



Instituto Nacional de Matemática Pura e Aplicada

Risco e Seleção de Portfólios com a Medida CVaR e o Modelo GO-GARCH

Sérgio Alvares R. de S. Maffra

Dr. Jorge P. Zubelli
Orientador

Rio de Janeiro
Agosto de 2013

Para os meus pais.

Agradecimentos

Ao meu orientador, professor Jorge Zubelli, por todo apoio e pela confiança em mim depositada. Sem estes, a conclusão deste trabalho não teria sido possível.

Aos professores Rama Cont e Luca Mertens, pelo incentivo, pelos comentários e pelas correções que ajudaram a melhorar a qualidade deste trabalho.

Aos professores e monitores do mestrado, por toda ajuda oferecida nas disciplinas.

Aos colegas de curso e do projeto de opções reais, pelo companheirismo que tornou a experiência mais proveitosa e agradável.

A Adriana Pereira, Diego Nehab, Fabíola Maffra, Luciana Blatter e Rafael Szmit pela ajuda com o texto.

À minha família, por tudo.

Resumo

Este trabalho trata de duas atribuições de um gestor de investimentos: a avaliação do risco de um portfólio de ativos financeiros e a seleção de novos portfólios.

Foram implementados algoritmos para: o cálculo de estimativas das medidas de risco, a seleção de um portfólio de risco mínimo, o ajuste do modelo de retornos a dados históricos e para a avaliação diária do risco de um portfólio. Nestes algoritmos, o risco dos ativos e portfólios é avaliado de forma objetiva através das medidas *Value at Risk* (VaR) e *Conditional Value at Risk* (CVaR). Além disso, os retornos financeiros são descritos pelos modelos *Vector Autoregressive* (VAR) e *Generalized Orthogonal GARCH* (GO-GARCH), que permitem representar retornos multivariados, não gaussianos e com volatilidade condicional.

Os algoritmos foram testados com séries históricas de preços de fechamento de ações negociadas na BM&FBovespa, o que permitiu comprovar a adequação das ferramentas desenvolvidas. Mais especificamente, foi verificado que os modelos VAR e GO-GARCH descrevem de forma apropriada os retornos financeiros das ações testadas, pois capturaram a dinâmica dos retornos corretamente. O estimador das medidas VaR e CVaR se mostrou sensível ao tamanho da amostra utilizada, sendo adequado nos casos em que grandes amostras estão disponíveis ou quando o tamanho das mesmas pode ser controlado. A avaliação diária do risco de um portfólio também se mostrou apropriada, pois as violações de VaR encontradas são independentes entre si e ocorreram com a frequência esperada. Finalmente, os portfólios de risco mínimo calculados apresentam as características esperadas. Isto é, o risco destes aumenta junto com o retorno esperado mínimo e também é possível observar o efeito benéfico da diversificação de um investimento.

Palavras chave: Value at Risk, Conditional Value at Risk, GO-GARCH, VAR, Seleção de Portfólios, Avaliação de Risco

Abstract

Two of the main tasks of a portfolio manager are studied in this work: the evaluation of risk for a portfolio of financial assets and the selection of new portfolios.

The study includes the development of algorithms for the estimation of model parameters for financial returns, for the selection of minimum risk portfolios and for the daily evaluation of portfolio risk. As objective measures of risk, the *Value at Risk* (VaR) and *Conditional Value at Risk* (CVaR) measures were used due to their popularity and properties. The *Vector Autoregressive* (VAR) and the *Generalized Orthogonal GARCH* (GO-GARCH) models were used to represent the financial returns, resulting in a non-Gaussian conditional volatility model for multivariate returns.

Using historical price series of a few stocks that are traded on BM&FBovespa, it was possible to verify the adequacy of the developed algorithms. More specifically, it was possible to verify that the VAR and GO-GARCH models were appropriate to represent the returns of the selected stocks, as the adopted model was able to capture the dynamics of the returns correctly. As expected, the VaR and CVaR estimates are sensitive to the size of the samples used in their computation. Therefore, these are adequate when a large sample of returns is available. The daily evaluation of portfolio risk was also appropriate, as the VaR violations observed in the backtesting procedures were found to be independent and to occur with the expected frequency. Finally, the minimum risk portfolios computed present the expected features. For example, their associated risk was found to increase with the minimum expected return and it was also possible to observe the beneficial effect of diversifying an investment.

Keywords: Value at Risk, Conditional Value at Risk, GO-GARCH, VAR, Portfolio Selection, Risk Evaluation

Sumário

1	Sobre o Trabalho	1
1.1	Introdução	1
1.2	Abordagem	1
1.3	Objetivo	2
1.4	Estrutura	3
I	Revisão da Teoria	5
2	Modelos para Retornos Financeiros	7
2.1	Definições Básicas	7
2.2	Modelos Seleccionados	9
2.3	Modelos GARCH e IGARCH	9
2.3.1	Simulação	10
2.4	Modelo VAR	12
2.4.1	Simulação	13
2.4.2	Ajuste do Modelo	14
2.5	Modelo GO-GARCH	15
2.5.1	Simulação	17
2.5.2	Ajuste do Modelo	18
2.5.3	Estimador ICA	20
2.5.4	Estimador MM2011	21
2.5.5	Resíduos Padronizados	25
2.5.6	Modelo O-GARCH	26
3	Portfólios com Risco Mínimo	27
3.1	Medidas de Risco	27
3.1.1	Value at Risk (VaR)	28
3.1.2	Conditional Value at Risk (CVaR)	30
3.2	Estimação do CVaR	32
3.3	Portfólios com CVaR Mínimo	35
II	Aplicações Práticas	37
4	Implementação e Avaliação dos Estimadores	39
4.1	Estimação de VaR e CVaR	39
4.2	Estimação de Modelos GO-GARCH	42
4.2.1	Estratégia de Simulação	42

