



Universidade Federal de Pernambuco
Centro de Ciências Exatas e da Natureza
Pós-graduação em Estatística

Rodrigo Gonçalves dos Santos

Refinamento para testes de hipóteses em modelos espaciais lineares Gaussianos com repetições

Recife
2020

Rodrigo Gonçalves dos Santos

Refinamento para testes de hipóteses em modelos espaciais lineares Gaussianos com repetições

Tese de Doutorado apresentada ao Programa de Pós-graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Doutor em Estatística.

Área de concentração: Estatística Aplicada

Orientadora: Profa. Dra. Audrey Helen Mariz de Aquino Cysneiros.

Coorientadora: Profa. Dra. Fernanda De Bastiani

Recife
2020

Catálogo na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

S237r Santos, Rodrigo Gonçalves dos
Refinamento para testes de hipóteses em modelos espaciais lineares
Gaussianos com repetições / Rodrigo Gonçalves dos Santos. – 2020.
55 f.: il., fig., tab.

Orientadora: Audrey Helen Mariz de Aquino Cysneiros.
Tese (Doutorado) – Universidade Federal de Pernambuco. CCEN,
Estatística, Recife, 2020.
Inclui referências e anexos.

1. Estatística aplicada. 2. Correção de Bartlett. I. Cysneiros, Audrey Helen
Mariz de Aquino (orientadora). II. Título.

310 CDD (23. ed.) UFPE- CCEN 2020 - 109

Rodrigo Gonçalves dos Santos

“Refinamento para testes de hipóteses em modelos espaciais lineares Gaussianos com repetições”

Aprovado em: 21/02/2020.

Orientadora: Profa. Dra. Audrey Helen Mariz de Aquino Cysneiros.

BANCA EXAMINADORA

Profa. Audrey Helen Mariz de Aquino Cysneiros
Centro de Ciências Exatas e da Natureza / UFPE

Prof. Aldo William Medina Garay
Centro de Ciências Exatas e da Natureza / UFPE

Prof. Vinícius Quintas Souto Maior
Centro de Ciências Exatas e da Natureza / UFPE

Prof. Miguel Angel Uribe Opazo
UNIOESTE-Paraná

Profa. Mariana Correia de Araújo
Universidade Federal do Rio Grande do Norte / UFRN

Aos meus Pais.

AGRADECIMENTOS

Inicialmente, agradeço à Deus pela saúde, paz e sabedoria que tem me dado para superar as dificuldades da vida.

Aos meus pais, **José e Maria**, pela boa criação e meu irmão **Rodolfo** que compartilhou diversos desses momentos comigo.

À minha esposa **Graça** pelo apoio incondicional desde sempre.

Aos colegas de Pós-Graduação pelos momentos de estudo em grupo e também pelos momentos de descontração. Em especial, agradeço ao casal **Wanessa e Márcio** pela ajuda oferecida no momento crucial deste curso.

Às minhas orientadora **Audrey Helen Mariz de Aquino Cysneiros** e coorientadora **Fernanda De Bastiani** pela oportunidade de trabalhar com um tema que gosto bastante. Agradeço todo suporte e compreensão. Em especial, agradeço à **Audrey** pela confiança e força repassadas.

À todos os professores do **Departamento de Estatística do CCEN** dos quais tive a oportunidade de conhecer e aprender um pouco mais.

RESUMO

Nesta tese, tratamos de refinamentos para testes de hipóteses em modelos espaciais lineares gaussianos com repetições. Nós obtivemos um fator de correção de Bartlett para a estatística da razão de verossimilhanças baseado na verossimilhança perfilada modificada proposta por Cox e Reid (1987). Além disso, desenvolvemos novos ajustes para a estatística da razão de verossimilhanças com base nos trabalhos de Skovgaard (SKOVGAARD et al., 1996; SKOVGAARD, 2001). Um outro teste de hipóteses em modelos lineares gaussianos com repetições foi apresentado, com o intuito de averiguar se a parte espacial do modelo é relevante, utilizando os desenvolvimentos dos ajustes para a estatística da razão de verossimilhanças com base nos trabalhos de Skovgaard. Estudos de simulação de Monte Carlo foram considerados para avaliar e comparar numericamente o desempenho dos testes propostos nesta tese, assim como o teste da razão de verossimilhanças tradicional e sua versão original corrigida por Bartlett (DE BASTIANI, 2016), em amostras finitas. Ademais, um exemplo utilizando dados reais foi ilustrado com o objetivo de utilizar todas as ferramentas apresentadas na tese.

Palavras-chaves: Correção de Bartlett. Geoestatística. Medidas repetidas. Teste da Razão de Verossimilhanças. Verossimilhança perfilada.

ABSTRACT

In this thesis, we deal with refinements for testing hypotheses in linear Gaussian spatial models with repetitions. We obtained a Bartlett correction factor for the likelihood ratio statistics based on the modified profiled likelihood proposed by Cox e Reid (1987). Also, we developed new adjustments to the likelihood ratio statistics based on the work of Skovgaard (SKOVGAARD et al., 1996; SKOVGAARD, 2001). Another test of hypotheses in linear Gaussian models with repetitions was presented, to verify if the spatial part of the model is relevant, using the developments of the adjustments for the likelihood ratio statistics based on the works of Skovgaard. Monte Carlo simulation studies were considered to numerically evaluate and compare the performance of the tests proposed in this thesis, as well as the traditional likelihood ratio test and its original version corrected by Bartlett (DE BASTIANI, 2016), in finite samples. Besides, an example using real data was illustrated in order to use all the tools presented in the thesis.

Keywords: Bartlett correction. Geostatistical. Likelihood ratio test. Profile likelihood. Repeated measures.

LISTA DE FIGURAS

- Figura 1 – (a), (b), (c), (d), (e) são histogramas dos dados de produtividade de soja para cada uma das cinco repetições (1998, 1999, 2000, 2001 e 2002, respectivamente) e (f) é um histograma de $\mathbf{Y}$ 14
- Figura 2 – Boxplot de produtividade de soja nos anos de 1998 ($Y_{(1)}$), 1999 ($Y_{(2)}$), 2000 ($Y_{(3)}$), 2001 ($Y_{(4)}$) e 2002 ($Y_{(5)}$). 15

LISTA DE TABELAS

Tabela 1	–	Análise descritiva da variável produtividade de soja.	13
Tabela 2	–	Testes de hipóteses usando diversas estatísticas baseadas na razão de verossimilhanças: usual (LR), baseada na verossimilhança perfilada ajustada (LR_{CR}), suas respectivas correções de Bartlett (LR^* e LR_{CR}^* e as versões ajustadas de Skovgaard (LR_{SK1} e LR_{SK2}), com p-valor (entre parêntesis). . .	32
Tabela 3	–	Tamanhos dos testes para tamanhos de amostras $n = 36$ e $n = 49$, nº de repetições $r = 5$ e nº de β 's em perturbação $q - p = 2$, para diferentes níveis nominais usuais e parâmetros de interesse $p = \{2, 3, 4, 5, 6\}$	35
Tabela 4	–	Tamanhos dos testes para tamanhos de amostras $n = 36$ e $n = 49$, nº de repetições $r = 5$ e nº de β 's de interesse $p = 2$, para diferentes níveis nominais usuais e parâmetros de perturbação $q - p = \{2, 3, 4\}$	36
Tabela 5	–	Média e variância para tamanhos de amostras $n = \{25, 36, 49\}$, nº de repetições $r = 5$, nº de β 's em perturbação $q - p = 2$, para diferentes níveis nominais usuais e parâmetros de interesse $p = \{2, 3, 4, 5, 6\}$	37
Tabela 6	–	Tamanhos dos testes para tamanho de amostra $n = 36$, nº de repetições $r = \{5, 7, 9\}$ e nº de β 's em perturbação $q - p = 2$, para diferentes níveis nominais usuais e parâmetros de interesse $p = \{3, 4\}$	37
Tabela 7	–	Tamanhos dos testes para tamanhos de amostras $n = 25$, nº de repetições $r = \{5, 7, 9, 11, 13, 20\}$ e nº de β 's de interesse $p = 4$, para diferentes níveis nominais usuais e parâmetros de perturbação $q - p = 2$	38
Tabela 8	–	Poder dos testes para tamanho de amostra $n = 25$, nº de repetição $r = 7$ e nº de β 's de interesse $p = 2$ e de perturbação $q - p = 2$, para diferentes níveis nominais usuais e diferentes valores de δ	39

SUMÁRIO

1	INTRODUÇÃO	11
2	DEFINIÇÃO DO MODELO MLEGR E ASPECTOS INFERENCIAIS	17
2.1	DEFINIÇÃO DO MODELO MLEGR	17
2.2	ASPECTOS INFERENCIAIS	19
2.3	ESTIMAÇÃO DO PARÂMETRO ϕ	23
2.4	TESTE DA RAZÃO DE VEROSSIMILHANÇAS NO MLEGR	23
3	CORREÇÃO DE BARTLETT	25
3.1	OBTENÇÃO DO FATOR DE CORREÇÃO DE BARTLETT NO MLEGR	26
3.2	APLICAÇÃO	30
4	RESULTADOS NUMÉRICOS	33
5	CONSIDERAÇÕES FINAIS	40
	REFERÊNCIAS	42
	ANEXO A – CÁLCULOS PARA OBTENÇÃO DE C	46
	ANEXO B – CÁLCULOS PARA OBTENÇÃO DE C_{CR}	51
	ANEXO C – CÁLCULOS PARA OBTENÇÃO DE ρ	53

1 INTRODUÇÃO

Modelos baseados em dados georreferenciados têm sido investigados por vários autores, não se limitando apenas à agronomia, geologia, meteorologia, climatologia, entre outros, mas também tem sido levado em consideração em outras áreas como a econometria espacial e epidemiologia, por exemplo.

Dois agrônomos, Youden e Mehlich (1938) viram na análise de variância uma ferramenta útil para mensurar a escala de variação espacial. Entretanto, esta mesma técnica foi redescoberta por outros autores, como por exemplo, Krumbein e Slack (1956) em geologia e Webster e Butler (1976) em ciência do solo. Matérn (1960) reconheceu as consequências da correlação espacial e derivou, teoricamente, seus efeitos nas estimativas globais. O engenheiro D. G. Krige percebeu que suas estimativas de graus de minérios em um bloco eram melhoradas se fossem levados em conta os graus de minérios em blocos vizinhos (KRIGE, 1951). O mesmo conceito de autocorrelação surgiu na França, praticamente ao mesmo tempo, sendo o matemático G. Matheron (MATHERON, 1963), o responsável por solucionar problemas de estimação da teoria de processos aleatórios.

Em sismologia, Veneziano e Dyck (1987) tentaram prever a localização e magnitudes de terremotos, através da estimação da taxas de terremotos de várias magnitudes ao redor de regiões de interesse. Kielland e Tubman (1994) discutiram as implicações da utilização do sistema de navegação geográfica na incerteza espacial tanto posicionamento da embarcação com relação ao mapa quanto no mapa de profundidade de superfícies, utilizado para planejar a rota de navegação.

Nesta tese, estudamos o efeito do erro de localização em estatística espacial, mais especificamente em um subcampo, a Geoestatística, que estuda a relação da variável resposta com variáveis explicativas levando em consideração a localização geográfica de cada ponto observado. A localização geográfica é fixa. Importantes trabalhos realizados estão em Krige (1951) e Matheron (1963) que são os precursores de ajustes de modelos levando em consideração essa localização. Mardia e Marshall (1984) descreveram o método de máxima verossimilhança em modelos lineares Gaussianos na situação específica de resíduos correlacionados.

Atkinson (1997), que derivou algumas propriedades de segundo momento em modelos Geoestatísticos. Chiles, Delfiner et al. (1999) derivaram a função de variância entre duas observações quaisquer cujas localizações sofreram alguma perturbação, sendo esse processo espacial com média constante. Zhu, Eickhoff e Yan (2005) propuseram inferência em uma classe de modelos espaço-temporais para dados multivariados através do algoritmo Monte Carlo EM.

Uribe-Opazo, Borssoi e Galea (2012) estudaram a influência de valores atípicos em modelos lineares espaciais Gaussianos. Nesse trabalho foram usadas técnicas de diagnóstico para avaliar a sensibilidade dos estimadores de máxima verossimilhança, funções de covariância e no preditor linear para pequenas perturbações ou mesmo nas hipóteses de um modelo espacial linear.

DE BASTIANI et al. (2017) analisaram o rendimento da soja usando o modelo linear espacial Gaussiano com repetição (MLEGr) e linear espaço-temporal Gaussiano, com dados de produtividade de soja cinco anos, com covariáveis. Como critério de seleção de modelos foi verificada a bondade do ajuste através do critério de informação de Akaike (AKAIKE, 1998) e também valores atípicos, conduzindo a uma análise de diagnóstico de influência. Considerando ainda o trabalho de DE BASTIANI et al. (2017), em um primeiro MLEGr ajustado para os dados de produtividade de soja, os autores verificaram que nenhuma das covariáveis eram significativas ao nível de significância de 5% através de testes de hipóteses formulados sobre o modelo.

A estatística de razão de verossimilhanças (LR), sob condições de regularidade e hipótese nula (H_0), converge em distribuição para qui-quadrado (χ_p^2) com erro da ordem n^{-1} em grandes amostras¹. Contudo, em amostras pequenas essas aproximações pela distribuição χ_p^2 pode não ser satisfatória, conduzindo a taxas de rejeição bem distorcidas. No entanto, faz-se necessário desenvolver técnicas de aperfeiçoamento das estatísticas de teste de tal forma que melhore a aproximação das distribuições dessas estatísticas pela distribuição χ_p^2 produzindo inferências mais precisas em amostras finitas e principalmente de tamanho pequeno ou moderado. Bartlett (1937) propôs uma versão aperfeiçoada da estatística LR, digamos LR^* , sendo estudada posteriormente por diversos autores como Lawley (1956), Cordeiro (1987), Barndorff-Nielsen e E. Hall (1988), Ferrari, Cysneiros e Cribari-Neto (2004), Cysneiros e Ferrari (2006) e Araújo, Cysneiros e Montenegro (2017), entre outros.

Na presença de parâmetros de perturbação, inferências podem ser baseadas na função de verossimilhança perfilada. Esta depende apenas do parâmetro de interesse, uma vez que é obtida da verossimilhança original em que os parâmetros de perturbação são substituídos por seus estimadores de máxima verossimilhança obtidos para cada valor fixado dos parâmetros de interesse. Vale salientar que a presença de um número considerável de parâmetros de perturbação também deteriora a qualidade das aproximações envolvidas nas inferências que se baseiam em resultados assintóticos. Vários ajustes para a função de verossimilhança perfilada e para estatísticas da razão de verossimilhanças derivadas de tal função têm sido propostos. Nesta tese enfocaremos o ajuste de Cox e Reid (1987). Para uma análise detalhada sobre os ajustes para a função de verossimilhança perfilada ver Pace e Salvan (1997) e Severini (2000).

Melo, Ferrari e Cribari-Neto (2009) obtiveram refinamentos (baseados em DiCiccio e Stern (1994)) para os testes da razão de verossimilhanças e razão de verossimilhanças perfiladas modificadas (baseada na proposta de Cox e Reid (1987)) em modelos lineares mistos. As estatísticas de razão de verossimilhanças modificadas desenvolvidas neste cenário são as que modificam o logaritmo da função de verossimilhança (log-verossimilhança), dita função de verossimilhança perfilada, proposta por Cox e Reid (1987) e sua versão corrigida por Bartlett (1937).

Duas vertentes da teoria assintótica de alta ordem serão utilizadas nesta tese, a saber: a obtenção de ajustes para verossimilhanças perfiladas (que reduzam o problema da deterioração da

¹ Se $\{a_n\}$ e $\{b_n\}$, $n \geq 1$, são duas seqüências de números reais, denotamos $a_n = o(b_n)$, ou seja, $\{a_n\}$ é de ordem menor que $\{b_n\}$, se $\lim_{n \rightarrow \infty} a_n/b_n = 0$ e denotamos por $\{a_n\} = O(b_n)$, ou seja, $\{a_n\}$ é de ordem no máximo igual a $\{b_n\}$, se existe um número real $M > 0$ tal que $|a_n/b_n| \geq M$.

qualidade das aproximações envolvidas nas inferências que se baseiam em resultados assintóticos) e a obtenção de ajustes para estatísticas de teste (que reduzam a distorção de tamanho do teste em amostras finitas através de uma aceleração da taxa de convergência do tamanho verdadeiro para o tamanho nominal (assintótico)).

DE BASTIANI (2016) obteve um fator de correção de Bartlett no MLEGr. Foram apresentados dois estudos de simulação, um sobre o modelo linear simples (uma covariável, um parâmetro de interesse a ser testado) e o outro o modelo linear múltiplo (com duas covariáveis, dois parâmetros de interesse a serem testados). Em ambas as simulações apresentadas, o número de réplicas e de observações foram fixados. Além disso, foi verificado ainda a proximidade da média e variância das estatísticas LR e LR* com as médias e variâncias teóricas da distribuição χ^2 .

Utilizaremos o conjunto de dados apresentado em DE BASTIANI (2016), como um exemplo motivacional que será reanalisado com a metodologia desenvolvida nesta tese. Os dados utilizados foram coletados numa área de $7,20 \times 7,20m$ em uma área experimental com $1,33ha$ no Centro de Pesquisa Eloy Gomes na Cooperativa Agropecuária de Desenvolvimento Tecnológico e Econômico Ltda (COODETEC), na cidade de Cascavel, Paraná - Brasil. Foram coletados dados de produtividade de soja e quatro atributos químicos do solo nos anos de 1998 ($Y_{(1)}$), 1999 ($Y_{(2)}$), 2000 ($Y_{(3)}$), 2001 ($Y_{(4)}$) e 2002 ($Y_{(5)}$). Em cada ano foram coletadas 253 observações referentes a produtividade da soja. Da mesma forma, os quatro atributos químicos do solo foram observados nos mesmos anos e são consideradas como variáveis explicativas do modelo: fósforo (P) [$mgdm^{-3}$], potássio (K) [$cmolcdm^{-3}$], cálcio (Ca) [$cmolcdm^{-3}$] e magnésio (Mg) [$cmolcdm^{-3}$]. O objetivo é entender a relação dos atributos químicos com a produtividade de soja. A Tabela 1 apresenta uma análise descritiva dos dados de produtividade de soja, divididos para cada correspondente a uma repetição no modelo espacial linear com repetições. As unidades amostrais foram observadas nos mesmos locais geográficos em cada repetição (ano). Nos anos de 1999 e 2002 a produtividade de soja destoa em alguns aspectos dos demais anos, como por exemplo, 1º e 3º quartis menores se comparados aos demais anos. Dentre os dois anos, 1999 e 2002, a média/mediana de produtividade de soja do ano de 2002 é bem menor. Além disso, o ano de 2002 apresenta maior variabilidade de produtividade de soja entre as observações dentre todas as repetições.

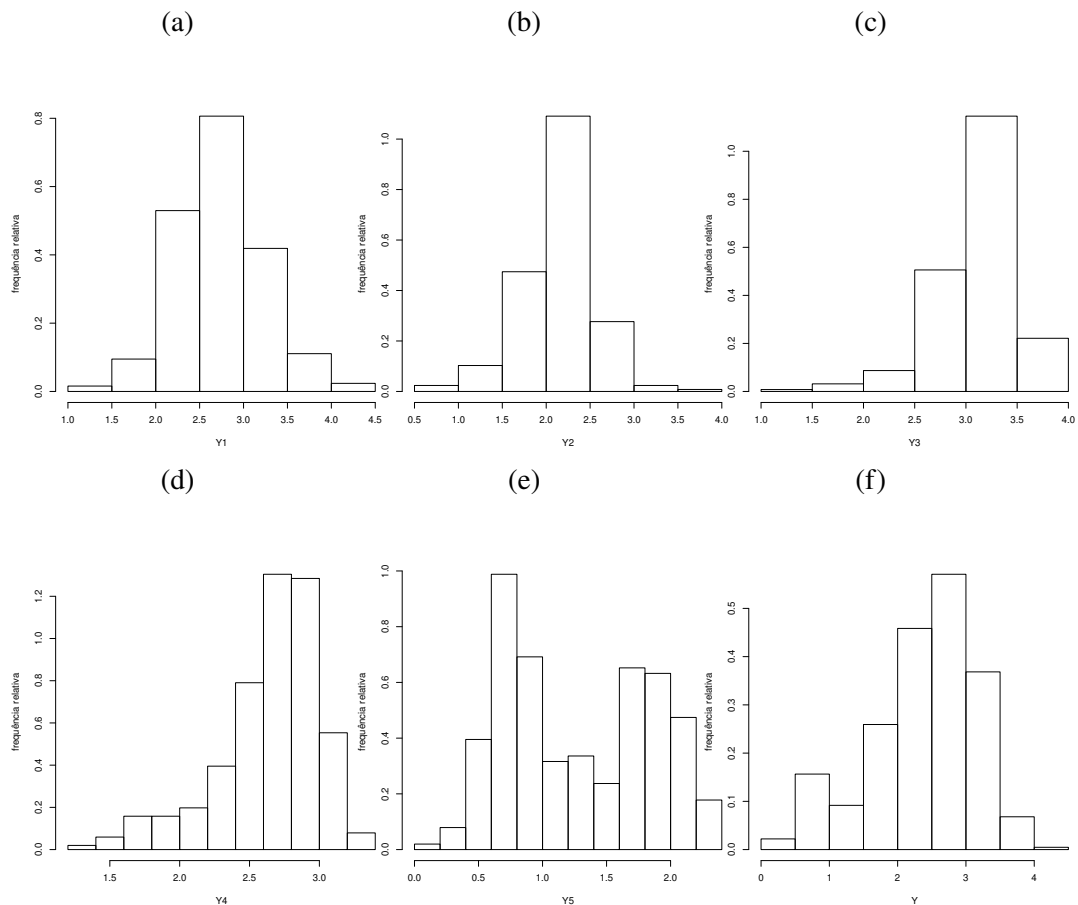
Tabela 1 – Análise descritiva da variável produtividade de soja.

