

Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Programa de Pós-Graduação em Estatística

Dissertação de Mestrado:

Testes de Hipótese Multivariados para Matrizes de
Covariâncias em Processos Autocorrelacionados com
Aplicações em Controle de Qualidade

Mestrando: Raphael Lennie Fernandes Ribeiro

Orientadora: Prof^a. Dr^a. Sueli Aparecida Mingoti

Belo Horizonte, Abril de 2010

Raphael Lennie Fernandes Ribeiro

Testes de Hipótese Multivariados para Matrizes de Covariâncias
em Processos Autocorrelacionados com Aplicações em Controle
de Qualidade

Dissertação apresentada ao Programa de Pós-Graduação em Estatística do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, como requisito para obtenção do título de Mestre em Estatística.

Orientadora: Prof^a. Dr^a. Sueli Aparecida Mingoti.

Belo Horizonte, Abril de 2010
Instituto de Ciências Exatas
Universidade Federal de Minas Gerais – UFMG

Agradecimentos

Agradeço aos meus Pais e irmãos, meus grandes incentivadores, que se fizeram presentes durante toda essa etapa.

À Prof^a. Dr^a. Sueli Aparecida Mingoti por me acompanhar e orientar desde a graduação, sempre com muita dedicação e comprometimento.

Aos Professores, Luiz H. Duczmal, Roberto da Costa Quinino e Daniel Furtado Ferreira pelas grandes contribuições que deram a este trabalho.

Aos Professores do Departamento de Estatística, pelos conhecimentos compartilhados.

A todos aqueles que contribuíram direta ou indiretamente para a realização deste trabalho fica meu eterno agradecimento.

RIBEIRO, R. L. F., MINGOTI, S. A. *Testes de Hipótese Multivariados para Matrizes de Covariâncias em Processos Autocorrelacionados com Aplicações em Controle de Qualidade*. Belo Horizonte: Departamento de Estatística, Universidade Federal de Minas Gerais, 2010, Dissertação (mestrado).

Resumo

Esta dissertação oferece algumas contribuições às áreas de testes de hipótese e de monitoramento de processos multivariados e controle de qualidade. Vários dos procedimentos clássicos de controle de processos multivariados são desenvolvidos assumindo que os vetores de observações são independentes com distribuição normal multivariada com vetor de médias μ e matriz de covariâncias Σ . Porém, na prática, é usual encontrar observações autocorrelacionadas. Deste modo, para melhorar o desempenho dos testes estatísticos e a qualidade do monitoramento do processo é necessário levar em consideração a autocorrelação dos vetores de observações. É essencial usar um modelo estatístico apropriado de séries temporais para modelar o comportamento dos vetores de observações do processo, tendo em conta a estrutura de dependência. Assim como é importante monitorar o vetor de médias, μ , do processo no caso multivariado, é também importante monitorar a variabilidade do processo. Dessa forma, nesta dissertação avaliamos inicialmente o comportamento de alguns testes estatísticos construídos para o monitoramento de matrizes de covariâncias, que são formulados sob a hipótese de independência entre as observações, quando essa suposição é violada.

Além disso, propomos algumas correções nestes testes para que sejam sensíveis a estrutura de autocorrelação das observações, correções estas que envolvem o ajuste de modelo de séries temporais multivariados. Os testes estatísticos avaliados nesta dissertação são: o teste da variância generalizada, o de razão de verossimilhança, o *Step down* proposto por Sullivan et al. (2007) e o teste VMAX proposto por Machado et al. (2008). A avaliação do impacto da autocorrelação nos níveis de significância dos testes que são formulados sob a suposição de independência, bem como a análise de poder para as correções dos testes que são propostos nesta dissertação foram realizados utilizando-se Simulação de Monte Carlo.

Palavras-chave: processos multivariados autocorrelacionados, testes estatísticos para matriz de covariâncias, modelos VAR(1).

RIBEIRO, R. L. F., MINGOTI, S. A. *Multivariate Hypothesis Tests for Covariance Matrices in Autocorrelated Processes with Applications in Quality Control*. Belo Horizonte: Departamento de Estatística, Universidade Federal de Minas Gerais, 2010, Dissertation (master degree).

Abstract

This dissertation brings some contributions to the areas of hypothesis testing, multivariate monitoring and quality control. Many of the classical procedures used to control multivariate processes are developed assuming that the vectors of observations are independent from a multivariate normal distribution with mean vector μ and covariance matrix Σ . However, in practice it is usual to find autocorrelated observations. Thus, to improve the performance of statistical tests and quality of monitoring the process is necessary to take into account the autocorrelation of the vectors of observations. It is essential to use a suitable time series statistical model to describe the behavior of the vectors of observations of the process, taking into account the dependence structure. Just as it is important to monitor the vector of means, μ , of the process in the multivariate case, it is also important to monitor the variability of the process. Thus, in this dissertation initially the behavior of some statistical tests built for monitoring the covariance matrices formulated under the assumption of independence among the observations, is evaluated when this assumption is violated.

Furthermore, we propose some corrections to these tests that are sensitive to the autocorrelation structure of the observations. These corrections are related to the estimation of the parameters of the multivariate time series model that generates the observations of the process. Statistical tests discussed in this dissertation are: the test of the generalized variance, the likelihood ratio, the Step Down suggested by Sullivan et al. (2007) and the VMAX test proposed by Machado et al. (2008). The impact of autocorrelation in the significance levels of the tests that are formulated under the assumption of independence, and the power analysis of the modified tests, that are proposed in this dissertation, were performed using Monte Carlo simulation.

Keywords: multivariate autocorrelated processes, statistical tests for covariance matrix, VAR(1) model.

Sumário

1	Introdução.....	01
1.1	Objetivos.....	05
1.2	Organização.....	06
2	Testes de hipóteses e gráficos de controle de processos.....	07
2.1	Conceitos básicos em teste de hipóteses.....	07
2.2	Base estatística do gráfico de controle.....	09
2.2.1	Exemplo de aplicação.....	14
3	Testes de hipóteses para matriz de covariâncias para amostras independentes.....	17
3.1	Distribuição normal multivariada.....	17
3.2	Teste do determinante.....	20
3.3	Teste do determinante proposto por Djauhari (2005).....	21
3.3.1	Exemplo de aplicação.....	23
3.4	Estudo da distribuição empírica de $ S $	25
3.5	Teste da Razão de Verossimilhança.....	28
3.5.1	Exemplo de aplicação.....	35
3.6	Teste <i>Step-down</i> proposto por Sullivan et al. (2007).....	36
3.6.1	Exemplo de aplicação.....	40
3.7	Teste VMAX proposto por Machado et. al (2008).....	41
3.7.1	Exemplo de aplicação.....	43
4	Processos multivariados autocorrelacionados.....	45
4.1	Modelo autorregressivo multivariado de ordem k , VAR(k).....	46
4.1.1	Estimadores de máxima verossimilhança de um modelo VAR(k).....	48

4.2 Modelo de média-móvel multivariado de ordem q , VMA(q).....	52
4.3 Matrizes de covariância (Γ) e correlação (ρ) cruzada.....	53
4.4 Estudo do comportamento de Σ , Γ_0 e $\rho(0)$ quando $\phi_{12} \neq 0$ e $\phi_{21} \neq 0$	56
4.5 Estudo da distribuição empírica de $ S $ e $ \hat{\Gamma}_0 $ em processos autocorrelacionados.....	60
4.6 Adaptação dos testes para matriz de covariâncias em processos autocorrelacionados.....	66
4.6.1 Modificação do teste do determinante tradicional e Djauhari (2005).....	67
4.6.2 Modificação do teste da razão de verossimilhança.....	68
4.6.3 Modificação teste Step-down proposto por Sullivan et al. (2007).....	69
4.6.4 Modificação teste VMAX proposto por Machado et. al (2008).....	70
5 Aspectos computacionais: Simulação e estimação dos parâmetros de um processo VAR(1)..	72
6 Avaliação do poder dos testes para matriz de covariâncias via simulação de Monte Carlo....	77
6.1 Estudo do poder dos testes - Situação 1.....	77
6.1.1 Estimação do tamanho dos testes na situação 1.....	81
6.1.2 Estimação do poder dos testes na situação 1.....	87
6.2 Estudo do poder dos testes - Situação 2.....	98
6.2.1 Estimação do tamanho dos testes na situação 1.....	102
6.2.2 Estimação do poder dos testes na situação 2.....	105
7 Conclusões finais e sugestões para trabalhos futuros.....	111
Apêndice A.....	115
Apêndice B.....	121
Apêndice C.....	123
Referências Bibliográficas.....	124

Lista de Figuras

Figura 2.1: Erros associados a testes de hipóteses.....	08
Figura 2.2: Processo no estado de controle estatístico.....	10
Figura 2.3: Processo no estado fora de controle.....	10
Figura 2.4: Exemplo de gráfico de controle.....	11
Figura 2.5: Relação entre gráfico de controle e teste de hipótese.....	12
Figura 2.6: Gráfico de controle \bar{x} para o diâmetro do anel de pistão	15
Figura 3.1: Função densidade de uma normal bivariada com correlação igual a 0,85.....	19
Figura 3.2: Função densidade de uma normal bivariada com correlação igual a 0.....	20
Figura 3.3: Interpretação geométrica do determinante.....	22
Figura 3.4: Proposta do teste de Sullivan et al. (2007)	36
Figura 5.1: Estratégia de Simulação.....	76

Lista de Quadros

Quadro 3.1: Medidas descritivas de $ S $ para amostras de tamanho $n = 25$	26
Quadro 3.2: Medidas descritivas de $ S $ para amostras de tamanho $n = 50$	26
Quadro 3.3: Medidas descritivas de $ S $ para amostras de tamanho $n = 100$	27
Quadro 3.4: Medidas descritivas de $ S $ para amostras de tamanho $n = 200$	27
Quadro 3.5: Medidas descritivas de $ S $ para amostras de tamanho $n = 1000$	28
Quadro 4.1: Medidas descritivas de $ S $ para amostras de tamanho $n = 25$	61
Quadro 4.2: Medidas descritivas de $ S $ para amostras de tamanho $n = 50$	61
Quadro 4.3: Medidas descritivas de $ S $ para amostras de tamanho $n = 100$	62
Quadro 4.4: Medidas descritivas de $ S $ para amostras de tamanho $n = 200$	62
Quadro 4.5: Medidas descritivas de $ S $ para amostras de tamanho $n = 1000$	63
Quadro 4.6: Medidas descritivas de $ \hat{\Gamma}_0 $ para amostras de tamanho $n = 25$	63
Quadro 4.7: Medidas descritivas de $ \hat{\Gamma}_0 $ para amostras de tamanho $n = 50$	64
Quadro 4.8: Medidas descritivas de $ \hat{\Gamma}_0 $ para amostras de tamanho $n = 100$	64
Quadro 4.9: Medidas descritivas de $ \hat{\Gamma}_0 $ para amostras de tamanho $n = 200$	65
Quadro 4.10: Medidas descritivas de $ \hat{\Gamma}_0 $ para amostras de tamanho $n = 1000$	65

Lista de Tabelas

Tabela 2.1: Valores de poder e ARL out of control para o exemplo de aplicação.....	16
Tabela 3.1: Matriz de covariâncias assintótica ($\Sigma_{\hat{\theta}_n}$) de $\hat{\theta}_n$ para uma distribuição normal bivar.....	38
Tabela 4.1: Alterações na matriz Γ_0 quando $\phi_{12} = \phi_{21} \geq 0$	56
Tabela 4.2: Alterações na matriz Γ_0 quando $\phi_{12} = \phi_{21} \leq 0$	57
Tabela 4.3: Alterações na matriz Γ_0 quando $\phi_{12} = 0$ e $\phi_{21} \geq 0$	57
Tabela 4.4: Alterações na matriz Γ_0 quando $\phi_{12} = 0$ e $\phi_{21} \leq 0$	57
Tabela 4.5: Alterações na matriz Γ_0 quando $\phi_{12} \geq 0$ e $\phi_{21} = 0$	58
Tabela 4.6: Alterações na matriz Γ_0 quando $\phi_{12} \leq 0$ e $\phi_{21} = 0$	58
Tabela 4.7: Alterações na matriz Γ_0 quando $\phi_{12} \neq 0$ e $\phi_{21} \neq 0$	59
Tabela 4.8: Valores médios de $ S $ e $ \hat{\Gamma}_0 $ para diferentes cenários de tamanhos de amostras.....	66
Tabela 6.1: Testes abordados na situação 1.....	80
Tabela 6.2: Limites críticos e estimativas do tamanho do teste na situação 1: n=25 - μ conhecido.....	81
Tabela 6.3: Limites críticos e estimativas do tamanho do teste na situação 1: n=25 - μ desconhecido.....	81
Tabela 6.4: Limites críticos e estimativas do tamanho do teste na situação 1: n=50 - μ conhecido.....	82
Tabela 6.5: Limites críticos e estimativas do tamanho do teste na situação 1: n=50 - μ desconhecido.....	82
Tabela 6.6: Limites críticos e estimativas do tamanho do teste na situação 1: n=100 - μ conhecido.....	82
Tabela 6.7: Limites críticos e estimativas do tamanho do teste na situação 1: n=100 - μ desconhecido.....	83
Tabela 6.8: Limites críticos e estimativas do tamanho do teste na situação 1: n=200 - μ conhecido.....	83
Tabela 6.9: Limites críticos e estimativas do tamanho do teste na situação 1: n=200 - μ desconhecido.....	83
Tabela 6.10: Limites críticos e estimativas do tamanho do teste na situação 1: n=1000 - μ conhecido.....	84
Tabela 6.11: Limites críticos e estimativas do tamanho do teste na situação 1: n=1000 - μ desconhecido.....	84

Tabela 6.12: Limites críticos e estimativas do tamanho do teste na situação 1: $n=5000$ - μ conhecido.....	84
Tabela 6.13: Limites críticos e estimativas do tamanho do teste na situação 1: $n=5000$ - μ desconhecido.....	85
Tabela 6.14: Testes selecionados para avaliação do poder na situação 1.....	86
Tabela 6.15: Modelos para avaliação do poder dos testes na situação 1.....	90
Tabela 6.16: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 1 na situação 1.....	91
Tabela 6.17: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 2 na situação 1.....	93
Tabela 6.18: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 3 na situação 1.....	94
Tabela 6.19: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 4 na situação 1.....	95
Tabela 6.20: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 5 na situação 1.....	97
Tabela 6.21: Testes abordados na situação 2.....	101
Tabela 6.22: Limites críticos e estimativas do tamanho do teste na situação 2, para $n=25$	102
Tabela 6.23: Limites críticos e estimativas do tamanho do teste na situação 2, para $n=50$	102
Tabela 6.24: Limites críticos e estimativas do tamanho do teste na situação 2, para $n=100$	103
Tabela 6.25: Limites críticos e estimativas do tamanho do teste na situação 2, para $n=200$	103
Tabela 6.26: Limites críticos e estimativas do tamanho do teste na situação 2, para $n=1000$	104
Tabela 6.27: Limites críticos e estimativas do tamanho do teste na situação 2, para $n=5000$	104
Tabela 6.28: Testes selecionados para avaliação do poder na situação 2.....	105
Tabela 6.29: Modelos para avaliação do poder dos testes na situação 2.....	107
Tabela 6.30: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 1 na situação 2.....	108
Tabela 6.31: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 2 na situação 2.....	108
Tabela 6.32: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 3 na situação 2.....	110
Tabela 6.33: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 4 na situação 2.....	110
Tabela B.1: Percentis da distribuição exata - Situação 1 – μ conhecido.....	121

Tabela B.2: Percentis da distribuição exata - Situação 1 – μ desconhecido.....	122
Tabela C.1: Percentis da distribuição exata - Situação 2 – μ desconhecido.....	123

Capítulo 1

Introdução

Qualidade sempre foi parte integrante de praticamente todos os processos de produtos e serviços. No entanto, a conscientização de sua importância e a introdução de métodos formais para o controle e melhoria da qualidade têm tido um desenvolvimento evolutivo (Montgomery, 2004). A definição de qualidade baseia-se no ponto de vista de que produtos e serviços devem apresentar medidas dentro das especificações exigidas por aqueles que os usam. Os métodos de melhoria da qualidade aplicam-se a qualquer área de uma companhia ou organização, incluindo manufatura, planejamento de engenharia, finanças, marketing, etc.

Atualmente a qualidade tornou-se um dos mais importantes fatores de decisão dos consumidores na seleção de produtos e serviços que competem entre si. Compreender e melhorar a qualidade é um fator-chave que conduz ao sucesso, crescimento e a uma melhor posição de competitividade de um negócio. Porém, no início do século XX os métodos estatísticos para melhoria da qualidade não eram tão utilizados nas indústrias. Em 1924, Walter A. Shewhart, desenvolveu o conceito estatístico de gráfico de controle, que é considerado, como o começo formal do controle estatístico da qualidade. Por ocasião da II Guerra Mundial, os EUA incentivaram a utilização dos métodos estatísticos de Shewhart pelos seus fornecedores ajudando a disseminar os novos métodos de controle no mundo. Após a II Guerra Mundial, a utilização das ferramentas do controle de qualidade foi bastante intensificada durante a implantação do Programa de Qualidade Total, no Japão, pelas mãos de Willian E. Deming e Joseph M. Juran. No final do século XX, a tolerância extremamente baixa a falhas em dispositivos tais como espaçonaves, usinas nucleares, medicamentos de última geração, entre outros, fizeram com que abordagens extremamente refinadas fossem criadas para o controle de qualidade. Montgomery (2004) apresenta uma linha do tempo de alguns marcos importantes do processo evolutivo do controle de qualidade.

Um dos problemas das cartas de controle desenvolvidas por Shewhart é que elas eram apropriadas apenas para a análise de uma variável do processo. Quando nesse sistema a qualidade do produto era medida por mais de uma variável, caso de processos

multivariados, várias cartas de controle deveriam ser utilizadas simultaneamente, e de forma independente, para que fosse feito o monitoramento do processo. Porém, quando as variáveis eram correlacionadas entre si, as cartas de controle não incorporavam essa correlação o que poderia gerar conclusões incorretas sobre a estabilidade do processo. Atualmente, os processos multivariados passaram a ser muito comuns uma vez que em geral, são várias as características que determinam a qualidade final de um produto (ou serviço) e que devem ser monitoradas ao longo do processo de produção. Shewhart, conhecido por ser o pioneiro em desenvolver gráficos de controle, destacou-se também por reconhecer a necessidade de considerar problemas de controle estatístico multivariado.

Recentemente, vários procedimentos estatísticos de controle de processos multivariados têm sido sugeridos. Para o controle do vetor de médias pode-se destacar o teste de hipótese de T^2 de Hotelling (1947) e o de Hayter e Tsui (1994), entre outros (ver Mason e Young, 2002). Já para o controle da matriz de covariâncias, pode-se destacar os testes propostos por Yeh et al. (2006); Alt (1985); Djauhari (2005); Montgomery e Wadsworth (1972), entre outros. No contexto do monitoramento da estrutura da variabilidade de processos independentes, dois testes muito conhecidos são o da variância generalizada (Montgomery, 2004) e o teste da razão de verossimilhança apresentado em Alt (1985) dentro do contexto de controle de qualidade, mas que já havia sido apresentado anteriormente dentro do contexto estatístico (ver Anderson, 1958, 2003). Em 2008, Machado et al. propuseram um gráfico de controle, para a matriz de covariâncias, denominado VMAX, que baseia-se nas variâncias amostrais padronizadas das características de qualidade. Os autores abordaram apenas o caso bivariado no artigo. A vantagem do gráfico VMAX, segundo os autores, é que através dele é mais fácil identificar a variável que teve sua variabilidade alterada pela ocorrência de uma causa especial. Outro método, proposto recentemente, é o apresentado no artigo de Sullivan et al. (2007). Esse procedimento é interessante, pois permite a comparação da matriz de covariâncias em todos os seus elementos sem a necessidade de reduzir a informação das matrizes em determinantes (variâncias generalizadas) e traços como acontece com outros testes. Uma revisão sobre alguns testes para matrizes de covariâncias é apresentada Yeh, et al. (2006).

O controle de processos multivariados é de particular importância nos dias de hoje, na medida em que os procedimentos de inspeção automática tornam relativamente fácil medir muitas variáveis em cada unidade do produto. A grande maioria dos métodos estatísticos para controle de processos multivariados assumem que os vetores de observações das características da qualidade são independentes. Porém, assumir

independência não é razoável em alguns processos. Um exemplo clássico de processos não independentes ocorre na indústria química. Num processo químico acontece um desgaste natural de alguns componentes, portanto, o que acontece no tempo t , dependendo da condição do equipamento, pode influenciar no que ocorrerá no tempo $t+h$, $h>0$. Assim, podemos introduzir a autocorrelação, que tem um efeito significativo no desempenho dos procedimentos de controle de qualidade. No caso univariado, quando detectamos a presença da autocorrelação, uma das técnicas utilizadas é ajustar um modelo de séries temporais ao conjunto de observações. Assim, como os resíduos são independentes, eles são usados no controle do processo. Porém, se o modelo de séries temporais não é adequado, os resíduos provavelmente não serão independentes e conseqüentemente haverá um grande número de alarmes falsos (Box e Luceno, 1997).

Na literatura há uma escassez de trabalhos publicados que tratam de controle multivariado de processos autocorrelacionados. A maioria dos artigos trata dos testes para comparação do vetor de médias (Kalgonda e Kulkarni, 2004; Mason e Young 1999; Lu e Reynolds 1991, entre outros) e poucos discutem testes para comparação da matriz de covariâncias (Chang e Zhang, 2007). Sabe-se que a análise de processos autocorrelacionados através de técnicas convencionais pode causar sérios impactos na carta de controle causado pelo fato da autocorrelação ter sido desconsiderada. Dessa forma, Kalgonda e Kulkarni (2004) estudaram processos bivariados autocorrelacionados de ordem 1 (VAR(1)) utilizando conceitos de séries temporais multivariadas (Brockweel e Davis, 1996; Lutkepohl, 1993; Tsay, 2002). Os autores mostram que quando levamos em consideração a autocorrelação do processo a matriz de covariâncias pode ser diferente daquela no qual a autocorrelação é ignorada. Dessa forma, é necessário ficar atento para a escolha da matriz de covariâncias que será utilizada no controle do processo. É, portanto, importante, antes de iniciar o monitoramento de um processo, identificar se ele produz observações independentes ou autocorrelacionadas. Essa influência da autocorrelação também é mostrada em Oliveira (2007) no que se refere a determinação da capacidade de processos multivariados. Esses fatos mostram a importância de se identificar a autocorrelação das observações e levá-la em consideração na construção dos testes estatísticos para monitoramento de parâmetros de processos.

Nessa dissertação será apresentado um procedimento de incorporação desta autocorrelação em alguns dos testes de hipótese para matrizes de covariâncias já existentes, que são construídos sob a suposição de independência entre as observações. Nesse contexto, faremos a avaliação do desempenho dos testes quando levamos em consideração a matriz de covariâncias cruzada, que contém toda a informação da variação temporal estimada através do sistema linear de Yule-Walker e quando

utilizamos a matriz de covariância amostral S para avaliar se o processo multivariado está sob controle estatístico no que se refere a estrutura de variabilidade.

1.1 Objetivos

O objetivo dessa dissertação é avaliar o efeito da autocorrelação em alguns testes estatísticos multivariados para matrizes de covariâncias que são construídos sob a suposição de independência entre as observações. Os testes a serem avaliados são: do determinante ou variância generalizada (Montgomery, 2004), do determinante modificado proposto por Djauhari (2005), da razão de verossimilhança (Alt, 1985), o *Step-down*, proposto por Sullivan et al. (2007) e finalmente o teste VMAX proposto por Machado et al. (2008). Assim, avaliaremos como a autocorrelação afeta o tamanho, o poder e a distribuição das estatísticas dos testes construídos para matriz de covariâncias sob a suposição de normalidade multivariada e independência entre as observações. Além disso, proporemos uma forma de correção nos testes para matriz de covariâncias quando levamos em consideração a autocorrelação entre as observações. Um outro ponto importante do nosso estudo será avaliar a aplicabilidade das distribuições assintóticas para as estatísticas de teste nos casos em que há autocorrelação. Como nosso estudo será focado nos modelos autorregressivo multivariado de ordem 1, em algumas partes do texto detalharemos os processos de séries temporais desse tipo em termos de seus aspectos teóricos e práticos.

1.2 Organização

O conteúdo dessa dissertação está organizado em sete capítulos e três apêndices. No Capítulo 2, apresentaremos alguns conceitos básicos utilizados em teste de hipóteses e em controle de qualidade que serão abordados nessa dissertação. No Capítulo 3, apresentaremos a função densidade da distribuição normal multivariada, já que grande parte dos testes que serão abordados vem do pressuposto de que os dados são gerados pela distribuição normal multivariada, e os testes para matriz de covariâncias, no contexto, em que temos independência entre as observações. Já no Capítulo 4, discutiremos os principais conceitos de processos autocorrelacionados, apresentaremos alguns modelos de séries temporais multivariados e mostraremos como os testes para matriz de covariâncias podem ser adaptados para o caso em que estamos em um processo autocorrelacionado. No Capítulo 5 mostraremos como serão gerados os vetores de observações autocorrelacionados e como será feita a estimação dos parâmetros dos modelos que serão objetos de estudo desta dissertação. Os resultados das simulações desenvolvidas para avaliarmos o desempenho dos testes para matriz de covariâncias quando estamos em um processo autocorrelacionado serão apresentados no Capítulo 6. No Capítulo 7 apresentaremos as conclusões e sugestões para futuros trabalhos. Finalmente, no Apêndice A será apresentado o cálculo da matriz de covariâncias cruzada, para um processo bivariado $VAR(I)$, considerando $\phi_{12} = \phi_{21}$ e nos Apêndices B e C os percentis das distribuições exatas dos testes em estudo.

Capítulo 2

Testes de hipótese e gráficos de controle de processos

Neste capítulo apresentaremos alguns conceitos básicos de testes de hipóteses, que serão tratados nesse projeto de dissertação. Além disso, abordaremos alguns pontos importantes em controle de qualidade, como, avaliação do desempenho do gráfico de controle através do *ARL*, erro do tipo I e do tipo II, e mostrar a relação existente entre o teste de hipótese e o gráfico de controle utilizado em controle estatístico de processos.

2.1 Conceitos básicos em teste de hipótese

O teste de hipótese é um procedimento de inferência estatística amplamente usado em várias áreas do conhecimento. Num teste de hipótese constrói-se hipóteses sobre os parâmetros com base no conhecimento ou nas especulações que se faz inicialmente sobre eles. Em estatística, uma hipótese é uma alegação, ou afirmação, sobre uma propriedade de uma população. Usualmente, as hipóteses são denominadas como hipótese nula (H_0) e hipótese alternativa (H_1). De uma forma geral, a hipótese nula é uma afirmação feita sobre os parâmetros de interesse, enquanto a hipótese alternativa é a afirmação contrária à hipótese nula. Como ilustração, considere θ o parâmetro de interesse e as seguintes hipóteses a serem testadas: $H_0: \theta \in \Omega_0$ versus $H_1: \theta \in \Omega_1$. Ao construirmos essas hipóteses estamos induzindo uma partição do espaço paramétrico (Ω) em dois subespaços denotados por Ω_0 e Ω_1 , onde $\Omega_0 \cup \Omega_1 = \Omega$ e $\Omega_0 \cap \Omega_1 = \emptyset$. Assim, o teste de hipótese nada mais é que um procedimento, φ , que nos permite decidir entre a hipótese nula (H_0) e a hipótese alternativa (H_1). Define-se φ da seguinte forma:

$$\varphi = \begin{cases} 1 & \text{rejeitamos } H_0 \\ 0 & \text{não rejeitamos } H_0 \end{cases} \quad (2.1)$$

ou seja, num teste de hipótese existe apenas duas decisões possíveis: rejeitar ou não H_0 .

De acordo com as hipóteses construídas, deve-se escolher entre $H_0: \theta \in \Omega_0$ ou $H_1: \theta \in \Omega_1$. Para isto, deve-se observar uma amostra de tamanho n de elementos de uma população cuja distribuição da variável de interesse envolva o parâmetro θ . Denote por S o conjunto de todos os resultados possíveis para a amostra. Isto é, seja S o espaço amostral $S = \{(X_1, \dots, X_n), X_i \in \mathfrak{X}, i = 1, 2, \dots, n\}$. Ao definir-se um procedimento de teste, φ , estamos definindo um procedimento que divide S em dois subconjuntos, um subconjunto que conduz à rejeição de H_0 e outro subconjunto que conduz a não-rejeição de H_0 . Assim, o subconjunto de valores de S que se observados, conduzirão a rejeição de H_0 é denominado região crítica, ζ_φ . Considere $X = (X_1, X_2, \dots, X_n)$ uma amostra aleatória de tamanho n da variável aleatória \mathbf{X} . De uma forma geral, definimos um procedimento de teste em função de uma estatística $T(X)$ da seguinte forma:

$$\varphi = \begin{cases} 1 & \text{se } T(X) \in \zeta_\varphi \\ 0 & \text{se } T(X) \notin \zeta_\varphi \end{cases} \quad (2.2)$$

Ao realizar um teste de hipóteses dois erros podem ser cometidos: rejeitar a hipótese H_0 quando tal hipótese é verdadeira, denominado *erro do tipo I*, ou não rejeitar a hipótese H_0 quando ela deveria ser rejeitada, denominado *erro do tipo II*. A Figura 2.1 resume as possíveis situações.

		Situação	
		H_0 verdadeira	H_0 falsa
Decisão	Rejeitar H_0	ERRO TIPO I	SEM ERRO
	Não Rejeitar H_0	SEM ERRO	ERRO TIPO II

Figura 2.1: Erros associados a testes de hipóteses

Em condições normais, só conseguimos controlar um tipo de erro, e o escolhido é o erro tipo I. Denominamos como nível de significância (α) a probabilidade do erro tipo I, isto é:

$$\begin{aligned} \alpha &= P(\text{Erro tipo I}) = P(\text{Rejeitar } H_0 \mid H_0 \text{ verdadeira}) \\ &= P(\varphi = 1 \mid \theta \in \Omega_0) \end{aligned} \quad (2.3)$$

Para avaliar a qualidade de um teste de hipótese uma ferramenta amplamente utilizada é a função poder. Seja φ um procedimento de teste e seja $X = (X_1, X_2, \dots, X_n)$ uma

amostra aleatória a ser observada. Considere ζ_φ como a região crítica induzida por φ . Chama-se função poder de φ a função π_φ definida de Ω em $[0,1]$ da seguinte forma:

$$\begin{aligned}\pi_\varphi &= P(X \in \zeta_\varphi | \theta) \\ &= P(\varphi = 1 | \theta) \\ &= P(\text{Rejeitar } H_0 | \theta)\end{aligned}\tag{2.4}$$

Quando $\theta \in \Omega_1$ estamos avaliando a probabilidade de que o teste venha a detectar que a hipótese nula é falsa. Quanto maior essa probabilidade melhor é o teste, ou seja, mais poderoso.

Para finalizar os conceitos usados em teste de hipótese, considere φ um procedimento de teste e seja π_φ a função poder do teste. O tamanho do teste é definido como:

$$\begin{aligned}\alpha_0 &= \sup_{\theta \in \Omega_0} \pi_\varphi \\ &= \sup_{\theta \in \Omega_0} P(\varphi = 1 | \theta)\end{aligned}\tag{2.5}$$

Note que o tamanho de um teste é o maior valor assumido pela função poder quando consideramos apenas os valores de $\theta \in \Omega_0$, ou seja, o tamanho do teste é o maior valor possível para o tamanho do erro tipo I quando o teste φ é construído. Para mais detalhes sobre teste de hipóteses ver Casella e Berger (2002).

2.2 Base estatística do gráfico de controle

Para que um produto corresponda às exigências de seus clientes, o mesmo deve ser produzido por um processo que seja estável, ou seja, o processo deve ser capaz de operar com pequena variabilidade em torno das dimensões-alvo. Porém, em qualquer processo de produção, independentemente de quão bem planejado ou cuidadosamente mantido ele seja, uma certa quantidade de variabilidade inerente ou natural sempre existirá. Trata-se da variabilidade natural do processo, que é fruto de uma série de pequenas perturbações, ou causas aleatórias, contra as quais pouco ou nada se pode fazer. Quando o processo apresenta apenas a variabilidade natural, devido às causas aleatórias, diz-se que ele está no estado de controle estatístico. A Figura 2.2 ilustra um

processo sujeito apenas a causas aleatórias. Observe que a característica X apresenta variabilidade que pode ser representada por uma distribuição de probabilidades que se mantém estável ao longo do tempo.

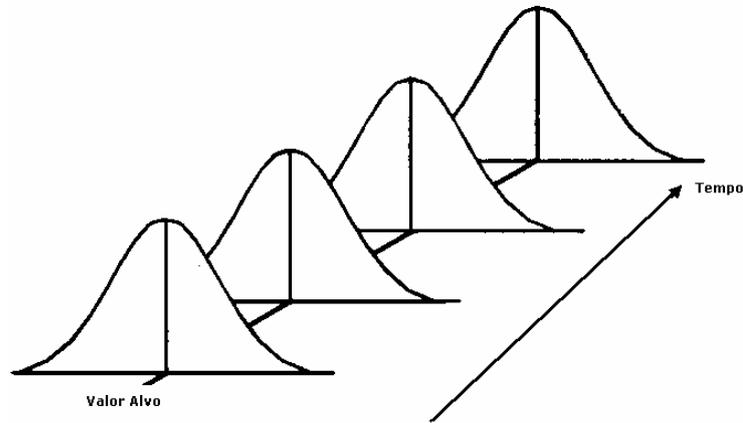


Figura 2.2: Processo no estado de controle estatístico

Fonte: Schissatti (1998)

Nenhum processo, porém, deixa de estar sujeito, também, à ocorrência ocasional de perturbações maiores, chamadas de causas especiais, que têm o efeito de deslocar a distribuição da variável aleatória X ou aumentar sua dispersão, ou ambos. Uma causa especial é um problema ou modo de operação anormal do processo, que pode, portanto, ser corrigido ou eliminado. Quando, além das causas aleatórias de variabilidade, causas especiais estiverem presentes, diz-se que o processo está fora de controle. A Figura 2.3 ilustra o caso em que uma causa especial afeta o processo.

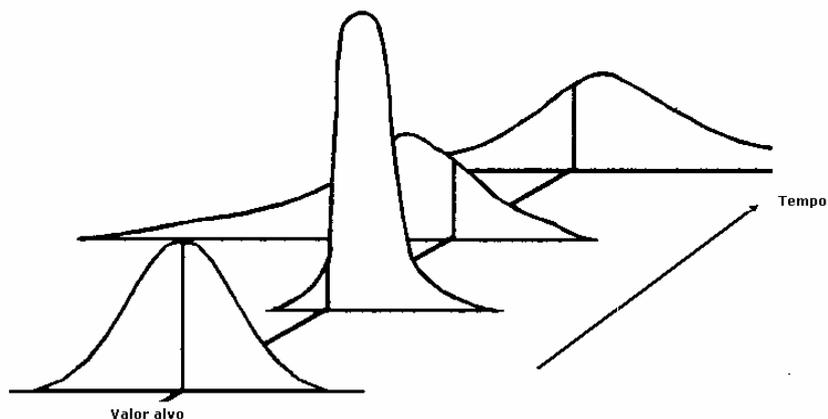


Figura 2.3: Processo no estado fora de controle

Fonte: Schissatti (1998)

Em geral, os processos de produção operarão em estado sob controle, produzindo itens aceitáveis por períodos de tempo relativamente longos. No entanto, causas especiais, ocorrerão, resultando em um deslocamento para um estado fora de controle

onde uma maior proporção de produtos do processo não corresponde às exigências. Um objetivo maior do controle estatístico do processo é detectar rapidamente a ocorrência de causas especiais que ocasionarão mudanças nos parâmetros do processo, de modo que a partir de uma investigação alguma ação corretiva possa ser realizada antes que muitas unidades não-conformes sejam produzidas.

Para verificar a estabilidade e a existência de causas especiais que atuam sobre o processo, é comum utilizar cartas de controle univariadas de Shewhart (Shewhart, 1929). Segundo Montgomery (2004), essas cartas são construídas a partir do valor de uma estatística $T(X)$ que é calculada para cada amostra do processo e logo em seguida os valores dessa estatística são colocados em um gráfico ou carta de controle. O gráfico de controle contém uma linha central, representando o valor médio da distribuição da característica da qualidade que corresponde ao estado sob controle. Duas outras linhas horizontais, chamadas de limite superior de controle (LSC) e limite inferior de controle (LIC), são mostradas no gráfico, como pode ser visto pela Figura 2.4. Esses limites de controle são escolhidos de modo que, se o processo está sob controle estatístico, praticamente todos os pontos amostrais estarão entre eles. No entanto, um ponto que caia fora dos limites de controle é interpretado como evidência de que o processo está fora de controle, e investigação e ações corretivas são necessárias para encontrar e eliminar as causas especiais responsáveis por esse comportamento.

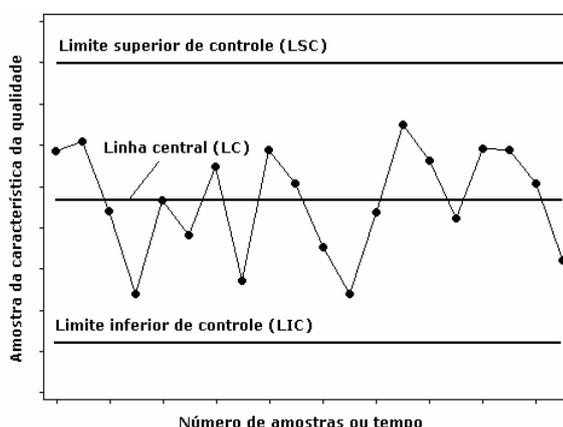


Figura 2.4: Exemplo de gráfico de controle

Os limites dos gráficos de controle são determinados numa fase denominada, Fase 1. Nessa fase, ocorre a estimação dos parâmetros do processo e ela só deve ser encerrada quando o processo estiver estável e ajustado. Vale lembrar que esses limites devem ser determinados somente quando o processo estiver isento de causas especiais. Uma vez identificados os limites de controle, podemos utilizá-los para monitorar o processo, o que é denominado de Fase 2. Para mais detalhes ver (Costa et al., 2005).

Há uma relação muito próxima entre gráficos de controle e testes de hipótese. Para ilustrar essa relação, considere X uma característica de qualidade com distribuição normal com média μ_0 e variância σ^2 . Suponha que o eixo horizontal central, na Figura 2.4, seja a média teórica μ_0 da distribuição de X . Agora, se o valor amostral da média \bar{x} se localizar entre os limites de controle, podemos concluir que a média do processo está sob controle; isto é, os dados amostrais são provavelmente provenientes da distribuição normal com média μ_0 (alvo). Por outro lado, se \bar{x} não se encontra dentro dos limites de controle, concluímos que o processo está fora de controle em relação a sua média; isto é, ela é igual a algum valor $\mu_1 \neq \mu_0$. Logo, o gráfico de controle é um teste de hipótese sendo H_0 : processo sob controle estatístico versus H_1 : processo fora de controle estatístico. Um valor de $T(\bar{X})$ que se localiza entre os limites de controle é equivalente à não rejeição da hipótese de controle estatístico, e um que se localiza fora dos limites de controle é equivalente à rejeição da hipótese de controle estatístico. A Figura 2.5 sintetiza a relação entre o gráfico de controle e o teste de hipótese no caso da distribuição normal.

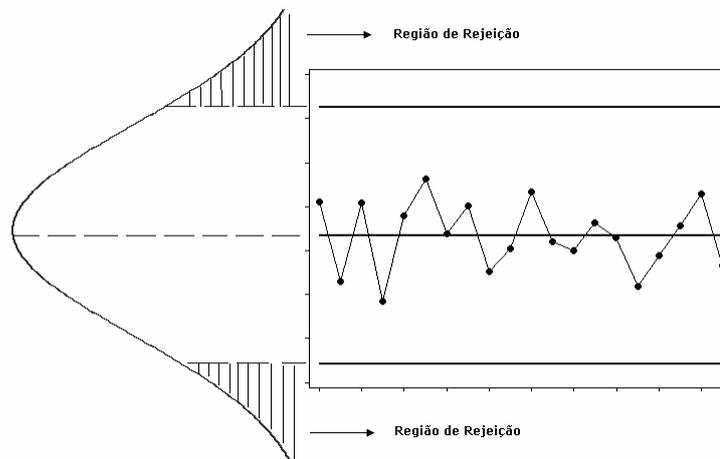


Figura 2.5: Relação entre gráfico de controle e teste de hipótese

Um lugar onde o esquema de teste de hipótese é útil é na análise do desempenho de um gráfico de controle. Por exemplo, podemos pensar na probabilidade de um erro do tipo I e do tipo II para o gráfico de controle. Se o processo estiver em controle (H_0 verdadeira), o erro do tipo I, α , representa a probabilidade de erroneamente considerar-se o processo como fora de controle, o que denominamos de alarme falso. A consequência de ordem prática associada ao erro do tipo I é intervir no processo sem necessidade, já que ele está isento de causas especiais, o que acarreta um custo e um risco de desajustar um processo que estava ajustado. Se o processo estiver fora de controle (H_1 verdadeira), o erro tipo II, β , representa a probabilidade de erroneamente

considerar-se o processo em controle quando ele está fora de controle, o que denominamos de não-deteção. Assim, a probabilidade de detectar-se na carta de controle uma mudança significativa no processo quando de fato ela existe, denominado alarme verdadeiro, é dada por $1 - \beta$, ou seja, o poder do teste. Cartas de controle com baixos valores de β devem ser preferidas e uma das formas de minimizar essa probabilidade é aumentando o tamanho das amostras observadas do processo (Montgomery, 2004), ou diminuindo as taxas de amostragem, isto é, coleta-se amostras em intervalos de tempo, h , menores.

Duas medidas de desempenho são utilizadas para avaliar a qualidade da carta de controle: *ARL in control* (*Average Run Length in control*) e o *ARL out of control* (*Average Run Length out of control*). O *ARL in control* representa o número médio de amostras observadas do processo até que um alarme falso seja detectado, enquanto o *ARL out of control* representa o número médio de amostras observadas do processo até que um alarme verdadeiro seja detectado. Essas medidas são definidas como em (2.6) e (2.7).

$$ARL \text{ in control} = \frac{1}{\alpha} \quad (2.6)$$

$$ARL \text{ out of control} = \frac{1}{1 - \beta} \quad (2.7)$$

Alguns autores traduzem ARL como NMA (número médio de amostras) ou CMS (comprimento médio da seqüência). Considerando-se a variável aleatória Y como o número de amostras aleatórias observadas do processo até que se obtenha o primeiro alarme falso, então Y tem uma distribuição geométrica com esperança e variância dadas como em (2.8) e (2.9).

$$E(Y) = \frac{1}{\alpha} \quad (2.8)$$

$$Var(Y) = \frac{1 - \alpha}{\alpha^2} \quad (2.9)$$

Da mesma forma, se Y representa o número de amostras aleatórias observadas do processo até que se obtenha o primeiro alarme verdadeiro, então Y tem uma distribuição geométrica com esperança e variância dadas como em (2.10) e (2.11).

$$E(Y) = \frac{1}{1 - \beta} \quad (2.10)$$

$$Var(Y) = \frac{\beta}{(1 - \beta)^2} \quad (2.11)$$

Assim, por meio dos valores do *ARL in control* e *ARL out of control* consegue-se medir a eficácia da carta de controle. Quanto maior for o *ARL in control* e menor o *ARL out of control* mais eficiente é a carta de controle. Para fins de exemplificação, da relação entre gráfico de controle e teste de hipótese, vamos considerar um exemplo adaptado de Montgomery (2004).

2.2.1 Exemplo de aplicação

Em uma fábrica de anéis de pistão para motores de automóveis, uma característica da qualidade é o diâmetro interno do anel. O processo sob controle requer um diâmetro médio interno de 74 mm, e um desvio padrão igual 0,01 mm. Agora, se são extraídas amostras de tamanho $n=5$, o desvio padrão da média amostral \bar{X} é:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0,01}{\sqrt{5}} = 0,0045$$

Portanto, se o processo está sob controle com um diâmetro médio de 74 mm, supondo-se que o diâmetro do anel tenha distribuição normal, esperaríamos que $100(1 - \alpha)\%$ dos diâmetros médios das amostras (\bar{x}), ficassem entre $74 + z_{\alpha/2}(0,0045)$ e $74 - z_{\alpha/2}(0,0045)$, sendo $z_{\alpha/2}$ o quantil da distribuição normal padronizada, tal que $P(Z \leq -z_{\alpha/2}) = \alpha/2$. Escolhemos a constante $z_{\alpha/2} = 3$ de modo que os limites superior e inferior se tornam:

$$LIC = 74 - 3(0,0045) = 73,9865$$

$$LSC = 74 + 3(0,0045) = 74,0135$$

Esses limites são tipicamente chamados limites de controle três-sigma, que correspondem a uma probabilidade igual a 0,9973 na distribuição normal padronizada. A Figura 2.6 mostra o gráfico de controle para o diâmetro médio do anel.