4.2.2	Métrica do Erro de Estimação	43
4.2.3	Testes e Resultados	43
5	Seleção de Portfólios	47
5.1	Implementação	47
5.2	Resultados	48
5.2.1	Parâmetros de Entrada	48
5.2.2	Análise de um Portfólio	49
5.2.3	Análise de Gráficos	49
5.3	Conclusão	51
6	Ajuste do Modelo VAR+GO-GARCH	55
6.1	Implementação	55
6.2	Resultados	56
6.2.1	Descrição do Ajuste	56
6.2.2	Diagnóstico do Ajuste	57
6.2.3	Análise da Covariância Condicional	64
6.3	Conclusão	66
7	Avaliação do Risco de um Portfólio	71
7.1	Implementação	71
7.2	Backtesting	72
7.3	Resultados	74
7.4	Conclusão	76
III	Conclusão e Complementos	83
8	Conclusão	85
8.1	Trabalhos Futuros	86
A	Análise de Componentes Independentes	87
A.1	Definição	87
A.2	Estimação	88
A.3	Centralização e Branqueamento	89
A.4	Função de Contraste	90
A.5	Algoritmo FastICA	93
A.5.1	Estimação de uma componente independente	93
A.5.2	FastICA	94
B	Análise de Componentes Principais Comuns	97
B.1	Definição	97
B.2	Estimação	98
B.3	Adaptação do algoritmo FG	102
C	Dados Históricos	107
D	Diagnóstico Modelos IGARCH	111
	Bibliografia	130

Lista de Figuras

3.1	VaR como um quantil da distribuição das perdas.	29
3.2	Valores de perda usado no cálculo de CVaR.	31
4.1	Teste de estimação para VaR e CVaR - Erro absoluto (RMSE)	40
4.2	Teste de estimação para VaR e CVaR - Erro relativo (RRMSE)	41
4.3	Densidade de probabilidade na cauda das distribuições	42
4.4	Parâmetros usados na simulação de processos GO-GARCH	44
4.5	Erros de estimação da matriz U usando os estimadores MM2011 e ICA	45
5.1	Composição x_1 do portfólio Π_1	50
5.2	Comparação de VaR/CVaR: portfólio Π_1 versus ações individuais	50
5.3	VaR/CVaR do portfólio em função do retorno médio	52
5.4	Composição do portfólio em função do retorno médio	53
6.1	Matrizes de correlação: r_t, x_t, z_t e seus quadrados	57
6.2	Testes de autocorrelação: r_t, x_t, z_t e seus quadrados	58
6.3	Teste da matriz de covariância dos retornos padronizados z_t	60
6.4	Q-Q Plot dos resíduos padronizados do modelo VAR+GO-GARCH	61
6.5	Matrizes do modelo GO-GARCH ajustado	62
6.6	Teste dos valores esperados de y_t, ϵ_t e h_t	64
6.7	Teste das matrizes de covariância de y_t e ϵ_t	64
6.8	Fatores do modelo GO-GARCH ajustado	65
6.9	Volatilidades condicionais do modelo GO-GARCH	66
6.10	Covariâncias condicionais do modelo GO-GARCH ajustado	67
6.11	Comparação de desvios padrão: GO-GARCH versus IGARCH	68
6.12	Comparação de volatilidades: GO-GARCH versus IGARCH	69
7.1	Previsões de VaR e CVaR	77
7.2	Previsões de VaR e CVaR (suavizadas)	78
7.3	Previsões de VaR e CVaR, com violações de VaR, para $\beta = 90\%$	79
7.4	Previsões de VaR e CVaR, com violações de VaR, para $\beta = 95\%$	80
7.5	Previsões de VaR e CVaR, com violações de VaR, para $\beta = 99\%$	81
C.1	Séries de preços de fechamento CMIG4 e EMBR3 antes e depois do ajuste	108
C.2	Média dos log-retornos diários	109
C.3	<i>Boxplot</i> dos log-retornos diários	109

Lista de Tabelas

4.1	VaR e CVaR para as distribuições normal padrão e t_5	40
5.1	Comparação do risco de portfólios	49
6.1	Teste do valor esperado dos resíduos x_t do ajuste VAR	59
6.2	Teste do valor esperado dos resíduos padronizados z_t	60
6.3	Teste de Kolmogorov-Smirnov: distribuição de z_t versus a distribuição de ϵ_t	62
6.4	Teste de estacionariedade dos fatores y_{it} do modelo GO-GARCH ajustado	63
7.1	Composição do portfólio Π_5	73
7.2	Avaliação das previsões de VaR feitas com modelo GO-GARCH (ICA)	74
7.3	Avaliação das previsões de VaR feitas com diversos modelos de retornos	75
C.1	Ações da BM&FBovespa selecionadas para estudo	107
C.2	Matriz de correlação dos log-retornos diários	109
D.1	Testes para os resíduos padronizados PETR4	112
D.2	Testes para o quadrado dos resíduos padronizados PETR4	112
D.3	Testes para os resíduos padronizados VALE5	115
D.4	Testes para o quadrado dos resíduos padronizados VALE5	115
D.5	Testes para os resíduos padronizados USIM5	118
D.6	Testes para o quadrado dos resíduos padronizados USIM5	118
D.7	Testes para os resíduos padronizados GGBR4	121
D.8	Testes para o quadrado dos resíduos padronizados GGBR4	121
D.9	Testes para os resíduos padronizados CSNA3	124
D.10	Testes para o quadrado dos resíduos padronizados CSNA3	124