Ano	Min.	1º Quartil	Mediana	Média	3º Quartil	Máx.	Var.
1998	1.190	2.420	2.740	2.755	3.050	4.140	0.237
1999	0.690	1.960	2.190	2.154	2.390	3.610	0.158
2000	1.260	2.930	3.150	3.106	3.330	3.980	0.143
2001	1.310	2.520	2.720	2.644	2.880	3.340	0.135
2002	0.190	0.770	1.220	1.283	1.820	2.340	0.330

Na Figura 1 estão apresentados os histogramas de cada repetição para dados de produtividade de soja. As Figuras 1(a), 1(b), 1(c) e 1(d), referente aos anos de 1998, 1999, 2000 e 2001, respec-

tivamente, nota-se unimodalidade. Os dados de produtividade de soja nos anos de 2000 e 2001 apresentam assimetria negativa, diferente dos anos de 1998 e 1999, um pouco mais simétricas. O ano de 2002, na Figura 1(e), apresenta dados de produtividade de soja com bimodalidade, novamente, o mais discrepante de todos os anos. Considere $\mathbf{Y} = \text{vec}(Y_{(1)}, Y_{(2)}, Y_{(3)}, Y_{(4)}, Y_{(5)})$. Assim, $\mathbf{Y}$ possui distribuição empírica vista na Figura 1(f). Nesta figura, nota-se uma distribuição assimétrica negativa. A Figura 2 (DE BASTIANI, 2016) mostra os boxplots para cada repetição. Na figura apenas o ano de 2002, novamente, difere dos demais anos. Todas repetições possuem observações que podem ser classificados como *outliers*, exceto 2002.

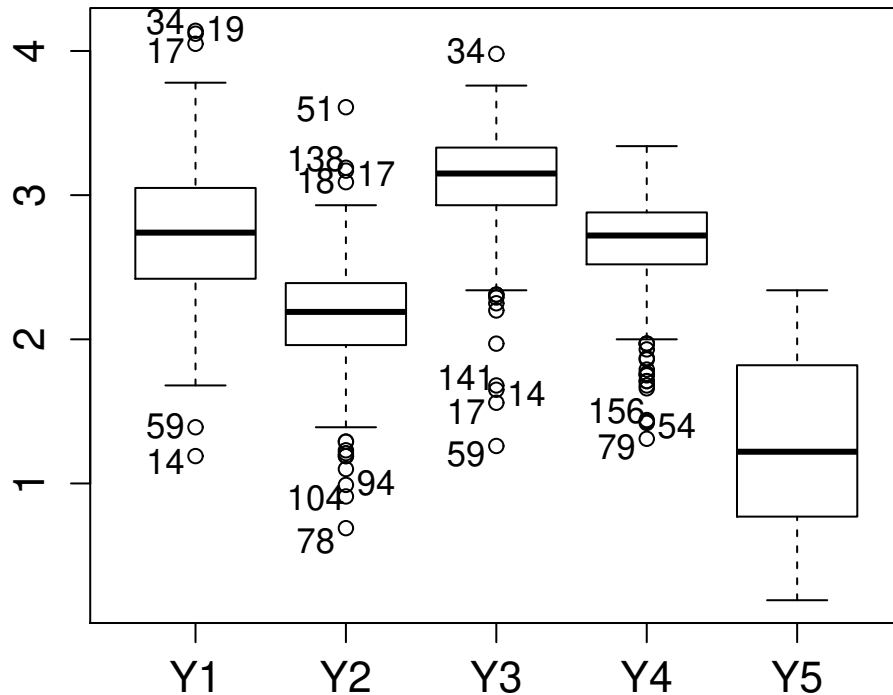
Figura 1 – (a), (b), (c), (d), (e) são histogramas dos dados de produtividade de soja para cada uma das cinco repetições (1998, 1999, 2000, 2001 e 2002, respectivamente) e (f) é um histograma de $\mathbf{Y}$.



Fonte: O AUTOR (2020).

A seguir, DE BASTIANI (2016) ajustou o MLEGr com estrutura de covariância Gaussiano e, utilizando as estatísticas LR e sua versão corrigida por Bartlett, executou diversos testes de hipóteses no modelo relacionando a produtividade de soja e os atributos químicos do solo. Nesta tese, utilizaremos outra estrutura de covariância no MLEGr, além de complementar com outras estatísticas de testes à mesma problemática. A discussão dessa aplicação será feita ao final do

Figura 2 – Boxplot de produtividade de soja nos anos de 1998 ($Y_{(1)}$), 1999 ($Y_{(2)}$), 2000 ($Y_{(3)}$), 2001 ($Y_{(4)}$) e 2002 ($Y_{(5)}$).



Fonte: DE BASTIANI (2016)

Capítulo 3.

Neste trabalho, complementamos o estudo de simulação de DE BASTIANI (2016) para avaliar o desempenho dos testes baseados na estatística da razão de verossimilhanças no MLEGr em diversos cenários de simulação, variando tamanhos de amostra, número de réplicas, além de parâmetros de interesse e perturbação. Foram utilizadas versões modificadas da estatística razão de verossimilhanças, em particular, sua correção de Bartlett. Considerando que o número de parâmetros de perturbação pode influenciar de maneira negativa nas inferências a cerca do modelo, utilizamos uma versão da estatística LR baseada na função de verossimilhança perfilada de Cox e Reid (1987), visando corrigir tal efeito dos parâmetros de perturbação, sendo necessário a ortogonalização de parâmetros. Como uma das contribuições da tese, desenvolvemos o fator de correção de Bartlett para a estatística LR baseada nessa função de verossimilhança perfilada.

Além dessas estatísticas utilizadas, recorreremos também a ajustes da estatísticas de razão de verossimilhanças. Skovgaard (SKOVGAARD et al., 1996; SKOVGAARD, 2001) propôs dois ajustes para a estatística LR, os quais foram utilizadas como alternativas para o estudo em testes

de hipóteses para o MLEGr. Diferente das estatísticas baseadas na correção de Bartlett, esta dupla de estatísticas apresenta relação com a estatística da razão de verossimilhanças sinalizada (BARNDORFF-NIELSEN, 1986, 1991), sendo esta uma versão para testes quando tem-se apenas um parâmetro de interesse. Portanto, o caso em que tem-se um vetor de parâmetros de interesse foi a extensão desenvolvida por Skovgaard. Ambas as estatísticas propostas são assintoticamente equivalentes e dependem de derivadas com relação ao espaço amostral, que podem ser aproximadas por quantidades que dependem do logaritmo da função de verossimilhança até a segunda ordem, ordem menor das correções de Bartlett que necessitam não só dessas derivadas até a quarta ordem, mas também suas esperanças. Ademais, ao contrário das estatísticas baseadas na função de verossimilhança perfilada propostas por Cox e Reid (1987), esses ajustes não necessitam de ortogonalização dos parâmetros.

Ferrari e Cysneiros (2008) desenvolveram novos ajustes à estatística da razão de verossimilhanças em modelos não lineares da família exponencial. Esses ajustes foram baseados na proposta de Skovgaard (2001).

Melo (2009) desenvolveu ajustes para a estatística da razão de verossimilhanças em modelo com erros nas variáveis multivariado baseado na proposta de Skovgaard (2001). Além disso, desenvolveu também um refinamento para teste de hipóteses no modelo estrutural elíptico baseado na proposta de Skovgaard et al. (1996).

O presente trabalho é dividido de forma a apresentar o MLEGr e suas características, em seguida, procedemos ao método de estimação dos parâmetros e por fim apresentamos os testes de razão de verossimilhança perfiladas modificadas (LR_{CR}) e usual no MLEGr no Capítulo 2. No Capítulo 3 apresentamos as principais contribuições do nosso trabalho, a saber: a obtenção do fator de correção de Bartlett (DICICCIO; STERN, 1994) à estatística LR_{CR} na proposta de Cox e Reid (1987), ajustes à estatística LR, LR_{SK1} e LR_{SK2} , baseadas na proposta de Skovgaard. Além disso, apresentamos o cálculo do fator de correção de Bartlett à estatística LR desenvolvido por DE BASTIANI (2016) e complementamos a análise a dados reais iniciadas neste capítulo. A comparação da performance dos diferentes testes apresentados nesta tese estão no Capítulo 4, os quais são analisados em função da proximidade das probabilidades de rejeição da hipótese nula. Além disso, temos como objetivo, analisar a influência do número de parâmetros de interesse, do número de parâmetros de perturbação, do tamanho da amostra e do número de réplicas nos desempenhos dos testes e compará-los entre si. Comparamos também a média e a variância das estatísticas com as quantidades correspondentes da distribuição de referência. Vale salientar que em relação ao estudo de simulação dos testes baseados nas estatísticas LR e LR^* , foram explorados outros cenários além dos que DE BASTIANI (2016) utilizou. Por fim, apresentamos algumas considerações finais sobre o nosso trabalho no Capítulo 5. Nos apêndices A, B e C detalhamos a obtenção dos fatores de correção de Bartlett para as estatísticas da razão de verossimilhanças e razão de verossimilhanças perfiladas modificada e também a obtenção das quantidades necessárias ao cálculo da estatística de razão de verossimilhanças ajustada de Skovgaard, respectivamente.

2 DEFINIÇÃO DO MODELO MLEGR E ASPECTOS INFERENCIAIS

A dependência espacial é uma característica que pode estar presente nos dados georreferenciados. Assim, essa dependência espacial mensurada de maneira eficaz se torna fundamental para revelar o nível de interdependência geográfica entre as variáveis, além da força desse relacionamento (SEIDEL; OLIVEIRA, 2014). Uma das formas de estudar a interdependência dos elementos amostrais é através da modelagem paramétrica da forma da matriz de covariância.

Dados com medidas repetidas têm sido estudados em diversas áreas como por exemplo biologia, medicina e agricultura, entretanto poucos trabalhos podem ser encontrados sobre o tema no campo da geoestatística. Neste campo, as observações são retiradas de diferentes unidades experimentais, cada uma com localização geográfica diferente, sendo cada variável observada mais de uma vez. Tendo como pressuposto independentes realizações do processo, pode-se verificar através de teste de hipóteses quais das variáveis explicativas devem fazer parte do modelo.

Quando o número de observações é pequena, é importante usar estratégias que tornem inferências mais precisas. O teste razão de verossimilhanças apresenta distorções de tamanho quando a amostra é pequena ou moderada, porque a aproximação da distribuição χ^2 limite possui erro de ordem n^{-1} . Dessa distribuição limite é que os valores críticos são obtidos.

Nosso principal objetivo é desenvolver testes de hipóteses baseados nas estatísticas da razão de verossimilhanças original e razão de verossimilhanças perfiladas modificada baseada proposta de Cox e Reid (1987) para o modelo linear espacial Gaussiano com repetição (MLEGr). Abordaremos também inferência para estimar os parâmetros da classe de modelo de família Matérn (1960). O capítulo segue com a apresentação do modelo linear espacial Gaussiano com repetições na Seção 2.1. Em seguida, na Seção 2.2, apresentamos alguns aspectos inferenciais da classe de modelos da família Mátern, a saber: a estimação dos parâmetros e os testes de hipóteses da razão de verossimilhanças e razão de verossimilhanças perfiladas modificada baseada na proposta de Cox e Reid (1987).

2.1 DEFINIÇÃO DO MODELO MLEGR

Considere um processo estocástico Gaussiano $\{\mathbf{Y}_{(i)}(\mathbf{s}), \mathbf{s} \in S\}$, em que S é um subconjunto de R^k , sendo k a dimensão do espaço euclidiano e $i = 1, \dots, r$, sendo r o número de realizações deste processo. Utilizaremos como notação o subscrito (i) quando nos referirmos a alguma repetição e diferir de outras quantidades que aparecerão no decorrer da tese. Ainda, utilizamos o espaço euclidiano bidimensional, ou seja, $k = 2$. Além disso, $\mathbf{Y}_{(i)} = \mathbf{Y}_{(i)}(\mathbf{s}) = (Y_{(i)}(\mathbf{s}_1), \dots, Y_{(i)}(\mathbf{s}_n))$ é observado em lugares conhecidos (DE BASTIANI et al., 2015), ou seja, as coordenadas geográficas as quais as unidades experimentais observadas são conhecidas. Seja $\mathbf{Y} = \text{vec}(\mathbf{Y}_{(1)}, \mathbf{Y}_{(2)}, \dots, \mathbf{Y}_{(r)})$ um vetor aleatório de r processos estocásticos independentes,

cujo cada processo tem tamanho n . O operador "vec" concatena as colunas da matriz como um vetor. Se, para $i \geq 1$, o processo $\mathbf{Y}_{(i)}$ é Gaussiano, isto é, $(Y_{(i)}(\mathbf{s}_1), \dots, Y_{(i)}(\mathbf{s}_n))$ tem distribuição normal multivariada, então este processo pertence à família Gaussiana de distribuições.

O modelo para este processo, na sua forma matricial, pode ser expresso como

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

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p, \beta_{p+1}, \dots, \beta_q)^\top$ é um vetor de dimensão $q \times 1$ de parâmetros desconhecidos, $\mathbf{X}(\mathbf{s}) = \mathbf{X} = \text{vec}(\mathbf{X}_{(1)}, \mathbf{X}_{(2)}, \dots, \mathbf{X}_{(r)})$ é a matriz modelo de dimensões $nr \times q$, cujos elementos de $\mathbf{X}_{(i)}$, as covariáveis observadas na repetição i , para $i = 1, 2, \dots, r$, são da forma $x_{jl} = x_l(s_j)$, para $j = 1, \dots, n$ e $l = 1, \dots, q$. Além disso, $\boldsymbol{\epsilon} = \text{vec}(\boldsymbol{\epsilon}_{(1)}, \boldsymbol{\epsilon}_{(2)}, \dots, \boldsymbol{\epsilon}_{(r)})$ é um erro estocástico de dimensões $nr \times 1$, com distribuição normal multivariada, tendo como média o vetor $\mathbf{0}$ e matriz de covariância $\boldsymbol{\Sigma}$.

A matriz de covariância de $\mathbf{Y}_{(i)}$, para $i = 1, \dots, r$, que supomos ser um processo estacionário, com o comportamento do processo sendo o mesmo para qualquer região escolhida, e isotrópico, é representada por $\boldsymbol{\Sigma}_{(i)} = [C_i(\mathbf{s}_u, \mathbf{s}_v)]$, uma matriz de dimensão $n \times n$, não-singular, simétrica e positiva definida. A isotropia é definida no sentido fraco (STEIN, 1999) quando se assume que existe uma constante m e uma função C em $[0, \infty)$ tal que

$$\begin{aligned} E[\mathbf{Y}(\mathbf{s})] &= m \quad \text{e} \\ \text{Cov}[\mathbf{Y}(\mathbf{s}_x), \mathbf{Y}(\mathbf{s}_y)] &= C(\|\mathbf{s}_x - \mathbf{s}_y\|), \end{aligned}$$

em que $\|\cdot\|$ indica a distância Euclidiana entre os pontos $\mathbf{s}_x$ e $\mathbf{s}_y$. Assim, a condição de isotropia estabelece que não existem razões para distinguir uma direção da outra no espaço considerado. Nós assumimos também que $\boldsymbol{\Sigma}_{(i)} = [C(\mathbf{s}_u, \mathbf{s}_v)]$, para $i = 1, \dots, r$, em que se omite o índice i da função C propositalmente. Além disso, veremos que a matriz bloco diagonal $\boldsymbol{\Sigma} = \text{diag}\{\boldsymbol{\Sigma}_{(1)}, \boldsymbol{\Sigma}_{(2)}, \dots, \boldsymbol{\Sigma}_{(r)}\}$ pode depender de um vetor de parâmetros, por exemplo, $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3, \phi_4)^\top$, e portanto $\boldsymbol{\Sigma}(\boldsymbol{\phi}) = \text{diag}\{\boldsymbol{\Sigma}_{(1)}(\boldsymbol{\phi}), \boldsymbol{\Sigma}_{(2)}(\boldsymbol{\phi}), \dots, \boldsymbol{\Sigma}_{(r)}(\boldsymbol{\phi})\}$, quando consideramos especificamente a função de covariância da classe Matérn.

De forma geral, a classe de funções de covariância apresenta a restrição da função ser positiva definida, sendo comum a utilização de funções que atendem a essa condição. Um exemplo da classe de funções de covariância é chamada função de covariância de Matérn, definida como

$$C(d_{uv}) = \begin{cases} \frac{\phi_2}{2^{\phi_4-1} \Gamma(\phi_4)} (d_{uv}/\phi_3)^{\phi_4} \mathbf{K}_{\phi_4}(d_{uv}/\phi_3), & d_{uv} > 0, \\ \phi_1 + \phi_2, & d_{uv} = 0, \end{cases}$$

em que $\mathbf{K}_w(u) = \frac{1}{2} \int_0^\infty x^{w-1} \exp\{-\frac{1}{2}u(x+x^{-1})\} dx$ é a função modificada de Bessel do terceiro tipo, de ordem w , $d_{uv} = \|\mathbf{s}_u - \mathbf{s}_v\|$ e $\phi_k \geq 0$, para $k = 1, 2, 3, 4$. Diversos comportamentos podem

ser obtidos dessa classe de funções de covariância. Por exemplo, para $\phi_4 = 1$, obtemos correlação elementar de Whittle (1954), quando $\phi_4 \rightarrow \infty$ temos a função de correlação Gaussiana, definida como

$$C(d_{uv}) = \begin{cases} \phi_2 \exp \{-(d_{uv}/\phi_3)^2\}, & d_{uv} > 0, \\ \phi_1 + \phi_2, & d_{uv} = 0, \end{cases}$$

enquanto que para $\phi_4 = 0.5$ é obtida uma estrutura de covariância exponencial.

A matriz de covariância é definida como

$$\Sigma_{(i)} = \phi_1 \mathbf{I}_n + \phi_2 \mathbf{R}(\phi_3, \phi_4), \quad (2.2)$$

em que $\mathbf{I}_n$ é a matriz identidade e $\mathbf{R}(\phi_3, \phi_4)$ é uma matriz simétrica, ambas com dimensões $n \times n$; ϕ_1 é um parâmetro conhecido pelo "efeito pepita", que representa a magnitude da aparente descontinuidade na origem, ϕ_2 é chamado "sill", ϕ_3 representa uma função do alcance, ou seja, o quão rapidamente a correlação decresce conforme a distância entre os pontos muda e ϕ_4 é um parâmetro de forma da família Mátern. Aqui, a matriz $\mathbf{R}$ é obtida através de um caso especial da função de covariância de Mátern, chamado exponencial, com função de covariância dada por

$$C(d_{uv}) = \begin{cases} \phi_2 \exp(-d_{uv}/\phi_3), & d_{uv} > 0 \\ \phi_1 + \phi_2, & d_{uv} = 0. \end{cases} \quad (2.3)$$

De acordo com Stein (1999), para um processo espaço-temporal observado em um local espacial fixo e pontos suficientemente distantes no tempo, é razoável assumir que as observações de diferentes tempos são independentes realizações de um campo aleatório. Goodall e Mardia (1994) explicam que podemos remover o efeito do tempo dos dados e então podemos ver os dados como medidas repetidas independentes no espaço.

O modelo em (2.1) pode ser reescrito como um modelo linear misto dado por

$$\mathbf{Y}_{(i)}(\mathbf{s}) = \boldsymbol{\mu}_{(i)}(\mathbf{s}) + \mathbf{b}_{(i)}(\mathbf{s}) + \boldsymbol{\tau}_{(i)}(\mathbf{s}), \text{ para } i = 1, \dots, r,$$

em que $\boldsymbol{\mu}_{(i)}(\mathbf{s}) = \mathbf{X}_{(i)}(\mathbf{s})\boldsymbol{\beta}$ é um vetor de dimensão $n \times 1$, $\mathbf{b}_{(j)}(\mathbf{s}) \sim N_n(\mathbf{0}, \phi_2 \mathbf{R})$ e $\boldsymbol{\tau}_{(i)}(\mathbf{s}) \sim N_n(\mathbf{0}, \phi_1 \mathbf{I}_n)$. Além disso, $\mathbf{b}_{(i)}(\mathbf{s})$ e $\boldsymbol{\tau}_{(i)}(\mathbf{s})$ são independentes. Note que na equação (2.1), $\boldsymbol{\epsilon}_{(i)}(\mathbf{s}) = \mathbf{b}_{(i)}(\mathbf{s}) + \boldsymbol{\tau}_{(i)}(\mathbf{s})$.

2.2 ASPECTOS INFERENCIAIS

Seja $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top)^\top$ um vetor $(q + 4) \times 1$ de parâmetros do modelo definido por (2.1), tendo em vista que q é a dimensão do vetor $\boldsymbol{\beta}$ e 4 é a dimensão do vetor $\boldsymbol{\phi}$. O logaritmo da função de verossimilhança (log-verossimilhança) pode ser expresso como

$$l(\boldsymbol{\theta}; \mathbf{Y}) = -\frac{nr}{2} \log(2\pi) - \frac{1}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}), \quad (2.4)$$

em que $|\cdot|$ denota o determinante de uma matriz. Assumimos que $l(\boldsymbol{\theta}; \mathbf{Y})$ seja regular com respeito aos componentes de $\boldsymbol{\theta}$ até a quarta ordem. Introduzimos a seguinte notação a fim de representar as derivadas do logaritmo da função de verossimilhança:

$$\begin{aligned} U_j &= \frac{\partial l(\boldsymbol{\theta})}{\partial \beta_j}, \\ U_{ja} &= \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_j \partial \phi_a}, \\ U_{jab} &= \frac{\partial^3 l(\boldsymbol{\theta})}{\partial \beta_j \partial \phi_a \partial \phi_b}, \end{aligned}$$

etc, em que os índices $j, l, k, \dots$ estão reservados para denotar os componentes do vetor $\boldsymbol{\beta}$ e os índices $a, b, c, \dots$ para denotar os componentes do vetor $\boldsymbol{\phi}$.

Consequentemente, os cumulantes conjuntos das derivadas da log-verossimilhança são definidos como:

$$\begin{aligned} \kappa_{jk} &= \mathbf{E}[U_{jk}], \\ \kappa_{jkl} &= \mathbf{E}[U_{jkl}], \end{aligned}$$

etc. A função escore total de $\boldsymbol{\theta}$ tem a forma $\mathbf{U}(\boldsymbol{\theta}) = (\mathbf{U}_\beta^\top, \mathbf{U}_\phi^\top)^\top$, sendo seus componentes $\mathbf{U}_\beta = (U_{\beta_1}, \dots, U_{\beta_q})^\top$ e $\mathbf{U}_\phi = (U_{\phi_1}, \dots, U_{\phi_4})^\top$. Em notação matricial, temos:

$$\mathbf{U}_\beta = \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = \mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}$$

e $\mathbf{U}_\phi = (U_{\phi_1}, \dots, U_{\phi_4})^\top$ em que $U_{\phi_j} = -\frac{1}{2} \text{tr}[\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_j] - \frac{1}{2} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^\top \boldsymbol{\Sigma}_j^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$, sendo $\boldsymbol{\Sigma}_j = \partial \boldsymbol{\Sigma} / \partial \phi_j$ e $\boldsymbol{\Sigma}_j^{-1} = \partial \boldsymbol{\Sigma}^{-1} / \partial \phi_j$ em que $j = \{1, 2, 3, 4\}$.

