k)})$ em relação à ϕ_j :

$$\begin{aligned} \dot{Q}_1(\boldsymbol{\tau}|\boldsymbol{\theta}^{(k)}) &= \frac{\partial Q_1(\boldsymbol{\tau}|\boldsymbol{\theta}^{(k)})}{\partial \phi_j} \\ &= -\frac{1}{2} \sum_{i=1}^r \left[\text{tr}(\boldsymbol{\Sigma}^{-1} \dot{\boldsymbol{\Sigma}}_j) - w_i^{(k)} \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \dot{\boldsymbol{\Sigma}}_j \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right], \end{aligned} \quad (4.35)$$

para $\dot{\boldsymbol{\Sigma}}_j = \partial \boldsymbol{\Sigma} / \partial \phi_j$ sendo $j = 1, 2$. Igualando a equação (4.35) a zero, tem-se:

$$\text{tr}(\boldsymbol{\Sigma}^{-1} \dot{\boldsymbol{\Sigma}}_j) = \text{tr}(\mathbf{E} \boldsymbol{\Sigma}^{-1} \dot{\boldsymbol{\Sigma}}_j \boldsymbol{\Sigma}^{-1}), \quad (4.36)$$

sendo $\mathbf{E} = \frac{1}{r} \sum_{i=1}^r w_i^{(k)} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top$, equivalente à equação (4.36),

$$\text{tr}(\boldsymbol{\Sigma}^{-1} \dot{\boldsymbol{\Sigma}}_j \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}) = \text{tr}(\mathbf{E} \boldsymbol{\Sigma}^{-1} \dot{\boldsymbol{\Sigma}}_j \boldsymbol{\Sigma}^{-1}). \quad (4.37)$$

Considerando-se $\Sigma = \sum_{l=1}^2 \phi_l \dot{\Sigma}_l$, em que $\dot{\Sigma}_1 = \mathbf{I}_n$ e $\dot{\Sigma}_2 = \mathbf{R}(\phi_3)$, aplicando-se em (4.37)

$$\sum_{l=1}^2 \text{tr}(\Sigma^{-1} \dot{\Sigma}_j \Sigma^{-1} \dot{\Sigma}_l) \phi_l = \text{tr}(\Sigma^{-1} \dot{\Sigma}_j \Sigma^{-1} \mathbf{E}). \quad (4.38)$$

Note-se que, em notação matricial, o sistema (4.38) pode ser escrito como:

$$\mathbf{A} \cdot \boldsymbol{\phi} = \mathbf{b},$$

sendo a matriz \mathbf{A} de ordem 2×2 , com elementos $a_{jl} = \text{tr}(\Sigma^{-1} \dot{\Sigma}_j \Sigma^{-1} \dot{\Sigma}_l)$ e \mathbf{b} um vetor (2×1) com elementos $b_j = \text{tr}(\Sigma^{-1} \dot{\Sigma}_j \Sigma^{-1} \mathbf{E})$, sendo correspondente à $a_{11} = \text{tr}(\Sigma^{-1} \Sigma^{-1})$, $a_{12} = a_{21} = \text{tr}(\Sigma^{-1} \Sigma^{-1} \mathbf{R}(\phi_3))$, $a_{22} = \text{tr}(\Sigma^{-1} \mathbf{R}(\phi_3) \Sigma^{-1} \mathbf{R}(\phi_3))$, $b_1 = \text{tr}(\Sigma^{-1} \Sigma^{-1} \mathbf{E})$ e $b_2 = \text{tr}(\Sigma^{-1} \mathbf{R}(\phi_3) \Sigma^{-1} \mathbf{E})$. O processo de estimação é dado pela resolução do sistema:

$$\hat{\phi}_k = \mathbf{A} \cdot \mathbf{b}, \quad k = 0, 1, 2, \dots \quad (4.39)$$

Para se obter a estimativa de ϕ_j em que $j = 3$, utiliza-se a equação (4.37) que pode ser reescrita da seguinte forma:

$$h(\phi_3) = \text{tr}[\Sigma^{-1} \dot{\Sigma}_3 \Sigma^{-1} (\Sigma - \mathbf{E})],$$

sendo $\Sigma = \phi_1 \mathbf{I} + \phi_2 \mathbf{R}(\phi_3)$, $\dot{\Sigma}_3 = \phi_2 \dot{\mathbf{R}}(\phi_3)$ e $h(\phi_3) = 0$. Obtém-se

$$h(\phi_3) = \text{tr}[\Sigma^{-1} \dot{\mathbf{R}}(\phi_3) \Sigma^{-1} (\phi_1 \mathbf{I} + \phi_2 \mathbf{R}(\phi_3) - \mathbf{E})].$$

Utilizando-se o método de Newton-Raphson, estima-se ϕ_3 dada por:

$$\hat{\phi}_3 = \phi_3 - \frac{h(\phi_3)}{h'(\phi_3)}, \quad (4.40)$$

em que $h'(\phi_3) = \phi_2 \text{tr} \left\{ \Sigma^{-1} \dot{\mathbf{R}}(\phi_3) \Sigma^{-1} \dot{\mathbf{R}}(\phi_3) \right\}$.

Verifica-se que $\dot{\Sigma}_3$, depende de cada modelo: exponencial, gaussiano e família Matérn. Porém, pode-se obter $\dot{\Sigma}_3$ utilizando somente o modelo da família Matérn, pois quando $k = 0, 5$ este corresponde ao modelo exponencial e quando $k \rightarrow \infty$ este refere-se ao modelo gaussiano. Para a estimativa de η , necessita-se aplicar a transformação em $\eta = 1/\nu$ sendo $\nu > 2$ em $Q_2(\eta | \boldsymbol{\theta}^{(k)})$, obtendo-se

$$\begin{aligned} Q_2(\nu | \boldsymbol{\theta}^{(k)}) &= r \left\{ \frac{\nu}{2} \log \left(\frac{\nu - 2}{2} \right) - \log \left(\Gamma \left(\frac{\nu}{2} \right) \right) \left(\frac{\nu - 2}{2} \right) \left[\psi \left(\frac{\nu^{(k)} + n}{2} \right) \right. \right. \\ &\quad \left. \left. + -\log \left(\frac{\nu^{(k)} + n}{2} \right) + \frac{1}{r} \sum_{i=1}^r \left(\log w_{iv}^{(k)} - w_{iv}^{(k)} \right) \right] \right\}, \end{aligned} \quad (4.41)$$

sendo $w_{iv}^{(k)} = \frac{\nu^{(k)} + n}{\nu^{(k)} - 2} \left(1 + \frac{\delta_i^2}{\nu^{(k)} - 2} \right)^{-1}$. Está transformação se faz necessária para melhorar a convergência computacional, a qual será utilizada na seção 4.25. A partir da equação (4.41),

maximiza-se $g(\nu)$ em que $g(\nu) = \left(\partial Q_2 \left(\nu | \boldsymbol{\theta}^{(k)} \right) / \partial \nu \right) = 0$ dada por:

$$g(\nu) = \log \left(\frac{\nu - 2}{2} \right) + \left(\frac{\nu}{\nu - 2} \right) - \psi \left(\frac{\nu}{2} \right) + c_\nu^{(k)},$$

com $c_\nu^{(k)} = \psi \left(\frac{\nu^{(k)} + n}{\nu^{(k)} - 2} \right) - \log \left(\frac{\nu^{(k)} + n}{2} \right) + \frac{1}{r} \sum_{i=1}^r \left(\log w_{i\nu}^{(k)} - w_{i\nu}^{(k)} \right)$. Logo, para se obter a estimativa de ν utilizou-se novamente o método de Newton-Raphson, tem-se que:

$$\hat{\nu} = \nu - \frac{g(\nu)}{g'(\nu)},$$

sendo $g'(\nu) = ((\nu - 4)/(\nu - 2)^2) - (1/2)\psi'(\nu/2)$. Obtendo-se $\hat{\nu}$ gera-se $\hat{\eta} = 1/\hat{\nu}$.

Na seção 4.25, segue a descrição do algoritmo iterativo parcial linear, descrito nesta seção. Em qual este algoritmo irá obter a estimativa dos parâmetros de $\boldsymbol{\theta}$ através de iterações, ou seja, irá obter $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(k)}$ quando atingir o critério de convergência na k -ésima iteração.

4.25 Algoritmo iterativo parcial linear

Baseados em Acosta et al. (2016), para obter a estimação dos parâmetros de $\boldsymbol{\theta}$, deve-se utilizar um algoritmo iterativo. Assim, seguem os seguintes procedimentos:

1º Iteração: Seja $k = 0$, em que k representa a fase de iteração:

1º passo: Neste passo, necessitamos definir um chute inicial para o parâmetro a ser estimado $\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^0$, em que $\boldsymbol{\theta}^0 = (\boldsymbol{\beta}^0, \phi^0, \eta^0)$. Neste estudo, optou-se em definir os parâmetros iniciais $\boldsymbol{\beta}^0$ e ϕ^0 , obtendo um modelo de regressão com distribuição normal e η^0 definido como 0,25.

2º passo: Calculam-se a partir dos parâmetros iniciais obtidos no 1º passo, as seguintes equações:

$$\begin{aligned} \mathbf{R}^{(k)} &= \mathbf{R}(\phi_3^{(k)}), \\ \dot{\mathbf{R}}^{(k)} &= \frac{\partial \mathbf{R}(\phi_3^{(k)})}{\partial \phi_3}, \\ \boldsymbol{\Sigma}^{(k)} &= \phi_1^{(k)} \mathbf{I} + \phi_2^{(k)} \mathbf{R}(\phi_3^{(k)}), \\ c(\eta^{(k)}) &= \eta^{(k)} / (1 - 2\eta^{(k)}), \\ q_i^{(k)} &= (1 + c(\eta^{(k)}) \delta_i^2)^{-1}, \\ \delta_i^2 &= (\boldsymbol{\varepsilon}_i^{(k)})^\top (\boldsymbol{\Sigma}^{(k)})^{-1} \boldsymbol{\varepsilon}_i^{(k)}, \\ \boldsymbol{\varepsilon}_i^{(k)} &= (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta}^{(k)}), \\ w_i^{(k)} &= [(1 + n\eta^{(k)}) / \eta^{(k)}] c(\eta^{(k)}) (q_i^{(k)})^{-1}, \\ \boldsymbol{\beta}^{(k)} &= (\sum_{i=1}^r w_i^{(k)} \mathbf{X}_i^\top (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{X}_i)^{-1} (\sum_{i=1}^r w_i^{(k)} \mathbf{X}_i^\top (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{Y}_i), \\ \mathbf{E}^{(k)} &= \frac{1}{r} \sum_{i=1}^r w_i^{(k)} \boldsymbol{\varepsilon}_i^{(k)} (\boldsymbol{\varepsilon}_i^{(k)})^\top, \end{aligned}$$

em que \mathbf{R} é a função de covariância que depende do modelo exponencial, gaussiano ou família Matérn, com \mathbf{R} e $\dot{\mathbf{R}}$ apresentados por Uribe-Opazo et al. (2012). Verifique que as equações calculadas no 2º passo estão sendo consideradas iniciais ou seja $k = 0$.

3º passo: A partir deste momento, atualizam-se os parâmetros para $(k + 1)$ em que $\boldsymbol{\theta}^{(k+1)} = (\boldsymbol{\beta}^{(k+1)}, \phi^{(k+1)}, \eta^{(k+1)})$. Considerando-se os seguintes procedimentos:

3.1 passo: Atualizar os parâmetros lineares ϕ_1 e ϕ_2 de Σ , pelo sistema linear:

$$\mathbf{A}^{(k)} \cdot \boldsymbol{\phi}^{(k+1)} = \mathbf{b}^{(k)},$$

sendo a matriz $\mathbf{A}^{(k)}$ de ordem 2×2 , com elementos $a_{jl}^{(k)} = \text{tr}((\boldsymbol{\Sigma}^{(k)})^{-1} \dot{\boldsymbol{\Sigma}}_j (\boldsymbol{\Sigma}^{(k)})^{-1} \dot{\boldsymbol{\Sigma}}_l)$ e $\mathbf{b}^{(k)}$ um vetor (2×1) com elementos $b_j^{(k)} = \text{tr}((\boldsymbol{\Sigma}^{(k)})^{-1} \dot{\boldsymbol{\Sigma}}_j (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{E}^{(k)})$, em que $j, l = 1, 2$ sendo correspondente à $a_{11}^{(k)} = \text{tr}((\boldsymbol{\Sigma}^{(k)})^{-1} (\boldsymbol{\Sigma}^{(k)})^{-1})$, $a_{12}^{(k)} = a_{21}^{(k)} = \text{tr}((\boldsymbol{\Sigma}^{(k)})^{-1} (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{R}(\phi_3^{(k)}))$, $a_{22}^{(k)} = \text{tr}((\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{R}(\phi_3^{(k)}) (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{R}(\phi_3^{(k)}))$, para o vetor $b_1^{(k)} = \text{tr}((\boldsymbol{\Sigma}^{(k)})^{-1} (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{E}^{(k)})$ e $b_2^{(k)} = \text{tr}((\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{R}(\phi_3^{(k)}) (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{E}^{(k)})$.

3.2 passo: Para atualizar ϕ_3 de Σ . Resolve-se a equação não linear $h(\phi_3^{(k)}) = 0$, para obter $\phi_3^{(k+1)}$ com:

$$h(\phi_3^{(k)}) = \text{tr} \left\{ (\boldsymbol{\Sigma}^{(k)})^{-1} \dot{\mathbf{R}}(\phi_3^{(k)}) (\boldsymbol{\Sigma}^{(k)})^{-1} \left[\phi_1^{(k+1)} \mathbf{I} + \phi_2^{(k+1)} \mathbf{R}(\phi_3^{(k)}) - \mathbf{E}^{(k)} \right] \right\}.$$

O algoritmo Newton-Raphson fornece uma expressão recursiva para $\phi_3^{(k+1)}$ dada por:

$$\phi_3^{(k+1)} = \phi_3^{(k)} - \frac{h(\phi_3^{(k)})}{h'(\phi_3^{(k)})},$$

com $h'(\phi_3^{(k)}) = \phi_2^{(k+1)} \text{tr} \left\{ (\boldsymbol{\Sigma}^{(k)})^{-1} \dot{\mathbf{R}}(\phi_3^{(k)}) (\boldsymbol{\Sigma}^{(k)})^{-1} \dot{\mathbf{R}}(\phi_3^{(k)}) \right\}$.

3.3 passo: Ao se obter $\phi^{(k+1)}$ dos passos (3.1) e (3.2), atualiza-se $\Sigma^{(k+1)}$, a que será utilizada para atualizar o parâmetro $\beta^{(k+1)}$, pela expressão dada por:

$$\boldsymbol{\beta}^{(k+1)} = \left(\sum_{i=1}^r w_i^{(k)} \mathbf{X}_i^\top (\boldsymbol{\Sigma}^{(k+1)})^{-1} \mathbf{X}_i \right)^{-1} \left(\sum_{i=1}^r w_i^{(k)} \mathbf{X}_i^\top (\boldsymbol{\Sigma}^{(k+1)})^{-1} \mathbf{Y}_i \right),$$

3.4 passo: Para atualizar $\eta^{(k+1)}$, utiliza-se $Q_2(\eta | \boldsymbol{\theta}^{(k)})$ apresentada em (4.35), transformada em $Q_2(\nu | \boldsymbol{\theta}^{(k)})$ em que $\eta = 1/\nu$ e $\nu > 2$. A partir $g(\nu^{(k)}) = \partial Q_2(\nu | \boldsymbol{\theta}^{(k)}) / \partial \nu = 0$ tem-se $\nu^{(k)}$. Em que $g(\nu^{(k)})$ é dada por:

$$g(\nu^{(k)}) = \log \left(\frac{\nu - 2}{2} \right) + \left(\frac{\nu}{\nu - 2} \right) - \psi \left(\frac{\nu}{2} \right) + c_\nu^{(k)},$$

sendo $c_\nu^{(k)} = \psi \left(\frac{\nu^{(k)} + n}{\nu^{(k)} - 2} \right) - \log \left(\frac{\nu^{(k)} + n}{2} \right) + \frac{1}{r} \sum_{i=1}^r \left(\log w_{i\nu}^{(k)} - w_{i\nu}^{(k)} \right)$ e $w_{i\nu}^{(k)} = \left(\frac{\nu^{(k)} + n}{\nu^{(k)} - 2} \right) \left(1 + \frac{\delta_i^2}{\nu^{(k)} - 2} \right)^{-1}$. Para se obter $\nu^{(k+1)}$, utilizou-se novamente o algoritmo de Newton-Raphson dado por:

$$\nu^{(k+1)} = \nu^{(k)} - \frac{g(\nu^{(k)})}{g'(\nu^{(k)})}, \quad k = 0, 1, 2, \dots$$

sendo $g'(\nu) = ((\nu - 4)/(\nu - 2)^2) - (1/2)\psi'(\nu/2)$. Obtendo-se $\nu^{(k+1)}$ e determina-se $\eta^{(k+1)} = 1/\nu^{(k+1)}$.

4º passo: Finaliza-se a iteração, definindo-se $\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(1)}$, obtida pelas atualizações (3.1), (3.2), (3.3) e (3.4). A partir de $\boldsymbol{\theta}^{(1)}$, aplica-se o critério de convergência, que para este algoritmo é definido: se $\|\boldsymbol{\theta}^{(k+1)} - \boldsymbol{\theta}^{(k)}\| < e_1$ ou $|\mathcal{L}_c(\boldsymbol{\theta}^{(k+1)}) - \mathcal{L}_c(\boldsymbol{\theta}^{(k)})| < e_2$ pare e defina $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(k+1)}$, caso contrário $k = k + 1$ retorne para a 2º passo. Como e_1 e e_2 são tolerâncias constantes, em

geral, os valores típicos para as tolerâncias são 10^{-3} e 10^{-6} respectivamente.

4.26 Critério de escolha para o parâmetro de forma η

A utilização do algoritmo iterativo parcial linear é calculada para diferentes modelos da família Matérn e implementa-se o critério validação cruzada e traço para escolher o melhor modelo em função do parâmetro η . Estes métodos foram estudados por: Zellner (1976), o qual observou que, para o caso do modelo de regressão linear t-Student usual, a função de probabilidade do modelo t-Student é uma função crescente em v e não pode ser usada para selecionar o grau de liberdade da distribuição, propondo trabalhar com a validação cruzada. Kano et al. (1993), dentro da classe de distribuições elípticas, propõem como alternativa um método denominado de traço, que consiste em calcular o traço da matriz de covariância assintótica da média estimada. Baseado nestes autores, De Bastiani et al. (2015) propuseram a utilização da validação cruzada e o traço para selecionar o grau de liberdade v do modelo espacial linear t-Student. Schemmer et al. (2017), considerando o modelo t-Student reparametrizado, também utilizam estes dois critérios para selecionar o parâmetro de forma η . Fagundes et al. (2018b) utilizaram estes critérios no modelo espacial linear Slash para determinar os parâmetros de ajuste da curtose η e de suavização κ do modelo da família Matérn e justificam a utilização destes critérios para evitar o problema de identificabilidade nas estimativas de máxima verossimilhança, quando tem-se uma única repetição.

Neste estudo, tem-se várias repetições e o problema de identificabilidade dos parâmetros não ocorre. Porém, a validação cruzada e o traço também podem ser utilizados, como apresentado por De Bastiani (2016), a qual utiliza estes critérios para selecionar o parâmetro ϕ_4 . Seguindo este raciocínio, os critérios foram utilizados para escolher o melhor η para o modelo espacial linear t-Student reparametrizado com repetição a validação cruzada dada por (4.42):

$$CV = CV(\eta) = \frac{1}{nr} \sum_{i=1}^r \sum_{j=1}^n \left(\frac{y_i(s_j) - \hat{y}_{i(j)}(s_j)}{1 - h_{i(jj)}^*} \right)^2, \quad (4.42)$$

em que $\hat{y}_{i(j)}(s_j) = \mathbf{x}_j^\top \hat{\boldsymbol{\beta}}_{(j)}$, com $\mathbf{x}_j^\top = (x_{j1}, \dots, x_{jp})$ sendo a j -ésima linha da matriz \mathbf{X}_i , é a predição na localização s_j sem considerar esta observação (y_j, \mathbf{x}_j^\top) , $\hat{\boldsymbol{\beta}}_{(j)}$ é o estimador de máxima verossimilhança para $\boldsymbol{\beta}$ sem considerar a j -ésima observação e $h_{i(jj)}^*$ é o j -ésimo elemento da diagonal da matriz Hat (\mathbf{H}_i), também chamada de matriz projeção, a matriz projeção é definida por $\mathbf{H}_i = \mathbf{X}_i (\mathbf{X}_i^\top \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{X}_i)^{-1} \mathbf{X}_i^\top \hat{\boldsymbol{\Sigma}}^{-1}$, em que cada elemento $h_{i(jj)}^*$ da diagonal da matriz \mathbf{H}_i é chamado de alavanca, descreve a influência que cada valor da variável resposta tem sobre o valor estimado para a mesma observação, sendo $j = 1, \dots, n$ e $i = 1, \dots, r$. Para simplificar os cálculos, utiliza-se a aproximação dada por:

$$\hat{\boldsymbol{\beta}}_{(j)} \approx \hat{\boldsymbol{\beta}} + \mathbf{I}_{i\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}}^{-1}(\boldsymbol{\theta}) \mathbf{U}_{i(j)}(\hat{\boldsymbol{\beta}}),$$

em que:

$$\mathcal{I}_{i\hat{\beta}\hat{\beta}} = c_{\beta}(\eta) \mathbf{X}_i^T \hat{\Sigma}^{-1} \mathbf{X}_i,$$

e

$$U_{i(j)}(\hat{\beta}) = w_{i(j)} \mathbf{X}_{i(j)}^T \hat{\Sigma}^{-1} \varepsilon_{i(j)}, \quad (4.43)$$

em que $w_{i(j)} = c(\eta) \left(\frac{1+n\eta}{\eta} \right) / (1 + c(\eta)\delta_{i(j)}^2)$, $\delta_{i(j)}^2 = (\mathbf{Y}_{i(j)} - \mathbf{X}_{i(j)}\hat{\beta})^T \hat{\Sigma}^{-1} (\mathbf{Y}_{i(j)} - \mathbf{X}_{i(j)}\hat{\beta})$, com $\mathbf{X}_{i(j)}$ de ordem $(n-1) \times p$, a matriz \mathbf{X}_i sem a j -ésima linha \mathbf{x}_{ij}^T e $\mathbf{Y}_{i(j)}$ de ordem $(n-1) \times 1$, denotado pelo o vetor \mathbf{Y}_i sem a resposta $y_i(s_j)$, para $j = 1, \dots, n$ e $i = 1, \dots, r$. Note que na equação (4.43), a matriz $\hat{\Sigma}^{-1}$ tem ordem $(n-1) \times (n-1)$ e $\varepsilon_{i(j)} = (\mathbf{Y}_{i(j)} - \mathbf{X}_{i(j)}\hat{\beta})$. Para o critério traço, obtém-se a seguinte equação:

$$T_r(\eta) = \left(\frac{(1-2\eta)(1+(n+2)\eta)}{r(1+n\eta)} \right) \sum_{i=1}^r \text{tr} \left[(\mathbf{X}_i^T \mathbf{X}_i) (\mathbf{X}_i^T \hat{\Sigma}^{-1} \mathbf{X}_i)^{-1} \right].$$

Para mais detalhes, De Bastiani (2016) e Schemmer et al. (2017).

4.27 Teste de hipóteses

Na construção de um modelo espacial, é importante a aplicação de testes de hipóteses para verificar a significância dos parâmetros estimados, principalmente quando envolvem covariáveis. Assumindo-se que os dados $\mathbf{y}_1, \dots, \mathbf{y}_n$ são provenientes de uma amostra aleatória com distribuição t-Student reparametrizada $\mathcal{T}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \eta)$, com $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \boldsymbol{\phi}^T, \eta) \in \Theta$, sendo Θ o espaço paramétrico, $\Theta \subset \mathbb{R}^{p+4}$, em que $p+4$ representa a quantidade de parâmetros envolvidos. Segundo Mardia et al. (1992) e Anderson (2003), em uma análise multivariada clássica, isto é, sob a hipótese de normalidade, as hipóteses de interesse estão principalmente relacionadas com o contrastes sobre o vetor médio $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ e sobre os elementos da matriz de covariância $\boldsymbol{\Sigma}$. Neste estudo, especificamente, estamos interessados na hipótese linear definida como:

$$H_0 : \mathbf{H}\boldsymbol{\beta} = \mathbf{h},$$

em que \mathbf{H} é a matriz de ordem $q \times p$ com posto q e \mathbf{h} é o vetor $q \times 1$ ambos conhecidos. O modelo reduzido é definido por $\mathbf{Y}_i \sim \mathcal{T}_n(\mathbf{X}_H \boldsymbol{\beta}_H, \boldsymbol{\Sigma}, \eta)$, em que \mathbf{X}_H é a matriz sobre a hipótese H_0 , $\boldsymbol{\beta}_H$ é o vetor de parâmetros sobre a hipótese H_0 . Em que o estimador de $\boldsymbol{\theta}$ sobre a hipótese H_0 é dado por $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\beta}}_H, \tilde{\boldsymbol{\phi}}, \tilde{\eta})$, definido como estimador restrito de $\boldsymbol{\theta}$. Já o modelo completo é dado por $\mathbf{Y}_i \sim \mathcal{T}_n(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}, \eta)$, em que o estimador irrestrito de $\boldsymbol{\theta}$ é definido por $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\phi}}, \hat{\eta})$.

Para se obter $\boldsymbol{\theta}$ sobre a hipótese H_0 , dado $\tilde{\boldsymbol{\theta}}$, utilizam-se o mesmo algoritmo definido na seção 4.25, que obtém a estimativa para $\boldsymbol{\theta}$ no modelo completo. Porém, neste caso, utiliza-se as restrições definidas pela hipótese de interesse. Vários métodos assintóticos podem ser utilizados para testar a hipótese de interesse, neste caso, utilizam-se os testes para amostras grandes comumente usadas no contexto da análise de probabilidade. Sendo estes a razão

de verossimilhanças (RV), Wald (W), estatística de Rao score (S) denominado também de multiplicador de Lagrange e Gradiente (G), cujas distribuições assintóticas foram estudadas em Wilks (1938), Wald (1943), Rao (1948) e Terrell (2002), respectivamente. Segundo Fagundes (2017), estes testes exibem o mesmo comportamento assintótico sobre certas condições gerais e assumindo que H_0 é verdadeiro. A estatística RV mensura o quanto as evidências estatísticas corroboram com valores para o vetor de parâmetros diferentes daqueles especificados por H_0 , ou seja, altos valores da razão de verossimilhanças não apoiam a plausibilidade da hipótese nula (WILKS, 1938).

Por outro lado, Wald (1943) e Rao (1948) apresentaram as estatísticas Wald (W) e Rao (S), as quais comumente são utilizadas para medir a distância entre as hipóteses nula e alternativa. Pode-se observar em (4.45) que a vantagem da escolha da estatística W está relacionada principalmente à conveniência computacional, pois esse requer para seu cálculo somente os estimadores irrestritos do modelo, assim como sua matriz de covariância estimada. Todavia, os métodos RV e S definidos em (4.44) e (4.46) exigem a reestimativa do modelo no âmbito de cada restrição testada. Porém, ao contrário do teste RV, o teste W é não invariante perante reparametrizações ou restrições equivalentes, e seu desempenho para amostras pequenas não apresenta boa aproximação assintótica quando comparado com outros testes (TSAI; KOZIOL, 1988; CASTRO et al., 2008). Já a estatística S requer somente o estimador sobre o espaço restrito e este permanece invariante dependendo da forma como a matriz de informação de Fisher é estimada Dagenais e Dufour (1991).

Terrell (2002) propõe uma estatística de teste alternativo notavelmente simples, denominada estatística gradiente (G). Em contraste com as estatísticas de W e S, o teste de gradiente não necessita obter, estimar ou inverter a matriz de informação esperada, portanto, é uma alternativa bastante interessante devido à sua facilidade de cálculo (LEMONTE; FERRARI, 2012).

Para testar a hipótese de interesse, utilizam-se os testes de razão de verossimilhanças, Wald, Rao score e gradiente, definidos, respectivamente, da seguinte forma:

$$RV = 2 \left\{ \mathcal{L}(\hat{\theta}) - \mathcal{L}(\tilde{\theta}) \right\}, \quad (4.44)$$

$$W = r(\mathbf{H}\hat{\beta} - \mathbf{h})^\top \left\{ \mathbf{H}\mathcal{I}^{-1}(\hat{\theta})\mathbf{H}^\top \right\} (\mathbf{H}\hat{\beta} - \mathbf{h}), \quad (4.45)$$

$$S = \frac{1}{r} \mathbf{U}^\top(\tilde{\theta})\mathcal{I}^{-1}(\tilde{\theta})\mathbf{U}(\tilde{\theta}), \quad (4.46)$$

$$G = \mathbf{U}^\top(\tilde{\theta})(\hat{\theta} - \tilde{\theta}), \quad (4.47)$$

em que $\hat{\theta} = (\hat{\beta}^\top, \hat{\phi}^\top, \hat{\eta}^\top)^\top$ e $\tilde{\theta} = (\tilde{\beta}^\top, \tilde{\phi}^\top, \tilde{\eta}^\top)^\top$ são os estimadores de máximo verossimilhança irrestritos e restritos de θ . Sob condições de regularidade e sob hipótese nula H_0 , as estatísticas descritas por (4.44), (4.45), (4.46) e (4.47) têm distribuição assintótica Qui-quadrado.

4.27.1 Teste de hipótese para o modelo t-Student reparametrizado com r repetições independentes

Dentre os testes de hipóteses citados na seção 4.27, optou-se em se trabalhar com a estatística de razão de verossimilhanças. Dagenais e Dufour (1991) citam as vantagens deste teste como sendo a invariância perante a reparametrização e seu desempenho para amostras pequenas apresentando boa aproximação assintótica quando comparados com os demais testes. Define-se o teste da razão da verossimilhanças (RV) para o modelo t-Student reparametrizado com r repetições independentes como segue:

Seja $\mathbf{Y} = \text{vec}(Y_1(s), \dots, Y_r(s))$ o vetor aleatório $nr \times 1$ em que cada $Y_i(s) = (Y_i(s_1), \dots, Y_i(s_n))^T$. Suponha-se $Y_i(s) = \mathbf{X}_i(s)\boldsymbol{\beta} + \boldsymbol{\epsilon}_i(s)$, em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ é um vetor de ordem $p \times 1$ de parâmetros desconhecidos, $\mathbf{X}_i(s) = \mathbf{X}_i$ é a matriz $n \times p$. De Bastiani et al. (2017) apresentaram um caso particular $\mathbf{X}_i = \mathbf{X}$.

Seja $\mathcal{L}(\boldsymbol{\theta})$ o logaritmo da função verossimilhança total dado \mathbf{Y} que depende do vetor de parâmetros, $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \boldsymbol{\phi}^T)^T$ com $p+m$ componentes, em que $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T)^T$ pode particionar-se sendo $\boldsymbol{\beta}_1 = (\beta_1, \dots, \beta_q)^T$ de q dimensões e $\boldsymbol{\beta}_2 = (\beta_{q+1}, \dots, \beta_p)^T$ de dimensão $(p-q)$, em que $\boldsymbol{\beta}_1$ é o vetor de parâmetro de interesse, $\boldsymbol{\beta}_2$ é o vetor de parâmetros de perturbação, conseqüentemente a matriz $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2]$ é particionada, em que \mathbf{X}_1 é uma matriz $n \times q$ formada por q colunas de \mathbf{X} e \mathbf{X}_2 é uma matriz $n \times (p-q)$, formada por $(p-q)$ colunas de \mathbf{X} .