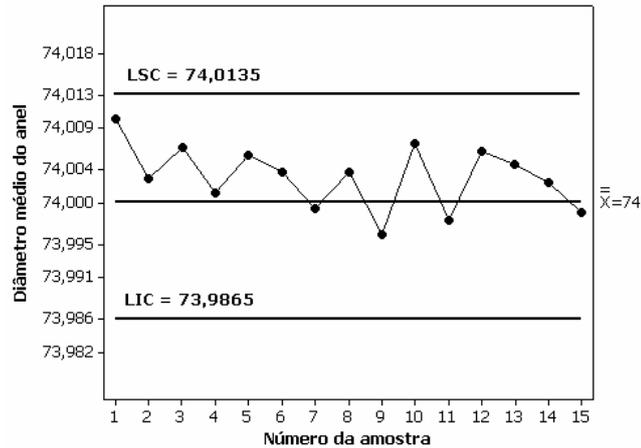


Figura 2.6: Gráfico de controle \bar{x} para o diâmetro do anel de pistão

Note que a escolha dos limites de controle é equivalente ao estabelecimento da região crítica para o teste de hipótese:

$$\begin{cases} H_0 : \mu = 74 \\ H_1 : \mu \neq 74 \end{cases}$$

onde $\sigma = 0,01$ é conhecido. Resumidamente, o gráfico de controle testa a hipótese nula repetidamente, em pontos diferentes do tempo.

Nesse exemplo, onde se usou os limites de controle três-sigma, a probabilidade do erro do tipo I é 0,0027. Dessa forma, o valor do *ARL in control* é dado por:

$$ARL \text{ in control} = \frac{1}{\alpha} = \frac{1}{0,0027} \approx 370$$

isto é, mesmo que o processo permaneça sob controle, um alarme falso será emitido a cada 370 amostras, em média.

Agora, suponha que a média do processo tivesse de fato se alterado de 74 para $74 + k \sigma = 74 + k \times 0,01$. Então o poder do teste seria dado por:

$$Poder = P(\bar{X} > LSC \mid \mu = 74 + k \times 0,01) + P(\bar{X} < LIC \mid \mu = 74 + k \times 0,01)$$

A Tabela 2.1 sumariza os valores de poder para diferentes valores de k .

Tabela 2.1: Valores de poder e *ARL out of control* para o exemplo de aplicação

k	<i>Poder</i>	<i>ARL out of control</i>
0,25	0,007	142,857
0,50	0,029	34,483
1	0,218	4,587
1,5	0,631	1,584
2,0	0,926	1,079

Observe que à medida que o valor de μ se distancia do valor alvo μ_0 , maior é a rapidez de detecção de descontroles no processo, como pode ser visto pela Tabela 2.1. Por exemplo, para $k=1$ teríamos que observar aproximadamente 5 amostras do processo para que o alarme verdadeiro fosse detectado. Para o caso em que $k=0$, o poder do teste é igual ao nível de significância do teste (0,0027):

$$\begin{aligned}
 \text{Poder} &= P(\bar{X} > LSC \mid \mu = 74 + k \times 0,01) + P(\bar{X} < LIC \mid \mu = 74 + k \times 0,01) \\
 &= P(\bar{X} > 74,0135 \mid \mu = 74) + P(\bar{X} < 73,9865 \mid \mu = 74) \\
 &= P(Z > 3) + P(Z < -3) = 0,0027
 \end{aligned}$$

Capítulo 3

Testes de hipótese para matriz de covariâncias para amostras independentes

Neste capítulo apresentaremos os testes de hipótese para a matriz de covariâncias que serão tratados nesta dissertação. Como todos estão fundamentados na suposição de normalidade multivariada, esta será introduzida na Seção 3.1, sendo os testes apresentados nas Seções 3.2 a 3.7. Todos os testes tratados neste capítulo foram formulados para processos onde os vetores de observações são independentes.

3.1 Distribuição normal multivariada

A função densidade da distribuição normal multivariada é uma generalização do caso univariado, porém, no caso multivariado trabalhamos com duas ou mais variáveis aleatórias simultaneamente. Para o caso univariado a função densidade de probabilidade de uma variável com distribuição normal, com média μ e variância σ^2 é dada por:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right\}, \quad -\infty < x < \infty; \mu \in (-\infty; \infty); \sigma > 0 \quad (3.1)$$

Seja $\mathbf{X} = (X_1, X_2, \dots, X_p)^T$ o vetor aleatório de dimensão p . Dizemos que o vetor \mathbf{X} tem uma distribuição normal p -variada com parâmetros $\boldsymbol{\mu}$ e $\Sigma_{p \times p}$, e denota-se por $\mathbf{X} \sim N_p(\boldsymbol{\mu}; \Sigma_{p \times p})$, se a função densidade de \mathbf{X} for dada por:

$$f(x_1, x_2, \dots, x_p) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (x - \boldsymbol{\mu})^T \Sigma^{-1} (x - \boldsymbol{\mu}) \right\} \quad (3.2)$$

onde $-\infty < x_j < \infty$, $-\infty < \mu_j < \infty$, $\boldsymbol{\mu} = (\mu_1, \mu_2 \dots \mu_p)^T$ representa o vetor de médias de \mathbf{X} e $\Sigma_{p \times p}$ a matriz de covariâncias de \mathbf{X} , positiva definida; $j=1, 2, \dots, p$ (“ T ” denota vetor transposto).

A distribuição normal bivariada ($p=2$) é um caso particular da distribuição normal multivariada e um bom exemplo para ilustrarmos algumas de suas propriedades. Considere X_1 e X_2 duas variáveis aleatórias com matriz de covariâncias $\Sigma_{2 \times 2}$ e vetor de médias $\boldsymbol{\mu}$, como mostrado em (3.3).

$$\Sigma_{2 \times 2} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} ; \boldsymbol{\mu} = (\mu_1, \mu_2)^T \quad (3.3)$$

onde σ_{jj} representa a variância da variável X_j e σ_{kj} representa a covariância entre as variáveis X_k e X_j , $k \neq j$.

Podemos especificar ainda a covariância em função do coeficiente de correlação ρ_{12} (ou simplesmente ρ) da seguinte forma: $\sigma_{12} = \rho \sqrt{\sigma_{11} \sigma_{22}}$. Antes de apresentarmos a distribuição normal bivariada de X_1 e X_2 devemos obter algumas quantidades mostradas a seguir. Tem-se que:

$$\Sigma_{2 \times 2} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \rho \sqrt{\sigma_{11} \sigma_{22}} \\ \rho \sqrt{\sigma_{11} \sigma_{22}} & \sigma_{22} \end{pmatrix} \quad (3.4)$$

sendo seu determinante igual a $|\Sigma| = \sigma_{11} \sigma_{22} (1 - \rho^2)$. A matriz inversa, Σ^{-1} , é dada por:

$$\Sigma^{-1} = \frac{1}{1 - \rho^2} \begin{pmatrix} 1/\sigma_{11} & -\rho/\sqrt{\sigma_{11} \sigma_{22}} \\ -\rho/\sqrt{\sigma_{11} \sigma_{22}} & 1/\sigma_{22} \end{pmatrix} \quad (3.5)$$

A forma quadrática $(x - \boldsymbol{\mu})^T \Sigma_{2 \times 2}^{-1} (x - \boldsymbol{\mu})$ pode ser expandida por:

$$(x - \boldsymbol{\mu})^T \Sigma^{-1} (x - \boldsymbol{\mu}) = \frac{1}{1 - \rho^2} \left[\left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}} \right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}} \right) \left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}} \right) + \left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}} \right)^2 \right] \quad (3.6)$$

Utilizando as informações anteriores temos que a função densidade da distribuição normal bivariada é dada por:

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho^2)}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}} \right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}} \right) \left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}} \right) + \left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}} \right)^2 \right] \right\} \quad (3.7)$$

onde $x_1, x_2 \in (-\infty, \infty)$, $\mu \in \mathfrak{R}^2$, $\sigma_{ii} \in \mathfrak{R}_+$ para $i=1,2$; e $|\rho| < 1$.

Ilustramos o gráfico da função densidade da normal bivariada na Figura 3.1 para uma situação particular na qual há alta correlação entre as duas variáveis aleatórias ($\rho = 0,85$). O vetor de médias da distribuição é $\mu=(0,0)^T$ e as variâncias são iguais a $\sigma_{11} = \sigma_{22} = 1$.

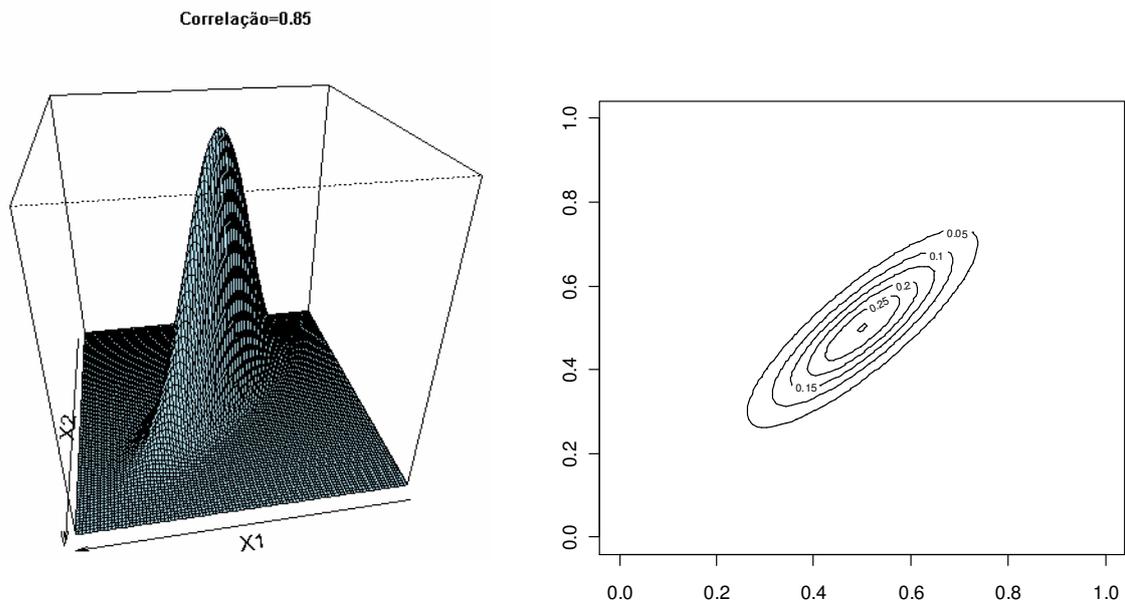


Figura 3.1: Função densidade de uma normal bivariada com correlação igual a 0,85

Observe que a correlação entre as variáveis é um fator que influencia na forma geométrica da função da distribuição normal bivariada. Note que existe uma concentração da densidade sobre a linha reta imaginária que passa sobre a diagonal do primeiro quadrante do plano cartesiano definido pelas variáveis X_1 e X_2 . Esse fato mostra que quando ρ é próximo de 1 ou -1, a superfície fica mais concentrada ao redor do vetor de médias da distribuição. À medida que o coeficiente de correlação entre as variáveis tende a zero, a superfície fica mais dispersa ao redor do vetor de médias, como pode ser observado pela Figura 3.2.

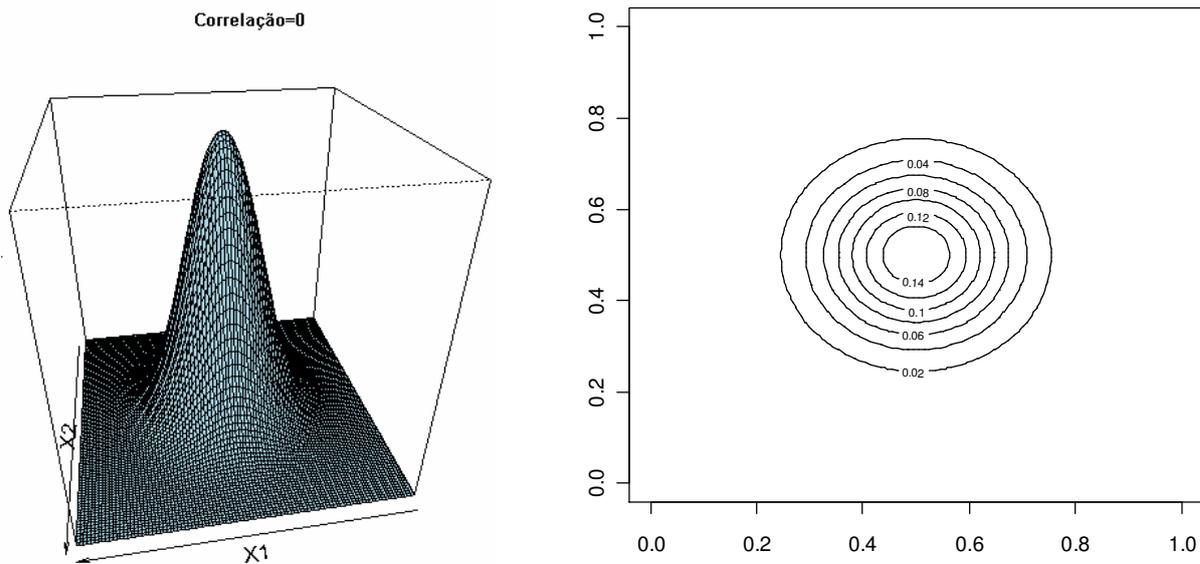


Figura 3.2: Função densidade de uma normal bivariada com correlação igual a 0

A distribuição normal multivariada é uma suposição dos testes que apresentaremos nas próximas seções. Existem alguns testes estatísticos e procedimentos gráficos para avaliar a suposição de normalidade multivariada, porém, não serão abordados nessa dissertação. Para mais detalhes desses procedimentos ver Mingoti (2005) e Jobson (1992).

3.2 Teste do determinante

Sabemos que a variabilidade de um processo multivariado é resumida pela matriz de covariâncias Σ_{pxp} . O teste do determinante baseia-se na variância amostral generalizada $|S|$, onde S é a matriz de covariâncias amostral. Essa estatística, que é o determinante de S , é uma medida amplamente usada para avaliar a dispersão multivariada. Suponha que estamos interessados em testar a hipótese $H_0: \Sigma = \Sigma_0$ versus $H_1: \Sigma \neq \Sigma_0$, onde Σ_0 é uma matriz de covariâncias pré-especificada (por exemplo, obtida na Fase 1 de controle de qualidade). Dessa forma, o gráfico de controle para o determinante da matriz de covariâncias teórica $|\Sigma|$ é definido por (Montgomery, 2004):

$$\begin{aligned}
 LSC &= |\Sigma_0| (b_1 + 3\sqrt{b_2}) \\
 LC &= b_1 |\Sigma_0| \\
 LIC &= \max(0; |\Sigma_0| (b_1 - 3\sqrt{b_2}))
 \end{aligned}
 \tag{3.8}$$

onde:

$$b_1 = \frac{1}{(n-1)^p} \prod_{i=1}^p (n-i)$$

$$b_2 = \frac{1}{(n-1)^{2p}} \prod_{i=1}^p (n-i) \left[\prod_{j=1}^p (n-j+2) - \prod_{j=1}^p (n-j) \right]$$

Vale ressaltar que os limites de controle em (3.8) são obtidos pelo uso da distribuição normal assintótica de $|S|$. Djauhari (2005) mostra que:

$$E(|S|) = b_1 |\Sigma| \quad e \quad Var(|S|) = b_2 |\Sigma|^2 \quad (3.9)$$

Assim, $|S|/b_1$ é um estimador não viciado de $|\Sigma|$, onde $|S|$ é o determinante da matriz de covariâncias amostral. Na próxima seção, apresentaremos uma modificação no teste do determinante proposta por Djauhari (2005). Na seção 3.4 mostraremos um estudo no qual a distribuição exata de $|S|$ é comparada com a distribuição assintótica.

3.3 Teste do determinante proposto por Djauhari (2005)

Djauhari (2005) propõe uma correção na forma de determinar os limites de controle para o teste do determinante. Segundo os autores essas modificações resultam numa maior sensibilidade em identificar mudanças na matriz de covariâncias do processo. Os limites propostos por Djauhari (2005) são dados em (3.10).

$$\begin{aligned} LSC &= |\Sigma_0| \left(\frac{b_1}{b_3} + 3 \sqrt{\frac{b_2}{b_3^2 + b_4}} \right) \\ LC &= |\Sigma_0| \frac{b_1}{b_3} \\ LIC &= \max \left(0; |\Sigma_0| \left(\frac{b_1}{b_3} - 3 \sqrt{\frac{b_2}{b_3^2 + b_4}} \right) \right) \end{aligned} \quad (3.10)$$

onde:

$$b_3 = \frac{1}{(n-1)^p} \prod_{i=1}^p \{(n-1) - i + 1\}$$

$$b_4 = \frac{1}{(n-1)^{2p}} \prod_{i=1}^p \{(n-1) - i + 1\} \left[\prod_{j=1}^p \{(n-1) - j + 3\} - \prod_{j=1}^p \{(n-1) - j + 1\} \right]$$

e b_1 e b_2 são dados como anteriormente.

Embora a proposta de Djauhari (2005) seja a de melhorar a qualidade do teste do determinante, vale ressaltar que o uso de $|S|$ como uma medida de dispersão multivariada tem algumas limitações. Em geral, algumas mudanças na estrutura da matriz de covariâncias não acarretam mudança da variância generalizada. Ainda que a variância amostral generalizada seja uma medida da dispersão multivariada amplamente usada, lembre-se de que é uma representação escalar relativamente simplista de um problema multivariado complexo, ou seja, resumimos toda a estrutura da matriz de covariâncias em um único valor. Em muitos casos essa sumarização pode nos levar a resultados errôneos, já que existem matrizes de covariâncias completamente diferentes, mas com mesmo determinante.

Geometricamente, o valor absoluto do determinante é igual volume do elipsóide delimitado pelos vetores que o compõe. Por exemplo, no espaço bidimensional o determinante é igual a área do paralelogramo determinado pelos vetores $\mathbf{u}=(u_1,u_2)$ e $\mathbf{v}=(v_1,v_2)$, ou ainda, igual ao volume do paralelepípedo no espaço tridimensional determinado pelos vetores $\mathbf{u}=(u_1,u_2,u_3)$, $\mathbf{v}=(v_1,v_2,v_3)$, $\mathbf{w}=(w_1,w_2,w_3)$, como mostra a Figura 3.3. Dessa forma, podemos ter vetores diferentes, mas que produzem a mesma área ou volume. Nessa situação, teríamos matrizes de covariâncias completamente diferentes, mas que produzem o mesmo determinante.

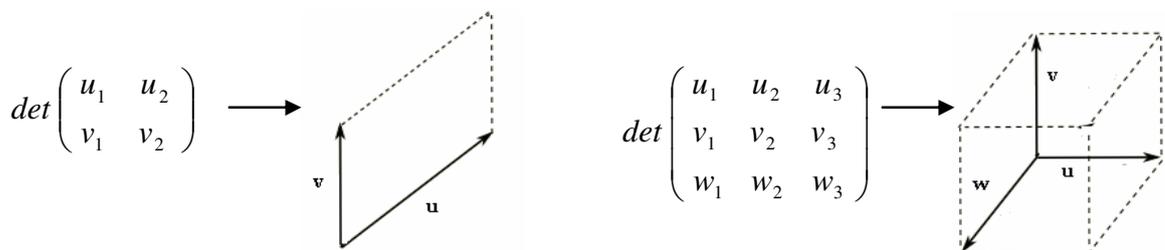


Figura 3.3: Interpretação geométrica do determinante

Como exemplo, observe as seguintes matrizes de covariâncias (Montgomery, 2004):

$$S_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad S_2 = \begin{pmatrix} 2,32 & 0,40 \\ 0,40 & 0,50 \end{pmatrix}$$

Observe que $|S_1|=|S_2|=1$, e, no entanto as duas matrizes transmitem informações consideravelmente diferentes sobre a variabilidade do processo e sobre a correlação entre as duas variáveis. Note que para S_1 a correlação entre as duas variáveis é zero, enquanto que para S_2 é 0,371. Por isso, deve-se ter muito cuidado ao utilizar essa metodologia para monitorar a variabilidade do processo.

Note que até o momento abordamos o caso em que o vetor de médias é desconhecido, ou seja, a matriz S é estimada com base na amostra de observações, sem nenhum conhecimento da média do processo. Porém, em algumas situações práticas o vetor de médias do processo já é conhecido. Dessa forma, ao invés de utilizarmos a matriz de covariância amostral S para manter o processo sob controle utilizamos a matriz S_μ que é dada por:

$$S_\mu = \frac{\sum_{j=1}^n (x_j - \mu) (x_j - \mu)^T}{n} \quad (3.11)$$

Nessa dissertação abordaremos os testes estatísticos para matriz de covariâncias sempre considerando os casos em que o vetor de médias é conhecido e aqueles em que é desconhecido.

Para ilustrar o controle da estrutura de variabilidade em um processo multivariado, pelo método do determinante, utilizaremos um exemplo adaptado de Montgomery (2004) e que também será usado em exemplos das próximas seções.

3.3.1 Exemplo de aplicação

A força de resistência (X_1) e o diâmetro (X_2) de uma fibra têxtil são duas características da qualidade importantes que devem ser controladas conjuntamente. Com o intuito de manter a variabilidade do processo sob controle, as seguintes hipóteses foram construídas:

$$\begin{cases} H_0 : \Sigma = \Sigma_0 = \begin{pmatrix} 1,23 & 0,79 \\ 0,79 & 0,83 \end{pmatrix} \\ H_1 : \Sigma \neq \Sigma_0 \end{cases}$$

onde o determinante de Σ_0 é: $|\Sigma_0| = 0,3968$. Foi selecionada, do processo de produção, uma amostra de tamanho $n=10$, obtendo-se a seguinte matriz de covariâncias amostral:

$$S = \begin{pmatrix} 1,25 & 0,80 \\ 0,80 & 0,87 \end{pmatrix}$$

de modo que $|S| = 0,4475$. As constantes b_1 , b_2 , b_3 e b_4 são dadas por:

$$b_1 = \frac{1}{(n-1)^p} \prod_{i=1}^p (n-i) = \frac{1}{9^2} (9)(8) = 0,8889$$

$$b_2 = \frac{1}{(n-1)^{2p}} \prod_{i=1}^p (n-i) \left[\prod_{j=1}^p (n-j+2) - \prod_{j=1}^p (n-j) \right] = \frac{1}{9^4} (9)(8) [(11)(10) - (9)(8)] = 0,4170$$

$$b_3 = \frac{1}{(n-1)^p} \prod_{i=1}^p \{(n-1) - i + 1\} = \frac{1}{9^2} (9)(8) = 0,8889$$

$$b_4 = \frac{1}{(n-1)^{2p}} \prod_{i=1}^p \{(n-1) - i + 1\} \left[\prod_{j=1}^p \{(n-1) - j + 3\} - \prod_{j=1}^p \{(n-1) - j + 1\} \right] = 0,4170$$

Com base nos limites de controle propostos por Montgomery (2004) temos que:

$$LSC = |\Sigma_0| (b_1 + 3\sqrt{b_2}) = 0,3968 (0,8889 + 3\sqrt{0,4170}) = 1,1214$$

$$LC = b_1 |\Sigma_0| = 0,8889 \times 0,3968 = 0,3527$$

$$LIC = \max(0; |\Sigma_0| (b_1 - 3\sqrt{b_2})) = \max(0; 0,3968 (0,8889 - 3\sqrt{0,4170})) = 0$$

Como o valor de $|S|=0,4475$ está dentro dos limites de controle temos evidências de que a estrutura de variabilidade do processo está de acordo com o que foi postulado em H_0 .

Quando construímos os limites de controle baseados em Djauhari (2005) temos os seguintes resultados:

$$LSC = |\Sigma_0| \left(\frac{b_1}{b_3} + 3\sqrt{\frac{b_2}{b_3^2 + b_4}} \right) = 0,3968 \left(1 + 3\sqrt{\frac{0,4170}{(0,8889)^2 + 0,4170}} \right) = 1,0965$$

$$LC = |\Sigma_0| \frac{b_1}{b_3} = 0,3968$$

$$LIC = \max \left(0; |\Sigma_0| \left(\frac{b_1}{b_3} - 3\sqrt{\frac{b_2}{b_3^2 + b_4}} \right) \right) = \max \left(0; 0,3968 \left(1 - 3\sqrt{\frac{0,4170}{(0,8889)^2 + 0,4170}} \right) \right) = 0$$

Como no caso anterior, o valor de $|S|=0,4475$ está dentro dos limites de controle, dessa forma, podemos concluir que a estrutura de variabilidade do processo está sob controle.

3.4 Estudo da distribuição empírica de |S|

Os limites de controle dos testes do determinante tradicional e do proposto em Djauhari (2005) são construídos usando a aproximação da distribuição de |S| pela distribuição normal. No entanto, se o vetor X tem distribuição normal p-variada com vetor de médias μ e matriz de covariâncias $\Sigma_{p \times p}$, a distribuição exata de |S| é dada por:

$$\left[\frac{|\Sigma|}{(n-1)^p} \right] \prod_{k=1}^p Z_k \quad (3.12)$$

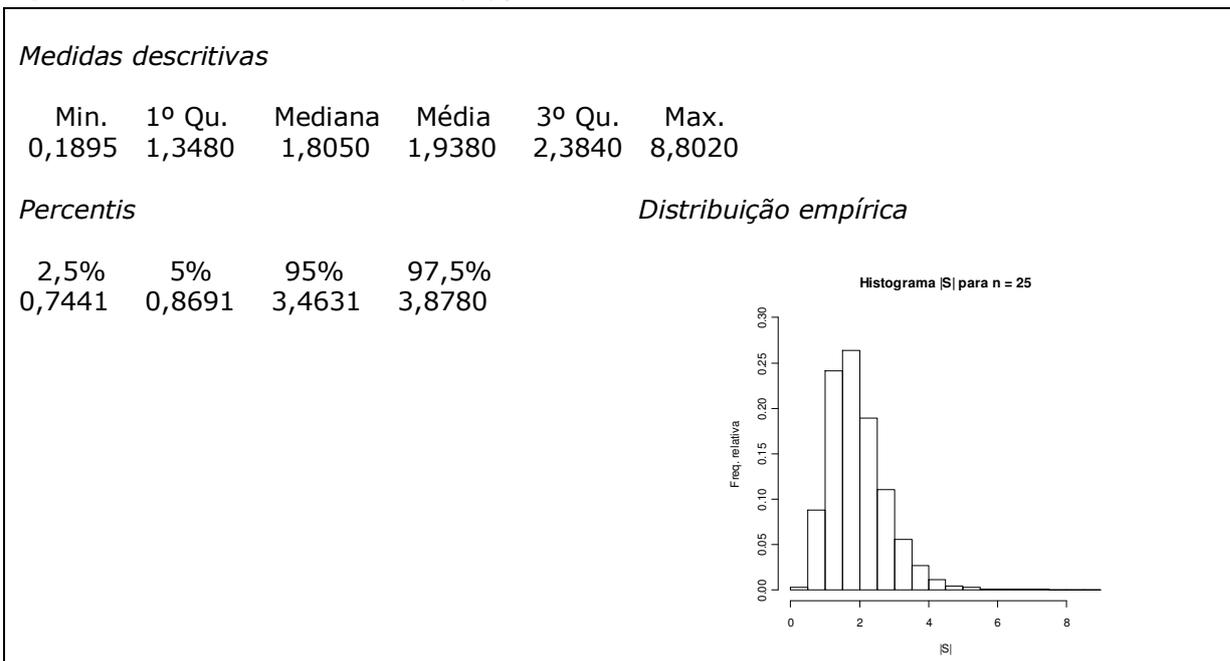
onde Z_k são independentes e $Z_k \sim \chi_{n-k}^2$; $k=1, 2, \dots, p$. Para mais detalhes sobre a distribuição exata de |S| ver Anderson, (1958;2003). Deste modo, decidimos avaliar a qualidade da aproximação normal para |S| no caso de observações independentes. Vale ressaltar a importância da verificação da distribuição do determinante da matriz de covariâncias amostral (|S|) uma vez que a utilização da distribuição assintótica normal para amostras pequenas pode gerar erros na determinação dos limites de controle.

No processo de simulação utilizamos como suporte o software R 2.9.0 juntamente com pacote MASS. O pacote MASS possui funções para análise multivariada de dados. No contexto que estamos trabalhando utilizamos a função *mvnorm()* para gerarmos vetores provenientes de uma distribuição normal p-variada. Para efeitos de simulação geramos 100.000 amostras independentes de tamanho n de uma distribuição normal bivariada com os seguintes parâmetros:

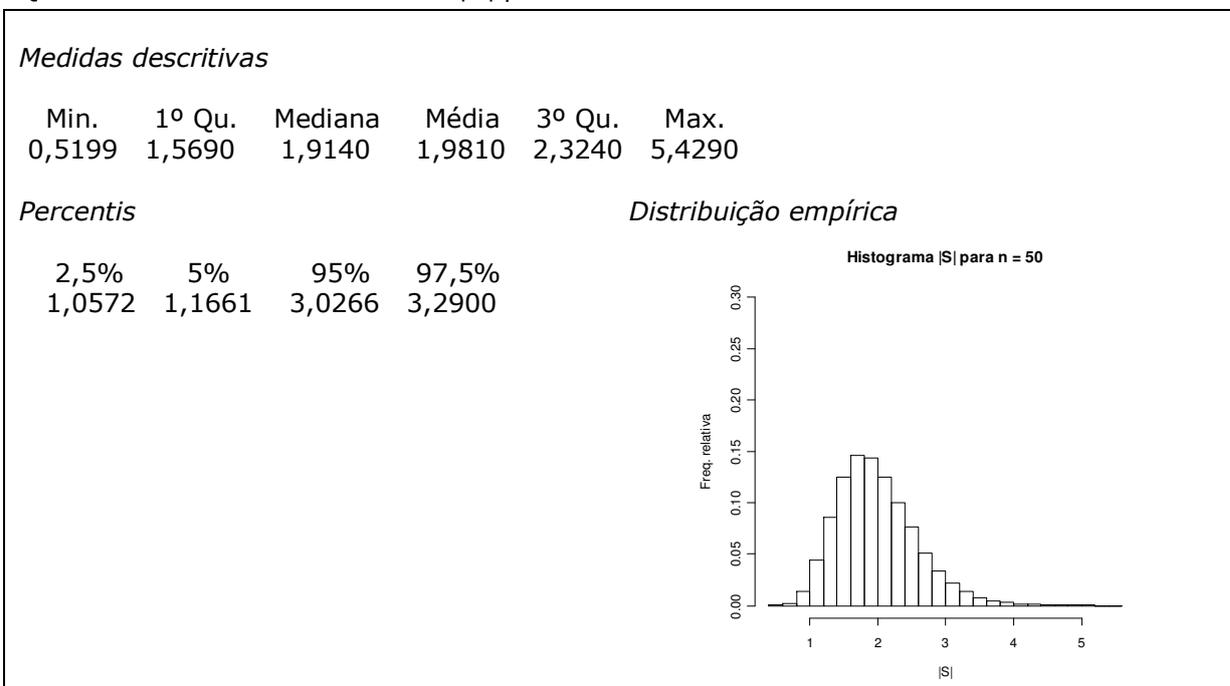
$$\mu = (0; 0)^T \text{ e } \Sigma = \begin{pmatrix} 1,3333 & 0,7692 \\ 0,7692 & 1,9608 \end{pmatrix}, \text{ sendo } |\Sigma| = 2,0227.$$

Para cada amostra determinou-se o valor de |S| construindo-se no final a distribuição empírica de |S|. Foram consideradas amostras de tamanhos 25, 50, 100, 200 e 1000. Para cada uma das situações apresentamos algumas medidas descritivas dos valores de |S|, além dos percentis e do histograma da distribuição empírica. Os resultados obtidos encontram-se nos Quadros 3.1 a 3.5. Observe que à medida que aumentamos o valor de n obtemos melhores resultados para o estimador |S| e uma maior aproximação da distribuição normal. Note que para $n = 200$ já obtemos bons resultados.

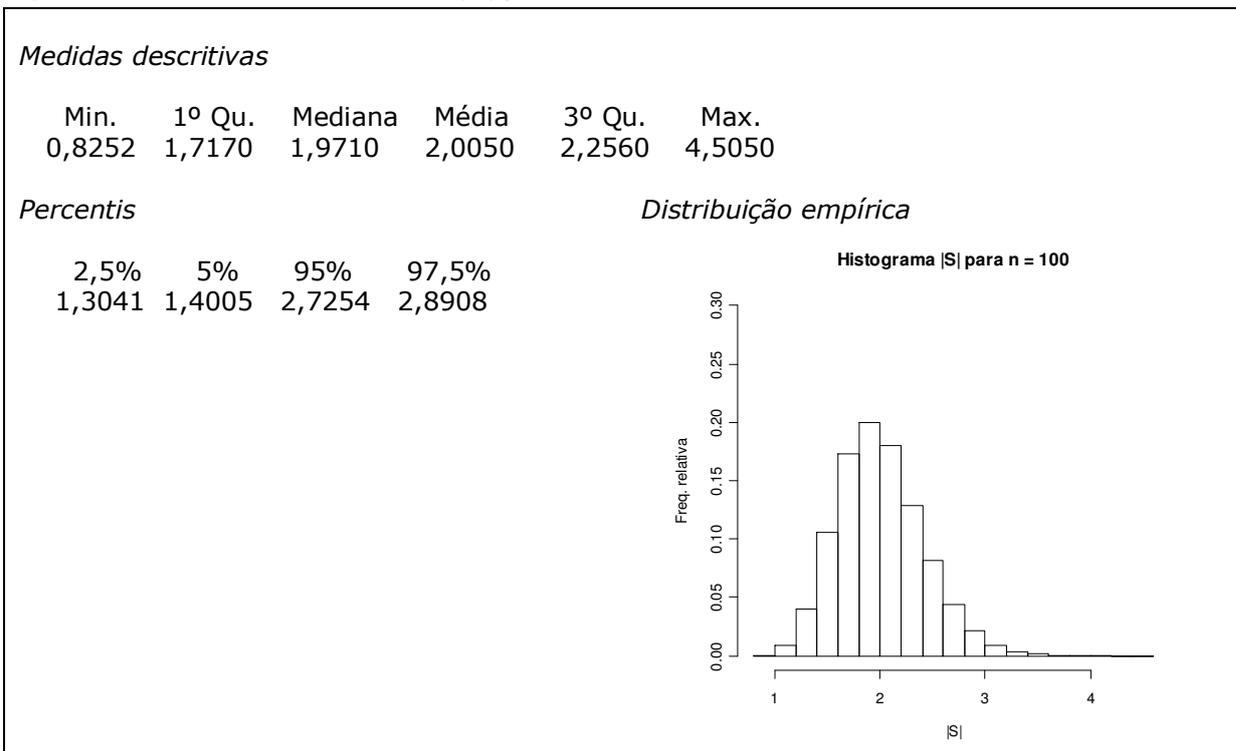
Quadro 3.1: Medidas descritivas de |S| para amostras de tamanho n = 25



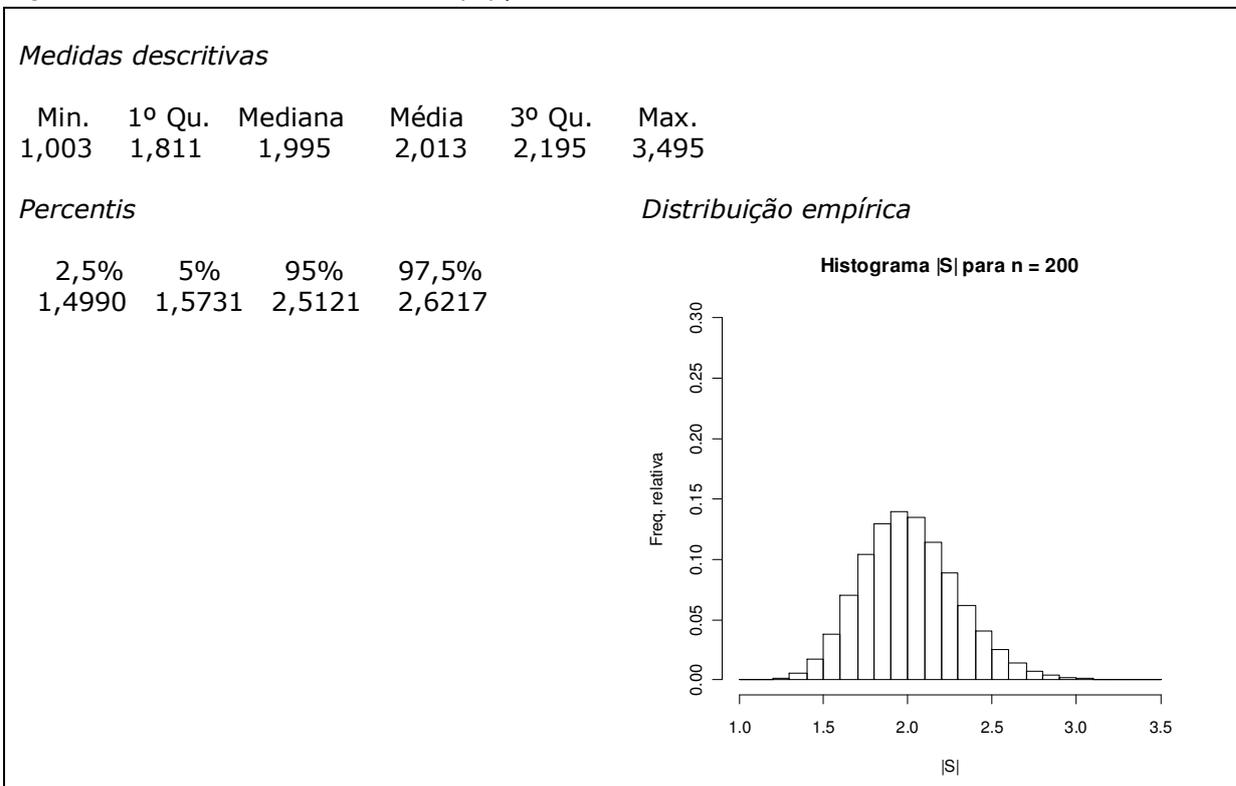
Quadro 3.2: Medidas descritivas de |S| para amostras de tamanho n = 50



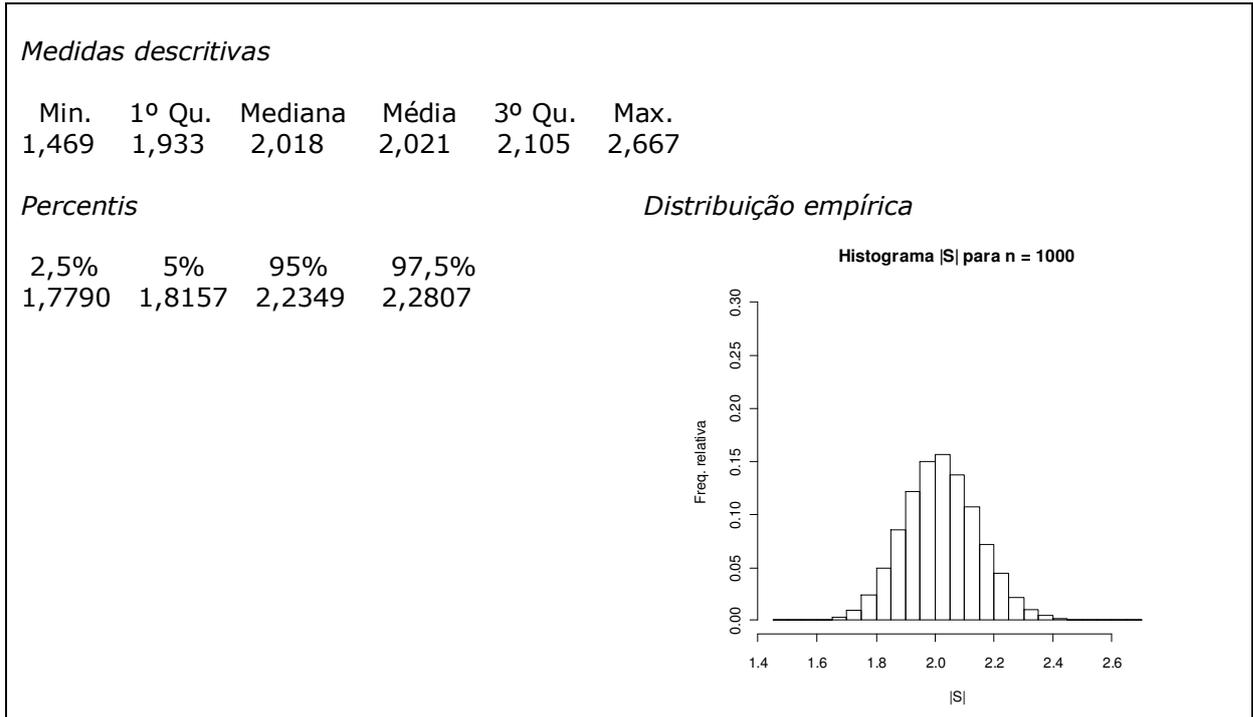
Quadro 3.3: Medidas descritivas de |S| para amostras de tamanho n = 100



Quadro 3.4: Medidas descritivas de |S| para amostras de tamanho n = 200



Quadro 3.5: Medidas descritivas de |S| para amostras de tamanho n = 1000



3.5 Teste da razão de verossimilhança

Seja uma amostra aleatória de tamanho n , X_1, X_2, \dots, X_n de uma distribuição normal p -variada com vetor de médias $\boldsymbol{\mu}$ e matriz de covariâncias $\Sigma_{p \times p}$, sendo ambos desconhecidos, e $\mathbf{X}_j = (X_{j1}, X_{j2}, \dots, X_{jp})^T$, $j=1, 2, \dots, n$. Estamos interessados em testar a hipótese: $H_0: \Sigma = \Sigma_0$ versus $H_1: \Sigma \neq \Sigma_0$, sendo que a matriz Σ_0 (alvo) é completamente especificada em H_0 . O primeiro passo para construirmos o teste da razão de verossimilhança será encontrar os estimadores de máxima verossimilhança de $\boldsymbol{\mu}$ e Σ . Sabemos que os vetores aleatórios \mathbf{X}_j ($j = 1, 2, \dots, n$) são independentes e identicamente distribuídos. Assim, a distribuição conjunta desses vetores é dada pelo produto das funções individuais (ver Ferreira, 2008). Essa função de densidade é conhecida como função de verossimilhança. A função de verossimilhança para o modelo normal multivariado é dada como em (3.13).

$$\begin{aligned}
L(X; \mu; \Sigma) &= \prod_{j=1}^n f_X(X_j) \\
&= \prod_{j=1}^n \left[\frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (x_j - \mu)^T \Sigma^{-1} (x_j - \mu) \right\} \right] \\
&= (2\pi)^{-\frac{np}{2}} |\Sigma|^{\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{j=1}^n (x_j - \mu)^T \Sigma^{-1} (x_j - \mu) \right\}
\end{aligned} \tag{3.13}$$

O máximo dessa função pode ser obtido tomando-se as derivadas parciais de primeira ordem em relação a cada parâmetro, igualando as funções obtidas a zero e resolvendo o sistema de equações formado. Os estimadores resultantes são os estimadores de máxima verossimilhança. Porém, para facilitar os cálculos vamos trabalhar com a função de log-verossimilhança. Vale ressaltar que o máximo das funções de verossimilhança e log-verossimilhança, em relação aos estimadores obtidos, são exatamente os mesmos. Isto decorre do fato da função logarítmica ser monótona crescente. Para esse caso denominaremos a função de log-verossimilhança por g , que é dada por:

$$g(X; \mu; \Sigma) = -\frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln|\Sigma| - \frac{1}{2} \sum_{j=1}^n (x_j - \mu)^T \Sigma^{-1} (x_j - \mu) \tag{3.14}$$

Primeiramente vamos tomar a derivada de primeira ordem de g em relação à μ . Antes disso, vale lembrar da seguinte propriedade de derivada matricial:

$$\frac{\partial B^T A B}{\partial B} = 2AB, \text{ sendo } A \text{ e } B \text{ duas matrizes quaisquer.} \tag{3.15}$$

Assim, utilizaremos esse resultado para encontrarmos o estimador de máxima verossimilhança de μ :

$$\frac{\partial g(X; \mu; \Sigma)}{\partial \mu} = \sum_{j=1}^n \Sigma^{-1} (x_j - \mu) \tag{3.16}$$

onde x_j é um vetor de dimensão $p \times 1$.

Igualando a zero e resolvendo a equação resultante obtemos facilmente o estimador de máxima verossimilhança de $\boldsymbol{\mu}$ que é dado por:

$$\hat{\boldsymbol{\mu}} = \frac{\sum_{j=1}^n x_j}{n} = \bar{X} \quad (3.17)$$

sendo \bar{X} um vetor $p \times 1$.