Capítulo 1

Sobre o Trabalho

1.1 Introdução

Este trabalho tem como tema a avaliação do risco de mercado associado a um portfólio de ativos financeiros, ou seja, a determinação do potencial das perdas futuras de um portfólio devido a variações no valor dos ativos que o compõem.

Aqui são tratadas duas das principais atribuições de um gestor de investimentos: avaliar o risco de um portfólio existente e selecionar novos portfólios.

A avaliação do risco tem um papel fundamental na gestão de capital, pois permite a um gestor tomar ações preventivas que podem proteger o seu patrimônio. Muitas vezes, as decisões que levam a essas ações partem da avaliação de cenários aos quais são atribuídas diferentes probabilidades, isto é, da análise de um modelo probabilístico criado para o valor futuro dos ativos. Assim, aqui também é abordada a criação destes modelos, que geralmente são construídos com a utilização de dados históricos de mercado.

A seleção de um portfólio consiste em determinar a fração do capital disponível que deve ser investida em cada ativo. Esta seleção pode ser bastante complexa, dada a grande quantidade de variáveis que podem ser exploradas (liquidez, risco, retorno, etc.). Desta forma, neste trabalho é tratado apenas um caso: a seleção de um portfólio que tenha risco mínimo em função de uma rentabilidade mínima desejada.

Para que o risco mínimo possa ser definido, é preciso que este seja avaliado objetivamente, ou seja, através de uma medida (ou métrica) de risco. Apenas desta forma é possível construir algoritmos que o minimizem, como é feito para a seleção de portfólios. Assim, este é mais um dos tópicos aqui abordados.

1.2 Abordagem

Todo o desenvolvimento feito neste trabalho tem como base as ferramentas escolhidas para:

- (i) Medir o risco de um portfólio de forma objetiva,
- (ii) Calcular a estimativa da medida de risco,
- (iii) Modelar os preços (ou retornos) futuros de um ativo.

Para a avaliação objetiva do risco de um portfólio foi escolhida a medida CVaR (*Conditional Value at Risk*), que indica o valor esperado das perdas financeiras condicionado ao fato destas serem

superiores a um limiar. Este limiar, como será mostrado no Capítulo 3, corresponde ao VaR (*Value at Risk*), que é uma medida de risco bastante difundida no mercado.

A medida CVaR foi escolhida devido às suas boas propriedades. Em primeiro lugar, CVaR é uma função convexa, o que a torna apropriada para a construção de problemas de otimização. Em segundo lugar, é uma *medida coerente de risco*, o que a indica como apropriada para mensurar o risco de perdas financeiras.

Além disso, o estimador para a medida CVaR, que foi apresentado em [49], pode ser adaptado para resolver o problema de seleção de portfólios de risco mínimo sem grandes dificuldades. Este estimador permite calcular o CVaR de um portfólio a partir de uma amostra de retornos (ou perdas) dos seus ativos, o que traz uma série de vantagens. Ao utilizar esta estratégia, por exemplo, é possível reduzir de forma significativa o risco de modelo, ou seja, o risco causado pelo uso de um modelo matemático inadequado; pois as estimativas de risco podem ser calculadas diretamente sobre dados históricos. Esta estratégia de estimação também permite calcular previsões de perdas futuras com facilidade, já que estas podem ser estimadas sobre uma amostra de retornos futuros. Portfólios heterogêneos podem ser tratados trivialmente, desde que o retorno dos ativos que os compõem estejam disponíveis.

Quanto ao modelo dos retornos, este foi escolhido tendo em vista a sua adequação aos fatos conhecidos sobre os retornos financeiros que podem ser observados na prática. Assim, foi escolhido um modelo não gaussiano, multivariado e de volatilidade condicional. Esta escolha pode ser facilmente justificada ao analisar as falhas de suas alternativas. A distribuição normal de probabilidade não é uma boa escolha para modelos de retornos por não apresentar assimetria e caudas pesadas. Modelos univariados são incapazes de representar a dinâmica de um conjunto de ativos. Desta forma, são inadequados para aplicações envolvendo portfólios. Além disso, os retornos financeiros têm sua dinâmica explicada principalmente pela volatilidade, o que explica a necessidade de um modelo de volatilidade condicional.

Dentre todas as opções disponíveis, o modelo GO-GARCH (*Generalized Orthogonal GARCH*) pareceu a mais promissora. Apesar do nome, o modelo GO-GARCH é, na verdade, uma família de modelos, na qual cada modelo de processo estocástico m -dimensional é a combinação linear de um conjunto de n fatores unidimensionais não correlacionados (ou independentes) entre si. Em geral, cada um destes fatores é um processo estocástico não gaussiano e de volatilidade condicional, o que mostra a adequação do modelo GO-GARCH aos problemas aqui tratados.