O vetor de parâmetros $\boldsymbol{\beta}$ considerado pode ser decomposto como $\boldsymbol{\beta} = (\boldsymbol{\Psi}_0^\top, \boldsymbol{\Psi}_1^\top)^\top$ sendo $\boldsymbol{\Psi}_0 = (\beta_1, \dots, \beta_p)^\top$ é o vetor contendo os primeiros p elementos de $\boldsymbol{\beta}$, $\boldsymbol{\Psi}_1^\top = (\beta_{p+1}, \beta_{p+2}, \dots, \beta_q)^\top$ é o vetor contendo os últimos $q - p$ elementos de $\boldsymbol{\beta}$ e $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3, \phi_4)^\top$ são vetores de parâmetros de perturbação, respectivamente. Da mesma forma, consideramos a decomposição da matriz modelo $\mathbf{X} = (\mathbf{X}_0, \mathbf{X}_1)$, em que $\mathbf{X}_0$ é uma matriz $rn \times p$ formada pelas primeiras p colunas de $\mathbf{X}$ e $\mathbf{X}_1$ é uma matriz $rn \times (q - p)$ formada pelas últimas $q - p$ colunas de $\mathbf{X}$.

Nesta tese, vamos fazer inferência em relação ao vetor de parâmetros de interesse $\boldsymbol{\Psi}_0$. Na presença de parâmetros de perturbação, inferências podem ser baseadas na função de verossimilhança perfilada. Esta depende apenas do parâmetro de interesse, uma vez que é obtida da

verossimilhança original em que os parâmetros de perturbação são substituídos por seus estimadores de máxima verossimilhança obtidos para cada valor fixado dos parâmetros de interesse. Embora a verossimilhança perfilada goze de algumas propriedades semelhantes às que são válidas para a verossimilhança genuína, ela pode apresentar, por exemplo, vícios na função escore e na informação.

Cox e Reid (1987) definiram uma versão modificada da verossimilhança perfilada explorando as consequências da ortogonalidade entre os parâmetros de interesse e os de perturbação em que um termo de ajuste é incorporado à verossimilhança perfilada anteriormente à estimação. Esta é uma versão penalizada do logaritmo da função de verossimilhança perfilada em que a penalização leva em conta a informação sobre o parâmetro de perturbação para cada valor do parâmetro de interesse.

Como proposto por Zucker, Lieberman e Manor (2000) (Equação (2), página 829), os parâmetros de interesse Ψ_0 e os parâmetros de perturbação $\boldsymbol{\nu}_1 = (\Psi_1^\top, \phi^\top)^\top$ foram ortogonalizados, resumindo-se na reparametrização $\boldsymbol{\nu} = (\Psi_0^\top, \Psi_1^\top, \phi^\top)^\top$ em $\boldsymbol{\iota} = (\Psi_0^\top, \zeta^\top, \phi^\top)^\top$, em que $\zeta = \Psi_1 + (\mathbf{X}_1^\top \Sigma^{-1} \mathbf{X}_1)^{-1} \mathbf{X}_1^\top \Sigma^{-1} \mathbf{X}_0 \Psi_0$. Podemos reescrever $\mathbf{X}\beta = \mathbf{X}'_0 \Psi_0 + \mathbf{X}_1 \zeta$, com $\mathbf{X}'_0 = [\mathbf{I}_{nr} - \mathbf{X}_1 (\mathbf{X}_1^\top \Sigma^{-1} \mathbf{X}_1)^{-1} \mathbf{X}_1^\top \Sigma^{-1}] \mathbf{X}_0$, em que $\mathbf{I}_{nr}$ representa a matriz identidade de ordem nr .

O logaritmo da função de verossimilhança perfilada de Ψ_0 é dado por:

$$\begin{aligned} l_p(\Psi_0; \mathbf{Y}) &= l(\Psi_0; \hat{\zeta}(\Psi_0), \hat{\phi}(\Psi_0), \mathbf{Y}) \\ &= -\frac{nr}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^r \log |\Sigma_{(i)}| + \mathbf{Z}_{(i)}^\top \Sigma_{(i)}^{-1} \mathbf{Z}_{(i)} \\ &= -\frac{nr}{2} \log 2\pi - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \mathbf{Z}^\top \Sigma^{-1} \mathbf{Z}, \end{aligned} \quad (2.5)$$

sendo $\mathbf{Z} = \mathbf{Y} - \mathbf{X}'_0 \Psi_0 - \mathbf{X}_1 \hat{\zeta}(\Psi_0)$ e $\hat{\phi}(\Psi_0)$ são os estimadores de máxima verossimilhança de ζ e ϕ , respectivamente, para um valor fixo Ψ_0 . Para a obtenção do EMV de Ψ_0 , $\hat{\Psi}_0$, é necessário maximizar a função na Equação (2.5) com às restrições $\mathbf{U}_\zeta = \mathbf{0}$ e $\mathbf{U}_\phi = \mathbf{0}$ (WRIGHT; NOCEDAL, 1999).

A função escore total de $\boldsymbol{\iota}$ tem a forma $\mathbf{U}(\boldsymbol{\iota}) = \partial l(\boldsymbol{\iota}) / \partial \boldsymbol{\iota} = (\mathbf{U}_{\Psi_0}^\top, \mathbf{U}_\zeta^\top, \mathbf{U}_\phi^\top)^\top$. Em notação matricial, temos:

$$\begin{aligned} \mathbf{U}_{\Psi_0} &= \mathbf{X}'_0{}^\top \Sigma^{-1} \mathbf{Z}, \\ \mathbf{U}_\zeta &= \mathbf{X}_1^\top \Sigma^{-1} \mathbf{Z} \end{aligned}$$

e $\mathbf{U}_\phi = (U_{\phi_1}, \dots, U_{\phi_4})^\top$ que possui elementos da forma

$$U_{\phi_j} = -\frac{1}{2} \text{tr} [\Sigma^{-1} \Sigma_j] + \Psi_0^\top \mathbf{X}'_{0j}{}^\top \Sigma^{-1} \mathbf{Z} - \frac{1}{2} \mathbf{Z}^\top \Sigma_j^{-1} \mathbf{Z}$$

e $\mathbf{X}'_{0j} = \partial \mathbf{X}'_0 / \partial \phi_j$, para $j = \{1, 2, 3, 4\}$.

As segundas derivadas da log-verossimilhança dada em (2.5) são determinadas pelas expressões dadas abaixo:

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\Psi}_0} &= -\mathbf{X}'_0{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_0, \\ \frac{\partial^2 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\zeta}} &= 0, \\ \frac{\partial^2 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \phi_j} &= \mathbf{X}'_{0j}{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} + \mathbf{X}'_0{}^\top \boldsymbol{\Sigma}_j^{-1} \mathbf{Z}, \\ \frac{\partial^2 l(\boldsymbol{\iota})}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta}} &= -\mathbf{X}_1{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_1, \\ \frac{\partial^2 l(\boldsymbol{\iota})}{\partial \boldsymbol{\zeta} \partial \phi_j} &= \mathbf{X}_1{}^\top \boldsymbol{\Sigma}_j^{-1} (\mathbf{Y} - \mathbf{X}_1 \boldsymbol{\zeta})\end{aligned}$$

e

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\iota})}{\partial \phi_j \partial \phi_k} &= -\frac{1}{2} \text{tr} [\boldsymbol{\Sigma}_j^{-1} \boldsymbol{\Sigma}_k] - \boldsymbol{\Psi}_0{}^\top \mathbf{X}'_{0k}{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}'_{0j} \boldsymbol{\Psi}_0 + \boldsymbol{\Psi}_0{}^\top (\mathbf{X}'_{0jk}{}^\top \boldsymbol{\Sigma}^{-1} + \mathbf{X}'_{0k}{}^\top \boldsymbol{\Sigma}_j^{-1} + \mathbf{X}'_{0j}{}^\top \boldsymbol{\Sigma}_k^{-1}) \mathbf{Z} \\ &\quad - \frac{1}{2} \mathbf{Z}{}^\top \boldsymbol{\Sigma}_{jk}^{-1} \mathbf{Z},\end{aligned}$$

em que $\mathbf{X}'_{0jk} = \partial \mathbf{X}'_{0j} / \partial \phi_k$.

A matriz de informação esperada de Fisher para o vetor $\boldsymbol{\iota}$

$$\mathbf{K}_{\boldsymbol{\iota}} = -\mathbf{E}[\partial^2 l(\boldsymbol{\iota}) / \partial \boldsymbol{\iota} \partial \boldsymbol{\iota}{}^\top] = \begin{bmatrix} \mathbf{K}_{\boldsymbol{\Psi}_0 \boldsymbol{\Psi}_0} & 0 \\ 0 & \mathbf{K}_{\boldsymbol{\varphi} \boldsymbol{\varphi}} \end{bmatrix} = \begin{bmatrix} \mathbf{K}_{\boldsymbol{\Psi}_0 \boldsymbol{\Psi}_0} & 0 & 0 \\ 0 & \mathbf{K}_{\boldsymbol{\zeta} \boldsymbol{\zeta}} & \mathbf{K}_{\boldsymbol{\zeta} \boldsymbol{\phi}} \\ 0 & \mathbf{K}_{\boldsymbol{\phi} \boldsymbol{\zeta}} & \mathbf{K}_{\boldsymbol{\phi} \boldsymbol{\phi}} \end{bmatrix},$$

em que as componentes são dadas por

$$\mathbf{K}_{\boldsymbol{\Psi}_0 \boldsymbol{\Psi}_0} = -\mathbf{X}'_0{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}'_0,$$

$$\mathbf{K}_{\boldsymbol{\zeta} \boldsymbol{\zeta}} = -\mathbf{X}_1{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_1,$$

$\mathbf{K}_{\boldsymbol{\zeta} \boldsymbol{\phi}}$ tem elementos $\mathbf{K}_{\boldsymbol{\zeta} \phi_j} = \mathbf{X}_1{}^\top \boldsymbol{\Sigma}_j^{-1} \mathbf{X}'_0 \boldsymbol{\Psi}_0$, para $j = \{1, 2, 3, 4\}$ e $\mathbf{K}_{\boldsymbol{\phi} \boldsymbol{\phi}}$ tem elementos $K_{\phi_j \phi_k} = \frac{1}{2} \text{tr} [\boldsymbol{\Sigma}_j^{-1} \boldsymbol{\Sigma}_k] - \boldsymbol{\Psi}_0{}^\top \mathbf{X}'_{0j}{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}'_{0k} \boldsymbol{\Psi}_0$, para j e k variando em $\{1, 2, 3, 4\}$. A matriz $\mathbf{K}_{\boldsymbol{\Psi}_0 \boldsymbol{\Psi}_0}$ possui dimensão $p \times p$, $\mathbf{K}_{\boldsymbol{\zeta} \boldsymbol{\zeta}}$ tem dimensão $(q - p) \times (q - p)$, $\mathbf{K}_{\boldsymbol{\zeta} \boldsymbol{\phi}} = \mathbf{K}_{\boldsymbol{\phi} \boldsymbol{\zeta}}{}^\top$ de dimensão $(q - p) \times 4$ e $\mathbf{K}_{\boldsymbol{\phi} \boldsymbol{\phi}}$ com dimensão 4×4 . Importante observar que devido ao processo de ortogonalização efetuado, a matriz $\mathbf{K}_{\boldsymbol{\iota}}$ é bloco diagonal. A ortogonalidade de parâmetros é definida a partir da esperança do produto cruzados entre $\theta_s \in \boldsymbol{\Psi}_0$ e $\theta_r \in \boldsymbol{\varphi}$ são nulos, ou seja,

$$\mathbf{E} \left[\frac{\partial l}{\partial \theta_s} \frac{\partial l}{\partial \theta_r} \right] = \mathbf{E} \left[\frac{\partial^2 l}{\partial \theta_s \partial \theta_r} \right] = 0,$$

que tem reflexo na matriz de informação $\mathbf{K}$, cujo elementos referentes a θ_s e θ_r serão nulos.

2.3 ESTIMAÇÃO DO PARÂMETRO ϕ

O estimador de máxima verossimilhança (EMV) de β é obtido como solução da equação

$$\mathbf{U}_\beta = \frac{\partial l(\boldsymbol{\theta})}{\partial \beta} = \mathbf{0},$$

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

Entretanto, o EMV de ϕ não possui solução em forma fechada. O logaritmo da função de verossimilhança perfilada é dado por

$$\begin{aligned} l_p(\phi; \mathbf{Y}) &= l(\phi; \hat{\beta}(\phi), \mathbf{Y}) \\ &= -\frac{nr}{2} \log 2\pi - \log |\boldsymbol{\Sigma}| - \frac{1}{2} \boldsymbol{\rho}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\rho}, \end{aligned}$$

sendo $\boldsymbol{\rho} = \mathbf{Y} - \mathbf{X}(\mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Y}$, que depende apenas de ϕ .

A aplicação de teoria assintótica a dados espaciais é complicada pelo fato que existem dois diferentes quadro assintóticos: aumento do domínio, na qual a mínima distância entre dois pontos amostrais é limitado longe do zero e portanto esse domínio espacial da observação é ilimitado e o preenchimento assintótico, na qual as observações são tomadas cada vez mais densamente em um domínio fixo e limitado. Zhang e Zimmerman (2005) compararam a estimação de máxima verossimilhança sobre esses dois quadros diferentes na família de covariância Matérn e encontraram que o comportamento de amostragem finita tendem a concordar com as inferências de preenchimento assintótico. Assim, as estimativas de máxima verossimilhança dos parâmetros de covariância são tipicamente consistentes e assintoticamente normais (MARDIA; MARSHALL, 1984), entretanto nem todos os parâmetros podem ser estimados de maneira consistente sob estrutura de domínio fixo (CHEN; SIMPSON; YING, 2000). Stein (1999) é uma das referências para mais detalhes relacionado à domínio fixo, enquanto Zhang e Zimmerman (2005) discutem qual o quadro mais apropriado dos dois diferentes quadros assintóticos relacionados anteriormente.

2.4 TESTE DA RAZÃO DE VEROSSIMILHANÇAS NO MLEG_r

Nosso objetivo é testar a hipótese nula $H_0 : \boldsymbol{\Psi}_0 = \boldsymbol{\Psi}_0^{(0)}$ contra $H_1 : \boldsymbol{\Psi}_0 \neq \boldsymbol{\Psi}_0^{(0)}$, em que $\boldsymbol{\Psi}_0^{(0)}$ é um vetor conhecido. A estatística razão de verossimilhanças (LR) para testar H_0 versus H_1 é dada por

$$\text{LR} = 2 \left[l_p(\hat{\boldsymbol{\Psi}}_0) - l_p(\boldsymbol{\Psi}_0^{(0)}) \right], \quad (2.6)$$

em que $\hat{\boldsymbol{\iota}} = (\hat{\boldsymbol{\Psi}}_0^\top, \hat{\boldsymbol{\zeta}}^\top, \hat{\boldsymbol{\phi}}^\top)^\top$ é o EMV irrestrito de $\boldsymbol{\iota}$ e $\tilde{\boldsymbol{\iota}} = (\boldsymbol{\Psi}_0^{(0)\top}, \tilde{\boldsymbol{\zeta}}^\top, \tilde{\boldsymbol{\phi}}^\top)^\top$ é o EMV restrito para $\boldsymbol{\iota}$ em que $\tilde{\boldsymbol{\zeta}}$ e $\tilde{\boldsymbol{\phi}}$ são o EMV restrito de $\boldsymbol{\zeta}$ e $\boldsymbol{\phi}$ sob H_0 , em que l_p é o logaritmo da função de verossimilhança perfilada apresentada na equação (2.5) e $\hat{\boldsymbol{\Psi}}_0$ denota o EMV de $\boldsymbol{\Psi}_0$. Assintoticamente

e sob H_0 , a estatística LR tem distribuição χ_p^2 .

Cox e Reid (1987) definiram uma versão modificada da verossimilhança perfilada, em que um termo de ajuste é incorporado à verossimilhança perfilada anteriormente à estimação. Considerando, portanto, os parâmetros de perturbação $\varphi = (\zeta^\top, \phi^\top)^\top$ e os parâmetros de interesse ortogonais. A log-verossimilhança perfilada modificada correspondente para Ψ_0 é dada por

$$l_{CR}(\Psi_0) = l_p(\Psi_0) - \frac{1}{2} \log |-l_{\varphi\varphi}(\hat{\varphi}(\Psi_0))|,$$

em que $l_{\varphi\varphi}$ denota a matriz de segundas derivadas de l com respeito a φ .

A estatística da razão de verossimilhanças perfiladas modificada (LR_{CR}) para o teste de H_0 versus H_1 é dada por

$$LR_{CR} = 2 \left[l_{CR}(\tilde{\Psi}_0) - l_{CR}(\Psi_0^{(0)}) \right], \quad (2.7)$$

em que $\tilde{\Psi}_0$ é o valor de Ψ_0 que maximiza a função l_{CR} . Assintoticamente e sob H_0 , a estatística LR_{CR} , tem distribuição χ_p^2 .

3 CORREÇÃO DE BARTLETT PARA A ESTATÍSTICA DA RAZÃO DE VEROSSIMILHANÇAS

Em problemas regulares (COX; HINKLEY, 1979), a estatística da razão de verossimilhanças (LR) tem, sob a hipótese nula (H_0), uma distribuição qui-quadrado (χ^2) aproximadamente, em amostras grandes, ou seja, χ^2 com p graus de liberdade, em que p é a diferença entre as dimensões dos espaços paramétricos sob as hipóteses alternativa (H_1) e H_0 . O teste da razão de verossimilhanças envolve estimação dos parâmetros sob as hipóteses nula e alternativa. Para a estatística LR o erro dessa aproximação é de ordem n^{-1} , isto é, $O(n^{-1})$, em que n é o tamanho da amostra (BARNDORFF-NIELSEN; E. HALL, 1988).

Os testes são usualmente baseados na comparação das estatísticas com o valor crítico obtido da distribuição χ^2 de referência para um nível de significância nominal fixado. Em geral, há grande dificuldade em se determinar a distribuição exata da estatísticas LR. Esta é a razão pela qual os testes têm sido construídos com base em resultados assintóticos. No entanto, as aproximações resultantes podem não ser boas para amostras pequenas ou mesmo de tamanho moderado pois os testes podem apresentar taxas de rejeição bastante liberal rejeitando muito mais a hipóteses H_0 do que deveria.

A correção de Bartlett é utilizada para melhorar a qualidade da aproximação da estatística da razão de verossimilhanças pela distribuição χ^2 . A ideia de modificar a estatística LR por um fator de correção, visando a produzir uma estatística modificada (LR^*) com o primeiro momento igual ao da distribuição χ^2 , é devida a Bartlett (1937). Deve-se destacar que tal fator não depende do valor da estatística LR, mas podem depender de parâmetros desconhecidos e, neste caso, estes devem ser substituídos por suas respectivas estimativas de máxima verossimilhança sob H_0 .

Em problemas regulares, a estatística LR^* tem distribuição assintótica χ_p^2 com erro de ordem n^{-2} (ver Barndorff-Nielsen e E. Hall (1988)). Assim, temos que $\Pr[LR > x_\alpha] = \alpha + O(n^{-1})$, tem-se $\Pr[LR^* > x_\alpha] = \alpha + O(n^{-2})$, em que x_α é o quantil $(1 - \alpha)$ da distribuição χ_p^2 . Um método geral de obtenção da constante que aparece no fator de correção foi desenvolvido por Lawley (1956) e uma revisão sobre correção de Bartlett à estatística LR pode ser encontrada em Cordeiro e Cribari-Neto (2014).

Melo, Ferrari e Cribari-Neto (2009) desenvolveram os fatores de correção de Bartlett para as estatísticas da razão de verossimilhanças usual e razão de verossimilhanças perfiladas modificada baseada na proposta de Cox e Reid (1987) em modelos lineares mistos.

Pires et al. (2018) propuseram refinamento de inferência para a distribuição Pareto generalizada baseadas nas propostas de Cox e Reid (1987, 1993), Barndorff-Nielsen (1983), Fraser e Reid (1995) e Severini (1999).

Araújo, Cysneiros e Montenegro (2017) propuseram um refinamento do teste de heteroscedasticidade baseado na estatística da razão de verossimilhanças perfiladas modificada proposta

por Cox e Reid (1987) nos modelos não lineares simétricos heteroscedásticos.

Desta forma, considere o modelo em estudo, sendo o vetor de parâmetros $\boldsymbol{\iota}$ particionado da seguinte forma $\boldsymbol{\iota} = (\boldsymbol{\Psi}_0^\top, \boldsymbol{\zeta}^\top, \boldsymbol{\phi}^\top)^\top$, com $\boldsymbol{\Psi}_0$ o vetor $p \times 1$ de parâmetros de interesse e $(\boldsymbol{\zeta}^\top, \boldsymbol{\phi}^\top)$ são os vetores de parâmetros de perturbação de dimensões $(q - p) \times 1$ e 4×1 , respectivamente.

Nosso interesse é testar a hipótese $H_0 : \boldsymbol{\Psi}_0 = \boldsymbol{\Psi}_0^{(0)}$ contra $H_1 : \boldsymbol{\Psi}_0 \neq \boldsymbol{\Psi}_0^{(0)}$, em que $\boldsymbol{\Psi}_0^{(0)}$ é um vetor p dimensional de constantes especificadas, no modelo linear espacial gaussiano com repetições, utilizaremos inferência assintótica de segunda ordem.

Para grandes amostras, a distribuição da estatística da razão de verossimilhanças, LR, pode ser aproximada pela distribuição χ_p^2 . No entanto, a aproximação resultante pode não ser satisfatória para amostras pequenas, conduzindo assim a taxas de rejeição bastante distorcidas. Bartlett (1937) propôs um fator de correção $(1 + c/p)^{-1}$ que é multiplicado à estatística LR, produzindo uma estatística da razão de verossimilhanças modificada (LR*) dada por

$$\text{LR}^* = \frac{\text{LR}}{1 + c/p}, \quad (3.1)$$

com média mais próxima do valor esperado da distribuição χ^2 . A constante c é obtida a partir de cumulantes da log-verossimilhança e será discutida a seguir.

3.1 OBTENÇÃO DO FATOR DE CORREÇÃO DE BARTLETT NO MLEGR

Assim como Melo, Ferrari e Cribari-Neto (2009), uma das contribuições desta tese está na obtenção do fator de correção de Bartlett para o modelo geoestatístico utilizando a verossimilhança perfilada.