Como já era esperado, o estimador de máxima verossimilhança do vetor de médias é o vetor de médias amostral. Antes de tomarmos a derivada de primeira ordem com relação à matriz de covariâncias Σ , vamos rearranjar o expoente do número neperiano na equação da função de verossimilhança. Assim temos:

$$\begin{aligned} & \sum_{j=1}^n (x_j - \boldsymbol{\mu})^T \Sigma^{-1} (x_j - \boldsymbol{\mu}) = \sum_{j=1}^n (x_j - \bar{x} + \bar{x} - \boldsymbol{\mu})^T \Sigma^{-1} (x_j - \bar{x} + \bar{x} - \boldsymbol{\mu}) \\ = & \sum_{j=1}^n (x_j - \bar{x})^T \Sigma^{-1} (x_j - \bar{x}) + \sum_{j=1}^n (x_j - \bar{x})^T \Sigma^{-1} (\bar{x} - \boldsymbol{\mu}) + \sum_{j=1}^n (\bar{x} - \boldsymbol{\mu})^T \Sigma^{-1} (x_j - \bar{x}) + \sum_{j=1}^n (\bar{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\bar{x} - \boldsymbol{\mu}) \\ = & \sum_{j=1}^n (x_j - \bar{x})^T \Sigma^{-1} (x_j - \bar{x}) + 2 \sum_{j=1}^n (x_j - \bar{x})^T \Sigma^{-1} (\bar{x} - \boldsymbol{\mu}) + \sum_{j=1}^n (\bar{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\bar{x} - \boldsymbol{\mu}) \end{aligned} \quad (3.18)$$

Como os elementos dessa soma são formas quadráticas e, portanto, escalares, podemos aplicar o traço, como em (3.19).

$$= \sum_{j=1}^n \text{tr}((x_j - \bar{x})^T \Sigma^{-1} (x_j - \bar{x})) + 2 \sum_{j=1}^n \text{tr}((x_j - \bar{x})^T \Sigma^{-1} (\bar{x} - \boldsymbol{\mu})) + \sum_{j=1}^n \text{tr}((\bar{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\bar{x} - \boldsymbol{\mu})) \quad (3.19)$$

Levando em consideração a propriedade: $\text{tr}(AB) = \text{tr}(BA)$, onde A e B são duas matrizes, tem-se o seguinte resultado:

$$= \sum_{j=1}^n \text{tr}[\Sigma^{-1} (x_j - \bar{x}) (x_j - \bar{x})^T] + 2 \sum_{j=1}^n \text{tr}[\Sigma^{-1} (x_j - \bar{x}) (\bar{x} - \boldsymbol{\mu})^T] + \sum_{j=1}^n \text{tr}[\Sigma^{-1} (\bar{x} - \boldsymbol{\mu}) (\bar{x} - \boldsymbol{\mu})^T] \quad (3.20)$$

Note que o termo central é uma matriz nula, assim temos que:

$$\sum_{j=1}^n (x_j - \boldsymbol{\mu})^T \Sigma^{-1} (x_j - \boldsymbol{\mu}) = \sum_{j=1}^n \text{tr}[\Sigma^{-1} (x_j - \bar{x}) (x_j - \bar{x})^T] + \sum_{j=1}^n \text{tr}[\Sigma^{-1} (\bar{x} - \boldsymbol{\mu}) (\bar{x} - \boldsymbol{\mu})^T] \quad (3.21)$$

Utilizando a propriedade de que a soma de traços é igual ao traço da soma, podemos desenvolver ainda mais a expressão anterior:

$$\sum_{j=1}^n (x_j - \mu)^T \Sigma^{-1} (x_j - \mu) = \text{tr} \left[\Sigma^{-1} \sum_{j=1}^n (x_j - \bar{x}) (x_j - \bar{x})^T \right] + \text{tr} \left[n \Sigma^{-1} (\bar{x} - \mu) (\bar{x} - \mu)^T \right] \quad (3.22)$$

Assim, podemos reescrever a função de log-verossimilhança utilizando o fato que:

$$(n-1) S = \sum_{j=1}^n (x_j - \bar{x}) (x_j - \bar{x})^T \quad (3.23)$$

Finalmente a função de log-verossimilhança pode ser reescrita como em (3.24).

$$g(X; \mu; \Sigma) = -\frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln |\Sigma| - \frac{1}{2} \text{tr} \left[(n-1) \Sigma^{-1} S \right] - \frac{1}{2} \text{tr} \left[n \Sigma^{-1} (\bar{x} - \mu) (\bar{x} - \mu)^T \right] \quad (3.24)$$

Tomando a derivada de primeira ordem da função de log-verossimilhança em relação à Σ temos:

$$\frac{\partial g(X; \mu; \Sigma)}{\partial \Sigma} = -\frac{n}{2} (\Sigma^{-1})^T + \frac{(n-1)}{2} \Sigma^{-1} S \Sigma^{-1} + \frac{n}{2} \Sigma^{-1} (\bar{x} - \mu) (\bar{x} - \mu)^T \Sigma^{-1} \quad (3.25)$$

$$= -\frac{n}{2} \Sigma^{-1} + \frac{(n-1)}{2} \Sigma^{-1} S \Sigma^{-1} + \frac{n}{2} \Sigma^{-1} (\bar{x} - \mu) (\bar{x} - \mu)^T \Sigma^{-1} \quad (3.26)$$

Igualando a zero e substituindo o estimador de μ encontrado anteriormente temos que o estimador de Σ é dado como em (3.27).

$$-\frac{n}{2} \hat{\Sigma}^{-1} + \frac{(n-1)}{2} \hat{\Sigma}^{-1} S \hat{\Sigma}^{-1} + \frac{n}{2} \hat{\Sigma}^{-1} (\bar{x} - \bar{x}) (\bar{x} - \bar{x})^T \hat{\Sigma}^{-1} = 0$$

$$-\frac{n}{2} \hat{\Sigma}^{-1} + \frac{(n-1)}{2} \hat{\Sigma}^{-1} S \hat{\Sigma}^{-1} = 0$$

$$-n \hat{\Sigma}^{-1} + (n-1) \hat{\Sigma}^{-1} S \hat{\Sigma}^{-1} = 0$$

$$-n \hat{\Sigma} \hat{\Sigma}^{-1} \hat{\Sigma} + (n-1) \hat{\Sigma} \hat{\Sigma}^{-1} S \hat{\Sigma}^{-1} \hat{\Sigma} = 0$$

$$n \hat{\Sigma} = (n-1) S \quad \text{sendo} \quad \hat{\Sigma} = \frac{(n-1) S}{n} = S_n \quad (3.27)$$

Assim, temos que os estimadores do vetor de médias e da matriz de covariâncias são \bar{X} e S_n , respectivamente. Uma vez identificado os estimadores de máxima verossimilhança para μ e Σ o próximo passo será construir o teste da razão de verossimilhança.

Considere $L_{\Omega_0}(X; \mu; \Sigma_0)$ o máximo da função de verossimilhança para o espaço restrito e $L_{\Omega}(X; \hat{\mu}; \hat{\Sigma})$, o máximo da função de verossimilhança para o espaço irrestrito. Vale lembrar que Ω corresponde ao espaço paramétrico completo e Ω_0 o espaço paramétrico sob H_0 . Dessa forma, podemos utilizar a razão de verossimilhança dada por:

$$\Lambda = \frac{L_{\Omega_0}(X; \mu; \Sigma)}{L_{\Omega}(X; \hat{\mu}; \hat{\Sigma})} \quad (3.28)$$

O teste da razão de verossimilhança baseia-se no critério de decisão sobre as hipóteses H_0 ou H_1 . Se a razão de verossimilhança em seu máximo é grande, H_0 é mais provável que H_1 e se a razão de verossimilhança for pequena H_1 deve ser escolhida. O primeiro passo para construirmos o teste descrito anteriormente será encontrar o máximo da função de verossimilhança para o modelo completo.

$$L_{\Omega}(X; \hat{\mu}; \hat{\Sigma}) = (2\pi)^{\frac{np}{2}} |S_n|^{-\frac{n}{2}} \exp \left\{ \frac{-tr(n S_n^{-1} S_n) - tr(n S_n^{-1} (\bar{x} - \bar{x}) (\bar{x} - \bar{x})^T)}{2} \right\} \quad (3.29)$$

$$L_{\Omega}(X; \hat{\mu}; \hat{\Sigma}) = (2\pi)^{\frac{np}{2}} |S_n|^{-\frac{n}{2}} \exp \left\{ \frac{-np}{2} \right\} \quad (3.30)$$

Sabemos que sob H_0 a função de verossimilhança pode ser escrita como:

$$L_{\Omega_0}(X; \mu; \Sigma_0) = (2\pi)^{\frac{np}{2}} |\Sigma_0|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} tr[(n-1) \Sigma_0^{-1} S] - \frac{1}{2} tr[n \Sigma_0^{-1} (\bar{x} - \mu) (\bar{x} - \mu)^T] \right\} \quad (3.31)$$

Como estamos restringindo o espaço paramétrico somente a $\Sigma = \Sigma_0$ temos que o estimador de máxima verossimilhança para μ é \bar{X} . Dessa forma, a função de máxima verossimilhança restrita é dada por:

$$L_{\Omega_0}(X; \mu; \Sigma_0) = (2\pi)^{\frac{np}{2}} |\Sigma_0|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} tr[n \Sigma_0^{-1} S_n] \right\} \quad (3.32)$$

Finalmente, a estatística do teste da razão de verossimilhança é dada por:

$$\Lambda = \frac{L_{\Omega_0}(X; \mu; \Sigma_0)}{L_{\Omega}(X; \hat{\mu}; \hat{\Sigma})} = \frac{(2\pi)^{-\frac{np}{2}} |\Sigma_0|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2} \text{tr}[n \Sigma_0^{-1} S_n]\right\}}{(2\pi)^{-\frac{np}{2}} |S_n|^{-\frac{n}{2}} \exp\left\{\frac{-np}{2}\right\}} \quad (3.33)$$

$$\Lambda = |\Sigma_0^{-1} S_n|^{-\frac{n}{2}} \exp\left\{\frac{-n \text{tr}(\Sigma_0^{-1} S_n) + np}{2}\right\} \quad (3.34)$$

sendo $S_n = \frac{(n-1)S}{n}$.

A determinação da distribuição nula exata de Λ é muito complexa, assim, uma alternativa é utilizar a aproximação qui-quadrado assintótica $-2 \ln(\Lambda)$. Como no modelo irrestrito são estimados $p(p+3)/2$ parâmetros e no modelo restrito p , a estatística (3.35) possui, sob H_0 , distribuição assintótica qui-quadrado com $p(p+1)/2$ graus de liberdade. Sendo assim, a estatística de teste será:

$$\begin{aligned} \chi^2 &= -2 \left\{ \frac{n \ln |\Sigma_0^{-1} S_n|}{2} - \frac{n \text{tr}(\Sigma_0^{-1} S_n) + np}{2} \right\} \\ &= n \text{tr}(\Sigma_0^{-1} S_n) - n \ln |\Sigma_0^{-1} S_n| - np \end{aligned} \quad (3.35)$$

Valores elevados de χ^2 são favoráveis à rejeição de H_0 . Alguns autores apresentam uma modificação do teste de razão de verossimilhança. Korin (1968) propôs uma correção para a estatística apresentada em (3.35) com o objetivo de melhorar a aderência para a distribuição qui-quadrado, sob H_0 , principalmente para tamanhos de amostras menores. A estatística modificada é dada por (ver Ferreira, 2008):

$$\chi^2 = \left[(n-1) - \frac{1}{6} \left(2p+1 - \frac{2}{p+1} \right) \right] [\text{tr}(\Sigma_0^{-1} S) - \ln |\Sigma_0^{-1} S| - p] \quad (3.36)$$

que possui distribuição qui-quadrado assintótica resultante, sob H_0 , com $p(p+1)/2$ graus de liberdade.

Como foi discutido na seção onde abordamos o teste do determinante, é de grande importância considerar o caso em que o vetor de médias é conhecido. Todo o desenvolvimento anterior foi feito considerando o vetor de médias desconhecido. Quando o vetor de médias é conhecido, o estimador de máxima verossimilhança da matriz de covariâncias, sob normalidade multivariada, é dado por:

$$\hat{\Sigma} = \frac{\sum_{j=1}^n (x_j - \mu)(x_j - \mu)^T}{n} = S_{\mu} \quad (3.37)$$

Considere $L_{\Omega_0}(X; \mu; \Sigma_0)$ o máximo da função de verossimilhança para o espaço restrito e $L_{\Omega}(X; \mu; \hat{\Sigma})$, o máximo da função de verossimilhança para o espaço irrestrito. Dessa forma, a estatística do teste da razão de verossimilhança, considerando μ conhecido, é dada por:

$$\Lambda = \frac{L_{\Omega_0}(X; \mu; \Sigma_0)}{L_{\Omega}(X; \hat{\mu}; \hat{\Sigma})} = \frac{(2\pi)^{-\frac{np}{2}} |\Sigma_0|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2} \text{tr}[n \Sigma_0^{-1} S_{\mu}]\right\}}{(2\pi)^{-\frac{np}{2}} |S_{\mu}|^{-\frac{n}{2}} \exp\left\{\frac{-np}{2}\right\}} \quad (3.38)$$

$$\Lambda = |\Sigma_0^{-1} S_{\mu}|^{\frac{n}{2}} \exp\left\{\frac{-n \text{tr}(\Sigma_0^{-1} S_{\mu}) + np}{2}\right\} \quad (3.39)$$

De forma análoga ao caso em que o vetor de médias é desconhecido, uma alternativa é utilizar a aproximação qui-quadrado assintótica $-2 \ln(\Lambda)$. Dessa forma temos:

$$\begin{aligned} \chi^2 &= -2 \left\{ \frac{n \ln |\Sigma_0^{-1} S_{\mu}|}{2} - \frac{n \text{tr}(\Sigma_0^{-1} S_{\mu}) + np}{2} \right\} \\ &= n \text{tr}(\Sigma_0^{-1} S_{\mu}) - n \ln |\Sigma_0^{-1} S_{\mu}| - np \end{aligned} \quad (3.40)$$

Como no modelo irrestrito são estimados $p(p+1)/2$ parâmetros e no modelo restrito nenhum parâmetro é estimado, visto que o vetor de médias e a matriz de covariâncias são conhecidos, a estatística (3.40) possui, sob H_0 , distribuição assintótica qui-quadrado com $p(p+1)/2$ graus de liberdade.

3.5.1 Exemplo de aplicação

Considere novamente o exemplo da subseção 3.3.1, página 23, no qual tínhamos $n = 10$ espécies de fibras e X_1 representava a força de resistência e X_2 o diâmetro. Com base na amostra de estudo tem-se que:

$$S = \begin{pmatrix} 1,25 & 0,80 \\ 0,80 & 0,87 \end{pmatrix}$$

Estamos interessados em testar a seguinte hipótese:

$$\begin{cases} H_0 : \Sigma = \Sigma_0 = \begin{pmatrix} 1,23 & 0,79 \\ 0,79 & 0,83 \end{pmatrix} \\ H_1 : \Sigma \neq \Sigma_0 \end{cases}$$

Sabemos que $S_n = \frac{(n-1)S}{n} = \begin{pmatrix} 1,125 & 0,720 \\ 0,720 & 0,783 \end{pmatrix}$, portanto:

$$\Sigma_0^{-1} S_n = \begin{pmatrix} 0,919 & -0,053 \\ -0,008 & 0,994 \end{pmatrix}, \text{ sendo } \begin{cases} |\Sigma_0^{-1} S_n| = 0,913 \\ \text{tr}(\Sigma_0^{-1} S_n) = 1,913 \end{cases}$$

Pela estatística (3.35) tem-se que:

$$\begin{aligned} \chi_1^2 &= n \text{tr}(\Sigma_0^{-1} S_n) - n \ln |\Sigma_0^{-1} S_n| - np \\ &= 10 \times 1,913 - 10 \ln(0,913) - 10 \times 2 = 0,040 \end{aligned}$$

Sob H_0 , essa estatística segue aproximadamente uma distribuição qui-quadrado com 3 graus de liberdade. Assumindo $\alpha = 0,05$, o limite crítico é dado por $\chi_{3;0,05}^2 = 7,8147$.

Como o valor obtido (0,040) pela estatística é inferior a $\chi_{3;0,05}^2 = 7,8147$ não existem evidências de que a matriz de covariâncias geradora dos dados amostrais seja diferente de Σ_0 . Usando a correção proposta por Korin (1968) obtemos o seguinte valor da estatística de teste:

$$\begin{aligned} \chi^2 &= \left[(n-1) - \frac{1}{6} \left(2p + 1 - \frac{2}{p+1} \right) \right] [\text{tr}(\Sigma_0^{-1} S) - \ln |\Sigma_0^{-1} S| - p] \\ &= \left[9 - \frac{1}{6} \left(4 + 1 - \frac{2}{3} \right) \right] [1,913 - \ln(0,913) - 2] = 0,033 \end{aligned}$$

o que nos leva também a não rejeição de H_0 já que $0,033 < 7,8147$.

3.6 Teste *Step-down* proposto por Sullivan et al. (2007)

Sullivan et al. (2007) propuseram um teste denominado “*step-down*” para detectar mudanças nos parâmetros da matriz de covariâncias e em outros parâmetros como o vetor de médias. Suponha que se tenha duas amostras A e B, sendo que A representa o processo antes da mudança de alguns parâmetros e B representa o processo depois da mudança como mostra a Figura 3.4.

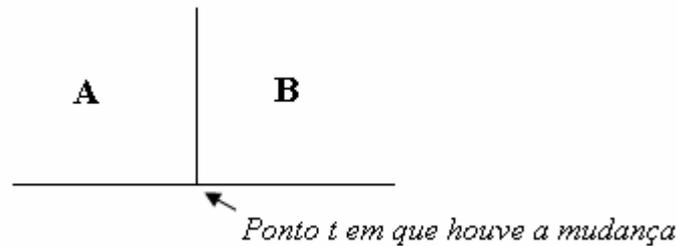


Figura 3.4: Proposta do teste de Sullivan et al. (2007)

O principal propósito desse método é aplicar o teste para comparação de parâmetros em dois grupos homogêneos consecutivos, o grupo antes e depois do ponto de mudança (A e B). O método “*step-down*” pode ser aplicado a qualquer modelo paramétrico onde a estimação de máxima verossimilhança pode ser utilizada para todos os parâmetros da distribuição de ambos os lados do ponto de mudança. Além disso, certas condições de regularidade têm que ser verificadas (ver Casella e Berger, 2002), de modo que a matriz de informação de Fisher possa ser calculada, já que esse teste depende dessa matriz.

Suponha que as distribuições antes e depois da mudança sejam completamente descritas pelo vetor de parâmetros θ , tendo q elementos. Para dados com distribuição normal multivariada, θ poderá consistir de elementos do vetor de médias e de elementos da matriz de covariâncias. Por exemplo, uma parametrização para distribuição normal bivariada pode ser dada como em (3.41) que é equivalente a (3.42).

$$\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \sigma_{12})^T \quad (3.41)$$

$$\theta = (\mu_1, \mu_2, \sigma_1, \rho, \sigma_2)^T \quad (3.42)$$

sendo σ_1^2 e σ_2^2 as variâncias de X_1 e X_2 e σ_1, σ_2 os respectivos desvios-padrão.

Podemos considerar o vetor θ_A proveniente do grupo A, com parâmetros estimados via máxima verossimilhança e θ_B para o grupo B. Os resultados podem ser mais bem entendidos se considerarmos a forma de parametrização (3.42), ou seja, utilizando desvio-padrão e correlação no lugar de variância e covariância. Por exemplo, suponha que ocorra uma mudança no desvio-padrão, mas não na correlação. Nessa situação todas as covariâncias envolvidas na mudança do desvio-padrão também serão influenciadas, porque cada covariância é o produto do coeficiente de correlação, que não muda, e dos desvios-padrão. Resumidamente, temos que, considerando $\rho \sigma_1 \sigma_2 = \sigma_{12}$, se σ_1 mudar e ρ e σ_2 não mudarem, σ_{12} terá mudado e se ρ não mudar e σ_1 mudar, por exemplo, a covariância tem que mudar automaticamente. É possível mudar a covariância e não mudar o desvio-padrão o que automaticamente acarreta na mudança no ρ . Como estamos trabalhando com uma distribuição normal multivariada existem duas formas de parametrizar a distribuição. Porém, utilizaremos desvios-padrão e coeficiente de correlação porque em alguns passos de nosso problema essa forma sintetiza e simplifica os cálculos. Até o momento discutimos o objetivo do teste. O próximo passo será apresentar algumas propriedades de inferência estatística do vetor $\hat{\theta}$. Seja $\hat{\theta}_n$ o estimador de máxima verossimilhança do parâmetro θ . Para n grande temos que:

$$\hat{\theta}_n \xrightarrow{d} N \left(\theta; \frac{1}{n} E_{\theta} \left(\left(\frac{\partial \ln f(x, \theta)}{\partial \theta} \right)^2 \right)^{-1} \right) \quad (3.43)$$

onde $E_{\theta} \left(\left(\frac{\partial \ln f(x, \theta)}{\partial \theta} \right)^2 \right)$ é denominado matriz Informação de Fisher (E_{θ} denota a esperança matemática). Casella e Berger (2002) mostram que se θ for um vetor $qx1$ e $\hat{\theta}_n$ o estimador de máxima verossimilhança de θ os seguintes resultados, para n grande, são verificados:

$$\hat{\theta}_n \xrightarrow{d} N_q \left(\theta_{qx1}; \Sigma_{\hat{\theta}_n} \right) \quad \text{onde} \quad \Sigma_{\hat{\theta}_n} = - \left(n E_{\theta} \left(\frac{\partial^2 \ln f(x, \theta)}{\partial \theta \partial \theta^T} \right) \right)^{-1} \quad (3.44)$$

Além disso, os autores mostram que $\hat{\theta}_n$ é assintoticamente eficiente para θ , além de ser consistente e assintoticamente não viesado. Assim, como é mostrado em Beyer (1978) o (i,k) -ésimo elemento da matriz de Informação de Fisher para uma distribuição normal multivariada ($X \sim N(\mu(\theta), \Sigma(\theta))$) é dado como em (3.45).

$$I_{i,k} = \frac{\partial \mu}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu^T}{\partial \theta_k} + \frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_k} \right) \quad (3.45)$$

onde $\frac{\partial \mu}{\partial \theta_i} = \begin{bmatrix} \frac{\partial \mu_1}{\partial \theta_i} & \frac{\partial \mu_2}{\partial \theta_i} & \dots & \frac{\partial \mu_q}{\partial \theta_i} \end{bmatrix}$ e $\frac{\partial \Sigma}{\partial \theta_i} = \begin{bmatrix} \frac{\partial \Sigma_{11}}{\partial \theta_i} & \frac{\partial \Sigma_{12}}{\partial \theta_i} & \dots & \frac{\partial \Sigma_{1q}}{\partial \theta_i} \\ \frac{\partial \Sigma_{21}}{\partial \theta_i} & \frac{\partial \Sigma_{22}}{\partial \theta_i} & \dots & \frac{\partial \Sigma_{2q}}{\partial \theta_i} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \Sigma_{N1}}{\partial \theta_i} & \frac{\partial \Sigma_{N2}}{\partial \theta_i} & \dots & \frac{\partial \Sigma_{qq}}{\partial \theta_i} \end{bmatrix}$.

Para uma amostra de tamanho n , Sullivan et al. (2007) mostram que sob H_0 , e considerando uma distribuição normal bivariada, a matriz de covariâncias de $\hat{\theta}_n$ obtida por (3.44) e (3.45) é dada como na Tabela 3.1.

Tabela 3.1: Matriz de covariâncias assintótica ($\Sigma_{\hat{\theta}_n}$) de $\hat{\theta}_n$ para uma distribuição normal bivariada

	μ_1	μ_2	σ_1	ρ	σ_2
μ_1	$\frac{\sigma_1^2}{n}$	$\frac{\rho^2 \sigma_1 \sigma_2}{n}$	0	0	0
μ_2	$\frac{\rho^2 \sigma_1 \sigma_2}{n}$	$\frac{\sigma_2^2}{n}$	0	0	0
σ_1	0	0	$\frac{\sigma_1^2}{2n}$	$-\frac{\rho(\rho^2-1)\sigma_1}{2n}$	$\frac{\rho^2 \sigma_1 \sigma_2}{2n}$
ρ	0	0	$-\frac{\rho(\rho^2-1)\sigma_1}{2n}$	$\frac{(\rho^2-1)^2}{n}$	$-\frac{\rho(\rho^2-1)\sigma_2}{2n}$
σ_2	0	0	$\frac{\rho^2 \sigma_1 \sigma_2}{2n}$	$-\frac{\rho(\rho^2-1)\sigma_2}{2n}$	$\frac{\sigma_2^2}{2n}$

Uma vez identificadas as propriedades dos estimadores de máxima verossimilhança vamos construir o teste de hipótese para o vetor θ . Considere a diferença no vetor de parâmetros como $\delta = (\theta_A - \theta_B)$ o qual é estimado por $\hat{\delta} = (\hat{\theta}_A - \hat{\theta}_B)$. Dessa forma, o objetivo é avaliar se δ é igual a zero. Como o foco dessa dissertação é averiguar mudanças na matriz de covariâncias vamos manter nossa atenção somente à parte que contém variação, fixando o vetor μ . Dessa forma, estamos interessados em testar a

seguinte hipótese: $H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$, onde $\theta = (\sigma_1, \rho, \sigma_2)^T$ e $\theta_0 = (\sigma_{1_0}, \rho_0, \sigma_{2_0})^T$. De acordo com o artigo de Sullivan et al. (2007), dado uma amostra de tamanho n , a estatística de teste é dada por:

$$T^2 = \delta^T \Sigma_{\hat{\delta}}^{-1} \delta \quad \text{onde} \quad \delta = (\hat{\theta} - \theta_0). \quad (3.46)$$

Sob H_0 , a estatística (3.46), tem distribuição assintótica qui-quadrado com q graus de liberdade, onde q é o número de parâmetros sob H_0 . Como estamos trabalhando somente com a parte que contém a variação de X , a matriz de covariâncias $\Sigma_{\hat{\delta}}$ passa a ter a seguinte forma:

$$\Sigma_{\hat{\delta}} = \begin{pmatrix} \frac{\sigma_1^2}{2n} & -\frac{\rho(\rho^2-1)\sigma_1}{2n} & \frac{\rho^2\sigma_1\sigma_2}{2n} \\ -\frac{\rho(\rho^2-1)\sigma_1}{2n} & \frac{(\rho^2-1)^2}{n} & -\frac{\rho(\rho^2-1)\sigma_2}{2n} \\ \frac{\rho^2\sigma_1\sigma_2}{2n} & -\frac{\rho(\rho^2-1)\sigma_2}{2n} & \frac{\sigma_2^2}{2n} \end{pmatrix} \quad (3.47)$$

Para o caso em que o vetor de médias é conhecido os elementos do vetor $\hat{\theta}$ serão aqueles provenientes da matriz S_{μ} , definida em (3.37). Vale lembrar que estamos interessados somente na parte que contém variação. Portanto, na situação em que o vetor de médias é conhecido o vetor $\hat{\theta}$ será dado por: $\hat{\theta} = (s_{\mu_1}, \hat{\rho}, s_{\mu_2})$, onde s_{μ_1} e s_{μ_2} são os desvios-padrão amostrais de X_1 e X_2 obtidos através da matriz S_{μ} (ver página 34). Nesse caso, sob H_0 , a estatística apresentada (3.46) tem distribuição assintótica qui-quadrado com 3 graus de liberdade, visto que na hipótese nula tem-se 3 (três) parâmetros de comparação σ_1, ρ e σ_2 .

Em Pinto e Mingoti (2009a) e Pinto (2009b) é mostrado que o teste proposto em Sullivan et al. (2007) é mais poderoso que os testes da variância generalizada tradicional e com modificação de Djauhari (2005) para dados independentes. Esse resultado é devido à própria construção do teste, já que ele leva em consideração todas as informações da matriz de covariâncias, não reduzindo-a em um único valor, como no teste do determinante. No contexto da distribuição normal multivariada o artigo de Sullivan et al. (2007) considera apenas o caso bivariado.

No artigo de Sullivan et al. (2007) é proposto um método para se identificar qual, ou quais, parâmetros são responsáveis pela rejeição de H_0 quando isto ocorre. Os autores denominam esse método como "step-down". Para mais detalhes sobre o método recomenda-se a leitura do artigo, já que esse tópico não é objetivo dessa dissertação. É interessante notar que o teste de Sullivan et al. (2007) pode ser aplicado a situações nas quais os dados não são provenientes de uma distribuição normal multivariada, desde que a matriz de Informação de Fisher esteja de acordo com a distribuição correspondente dos dados amostrais.

3.6.1 Exemplo de aplicação

Considere novamente o exemplo da subseção 3.3.1, página 23, na qual se tem $n=10$ espécies de fibras e X_1 é a força de resistência e X_2 o diâmetro. Estamos interessados em testar a seguinte hipótese:

$$\begin{cases} H_0 : \Sigma = \Sigma_0 = \begin{pmatrix} 1,23 & 0,79 \\ 0,79 & 0,83 \end{pmatrix} \\ H_1 : \Sigma \neq \Sigma_0 \end{cases}$$

A matriz de covariâncias amostral é dada por:

$$S = \begin{pmatrix} 1,25 & 0,80 \\ 0,80 & 0,87 \end{pmatrix}$$

Vale ressaltar, que no teste proposto por Sullivan et al. (2007), devemos utilizar a matriz de covariâncias amostral estimada via máxima verossimilhança. Dessa forma:

$$\frac{(n-1)S}{n} = \begin{pmatrix} 1,125 & 0,720 \\ 0,720 & 0,783 \end{pmatrix}, \text{ sendo } \begin{cases} \hat{\theta} = (\hat{\sigma}_1, \hat{\rho}, \hat{\sigma}_2)^T = (1,061 \quad 0,767 \quad 0,885)^T \\ \theta_0 = (\sigma_{1_0}, \rho_0, \sigma_{2_0})^T = (1,109 \quad 0,782 \quad 0,911)^T \\ \hat{\delta} = (\hat{\theta} - \theta_0)^T = (-0,048 \quad -0,015 \quad -0,026)^T \end{cases}$$

Utilizando a informação dada em (3.47) temos:

$$\Sigma_{\hat{\delta}} = \begin{pmatrix} 0,062 & 0,017 & 0,031 \\ & 0,015 & 0,014 \\ & & 0,041 \end{pmatrix}, \text{ e conseqüentemente } \Sigma_{\hat{\delta}}^{-1} = \begin{pmatrix} 29,047 & & \\ -18,138 & 106,661 & \\ -15,566 & -22,080 & 43,046 \end{pmatrix}$$

Portanto;

$$T^2 = \delta^T \Sigma_{\delta}^{-1} \delta =$$

$$= (-0,048 \quad -0,015 \quad -0,026) \begin{pmatrix} 29,05 & & \\ -18,14 & 106,661 & \\ -15,57 & -22,08 & 43,05 \end{pmatrix} \begin{pmatrix} -0,048 \\ -0,015 \\ -0,026 \end{pmatrix} = 0,0383$$

Sob H_0 , T^2 tem distribuição qui-quadrado com 3 graus de liberdade. Dessa forma, assumindo $\alpha = 0,05$, o limite crítico é dado por $\chi_{3;0,05}^2 = 7,8147$. Como o valor observado da estatística T^2 é inferior a 7,8141 não existem evidências de diferença significativa das matrizes de covariâncias. Com isso, não temos evidências para rejeitar a hipótese nula.

3.7 Teste VMAX proposto por Machado et al. (2008)

Nesta seção apresenta-se o teste VMAX para o monitoramento da matriz de covariâncias $\Sigma_{2 \times 2}$ de um processo bivariado. Como o foco dessa dissertação é averiguar mudanças na matriz de covariâncias, vamos manter nossa atenção somente à parte que contém variação, fixando o vetor μ . O teste proposto por Machado et al. (2008), denominado de VMAX, baseia-se no maior valor das variâncias amostrais, ou seja, em um processo bivariado, os pontos plotados no gráfico de controle VMAX correspondem ao valor da maior variância padronizada. Segundo os autores, a principal vantagem do teste VMAX é que ele possui uma melhor capacidade de diagnóstico, ou seja, com ele é mais fácil identificar a variável que teve sua variabilidade alterada pela ocorrência de uma causa especial. Machado et al. (2008) mostram que o teste VMAX tem um desempenho superior ao teste do determinante, que é baseado na variância amostral generalizada $|S|$.

Considere um processo bivariado no qual as duas características de qualidade são representadas pelas variáveis X_1 e X_2 que seguem uma distribuição normal bivariada com vetor de médias conhecido $\mu = (\mu_1, \mu_2)^T$ e matriz de covariâncias $\Sigma_{2 \times 2}$. Sem perda de generalidade, suponha que estamos interessados em testar a hipótese $H_0: \Sigma = \Sigma_0$ versus $H_1: \Sigma \neq \Sigma_0$, onde sob H_0 temos o processo sob controle. A estatística de monitoramento corresponde ao maior valor das variâncias amostrais padronizadas, $VMAX = \max\{s_{X_1}^2; s_{X_2}^2\}$, onde $s_{X_1}^2$ e $s_{X_2}^2$ são dadas como em (3.48).

$$s_{X_1}^2 = \frac{\sum_{i=1}^n x_{1i}^2}{n} \quad s_{X_2}^2 = \frac{\sum_{i=1}^n x_{2i}^2}{n} \quad (3.48)$$

onde $x_{1i} = \frac{X_{1i} - \mu_{X_1}}{\sigma_1}$ e $x_{2i} = \frac{X_{2i} - \mu_{X_2}}{\sigma_2}$, σ_1 e σ_2 postulados sob H_0 , $(X_{1i}, X_{2i})^T$ o vetor de observações da unidade amostral i , $i=1, 2, \dots, n$.

Quando o teste VMAX está em uso, amostras de tamanho n são selecionadas do processo em intervalos de tempo regulares. Duas características de qualidade $(X_{1i}; X_{2i})$ das n unidades são medidas e a estatística $VMAX = \max\{S_{X_1}^2; S_{X_2}^2\}$ é calculada. Se a estatística VMAX for maior que o limite de controle (LC), calculado através da equação (3.49), existem evidências de que o processo está fora de controle.

$$\alpha = 1 - \int_0^{nLC} P \left[\chi_{n, \left(\frac{t\rho^2}{1-\rho^2} \right)}^2 < \frac{nLC}{(1-\alpha^2)} \right] \frac{1}{2^{n/2} \Gamma(n/2)} e^{-t/2} t^{(n/2)-1} dt \quad (3.49)$$

onde α é o nível de significância do teste, LC o valor do limite de controle e $\chi_{n, (t\rho^2 / 1-\rho^2)}^2$ representa uma distribuição qui-quadrado não-central com n graus de liberdade e com o parâmetro de não-centralidade dado por $(t\rho^2 / 1-\rho^2)$.

No artigo de Machado et al. (2008) apenas o caso em que o vetor μ é postulado (conhecido) foi considerado. No entanto, é fácil perceber que o teste VMAX pode ser estendido para as situações em que μ é desconhecido, bastando que as variâncias amostrais sejam calculadas em relação às estimativas de μ_i , $i=1,2$, e que a distribuição da estatística de teste, sob H_0 , seja ajustada para acomodar essa nova situação. Considerando como estimativas de μ_i a média amostral \bar{x}_i , temos:

$$s_{X_1}^2 = \frac{\sum_{i=1}^n x_{1i}^2}{n-1} \quad s_{X_2}^2 = \frac{\sum_{i=1}^n x_{2i}^2}{n-1} \quad (3.50)$$

onde $x_{1i} = \frac{X_{1i} - \bar{x}_1}{\sigma_1}$ e $x_{2i} = \frac{X_{2i} - \bar{x}_2}{\sigma_2}$, σ_1 e σ_2 postulados sob H_0 . Lembrando que o teste é baseado na maior variância das duas variáveis ($VMAX = \max\{s_{X_1}^2; s_{X_2}^2\}$), onde $s_{X_1}^2$ e $s_{X_2}^2$ são dadas como em (3.50).

No artigo de Machado et al. (2008) mostra-se que o teste VMAX é mais eficiente do que o teste do determinante ($|S|$) para alguns modelos de variabilidade, no entanto, apenas o caso em que $p=2$ foi considerado pelos autores. Como em muitas situações práticas tem-se mais de 2 variáveis ($p>2$) de interesse é interessante estender o teste VMAX para $p>2$, o que foi feito no trabalho de Pereira e Mingoti (2009) tanto para o caso em que o vetor de médias é conhecido quanto desconhecido. Neste trabalho de Pereira e Mingoti (2009) a distribuição exata da estatística VMAX é obtida via simulação para dados multivariados independentes. Do ponto de vista prático a simulação é importante em vista da dificuldade de se utilizar a integração numérica para o caso em que temos mais de duas variáveis ($p>2$).

3.7.1 Exemplo de aplicação

Considere o exemplo da subseção 3.3.1, página 23. Quando o processo está sob controle a matriz de covariância, Σ_0 , do processo é conhecida e o vetor de médias é dado por $\mu_0=(115,59; 1,06)$. Assim, a seguinte hipótese será testada:

$$\begin{cases} H_0 : \Sigma = \Sigma_0 = \begin{pmatrix} 1,23 & 0,79 \\ 0,79 & 0,83 \end{pmatrix} \\ H_1 : \Sigma \neq \Sigma_0 \end{cases}$$

Sob H_0 , temos que a variabilidade do processo está sob controle. Através de métodos de simulação, e considerando $\alpha = 0,05$, obtemos o limite de controle ($LC = 2,002$), para a situação em estudo. Para verificar se o processo está sob controle, selecionou-se uma amostra de tamanho 10 obtendo-se as seguintes variâncias padronizadas:

$$s_{X_1}^2 = 1,1418 \text{ e } s_{X_2}^2 = 0,9248$$

Vale lembrar que $s_{X_1}^2$ e $s_{X_2}^2$ foram calculadas como em (3.48), onde μ é o vetor postulado em H_0 . Portanto, $VMAX = \max\{s_{X_1}^2; s_{X_2}^2\} = 1,1418$. Como o valor de $VMAX < LC$ não temos evidências para rejeitar a hipótese nula e, portanto podemos concluir que o processo está sob controle. É importante observar que o teste $VMAX$ é função apenas das variâncias amostrais e da comparação destas com as respectivas variâncias teóricas, ou seja, é um teste para testar a variabilidade das variáveis conjuntamente, mas não exatamente para a matriz de covariâncias, já que não depende da covariância entre as variáveis.

Capítulo 4

Processos Multivariados Autocorrelacionados

Os procedimentos clássicos de controle de processos multivariados são desenvolvidos assumindo que os vetores de observações são independentes com distribuição normal multivariada com vetor de médias μ e matriz de covariâncias Σ . Porém, na prática, muitas das operações apresentam autocorrelação. Deste modo, para aumentar a qualidade do monitoramento do processo é necessário levar em consideração a autocorrelação dos vetores de observações. Antes de introduzirmos o conceito de processo multivariado autocorrelacionado, vamos apresentar a forma de um processo em que não há autocorrelação. Seja $Y_t = [Y_{t1}, Y_{t2}, \dots, Y_{tp}]^T$ o vetor aleatório de dimensão $px1$ no tempo t , $\mu = [\mu_1, \mu_2, \dots, \mu_p]^T$ é o vetor de médias de dimensão $px1$, e $\varepsilon_t = [\varepsilon_{t1}, \varepsilon_{t2}, \dots, \varepsilon_{tp}]^T$ é um vetor aleatório de dimensão $px1$ para o tempo t proveniente de uma distribuição normal multivariada com vetor de médias nulo e matriz de covariâncias igual a Σ_ε , e que pode ser representado por:

$$Y_t = \mu + \varepsilon_t, \quad t = 1, 2, \dots \quad (4.1)$$

Neste caso, tem-se um processo de observações independentes.

Um processo multivariado autocorrelacionado pode ser representado por diversos modelos de séries temporais estacionários. A classe mais conhecida é a de modelos estacionários autorregressivos e de média móvel VARMA(k,q), cujo modelo é definido genericamente como:

$$Y_t = \mu + \varepsilon_t + \sum_{j=1}^k \Phi_j (Y_{t-j} - \mu) - \sum_{j=1}^q \Theta_j \varepsilon_{t-j}, \quad t = 1, 2, \dots \quad (4.2)$$

sendo Φ_j e Θ_j as matrizes de dimensão pxp que contém os parâmetros autorregressivos e de médias móveis do modelo de séries temporais multivariado e $\mu_t = [\mu_{t1}, \mu_{t2}, \dots, \mu_{tp}]^T$ o

vetor de médias de Y_t no tempo t . Considerando $\mu=0$, sem perda de generalidade, podemos reescrever o modelo VARMA(k,q) da seguinte forma:

$$\Phi(B)Y_t = \Theta(B)\varepsilon_t \quad (4.3)$$

onde $\Phi(B) = (I - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_k B^k)$, $\Theta(B) = (I - \Theta_1 B - \Theta_2 B^2 - \dots - \Theta_q B^q)$ são matrizes $p \times p$ polinomiais. Vale lembrar que B é um operador de translação para o passado. Para mais detalhes dos modelos VARMA(k,q) ver Lutkepohl (1993) e Tsay (2002).

Observe pela equação (4.2) que quando Θ_j é uma matriz nula caímos na classe dos modelos autorregressivos multivariados de ordem k , denominados VAR(k). Quando a matriz Φ_j é nula têm-se os modelos multivariados de médias móveis de ordem q , representados por VMA(q).

É essencial usar um modelo apropriado de séries temporais para modelar os vetores de observações do processo, tendo em conta a estrutura de dependência. Para isso, um dos primeiros passos na análise das séries de dados é avaliar a existência de tendência, sazonalidade, valores discrepantes (*outliers*), descontinuidade, através de um gráfico de séries temporais. Nas próximas seções exploraremos com mais detalhes os dois principais modelos utilizados, VAR(k) e VMA(q).

4.1 Modelo autorregressivo multivariado de ordem k , VAR(k)

Um dos modelos mais importantes, e que será mais utilizado nessa dissertação, é conhecido como modelo autorregressivo multivariado de ordem 1, VAR(1). Para um processo autocorrelacionado o modelo VAR(1) é dado por:

$$Y_t = \mu_t + \Phi(Y_{t-1} - \mu_t) + \varepsilon_t \quad (4.4)$$

onde μ_t é o vetor de médias no tempo t , ε_t é um vetor de variáveis aleatórias independentes com vetor de médias zero e matriz de covariâncias Σ_ε positiva definida. Além disso, Φ é uma matriz $p \times p$ com os parâmetros de autocorrelação. Se assumirmos que Y_t é estacionário, μ_t é constante no tempo, portanto podemos reescrever modelo anterior da seguinte forma:

$$Y_t = \mu + \Phi(Y_{t-1} - \mu) + \varepsilon_t \quad (4.5)$$

Ainda considerando o modelo VAR(1) temos que se a matriz Φ é nula o modelo anterior se reduz a $Y_t = \mu + \varepsilon_t$, caso contrário, Φ afeta a matriz de covariâncias de Y_t . Sem perda de generalidade vamos supor que $\mu=0$ e que estamos em um processo bivariado. Dessa forma temos o seguinte modelo VAR(1):

$$Y_t = \Phi Y_{t-1} + \varepsilon_t \quad (4.6)$$

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \quad (4.7)$$

$$Y_{1t} = \phi_{11} Y_{1,t-1} + \phi_{12} Y_{2,t-1} + \varepsilon_{1t} \quad (4.8)$$

$$Y_{2t} = \phi_{21} Y_{1,t-1} + \phi_{22} Y_{2,t-1} + \varepsilon_{2t} \quad (4.9)$$

onde ϕ_{ij} é o (i,j) -ésimo elemento de Φ . Baseado na equação (4.8), temos que ϕ_{12} representa a dependência linear de Y_{1t} sobre $Y_{2,t-1}$ na presença de $Y_{1,t-1}$. Portanto, ϕ_{12} é o efeito condicional de $Y_{2,t-1}$ sobre Y_{1t} dado $Y_{1,t-1}$. Se $\phi_{12}=0$, temos que Y_{1t} não depende de $Y_{2,t-1}$, e o modelo mostra que Y_{1t} depende apenas do seu próprio passado. Da mesma forma, se $\phi_{21}=0$, a equação (4.9) mostra que Y_{2t} não depende de $Y_{1,t-1}$ quando $Y_{1,t-1}$ é dado. Para mais detalhes ver Tsay (2002).

Considere as duas equações (4.8) e (4.9) conjuntamente. Se $\phi_{12}=0$ e $\phi_{21} \neq 0$, então existe uma relação unidirecional de Y_{1t} para Y_{2t} . Se $\phi_{12} = \phi_{21} = 0$, então Y_{1t} e Y_{2t} são não relacionadas. Por fim, se $\phi_{12} \neq 0$ e $\phi_{21} \neq 0$ então existe um relacionamento entre as duas séries de estudo.