Entretanto, também é necessário adotar um modelo para a média dos retornos, pois os modelos GO-GARCH assumem que os mesmos têm média nula. Assim, devido à sua simplicidade, também foi adotado o modelo VAR (*Vector Autoregressive*).

Resumindo, é possível dizer que a abordagem deste trabalho consiste em calcular as previsões dos retornos usando os modelos VAR e GO-GARCH e mensurar o risco através da medida CVaR, estimada sobre uma amostra de retornos.

1.3 Objetivo

Este trabalho tem como objetivo o desenvolvimento de algoritmos cobrindo os tópicos mencionados na Seção 1.1, ou seja, desenvolver algoritmos para:

- (i) O cálculo de medidas de risco,
- (ii) A previsão do valor futuro de um conjunto de ativos,
- (iii) A previsão de perdas futuras (combinando os algoritmos dos dois itens anteriores),
- (iv) A seleção de portfólios de risco mínimo em função de uma rentabilidade mínima.

As ferramentas matemáticas selecionadas na Seção 1.2 foram utilizadas no desenvolvimento destes algoritmos, buscando detalhar a justificativa teórica e a implementação dos mesmos. Também é dada ênfase à utilização de implementações próprias, em detrimento ao uso de pacotes prontos.

Finalmente, uma vez definidos os algoritmos a desenvolver e as ferramentas matemáticas que serão utilizadas, um conjunto de questões surge naturalmente:

- (i) O quão difícil é a implementação dos algoritmos?
- (ii) O quão difícil é o ajuste de um modelo GO-GARCH a dados históricos de mercado?
- (iii) O modelo GO-GARCH é realmente um bom modelo para os retornos financeiros?
- (iv) O estimador de CVaR é capaz de avaliar o risco de maneira apropriada?
- (v) O modelo GO-GARCH pode ser usado em conjunto com o estimador de CVaR para gerenciar o risco de um portfólio?
- (vi) Quais as propriedades dos portfólios de risco mínimo baseados nos estimadores de CVaR?

1.4 Estrutura

Este trabalho foi dividido em três partes. Na primeira delas, é feita uma *Revisão da Teoria*, que é composta dos capítulos 2 e 3. No Capítulo 2, são apresentados os modelos que foram utilizados para a modelagem de retornos financeiros. Estes são os modelos GARCH, IGARCH, VAR e GO-GARCH. Em seguida, no Capítulo 3, é feita uma revisão sobre *medidas coerentes de risco*, sobre as medidas de risco VaR e CVaR, sobre seus estimadores e sobre como pode ser feita a seleção de um portfólio de risco mínimo.

Na segunda parte são descritas as *Aplicações Práticas* desenvolvidas neste trabalho. Nesta parte são apresentados detalhes de implementação e os resultados dos testes de cada um dos algoritmos implementados. O Capítulo 4 trata dos estimadores desenvolvidos para VaR e CVaR e para o ajuste de modelos GO-GARCH. O Capítulo 5 apresenta o algoritmo para a seleção de portfólios de risco mínimo. Em seguida, o Capítulo 6 mostra como os modelos VAR e GO-GARCH são ajustados aos dados históricos. Concluindo a segunda parte do trabalho, o modelo de retornos adotado e os estimadores de VaR/CVaR são utilizados em conjunto no Capítulo 7 com o objetivo de avaliar o risco associado a um portfólio, ou seja, com o objetivo de calcular previsões para perdas futuras.

Finalmente, na terceira parte são apresentadas as conclusões, no Capítulo 8, e os complementos dos capítulos anteriores. A análise de componentes independentes (ICA) e a análise de componentes principais comuns são apresentadas, respectivamente, nos apêndices A e B. No Apêndice C é apresentada a descrição dos dados históricos utilizados e o processamento realizado sobre os mesmos. Por último, no Apêndice D, são apresentados os resultados dos diagnósticos de ajustes IGARCH realizados.

Parte I
Revisão da Teoria

Capítulo 2

Modelos para Retornos Financeiros

Dada a vasta literatura existente sobre a modelagem de retornos financeiros e a quantidade de modelos experimentados neste trabalho, a apresentação feita neste capítulo se limita aos principais modelos empregados.

Inicialmente, na Seção 2.1, são apresentadas definições básicas, comuns a todos os modelos. Em seguida, na Seção 2.2, são mostrados os modelos utilizados no trabalho. A Seção 2.3 trata dos modelos de volatilidade condicional GARCH e IGARCH, enquanto as seções 2.4 e 2.5 mostram, respectivamente, os modelos multivariados VAR e GO-GARCH.

Além de definir os modelos utilizados, este capítulo apresenta técnicas para simulação e ajuste dos modelos a dados reais.