DE BASTIANI (2016) desenvolveu uma expressão para c , sendo c função de cumulantes e de derivadas de cumulantes, de ordem n^{-1} , expresso da seguinte forma

$$c = \sum_{\boldsymbol{\Psi}_0, \boldsymbol{\zeta}, \boldsymbol{\phi}} (\lambda_{rstu} - \lambda_{rstuvw}) - \sum_{\boldsymbol{\zeta}, \boldsymbol{\phi}} (\lambda_{rstu} - \lambda_{rstuvw}), \quad (3.2)$$

em que

$$\lambda_{rstu} = \kappa^{rs} \kappa^{tu} \left[\frac{1}{4} \kappa_{rstu} - (\kappa_{rst})_u - (\kappa_{rt})_{su} \right],$$

e

$$\begin{aligned} \lambda_{rstuvw} = & \kappa^{rs} \kappa^{tu} \kappa^{vw} \left\{ \kappa_{rtv} \left(\frac{1}{6} \kappa_{suw} - (\kappa_{sw})_u \right) + \left(\frac{1}{4} \kappa_{svw} - (\kappa_{sw})_v \right) \right. \\ & \left. + (\kappa_{rt})_v (\kappa_{sw})_u + (\kappa_{rt})_u (\kappa_{sw})_v \right\}. \end{aligned}$$

com $\kappa^{rs} = \kappa^{\{r,s\}}$ representando o elemento (r, s) da inversa da matriz de informação de Fisher $\mathbf{K}$ de $\boldsymbol{\iota}$, sendo as derivadas dos cumulantes em relação aos componentes do vetor $\boldsymbol{\iota}$ dadas por enquanto $(\kappa_{sw})_u = \partial \kappa_{sw} / \partial \boldsymbol{\iota}_u$ e $(\kappa_{rt})_{su} = \partial^2 \kappa_{rt} / \partial \boldsymbol{\iota}_s \partial \boldsymbol{\iota}_u$. Na Equação (3.2), enquanto o primeiro somatório r, s, t, u, v, w variam em todos os parâmetros do modelo $(\boldsymbol{\Psi}_0^\top, \boldsymbol{\zeta}^\top, \boldsymbol{\phi}^\top)^\top$, no segundo

somatório os mesmos índices variam apenas nos parâmetros de perturbação $(\zeta^\top, \phi^\top)^\top$. Os terceiros e quartos cumulantes e derivadas dos cumulantes podem ser encontrados no Apêndice A.

DE BASTIANI (2016) obteve o fator de correção de Bartlett para a estatística LR no MLEGr, a Equação 3.2 resume-se a

$$c = \sum \left[-\frac{1}{2} \kappa^{ab} \kappa^{jk} \kappa_{abjk} + \frac{1}{4} \kappa^{ab} \kappa^{cd} \kappa^{jk} \kappa_{abj} \kappa_{cdk} - \frac{1}{2} \kappa^{ab} \kappa^{jk} \kappa^{lo} \kappa_{abj} (\kappa_{lok} - 2(\kappa_{lo})_k) \right. \\ \left. + \kappa^{ab} \kappa^{fk} \kappa^{jl} \kappa_{abj} \left(2(\kappa_{fk})_l - \frac{3}{2} \kappa_{fkl} \right) - \frac{1}{2} \kappa^{ab} \kappa^{fg} \kappa^{jk} \kappa_{abj} \kappa_{fgk} \right],$$

em que os índices a, b, c e d variam em elementos de Ψ_0 , f e g em elementos de ζ e j, k, l e o em elementos de ϕ . Em notação matricial, temos

$$c = \text{tr} \left(\frac{1}{r} \mathbf{D}^{-1} \left(-\frac{1}{2} \mathbf{M} + \frac{1}{4} \mathbf{P} - \frac{1}{2} (\boldsymbol{\varsigma} + \boldsymbol{\nu}) \boldsymbol{\gamma}^\top \right) \right), \quad (3.3)$$

em que $\mathbf{D}$, $\mathbf{M}$, $\mathbf{P}$ são matrizes de dimensão 4×4 dadas por

$$\mathbf{D} = \frac{1}{2} \text{tr} \left(\boldsymbol{\Sigma}_k^{-1} \boldsymbol{\Sigma}_l \right), \\ \mathbf{M} = \text{tr} \left[\left(\mathbf{X}_0'^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_0' \right)^{-1} \left(\mathbf{X}_0'^\top \boldsymbol{\Sigma}_{kl}^{-1} \mathbf{X}_0' + 2 \mathbf{X}_{0l}'^\top \boldsymbol{\Sigma}_k^{-1} \mathbf{X}_0' \right) \right], \\ \mathbf{P} = \text{tr} \left[\left(\mathbf{X}_0'^\top \boldsymbol{\Sigma}_k^{-1} \mathbf{X}_0' \right) \left(\mathbf{X}_0'^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_0' \right)^{-1} \left(\mathbf{X}_0'^\top \boldsymbol{\Sigma}_l^{-1} \mathbf{X}_0' \right) \left(\mathbf{X}_0'^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_0' \right)^{-1} \right],$$

e $\boldsymbol{\varsigma}$, $\boldsymbol{\gamma}$ e $\boldsymbol{\nu}$ são vetores de dimensão 4 cujo k -ésimos elementos são $\text{tr} [\mathbf{D}^{-1} \mathbf{A}_k]$, $\text{tr} \left[\left(\mathbf{X}_0'^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_0' \right)^{-1} \left(\mathbf{X}_0'^\top \boldsymbol{\Sigma}_k^{-1} \mathbf{X}_0' \right) \right]$ e $\text{tr} \left[\left(\mathbf{X}_1'^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_1' \right)^{-1} \left(\mathbf{X}_1'^\top \boldsymbol{\Sigma}_k^{-1} \mathbf{X}_1' \right) \right]$, respectivamente, em que

$$\mathbf{A}_k = \frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}_i^{-1} \boldsymbol{\Sigma}_{kl} \right] - \frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}_l^{-1} \boldsymbol{\Sigma}_{ki} \right] - \frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}_k^{-1} \boldsymbol{\Sigma}_{il} \right],$$

em que k, i e l variam nos índices $\{1, 2, 3, 4\}$. Introduzimos a seguinte notação $\boldsymbol{\Sigma}_k = \partial \boldsymbol{\Sigma} / \partial \phi_k$, $\boldsymbol{\Sigma}_{kl} = \partial^2 \boldsymbol{\Sigma} / \partial \phi_k \partial \phi_l$, $\boldsymbol{\Sigma}_k^{-1} = \partial \boldsymbol{\Sigma}^{-1} / \partial \phi_k$, $\boldsymbol{\Sigma}_{kl}^{-1} = \partial^2 \boldsymbol{\Sigma}^{-1} / \partial \phi_k \partial \phi_l$ e $\mathbf{X}_{0k}' = \partial \mathbf{X}_0' / \partial \phi_k$, para $(k, l) \in \{1, 2, 3, 4\}$. No Apêndice A estão apresentados os desenvolvimentos para a obtenção da quantidade c .

É importante salientar que a fórmula 3.3 somente envolve operações simples de matrizes e pode ser facilmente implementada em pacotes de computação simbólica e linguagens que permitam executar operações simples de álgebra linear, tais como Ox, R, MAPLE, MATHEMATICA, etc. Estes resultados são similares aos obtidos por Melo, Ferrari e Cribari-Neto (2009).

A seguir, apresentaremos algumas contribuições importantes para testes de hipóteses nos modelos geoestatísticos. Um estudo extenso de simulação foi realizado nesta tese com o intuito de analisar o efeito da influência do número parâmetros de perturbação, do número parâmetros de interesse, além do número de repetições no MLEGr como contribuição, uma vez que em DE BASTIANI (2016) só foram apresentados dois simples resultados numéricos.

No Capítulo 2 foi mostrado como fica o teste da razão de verossimilhanças perfiladas modificado (LR_{CR}) baseado no ajuste de Cox e Reid (1987). Este ajuste de Cox e Reid para a

verossimilhança perfilada, reduz o impacto de parâmetros de perturbação na inferência resultante, conduzindo a testes que produzem taxas de rejeição abaixo dos níveis nominais. Contudo, a partir dos resultados de DiCiccio e Stern (1994), Cysneiros e Ferrari (2006) desenvolveram o refinamento de um teste de heteroscedasticidade baseado em verossimilhança perfilada modificada em modelos não lineares da família exponencial, baseado na proposta de Cox e Reid (1987).

Ferrari, Cysneiros e Cribari-Neto (2004) propuseram o refinamento de um teste de heteroscedasticidade baseado em verossimilhança perfilada modificada no modelo normal linear, estendendo os resultados de Ferrari e Cribari-Neto (2002) para abordar o caso em que o parâmetro de interesse (heteroscedasticidade) é multidimensional.

Definimos a estatística da razão de verossimilhanças perfiladas modificadas corrigida como

$$\text{LR}_{\text{CR}}^* = \frac{\text{LR}_{\text{CR}}}{1 + c_{\text{CR}}/p}, \quad (3.4)$$

em que a estatística LR_{CR} foi definida em (2.7) de tal forma que, em problemas regulares, a $\Pr[\text{LR}_{\text{CR}}^* > x_\alpha] = \alpha + O(n^{-2})$ em que x_α é o quantil $(1 - \alpha)$ da distribuição χ_p^2 , enquanto que $\Pr[\text{LR}_{\text{CR}} > x_\alpha] = \alpha + O(n^{-1})$. Isto é, o erro de aproximação por χ^2 para a distribuição nula de LR_{CR} é de ordem n^{-1} , enquanto que o erro desta aproximação para a distribuição de LR_{CR}^* é reduzido à ordem n^{-2} . DiCiccio e Stern (1994, p. 404, equação (25)) definiram uma expressão geral para c_{CR} que pode ser escrita como

$$\begin{aligned} c_{\text{CR}} = & \left\{ \frac{1}{4} \tau^{ru} \tau^{st} \kappa_{rstu} - \kappa^{ru} \tau^{st} (\kappa_{rst})_u + (\kappa^{ru} \kappa^{st} - \nu^{ru} \nu^{st}) (\kappa_{rs})_{tu} - \right. \\ & - \left(\frac{1}{4} \kappa^{ru} \tau^{st} \tau^{vw} + \frac{1}{2} \kappa^{tu} \tau^{sw} \tau^{tv} - \frac{1}{3} \tau^{tu} \tau^{sw} \tau^{tv} \right) \kappa_{rst} \kappa_{uvw} + \\ & + (\kappa^{ru} \kappa^{st} \kappa^{vw} + \kappa^{ru} \kappa^{sw} \kappa^{tv} - \nu^{ru} \kappa^{sw} \nu^{tv}) \kappa_{rst} (\kappa_{uv})_w - \\ & \left. - (\kappa^{tu} \kappa^{st} \kappa^{vw} - \nu^{ru} \nu^{st}) (\kappa_{rs})_t (\kappa_{uv})_w - (\kappa^{ru})_w - (\kappa^{ru} \kappa^{sw} \kappa^{vw} - \nu^{ru} \nu^{st} \nu^{vw}) (\kappa_{rs})_t (\kappa_{uv})_w \right\}, \end{aligned}$$

em que $\nu^{rs} = \kappa^{rs} - \tau^{rs}$, $\tau^{rs} = \kappa_{rb} \kappa^{sa} \omega_{ab}$, sendo ω_{ab} elemento da matriz $K^{\Psi_0 \Psi_0}$, parte referente a Ψ_0 de $\mathbf{K}^{-1}$. Os índices r, s, t, u, v, w percorrem todos os elementos do vetor $\boldsymbol{\iota} = (\Psi_0^\top, \zeta^\top, \phi^\top)^\top$. Da mesma forma que feito anteriormente, todas as quantidades são obtidas individualmente, analisados os produtos e parcelas e, por fim, simplificadas. Dessa simplificação, resulta

$$\begin{aligned} c_{\text{CR}} = & \sum \left\{ -\kappa^{ab} \kappa^{jk} \kappa_{abjk} + \frac{1}{4} \kappa^{ab} \kappa^{cd} \kappa^{jk} \kappa_{abj} \kappa_{cdk} + \kappa^{ab} \kappa^{jk} \kappa^{lo} \kappa_{abj} (\kappa_{kl})_o + \kappa^{ab} \kappa^{fj} \kappa^{gk} \kappa_{abj} \kappa_{fgk} + \right. \\ & \left. + 2 \kappa^{ab} \kappa^{fj} \kappa^{kl} \kappa_{abj} (\kappa_{fk})_l \right\}, \end{aligned}$$

sendo a, b, c, d índices para se referir aos componentes de Ψ_0 , f e g aos componentes de ζ e j, k, l e m aos componentes de ϕ . A partir da última simplificação foi obtida a expressão geral, em notação matricial, da constante da correção de Bartlett para a estatística de razão de verossimilhanças baseada na função log-verossimilhança perfilada ajustada, definida por

$$c_{\text{CR}} = \text{tr} \left[\mathbf{D}^{-1} \left(-\mathbf{M} + \frac{1}{4} \mathbf{P} + \mathbf{s}_1 \boldsymbol{\gamma}^\top \right) \right] + \boldsymbol{\gamma}^\top \mathbf{K}^{\phi \phi} \boldsymbol{\eta}^*, \quad (3.5)$$

em que $\mathbf{D}$, $\mathbf{M}$, $\mathbf{P}$ e γ são definidos acima, enquanto ς_1 é $\text{tr} [\mathbf{D}^{-1} \mathbf{E}_k]$, sendo

$$\mathbf{E}_k = -\text{tr} [\boldsymbol{\Sigma}_l^{-1} \boldsymbol{\Sigma}_k \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_i^{-1}] + 1/2 \text{tr} [\boldsymbol{\Sigma}_k^{-1} \boldsymbol{\Sigma}_{li}] + 1/2 \text{tr} [\boldsymbol{\Sigma}_l^{-1} \boldsymbol{\Sigma}_{ki}]$$

uma matriz de dimensão 4×4 . No Apêndice B estão apresentados os desenvolvimentos para a obtenção da quantidade c_{CR} . É importante salientar que a fórmula (3.5) somente envolve operações simples de matrizes e pode ser facilmente implementada em pacotes de computação simbólica e linguagens que permitam executar operações simples de álgebra linear, tais como Ox, R, MAPLE, MATHEMATICA, etc. Estes resultados são similares aos obtidos por Melo, Ferrari e Cribari-Neto (2009).

Uma nova contribuição deste trabalho é a obtenção de ajustes para a estatística da razão de verossimilhanças (LR) com base nos resultados de Skovgaard (2001) nos modelos lineares espaciais com repetições. Estes ajustes têm ampla aplicabilidade, pois são válidos para testar quaisquer hipóteses nulas que estabeleçam um vetor de valores fixados para uma parte do vetor de parâmetros do modelo. Estes novos ajustes apresentam uma vantagem em relação ao ajuste de Cox e Reid (1987) por serem invariantes segundo reparametrizações da forma $(\nu, \psi) \rightarrow (\nu, \zeta(\nu, \psi))$, sendo ν o vetor de parâmetros de interesse, ψ o vetor de parâmetros de perturbação e ζ uma função de ν e ψ . Outra vantagem é a aplicação direta à estatística da razão de verossimilhanças, enquanto que o ajuste de Cox e Reid, ao contrário, é aplicado à função de verossimilhança perfilada, a qual deve ser maximizada segundo as hipóteses nula e alternativa para, posteriormente, obter-se a estatística ajustada. Outra vantagem adicional é que não é necessária a ortogonalização entre os parâmetros de interesse e os de perturbação. Estes novos ajustes apresentam uma vantagem em relação ao fator de correção de Bartlett uma vez que são mais simples de serem obtidos pois requerem cálculos dos cumulantes de derivadas do logaritmo da função de verossimilhanças até segunda ordem, enquanto que o cálculo do fator de correção de Bartlett é mais complexo e requer a obtenção dos cumulantes de derivadas do logaritmo da função de verossimilhanças até a quarta ordem.

Uma generalização da estatística da razão de verossimilhanças sinalizada modificada de Barndorff-Nielsen para o caso em que o parâmetro de interesse é multidimensional com propriedades similares ao caso unidimensional é proposta por Skovgaard (2001).

A estatística da razão de verossimilhanças ajustada para o teste de H_0 é dada por

$$\text{LR}_{\text{SK1}} = \text{LR} \left(1 - \frac{1}{\text{LR}} \log \rho \right)^2, \quad (3.6)$$

ou sua versão assintoticamente equivalente

$$\text{LR}_{\text{SK2}} = \text{LR} - 2 \log \rho, \quad (3.7)$$

em que ρ é definido por

$$\rho = |\hat{\mathbf{J}}|^{1/2} |\tilde{\mathbf{U}}'|^{-1} |\tilde{\mathbf{J}}_{\varphi\varphi}|^{1/2} |\tilde{\mathbf{J}}_{\varphi\varphi}|^{-1/2} |\tilde{\mathbf{J}}|^{1/2} \frac{(\tilde{\mathbf{U}}' \tilde{\mathbf{J}}^{-1} \tilde{\mathbf{U}})^{p/2}}{\text{LR}^{p/2-1} (\hat{\mathbf{L}}' - \tilde{\mathbf{L}}')^\top (\tilde{\mathbf{U}}')^{-1} \tilde{\mathbf{U}}}, \quad (3.8)$$

sendo $\mathbf{U} = (\mathbf{U}_{\Psi_0}^\top, \mathbf{U}_\varphi^\top)^\top$, $\mathbf{U}_{\Psi_0} = \partial l(\Psi_0, \varphi) / \partial \Psi_0$ e $\mathbf{U}_\varphi = \partial l(\Psi_0, \varphi) / \partial \varphi$, $\mathbf{J} = \partial^2 l(\Psi_0, \varphi) / \partial \omega \partial \omega^\top$, a matriz de segundas derivadas, com $\omega \in \{\Psi_0, \varphi\}$. Aqui, $\mathbf{J}_{\varphi\varphi}$ representa a matriz de segundas derivadas referente aos parâmetros de perturbação, com dimensão $(q - p + 4) \times (q - p + 4)$, $\widehat{\mathbf{U}} = \mathbf{U}(\widehat{\iota})$, $\widetilde{\mathbf{U}} = \mathbf{U}(\widetilde{\iota})$, $\widehat{\mathbf{J}} = \mathbf{J}(\widehat{\iota})$, $\widetilde{\mathbf{J}} = \mathbf{J}(\widetilde{\iota})$. Além disso, $\widetilde{\mathbf{J}} \approx \widetilde{\mathbf{I}} \widehat{\mathbf{T}}^{-1} \widehat{\mathbf{J}} \widehat{\mathbf{I}}^{-1} \widehat{\mathbf{T}}$, $\widetilde{\mathbf{U}}' \approx \widehat{\mathbf{T}} \widehat{\mathbf{I}}^{-1} \widehat{\mathbf{J}} \widehat{\mathbf{L}}' - \widetilde{\mathbf{L}}' \approx \widehat{h} \widehat{\mathbf{I}}^{-1} \widehat{\mathbf{J}}$ são aproximações de Skovgaard et al. (1996) e Skovgaard (2001), respectivamente, sendo $\mathbf{I} = \mathbf{E}[-\mathbf{J}]$, $\widehat{h} = \text{cov}_{\iota_0} [U(\iota_0), l(\iota_0) - l(\iota)] \mid_{\iota_0=\widehat{\iota}, \iota=\widehat{\iota}}$ e $\widehat{\mathbf{T}} = \text{cov}_{\iota_0} [U(\iota_0), U(\iota)] \mid_{\iota_0=\widehat{\iota}, \iota=\widehat{\iota}}$.

Sob H_0 , as estatísticas LR_{SK1} e LR_{SK2} são aproximadamente distribuídas como uma χ_p^2 com alto grau de precisão (SKOVGAARD, 2001, pág. 7). Diferentemente de LR_{SK2} , LR_{SK1} é sempre não-negativa, além de ter como caso particular a estatística R^{*2} , quando se testa apenas um parâmetro em H_0 , ou seja, $p = 1$, sendo R^* uma versão modificada da estatística de verossimilhança sinalizada (BARNDORFF-NIELSEN, 1986, 1991). Sob H_0 , R^* tem distribuição normal padrão com erro da ordem de $n^{-3/2}$.

Ferrari e Cysneiros (2008) desenvolveram estatísticas da razão de verossimilhanças ajustadas segundo a proposta de Skovgaard (2001) para os modelos não-lineares da família exponencial. É de se destacar que, neste contexto, os resultados não se limitam a testes de heteroscedasticidade, mas abrangem testes de quaisquer hipóteses nulas que estabelecem um vetor de valores fixados para uma parte do vetor de parâmetros do modelo. Dentre todas as estatísticas apresentadas, àquelas que produziram melhores testes de heteroscedasticidade foram LR_{CR}^* e LR_{SK2} .

Todos os desenvolvimentos para a obtenção das quantidades nas equações (3.6) e (3.7) foram feitas no Apêndice C.

É de se destacar aqui, que os nossos resultados não se limitam apenas a testar $H_0 : \Psi_0 = 0$ mas abrangem testes de quaisquer hipóteses nulas que estabelecem um vetor de valores fixados para uma parte do vetor de parâmetros do modelo. Por exemplo, poderíamos testar a seguinte hipótese $H_0 : \phi_2 = \phi_2^{(0)}$, sendo $\phi_2^{(0)}$.

Observando a Equação (3.8), nota-se uma sequência de produtos de determinantes. Nesta sequência apenas os determinantes de $\widetilde{\mathbf{J}}_{\varphi\varphi}$ e $\widetilde{\mathbf{J}}_{\varphi\varphi}$ seriam diferentes, já que o vetor de perturbação para o novo teste de hipótese seria $\varphi = (\beta^\top, \phi_1)^\top$, em que $\beta = (\Psi_0^\top, \zeta^\top)$, embora as matrizes $\widetilde{\mathbf{J}}$ e $\widetilde{\mathbf{J}}$ se mantenham com a mesma estrutura já discutida. Continuando a análise, de toda a fração apenas LR seria alterado, ou seja,

$$LR = LR(\phi_2^{(0)}) = 2 \left[l_p(\widehat{\phi}_2) - l_p(\phi_2^{(0)}) \right],$$

em que $l_p(\phi_2) = L(\phi_2; \widehat{\Psi}_0(\phi_2), \widehat{\zeta}(\phi_2), \widehat{\phi}_1(\phi_2), \mathbf{Y})$ é a função de log-verossimilhança perfilada. Dessa forma, todos os desenvolvimentos no Apêndice C continuam úteis para executar este novo teste de hipótese.