Seja $\boldsymbol{\beta}_1$ o vetor de parâmetros de interesse similarmente como Zucker et al. (2000), Melo et al. (2009) e De Bastiani et al. (2017), transformamos o vetor $\boldsymbol{\theta} = (\boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T, \boldsymbol{\phi}^T)^T$ em $\boldsymbol{\psi} = (\boldsymbol{\beta}_1^T, \boldsymbol{\xi}^T, \boldsymbol{\phi}^T)^T$ em que $\boldsymbol{\xi} = \boldsymbol{\beta}_2 + (\mathbf{X}_2^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_2)^{-1} \mathbf{X}_2^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_1 \boldsymbol{\beta}_1$, usando esta transformação $\boldsymbol{\beta}_1^T$ e $\boldsymbol{\nu} = (\boldsymbol{\xi}^T, \boldsymbol{\phi}^T)$ são ortogonais. Isto é

$$E \left[\frac{\partial^2 \mathcal{L}(\boldsymbol{\psi})}{\partial \boldsymbol{\beta}_1 \partial \boldsymbol{\xi}^T} \right] = 0 \quad \text{e} \quad E \left[\frac{\partial^2 \mathcal{L}(\boldsymbol{\psi})}{\partial \boldsymbol{\beta}_1 \partial \boldsymbol{\phi}^T} \right] = 0$$

Seja $\mathcal{L}(\boldsymbol{\theta})$ o logaritmo da função verossimilhança:

$$\mathcal{L}(\boldsymbol{\theta}) = \log \prod_{i=1}^r k_n(\eta) |\boldsymbol{\Sigma}|^{-1/2} [1 + c(\eta) \delta_i^2]^{-\frac{(1+n\eta)}{2\eta}}$$

e

$$\mathcal{L}(\boldsymbol{\psi}) = \log \prod_{i=1}^r k_n(\eta) |\boldsymbol{\Sigma}|^{-1/2} [1 + c(\eta) \delta_i^2]^{-\frac{(1+n\eta)}{2\eta}},$$

em que $\delta_i^2 = (\mathbf{Y}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y}_i - \boldsymbol{\mu})$ como $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} = \mathbf{X}_1 \boldsymbol{\beta}_1 + \mathbf{X}_2 \boldsymbol{\beta}_2$ tem-se que $\mathbf{X}\boldsymbol{\beta} = \mathbf{X}_1^* \boldsymbol{\beta}_1 + \mathbf{X}_2 \boldsymbol{\xi}$, sendo $\mathbf{X}_1^* = \mathbf{X}_1 - \mathbf{X}_2 (\mathbf{X}_2^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_2)^{-1} \mathbf{X}_2^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_1$. Assim,

$$\mathcal{L}(\boldsymbol{\psi}) = r \log(k_n(\eta)) - \frac{r}{2} \log |\boldsymbol{\Sigma}| - \frac{(1+n\eta)}{2\eta} \sum_{i=1}^r \log(1 + c(\eta) \delta_i^2),$$

em que $\delta_i^2 = (\mathbf{Y}_i - \mathbf{X}_1^* \boldsymbol{\beta}_1 - \mathbf{X}_2 \boldsymbol{\xi})^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y}_i - \mathbf{X}_1^* \boldsymbol{\beta}_1 - \mathbf{X}_2 \boldsymbol{\xi})$. Seja a hipótese de interesse

definida por:

$$\begin{cases} H_0 : \beta_1 = \beta_1^0, \text{ em que } \beta_1^0 \text{ é fixo, com } q \leq p \\ \text{versus} \\ H_1 : \beta_1 \neq \beta_1^0. \end{cases}$$

Sendo $\hat{\psi} = (\hat{\beta}_1^\top, \hat{\xi}^\top, \hat{\phi}^\top)^\top$ é o estimador ML irrestrito e $\tilde{\psi} = (\beta_1^{0\top}, \tilde{\xi}^\top, \tilde{\phi}^\top)^\top$ é o estimador ML restrito sob H_0 . Assim, a estatística da razão da verossimilhanças para testar a hipótese H_0 versus H_1 define-se por:

$$RV = 2 \left(\mathcal{L}(\hat{\psi}) - \mathcal{L}(\tilde{\psi}) \right),$$

em que $\mathcal{L}(\hat{\psi}) = r \log(k_n(\eta)) - \frac{r}{2} \log \left| \hat{\Sigma} \right| - \frac{(1+n\eta)}{2\eta} \sum_{i=1}^r \log(1 + c(\eta)\hat{\delta}_i^2)$, com $\hat{\delta}_i^2 = (\mathbf{Y}_i - \hat{\mathbf{X}}_1^* \hat{\beta}_1 - \mathbf{X}_2 \hat{\xi})^\top \hat{\Sigma}^{-1} (\mathbf{Y}_i - \hat{\mathbf{X}}_1^* \hat{\beta}_1 - \mathbf{X}_2 \hat{\xi})$ e $\mathcal{L}(\tilde{\psi}) = r \log(k_n(\eta)) - \frac{r}{2} \log \left| \tilde{\Sigma} \right| - \frac{(1+n\eta)}{2\eta} \sum_{i=1}^r \log(1 + c(\eta)\tilde{\delta}_i^2)$ com $\tilde{\delta}_i^2 = (\mathbf{Y}_i - \tilde{\mathbf{X}}_1^* \beta_1^0 - \mathbf{X}_2 \tilde{\xi})^\top \tilde{\Sigma}^{-1} (\mathbf{Y}_i - \tilde{\mathbf{X}}_1^* \beta_1^0 - \mathbf{X}_2 \tilde{\xi})$. Logo

$$\begin{aligned} RV &= 2 \left[-\frac{r}{2} \log \left| \hat{\Sigma} \right| + \frac{r}{2} \log \left| \tilde{\Sigma} \right| - \frac{(1+n\eta)}{2\eta} \sum_{i=1}^r \log(1 + c(\eta)\hat{\delta}_i^2) + \frac{(1+n\eta)}{2\eta} \sum_{i=1}^r \log(1 + c(\eta)\tilde{\delta}_i^2) \right] \\ &= 2 \left[\frac{r}{2} \log \left(\frac{\left| \tilde{\Sigma} \right|}{\left| \hat{\Sigma} \right|} \right) + \frac{(1+n\eta)}{2\eta} \sum_{i=1}^r \log \left(\frac{(1 + c(\eta)\tilde{\delta}_i^2)}{(1 + c(\eta)\hat{\delta}_i^2)} \right) \right]. \end{aligned}$$

Portanto, a razão de verossimilhança para a distribuição t-Student reparametrizada com repetição é dada por:

$$RV = r \log \left(\frac{\left| \tilde{\Sigma} \right|}{\left| \hat{\Sigma} \right|} \right) + \frac{(1+n\eta)}{\eta} \sum_{i=1}^r \log \left(\frac{(1 + c(\eta)\tilde{\delta}_i^2)}{(1 + c(\eta)\hat{\delta}_i^2)} \right)$$

em que $\hat{\delta}_i^2 = (\mathbf{Y}_i - \hat{\mathbf{X}}_1^* \hat{\beta}_1 - \mathbf{X}_2 \hat{\xi})^\top \hat{\Sigma}^{-1} (\mathbf{Y}_i - \hat{\mathbf{X}}_1^* \hat{\beta}_1 - \mathbf{X}_2 \hat{\xi})$, $\tilde{\delta}_i^2 = (\mathbf{Y}_i - \tilde{\mathbf{X}}_1^* \beta_1^0 - \mathbf{X}_2 \tilde{\xi})^\top \tilde{\Sigma}^{-1} (\mathbf{Y}_i - \tilde{\mathbf{X}}_1^* \beta_1^0 - \mathbf{X}_2 \tilde{\xi})$, $\hat{\mathbf{X}}_1^* = \mathbf{X}_1 - \mathbf{X}_2 (\mathbf{X}_2^\top \hat{\Sigma}^{-1} \mathbf{X}_2) \mathbf{X}_2^\top \hat{\Sigma}^{-1} \mathbf{X}_1$, $\tilde{\mathbf{X}}_1^* = \mathbf{X}_1 - \mathbf{X}_2 (\mathbf{X}_2^\top \tilde{\Sigma}^{-1} \mathbf{X}_2) \mathbf{X}_2^\top \tilde{\Sigma}^{-1} \mathbf{X}_1$.

4.28 Análise de diagnósticos

Um estudo de grande importância na análise de diagnósticos é a detecção de observações influentes, isto é, pontos que exercem um peso desproporcional nas estimações dos parâmetros do modelo ou até mesmo na significância dos parâmetros. A detecção de pontos, talvez seja a técnica mais conhecida para avaliar o impacto da retirada de uma observação particular nas estimativas da regressão.

Diversos trabalhos vêm sendo desenvolvidos utilizando procedimentos que detectam a presença de observações que exercem grande influência nos ajustes do modelos, por meio das técnicas de diagnósticos. Uribe-Opazo et al. (2012) utilizaram técnicas de diagnósticos para avaliar a sensibilidade dos estimadores de máxima verossimilhança, funções de covariância e preditor linear a pequenas perturbações nos dados e nos pressupostos do modelo linear espacial. Prado et al. (2013) trabalharam com influência local na variável resposta utilizando a

distribuição t -Student. Assumpção et al. (2014) apresentaram o estudo da influência local por meio de dois métodos conhecidos como deslocamento de verossimilhança e Q-deslocamento de verossimilhança, ambos utilizando a distribuição t -Student. De Bastiani et al. (2015) estudaram os métodos de influência na família das distribuições elípticas. Garcia-Papani et al. (2018) apresentaram técnicas de diagnósticos para avaliar a influência de observações atípicas nas estimativas de máxima verossimilhança de seus parâmetros com modelo de regressão espacial Birnbaum-Saunders. Entre os trabalhos de análise de diagnósticos, há os que apresentam dados com repetição. Smith (2001) apresentou o modelo de repetições múltiplas de um processo espacial e processo de estimação de parâmetros, De Bastiani et al. (2017) e De Bastiani et al. (2018) trabalharam com técnicas de diagnósticos para modelos espaciais gaussianos com repetição.

São duas as linhas de diagnósticos de influência: a influência global que consiste em avaliar o grau de sensibilidade do modelo ou da estimação dos parâmetros, quando uma observação considerada influente é eliminada do conjunto de dados, verifica-se o impacto dessa retirada; e a influência local propõe avaliar a influência conjunta das observações sob pequenas perturbações no modelo ou nos dados, usando uma medida de influência apropriada.

4.28.1 Influência global

A análise de influência global fundamenta-se na eliminação de uma ou mais observações consideradas influentes no conjunto de dados e assim, avaliar o impacto na estimação dos parâmetros. As técnicas de diagnósticos são discutidas por Cook e Weisberg (1982), Chatterjee e Hadi (2009) e Pan et al. (2014). Uma das medidas mais utilizadas para mensurar as mudanças dos parâmetros estimados, após excluídas as observações, é denominada de distância de Cook (COOK, 1977). Proposta inicialmente para modelos normais e rapidamente expandida para diversas classes de modelos. Denotada por:

$$D_{i\theta} = (\hat{\theta}_{[i]} - \hat{\theta})^\top M (\hat{\theta}_{[i]} - \hat{\theta}), \quad i = 1, \dots, r, \quad (4.48)$$

sendo $\hat{\theta}$ e $\hat{\theta}_{[i]}$ são os estimadores do vetor de parâmetros θ , com e sem a i -ésima observação respectivamente, e M é uma matriz definida positiva devidamente escolhida. A escolha está baseada diretamente no método de estimação, sendo relevante ao calcular $D_{i\theta}$.

4.28.2 Influência global baseada na verossimilhança

Seja $\mathcal{L}(\theta)$ o logaritmo da função verossimilhança denotado na equação (4.26). Denota-se $\mathcal{L}_{[i]}(\theta)$ como o logaritmo da função verossimilhança sem a i -ésima observação. Admite-se que $\hat{\theta}$ e $\hat{\theta}_{[i]}$ são, respectivamente, os estimadores de máxima verossimilhança de θ sobre $\mathcal{L}(\theta)$ e $\mathcal{L}_{[i]}(\theta)$. É possível calcular uma aproximação para $\hat{\theta}_{[i]}$ utilizando o método de Newton-Raphson a um passo, considera-se a seguinte equação atualizada.

$$\hat{\theta}_{[i]}^1 = \hat{\theta} + \left[-\ddot{\mathcal{L}}_{[i]}(\hat{\theta}) \right] \dot{\mathcal{L}}_{[i]}(\hat{\theta}), \quad i = 1, \dots, r, \quad (4.49)$$

em que $\dot{\mathcal{L}}_{[i]}(\hat{\theta})$ e $\ddot{\mathcal{L}}_{[i]}(\hat{\theta})$ são a primeira e segunda derivada da função log verossimilhança em relação ao parâmetro θ sem a i -ésima observação. Sendo

$$\mathcal{L}_{[i]}(\theta) = \sum_{\substack{k=1 \\ k \neq i}}^r \log K_n(\eta) - \frac{1}{2} \log |\Sigma_{[i]}| - \frac{1}{2\eta} (1 + n\eta) \log \left(1 + c(\eta) \delta_{[i]}^2 \right),$$

em que $\delta_{[i]}^2$ é a distância de Mahalanobis, $\Sigma_{[i]}$ é a matriz de covariância ambas sem i -ésima observação e $i = 1, \dots, r$ refere-se às repetições. Ao se substituir a equação (4.49), na equação (4.48) tem-se

$$D_{i\theta}^1 = \left[\dot{\mathcal{L}}_{[i]}(\hat{\theta}) \right]^\top \left[-\ddot{\mathcal{L}}_{[i]}(\hat{\theta}) \right]^{-1} \left[\dot{\mathcal{L}}_{[i]}(\hat{\theta}) \right]. \quad (4.50)$$

A equação (4.50) pode ser reescrita da forma:

$$D_{i\theta}^1 = \left[U_{[i]}(\hat{\theta}) \right]^\top \left[-L_{[i]}(\hat{\theta}) \right]^{-1} \left[U_{[i]}(\hat{\theta}) \right],$$

$M = \left[-L_{[i]}(\hat{\theta}) \right]$ é a matriz de covariância assintótica e $U_{[i]}(\hat{\theta})$ é a função escore, ambas sem a i -ésima observação. Pan et al. (2014) mostram que é possível substituir $L_{[i]}(\hat{\theta})$ por $L(\hat{\theta})$, e é uma alternativa viável para que se reduza significativamente a carga computacional. Pois, a estimação da matriz de covariância assintótica $L_{[i]}(\hat{\theta})$ torna-se computacionalmente pesada para amostras suficientemente grandes. Assim, a distância de Cook é representada por

$$D_{i\theta} = \left[U_{[i]}(\hat{\theta}) \right]^\top \left[-L(\hat{\theta}) \right]^{-1} \left[U_{[i]}(\hat{\theta}) \right]. \quad (4.51)$$

Porém, observa-se que esta substituição não garante que a matriz seja bloco diagonal, o que pode gerar dificuldades para decompor $D_{i\theta}$ em termos das componentes de interesse. Segundo Pan et al. (2014), para solucionar este problema, é possível utilizar a matriz de informação esperada de Fisher, ou seja, substituir $-L(\hat{\theta})$ por $\mathcal{I}(\hat{\theta}) = E[-L(\hat{\theta})]$ em (4.51), obtendo-se assim:

$$D_{i\theta}^* = \left[U_{[i]}(\hat{\theta}) \right]^\top \left[\mathcal{I}(\hat{\theta}) \right]^{-1} \left[U_{[i]}(\hat{\theta}) \right].$$

A função escore sem a i -ésima observação, avaliada em torno de $\theta = \hat{\theta}$, pode ser decomposta por:

$$U_{[i]}(\hat{\theta}) = \left(U_{[i]}(\hat{\beta}), U_{[i]}(\hat{\phi}), U_{[i]}(\hat{\eta}) \right)^\top.$$

Logo, ao assumir que os dados se ajustavam à distribuição t-Student reparametrizada multivariada à estatística de Cook pode ser escrita decompondo-se $D_{i\theta}^*$ como a soma das medidas de diagnósticos dos efeitos fixos β , componentes de variância ϕ e parâmetro de forma η . Dado por

$$D_{i\theta}^* = D_{i\beta}^* + D_{i\phi\eta}^*,$$

para $i = 1, \dots, r$, sendo

$$D_{i\beta}^* = \left[U_{[i]}(\hat{\beta}) \right]^\top \left[\mathcal{I}_{11}(\hat{\theta}) \right]^{-1} \left[U_{[i]}(\hat{\beta}) \right],$$

$$D_{i(\phi,\eta)}^* = \left[U_{[i]}(\hat{\phi})^\top, U_{[i]}(\hat{\eta})^\top \right] \left[\mathcal{I}_{22}(\hat{\theta}) \right]^{-1} \left[U_{[i]}(\hat{\phi}), U_{[i]}(\hat{\eta}) \right]^\top,$$

com $\mathcal{I}_{11}(\theta)$ e $\mathcal{I}_{22}(\theta)$ são apresentados na seção 4.23. Verifica-se que $\mathcal{I}_{11}(\hat{\theta}) = \mathcal{I}_{11}(\theta)|_{\theta=\hat{\theta}}$ e $\mathcal{I}_{22}(\hat{\theta}) = \mathcal{I}_{22}(\theta)|_{\theta=\hat{\theta}}$.

4.28.3 Influência global baseada na Q-function

Pan et al. (2014) propuseram uma abordagem de deleção de casos para identificar indivíduos influentes e observações influentes em modelos mistos lineares com base na Q-function. De Bastiani et al. (2018) também apresentam a Q-function na aplicação de diagnósticos de exclusão de caso para modelos lineares espaciais com repetição. Zhu e Lee (2001) estenderam a distância de Cook a partir do uso da esperança condicional do logaritmo da função verossimilhança completa, denominada Q-function, obtida no passo E do algoritmo EM, em que definem a matriz $M = -\ddot{Q}(\hat{\theta}|\hat{\theta})$ sendo $\ddot{Q}(\hat{\theta}|\hat{\theta}) = \frac{\partial^2 Q(\theta|\hat{\theta})}{\partial\theta\partial\theta^T}$ avaliado em $\theta = \hat{\theta}$. Propuseram a estimativa $\hat{\theta}_{[i]}$ de θ , pela seguinte aproximação

$$\hat{\theta}_{[i]} = \hat{\theta} + \left[-\ddot{Q}_{[i]}(\hat{\theta}|\hat{\theta}) \right] - \dot{Q}_{[i]}(\hat{\theta}|\hat{\theta}),$$

em que $Q_{[i]}(\hat{\theta}|\hat{\theta}) = Q(\theta|\hat{\theta})|_{\theta=\hat{\theta}}$ é a função Q-function formada sem a i -ésima observação, avaliada sobre o estimador de máxima verossimilhança em relação à θ , em que a primeira e segunda derivada da função são representadas respectivamente por:

$$\dot{Q}_{[i]}(\hat{\theta}|\hat{\theta}) = \left[\frac{\partial Q_{[i]}(\theta|\hat{\theta})}{\partial\theta} \right]_{\theta=\hat{\theta}} \quad e \quad \ddot{Q}_{[i]}(\hat{\theta}|\hat{\theta}) = \left[\frac{\partial^2 Q_{[i]}(\theta|\hat{\theta})}{\partial\theta\partial\theta^T} \right]_{\theta=\hat{\theta}}.$$

Semelhante ao caso discutido para verossimilhança, Zhu e Lee (2001) sugeriram substituir $\ddot{Q}_{[i]}(\hat{\theta}|\hat{\theta})$ por $\ddot{Q}(\hat{\theta}|\hat{\theta})$, reduzindo, desta forma, o tempo computacional. Assim, a estatística de Cook é expressa por

$$QD_{[i]}(\theta) = \left[\dot{Q}_{[i]}(\hat{\theta}|\hat{\theta}) \right]^T \left[-\ddot{Q}(\hat{\theta}|\hat{\theta}) \right]^{-1} \left[\dot{Q}_{[i]}(\hat{\theta}|\hat{\theta}) \right]. \quad (4.52)$$

O teorema proposto por Pan et al. (2014), demonstra que é possível substituir $-\ddot{Q}(\hat{\theta}|\hat{\theta})$ por $E[-\ddot{Q}(\hat{\theta}|\hat{\theta})]$ em (4.52), que representa a esperança da derivada de segunda ordem da Q-function. Esta substituição mantém a mesma acurácia para a aproximação da estimação de $\hat{\theta}_{[i]}$. Deste modo, a expressão modificada da estatística de Cook é dada por

$$QD_{[i]}^*(\theta) = \left[\dot{Q}_{[i]}(\hat{\theta}|\hat{\theta}) \right]^T \left[E \left[-\ddot{Q}(\hat{\theta}|\hat{\theta}) \right] \right]^{-1} \left[\dot{Q}_{[i]}(\hat{\theta}|\hat{\theta}) \right], \quad (4.53)$$

em que

$$\dot{Q}_{[i]}(\hat{\theta}|\hat{\theta}) = \begin{pmatrix} \dot{Q}_{i\beta} \\ \dot{Q}_{i\phi} \\ \dot{Q}_{i\eta} \end{pmatrix},$$

em que cada elemento de $\dot{Q}_{[i]}(\hat{\theta}|\hat{\theta})$ é representado por $\dot{Q}_{i\beta}$, $\dot{Q}_{i\phi}$ e $\dot{Q}_{i\eta}$ sendo a primeira derivada da função Q-function em relação ao parâmetro β , ϕ e η , porém, sem a i -ésima observação.

Para $E[-\ddot{Q}_{[i]}(\hat{\theta}|\hat{\theta})]$ avaliado em $\theta = \hat{\theta}$, definido por

$$E[-\ddot{Q}_{[i]}(\hat{\theta}|\hat{\theta})] = \begin{pmatrix} E[-\ddot{Q}_{i\beta\beta}] & E[-\ddot{Q}_{i\beta\phi}] & E[-\ddot{Q}_{i\beta\eta}] \\ E[-\ddot{Q}_{i\phi\beta}] & E[-\ddot{Q}_{i\phi\phi}] & E[-\ddot{Q}_{i\phi\eta}] \\ E[-\ddot{Q}_{i\eta\beta}] & E[-\ddot{Q}_{i\eta\phi}] & E[-\ddot{Q}_{i\eta\eta}] \end{pmatrix}.$$

Cada elemento da matriz $E[-\ddot{Q}_{[i]}(\hat{\theta}|\hat{\theta})]$ é definido por;

$$E[-\ddot{Q}_{i\beta\beta}] = w_i^k \mathbf{X}_i^\top \Sigma^{-1} \mathbf{X}_i,$$

$$E[-\ddot{Q}_{i\beta\phi}] = E[-\ddot{Q}_{i\phi\beta}] = \mathbf{0},$$

$$E[-\ddot{Q}_{i\beta\eta}] = E[-\ddot{Q}_{i\eta\beta}] = \mathbf{0},$$

$$E[-\ddot{Q}_{i\phi\phi}] = -\frac{1}{2} \text{tr} \left[\Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_k} - \Sigma^{-1} \frac{\partial^2 \Sigma}{\partial \phi_k \partial \phi_j} \right] + \frac{w_i^k}{2} \text{tr} \left[\Sigma^{-1} \left(\frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_k} - \frac{\partial^2 \Sigma}{\partial \phi_k \partial \phi_j} + \frac{\partial \Sigma}{\partial \phi_k} \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \right) \right],$$

$$E[-\ddot{Q}_{i\phi\eta}] = E[-\ddot{Q}_{i\eta\phi}] = \mathbf{0},$$

$$E[-\ddot{Q}_{i\eta\eta}] = -\left\{ \frac{1}{\eta^3} \log \left(\frac{1}{2c(\eta)} \right) + \frac{c(\eta)}{2\eta^4} - \frac{1}{\eta^3} \psi \left(\frac{1}{2\eta} \right) - \frac{1}{4\eta^4} \psi' \left(\frac{1}{2\eta} \right) + \frac{\eta - 3\eta^2}{(\eta^2(1-2\eta))^2} + \frac{c_i^{(k)}}{\eta^3} \right\}.$$

Mais detalhes são apresentados no Apêndice (4.37). Verifica-se que a matriz $E[-\ddot{Q}_{[i]}(\hat{\theta}|\hat{\theta})]$ é sempre bloco diagonal em relação aos parâmetros β , ϕ e η . Assim, a estatística modificada de Cook, baseada na Q-function apresentada na equação (4.53), pode ser decomposta na soma de três componentes

$$QD_{[i]}^*(\theta) = QD_{[i]}^*(\beta) + QD_{[i]}^*(\phi) + QD_{[i]}^*(\eta),$$

sendo

$$QD_{[i]}^*(\beta) = \left[\dot{Q}_{i\beta} \right]^\top \left[E[-\ddot{Q}_{i\beta\beta}] \right]^{-1} \left[\dot{Q}_{i\beta} \right],$$

$$QD_{[i]}^*(\phi) = [\dot{Q}_{i\phi}]^\top [E [-\ddot{Q}_{i\phi\phi}]]^{-1} [\dot{Q}_{i\phi}]$$

e

$$QD_{[i]}^*(\eta) = [\dot{Q}_{i\eta}]^\top [E [-\ddot{Q}_{i\eta\eta}]]^{-1} [\dot{Q}_{i\eta}].$$

Indicando que a medida de diagnósticos de θ , é dada da forma independente, a partir da soma das medidas de diagnósticos de cada parâmetro β , ϕ e η .

4.28.4 Influência local

Um problema que pode ocorrer com a exclusão individual de pontos é denominado de *masking effect*, consiste em deixar de detectar pontos conjuntamente influentes (PAULA, 2010). Uma alternativa foi apresentada por Cook (1986) quando propôs um enfoque geométrico para verificação das suposições do modelo assim como a identificação de pontos influentes, para analisar o efeito de introduzir pequenas perturbações no modelo ou nos dados, ao invés da avaliação da retirada individual ou conjunta de pontos. Este método foi denominado influência local.

A técnica de influência local tem se constituído em uma ferramenta útil para realizar análises de sensibilidades de modelos estatísticos, sendo amplamente utilizada em modelos de regressão linear e não linear. Em particular, para os modelos de regressão linear o método de influência local baseia-se no afastamento da verossimilhança e na função Q-afastamento (ZHU e LEE, 2001).

Galea et al. (1997) apresentaram estudos de influência local em modelos de regressão linear elíptica, baseados no deslocamento da verossimilhança. Zhu e Lee (2001) propuseram um método via algoritmo EM para avaliar a influência local em dados incompletos e definiram um ponto de corte para observações influentes, ao estudarem o comportamento local da Q-function. Liu (2002) e Galea et al. (2005) aplicaram a técnica de influência em modelos lineares elípticos multivariados sobre a função de verossimilhança, Osorio et al. (2007) aplicaram a influência local baseada no deslocamento da função verossimilhança, com perturbações na matriz escala e na variável resposta em modelos lineares elípticos com estrutura longitudinal. Osorio et al. (2009) apresentaram a estimação e a influência local, baseada na função Q-deslocamento para o modelo de Grubbs para distribuições com caudas pesadas.

Zeller et al. (2018) afirmam que a metodologia de influência local proposta por Cook (1986), em alguns casos, mostra que as aplicações dessa abordagem envolve extensas manipulações algébricas e intenso trabalho computacional. Assim, uma alternativa interessante foi proposta por Zhu e Lee (2001), os quais trabalham com medidas de influência baseada na Q-function. Contudo, pouco tem sido investigado a respeito da obtenção de medidas de influência local baseada na função Q-afastamento, denominada também como Q-function aplicada em modelos lineares com distribuição t-Student reparametrizado. Assim baseado no trabalho de Zeller et al. (2018) recomendamos o uso da Q-function.

4.28.5 Influência local baseada na Q-function

A abordagem geral desenvolvida em Zhu e Lee (2001) para a análise de influência local de modelos estatísticos gerais com dados perdidos, será utilizada para obter as medidas de diagnósticos para o modelo linear t-Student reparametrizado. Considera-se um vetor de perturbação ω , g -dimensional, o qual assume valores em um subconjunto aberto $\Omega \subseteq \mathbb{R}^g$ e o seguinte modelo estatístico perturbado $\mathcal{M} = \{f(\mathbf{Y}_c, \boldsymbol{\theta}, \omega) : \omega \in \Omega\}$, em que $f(\mathbf{Y}_c, \boldsymbol{\theta}, \omega)$ é a função de densidade de probabilidade para os dados completos, \mathbf{Y}_c , perturbado por ω e $\mathcal{L}_c(\boldsymbol{\theta}, \omega | \mathbf{Y}_c) = \log f(\mathbf{Y}_c, \boldsymbol{\theta}, \omega)$ correspondente à função de log-verossimilhança completa. Assume-se que existe um ω_0 tal que $\mathcal{L}_c(\boldsymbol{\theta}, \omega_0 | \mathbf{Y}_c) = \mathcal{L}_c(\boldsymbol{\theta} | \mathbf{Y}_c)$ para todo $\boldsymbol{\theta}$. Seja $\hat{\boldsymbol{\theta}}(\omega)$ o máximo da função $Q(\boldsymbol{\theta}, \omega | \hat{\boldsymbol{\theta}}) = E[\mathcal{L}_c(\boldsymbol{\theta}, \omega | \mathbf{Y}_c) | \mathbf{Y}, \hat{\boldsymbol{\theta}}]$. Para realizar o estudo de influência local, utiliza-se a função deslocamento, denominada Q-displacement. Esta função é definida por:

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

Seguindo a abordagem desenvolvida por Cook (1986) e Zhu e Lee (2001), o gráfico de influência local de $f_Q(\omega)$ é definido como $\alpha(\omega) = (\omega^\top, f_Q(\omega))^\top$. Neste caso, a curva normal $C_{f_Q, d}$ de $\alpha(\omega)$ em ω_0 na direção de algum vetor unitário d pode ser usada para resumir o comportamento local de $f_Q(\omega)$. A curva normal $C_{f_Q, d}$ para $\alpha(\omega)$ e ω_0 é definida como:

$$C_{f_Q, d} = -2d^\top \ddot{Q}_{\omega_0} d,$$

com

$$-\ddot{Q}_{\omega_0} = \Delta_{\omega_0}^\top \left[-\ddot{Q}(\hat{\boldsymbol{\theta}}) \right]^{-1} \Delta_{\omega_0},$$

sendo $\ddot{Q}(\hat{\boldsymbol{\theta}})$ e Δ_{ω_0} dado por:

$$\ddot{Q}(\hat{\boldsymbol{\theta}}) = \left. \frac{\partial^2 Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right|_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}} \quad \text{e} \quad \Delta_{\omega_0} = \left. \frac{\partial^2 Q(\boldsymbol{\theta}, \omega | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \omega^\top} \right|_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}, \omega = \omega_0}.$$

Cook (1986) e Zeller et al. (2018) destacam que a matriz simétrica $-2\ddot{Q}_{\omega_0}$ é fundamental para a detecção de observações influentes, e sua decomposição espectral é dada por $-2\ddot{Q}_{\omega_0} = \sum_{j=1}^g \zeta_j \epsilon_j \epsilon_j^\top$, em que $\{(\zeta_j, \epsilon_j), j = 1, \dots, g\}$ são pares de autovalores e autovetores da matriz $-2\ddot{Q}_{\omega_0}$, com $\zeta_1 \geq \dots \geq \zeta_{r^*} > \zeta_{r^*+1} = \dots = 0$ e autovetores ortonormais $\{\epsilon_j, j = 1, \dots, g\}$. Lesaffre e Verbeke (1998) e Zhu e Lee (2001) propuseram inspecionar todos os autovetores correspondentes a autovalores não nulos para informações mais reveladoras. Zhu e Lee (2001) e Zeller et al. (2018) consideraram o vetor de contribuição agregado de todos os autovetores correspondentes a autovalores não nulos, sendo $\tilde{\zeta}_j = \zeta_j / (\zeta_1 + \dots + \zeta_{r^*})$, $\epsilon_j^2 = (\epsilon_{j1}^2, \dots, \epsilon_{jg}^2)^\top$ e $M(0) = \sum_{j=1}^{r^*} \tilde{\zeta}_j \epsilon_j^2$, em que o l -ésimo componente de $M(0)$, definido por $M(0)_l$, é igual a $\sum_{j=1}^{r^*} \tilde{\zeta}_j \epsilon_{jl}^2$. A avaliação de casos influente é baseada na inspeção visual do $\{M(0)_l, l = 1, \dots, g\}$ plotado contra o índice l . O caso l -ésimo pode ser considerado como influente se $M(0)_l$ for maior que o referência.