Em geral, a matriz de coeficientes Φ mede a dinâmica de dependência de Y_t . A relação existente entre Y_{1t} e Y_{2t} é indicada pelos elementos da diagonal secundária da matriz de covariâncias Σ de ε_t . Se $\sigma_{12} = 0$, então não existe uma relação linear entre os dois componentes das séries.

A generalização do modelo VAR(1) para o VAR(k) nos fornece o seguinte modelo:

$$(Y_t - \mu) = \Phi_1 (Y_{t-1} - \mu) + \dots + \Phi_k (Y_{t-k} - \mu) + \varepsilon_t, \quad k \geq 1 \quad (4.10)$$

onde Φ_j são matrizes de coeficientes $p \times p$.

Utilizando o operador de translação para o passado, denotado por B, podemos reescrever o modelo (4.10) como em (4.11).

$$(I - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_k B^k)(Y_t - \mu) = \varepsilon_t$$

$$\Phi(B)(Y_t - \mu) = \varepsilon_t \quad (4.11)$$

onde I é uma matriz identidade $p \times p$ e $\Phi(B) = (I - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_k B^k)$ é uma matriz polinomial. Quando $\mu = 0$ a equação (4.10) torna-se:

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \dots + \Phi_k Y_{t-k} + \varepsilon_t, \quad k \geq 1 \quad (4.12)$$

4.1.1 Estimadores de máxima verossimilhança de um modelo VAR(k)

Antes de iniciarmos os cálculos para determinação dos estimadores de máxima verossimilhança de um modelo VAR(k), vamos apresentar a notação que será utilizada para representação do modelo VAR(k). Essa notação está de acordo com aquela empregada em Lutkepohl (1993). Considere a seguinte notação:

(a) $Y^0 = (y_1 - \mu, \dots, y_T - \mu)$ uma matriz com dimensão $(p \times T)$

(b) $\Phi = (\Phi_1, \dots, \Phi_k)$ uma matriz com dimensão $(p \times pk)$

(c) $Y_t^0 = \begin{bmatrix} y_t - \mu \\ \vdots \\ y_{t-k+1} - \mu \end{bmatrix}$ um vetor com dimensão $(pk \times 1)$

- (d) $E = (\varepsilon_1, \dots, \varepsilon_T)$ uma matriz com dimensão $(p \times T)$
- (e) $X = (Y_0^0, \dots, Y_{T-1}^0)$ uma matriz com dimensão $(pk \times T)$
- (f) $y^0 = \text{vec}(Y^0)$ um vetor com dimensão $(pT \times 1)$
- (g) $\alpha = \text{vec}(\Phi)$ um vetor com dimensão $(p^2k \times 1)$
- (h) $u = \text{vec}(E)$ um vetor com dimensão $(pT \times 1)$

onde $y_1, y_2 \dots y_T$ são observações da série e $\text{vec}(\cdot)$ é um operador de vetorização de matrizes. De acordo com a notação assim definida pode-se reescrever o modelo VAR(k) apresentado em (4.10), da seguinte forma:

$$Y^0 = \Phi X + E \quad \text{ou} \quad y^0 = (X' \otimes I_{p \times p}) \alpha + u \quad (4.13)$$

onde $I_{p \times p}$ é uma matriz identidade de ordem p , e \otimes o produto de Kronecker.

Para obtermos os estimadores de máxima verossimilhança vamos assumir que o processo VAR(k) é gaussiano. Mais precisamente temos:

$$u = \text{vec}(E) = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_T \end{bmatrix} \sim N(0; I_{T \times T} \otimes \Sigma_\varepsilon) \quad (4.14)$$

Em outras palavras, a função densidade de probabilidade de u é dada por:

$$f_u(u) = \frac{1}{(2\pi)^{pT/2}} |I_{T \times T} \otimes \Sigma_\varepsilon|^{-1/2} \exp \left[-\frac{1}{2} u' (I_{T \times T} \otimes \Sigma_\varepsilon^{-1}) u \right] \quad (4.15)$$

Além disso, temos:

$$u = \begin{bmatrix} I_{p \times p} & 0 & \dots & 0 & \dots & \dots & 0 \\ -\Phi_1 & I_{p \times p} & \dots & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & & & \vdots \\ -\Phi_k & -\Phi_{k-1} & \dots & I_{p \times p} & & & 0 \\ 0 & -\Phi_k & & \ddots & & & \vdots \\ \vdots & & \ddots & & \ddots & & \vdots \\ 0 & 0 & \dots & -\Phi_k & \dots & \dots & I_{p \times p} \end{bmatrix} (y - \mu^*) + \begin{bmatrix} -\Phi_1 & -\Phi_2 & \dots & -\Phi_k \\ -\Phi_2 & -\Phi_3 & \dots & 0 \\ \vdots & & & \vdots \\ -\Phi_k & 0 & \dots & 0 \\ & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} (Y_0 - \mu) \quad (4.16)$$

onde $y = \text{vec}(Y)$ e $\boldsymbol{\mu}^* = (\mu', \dots, \mu')$ são vetores de dimensão $Tpx1$ e $\boldsymbol{\mu} = (\mu', \dots, \mu')$ e $Y^0 = (y_0' \dots y_{-k+1}')$ são vetores de dimensão $pkx1$. Consequentemente, $\partial u / \partial y'$ é uma matriz triangular inferior com diagonal e determinante igual a 1 (um). Daí usando que $u = y - \boldsymbol{\mu}^* - (X' \otimes I_{p \times p}) \boldsymbol{\alpha}$ temos:

$$\begin{aligned} f_y(y) &= \left| \frac{\partial u}{\partial y'} \right| f_u(u) \\ &= \frac{1}{(2\pi)^{pT/2}} |I_{T \times T} \otimes \Sigma_\epsilon|^{-1/2} \times \\ &\quad \times \exp \left[-\frac{1}{2} (y - \boldsymbol{\mu}^* - (X' \otimes I_{p \times p}) \boldsymbol{\alpha})' (I_{T \times T} \otimes \Sigma_\epsilon^{-1}) (y - \boldsymbol{\mu}^* - (X' \otimes I_{p \times p}) \boldsymbol{\alpha}) \right] \end{aligned} \quad (4.17)$$

onde X e $\boldsymbol{\alpha}$ são expressos como na notação definida anteriormente. Dessa forma, a função de log-verossimilhança é dada por:

$$\begin{aligned} \ln l(\boldsymbol{\mu}, \boldsymbol{\alpha}, \Sigma_\epsilon) &= -\frac{pT}{2} \ln 2\pi - \frac{T}{2} \ln |\Sigma_\epsilon| - \frac{1}{2} [y - \boldsymbol{\mu}^* - (X' \otimes I_{p \times p}) \boldsymbol{\alpha}]' (I_{T \times T} \otimes \Sigma_\epsilon^{-1}) [y - \boldsymbol{\mu}^* - (X' \otimes I_{p \times p}) \boldsymbol{\alpha}] \\ &= -\frac{pT}{2} \ln 2\pi - \frac{T}{2} \ln |\Sigma_\epsilon| - \frac{1}{2} \sum_{t=1}^T \left[(y_t - \boldsymbol{\mu}) - \sum_{i=1}^k \Phi_i (y_{t-i} - \boldsymbol{\mu}) \right]' \Sigma_\epsilon^{-1} \left[(y_t - \boldsymbol{\mu}) - \sum_{i=1}^k \Phi_i (y_{t-i} - \boldsymbol{\mu}) \right] \\ &= -\frac{pT}{2} \ln 2\pi - \frac{T}{2} \ln |\Sigma_\epsilon| - \frac{1}{2} \sum_t \left(y_t - \sum_i \Phi_i y_{t-i} \right)' \Sigma_\epsilon^{-1} \left(y_t - \sum_i \Phi_i y_{t-i} \right) + \\ &\quad + \boldsymbol{\mu}' \left(I_{p \times p} - \sum_i \Phi_i \right)' \Sigma_\epsilon^{-1} \sum_t \left(y_t - \sum_i \Phi_i y_{t-i} \right) - \frac{T}{2} \boldsymbol{\mu}' \left(I_{p \times p} - \sum_i \Phi_i \right)' \Sigma_\epsilon^{-1} \left(I_{p \times p} - \sum_i \Phi_i \right) \boldsymbol{\mu} \\ &= -\frac{pT}{2} \ln 2\pi - \frac{T}{2} \ln |\Sigma_\epsilon| - \frac{1}{2} \text{tr} [(Y^0 - \Phi X)' \Sigma_\epsilon^{-1} (Y^0 - \Phi X)] \end{aligned} \quad (4.18)$$

onde Y^0 e Φ são expressos como na notação definida anteriormente na página 48.

Uma vez identificada a função de log-verossimilhança o próximo passo será determinar os estimadores de máxima verossimilhança de $\boldsymbol{\mu}$, $\boldsymbol{\alpha}$ e Σ_ϵ . Dessa forma, temos que as derivadas da função de log-verossimilhança em função dos parâmetros de interesse serão dadas por:

(a) Em função de μ

$$\begin{aligned} \frac{\partial \ln l}{\partial \mu} &= \left(I_{p \times p} - \sum_i \Phi_i \right)' \Sigma_\varepsilon^{-1} \sum_t \left(y_t - \sum_i \Phi_i y_{t-i} \right) - T \left(I_{p \times p} - \sum_i \Phi_i \right)' \Sigma_\varepsilon^{-1} \left(I_{p \times p} - \sum_i \Phi_i \right) \mu \\ &= [I_{p \times p} - \Phi(j \otimes I_{p \times p})]' \Sigma_\varepsilon^{-1} \left[\sum_t (y_t - \mu - \Phi Y_{t-1}^0) \right] \end{aligned} \quad (4.19)$$

onde Y_t^0 é expresso como na notação definida anteriormente e $j = (1, \dots, 1)'$ é um vetor de 1 com dimensão $(k \times 1)$.

(b) Em função de α

$$\begin{aligned} \frac{\partial \ln l}{\partial \alpha} &= (X \otimes I_{p \times p}) (I_{T \times T} \otimes \Sigma_\varepsilon^{-1}) [y - \mu^* - (X' \otimes I_{p \times p}) \alpha] \\ &= (X \otimes \Sigma_\varepsilon^{-1}) (y - \mu^*) - (XX' \otimes \Sigma_\varepsilon^{-1}) \alpha \end{aligned} \quad (4.20)$$

(c) Em função de Σ_ε

$$\frac{\partial \ln l}{\partial \Sigma_\varepsilon} = -\frac{T}{2} \Sigma_\varepsilon^{-1} + \frac{1}{2} \Sigma_\varepsilon^{-1} (Y^0 - \Phi X) (Y^0 - \Phi X)' \Sigma_\varepsilon^{-1} \quad (4.21)$$

Igualando as derivadas a zero obtemos os seguintes estimadores:

$$\tilde{\mu} = \frac{1}{T} \left(I_{p \times p} - \sum_i \tilde{\Phi}_i \right)^{-1} \sum_t \left(y_t - \sum_i \tilde{\Phi}_i y_{t-i} \right) \quad (4.22)$$

$$\tilde{\alpha} = ((\tilde{X}\tilde{X}')^{-1} \tilde{X} \otimes I_{p \times p}) (y - \tilde{\mu}^*) \quad (4.23)$$

$$\tilde{\Sigma}_\varepsilon = \frac{1}{T} (\tilde{Y}^0 - \tilde{\Phi}\tilde{X}) (\tilde{Y}^0 - \tilde{\Phi}\tilde{X})' \quad (4.24)$$

onde \tilde{Y}^0 e \tilde{X} são obtidos de Y^0 e X , respectivamente, substituindo μ por $\tilde{\mu}$. Um fato a ser considerado é que os estimadores de μ e α são idênticos àqueles obtidos pelo método de mínimos quadrados, além disso, os estimadores $\tilde{\alpha}$ e $\tilde{\mu}$ são consistentes se Y_t for um processo gaussiano estacionário. Para mais detalhes das propriedades dos estimadores de máxima verossimilhança e mínimos quadrados ver Lutkepohl (1993).

4.2 Modelo de média-móvel multivariado de ordem q , $VMA(q)$

O modelo de média-móvel multivariado de ordem q , ou $VMA(q)$, é dado da seguinte forma:

$$Y_t - \mu = \varepsilon_t - \Theta_1 \varepsilon_{t-1} - \Theta_2 \varepsilon_{t-2} - \dots - \Theta_q \varepsilon_{t-q}$$

$$Y_t - \mu = \Theta(B) \varepsilon_t \quad (4.25)$$

onde Θ_i são matrizes $p \times p$, $i=1,2,\dots,q$, e $\Theta(B) = (I - \Theta_1 B - \Theta_2 B^2 - \dots - \Theta_q B^q)$ é uma matriz polinomial levando em conta o operador de translação para o passado, B . Como ilustração, para uma melhor compreensão do processo VMA , vamos considerar o caso bivariado do processo $VMA(1)$, como é mostrado em (4.26). Sem perda de generalidade vamos considerar $\mu=0$.

$$Y_t = \varepsilon_t - \Theta \varepsilon_{t-1} \quad (4.26)$$

Podemos reescrever o modelo (4.26) da seguinte forma:

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} - \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{pmatrix} \quad (4.27)$$

$$Y_{1t} = \varepsilon_{1t} - \theta_{11} \varepsilon_{1,t-1} - \theta_{12} \varepsilon_{2,t-1} \quad (4.28)$$

$$Y_{2t} = \varepsilon_{2t} - \theta_{21} \varepsilon_{1,t-1} - \theta_{22} \varepsilon_{2,t-1} \quad (4.29)$$

Considere a equação (4.28) para Y_{1t} . O parâmetro θ_{12} denota a relação de dependência linear de Y_{1t} sobre $\varepsilon_{2,t-1}$ em presença de $\varepsilon_{1,t-1}$. Se $\theta_{12} = 0$, então Y_{1t} não depende de valores defasados de ε_{2t} nem de Y_{2t} . Da mesma forma, se $\theta_{21} = 0$ então Y_{2t} não depende de valores passados de Y_{1t} . Os elementos da diagonal secundária da matriz Θ representam a dinâmica de dependência entre os componentes da série. Para um modelo $VMA(1)$, podemos classificar a relação entre Y_{1t} e Y_{2t} da seguinte forma:

1. Se $\theta_{12} = \theta_{21} = 0$, então Y_{1t} e Y_{2t} não são relacionadas;

2. Existe uma relação unidirecional de Y_{1t} para Y_{2t} se $\theta_{12} = 0$, mas $\theta_{21} \neq 0$. A relação unidirecional oposta ocorre quando $\theta_{21} = 0$, mas $\theta_{12} \neq 0$.
3. Existe uma relação de dependência entre Y_{1t} e Y_{2t} se $\theta_{12} \neq 0$ e $\theta_{21} \neq 0$.

Em Lutkepohl (1993), os estimadores de máxima verossimilhança de um modelo VMA(q) são mostrados detalhadamente como foi feito na seção 4.1.1 para um modelo VAR(k). Além disso, vale ressaltar que existem os processos VARMA(k,q) que são a generalização dos processos vistos, mas que não são objetivos de estudo dessa dissertação. Para mais detalhes dos processos VARMA(k,q) ver Brockweel e Davis (1996); Lutkepohl (1993) e Tsay (2002). Uma vez que apresentamos os modelos de séries temporais multivariados, o próximo passo será identificar as matrizes de covariâncias e correlação cruzadas.

4.3 Matrizes de covariâncias (Γ) e correlação (ρ) cruzadas

Considere $\Gamma(t, t+h)$ como a matriz de covariâncias cruzada entre Y_t e Y_{t+h} e o (i,j) -ésimo elemento dessa matriz como $\gamma_{ij}(h)$ onde:

$$\gamma_{ij}(h) = E((Y_{it} - \mu_{it})(Y_{j,t+h} - \mu_{j,t+h})) \quad (4.30)$$

Sob o pressuposto de estacionariedade μ_t será constante (μ) enquanto $\Gamma(t, t+h)$ será uma função somente do lag h , e será escrito como $\Gamma(h)$. Assim, a matriz de covariâncias cruzada para o lag 0 é representada como em (4.31).

$$\Gamma_0 = \Gamma_{ij}(0) = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1p} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{p1} & \gamma_{p2} & \cdots & \gamma_{pp} \end{pmatrix} = E[(Y_t - \mu_t)(Y_t - \mu_t)^T] \quad (4.31)$$

onde Γ_0 é a matriz de covariâncias $p \times p$ de Y_t . O i -ésimo elemento da diagonal principal da matriz Γ_0 é a variância de Y_{it} , e o (i,j) -ésimo elemento é a covariância entre Y_{it} e Y_{jt} . Considere D uma matriz diagonal consistindo dos elementos:

$$D = \text{diag}(\sqrt{\gamma_{11}}, \sqrt{\gamma_{22}}, \dots, \sqrt{\gamma_{pp}}) \quad (4.32)$$

Dessa forma, Tsay (2002) mostra que a matriz de correlação cruzada no *lag* 0 é dada por:

$$\rho_0 \equiv [\rho_{ij}(0)] = D^{-1} \Gamma_0 D^{-1} \quad (4.33)$$

mais especificadamente temos:

$$\rho_{ij}(0) = \frac{\gamma_{ij}}{\sqrt{\gamma_{ii} \gamma_{jj}}} \quad (4.34)$$

que representa o coeficiente de correlação entre Y_{it} e Y_{jt} . Além disso, é fácil verificar que:

- (i) $\rho_{ij}(0) = \rho_{ji}(0), \forall (i, j)$
- (ii) $-1 \leq \rho_{ij}(0) \leq 1, \forall (i, j)$
- (iii) $\rho_{ii}(0) = 1, \forall i$

Com o intuito de determinar a matriz de covariâncias cruzada para um processo VAR(1), Kalgonda e Kulkarni (2004) usaram as equações de *Yule-Walker*, ver (Moretтин e Toloι, 2006), e obtiveram o resultado apresentado em (4.35).

$$\Gamma_0 = \Phi \Gamma_0 \Phi' + \Sigma_\varepsilon \quad (4.35)$$

onde Φ e Σ_ε são as matrizes com os parâmetros do modelo temporal.

Como ilustração, vamos considerar o exemplo apresentado em Kalgonda e Kulkarni (2004). Seja um vetor bivariado $Y_t = (Y_{1t}, Y_{2t})^T$ onde se identifica um modelo VAR(1). Considere as seguintes informações:

$$\Sigma_\varepsilon = \begin{pmatrix} 1 & 0,5 \\ 0,5 & 1 \end{pmatrix} \quad \Phi = \begin{pmatrix} 0,5 & 0 \\ 0 & 0,7 \end{pmatrix} \quad \mu = (0,0)^T$$

Assim, utilizando o resultado apresentado em (4.35) tem-se que:

$$\Gamma_0 = \Phi \Gamma_0 \Phi' + \Sigma_\varepsilon$$

$$\begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} 0,5 & 0 \\ 0 & 0,7 \end{pmatrix} \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} 0,5 & 0 \\ 0 & 0,7 \end{pmatrix} + \begin{pmatrix} 1 & 0,5 \\ 0,5 & 1 \end{pmatrix}$$

Desenvolvendo as equações anteriores temos que:

$$\Gamma_0 = \begin{pmatrix} 1,3333 & 0,7692 \\ 0,7692 & 1,9608 \end{pmatrix} \text{ e } \rho(0) = \begin{pmatrix} 1 & 0,4757 \\ 0,4757 & 1 \end{pmatrix}$$

No Apêndice A apresentamos os cálculos para obtenção da matriz Γ_0 proveniente de um processo bivariado VAR(1). Através desses cálculos é possível verificar como a obtenção de Γ_0 está ligada aos parâmetros do processo e a matriz Σ_ε .

Suponha que em uma indústria a qualidade de um produto seja determinada por duas características Y_1 e Y_2 , e estamos interessados em manter a estrutura de variabilidade desse processo multivariado sob controle. Além disso, considere que foi identificado autocorrelação nos vetores de observações e foi ajustado um modelo VAR(1). Dessa forma, deveríamos utilizar a matriz Γ_0 para manter a estrutura de variabilidade do processo sob controle ao invés da matriz de covariância amostral S . Muitos pesquisadores desconhecem a importância de detectar a autocorrelação nos vetores de observações e acabam utilizando a matriz S para manter o processo sob controle, e na maioria das vezes os resultados obtidos são enganosos, já que as matrizes Γ_0 e S são diferentes.

O exemplo anterior aborda o caso onde temos um processo VAR(1), porém, Oliveira (2007) obteve as fórmulas para se encontrar as matrizes de covariâncias cruzadas, para os modelos VAR(2) e VARMA(1,1). No artigo de Kalgonda e Kulkarni (2004) é abordado somente o caso onde os parâmetros da diagonal secundária de Φ , do modelo VAR(1), são nulos, ou seja, $\phi_{12} = \phi_{21} = 0$. Dessa forma, na próxima seção apresentaremos o impacto na matriz Γ_0 , quando, além da autocorrelação temporal nas séries, existe uma relação de dependência entre as duas séries temporais, ou seja, $\phi_{12} \neq 0$ e $\phi_{21} \neq 0$. Vale ressaltar que esse estudo será feito considerando um processo bivariado VAR(1).

4.4 Estudo do comportamento de Γ_0 e $\rho(0)$ quando $\phi_{12} \neq 0$ e $\phi_{21} \neq 0$

Com o auxílio do *software* R 2.9.0, vamos avaliar o comportamento das matrizes Γ_0 , $\rho(0)$, de um processo bivariado VAR(1), quando além da dependência temporal, temos correlação entre as duas séries, ou seja, $\phi_{12} \neq 0$ e $\phi_{21} \neq 0$. Como ilustração, no Apêndice A, apresentamos o cálculo da matriz de covariâncias cruzada, para um processo bivariado VAR(1), considerando $\phi_{12} = \phi_{21}$. As equações para um processo bivariado VAR(1) são dadas por:

$$Y_{1t} = \phi_{11} Y_{1,t-1} + \phi_{12} Y_{2,t-1} + \varepsilon_{1t}$$

$$Y_{2t} = \phi_{21} Y_{1,t-1} + \phi_{22} Y_{2,t-1} + \varepsilon_{2t}$$

Em todos os casos vamos considerar $\Sigma_\varepsilon = \begin{pmatrix} 1 & 0,5 \\ 0,5 & 1 \end{pmatrix}$, $\Phi = \begin{pmatrix} 0,5 & \phi_{12} \\ \phi_{21} & 0,7 \end{pmatrix}$, $\mu = (0,0)'$.

Kalgonda e Kulkarni (2004) mostraram que se $\phi_{12} = \phi_{21} = 0$ tem-se que:

$$\Gamma_0 = \begin{pmatrix} 1,3333 & 0,7692 \\ 0,7692 & 1,9608 \end{pmatrix} \text{ e } \rho(0) = \begin{pmatrix} 1 & 0,4757 \\ 0,4757 & 1 \end{pmatrix}$$

No artigo, não foi construído um exemplo para o caso em que $\phi_{12} \neq 0$ e $\phi_{21} \neq 0$. Dessa forma, para avaliar qual o comportamento de Γ_0 e $\rho(0)$, essas matrizes foram construídas para vários valores de ϕ_{12} e ϕ_{21} . Os resultados são apresentados a seguir.

Caso 1: $\phi_{12} = \phi_{21} \geq 0$

Tabela 4.1: Alterações na matriz Γ_0 quando $\phi_{12} = \phi_{21} \geq 0$

Parâmetros	$\phi_{12} = \phi_{21}$				
	0,0	0,10	0,20	0,30	0,38
γ_{11}	1,3333	1,5177	1,9994	3,7190	39,7829
γ_{22}	1,9608	2,3068	3,1437	6,0371	66,1603
$\gamma_{12} = \gamma_{21}$	0,7692	1,1521	1,8689	4,1529	50,7465
$\rho(0)$	0,4757	0,6157	0,7455	0,8765	0,9891

Caso 2: $\phi_{12} = \phi_{21} \leq 0$

Tabela 4.2: Alterações na matriz Γ_0 quando $\phi_{12} = \phi_{21} \leq 0$

Parâmetros	$\phi_{12} = \phi_{21}$				
	0,0	-0,10	-0,20	-0,30	-0,38
γ_{11}	1,3333	1,2944	1,4052	1,9721	14,4952
γ_{22}	1,9608	1,8552	1,9994	2,9060	23,7489
$\gamma_{12} = \gamma_{21}$	0,7692	0,4772	0,1304	-0,7251	-16,9527
$\rho(0)$	0,4757	0,3079	0,0778	-0,3029	-0,9137

Caso 3: $\phi_{12} = 0$ e $\phi_{21} \geq 0$

Tabela 4.3: Alterações na matriz Γ_0 quando $\phi_{12} = 0$ e $\phi_{21} \geq 0$

Param.	ϕ_{21}										
	0,0	0,10	0,20	0,30	0,40	0,50	0,60	0,70	0,80	0,90	1,0
γ_{11}	1,3333	1,3333	1,3333	1,3333	1,3333	1,3333	1,3333	1,3333	1,3333	1,3333	1,3333
γ_{22}	1,9608	2,2262	2,6003	3,0830	3,6742	4,3741	5,1825	6,0995	7,1252	8,2594	9,5023
$\gamma_{12} = \gamma_{21}$	0,7692	0,8718	0,9744	1,0769	1,1795	1,2821	1,3846	1,4872	1,5897	1,6923	1,7949
$\rho(0)$	0,4757	0,5060	0,5233	0,5312	0,5329	0,5309	0,5267	0,5215	0,5158	0,5100	0,5043

Caso 4: $\phi_{12} = 0$ e $\phi_{21} < 0$

Tabela 4.4: Alterações na matriz Γ_0 quando $\phi_{12} = 0$ e $\phi_{21} < 0$

Param,	ϕ_{21}										
	0,0	-0,10	-0,20	-0,30	-0,40	-0,50	-0,60	-0,70	-0,80	-0,90	-1,0
γ_{11}	1,3333	1,3333	1,3333	1,3333	1,3333	1,3333	1,3333	1,3333	1,3333	1,3333	1,3333
γ_{22}	1,9608	1,8039	1,7557	1,8160	1,9849	2,2624	2,6486	3,1433	3,7466	4,4585	5,279
$\gamma_{12} = \gamma_{21}$	0,7692	0,6667	0,5641	0,4615	0,359	0,2564	0,1538	0,0513	-0,0513	-0,1538	-0,2564
$\rho(0)$	0,4757	0,4299	0,3687	0,2966	0,2207	0,1476	0,0819	0,025	-0,0229	-0,0631	-0,0966

Caso 5: $\phi_{12} \geq 0$ e $\phi_{21} = 0$

Tabela 4.5: Alterações na matriz Γ_0 quando $\phi_{12} \geq 0$ e $\phi_{21} = 0$

Parâmetros	ϕ_{12}										
	0,0	0,10	0,20	0,30	0,40	0,50	0,60	0,70	0,80	0,90	1,0
γ_{11}	1,3333	1,4902	1,7557	2,1297	2,6124	3,2036	3,9035	4,7119	5,6290	6,6546	7,7888
γ_{22}	1,9608	1,9608	1,9608	1,9608	1,9608	1,9608	1,9608	1,9608	1,9608	1,9608	1,9608
$\gamma_{12} = \gamma_{21}$	0,7692	0,9804	1,1916	1,4027	1,6139	1,8250	2,0362	2,2474	2,4585	2,6697	2,8808
$\rho(0)$	0,4757	0,5735	0,6422	0,6864	0,7131	0,7282	0,7360	0,7394	0,7400	0,7391	0,7372

Caso 6.: $\phi_{12} \leq 0$ e $\phi_{21} = 0$

Tabela 4.6: Alterações na matriz Γ_0 quando $\phi_{12} \leq 0$ e $\phi_{21} = 0$

Parâmetros	ϕ_{12}										
	0,0	-0,10	-0,20	-0,30	-0,40	-0,50	-0,60	-0,70	-0,80	-0,90	-1,0
γ_{11}	1,3333	1,2851	1,3454	1,5143	1,7919	2,178	2,6727	3,276	3,9879	4,8084	5,7376
γ_{22}	1,9608	1,9608	1,9608	1,9608	1,9608	1,9608	1,9608	1,9608	1,9608	1,9608	1,9608
$\gamma_{12} = \gamma_{21}$	0,7692	0,5581	0,3469	0,1357	-0,0754	-0,2866	-0,4977	-0,7089	-0,9201	-1,1312	-1,3424
$\rho(0)$	0,4757	0,3516	0,2136	0,0788	-0,0402	-0,1387	-0,2174	-0,2797	-0,329	-0,3684	-0,4002

Caso 7: $\phi_{12} \neq 0$ e $\phi_{21} \neq 0$

Tabela 4.7: Alterações na matriz Γ_0 quando $\phi_{12} \neq 0$ e $\phi_{21} \neq 0$

Parâmetros	$\phi_{12} \neq 0 ; \phi_{21} \neq 0$										
	$\phi_{12} = 0$ $\phi_{21} = 0$	$\phi_{12} = 0,15$ $\phi_{21} = 0,10$	$\phi_{12} = 0,20$ $\phi_{21} = 0,10$	$\phi_{12} = 0,30$ $\phi_{21} = 0,10$	$\phi_{12} = -0,15$ $\phi_{21} = -0,10$	$\phi_{12} = -0,20$ $\phi_{21} = -0,10$	$\phi_{12} = -0,30$ $\phi_{21} = -0,10$	$\phi_{12} = -0,40$ $\phi_{21} = 0,10$	$\phi_{12} = 0,50$ $\phi_{21} = -0,10$	$\phi_{12} = -0,30$ $\phi_{21} = 0,20$	$\phi_{12} = 0,30$ $\phi_{21} = -0,40$
γ_{11}	1,3333	1,6654	1,8546	2,3751	1,3154	1,3664	1,5698	1,7413	2,7756	1,4964	1,8161
γ_{22}	1,9608	2,3524	2,4019	2,5152	1,8844	1,9164	1,9900	2,0051	1,6474	2,2200	1,7973
$\gamma_{12} = \gamma_{21}$	0,7692	1,3075	1,4746	1,8499	0,3722	0,2593	0,0058	0,0371	1,3397	0,2584	0,6678
$\rho(0)$	0,4757	0,6606	0,6987	0,7569	0,2364	0,1603	0,0033	0,0199	0,6265	0,1418	0,3696

4.5 Estudo da distribuição empírica de $|S|$ e $|\hat{\Gamma}_0|$ em processos autocorrelacionados

Com o intuito de avaliarmos a distribuição empírica de $|S|$ e $|\hat{\Gamma}_0|$, em um processo bivariado VAR(1), geramos 100.000 amostras provenientes desse processo, considerando diferentes cenários de tamanhos de amostra ($n = 25, 50, 100, 200$ e 1000) e o vetor de médias desconhecido, com os seguintes parâmetros:

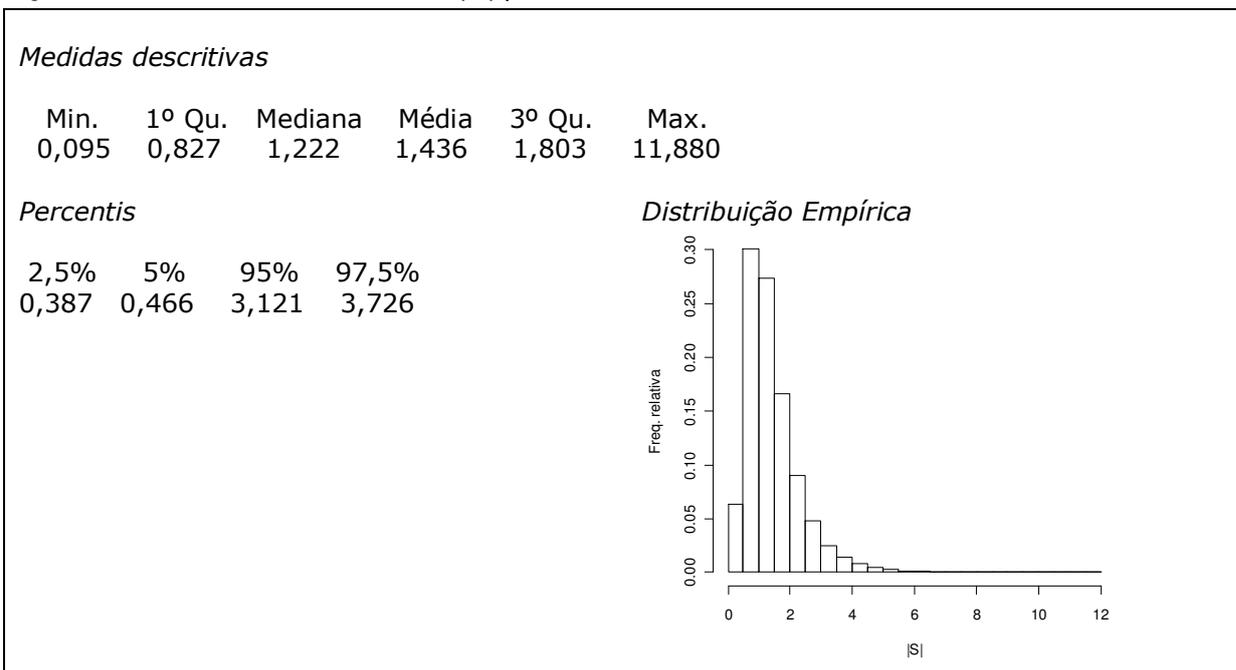
$$\Phi = \begin{pmatrix} \phi_{11} & 0 \\ 0 & \phi_{22} \end{pmatrix} = \begin{pmatrix} 0,5 & 0 \\ 0 & 0,7 \end{pmatrix} \text{ e } \Sigma_\varepsilon = \begin{pmatrix} 1 & 0,5 \\ 0,5 & 1 \end{pmatrix}$$

Por Kalgonda e Kulkarni (2004) espera-se que com esses valores obtenhamos a seguinte matriz de covariâncias cruzada:

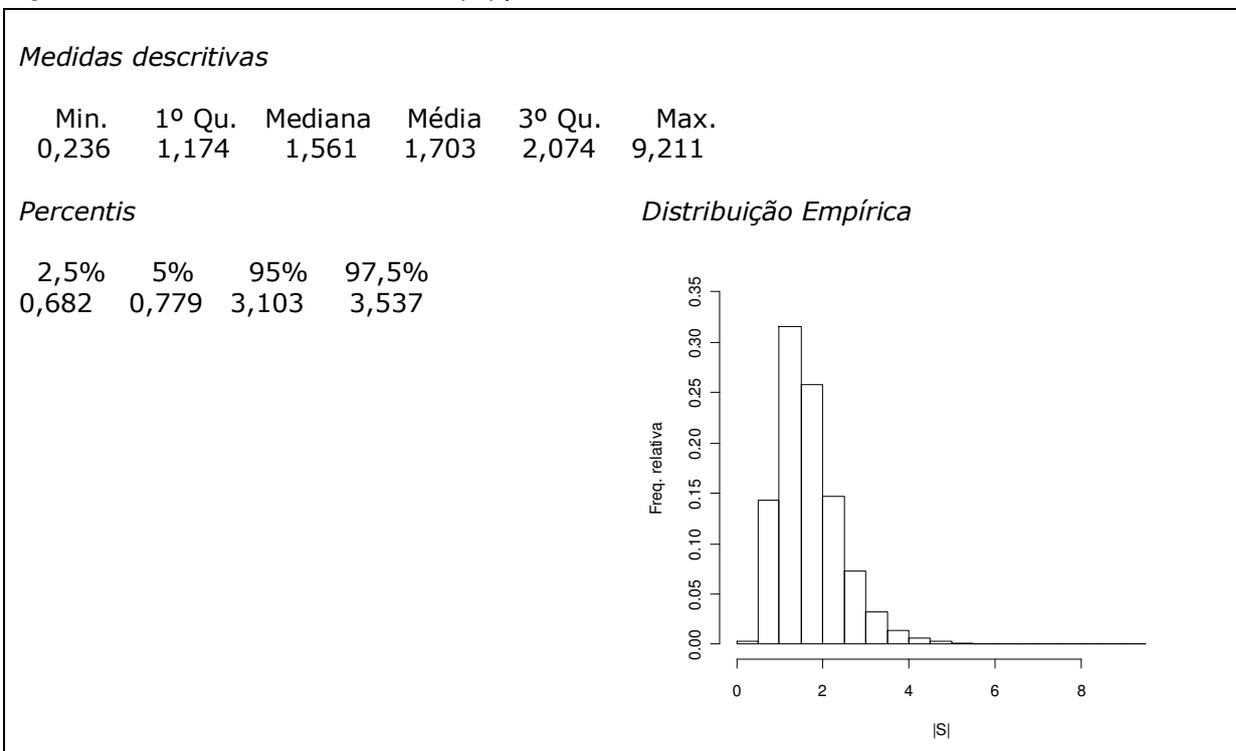
$$\Gamma_0 = \begin{pmatrix} 1,3333 & 0,7692 \\ 0,7692 & 1,9608 \end{pmatrix}; \text{ sendo } |\Gamma_0| = 2,020$$

Vale ressaltar que para gerarmos um processo bivariado VAR (1) utilizamos a proposta de simulação que será apresentada no Capítulo 5. A estimação dos parâmetros foi feita pelo método da máxima verossimilhança apresentado na seção 4.1.1. Utilizamos como suporte nas simulações o software R 2.9.0. Para cada amostra determinou-se o valor de $|S|$ e $|\hat{\Gamma}_0|$ construindo-se no final a distribuição empírica de $|S|$ e $|\hat{\Gamma}_0|$. Para cada uma das situações apresentamos algumas medidas descritivas dos valores de $|S|$ e $|\hat{\Gamma}_0|$, além dos percentis e do histograma da distribuição empírica. Os resultados obtidos para $|S|$ encontram-se nos Quadros 4.1 a 4.5, enquanto que para $|\hat{\Gamma}_0|$ nos Quadros 4.6 a 4.10.

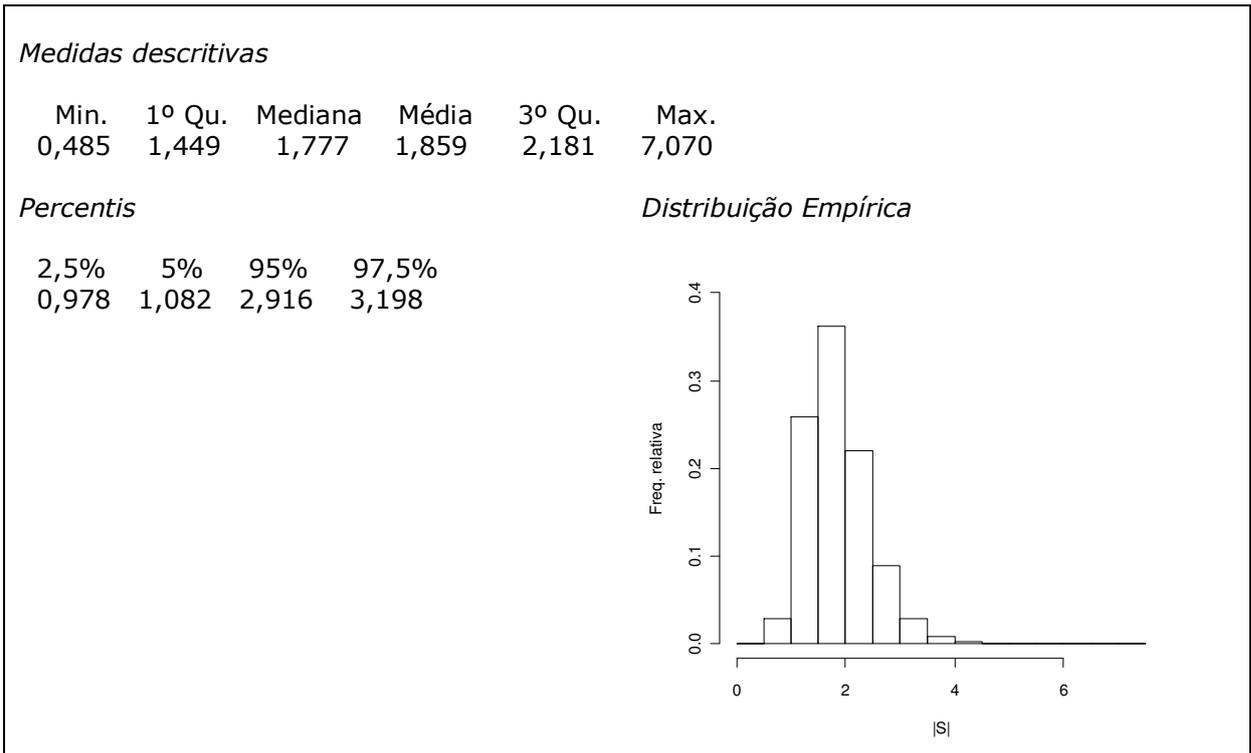
Quadro 4.1: Medidas descritivas de $|S|$ para amostras de tamanho $n = 25$



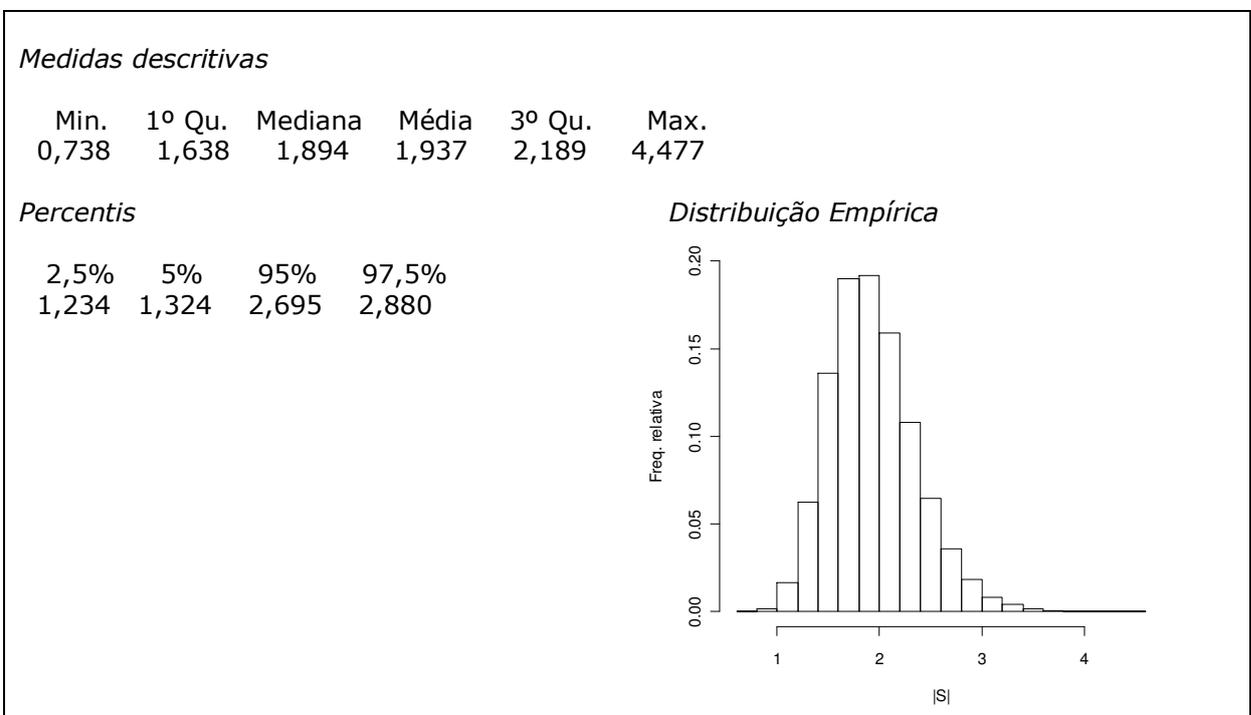
Quadro 4.2: Medidas descritivas de $|S|$ para amostras de tamanho $n = 50$