3.2 APLICAÇÃO

Nesta seção, continuamos a analisar o exemplo descrito no Capítulo 1 com a finalidade de ilustrar a teoria apresentada nesta tese. É importante mencionar que DE BASTIANI et al. (2017) também analisou este conjunto de dados. Nele, analisaram que tipo de estrutura de covariância era a mais

adequada aos dados e as estruturas de covariâncias espaciais (Gaussiano, exponencial e Mátern com $\phi_4 = 1$) apresentaram menor AIC do que as estruturas de covariância espaço-temporais e portanto, são indicadas a serem utilizadas. Os autores acreditam que a independência temporal entre as repetições se deve ao fato de que o período de observação entre uma realização e outra é muito longo.

Considere o modelo na Equação 2.1, com as quatro variáveis explicativas P, K, Ca e Mg relacionadas à $\beta_1, \beta_2, \beta_3$ e β_4 , respectivamente, além do intercepto β_0 . A mesma função de covariância exponencial descrita no capítulo anterior foi selecionada.

O objetivo é efetuar diversos testes de hipóteses a dados reais utilizando as diversas estatísticas de teste apresentadas: razão de verossimilhanças usual (LR), razão de verossimilhança baseada no ajuste de Cox e Reid (1987), suas versões corrigidas por Bartlett (LR* e LR_{CR}*, respectivamente), além dos ajustes de Skovgaard (2001) (LR_{SK1} e LR_{SK2}).

A Tabela 2 apresenta um resumo dos testes de hipóteses executados, com a primeira coluna identificando a hipótese nula. Na tabela são encontrados os valores observados das estatísticas de teste e seu respectivo p-valor, entre parêntesis.

Para esses dados de produtividade de soja apenas o primeiro teste, $H_0 : \beta_1 = 0$, apresentou p-valores maiores que zero. Neste teste, apenas ao nível de significância 1% não rejeitamos H_0 , diferente dos níveis 5% e 10%, nos quais rejeitamos H_0 . Para todos os demais testes, rejeitamos a hipótese nula para todos os níveis de significância usuais 1%, 5% e 10%.

Tabela 2 – Testes de hipóteses usando diversas estatísticas baseadas na razão de verossimilhanças: usual (LR), baseada na verossimilhança perfilada ajustada (LR_{CR}), suas respectivas correções de Bartlett (LR^* e LR_{CR}^* e as versões ajustadas de Skovgaard (LR_{SK1} e LR_{SK2}), com p-valor (entre parêntesis).

H_0	LR	LR^*	LR_{CR}	LR_{CR}^*	LR_{SK1}	LR_{SK2}
$\beta_1 = 0$	4,96 (0,03)	4,74 (0,03)	4,15 (0,04)	4,37 (0,04)	4,06 (0,04)	4,02 (0,04)
$\beta_2 = 0$	26,23 (0,00)	26,16 (0,00)	25,74 (0,00)	26,11 (0,00)	25,59 (0,00)	25,59 (0,00)
$\beta_3 = 0$	15,56 (0,00)	15,28 (0,00)	14,81 (0,00)	15,23 (0,00)	14,75 (0,00)	14,74 (0,00)
$\beta_4 = 0$	22,61 (0,00)	21,85 (0,00)	20,47 (0,00)	21,38 (0,00)	20,39 (0,00)	20,33 (0,00)
$\beta_1 = \beta_2 = 0$	28,67 (0,00)	27,91 (0,00)	27,66 (0,00)	28,05 (0,00)	25,05 (0,00)	24,92 (0,00)
$\beta_1 = \beta_3 = 0$	22,73 (0,00)	22,08 (0,00)	21,26 (0,00)	21,83 (0,00)	20,26 (0,00)	20,19 (0,00)
$\beta_1 = \beta_4 = 0$	22,67 (0,00)	21,88 (0,00)	20,59 (0,00)	21,24 (0,00)	20,23 (0,00)	20,16 (0,00)
$\beta_2 = \beta_3 = 0$	33,09 (0,00)	32,17 (0,00)	30,02 (0,00)	30,65 (0,00)	29,24 (0,00)	29,12 (0,00)
$\beta_2 = \beta_4 = 0$	37,22 (0,00)	35,97 (0,00)	33,71 (0,00)	34,12 (0,00)	33,38 (0,00)	33,27 (0,00)
$\beta_3 = \beta_4 = 0$	29,47 (0,00)	28,86 (0,00)	26,98 (0,00)	27,36 (0,00)	25,81 (0,00)	25,69 (0,00)
$\beta_1 = \beta_2 = \beta_3 = 0$	34,71 (0,00)	33,52 (0,00)	32,07 (0,00)	32,58 (0,00)	30,88 (0,00)	30,77 (0,00)
$\beta_1 = \beta_2 = \beta_4 = 0$	41,11 (0,00)	39,86 (0,00)	37,54 (0,00)	37,91 (0,00)	37,09 (0,00)	36,99 (0,00)
$\beta_2 = \beta_3 = \beta_4 = 0$	43,42 (0,00)	41,60 (0,00)	38,52 (0,00)	39,04 (0,00)	39,32 (0,00)	39,21 (0,00)
$\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$	47,16 (0,00)	45,92 (0,00)	41,44 (0,00)	41,82 (0,00)	42,83 (0,00)	42,73 (0,00)

4 RESULTADOS NUMÉRICOS

Neste capítulo, nosso objetivo é comparar o desempenho (ou a performance) dos diferentes testes apresentados anteriormente baseados nas seguintes estatísticas, a saber: razão de verossimilhanças usual (LR), razão de verossimilhanças perfiladas modificada (LR_{CR}), suas respectivas versões corrigidas (LR^* e LR_{CR}^*), razão de verossimilhanças ajustada (LR_{SK1}) e sua versão assintoticamente equivalente (LR_{SK2}). Os desempenhos dos testes são avaliados em função da proximidade das probabilidades de rejeição da hipótese nula, sendo esta verdadeira (probabilidade do erro tipo I), aos respectivos níveis nominais dos testes.

As simulações realizadas são baseadas no modelo linear espacial gaussiano com repetição apresentado em (2.1). A hipótese nula considerada é $H_0 : \Psi_0 = \Psi_0^{(0)}$, em que $\Psi_0^{(0)}$ é o vetor nulo p -dimensional.

No modelo linear apresentado na equação (2.1), com matriz de covariância $\Sigma = \text{vec}(\Sigma_{(1)}, \dots, \Sigma_{(r)})$, sendo $\Sigma_{(i)}$, $i = 1, \dots, r$, tem forma dada pela Equação (2.2) e $R(\phi_3, \phi_4)$ dada pela Equação (2.3), a função de covariância exponencial. O valor dos parâmetros considerados no estudo foi $\beta_k = 0,00$ para $k = 1, \dots, p$ e $\beta_l = 0,50$ para $l = p + 1, \dots, q$, $\phi_1 = 0,55$ e $\phi_2 = 2,85$. Como ϕ_3 representa uma função do alcance da correlação como discutido no Capítulo 2 e por simplicidade, este parâmetro foi fixado como 30% da distância Euclidiana máxima. O número de repetições e tamanhos amostrais considerados foram, respectivamente, $r = \{5, 7, 9, 11, 13\}$ e $n = \{25, 36, 49\}$. Os tamanhos da amostra foram escolhidos de forma a termos uma grade regular de observações, ou seja, observações dispostas da forma $\sqrt{n} \times \sqrt{n}$. Essa não é uma condição necessária para utilização do modelo, pois as unidades amostrais podem ser observadas em qualquer ponto no espaço. Devido ao alto custo computacional, apenas para alguns casos específicos foi considerado também o número de repetições $r = 20$. O custo computacional reside no fato de vetores e matrizes de dimensões n serem impactados pelas r repetições, implicando em dimensões de ordem nr quando considerado um modelo matricial.

A matriz de covariáveis $X_{(1)}$ da ordem $n \times q$ para a primeira réplica foi gerada como amostras aleatórias da distribuição uniforme entre 0 e 1. Em seguida, consideramos a construção $\mathbf{X} = \text{vec}(X_{(1)}, X_{(2)}, \dots, X_{(r)})$, com $X_{(k)} = X_{(1)}, \forall k = 2, \dots, r$.

O número de réplicas considerado foi 10000 e os níveis nominais utilizados foram $\alpha = \{1\%, 5\%, 10\%\}$.

Para realizar as simulações foi utilizada a linguagem de programação matricial Ox (DOORNIK, 2009). Para cada tamanho de amostra e cada nível nominal considerado, calculamos as taxas de rejeição de cada teste, isto é, estimamos, via simulação, $\Pr[LR > x_\alpha]$, $\Pr[LR_{CR} > x_\alpha]$, $\Pr[LR^* > x_\alpha]$, $\Pr[LR_{CR}^* > x_\alpha]$, $\Pr[LR_{SK1} > x_\alpha]$ e $\Pr[LR_{SK2} > x_\alpha]$, em que x_α é o quantil $(1 - \alpha)$ apropriado da distribuição χ_p^2 . Todas as entradas das tabelas apresentadas são porcentagens. Ao final do Capítulo, realizamos a análise do poder dos testes, no qual consideramos uma perturbação δ , variando entre $\{0, 1; 1, 5\}$, em cada componente do vetor de interesse Ψ_0 .

Inicialmente, temos como objetivo analisar a influência do número de parâmetros de interesse, do número de parâmetros de perturbação, do tamanho da amostra e do número de réplicas nos desempenhos dos testes e compará-los entre si. Além disso, comparamos a média e a variância das estatísticas com as quantidades correspondentes da distribuição de referência.

A Tabela 3 apresenta resultados de simulação para $q - p = 2$ (parâmetros de perturbação), $r = 5$ (número de réplicas) e diferentes valores de n e p . Variamos p para analisar o efeito do número de parâmetros de interesse nos diferentes testes e variamos o n para avaliar o efeito do tamanho da amostra nos diferentes testes.

Para $n = 49$, observamos que o teste da razão de verossimilhanças usual (LR) é liberal. Por exemplo, quando $p = 5$ e $\alpha = 10\%$, a taxa de rejeição do teste é 10,48%. A tendência do teste em rejeitar com frequência demasiada a hipótese nula é atenuada pela correção de Bartlett de forma que o teste da razão de verossimilhanças corrigido (LR*) apresenta distorções de tamanho menores. Por exemplo, sua taxa de rejeição para a mesma situação é 9,86%.

O desempenho do teste baseado na estatística da razão de verossimilhanças perfiladas modificadas (LR_{CR}) é pior que a sua versão corrigida, a razão de verossimilhanças perfiladas modificadas corrigida (LR_{CR}*), que também apresenta desempenho bem melhor que o teste LR, pois sua taxa de rejeição está mais próxima do nível nominal. O desempenho dos testes LR* e LR_{CR}* são bem parecidos nesse cenário, as taxas mais próximas. Novamente, quando $p = 5$ e $\alpha = 10\%$, as taxas de rejeição de LR e LR* são, respectivamente, 9,65% e 9,82%.

Os testes baseados nas estatísticas da razão de verossimilhanças ajustadas LR_{SK1} e sua versão equivalente assintoticamente, LR_{SK2}, tiveram pior desempenho do que o teste LR, apresentando taxas de rejeição mais distantes dos níveis nominais, porém conservativas na maioria dos casos. Notamos que o impacto do número de parâmetros de interesse foi mais marcante nos testes baseados nas estatísticas LR, LR_{SK1} e LR_{SK2}. Para estes testes, quando o número de parâmetros de interesse aumenta, as taxas de rejeição desses testes permanecem mais instáveis em relação aos respectivos níveis nominais do que quando os testes são baseados nas outras estatísticas.

Em suma, para o estudo de simulação apresentado na Tabela 3, temos que os testes LR*, LR_{CR} e LR_{CR}* apresentaram melhores desempenhos.

As conclusões anteriores também se verificam para outros níveis nominais, até mesmo para níveis nominais muito pequenos ($\alpha = 1\%$). A Tabela 4 apresenta resultados de simulação para $p = 2$ (parâmetros de interesse), $r = 5$ (número de réplicas) e diferentes valores de n e $q - p$. Variamos $q - p$ para analisar o efeito do número de parâmetros de perturbação nos diferentes testes e variamos o n para avaliar o efeito do tamanho da amostra nos diferentes testes.

Podemos observar que quando temos poucos parâmetros de perturbação, observamos que o teste LR (baseado na estatística da razão de verossimilhanças original) apresenta taxas de rejeição acima do nível nominal considerado. Por exemplo, quando $\alpha = 10\%$, $q - p = 3$ e $n = 36$, a taxa de rejeição do teste LR é 11,18%. Já o teste LR* (baseado na estatística da razão de verossimilhanças corrigida, respectivamente) foi eficaz em reduzir a taxa de rejeição do teste LR exibindo taxa de rejeição menor e mais próxima do nível nominal considerado, a saber

Tabela 3 – Tamanhos dos testes para tamanhos de amostras $n = 36$ e $n = 49$, nº de repetições $r = 5$ e nº de β 's em perturbação $q - p = 2$, para diferentes níveis nominais usuais e parâmetros de interesse $p = \{2, 3, 4, 5, 6\}$.

$\alpha = 10\%$	$n = 36$						$n = 49$					
p	LR	LR*	LR _{CR}	LR* _{CR}	LR _{SK1}	LR _{SK2}	LR	LR*	LR _{CR}	LR* _{CR}	LR _{SK1}	LR _{SK2}
2	10,85	10,40	10,23	10,37	10,17	10,16	10,77	10,23	10,04	10,21	10,00	10,00
3	11,27	10,64	10,42	10,59	9,91	9,91	10,62	10,12	9,89	10,06	9,83	9,81
4	10,96	10,25	10,07	10,20	9,69	9,66	10,60	9,98	9,72	9,94	9,46	9,46
5	11,15	10,44	10,29	10,42	9,62	9,59	10,48	9,86	9,65	9,82	9,25	9,25
$\alpha = 5\%$	$n = 36$						$n = 49$					
p	LR	LR*	LR _{CR}	LR* _{CR}	LR _{SK1}	LR _{SK2}	LR	LR*	LR _{CR}	LR* _{CR}	LR _{SK1}	LR _{SK2}
2	5,78	5,44	5,33	5,46	5,20	5,19	5,42	5,15	5,10	5,15	5,10	5,10
3	5,64	5,27	5,16	5,28	4,83	4,81	5,55	5,26	5,15	5,25	5,03	5,03
4	5,75	5,32	5,23	5,25	4,81	4,80	5,47	5,06	4,98	5,09	4,73	4,72
5	5,79	5,32	5,18	5,27	4,63	4,62	5,22	4,78	4,69	4,78	4,44	4,41
$\alpha = 1\%$	$n = 36$						$n = 49$					
p	LR	LR*	LR _{CR}	LR* _{CR}	LR _{SK1}	LR _{SK2}	LR	LR*	LR _{CR}	LR* _{CR}	LR _{SK1}	LR _{SK2}
2	1,19	1,06	1,05	1,07	1,05	1,05	1,25	1,15	1,09	1,15	1,09	1,09
3	1,27	1,12	1,07	1,09	1,02	1,01	1,18	1,08	1,02	1,06	0,99	0,99
4	1,26	1,12	1,05	1,10	0,99	0,98	1,21	1,02	0,95	1,01	0,93	0,93
5	1,08	0,91	0,89	0,90	0,77	0,77	1,03	0,89	0,86	0,90	0,83	0,81

10,67%, embora ainda ligeiramente liberal. Os desempenhos dos testes LR_{CR} e LR*_{CR} (baseados nas estatísticas da razão de verossimilhanças perfiladas modificadas e razão de verossimilhanças perfiladas modificadas corrigidas) são bem melhores que os dos testes LR e LR* pois suas taxas de rejeição estão mais próximas dos níveis nominais. Para a mesma situação acima, temos que as taxas rejeição dos testes LR_{CR} e LR*_{CR} são 10,21% e 10,38%, respectivamente.

Dentre todos os testes considerados, os testes LR_{SK1} e LR_{SK2} (baseados nas estatísticas ajustadas de Skovgaard) foram os que apresentaram, em geral, taxas de rejeição mais próximas dos níveis nominais. Para a mesma situação acima, temos que as taxas de rejeição dos testes são 9,94% (LR_{SK1}) e 9,93% (LR_{SK2}).

É interessante notar que quando aumentamos o número de parâmetros de perturbação o impacto é bem mais marcante nos testes LR e LR* do que para os testes LR_{CR}, LR*_{CR}, LR_{SK1} e LR_{SK2}. Também observa-se (nestas situações apresentadas) que o aumento no número de parâmetros de perturbação impacta no desempenho de LR*_{CR}, que não foi efetiva, a deixando mais liberal quando comparada com LR_{CR}.

Além disso, notamos também que para o número de parâmetros de perturbação fixo, à medida que aumentamos o tamanho da amostra, quase todos os testes têm taxas de rejeição mais próximas dos níveis nominais, excetuando os casos dos testes LR_{SK1} e LR_{SK2} nos níveis nominais $\alpha = 10\%$ e $\alpha = 5\%$. Os testes LR e LR* ainda são liberais, apresentando taxas de rejeição acima dos níveis nominais.

Os mesmos comportamentos podem ser verificados para os demais níveis de significância.

Em suma, para o estudo de simulação apresentado na Tabela 4, temos que os testes LR_{SK1}, LR_{SK2}, LR_{CR} e LR*_{CR} apresentaram melhores desempenhos, nesta ordem.

Na Tabela 5 apresentamos comparações da média e da variância das estatísticas LR, LR*, LR_{CR}, LR*_{CR}, LR_{SK1}, LR_{SK2} e da distribuição χ^2_p , para o caso em que $q - p = 2$ (número de parâmetros de perturbação), $r = 5$ (número de réplicas) e diferentes valores de p (número de

Tabela 4 – Tamanhos dos testes para tamanhos de amostras $n = 36$ e $n = 49$, nº de repetições $r = 5$ e nº de β 's de interesse $p = 2$, para diferentes níveis nominais usuais e parâmetros de perturbação $q - p = \{2, 3, 4\}$.

$\alpha = 10\%$	$n = 36$						$n = 49$					
q	LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}	LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}
2	10,85	10,40	10,23	10,37	10,17	10,16	10,77	10,23	10,04	10,21	10,00	10,00
3	11,18	10,67	10,21	10,38	9,94	9,93	11,02	10,41	10,00	10,12	9,86	9,85
4	11,37	10,81	10,32	10,43	9,73	9,72	11,31	10,72	10,08	10,22	9,71	9,71
$\alpha = 5\%$	$n = 36$						$n = 49$					
q	LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}	LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}
2	5,78	5,44	5,33	5,46	5,20	5,19	5,42	5,15	5,10	5,15	5,10	5,10
3	5,90	5,56	5,29	5,41	5,01	4,98	5,68	5,32	5,03	5,11	4,97	4,96
4	6,23	5,78	5,43	5,51	4,92	4,87	5,91	5,51	5,07	5,17	4,88	4,88
$\alpha = 1\%$	$n = 36$						$n = 49$					
q	LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}	LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}
2	1,19	1,06	1,05	1,07	1,05	1,05	1,20	1,08	1,05	1,09	1,09	1,09
3	1,28	1,18	1,02	1,05	0,98	0,96	1,26	1,16	1,08	1,13	1,01	1,00
4	1,41	1,32	1,09	1,12	0,89	0,87	1,37	1,28	1,10	1,16	0,97	0,96

parâmetros de interesse) e n .

Para $n = 25$, os resultados da Tabela 5 mostram que a medida que aumentamos o número de parâmetros de interesse, as estatísticas LR_{CR}*, LR_{CR}, LR*, apresentam médias e variâncias mais próximas da média e da variância da distribuição de referência (χ_p^2) do que as estatísticas LR_{SK1}, LR_{SK2} e LR, nesta ordem. Por outro lado, essas medidas para LR excedem as da distribuição χ_p^2 , enquanto que essas medidas para LR_{SK1} e LR_{SK2}, ficaram abaixo às da distribuição χ_p^2 .

Ainda na mesma tabela, para $n = 36$, os resultados mostram que a medida que aumentamos o número de parâmetros de interesse, as estatísticas LR_{CR}, LR_{CR}* e LR*, LR_{SK1}, LR_{SK2} e LR, nesta ordem, apresentam médias e variâncias mais próximas da média e da variância da distribuição de referência (χ_p^2) do que a estatística LR. Essas medidas para LR excedem às da distribuição χ_p^2 .

Por fim, para $n = 49$, os resultados mostram que a medida que aumentamos o número de parâmetros de interesse, as estatísticas LR_{CR}, LR_{SK1} e LR_{SK2}, nesta ordem, apresentam médias e variâncias mais próximas da média e da variância da distribuição de referência (χ_p^2) do que a estatística LR_{CR}*, LR* e LR. Essas medidas para LR excedem às da distribuição χ_p^2 .

A Tabela 6 apresenta resultados de simulação para $n = 36$, $q - p = 2$ e diferentes valores de r e p . Variamos r para analisar o efeito do número de réplicas nos diferentes testes e variamos p para analisar o efeito do número de parâmetros de interesse.