Zeller et al. (2018) destaca a inconveniência envolvida no uso da curvatura normal para decidir sobre a influência das observações, em que $C_{f_Q, d}(\boldsymbol{\theta})$ pode assumir qualquer valor,

também não é invariante sob uma mudança uniforme de escala. Zhu e Lee (2001) consideraram a seguinte curvatura normal conformal $B_{f_Q,d}(\boldsymbol{\theta}) = C_{f_Q,d}(\boldsymbol{\theta})/tr[-2\ddot{Q}_{\omega_0}]$, que tem uma propriedade interessante $0 \leq B_{f_Q,d}(\boldsymbol{\theta}) \leq 1$, para qualquer direção unitária d . Isso permite a comparação da curvatura entre diferentes modelos (ASSUMPÇÃO et al., 2014).

Baseados em De Bastiani et al. (2017), denotamos a curvatura normal conformal $B_i = 2|b_{jj}|$, na direção unitária com a j -ésima entrada 1 e as demais entradas 0, em que b_{jj} são os elementos da diagonal principal da matriz $B_{f_Q,d}(\boldsymbol{\theta})$. De acordo com Zhu e Lee (2001), a j -ésima observação é potencialmente influente se $B_i > \bar{B} + 2sd(B)$, em que $\bar{B} = \sum_{i=1}^n B_i/n$ e $sd(B)$ é o desvio padrão de B_1, \dots, B_n . Define-se d_{Qmax} o autovetor, normalizado, associado ao maior autovalor, em módulo da matriz $B_{f_Q,d}(\boldsymbol{\theta})$, sendo esse autovetor utilizado na construção do gráfico dos elementos $|d_{Qmax}|$ versus i (ordem dos dados).

A seguir, apresenta-se uma breve discussão sobre os esquemas de perturbação empregados neste estudo.

4.28.6 Seleção de esquemas de perturbação

Uma questão fundamental ao ser aplicada a análise de influência local é a seleção do esquema de perturbação, pois perturbar um modelo de forma arbitrária pode induzir a inferências enganosas. Chen et al. (2009) propuseram um método para selecionar esquemas de perturbação em modelos estatísticos para dados completos, com base na matriz de informação de Fisher para $\boldsymbol{\omega}$, que é $G(\boldsymbol{\omega}) = (g_{il}(\boldsymbol{\omega})) = E_{\boldsymbol{\omega}}(\mathbf{U}_{\boldsymbol{\omega}}\mathbf{U}_{\boldsymbol{\omega}}^{\top}) = -E_{\boldsymbol{\omega}}\{\partial^2 \mathcal{L}_c(\boldsymbol{\theta}, \boldsymbol{\omega} | \mathbf{Y}_c)/\partial \boldsymbol{\omega} \partial \boldsymbol{\omega}^{\top}\}$, com

$$g_{il}(\boldsymbol{\omega}) = E_{\boldsymbol{\omega}} \left\{ \frac{\partial}{\partial \omega_i} \mathcal{L}_c(\boldsymbol{\theta}, \boldsymbol{\omega} | \mathbf{Y}_c) \frac{\partial}{\partial \omega_l} \mathcal{L}_c(\boldsymbol{\theta}, \boldsymbol{\omega} | \mathbf{Y}_c) \right\}, \quad i, l = 1, \dots, g,$$

em que $E_{\boldsymbol{\omega}}$ denota a esperança em relação à densidade do modelo perturbado $f(\mathbf{Y}_c, \boldsymbol{\theta}, \boldsymbol{\omega})$, e $\mathbf{U}_{\boldsymbol{\omega}} = \partial \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega} | \mathbf{Y}_c)/\partial \boldsymbol{\omega}$ a função escore perturbada. Os elementos diagonais $\{g_{ii}(\boldsymbol{\omega}), i = 1, \dots, g\}$ de $G(\boldsymbol{\omega})$ são as variâncias da função escore em relação aos elementos de $\boldsymbol{\omega}$, que indicam a quantidade de perturbação do componente i -ésimo de $\boldsymbol{\omega}$. Os elementos fora da diagonal de $G(\boldsymbol{\omega})$ são as covariâncias da função escore em relação aos componentes de $\boldsymbol{\omega}$ e representam a associação linear entre os diferentes componentes de $\boldsymbol{\omega}$. Note-se que $G(\boldsymbol{\omega})$ é uma matriz diagonal, os componentes de $\boldsymbol{\omega}$ são ortogonais, sendo esta definição de ortogonalidade apresentada em (COX; REID, 1987).

A perturbação $\boldsymbol{\omega}$ é considerada apropriada se satisfaz $G(\boldsymbol{\omega}_0) = c\mathbf{I}_g$, em que $c > 0$. Isso garante que os componentes de $\boldsymbol{\omega}$ sejam assintoticamente independentes e, se $\boldsymbol{\omega}$ satisfazem $G(\boldsymbol{\omega}_0) = \text{Diag}(g_{11}(\boldsymbol{\omega}_0), \dots, g_{gg}(\boldsymbol{\omega}_0))$, podemos sempre escolher um novo esquema de perturbação $\tilde{\boldsymbol{\omega}}$, definido por $\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega}_0 + c^{-1/2}G(\boldsymbol{\omega}_0)^{1/2}(\boldsymbol{\omega} - \boldsymbol{\omega}_0)$, tal que $G(\tilde{\boldsymbol{\omega}})$ avaliado em $\boldsymbol{\omega}_0$ é igual a $c\mathbf{I}_g$. Se os esquemas de perturbação são considerados, a densidade perturbada assume a seguinte forma:

$$f(\mathbf{Y}_c, \boldsymbol{\theta}, \boldsymbol{\omega}) = \prod_{i=1}^r f(\mathbf{Y}_{ic}, \boldsymbol{\theta}, \omega_i), \quad (4.54)$$

em que $\boldsymbol{\omega} = (\omega_1, \dots, \omega_r)$ e $f(\mathbf{Y}_{ic}, \boldsymbol{\theta}, \omega_i)$ para $i = 1, \dots, r$, com as mesmas formas funcionais.

Entre esses esquemas, pode-se citar ponderação de casos, perturbação na variável resposta e perturbação na matriz escala.

Lema 1. *Se um esquema de perturbação é tal que a função de densidade perturbada pode ser escrita como (4.54), então*

$$\mathbf{G}(\omega_0) = c\mathbf{I}_n,$$

em que $c = E_{\omega} \{ \partial \log f(\mathbf{Y}_{ic}, \boldsymbol{\theta}, \omega_i) / \partial \omega_i \}^2$, avaliado em $\omega = \omega_0$.

Prova: Sob o modelo perturbado definido por (4.54), tem-se

$$g_{il}(\omega) = \begin{cases} E_{\omega} \{ \partial \log f(\mathbf{Y}_{ic}, \boldsymbol{\theta}, \omega_i) / \partial \omega_i \} E_{\omega} \{ \partial \log f(\mathbf{Y}_{lc}, \boldsymbol{\theta}, \omega_l) / \partial \omega_l \}, & \text{se } i \neq l, \\ E_{\omega} \{ \partial \log f(\mathbf{Y}_{ic}, \boldsymbol{\theta}, \omega_i) / \partial \omega_i \}^2, & \text{se } i = l. \end{cases}$$

Como $E_{\omega} \{ \partial \log f(\mathbf{Y}_{ic}, \boldsymbol{\theta}, \omega_i) / \partial \omega_i \} = 0$, $i = 1, \dots, r$, então $g_{il}(\omega) = c(\omega_i)\delta_{il}$ e $g_{il}(\omega_0) = c\delta_{il}$, em que δ_{il} denota o delta de Kronecker, $c(\omega_i) = E_{\omega} \{ \partial \log f(\mathbf{Y}_{ic}, \boldsymbol{\theta}, \omega_i) / \partial \omega_i \}^2$, e $c = c(\omega_i)$, avaliado em $\omega = \omega_0$. Assim, para os três esquemas de perturbação anteriores, temos que $\mathbf{G}(\omega_0) = c\mathbf{I}_n$. Para mais detalhes, veja Chen et al. (2009).

4.28.7 Esquemas de perturbação

Os esquemas de perturbação em estudo são baseados em Zeller et al. (2018). Consideram-se três diferentes esquemas de perturbação para o modelo linear t-Student reparametrizado: ponderação de casos, perturbação na variável resposta e perturbação na matriz escala.

4.28.8 Ponderação de casos

Este esquema de perturbação permite identificar as observações que exercem forte influência no processo de estimação e conseqüentemente nas estimativas dos parâmetros. A Q-function perturbada para a ponderação de casos é definida por:

$$Q(\boldsymbol{\theta}, \omega | \hat{\boldsymbol{\theta}}) = \sum_{i=1}^r \omega_i Q_i(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}) = \sum_{i=1}^r \omega_i \left\{ Q_{1i}(\boldsymbol{\theta}_1 | \hat{\boldsymbol{\theta}}) + Q_{2i}(\boldsymbol{\theta}_2 | \hat{\boldsymbol{\theta}}) \right\} \quad (4.55)$$

em que $\omega = (\omega_1, \dots, \omega_r)^\top$ denota o vetor de perturbação que pertence ao subconjunto aberto $\Omega \subseteq \mathbb{R}^r$, com $0 \leq \omega_i \leq 1$ em que $i = 1, \dots, r$. Para este caso, $\omega_0 = \mathbf{1}$ é o vetor não perturbado tal que $Q(\boldsymbol{\theta}, \omega | \hat{\boldsymbol{\theta}}) = Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}})$. Para a Q-function perturbada definida em (4.55), considera-se $\boldsymbol{\theta}_1 = (\boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top)^\top$ e $\boldsymbol{\theta}_2 = \eta$, em que $Q_{1i}(\boldsymbol{\theta}_1 | \hat{\boldsymbol{\theta}})$ e $Q_{2i}(\boldsymbol{\theta}_2 | \hat{\boldsymbol{\theta}})$ são dados por:

$$Q_{1i}(\boldsymbol{\theta}_1 | \hat{\boldsymbol{\theta}}) = -\frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{w_i^k}{2} \delta_i^2$$

e

$$Q_{2i}(\boldsymbol{\theta}_2 | \hat{\boldsymbol{\theta}}) = +\frac{1}{2\eta} \log \left(\frac{1}{2c(\eta)} \right) - \log \left(\Gamma \left(\frac{1}{2\eta} \right) \right) + \frac{1}{2c(\eta)} c_i^{(k)}$$

Alternativamente, podemos pensar em outras duas sub-perturbações: i) $Q(\theta, \omega|\hat{\theta}) = \sum_{i=1}^r \left\{ \omega_i Q_{1i}(\theta_1|\hat{\theta}) + Q_{2i}(\theta_2|\hat{\theta}) \right\}$ e ii) $Q(\theta, \omega|\hat{\theta}) = \sum_{i=1}^r \left\{ Q_{1i}(\theta_1|\hat{\theta}) + \omega_i Q_{2i}(\theta_2|\hat{\theta}) \right\}$, que são avaliadas com base no $C_{f_{Q,d}}(\theta_1)$ e $C_{f_{Q,d}}(\theta_2)$, respectivamente. Sob este esquema de perturbação, as matrizes $\Delta_{1\beta}$, $\Delta_{1\phi}$ e $\Delta_{2\eta}$ são dadas por $\Delta_{1\beta} = (\Delta_{11\beta}, \dots, \Delta_{1r\beta})^\top$ com dimensão $p \times r$, $\Delta_{1\phi} = (\Delta_{11\phi}, \dots, \Delta_{1r\phi})^\top$ com dimensão $3 \times r$ e $\Delta_{2\eta} = (\Delta_{21\eta}, \dots, \Delta_{2r\eta})^\top$ com dimensão $1 \times r$, sendo

$$\Delta_{1\beta} = \frac{\partial^2 Q(\theta, \omega|\hat{\theta})}{\partial \beta \partial \omega^\top}, \quad \Delta_{1\phi} = \frac{\partial^2 Q(\theta, \omega|\hat{\theta})}{\partial \phi \partial \omega^\top} \quad \text{e} \quad \Delta_{2\eta} = \frac{\partial^2 Q(\theta, \omega|\hat{\theta})}{\partial \eta \partial \omega^\top}.$$

Como $\Delta_{1\beta}$, $\Delta_{1\phi}$ e $\Delta_{2\eta}$ são avaliadas em $\omega = \omega_0$, $\theta = \hat{\theta}$, tendo como elementos

$$\Delta_{1i\beta} = w_i^k \mathbf{X}_i^\top \Sigma^{-1} \varepsilon_i, \\ \Delta_{1i\phi} = \frac{1}{2} D_\phi^\top \text{vec}[(w_i^k \Sigma^{-1} \varepsilon_i \varepsilon_i^\top \Sigma^{-1}) - \Sigma^{-1}], \quad \Delta_{2i\eta} = -\frac{1}{2\eta^2} \left[\log\left(\frac{1}{2c(\eta)}\right) + \frac{c(\eta)}{\eta} - \psi\left(\frac{1}{2\eta}\right) + c_i^{(k)} \right],$$

sendo $\Delta_{1i\beta}$ com dimensão $p \times 1$, $\Delta_{1i\phi}$ com dimensão 3×1 e $\Delta_{2i\eta}$ é escalar. Os cálculos são apresentados no Apêndice 4.38.1. A curvatura normal $C_{f_{Q,d}}$ para o caso da ponderação de casos é dada por:

$$C_{f_{Q,d}} = C_{f_{Q,d}}(\theta_1) + C_{f_{Q,d}}(\theta_2) \quad (4.56)$$

como $\theta_1 = (\beta, \phi)^\top$ e $\theta_2 = \eta$, em que cada elemento da equação (4.56) é definido por $C_{f_{Q,d}}(\theta_1) = 2\mathbf{d}^\top \Delta_{1\theta_1}^\top (\ddot{Q}_{11}) \Delta_{1\theta_1} \mathbf{d}$, e $C_{f_{Q,d}}(\theta_2) = 2\mathbf{d}^\top \Delta_{2\theta_2}^\top (\ddot{Q}_{22}) \Delta_{2\theta_2} \mathbf{d}$, sendo \ddot{Q}_{11} e \ddot{Q}_{22} apresentados no Apêndice 4.38.7

4.28.9 Perturbação na variável resposta

Este esquema de perturbação pode ser utilizado quando o objetivo for avaliar a sensibilidade das estimativas ao serem introduzidas pequenas perturbações nos componentes de cada vetor de resposta (BORSSOI, 2013).

Sob esta perturbação, um ou todos os vetores \mathbf{Y}_i são modificados considerando o esquema de perturbação aditiva. A perturbação da variável resposta $\mathbf{Y}_i = (\mathbf{Y}_1^\top, \dots, \mathbf{Y}_r^\top)^\top$ é introduzida pela substituição \mathbf{Y}_i por $\mathbf{Y}_i(\omega)$. Aqui, considera-se a perturbação na variável produtividade $\mathbf{Y}_i(\omega) = \mathbf{Y}_i + \mathbf{A}\omega_i$ em que \mathbf{A} é uma matriz de ordem $n \times n$, que não depende de β nem de ω_i . Para este caso, o vetor não perturbado é definido por $\omega_0 = \mathbf{0}$. A função de log-verossimilhança perturbada para os dados completos é dada por

$$Q(\theta, \omega|\hat{\theta}) = \sum_{i=1}^r \left\{ Q_{1i}(\theta_1, \omega_i|\hat{\theta}) + Q_{2i}(\theta_2|\hat{\theta}) \right\},$$

em que

$$Q_{1i}(\theta_1, \omega_i|\hat{\theta}) = -\frac{1}{2} \log |\Sigma| - \frac{w_i^k}{2} \delta_{i\omega}^2$$

com $\delta_{i\omega}^2 = \varepsilon_{i\omega}^\top \Sigma^{-1} \varepsilon_{i\omega}$ sendo $\varepsilon_{i\omega} = [(\mathbf{Y}_i + \mathbf{A}\omega_i) - \mathbf{X}_i\beta]$ e $\mathbf{A} = \Sigma^{1/2}$. Sabe-se que $Q_{2i}(\theta_2|\hat{\theta})$ não depende de ω . O procedimento de influência local é baseado em $C_{f_{Q,d}}(\theta_1)$, em que

$\Delta_{\theta_1} = (\Delta_{\beta}^{\top}, \Delta_{\phi}^{\top})$ em que $\Delta_{\beta} = (\Delta_{1\beta}, \dots, \Delta_{r\beta})$ é uma matriz $p \times nr$ e $\Delta_{\phi} = (\Delta_{1\phi}, \dots, \Delta_{r\phi})$ é uma matriz $3 \times nr$, temos que

$$\Delta_{i\beta} = w_i X_i \Sigma^{-1} \text{ e } \Delta_{i\phi_j} = w_i \varepsilon_i^{\top} \Sigma^{-1/2} \frac{\partial \Sigma^{1/2}}{\partial \phi_j} \Sigma^{-1/2}.$$

Sendo $\Delta_{i\beta}$ e $\Delta_{i\phi_j}$ com dimensões $p \times n$ e $3 \times n$, respectivamente. Para $\Delta_{i\phi_j}$, tem-se $j = 1, 2$ e 3 , logo:

$$\Delta_{i\phi_1} = w_i^{\top} \varepsilon_i^{\top} \Sigma^{-1/2} P (J_1 \odot Q) P^{\top} \Sigma^{-1/2},$$

$$\Delta_{i\phi_2} = w_i^{\top} \varepsilon_i^{\top} \Sigma^{-1/2} P (J_2 \odot Q) P^{\top} \Sigma^{-1/2},$$

$$\Delta_{i\phi_3} = w_i^{\top} \varepsilon_i^{\top} \Sigma^{-1/2} P (J_3 \odot Q) P^{\top} \Sigma^{-1/2}.$$

em que $\frac{\partial \Sigma^{1/2}}{\partial \phi_j} = P (J_j \odot Q) P^{\top}$ é apresentado no Apêndice 4.38.4. O estudo da curvatura normal para perturbação na variável resposta é dado por:

$$C_{f_Q, d}(\theta_1) = 2d^{\top} \Delta_{\theta_1}^{\top} (\ddot{Q}_{11}) \Delta_{\theta_1} d,$$

em que \ddot{Q}_{11} é apresentado no Apêndice 4.38.7.

4.28.10 Perturbação na matriz escala

O interesse do esquema de perturbação da matriz escala é revelar observações que interferem na modelagem e na estrutura de variância bem como indicar os elementos influentes na estimação do parâmetro ϕ (OSORIO et al., 2009).

Para analisar a sensibilidade da estimativa ML, implementou-se um esquema de perturbação na matriz escala, substituindo-se Σ por $\omega_i^{-1} \Sigma$, em que $\omega = (\omega_1, \dots, \omega_r)^{\top}$ em que ω é o vetor de dimensão $rn \times 1$, sendo ω_i de ordem $n \times 1$ com $i = 1, \dots, r$. O modelo não perturbado é obtido quando $\omega_0 = \mathbf{1}$, sendo ω_0 o vetor de dimensão $n \times 1$. Além disso, a Q-function assume a forma:

$$Q(\theta, \omega | \hat{\theta}) = \sum_{i=1}^r Q_i(\theta, \omega | \hat{\theta}) = \sum_{i=1}^r \left\{ Q_{1i}(\theta_1, \omega | \hat{\theta}) + Q_{2i}(\theta_2, \hat{\theta}) \right\}$$

em que

$$Q_{1i}(\theta_1, \omega | \hat{\theta}) = -\frac{1}{2} \log |\omega_i^{-1} \Sigma| - \frac{w_i^k}{2} \delta_{i\omega}^2 = -\frac{1}{2} \log |\Sigma| + \frac{n}{2} \log(\omega_i^{-1}) - \frac{w_i^k}{2} \delta_{i\omega}^2$$

sendo $\delta_{i\omega}^2 = \omega_i \varepsilon_i^{\top} \Sigma^{-1} \varepsilon_i$. Pode-se verificar que, para este esquema de perturbação, a função $Q_{2i}(\theta_2, \omega | \hat{\theta})$ não depende de ω . Segundo Osorio (2006), os esquemas de perturbação da variável resposta e matriz escala são semelhantes.

4.29 Material e Métodos

4.29.1 Área de estudo e dados

Utilizou-se o modelo espacial linear t-Student reparametrizado em que as repetições são consideradas independentes no espaço-tempo. Segundo Goodal e Mardia (1994), pode-se remover o efeito do tempo nos dados, portanto, podem-se considerar os dados como medições repetidas independentes. De acordo com Stein (1999), pode-se supor que as observações de diferentes tempos são realizações independentes em um campo aleatório, sendo que o processo de espaço-tempo observado seja um conjunto em locais espaciais, em tempo suficiente e pontos distantes no tempo. De Bastiani et al. (2017) afirmam que a independência dos dados vem do longo período considerado de uma realização e outra. Também, quando se trabalha com dados agrícolas, tem-se a aplicação de insumos para o melhor preparo do solo a cada ano. Baseados em Goodal e Mardia (1994), Stein (1999) e De Bastiani et al. (2017) definiram-se as repetições independentes em cinco anos agrícolas sendo 2012/2013, 2013/2014, 2014/2015, 2015/2016 e 2016/2017. A variável resposta em estudo é a produtividade da soja e as covariáveis são Cálcio (Ca), Magnésio (Mg), Potássio (K) e Fósforo (P), consideradas atributos químicos do solo.

A área comercial de produção de grãos em estudo contém 127,18 ha, em que o sistema de plantio direto é praticado desde 1994. Está localizada no município de Cascavel, cuja posição geográfica é 24,95° Sul de latitude, 53,57° Oeste de longitude. O solo é classificado como Latossolo Vermelho Distroférrico (SOLOS, 2013). O clima da região é classificado como temperado mesotérmico e superúmido, tipo climático Cfa (Köppen). Na Figura 28 estão ilustradas a área experimental e a grade amostral, totalizando 74 observações, em que cada ponto representa uma amostra. Todas as amostras foram georreferenciadas com o auxílio de um receptor de dados GPS (*Global Positioning System*) GEOEXPLORE 3, em um sistema de coordenadas (UTM).

É importante ressaltar que esta área vem sendo estudada desde do ano 2004, pelos grupos de pesquisa do Laboratório de Estatística Espacial (LEE) e do Laboratório de Estatística Aplicada (LEA) UNIOESTE, Campus Cascavel. Assim, as variáveis em estudo, produtividade da soja (Prod) [$t\ ha^{-1}$] e os dados químicos Cálcio (Ca) [$cmolc\ dm^{-3}$], Magnésio (Mg) [$cmolc\ dm^{-3}$], Potássio (K) [$cmolc\ dm^{-3}$] e Fósforo (P) [$mg\ dm^{-3}$], são referentes aos cinco anos agrícolas citados acima.

4.29.2 Coleta de dados

A coleta de dados durante os anos agrícolas 2012/2013, 2013/2014, 2014/2015, 2015/2016 e 2016/2017 (vale ressaltar que o ano agrícola da soja na região Oeste do Paraná está definida de setembro a fevereiro), ocorreu da seguinte forma:

A produtividade da soja foi estimada considerando-se a quantidade de soja colhida de todas as plantas distribuídas em duas leiras ao longo de um metro de comprimento, representando a parcela. Após a trilha, foi feita a pesagem de grãos para cada parcela e

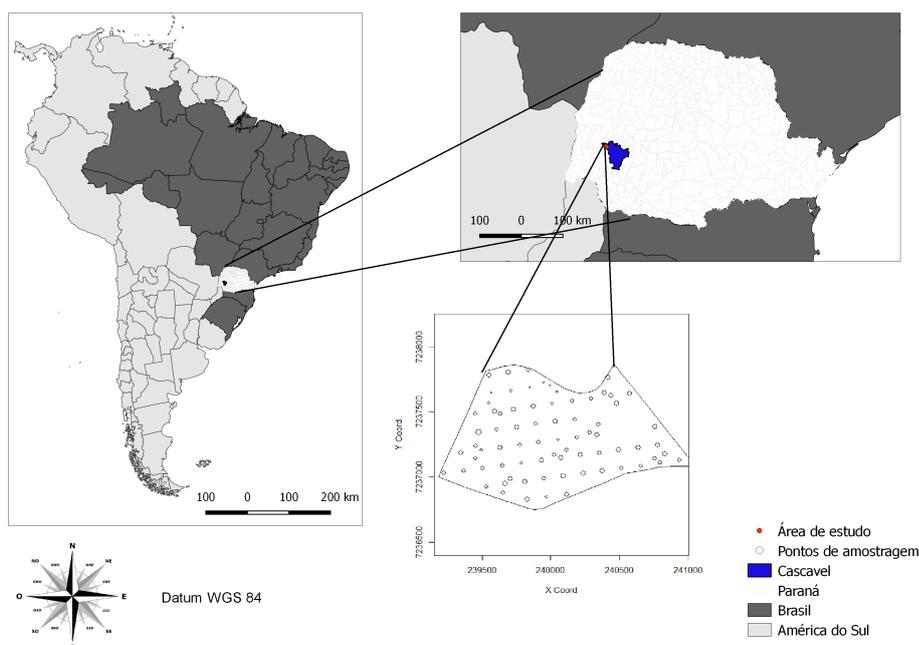


Figura 28 Localização da área em estudo.

verificado o teor de água para posterior correção para 13%. O valor foi também convertido em $t\ ha^{-1}$.

Os atributos químicos foram obtidos a partir de amostragem do solo para determinação dos níveis de cada atributo e realizada em cada ponto demarcado na Figura 28. Foram coletadas quatro subamostras de solo de 0,00 a 0,20 m de profundidade nas proximidades de cada ponto, misturadas e colocadas em sacolas plásticas, com aproximadamente 500 g, compondo assim uma amostra representativa da parcela. As amostras foram encaminhadas ao laboratório da Cooperativa Central de Desenvolvimento Tecnológico e Econômico Ltda (COODETEC-Brasil), em que se procedeu a análise química.

A utilização dos atributos químicos como variáveis explicativas é uma prática comum em estudos de campo, veja Garcia-Papani et al. (2017), Schemmer et al. (2017), Garcia-Papani et al. (2018), Fagundes et al. (2018a), Fagundes et al. (2018b), Kbakural et al. (1999) explicaram esta utilização, pois as variações nas propriedades físicas e químicas do solo são responsáveis por grande parte das variações do rendimento da cultura. Reichert et al. (2008) destacam que o estudo do comportamento das propriedades químicas do solo é um fator importante tanto para o planejamento como para a avaliação de lavouras comerciais.

4.29.3 Resultados

Foram analisados os dados de produtividade da soja e quatro atributos químicos, considerando o seguinte caso: o ano agrícola 2012/2013 refere-se a Y_1 representando a 1ª repetição; 2013/2014 refere-se a Y_2 representando a 2ª repetição; 2014/2015 refere-se a Y_3 representando a 3ª repetição; 2015/2016 refere-se a Y_4 representando a 4ª repetição e 2016/2017 refere-se a Y_5 representando a 5ª repetição. Cada repetição contém 74 amostras, as

quais estão posicionadas na mesma localização para cada repetição, como apresentado na Figura 28.

A Tabela 12 apresenta uma análise descritiva do conjunto de dados da produtividade, e observa-se que a menor média da produtividade da soja é apresentada em Y_3 correspondente a $2,376 t ha^{-1}$ e a maior em Y_4 é referente a $4,227 t ha^{-1}$. As maiores amplitudes entre os dados estão em Y_2, Y_5 e Y_1 como pode-se constatar no Boxplot da Figura 30 (a), também na apresentação das Figuras 29 (b), (e) e (a) referentes ao Scatter plot dos mesmos. Porém, ao visualizar os Scatter plots, atente-se à escala de cada gráfico.

Tabela 12 Análise descritiva da produtividade da soja referente a cada repetição

Prod	Mín	Q_1	Mediana	Média	Q_3	Max
Y_1	2,239	2,948	3,261	3,259	3,530	4,511
Y_2	2,913	3,841	4,212	4,227	4,527	5,766
Y_3	1,874	2,179	2,340	2,376	2,557	3,176
Y_4	0,674	2,309	2,509	2,458	2,606	2,834
Y_5	1,518	2,799	3,151	3,057	3,438	4,030

Mín: mínimo, Q_1 : primeiro quartil, Q_3 : terceiro quartil, Max: máximo

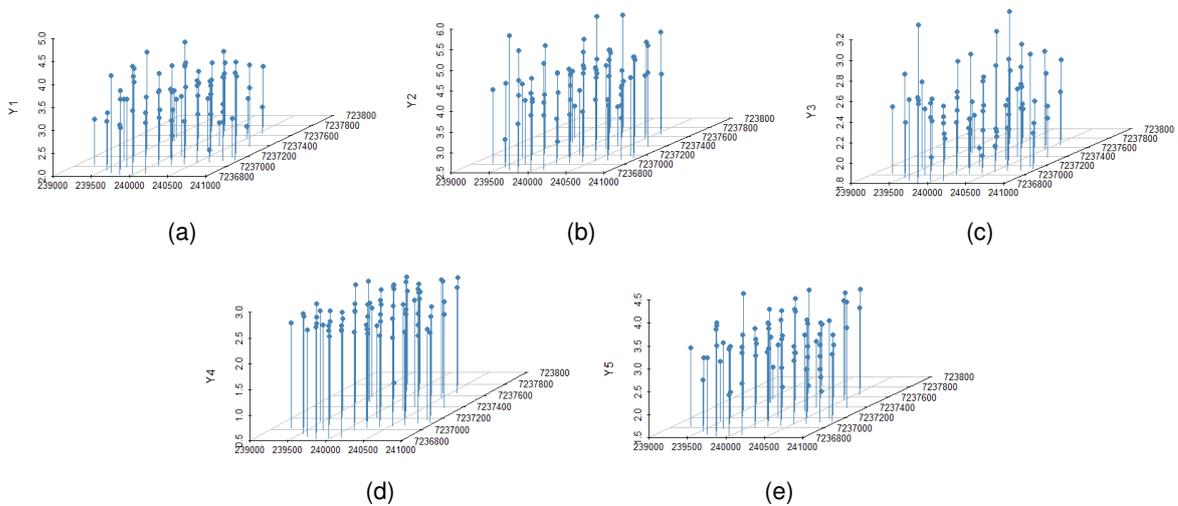


Figura 29 Scatter plot da produtividade da soja referente a cada repetição

Os *Boxplots* das variáveis estão apresentados na Figura 30. Neste estudo, considerou-se a produtividade da soja como variável resposta e as demais como covariáveis. Pode-se verificar que as localizações 16 e 39 aparecem quatro vezes em destaque, referente, à covariável Mg nos primeiros 4 anos agrícolas. Enquanto a localização 39 aparece na covariável Mg três vezes e uma vez em K. Tem-se também a localização 4, destaque três vezes em Mg. Nas demais variáveis, algumas localizações se repetem no máximo duas vezes.