2.1 Definições Básicas

Seja $\{\eta_t\}_{t \geq 1}$ uma série temporal m -variada. A série $\{\eta_t\}_{t \geq 1}$ é *estritamente estacionária* quando $A = (\eta_{t_1}, \eta_{t_2}, \dots, \eta_{t_k})$ e $B = (\eta_{t_1+\delta}, \eta_{t_2+\delta}, \dots, \eta_{t_k+\delta})$ têm a mesma distribuição conjunta, ou seja, a distribuição de A é invariante em relação a deslocamentos de tempo δ . Como esta condição é bastante restritiva, muitas vezes recorre-se à *estacionariedade fraca*, que pode ser verificada com mais facilidade na prática. A série $\{\eta_t\}_{t \geq 1}$ é *fracamente estacionária* quando seu vetor média $\mu = \mathbb{E}[\eta_t]$ e suas matrizes de covariância Γ_k e de correlação ρ_k , são invariantes no tempo. Estes são definidos como

$$\mu = \mathbb{E}[\eta_t], \quad (2.1.1)$$

$$\Gamma_k = \text{Cov}(\eta_t, \eta_{t-k}) = \mathbb{E}[(\eta_t - \mu)(\eta_{t-k} - \mu)'], \quad (2.1.2)$$

e

$$\rho_k = \text{Corr}(\eta_t, \eta_{t-k}) = D^{-1/2} \Gamma_k D^{-1/2}, \quad (2.1.3)$$

onde $k \in \mathbb{Z}$ e $D = \text{diag}(\sigma_1^2, \dots, \sigma_m^2)$ é a matriz diagonal cujos elementos correspondem às variâncias das componentes de η_t .

Geralmente, Γ_k e ρ_k não são matrizes simétricas. Na verdade, são válidas as relações

$$\Gamma_k' = \Gamma_{-k} \quad (2.1.4)$$

e

$$\rho_k' = \rho_{-k}. \quad (2.1.5)$$

Quando a distribuição de η_t é desconhecida, μ , Γ_k e ρ_k podem ser estimados, a partir de uma

amostra $A = \{\eta_1, \dots, \eta_T\}$, pelos estimadores

$$\hat{\mu} = \frac{1}{T} \sum_{i=1}^T \eta_i, \quad (2.1.6)$$

$$\hat{\Gamma}_k = \frac{1}{T} \sum_{i=1}^T (\eta_i - \hat{\mu})(\eta_{i-k} - \hat{\mu})', \quad (2.1.7)$$

e

$$\hat{\rho}_k = \hat{D}^{-1/2} \hat{\Gamma}_k \hat{D}^{-1/2}, \quad (2.1.8)$$

onde temos $k \in \mathbb{Z}$ e $\hat{D} = \text{diag}(\hat{\sigma}_1^2, \dots, \hat{\sigma}_m^2)$. Com $\hat{\sigma}_j^2 = 1/T \sum_{i=1}^T (\eta_{j,i} - \hat{\mu}_j)^2$ sendo a variância amostral da série formada pela j -ésima componente dos elementos da amostra A .

Os estimadores acima (e suas versões univariadas) são fundamentais na análise de séries temporais, pois podem ser utilizados tanto na construção de estratégias para o ajuste de séries a um determinado conjunto de dados, quanto na construção de diagnósticos de modelos ajustados.

Neste trabalho são utilizados dois testes estatísticos baseados nos estimadores acima: o teste de Ljung-Box [38] e sua generalização para m dimensões, o teste de Hosking [28].

Ambos os testes, definidos a seguir, testam a existência de correlação serial nos primeiros n lags de uma série temporal.

Definição 1. *Teste de Ljung-Box*

$$\begin{aligned} H_0: & \quad \rho_i = 0 \text{ para todo } i \in \{1, \dots, n\} \\ H_a: & \quad \rho_i \neq 0 \text{ para algum } i \in \{1, \dots, n\} \\ \text{Estatística:} & \quad Q(n) \sim \chi_n^2 \\ \text{Regra:} & \quad H_0 \text{ é rejeitada em favor de } H_a \text{ quando } Q(n) > \chi_n^2(\alpha) \end{aligned}$$

Definição 2. *Teste de Hosking*

$$\begin{aligned} H_0: & \quad \rho_i = 0 \text{ para todo } i \in \{1, \dots, n\} \\ H_a: & \quad \rho_i \neq 0 \text{ para algum } i \in \{1, \dots, n\} \\ \text{Estatística:} & \quad Q_m(n) \sim \chi_{m^2 n}^2 \\ \text{Regra:} & \quad H_0 \text{ é rejeitada em favor de } H_a \text{ quando } Q_m(n) > \chi_{m^2 n}^2(\alpha) \end{aligned}$$

Nas Definições 1 e 2, $\chi_q^2(\alpha)$ representa o quantil de nível α da distribuição chi-quadrada com q graus de liberdade. Além disso, as estatísticas $Q(n)$ e $Q_m(n)$ são dadas por

$$Q(n) = T(T+2) \sum_{l=1}^n \frac{\hat{\rho}_l^2}{T-l} \quad (2.1.9)$$

e

$$Q_m(n) = T^2 \sum_{l=1}^n \frac{1}{T-l} \text{tr} \left(\hat{\Gamma}_l' \hat{\Gamma}_0^{-1} \hat{\Gamma}_l \hat{\Gamma}_0^{-1} \right). \quad (2.1.10)$$

Estes testes também podem ser utilizados como diagnóstico do ajuste de um modelo ARMA(p, q). Para este fim, os testes são aplicados à série dos resíduos e os grau de liberdade de $Q(n)$ e $Q_m(n)$ são ajustados de forma que $Q(n) \sim \chi_{n-p-q}^2$ e $Q_m(n) \sim \chi_{m^2(n-p-q)}^2$. Quando a hipótese nula H_0 não é rejeitada, os testes indicam que os ajustes são apropriados.