Podemos observar que para um número pequeno de réplicas, o teste LR apresenta-se bem liberal, isto é, apresentando taxas de rejeição acima do nível nominal considerado. Especialmente quando o número de parâmetros de interesse diminui. Por exemplo, quando $\alpha = 10\%$ e $r = 5$, a taxa de rejeição do teste é superior a 10%, isto é, 11,27%. O teste LR* foi eficaz em reduzir a taxa de rejeição do teste LR exibindo taxa de rejeição menor e mais próxima do nível nominal considerado, a saber 10,64%. Observamos também que o teste LR_{CR} apresenta taxas de rejeição mais próxima do nível nominal do que a sua versão corrigida (LR_{CR}*). Para a situação acima, as taxas de rejeição dos testes são 10,42% (LR_{CR}) e 10,59% (LR_{CR}*). É interessante notar que dentre todos os testes, os testes LR_{SK1} e LR_{SK2} foram os que apresentaram, de forma geral,

Tabela 5 – Média e variância para tamanhos de amostras $n = \{25, 36, 49\}$, nº de repetições $r = 5$, nº de β 's em perturbação $q - p = 2$, para diferentes níveis nominais usuais e parâmetros de interesse $p = \{2, 3, 4, 5, 6\}$.

n	p	média							variância						
		χ_p^2	LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}	χ_p^2	LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}
25	2	2	1,98	1,92	1,92	1,91	1,83	1,80	4	3,81	3,57	3,57	3,55	3,58	3,51
	3	3	3,05	2,95	2,94	2,94	2,91	2,89	6	6,07	5,69	5,63	5,65	5,78	5,77
	4	4	4,07	3,94	3,92	3,93	3,89	3,88	8	8,10	7,60	7,50	7,54	7,52	7,52
	5	5	5,18	5,01	4,97	4,99	4,71	4,68	10	10,62	9,95	9,77	9,87	9,61	9,60
	6	6	6,24	6,04	5,99	6,02	5,67	5,63	12	12,95	12,11	11,89	12,02	11,76	11,67
36	2	2	2,08	2,03	2,02	2,03	1,90	1,87	4	4,29	4,12	4,05	4,10	3,98	3,95
	3	3	3,12	3,05	3,03	3,04	2,95	2,93	6	6,52	6,24	6,14	6,21	6,02	6,00
	4	4	4,09	4,00	3,98	4,00	3,97	3,96	8	8,52	8,16	8,07	8,13	7,99	7,98
	5	5	5,14	5,03	5,00	5,02	4,98	4,97	10	10,60	10,15	10,03	10,11	9,98	9,97
	6	6	6,16	6,02	5,98	6,01	5,99	5,99	12	12,62	12,07	11,91	12,02	11,99	11,99
49	2	2	2,09	2,05	2,03	2,05	1,99	1,98	4	4,30	4,14	4,08	4,13	4,04	4,04
	3	3	3,05	2,99	2,97	2,99	2,94	2,92	6	6,37	6,15	6,06	6,14	6,03	6,02
	4	4	4,10	4,02	4,00	4,02	3,96	3,95	8	8,33	8,04	7,94	8,03	8,01	8,00
	5	5	5,09	5,00	4,97	4,99	4,98	4,98	10	10,17	9,82	9,71	9,81	9,96	9,96
	6	6	6,13	6,02	6,00	6,02	5,99	5,99	12	12,18	11,76	11,64	11,74	11,93	11,93

taxas de rejeição mais próximas do nível nominal, a saber 9,91% (LR_{SK1}) e 9,91% (LR_{SK2}). É interessante notar que quando aumentamos o número de réplicas as taxas de rejeição de todos os testes aproximam-se do nível nominal, como era o esperado.

Vale ressaltar aqui também que o aumento do número de parâmetros de interesse fez com que todos os testes tivessem taxas de rejeição mais próximas do nível nominal, embora o teste LR ainda apresente uma taxa de rejeição acima do nível nominal. O mesmo comportamento pode ser verificado para os demais níveis de significância.

Em suma, para o estudo de simulação apresentado na Tabela 6, temos que os testes LR_{SK1}, LR_{SK2}, LR_{CR}, LR_{CR}* e LR* apresentaram melhores desempenhos, nesta ordem.

Tabela 6 – Tamanhos dos testes para tamanho de amostra $n = 36$, nº de repetições $r = \{5, 7, 9\}$ e nº de β 's em perturbação $q - p = 2$, para diferentes níveis nominais usuais e parâmetros de interesse $p = \{3, 4\}$.

r	$p=3$							$p=4$						
	$\alpha=10\%$							$\alpha=10\%$						
	LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}		LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}	
5	11,27	10,64	10,42	10,59	9,91	9,91		10,96	10,25	10,07	10,20	9,69	9,66	
7	10,78	10,45	10,24	10,36	10,01	10,01		10,67	10,17	10,04	10,14	9,65	9,65	
9	10,58	10,13	10,05	10,08	10,02	10,01		10,33	10,02	10,02	10,08	9,79	9,79	
5	$\alpha=5\%$							$\alpha=5\%$						
	LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}		LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}	
	5,64	5,27	5,16	5,28	4,83	4,81		5,75	5,32	5,23	5,25	4,81	4,80	
7	5,73	5,52	5,41	5,45	5,17	5,17		5,30	5,01	4,95	5,01	4,76	4,76	
9	5,18	4,99	5,02	5,05	4,98	4,98		5,17	4,98	4,91	4,99	4,90	4,89	
5	$\alpha=1\%$							$\alpha=1\%$						
	LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}		LR	LR*	LR _{CR}	LR _{CR} *	LR _{SK1}	LR _{SK2}	
	1,27	1,12	1,07	1,09	1,02	1,01		1,26	1,12	1,05	1,10	0,99	0,98	
7	1,17	1,06	1,04	1,04	1,03	1,03		1,14	1,05	1,07	1,10	0,96	0,96	
9	1,11	1,02	1,01	1,01	1,02	1,01		1,10	0,99	1,03	1,03	0,98	0,98	

A Tabela 7 apresenta resultados de simulação para $n = 25$, $q - p = 2$, $p = 4$ e diferentes valores de r . Variamos r para analisar o efeito do número de repetições nos diferentes testes.

Podemos observar que para um número pequeno de repetições, o teste LR apresenta taxas de rejeição mais próximas dos níveis nominais enquanto que os demais testes são conservativos.

Por exemplo, quando $\alpha = 10\%$ e $r = 5$, o teste LR tem taxa de rejeição igual a 10,35% enquanto que os testes LR^* , LR_{CR} , LR_{CR}^* , LR_{SK1} e LR_{SK2} apresentam as seguintes taxas de rejeição 9,25, 9,10, 9,18, 8,25 e 8,24, respectivamente. Porém, conforme vai aumentando o número de réplicas, observamos que todos os testes têm taxas de rejeição mais próximas dos níveis nominais. Novamente, para $\alpha = 10\%$ e $r = 20$ as taxas de rejeição dos testes LR, LR^* , LR_{CR} , LR_{CR}^* , LR_{SK1} e LR_{SK2} , são respectivamente iguais a, 10,06, 9,97, 9,94, 10,04, 9,94 e 9,94. Ressaltamos aqui que não há um favoritismo entre os testes LR_{CR} e LR_{CR}^* , embora o teste LR_{CR}^* teve uma taxa de rejeição mais próxima do nível nominal considerado.

Observamos também que o impacto do número de réplicas para estes testes, quando o número de réplicas aumenta, as taxas de rejeição dos testes LR, LR^* , LR_{CR} e LR_{CR}^* permanecem mais estáveis em relação aos respectivos níveis nominais do que quando os testes são baseados nas estatísticas LR_{SK1} e LR_{SK2} .

O mesmo comportamento pode ser verificado para os demais níveis de significância.

Tabela 7 – Tamanhos dos testes para tamanhos de amostras $n = 25$, nº de repetições $r = \{5, 7, 9, 11, 13, 20\}$ e nº de β 's de interesse $p = 4$, para diferentes níveis nominais usuais e parâmetros de perturbação $q - p = 2$.

r	$\alpha = 10\%$					
	LR	LR^*	LR_{CR}	LR_{CR}^*	LR_{SK1}	LR_{SK2}
5	10,35	9,25	9,10	9,18	8,25	8,24
7	10,28	9,52	9,45	9,51	8,76	8,73
9	10,75	10,19	10,09	10,21	9,53	9,52
11	10,19	9,89	9,81	9,92	9,62	9,61
13	10,15	9,91	9,90	10,02	9,70	9,69
20	10,06	9,97	9,94	10,04	9,94	9,94
r	$\alpha = 5\%$					
	LR	LR^*	LR_{CR}	LR_{CR}^*	LR_{SK1}	LR_{SK2}
5	5,18	4,63	4,51	4,56	3,96	3,95
7	5,05	4,58	4,50	4,55	4,13	4,12
9	5,39	4,98	4,92	4,95	4,61	4,59
11	5,11	4,89	4,75	4,87	4,53	4,52
13	5,08	4,94	4,81	4,95	4,72	4,72
20	5,01	4,97	4,91	5,01	4,92	4,92
r	$\alpha = 1\%$					
	LR	LR^*	LR_{CR}	LR_{CR}^*	LR_{SK1}	LR_{SK2}
5	1,09	0,91	0,92	0,92	0,71	0,70
7	1,13	0,97	0,98	0,98	0,79	0,79
9	1,22	1,10	1,10	1,10	0,98	0,98
11	1,15	1,08	1,08	1,08	0,84	0,84
13	1,10	1,06	1,06	1,06	0,89	0,89
20	1,06	1,02	1,03	1,03	0,96	0,96

A Tabela 8 apresenta resultados de simulação obtidos sob a hipótese alternativa para $n = 25$, $r = 7$, $p = 2$, $q - p = 2$, $\alpha = 10\%$ e diferentes valores de $\delta_1 = \delta_2 = \delta$ com δ variando de 0,1 a 1,5. São desconsiderados os testes que nas simulações de tamanho mostraram-se liberais (ou conservativos), ou seja, aqueles que apresentaram, sob H_0 , taxas de rejeição superiores (ou inferiores) aos níveis de significância nominal. Os testes desconsiderados são baseados nas estatísticas LR, LR_{SK1} e LR_{SK2} . Então, comparamos os poderes dos testes e observamos que eles

são bem semelhantes. Notamos uma leve vantagem para os testes baseados nas estatísticas LR_{CR} e LR_{CR}^* , uma vez que apresentaram menor distorção de tamanho.

Tabela 8 – Poder dos testes para tamanho de amostra $n = 25$, nº de repetição $r = 7$ e nº de β 's de interesse $p = 2$ e de perturbação $q - p = 2$, para diferentes níveis nominais usuais e diferentes valores de δ .

δ	$\alpha = 10\%$		
	LR^*	LR_{CR}	LR_{CR}^*
0,1	10,64	10,60	10,62
0,2	14,31	14,20	14,20
0,3	20,80	20,74	20,78
0,4	30,31	30,12	30,14
0,5	41,38	41,29	41,35
0,6	53,61	53,46	53,48
0,7	65,83	65,65	65,72
0,8	76,70	76,47	76,51
0,9	85,96	85,73	85,74
1,0	92,03	92,05	92,05
1,5	99,91	99,91	99,91

5 CONSIDERAÇÕES FINAIS

Resumimos as principais contribuições teóricas desta tese nos seguintes itens:

- (i) Apresentamos no Capítulo 2 uma breve introdução ao modelo linear espacial Gaussiano com repetição (MLEGr), métodos de estimação, aspectos inferenciais e dois testes a saber: o teste da razão de verossimilhanças e o teste da razão de verossimilhanças perfiladas modificadas.
- (ii) No Capítulo 3, revisamos a obtenção do fator de correção de Bartlett derivado por DE BASTIANI (2016) para a estatística da razão de verossimilhanças no MLEGr. Além disso, neste trabalho derivamos uma expressão para o fator de correção de Bartlett (proposto por DiCiccio e Stern (1994)) para a estatística da razão de verossimilhanças baseada em uma verossimilhança perfilada modificada proposta por Cox e Reid (1987) no MLEGr. Obtemos então um teste da razão de verossimilhanças perfiladas modificadas corrigido, que é preciso até segunda ordem. A fórmula da correção é dada, em notação matricial, e pode ser implementada facilmente em qualquer sistema de computação algébrica. Apresentamos uma aplicação a dados reais.
- (iii) No Capítulo 3, foi utilizada a abordagem da verossimilhança perfilada modificada proposta por Cox e Reid (1987), que requer uma parametrização ortogonal entre os parâmetros de interesse e os de perturbação. Desenvolvemos estatísticas da razão de verossimilhanças ajustadas baseadas na proposta de Skovgaard (2001) para o MLEGr. Vale salientar aqui que, neste contexto, nossos resultados não se limitam apenas em testar um subconjunto do vetor de parâmetros β , mas abrangem testes de quaisquer hipóteses nulas que estabeleçam um vetor de valores fixados para uma parte do vetor de parâmetros do modelo. Isto é, podemos estar interessados em testar um subconjunto do vetor de parâmetros γ . Apresentamos uma aplicação a dados reais.
- (iv) No Capítulo 4, realizamos um extenso estudo de simulação de Monte Carlo para analisar o desempenho dos diferentes testes apresentados na literatura e neste trabalho. Os desempenhos dos testes são avaliados em função da proximidade das probabilidades de rejeição da hipótese nula, sendo esta verdadeira (probabilidade do erro tipo I), aos respectivos níveis nominais dos testes. Analisamos também a influência do número de parâmetros de interesse, do número de parâmetros de perturbação, do tamanho da amostra e do número de réplicas nos desempenhos dos testes. Além disso, comparamos a média e a variância das estatísticas com as quantidades correspondentes da distribuição χ^2 de referência e realizamos também um estudo do poder dos testes.

Além dessas contribuições, concluímos o seguinte:

- a) Em relação aos estudos de simulação dos tamanhos dos testes concluímos que os testes LR^* , LR_{CR} e LR_{CR}^* apresentaram melhores desempenhos do que os demais testes.
- b) Em relação aos estudos de simulação realizados com o objetivo de analisar a influência do número de parâmetros de perturbação no desempenho dos testes, concluímos que os testes LR_{SK1} , LR_{SK2} , LR_{CR} e LR_{CR}^* apresentaram melhores desempenhos em relação aos demais testes.
- c) Em relação aos estudos de simulação realizados com o objetivo de analisar a influência do número de parâmetros de interesse no desempenho dos testes, concluímos que os testes LR^* , LR_{CR} e LR_{CR}^* apresentaram melhores desempenhos em relação aos outros testes.
- d) Em relação aos estudos de simulação realizados com o objetivo de analisar a da variação do número de réplicas, os testes LR_{CR} e LR_{CR}^* apresentaram o melhores desempenhos.
- e) Nas simulações envolvendo o poder dos testes, os testes LR_{CR} e LR_{CR}^* apresentaram vantagens em relação ao teste LR^* , uma vez que apresentaram menor distorção de tamanho.

Em suma, os testes ajustados apresentam desempenhos melhores em amostras de tamanho pequeno do que os testes originais.

REFERÊNCIAS

- AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: *Selected papers of hirotugu akaike*. [S.l.]: Springer, 1998. p. 199–213.
- ARAÚJO, M. C.; CYSNEIROS, A. H.; MONTENEGRO, L. C. Improved heteroskedasticity likelihood ratio tests in symmetric nonlinear regression models. *Statistical Papers*, Springer, p. 1–22, 2017.
- ATKINSON, P. Simulating locational error in field-based measurements of reflectance. In: *geoENV I—Geostatistics for Environmental Applications*. [S.l.]: Springer, 1997. p. 297–308.
- BARNDORFF-NIELSEN; E. HALL, P. O. On the level-error after bartlett adjustment of the likelihood ratio statistic. *Biometrika*, 1988.
- BARNDORFF-NIELSEN, O. On a formula for the distribution of the maximum likelihood estimator. *Biometrika*, Oxford University Press, v. 70, n. 2, p. 343–365, 1983.
- BARTLETT, M. S. Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences*, The Royal Society London, v. 160, n. 901, p. 268–282, 1937.
- CHEN, H.-S.; SIMPSON, D. G.; YING, Z. Infill asymptotics for a stochastic process model with measurement error. *Statistica Sinica*, JSTOR, p. 141–156, 2000.
- CHILES, J.-P.; DELFINER, P. et al. Modeling spatial uncertainty. *Geostatistics, Wiley series in probability and statistics*, 1999.
- CORDEIRO, G. M. On the corrections to the likelihood ratio statistics. *Biometrika*, Oxford University Press, v. 74, n. 2, p. 265–274, 1987.
- CORDEIRO, G. M.; CRIBARI-NETO, F. *An introduction to Bartlett correction and bias reduction*. [S.l.]: Springer, 2014.
- COX, D.; REID, N. A note on the calculation of adjusted profile likelihood. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 55, n. 2, p. 467–471, 1993.
- COX, D. R.; HINKLEY, D. V. *Theoretical statistics*. [S.l.]: CRC Press, 1979.
- COX, D. R.; REID, N. Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 49, n. 1, p. 1–18, 1987.
- CYSNEIROS, A. H.; FERRARI, S. L. An improved likelihood ratio test for varying dispersion in exponential family nonlinear models. *Statistics & probability letters*, Elsevier, v. 76, n. 3, p. 255–265, 2006.
- DE BASTIANI, F. *Inference and diagnostics in spatial models*. Tese (Doutorado) — Universidade Federal de Pernambuco, 2016.

- DE BASTIANI, F.; CYSNEIROS, A. H. M. de A.; URIBE-OPAZO, M. A.; GALEA, M. Influence diagnostics in elliptical spatial linear models. *Test*, Springer, v. 24, n. 2, p. 322–340, 2015.
- DE BASTIANI, F.; GALEA, M.; CYSNEIROS, A.; URIBE-OPAZO, M. Gaussian spatial linear models with repetitions: An application to soybean productivity. *Spatial Statistics*, Elsevier, v. 21, p. 319–335, 2017.
- DICICCIO, T. J.; STERN, S. E. Frequentist and bayesian bartlett correction of test statistics based on adjusted profile likelihoods. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 56, n. 2, p. 397–408, 1994.
- DOORNIK, J. A. An object-oriented matrix programming language ox 6. 2009.
- FERRARI, S. L.; CRIBARI-NETO, F. Corrected modified profile likelihood heteroskedasticity tests. *Statistics & probability letters*, Elsevier, v. 57, n. 4, p. 353–361, 2002.
- FERRARI, S. L.; CYSNEIROS, A. H. Skovgaard's adjustment to likelihood ratio tests in exponential family nonlinear models. *Statistics & probability letters*, Elsevier, v. 78, n. 17, p. 3047–3055, 2008.
- FERRARI, S. L.; CYSNEIROS, A. H.; CRIBARI-NETO, F. An improved test for heteroskedasticity using adjusted modified profile likelihood inference. *Journal of statistical planning and inference*, Elsevier, v. 124, n. 2, p. 423–437, 2004.
- FRASER, D.; REID, N. Ancillaries and third order significance. *Utilitas Mathematica*, UTILITAS MATHEMATICA PUBLISHING INCORPORATED, v. 7, p. 33–53, 1995.
- GOODALL, C.; MARDIA, K. V. Challenges in multivariate spatio-temporal modeling. In: *Proceedings of the XVIIth international biometric conference*. [S.l.: s.n.], 1994. p. 1–17.
- KIELLAND, P.; TUBMAN, T. On estimating map model errors and gps position errors: Applying more science to the art of navigation. *Navigation*, Wiley Online Library, v. 41, n. 4, p. 478–499, 1994.
- KRIGE, D. G. A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, Southern African Institute of Mining and Metallurgy, v. 52, n. 6, p. 119–139, 1951.
- KRUMBEIN, W.; SLACK, H. Statistical analysis of low-level radioactivity of pennsylvanian black fissile shale in illinois. *Geological Society of America Bulletin*, 1956.
- LAWLEY, D. N. A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika*, JSTOR, v. 43, n. 3/4, p. 295–303, 1956.
- MARDIA, K. V.; MARSHALL, R. J. Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika*, Oxford University Press, v. 71, n. 1, p. 135–146, 1984.
- MATÉRN, B. Spatial variation: Meddelanden fran statens skogsforskningsinstitut. *Lecture Notes in Statistics*, v. 36, p. 21, 1960.
- MATHERON, G. Principles of geostatistics. *Economic geology*, Society of Economic Geologists, v. 58, n. 8, p. 1246–1266, 1963.

- MELO, F. *Refinamentos para testes de hipóteses em modelos lineares mistos e modelos lineares com erros nas variáveis*. Tese (Doutorado) — Universidade de São Paulo, 2009.
- 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.
- PACE, L.; SALVAN, A. *Principles of statistical inference: from a Neo-Fisherian perspective*. [S.l.]: World scientific, 1997. v. 4.
- PETERSEN, K. B.; PEDERSEN, M. S. et al. The matrix cookbook. *Technical University of Denmark*, v. 7, n. 15, p. 510, 2008.
- PIRES, J. F.; CYSNEIROS, A. H.; CRIBARI-NETO, F. et al. Improved inference for the generalized pareto distribution. *Brazilian Journal of Probability and Statistics*, Brazilian Statistical Association, v. 32, n. 1, p. 69–85, 2018.
- RAO, C. R.; RAO, C. R.; STATISTIKER, M.; RAO, C. R.; RAO, C. R. *Linear statistical inference and its applications*. [S.l.]: Wiley New York, 1973. v. 2.
- SEIDEL, E. J.; OLIVEIRA, M. S. d. Novo índice geoestatístico para a mensuração da dependência espacial. *Revista Brasileira de Ciência do Solo*, SciELO Brasil, v. 38, n. 3, p. 699–705, 2014.
- SEVERINI, T. A. An empirical adjustment to the likelihood ratio statistic. *Biometrika*, Oxford University Press, v. 86, n. 2, p. 235–247, 1999.
- SEVERINI, T. A. *Likelihood methods in statistics*. [S.l.]: Oxford University Press, 2000.
- SKOVGAARD, I. M. Likelihood asymptotics. *Scandinavian Journal of Statistics*, Wiley Online Library, v. 28, n. 1, p. 3–32, 2001.
- SKOVGAARD, I. M. et al. An explicit large-deviation approximation to one-parameter tests. *Bernoulli*, Bernoulli Society for Mathematical Statistics and Probability, v. 2, n. 2, p. 145–165, 1996.
- STEIN, M. L. *Interpolation of spatial data: some theory for kriging*. [S.l.], 1999.
- URIBE-OPAZO, M. A.; BORSSOI, J. A.; GALEA, M. Influence diagnostics in gaussian spatial linear models. *Journal of Applied Statistics*, Taylor & Francis, v. 39, n. 3, p. 615–630, 2012.
- VENEZIANO, D.; DYCK, J. V. Statistical analysis of earthquake catalogs for seismic hazard. In: *Stochastic Approaches in Earthquake Engineering*. [S.l.]: Springer, 1987. p. 385–427.
- WEBSTER, R.; BUTLER, B. Soil classification and survey studies at ginninderra. *Soil Research*, CSIRO, v. 14, n. 1, p. 1–24, 1976.
- WHITTLE, P. On stationary processes in the plane. *Biometrika*, JSTOR, p. 434–449, 1954.
- WRIGHT, S.; NOCEDAL, J. Numerical optimization. *Springer Science*, v. 35, n. 67-68, p. 7, 1999.
- YAUDEN, W.; MEHLICH, A. Selection of efficient methods for soil sampling 1. *Soil Science Society of America Journal*, Soil Science Society of America, v. 2, n. C, p. 399–399, 1938.

ZHANG, H.; ZIMMERMAN, D. L. Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika*, Oxford University Press, v. 92, n. 4, p. 921–936, 2005.

ZHU, J.; EICKHOFF, J.; YAN, P. Generalized linear latent variable models for repeated measures of spatially correlated multivariate data. *Biometrics*, Wiley Online Library, v. 61, n. 3, p. 674–683, 2005.

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.

ANEXO A – CÁLCULOS PARA OBTENÇÃO DE C

Neste Apêndice estão descritos os passos para obtenção do fator de correção de Bartlett para a estatística de razão de verossimilhanças no modelo linear espacial Gaussiano com repetições. Este fator foi obtido por DE BASTIANI (2016).