A análise espacial dos dados pela técnica geoestatística, para identificar a estrutura de dependência espacial, realizou-se por meio do algoritmo iterativo parcial linear (Seção 4.25), em que foram estimados os parâmetros $\theta = (\beta, \phi, \eta)^T$. O valor inicial de cada estimaco foi definido da seguinte forma: para o vetor de parâmetro β a partir do modelo normal, para o vetor

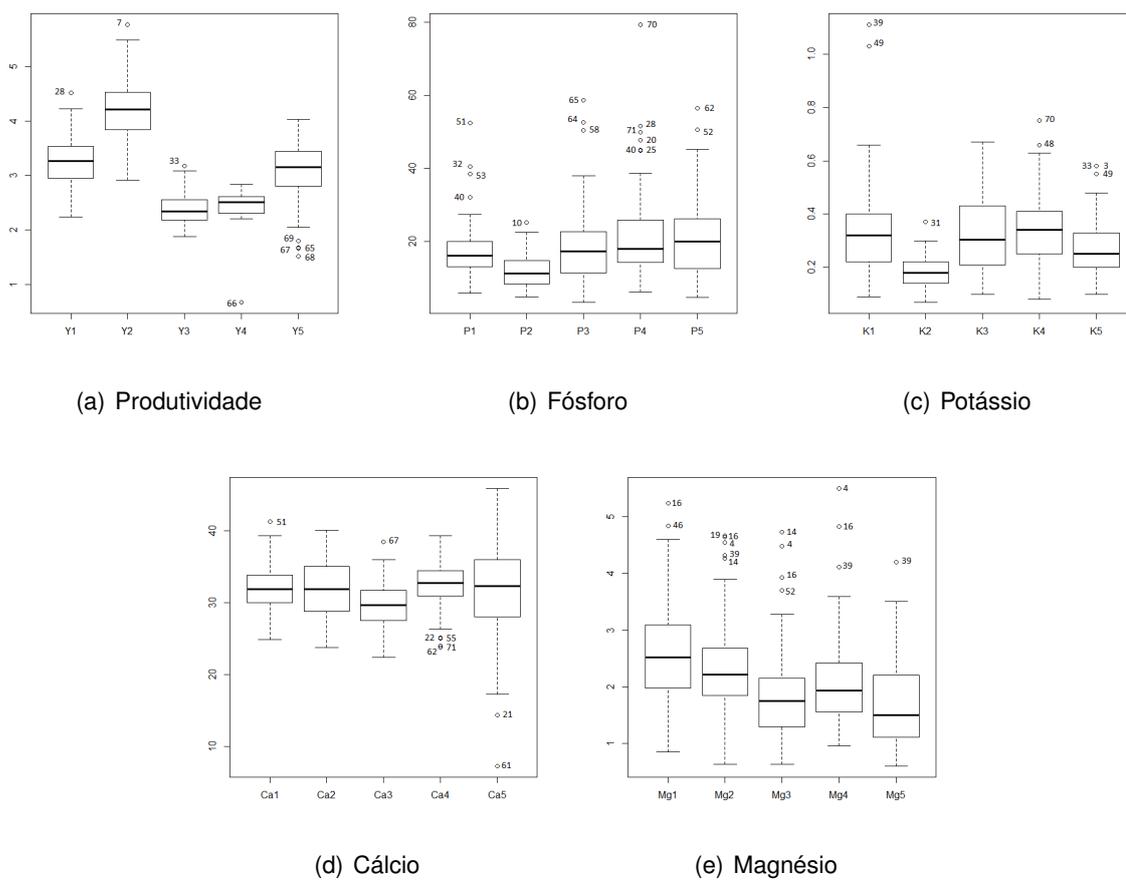


Figura 30 *Boxplot* da variável explicativa produtividade da soja e das covariáveis Fósforo, Potássio, Cálcio e Magnésio nos anos agrícolas 2012/2013, 2013/2014, 2014/2015, 2015/2016 e 2016/2017

de parâmetro ϕ , obteve-se por sorteio dentro de um intervalo para cada ϕ_j , em que $j = 1, 2$ e 3 em relação ao semivariograma dos dados em estudo. O parâmetro de forma η foi sorteado dentro do intervalo $(0, 1/2)$ e o parâmetro do modelo da família Matérn κ foi sorteado entre os intervalos $(1/2, 3)$. Vale ressaltar que o parâmetro κ é fixo, ou seja, o valor inicial sorteado será sempre igual ao valor de κ fixo.

Foram realizadas 1500 estimações com sucesso, em que o algoritmo somente armazenava as respostas em que obtiveram-se os padrões pré-definidos, sendo $\phi_j \geq 0$ em que $j = 1, 2, 3$ e $0 < \eta < 1/2$. Caso contrário, o algoritmo reiniciava a estimativa descartando a estimativa anterior. Isto é de extrema necessidade, pois trata-se de um método iterativo o qual pode trazer soluções corretas, porém não adequadas ao estudo. Dentre todas estas estimativas de sucesso, determinou-se o melhor modelo pelo critério Traço e Validação Cruzada em função do parâmetro η , para mais detalhes, veja 4.26. Vale ressaltar que o valor de κ é fixo no chute inicial, determinado por um sorteio aleatório dentro de um intervalo previamente definido.

Na Tabela 13, apresentam-se as estimativas dos parâmetros do modelo escolhido e os respectivos erros padrões entre parênteses. Nota-se que Ca tem relação inversamente proporcional à média da produtividade, ou seja, após determinado valor, à medida que aumenta o teor de Ca, a média da produtividade diminui. O oposto ocorre em P, K e Mg. Dentre esses, o valor do teor potássio (K) destaca-se como o mais elevado, indicando que esta variável é a que mais contribui para o aumento da média da produtividade da soja. Este resultado contribui com os estudos Dos Passos et al. (2015), os quais relatam que o potássio é o segundo macronutriente mais extraído pela soja.

Tabela 13 Estimativa dos parâmetros do modelo escolhido para a produtividade da soja considerando a função de covariância da família Matérn com parâmetro de forma $\kappa = 2,9315$ com os respectivos erros padrões entre parênteses

$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	η
2,6941	0,0007	0,1080	-0,0013	0,0195	0,0599	0,1400	4,7904	0,4689
(0,1089)	(0,0008)	(0,0608)	(0,0018)	(0,0092)	(0,0597)	(0,2424)	(2,5845)	(0,0310)

$\hat{\beta}_1$: Intercepto; $\hat{\beta}_2$: Estimador do parâmetro associado à variável P; $\hat{\beta}_3$: Estimador do parâmetro associado à variável K; $\hat{\beta}_4$: Estimador do parâmetro associado à variável Ca; $\hat{\beta}_5$: Estimador do parâmetro associado à variável Mg; $\hat{\phi}_1$: Estimador do efeito pepita; $\hat{\phi}_2$: Estimador da contribuição; $\hat{\phi}_3$: Estimador do parâmetro que define o alcance; η : Parâmetro de forma da distribuição t-Student

O teste de hipótese sobre o vetor de parâmetros β foi aplicado individualmente. Na Tabela 14 apresentam os valores da estatística de teste RV e seus respectivos p-valores. Verifica-se que em todos os testes é rejeitada ao nível de 5% de significância a hipótese nula H_0 . Portanto as covariáveis P, K, Mg e Ca devem ser mantidas na formulação do modelo.

Para a representação gráfica dos diagnósticos de influência: local perturbação na variável resposta, global baseada na verossimilhança e na Q-function. Considerou-se Y_1 com seus pontos enumerados de 1 até 74, Y_2 de 75 até 148, Y_3 de 149 até 222, Y_4 de 223 até 296 e Y_5 de 297 até 370.

Observa-se nas Figuras 31 (a) e (b), referentes aos gráficos de influência local

Tabela 14 Teste da razão de verossimilhanças (RV) sob o vetor de parâmetros β do modelo espacial linear t-Student reparametrizado com repetição

Hipótese	RV	p-valor
$\beta_1 = 0$	5,472	0,019*
$\beta_2 = 0$	4,110	0,043*
$\beta_3 = 0$	21,704	0,000*
$\beta_4 = 0$	14,709	0,000*

* Significativo ao nível de 5% de significância

com perturbação na variável resposta, que os pontos considerados como influentes estão relacionados a Y_3 e Y_4 referentes aos anos agrícolas 2014/2015 e 2015/2016, respectivamente.

Nas Figuras 32 (a) e (b), apresentam-se os gráficos de influência local aplicando a ponderação de casos. Nota-se também que Y_3 , referente ao ano agrícola 2014/2015, encontra-se em destaque quando comparado aos demais, logo, o ano agrícola 2014/2015 exerce forte influência nas estimativas dos parâmetros.

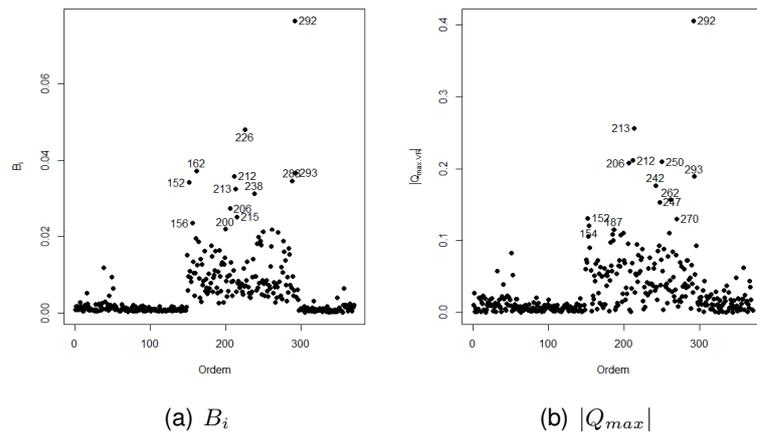


Figura 31 Gráficos de diagnósticos de influência local perturbação na variável resposta

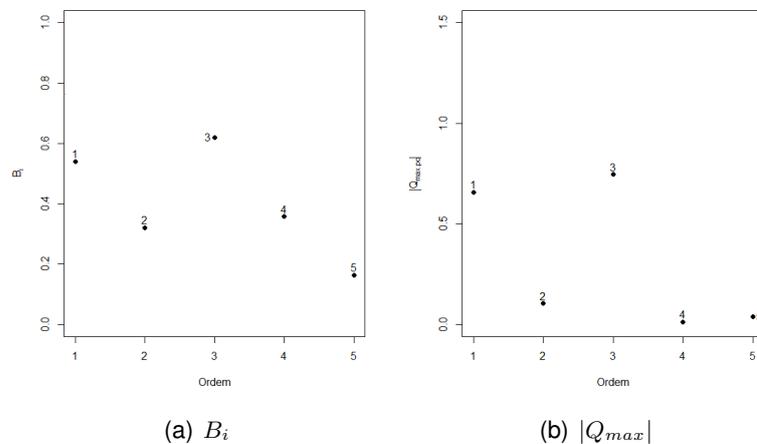


Figura 32 Gráficos de diagnósticos de influência local e ponderação de casos

Na Figura 33, apresentam-se os gráficos de diagnósticos de influência global $D_{i\theta}$,

$D_{i\beta}$ e $D_{i\phi\eta}$, em que os dois últimos somados obtém-se $D_{i\theta}$. Ao serem avaliados os pontos discrepantes, tem-se em destaque: #98 referente a Y_2 no ano agrícola 2013/2014, com sua localização no ponto 24, que corresponde ao valor de produtividade $2,913 t ha^{-1}$, como pode-se notar na Tabela 12, que refere-se ao valor mínimo. O ponto #81 refere-se ao mesmo ano agrícola, sua localização corresponde ao ponto 7, com valor de produtividade igual a $5,766 t ha^{-1}$ equivalente ao valor máximo em Y_2 apresentado na Tabela 12.

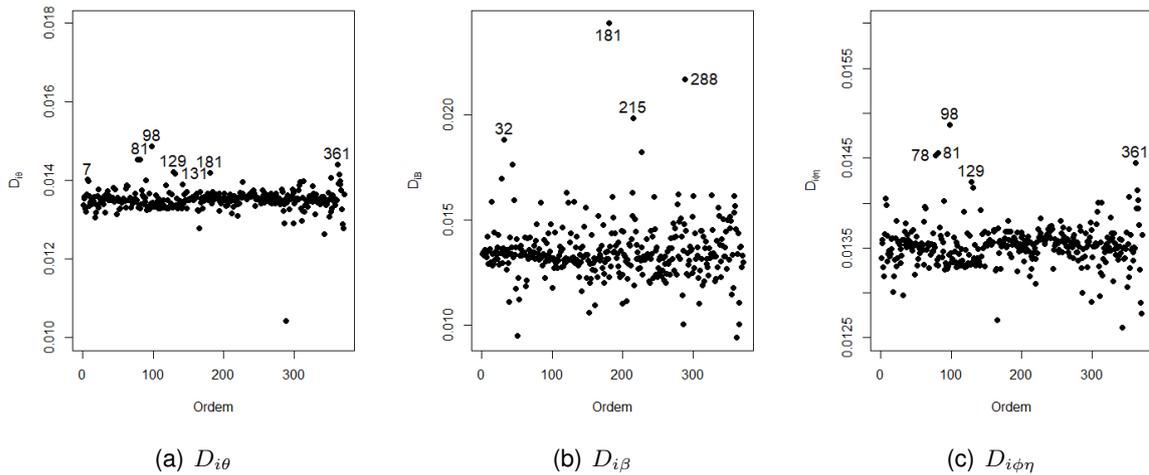


Figura 33 Gráficos de diagnósticos de influência global baseada na verossimilhança

Na Figura 34, apresentam-se os gráficos de diagnósticos de influência global, baseados na Q-function $QD_{i\theta}$, $QD_{i\beta}$, $QD_{i\phi}$ e $QD_{i\eta}$, em que os três últimos somados obtém-se $QD_{i\theta}$. Observam-se os pontos destaques no gráfico de $QD_{i\theta}$ sendo #288 correspondente a Y_4 , referente ao ano agrícola 2015/2016, com sua localização no ponto 66 tendo valor de produtividade igual a $2,26 t ha^{-1}$. Já o ponto #358 apresenta-se em Y_5 no ano agrícola 2016/2017, localizado no ponto 62 com valor referente a $3,46 t ha^{-1}$.

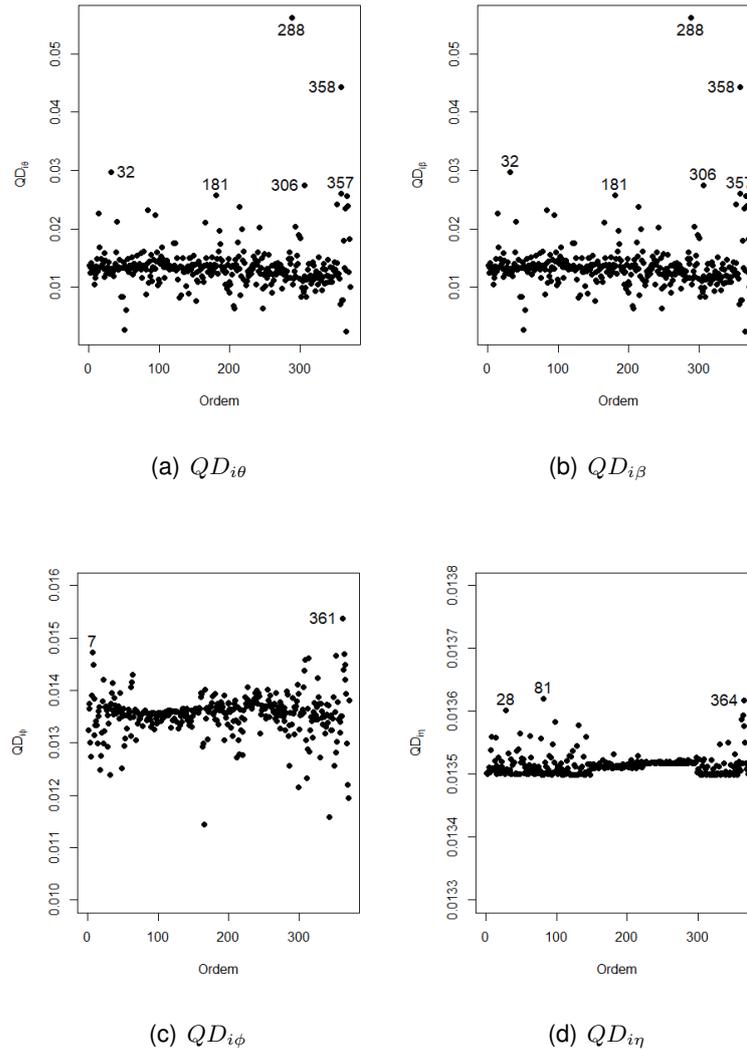


Figura 34 Gráficos de diagnósticos de influência global baseada na Q-function

4.30 Conclusões

Neste artigo, desenvolveu-se o modelo espacial linear t-Student reparametrizado com repetições múltiplas, a função escore e a matriz informação esperada. Apresentou-se o Algoritmo Iterativo parcial linear, que também foi responsável por estimar o parâmetro de forma η . Propôs-se a utilização dos critérios VC e T_r para a escolha do melhor modelo. Avaliou-se a significância dos parâmetros β_i pelo teste de razão de verossimilhança em que as covariáveis foram mantidas no modelo de regressão. A avaliação da estabilidade do modelo ajustado em um conjunto de dados deve fazer parte das análises estatísticas, pois auxiliam nas escolhas das covariáveis adequadas ao modelo escolhido. Foram apresentadas as análises de diagnósticos de influência global e influência local, e para isso, aplicou-se o esquema de perturbação de casos e perturbação na variável resposta. Torna-se um artigo de suma importância, pois todos esses procedimentos citados acima foram desenvolvidos com repetição. Realizou-se uma aplicação de toda a metodologia a um conjunto de dados agrícolas com covariáveis.

4.31 Apêndice

4.32 Função escore

A partir do logaritmo da função de verossimilhança definido na Equação (4.27), aplica-se a diferenciação em função de $\theta = (\beta^\top, \phi^\top, \eta)^\top$, é possível obter a função escore $U(\theta) = \sum_{i=1}^r U_i(\theta)$ em que cada elemento pode ser escrito como:

$$\begin{aligned}
 U_i(\beta) &= \frac{\partial \mathcal{L}_i(\theta)}{\partial \beta}, \\
 &= \frac{\partial}{\partial \beta} \left[\log K_n(\eta) - \frac{1}{2} \log |\Sigma| - \frac{1}{2\eta} (1 + n\eta) \log (1 + c(\eta) \delta_i^2) \right], \\
 &= \frac{\partial}{\partial \beta} \left[-\frac{1}{2\eta} (1 + n\eta) \log (1 + c(\eta) \delta_i^2) \right], \\
 &= -\frac{1}{2\eta} (1 + n\eta) \frac{1}{(1 + c(\eta) \delta_i^2)} \frac{\partial}{\partial \beta} [(1 + c(\eta) \delta_i^2)], \\
 &= -\frac{1}{2\eta} (1 + n\eta) \frac{1}{(1 + c(\eta) \delta_i^2)} \left[-2\mathbf{X}_i^\top \Sigma^{-1} (\mathbf{Y}_i - \mathbf{X}_i \beta) \right] c(\eta), \\
 &= \left(\frac{1 + n\eta}{\eta} \right) \frac{1}{(1 + c(\eta) \delta_i^2)} c(\eta) \left[\mathbf{X}_i^\top \Sigma^{-1} (\mathbf{Y}_i - \mathbf{X}_i \beta) \right], \\
 U_i(\beta) &= w_i \mathbf{X}_i^\top \Sigma^{-1} \varepsilon_i.
 \end{aligned}$$

sendo $w_i = ((1 + n\eta)/\eta) q^{-1} c(\eta)$, $q = (1 + c(\eta) \delta_i^2)$, $\varepsilon_i = (\mathbf{Y}_i - \mathbf{X}_i \beta)$ e $i = 1, \dots, r$.

Pode-se escrever $U_i(\phi)$ em dois formatos, são eles:

Formato 1:

$$\begin{aligned}
 U_i(\phi) &= \frac{\partial \mathcal{L}_i(\theta)}{\partial \phi}, \\
 &= \frac{\partial}{\partial \phi} \left[\log K_n(\eta) - \frac{1}{2} \log |\Sigma| - \frac{1}{2\eta} (1 + n\eta) \log (1 + c(\eta) \delta_i^2) \right], \\
 &= -\frac{1}{2} \frac{\partial}{\partial \phi} [\log |\Sigma|] - \frac{1}{2\eta} (1 + n\eta) \frac{\partial}{\partial \phi} [\log (1 + c(\eta) \delta_i^2)], \\
 &= -\frac{1}{2} \left[\frac{\partial \text{vec}^\top(\Sigma)}{\partial \phi} \text{vec}(\Sigma^{-1}) \right] - \frac{1}{2\eta} (1 + n\eta) \frac{1}{(1 + c(\eta) \delta_i^2)} \frac{\partial}{\partial \phi} [(1 + c(\eta) \delta_i^2)], \\
 &= -\frac{1}{2} \left[\frac{\partial \text{vec}^\top(\Sigma)}{\partial \phi} \text{vec}(\Sigma^{-1}) \right] + \frac{(1 + n\eta) c(\eta)}{2\eta (1 + c(\eta) \delta_i^2)} \frac{\partial \text{vec}^\top(\Sigma)}{\partial \phi} \left[\text{vec}(\Sigma^{-1} \varepsilon_i \varepsilon_i^\top \Sigma^{-1}) \right], \\
 &= -\frac{1}{2} D_\phi^\top \text{vec}(\Sigma^{-1}) + \frac{1}{2} w_i D_\phi^\top \left[\text{vec}(\Sigma^{-1} \varepsilon_i \varepsilon_i^\top \Sigma^{-1}) \right], \\
 &= \frac{1}{2} D_\phi^\top \text{vec} \left[w_i \Sigma^{-1} \varepsilon_i \varepsilon_i^\top \Sigma^{-1} - \Sigma^{-1} \right],
 \end{aligned}$$

em que $D_\phi^\top = \frac{\partial \text{vec}^\top(\Sigma)}{\partial \phi}$.

Formato 2:

$U_i(\phi) = \frac{\partial \mathcal{L}_i(\theta)}{\partial \phi}$, que tem como elementos para $j = 1, 2, 3$.

$$\begin{aligned}
U_i(\phi_j) &= \frac{\partial}{\partial \phi} \left[\log K_n(\eta) - \frac{1}{2} \log |\Sigma| - \frac{1}{2\eta} (1 + n\eta) \log (1 + c(\eta) \delta_i^2) \right], \\
&= -\frac{1}{2} \frac{\partial}{\partial \phi_j} [\log |\Sigma|] - \frac{1}{2\eta} (1 + n\eta) \frac{\partial}{\partial \phi_j} [\log (1 + c(\eta) \delta_i^2)], \\
&= -\frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \right) - \frac{1}{2\eta} (1 + n\eta) \frac{c(\eta)}{(1 + c(\eta) \delta_i^2)} \left(\frac{\partial}{\partial \phi_j} \delta_i^2 \right), \\
&= -\frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \right) - \frac{1}{2} \left(\frac{1 + n\eta}{\eta} \right) c(\eta) q_i^{-1} \left(-\varepsilon_i^\top \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \varepsilon_i \right), \\
U_i(\phi_j) &= -\frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \right) + \frac{1}{2} \left(w_i \varepsilon_i^\top \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \varepsilon_i \right) \text{ para } j = 1, 2, 3.
\end{aligned}$$

$$\begin{aligned}
U_i(\eta) &= \frac{\partial \mathcal{L}_i(\theta)}{\partial \eta}, \\
&= \frac{\partial}{\partial \eta} \left[\log K_n(\eta) - \frac{1}{2} \log |\Sigma| - \frac{1}{2\eta} (1 + n\eta) \log (1 + c(\eta) \delta_i^2) \right], \\
&= \underbrace{\frac{\partial}{\partial \eta} [\log K_n(\eta)]}_{\text{Parte A}} + \underbrace{\frac{\partial}{\partial \eta} \left[-\frac{1}{2\eta} (1 + n\eta) \log (1 + c(\eta) \delta_i^2) \right]}_{\text{Parte B}},
\end{aligned}$$

dividi-se a equação 4.57 em duas partes. Inicialmente, apresentamos a resolução da parte A, sendo:

$$\begin{aligned}
\frac{\partial}{\partial \eta} [\log K_n(\eta)] &= \frac{\partial}{\partial \eta} \left[\frac{n}{2} \log \left(\frac{c(\eta)}{\pi} \right) + \log \Gamma \left(\frac{1 + n\eta}{2\eta} \right) - \log \Gamma \left(\frac{1}{2\eta} \right) \right], \\
&= \frac{n}{2} \frac{1}{\left(\frac{c(\eta)}{\pi} \right)} \frac{\partial}{\partial \eta} \left(\frac{c(\eta)}{\pi} \right) + \psi \left(\frac{1 + n\eta}{2\eta} \right) \frac{\partial}{\partial \eta} \left(\frac{1 + n\eta}{2\eta} \right) - \psi \left(\frac{1}{2\eta} \right) \frac{\partial}{\partial \eta} \left(\frac{1}{2\eta} \right), \\
&= \frac{n}{2} \frac{1}{c(\eta)} \frac{1}{(1 + 2\eta)^2} - \frac{1}{2\eta^2} \psi \left(\frac{1 + n\eta}{2\eta} \right) + \frac{1}{2\eta^2} \psi \left(\frac{1}{2\eta} \right).
\end{aligned}$$

Para a resolução da parte B, tem-se:

$$\begin{aligned}
&\frac{\partial}{\partial \eta} \left[-\frac{1}{2\eta} (1 + n\eta) \log (1 + c(\eta) \delta_i^2) \right] = \\
&= - \left[\frac{\partial}{\partial \eta} \left(\frac{1 + n\eta}{2\eta} \right) \log (1 + c(\eta) \delta_i^2) + \left(\frac{1 + n\eta}{2\eta} \right) \frac{\partial}{\partial \eta} \log (1 + c(\eta) \delta_i^2) \right], \\
&= - \left[-\frac{1}{2\eta^2} \log (1 + c(\eta) \delta_i^2) + \left(\frac{1 + n\eta}{2\eta} \right) \frac{1}{(1 + c(\eta) \delta_i^2)} \frac{\delta_i^2}{(1 - 2\eta)^2} \right], \\
&= \frac{1}{2\eta^2} \log (1 + c(\eta) \delta_i^2) - \left(\frac{1 + n\eta}{2\eta} \right) \frac{1}{(1 + c(\eta) \delta_i^2)} \frac{\delta_i^2}{(1 - 2\eta)^2}.
\end{aligned}$$

A partir da resolução da parte A e B, obtém-se $U_i(\eta)$

$$U_i(\eta) = \underbrace{\frac{n}{2} \frac{1}{c(\eta)} \frac{1}{(1+2\eta)^2}}_{\text{Parte C}} - \frac{1}{2\eta^2} \psi\left(\frac{1+n\eta}{2\eta}\right) + \frac{1}{2\eta^2} \psi\left(\frac{1}{2\eta}\right) + \frac{1}{2\eta^2} \log(1+c(\eta)\delta_i^2) - \underbrace{\left(\frac{1+n\eta}{2\eta}\right) \frac{1}{(1+c(\eta)\delta_i^2)} \frac{\delta_i^2}{(1-2\eta)^2}}_{\text{Parte D}}.$$

Reescrevendo as partes C e D, tem-se:

$$\begin{aligned} \frac{n}{2\eta(1-2\eta)} - \frac{(1+n\eta)\delta_i^2}{2\eta(1+c(\eta)\delta_i^2)(1-2\eta)^2} &= \frac{1}{2\eta^2} \left[\frac{n\eta}{(1-2\eta)} - \frac{\eta(1+n\eta)\delta_i^2}{(1+c(\eta)\delta_i^2)(1-2\eta)^2} \right], \\ &= \frac{1}{2\eta^2} \left[nc(\eta) - c(\eta) \frac{(1+n\eta)\delta_i^2}{(1+c(\eta)\delta_i^2)(1-2\eta)} \right], \\ &= \frac{1}{2\eta^2} c(\eta) [n - w_i \delta_i^2]. \end{aligned}$$

Logo, tem-se $U_i(\eta)$ como:

$$U_i(\eta) = -\frac{1}{2\eta^2} \psi\left(\frac{1+n\eta}{2\eta}\right) + \frac{1}{2\eta^2} \psi\left(\frac{1}{2\eta}\right) + \frac{1}{2\eta^2} \log(1+c(\eta)\delta_i^2) + \frac{1}{2\eta^2} c(\eta) [n - w_i \delta_i^2],$$

portanto,

$$U_i(\eta) = \frac{1}{2\eta^2} \left[\psi\left(\frac{1}{2\eta}\right) - \psi\left(\frac{1+n\eta}{2\eta}\right) + c(\eta) (n - w_i \delta_i^2) + \log(q_i) \right],$$

em que $\psi(z) = \frac{\partial \log \Gamma(z)}{\partial(z)}$ denominada função digamma e $q_i = (1+c(\eta)\delta_i^2)$.

4.33 Derivada da matriz de covariância

Seja a estrutura da matriz de covariância conforme definida na equação (4.25), em que $\phi = (\phi_1, \phi_2, \phi_3)^\top$ são parâmetros desconhecidos, \mathbf{I}_n é a matriz identidade de ordem $n \times n$ e a matriz \mathbf{R} de ordem $n \times n$, simétrica é descrita pelo modelo da família Matérn, com parâmetro de forma κ fixo. Aplicando-se o operador “vec”, tem-se que $\text{vec}^\top(\Sigma) = \phi_1 \text{vec}^\top(\mathbf{I}) + \phi_2 \text{vec}^\top(\mathbf{R}(\phi_3))$. Aplicando-se a derivada em relação à ϕ tem-se:

$$\frac{\partial \text{vec}^\top(\Sigma)}{\partial \phi} = \frac{\partial \phi_1 \text{vec}^\top(\mathbf{I})}{\partial \phi} + \frac{\partial \phi_2 \text{vec}^\top(\mathbf{R}(\phi_3))}{\partial \phi}. \quad (4.57)$$

Desenvolve-se a primeira parte da equação (4.57) depois da igualdade, e tem-se:

$$\begin{aligned} \frac{\partial \phi_1 \text{vec}^\top(\mathbf{I})}{\partial \phi} &= \left(\frac{\partial \phi_1 \text{vec}^\top(\mathbf{I})}{\partial \phi_1}, \frac{\partial \phi_1 \text{vec}^\top(\mathbf{I})}{\partial \phi_2}, \frac{\partial \phi_1 \text{vec}^\top(\mathbf{I})}{\partial \phi_3} \right)^\top \\ &= \left(\text{vec}^\top(\mathbf{I}), 0, 0 \right)^\top, \end{aligned}$$

na segunda parte da equação (4.57):

$$\begin{aligned} \frac{\partial \phi_2 \text{vec}^\top(\mathbf{R}(\phi_3))}{\partial \phi} &= \left(\frac{\partial \phi_2 \text{vec}^\top(\mathbf{R}(\phi_3))}{\partial \phi_1}, \frac{\partial \phi_2 \text{vec}^\top(\mathbf{R}(\phi_3))}{\partial \phi_2}, \frac{\partial \phi_2 \text{vec}^\top(\mathbf{R}(\phi_3))}{\partial \phi_3} \right)^\top \\ &= \left(0, \text{vec}^\top(\mathbf{R}(\phi_3)), \phi_2 \text{vec}^\top \left(\frac{\partial \mathbf{R}(\phi_3)}{\partial \phi_3} \right) \right)^\top. \end{aligned}$$

Portanto, a solução da equação (4.57) é a matriz de ordem $3 \times n^2$, dada por:

$$\frac{\partial \text{vec}^\top(\boldsymbol{\Sigma})}{\partial \phi} = \left(\text{vec}^\top(\mathbf{I}), \text{vec}^\top(\mathbf{R}(\phi_3)), \phi_2 \text{vec}^\top \left(\frac{\partial \mathbf{R}(\phi_3)}{\partial \phi_3} \right) \right)^\top.$$

Os valores de $\partial \mathbf{R}(\phi_3)/\partial \phi_3$, para a função covariância da família Matérn, são apresentadas em Uribe-Opazo et al., 2012.