2.2 Modelos Seleccionados

Seja $\{r_t\}_{t \geq 1}$ uma série temporal m -variada que representa os retornos diários para m ativos financeiros e que pode ser decomposto em duas outras séries: $\{\mu_t\}_{t \geq 1}$ e $\{x_t\}_{t \geq 1}$, correspondentes à média e à volatilidade de $\{r_t\}_{t \geq 1}$, ou seja,

$$r_t = \mu_t + x_t, \quad t \geq 1. \quad (2.2.1)$$

No caso do modelo multivariado adotado, temos μ_t como um modelo VAR (*Vector Autoregressive*) e x_t como um modelo GO-GARCH (*Generalized Orthogonal GARCH*), o que explica a denominação VAR+GO-GARCH. Veja [25, 52, 27, 53, 7, 44].

Com o objetivo principal de permitir a avaliação dos resultados obtidos para o modelo multivariado, foram testados diversos modelos univariados para μ_t e x_t , totalizando um número de combinações superior a 800. Para μ_t foram adotados os modelos ARMA e para x_t os modelos GARCH, IGARCH, EGARCH, GJR-GARCH, APARCH, FGARCH (com seus diversos submodelos) e CSGARCH, ou seja, foram testados todos os modelos para volatilidade disponíveis no pacote *rugarch* [24]. Para uma descrição destes modelos, veja [52] e [24].

Dentre todos os modelos univariados testados, o modelo IGARCH(1,1) foi escolhido por dois motivos. Em primeiro lugar, apresentou um excelente desempenho sob o critério BIC (Critério de Informação Bayesiano, veja [51, 52]). Além disso, é o modelo adotado pela bem conhecida metodologia RiskMetricsTM. Veja [48, 52].

2.3 Modelos GARCH e IGARCH

Nesta seção é feita uma breve revisão do bem conhecido modelo GARCH e de um caso particular, o modelo IGARCH. Como o foco deste trabalho são os modelos multivariados, apenas as principais características desses modelos são apresentadas. Assim, também são omitidos os métodos para ajustes a dados reais. Para uma melhor referência sobre os mesmos, veja [52, 56].

O modelo GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*) é um modelo para a volatilidade condicional dos retornos r_t . Assim, considerando a Equação (2.2.1), é um modelo para a série $\{x_t\}_{t \geq 1}$, cujos valores são dados por

$$x_t = \sigma_t \epsilon_t, \quad (2.3.1)$$

onde a volatilidade condicional σ_t é dada pela equação

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i x_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \quad (2.3.2)$$

e $\{\epsilon_t\}_{t \geq 1}$ é uma sequência de variáveis aleatórias iid com $\mathbb{E}[\epsilon_t] = 0$ e $\text{Var}(\epsilon_t) = 1$.

Considere a sequência martingal de diferenças $\{\eta_t\}_{t \geq 1}$ (veja a Definição 3), dada por

$$\eta_t = x_t^2 - \sigma_t^2. \quad (2.3.3)$$

Tomando $\sigma_t^2 = x_t^2 - \eta_t$ e substituindo o valor de σ_t na Equação 2.3.2, é possível caracterizar um processo GARCH(m,s) como um processo ARMA(m,s) para a série $\{x_t^2\}_{t \geq 1}$, ou seja,

$$x_t^2 = \alpha_0 + \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) x_{t-i}^2 + \eta_t + \sum_{j=1}^s \beta_j \eta_{t-j}. \quad (2.3.4)$$

Definição 3. *Sequência Martingal de Diferenças*

Uma sequência martingal de diferenças é um sequência $\{X_t\}_{t=1}^{\infty}$ de variáveis aleatórias tal que $\mathbb{E}[X_t] = 0$ e $\mathbb{E}[X_t | \mathcal{F}_{t-1}] = 0$ para todo t .

A Equação (2.3.4) também pode ser escrita utilizando o operador de lag L , o que leva à expressão

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_k L^k) x_t^2 = \alpha_0 + (1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_s L^s) \eta_t, \quad (2.3.5)$$

onde $\phi_i = \alpha_i + \beta_i$ e $k = \max(m, s)$. Ou, de forma mais compacta, é possível escrever

$$\Phi(L) x_t^2 = \alpha_0 + B(L) \eta_t, \quad (2.3.6)$$

onde $\Phi(L)$ e $B(L)$ são polinômios em L que recebem, respectivamente, os nomes de *polinômio AR* e *polinômio MA*.

O modelo IGARCH (ou *Integrated GARCH*) pode ser visto como um caso especial do modelo GARCH, caracterizado pela presença de raízes unitárias no polinômio $\Phi(L)$, ou seja, quando $\Phi(L)$ tem a unidade como uma de suas raízes, tem-se em mãos um modelo IGARCH.

Entretanto, existe uma diferença significativa entre os dois modelos, já que no modelo IGARCH a variância não condicional $\text{Var}(x_t)$ não está bem definida. Como

$$\text{Var}(x_t) = \frac{\alpha_0}{\Phi(1)}, \quad (2.3.7)$$

a mesma é bem definida quando $\Phi(1) \neq 0$. No caso do modelo IGARCH, $\Phi(1) = 0$, pois a unidade é uma das raízes do polinômio $\Phi(L)$.

Os modelos GARCH(1,1) e IGARCH(1,1) são de particular interesse. O primeiro destes é utilizado na modelagem dos fatores que controlam a dinâmica dos modelos GO-GARCH e o segundo apresentou os melhores resultados nos testes de modelos univariados aqui realizados, como descrito na Seção 2.2.