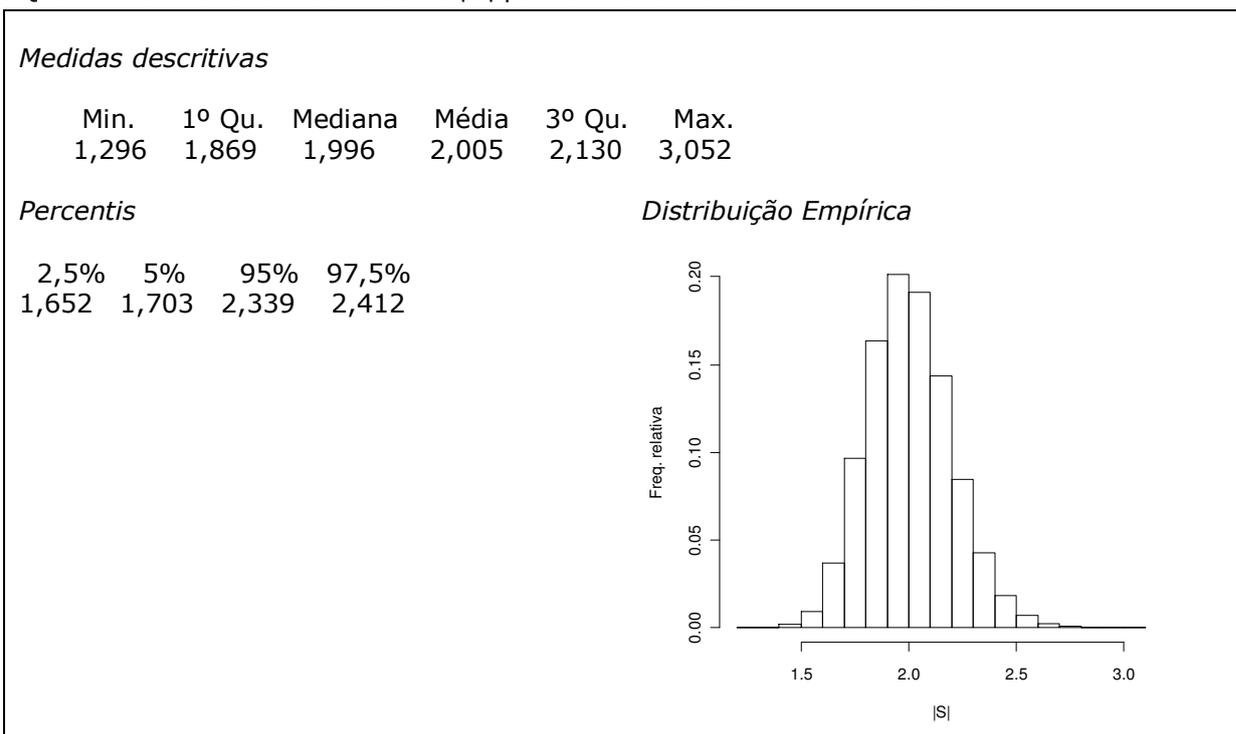
Quadro 4.3: Medidas descritivas de |S| para amostras de tamanho n = 100



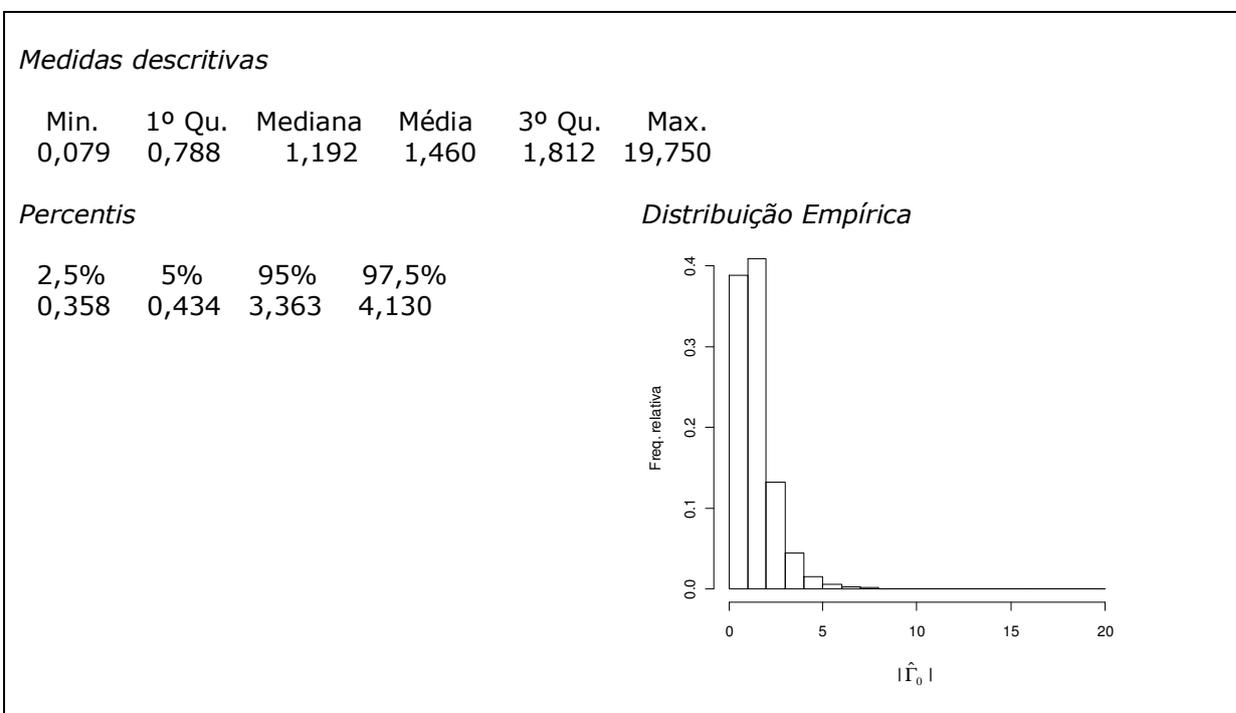
Quadro 4.4: Medidas descritivas de |S| para amostras de tamanho n = 200



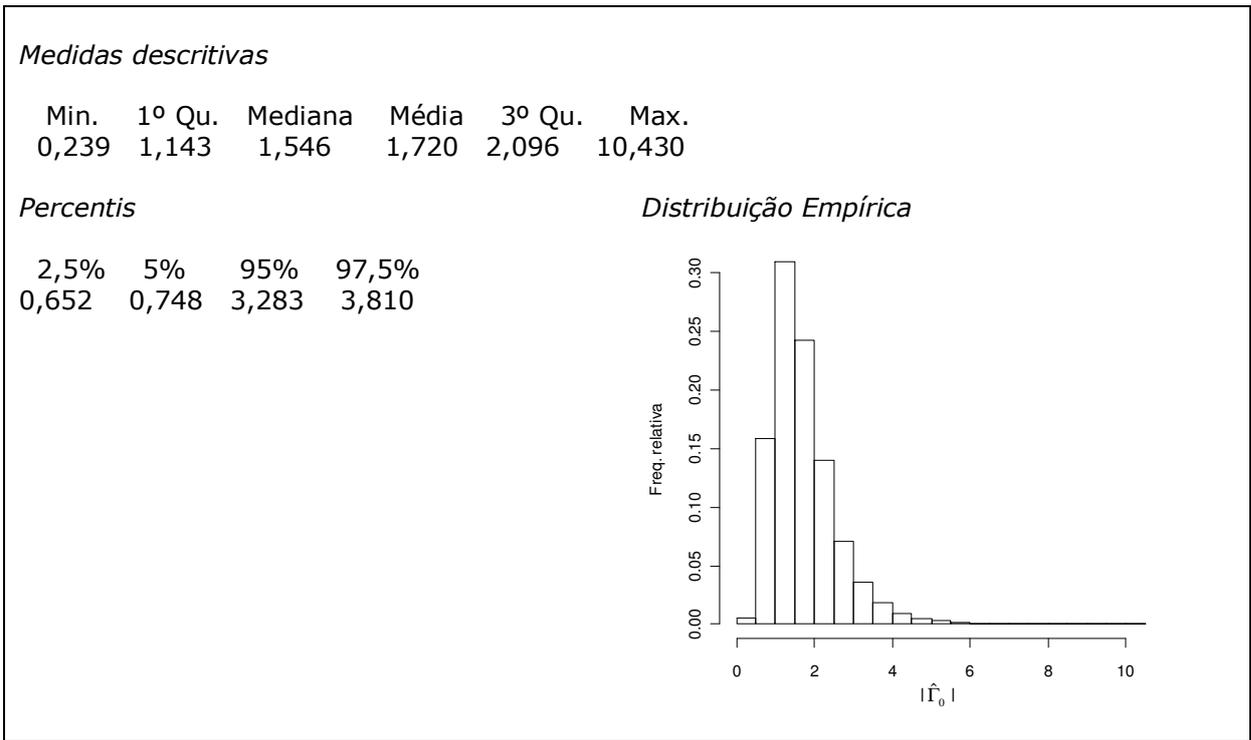
Quadro 4.5: Medidas descritivas de $|S|$ para amostras de tamanho $n = 1000$



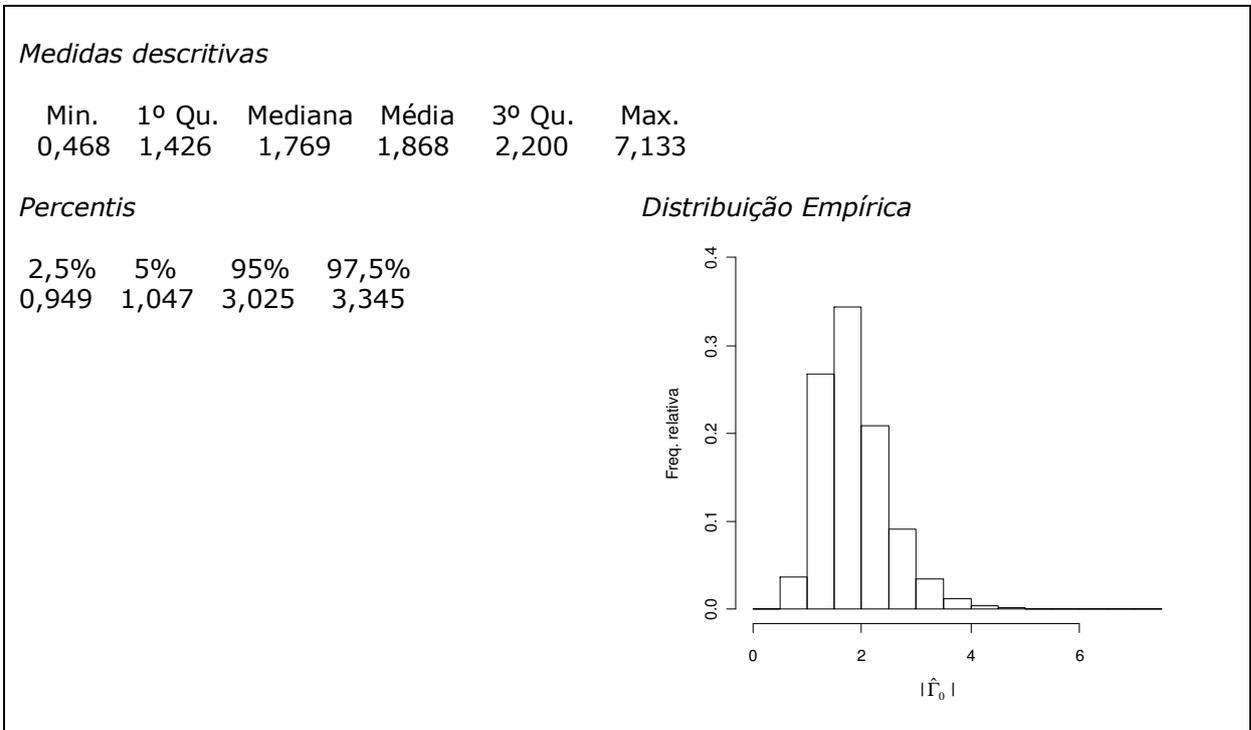
Quadro 4.6: Medidas descritivas de $|\hat{\Gamma}_0|$ para amostras de tamanho $n = 25$



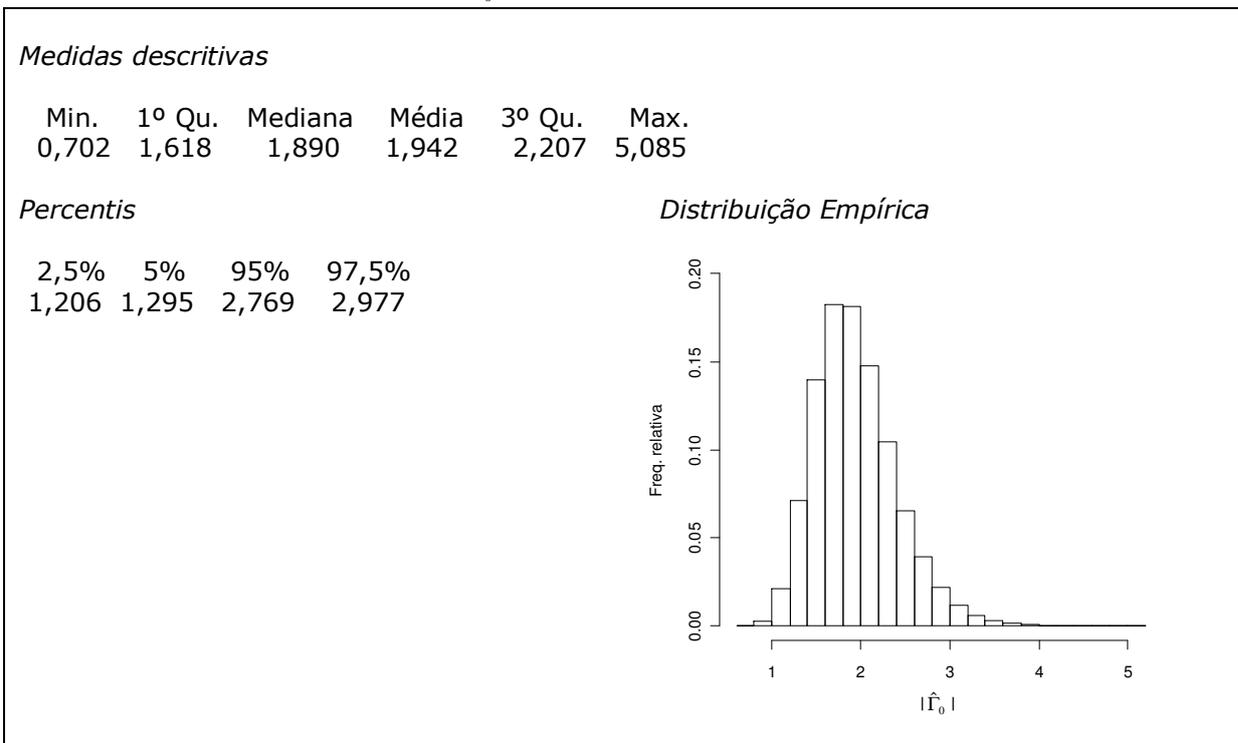
Quadro 4.7: Medidas descritivas de $|\hat{\Gamma}_0|$ para amostras de tamanho $n = 50$



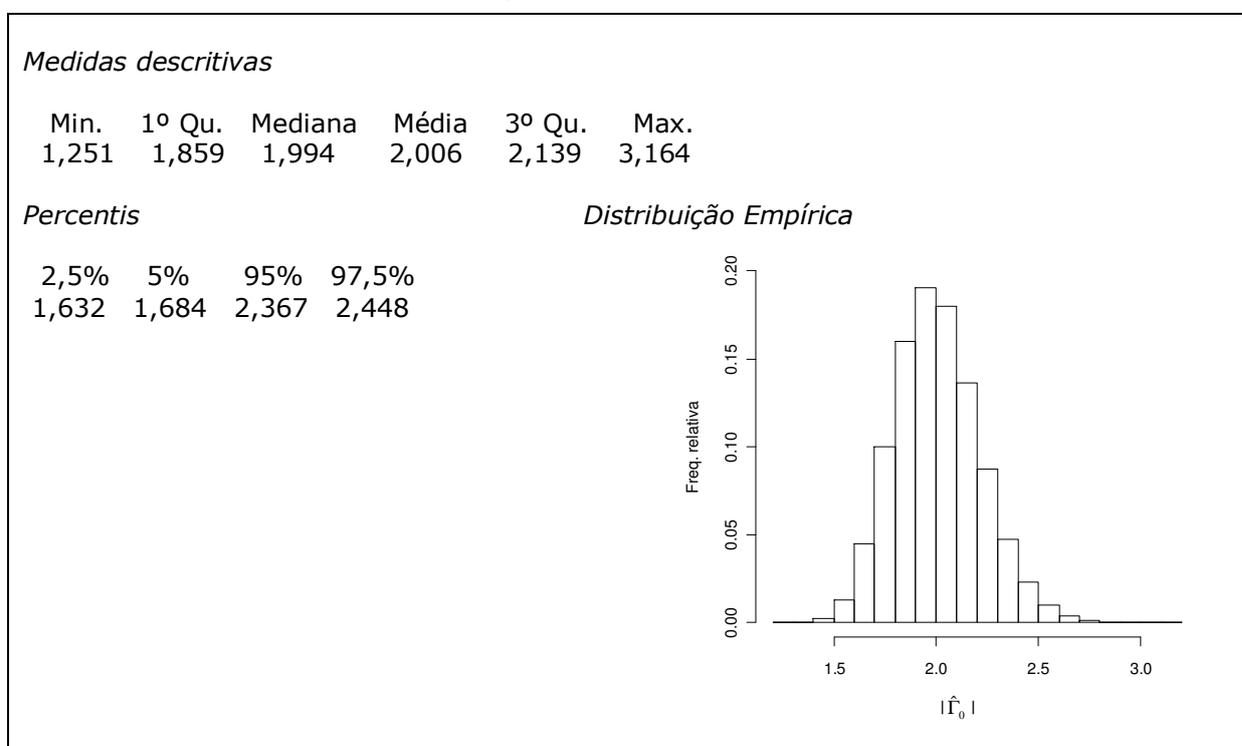
Quadro 4.8: Medidas descritivas de $|\hat{\Gamma}_0|$ para amostras de tamanho $n = 100$



Quadro 4.9: Medidas descritivas de $|\hat{\Gamma}_0|$ para amostras de tamanho $n = 200$



Quadro 4.10: Medidas descritivas de $|\hat{\Gamma}_0|$ para amostras de tamanho $n = 1000$



Através da análise dos histogramas, para pequenas amostras, poderíamos pensar em uma aproximação para a distribuição qui-quadrado, tanto para $|S|$ quanto para $|\hat{\Gamma}_0|$. Além disso, nota-se que para tamanhos amostrais grandes ($n=1000$) as distribuições de $|S|$ e $|\hat{\Gamma}_0|$ são bem próximas, o que nos leva a concluir a existência de uma convergência dessas duas matrizes. Um outro ponto a ser discutido é que o valor médio de $|\hat{\Gamma}_0|$ converge mais lentamente para o verdadeiro valor de $|\Gamma_0|=2,02$ do que $|S|$ como pode ser visto pela Tabela 4.8.

Tabela 4.8: Valores médios de $|S|$ e $|\hat{\Gamma}_0|$ para diferentes cenários de tamanhos de amostras

n	$ S $	$ \hat{\Gamma}_0 $
25	1,436	1,460
50	1,703	1,720
100	1,859	1,868
200	1,937	1,942
1000	2,005	2,006

Observe que para amostras pequenas ($n=25$) os valores médios de $|S|$ e $|\hat{\Gamma}_0|$ são diferentes, o que nos leva a reforçar a idéia de que quando estamos em um processo autocorrelacionado, o uso da matriz S para monitorar a variabilidade do processo pode nos levar a resultados errôneos.

Nas próximas seções apresentaremos a adaptação dos testes apresentados no Capítulo 3 para o caso em que temos autocorrelação, ou seja, mostraremos como a matriz $\hat{\Gamma}_0$ pode ser utilizada nos testes estatísticos para matriz de covariâncias.

4.6 Adaptação dos testes para matriz de covariâncias em processos autocorrelacionados

Se em um processo a autocorrelação for detectada é essencial usar um modelo apropriado de séries temporais para modelar os vetores de observações, tendo em conta a estrutura de dependência temporal. Uma vez identificado o modelo que melhor se ajusta ao conjunto de observações, (ver Lutkepohl, 1993), o próximo passo é obter a matriz de covariâncias cruzadas ($\hat{\Gamma}_0$) do processo e incorporá-la nos testes de hipótese

utilizados para monitorar a matriz de covariâncias do processo. No Capítulo 3 apresentamos os testes para matriz de covariâncias considerando independência entre os vetores de observações. Nas seções 4.6.1 a 4.6.4 apresentaremos uma modificação desses testes para o caso em que a autocorrelação está presente nos dados. Como feito no Capítulo 3 consideraremos os casos em que o vetor de médias é conhecido e desconhecido.

4.6.1 Modificação do Teste do Determinante Tradicional e Djauhari (2005)

Suponha que em um processo autocorrelacionado nosso foco seja testar a hipótese $H_0: \Gamma = \Gamma_0$ versus $H_1: \Gamma \neq \Gamma_0$, onde Γ_0 é uma matriz de covariâncias pré-especificada. Para isso os seguintes limites de controle serão considerados:

<i>Teste Tradicional</i>	<i>Teste Djauhari(2005)</i>
$LSC = \Gamma_0 (b_1 + 3\sqrt{b_2})$ $LC = b_1 \Gamma_0 $ $LIC = \max(0; \Gamma_0 (b_1 - 3\sqrt{b_2}))$	$LSC = \Gamma_0 \left(\frac{b_1}{b_3} + 3\sqrt{\frac{b_2}{b_3^2 + b_4}} \right)$ $LC = \Gamma_0 \frac{b_1}{b_3}$ $LIC = \max \left(0; \Gamma_0 \left(\frac{b_1}{b_3} - 3\sqrt{\frac{b_2}{b_3^2 + b_4}} \right) \right)$

onde b_1, b_2, b_3 e b_4 são aqueles apresentados como no Capítulo 3 (seções 3.2 e 3.3).

Para o teste do determinante proposto por Djauhari (2005) a mesma correção é feita, ou seja, a troca da matriz Σ , sob H_0 , pela matriz Γ_0 . Para o caso em que o vetor de médias do processo é conhecido o modelo de séries temporais é ajustado aos dados sem estimar o vetor de médias e o determinante da estimativa da matriz Γ_0 , aqui denotado por $|\hat{\Gamma}_0^*|$, é utilizado para monitorar o processo. No caso em que o vetor de médias é desconhecido ajustamos o modelo de séries temporais estimando todos os parâmetros, inclusive o vetor de médias, e utilizamos o determinante da estimativa da matriz Γ_0 , denotada nesse caso por $|\hat{\Gamma}_0|$, para testar a hipótese sobre a matriz de covariâncias. Vale ressaltar que além dos limites de controle obtidos para o teste tradicional via aproximação pela distribuição normal e o proposto por Djauhari (2005), que também se utiliza dessa aproximação, também utilizaremos os limites exatos obtidos via simulação.

4.6.2 Modificação do Teste da razão de verossimilhança

Suponha que estejamos interessados em testar a hipótese: $H_0: \Gamma = \Gamma_0$ versus $H_1: \Gamma \neq \Gamma_0$, sendo que a matriz Γ_0 (alvo) é completamente especificada em H_0 . No caso em que há autocorrelação entre as observações, é possível mostrar, usando procedimentos similares aos do Capítulo 3 (seção 3.5), que para o caso em que o vetor de médias é conhecido, a estatística do teste de razão de verossimilhança com a aproximação qui-quadrado assintótica ($-2 \ln(\Lambda)$) é dada por:

$$-2 \ln \Lambda = -2 \ln \left(\frac{(2\pi)^{-\frac{np}{2}} |\Gamma_0|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{j=1}^n (x_j - \mu)^T \Gamma_0^{-1} (x_j - \mu) \right\}}{(2\pi)^{-\frac{np}{2}} |\hat{\Gamma}_0|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{j=1}^n (x_j - \mu)^T \hat{\Gamma}_0^{-1} (x_j - \mu) \right\}} \right) \quad (4.36)$$

onde $\hat{\Gamma}_0$ é a estimativa de máxima verossimilhança de Γ_0 . Quando fazemos a adaptação do teste da razão de verossimilhança para o caso em que temos autocorrelação o número de graus de liberdade na distribuição qui-quadrado se altera, visto que temos um maior número de parâmetros a ser estimado. Considerando o caso em que o vetor de médias é conhecido e estamos em um processo bivariado VAR(1) utilizamos 7 graus de liberdade. Isso porque, quando estamos no espaço paramétrico irrestrito estimamos os 7 parâmetros do modelo VAR(1), isto é, $(\sigma_{11}, \sigma_{22}, \sigma_{21}, \phi_{11}, \phi_{22}, \phi_{12}$ e $\phi_{21})$. À medida que a ordem do modelo ajustado se altera o número de graus de liberdade também muda.

Para o caso em que o vetor de médias é desconhecido a estatística do teste de razão de verossimilhança com a aproximação qui-quadrado assintótica ($-2 \ln(\Lambda)$) será dada por:

$$-2 \ln \Lambda = -2 \ln \left(\frac{(2\pi)^{-\frac{np}{2}} |\Gamma_0|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{j=1}^n (x_j - \tilde{\mu}^*)^T \Gamma_0^{-1} (x_j - \tilde{\mu}^*) \right\}}{(2\pi)^{-\frac{np}{2}} |\hat{\Gamma}_0|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{j=1}^n (x_j - \tilde{\mu})^T \hat{\Gamma}_0^{-1} (x_j - \tilde{\mu}) \right\}} \right) \quad (4.37)$$

onde $\tilde{\mu}$ é a estimativa de máxima verossimilhança de μ considerando a matriz de parâmetros estimados $\hat{\Phi}$ e $\tilde{\mu}^*$ é a estimativa de máxima verossimilhança de μ considerando a matriz de parâmetros Φ sob H_0 (teórica). Considerando que o vetor de médias seja desconhecido e que estamos em um processo bivariado VAR(1) utilizamos 7 graus de liberdade na distribuição assintótica qui-quadrado. Isso porque, sob o espaço paramétrico irrestrito estimamos os 9 parâmetros do modelo, ou seja, $(\mu_1, \mu_2, \sigma_{11}, \sigma_{22}, \sigma_{21}, \phi_{11}, \phi_{22}, \phi_{12}$ e $\phi_{21})$ e sob o espaço paramétrico restrito estimamos o vetor de médias μ de dimensão 2×1 . Observe que é necessário fazer um estudo do número de graus de liberdade em cada situação, levando em consideração o modelo ajustado, número de variáveis, etc.

Nesta dissertação, além das distribuições assintóticas, consideraremos as distribuições exatas geradas via simulação.

4.6.3 Modificação do Teste Step-down proposto por Sullivan et al. (2007)

Suponha que estamos interessados em testar a seguinte hipótese:

$$\begin{cases} H_0 : \Gamma = \Gamma_0 = \begin{pmatrix} \gamma_{0,11} & \gamma_{0,12} \\ \gamma_{0,21} & \gamma_{0,22} \end{pmatrix} \\ H_1 : \Gamma \neq \Gamma_0 \end{cases}$$

onde $\gamma_{0,11}, \gamma_{0,12}, \gamma_{0,21}, \gamma_{0,22}$, são os componentes da matriz Γ_0 , sob H_0 .

Além disso, considere que o vetor de médias (μ) do processo seja conhecido. Portanto, a matriz de covariância cruzada estimada ($\hat{\Gamma}_0^*$) é dada por:

$$\hat{\Gamma}_0^* = \begin{pmatrix} \hat{\gamma}_{11}^* & \hat{\gamma}_{12}^* \\ \hat{\gamma}_{21}^* & \hat{\gamma}_{22}^* \end{pmatrix} \quad (4.38)$$

A modificação do teste de Sullivan et al. (2007) está baseada no uso da estatística dada em (4.39), que sob H_0 , possui distribuição assintótica qui-quadrado. Como agora estamos considerando a matriz $\hat{\Gamma}_0^*$ como estimador da matriz de covariâncias do processo temos o seguinte vetor de parâmetros $\hat{\theta} = (\hat{\gamma}_{11}^*, \hat{\rho}^*, \hat{\gamma}_{22}^*)^T$.

$$T^2 = \delta^T \Sigma_{\hat{\delta}}^{-1} \delta \quad \text{onde} \quad \delta = (\hat{\theta} - \theta_0) \quad (4.39)$$

Dessa forma, a idéia do teste será comparar as estimativas dos parâmetros de $\hat{\theta} = (\hat{\gamma}_{11}^*, \hat{\rho}^*, \hat{\gamma}_{22}^*)^T$ com aqueles sob H_0 , $\theta_0 = (\gamma_{0,11}, \rho_0, \gamma_{0,22})^T$. Vale lembrar que a matriz $\Sigma_{\hat{\delta}}$ é obtida utilizando a matriz de Informação de Fisher como apresentada no Capítulo 3.

No caso em que o vetor de médias do processo é desconhecido utilizamos a estimativa de máxima verossimilhança $\hat{\Gamma}_0$, que é feita considerando o vetor de médias desconhecido. Nesse caso teríamos a comparação entre os vetores $\theta_0 = (\gamma_{0,11}, \rho_0, \gamma_{0,22})^T$ e $\hat{\theta} = (\hat{\gamma}_{11}, \hat{\rho}, \hat{\gamma}_{22})^T$. Como estamos interessados somente na parte que contém variação os vetores possuem dimensão 3x1, visto que o vetor de médias não está sendo considerado.

Quanto ao número de graus de liberdade na distribuição assintótica qui-quadrado, deve-se fazer um estudo detalhado de quais parâmetros estão sendo estimados, uma vez que o número de graus de liberdade pode alterar de acordo com o modelo ajustado. Suponha que tenhamos ajustado um modelo bivariado VAR(1). Se considerarmos o vetor de médias conhecido, usamos a aproximação para a distribuição qui-quadrado com 7 graus de liberdade (7 parâmetros estimados no espaço irrestrito). Para o caso em que consideramos o vetor de médias desconhecido, podemos usar 7 graus de liberdade, se seguirmos a idéia original do artigo de Sullivan et al. (2007), ou 9 graus de liberdade se pensarmos na estimação do vetor de médias, dado que para se obter as estimativas das variâncias e covariâncias é necessário estimar o vetor de médias.

Nesta dissertação, além das distribuições assintóticas também consideraremos a distribuição exata da estatística desse teste.

4.6.4 Modificação do Teste VMAX proposto por Machado et. al (2008)

O teste proposto por Machado et al. (2008), denominado de VMAX, baseia-se no maior valor das variâncias amostrais padronizadas. Suponha que em um processo autocorrelacionado estejamos interessados em testar a hipótese: $H_0: \Gamma = \Gamma_0$ versus $H_1: \Gamma \neq \Gamma_0$, sendo que a matriz Γ_0 é completamente especificada em H_0 . Para o caso em que o vetor de médias é conhecido, ajustamos o modelo de séries temporais adequado às

observações sem estimar o vetor de médias e a partir disto obtemos a matriz $\hat{\Gamma}_0^*$. Sabemos que o teste VMAX está baseado no valor da maior variância padronizada. Diante disso, a estatística de teste será dada por:

$$VMAX = \max \left\{ \frac{\hat{\gamma}_{11}^*}{\gamma_{0,11}}; \frac{\hat{\gamma}_{22}^*}{\gamma_{0,22}} \right\} \quad (4.40)$$

Quando o vetor de médias é desconhecido basta trabalhar com a matriz $\hat{\Gamma}_0$ que é estimada sem nenhuma restrição. Ou seja, nesse caso o vetor de médias do modelo também é estimado. Um fato importante é que independentemente se o vetor de médias é conhecido ou não, sempre padronizaremos as variâncias com $\gamma_{0,11}$ e $\gamma_{0,22}$, sob H_0 .

Ao contrário do que ocorre com o teste da razão de verossimilhança e com o teste proposto por Sullivan et al. (2007) o teste VMAX não possui nenhuma distribuição aproximada, diante disso optamos por trabalhar somente com a distribuição exata que foi obtida via simulação.

Capítulo 5

Aspectos Computacionais: Simulação e Estimação dos Parâmetros de um Processo VAR(1)

O R é um software livre para ambiente estatístico e computação gráfica e está disponível no *site* (<http://cran.r-project.org/>). Nesta dissertação o utilizaremos como ferramenta para as simulações que serão realizadas juntamente com o pacote *mAr*, versão 1.1-1. O pacote *mAr* foi desenvolvido por Barbosa (2009) e oferece várias funções para análise de modelos multivariados autorregressivos. Dessa forma, será possível verificar o comportamento de vários testes de hipótese para matriz de covariâncias no contexto em que temos autocorrelação. Realizaremos simulações com diferentes cenários de tamanhos de amostra, com o intuito de verificar até que ponto o tamanho amostral pode interferir nos resultados.

Em um primeiro momento focalizaremos nossos estudos na análise do poder dos testes quando temos um modelo bivariado VAR(1) onde as duas séries não são correlacionadas ($\phi_{12} = \phi_{21} = 0$), ou seja, temos somente a dependência temporal, $\phi_{11} \neq 0$, $\phi_{22} \neq 0$. Denominaremos esse primeiro estudo como situação 1. Uma desvantagem do uso do pacote *mAr*, mas especificamente da função *mAr.est*, na situação 1, é que essa função sempre estimará os parâmetros da diagonal secundária, que representam a relação entre as duas séries, mesmo que teoricamente essa relação não exista, ou seja, $\phi_{12} = \phi_{21} = 0$. Dessa forma, nessa primeira situação, optamos por não usar o pacote *mAr*, uma vez que estamos interessados no caso em que $\phi_{12} = \phi_{21} = 0$ e o uso do pacote nos forneceria resultados pouco confiáveis. Uma forma de contornar esse problema foi criar nosso próprio algoritmo de simulação e estimação. O processo de geração dos vetores de observações, provenientes de um processo bivariado VAR(1), seguiu as seguintes etapas:

1. Gerar $2n$ vetores de $\mathcal{E} = \begin{pmatrix} \mathcal{E}_1 \\ \mathcal{E}_2 \end{pmatrix}$ proveniente de uma $N_p(0, \Sigma_{\mathcal{E}})$;
2. Gerar um vetor de observação $Y_0 \sim N_p(0, \Sigma_{\mathcal{E}})$ para iniciar o processo de geração;

3. Por exemplo, para o caso bivariado, temos:

$$\begin{bmatrix} Y_{1,t} = \phi_{11}Y_{10} + \varepsilon_{1,t} & Y_{2,t} = \phi_{22}Y_{20} + \varepsilon_{2,t} \\ Y_{1,t+1} = \phi_{11}Y_{1,t} + \varepsilon_{1,t+1} & Y_{2,t+1} = \phi_{22}Y_{2,t} + \varepsilon_{2,t+1} \\ \vdots & \vdots \\ Y_{1,t+h} = \phi_{11}Y_{1,(t+h-1)} + \varepsilon_{1,t+h} & Y_{2,t+h} = \phi_{22}Y_{2,(t+h-1)} + \varepsilon_{2,t+h} \end{bmatrix}$$

4. Selecionar somente os n últimos vetores gerados.

Através dos quatro itens descritos anteriormente será gerado um processo bivariado VAR(1). Observe que selecionaremos somente os n últimos vetores de observações, isso porque, no começo do processo de geração, os vetores estão influenciados pelo valor inicial \mathbf{Y}_0 e podem não captar a autocorrelação real existente na série. Assim, se selecionarmos somente os n últimos vetores, obteremos melhores resultados. Quanto a estimação dos parâmetros do modelo, implementamos o procedimento de estimação de máxima verossimilhança como descrito na seção 4.1.1. Vale lembrar que como nosso objetivo é estudar testes de hipótese para matriz de covariâncias consideraremos as situações em que o vetor de médias é conhecido e quando é desconhecido. Para as simulações estipulamos $\boldsymbol{\mu} = (0,0)'$.

Posteriormente, avaliaremos o poder dos testes quando temos um modelo bivariado VAR(1) onde identificamos a correlação temporal e entre as duas séries. Em termos teóricos vamos considerar o caso em que $\phi_{11} \neq 0$, $\phi_{22} \neq 0$, $\phi_{12} \neq 0$ e $\phi_{21} \neq 0$. A opção de trabalharmos com essa situação, que denominaremos de situação 2, é que os trabalhos publicados até o momento na área de controle de qualidade para processos autocorrelacionados em geral se utilizam do modelo de séries temporais VAR(1) com matriz de parâmetros Φ diagonal. Assim, é nossa intenção explorar modelos mais gerais. Na situação 2 utilizaremos como ferramenta a função *mAr.est* e a função *mAr.sim*, do pacote *mAr*. A função *mAr.sim* nos permite gerar processos multivariados autoregressivos de ordem k , bastando para isso fornecer os parâmetros e a ordem do modelo desejado. Para estimar os parâmetros do modelo utilizaremos a função *mAr.est*, que está fundamentada no algoritmo de Neumaier e Schneider (2001). Um fato a ser considerado é que os as estimativas obtidas via máxima verossimilhança e pelo algoritmo de Neumaier e Schneider (2001) são bem próximas. Ao contrário da situação 1, nessa segunda parte consideraremos somente o caso em que o vetor de médias é

desconhecido, isso porque, a função *mAr.est* não nos dá a opção de declararmos o vetor de médias como conhecido, ou seja, a função sempre estimará o vetor de médias.

A avaliação da qualidade dos testes que serão propostos para processos autocorrelacionados será realizada através da determinação do tamanho do teste e de seu poder via simulações de Monte Carlo. Vale ressaltar que os testes descritos nessa dissertação serão avaliados com as distribuições aproximadas e exatas. Para obtermos os limites críticos das distribuições exatas dos testes geramos um processo bivariado VAR(1) e encontramos os percentis da distribuição empírica considerando vários tamanhos de amostra. Os percentis das distribuições exatas, para $n = 25, 50, 100, 200, 1000$ e 5000 , estão apresentados nos Apêndices B e C dessa dissertação. Finalmente, avaliaremos como a autocorrelação interfere nos resultados dos testes estatísticos para matriz de covariâncias formulados teoricamente para observações independentes.

Um estudo do pacote *mAr*, antes de iniciarmos as simulações, nos indicou que teríamos um grande problema ao trabalhar com amostras pequenas. Esse estudo foi feito com o objetivo de avaliarmos a qualidade do pacote *mAr* em gerar (*mAr.sim*) e estimar (*mAr.est*) os parâmetros de um modelo VAR(k), considerando diferentes cenários de tamanhos de amostra. Esse estudo nos mostrou que em situações onde temos pequenas amostras a função *mAr.sim* não é consistente, ou seja, ela gera observações que não são condizentes com o modelo e parâmetros especificados. Isso poderia nos levar a resultados incorretos visto que estaríamos gerando um processo diferente daquele postulado. O desejo de avaliar o desempenho dos testes para tamanhos de amostras pequenas veio do fato de que em controle de qualidade é usual trabalhar com amostras pequenas ($n < 10$), porém, diante de um problema multivariado autocorrelacionado não achamos conveniente trabalhar com amostras tão pequenas. Em algumas situações será necessário estimar até 9 parâmetros ($\mu_1, \mu_2, \sigma_{11}, \sigma_{22}, \sigma_{21}, \phi_{11}, \phi_{22}, \phi_{12}$ e ϕ_{21}), por isso, optamos por trabalhar com amostras de no mínimo 25 observações. Mesmo para 25 observações o algoritmo de simulação do pacote *mAr* ainda é muito inconsistente, portanto, para contornar esse problema uma alternativa foi criar restrições no processo de geração das amostras, ou seja, aceitaríamos somente as amostras que satisfizessem nossos objetivos (modelo e parâmetros). O procedimento de restrição foi feito da seguinte forma: suponha que estejamos interessados em gerar 25 observações, provenientes de um modelo bivariado VAR(1) com parâmetros $\phi_{11} = 0,5, \phi_{22} = 0,7, \phi_{12} = 0$ e $\phi_{21} = 0$. Se não criássemos nenhuma restrição no processo de simulação obteríamos amostras de modelos com parâmetros completamente diferentes do teórico simulado. Diante disso optamos por aceitar somente as amostras que satisfiziam o

critério $\phi_{11} - erro \leq \hat{\phi}_{11} \leq \phi_{11} + erro$, onde o erro era da ordem de 0,05. Essa restrição foi feita nos quatro parâmetros da matriz Φ . A restrição imposta no processo de simulação foi satisfatória, porém, aumentou consideravelmente o tempo de simulação, visto a grande inconsistência da função *mAr.sim* para gerar amostras autocorrelacionadas em cenários com um pequeno número de observações.

A escolha dos testes para os quais iremos avaliar o poder será feita através do tamanho do teste. Ou seja, simularemos n observações, sob H_0 , de um processo bivariado VAR(1) 100.000 vezes e contabilizaremos a quantidade de amostras que levarão a rejeição de H_0 . O tamanho do teste será determinado pela proporção de amostras que levaram a rejeição de H_0 . Como estamos simulando sob H_0 , espera-se que esses valores sejam bem próximos do valor do nível de significância nominal escolhido para o teste (5%). Vale lembrar que consideraremos os limites críticos das distribuições aproximadas e exatas. Uma vez identificado a estimativa do tamanho de cada teste selecionaremos somente aqueles testes que fornecerem valores de $\hat{\alpha}$ em torno de 5%.

Para avaliação do poder dos testes simularemos n observações, sob H_1 , 20.000 vezes e estimaremos a proporção de amostras que levarão a rejeição de H_0 . Altos valores dessa proporção indicam uma eficácia dos testes em detectar alterações na matriz de covariâncias, visto que estamos simulando observações sob H_1 . Valores baixos dessa proporção indicam que os testes não foram sensíveis a tais alterações, já que estamos simulando observações de um modelo diferente daquele postulado sob H_0 . Essa proporção de rejeição é denominada poder do teste. Esse processo será repetido 25 vezes para cada teste, dessa forma, teremos 25 estimativas de poder para cada teste, que é equivalente a uma corrida de 500.000 repetições. Isso será feito com o intuito de termos uma maior confiabilidade nos resultados. Os valores médios dos poderes de cada teste serão apresentados nas próximas seções. A Figura 5.1 mostra, resumidamente, a estratégia de simulação. A contrário do que foi feito na fase determinação do poder dos testes, onde o processo foi repetido 25 vezes, na fase de estimação do tamanho dos testes o processo foi feito somente uma vez.

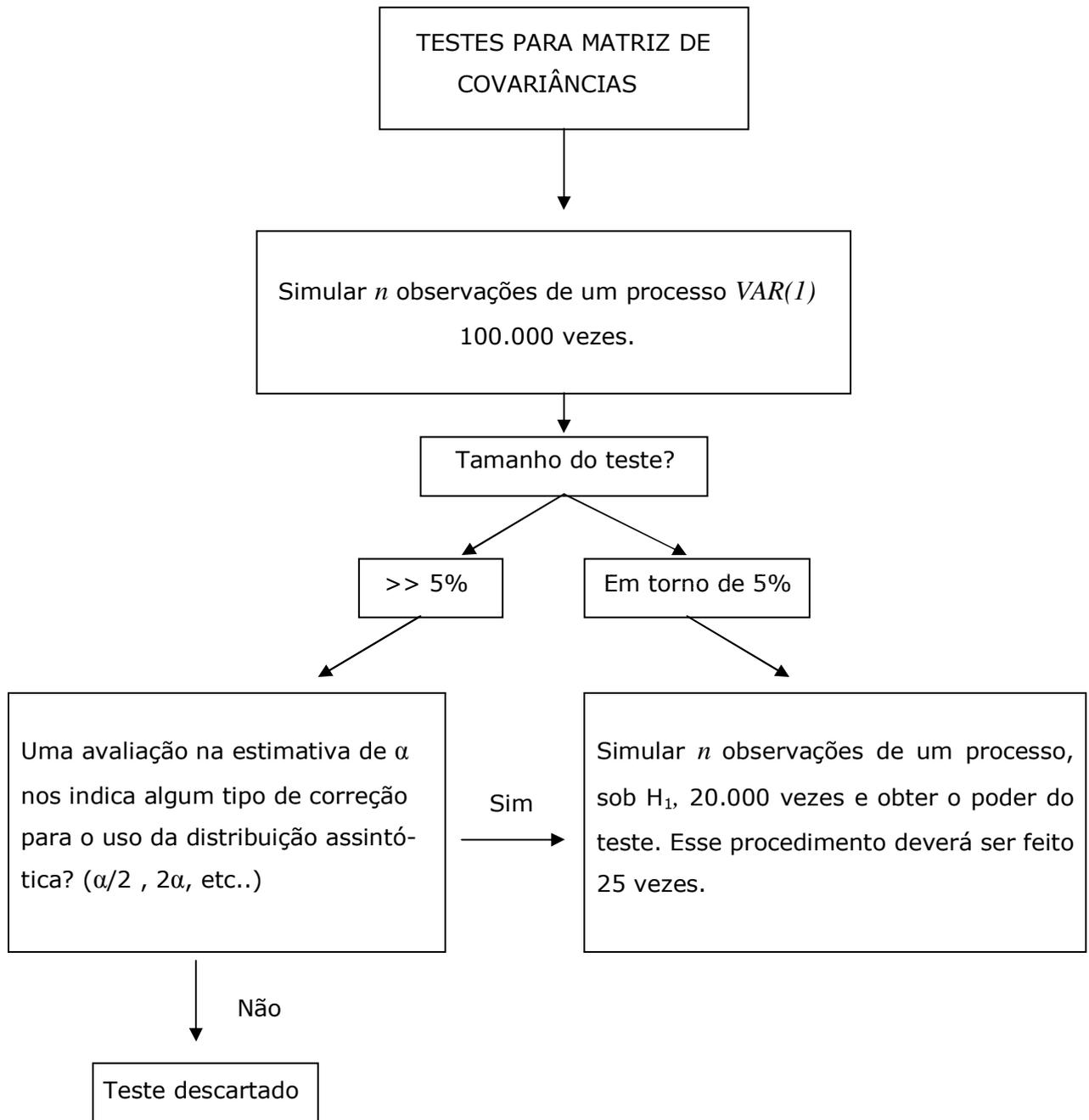


Figura 5.1: Estratégia de Simulação

Capítulo 6

Avaliação do Poder dos Testes para Matriz de Covariâncias via Simulação de Monte Carlo

Neste capítulo apresentaremos os resultados das simulações desenvolvidas onde incorporamos a autocorrelação nos testes de hipótese para matrizes de covariâncias já existentes, que são construídos sob a suposição de independência entre as observações, mas especificamente, os testes do determinante, o de razão de verossimilhança, o de Sullivan et al. (2007) e o VMAX de Machado et al. (2008) discutidos no Capítulo 3. Para efeito de estudo, consideraremos duas situações para análise. Na Situação 1 abordaremos o caso onde temos um modelo VAR(1) e as duas séries não são relacionadas ($\phi_{12} = \phi_{21} = 0$), ou seja, temos somente a dependência temporal $\phi_{11} \neq 0$, $\phi_{22} \neq 0$. Já na Situação 2, consideraremos novamente um modelo VAR(1), mas avaliaremos o caso onde identificamos correlação temporal e entre as duas séries ($\phi_{11} \neq 0$, $\phi_{22} \neq 0$, $\phi_{12} \neq 0$ e $\phi_{21} \neq 0$).