4.34 Matriz de informação observada

A partir da função escore, obtém-se a matriz de informação observada $\mathbf{L}(\boldsymbol{\theta}) = \sum_{i=1}^r \mathbf{L}_i(\boldsymbol{\theta})$ de ordem (3×3) . Dada por:

$$\mathbf{L}_i(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{L}_{i\beta\beta}(\boldsymbol{\theta}) & \mathbf{L}_{i\beta\phi}(\boldsymbol{\theta}) & \mathbf{L}_{i\beta\eta}(\boldsymbol{\theta}) \\ \mathbf{L}_{i\phi\beta}(\boldsymbol{\theta}) & \mathbf{L}_{i\phi\phi}(\boldsymbol{\theta}) & \mathbf{L}_{i\phi\eta}(\boldsymbol{\theta}) \\ \mathbf{L}_{i\eta\beta}(\boldsymbol{\theta}) & \mathbf{L}_{i\eta\phi}(\boldsymbol{\theta}) & \mathbf{L}_{i\eta\eta}(\boldsymbol{\theta}) \end{pmatrix},$$

Os elementos da matriz informação observada são definidos por:

$$\begin{aligned} \mathbf{L}_{i\beta\beta}(\boldsymbol{\theta}) &= \frac{\partial^2 \mathcal{L}_i(\boldsymbol{\theta})}{\partial \beta \partial \beta^\top}, \\ &= \frac{\partial}{\partial \beta^\top} \mathbf{U}_i(\beta), \\ &= \frac{\partial}{\partial \beta^\top} (w_i \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i), \\ &= \left[\frac{\partial}{\partial \beta^\top} (w_i) \right] (\mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i) + w_i \left[\frac{\partial}{\partial \beta^\top} (\mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i) \right], \end{aligned}$$

como

$$\frac{\partial}{\partial \beta^\top} (w_i) = 2w_i^2 \left(\frac{\eta}{1+n\eta} \right) \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i,$$

e

$$\frac{\partial}{\partial \beta^\top} (\mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i) = -\mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_i,$$

portanto,

$$\mathbf{L}_{i\beta\beta}(\boldsymbol{\theta}) = w'_\beta (\mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i)^\top - w_i \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_i.$$

$$\text{sendo } w'_\beta = 2w_i^2 \left(\frac{\eta}{1+n\eta} \right) \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i.$$

Para obter $\mathbf{L}_{i\beta\phi}$, tem-se:

$L_{i\beta\phi}(\boldsymbol{\theta}) = \frac{\partial^2 \mathcal{L}_i(\boldsymbol{\theta})}{\partial \beta \partial \phi^\top}$, que tem como elementos

$$\frac{\partial^2 \mathcal{L}_i(\boldsymbol{\theta})}{\partial \beta \partial \phi_j} = \frac{\partial}{\partial \phi_j} U_i(\boldsymbol{\beta}) = \frac{\partial}{\partial \phi_j} (w_i \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i), \text{ para } j = 1, 2 \text{ e } 3.$$

$$L_{i\beta\phi_j}(\boldsymbol{\theta}) = \left[\frac{\partial}{\partial \phi_j} (w_i) \right] (\mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i) + w_i \left[\frac{\partial}{\partial \phi_j} (\mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i) \right]$$

como

$$\frac{\partial}{\partial \phi_j} (w_i) = w_i^2 \left(\frac{\eta}{1+n\eta} \right) (\boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i),$$

e

$$\frac{\partial}{\partial \phi_j} (\mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i) = -\mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i,$$

assim,

$$L_{i\beta\phi_j}(\boldsymbol{\theta}) = w'_{\phi_j} \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i - w_i \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\varepsilon}_i,$$

$$\text{com } w'_{\phi_j} = w_i^2 \left(\frac{\eta}{1+n\eta} \right) (\boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i).$$

Logo, $L_{i\phi\beta}(\boldsymbol{\theta}) = L_{i\beta\phi}(\boldsymbol{\theta})$.

Para se determinar $L_{i\beta\eta}(\boldsymbol{\theta})$, tem-se que:

$$\begin{aligned} L_{i\beta\eta}(\boldsymbol{\theta}) &= \frac{\partial^2 \mathcal{L}_i(\boldsymbol{\theta})}{\partial \beta \partial \eta}, \\ &= \frac{\partial}{\partial \eta} U_i(\boldsymbol{\beta}), \\ &= \frac{\partial}{\partial \eta} (w_i \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i), \\ &= \left[\frac{\partial}{\partial \eta} (w_i) \right] (\mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i), \\ &= (1-2\eta)^{-3} [(n+2)(1-2\eta)q_i^{-1} - (1+n\eta)q_i^{-2}\delta_i^2] \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i. \end{aligned}$$

Para se determinar $L_{i\phi\phi}(\boldsymbol{\theta})$, tem-se:

$L_{i\phi\phi}(\boldsymbol{\theta}) = \frac{\partial^2 \mathcal{L}_i(\boldsymbol{\theta})}{\partial \phi \partial \phi^\top}$ que tem como elementos

$$L_{i\phi_l \phi_j}(\boldsymbol{\theta}) = \frac{\partial^2 \mathcal{L}_i(\boldsymbol{\theta})}{\partial \phi_j \partial \phi_l} = \frac{\partial}{\partial \phi_l} (U_i(\phi_j)), \text{ para } j, l = 1, 2, 3.$$

$$\begin{aligned}
L_{i\phi_l\phi_j}(\boldsymbol{\theta}) &= \frac{\partial}{\partial\phi_l} \left[-\frac{1}{2}tr \left(\boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} \right) + \frac{1}{2} \left(w_i \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right) \right] \\
&= -\frac{1}{2}tr \left[\frac{\partial}{\partial\phi_l} \left(\boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} \right) \right] + \frac{1}{2} \left[\frac{\partial}{\partial\phi_l} \left(w_i \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right) \right], \\
&= -\frac{1}{2}tr \left[\frac{\partial\boldsymbol{\Sigma}^{-1}}{\partial\phi_l} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} + \boldsymbol{\Sigma}^{-1} \frac{\partial^2\boldsymbol{\Sigma}}{\partial\phi_j\partial\phi_l} \right] + \frac{1}{2} \left[\left(\frac{\partial w_i}{\partial\phi_l} \right) \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right. \\
&\quad \left. + w_i \left(\frac{\partial}{\partial\phi_l} \left(\boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right) \right) \right], \\
&= \underbrace{-\frac{1}{2}tr \left[\frac{\partial\boldsymbol{\Sigma}^{-1}}{\partial\phi_l} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} + \boldsymbol{\Sigma}^{-1} \frac{\partial^2\boldsymbol{\Sigma}}{\partial\phi_j\partial\phi_l} \right]}_{\text{Parte E}} + \underbrace{\frac{1}{2} \left[\left(\frac{\partial w_i}{\partial\phi_l} \right) \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right]}_{\text{Parte F}} \\
&\quad + \underbrace{\frac{1}{2} \left[w_i \left(\frac{\partial}{\partial\phi_l} \left(\boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right) \right) \right]}_{\text{Parte G}},
\end{aligned} \tag{4.58}$$

sabe-se que

$$\frac{\partial w_i}{\partial\phi_l} = w_i^2 \left(\frac{\eta}{1+n\eta} \right) \left(\boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_l} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right),$$

e

$$\frac{\partial\boldsymbol{\Sigma}^{-1}}{\partial\phi_l} = -\boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_l} \boldsymbol{\Sigma}^{-1}$$

assim, resolvendo-se a equação (4.58) separadamente, tem-se:

$$\begin{aligned}
\text{Parte E} &= -\frac{1}{2}tr \left[\frac{\partial\boldsymbol{\Sigma}^{-1}}{\partial\phi_l} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} + \boldsymbol{\Sigma}^{-1} \frac{\partial^2\boldsymbol{\Sigma}}{\partial\phi_j\partial\phi_l} \right], \\
&= -\frac{1}{2}tr \left[-\boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_l} \boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} + \boldsymbol{\Sigma}^{-1} \frac{\partial^2\boldsymbol{\Sigma}}{\partial\phi_j\partial\phi_l} \right], \\
&= -\frac{1}{2}tr \left[\boldsymbol{\Sigma}^{-1} \left(\frac{\partial^2\boldsymbol{\Sigma}}{\partial\phi_j\partial\phi_l} - \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_l} \boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} \right) \right],
\end{aligned} \tag{4.59}$$

$$\begin{aligned}
\text{Parte F} &= \frac{1}{2} \left[\left(\frac{\partial w_i}{\partial\phi_l} \right) \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right], \\
&= \frac{1}{2} w_i^2 \left(\frac{\eta}{1+n\eta} \right) \left(\boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_l} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right) \left(\boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial\boldsymbol{\Sigma}}{\partial\phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right),
\end{aligned} \tag{4.60}$$

$$\begin{aligned}
\text{Parte G} &= \frac{1}{2} \left[w_i \left(\frac{\partial}{\partial \phi_l} \left(\varepsilon_i^\top \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \varepsilon_i \right) \right) \right], \\
&= \frac{1}{2} w_i \left\{ \varepsilon_i^\top \Sigma^{-1} \left[\frac{\partial}{\partial \phi_l} \left(\frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \varepsilon_i \right) \right] + \left[\frac{\partial}{\partial \phi_l} \left(\varepsilon_i^\top \Sigma^{-1} \right) \right] \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \varepsilon_i \right\}, \\
&= \frac{1}{2} w_i \left\{ \varepsilon_i^\top \Sigma^{-1} \left[\frac{\partial \Sigma}{\partial \phi_j} \left(\frac{\partial \Sigma^{-1}}{\partial \phi_l} \varepsilon_i \right) + \frac{\partial^2 \Sigma}{\partial \phi_l \partial \phi_j} \left(\Sigma^{-1} \varepsilon_i \right) \right] + \varepsilon_i^\top \frac{\partial \Sigma^{-1}}{\partial \phi_l} \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \varepsilon_i \right\}, \\
&= \frac{1}{2} w_i \left\{ \varepsilon_i^\top \Sigma^{-1} \left[\frac{\partial \Sigma}{\partial \phi_j} \left(-\Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_l} \Sigma^{-1} \right) \varepsilon_i + \frac{\partial^2 \Sigma}{\partial \phi_l \partial \phi_j} \Sigma^{-1} \varepsilon_i \right] \right. \\
&\quad \left. + \varepsilon_i^\top \left(-\Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_l} \Sigma^{-1} \right) \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \varepsilon_i \right\}, \\
&= \frac{1}{2} w_i \left\{ \varepsilon_i^\top \Sigma^{-1} \left[-\frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_l} \Sigma^{-1} \varepsilon_i + \frac{\partial^2 \Sigma}{\partial \phi_l \partial \phi_j} \Sigma^{-1} \varepsilon_i \right] \right. \\
&\quad \left. - \varepsilon_i^\top \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_l} \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \varepsilon_i \right\},
\end{aligned} \tag{4.61}$$

Substituindo as equações (4.59), (4.60), (4.61) em (4.58), tem-se

$$\begin{aligned}
L_{i\phi_j\phi_l}(\boldsymbol{\theta}) &= -\frac{1}{2} \text{tr} \left[\Sigma^{-1} \left(\frac{\partial^2 \Sigma}{\partial \phi_j \partial \phi_l} - \frac{\partial \Sigma}{\partial \phi_l} \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \right) \right] + \frac{1}{2} w'_{\phi_l} \left(\varepsilon_i^\top \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \varepsilon_i \right) \\
&\quad + \frac{1}{2} w_i \varepsilon_i^\top \Sigma^{-1} \left(\frac{\partial^2 \Sigma}{\partial \phi_l \partial \phi_j} - 2 \frac{\partial \Sigma}{\partial \phi_l} \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \right) \Sigma^{-1} \varepsilon_i.
\end{aligned}$$

Em que w'_{ϕ_l} é definido por:

$$w'_{\phi_l} = w_i^2 \left(\frac{\eta}{1 + n\eta} \right) \left(\varepsilon_i^\top \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_l} \Sigma^{-1} \varepsilon_i \right).$$

Para se determinar $L_{i\phi\eta}(\boldsymbol{\theta})$, tem-se:

$$\begin{aligned}
L_{i\phi\eta}(\boldsymbol{\theta}) &= \frac{\partial^2 \mathcal{L}_i(\boldsymbol{\theta})}{\partial \phi \partial \eta}, \\
&= \frac{\partial}{\partial \phi} [U_i(\eta)], \\
&= \frac{\partial}{\partial \phi} \left\{ \frac{1}{2\eta^2} \left[\psi \left(\frac{1}{2\eta} \right) - \psi \left(\frac{1+n\eta}{2\eta} \right) + c(\eta) (n - w_i \delta_i^2) + \log(q_i) \right] \right\}, \\
&= \frac{\partial}{\partial \phi} \left\{ \left[\frac{1}{2\eta^2} c(\eta) (n - w_i \delta_i^2) + \frac{1}{2\eta^2} \log(q_i) \right] \right\}, \\
&= \underbrace{\frac{\partial}{\partial \phi} \left[\frac{1}{2\eta^2} c(\eta) (n - w_i \delta_i^2) \right]}_{\text{Parte H}} + \underbrace{\frac{\partial}{\partial \phi} \left[\frac{1}{2\eta^2} \log(q_i) \right]}_{\text{Parte I}},
\end{aligned} \tag{4.62}$$

resolvendo separadamente a Parte H, tem-se:

$$\begin{aligned}
\text{Parte H} &= \frac{\partial}{\partial \phi} \left[\frac{1}{2\eta^2} c(\eta) (n - w_i \delta_i^2) \right] \\
&= \frac{\partial}{\partial \phi} \left[\frac{1}{2\eta^2} c(\eta) n \right] - \frac{\partial}{\partial \phi} \left[\frac{1}{2\eta^2} c(\eta) w_i \delta_i^2 \right] \\
&= -\frac{\partial}{\partial \phi} \left[\frac{1}{2\eta^2} c(\eta) w_i \delta_i^2 \right] \\
&= -\frac{\partial}{\partial \phi} \left\{ \frac{1}{2\eta^2} \left[c(\eta) \left(\frac{1+n\eta}{\eta} \right) c(\eta) q^{-1} \delta_i^2 \right] \right\} \\
&= -\frac{1}{2\eta^2} c^2(\eta) \left(\frac{1+n\eta}{\eta} \right) \frac{\partial}{\partial \phi} [q^{-1} \delta_i^2] \\
&= -\frac{1}{2\eta^2} c^2(\eta) \left(\frac{1+n\eta}{\eta} \right) \left\{ \frac{\partial}{\partial \phi} [q^{-1}] \delta_i^2 + q^{-1} \frac{\partial}{\partial \phi} [\delta_i^2] \right\} \\
&= -\frac{1}{2\eta^2} c^2(\eta) \left(\frac{1+n\eta}{\eta} \right) \left[(c(\eta) q^{-2} \delta_i^2 - q^{-1}) \left(D_\phi^\top \text{vec}(\Sigma^{-1} \epsilon_i \epsilon_i^\top \Sigma^{-1}) \right) \right] \\
&= -\frac{1}{2\eta^2} \left[(w_i c^2(\eta) q^{-1} \delta_i^2 - w_i c(\eta)) \left(D_\phi^\top \text{vec}(\Sigma^{-1} \epsilon_i \epsilon_i^\top \Sigma^{-1}) \right) \right],
\end{aligned} \tag{4.63}$$

resolvendo separadamente Parte I, tem-se:

$$\begin{aligned}
\text{Parte I} &= \frac{\partial}{\partial \phi} \left[\frac{1}{2\eta^2} \log(q_i) \right], \\
&= \frac{1}{2\eta^2} \left\{ \frac{\partial}{\partial \phi} [\log(q_i)] \right\}, \\
&= \frac{1}{2\eta^2} \left(\frac{1}{q_i} \right) \left[\frac{\partial}{\partial \phi} (q_i) \right], \\
&= \frac{1}{2\eta^2} \left(\frac{1}{(1+c(\eta)\delta_i^2)} \right) \left[\frac{\partial}{\partial \phi} (1+c(\eta)\delta_i^2) \right], \\
&= \frac{1}{2\eta^2} \left(\frac{1}{(1+c(\eta)\delta_i^2)} \right) c(\eta) \left[\frac{\partial}{\partial \phi} (\delta_i^2) \right], \\
&= \frac{1}{2\eta^2} \left[(q_i^{-1} c(\eta)) \left(-D_\phi^\top \text{vec}(\Sigma^{-1} \epsilon_i \epsilon_i^\top \Sigma^{-1}) \right) \right], \\
&= -\frac{1}{2\eta^2} \left[q_i^{-1} c(\eta) D_\phi^\top \text{vec}(\Sigma^{-1} \epsilon_i \epsilon_i^\top \Sigma^{-1}) \right],
\end{aligned} \tag{4.64}$$

quando os resultados obtidos nas equações (4.63) e (4.64) são substituídas na equação (4.62), tem-se:

$$L_{i\phi\eta}(\boldsymbol{\theta}) = -\frac{1}{2\eta^2} \left[(w_i c^2(\eta) q^{-1} \delta_i^2 - w_i c(\eta) + q_i^{-1} c(\eta)) \left(D_\phi^\top \text{vec}(\Sigma^{-1} \epsilon_i \epsilon_i^\top \Sigma^{-1}) \right) \right].$$

Para se determinar $L_{i\eta\eta}(\theta)$, tem-se:

$$\begin{aligned}
L_{i\eta\eta}(\theta) &= \frac{\partial^2 \mathcal{L}_i(\theta)}{\partial \eta \partial \eta}, \\
&= \frac{\partial}{\partial \eta} [U_i(\eta)], \\
&= \frac{\partial}{\partial \eta} \left[\frac{1}{2\eta^2} \left[\psi\left(\frac{1}{2\eta}\right) - \psi\left(\frac{1+n\eta}{2\eta}\right) + c(\eta)(n - w_i \delta_i^2) + \log(q_i) \right] \right], \\
&= \underbrace{\frac{\partial}{\partial \eta} \left[\frac{1}{2\eta^2} \psi\left(\frac{1}{2\eta}\right) \right]}_{\text{Parte J}} + \underbrace{\frac{\partial}{\partial \eta} \left[\frac{1}{2\eta^2} c(\eta)(n - w_i \delta_i^2) \right]}_{\text{Parte L}} \\
&\quad - \underbrace{\frac{\partial}{\partial \eta} \left[\frac{1}{2\eta^2} \psi\left(\frac{1+n\eta}{2\eta}\right) \right]}_{\text{Parte M}} + \underbrace{\frac{\partial}{\partial \eta} \left[\frac{1}{2\eta^2} \log(q_i) \right]}_{\text{Parte N}}, \tag{4.65}
\end{aligned}$$

Ao se resolver separadamente cada parte definida na equação (4.65), tem-se:

Parte J:

$$\begin{aligned}
\text{Parte J} &= \frac{\partial}{\partial \eta} \left[\frac{1}{2\eta^2} \psi\left(\frac{1}{2\eta}\right) \right], \\
&= \left[\frac{\partial}{\partial \eta} \left(\frac{1}{2\eta^2} \right) \right] \psi\left(\frac{1}{2\eta}\right) + \frac{1}{2\eta^2} \left[\frac{\partial}{\partial \eta} \left(\psi\left(\frac{1}{2\eta}\right) \right) \right], \\
&= -\frac{1}{\eta^3} \psi\left(\frac{1}{2\eta}\right) - \frac{1}{4\eta^4} \psi'\left(\frac{1}{2\eta}\right). \tag{4.66}
\end{aligned}$$

Parte L:

$$\begin{aligned}
\text{Parte L} &= \frac{\partial}{\partial \eta} \left[\frac{1}{2\eta^2} c(\eta)(n - w_i \delta_i^2) \right] \\
&= \frac{\partial}{\partial \eta} \left[\frac{1}{2\eta^2} c(\eta)n \right] - \frac{\partial}{\partial \eta} \left[\frac{1}{2\eta^2} c(\eta)w_i \delta_i^2 \right] \\
&= n \left\{ \left[\frac{\partial}{\partial \eta} \left(\frac{1}{2\eta^2} \right) \right] c(\eta) + \frac{1}{2\eta^2} \left[\frac{\partial}{\partial \eta} c(\eta) \right] \right\} \\
&\quad - \delta_i^2 \left\{ \left[\frac{\partial}{\partial \eta} \left(\frac{1}{2\eta^2} \right) \right] c(\eta) w_i + \frac{1}{2\eta^2} c(\eta) \left[\frac{\partial}{\partial \eta} (w_i) \right] \right\} \\
&= n \left\{ \left[\frac{\partial}{\partial \eta} \left(\frac{1}{2\eta^2} \right) \right] c(\eta) + \frac{1}{2\eta^2} \left[\frac{\partial}{\partial \eta} c(\eta) \right] \right\} \\
&\quad - \delta_i^2 \left\{ \left[\left[\frac{\partial}{\partial \eta} \left(\frac{1}{2\eta^2} \right) \right] c(\eta) + \left(\frac{1}{2\eta^2} \right) \left[\frac{\partial}{\partial \eta} c(\eta) \right] \right] w_i + \frac{1}{2\eta^2} c(\eta) \left[\frac{\partial}{\partial \eta} (w_i) \right] \right\}. \tag{4.67}
\end{aligned}$$

Inicialmente, para se resolver a Parte L, precisamos derivar w_i :

$$\begin{aligned}
\frac{\partial w_i}{\partial \eta} &= \frac{\partial}{\partial \eta} \left[\left(\frac{1+n\eta}{\eta} \right) c(\eta) q_i^{-1} \right], \\
&= \frac{\partial}{\partial \eta} \left[\left(\frac{1+n\eta}{1-2\eta} \right) \frac{1}{1+c(\eta)\delta_i^2} \right], \\
&= \left\{ \left[\frac{\partial}{\partial \eta} \left(\frac{1+n\eta}{1-2\eta} \right) \right] \frac{1}{1+c(\eta)\delta_i^2} + \left(\frac{1+n\eta}{1-2\eta} \right) \left[\frac{\partial}{\partial \eta} \left(\frac{1}{1+c(\eta)\delta_i^2} \right) \right] \right\} \\
&= (1-2\eta)^{-3} [(n+2)(1-2\eta)q_i^{-1} - (1+n\eta)q_i^{-2}\delta_i^2],
\end{aligned}$$

resolvendo as derivadas da equação (4.67), obtém-se:

$$\begin{aligned}
\text{Parte L} &= -\frac{1}{\eta^3}nc(\eta) + \frac{n}{2\eta^2(1-2\eta)^2} + \frac{1}{\eta^3}\delta_i^2c(\eta)w_i - \frac{1}{2\eta^2(1-2\eta)^2}\delta_i^2w_i \\
&- \frac{1}{2\eta^2}c(\eta)(1-2\eta)^{-3}[(n+2)(1-2\eta)q_i^{-1} - (1+n\eta)q_i^{-2}\delta_i^2]\delta_i^2, \\
&= \frac{1}{\eta^3}c(\eta)[\delta_i^2w_i - n] + \frac{1}{2\eta^2(1-2\eta)^2}[n - \delta_i^2w_i] \\
&- \frac{1}{2\eta^2}c(\eta)(1-2\eta)^{-3}[(n+2)(1-2\eta)q_i^{-1} - (1+n\eta)q_i^{-2}\delta_i^2]\delta_i^2. \tag{4.68}
\end{aligned}$$

Parte M:

$$\begin{aligned}
\text{Parte M} &= -\frac{\partial}{\partial\eta}\left[\frac{1}{2\eta^2}\psi\left(\frac{1+n\eta}{2\eta}\right)\right] \\
&= -\left\{\left[\frac{\partial}{\partial\eta}\left(\frac{1}{2\eta^2}\right)\right]\psi\left(\frac{1+n\eta}{2\eta}\right) + \frac{1}{2\eta^2}\left[\frac{\partial}{\partial\eta}\left(\psi\left(\frac{1+n\eta}{2\eta}\right)\right)\right]\right\} \\
&= \frac{1}{\eta^3}\psi\left(\frac{1+n\eta}{2\eta}\right) + \frac{1}{4\eta^4}\psi'\left(\frac{1+n\eta}{2\eta}\right) \tag{4.69}
\end{aligned}$$

Parte N:

$$\begin{aligned}
\text{Parte N} &= \frac{\partial}{\partial\eta}\left[\frac{1}{2\eta^2}\log(q_i)\right], \\
&= \frac{\partial}{\partial\eta}\left[\frac{1}{2\eta^2}\log(1+c(\eta)\delta_i^2)\right], \\
&= \left\{\left[\frac{\partial}{\partial\eta}\left(\frac{1}{2\eta^2}\right)\right]\log(1+c(\eta)\delta_i^2) + \frac{1}{2\eta^2}\left[\frac{\partial}{\partial\eta}(\log(1+c(\eta)\delta_i^2))\right]\right\}, \tag{4.70}
\end{aligned}$$

com

$$\begin{aligned}
\frac{\partial}{\partial\eta}[\log(1+c(\eta)\delta_i^2)] &= \left(\frac{1}{(1+c(\eta)\delta_i^2)}\right)\delta_i^2\left[\frac{\partial}{\partial\eta}c(\eta)\right], \\
&= \frac{\delta_i^2}{(1-2\eta)^2(1+c(\eta)\delta_i^2)},
\end{aligned}$$

Logo, a equação (4.70) é definida por:

$$\text{Parte N} = -\frac{1}{\eta^3}\log(q_i) + \frac{\delta_i^2q_i^{-1}}{2\eta^2(1-2\eta)^2}. \tag{4.71}$$

Quando as equações obtidas em (4.66), (4.68), (4.69) e (4.71) são substituídas em (4.65), tem-se que $\mathbf{L}_{i\eta\eta}(\boldsymbol{\theta})$ é dado por:

$$\begin{aligned}
\mathbf{L}_{i\eta\eta}(\boldsymbol{\theta}) &= -\frac{1}{\eta^3}\psi\left(\frac{1}{2\eta}\right) - \frac{1}{4\eta^4}\psi'\left(\frac{1}{2\eta}\right) + \frac{1}{\eta^3}\psi\left(\frac{1+n\eta}{2\eta}\right) + \frac{1}{4\eta^4}\psi'\left(\frac{1+n\eta}{2\eta}\right) \\
&+ \frac{1}{\eta^3}c(\eta)(\delta_i^2w_i - n) + \frac{1}{2\eta^2(1-2\eta)^2}(n - \delta_i^2w_i) - \frac{1}{\eta^3}\log(q_i) + \frac{\delta_i^2q_i^{-1}}{2\eta^2(1-2\eta)^2} \\
&- \frac{1}{2\eta^2}\delta_i^2c(\eta)(1-2\eta)^{-3}[(n+2)(1-2\eta)q_i^{-1} - (1+n\eta)q_i^{-2}\delta_i^2].
\end{aligned}$$

4.35 Matriz de informação esperada de Fisher

Antes de se obter a matriz informação esperada de Fisher, é preciso definir algumas propriedades, apresentadas por Osorio e Galea (2015). Para a distribuição t-Student reparametrizada com repetição, o lema é dado por:

Lema 2. *Seja $\varepsilon_i \sim \mathcal{T}_n(\mathbf{0}, \Sigma, \eta)$ em que $\varepsilon_i = (\mathbf{Y}_i - \mathbf{X}_i\beta)$, $q_i = (1 + c(\eta)\delta_i^2)$ em que $\delta_i^2 = (\mathbf{Y}_i - \mathbf{X}_i\beta)^\top \Sigma^{-1} (\mathbf{Y}_i - \mathbf{X}_i\beta) = \varepsilon_i^\top \Sigma^{-1} \varepsilon_i$. Então segue que:*

- (i) $E(q_i^{-1}) = (1 + n\eta)^{-1}$,
- (ii) $E(q_i^{-1}\varepsilon_i) = \mathbf{0}$,
- (iii) $E(q_i^{-1}\varepsilon_i\varepsilon_i^\top) = [(1 - 2\eta)/(1 + n\eta)] \Sigma$,
- (iv) $E(q_i^{-1}\delta_i^2) = n(1 - 2\eta)/(1 + n\eta)$,
- (v) $E(q_i^{-2}) = (1 + 2\eta)/(1 + n\eta)(1 + (n + 2)\eta)$,
- (vi) $E(q_i^{-2}\varepsilon_i) = \mathbf{0}$,
- (vii) $E(q_i^{-2}\varepsilon_i\varepsilon_i^\top) = [(1 - 2\eta)/(1 + n\eta)(1 + (n + 2)\eta)] \Sigma$,
- (viii) $E(q_i^{-2}\delta_i^2) = nc^{-1}(\eta)\eta/(1 + n\eta)(1 + (n + 2)\eta)$,
- (ix) $E(q_i^{-2}\delta_i^2\varepsilon_i) = \mathbf{0}$,
- (x) $E(q_i^{-2}\delta_i^2\varepsilon_i\varepsilon_i^\top) = [(1 - 2\eta)^2(n + 2)/(1 + n\eta)(1 + (n + 2)\eta)] \Sigma$,
- (xi) $E(q_i^{-2}\delta_i^4) = n(n + 2)(1 - 2\eta)^2/(1 + n\eta)(1 + (n + 2)\eta)$,
- (xii) $E(\log(q_i)) = \psi\left(\frac{1+n\eta}{2\eta}\right) - \psi\left(\frac{1}{2\eta}\right)$.

Osorio e Galea (2015) apresentam lemas relacionados à matriz de comutação, os quais são aplicados neste estudo para distribuição t-Student reparametrizada com repetição, definidos por:

Lema 3. *Assume-se que $\mathbf{Z}_i \sim \mathcal{N}_n(\mathbf{0}, \Sigma)$, então*

$$E(\mathbf{Z}_i\mathbf{Z}_i^\top \otimes \mathbf{Z}_i\mathbf{Z}_i^\top) = 2(\Sigma \otimes \Sigma)\mathbf{N}_n + (\text{vec}\Sigma)(\text{vec}\Sigma)^\top,$$

em que $\mathbf{N}_n = \frac{1}{2}(\mathbf{I}_{n^2} + \mathbf{K}_n)$, com \mathbf{K}_n a matriz de comutação de ordem n e \mathbf{I}_{n^2} matriz identidade de ordem $n^2 \times n^2$.