Assim, aplicando a Equação (2.3.2), a volatilidade condicional do modelo GARCH(1,1) é dada por

$$\sigma_t = \alpha_0 + \alpha_1 x_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (2.3.8)$$

e, como

$$\Phi(1) = 0 \longrightarrow \alpha_1 = 1 - \beta_1, \quad (2.3.9)$$

a do modelo IGARCH(1,1) por

$$\sigma_t = \alpha_0 + (1 - \beta_1) x_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \quad (2.3.10)$$

2.3.1 Simulação

A simulação de uma série temporal entre os períodos t e $t + h$ de acordo com o modelo GARCH(1,1) pode ser feita de forma iterativa. Os valores da volatilidade e do retorno para o tempo $t + 1$ são calculados pela aplicação direta das Equações (2.3.8) e (2.3.1). Incorporando estes novos valores à série temporal existente no tempo t , é possível calcular os valores para $t + 2$ e, procedendo da mesma forma, para todos os tempos até o tempo $t + h$. O mesmo processo pode ser usado também para o modelo IGARCH(1,1).

Como exemplo, considere as Equações (2.3.11) a (2.3.16).

$$\sigma_{t+1}^2 | \mathcal{F}_t = \alpha_0 + \alpha_1 x_t^2 + \beta_1 \sigma_t^2, \quad (2.3.11)$$

$$\sigma_{t+2}^2 | \mathcal{F}_t = \alpha_0 + \alpha_1 \{x_{t+1} | \mathcal{F}_t\}^2 + \beta_1 \{\sigma_{t+1}^2 | \mathcal{F}_t\}, \quad (2.3.12)$$

⋮

$$\sigma_{t+h}^2 | \mathcal{F}_t = \alpha_0 + \alpha_1 \{x_{t+h-1} | \mathcal{F}_t\}^2 + \beta_1 \{\sigma_{t+h-1}^2 | \mathcal{F}_t\}. \quad (2.3.13)$$

$$x_{t+1} | \mathcal{F}_t = \epsilon_{t+1} \sqrt{\sigma_{t+1}^2 | \mathcal{F}_t}, \quad (2.3.14)$$

$$x_{t+2} | \mathcal{F}_t = \epsilon_{t+2} \sqrt{\sigma_{t+2}^2 | \mathcal{F}_t}, \quad (2.3.15)$$

⋮

$$x_{t+h} | \mathcal{F}_t = \epsilon_{t+h} \sqrt{\sigma_{t+h}^2 | \mathcal{F}_t}. \quad (2.3.16)$$

A cada passo, os valores da série $\{\epsilon_t\}_{t \geq 1}$ também devem ser gerados para calcular os retornos x_t . Sendo esta uma sequência iid de variáveis aleatórias com distribuição conhecida, os valores da sequência são dados por realizações independentes de uma variável aleatória com a mesma distribuição.

Considere agora o valor da volatilidade no tempo $t + h$. Tomando $x_t = \sigma_t \epsilon_t$ na Equação (2.3.8) é possível descrever a volatilidade σ_{t+h} apenas em função das volatilidades passadas, ou seja,

$$\begin{aligned} \sigma_{t+h}^2 &= \alpha_0 + \alpha_1 x_{t+h-1}^2 + \beta_1 \sigma_{t+h-1}^2 \\ &= \alpha_0 + \alpha_1 (\sigma_{t+h-1}^2 \epsilon_{y+h-1}^2) + \beta_1 \sigma_{t+h-1}^2 \\ &= \alpha_0 + (\alpha_1 + \beta_1) \sigma_{t+h-1}^2 + (\epsilon_{t+h-1}^2 - 1) \alpha_1 \sigma_{t+h-1}^2. \end{aligned} \quad (2.3.17)$$

Desta forma, a previsão da volatilidade $\sigma_t^2(h)$ para o tempo $t + h$, com base na informação disponível até o tempo t , é dada por

$$\begin{aligned} \sigma_t^2(h) &= \mathbb{E} [\sigma_{t+h}^2 | \mathcal{F}_t] \\ &= \alpha_0 + (\alpha_1 + \beta_1) \mathbb{E} [\sigma_{t+h-1}^2 | \mathcal{F}_t] + \alpha_1 \mathbb{E} [(\epsilon_{t+h-1}^2 - 1) \sigma_{t+h-1}^2 | \mathcal{F}_t] \\ &= \alpha_0 + (\alpha_1 + \beta_1) \mathbb{E} [\sigma_{t+h-1}^2 | \mathcal{F}_t] + \alpha_1 \mathbb{E} [\epsilon_{t+h-1}^2 - 1] \mathbb{E} [\sigma_{t+h-1}^2 | \mathcal{F}_t] \\ &= \alpha_0 + (\alpha_1 + \beta_1) \mathbb{E} [\sigma_{t+h-1}^2 | \mathcal{F}_t] + 0 \\ &= \alpha_0 + (\alpha_1 + \beta_1) \sigma_t^2(h-1), \end{aligned} \quad (2.3.18)$$

já que $\mathbb{E} [\epsilon_{t+h-1}^2 - 1] = 0$ e ϵ_{t+h-1} é independente de σ_{t+h-1} .