6.1 Estudo do Poder dos Testes - Situação 1

Nessa seção nosso objetivo será avaliar o poder dos testes para matriz de covariâncias, formulados sob a suposição de independência entre as observações, quando as duas séries não são correlacionadas ($\phi_{12} = \phi_{21} = 0$), mas a dependência temporal, $\phi_{11} \neq 0$, $\phi_{22} \neq 0$, é verificada. Para isso, consideraremos as distribuições exatas e aproximadas dos testes em questão. Para obtermos os percentis das distribuições exatas algumas simulações foram feitas como descrito no Capítulo 5. Os resultados dessas simulações estão apresentados no Apêndice B. Vale lembrar que consideraremos os casos onde o vetor de médias é conhecido ($\mu = (0,0)'$) e quando é desconhecido.

Tanto na avaliação do tamanho dos testes quanto do poder utilizaremos a seguinte hipótese:

$$\begin{cases} H_0 : \Gamma = \Gamma_0 = \begin{pmatrix} 1,3333 & 0,7692 \\ 0,7692 & 1,9608 \end{pmatrix} \\ H_1 : \Gamma \neq \Gamma_0 \end{cases}$$

onde para a obtenção do modelo postulado sob H_0 , consideraremos um processo VAR(1) com os seguintes parâmetros:

$$\Phi = \begin{pmatrix} 0,5 & 0,0 \\ 0,0 & 0,7 \end{pmatrix}, \quad \Sigma_\varepsilon = \begin{pmatrix} 1,0 & 0,5 \\ 0,5 & 1,0 \end{pmatrix} \text{ e consequentemente } \Gamma_0 = \begin{pmatrix} 1,3333 & 0,7692 \\ 0,7692 & 1,9608 \end{pmatrix}.$$

Para a matriz Γ_0 temos as seguintes informações: $|\Gamma_0| = 2,0226$; $tr(\Gamma_0) = 3,2941$; autovalores: $\lambda_1 = 2,4777$, $\lambda_2 = 0,8163$, sendo a correlação entre as variáveis é igual a 0,475. Observe que como nessa primeira situação abordaremos o caso onde as duas séries não são correlacionadas ($\phi_{12} = \phi_{21} = 0$) a diagonal secundária da matriz Φ é nula.

A seleção dos testes para os quais avaliaremos o poder foi feita através da estimação do tamanho do teste. Para isso, usando o algoritmo descrito no Capítulo 5 para a situação 1, simulamos um modelo VAR(1), sob H_0 , e determinamos o tamanho dos testes. Os testes escolhidos para avaliação do poder foram aqueles que forneceram estimativas de α em torno de 5%. Na determinação do tamanho do teste consideramos amostras de tamanho 25, 50, 100, 200, 1000 e 5000. Os tamanhos amostrais, $n=1000$ e $n=5000$, foram considerados nesse caso com o objetivo de verificarmos a convergência de estimativas do tamanho dos testes.

Como visto no Capítulo 3, no caso do teste da razão de verossimilhança, considerando observações independentes, a distribuição assintótica da estatística de teste é qui-quadrado com $p(p+1)/2$ graus de liberdade tanto para os casos em que μ é conhecido quanto desconhecido. Como vamos tratar o processo bivariado teremos uma qui-quadrado com 3 graus de liberdade para os dois casos.

Já no caso em que consideramos autocorrelação, ou seja, usamos o teste como explicado na seção 4.6.2 do Capítulo 4, a distribuição assintótica da estatística de teste segue uma distribuição qui-quadrado com 5 graus de liberdade se μ é conhecido. Isso

porque, quando estamos no espaço paramétrico restrito por H_0 , não estimamos nenhum parâmetro e sob o espaço paramétrico irrestrito estimamos os 5 parâmetros do modelo VAR(1). Note que nessa situação 1 estamos considerando o caso em que $\phi_{12} = \phi_{21} = 0$, dessa forma existem somente 7 parâmetros a serem estimados, incluindo o vetor de médias. Finalmente, quando consideramos o vetor de médias desconhecido, utilizamos a distribuição assintótica qui-quadrado com 5 graus de liberdade. Porém, um estudo da distribuição exata do teste da razão de verossimilhança, considerando $\hat{\Gamma}_0$ para o caso em que o vetor de médias é desconhecido, nos mostrou que quando considerávamos 7 graus de liberdade na distribuição aproximada usando a correção $\alpha/2$ obtínhamos valores bem próximos da distribuição exata. Diante dessa situação também optamos por avaliar o tamanho do teste, para o caso em que estamos trabalhando com $\hat{\Gamma}_0$, com vetor de médias desconhecido, considerando uma distribuição qui-quadrado com 7 graus de liberdade, mas com um nível de significância nominal de $\alpha/2$. O valor estipulado de α para todos os testes estatísticos foi 0,05.

Para o teste de Sullivan et al. (2007), descrito no Capítulo 3, seção 3.6, caso em que assumimos independência entre as observações, a distribuição assintótica da estatística de teste é qui-quadrado com 3 graus de liberdade quando μ é conhecido, visto que na hipótese nula temos três parâmetros de comparação (σ_1, σ_2, ρ). Vale lembrar que estamos considerando o número de graus de liberdade para um processo bivariado. Para o caso em que o vetor de médias é desconhecido temos duas possibilidades de graus de liberdade a usar. Se pensarmos na idéia apresentada originalmente no artigo de Sullivan et al. (2007) o correto seria usar 3 graus de liberdade, porém, podemos considerar 5 graus de liberdade se pensarmos também na estimação das médias, já que para obter as estimativas de σ_1 e σ_2 é necessário estimar as médias (μ_1 e μ_2). Como estamos em um estudo exploratório para avaliarmos qual é a melhor aproximação para a distribuição da estatística de teste no caso autocorrelacionado, optamos por testar as duas opções de graus de liberdade e verificar o que ocorre com o tamanho do teste.

Quando trabalhamos com o teste de Sullivan et al. (2007), considerando autocorrelação, como descrito no Capítulo 4, seção 4.6.3, com o vetor de médias conhecido, usamos a aproximação para a distribuição qui-quadrado com 5 graus de liberdade. Para o caso em que consideramos $\hat{\Gamma}_0$, com vetor de médias desconhecido, podemos usar 5 graus de liberdade, se seguirmos a idéia original do artigo de Sullivan et al. (2007), ou 7 graus de liberdade se pensarmos na estimação do vetor de médias. Dessa forma, optamos por usar as duas propostas. Da mesma maneira que verificamos

no teste da razão de verossimilhança testamos uma correção para a distribuição qui-quadrado assintótica, considerando $\alpha/2$, quando usamos $\hat{\Gamma}_0$. O mesmo foi feito no teste de Sullivan et al. (2007). Porém, no caso desse teste, obtivemos boas aproximações, tanto para o caso em que o vetor de médias era conhecido quanto para o caso em que era desconhecido. Dessa forma, também consideraremos essa situação na avaliação do tamanho do teste.

No teste do determinante, tanto quando utiliza-se a matriz quando se usa a matriz $\hat{\Gamma}_0$, e considerando os casos onde o vetor de médias era conhecido e desconhecido, utilizamos o teste exato, o teste tradicional, que usa a distribuição normal como aproximação para a distribuição do determinante de S e $\hat{\Gamma}_0$, e a modificação proposta por Djauhari (2005), que também usa a aproximação normal para os determinantes de S e $\hat{\Gamma}_0$. Finalmente para o teste VMAX, proposto por Machado et al. (2008), descritos nos Capítulos 3 e 4, consideramos somente a distribuição exata. A Tabela 6.1 apresenta sumariamente os testes para os quais avaliaremos o seu tamanho especificando o número de graus de liberdade da distribuição assintótica de cada um deles.

Tabela 6.1: Testes abordados na situação 1

Testes			
<i>μ conhecido</i>		<i>μ desconhecido</i>	
S	$\hat{\Gamma}_0$	S	$\hat{\Gamma}_0$
Determinante Exato	Determinante - Exato	Determinante Exato	Determinante - Exato
Determinante Trad.	Determinante - Trad.	Determinante Trad.	Determinante - Trad.
Determinante Djauhari	Determinante - Djauhari	Determinante Djauhari	Determinante - Djauhari
Razão de V. - Exato	Razão de V. - Exato	Razão de V. - Exato	Razão de V. - Exato
Razão de V. - $X^2_{(3q,l,\alpha)}$	Razão de V. - $X^2_{(5q,l,\alpha)}$	Razão de V. - $X^2_{(3q,l,\alpha)}$	Razão de V. - $X^2_{(5q,l,\alpha)}$
Sullivan - Exato	Razão de V. - $X^2_{(5q,l,\alpha/2)}$	Sullivan - Exato	Razão de V. - $X^2_{(7q,l,\alpha)}$
Sullivan - $X^2_{(3q,l,\alpha)}$	Sullivan - Exato	Sullivan - $X^2_{(3q,l,\alpha)}$	Razão de V. - $X^2_{(7q,l,\alpha/2)}$
VMAX - Exato	Sullivan - $X^2_{(5q,l,\alpha)}$	Sullivan - $X^2_{(5q,l,\alpha)}$	Sullivan - Exato
	Sullivan - $X^2_{(7q,l,\alpha)}$	VMAX - Exato	Sullivan - $X^2_{(5q,l,\alpha)}$
	Sullivan - $X^2_{(7q,l,\alpha/2)}$		Sullivan - $X^2_{(7q,l,\alpha)}$
	VMAX - Exato		Sullivan - $X^2_{(7q,l,\alpha/2)}$
			VMAX - Exato

Note que todos os graus de liberdade das distribuições assintóticas foram determinados com base na situação 1, ou seja, temos um modelo bivariado VAR(1) onde $\phi_{12} = \phi_{21} = 0$. Na próxima seção apresentaremos as estimativas do tamanho dos testes apresentados na Tabela 6.1.

6.1.1 Estimação do tamanho dos testes na situação 1

Os resultados obtidos via simulação de Monte Carlo para o tamanho dos testes são apresentados nas Tabelas 6.2 a 6.13. Observe que as Tabelas estão discriminadas para os casos em que o vetor de médias é conhecido e para aqueles em que é desconhecido. Além disso, em cada uma das Tabelas fornecemos os valores críticos das distribuições assintóticas e exatas (considerando $\alpha = 0,05$) e a estimativa do tamanho do teste.

Tabela 6.2: Limites críticos e estimativas do tamanho do teste na situação 1: $n=25 - \mu$ conhecido

Teste	$n = 25 - \mu$ conhecido			
	S_{μ}		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(0,498; 4,874)	0,0501	(0,501; 5,944)	0,0489
Determinante - Trad.	(0,338; 3,539)	0,0860	(0,338; 3,539)	0,1272
Determinante - Djauhari	(0,483; 3,562)	0,1030	(0,483; 3,562)	0,1417
Razão de V. - Exato	15,878	0,0490	13,820	0,0498
Razão de V. - $X^2_{(3q,l,\alpha)}$	7,815	0,2811	-	-
Razão de V. - $X^2_{(5q,l,\alpha)}$	-	-	11,071	0,0944
Razão de V. - $X^2_{(5q,l,\alpha/2)}$	-	-	12,833	0,0628
Sullivan - Exato	16,095	0,0493	15,086	0,0504
Sullivan - $X^2_{(3q,l,\alpha)}$	7,815	0,2568	-	-
Sullivan - $X^2_{(5q,l,\alpha)}$	-	-	11,070	0,1088
Sullivan - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0606
Sullivan - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0430
VMAX - Exato	2,008	0,0482	2,022	0,0493

Tabela 6.3: Limites críticos e estimativas do tamanho do teste na situação 1: $n=25 - \mu$ desconhecido

Teste	$n = 25 - \mu$ desconhecido			
	S_n		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(0,387; 3,748)	0,0495	(0,355; 4,185)	0,0483
Determinante - Trad.	(0,338; 3,539)	0,0458	(0,338; 3,539)	0,0623
Determinante - Djauhari	(0,483; 3,562)	0,0867	(0,483; 3,562)	0,1117
Razão de V. - Exato	16,523	0,0505	15,222	0,0500
Razão de V. - $X^2_{(3q,l,\alpha)}$	7,815	0,3510	-	-
Razão de V. - $X^2_{(5q,l,\alpha)}$	-	-	11,070	0,1373
Razão de V. - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0673
Razão de V. - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0409
Sullivan - Exato	15,109	0,0503	14,122	0,0495
Sullivan - $X^2_{(3q,l,\alpha)}$	7,815	0,2983	-	-
Sullivan - $X^2_{(5q,l,\alpha)}$	11,070	0,1375	11,070	0,1137
Sullivan - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0505
Sullivan - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0294
VMAX - Exato	1,742	0,0480	1,699	0,0488

Tabela 6.4: Limites críticos e estimativas do tamanho do teste na situação 1: n=50 - μ conhecido

Teste	n = 50 - μ conhecido			
	S_{μ}		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(0,775; 4,027)	0,0506	(0,772; 4,520)	0,0501
Determinante - Trad.	(0,855; 3,108)	0,1349	(0,855; 3,108)	0,1673
Determinante - Djauhari	(0,918; 3,127)	0,1488	(0,918; 3,127)	0,1812
Razão de V. - Exato	16,678	0,0503	14,650	0,0506
Razão de V. - X^2 (3q,l,a)	7,815	0,2975	-	-
Razão de V. - X^2 (5q,l,a)	-	-	11,071	0,1092
Razão de V. - X^2 (5q,l,a/2)	-	-	12,833	0,0744
Sullivan - Exato	16,710	0,0505	15,629	0,0510
Sullivan - X^2 (3q,l,a)	7,815	0,2850	-	-
Sullivan - X^2 (5q,l,a)	-	-	11,070	0,1211
Sullivan - X^2 (7q,l,a)	-	-	14,067	0,0672
Sullivan - X^2 (7q,l,a/2)	-	-	16,013	0,0475
VMAX - Exato	1,689	0,0497	1,692	0,0496

Tabela 6.5: Limites críticos e estimativas do tamanho do teste na situação 1: n=50 - μ desconhecido

Teste	n = 50 - μ desconhecido			
	S_n		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(0,681; 3,542)	0,0497	(0,652; 3,802)	0,0504
Determinante - Trad.	(0,855; 3,108)	0,1272	(0,855; 3,108)	0,1530
Determinante - Djauhari	(0,918; 3,127)	0,1533	(0,918; 3,127)	0,1820
Razão de V. - Exato	17,153	0,0496	15,585	0,0504
Razão de V. - X^2 (3q,l,a)	7,815	0,3350	-	-
Razão de V. - X^2 (5q,l,a)	-	-	11,070	0,1334
Razão de V. - X^2 (7q,l,a)	-	-	14,067	0,0696
Razão de V. - X^2 (7q,l,a/2)	-	-	16,013	0,0456
Sullivan - Exato	16,330	0,0485	15,193	0,0486
Sullivan - X^2 (3q,l,a)	7,815	0,3085	-	-
Sullivan - X^2 (5q,l,a)	11,070	0,1564	11,070	0,1248
Sullivan - X^2 (7q,l,a)	-	-	14,067	0,0624
Sullivan - X^2 (7q,l,a/2)	-	-	16,013	0,0407
VMAX - Exato	1,569	0,0493	1,546	0,0491

Tabela 6.6: Limites críticos e estimativas do tamanho do teste na situação 1: n=100 - μ conhecido

Teste	n = 100 - μ conhecido			
	S_{μ}		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(1,046; 3,413)	0,0502	(1,029; 3,689)	0,0480
Determinante - Trad.	(1,207; 2,797)	0,1666	(1,207; 2,797)	0,1986
Determinante - Djauhari	(1,236; 2,810)	0,1745	(1,236; 2,810)	0,2066
Razão de V. - Exato	17,093	0,0487	15,087	0,0481
Razão de V. - X^2 (3q,l,a)	7,815	0,3080	-	-
Razão de V. - X^2 (5q,l,a)	-	-	11,071	0,1142
Razão de V. - X^2 (5q,l,a/2)	-	-	12,833	0,0800
Sullivan - Exato	17,185	0,0496	16,086	0,0491
Sullivan - X^2 (3q,l,a)	7,815	0,2969	-	-
Sullivan - X^2 (5q,l,a)	-	-	11,070	0,1273
Sullivan - X^2 (7q,l,a)	-	-	14,067	0,0714
Sullivan - X^2 (7q,l,a/2)	-	-	16,013	0,0498
VMAX - Exato	1,471	0,0506	1,469	0,0515

Tabela 6.7: Limites críticos e estimativas do tamanho do teste na situação 1: n=100 - μ desconhecido

Teste	n = 100 - μ desconhecido			
	S_n		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(0,974; 3,206)	0,0486	(0,942; 3,364)	0,0478
Determinante - Trad.	(1,207; 2,797)	0,1687	(1,207; 2,797)	0,1964
Determinante - Djauhari	(1,236; 2,810)	0,1809	(1,236; 2,810)	0,2096
Razão de V. - Exato	17,456	0,0507	15,656	0,0507
Razão de V. - X^2 (3q,l,a)	7,815	0,3270	-	-
Razão de V. - X^2 (5q,l,a)	-	-	11,070	0,1300
Razão de V. - X^2 (7q,l,a)	-	-	14,067	0,0700
Razão de V. - X^2 (7q,l,a/2)	-	-	16,013	0,0470
Sullivan - Exato	17,048	0,0493	15,821	0,0496
Sullivan - X^2 (3q,l,a)	7,815	0,3101	-	-
Sullivan - X^2 (5q,l,a)	11,070	0,1634	11,070	0,1300
Sullivan - X^2 (7q,l,a)	-	-	14,067	0,0709
Sullivan - X^2 (7q,l,a/2)	-	-	16,013	0,0478
VMAX - Exato	1,418	0,0514	1,405	0,0508

Tabela 6.8: Limites críticos e estimativas do tamanho do teste na situação 1: n=200 - μ conhecido

Teste	n = 200 - μ conhecido			
	S_μ		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(1,279; 2,967)	0,0509	(1,266; 3,111)	0,0515
Determinante - Trad.	(1,451; 2,574)	0,1858	(1,451; 2,574)	0,2160
Determinante - Djauhari	(1,464; 2,581)	0,1894	(1,464; 2,581)	0,2202
Razão de V. - Exato	17,414	0,0495	15,359	0,0482
Razão de V. - X^2 (3q,l,a)	7,815	0,3060	-	-
Razão de V. - X^2 (5q,l,a)	-	-	11,071	0,1140
Razão de V. - X^2 (5q,l,a/2)	-	-	12,833	0,0785
Sullivan - Exato	17,440	0,0489	16,389	0,0502
Sullivan - X^2 (3q,l,a)	7,815	0,3044	-	-
Sullivan - X^2 (5q,l,a)	-	-	11,070	0,1326
Sullivan - X^2 (7q,l,a)	-	-	14,067	0,0758
Sullivan - X^2 (7q,l,a/2)	-	-	16,013	0,0531
VMAX - Exato	1,326	0,0490	1,326	0,0486

Tabela 6.9: Limites críticos e estimativas do tamanho do teste na situação 1: n=200 - μ desconhecido

Teste	n = 200 - μ desconhecido			
	S_n		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(1,238; 2,886)	0,0504	(1,210; 2,979)	0,0508
Determinante - Trad.	(1,451; 2,574)	0,1852	(1,451; 2,574)	0,2161
Determinante - Djauhari	(1,464; 2,581)	0,1912	(1,464; 2,581)	0,2222
Razão de V. - Exato	17,606	0,0490	15,642	0,0499
Razão de V. - X^2 (3q,l,a)	7,815	0,3170	-	-
Razão de V. - X^2 (5q,l,a)	-	-	11,070	0,1240
Razão de V. - X^2 (7q,l,a)	-	-	14,067	0,0681
Razão de V. - X^2 (7q,l,a/2)	-	-	16,013	0,0463
Sullivan - Exato	17,388	0,0489	16,271	0,0496
Sullivan - X^2 (3q,l,a)	7,815	0,3122	-	-
Sullivan - X^2 (5q,l,a)	11,070	0,1689	11,070	0,1347
Sullivan - X^2 (7q,l,a)	-	-	14,067	0,0754
Sullivan - X^2 (7q,l,a/2)	-	-	16,013	0,0523
VMAX - Exato	1,301	0,0483	1,294	0,0485

Tabela 6.10: Limites críticos e estimativas do tamanho do teste na situação 1: n=1000 - μ conhecido

Teste	n = 1000 - μ conhecido			
	S_{μ}		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(1,666; 2,432)	0,0502	(1,647; 2,468)	0,0500
Determinante - Trad.	(1,770; 2,271)	0,1978	(1,770; 2,271)	0,2294
Determinante - Djahari	(1,772; 2,273)	0,1984	(1,772; 2,273)	0,2304
Razão de V. - Exato	17,957	0,0490	15,570	0,0516
Razão de V. - X^2 (3q,l,a)	7,815	0,3061	-	-
Razão de V. - X^2 (5q,l,a)	-	-	11,071	0,1199
Razão de V. - X^2 (5q,l,a/2)	-	-	12,833	0,0862
Sullivan - Exato	17,741	0,0491	16,666	0,0497
Sullivan - X^2 (3q,l,a)	7,815	0,3131	-	-
Sullivan - X^2 (5q,l,a)	-	-	11,070	0,1379
Sullivan - X^2 (7q,l,a)	-	-	14,067	0,0796
Sullivan - X^2 (7q,l,a/2)	-	-	16,013	0,0560
VMAX - Exato	1,139	0,0497	1,139	0,0499

Tabela 6.11: Limites críticos e estimativas do tamanho do teste na situação 1: n=1000- μ desconhec

Teste	n = 1000 - μ desconhecido			
	S_n		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(1,652; 2,412)	0,0502	(1,632; 2,448)	0,0499
Determinante - Trad.	(1,770; 2,271)	0,1990	(1,770; 2,271)	0,2295
Determinante - Djahari	(1,772; 2,273)	0,2002	(1,772; 2,273)	0,2305
Razão de V. - Exato	17,553	0,0516	15,537	0,0504
Razão de V. - X^2 (3q,l,a)	7,815	0,3134	-	-
Razão de V. - X^2 (5q,l,a)	-	-	11,070	0,1222
Razão de V. - X^2 (7q,l,a)	-	-	14,067	0,0672
Razão de V. - X^2 (7q,l,a/2)	-	-	16,013	0,0460
Sullivan - Exato	17,741	0,0488	16,634	0,0503
Sullivan - X^2 (3q,l,a)	7,815	0,3158	-	-
Sullivan - X^2 (5q,l,a)	11,070	0,1712	11,070	0,1382
Sullivan - X^2 (7q,l,a)	-	-	14,067	0,0790
Sullivan - X^2 (7q,l,a/2)	-	-	16,013	0,0561
VMAX - Exato	1,134	0,0505	1,133	0,0503

Tabela 6.12: Limites críticos e estimativas do tamanho do teste na situação 1: n=5000 - μ conhecido

Teste	n = 5000 - μ conhecido			
	S_{μ}		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(1,857; 2,201)	0,0508	(1,846; 2,215)	0,0498
Determinante - Trad.	(1,910; 2,134)	0,2010	(1,910; 2,134)	0,2305
Determinante - Djahari	(1,911; 2,135)	0,2012	(1,911; 2,135)	0,2307
Razão de V. - Exato	17,545	0,0560	15,675	0,0531
Razão de V. - X^2 (3q,l,a)	7,815	0,3169	-	-
Razão de V. - X^2 (5q,l,a)	-	-	11,071	0,1270
Razão de V. - X^2 (5q,l,a/2)	-	-	12,833	0,0907
Sullivan - Exato	17,743	0,0499	16,721	0,0500
Sullivan - X^2 (3q,l,a)	7,815	0,3138	-	-
Sullivan - X^2 (5q,l,a)	-	-	11,070	0,1390
Sullivan - X^2 (7q,l,a)	-	-	14,067	0,0802
Sullivan - X^2 (7q,l,a/2)	-	-	16,013	0,0564
VMAX - Exato	1,061	0,0485	1,061	0,0486

Tabela 6.13: Limites críticos e estimativas do tamanho do teste na situação 1: $n=5000$ - μ desconhecido

Teste	$n = 5000 - \mu$ desconhecido			
	S_n		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(1,854; 2,197)	0,0507	(1,843; 2,211)	0,0501
Determinante - Trad.	(1,910; 2,134)	0,2019	(1,910; 2,134)	0,2311
Determinante - Djauhari	(1,911; 2,135)	0,2021	(1,911; 2,135)	0,2314
Razão de V. - Exato	17,672	0,0522	15,598	0,0515
Razão de V. - $X^2_{(3q,l,\alpha)}$	7,815	0,3140	-	-
Razão de V. - $X^2_{(5q,l,\alpha)}$	-	-	11,070	0,1232
Razão de V. - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0689
Razão de V. - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0479
Sullivan - Exato	17,734	0,0500	16,722	0,0498
Sullivan - $X^2_{(3q,l,\alpha)}$	7,815	0,3144	-	-
Sullivan - $X^2_{(5q,l,\alpha)}$	11,070	0,1741	11,070	0,1389
Sullivan - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0799
Sullivan - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0564
VMAX - Exato	1,060	0,0493	1,060	0,0487

Diante dos resultados obtidos via simulação, selecionamos para avaliação do poder aqueles testes que forneceram uma estimativa de α em torno de 5%. Observe que para o teste proposto por Sullivan et al. (2007), quando usamos a matriz $\hat{\Gamma}_0$, tanto para o caso em que o vetor de médias é conhecido quanto desconhecido, obtivemos bons resultados do tamanho do teste na aproximação para a distribuição qui-quadrado levando em consideração a correção com $\alpha/2$. Esse resultado é de grande importância, visto que na prática seria muito trabalhoso utilizar a distribuição exata, e essa aproximação, pode ser de grande utilidade. Essa correção também é útil quando consideramos o teste da razão de verossimilhança. Note que ao aumentarmos o tamanho amostral notamos certa estabilidade das estimativas dos tamanhos dos testes. Quando trabalhamos com a matriz de covariâncias amostral S , não obtivemos nenhum sucesso nas aproximações para a distribuição qui-quadrado. Dessa forma, somente as distribuições exatas, para o caso em que trabalhamos com a matriz S , serão consideradas na avaliação do poder. No teste do determinante, tanto para S quanto para $\hat{\Gamma}_0$, somente os resultados obtidos com a distribuição exata foram satisfatórios. Pelos resultados obtidos os testes do determinante tradicional e da correção proposta por Djauhari (2005) não tiveram bom desempenho. Finalmente para o teste VMAX avaliaremos o poder considerando somente as distribuições exatas. Lembrando que para a obtenção dos limites críticos do teste VMAX utilizamos simulação, ao invés de integração numérica.

Considere uma situação prática onde estamos trabalhando com um processo onde existe autocorrelação entre as observações, mas que por desconhecimento essa autocorrelação seja desconsiderada, ou seja, assumimos independência entre as observações e usamos a matriz de covariância amostral S para verificar possíveis

alterações na matriz postulada sob H_0 , considerando o nível de significância nominal $\alpha = 0,05$. Dessa forma, se usarmos os testes para matrizes de covariâncias, que foram apresentados no Capítulo 3, assumindo independência entre as observações e aproximação para a distribuição qui-quadrado, teremos conclusões incorretas em vista do fato que esses testes têm um tamanho acima do α nominal estipulado, mesmo para amostras maiores. Como vimos no Capítulo 3, quando as observações são independentes é coerente usar a aproximação das estatísticas de teste para a distribuição qui-quadrado, o que não acontece no caso autocorrelacionado quando usamos a matriz de covariância amostral S . Quando estamos em um processo autocorrelacionado nenhuma aproximação das estatísticas de teste, quando usamos a matriz de covariância amostral S , para a distribuição qui-quadrado, foi satisfatória, como pode ser visto pelas Tabelas 6.2 a 6.13. Assim, se em um processo autocorrelacionado ignorarmos esse fato, e usarmos a matriz de covariância amostral S , juntamente com os testes propostos no Capítulo 3, assumindo uma aproximação para a distribuição qui-quadrado estaremos cometendo um grave engano. O mesmo vale para os testes do determinante tradicional e Djauhari (2005) no qual a aproximação normal para a distribuição da estatística de teste não funciona bem quando se usa a matriz S nem quando se usa a matriz $\hat{\Gamma}_0$. A Tabela 6.14 apresenta resumidamente os testes que foram selecionados para a fase de determinação do poder.

Tabela 6.14: Testes selecionados para avaliação do poder na situação 1

Testes	Média conhecida		Média desconhecida	
	S_μ	$\hat{\Gamma}_0$	S_n	$\hat{\Gamma}_0$
Determinante	Exato	Exato	Exato	Exato
VMAX	Exato	Exato	Exato	Exato
Razão de Verossim.	Exato	Exato	Exato	Exato $\chi^2_{7\text{gl.},\alpha}$ $\chi^2_{7\text{gl.},\alpha/2}$
Step-Down (Sullivan et. al)	Exato	Exato $\chi^2_{7\text{gl.},\alpha}$ $\chi^2_{7\text{gl.},\alpha/2}$	Exato	Exato $\chi^2_{7\text{gl.},\alpha}$ $\chi^2_{7\text{gl.},\alpha/2}$

6.1.2 Estimação do poder dos testes na situação 1

Na avaliação do poder dos testes, descritos na Tabela 6.14, optamos por trabalhar com amostras de tamanho 25, 50, 100 e 200, visto que na prática é menos usual usar tamanhos amostrais superiores a esses valores. Para avaliação do poder dos testes simulamos n observações, sob H_1 , 20.000 vezes e determinamos o poder de cada teste. Esse processo foi repetido 25 vezes para cada teste, dessa forma, temos 25 estimativas de poder para cada teste. Isso foi feito com o intuito de termos uma maior confiabilidade nos resultados. Na avaliação de poder dos testes utilizamos as seguintes hipóteses nula e alternativa:

$$\begin{cases} H_0 : \Gamma = \Gamma_0 = \begin{pmatrix} 1,3333 & 0,7692 \\ 0,7692 & 1,9608 \end{pmatrix} \\ H_1 : \Gamma \neq \Gamma_0 \end{cases}$$

Para o estudo do poder dos testes consideramos 5 modelos distintos para avaliar a capacidade dos testes em detectar alterações na matriz de covariâncias. Em alguns cenários foram feitas pequenas mudanças na estrutura da matriz de covariâncias, para verificar se os testes eram sensíveis a pequenos deslocamentos em relação a matriz estipulada em H_0 . Em outros cenários alteramos somente a covariância entre as variáveis para ver se os testes conseguiam perceber tal alteração. De uma forma geral, tentamos explorar várias mudanças possíveis na matriz de covariâncias. A Tabela 6.15 apresenta os 5 modelos para os quais avaliamos o poder dos testes, com seus respectivos parâmetros. Algumas medidas como razão entre determinante e traços das matrizes de covariâncias sob H_1 e H_0 são apresentadas. Vale lembrar que todos os modelos apresentados correspondem a um processo VAR(1). O valor médio das 25 estimativas de poder de cada teste estão apresentados nas Tabelas 6.16 a 6.20.

Ao formularmos o modelo 1, nosso objetivo era verificar se os testes em questão eram capazes de detectar uma mudança na covariância das variáveis. Observe que as variâncias se mantiveram constantes, somente houve uma diminuição na covariância. Na Tabela 6.15 fornecemos algumas medidas importantes como $\Lambda_{\det} = 1,1872$ e $\Lambda_{tr} = 1,000$. Esses resultados já nos davam indícios de que o teste do determinante e o teste VMAX teriam dificuldades para captar a alteração na matriz de covariâncias. Observe pela Tabela 6.16 que o poder do teste VMAX, para qualquer tamanho amostral, fica bem próximo de 0,05. Isso já era esperado visto que o teste VMAX se baseia somente nas variâncias amostrais e como a alteração se deu somente na covariância o teste não foi capaz de captar tal alteração. Além disso, observe que a razão dos determinantes, da

matriz postulada sob a hipótese alternativa e sob a hipótese nula, forneceu um valor próximo de um, indicando que os testes fundamentados no determinante teriam dificuldades em detectar tal alteração. Pela Tabela 6.16 podemos observar que mesmo para tamanhos amostrais grandes ($n=200$) o poder do teste do determinante é bem baixo quando comparado com os demais.

Ainda no contexto do modelo 1, ao compararmos o desempenho dos testes exatos notamos uma superioridade do teste de Sullivan et al. (2007). Em algumas situações o poder desse teste é praticamente o dobro daquele obtido pelo teste da razão de verossimilhança. Observe que somente para $n=100$ o teste da razão de verossimilhança consegue atingir o poder que o teste de Sullivan et al. (2007) possui com $n=25$. Esse fato mostra a superioridade do teste de Sullivan et al. (2007) em relação aos demais. De uma forma geral, no modelo 1, considerando os testes exatos, podemos destacar o bom desempenho do teste de Sullivan et al. (2007) e do teste da razão de verossimilhança, ressaltando que o desempenho desses testes melhorou à medida que aumentamos o tamanho amostral.

Mudando o foco da análise para os testes em que consideramos as distribuições assintóticas, observe que tanto no teste de Sullivan et al. (2007) quanto no teste da razão de verossimilhança, a aproximação para a distribuição qui-quadrado com 7 graus de liberdade, considerando a correção $\alpha/2$, forneceu valores bem próximos daqueles obtidos pela distribuição exata. Esse fato é de grande importância visto o ganho que se tem na prática ao se trabalhar com uma distribuição aproximada ao invés da exata, já que para a obtenção da distribuição exata da estatística de teste é necessário utilizar-se o recurso da simulação o que demanda tempo computacional.

É importante observar que mesmo para $n=200$ não foi possível obter um poder próximo de 1. O valor de poder neste caso está entre 0,54 e 0,60. Sendo assim, seria necessário um tamanho maior que 200 caso fosse de interesse detectar alguma mudança similar à do modelo 1.

Um ponto a ser observado é que quando estamos no teste da razão de verossimilhança e usamos a matriz $\hat{\Gamma}_0$, o poder do teste é superior ao que obtemos quando usamos a matriz de covariância amostral S . Ou seja, se levarmos em consideração a autocorrelação existente nos dados e conseqüentemente usarmos a matriz $\hat{\Gamma}_0$, teremos resultados mais satisfatórios. No caso do teste de Sullivan et al. (2007) nem sempre acontece a superioridade quando consideramos $\hat{\Gamma}_0$, ou seja, em

alguns momentos o uso da matriz S forneceu valores de poder superiores. Porém, uma desvantagem do uso da matriz S é que sempre teremos que trabalhar com a distribuição exata, o que não acontece nas situações em que usamos a matriz $\hat{\Gamma}_0$. Isso porque, verificamos que existe uma aproximação para distribuição qui-quadrado, considerando $\alpha/2$, quando estamos trabalhando com a matriz $\hat{\Gamma}_0$.

Tabela 6.15: Modelos para avaliação do poder dos testes na situação 1

Modelos	Parâmetros e Medidas
<i>Modelo sob H₀</i>	Sob H ₀ : $\Gamma_0 = \begin{pmatrix} 1,3333 & 0,7692 \\ 0,7692 & 1,9608 \end{pmatrix}$ onde $\Phi = \begin{pmatrix} 0,5 & 0,0 \\ 0,0 & 0,7 \end{pmatrix}$, $\Sigma_\varepsilon = \begin{pmatrix} 1,0 & 0,5 \\ 0,5 & 1,0 \end{pmatrix}$ $ \Gamma_0 = 2,0226$, $tr(\Gamma_0) = 3,2941$ e $\rho = 0,476$
<i>Modelo 1</i>	Sob H ₁ : $\Gamma = \begin{pmatrix} 1,3333 & 0,4615 \\ 0,4615 & 1,9608 \end{pmatrix}$ onde $\Phi = \begin{pmatrix} 0,5 & 0,0 \\ 0,0 & 0,7 \end{pmatrix}$, $\Sigma_\varepsilon = \begin{pmatrix} 1,0 & 0,3 \\ 0,3 & 1,0 \end{pmatrix}$ $\Lambda_{det} = \frac{ \Gamma }{ \Gamma_0 } = \frac{2,4014}{2,0226} = 1,1872$ e $\Lambda_{tr} = \frac{tr(\Gamma)}{tr(\Gamma_0)} = 1,0000$ e $\rho = 0,29$
<i>Modelo 2</i>	Sob H ₁ : $\Gamma = \begin{pmatrix} 2,7778 & 1,8182 \\ 1,8182 & 1,9608 \end{pmatrix}$ onde $\Phi = \begin{pmatrix} 0,8 & 0,0 \\ 0,0 & 0,7 \end{pmatrix}$, $\Sigma_\varepsilon = \begin{pmatrix} 1,0 & 0,8 \\ 0,8 & 1,0 \end{pmatrix}$, $\Lambda_{det} = \frac{ \Gamma }{ \Gamma_0 } = \frac{2,1409}{2,0226} = 1,0585$ e $\Lambda_{tr} = \frac{tr(\Gamma)}{tr(\Gamma_0)} = 1,4385$ e $\rho = 0,78$
<i>Modelo 3</i>	Sob H ₁ : $\Gamma = \begin{pmatrix} 2,6667 & 0,4615 \\ 0,4615 & 2,9412 \end{pmatrix}$ onde $\Phi = \begin{pmatrix} 0,5 & 0,0 \\ 0,0 & 0,7 \end{pmatrix}$, $\Sigma_\varepsilon = \begin{pmatrix} 2,0 & 0,3 \\ 0,3 & 1,5 \end{pmatrix}$, $\Lambda_{det} = \frac{ \Gamma }{ \Gamma_0 } = \frac{7,6303}{2,0226} = 3,7725$ e $\Lambda_{tr} = \frac{tr(\Gamma)}{tr(\Gamma_0)} = 1,7024$ e $\rho = 0,16$
<i>Modelo 4</i>	Sob H ₁ : $\Gamma = \begin{pmatrix} 5,3333 & 1,5385 \\ 1,5385 & 1,9608 \end{pmatrix}$ onde $\Phi = \begin{pmatrix} 0,5 & 0,0 \\ 0,0 & 0,7 \end{pmatrix}$, $\Sigma_\varepsilon = \begin{pmatrix} 4,0 & 1,0 \\ 1,0 & 1,0 \end{pmatrix}$, $\Lambda_{det} = \frac{ \Gamma }{ \Gamma_0 } = \frac{8,0906}{2,0226} = 3,9999$ e $\Lambda_{tr} = \frac{tr(\Gamma)}{tr(\Gamma_0)} = 2,2143$ e $\rho = 0,48$
<i>Modelo 5</i>	Sob H ₁ : $\Gamma = \begin{pmatrix} 2,7778 & 1,5909 \\ 1,5909 & 1,9608 \end{pmatrix}$ onde $\Phi = \begin{pmatrix} 0,8 & 0,0 \\ 0,0 & 0,7 \end{pmatrix}$, $\Sigma_u = \begin{pmatrix} 1,0 & 0,7 \\ 0,7 & 1,0 \end{pmatrix}$, $\Lambda_{det} = \frac{ \Gamma }{ \Gamma_0 } = \frac{2,9157}{2,0226} = 1,4415$ e $\Lambda_{tr} = \frac{tr(\Gamma)}{tr(\Gamma_0)} = 1,4385$ e $\rho = 0,68$

Nota: ρ é o coeficiente de correlação entre as duas variáveis X_1 e X_2 .

Tabela 6.16: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 1 na situação 1

n	μ	Estimador	Testes							
			Razão V. Exato	Razão V. $X^2_{(7g,l,\alpha)}$	Razão V. $X^2_{(7g,l,\alpha/2)}$	Sullivan Exato	Sullivan $X^2_{(7g,l,\alpha)}$	Sullivan $X^2_{(7g,l,\alpha/2)}$	VMAX Exato	Determin. Exato
25	$\mu_{conhecido}$	S_μ	0,1020	-	-	0,1524	-	-	0,0497	0,0594
		$\hat{\Gamma}_0$	0,1121	-	-	0,1319	0,1533	0,1156	0,0500	0,0461
	$\mu_{desconhecido}$	S_n	0,0721	-	-	0,1346	-	-	0,0502	0,0582
		$\hat{\Gamma}_0$	0,0742	0,0958	0,0620	0,1137	0,1150	0,0792	0,0504	0,0464
50	$\mu_{conhecido}$	S_μ	0,1501	-	-	0,2143	-	-	0,0510	0,0695
		$\hat{\Gamma}_0$	0,1653	-	-	0,1929	0,2360	0,1835	0,0507	0,0536
	$\mu_{desconhecido}$	S_n	0,1171	-	-	0,1912	-	-	0,0514	0,0677
		$\hat{\Gamma}_0$	0,1252	0,1609	0,1165	0,1654	0,1968	0,1457	0,0510	0,0535
100	$\mu_{conhecido}$	S_μ	0,2534	-	-	0,3285	-	-	0,0509	0,0864
		$\hat{\Gamma}_0$	0,2830	-	-	0,3279	0,4031	0,3306	0,0516	0,0641
	$\mu_{desconhecido}$	S_n	0,2182	-	-	0,3035	-	-	0,0511	0,0844
		$\hat{\Gamma}_0$	0,2415	0,2917	0,2313	0,2985	0,3655	0,2919	0,0511	0,0662
200	$\mu_{conhecido}$	S_μ	0,4698	-	-	0,5453	-	-	0,0499	0,1288
		$\hat{\Gamma}_0$	0,5144	-	-	0,6072	0,6967	0,6214	0,0499	0,0998
	$\mu_{desconhecido}$	S_n	0,4363	-	-	0,5234	-	-	0,0502	0,1238
		$\hat{\Gamma}_0$	0,4786	0,5357	0,4656	0,5824	0,6714	0,5929	0,0507	0,1001

MODELO 1: Sob H_1 : $\Gamma = \begin{pmatrix} 1,3333 & 0,4615 \\ 0,4615 & 1,9608 \end{pmatrix}$

Nos modelos 2 e 4 alteramos a variância de uma das variáveis e a covariância, sendo que no modelo 4 a mudança foi mais acentuada. Ao observarmos as medidas fornecidas na Tabela 6.15, para o modelo 4, verificamos que a razão entre os determinantes, da matriz postulada sob a hipótese alternativa e sob a hipótese nula, é extremamente grande ($\Lambda_{\text{det}}=3,9999$). Essa grande diferença já nos mostra que o desempenho do teste do determinante será melhor do que o obtido quando usamos o modelo 1. Como a alteração na estrutura da matriz de covariâncias no modelo 4 foi grande, para tamanhos amostrais pequenos ($n=25$) já conseguimos bom resultados (poderes $> 0,90$), como pode ser visto na Tabela 6.19. Observe que para $n = 100$, praticamente todos os testes já alcançam poder 1,0 (um), ou seja, para mudanças similares a tratadas no modelo 4 qualquer um dos testes seria adequado tanto com a matriz S quanto com a matriz $\hat{\Gamma}_0$, se $n \geq 50$. Já para $n=25$ o teste do determinante exato não seria indicado por ter um poder inferior aos outros testes (valor máximo 0,6756).

Comparando os testes exatos no modelo 2, notamos uma superioridade do teste da razão de verossimilhança, para tamanhos de amostras pequenas ($n=25$ e $n=50$). Para tamanhos amostrais maiores ($n > 100$) o teste de Sullivan et al. (2007) e o teste VMAX conseguem atingir valores bem próximos, ou até superiores, daqueles obtidos pelo teste da razão de verossimilhança. No modelo 2, para $n=100$, todos os testes já apresentam valores de poder altos, exceto o teste do determinante. Isso se deve ao fato de que a razão entre os determinantes das matrizes teóricas sob H_1 e H_0 é $\Lambda_{\text{det}} = 1,0585$, ou seja, a diferença do determinante das duas matrizes é bem pequena, e já era esperado a inferioridade do teste do determinante. Observe que mesmo para $n=200$, onde a maioria das estimativas do poder dos testes aproximam-se de 1 (um), o teste do determinante fornece valores de poder inferiores a 0,20.