A partir dos lemas 2 e 3, pode-se estabelecer o seguinte resultado.

Lema 4. *Seja $\varepsilon_i \sim \mathcal{T}_n(\mathbf{0}, \Sigma, \eta)$ e $q = 1 + c(\eta)\delta_i^2$, então*

$$E(q^{-2}\varepsilon_i\varepsilon_i^\top \otimes \varepsilon_i\varepsilon_i^\top) = \frac{(1 - 2\eta)^2}{(1 + n_i\eta)(1 + (n_i + 2)\eta)} \left\{ 2(\Sigma \otimes \Sigma)\mathbf{N}_n + (\text{vec}\Sigma)(\text{vec}\Sigma)^\top \right\},$$

porque $1 + c(\eta)\delta_i^2 \stackrel{d}{=} 1 + \eta n F(n, 1/\eta)$.

Apresenta-se a resolução dos elementos da matriz informação esperada de Fisher, definida por:

$$\mathcal{I}(\boldsymbol{\theta}) = \begin{pmatrix} \mathcal{I}_{\beta\beta}(\boldsymbol{\theta}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathcal{I}_{\phi\phi}(\boldsymbol{\theta}) & \mathcal{I}_{\phi\eta}(\boldsymbol{\theta}) \\ \mathbf{0} & \mathcal{I}_{\eta\phi}(\boldsymbol{\theta}) & \mathcal{I}_{\eta\eta}(\boldsymbol{\theta}) \end{pmatrix},$$

em que

$$\begin{aligned} \mathcal{I}_{\beta\beta}(\boldsymbol{\theta}) &= \frac{1}{r} \sum_{i=1}^r E \left\{ U_i(\boldsymbol{\beta}) U_i^\top(\boldsymbol{\beta}) \right\} \\ &= \frac{1}{r} \sum_{i=1}^r E \left[\left(w_i \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right) \left(w_i \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right)^\top \right] \\ &= \frac{1}{r} \sum_{i=1}^r E \left[w_i^2 \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right] \\ &= \frac{1}{r} \sum_{i=1}^r E \left[\left(\frac{1+n\eta}{\eta} \right)^2 c^2(\eta) q^{-2} \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right] \\ &= \frac{1}{r} \sum_{i=1}^r \left(\frac{1+n\eta}{\eta} \right)^2 c^2(\eta) E \left[\mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} q^{-2} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right] \\ &= \frac{1}{r} \sum_{i=1}^r \left(\frac{1+n\eta}{\eta} \right)^2 c^2(\eta) \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} E \left[q^{-2} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \right] \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \end{aligned}$$

Aplica-se a propriedade (vii) referente ao lema 2 e tem-se que:

$$\begin{aligned} \mathcal{I}_{\beta\beta}(\boldsymbol{\theta}) &= \frac{1}{r} \sum_{i=1}^r \left(\frac{1+n\eta}{\eta} \right)^2 \left(\frac{\eta}{1-2\eta} \right)^2 \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \left[\frac{(1-2\eta)}{(1+n\eta)(1+(n+2)\eta)} \boldsymbol{\Sigma} \right] \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \\ &= \frac{1}{r} \sum_{i=1}^r \frac{(1+n\eta)}{(1-2\eta)(1+(n+2)\eta)} \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \\ &= \frac{c_\beta(\eta)}{r} \sum_{i=1}^r \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \end{aligned}$$

em que $c_\beta(\eta) = (1+n\eta)/(1-2\eta)(1+(n+2)\eta)$.

$$\begin{aligned} \mathcal{I}_{\phi\phi}(\boldsymbol{\theta}) &= E \left\{ U_i(\boldsymbol{\phi}) U_i^\top(\boldsymbol{\phi}) \right\} \\ &= E \left[\left(\frac{1}{2} D_\phi^\top \text{vec} \left[\left(w_i \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \right) - \boldsymbol{\Sigma}^{-1} \right] \right) \left(\frac{1}{2} D_\phi^\top \text{vec} \left[\left(w_i \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \right) - \boldsymbol{\Sigma}^{-1} \right] \right)^\top \right] \\ &= \frac{1}{4} E \left[\left(D_\phi^\top \text{vec} \left(w_i \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \right) \right) \left(D_\phi \text{vec}^\top \left(w_i \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \right) \right) \right. \\ &\quad - \left(D_\phi^\top \text{vec} \left(w_i \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \right) \right) \left(D_\phi \text{vec}^\top \left(\boldsymbol{\Sigma}^{-1} \right) \right) \\ &\quad - \left(D_\phi^\top \text{vec} \left(\boldsymbol{\Sigma}^{-1} \right) \right) \left(D_\phi \text{vec}^\top \left(w_i \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \right) \right) \\ &\quad \left. + \left(D_\phi^\top \text{vec} \left(\boldsymbol{\Sigma}^{-1} \right) \right) \left(D_\phi \text{vec}^\top \left(\boldsymbol{\Sigma}^{-1} \right) \right) \right] \end{aligned}$$

Define-se $\mathcal{I}_{\phi\phi}(\boldsymbol{\theta}) = (1/4) [E(S_1) + E(S_2) + E(S_3) + E(S_4)]$, assim calcula-se separadamente cada $E(S_i)$. Inicialmente, são necessárias algumas considerações em que:

1. $D_\phi = \frac{\partial \text{vec}(\Sigma)}{\partial \phi^\top}$;
2. $w_i = c(\eta) \left(\frac{1+n\eta}{\eta} \right) q^{-1}$;
3. $c(\eta) = \frac{\eta}{1-2\eta}$;
4. $c_\phi(\eta) = \frac{1+n\eta}{(1+(n+2)\eta)}$;
5. $\text{vec}(\Sigma^{-1} w_i \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \Sigma^{-1}) = \text{vec}(\Sigma^{-1} \Sigma^{-1}) \text{vec}(w_i \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top) = (\Sigma^{-1} \otimes \Sigma^{-1}) \text{vec}(w_i \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top)$;
6. $\text{vec}(\Sigma^{-1} w_i \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \Sigma^{-1}) \text{vec}^\top(\Sigma^{-1} w_i \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \Sigma^{-1}) = (\Sigma^{-1} \otimes \Sigma^{-1}) (w_i \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \otimes w_i \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top) (\Sigma^{-1} \otimes \Sigma^{-1})$;
7. $\text{vec}(ABC) = ((C^\top \otimes A) \text{vec}(B))$;
8. $E(\text{vec}(A)) = \text{vec}(E(A))$.

Resolve-se separadamente $\mathcal{I}_{\phi\phi}(\boldsymbol{\theta})$ e tem-se que:

$$\begin{aligned} E(S_4) &= E \left[\left(D_\phi^\top \text{vec}(\Sigma^{-1}) \right) \left(\text{vec}^\top(\Sigma^{-1}) D_\phi \right) \right], \\ &= D_\phi^\top \text{vec}(\Sigma^{-1}) \text{vec}^\top(\Sigma^{-1}) D_\phi. \end{aligned}$$

$$\begin{aligned} E(S_3) &= E \left[- \left(D_\phi^\top \text{vec}(\Sigma^{-1}) \right) \left(D_\phi \text{vec}^\top \left(w_i \Sigma^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \Sigma^{-1} \right) \right) \right], \\ &= -D_\phi^\top \text{vec}(\Sigma^{-1}) E \left[\text{vec}^\top \left(w_i \Sigma^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \Sigma^{-1} \right) \right] D_\phi, \\ &= -D_\phi^\top \text{vec}(\Sigma^{-1}) E \left[(\Sigma^{-1} \otimes \Sigma^{-1}) \text{vec}^\top \left(w_i \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \right) \right] D_\phi, \\ &= -D_\phi^\top \text{vec}(\Sigma^{-1}) (\Sigma^{-1} \otimes \Sigma^{-1}) c(\eta) \left(\frac{1+n\eta}{\eta} \right) E \left[\text{vec}^\top \left(q_i^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top \right) \right] D_\phi, \\ &= -D_\phi^\top \text{vec}(\Sigma^{-1}) (\Sigma^{-1} \otimes \Sigma^{-1}) c(\eta) \left(\frac{1+n\eta}{\eta} \right) \left(\frac{1-2\eta}{1+n\eta} \right) \text{vec}^\top(\Sigma) D_\phi, \\ &= -D_\phi^\top \text{vec}(\Sigma^{-1}) (\Sigma^{-1} \otimes \Sigma^{-1}) \text{vec}^\top(\Sigma) D_\phi, \\ &= -D_\phi^\top \text{vec}(\Sigma^{-1}) \text{vec}^\top(\Sigma^{-1} \Sigma \Sigma^{-1}) D_\phi, \end{aligned}$$

então,

$$E(S_3) = -D_\phi^\top \text{vec}(\Sigma^{-1}) \text{vec}^\top(\Sigma^{-1}) D_\phi.$$

Tem-se que $E(S_3) = E(S_2)$

$$\begin{aligned}
E(S_1) &= E \left[\left(D_\phi^\top \text{vec} \left(w_i \Sigma^{-1} \varepsilon_i \varepsilon_i^\top \Sigma^{-1} \right) \right) \left(D_\phi \text{vec}^\top \left(w_i \Sigma^{-1} \varepsilon_i \varepsilon_i^\top \Sigma^{-1} \right) \right) \right], \\
&= D_\phi^\top E \left[\text{vec} \left(w_i \Sigma^{-1} \varepsilon_i \varepsilon_i^\top \Sigma^{-1} \right) \text{vec}^\top \left(w_i \Sigma^{-1} \varepsilon_i \varepsilon_i^\top \Sigma^{-1} \right) \right] D_\phi, \\
&= D_\phi^\top (\Sigma^{-1} \otimes \Sigma^{-1}) E \left(w_i \varepsilon_i \varepsilon_i^\top \otimes w_i \varepsilon_i \varepsilon_i^\top \right) (\Sigma^{-1} \otimes \Sigma^{-1}) D_\phi, \\
&= (c(\eta))^2 \left(\frac{1+n\eta}{\eta} \right)^2 D_\phi^\top (\Sigma^{-1} \otimes \Sigma^{-1}) E \left(q_i^2 \varepsilon_i \varepsilon_i^\top \otimes \varepsilon_i \varepsilon_i^\top \right) (\Sigma^{-1} \otimes \Sigma^{-1}) D_\phi, \\
&= (c(\eta))^2 \left(\frac{1+n\eta}{\eta} \right)^2 \frac{(1-2\eta)^2}{(1+n\eta)(1+(n+2)\eta)} D_\phi^\top (\Sigma^{-1} \otimes \Sigma^{-1}) (2(\Sigma \otimes \Sigma) N_n \\
&\quad + \text{vec}(\Sigma) \text{vec}^\top(\Sigma)) (\Sigma^{-1} \otimes \Sigma^{-1}) D_\phi, \\
&= c_\phi(\eta) D_\phi^\top (\Sigma^{-1} \otimes \Sigma^{-1}) \left[2(\Sigma \otimes \Sigma) N_n + \text{vec}(\Sigma) \text{vec}^\top(\Sigma) \right] (\Sigma^{-1} \otimes \Sigma^{-1}) D_\phi, \\
&= c_\phi(\eta) D_\phi^\top \left[2(\Sigma^{-1} \otimes \Sigma^{-1}) (\Sigma \otimes \Sigma) N_n \right. \\
&\quad \left. + (\Sigma^{-1} \otimes \Sigma^{-1}) \text{vec}(\Sigma) \text{vec}^\top(\Sigma) \right] (\Sigma^{-1} \otimes \Sigma^{-1}) D_\phi, \\
&= c_\phi(\eta) D_\phi^\top \left[2(\Sigma^{-1} \Sigma \otimes \Sigma^{-1} \Sigma) N_n (\Sigma^{-1} \otimes \Sigma^{-1}) \right. \\
&\quad \left. + (\Sigma^{-1} \otimes \Sigma^{-1}) \text{vec}(\Sigma) \text{vec}^\top(\Sigma) (\Sigma^{-1} \otimes \Sigma^{-1}) \right] D_\phi, \\
&= c_\phi(\eta) D_\phi^\top \left[2(\Sigma^{-1} \otimes \Sigma^{-1}) N_n + \text{vec}(\Sigma^{-1} \Sigma \Sigma^{-1}) \text{vec}^\top(\Sigma^{-1} \Sigma \Sigma^{-1}) \right] D_\phi,
\end{aligned}$$

então,

$$E(S_1) = c_\phi(\eta) D_n^\top \left[2(\Sigma^{-1} \otimes \Sigma^{-1}) N_n + \text{vec}(\Sigma^{-1}) \text{vec}^\top(\Sigma^{-1}) \right] D_\phi.$$

Em que $N_n = \frac{1}{2}(I_{n^2} + K_n)$, K_n é a matriz de comutação de ordem n (MAGNUS; NEUDECKER, 1999), I_{n^2} é a matriz identidade de ordem $n^2 \times n^2$.

Portanto,

$$\begin{aligned}
\mathcal{I}_{\phi\phi}(\boldsymbol{\theta}) &= \frac{1}{4} D_\phi^\top \left[2c_\phi(\eta) (\Sigma^{-1} \otimes \Sigma^{-1}) N_n + \right. \\
&\quad \left. (c_\phi(\eta) - 1) \text{vec}(\Sigma^{-1}) \text{vec}^\top(\Sigma^{-1}) \right] D_\phi,
\end{aligned}$$

Para se determinar $\mathcal{I}_{\phi\eta}(\boldsymbol{\theta})$, define-se da seguinte forma:

$$\begin{aligned}
\mathcal{I}_{\phi\eta}(\boldsymbol{\theta}) &= E\{-\mathbf{L}_{\phi\eta}(\boldsymbol{\theta})\}, \\
&= E\left[-\frac{1}{2\eta^2}\left[(w_i c^2(\eta)q_i^{-1}\delta_i^2 - w_i c(\eta) + q_i^{-1}c(\eta))\left(D_\phi^\top \text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^\top\boldsymbol{\Sigma}^{-1})\right)\right]\right], \\
&= -E\left[\underbrace{\frac{1}{2\eta^2}w_i c^2(\eta)q_i^{-1}\delta_i^2 D_\phi^\top \text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^\top\boldsymbol{\Sigma}^{-1})}_{\text{Parte O}}\right. \\
&\quad \left.-\underbrace{\frac{1}{2\eta^2}w_i c(\eta)D_\phi^\top \text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^\top\boldsymbol{\Sigma}^{-1})}_{\text{Parte P}}\right. \\
&\quad \left.+\underbrace{\frac{1}{2\eta^2}q_i^{-1}c(\eta)D_\phi^\top \text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^\top\boldsymbol{\Sigma}^{-1})}_{\text{Parte Q}}\right], \tag{4.72}
\end{aligned}$$

ao se resolver a parte O, tem-se:

$$\begin{aligned}
\text{Parte O} &= E\left[\frac{1}{2\eta^2}w_i c^2(\eta)q_i^{-1}\delta_i^2 D_\phi^\top \text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^\top\boldsymbol{\Sigma}^{-1})\right], \\
&= E\left[\frac{1}{2\eta^2}\left(\frac{1+n\eta}{\eta}\right)c(\eta)q_i^{-1}c^2(\eta)q_i^{-1}\delta_i^2 D_\phi^\top \text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^\top\boldsymbol{\Sigma}^{-1})\right], \\
&= \frac{1}{2\eta^2}\left(\frac{1+n\eta}{\eta}\right)c^3(\eta)D_\phi^\top \text{vec}\left[E\left(q_i^{-2}\delta_i^2\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^\top\right)\right]\text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^{-1}), \\
&= \frac{1}{2\eta^2}\left(\frac{1+n\eta}{\eta}\right)\frac{\eta^3}{(1-2\eta)^3}D_\phi^\top \frac{(1-2\eta)^2(n+2)}{(1+n\eta)(1+(n+2)\eta)}\text{vec}(\boldsymbol{\Sigma})\text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^{-1}), \\
&= \frac{1}{2}\frac{(n+2)}{(1-2\eta)(1+(n+2)\eta)}D_\phi^\top \text{vec}(\boldsymbol{\Sigma}^{-1}), \tag{4.73}
\end{aligned}$$

para resolver a parte P, tem-se:

$$\begin{aligned}
\text{Parte P} &= E\left[-\frac{1}{2\eta^2}w_i c(\eta)D_\phi^\top \text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^\top\boldsymbol{\Sigma}^{-1})\right], \\
&= E\left[-\frac{1}{2\eta^2}\left(\frac{1+n\eta}{\eta}\right)c(\eta)q_i^{-1}c(\eta)D_\phi^\top \text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^\top\boldsymbol{\Sigma}^{-1})\right], \\
&= -\frac{1}{2\eta^2}\left(\frac{1+n\eta}{\eta}\right)c^2(\eta)D_\phi^\top \text{vec}\left[E\left(q_i^{-1}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^\top\right)\right]\text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^{-1}), \\
&= -\frac{1}{2\eta^2}\left(\frac{1+n\eta}{\eta}\right)\frac{\eta^2}{(1-2\eta)^2}D_\phi^\top \frac{(1-2\eta)}{(1+n\eta)}\text{vec}(\boldsymbol{\Sigma})\text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^{-1}), \\
&= -\frac{1}{2}\frac{1}{\eta(1-2\eta)}D_\phi^\top \text{vec}(\boldsymbol{\Sigma}^{-1}), \tag{4.74}
\end{aligned}$$

e para se resolver a parte Q, tem-se:

$$\begin{aligned}
\text{Parte Q} &= E \left[\frac{1}{2\eta^2} q_i^{-1} c(\eta) D_\phi^\top \text{vec}(\Sigma^{-1} \varepsilon_i \varepsilon_i^\top \Sigma^{-1}) \right], \\
&= \frac{1}{2\eta^2} c(\eta) D_\phi^\top \text{vec} \left[E(q_i^{-1} \varepsilon_i \varepsilon_i^\top) \right] \text{vec}(\Sigma^{-1} \Sigma^{-1}), \\
&= \frac{1}{2\eta^2} \frac{\eta}{(1-2\eta)} D_\phi^\top \frac{(1-2\eta)}{(1+n\eta)} \text{vec}(\Sigma) \text{vec}(\Sigma^{-1} \Sigma^{-1}), \\
&= \frac{1}{2} \frac{1}{\eta(1+n\eta)} D_\phi^\top \text{vec}(\Sigma^{-1}), \tag{4.75}
\end{aligned}$$

assim, quando as equações (4.73), (4.74), (4.75) são substituídas na equação (4.72), tem-se:

$$\mathcal{I}_{\phi\eta}(\boldsymbol{\theta}) = - \left[\left(\frac{(n+2)}{2(1-2\eta)(1+(n+2)\eta)} - \frac{1}{2\eta(1-2\eta)} + \frac{1}{2\eta(1+n\eta)} \right) \left(D_\phi^\top \text{vec}(\Sigma^{-1}) \right) \right],$$

cuja resolução dos cálculos resulta em:

$$\mathcal{I}_{\phi\eta}(\boldsymbol{\theta}) = - \left[\frac{c(\eta)(n+2)}{(1+(n+2)\eta)(1+n\eta)} \right] D_\phi^\top \text{vec}(\Sigma^{-1}),$$

Para se determinar $\mathcal{I}_{\eta\eta}(\boldsymbol{\theta})$, define-se da seguinte forma:

$$\begin{aligned}
\mathcal{I}_{\eta\eta}(\boldsymbol{\theta}) &= E\{-\mathbf{L}_{\eta\eta}(\boldsymbol{\theta})\}, \\
&= -E\left[-\frac{1}{\eta^3}\psi\left(\frac{1}{2\eta}\right) - \frac{1}{4\eta^4}\psi'\left(\frac{1}{2\eta}\right) + \frac{1}{\eta^3}\psi\left(\frac{1+n\eta}{2\eta}\right) + \frac{1}{4\eta^4}\psi'\left(\frac{1+n\eta}{2\eta}\right)\right. \\
&\quad + \frac{1}{\eta^3}c(\eta)(\delta_i^2 w_i - n) + \frac{1}{2\eta^2(1-2\eta)^2}(n - \delta_i^2 w_i) - \frac{1}{\eta^3}\log(q_i) + \frac{\delta_i^2 q_i^{-1}}{2\eta^2(1-2\eta)^2} \\
&\quad \left. - \frac{1}{2\eta^2}\delta_i^2 c(\eta)(1-2\eta)^{-3}[(n+2)(1-2\eta)q_i^{-1} - (1+n\eta)q_i^{-2}\delta_i^2]\right], \\
&= -\left\{-\frac{1}{\eta^3}\psi\left(\frac{1}{2\eta}\right) - \frac{1}{4\eta^4}\psi'\left(\frac{1}{2\eta}\right) + \frac{1}{\eta^3}\psi\left(\frac{1+n\eta}{2\eta}\right) + \frac{1}{4\eta^4}\psi'\left(\frac{1+n\eta}{2\eta}\right)\right. \\
&\quad - \frac{1}{\eta^3}c(\eta)n + \frac{1}{\eta^3}c(\eta)\left(\frac{1+n\eta}{\eta}\right)c(\eta)E[q_i^{-1}\delta_i^2] + \frac{n}{2\eta^2(1-2\eta)^2} \\
&\quad - \frac{1}{2\eta^2(1-2\eta)^2}\left(\frac{1+n\eta}{\eta}\right)c(\eta)E[q_i^{-1}\delta_i^2] - \frac{1}{\eta^3}E[\log(q_i)] \\
&\quad + \frac{1}{2\eta^2(1-2\eta)^2}E[q_i^{-1}\delta_i^2] - \frac{1}{2\eta^2}c(\eta)(1-2\eta)^{-3}(n+2)(1-2\eta)E[q_i^{-1}\delta_i^2] \\
&\quad \left. + \frac{1}{2\eta^2}c(\eta)(1-2\eta)^{-3}(1+n\eta)E[q_i^{-2}\delta_i^4]\right\} \\
&= -\left\{-\frac{1}{\eta^3}\psi\left(\frac{1}{2\eta}\right) - \frac{1}{4\eta^4}\psi'\left(\frac{1}{2\eta}\right) + \frac{1}{\eta^3}\psi\left(\frac{1+n\eta}{2\eta}\right) + \frac{1}{4\eta^4}\psi'\left(\frac{1+n\eta}{2\eta}\right)\right. \\
&\quad - \frac{1}{\eta^3}c(\eta)n + \frac{1}{\eta^3}c^2(\eta)\left(\frac{1+n\eta}{\eta}\right)\left(\frac{n(1-2\eta)}{(1+n\eta)}\right) + \frac{n}{2\eta^2(1-2\eta)^2} \\
&\quad - \frac{c(\eta)}{2\eta^2(1-2\eta)^2}\left(\frac{1+n\eta}{\eta}\right)\left(\frac{n(1-2\eta)}{(1+n\eta)}\right) - \frac{1}{\eta^3}\psi\left(\frac{1+n\eta}{2\eta}\right) + \frac{1}{\eta^3}\psi\left(\frac{1}{2\eta}\right) \\
&\quad + \frac{1}{2\eta^2(1-2\eta)^2}\left(\frac{n(1-2\eta)}{(1+n\eta)}\right) - \frac{1}{2\eta^2}c(\eta)\frac{(n+2)}{(1-2\eta)^3}(1-2\eta)\left(\frac{n(1-2\eta)}{(1+n\eta)}\right) \\
&\quad \left. + \frac{1}{2\eta^2}c(\eta)\frac{(1+n\eta)}{(1-2\eta)^3}\left(\frac{n(n+2)(1-2\eta)^2}{(1+(n+2)\eta)}\right)\right\} \\
&= -\left\{-\frac{1}{4\eta^4}\psi'\left(\frac{1}{2\eta}\right) + \frac{1}{4\eta^4}\psi'\left(\frac{1+n\eta}{2\eta}\right) - \frac{n(n+2)}{2\eta(1+n\eta)(1-2\eta)^2}\right. \\
&\quad \left. + \frac{n(n+2)}{2\eta(1+(n+2)\eta)} + \frac{n}{2\eta^2(1-2\eta)(1+n\eta)}\right\} \\
&= -\left\{\frac{n}{2\eta^2}\left[-\frac{\eta(n+2)}{(1+n\eta)(1-2\eta)^2} + \frac{\eta(n+2)}{(1+(n+2)\eta)} + \frac{1}{(1-2\eta)(1+n\eta)}\right]\right. \\
&\quad \left. - \frac{1}{4\eta^4}\psi'\left(\frac{1}{2\eta}\right) + \frac{1}{4\eta^4}\psi'\left(\frac{1+n\eta}{2\eta}\right)\right\} \\
&= -\left\{\frac{1}{2\eta^2}\left[\frac{n}{(1-2\eta)^2}\left(\frac{1+n\eta(1-4\eta)-8\eta^2}{(1+n\eta)(1+(n+2)\eta)}\right) - \frac{1}{2\eta^2}\psi'\left(\frac{1}{2\eta}\right)\right.\right. \\
&\quad \left. \left. + \frac{1}{2\eta^2}\psi'\left(\frac{1+n\eta}{2\eta}\right)\right]\right\}.
\end{aligned}$$

Portanto,

$$\mathcal{I}_{\eta\eta}(\boldsymbol{\theta}) = -\frac{1}{2\eta^2}\left[\frac{n}{(1-2\eta)^2}\left(\frac{1+n\eta(1-4\eta)-8\eta^2}{(1+n\eta)(1+(n+2)\eta)}\right) - \mathcal{B}'(\eta)\right],$$

em que $B'(\eta)$ é dado por:

$$B'(\eta) = -\frac{1}{2\eta^2} \left(\psi' \left(\frac{1+n\eta}{2\eta} \right) - \psi' \left(\frac{1}{2\eta} \right) \right).$$

4.36 Q-function

Sabemos que $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) = \sum_{i=1}^r Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$ e, para se obter $Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$, calcula-se a esperança condicional do logaritmo da função verossimilhança completa, definida por $Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) = E[\mathcal{L}_{ci}(\boldsymbol{\theta}|\mathbf{Y}_c)V, \boldsymbol{\theta}^{(k)}]$. Assim, tem-se que:

$$\begin{aligned} Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) &= E[\mathcal{L}_{ci}(\boldsymbol{\theta}, \mathbf{Y}_c)|V, \boldsymbol{\theta}^{(k)}] \\ &= -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2} \delta_i^2 E[V|\mathbf{Y}, \boldsymbol{\theta}^{(k)}] \\ &\quad + \frac{1}{2\eta} \log\left(\frac{1}{2c(\eta)}\right) - \log\left(\Gamma\left(\frac{1}{2\eta}\right)\right) + \frac{n}{2} E[\log(V|\mathbf{Y}, \boldsymbol{\theta}^{(k)})] \\ &\quad + \frac{1}{2c(\eta)} E[\log(V|\mathbf{Y}, \boldsymbol{\theta}^{(k)})] - \frac{1}{2c(\eta)} E[(V|\mathbf{Y}, \boldsymbol{\theta}^{(k)})], \end{aligned}$$

$$\begin{aligned} Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) &= -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2} \delta_i^2 w_i^{(k)} \\ &\quad + \frac{1}{2\eta} \log\left(\frac{1}{2c(\eta)}\right) - \log\left(\Gamma\left(\frac{1}{2\eta}\right)\right) + \frac{n}{2} ac_i^{(k)} \\ &\quad + \frac{1}{2c(\eta)} [ac_i^{(k)} - w_i^{(k)}], \end{aligned} \tag{4.76}$$

em que

$$w_i^{(k)} = E[V|\mathbf{Y}, \boldsymbol{\theta}^{(k)}] = \frac{\left(\frac{1+n\eta^{(k)}}{2\eta^{(k)}}\right)}{\left(\frac{1+c(\eta^{(k)})\delta_i^2}{2c(\eta^{(k)})}\right)} = \left(\frac{1+n\eta^{(k)}}{\eta^{(k)}}\right) c(\eta^{(k)}) (q^{(k)})^{-1},$$

$$q^{(k)} = 1 + c(\eta^{(k)})\delta_i^2$$

e

$$ac_i^{(k)} = E[\log(V|\mathbf{Y}, \boldsymbol{\theta}^{(k)})] = \psi\left(\frac{1+n\eta^{(k)}}{2\eta^{(k)}}\right) - \log\left(\frac{1+c(\eta^{(k)})\delta_i^2}{2c(\eta^{(k)})}\right).$$

A partir das propriedades de McLachlan e Krishnan (2007), a equação (4.76) pode ser simplificada por:

$$\begin{aligned} Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) &= -\frac{1}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2} \delta_i^2 w_i^{(k)} \\ &\quad + \frac{1}{2\eta} \log\left(\frac{1}{2c(\eta)}\right) - \log\left(\Gamma\left(\frac{1}{2\eta}\right)\right) + \frac{1}{2c(\eta)} c_i^{(k)} \end{aligned}$$

sendo $c_i^{(k)} = \psi\left(\frac{1+n\eta^{(k)}}{2\eta^{(k)}}\right) - \log\left(\frac{1+n\eta^{(k)}}{2\eta^{(k)}}\right) + \log w_i^{(k)} - (w_i^{(k)})$. Considera-se $\boldsymbol{\theta} = (\boldsymbol{\tau}^\top, \eta)^\top$ em que $\boldsymbol{\tau} = (\boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top)^\top$, pode-se escrever $Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$ como uma soma representada por

$$Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) = Q_{i1}(\boldsymbol{\tau}|\boldsymbol{\theta}^{(k)}) + Q_{i2}(\eta|\boldsymbol{\theta}^{(k)}), \tag{4.77}$$

com

$$Q_{i1}(\boldsymbol{\tau}|\boldsymbol{\theta}^{(k)}) = -\frac{1}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2} \delta_i^2 w_i^{(k)}$$

e

$$Q_{i2}(\eta|\boldsymbol{\theta}^{(k)}) = \frac{1}{2\eta} \log\left(\frac{1}{2c(\eta)}\right) - \log\left(\Gamma\left(\frac{1}{2\eta}\right)\right) + \frac{1}{2c(\eta)} c_i^{(k)}.$$

Portanto, a esperança condicional do logaritmo da função verossimilhança completa para cada $i = 1, \dots, r$ é representada pela equação (4.77). Sendo também definida para todas as repetições como:

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) = -\frac{r}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{i=1}^r \delta_i^2 w_i^{(k)} + r \left\{ \frac{1}{2\eta} \log\left(\frac{1}{2c(\eta)}\right) - \log\left(\Gamma\left(\frac{1}{2\eta}\right)\right) + \frac{1}{2c(\eta)} \left[\frac{1}{r} \sum_{i=1}^r c_i^{(k)} \right] \right\}. \quad (4.78)$$

A separação como soma também é válida para (4.78), sendo $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) = Q_1(\boldsymbol{\tau}|\boldsymbol{\theta}^{(k)}) + Q_2(\eta|\boldsymbol{\theta}^{(k)})$.

4.37 Derivada Q-function

Utiliza-se a função $Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^k)$ definida em (4.77). Para se obter a derivada de primeira e segunda ordem, respectivamente.