Considere inicialmente o logaritmo da função de verossimilhança dada por

$$l = l(\boldsymbol{\iota}; \mathbf{Y}) = -\frac{nr}{2} \log(2\pi) - \frac{1}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2} \mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z},$$

em que $\boldsymbol{\iota} = (\boldsymbol{\Psi}_0, \boldsymbol{\zeta}, \boldsymbol{\phi})$ e $\mathbf{Z} = \mathbf{Y} - \mathbf{X}'_0 \boldsymbol{\Psi}_0 - \mathbf{X}_1 \boldsymbol{\zeta}$.

Considere o vetor $\boldsymbol{\iota} = (\boldsymbol{\Psi}_0^\top, \boldsymbol{\zeta}^\top, \boldsymbol{\phi}^\top)^\top$, em que ι_r é o r -ésimo elemento de $\boldsymbol{\iota}$. A notação de tensor foi adotada para os cumulantes da log-verossimilhança: $\kappa_{rs} = \mathbf{E}[\partial^2 l(\boldsymbol{\iota}) / \partial \iota_r \partial \iota_s]$, $\kappa_{rst} = \mathbf{E}[\partial^3 l(\boldsymbol{\iota}) / \partial \iota_r \partial \iota_s \partial \iota_t]$, $\dots$, e também para as derivadas dos cumulantes: $(\kappa_{rs})_t = \partial \kappa_{rs} / \partial \iota_t$, $(\kappa_{rs})_{tu} = \partial^2 \kappa_{rs} / \partial \iota_t \partial \iota_u$, $\dots$. A matriz de informação tem elementos $-\kappa_{rs}$, sendo elementos $-\kappa^{rs}$ elementos de sua inversa. $\tau^{rs} = \kappa_{rb} \kappa^{sa} \omega_{ab}$, sendo ω_{ab} elemento da matriz inversa de κ^{rs} , para $(r, s) \in \boldsymbol{\Psi}_0$. Os índices a, b, c, d em referencia aos componentes $\boldsymbol{\Psi}_0$, f, g em referencia $\boldsymbol{\zeta}$ e índices j, k, l, o para elementos $\boldsymbol{\phi}$.

Lawley (1956) desenvolveu uma fórmula para o fator de correção de Bartlett, c , que é dado por

$$c = \sum_{\boldsymbol{\Psi}_0, \boldsymbol{\zeta}, \boldsymbol{\phi}} (\lambda_{rstu} - \lambda_{rstuvw}) - \sum_{\boldsymbol{\zeta}, \boldsymbol{\phi}} (\lambda_{rstu} - \lambda_{rstuvw}),$$

em que

$$\lambda_{rstu} = \lambda^{rs} \lambda^{tu} \left[\frac{1}{4} \lambda_{rstu} - (\lambda_{rst})_u - (\lambda_{rt})_{su} \right],$$

e

$$\begin{aligned} \lambda_{rstuvw} = & \lambda^{rs} \lambda^{tu} \lambda^{vw} \left\{ \lambda_{rtv} \left(\frac{1}{6} \lambda_{suw} - (\lambda_{sw})_u \right) + \left(\frac{1}{4} \lambda_{svw} - (\lambda_{sw})_v \right) \right. \\ & \left. + (\lambda_{rt})_v (\lambda_{sw})_u + (\lambda_{rt})_u (\lambda_{sw})_v \right\}. \end{aligned}$$

A seguinte notação e resultados foram usados para obter c .

$$\begin{aligned}
\mathbf{Z} &= \mathbf{Y} - \mathbf{X}'_0 \boldsymbol{\Psi}_0 - \mathbf{X}_1 \boldsymbol{\zeta}, \\
\frac{\partial \mathbf{X}'_0}{\partial \boldsymbol{\Psi}_0} &= \mathbf{0}, \\
\frac{\partial \mathbf{X}'_0}{\partial \boldsymbol{\zeta}} &= \mathbf{0}, \\
\frac{\partial \mathbf{X}'_0}{\partial \phi_j} &= n X'_{0j} = -\mathbf{X}_1 (\mathbf{X}_1^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_1)^{-1} \mathbf{X}_1^\top \boldsymbol{\Sigma}_j^{-1} \mathbf{X}'_0, \\
\frac{\mathbf{X}'_{0j}}{\partial \phi_k} &= \mathbf{X}'_{1jk} = 2\mathbf{X}_1 (\mathbf{X}_1^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_1)^{-1} \mathbf{X}_1^\top \boldsymbol{\Sigma}_k^{-1} \mathbf{X}_1 (\mathbf{X}_1^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_1)^{-1} \mathbf{X}_1^\top \boldsymbol{\Sigma}_j^{-1} \mathbf{X}'_0 \\
&\quad - \mathbf{X}_1 (\mathbf{X}_1^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_1)^{-1} \mathbf{X}_1^\top \boldsymbol{\Sigma}_{jk}^{-1} \mathbf{X}'_0, \\
\mathbf{X}'_{0j}{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}'_0 &= \mathbf{0}, \\
\mathbf{X}_1^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}'_0 &= \mathbf{0}, \\
\mathbf{X}'_{0k}{}^\top \boldsymbol{\Sigma}_l^{-1} \mathbf{X}'_{0j} &= \mathbf{0}.
\end{aligned}$$

As derivadas de primeira ordem da log-verossimilhança são:

$$\begin{aligned}
\frac{\partial l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0} &= \mathbf{X}'_0{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \\
\frac{\partial l(\boldsymbol{\iota})}{\partial \boldsymbol{\zeta}} &= \mathbf{X}_1^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \\
\frac{\partial l(\boldsymbol{\iota})}{\partial \phi_j} &= -\frac{1}{2} \text{tr} [\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_j] + \boldsymbol{\Psi}_0^\top \mathbf{X}'_{0j}{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} - \frac{1}{2} \mathbf{Z}^\top \boldsymbol{\Sigma}_j^{-1} \mathbf{Z},
\end{aligned}$$

em que $\boldsymbol{\Sigma}_j = \partial \boldsymbol{\Sigma} / \partial \phi_j$, $\mathbf{X}'_{0j} = \partial \mathbf{X}'_0 / \partial \phi_j$ e $\boldsymbol{\Sigma}_j^{-1} = \partial \boldsymbol{\Sigma}^{-1} / \partial \phi_j$.

As derivadas de segunda ordem são

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\Psi}_0} &= -\mathbf{X}'_0{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}_0, \\
\frac{\partial^2 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\zeta}} &= \mathbf{0}, \\
\frac{\partial^2 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \phi_j} &= \mathbf{X}'_{0j}{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} + \mathbf{X}'_0{}^\top \boldsymbol{\Sigma}_j^{-1} \mathbf{Z}, \\
\frac{\partial^2 l(\boldsymbol{\iota})}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta}} &= -\mathbf{X}_1 \boldsymbol{\Sigma}^{-1} \mathbf{X}_1, \\
\frac{\partial^2 l(\boldsymbol{\iota})}{\partial \boldsymbol{\zeta} \partial \phi_j} &= \mathbf{X}_1^\top \boldsymbol{\Sigma}_j^{-1} (\mathbf{Y} - \mathbf{X}_1 \boldsymbol{\zeta}), \\
\frac{\partial^2 l(\boldsymbol{\iota})}{\partial \phi_j \partial \phi_k} &= -\frac{1}{2} \text{tr} [\boldsymbol{\Sigma}_j^{-1} \boldsymbol{\Sigma}_k] - \boldsymbol{\Psi}_0^\top \mathbf{X}'_{0k}{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}'_{0j} \boldsymbol{\Psi}_0 + \boldsymbol{\Psi}_0^\top (\mathbf{X}'_{0jk}{}^\top \boldsymbol{\Sigma}^{-1} + \mathbf{X}'_{0k}{}^\top \boldsymbol{\Sigma}_j^{-1} + \mathbf{X}'_{0j}{}^\top \boldsymbol{\Sigma}_k^{-1}) \mathbf{Z} \\
&\quad - \frac{1}{2} \mathbf{Z}^\top \boldsymbol{\Sigma}_{jk}^{-1} \mathbf{Z},
\end{aligned}$$

em que $\mathbf{X}'_{0jk} = \partial \mathbf{X}'_{0j} / \partial \phi_k$.

As derivadas de terceira ordem são

$$\begin{aligned}
\frac{\partial^3 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\Psi}_0} &= \mathbf{0}, \\
\frac{\partial^3 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\zeta}} &= \mathbf{0}, \\
\frac{\partial^3 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\Psi}_0 \partial \phi_j} &= -\mathbf{X}'_0{}^\top \boldsymbol{\Sigma}_j \mathbf{X}'_0, \\
\frac{\partial^3 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta}} &= \frac{\partial^3 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\zeta} \partial \phi_j} = \mathbf{0}, \\
\frac{\partial^3 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \phi_j \partial \phi_k} &= \left(\mathbf{X}'_{0jk}{}^\top \boldsymbol{\Sigma}^{-1} + \mathbf{X}'_{0j}{}^\top \boldsymbol{\Sigma}_k^{-1} + \mathbf{X}'_{0k}{}^\top \boldsymbol{\Sigma}_l^{-1} + \mathbf{X}'_0{}^\top \boldsymbol{\Sigma}_{jk}^{-1} \right) \mathbf{Z}, \\
\frac{\partial^3 l(\boldsymbol{\iota})}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta}} &= \mathbf{0}, \\
\frac{\partial^3 l(\boldsymbol{\iota})}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta} \partial \phi_j} &= -\mathbf{X}_1{}^\top \boldsymbol{\Sigma}_j^{-1} \mathbf{X}_1, \\
\frac{\partial^3 l(\boldsymbol{\iota})}{\partial \boldsymbol{\zeta} \partial \phi_j \partial \phi_k} &= -\mathbf{X}_1{}^\top \boldsymbol{\Sigma}_{jk}^{-1} (\mathbf{Y} - \mathbf{X}_1 \boldsymbol{\zeta}) \text{ e} \\
\frac{\partial^3 l(\boldsymbol{\iota})}{\partial \phi_j \partial \phi_k \partial \phi_l} &= -\frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}_{kl}^{-1} \boldsymbol{\Sigma}_j \right] - \frac{1}{2} \mathbf{Z}^\top \boldsymbol{\Sigma}_{jkl}^{-1} \mathbf{Z} + \boldsymbol{\Psi}_0{}^\top \left(\mathbf{X}'_{0lk}{}^\top \boldsymbol{\Sigma}_j^{-1} + \mathbf{X}'_{0k}{}^\top \boldsymbol{\Sigma}_{lj}^{-1} + \mathbf{X}'_{0jk}{}^\top \boldsymbol{\Sigma}_l^{-1} + \right. \\
&\quad \left. + \mathbf{X}'_{0lj}{}^\top \boldsymbol{\Sigma}_k^{-1} + \mathbf{X}'_{0j}{}^\top \boldsymbol{\Sigma}_{lk}^{-1} + \mathbf{X}'_{0l}{}^\top \boldsymbol{\Sigma}_{kj}^{-1} + \mathbf{X}'_{0jkl}{}^\top \boldsymbol{\Sigma}^{-1} \right) \mathbf{Z} - \frac{1}{2} \mathbf{Z}^\top \boldsymbol{\Sigma}_{jkl}^{-1} \mathbf{Z},
\end{aligned}$$

$\boldsymbol{\Sigma}_{jk}^{-1} = \partial \boldsymbol{\Sigma}^{-1} / \partial \phi_k$, $\boldsymbol{\Sigma}_{jkl}^{-1} = \partial \boldsymbol{\Sigma}_{jk}^{-1} / \partial \phi_l$ e $\mathbf{X}'_{0jkl} = \partial \mathbf{X}'_{0jk} / \partial \phi_l$.

As derivadas de quarta ordem são dadas por

$$\begin{aligned}
\frac{\partial^4 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\Psi}_0 \partial \phi_j \partial \phi_k} &= -2 \mathbf{X}'_{0k}{}^\top \boldsymbol{\Sigma}_j \mathbf{X}'_0 - \mathbf{X}'_0{}^\top \boldsymbol{\Sigma}_{jk} \mathbf{X}'_0, \\
\frac{\partial^4 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta}} &= \frac{\partial^4 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\zeta} \partial \phi_j} = \mathbf{0}.
\end{aligned}$$

Calculando as esperanças das derivadas obtidas acima, obtemos

$$\mathbf{K}_{\boldsymbol{\Psi}_0 \boldsymbol{\Psi}_0} = \mathbf{E} \left[\frac{\partial^2 l(\boldsymbol{\iota})}{\partial \boldsymbol{\Psi}_0 \partial \boldsymbol{\Psi}_0} \right] = -\mathbf{X}'_0{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}'_0,$$

e de maneira similar

$$\begin{aligned}
\mathbf{K}_{\Psi_0\zeta} &= \mathbf{0}, \\
\mathbf{K}_{\Psi_0\phi_j} &= \mathbf{0}, \\
\mathbf{K}_{\zeta\zeta} &= -\mathbf{X}_1^\top \Sigma^{-1} \mathbf{X}_1, \\
\mathbf{K}_{\zeta\phi_j} &= \mathbf{X}_1^\top \Sigma_j^{-1} \mathbf{X}'_0 \Psi_0, \\
\mathbf{K}_{\Psi_0\Psi_0\phi_j} &= -\mathbf{X}'_0{}^\top \Sigma_j^{-1} \mathbf{X}'_0, \\
\mathbf{K}_{\Psi_0\phi_j\phi_k} &= \mathbf{0}, \\
\mathbf{K}_{\zeta\zeta\phi_j} &= -\mathbf{X}_1^\top \Sigma_j^{-1} \mathbf{X}_1, \\
\mathbf{K}_{\zeta\phi_j\phi_k} &= \mathbf{X}_1^\top \Sigma_{jk}^{-1} \mathbf{X}'_0 \Psi_0, \\
\mathbf{K}_{\Psi_0\Psi_0\zeta\zeta} &= \mathbf{0}, \\
\mathbf{K}_{\Psi_0\Psi_0\zeta\phi_k} &= \mathbf{0}, \\
\mathbf{K}_{\Psi_0\Psi_0\phi_j\phi_k} &= -2\mathbf{X}'_0{}^\top \Sigma_j^{-1} \mathbf{X}'_0 - \mathbf{X}'_0{}^\top \Sigma_{jk}^{-1} \mathbf{X}'_0, \\
\mathbf{K}_{\phi_j\phi_k} &= -\frac{1}{2} \text{tr} [\Sigma_j^{-1} \Sigma_k] - \Psi_0 \mathbf{X}'_0{}^\top \Sigma_j^{-1} \mathbf{X}'_0 \Psi_0, \\
\mathbf{K}_{\phi_l\phi_j\phi_k} &= -2\text{tr} [\Sigma_l^{-1} \Sigma_k \Sigma^{-1} \Sigma_j] + \frac{1}{2} \text{tr} [\Sigma_j^{-1} \Sigma_{lk}^{-1}] + \frac{1}{2} \text{tr} [\Sigma_k^{-1} \Sigma_{lj}^{-1}] \\
&\quad + \frac{1}{2} \text{tr} [\Sigma_l^{-1} \Sigma_{jk}^{-1}].
\end{aligned}$$

Menos a matriz de informação de Fisher é dada por

$$\begin{pmatrix} \mathbf{K}_{\Psi_0\Psi_0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{\zeta\zeta} & \mathbf{K}_{\zeta\phi} \\ \mathbf{0} & \mathbf{K}_{\zeta\phi}^\top & \mathbf{K}_{\phi\phi} \end{pmatrix}.$$

Usando a regra de inversão de matrizes de Rao et al. (1973), as matrizes inversas formadas são:

$\mathbf{K}^{\Psi_0\Psi_0} = \mathbf{K}_{\Psi_0\Psi_0}^{-1}$, $\mathbf{K}^{\phi\phi} = (\mathbf{K}_{\phi\phi} - \mathbf{K}_{\zeta\phi}^\top \mathbf{K}_{\zeta\zeta}^{-1} \mathbf{K}_{\zeta\phi})^{-1}$, $\mathbf{K}^{\zeta\zeta} = \mathbf{K}_{\zeta\zeta}^{-1} + \mathbf{K}_{\zeta\zeta}^{-1} \mathbf{K}_{\zeta\phi} \mathbf{K}^{\phi\phi} \mathbf{K}_{\zeta\phi}^\top \mathbf{K}_{\zeta\zeta}^{-1}$ e $\mathbf{K}^{\zeta\phi} = -\mathbf{K}_{\zeta\zeta}^{-1} \mathbf{K}_{\zeta\phi} \mathbf{K}^{\phi\phi}$. As derivadas dos cumulantes são

$$\begin{aligned}
(\mathbf{K}_{\Psi_0\Psi_0})_j &= -\mathbf{X}'_0{}^\top \Sigma_j^{-1} \mathbf{X}'_0, \\
(\mathbf{K}_{\Psi_0\Psi_0})_{jk} &= -2\mathbf{X}'_0{}^\top \Sigma_j^{-1} \mathbf{X}'_0 \Psi_0 - \mathbf{X}'_0{}^\top \Sigma_{jk}^{-1} \mathbf{X}'_0, \\
(\mathbf{K}_{\zeta\zeta})_j &= -\mathbf{X}_1^\top \Sigma_j^{-1} \mathbf{X}_1, \\
(\mathbf{K}_{\zeta\phi_j})_k &= \mathbf{X}_1^\top \Sigma_{jk}^{-1} \mathbf{X}'_0 \Psi_0 + \mathbf{X}_1^\top \Sigma_j^{-1} \mathbf{X}'_0 \Psi_0, \\
(\mathbf{K}_{\phi_j\phi_l})_k &= -\text{tr} [\Sigma_l^{-1} \Sigma_j \Sigma^{-1} \Sigma_k] + \frac{1}{2} \text{tr} [\Sigma_l^{-1} \Sigma_{jk}] \\
&\quad + \frac{1}{2} \text{tr} [\Sigma_j^{-1} \Sigma_{kl}].
\end{aligned}$$

Da ortogonalidade entre Ψ_0 e $(\zeta^\top, \phi^\top)^\top$ temos $\kappa^{af} = \kappa^{aj} = (\kappa_{af})_{jb} = \kappa_{jfa} = \kappa_{jfab} = 0$. E portanto é possível mostrar que $\lambda_{abcd} = \lambda_{abfg} = \lambda_{abjf} = \lambda_{fgab} = 0$ e também $\kappa_{abjk} = (\kappa_{abj})_k =$

κ_{jkab} . Após algumas manipulações algébricas, o fator de correção de Bartlett reduz-se à

$$c = \sum \left[-\frac{1}{2} \kappa^{ab} \kappa^{jk} \kappa_{abjk} + \frac{1}{4} \kappa^{ab} \kappa^{cd} \kappa^{jk} \kappa_{abj} \kappa_{cdk} - \frac{1}{2} \kappa^{ab} \kappa^{jk} \kappa^{lo} \kappa_{abj} (\kappa_{lok} - 2(\kappa_{lo})_k) \right. \\ \left. + \kappa^{ab} \kappa^{fk} \kappa^{jl} \kappa_{abj} \left(2(\kappa_{fk})_l - \frac{3}{2} \kappa_{fkl} \right) - \frac{1}{2} \kappa^{ab} \kappa^{fg} \kappa^{jk} \kappa_{abj} \kappa_{fgk} \right].$$

Em notação matricial, o fator de correção é dado por

$$c = \text{tr} \left\{ \mathbf{K}^{\phi\phi} \left[-\frac{1}{2} \mathbf{M} + \frac{1}{4} \mathbf{P} - \left(\frac{1}{2} \boldsymbol{\varsigma} - \boldsymbol{\delta} + \frac{1}{2} \boldsymbol{\nu} \right) \boldsymbol{\tau}^\top \right] \right\},$$

em que $\boldsymbol{\varsigma}$, $\boldsymbol{\delta}$ e $\boldsymbol{\nu}$ são vetores de dimensão m cujo j -ésimo elementos são, respectivamente, $\text{tr} [\mathbf{K}^{\phi\phi} \mathbf{A}_j]$, $\text{tr} [\mathbf{K}^{\zeta\phi^\top} \mathbf{B}_j]$ e $\text{tr} [-\mathbf{K}^{\zeta\zeta} (\mathbf{X}_1^\top \boldsymbol{\Sigma}_j^{-1} \mathbf{X}_1)]$.

A matriz $\mathbf{B}_j$ contém o m -ésimo vetor coluna $\left(\frac{1}{2} \mathbf{X}_1^\top \boldsymbol{\Sigma}_{jk}^{-1} \mathbf{X}'_0 + 2 \mathbf{X}_1^\top \boldsymbol{\Sigma}_j^{-1} \mathbf{X}'_{0k} \right) \boldsymbol{\Psi}_0$. Estes resultados são similares aos obtidos por Melo, Ferrari e Cribari-Neto (2009).

ANEXO B – CÁLCULOS PARA OBTENÇÃO DE c_{CR}

Nesta seção estarão os cálculos necessários para obtenção de c_{CR} . Considere a expressão proposta por DiCiccio e Stern (DICICCIO; STERN, 1994) dada por

$$c_{CR} = \sum_{\Psi_0, \zeta, \phi} \left\{ \frac{1}{4} \tau^{ru} \tau^{st} \kappa_{rstu} - \kappa^{ru} \tau^{st} (\kappa_{rst})_u + (\kappa^{ru} \kappa^{st} - \nu^{ru} \nu^{st}) (\kappa_{rs})_{tu} - \right. \\ \left. - \left(\frac{1}{4} \kappa^{ru} \tau^{st} \tau^{vw} + \frac{1}{2} \kappa^{tu} \tau^{sw} \tau^{tv} - \frac{1}{3} \tau^{tu} \tau^{sw} \tau^{tv} \right) \kappa_{rst} \kappa_{uvw} + \right. \\ \left. + (\kappa^{ru} \kappa^{st} \kappa^{vw} + \kappa^{ru} \kappa^{sw} \kappa^{tv} - \nu^{ru} \kappa^{sw} \nu^{tv}) \kappa_{rst} (\kappa_{uv})_w - \right. \\ \left. - (\kappa^{tu} \kappa^{st} \kappa^{vw} - \nu^{ru} \nu^{st}) (\kappa_{rs})_t (\kappa_{uv})_w - (\kappa^{ru})_w - (\kappa^{ru} \kappa^{sw} \kappa^{vw} - \nu^{ru} \nu^{st} \nu^{vw}) (\kappa_{rs})_t (\kappa_{uv})_w \right\}, \quad (B.1)$$

em que $\nu^{rs} = \kappa^{rs} - \tau^{rs}$, $\tau^{rs} = \kappa_{rb} \kappa^{sa} \omega_{ab}$, sendo ω_{ab} elemento da matriz inversa $K^{\Psi_0 \Psi_0}$, conforme apresentado no apêndice anterior. Os índices r, s, t, u, v, w percorrem todos os componentes de $\iota = (\Psi_0, \zeta, \phi)$.