Como no modelo 2 houve uma mudança acentuada na variância da primeira variável (é mais que o dobro da variância sob H_0) o teste VMAX teve um bom desempenho, similar aos testes da razão de verossimilhança e do proposto por Sullivan et al. (2007), perdendo muito pouco para o teste da razão de verossimilhança e ganhando algumas vezes do teste de Sullivan et al. (2007). É interessante observar o comportamento do teste VMAX neste caso. Como no modelo 4 há uma alteração brusca da variância de X_1 , e o teste VMAX leva em consideração somente alterações nas variâncias, nessa situação já era esperado seu bom desempenho. Porém, observe que a diferença do poder do teste VMAX quando comparado com o teste exato de Sullivan et al. (2007), é mínima.

Tabela 6.17: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 2 na situação 1

n	μ	Estimador	Testes							
			Razão V. Exato	Razão V. $X^2_{(7g,l,a)}$	Razão V. $X^2_{(7g,l,a/2)}$	Sullivan Exato	Sullivan $X^2_{(7g,l,a)}$	Sullivan $X^2_{(7g,l,a/2)}$	VMAX Exato	Determin. Exato
25	$\mu_{\text{conhecido}}$	S_{μ}	0,5710	-	-	0,3906	-	-	0,4236	0,1142
		$\hat{\Gamma}_0$	0,5750	-	-	0,4083	0,4425	0,3806	0,4264	0,1534
	$\mu_{\text{desconhecido}}$	S_n	0,5248	-	-	0,3061	-	-	0,3085	0,1315
		$\hat{\Gamma}_0$	0,5454	0,5943	0,5118	0,3371	0,3398	0,2563	0,3122	0,1489
50	$\mu_{\text{conhecido}}$	S_{μ}	0,8074	-	-	0,6932	-	-	0,6242	0,1242
		$\hat{\Gamma}_0$	0,8091	-	-	0,7550	0,8133	0,7403	0,6257	0,1704
	$\mu_{\text{desconhecido}}$	S_n	0,7894	-	-	0,6546	-	-	0,5442	0,1331
		$\hat{\Gamma}_0$	0,7967	0,8275	0,7876	0,7295	0,7819	0,6903	0,5452	0,1689
100	$\mu_{\text{conhecido}}$	S_{μ}	0,9744	-	-	0,9553	-	-	0,8503	0,1304
		$\hat{\Gamma}_0$	0,9536	-	-	0,9913	0,9962	0,9916	0,8526	0,1764
	$\mu_{\text{desconhecido}}$	S_n	0,9707	-	-	0,9504	-	-	0,8147	0,1393
		$\hat{\Gamma}_0$	0,9479	0,9554	0,9461	0,9909	0,9957	0,9901	0,8158	0,1733
200	$\mu_{\text{conhecido}}$	S_{μ}	0,9998	-	-	0,9996	-	-	0,9796	0,1371
		$\hat{\Gamma}_0$	0,9951	-	-	1,0000	1,0000	1,0000	0,9798	0,1904
	$\mu_{\text{desconhecido}}$	S_n	0,9998	-	-	0,9996	-	-	0,9742	0,1407
		$\hat{\Gamma}_0$	0,9952	0,9959	0,9950	1,0000	1,0000	1,0000	0,9745	0,1841

$$\text{MODELO 2: Sob } H_1 : \Gamma = \begin{pmatrix} 2,7778 & 1,8182 \\ 1,8182 & 1,9608 \end{pmatrix}$$

Tabela 6.18: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 3 na situação 1

n	μ	Estimador	Testes							
			Razão V. Exato	Razão V. $X^2_{(7g,l,a)}$	Razão V. $X^2_{(7g,l,a/2)}$	Sullivan Exato	Sullivan $X^2_{(7g,l,a)}$	Sullivan $X^2_{(7g,l,a/2)}$	VMAX Exato	Determin. Exato
25	$\mu_{\text{conhecido}}$	S_{μ}	0,6738	-	-	0,7160	-	-	0,5421	0,6391
		$\hat{\Gamma}_0$	0,6970	-	-	0,7559	0,7797	0,7335	0,5366	0,5383
	$\mu_{\text{desconhecido}}$	S_n	0,4923	-	-	0,6033	-	-	0,5833	0,6408
		$\hat{\Gamma}_0$	0,5216	0,5576	0,4976	0,6457	0,6474	0,5887	0,5652	0,5377
50	$\mu_{\text{conhecido}}$	S_{μ}	0,9210	-	-	0,9354	-	-	0,7927	0,8893
		$\hat{\Gamma}_0$	0,9284	-	-	0,9573	0,9677	0,9546	0,7901	0,8389
	$\mu_{\text{desconhecido}}$	S_n	0,8585	-	-	0,8956	-	-	0,8168	0,8876
		$\hat{\Gamma}_0$	0,8716	0,8940	0,8650	0,9288	0,9407	0,9195	0,8098	0,8380
100	$\mu_{\text{conhecido}}$	S_{μ}	0,9978	-	-	0,9984	-	-	0,9699	0,9930
		$\hat{\Gamma}_0$	0,9976	-	-	0,9995	0,9997	0,9995	0,9702	0,9865
	$\mu_{\text{desconhecido}}$	S_n	0,9955	-	-	0,9969	-	-	0,9736	0,9927
		$\hat{\Gamma}_0$	0,9957	0,9970	0,9954	0,9990	0,9994	0,9990	0,9730	0,9871
200	$\mu_{\text{conhecido}}$	S_{μ}	1,0000	-	-	1,0000	-	-	0,9997	1,0000
		$\hat{\Gamma}_0$	1,0000	-	-	1,0000	1,0000	1,0000	0,9997	1,0000
	$\mu_{\text{desconhecido}}$	S_n	1,0000	-	-	1,0000	-	-	0,9997	1,0000
		$\hat{\Gamma}_0$	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9997	1,0000

$$\text{MODELO 3: Sob } H_1: \Gamma = \begin{pmatrix} 2,6667 & 0,4615 \\ 0,4615 & 2,9412 \end{pmatrix}$$

Tabela 6.19: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 4 na situação 1

n	μ	Estimador	Testes							
			Razão V. Exato	Razão V. $X^2_{(7g,l,a)}$	Razão V. $X^2_{(7g,l,a/2)}$	Sullivan Exato	Sullivan $X^2_{(7g,l,a)}$	Sullivan $X^2_{(7g,l,a/2)}$	VMAX Exato	Determin. Exato
25	$\mu_{\text{conhecido}}$	S_{μ}	0,9393	-	-	0,9506	-	-	0,9559	0,6741
		$\hat{\Gamma}_0$	0,9476	-	-	0,9555	0,9620	0,9493	0,9519	0,5838
	$\mu_{\text{desconhecido}}$	S_n	0,8994	-	-	0,9288	-	-	0,9675	0,6756
		$\hat{\Gamma}_0$	0,9019	0,9157	0,8920	0,9345	0,9351	0,9151	0,9612	0,5826
50	$\mu_{\text{conhecido}}$	S_{μ}	0,9982	-	-	0,9990	-	-	0,9991	0,9106
		$\hat{\Gamma}_0$	0,9988	-	-	0,9993	0,9995	0,9992	0,9990	0,8628
	$\mu_{\text{desconhecido}}$	S_n	0,9973	-	-	0,9983	-	-	0,9994	0,9092
		$\hat{\Gamma}_0$	0,9975	0,9982	0,9972	0,9987	0,9991	0,9985	0,9993	0,8620
100	$\mu_{\text{conhecido}}$	S_{μ}	1,0000	-	-	1,0000	-	-	1,0000	0,9956
		$\hat{\Gamma}_0$	1,0000	-	-	1,0000	1,0000	1,0000	1,0000	0,9896
	$\mu_{\text{desconhecido}}$	S_n	1,0000	-	-	1,0000	-	-	1,0000	0,9954
		$\hat{\Gamma}_0$	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9902
200	$\mu_{\text{conhecido}}$	S_{μ}	1,0000	-	-	1,0000	-	-	1,0000	1,0000
		$\hat{\Gamma}_0$	1,0000	-	-	1,0000	1,0000	1,0000	1,0000	1,0000
	$\mu_{\text{desconhecido}}$	S_n	1,0000	-	-	1,0000	-	-	1,0000	1,0000
		$\hat{\Gamma}_0$	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

$$\text{MODELO 4: Sob } H_1 : \Gamma = \begin{pmatrix} 5,3333 & 1,5385 \\ 1,5385 & 1,9608 \end{pmatrix}$$

Finalmente, no modelo 3, alteramos toda a matriz de covariâncias, e no modelo 5 mantivemos inalterada uma das variâncias e variamos os demais valores, como pode ser visto pela Tabela 6.15. No modelo 3 (ver Tabela 6.18), quando comparamos os testes exatos, novamente verificamos a superioridade do teste de Sullivan et al. (2007). Ao contrário do que ocorreu no modelo 1, nesse caso o teste do determinante conseguiu captar tal alteração na matriz de covariâncias. Isso deve-se ao fato de que razão dos determinantes, da matriz postulada sob a hipótese alternativa e sob a hipótese nula, é grande ($\Lambda_{\text{det}}=3,7725$)., indicando que o teste do determinante seria capaz perceber tal alteração. Note que para amostras grandes ($n>100$), o desempenho do teste do determinante é equiparável com o teste da razão de verossimilhança e do teste de Sullivan et al. (2007)

Observe, no modelo 3, Tabela 6.18, que para tamanhos amostrais grandes ($n=200$) a maioria dos testes alcança poder 1 (um). Um fato a ser considerado é que as correções nas distribuições assintóticas das estatísticas de teste foram muito eficientes. Pela Tabela 6.18, observa-se que a aproximação para a distribuição $\chi^2_{7,\alpha/2}$ forneceu valores de poder bem próximos daqueles obtidos pela distribuição exata. Para $n=100$, considerando μ desconhecido, e $\hat{\Gamma}_0$ como o estimador da matriz de covariâncias obtemos um poder de 0,9957 para o teste exato da razão de verossimilhança e um poder de 0,9954 quando usamos esse teste considerando a distribuição assintótica $\chi^2_{7,\alpha/2}$. Esse fato também é verificado quando consideramos o teste de Sullivan et al. (2007). Nesse mesmo contexto o teste exato de Sullivan et al. (2007) fornece um poder de 0,9990 e a distribuição assintótica $\chi^2_{7,\alpha/2}$, considerando a correção de $\alpha/2$, fornece o mesmo valor (0,9990). Isso mostra que a correção proposta pode ser de grande utilidade, uma vez que o uso da distribuição exata requer práticas de simulação o que demandaria mais tempo que o uso da distribuição assintótica.

O modelo 5 se destaca pelo baixo desempenho do teste exato do determinante quando comparado com os demais testes exatos, sendo que somente com $n=200$, esse teste consegue um poder próximo daqueles obtidos para os demais testes com tamanho amostral igual a 25. Ou seja, o desempenho dos demais testes exatos, com $n=25$, é igual ao desempenho do teste do determinante considerando $n=200$. O teste VMAX obteve estimativas de poder bem próximas ao teste exato da razão de verossimilhança e o teste exato de Sullivan et al. (2007), isso porque no modelo 5 uma das variâncias alterou significativamente e o teste VMAX foi capaz de sinalizar tal mudança (valores de poder entre 0,50 e 0,60 para $n=50$; entre 0,80 e 0,86 para $n=100$ e 0,95 para $n=200$).

Tabela 6.20: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 5 na situação 1

n	μ	Estimador	Testes							
			Razão V. Exato	Razão V. $X^2_{(7g,l,a)}$	Razão V. $X^2_{(7g,l,a/2)}$	Sullivan Exato	Sullivan $X^2_{(7g,l,a)}$	Sullivan $X^2_{(7g,l,a/2)}$	VMAX Exato	Determin. Exato
25	$\mu_{\text{conhecido}}$	S_{μ}	0,4412	-	-	0,3759	-	-	0,4262	0,1285
		$\hat{\Gamma}_0$	0,4230	-	-	0,3841	0,4086	0,3637	0,4287	0,1622
	$\mu_{\text{desconhecido}}$	S_n	0,3244	-	-	0,2385	-	-	0,3107	0,0959
		$\hat{\Gamma}_0$	0,3405	0,3872	0,3109	0,2548	0,2567	0,2026	0,3142	0,1239
50	$\mu_{\text{conhecido}}$	S_{μ}	0,6516	-	-	0,6019	-	-	0,6259	0,1780
		$\hat{\Gamma}_0$	0,6528	-	-	0,6181	0,6628	0,6078	0,6280	0,2047
	$\mu_{\text{desconhecido}}$	S_n	0,5644	-	-	0,4986	-	-	0,5461	0,1350
		$\hat{\Gamma}_0$	0,5787	0,6256	0,5659	0,5229	0,5665	0,4939	0,5470	0,1647
100	$\mu_{\text{conhecido}}$	S_{μ}	0,8854	-	-	0,8601	-	-	0,8500	0,2558
		$\hat{\Gamma}_0$	0,8752	-	-	0,8960	0,9250	0,8971	0,8522	0,2634
	$\mu_{\text{desconhecido}}$	S_n	0,8441	-	-	0,8200	-	-	0,8150	0,2091
		$\hat{\Gamma}_0$	0,8397	0,8621	0,8344	0,8685	0,9017	0,8647	0,8161	0,2276
200	$\mu_{\text{conhecido}}$	S_{μ}	0,9907	-	-	0,9884	-	-	0,9799	0,4025
		$\hat{\Gamma}_0$	0,9803	-	-	0,9970	0,9986	0,9974	0,9800	0,3843
	$\mu_{\text{desconhecido}}$	S_n	0,9867	-	-	0,9848	-	-	0,9746	0,3542
		$\hat{\Gamma}_0$	0,9750	0,9788	0,9740	0,9962	0,9981	0,9965	0,9749	0,3479

$$\text{MODELO 5: Sob } H_1: \Gamma = \begin{pmatrix} 2,7778 & 1,5909 \\ 1,5909 & 1,9608 \end{pmatrix}$$

De um modo geral, quando usamos o teste *Step-Down*, proposto por Sullivan et al. (2007), exceto no modelo 1, o poder do teste, quando usamos a matriz $\hat{\Gamma}_0$ é superior aquele em que usamos a matriz S. Esse fato mostra a importância de detectarmos a autocorrelação das observações e ajustarmos um modelo coerente. Já para os testes da razão de verossimilhança, determinante e VMAX, não verificamos a superioridade predominante quando usamos a matriz $\hat{\Gamma}_0$. Ou seja, existe uma alternância de acordo com o modelo estudado. Em algumas situações o uso da matriz S fornece maior poder do que quando usamos a matriz $\hat{\Gamma}_0$, mas sempre valores bem próximos. Como discutido anteriormente, uma vantagem do uso da matriz $\hat{\Gamma}_0$ é que podemos usar aproximações para a distribuição qui-quadrado, o que não ocorre quando usamos a matriz S. Assim, se optarmos por trabalhar com a matriz de covariância amostral S teremos que trabalhar com a distribuição exata, que em termos práticos dificultaria bastante o trabalho.

Em todos os modelos estudados, quando a alteração na matriz de covariâncias é pequena, todos os testes têm dificuldade em detectar tal alteração, porém, com uma superioridade do teste proposto por Sullivan et al. (2007). De uma forma geral, na situação em que temos pequenas alterações, quando aumentamos o tamanho amostral os testes conseguem perceber tal alteração, como é esperado.

Até o momento focalizamos nosso estudo no caso em que temos um modelo VAR(1) e as duas séries não são relacionadas ($\phi_{12} = \phi_{21} = 0$), ou seja, temos somente a dependência temporal $\phi_{11} \neq 0$, $\phi_{22} \neq 0$. Na próxima seção nosso objetivo será avaliar o poder dos testes quando temos um modelo bivariado VAR(1) e detectamos correlação temporal e entre as duas séries ($\phi_{11} \neq 0, \phi_{22} \neq 0, \phi_{12} \neq 0$ e $\phi_{21} \neq 0$).

6.2 Estudo do Poder dos Testes - Situação 2

Nessa seção avaliaremos o tamanho e o poder dos testes para matriz de covariâncias, formulados sob a suposição de independência entre as observações, quando detectamos correlação temporal e entre as duas séries, como também das correções dos testes propostos no Capítulo 4. Mantendo o mesmo padrão da seção anterior, consideraremos as distribuições exatas e aproximadas dos testes em questão. Os percentis das distribuições exatas foram obtidos através de simulação e estão apresentados no Apêndice C. Vale lembrar que nessa seção utilizaremos como

ferramenta a função *mAr.est* e a função *mAr.sim*, do pacote *mAr*, como discutimos no Capítulo 5, uma vez que estamos abordando o caso em que ($\phi_{11} \neq 0, \phi_{22} \neq 0, \phi_{12} \neq 0$ e $\phi_{21} \neq 0$) e o uso do pacote *mAr* facilitará o processo de simulação. Ao contrário da situação abordada na seção 6.1, nessa segunda parte consideraremos somente o caso em que o vetor de médias é desconhecido, isso porque, a função *mAr.est* não nos dá a opção de declararmos o vetor de médias como conhecido, ou seja, a função sempre estimará o vetor de médias. Tanto na avaliação do tamanho dos testes quanto do poder utilizaremos a seguinte hipótese:

$$\begin{cases} H_0 : \Gamma = \Gamma_0 = \begin{pmatrix} 1,3333 & 0,9744 \\ 0,9744 & 2,6003 \end{pmatrix} \\ H_1 : \Gamma \neq \Gamma_0 \end{cases}$$

onde, sob H_0 , consideraremos um processo VAR(1) com os seguintes parâmetros:

$$\Phi = \begin{pmatrix} 0,5 & 0,0 \\ 0,2 & 0,7 \end{pmatrix}, \quad \Sigma_\varepsilon = \begin{pmatrix} 1,0 & 0,5 \\ 0,5 & 1,0 \end{pmatrix} \text{ e consequentemente } \Gamma_0 = \begin{pmatrix} 1,3333 & 0,9744 \\ 0,9744 & 2,6003 \end{pmatrix}$$

Para a matriz Γ_0 temos as seguintes informações: $|\Gamma_0| = 2,5175$; $tr(\Gamma_0) = 3,9336$; autovalores: $\lambda_1 = 3,1290$, $\lambda_2 = 0,8046$, sendo a correlação entre as duas variáveis igual a 0,523.

No Capítulo 4, seção 4.1, vimos que se $\phi_{12} = 0$ e $\phi_{21} \neq 0$, que é o caso em que estamos abordando nessa seção, existe uma relação unidirecional de Y_{1t} para Y_{2t} . Se $\phi_{12} = \phi_{21} = 0$, caso abordado na seção 6.1, as variáveis Y_{1t} e Y_{2t} são não relacionadas. Por fim, se $\phi_{12} \neq 0$ e $\phi_{21} \neq 0$ existe um relacionamento entre as duas séries de estudo. Em geral, a matriz de coeficientes Φ mede a dinâmica de dependência de Y_t . Na situação abordada nessa seção consideraremos uma dependência temporal na ordem de $\phi_{11} = 0,5$ e $\phi_{22} = 0,7$ para as variáveis e uma relação unidirecional de Y_{1t} para Y_{2t} na ordem de $\phi_{21} = 0,2$. A relação entre Y_{1t} e Y_{2t} é considerado unidirecional pelo fato de que nessa situação estamos considerando $\phi_{12} = 0$.

A seleção dos testes para os quais avaliaremos o poder foi feita através da estimação do tamanho dos mesmos. O procedimento utilizado para a determinação do tamanho dos testes, nessa situação, foi similar àquela feita para o caso abordado na

seção 6.1. A única diferença é que na situação 1 usamos nosso próprio algoritmo de simulação e estimação, e já nessa situação usamos funções do pacote *mAr*. O processo de determinação do tamanho dos testes seguiu as seguintes etapas: simulamos n observações, sob H_0 , de um processo bivariado VAR(1) 100.000 vezes e contabilizamos a quantidade de amostras que levaram a rejeição de H_0 . A estimativa do tamanho do teste foi determinada pela proporção de amostras que levaram a rejeição de H_0 . Os testes escolhidos para avaliação do poder foram aqueles que fornecerem estimativas de α em torno de 5%. Semelhante a situação 1, consideramos amostras de tamanhos 25, 50, 100, 200, 1000 e 5000. Vale lembrar que em situações onde temos pequenas amostras a função *mAr.sim* não é consistente, ou seja, ela gera observações que não são condizentes com o modelo e parâmetros especificados. Diante disso, como discutido no Capítulo 5, criamos certas restrições no processo de simulação, de forma que somente as amostras que satisfizessem nossos objetivos (modelo e parâmetros) fossem selecionadas. Essas restrições aumentaram bruscamente o tempo de simulação, mas foi a única alternativa encontrada para trabalharmos com amostras pequenas.

Quanto ao número de graus de liberdade, quando estamos trabalhando com as distribuições aproximadas, houve uma pequena modificação para o caso em que trabalhamos com a matriz $\hat{\Gamma}_0$, visto que nessa situação temos que estimar os parâmetros da diagonal secundária da matriz Φ , o que não acontecia na situação 1. Observe que quando trabalhamos com a matriz de covariâncias amostral S , mantivemos os mesmos graus de liberdade da situação 1, visto que a quantidade de parâmetros a serem estimados em ambas as situações são as mesmas. A única mudança foi para os casos onde trabalhamos com a matriz $\hat{\Gamma}_0$. Para o teste da razão de verossimilhança e para o teste *Step-Down* proposto por Sullivan et al. (2007), quando trabalhamos com a matriz $\hat{\Gamma}_0$, considerando o vetor de médias desconhecido, usamos uma aproximação para a distribuição qui-quadrado com 7 graus de liberdade. Isso porque, sob o espaço paramétrico restrito por H_0 , temos que estimar o vetor de médias e sob o espaço paramétrico irrestrito temos que estimar os 9 parâmetros do modelo VAR(1) ($\mu_1, \mu_2, \sigma_{11}, \sigma_{22}, \sigma_{21}, \phi_{11}, \phi_{22}, \phi_{12}$ e ϕ_{21}).

Um estudo da distribuição exata do teste da razão de verossimilhança e do teste *Step-Down* de Sullivan et al. (2007) indicou uma proximidade com os valores da distribuição qui-quadrado com 7 graus de liberdade, considerando $\alpha/2$, e com 9 graus de liberdade, considerando o nível de significância α de 5%. Encontrar uma distribuição aproximada para as estatísticas de testes, quando estamos em um processo

autocorrelacionado é um grande ganho, visto que temos uma alternativa de distribuição a trabalhar além da distribuição exata. Como estamos em um estudo exploratório para avaliarmos qual é a melhor aproximação para a distribuição da estatística de teste no caso autocorrelacionado, optamos por testar essas outras duas opções de graus de liberdade e verificar o que ocorre com o tamanho dos testes. No teste do determinante, tanto para S quanto para $\hat{\Gamma}_0$, considerando o caso onde o vetor de médias é desconhecido, utilizamos o teste exato, o teste tradicional da variância generalizada e a modificação proposta por Djauhari (2005). Para o teste VMAX, consideramos somente a distribuição exata, que foi obtida via simulação. A Tabela 6.21 apresenta sumariamente os testes para os quais avaliaremos o tamanho especificando o número de graus de liberdade da distribuição assintótica de cada um deles.

Tabela 6.21: Testes abordados na situação 2

Testes	
<i>μ desconhecido</i>	
S	$\hat{\Gamma}_0$
Determinante - Exato	Determinante - Exato
Determinante - Trad.	Determinante - Trad.
Determinante - Djauhari	Determinante - Djauhari
Razão de V. - Exato	Razão de V. - Exato
Razão de V. - X^2 (3 g.l.g)	Razão de V. - X^2 (7 g.l.g)
Sullivan - Exato	Razão de V. - X^2 (7 g.l.g/2)
Sullivan - X^2 (3 g.l.g)	Razão de V. - X^2 (9 g.l.g)
Sullivan - X^2 (5 g.l.g)	Sullivan - Exato
VMAX - Exato	Sullivan - X^2 (7 g.l.g)
	Sullivan - X^2 (7 g.l.g/2)
	Sullivan - X^2 (9 g.l.g)
	VMAX - Exato

Observe que os graus de liberdade das distribuições assintóticas foram determinados levando em consideração um modelo bivariado VAR(1) onde $\phi_{21} \neq 0$ e $\phi_{12} \neq 0$, ou seja, existem mais parâmetros a serem estimados do que na situação 1. Na próxima seção apresentaremos as estimativas do tamanho dos testes apresentados na Tabela 6.21.

6.2.1 Estimação do tamanho dos testes na situação 1

Os resultados obtidos via simulação de Monte Carlo para as estimativas dos tamanhos dos testes são apresentados nas Tabelas 6.22 a 6.27. Em cada uma das Tabelas apresentamos os valores críticos das distribuições assintóticas e exatas (considerando $\alpha = 0,05$) e a estimativa do tamanho do teste.

Tabela 6.22: Limites críticos e estimativas do tamanho do teste na situação 2, para $n=25$

Teste	n = 25 - μ desconhecido			
	S_n		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(0,835 ; 5,892)	0,0493	(0,976 ; 7,327)	0,0508
Determinante - Trad.	(0,420 ; 4,405)	0,0901	(0,420 ; 4,405)	0,1786
Determinante - Djauhari	(0,602 ; 4,433)	0,0923	(0,602 ; 4,433)	0,1771
Razão de V. - Exato	11,602	0,0505	14,928	0,0508
Razão de V. - $X^2_{(3q,l,\alpha)}$	7,815	0,1553	-	-
Razão de V. - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0597
Razão de V. - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0420
Razão de V. - $X^2_{(9q,l,\alpha)}$	-	-	16,919	0,0359
Sullivan - Exato	12,874	0,0511	17,137	0,0505
Sullivan - $X^2_{(3q,l,\alpha)}$	7,815	0,1708	-	-
Sullivan - $X^2_{(5q,l,\alpha)}$	11,070	0,0763	-	-
Sullivan - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0814
Sullivan - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0601
Sullivan - $X^2_{(9q,l,\alpha)}$	-	-	16,919	0,0521
VMAX - Exato	1,848	0,0491	2,103	0,0497

Tabela 6.23: Limites críticos e estimativas do tamanho do teste na situação 2, para $n=50$

Teste	n = 50 - μ desconhecido			
	S_n		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(1,199 ; 4,726)	0,0508	(1,298 ; 5,186)	0,0513
Determinante - Trad.	(1,064 ; 3,869)	0,0966	(1,064 ; 3,869)	0,1393
Determinante - Djauhari	(1,143 ; 3,892)	0,1009	(1,143 ; 3,892)	0,1397
Razão de V. - Exato	11,962	0,0516	13,407	0,0503
Razão de V. - $X^2_{(3q,l,\alpha)}$	7,815	0,1584	-	-
Razão de V. - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0434
Razão de V. - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0280
Razão de V. - $X^2_{(9q,l,\alpha)}$	-	-	16,919	0,0230
Sullivan - Exato	12,689	0,0519	14,567	0,0506
Sullivan - $X^2_{(3q,l,\alpha)}$	7,815	0,1662	-	-
Sullivan - $X^2_{(5q,l,\alpha)}$	11,070	0,0758	-	-
Sullivan - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0560
Sullivan - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0388
Sullivan - $X^2_{(9q,l,\alpha)}$	-	-	16,919	0,0330
VMAX - Exato	1,596	0,0496	1,683	0,0493

Tabela 6.24: Limites críticos e estimativas do tamanho do teste na situação 2, para n=100

Teste	n = 100 - μ desconhecido			
	S_n		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(1,492 ; 4,021)	0,0503	(1,553 ; 4,206)	0,0505
Determinante - Trad.	(1,503 ; 3,481)	0,1100	(1,503 ; 3,481)	0,1312
Determinante - Djauhari	(1,538 ; 3,497)	0,1136	(1,538 ; 3,497)	0,1325
Razão de V. - Exato	12,527	0,0503	13,164	0,0502
Razão de V. - $X^2_{(3q,l,\alpha)}$	7,815	0,1726	-	-
Razão de V. - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0400
Razão de V. - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0257
Razão de V. - $X^2_{(9q,l,\alpha)}$	-	-	16,919	0,0208
Sullivan - Exato	12,982	0,0494	13,828	0,0493
Sullivan - $X^2_{(3q,l,\alpha)}$	7,815	0,1757	-	-
Sullivan - $X^2_{(5q,l,\alpha)}$	11,070	0,0782	-	-
Sullivan - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0469
Sullivan - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0317
Sullivan - $X^2_{(9q,l,\alpha)}$	-	-	16,919	0,0266
VMAX - Exato	1,411	0,0509	1,445	0,0507

Tabela 6.25: Limites críticos e estimativas do tamanho do teste na situação 2, para n=200

Teste	n = 200 - μ desconhecido			
	S_n		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(1,716 ; 3,579)	0,0500	(1,749 ; 3,658)	0,0498
Determinante - Trad.	(1,806 ; 3,204)	0,1313	(1,806 ; 3,204)	0,1414
Determinante - Djauhari	(1,822 ; 3,213)	0,1336	(1,822 ; 3,213)	0,1425
Razão de V. - Exato	13,238	0,0511	13,568	0,0513
Razão de V. - $X^2_{(3q,l,\alpha)}$	7,815	0,1900	-	-
Razão de V. - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0456
Razão de V. - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0285
Razão de V. - $X^2_{(9q,l,\alpha)}$	-	-	16,919	0,0233
Sullivan - Exato	13,393	0,0511	13,873	0,0501
Sullivan - $X^2_{(3q,l,\alpha)}$	7,815	0,1906	-	-
Sullivan - $X^2_{(5q,l,\alpha)}$	11,070	0,0885	-	-
Sullivan - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0485
Sullivan - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0320
Sullivan - $X^2_{(9q,l,\alpha)}$	-	-	16,919	0,0268
VMAX - Exato	1,291	0,0505	1,307	0,0500

Tabela 6.26: Limites críticos e estimativas do tamanho do teste na situação 2, para n=1000

Teste	n = 1000 - μ desconhecido			
	S_n		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(2,057 ; 3,030)	0,0513	(2,065 ; 3,043)	0,0512
Determinante - Trad.	(2,203 ; 2,827)	0,2095	(2,203 ; 2,827)	0,2104
Determinante - Djauhari	(2,206 ; 2,829)	0,2107	(2,206 ; 2,829)	0,2109
Razão de V. - Exato	16,804	0,0501	16,919	0,0502
Razão de V. - $X^2_{(3q,l,\alpha)}$	7,815	0,2778	-	-
Razão de V. - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0863
Razão de V. - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0593
Razão de V. - $X^2_{(9q,l,\alpha)}$	-	-	16,919	0,0502
Sullivan - Exato	16,806	0,0501	16,857	0,0505
Sullivan - $X^2_{(3q,l,\alpha)}$	7,815	0,2776	-	-
Sullivan - $X^2_{(5q,l,\alpha)}$	11,070	0,1496	-	-
Sullivan - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,0850
Sullivan - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0589
Sullivan - $X^2_{(9q,l,\alpha)}$	-	-	16,919	0,0499
VMAX - Exato	1,140	0,0511	1,143	0,0503

Tabela 6.27: Limites críticos e estimativas do tamanho do teste na situação 2, para n=5000

Teste	n = 5000 - μ desconhecido			
	S_n		$\hat{\Gamma}_0$	
	Valor Crítico	Tam. Teste	Valor Crítico	Tam. Teste
Determinante - Exato	(2,291 ; 2,751)	0,0508	(2,293 ; 2,753)	0,0509
Determinante - Trad.	(2,377 ; 2,657)	0,2376	(2,377 ; 2,657)	0,2383
Determinante - Djauhari	(2,378 ; 2,657)	0,2381	(2,378 ; 2,657)	0,2381
Razão de V. - Exato	18,289	0,0503	18,316	0,0504
Razão de V. - $X^2_{(3q,l,\alpha)}$	7,815	0,3060	-	-
Razão de V. - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,1030
Razão de V. - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0746
Razão de V. - $X^2_{(9q,l,\alpha)}$	-	-	16,919	0,0636
Sullivan - Exato	18,267	0,0503	18,277	0,0504
Sullivan - $X^2_{(3q,l,\alpha)}$	7,815	0,3056	-	-
Sullivan - $X^2_{(5q,l,\alpha)}$	11,070	0,1711	-	-
Sullivan - $X^2_{(7q,l,\alpha)}$	-	-	14,067	0,1022
Sullivan - $X^2_{(7q,l,\alpha/2)}$	-	-	16,013	0,0740
Sullivan - $X^2_{(9q,l,\alpha)}$	-	-	16,919	0,0630
VMAX - Exato	1,065	0,0480	1,065	0,0495

Diante dos resultados obtidos (Tabelas 6.22-6.27), selecionamos para avaliação do poder aqueles que forneceram estimativas de α em torno de 5%. Um fato a ser observado é que nos casos em que consideramos a distribuição assintótica $\chi^2_{7gl,\alpha}$ obtivemos estimativas de α próximas a 5%.

Note que quando trabalhamos com amostras de tamanho 50, 100 e 200 e usamos a distribuição aproximada qui-quadrado com 9 graus de liberdade, tanto para o teste da razão de verossimilhança quanto para o teste Step-Down de Sullivan et al. (2007),

obtemos valores de $\hat{\alpha}$ inferiores ao α nominal (0,05). Diante desse fato, nos surgiu o interesse de avaliar o poder dos testes quando usamos a distribuição aproximada qui-quadrado com 9 graus de liberdade, porém, considerando 2α ao invés de α . Novamente esse fato é de grande importância prática, visto a dificuldade que se tem para trabalhar com a distribuição exata já que ela é simulada.

Da mesma forma que verificamos na situação abordada na seção 6.1, quando trabalhamos com a matriz de covariâncias amostral S , não obtivemos nenhum sucesso nas aproximações para a distribuição qui-quadrado. Diante desse fato, somente as distribuições exatas, quando se utilizava a matriz S , foram consideradas na avaliação do poder. A Tabela 6.28 apresenta resumidamente os testes que foram selecionados para a fase de determinação do poder.

Tabela 6.28: Testes selecionados para avaliação do poder na situação 2

Testes	Média desconhecida	
	S_n	$\hat{\Gamma}_0$
Determinante	Exato	Exato
VMAX	Exato	Exato
Razão de Verossim.	Exato	Exato - $\chi^2_{7gl.,\alpha}$ - $\chi^2_{9gl.,\alpha}$ - $\chi^2_{9gl.,2\alpha}$
Step-Down (Sullivan et. al)	Exato	Exato - $\chi^2_{7gl.,\alpha}$ - $\chi^2_{9gl.,\alpha}$ - $\chi^2_{9gl.,2\alpha}$

6.2.2 Estimação do poder dos testes na situação 2

Na avaliação do poder dos testes, consideramos amostras de tamanho 25, 50, 100 e 200. Para obter o poder dos testes simulamos n observações, sob H_1 , de um modelo VAR(1) 20.000 vezes e determinamos o poder de cada teste. Esse processo foi repetido 25 vezes para cada teste, como feito na seção 6.1. Vale lembrar que tanto para avaliação do tamanho quanto para a determinação do poder dos testes utilizamos a seguinte hipóteses nula e alternativa:

$$\begin{cases} H_0 : \Gamma = \Gamma_0 = \begin{pmatrix} 1,3333 & 0,9744 \\ 0,9744 & 2,6003 \end{pmatrix} \\ H_1 : \Gamma \neq \Gamma_0 \end{cases}$$

Na determinação do poder dos testes consideramos 4 modelos distintos (ver Tabela 6.29) para avaliar a capacidade dos testes em detectar alterações na matriz de covariâncias. De uma forma geral, buscamos selecionar modelos que abrangessem vários tipos de alterações na estrutura da matriz de covariâncias. Em todos os modelos apresentados na Tabela 6.29 estamos considerando um processo bivariado VAR(1). O valor médio das 25 estimativas de poder de cada teste estão apresentados nas Tabelas 6.30 a 6.33.

Com os modelos propostos na Tabela 6.29 avaliamos o desempenho dos testes quando tínhamos pequenas e grandes alterações na estrutura da matriz de covariâncias. Na situação abordada na seção 6.1 vimos que quanto menor a alteração na estrutura da matriz maior era a dificuldade dos testes em captar tal mudança. Diante disso, a única alternativa para detecção da não veracidade de H_0 , era aumentar o tamanho amostral. Esse mesmo fato também foi verificado na situação 2, onde além da dependência temporal temos uma relação entre as duas séries, isto é, todos os testes apresentaram dificuldades em captar mudanças na estrutura da matriz quando essas eram pequenas.

No modelo 1, apresentado na Tabela 6.29, alteramos minimamente o parâmetro ϕ_{21} da matriz Φ . Observe que para $n=25$, todos os testes apresentaram um baixo desempenho ($\max \approx 0,14$ para o teste da razão de verossimilhança). Mesmo quando aumentamos o tamanho amostral para $n=200$ o poder dos testes não apresentam grandes ganhos em relação ao que acontece para n menor. Nesse caso, a alteração na matriz de covariâncias foi tão pequena que os testes não conseguiram captá-la.

No modelo 2, também alteramos somente o parâmetro ϕ_{21} da matriz Φ , porém, a mudança na matriz de covariâncias foi mais acentuada que no modelo 1. Como tivemos uma maior alteração na matriz de covariâncias os testes conseguiram sinalizar com maior facilidade a não veracidade de H_0 do que no caso do modelo 1. Observe que nesse modelo 2 a variabilidade da segunda variável aumentou consideravelmente em relação ao postulado sob H_0 . Diante disso, o teste VMAX, que leva em consideração somente alterações na variância das variáveis, teve valores de poder superiores daqueles obtidos pelo teste da razão de verossimilhança e pelo teste proposto por Sullivan et al. (2007). Ainda no contexto do modelo 2, o teste com o pior desempenho foi o teste do determinante ($\max \approx 0,42$, para $n=200$). No entanto, embora os valores de poder tenham sido maiores que os observados para o modelo 1, o valor máximo alcançado não foi próximo de 1 (poder máximo $\approx 0,66$ para $n=200$, considerando o teste VMAX). Assim, seria necessário aumentar o tamanho da amostra para obter um valor maior de poder.

Tabela 6.29: Modelos para avaliação do poder dos testes na situação 2

Modelos	Parâmetros e Medidas
<i>Modelo sob H_0</i>	Sob H_0 : $\Gamma_0 = \begin{pmatrix} 1,3333 & 0,9744 \\ 0,9744 & 2,6003 \end{pmatrix}$ onde $\Phi = \begin{pmatrix} 0,5 & 0,0 \\ 0,2 & 0,7 \end{pmatrix}$, $\Sigma_\varepsilon = \begin{pmatrix} 1,0 & 0,5 \\ 0,5 & 1,0 \end{pmatrix}$, $ \Gamma_0 = 2,5175$, $tr(\Gamma_0) = 3,9336$, $\rho = 0,52$.
<i>Modelo 1</i>	Sob H_1 : $\Gamma = \begin{pmatrix} 1,3333 & 0,8718 \\ 0,8718 & 2,2262 \end{pmatrix}$ onde $\Phi = \begin{pmatrix} 0,5 & 0,0 \\ 0,1 & 0,7 \end{pmatrix}$, $\Sigma_\varepsilon = \begin{pmatrix} 1,0 & 0,5 \\ 0,5 & 1,0 \end{pmatrix}$, $\Lambda_{det} = \frac{ \Gamma }{ \Gamma_0 } = \frac{2,2082}{2,5175} = 0,8771$ e $\Lambda_{tr} = \frac{tr(\Gamma)}{tr(\Gamma_0)} = 0,9049$, $\rho = 0,51$.
<i>Modelo 2</i>	Sob H_1 : $\Gamma = \begin{pmatrix} 1,3333 & 1,1795 \\ 1,1795 & 3,6742 \end{pmatrix}$ onde $\Phi = \begin{pmatrix} 0,5 & 0,0 \\ 0,4 & 0,7 \end{pmatrix}$, $\Sigma_\varepsilon = \begin{pmatrix} 1,0 & 0,5 \\ 0,5 & 1,0 \end{pmatrix}$, $\Lambda_{det} = \frac{ \Gamma }{ \Gamma_0 } = \frac{3,5076}{2,5175} = 1,3933$ e $\Lambda_{tr} = \frac{tr(\Gamma)}{tr(\Gamma_0)} = 1,2730$, $\rho = 0,53$.
<i>Modelo 3</i>	Sob H_1 : $\Gamma = \begin{pmatrix} 1,3333 & 1,7692 \\ 1,7692 & 2,6078 \end{pmatrix}$ onde $\Phi = \begin{pmatrix} 0,5 & 0,0 \\ 0,0 & 0,7 \end{pmatrix}$, $\Sigma_\varepsilon = \begin{pmatrix} 1,00 & 1,15 \\ 1,15 & 1,33 \end{pmatrix}$, $\Lambda_{det} = \frac{ \Gamma }{ \Gamma_0 } = \frac{0,3469}{2,5175} = 0,1378$ e $\Lambda_{tr} = \frac{tr(\Gamma)}{tr(\Gamma_0)} = 1,0019$, $\rho = 0,95$.
<i>Modelo 4</i>	Sob H_1 : $\Gamma = \begin{pmatrix} 3,1216 & 2,2925 \\ 2,2925 & 2,6513 \end{pmatrix}$ onde $\Phi = \begin{pmatrix} 0,5 & 0,4 \\ 0,1 & 0,7 \end{pmatrix}$, $\Sigma_\varepsilon = \begin{pmatrix} 1,0 & 0,5 \\ 0,5 & 1,0 \end{pmatrix}$, $\Lambda_{det} = \frac{ \Gamma }{ \Gamma_0 } = \frac{3,0207}{2,5175} = 1,1999$ e $\Lambda_{tr} = \frac{tr(\Gamma)}{tr(\Gamma_0)} = 1,4676$, $\rho = 0,80$.

Nota: ρ é o coeficiente de correlação entre as duas variáveis X_1 e X_2 .

Tabela 6.30: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 1 na situação 2

n	μ	Estimador	Testes									
			Razão V. Exato	Razão V. $X^2_{(7g,l,a)}$	Razão V. $X^2_{(9g,l,a)}$	Razão V. $X^2_{(9g,l,2a)}$	Sullivan Exato	Sullivan $X^2_{(7g,l,a)}$	Sullivan $X^2_{(9g,l,a)}$	Sullivan $X^2_{(9g,l,2a)}$	VMAX Exato	Determin. Exato
25	$\mu_{\text{desconhecido}}$	S_n	0,0590	-	-	-	0,0550	-	-	-	0,0258	0,0449
		$\hat{\Gamma}_0$	0,0505	0,0599	0,0347	0,0529	0,0430	0,0718	0,0444	0,0645	0,0257	0,0436
50	$\mu_{\text{desconhecido}}$	S_n	0,0697	-	-	-	0,0642	-	-	-	0,0182	0,0516
		$\hat{\Gamma}_0$	0,0604	0,0514	0,0252	0,0444	0,0492	0,0545	0,0313	0,0482	0,0181	0,0520
100	$\mu_{\text{desconhecido}}$	S_n	0,0902	-	-	-	0,0816	-	-	-	0,0138	0,0664
		$\hat{\Gamma}_0$	0,0841	0,0680	0,0339	0,0585	0,0642	0,0611	0,0321	0,0534	0,0137	0,0660
200	$\mu_{\text{desconhecido}}$	S_n	0,1446	-	-	-	0,1351	-	-	-	0,0100	0,0928
		$\hat{\Gamma}_0$	0,1394	0,1268	0,0729	0,1127	0,1139	0,1097	0,0616	0,0971	0,0098	0,0924

Tabela 6.31: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 2 na situação 2

n	μ	Estimador	Testes									
			Razão V. Exato	Razão V. $X^2_{(7g,l,a)}$	Razão V. $X^2_{(9g,l,a)}$	Razão V. $X^2_{(9g,l,2a)}$	Sullivan Exato	Sullivan $X^2_{(7g,l,a)}$	Sullivan $X^2_{(9g,l,a)}$	Sullivan $X^2_{(9g,l,2a)}$	VMAX Exato	Determin. Exato
25	$\mu_{\text{desconhecido}}$	S_n	0,1275	-	-	-	0,1300	-	-	-	0,1964	0,1300
		$\hat{\Gamma}_0$	0,1386	0,1543	0,1082	0,1431	0,1426	0,1966	0,1455	0,1841	0,1789	0,1257
50	$\mu_{\text{desconhecido}}$	S_n	0,2130	-	-	-	0,2189	-	-	-	0,2929	0,1853
		$\hat{\Gamma}_0$	0,2094	0,1936	0,1370	0,1799	0,2411	0,2527	0,1917	0,2381	0,2842	0,1837
100	$\mu_{\text{desconhecido}}$	S_n	0,3625	-	-	-	0,3715	-	-	-	0,4530	0,2759
		$\hat{\Gamma}_0$	0,3571	0,3284	0,2516	0,3100	0,3964	0,3889	0,3127	0,3712	0,4474	0,2724
200	$\mu_{\text{desconhecido}}$	S_n	0,5952	-	-	-	0,6071	-	-	-	0,6588	0,4195
		$\hat{\Gamma}_0$	0,5926	0,5760	0,4861	0,5556	0,6235	0,6177	0,5353	0,5989	0,6534	0,4173

No modelo 3, alteramos a matriz Φ e a matriz Σ_ε de forma que somente a covariância entre as duas variáveis fosse diferente do valor da matriz Γ_0 . Pela Tabela 6.29, observa-se que para o modelo 3 temos que $\Lambda_{\det} = 0,1378$, ou seja, o determinante da matriz postulada sob H_0 e da matriz sob H_1 , são bem diferentes. Diante desse fato, observe-se que para $n=25$ o teste do determinante apresentou um ótimo desempenho (min $\approx 0,95$ para $n=25$ e max $\approx 1,0$ para $n\geq 100$), assim como o teste da razão de verossimilhança. Embora na maioria dos modelos apresentados, tanto nessa seção, quanto na seção 6.1, o teste da razão de verossimilhança e o teste Step-Down tenham apresentado resultados bem próximos, para o modelo 3 o desempenho do teste proposto por Sullivan et al. (2007) foi um pouco inferior para amostras de tamanho $n=25$, embora tenha sido muito satisfatório (poder mínimo de 0,72). Uma possível explicação para esse fato é que o teste da razão de verossimilhança leva em consideração os determinantes das matrizes postuladas sob H_0 e sob H_1 , o que não ocorre no teste proposto por Sullivan et al. (2007). Como o determinante das duas matrizes são bem diferentes, o teste da razão de verossimilhança conseguiu sinalizar com maior facilidade. Porém, à medida que aumentamos o tamanho amostral o poder de ambos os testes se aproximam. No caso do teste VMAX, como este é sensível apenas a mudanças nas variâncias das variáveis, o seu desempenho foi baixo para o modelo 3. Observe que para $n\geq 50$ alguns testes já atingem poder 1 (ver Tabela 6.32)

Finalmente, no modelo 4, alteramos ambos os parâmetros ϕ_{12} e ϕ_{21} o que ocasiona uma grande mudança na matriz de covariâncias. Para $n=50$, já obtemos valores elevados de poder para os testes em estudo (valores de poder acima de 0,70), exceto para o teste do determinante, que mesmo para $n=200$ apresentou um baixo poder, da ordem de 0,17. Isso é explicado pelo fato de que a razão entre os determinantes (Λ_{\det}) é próxima de 1 ($\Lambda_{\det} = 1,1999$). Os menores valores de poder ocorrerão para $n=25$, como era de se esperar, mas valores até expressivos considerando-se o tamanho da amostra (min $\approx 0,38$ e max $\approx 0,65$), com exceção do teste do determinante (valor do poder em torno de 0,08). Ainda considerando o modelo 4, para tamanhos amostrais maiores ($n\geq 100$) os valores de poder se aproximam de 1, exceto para o teste do determinante pelo fato descrito anteriormente.