4.37.1 Derivada de primeira ordem da Q-function

Com base na equação (4.77), a derivada de primeira ordem de $Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^k)$ com respeito a $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top, \eta)^\top$ é definida pelo vetor $\dot{Q}_i(\boldsymbol{\theta}|\boldsymbol{\theta}^k)$ de ordem 3×1 , dado por:

$$\dot{Q}_i(\boldsymbol{\theta}|\boldsymbol{\theta}^k) = \begin{pmatrix} \dot{Q}_{i\beta} \\ \dot{Q}_{i\phi} \\ \dot{Q}_{i\eta} \end{pmatrix}, \quad (V1)$$

em que cada elemento do vetor (V1) é calculado da seguinte forma:

Primeira derivada de $Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^k)$ em relação à $\boldsymbol{\beta}$:

$$\begin{aligned} \dot{Q}_{i\beta} &= \frac{\partial Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^k)}{\partial \boldsymbol{\beta}}, \\ &= \frac{\partial}{\partial \boldsymbol{\beta}} \left[-\frac{1}{2} \delta_i^2 w_i^k \right], \\ &= -\frac{1}{2} w_i^k \left[\frac{\partial}{\partial \boldsymbol{\beta}} \delta_i^2 \right], \\ &= -\frac{1}{2} w_i^k \left[-2 \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta}) \right], \\ &= w_i^k \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i, \end{aligned}$$

com $\varepsilon_i = (Y_i - X_i\beta)$.

Primeira derivada de $Q_i(\theta|\theta^k)$ em relação à ϕ :

$$\begin{aligned}
\dot{Q}_{i\phi} &= \frac{\partial Q_i(\theta|\theta^k)}{\partial \phi}, \\
&= \frac{\partial}{\partial \phi} \left[-\frac{1}{2} \log |\Sigma| - \frac{1}{2} \delta_i^2 w_i^k \right], \\
&= -\frac{1}{2} \left[\frac{\partial}{\partial \phi} \log |\Sigma| \right] - \frac{1}{2} w_i^k \left[\frac{\partial}{\partial \phi} \delta_i^2 \right], \\
&= -\frac{1}{2} \frac{\partial \text{vec}^\top(\Sigma)}{\partial \phi} \text{vec}(\Sigma^{-1}) - \frac{1}{2} w_i^k \left[-\frac{\partial \text{vec}^\top(\Sigma)}{\partial \phi} \left(\text{vec}(\Sigma^{-1} \varepsilon_i \varepsilon_i^\top \Sigma^{-1}) \right) \right], \\
&= -\frac{1}{2} \frac{\partial \text{vec}^\top(\Sigma)}{\partial \phi} \text{vec}(\Sigma^{-1}) + \frac{1}{2} w_i^k \frac{\partial \text{vec}^\top(\Sigma)}{\partial \phi} \left(\text{vec}(\Sigma^{-1} \varepsilon_i \varepsilon_i^\top \Sigma^{-1}) \right), \\
&= -\frac{1}{2} D_\phi^\top \left[\text{vec}(\Sigma^{-1}) - w_i^k \left(\text{vec}(\Sigma^{-1} \varepsilon_i \varepsilon_i^\top \Sigma^{-1}) \right) \right],
\end{aligned}$$

em que $D_\phi^\top = \frac{\partial \text{vec}^\top(\Sigma)}{\partial \phi}$. Pode-se obter também a primeira derivada de $Q_i(\theta|\theta^k)$ em relação à ϕ , da seguinte maneira:

$$\dot{Q}_{i\phi} = \frac{\partial Q_i(\theta|\theta^k)}{\partial \phi}, \text{ com elementos}$$

$$\begin{aligned}
\frac{\partial Q_i(\theta|\theta^k)}{\partial \phi_j} &= \frac{\partial}{\partial \phi_j} \left[-\frac{1}{2} \log |\Sigma| - \frac{1}{2} \delta_i^2 w_i^k \right], \\
&= -\frac{1}{2} \left[\text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \right) \right] - \frac{1}{2} w_i^k \left[-\varepsilon_i^\top \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \varepsilon_i \right], \\
&= -\frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \right) + \frac{1}{2} w_i^k \varepsilon_i^\top \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1} \varepsilon_i,
\end{aligned} \tag{4.79}$$

em que $j = 1, 2, 3$. Primeira derivada de $Q_i(\theta|\theta^k)$ em relação à η :

$$\begin{aligned}
\dot{Q}_{i\eta} &= \frac{\partial Q_i(\theta|\theta^k)}{\partial \eta}, \\
&= \frac{\partial}{\partial \eta} \left\{ \frac{1}{2\eta} \log \left(\frac{1}{2c(\eta)} \right) - \log \left(\Gamma \left(\frac{1}{2\eta} \right) \right) + \frac{1}{2c(\eta)} c_i^{(k)} \right\}.
\end{aligned}$$

Logo, quando se resolve cada expressão da equação (4.80), tem-se:

$$\frac{\partial}{\partial \eta} \left[\frac{1}{2\eta} \log \left(\frac{1}{2c(\eta)} \right) \right] = \frac{1}{2} \left[-\frac{1}{\eta^2} \log \left(\frac{1}{2c(\eta)} \right) - \frac{c(\eta)}{\eta^3} \right] \tag{4.80}$$

$$\frac{\partial}{\partial \eta} \left[-\log \left(\Gamma \left(\frac{1}{2\eta} \right) \right) \right] = \frac{1}{2\eta^2} \psi \left(\frac{1}{2\eta} \right) \tag{4.81}$$

$$\frac{\partial}{\partial \eta} \left[\frac{1}{2c(\eta)} c_i^{(k)} \right] = -\frac{1}{2\eta^2} c_i^{(k)} \tag{4.82}$$

A partir das resoluções (4.80), (4.81) e (4.82), tem-se

$$\dot{Q}_{i\eta} = -\frac{1}{2\eta^2} \left[\log \left(\frac{1}{2c(\eta)} \right) + \frac{c(\eta)}{\eta} - \psi \left(\frac{1}{2\eta} \right) + c_i^{(k)} \right].$$

4.37.2 Derivada de segunda ordem da Q-function

Com base nas derivadas de primeira ordem, a derivada de segunda ordem de $Q_i(\theta|\theta^k)$ com respeito à $\theta = (\beta^\top, \phi^\top, \eta)^\top$ pode ser obtida da seguinte forma:

$$\ddot{Q}_i(\theta|\theta^k) = \frac{\partial^2 Q_i(\theta|\theta^k)}{\partial \theta \partial \theta^\top},$$

e representada pela matriz de ordem 3×3 , da forma:

$$\ddot{Q}_i(\theta|\theta^k) = \begin{pmatrix} \ddot{Q}_{i\beta\beta} & \ddot{Q}_{i\beta\phi} & \ddot{Q}_{i\beta\eta} \\ \ddot{Q}_{i\phi\beta} & \ddot{Q}_{i\phi\phi} & \ddot{Q}_{i\phi\eta} \\ \ddot{Q}_{i\eta\beta} & \ddot{Q}_{i\eta\phi} & \ddot{Q}_{i\eta\eta} \end{pmatrix}, \quad (\text{M3})$$

em que cada elemento da matriz (M3) é definido por:

$$\begin{aligned} \ddot{Q}_{i\beta\beta} &= \frac{\partial^2 Q_i(\theta|\theta^k)}{\partial \beta \partial \beta^\top} \\ &= \frac{\partial \dot{Q}_{i\beta}}{\partial \beta^\top} \\ &= \frac{\partial}{\partial \beta^\top} \left(w_i^k \mathbf{X}_i^\top \Sigma^{-1} \varepsilon_i \right) \\ &= w_i^k \mathbf{X}_i^\top \Sigma^{-1} \left[\frac{\partial}{\partial \beta^\top} (\mathbf{Y}_i - \mathbf{X}_i \beta) \right] \\ &= -w_i^k \mathbf{X}_i^\top \Sigma^{-1} \mathbf{X}_i. \end{aligned}$$

Para $\ddot{Q}_{i\beta\phi}$, sabemos que $\ddot{Q}_{i\beta\phi} = \ddot{Q}_{i\phi\beta}$. Logo, iremos calcular $\ddot{Q}_{i\phi\beta}$ assim, tem-se que:

$$\begin{aligned} \ddot{Q}_{i\phi\beta} &= \frac{\partial^2 Q_i(\theta|\theta^k)}{\partial \phi \partial \beta} \\ &= \frac{\partial \dot{Q}_{i\beta}}{\partial \phi} \\ &= \frac{\partial}{\partial \phi} \left(w_i^k \mathbf{X}_i^\top \Sigma^{-1} \varepsilon_i \right) \\ &= w_i^k \mathbf{X}_i^\top \left(\frac{\partial \Sigma^{-1}}{\partial \phi} \right) \varepsilon_i \\ &= -w_i^k \mathbf{X}_i^\top \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi} \Sigma^{-1} \varepsilon_i. \end{aligned}$$

Portanto, $\ddot{Q}_{i\phi\beta} = \ddot{Q}_{i\beta\phi} = -w_i^k \mathbf{X}_i^\top \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi} \Sigma^{-1} \varepsilon_i$.

Para $\ddot{Q}_{i\beta\eta}$, sabe-se que $\ddot{Q}_{i\beta\eta} = \ddot{Q}_{i\eta\beta}$. Logo, resolveu-se somente $\ddot{Q}_{i\eta\beta}$, e tem-se que:

$$\begin{aligned} \ddot{Q}_{i\eta\beta} &= \frac{\partial^2 Q_i(\theta|\theta^k)}{\partial \eta \partial \beta} \\ &= \frac{\partial \dot{Q}_{i\beta}}{\partial \eta} \\ &= \frac{\partial}{\partial \eta} \left(w_i^k \mathbf{X}_i^\top \Sigma^{-1} \varepsilon_i \right) \\ &= 0. \end{aligned}$$

Portanto, $\ddot{Q}_{i\beta\eta} = \ddot{Q}_{i\eta\beta} = 0$. Para se obter a $\ddot{Q}_{i\phi\phi}$ com elementos (j, l) , utilizou-se a $\dot{Q}_{i\phi\phi}$ apresentada na equação (4.79), em que

$$\begin{aligned}\ddot{Q}_{i\phi\phi} &= \left[\left(\frac{\partial^2 Q_i(\boldsymbol{\theta} | \boldsymbol{\theta}^k)}{\partial \phi_j \partial \phi_l} \right) \right], \\ &= \frac{\partial \dot{Q}_{i\phi}}{\partial \phi_l}, \\ &= \frac{\partial}{\partial \phi_l} \left[-\frac{1}{2} \text{tr} \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \right) + \frac{1}{2} w_i^k \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right],\end{aligned}\quad (4.83)$$

Para resolver separadamente a equação (4.83), tem-se:

$$\begin{aligned}\frac{\partial}{\partial \phi_l} \left[-\frac{1}{2} \text{tr} \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \right) \right] &= -\frac{1}{2} \text{tr} \left[\frac{\partial}{\partial \phi_l} \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \right) \right] \\ &= -\frac{1}{2} \text{tr} \left[\frac{\partial \boldsymbol{\Sigma}^{-1}}{\partial \phi_l} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} + \boldsymbol{\Sigma}^{-1} \frac{\partial^2 \boldsymbol{\Sigma}}{\partial \phi_j \partial \phi_l} \right] \\ &= -\frac{1}{2} \text{tr} \left[-\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} + \boldsymbol{\Sigma}^{-1} \frac{\partial^2 \boldsymbol{\Sigma}}{\partial \phi_j \partial \phi_l} \right] \\ &= \frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} - \boldsymbol{\Sigma}^{-1} \frac{\partial^2 \boldsymbol{\Sigma}}{\partial \phi_j \partial \phi_l} \right]\end{aligned}\quad (4.84)$$

e

$$\begin{aligned}\frac{\partial}{\partial \phi_l} \left[\frac{1}{2} w_i^k \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right] &= \frac{w_i^k}{2} \frac{\partial}{\partial \phi_l} \left[\boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right], \\ &= \frac{w_i^k}{2} \left\{ \left[\frac{\partial}{\partial \phi_l} \left(\boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \right) \right] \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right. \\ &\quad \left. + \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \left[\frac{\partial}{\partial \phi_l} \left(\frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right) \right] \right\}, \\ &= \frac{w_i^k}{2} \left[\boldsymbol{\varepsilon}_i^\top \frac{\partial \boldsymbol{\Sigma}^{-1}}{\partial \phi_l} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right. \\ &\quad \left. + \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \left(\frac{\partial^2 \boldsymbol{\Sigma}}{\partial \phi_j \partial \phi_l} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i + \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \frac{\partial \boldsymbol{\Sigma}^{-1}}{\partial \phi_l} \boldsymbol{\varepsilon}_i \right) \right], \\ &= \frac{w_i^k}{2} \left[-\boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right. \\ &\quad \left. + \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1} \left(\frac{\partial^2 \boldsymbol{\Sigma}}{\partial \phi_j \partial \phi_l} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i - \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \right) \right], \\ &= -\frac{1}{2} \left\{ w_i^k \boldsymbol{\varepsilon}_i^\top \left[\boldsymbol{\Sigma}^{-1} \left(\frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} - \frac{\partial^2 \boldsymbol{\Sigma}}{\partial \phi_j \partial \phi_l} \right) \right. \right. \\ &\quad \left. \left. + \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \boldsymbol{\Sigma}^{-1} \right] \boldsymbol{\varepsilon}_i \right\}.\end{aligned}\quad (4.85)$$

Quando as equações (4.84) e (4.85) são substituídas em (4.83), obtém-se $\ddot{Q}_{i\phi\phi}$:

$$\begin{aligned}\ddot{Q}_{i\phi\phi} &= \frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} - \boldsymbol{\Sigma}^{-1} \frac{\partial^2 \boldsymbol{\Sigma}}{\partial \phi_j \partial \phi_l} \right] \\ &\quad - \frac{1}{2} \left\{ w_i^k \boldsymbol{\varepsilon}_i^\top \left[\boldsymbol{\Sigma}^{-1} \left(\frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} - \frac{\partial^2 \boldsymbol{\Sigma}}{\partial \phi_j \partial \phi_l} + \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi_l} \right) \boldsymbol{\Sigma}^{-1} \right] \boldsymbol{\varepsilon}_i \right\}.\end{aligned}$$

Para $\ddot{Q}_{i\phi\eta}$, sabe-se que $\ddot{Q}_{i\phi\eta} = \ddot{Q}_{i\eta\phi}$. Portanto, resolveu-se somente uma derivada $\ddot{Q}_{i\eta\phi}$, logo, tem-se:

$$\begin{aligned}\ddot{Q}_{i\eta\phi} &= \frac{\partial^2 Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^k)}{\partial\eta\partial\phi}, \\ &= \frac{\partial\dot{Q}_{i\phi}}{\partial\eta}, \\ &= \frac{\partial}{\partial\eta} \left\{ -\frac{1}{2} D_{\phi}^{\top} \left[\text{vec}(\boldsymbol{\Sigma}^{-1}) - w_i^k (\text{vec}(\boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^{\top} \boldsymbol{\Sigma}^{-1})) \right] \right\}, \\ &= 0.\end{aligned}$$

Portanto, $\ddot{Q}_{i\eta\phi} = \ddot{Q}_{i\phi\eta} = 0$.

$$\begin{aligned}\ddot{Q}_{i\eta\eta} &= \frac{\partial^2 Q_i(\boldsymbol{\theta}|\boldsymbol{\theta}^k)}{\partial\eta\partial\eta^{\top}}, \\ &= \frac{\partial\dot{Q}_{i\eta}}{\partial\eta}, \\ &= \frac{\partial}{\partial\eta} \left\{ -\frac{1}{2\eta^2} \left[\log\left(\frac{1}{2c(\eta)}\right) + \frac{c(\eta)}{\eta} - \psi\left(\frac{1}{2\eta}\right) + c_i^{(k)} \right] \right\}\end{aligned}\tag{4.86}$$

Para se resolver separadamente a equação (4.86), tem-se

$$\frac{\partial}{\partial\eta} \left[-\frac{1}{2\eta^2} \log\left(\frac{1}{2c(\eta)}\right) \right] = \frac{1}{\eta^3} \log\left(\frac{1}{2c(\eta)}\right) + \frac{c(\eta)}{2\eta^4},\tag{4.87}$$

$$\frac{\partial}{\partial\eta} \left[\frac{1}{2\eta^2} \psi\left(\frac{1}{2\eta}\right) \right] = -\frac{1}{\eta^3} \psi\left(\frac{1}{2\eta}\right) - \frac{1}{4\eta^4} \psi'\left(\frac{1}{2\eta}\right)\tag{4.88}$$

$$\frac{\partial}{\partial\eta} \left[-\frac{1}{2\eta^2} \frac{c(\eta)}{\eta} \right] = \frac{\eta - 3\eta^2}{(\eta^2(1 - 2\eta))^2}\tag{4.89}$$

$$\frac{\partial}{\partial\eta} \left[-\frac{1}{2\eta^2} c_i^{(k)} \right] = \frac{c_i^{(k)}}{\eta^3}\tag{4.90}$$

A partir das resoluções (4.87), (4.88), (4.89), e (4.90). Obtém-se $\ddot{Q}_{i\eta\eta}$, dado por:

$$\begin{aligned}\ddot{Q}_{i\eta\eta} &= \frac{1}{\eta^3} \log\left(\frac{1}{2c(\eta)}\right) + \frac{c(\eta)}{2\eta^4} - \frac{1}{\eta^3} \psi\left(\frac{1}{2\eta}\right) - \frac{1}{4\eta^4} \psi'\left(\frac{1}{2\eta}\right) \\ &\quad + \frac{\eta - 3\eta^2}{(\eta^2(1 - 2\eta))^2} + \frac{c_i^{(k)}}{\eta^3}.\end{aligned}$$

4.37.3 Esperança da Q-function

Para se obter esperança condicional de $\ddot{Q}_i(\theta|\theta^k)$, aplica-se a esperança em cada elemento da matriz (M3). Assim, obtem-se a seguinte matriz:

$$E[-\ddot{Q}_i(\theta|\theta^k)] = \begin{pmatrix} E[-\ddot{Q}_{i\beta\beta}] & E[-\ddot{Q}_{i\beta\phi}] & E[-\ddot{Q}_{i\beta\eta}] \\ E[-\ddot{Q}_{i\phi\beta}] & E[-\ddot{Q}_{i\phi\phi}] & E[-\ddot{Q}_{i\phi\eta}] \\ E[-\ddot{Q}_{i\eta\beta}] & E[-\ddot{Q}_{i\eta\phi}] & E[-\ddot{Q}_{i\eta\eta}] \end{pmatrix}, \quad (\text{M4})$$

cada elemento da matriz (M6) é dado por:

$$\begin{aligned} E[-\ddot{Q}_{i\beta\beta}] &= E\left[-\left(-w_i^k \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right)\right] \\ &= w_i^k \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_i. \end{aligned}$$

$$\begin{aligned} E[-\ddot{Q}_{i\beta\phi}] = E[-\ddot{Q}_{i\phi\beta}] &= E\left[-\left(-w_i^k \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_i\right)\right] \\ &= w_i^k \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \phi} \boldsymbol{\Sigma}^{-1} E[\boldsymbol{\varepsilon}_i] = 0. \end{aligned}$$

$$E[-\ddot{Q}_{i\beta\eta}] = E[-\ddot{Q}_{i\eta\beta}] = 0.$$

$$\begin{aligned}
E[-\ddot{Q}_{i\phi\phi}] &= -E\left\{\frac{1}{2}\text{tr}\left[\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_l}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j}-\Sigma^{-1}\frac{\partial^2\Sigma}{\partial\phi_j\partial\phi_l}\right],\right. \\
&\quad \left.-\frac{1}{2}\left\{w_i^k\varepsilon_i^\top\left[\Sigma^{-1}\left(\frac{\partial\Sigma}{\partial\phi_l}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j}-\frac{\partial^2\Sigma}{\partial\phi_j\partial\phi_l}+\frac{\partial\Sigma}{\partial\phi_j}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_l}\right)\Sigma^{-1}\right],\varepsilon_i\right\}\right\}, \\
&= -\frac{1}{2}\text{tr}\left[\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_l}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j}-\Sigma^{-1}\frac{\partial^2\Sigma}{\partial\phi_j\partial\phi_l}\right] \\
&\quad +\frac{w_i^k}{2}E\left[\varepsilon_i^\top\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_l}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j}\Sigma^{-1}\varepsilon_i\right] \\
&\quad -\frac{w_i^k}{2}E\left[\varepsilon_i^\top\Sigma^{-1}\frac{\partial^2\Sigma}{\partial\phi_j\partial\phi_l}\Sigma^{-1}\varepsilon_i\right] \\
&\quad +\frac{w_i^k}{2}E\left[\varepsilon_i^\top\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_l}\Sigma^{-1}\varepsilon_i\right], \\
&= -\frac{1}{2}\text{tr}\left[\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_l}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j}-\Sigma^{-1}\frac{\partial^2\Sigma}{\partial\phi_j\partial\phi_l}\right] \\
&\quad +\frac{w_i^k}{2}\text{tr}\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_l}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j}\Sigma^{-1}\Sigma\right) \\
&\quad -\frac{w_i^k}{2}\text{tr}\left(\Sigma^{-1}\frac{\partial^2\Sigma}{\partial\phi_j\partial\phi_l}\Sigma^{-1}\Sigma\right) \\
&\quad +\frac{w_i^k}{2}\text{tr}\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_l}\Sigma^{-1}\Sigma\right), \\
&= -\frac{1}{2}\text{tr}\left[\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_l}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j}-\Sigma^{-1}\frac{\partial^2\Sigma}{\partial\phi_j\partial\phi_l}\right] \\
&\quad +\frac{w_i^k}{2}\text{tr}\left[\Sigma^{-1}\left(\frac{\partial\Sigma}{\partial\phi_l}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j}-\frac{\partial^2\Sigma}{\partial\phi_j\partial\phi_l}+\frac{\partial\Sigma}{\partial\phi_j}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_l}\right)\right].
\end{aligned}$$

$$E[-\ddot{Q}_{i\phi\eta}] = E[-\ddot{Q}_{i\eta\phi}] = 0.$$

$$\begin{aligned}
E[-\ddot{Q}_{i\eta\eta}] &= -\left\{\frac{1}{\eta^3}\log\left(\frac{1}{2c(\eta)}\right)+\frac{c(\eta)}{2\eta^4}-\frac{1}{\eta^3}\psi\left(\frac{1}{2\eta}\right)-\frac{1}{4\eta^4}\psi'\left(\frac{1}{2\eta}\right)\right. \\
&\quad \left.+\frac{\eta-3\eta^2}{(\eta^2(1-2\eta))^2}+\frac{c_i^{(k)}}{\eta^3}\right\}.
\end{aligned}$$

4.38 Esquemas de perturbação para a influência local

4.38.1 Ponderação de casos

A Q -function perturbada para a ponderação de casos é definida por:

$$Q(\boldsymbol{\theta}, \omega|\hat{\boldsymbol{\theta}}) = \sum_{i=1}^r \omega_i Q_i(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}) = \sum_{i=1}^r \omega_i \left\{ Q_{1i}(\boldsymbol{\theta}_1|\hat{\boldsymbol{\theta}}) + Q_{2i}(\boldsymbol{\theta}_2|\hat{\boldsymbol{\theta}}) \right\},$$

em que $Q_{1i}(\theta_1|\hat{\theta})$ e $Q_{2i}(\theta_1|\hat{\theta})$ são definidos da seguinte forma:

$$Q_{1i}(\theta_1|\hat{\theta}) = -\frac{1}{2} \log |\Sigma| - \frac{w_i^k}{2} \delta_i^2, \quad (4.91)$$

e

$$Q_{2i}(\theta_2|\hat{\theta}) = +\frac{1}{2\eta} \log \left(\frac{1}{2c(\eta)} \right) - \log \left(\Gamma \left(\frac{1}{2\eta} \right) \right) + \frac{1}{2c(\eta)} c_i^{(k)}. \quad (4.92)$$

sendo $\delta_i^2 = (\mathbf{Y}_i - \mathbf{X}_i\beta)^\top \Sigma^{-1} (\mathbf{Y}_i - \mathbf{X}_i\beta)$. Verifica-se que $Q_{1i}(\theta_1|\hat{\theta})$ dada em (4.91) está em função de β e ϕ , pois define-se $\theta_1 = (\beta^\top, \phi^\top)^\top$, já para $Q_{2i}(\theta_2|\hat{\theta})$ dada em (4.92) está em função de $\theta_2 = \eta$. Para se obter a derivada de $Q(\theta, \omega|\hat{\theta})$ em relação à ω , considera-se

$$\begin{aligned} \dot{Q}(\theta, \omega|\hat{\theta}) &= \frac{\partial Q(\theta, \omega|\hat{\theta})}{\partial \omega^\top}, \\ &= \frac{\partial \sum_{i=1}^r \omega_i Q_i(\theta|\hat{\theta})}{\partial \omega^\top}, \\ &= \sum_{i=1}^r \left[\frac{\partial \omega_i Q_i(\theta|\hat{\theta})}{\partial \omega^\top} \right], \\ &= \sum_{i=1}^r Q_i(\theta|\hat{\theta}) e_{ir}^\top. \end{aligned}$$

Seja e_{ir} o vetor de ordem $(r \times 1)$, somente na i -ésima posição tem-se $\partial \omega_i / \partial \omega_i = 1$, as demais posições são iguais a 0. Tem-se que $\Delta = (\Delta_{1\beta}^\top, \Delta_{1\phi}^\top, \Delta_{2\eta}^\top)^\top$, em que $\Delta_{1\beta} = (\Delta_{11\beta}^\top, \dots, \Delta_{1r\beta}^\top)^\top$ com dimensão $p \times r$, $\Delta_{1\phi} = (\Delta_{11\phi}^\top, \dots, \Delta_{1r\phi}^\top)^\top$ com dimensão $3 \times r$ e $\Delta_{2\eta} = (\Delta_{21\eta}^\top, \dots, \Delta_{2r\eta}^\top)^\top$ com dimensão $1 \times r$, os quais são obtidos da seguinte forma:

$$\begin{aligned} \Delta_{1\beta} &= \frac{\partial^2 Q(\theta, \omega|\hat{\theta})}{\partial \beta \partial \omega^\top}, \\ &= \frac{\partial \dot{Q}(\theta, \omega|\hat{\theta})}{\partial \beta}, \\ &= \sum_{i=1}^r \dot{Q}_{1i\beta} e_{ir}^\top, \\ &= (\dot{Q}_{11\beta}, \dots, \dot{Q}_{1r\beta})^\top, \end{aligned}$$

em que $\dot{Q}_{1i\beta} = w_i^k \mathbf{X}_i^\top \Sigma^{-1} \varepsilon_i$ com dimensão $p \times 1$. Para se obter $\Delta_{1\phi}$, calcula-se

$$\begin{aligned} \Delta_{1\phi} &= \frac{\partial^2 Q(\theta, \omega|\hat{\theta})}{\partial \phi \partial \omega^\top}, \\ &= \frac{\partial^2 \dot{Q}(\theta, \omega|\hat{\theta})}{\partial \phi}, \\ &= \sum_{i=1}^r \dot{Q}_{1i\phi} e_{ir}^\top, \\ &= (\dot{Q}_{11\phi}, \dots, \dot{Q}_{1r\phi})^\top, \end{aligned}$$

com $\dot{Q}_{1i\phi} = \frac{1}{2}D_{\phi}^{\top} \text{vec} [(w_i^k \Sigma^{-1} \varepsilon_i \varepsilon_i^{\top} \Sigma^{-1}) - \Sigma^{-1}]$ com dimensão 3×1 . Para calcular $\Delta_{2\eta}$ tem-se:

$$\begin{aligned} \Delta_{2\eta} &= \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \eta \partial \boldsymbol{\omega}^{\top}}, \\ &= \frac{\partial^2 \dot{Q}(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \eta}, \\ &= \sum_{i=1}^r \dot{Q}_{2i\eta} \mathbf{e}_{ir}^{\top}, \\ &= (\dot{Q}_{21\eta}, \dots, \dot{Q}_{2r\eta})^{\top}, \end{aligned}$$

sendo $\dot{Q}_{2i\eta}$ um escalar da forma:

$$\dot{Q}_{2i\eta} = -\frac{1}{2\eta^2} \left[\log \left(\frac{1}{2c(\eta)} \right) + \frac{c(\eta)}{\eta} - \psi \left(\frac{1}{2\eta} \right) + c_i^{(k)} \right].$$

4.38.2 Função escore perturbada para encontrar a perturbação apropriada para a variável resposta

Um esquema de perturbação é tal que a função de densidade de perturbação pode ser escrita por:

$$\mathbf{G}(\boldsymbol{\omega}_0) = c\mathbf{I}_n$$

com $c > 0$, $\mathbf{G}(\boldsymbol{\omega}) = E_{\boldsymbol{\omega}} [\dot{Q}_{\boldsymbol{\omega}} \dot{Q}_{\boldsymbol{\omega}}^{\top}]$ em que $E_{\boldsymbol{\omega}}$ é denotada a esperança em relação à densidade do modelo perturbado $f(\mathbf{Y}_c, \hat{\boldsymbol{\theta}}, \boldsymbol{\omega})$, sendo $\dot{Q}_{\boldsymbol{\omega}} = \frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\omega}}$ avaliado em $\boldsymbol{\omega} = \boldsymbol{\omega}_0$. Sabemos que

$$Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}}) = \sum_{i=1}^r [Q_{1i}(\boldsymbol{\theta}_1, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}}) + Q_{2i}(\boldsymbol{\theta}_2, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})], \quad (4.93)$$

em que

$$Q_{1i}(\boldsymbol{\theta}_1, \boldsymbol{\omega}_i | \hat{\boldsymbol{\theta}}) = -\frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{w_i^k}{2} \delta_{i\boldsymbol{\omega}}^2$$

com $\delta_{i\boldsymbol{\omega}}^2 = \boldsymbol{\varepsilon}_{i\boldsymbol{\omega}}^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_{i\boldsymbol{\omega}}$ sendo $\boldsymbol{\varepsilon}_{i\boldsymbol{\omega}} = [(\mathbf{Y}_i + \mathbf{A}\boldsymbol{\omega}_i) - \mathbf{X}_i \boldsymbol{\beta}] = (\boldsymbol{\varepsilon}_i + \mathbf{A}\boldsymbol{\omega}_i)$. Observa-se que $Q_{2i}(\boldsymbol{\theta}_2 | \hat{\boldsymbol{\theta}})$ não depende de $\boldsymbol{\omega}$. Ao se derivar a equação (4.93) em relação ao vetor $\boldsymbol{\omega}$, obtém-se:

$$\begin{aligned} \dot{Q}_{\boldsymbol{\omega}} &= \frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\omega}}, \\ &= \frac{\partial \sum_{i=1}^r [Q_{1i}(\boldsymbol{\theta}_1, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}}) + Q_{2i}(\boldsymbol{\theta}_2, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})]}{\partial \boldsymbol{\omega}}, \\ &= \sum_{i=1}^r \frac{\partial Q_{1i}(\boldsymbol{\theta}_1, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\omega}}, \\ &= \sum_{i=1}^r \left[\frac{\partial}{\partial \boldsymbol{\omega}} \left(-\frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{w_i^k}{2} \delta_{i\boldsymbol{\omega}}^2 \right) \right], \\ &= \sum_{i=1}^r \left(-w_i^k \mathbf{A}^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_{i\boldsymbol{\omega}} \right) \end{aligned} \quad (4.94)$$

A partir da equação (4.94), deve-se validar $G(\omega)$. Assim,

$$\begin{aligned}
 G(\omega) &= E_{\omega} \left[\dot{Q}_{\omega} \dot{Q}_{\omega}^{\top} \right], \\
 &= E_{\omega} \left\{ \left[\sum_{i=1}^r \left(-w_i^k \mathbf{A}^{\top} \Sigma^{-1} \varepsilon_{i\omega} \right) \right] \left[\sum_{i=1}^r \left(-w_i^k \mathbf{A}^{\top} \Sigma^{-1} \varepsilon_{i\omega} \right) \right]^{\top} \right\}, \\
 &= \sum_{i=1}^r E_{\omega} \left[(w_i^k)^2 \mathbf{A}^{\top} \Sigma^{-1} \varepsilon_{i\omega} \varepsilon_{i\omega}^{\top} \Sigma^{-1} \mathbf{A} \right], \\
 &= \sum_{i=1}^r (w_i^k)^2 \mathbf{A}^{\top} \Sigma^{-1} \Sigma \Sigma^{-1} \mathbf{A}, \\
 &= \sum_{i=1}^r (w_i^k)^2 \mathbf{A}^{\top} \Sigma^{-1} \mathbf{A}.
 \end{aligned}$$

Considera-se $c = (w_i^k)^2$ uma constante positiva, tem-se $G(\omega_0) = c \mathbf{A}^{\top} \Sigma^{-1} \mathbf{A}$ para que $G(\omega_0) = c \mathbf{I}_n$, ao se considerar $\mathbf{A} = \Sigma^{1/2}$. Logo, a perturbação $\mathbf{Y}_{\omega} = \mathbf{Y}_i + \Sigma^{1/2} \omega_i$ é uma perturbação adequada.