Considere inicialmente o logaritmo da função de verossimilhança dada por

$$l = l(\iota; \mathbf{Y}) = -\frac{nr}{2} \log(2\pi) - \frac{1}{2} \log|\Sigma| - \frac{1}{2} \mathbf{Z}^\top \Sigma^{-1} \mathbf{Z},$$

em que $\iota = (\Psi_0, \zeta, \phi)$ e $\mathbf{Z} = \mathbf{Y} - \mathbf{X}_0' \Psi_0 - \mathbf{X}_1 \zeta$.

As derivadas de primeira, segunda, terceira e quarta ordens da função de log-verossimilhança já foram apresentadas no Apêndice anterior A, não sendo necessário sua reprodução. Da mesma forma, as derivadas de cumulantes e algumas matrizes definidas anteriormente serão utilizadas neste apêndice.

Sejam a, b, c, d índices para se referir aos componentes de Ψ_0 , f e g ao componentes de ζ e j, k, l e m aos componentes de ϕ . Da ortogonalidade entre Ψ_0 e ϕ , temos que κ_{aj} , elementos da matriz $K_{\Psi_0 \phi}$, são 0 e portanto $\tau^{aj} = 0$. Note também que $\tau^{ab} = \kappa^{ab}$. Além disso

$$\nu^{ru} \nu^{tv} = (\kappa^{ru} - \tau^{ru})(\kappa^{tv} - \tau^{tv}) = \kappa^{ru} \kappa^{tv} + \kappa^{ru} \tau^{tv} - \tau^{ru} \kappa^{tv} + \tau^{ru} \tau^{tv},$$

e portanto $\kappa^{ru} \kappa^{tv} - \nu^{ru} \nu^{tv} = \tau^{ru} \kappa^{tv} - \tau^{ru} \tau^{tv}$. Ainda, $(\kappa_{ab})_k = \kappa_{abk}$ e $\kappa_{abcd} = \kappa_{abc} = \kappa_{abf} = (\kappa_{ab})_{bu} = (\kappa_{af})_{tu} = (\kappa_{aj})_{tu} = 0$. Assim, a equação (B.1) se resume à

$$c_{CR} = \sum \left\{ -\kappa^{ab} \kappa^{jk} \kappa_{abjk} + \frac{1}{4} \kappa^{ab} \kappa^{cd} \kappa^{jk} \kappa_{abj} \kappa_{cdk} + \kappa^{ab} \kappa^{jk} \kappa^{lo} \kappa_{abj} (\kappa_{kl})_o + \kappa^{ab} \kappa^{fj} \kappa^{gk} \kappa_{abj} \kappa_{fgk} + \right. \\ \left. + 2\kappa^{ab} \kappa^{fj} \kappa^{kl} \kappa_{abj} (\kappa_{fk})_l \right\},$$

que pode ser simplificado, em forma matricial, como

$$c_{CR} = \text{tr} \left[K^{\phi\phi} \left\{ -\mathbf{M} + \frac{1}{4} \mathbf{P} + (\mathbf{s}_1 + 2\delta^*) \boldsymbol{\gamma}^\top \right\} \right] + \boldsymbol{\gamma}^\top K^{\phi\zeta} \boldsymbol{\eta}^*,$$

em que γ^* é um vetor $\text{tr} [K^{\phi\phi} D^j]$, sendo a matriz $D^j = \Sigma_k^{-1} \Sigma_j \Sigma^{-1} \Sigma_l$, δ^* é um vetor definido como $\text{tr} [K^{\zeta\phi} F^j]$, sendo F^j a matriz definida como $(\mathbf{X}_1^\top \Sigma_{jk}^{-1} \mathbf{X}'_0 + \mathbf{X}_1^\top \Sigma_j^{-1} \mathbf{X}'_{0k}) \Psi_0$. O vetor η^* é definido como $\text{tr} [K^{\phi\zeta} G^j]$, sendo $G^j = -\mathbf{X}_1^\top \Sigma_j^{-1} \mathbf{X}_1$. No caso de se testar $H_0 : \Psi_0 = 0$, então

$$c_{CR} = \text{tr} \left[K^{\phi\phi} \left\{ -\mathbf{M} + \frac{1}{4} \mathbf{P} + \mathbf{s}_1 \boldsymbol{\gamma}^\top \right\} \right] + \boldsymbol{\gamma}^\top K^{\phi\zeta} \boldsymbol{\eta}^*.$$

ANEXO C – CÁLCULOS PARA OBTENÇÃO DE ρ

Nesta seção estarão os cálculos necessários para obtenção de ρ . Essa quantidade apresenta

$$\hat{h} = (h_{\Psi_0}^\top, h_{\zeta}^\top, h_{\phi}^\top) = \text{cov}_{\boldsymbol{\iota}_0} [U(\boldsymbol{\iota}_0), L(\boldsymbol{\iota}_0) - L(\boldsymbol{\iota})] |_{\boldsymbol{\iota}_0=\hat{\boldsymbol{\iota}}, \boldsymbol{\iota}=\hat{\boldsymbol{\iota}}} = E_{\boldsymbol{\iota}_0} [U(\boldsymbol{\iota}_0) (L(\boldsymbol{\iota}_0) - L(\boldsymbol{\iota}))] |_{\boldsymbol{\iota}_0=\hat{\boldsymbol{\iota}}, \boldsymbol{\iota}=\hat{\boldsymbol{\iota}}}, \quad (\text{C.1})$$

já que $E_{\boldsymbol{\iota}_0} [U(\boldsymbol{\iota}_0)] = 0$. Assim, procedemos a obtenção de $U(\boldsymbol{\iota}) = \partial l(\boldsymbol{\iota}) / \partial \boldsymbol{\iota}$ a partir de l dado na equação (2.4). Desta forma,

$$\begin{aligned} U_{\Psi_0} &= U_{\Psi_0}(\boldsymbol{\iota}) = \mathbf{X}'_0{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z}, \\ U_{\zeta} &= U_{\zeta}(\boldsymbol{\iota}) = \mathbf{X}_1^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \text{ e} \\ U_{\phi_j} &= U_{\phi_j}(\boldsymbol{\iota}) = -\frac{1}{2} \text{tr} [\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_j] - \frac{1}{2} \mathbf{Z}^\top \boldsymbol{\Sigma}_j^{-1} \mathbf{Z} + \boldsymbol{\Psi}_0^\top \mathbf{X}'_{0j}{}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \end{aligned}$$

em que $\mathbf{Z} = \mathbf{Y} - \mathbf{X}'_0 \boldsymbol{\Psi}_0 - \mathbf{X}_1 \boldsymbol{\zeta}$, sendo $\mathbf{X}'_0 = \mathbf{X}'_0(\boldsymbol{\phi})$ e $\mathbf{X}'_{0j} = \mathbf{X}'_{0j}(\boldsymbol{\phi}) = \partial \mathbf{X}'_0 / \partial \phi_j$. Importante observar também algumas esperanças que foram utilizadas. Considere antes, $\mathbf{K}$ e $\mathbf{F}$ duas matrizes sem relação com $\mathbf{Y}$, $\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\iota}) = \mathbf{X}'_0 \boldsymbol{\Psi}_0 + \mathbf{X}_1 \boldsymbol{\zeta}$, $\boldsymbol{\mu}_0 = \boldsymbol{\mu}(\boldsymbol{\iota}_0) = \mathbf{X}'_{00} \boldsymbol{\Psi}_{00} + \mathbf{X}_1 \boldsymbol{\zeta}_0$, sendo $\mathbf{X}'_{00} = \mathbf{X}'_0(\boldsymbol{\phi}_0)$ e $\mathbf{Z}_0 = \mathbf{Z}(\boldsymbol{\iota}_0) = \mathbf{Y} - \boldsymbol{\mu}_0$. Utilizamos algumas esperanças no decorrer dos desenvolvimentos. Elas foram obtidas de (PETERSEN; PEDERSEN et al., 2008) e algumas delas são

$$E_{\boldsymbol{\iota}_0} [\mathbf{K} U_{\boldsymbol{\iota}_0}(\boldsymbol{\iota}_0)] = \mathbf{K} E_{\boldsymbol{\iota}_0} [U_{\boldsymbol{\iota}_0}(\boldsymbol{\iota}_0)] = 0, \quad (\text{C.2})$$

$$E_{\boldsymbol{\iota}_0} [(\mathbf{K} \mathbf{Z}_0)^\top] = E_{\boldsymbol{\iota}_0} [\mathbf{Z}_0^\top] \mathbf{K}^\top = (\boldsymbol{\mu}_0 - \boldsymbol{\mu}_0)^\top \mathbf{K}^\top = \mathbf{0}, \quad (\text{C.3})$$

$$E_{\boldsymbol{\iota}_0} [(\mathbf{K} \mathbf{Z})^\top] = E_{\boldsymbol{\iota}_0} [\mathbf{Z}^\top] \mathbf{K}^\top = (\boldsymbol{\mu} - \boldsymbol{\mu}_0)^\top \mathbf{K}^\top, \quad (\text{C.4})$$

$$E_{\boldsymbol{\iota}_0} [\mathbf{Z}_0^\top \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_0] = \text{tr} [\boldsymbol{\Sigma}_0^{-1} \boldsymbol{\Sigma}_0] = N, \quad (\text{C.5})$$

$$E_{\boldsymbol{\iota}_0} [\mathbf{Z}^\top \mathbf{F} \mathbf{Z}] = \text{tr} [\mathbf{F} \boldsymbol{\Sigma}_0] + (\boldsymbol{\mu}_0 - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_0 - \boldsymbol{\mu}). \quad (\text{C.6})$$

Utilizamos também o intercâmbio entre derivada e integral, sob a suposição de que se trata de um caso de regularidade. Portanto, sendo $g(\mathbf{Y})$ uma função em $\mathbf{Y}$ e $l(\mathbf{Y}; \boldsymbol{\iota}_0) = \sum_i \ln [f(\mathbf{Y}; \boldsymbol{\iota}_0)]$ a função de log-verossimilhança de $\mathbf{Y}$, temos a relação

$$E_{\boldsymbol{\iota}_0} [g(\mathbf{Y})] = \int g(\mathbf{Y}) l(\mathbf{Y}; \boldsymbol{\iota}_0) d\mathbf{Y}$$

e sua derivada é

$$\frac{\partial E_{\boldsymbol{\iota}_0} [g(\mathbf{Y})]}{\partial \boldsymbol{\iota}_0} = \frac{\partial}{\partial \boldsymbol{\iota}_0} \int g(\mathbf{Y}) l(\mathbf{Y}; \boldsymbol{\iota}_0) d\mathbf{Y} = \int \frac{\partial}{\partial \boldsymbol{\iota}_0} \{g(\mathbf{Y}) l(\mathbf{Y}; \boldsymbol{\iota}_0)\} d\mathbf{Y} \quad (\text{C.7})$$

Para obtermos $\hat{h}$, separamos a esperança na equação (C.1) em duas, $E_{\boldsymbol{\iota}_0} [U(\boldsymbol{\iota}_0) L(\boldsymbol{\iota}_0)]$ e $E_{\boldsymbol{\iota}_0} [U(\boldsymbol{\iota}_0) L(\boldsymbol{\iota})]$. Importante ressaltar que estas esperanças ainda se resumem a

$$\mathbf{E}_{\boldsymbol{\iota}_0} [U(\boldsymbol{\iota}_0)L(\boldsymbol{\iota}_0)] = -\frac{1}{2} \mathbf{E}_{\boldsymbol{\iota}_0} \left[U(\boldsymbol{\iota}_0) \left(\mathbf{Z}_0^\top \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_0 \right) \right] \text{ e} \quad (\text{C.8})$$

$$\mathbf{E}_{\boldsymbol{\iota}_0} [U(\boldsymbol{\iota}_0)L(\boldsymbol{\iota})] = -\frac{1}{2} \mathbf{E}_{\boldsymbol{\iota}_0} \left[U(\boldsymbol{\iota}_0) \left(\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \right) \right]. \quad (\text{C.9})$$

Para resolvermos a equação (C.8), considere a equação (C.5) e a diferenciação em ambos os lados por $\boldsymbol{\Psi}_0$, ζ e ϕ

$$\begin{aligned} 0 &= \int \left(-2\mathbf{X}'_{00}{}^\top \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_0 \right) f(\mathbf{Y}; \boldsymbol{\iota}_0) d\mathbf{Y} + \int \left(\mathbf{Z}_0^\top \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_0 \right) \frac{\partial f(\mathbf{Y}; \boldsymbol{\iota}_0)}{\partial \boldsymbol{\Psi}_{00}} d\mathbf{Y} \\ 0 &= \int \left(-2\mathbf{X}_1^\top \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_0 \right) f(\mathbf{Y}; \boldsymbol{\iota}_0) d\mathbf{Y} + \int \left(\mathbf{Z}_0^\top \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_0 \right) \frac{\partial f(\mathbf{Y}; \boldsymbol{\iota}_0)}{\partial \zeta} d\mathbf{Y} \\ 0 &= \int \left(-2\boldsymbol{\Psi}_{00}^\top \mathbf{X}'_{00j}{}^\top \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_0 + \mathbf{Z}_0^\top \boldsymbol{\Sigma}_{0j}^{-1} \mathbf{Z}_0 \right) f(\mathbf{Y}; \boldsymbol{\iota}_0) d\mathbf{Y} + \int \left(\mathbf{Z}_0^\top \boldsymbol{\Sigma}_0^{-1} \mathbf{Z}_0 \right) \frac{\partial f(\mathbf{Y}; \boldsymbol{\iota}_0)}{\partial \phi_j} d\mathbf{Y}, \end{aligned}$$

das quais resultam, respectivamente,

$$\begin{aligned} \mathbf{E}_{\boldsymbol{\iota}_0} [U_{\boldsymbol{\Psi}_0}(\boldsymbol{\iota}_0)L(\boldsymbol{\iota}_0)] &= 0, \\ \mathbf{E}_{\boldsymbol{\iota}_0} [U_\zeta(\boldsymbol{\iota}_0)L(\boldsymbol{\iota}_0)] &= 0, \\ \mathbf{E}_{\boldsymbol{\iota}_0} [U_{\phi_j}(\boldsymbol{\iota}_0)L(\boldsymbol{\iota}_0)] &= -\frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}_{0j}^{-1} \boldsymbol{\Sigma}_0 \right], \end{aligned}$$

Da mesma forma, a solução para a equação (C.9) passa pela diferenciação de ambos os lados da equação (C.6), com relação a $\boldsymbol{\Psi}_0$, ζ e ϕ , sendo obtido

$$\begin{aligned} 2\mathbf{X}'_{00}{}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_0 - \boldsymbol{\mu}) &= \int \left(\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \right) \frac{\partial f(\mathbf{Y}; \boldsymbol{\iota}_0)}{\partial \boldsymbol{\Psi}_0} d\mathbf{Y} \\ 2\mathbf{X}_1^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_0 - \boldsymbol{\mu}) &= \int \left(\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \right) \frac{\partial f(\mathbf{Y}; \boldsymbol{\iota}_0)}{\partial \zeta} d\mathbf{Y}, \\ \text{tr} \left[\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{0j} \right] + 2(\mathbf{X}_{00j} \boldsymbol{\Psi}_{00})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_0 - \boldsymbol{\mu}) &= \int \left(\mathbf{Z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{Z} \right) \frac{\partial f(\mathbf{Y}; \boldsymbol{\iota}_0)}{\partial \phi_j} d\mathbf{Y}, \end{aligned}$$

das quais resultam, respectivamente

$$\begin{aligned} \mathbf{E}_{\boldsymbol{\iota}_0} [U_{\boldsymbol{\Psi}_0}(\boldsymbol{\iota}_0)L(\boldsymbol{\iota})] &= -\mathbf{X}'_{00}{}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_0 - \boldsymbol{\mu}), \\ \mathbf{E}_{\boldsymbol{\iota}_0} [U_\zeta(\boldsymbol{\iota}_0)L(\boldsymbol{\iota})] &= -\mathbf{X}_1^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_0 - \boldsymbol{\mu}) \text{ e} \\ \mathbf{E}_{\boldsymbol{\iota}_0} [U_{\phi_j}(\boldsymbol{\iota}_0)L(\boldsymbol{\iota})] &= -\frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{0j} \right] - (\mathbf{X}_{00j} \boldsymbol{\Psi}_0)^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_0 - \boldsymbol{\mu}). \end{aligned}$$

Portanto, $\hat{h}$ tem suas componentes definidas como

$$\begin{aligned} \hat{h}_{\boldsymbol{\Psi}_0} &= \mathbf{X}'_{00}{}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_0 - \boldsymbol{\mu}), \\ \hat{h}_\zeta &= \mathbf{X}_1^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_0 - \boldsymbol{\mu}) \\ \hat{h}_{\phi_j} &= -\frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}_{0j}^{-1} \boldsymbol{\Sigma}_0 \right] + \frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{0j} \right] + (\mathbf{X}_{00j} \boldsymbol{\Psi}_0)^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_0 - \boldsymbol{\mu}). \end{aligned}$$

$$\hat{\boldsymbol{\Upsilon}} = \text{cov}_{\boldsymbol{\iota}_0} [U(\boldsymbol{\iota}_0), U(\boldsymbol{\iota})^\top] |_{\boldsymbol{\iota}_0=\hat{\boldsymbol{\iota}}, \boldsymbol{\iota}=\hat{\boldsymbol{\iota}}} = \mathbf{E}_{\boldsymbol{\iota}_0} [U(\boldsymbol{\iota}_0) U(\boldsymbol{\iota})^\top] |_{\boldsymbol{\iota}_0=\hat{\boldsymbol{\iota}}, \boldsymbol{\iota}=\hat{\boldsymbol{\iota}}},$$

que tem forma

$$\hat{\Upsilon} = \begin{bmatrix} \Upsilon_{\Psi_0\Psi_0} & \Upsilon_{\Psi_0\zeta} & \Upsilon_{\Psi_0\phi} \\ \Upsilon_{\zeta\Psi_0} & \Upsilon_{\zeta\zeta} & \Upsilon_{\zeta\phi} \\ \Upsilon_{\phi\Psi_0} & \Upsilon_{\phi\zeta} & \Upsilon_{\phi\phi} \end{bmatrix}$$

sendo $\Upsilon_{ab} = E_{\boldsymbol{\iota}_0} [U_a(\boldsymbol{\iota}_0) U_b(\boldsymbol{\iota})^\top]$, para $\{a, b\} \in \{\Psi_0, \zeta, \phi\}$. Inicialmente, para obtermos os componentes por linha da matriz, calculamos a esperança $E_{\boldsymbol{\iota}_0} [U(\boldsymbol{\iota})^\top]$ para cada componente, utilizando as equações (C.2-C.6), obtendo

$$E_{\boldsymbol{\iota}_0} [U_{\Psi_0}(\boldsymbol{\iota})^\top] = (\boldsymbol{\mu}_0 - \boldsymbol{\mu})^\top \Sigma^{-1} \mathbf{X}'_0, \quad (\text{C.10})$$

$$E_{\boldsymbol{\iota}_0} [U_{\zeta}(\boldsymbol{\iota})^\top] = (\boldsymbol{\mu}_0 - \boldsymbol{\mu})^\top \Sigma^{-1} \mathbf{X}_1, \quad (\text{C.11})$$

$$E_{\boldsymbol{\iota}_0} [U_{\phi_j}(\boldsymbol{\iota})^\top] = -\frac{1}{2} \text{tr} [\Sigma^{-1} \Sigma_j] - \frac{1}{2} \text{tr} [\Sigma_j^{-1} \Sigma_0] - \frac{1}{2} (\boldsymbol{\mu}_0 - \boldsymbol{\mu})^\top \Sigma^{-1} (\boldsymbol{\mu}_0 - \boldsymbol{\mu}) + \quad (\text{C.12}) \\ + (\boldsymbol{\mu}_0 - \boldsymbol{\mu})^\top \Sigma^{-1} \mathbf{X}'_{0j} \Psi_0.$$

Derivando todas as equações acima com respeito a Ψ_0 e usando o argumento da equação (C.7) teremos a primeira linha de $\hat{\Upsilon}$, sendo, respectivamente,

$$\begin{aligned} \Upsilon_{\Psi_0\Psi_0} &= \mathbf{X}'_{00}{}^\top \Sigma^{-1} \mathbf{X}'_0, \\ \Upsilon_{\Psi_0\zeta} &= \mathbf{X}'_{00}{}^\top \Sigma^{-1} \mathbf{X}_1, \\ \Upsilon_{\Psi_0\phi_j} &= -\mathbf{X}'_{00}{}^\top \Sigma^{-1} (\boldsymbol{\mu}_0 - \boldsymbol{\mu}) + \mathbf{X}'_{00}{}^\top \Sigma^{-1} \mathbf{X}'_{0j} \Psi_0. \end{aligned}$$

Da mesma forma, para as segunda e terceira linhas da matriz $\hat{\Upsilon}$, basta derivarmos com relação a ζ e ϕ_j as equações (C.10-C.12), obtendo respectivamente,

$$\begin{aligned} \Upsilon_{\zeta\Psi_0} &= (\mathbf{X}'_{00})^\top \Sigma^{-1} \mathbf{X}'_0, \\ \Upsilon_{\zeta\zeta} &= \mathbf{X}_1^\top \Sigma^{-1} \mathbf{X}_1, \\ \Upsilon_{\zeta\phi_j} &= -\mathbf{X}_1^\top \Sigma^{-1} (\boldsymbol{\mu}_0 - \boldsymbol{\mu}) + \mathbf{X}_1^\top \Sigma^{-1} \mathbf{X}'_{0j} \Psi_0 \end{aligned}$$

e

$$\begin{aligned} \Upsilon_{\phi_j\Psi_0} &= \Psi_{00}^\top \mathbf{X}'_{00j}{}^\top \Sigma^{-1} \mathbf{X}'_0, \\ \Upsilon_{\phi_j\zeta} &= \Psi_{00}^\top \mathbf{X}'_{00j}{}^\top \Sigma^{-1} \mathbf{X}_1, \\ \Upsilon_{\phi_j\phi_k} &= -\frac{1}{2} \text{tr} [\Sigma^{-1} \Sigma_{0j}] - \Psi_{00}^\top \mathbf{X}'_{00j}{}^\top \Sigma^{-1} (\boldsymbol{\mu}_0 - \boldsymbol{\mu}) + \Psi_{00}^\top \mathbf{X}'_{00j}{}^\top \Sigma^{-1} \mathbf{X}'_{0k} \Psi_0. \end{aligned}$$