Os resultados obtidos usando a aproximação para a distribuição $\chi^2_{7gl,\alpha}$, quando trabalhamos com a matriz $\hat{\Gamma}_0$, foram bem próximos daqueles obtidos com as distribuições exatas, tanto para o teste da razão de verossimilhança tanto para o teste Step-Down proposto por Sullivan et al. (2007).

Tabela 6.32: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 3 na situação 2

n	μ	Estimador	Testes									
			Razão V. Exato	Razão V. $X^2_{(7g,l,a)}$	Razão V. $X^2_{(9g,l,a)}$	Razão V. $X^2_{(9g,l,2a)}$	Sullivan Exato	Sullivan $X^2_{(7g,l,a)}$	Sullivan $X^2_{(9g,l,a)}$	Sullivan $X^2_{(9g,l,2a)}$	VMAX Exato	Determin. Exato
25	$\mu_{\text{desconhecido}}$	S_n	0.9859	-	-	-	0.7642	-	-	-	0.0452	0.9510
		$\hat{\Gamma}_0$	0.9852	0.9869	0.9809	0.9857	0.7134	0.8289	0.7192	0.8006	0.0546	0.9475
50	$\mu_{\text{desconhecido}}$	S_n	1,0000	-	-	-	1,0000	-	-	-	0.0495	0.9992
		$\hat{\Gamma}_0$	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0.9882	1,0000	0.0560	0.9994
100	$\mu_{\text{desconhecido}}$	S_n	1,0000	-	-	-	1,0000	-	-	-	0.0603	1,0000
		$\hat{\Gamma}_0$	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0.0579	1,0000
200	$\mu_{\text{desconhecido}}$	S_n	1,0000	-	-	-	1,0000	-	-	-	0.0747	1,0000
		$\hat{\Gamma}_0$	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0.0723	1,0000

Tabela 6.33: Estimativas do poder dos testes para matriz de covariâncias considerando o Modelo 4 na situação 2

n	μ	Estimador	Testes									
			Razão V. Exato	Razão V. $X^2_{(7g,l,a)}$	Razão V. $X^2_{(9g,l,a)}$	Razão V. $X^2_{(9g,l,2a)}$	Sullivan Exato	Sullivan $X^2_{(7g,l,a)}$	Sullivan $X^2_{(9g,l,a)}$	Sullivan $X^2_{(9g,l,2a)}$	VMAX Exato	Determin. Exato
25	$\mu_{\text{desconhecido}}$	S_n	0,5430	-	-	-	0,4209	-	-	-	0,4807	0,0657
		$\hat{\Gamma}_0$	0,6273	0,6515	0,5688	0,6335	0,3827	0,4644	0,3879	0,4450	0,4499	0,0895
50	$\mu_{\text{desconhecido}}$	S_n	0,8887	-	-	-	0,8379	-	-	-	0,7641	0,0922
		$\hat{\Gamma}_0$	0,9210	0,9101	0,8568	0,8998	0,7997	0,8129	0,7387	0,7968	0,7490	0,1001
100	$\mu_{\text{desconhecido}}$	S_n	0,9971	-	-	-	0,9951	-	-	-	0,9570	0,1238
		$\hat{\Gamma}_0$	0,9984	0,9978	0,9945	0,9972	0,9929	0,9924	0,9831	0,9907	0,9541	0,1250
200	$\mu_{\text{desconhecido}}$	S_n	1,0000	-	-	-	1,0000	-	-	-	0,9987	0,1704
		$\hat{\Gamma}_0$	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9985	0,1708

Capítulo 7

Conclusões Finais e Sugestões para Trabalhos Futuros

Com base nas simulações desenvolvidas foi possível observar que se estamos em um processo autocorrelacionado é de extrema importância levarmos em consideração tal fato. Se a autocorrelação está presente nas observações, mas que por desconhecimento essa autocorrelação é desconsiderada, ou seja, assumimos independência entre as observações, e usamos a matriz de covariâncias amostral S , juntamente com os testes apresentados no Capítulo 3 desta dissertação, e com as aproximações para distribuição qui-quadrado para verificar possíveis alterações na matriz postulada sob H_0 , teremos conclusões incorretas.

Diante disso, quando verificamos autocorrelação entre as observações de um processo vimos que temos duas alternativas de análise. A primeira delas é trabalhar com a matriz de covariâncias amostral S . Porém, essa escolha pode ser mais trabalhosa, visto que quando optamos por utilizar o teste estatístico com a matriz S sempre teremos que usar as distribuições exatas dos testes. Se por um lado, ao usarmos a matriz de covariâncias amostral S , temos um ganho no sentido de não termos que ajustar um modelo de séries temporais às observações amostrais para posteriormente obter a matriz $\hat{\Gamma}_0$, de outro sempre que optarmos por trabalhar com a matriz S teremos que usar a distribuição exata. A segunda opção é trabalhar com a matriz $\hat{\Gamma}_0$. Para isso, um dos primeiros passos é identificar e ajustar um modelo de séries temporais às observações levando em consideração a existência de tendência, sazonalidade, valores discrepantes (*outliers*), descontinuidade, etc. Embora seja trabalhoso ajustar um modelo de séries temporais ao conjunto de observações, temos um ganho ao optarmos pelo uso da matriz $\hat{\Gamma}_0$. Através das simulações desenvolvidas vimos que em alguns testes, quando usamos a matriz $\hat{\Gamma}_0$, é possível utilizar a distribuição assintótica qui-quadrado para a estatística de teste, ou seja, não há necessidade de se trabalhar sempre com a distribuição exata como ocorre no caso onde usamos a matriz de covariância amostral S . Para o teste da razão de

verossimilhança e para o teste Step-Down proposto por Sullivan et al. (2007) vimos que em alguns contextos é possível usar a distribuição qui-quadrado para determinar os limites críticos dos testes.

Durante as simulações verificamos que o poder dos testes quando usamos a matriz $\hat{\Gamma}_0$ e a matriz S foram bem próximos quando se utiliza a distribuição exata. Não houve um predomínio de um caso em relação ao outro. Em alguns momentos o uso dos testes propostos considerando a matriz S forneceram poderes superiores aos casos em que usávamos a matriz $\hat{\Gamma}_0$. Porém, em outras situações verificamos o contrário. Embora o uso da matriz S tenha fornecido bons resultados, vale lembrar que para esse caso sempre teremos que trabalhar com a distribuição exata, como discutido anteriormente.

Com respeito a qual dos testes usados teve um melhor desempenho podemos destacar o teste da razão de verossimilhança e o teste Step-Down proposto por Sullivan et al. (2007). Não houve uma superioridade predominante de um desses testes em relação a outro. Já o teste do determinante e o teste VMAX não tiveram um desempenho muito satisfatório nas situações em que a matriz de covariâncias sob a hipótese alternativa não apresentava uma alteração acentuada em relação a H_0 . De uma forma geral, à medida que aumentamos o tamanho amostral o poder dos testes estudados aumenta. É importante lembrar que o bom desempenho do teste da razão de verossimilhança e do proposto por Sullivan et al. (2007) para dados autocorrelacionados, observados nessa dissertação, concorda com o que foi apresentado em Pinto (2009b), que mostra que esses dois testes são os melhores para dados independentes.

Como discutido no Capítulo 3, vimos que o uso do determinante como uma medida de dispersão multivariada tem algumas limitações. Em geral, algumas mudanças na estrutura da matriz de covariâncias acarretam mudanças na variância generalizada. Vale ressaltar que o determinante é uma representação escalar relativamente simplista de um problema multivariado complexo. O desempenho do teste do determinante só foi satisfatório nas situações nas quais o determinante da matriz postulada sob H_0 e daquela sob H_1 eram completamente diferentes. Nas situações onde tínhamos pequenas alterações nos determinantes das matrizes, o teste teve um baixo desempenho. No mesmo patamar do teste do determinante temos o teste VMAX. Como o teste VMAX capta alterações somente nas variâncias das variáveis, esse só teve um bom desempenho quando as alterações na matriz de covariâncias ocorriam nas variâncias. Nessas situações o teste VMAX conseguiu valores de poder mais elevados. De uma forma geral, em todos os modelos estudados, quando a alteração na matriz de covariâncias era pequena, todos os

testes tiveram dificuldades em detectar tal alteração, porém, o teste proposto por Sullivan et al. (2007) e o da razão de verossimilhança tiveram valores de poder maiores que os outros testes mesmo nessas situações, sendo que o aumento do tamanho amostral eleva a capacidade desses testes em detectar a não veracidade nesses casos.

Nessa dissertação nosso foco de estudo foi o modelo autorregressivo multivariado de ordem 1, VAR(1), sempre considerando o caso onde tínhamos duas variáveis ($p=2$). Embora o modelo VAR(1) seja um dos mais importantes modelos utilizados na prática sugerimos que os processos autocorrelacionados sejam mais bem estudados nos casos em que se tem autocorrelação de ordem maior que 1 (VAR(k), $k>1$). Devido ao tempo estipulado para o desenvolvimento desta dissertação não foi possível abordar outros modelos além do VAR(1). Diante disso, sugerimos também que o desempenho dos testes quando estamos em um processo do tipo VARMA(k,q) sejam investigados, embora esses modelos sejam menos usuais.

Um outro fator importante é que a correção dos graus de liberdade ou do nível de significância de α para $\alpha/2$ foi feita para $p=2$ e para a hipótese nula previamente postulada. Diante disso, uma outra proposta de estudo seria mudar a hipótese nula, ou o número de variáveis (p), para verificar se essas correções também se mantêm apropriadas nesses casos. O ideal é realizar um estudo matemático (ou via simulação) mais detalhado para se obter melhores aproximações para a distribuição das estatísticas dos testes.

Finalmente, em controle de qualidade, sabemos que atualmente é necessário acompanhar as condições do processo por meio do monitoramento simultâneo de muitas características ($p>2$) de qualidade. Diante disso seria um grande ganho estender os esses estudos dessa dissertação para os casos em que $p>2$.

Apêndice

Apêndice A: Cálculo da matriz de covariância cruzada para um processo bivariado VAR(1) com

$$\phi_{12} = \phi_{21}$$

Com o intuito de se obter os componentes γ_{ij} da matriz Γ_0 , de um processo bivariado VAR(1), para o caso em que $\phi_{12} = \phi_{21} \neq 0$, desenvolvemos as equações propostas por Kalgonda e Kulkarni (2004). Primeiramente temos as seguintes informações:

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \quad \Phi = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}$$

Vamos considerar o caso particular onde $\phi_{12} = \phi_{21} = \rho$. Portanto, $\Phi = \begin{pmatrix} \phi_1 & \rho \\ \rho & \phi_2 \end{pmatrix}$.

Assim, por Kalgonda e Kulkarni (2004) temos que a matriz de covariância cruzada para um modelo VAR(1) é dada por:

$$\Gamma(0) = \Phi \Gamma(0) \Phi' + \Sigma$$

$$\begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} \phi_1 & \rho \\ \rho & \phi_2 \end{pmatrix} \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} \phi_1 & \rho \\ \rho & \phi_2 \end{pmatrix} + \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

$$\begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} \phi_1 \gamma_{11} + \rho \gamma_{21} & \phi_1 \gamma_{12} + \rho \gamma_{22} \\ \rho \gamma_{11} + \phi_2 \gamma_{21} & \rho \gamma_{12} + \phi_2 \gamma_{22} \end{pmatrix} \begin{pmatrix} \phi_1 & \rho \\ \rho & \phi_2 \end{pmatrix} + \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

$$\begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} \phi_1^2 \gamma_{11} + \phi_1 \rho \gamma_{21} + \phi_1 \rho \gamma_{12} + \rho^2 \gamma_{22} & \phi_1 \rho \gamma_{11} + \rho^2 \gamma_{21} + \phi_1 \phi_2 \gamma_{12} + \phi_2 \rho \gamma_{22} \\ \phi_1 \rho \gamma_{11} + \phi_1 \phi_2 \gamma_{21} + \rho^2 \gamma_{12} + \phi_2 \rho \gamma_{22} & \rho^2 \gamma_{11} + \phi_2 \rho \gamma_{21} + \phi_2 \rho \gamma_{12} + \phi_2^2 \gamma_{22} \end{pmatrix} + \Sigma$$

$$\begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} \phi_1^2 \gamma_{11} + \phi_1 \rho \gamma_{21} + \phi_1 \rho \gamma_{12} + \rho^2 \gamma_{22} + \sigma_{11} & \phi_1 \rho \gamma_{11} + \rho^2 \gamma_{21} + \phi_1 \phi_2 \gamma_{12} + \phi_2 \rho \gamma_{22} + \sigma_{12} \\ \phi_1 \rho \gamma_{11} + \phi_1 \phi_2 \gamma_{21} + \rho^2 \gamma_{12} + \phi_2 \rho \gamma_{22} + \sigma_{21} & \rho^2 \gamma_{11} + \phi_2 \rho \gamma_{21} + \phi_2 \rho \gamma_{12} + \phi_2^2 \gamma_{22} + \sigma_{22} \end{pmatrix}$$

Observe que $\gamma_{21} = \gamma_{12}$, por isso vamos trabalhar somente com uma das duas variáveis (γ_{12}):

$$\gamma_{11} = \phi_1^2 \gamma_{11} + \phi_1 \rho \gamma_{21} + \phi_1 \rho \gamma_{12} + \rho^2 \gamma_{22} + \sigma_{11}$$

$$\gamma_{11} = \phi_1^2 \gamma_{11} + \phi_1 \rho \gamma_{12} + \phi_1 \rho \gamma_{12} + \rho^2 \gamma_{22} + \sigma_{11}$$

$$\gamma_{11} = \phi_1^2 \gamma_{11} + 2\phi_1 \rho \gamma_{12} + \rho^2 \gamma_{22} + \sigma_{11}$$

$$(1 - \phi_1^2) \gamma_{11} - 2\phi_1 \rho \gamma_{12} - \rho^2 \gamma_{22} = \sigma_{11}$$

$$\gamma_{12} = \phi_1 \rho \gamma_{11} + \rho^2 \gamma_{21} + \phi_1 \phi_2 \gamma_{12} + \phi_2 \rho \gamma_{22} + \sigma_{12}$$

$$\gamma_{12} = \phi_1 \rho \gamma_{11} + \rho^2 \gamma_{12} + \phi_1 \phi_2 \gamma_{12} + \phi_2 \rho \gamma_{22} + \sigma_{12}$$

$$\gamma_{12} = \phi_1 \rho \gamma_{11} + (\rho^2 + \phi_1 \phi_2) \gamma_{12} + \phi_2 \rho \gamma_{22} + \sigma_{12}$$

$$-\phi_1 \rho \gamma_{11} + (1 - \phi_1 \phi_2 - \rho^2) \gamma_{12} - \phi_2 \rho \gamma_{22} = \sigma_{12}$$

Finalmente temos;

$$\gamma_{22} = \rho^2 \gamma_{11} + \phi_2 \rho \gamma_{21} + \phi_2 \rho \gamma_{12} + \phi_2^2 \gamma_{22} + \sigma_{22}$$

$$\gamma_{22} = \rho^2 \gamma_{11} + \phi_2 \rho \gamma_{12} + \phi_2 \rho \gamma_{12} + \phi_2^2 \gamma_{22} + \sigma_{22}$$

$$\gamma_{22} = \rho^2 \gamma_{11} + 2\phi_2 \rho \gamma_{12} + \phi_2^2 \gamma_{22} + \sigma_{22}$$

$$-\rho^2 \gamma_{11} - 2\phi_2 \rho \gamma_{12} + (1 - \phi_2^2) \gamma_{22} = \sigma_{22}$$

De posse das equações anteriores temos a seguinte representação matricial:

$$\begin{pmatrix} 1 - \phi_1^2 & -2\phi_1 \rho & -\rho^2 \\ -\phi_1 \rho & 1 - \phi_1 \phi_2 - \rho^2 & -\phi_2 \rho \\ -\rho^2 & -2\phi_2 \rho & 1 - \phi_2^2 \end{pmatrix} \begin{pmatrix} \gamma_{11} \\ \gamma_{12} \\ \gamma_{22} \end{pmatrix} = \begin{pmatrix} \sigma_{11} \\ \sigma_{12} \\ \sigma_{22} \end{pmatrix} \quad \text{onde} \quad \begin{array}{l} -1 \leq \phi_1 \leq 1 \\ -1 \leq \phi_2 \leq 1 \\ -1 \leq \rho \leq 1 \\ \sigma_{ij} \geq 0, \forall i, j = 1, 2 \end{array}$$

Considere a seguinte matriz $A = \begin{pmatrix} 1 - \phi_1^2 & -2\phi_1\rho & -\rho^2 \\ -\phi_1\rho & 1 - \phi_1\phi_2 - \rho^2 & -\phi_2\rho \\ -\rho^2 & -2\phi_2\rho & 1 - \phi_2^2 \end{pmatrix}$.

Se $\det(A) \neq 0$ então:

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$$

onde denotamos $\text{adj}(A)$, como a transposta da matriz formada pelos cofatores de A. O próximo passo será encontrar a matriz de cofatores \tilde{A} .

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \tilde{a}_{13} \\ \tilde{a}_{21} & \tilde{a}_{22} & \tilde{a}_{23} \\ \tilde{a}_{31} & \tilde{a}_{32} & \tilde{a}_{33} \end{pmatrix}$$

$$\tilde{a}_{11} = (-1)^{1+1} \det \begin{pmatrix} 1 - \phi_1\phi_2 - \rho^2 & -\phi_2\rho \\ -2\phi_2\rho & 1 - \phi_2^2 \end{pmatrix} = (1 - \phi_1\phi_2 - \rho^2)(1 - \phi_2^2) - 2\phi_2^2\rho^2$$

$$\tilde{a}_{12} = (-1)^{1+2} \det \begin{pmatrix} -\phi_1\rho & -\phi_2\rho \\ -\rho^2 & 1 - \phi_2^2 \end{pmatrix} = (\phi_1\rho)(1 - \phi_2^2) + \phi_2\rho^3$$

$$\tilde{a}_{13} = (-1)^{1+3} \det \begin{pmatrix} -\phi_1\rho & 1 - \phi_1\phi_2 - \rho^2 \\ -\rho^2 & -2\phi_2\rho \end{pmatrix} = 2\phi_1\phi_2\rho^2 + (1 - \phi_1\phi_2 - \rho^2)\rho^2 = (1 + \phi_1\phi_2 - \rho^2)\rho^2$$

$$\tilde{a}_{21} = (-1)^{2+1} \det \begin{pmatrix} -2\phi_1\rho & -\rho^2 \\ -2\phi_2\rho & 1 - \phi_2^2 \end{pmatrix} = 2\phi_1\rho(1 - \phi_2^2) + 2\phi_2\rho^3 = 2\rho(\phi_1(1 - \phi_2^2) + \phi_2\rho^2)$$

$$\tilde{a}_{22} = (-1)^{2+2} \det \begin{pmatrix} 1 - \phi_1^2 & -\rho^2 \\ -\rho^2 & 1 - \phi_2^2 \end{pmatrix} = (1 - \phi_1^2)(1 - \phi_2^2) - \rho^4$$

$$\tilde{a}_{23} = (-1)^{2+3} \det \begin{pmatrix} 1-\phi_1^2 & -2\phi_1\rho \\ -\rho^2 & -2\phi_2\rho \end{pmatrix} = 2\rho(\phi_2(1-\phi_1^2) + \phi_1\rho^2)$$

$$\tilde{a}_{31} = (-1)^{3+1} \det \begin{pmatrix} -2\phi_1\rho & -\rho^2 \\ 1-\phi_1\phi_2 - \rho^2 & -\phi_2\rho \end{pmatrix} = 2\phi_1\phi_2\rho^2 + (1-\phi_1\phi_2 - \rho^2)\rho^2 = (1+\phi_1\phi_2 - \rho^2)\rho^2$$

$$\tilde{a}_{32} = (-1)^{3+2} \det \begin{pmatrix} 1-\phi_1^2 & -\rho^2 \\ -\phi_1\rho & -\phi_2\rho \end{pmatrix} = (\phi_2\rho)(1-\phi_1^2) + \phi_1\rho^3$$

$$\tilde{a}_{33} = (-1)^{3+3} \det \begin{pmatrix} 1-\phi_1^2 & -2\phi_1\rho \\ -\phi_1\rho & 1-\phi_1\phi_2 - \rho^2 \end{pmatrix} = (1-\phi_1\phi_2 - \rho^2)(1-\phi_1^2) - 2\phi_1^2\rho^2$$

Com isso, a matriz de cofatores \tilde{A} é dada por:

$$\tilde{A} = \begin{pmatrix} (1-\phi_1\phi_2 - \rho^2)(1-\phi_2^2) - 2\phi_2^2\rho^2 & (\phi_1\rho)(1-\phi_2^2) + \phi_2\rho^3 & (1+\phi_1\phi_2 - \rho^2)\rho^2 \\ 2\rho(\phi_1(1-\phi_2^2) + \phi_2\rho^2) & (1-\phi_1^2)(1-\phi_2^2) - \rho^4 & 2\rho(\phi_2(1-\phi_1^2) + \phi_1\rho^2) \\ (1+\phi_1\phi_2 - \rho^2)\rho^2 & (\phi_2\rho)(1-\phi_1^2) + \phi_1\rho^3 & (1-\phi_1\phi_2 - \rho^2)(1-\phi_1^2) - 2\phi_1^2\rho^2 \end{pmatrix}$$

Portanto, pela definição de matriz adjunta temos que:

$$adj(A) = \begin{pmatrix} (1-\phi_1\phi_2 - \rho^2)(1-\phi_2^2) - 2\phi_2^2\rho^2 & 2\rho(\phi_1(1-\phi_2^2) + \phi_2\rho^2) & (1+\phi_1\phi_2 - \rho^2)\rho^2 \\ (\phi_1\rho)(1-\phi_2^2) + \phi_2\rho^3 & (1-\phi_1^2)(1-\phi_2^2) - \rho^4 & (\phi_2\rho)(1-\phi_1^2) + \phi_1\rho^3 \\ (1+\phi_1\phi_2 - \rho^2)\rho^2 & 2\rho(\phi_2(1-\phi_1^2) + \phi_1\rho^2) & (1-\phi_1\phi_2 - \rho^2)(1-\phi_1^2) - 2\phi_1^2\rho^2 \end{pmatrix}$$

Lembrando que $A^{-1} = \frac{1}{\det(A)} adj(A)$ o próximo passo será encontrar o determinante da

matriz A. Assim, vamos calcular o determinante de A utilizando o desenvolvimento em cofatores do determinante de A em termos da 1ª linha.

$$\det(A) = a_{11} \tilde{a}_{11} + a_{12} \tilde{a}_{12} + \dots + a_{1n} \tilde{a}_{1n} = \sum_{j=1}^n a_{1j} \tilde{a}_{1j}$$

Como já temos os cofatores de A o determinante é dado pela seguinte forma:

$$\det(A) = (1 - \phi_1^2)[(1 - \phi_1\phi_2 - \rho^2)(1 - \phi_2^2) - 2\phi_2^2\rho^2] - 2\phi_1\rho[(\phi_1\rho)(1 - \phi_2^2) + \phi_2\rho^3] - \rho^2(1 + \phi_1\phi_2 - \rho^2)\rho^2$$

$$\det(A) = (1 - \phi_1^2)[(1 - \phi_1\phi_2 - \rho^2)(1 - \phi_2^2) - 2\phi_2^2\rho^2] - 2\phi_1^2\rho^2(1 - \phi_2^2) - 2\phi_1\phi_2\rho^4 - (1 + \phi_1\phi_2 - \rho^2)\rho^4$$

$$\det(A) = (1 - \phi_1^2)[(1 - \phi_1\phi_2 - \rho^2)(1 - \phi_2^2) - 2\phi_2^2\rho^2] - 2\phi_1^2\rho^2(1 - \phi_2^2) - (1 + 3\phi_1\phi_2 - \rho^2)\rho^4$$

Dessa forma;

$$\det(A) = \underbrace{(1 - \phi_1^2)[(1 - \phi_1\phi_2 - \rho^2)(1 - \phi_2^2) - 2\phi_2^2\rho^2] - 2\phi_1^2\rho^2(1 - \phi_2^2) - (1 + 3\phi_1\phi_2 - \rho^2)\rho^4}_k$$

$$A^{-1} = \frac{1}{K} \begin{pmatrix} (1 - \phi_1\phi_2 - \rho^2)(1 - \phi_2^2) - 2\phi_2^2\rho^2 & 2\rho(\phi_1(1 - \phi_2^2) + \phi_2\rho^2) & (1 + \phi_1\phi_2 - \rho^2)\rho^2 \\ (\phi_1\rho)(1 - \phi_2^2) + \phi_2\rho^3 & (1 - \phi_1^2)(1 - \phi_2^2) - \rho^4 & (\phi_2\rho)(1 - \phi_1^2) + \phi_1\rho^3 \\ (1 + \phi_1\phi_2 - \rho^2)\rho^2 & 2\rho(\phi_2(1 - \phi_1^2) + \phi_1\rho^2) & (1 - \phi_1\phi_2 - \rho^2)(1 - \phi_1^2) - 2\phi_1^2\rho^2 \end{pmatrix}$$

Inicialmente tínhamos o seguinte sistema;

$$\begin{pmatrix} 1 - \phi_1^2 & -2\phi_1\rho & -\rho^2 \\ -\phi_1\rho & 1 - \phi_1\phi_2 - \rho^2 & -\phi_2\rho \\ -\rho^2 & -2\phi_2\rho & 1 - \phi_2^2 \end{pmatrix} \begin{pmatrix} \gamma_{11} \\ \gamma_{12} \\ \gamma_{22} \end{pmatrix} = \begin{pmatrix} \sigma_{11} \\ \sigma_{12} \\ \sigma_{22} \end{pmatrix}$$

Assim a solução do sistema é dada por:

$$\begin{pmatrix} \gamma_{11} \\ \gamma_{12} \\ \gamma_{22} \end{pmatrix} = A^{-1} \begin{pmatrix} \sigma_{11} \\ \sigma_{12} \\ \sigma_{22} \end{pmatrix}$$

$$\begin{pmatrix} \gamma_{11} \\ \gamma_{12} \\ \gamma_{22} \end{pmatrix} = \frac{1}{K} \begin{pmatrix} [(1 - \phi_1\phi_2 - \rho^2)(1 - \phi_2^2) - 2\phi_2^2\rho^2]\sigma_{11} + [2\rho(\phi_1(1 - \phi_2^2) + \phi_2\rho^2)]\sigma_{12} + [(1 + \phi_1\phi_2 - \rho^2)\rho^2]\sigma_{22} \\ [(\phi_1\rho)(1 - \phi_2^2) + \phi_2\rho^3]\sigma_{11} + [(1 - \phi_1^2)(1 - \phi_2^2) - \rho^4]\sigma_{12} + [(\phi_2\rho)(1 - \phi_1^2) + \phi_1\rho^3]\sigma_{22} \\ [(1 + \phi_1\phi_2 - \rho^2)\rho^2]\sigma_{11} + [2\rho(\phi_2(1 - \phi_1^2) + \phi_1\rho^2)]\sigma_{12} + [(1 - \phi_1\phi_2 - \rho^2)(1 - \phi_1^2) - 2\phi_1^2\rho^2]\sigma_{22} \end{pmatrix}$$

Finalmente;

$$\gamma_{11} = \frac{[(1-\phi_1\phi_2 - \rho^2)(1-\phi_2^2) - 2\phi_2^2\rho^2]\sigma_{11} + [2\rho(\phi_1(1-\phi_2^2) + \phi_2\rho^2)]\sigma_{12} + [(1+\phi_1\phi_2 - \rho^2)\rho^2]\sigma_{22}}{(1-\phi_1^2)[(1-\phi_1\phi_2 - \rho^2)(1-\phi_2^2) - 2\phi_2^2\rho^2] - 2\phi_1^2\rho^2(1-\phi_2^2) - (1+3\phi_1\phi_2 - \rho^2)\rho^4}$$

$$\gamma_{12} = \gamma_{21} = \frac{[(\phi_1\rho)(1-\phi_2^2) + \phi_2\rho^3]\sigma_{11} + [(1-\phi_1^2)(1-\phi_2^2) - \rho^4]\sigma_{12} + [(\phi_2\rho)(1-\phi_1^2) + \phi_1\rho^3]\sigma_{22}}{(1-\phi_1^2)[(1-\phi_1\phi_2 - \rho^2)(1-\phi_2^2) - 2\phi_2^2\rho^2] - 2\phi_1^2\rho^2(1-\phi_2^2) - (1+3\phi_1\phi_2 - \rho^2)\rho^4}$$

$$\gamma_{22} = \frac{[(1+\phi_1\phi_2 - \rho^2)\rho^2]\sigma_{11} + [2\rho(\phi_2(1-\phi_1^2) + \phi_1\rho^2)]\sigma_{12} + [(1-\phi_1\phi_2 - \rho^2)(1-\phi_1^2) - 2\phi_1^2\rho^2]\sigma_{22}}{(1-\phi_1^2)[(1-\phi_1\phi_2 - \rho^2)(1-\phi_2^2) - 2\phi_2^2\rho^2] - 2\phi_1^2\rho^2(1-\phi_2^2) - (1+3\phi_1\phi_2 - \rho^2)\rho^4}$$

Apêndice B: Distribuições Exatas – Situação 1

Tabela B.1: Percentis da distribuição exata - Situação 1 - μ conhecido

n	Teste	Estimadores							
		S_{μ}				$\hat{\Gamma}_0$			
		2,5%	5,0%	95,0%	97,5%	2,5%	5,0%	95,0%	97,5%
25	Determinante	0,498	0,601	4,089	4,874	0,501	0,609	4,840	5,944
	Razão de V.	0,474	0,754	15,878	19,203	0,932	2,070	13,820	16,842
	Sullivan	0,465	0,741	16,095	20,312	0,369	0,602	15,086	19,542
	VMAX	0,587	0,650	2,008	2,257	0,585	0,649	2,022	2,273
50	Determinante	0,775	0,884	3,549	4,027	0,772	0,887	3,922	4,520
	Razão de V.	0,481	0,762	16,678	20,114	0,718	1,729	14,650	17,850
	Sullivan	0,461	0,753	16,710	20,858	0,367	0,594	15,629	19,763
	VMAX	0,692	0,745	1,689	1,843	0,692	0,745	1,692	1,843
100	Determinante	1,046	1,151	3,110	3,413	1,029	1,141	3,309	3,689
	Razão de V.	0,459	0,747	17,093	20,573	0,650	1,654	15,087	18,397
	Sullivan	0,460	0,760	17,185	21,002	0,374	0,608	16,086	20,049
	VMAX	0,771	0,814	1,471	1,565	0,771	0,814	1,469	1,565
200	Determinante	1,279	1,369	2,781	2,967	1,266	1,358	2,900	3,111
	Razão de V.	0,462	0,757	17,414	21,088	0,639	1,627	15,359	18,895
	Sullivan	0,447	0,734	17,440	21,214	0,362	0,597	16,389	20,388
	VMAX	0,835	0,868	1,326	1,387	0,835	0,868	1,326	1,387
1000	Determinante	1,666	1,717	2,359	2,432	1,647	1,699	2,388	2,468
	Razão de V.	0,488	0,790	17,957	21,591	0,501	1,404	15,570	19,214
	Sullivan	0,468	0,760	17,741	21,403	0,371	0,604	16,666	20,648
	VMAX	0,925	0,941	1,139	1,163	0,925	0,941	1,139	1,163
5000	Determinante	1,857	1,882	2,170	2,201	1,846	1,873	2,182	2,215
	Razão de V.	0,490	0,749	17,545	21,241	0,576	1,610	15,675	19,552
	Sullivan	0,467	0,759	17,743	21,391	0,370	0,603	16,721	20,656
	VMAX	0,966	0,973	1,061	1,071	0,966	0,973	1,061	1,071

Tabela B.2: Percentis da distribuição exata - Situação 1 – μ desconhecido

n	Teste	Estimadores							
		S_n				$\hat{\Gamma}_0$			
		2,5%	5,0%	95,0%	97,5%	2,5%	5,0%	95,0%	97,5%
25	Determinante	0,387	0,465	3,145	3,748	0,355	0,431	3,410	4,185
	Razão de V.	0,583	0,945	16,523	19,219	0,000	0,000	15,222	17,887
	Sullivan	0,582	0,946	15,109	17,833	0,447	0,719	14,122	16,696
	VMAX	0,522	0,577	1,742	1,939	0,496	0,551	1,699	1,906
50	Determinante	0,681	0,779	3,109	3,542	0,652	0,748	3,286	3,802
	Razão de V.	0,511	0,838	17,153	20,296	0,000	0,000	15,585	18,649
	Sullivan	0,511	0,848	16,330	19,567	0,398	0,659	15,193	18,227
	VMAX	0,651	0,702	1,569	1,708	0,636	0,686	1,546	1,682
100	Determinante	0,974	1,075	2,913	3,206	0,942	1,045	3,026	3,364
	Razão de V.	0,499	0,814	17,456	20,853	0,000	0,000	15,656	18,958
	Sullivan	0,497	0,800	17,048	20,446	0,394	0,640	15,821	19,214
	VMAX	0,749	0,789	1,418	1,508	0,741	0,781	1,405	1,495
200	Determinante	1,238	1,328	2,693	2,886	1,210	1,300	2,770	2,979
	Razão de V.	0,477	0,774	17,606	21,137	0,000	0,000	15,642	19,134
	Sullivan	0,464	0,767	17,388	20,885	0,376	0,618	16,271	19,994
	VMAX	0,822	0,855	1,301	1,360	0,818	0,851	1,294	1,353
1000	Determinante	1,652	1,703	2,339	2,412	1,632	1,684	2,367	2,448
	Razão de V.	0,465	0,754	17,553	21,275	0,000	0,000	15,537	19,027
	Sullivan	0,474	0,769	17,741	21,344	0,377	0,609	16,634	20,542
	VMAX	0,922	0,938	1,134	1,158	0,921	0,937	1,133	1,157
5000	Determinante	1,854	1,879	2,167	2,197	1,843	1,870	2,178	2,211
	Razão de V.	0,473	0,765	17,672	21,402	0,000	0,000	15,598	19,197
	Sullivan	0,465	0,760	17,734	21,379	0,372	0,602	16,722	20,598
	VMAX	0,965	0,973	1,060	1,070	0,965	0,972	1,060	1,070

Apêndice C: Distribuições Exatas – Situação 2

Tabela C.1: Percentis da distribuição exata - Situação 2 – μ desconhecido

n	Teste	Estimadores							
		S_n				$\hat{\Gamma}_0$			
		2,5%	5,0%	95,0%	97,5%	2,5%	5,0%	95,0%	97,5%
25	Determinante	0,835	0,990	5,054	5,892	0,976	1,167	6,260	7,327
	Razão de V.	0,316	0,520	11,602	14,003	0,000	0,000	14,928	19,113
	Sullivan	0,318	0,521	12,874	16,195	0,357	0,576	17,137	22,023
	VMAX	0,623	0,682	1,848	2,063	0,672	0,738	2,103	2,396
50	Determinante	1,199	1,343	4,257	4,726	1,298	1,460	4,681	5,186
	Razão de V.	0,323	0,531	11,962	14,580	0,000	0,094	13,407	16,399
	Sullivan	0,322	0,532	12,689	15,869	0,335	0,546	14,567	18,445
	VMAX	0,718	0,767	1,596	1,737	0,747	0,798	1,683	1,839
100	Determinante	1,492	1,612	3,719	4,021	1,553	1,678	3,889	4,206
	Razão de V.	0,322	0,525	12,527	15,222	0,130	0,387	13,164	16,073
	Sullivan	0,323	0,525	12,982	15,992	0,330	0,534	13,828	17,161
	VMAX	0,788	0,826	1,411	1,504	0,802	0,841	1,445	1,542
200	Determinante	1,716	1,817	3,375	3,579	1,749	1,854	3,448	3,658
	Razão de V.	0,341	0,553	13,238	16,145	0,280	0,515	13,568	16,617
	Sullivan	0,342	0,552	13,393	16,568	0,340	0,558	13,873	17,147
	VMAX	0,840	0,869	1,291	1,350	0,848	0,878	1,307	1,367
1000	Determinante	2,057	2,119	2,938	3,030	2,065	2,127	2,950	3,043
	Razão de V.	0,412	0,661	16,804	20,406	0,418	0,672	16,919	20,543
	Sullivan	0,412	0,660	16,806	20,407	0,413	0,662	16,857	20,523
	VMAX	0,916	0,932	1,140	1,167	0,918	0,934	1,143	1,170
5000	Determinante	2,291	2,325	2,712	2,751	2,293	2,326	2,714	2,753
	Razão de V.	0,437	0,717	18,289	22,442	0,441	0,719	18,316	22,480
	Sullivan	0,437	0,717	18,267	22,481	0,438	0,716	18,277	22,520
	VMAX	0,960	0,968	1,065	1,076	0,961	0,968	1,065	1,077

Referências Bibliográficas

- ALT, F. B. Multivariate quality control. *Encyclopedia of Statistical Sciences*. Willey, 1985.
- ANDERSON, T. W. *A Introduction to Multivariate Statistical Analysis*, John Wiley, 1958, 374 p.
- ANDERSON, T. W. *A Introduction to Multivariate Statistical Analysis*, John Wiley, 2003, 721 p.
- BARBOSA, S. M. *mAr: Multivariate AutoRegressive analysis*. R package version 1.1-1, 2009.
- BEYER, W. H. *Standard mathematical tables*. West Palm Beach: CRC, 1978.
- BOX, G. P., LUCENO, A. *Statistical Control by Monitoring and Feedback Adjustment*. New York: John Wiley & Sons, 1997.
- BROCKWELL, P. J, DAVIS, R. A. *Introduction to time series and forecasting*. New York: Springer, 1996, 420 p.
- CASELLA, G., BERGER, R. L. *Statistical inference*. 2a edição. Australia; Thomson Learning, 2002, 660 p.
- CHANG, S. I, ZHANG, K. Statistical process control for variance shift detections of multivariate autocorrelated processes. *Quality Technology of Quantitative Management*, vol. 4, pp 413-435, 2007.
- COSTA, A. F. B., EPPRECHT E. K., CARPINETTI, L. C. R. *Controle Estatístico de Qualidade*, 2a edição, São Paulo: Editora Atlas, 2005, 334 p.
- DJAUHARI, M. A. Improved Monitoring of Multivariate Process Variability, *Journal of Quality Technology*, vol. 39, No. 1, p. 32 – 39, 2005.
- FERREIRA, D. F. *Estatística Multivariada*. 1a. edição. Lavras: Editora UFLA, 2008, 661 p.
- HAYTER, A. J., TSUI, K. L. Identification and Quantification in Multivariate Quality Control Problems. *Journal of Applied Statistics*, vol. 26, pp 197-208, 1994.

HOTELLING, H. Multivariate Quality Control. *Techniques of statistical analysis*. New York, Mc Graw Hill, pp 111-184, 1947.

JOBSON, J. D. *Applied multivariate data analysis*. New York: Springer, vol. 2, 1992, 731 p.

KALGONDA, A. A., KULKARNI, S.R. Multivariate Control Chart for Autocorrelated Process, *Journal of Applied Statistics*, vol. 31, pp 317-327, 2004.

LU, C. W., REYNOLDS, M. R. Jr. Control charts for monitoring the mean and variance of autocorrelated process. *Journal of Quality Technology*, p. 259-273, 1991.

LUTKEPOHL, H. *Introduction to Multiple Time Series Analysis*. Springer-Verlag, 1993, 764 p.

MACHADO, M. A. G., MAGALHÃES, M. S., COSTA, A. F. B. Gráfico de controle de VMAX para o monitoramento da matriz de covariâncias. *Produção*, vol.18, nº 2, p. 222-239, 2008.

MASON, R. L., YOUNG, J. C. Multivariate statistical process control with industrial applications. Philadelphia: ASA, 2002, 263 p.

MINGOTI, S. A. *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Belo Horizonte: Editora UFMG, 2005. 295 p.

MONTGOMERY, D.C. *Introdução ao Controle Estatístico da Qualidade*. 4a. edição. Rio de Janeiro: LTC, 2004, 513 P.

MONTGOMERY, D.C. WADSWORTH, H. M. Some techniques for multivariate quality control applications. *ASQC Technical Conference Transactions*, Washington, DC, 1972, 427-435 pp.

MORETTIN, P. A., TOLOI, C. M. C. *Análise de séries temporais*. 2a. edição. São Paulo: Editora Blücher, 2006. 538p.

MULTIVARIATE NORMAL DISTRIBUTION. In Wikipédia: a enciclopédia livre. Wikimedia. Disponível em: <http://en.wikipedia.org/w/index.php?title=Multivariate_normal_distribution&oldid=287365281> Acesso em: 16 de maio de 2009.

NEUMAIER, A., SCHNEIDER, T. Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Transactions on Mathematical Software*, vol. 27, nº1, 27-57 pp, 2001.

OLIVEIRA, F. L. P. *Índices de capacidade para processos multivariados autocorrelacionados*. Belo Horizonte: Departamento de Estatística, Universidade Federal de Minas Gerais, 2007, Dissertação (mestrado).

PEREIRA, F. H., MINGOTI, S. A. *Estatística VMAX via Simulação de Monte Carlo*. Caderno de Resumo do 8º MGEST, Universidade Federal de Juiz de Fora, pg. 63, 2009.

PINTO, L. P., MINGOTI, S.A. *Monitoramento de processos normais multivariados*. In Anais do V EMEPRO realizado em Viçosa, MG, 2009a.

PINTO, L. P. *Testes de hipótese para matriz de covariâncias e suas aplicações em controle de qualidade*. Belo Horizonte: Departamento de Estatística, Universidade Federal de Minas Gerais, 2009b, Dissertação (mestrado).

R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>, 2008.

SCHISSATTI, M. L., BARCIA, R. M. *Uma metodologia de implantação de cartas de Shewhart para o controle estatístico de processos*. Florianópolis: Universidade Federal de Santa Catarina, 1998, Dissertação (mestrado).

SHEWHART, W. A. Quality control charts. *Bell System Technical Journal*, 1929.

SULLIVAN J. H, STOUMBOS Z. G., MASON R. L., YOUNG, J. C. Step-Down Analysis for Changes in the Covariance Matrix and Other Parameters, *Journal of quality technology*, vol. 39, nº1, pp. 66-84, 2007.

TSAY, R. S. *Analysis of financial time series*. New York: Wiley, 2002. 448 p.

YEH, A. B., LIN, D. K. J., MCGRATH, R. N. Multivariate control charts for monitoring covariance matrix: a review. *Quality Technology and Quantitative Management*, vol. 3, p. 415-436, 2006.