4.38.3 Perturbação na variável resposta

Para o esquema de perturbação na variável resposta baseada na função de log-verossimilhança, para os dados completos é definido por:

$$Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}}) = \sum_{i=1}^r \left\{ Q_{1i}(\boldsymbol{\theta}_1, \omega_i | \hat{\boldsymbol{\theta}}) + Q_{2i}(\boldsymbol{\theta}_2 | \hat{\boldsymbol{\theta}}) \right\},$$

em que

$$Q_{1i}(\boldsymbol{\theta}_1, \omega_i | \hat{\boldsymbol{\theta}}) = -\frac{1}{2} \log |\Sigma| - \frac{w_i^k}{2} \delta_{i\omega}^2$$

com $\delta_{i\omega}^2 = \varepsilon_{i\omega}^{\top} \Sigma^{-1} \varepsilon_{i\omega}$ sendo $\varepsilon_{i\omega} = [(\mathbf{Y}_i + \mathbf{A} \omega_i) - \mathbf{X}_i \boldsymbol{\beta}]$ em que $\mathbf{A} = \Sigma^{1/2}$. Sabemos que $Q_{2i}(\boldsymbol{\theta}_2 | \hat{\boldsymbol{\theta}})$ não depende de ω_i . Logo, a derivada de $Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})$ em relação a $\boldsymbol{\omega}$ para este

esquema de perturbação é obtida por:

$$\begin{aligned}
\dot{Q}_{1i\omega} &= \frac{\partial Q_{1i}(\boldsymbol{\theta}_1, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \omega_i^\top}, \\
&= \frac{\partial}{\partial \omega_i^\top} \left[-\frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{w_i^k}{2} \delta_{i\omega}^2 \right], \\
&= -\frac{w_i^k}{2} \left[\frac{\partial \delta_{i\omega}^2}{\partial \omega_i^\top} \right], \\
&= -\frac{w_i^k}{2} \left[\frac{\partial}{\partial \omega_i^\top} \left(\boldsymbol{\varepsilon}_{i\omega}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_{i\omega} \right) \right], \\
&= -\frac{w_i^k}{2} \left(2 \frac{\partial \boldsymbol{\varepsilon}_{i\omega}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_{i\omega}}{\partial \omega_i} \right)^\top, \\
&= -\frac{w_i^k}{2} \left\{ 2 \left[\frac{\partial (\mathbf{Y}_i + \boldsymbol{\Sigma}^{1/2} \boldsymbol{\omega}_i - \mathbf{X}_i \boldsymbol{\beta})^\top}{\partial \omega_i} \right] \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_{i\omega} \right\}^\top, \\
&= -w_i^k (\boldsymbol{\Sigma}^{1/2} \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_{i\omega})^\top, \\
&= -w_i^k \boldsymbol{\varepsilon}_{i\omega}^\top \boldsymbol{\Sigma}^{-1/2}.
\end{aligned}$$

Para se obter $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\beta^\top, \boldsymbol{\Delta}_\phi^\top)^\top$, com $\boldsymbol{\Delta}_\beta = (\boldsymbol{\Delta}_{1\beta}^\top, \dots, \boldsymbol{\Delta}_{r\beta}^\top)^\top$ com dimensão $p \times nr$, $\boldsymbol{\Delta}_\phi = (\boldsymbol{\Delta}_{1\phi}^\top, \dots, \boldsymbol{\Delta}_{r\phi}^\top)^\top$ com dimensão $3 \times nr$. Tem-se que $\boldsymbol{\Delta}_{i\beta}$ e $\boldsymbol{\Delta}_{i\phi}$, são definidos por:

$$\begin{aligned}
\boldsymbol{\Delta}_{i\beta} &= \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta} \partial \omega_i^\top}, \\
&= \frac{\partial \dot{Q}_{1i\omega}}{\partial \boldsymbol{\beta}}, \\
&= \frac{\partial}{\partial \boldsymbol{\beta}} \left(-w_i^k \boldsymbol{\varepsilon}_{i\omega}^\top \boldsymbol{\Sigma}^{-1/2} \right), \\
&= -w_i^k \left(\frac{\partial \boldsymbol{\varepsilon}_{i\omega}^\top}{\partial \boldsymbol{\beta}} \right) \boldsymbol{\Sigma}^{-1/2}, \\
&= -w_i^k \left[\frac{\partial (-\mathbf{X}_i \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right]^\top \boldsymbol{\Sigma}^{-1/2}, \\
&= w_i^k \mathbf{X}_i^\top \boldsymbol{\Sigma}^{-1/2}.
\end{aligned}$$

Com $\Delta_{i\beta}$ tem dimensão $p \times n$. Para se calcular $\Delta_{i\phi_j}$, tem-se:

$$\begin{aligned}
\Delta_{i\phi_j} &= \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \phi_j \partial \omega_i^\top}, \\
&= \frac{\partial \dot{Q}_{1i\omega}}{\partial \phi_j}, \\
&= \frac{\partial}{\partial \phi_j} \left(-w_i^k \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1/2} \right), \\
&= \frac{\partial}{\partial \phi_j} \left[-w_i^k \left(\boldsymbol{\varepsilon}_i + \boldsymbol{\Sigma}^{1/2} \boldsymbol{\omega}_i \right)^\top \boldsymbol{\Sigma}^{-1/2} \right], \\
&= \frac{\partial}{\partial \phi_j} \left[-w_i^k \left(\boldsymbol{\varepsilon}_i^\top + \boldsymbol{\omega}_i^\top \boldsymbol{\Sigma}^{1/2} \right) \boldsymbol{\Sigma}^{-1/2} \right], \\
&= -w_i^k \left[\frac{\partial}{\partial \phi_j} \left(\boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1/2} + \boldsymbol{\omega}_i^\top \boldsymbol{\Sigma}^{1/2} \boldsymbol{\Sigma}^{-1/2} \right) \right], \\
&= -w_i^k \left[\frac{\partial}{\partial \phi_j} \left(\boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1/2} + \boldsymbol{\omega}_i^\top \mathbf{I} \right) \right], \\
&= -w_i^k \boldsymbol{\varepsilon}_i^\top \left[\frac{\partial \boldsymbol{\Sigma}^{-1/2}}{\partial \phi_j} \right], \\
&= -w_i^k \boldsymbol{\varepsilon}_i^\top \left[-\boldsymbol{\Sigma}^{-1/2} \frac{\partial \boldsymbol{\Sigma}^{1/2}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1/2} \right], \\
&= w_i^k \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1/2} \frac{\partial \boldsymbol{\Sigma}^{1/2}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1/2},
\end{aligned}$$

Portanto,

$$\Delta_{i\phi_j} = w_i^k \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1/2} \frac{\partial \boldsymbol{\Sigma}^{1/2}}{\partial \phi_j} \boldsymbol{\Sigma}^{-1/2}, \text{ em que } j = 1, 2 \text{ e } 3.$$

Em que

$$\Delta_{i\phi_1} = w_i^k \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1/2} \frac{\partial \boldsymbol{\Sigma}^{1/2}}{\partial \phi_1} \boldsymbol{\Sigma}^{-1/2} = w_i^k \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1/2} \mathbf{P} (\mathbf{J}_1 \odot \mathbf{Q}) \mathbf{P}^\top \boldsymbol{\Sigma}^{-1/2}$$

$$\Delta_{i\phi_2} = w_i^k \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1/2} \frac{\partial \boldsymbol{\Sigma}^{1/2}}{\partial \phi_2} \boldsymbol{\Sigma}^{-1/2} = w_i^k \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1/2} \mathbf{P} (\mathbf{J}_2 \odot \mathbf{Q}) \mathbf{P}^\top \boldsymbol{\Sigma}^{-1/2}$$

$$\Delta_{i\phi_3} = w_i^k \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1/2} \frac{\partial \boldsymbol{\Sigma}^{1/2}}{\partial \phi_3} \boldsymbol{\Sigma}^{-1/2} = w_i^k \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Sigma}^{-1/2} \mathbf{P} (\mathbf{J}_3 \odot \mathbf{Q}) \mathbf{P}^\top \boldsymbol{\Sigma}^{-1/2}$$

com $\boldsymbol{\varepsilon}_i = (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta})$ e $\Delta_{i\phi_j} = (\Delta_{i\phi_1}, \Delta_{i\phi_2}, \Delta_{i\phi_3})^\top$ com dimensão $3 \times n$.

4.38.4 Derivada de $\boldsymbol{\Sigma}^{1/2}$

Correspondendo a qualquer matriz $\boldsymbol{\Sigma}$, $n \times n$ simétrica e não-negativa definida, existe uma matriz simétrica não-negativa definida como $\boldsymbol{\Sigma}^{1/2} = \mathbf{W}$, tal que $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{1/2} \boldsymbol{\Sigma}^{1/2} = \mathbf{W}^2$.

Além disso, W é único e pode ser expresso por

$$W = PA^{1/2}P^\top,$$

com $A^{1/2} = \text{diag}(\sqrt{\alpha_1}, \dots, \sqrt{\alpha_n})$, com $\alpha_1, \dots, \alpha_n$ os autovalores de Σ e P é uma matriz ortogonal $n \times n$ ($PP^\top = I_n$) tal que $P\Sigma P^\top = A$, com $A = \text{diag}(\alpha_1, \dots, \alpha_n)$. Então, derivado Σ em relação à ϕ_j , tem-se que

$$\frac{\partial \Sigma}{\partial \phi_j} = W \frac{\partial W}{\partial \phi_j} + \frac{\partial W}{\partial \phi_j} W, \text{ para } j = 1, 2, 3. \quad (4.95)$$

A equação (4.95) pode ser escrita como $\dot{\Sigma}_j = W \dot{W}_j + \dot{W}_j W$, em que $\dot{\Sigma}_j = \frac{\partial \Sigma}{\partial \phi_j}$ e $\dot{W}_j = \frac{\partial W}{\partial \phi_j}$, a qual tem sido extensivamente estudada na literatura por Jameson (1968).

Note que $\dot{\Sigma}_j$, W e \dot{W}_j são matrizes simétricas. Seja $J_j = P^\top \dot{\Sigma}_j P$ e $Q = [(q_{rs})]$ matrizes simétricas $n \times n$, com $q_{rs} = (\sqrt{\alpha_r} + \sqrt{\alpha_s})^{-1}$, para $r = s = 1, \dots, n$. Então, a solução para a equação (4.95) é dada por

$$\frac{\partial W}{\partial \phi_j} = \frac{\partial \Sigma^{1/2}}{\partial \phi_j} = P(J_j \odot Q)P^\top,$$

em que \odot denota o produto de Hadamard para $j = 1, 2, 3$ (DE BASTIANI, 2017).

4.38.5 Derivada da matriz de covariância

Sabemos que $\Sigma = \phi_1 I + \phi_2 R(\phi_3)$, derivando Σ em relação à ϕ_j em que $j = 1, 2$ e 3 . Obtém-se:

- Para $j = 1$

$$\frac{\partial \Sigma}{\partial \phi_1} = \frac{\partial}{\partial \phi_1} [\phi_1 I + \phi_2 R(\phi_3)] = I,$$

- Para $j = 2$

$$\frac{\partial \Sigma}{\partial \phi_2} = \frac{\partial}{\partial \phi_2} [\phi_1 I + \phi_2 R(\phi_3)] = R(\phi_3) = \frac{1}{\phi_2} (\Sigma - \phi_1 I),$$

- Para $j = 3$

$$\frac{\partial \Sigma}{\partial \phi_3} = \frac{\partial}{\partial \phi_3} [\phi_1 I + \phi_2 R(\phi_3)] = \phi_2 \frac{\partial R(\phi_3)}{\partial \phi_3} = \phi_2 R', \quad (4.96)$$

Verifica-se que para $j = 3$, R' depende de cada modelo: exponencial, gaussiano e família Matérn. Porém, pode-se obter R' utilizando somente o modelo da família Matérn, pois quando $k = 0,5$ corresponde ao modelo exponencial e quando $k \rightarrow \infty$ refere-se ao modelo gaussiano. Assim, para se obter $R' = R'_m(\phi_3)$, utiliza-se a função de correlação espacial para o modelo da família Matérn dado por:

$$R_m(\phi_3) = \left[\left(\frac{1}{2^{k-1} \Gamma(k)} \right) \left(\frac{h}{\phi_3} \right)^k K_k \left(\frac{h}{\phi_3} \right) \right]. \quad (4.97)$$

Aplica-se a derivada em relação à ϕ_3 na equação (4.97). Obtém-se:

$$\begin{aligned}
 R'_m(\phi_3) &= \frac{\partial}{\partial \phi_3} \left[\left(\frac{1}{2^{k-1}\Gamma(k)} \right) \left(\frac{h}{\phi_3} \right)^k K_k \left(\frac{h}{\phi_3} \right) \right] \\
 &= \left(\frac{1}{2^{k-1}\Gamma(k)} \right) \frac{\partial}{\partial \phi_3} \left[\left(\frac{h}{\phi_3} \right)^k K_k \left(\frac{h}{\phi_3} \right) \right] \\
 &= \left(\frac{1}{2^{k-1}\Gamma(k)} \right) \left\{ \left[\frac{\partial}{\partial \phi_3} \left(\frac{h^k}{\phi_3^k} \right) \right] K_k \left(\frac{h}{\phi_3} \right) + \left(\frac{h}{\phi_3} \right)^k \left[\frac{\partial}{\partial \phi_3} \left(K_k \left(\frac{h}{\phi_3} \right) \right) \right] \right\} \\
 &= -\frac{1}{\phi_3} \left[kR_m(\phi_3) + \frac{1}{2^{k-1}\Gamma(k)} \left(\frac{h}{\phi_3} \right)^{k+1} K'_k \left(\frac{h}{\phi_3} \right) \right] \\
 &= -\frac{1}{\phi_3} m.
 \end{aligned}$$

Portanto, a equação (4.96) é representada por:

$$\frac{\partial \Sigma}{\partial \phi_3} = -\frac{\phi_2}{\phi_3} m, \text{ em que } m = \left[kR_m(\phi_3) + \frac{1}{2^{k-1}\Gamma(k)} \left(\frac{h}{\phi_3} \right)^{k+1} K'_k \left(\frac{h}{\phi_3} \right) \right].$$

4.38.6 Perturbação na matriz escala

4.38.7 Representação $\ddot{Q}(\hat{\theta})$ para os esquemas de perturbação

A representação de $\ddot{Q}(\hat{\theta})$ para os esquemas de ponderação aplicados ao estudo de influência local, considera-se o resultado da matriz (M2) obtido no Apêndice 4.37.2, tem-se que

$$\ddot{Q}(\hat{\theta}) = \begin{pmatrix} \ddot{Q}_{11} & \mathbf{0} \\ \mathbf{0} & \ddot{Q}_{22} \end{pmatrix}, \quad (\text{M5})$$

com

$$\ddot{Q}_{11} = \begin{pmatrix} \ddot{Q}_{i\beta\beta} & \ddot{Q}_{i\beta\phi} \\ \ddot{Q}_{i\phi\beta} & \ddot{Q}_{i\phi\phi} \end{pmatrix} \text{ e } \ddot{Q}_{22} = \left(\ddot{Q}_{i\eta\beta} \right). \quad (\text{M6})$$

4.39 Referências

ACOSTA, J., OSORIO, F., VALLEJOS, R. Effective sample size for line transect sampling models with an application to marine macroalgae. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 21, p. 407- 425, 2016.

ANDERSON, T. **An Introduction to Multivariate Statistical Analysis**. 3. ed. New York: Wiley, 2003.

ASSUMPCÃO, R. A. B.; URIBE-OPAZO, M. A.; GALEA, M. Analysis of local influence in geostatistics using Student's t-distribution. **Journal of Applied Statistics**, v. 41, n. 11, p. 2323 - 2341, 2014.

BORSSOI, J. A. **Modelos mistos lineares elípticos com erros de medição**. Tese (Doutorado), Programa de Pós-Graduação em Estatística, Universidade de São Paulo, 2013.

CASTRO, M. D., GALEA, M., BOLFARINE, H. Hypothesis testing in an errors-in-variables model with heteroscedastic measurement errors. **Statistics in Medicine**, Wiley Online Library, v. 27, n. 25, p. 5217-5234, 2008.

COOK, R. D. Detection of influential observation in linear regression. **Technometrics**, [Taylor & Francis, Ltd., American Statistical Association, American Society for Quality], v. 19, n. 1, p. 15-18, 1977. ISSN 00401706. Disponível em: <<http://www.jstor.org/stable/1268249>>.

COOK, R. D. Assessment of local influence. **Journal of the Royal Statistical Society. Series B (Methodological)**, [Royal Statistical Society, Wiley], v. 48, n. 2, p. 133-169, 1986. ISSN 00359246. Disponível em: <<http://www.jstor.org/stable/2345711>>.

COOK, R. D.; WEISBERG, S. **Residuals and influence in regression**. [S.l.]: New York: Chapman and Hall, 1982.

COX, D. R., REID, N. Parameter orthogonality and approximate conditional inference. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 1-39, 1987.

CHATTERJEE, S., HADI, A. S. **Sensitivity analysis in linear regression**. [S.l.]: John Wiley & Sons, v. 327, 2009.

CHEN, F., ZHU, H.-T., LEE, S.-Y. Perturbation selection and local influence analysis for nonlinear structural equation model. **Psychometrika**, Springer, v. 74, n. 3, p. 493-516, 2009.

DAGENAIS, M. G., DUFOUR, J. M. Invariance, nonlinear models, and asymptotic tests. **Econometrica: Journal of the Econometric Society**, p. 1601-1615, 1991.

DE BASTIANI, F., URIBE-OPAZO, M., GALEA, M., CYSNEIROS, A. Case-deletion diagnostics for spatial linear mixed models. **Spatial statistics**, Elsevier, v. 28, p. 284-303, 2018.

DE BASTIANI, F., GALEA, M., CYSNEIROS, A. H. M. A., URIBE-OPAZO, M. A. Gaussian spatial linear models with repetitions: An application to soybean productivity. **Spatial Statistics**, v. 21, p. 319-335, 2017.

DE BASTIANI, F. **Inference and Diagnostics in spatial models**. Tese (Doutorado) Universidade Federal de Pernambuco, <https://repositorio.ufpe.br/handle/123456789/17304>, 2016.

DE BASTIANI, F., DE AQUINO CYSNEIROS, A. H. M., URIBE-OPAZO, M. A., GALEA, M. Influence diagnostics in elliptical spatial linear models. **Test**, v. 24, n. 2, p. 322-340, 2015.

DOS PASSOS, A. M. A., REZENDE, P. M.D., CARVALHO, E. R., ÁVILA, F. W. D. Biochar, farmyard manure and poultry litter on chemical attributes of a dystrophic cambissol and soybean crop. **Revista Brasileira de Ciências Agrárias**, Universidade Federal Rural de Pernambuco, v. 10, n. 3, p. 382-388, 2015.

FAGUNDES, R., URIBE-OPAZO, M., GALEA, M., GUEDES, L. Spatial variability in slash linear modeling with finite second moment. **Journal of Agricultural, Biological and Environmental Statistics**, Springer, v. 23, n. 2, p. 276-296, 2018a.

FAGUNDES, R. S., URIBE-OPAZO, M. A., GUEDES, L. P. C., GALEA, M. Slash spatial linear modeling: soybean yield variability as a function of soil chemical properties. **Revista Brasileira de Ciência do Solo**, v. 42, 2018b.

FAGUNDES, R. S. **Distribuição Slash Multivariada aplicada a dados agrícolas**. Programa de Pós-Graduação em Engenharia Agrícola, Tese (Doutorado)- UNIOESTE, <http://tede.unioeste.br/handle/tede/3087>, 2017.

FIORENTINI, G., SENTANA, E., CALZOLARI, G. Maximum likelihood estimation and inference in multivariate conditionally heteroscedastic dynamic regression models with Student t innovations. **Journal of Business & Economic Statistics**, v. 21, n. 4, p. 532-546, 2003.

GALEA, M., CADEMARTORI, D., CURCI, R., MOLINA, A. Robust Inference in the Capital Asset Pricing Model Using the Multivariate t-Distribution. **Journal of Risk and Financial Management**, v. 13, n. 6, p. 123, 2020.

GALEA, M., DE CASTRO, M. Robust inference in a linear functional model with replications using the t distribution. **Journal of Multivariate Analysis**, v. 160, p. 134-145, 2017.

GALEA, M., BOLFARINE, H., VILCA, F. Local influence in comparative calibration models under elliptical t-distributions. **Biometrical Journal: Journal of Mathematical Methods in Biosciences**, Wiley Online Library, v. 47, n. 5, p. 691-706, 2005.

GALEA, M., PAULA, G. A., BOLFARINE, H. Local influence in elliptical linear regression models. **Journal of the Royal Statistical Society: Series D (The Statistician)**, Wiley Online Library, v. 46, n. 1, p. 71-79, 1997.

GARCIA-PAPANI, F., LEIVA, V., URIBE-OPAZO, M. A., AYKROYD, R. G. Birnbaum-Saunders spatial regression models: Diagnostics and application to chemical data. **Chemometrics and Intelligent Laboratory Systems**, v. 177, p. 114-128, 2018.

GARCIA-PAPANI, F., URIBE-OPAZO, M. A., LEIVA, V., AYKROYD, R. G. BirnbaumSaunders spatial modelling and diagnostics applied to agricultural engineering data. **Stochastic Environmental Research and Risk Assessment**, v. 31, n. 1, p. 105-124, 2017.

GOODAL, C., MARDIA, K.V., Challenges in multivariate spatio-temporal modeling. In: **XVIIth International Biometric Conference**. Ontario, Canada, August, pp. 117, 1994.

JAMESON, A. Solution of the equation $ax + xb = c$ by inversion of an $m \times m$ or $n \times n$ matrix. **SIAM J Appl Math**, 1968.

KANO, Y., BERKANE, M., BENTLER, P. M. Statistical inference based on pseudo-maximum likelihood estimators in elliptical populations. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 88, n. 421, p. 135-143, 1993.

KBAKURAL, B., ROBERT, P., HUGGINS, D. Variability of corn/soybean yield and soil/landscape properties across a southwestern Minnesota landscape. In: WILEY ONLINE LIBRARY. **Proceedings of the fourth international conference on precision agriculture**. [S.l.], p.573-579, 1999.

- KOTZ, S., NADARAJAH, S. **Multivariate t-distributions and their applications**. Cambridge University Press, 2004.
- LANGE, K. L., LITTLE, R. J., TAYLOR, J. M. Robust statistical modeling using the t distribution. **Journal of the American Statistical Association**, v. 84, n. 408, p. 881-896, 1989.
- LEMONTE, A. J., FERRARI, S. L. The local power of the gradient test. **Annals of the Institute of Statistical Mathematics**, Springer, v. 64, n. 2, p. 373-381, 2012.
- LESAFFRE, E., VERBEKE, G. Local influence in linear mixed models. **Biometrics**, JSTOR, p. 570-582, 1998.
- LIU, S. Local influence in multivariate elliptical linear regression models. **Linear Algebra and Its Applications**, Elsevier, v. 354, n. 1-3, p. 159-174, 2002.
- MAGNUS, J. R., NEUDECKER, H. Matrix differential calculus with applications in statistics and econometrics, **Google Scholar**, p. 501-502, 1999.
- MARDIA, K., KENT, J., BIBBY, J. Multivariate analysis. **Academic Press**, London, 1992.
- MARDIA, K. V.; MARSHALL, R. J. Maximum likelihood estimation of models for residual covariance in spatial regression. **Biometrika**, [Oxford University Press, Biometrika Trust], v. 71, n. 1, p. 135-146, 1984. ISSN 00063444. Disponível em: <<http://www.jstor.org/stable/2336405>>.
- MCLACHLAN, G. J., KRISHNAN, T. **The EM algorithm and extensions**. John Wiley & Sons, 2007.
- MELO, T. F., FERRARI, S. L., CRIBARI-NETO, F. Improved testing inference in mixed linear models. **Computational Statistics & Data Analysis**, Elsevier, v. 53, n. 7, p. 2573-2582, 2009.
- OSORIO, F., GALEA, M. Statistical inference in multivariate analysis using the t-distribution. **Unpublished manuscript**, 2015.
- OSORIO, F., PAULA, G. A., GALEA, M. On estimation and influence diagnostics for the grubb's model under heavy-tailed distributions. **Computational Statistics & Data Analysis**, Elsevier, v. 53, n. 4, p. 1249-1263, 2009.
- OSORIO, F., PAULA, G. A., GALEA, M. Assessment of local influence in elliptical linear models with longitudinal structure. **Computational Statistics & Data Analysis**, v. 51, n. 9, p. 4354-4368, 2007. ISSN 0167-9473. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167947306001940>>.
- OSORIO, F. **Diagnóstico de influência em modelos elípticos de efeitos mistos**. Tese (Doutorado) Programa de Pós-Graduação em Estatística, Universidade de São Paulo, Brasil, 2006
- PAN, J., FEI, Y., FOSTER, P. Case-deletion diagnostics for linear mixed models. **Technometrics**, Taylor & Francis, v. 56, n. 3, p. 269-281, 2014. Disponível em: <<https://doi.org/10.1080/00401706.2013.810173>>.

PAULA, G. **Modelos de regressão com apoio computacional**. São Paulo-SP: Instituto de Matemática e Estatística (IME), Universidad de São Paulo, 2010.

PINHEIRO, J. C., LIU, C., WU, Y. N. Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution. **Journal of Computational and Graphical Statistics**, v. 10, n. 2, p. 249-276, 2001.

PRADO, N. V. D.; URIBE-OPAZO, M. A.; GALEA, M.; ASSUMPÇÃO, R. A. Influência local em um modelo espacial linear da produtividade da soja utilizando distribuição t-student. **Engenharia Agrícola**, SciELO Brasil, v. 33, n. 5, p. 1003-1016, 2013.

RAO, C. R. Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. In: CAMBRIDGE UNIVERSITY PRESS. **Mathematical Proceedings of the Cambridge Philosophical Society**. [S.l.], v. 44, n. 1, p. 50-57, 1948.

REICHERT, J. M., DARIVA, T. A., REINERT, D. J., SILVA, V. R. D. Variabilidade espacial de planos solo e produtividade de soja em várzea sistematizada: análise geoestatística e análise de regressão. **Ciência Rural**, SciELO Brasil, v. 38, n. 4, p. 981-988, 2008.

SOLOS, Embrapa. Sistema brasileiro de classificação de solos. **Centro Nacional de Pesquisa de Solos**: Rio de Janeiro, 2013.

SCHEMMER, R. C., URIBE-OPAZO, M. A., GALEA, M., ASSUMPÇÃO, R. A. Spatial variability of soybean yield through a reparameterized t-Student model. **Engenharia Agrícola**, v. 37, n. 4, p. 760-770, 2017.

SUTRADHAR, B. C. Score test for the covariance matrix of the elliptic t-distribution. **Journal of Multivariate Analysis**, v. 46, n. 1, p. 1-12, 1993.

SMITH, R. L. Environmental statistics. **Notes at the conference board of the mathematical sciences (cbms) course at university of washington**, 2001.

STEIN, M. L. **Interpolation of Spatial Data: Some Theory for Kriging**. Springer, New York, 1999.

TSAI, K.-T.; KOZIOL, J. A. Score and wald tests for the multivariate growth curve model with missing data. **Annals of the Institute of Statistical Mathematics**, Springer, v. 40, n. 1, p. 179-186, 1988.

TERRELL, G. The gradient statistic. **Computing Science and Statistics**, v. 34, p. 206-215, 2002.

URIBE-OPAZO, M. A., BORSSOI, J. A., GALEA, M. Influence diagnostics in Gaussian spatial linear models. **Journal of Applied Statistics**, v. 39, n. 3, p. 615-630, 2012.

URIBE-OPAZO, M. A., DE BASTIANI, F., GALEA, M., SCHEMMER, R. C., ASSUMPÇÃO, R. A. B. Influence diagnostics on a reparameterized t-Student spatial linear model. **Spatial Statistics**, v. 41, p. 100481, 2021.

WALD, A. Tests of statistical hypotheses concerning several parameters when the number of

observations is large. **Transactions of the American Mathematical society**, JSTOR, v. 54, n. 3, p. 426-482, 1943.

WILKS, S. S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. **The Annals of Mathematical Statistics**, JSTOR, v. 9, n. 1, p. 60-62, 1938.

ZELLER, C. B., VILCA, F., GALEA, M. Influence diagnostics for the structural sharpe model under normal/independent distributions. **Communications in Statistics-Theory and Methods**, Taylor & Francis, p. 1-21, 2018.

ZELLNER, A. Bayesian and non-bayesian analysis of the regression model with multivariate student-t error terms. **Journal of the American Statistical Association**, Taylor Francis Group, v. 71, n. 354, p. 400-405, 1976.

ZHU, H.-T., LEE, S.-Y. Local influence for incomplete data models. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 63, n. 1, p.111-126, 2001.

ZUCKER, D. M., LIEBERMAN, O., MANOR, O. Improved small sample inference in the mixed linear model: Bartlett correction and adjusted likelihood. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 62, n. 4, p. 827-838, 2000.

5 CONSIDERAÇÕES FINAIS

Os resultados desta tese foram apresentados em formato de artigos, com conclusões individualizadas para cada um deles. Assim, apresentam-se aqui as principais contribuições, verificadas durante o desenvolvimento deste trabalho:

- A construção dos algoritmos sendo o algoritmo Iterativo (Artigo 1 e Artigo 2), e o algoritmo Iterativo parcial linear (Artigo 3), que podem ser estendidos para outros modelos das classes das distribuições simétricas.
- A proposta de extensão do trabalho de Assumpção et al. (2014) para a distribuição t-Student reparametrizada, em que foram apresentados os diagnósticos de influência global, influência local com perturbação de casos e na variável resposta foram executados, e contribuíram para a publicação do Artigo 1, apresentado nesta tese.
- A proposta da extensão do trabalho de De Bastiani et al. (2017), para o modelo de regressão espacial linear t-Student reparametrizado com repetição, foi realizada e obteve sucesso no desenvolvimento do algoritmo iterativo parcial linear, que tornou-se mais rápido, com a obtenção de inúmeras iterações de sucesso. Esta pesquisa também, possibilitou o desenvolvimento da validação cruzada, traço, teste de razão de verossimilhança, diagnóstico de influência global, influência local com esquema de perturbação sendo ponderação de casos, na variável resposta e na matriz escala. A partir deste estudo, abre-se caminho para o desenvolvimento de modelos com repetição em distribuição simétricas.