A Equação (2.3.18) também pode ser aplicada no caso do modelo IGARCH(1,1). Como neste modelo $\alpha_1 + \beta_1 = 1$, após repetidas substituições, obtém-se que

$$\sigma_t^2(h) = (h-1) \alpha_0 + \sigma_t^2(1). \quad (2.3.19)$$

2.4 Modelo VAR

Considere a série temporal $\{\mu_t\}_{t \geq 1}$, m -variada, correspondente à média dos retornos de m ativos financeiros. $\{\mu_t\}_{t \geq 1}$ segue um modelo VAR(p) quando

$$\mu_t = \phi_0 + a_t + \sum_{i=1}^p \Phi_i \mu_{t-i}, \quad (2.4.1)$$

onde ϕ_0 é um vetor de dimensão m , Φ_k são matrizes $m \times m$ e $\{a_t\}_{t \geq 1}$ é uma sequência de vetores aleatórios que não apresenta correlações seriais. Além disso, $\mathbb{E}[a_t] = \bar{0}$ e $\mathbb{E}[a_t a_t'] = \Sigma$, onde Σ é uma matriz $m \times m$ positiva definida. É comum assumir que a_t segue uma distribuição normal multivariada, ou seja, $a_t \sim N(0, \Sigma)$.

Considerando que a série $\{\mu_t\}_{t \geq 1}$ é fracamente estacionária, vale que $\mathbb{E}[\mu_{t-i}] = \mu$ para $\mu \in \mathbb{R}^m$ e $i \in \mathbb{Z}$. Assim, ϕ_0 pode ser expresso em função de μ e a Equação (2.4.1) reescrita utilizando as variações da série em torno de sua média, ou seja,

$$\mu = \mathbb{E}[\mu_t] = \phi_0 + \sum_{i=1}^p \Phi_i \mathbb{E}[\mu_{t-i}] = \phi_0 + \sum_{i=1}^p \Phi_i \mu \longrightarrow \phi_0 = \left(I_m - \sum_{i=1}^p \Phi_i \right) \mu. \quad (2.4.2)$$

Substituindo o valor de ϕ_0 na Equação (2.4.1), obtém-se

$$(\mu_t - \mu) = a_t + \sum_{i=1}^p \Phi_i (\mu_{t-i} - \mu). \quad (2.4.3)$$

Finalmente, com a mudança de variável $v_t = \mu_t - \mu$, a Equação (2.4.1) pode ser reescrita sem ϕ_0 (como na Equação (2.4.4)), o que facilita a prova de algumas propriedades do modelo.

$$v_t = a_t + \sum_{i=1}^p \Phi_i v_{t-i} \quad (2.4.4)$$

São válidas as seguintes propriedades relativas a a_t e v_t

- (i) $\text{Cov}(v_t, v_{t-k}) = \Gamma_k = \sum_{i=1}^p \Phi_i \Gamma_{k-i}$,
- (ii) $\text{Corr}(v_t, v_{t-k}) = \rho_k = \sum_{i=1}^p \Lambda_i \rho_{k-i}$,
- (iii) $\text{Cov}(v_t, a_t) = \Sigma$,
- (iv) $\text{Cov}(v_t, a_{t-i}) = 0$,

onde $\Lambda_i = D^{-1/2} \Gamma_i D^{1/2}$ e a matriz D é definida como na Equação (2.1.3).

A propriedade (i), é comprovada na Equação (2.4.5), recorrendo à estacionariedade de $\{v_t\}_{t \geq 1}$ ($\mathbb{E}[v_{t-k}] = 0$ para todo k) e à linearidade da esperança. Isto é,

$$\Gamma_k = \mathbb{E}[v_t v_{t-k}'] = \mathbb{E} \left[\left(a_t + \sum_{i=1}^p \Phi_i v_{t-i} \right) v_{t-k}' \right] = 0 + \sum_{i=1}^p \Phi_i \mathbb{E}[v_{t-i} v_{t-k}'] = \sum_{i=1}^p \Phi_i \Gamma_{k-i}. \quad (2.4.5)$$

Através da propriedade (i) e do fato $I = D^{1/2} D^{-1/2}$, a propriedade (ii) é facilmente verificada,

como mostrado na Equação (2.4.6)

$$\begin{aligned}
\rho_k &= D^{-1/2} \Gamma_k D^{-1/2} \\
&= D^{-1/2} \left(\sum_{i=1}^p \Phi_i \Gamma_{k-i} \right) D^{-1/2} \\
&= \sum_{i=1}^p D^{-1/2} \Phi_i (D^{1/2} D^{-1/2}) \Gamma_{k-i} D^{-1/2} \\
&= \sum_{i=1}^p (D^{-1/2} \Phi_i D^{1/2}) (D^{-1/2} \Gamma_{k-i} D^{-1/2}) \\
&= \sum_{i=1}^p \Lambda_i \rho_{k-i}.
\end{aligned} \tag{2.4.6}$$

Conforme demonstrado em [25], v_t pode ser descrito em função dos valores de a_t , o que permite a demonstração de (iii) e (iv). Veja a Equação (2.4.7).

$$v_t = \sum_{i=0}^{\infty} (\Phi_1)^i a_{t-i} \tag{2.4.7}$$

Utilizando a Equação (2.4.7), a linearidade da esperança e a ausência de correlações seriais em $\{a_t\}_{t \geq 1}$, como feito nas Equações (2.4.8) e (2.4.9), são comprovadas as propriedades (iii) e (iv), pois a ausência de correlações seriais implica que $\mathbb{E}[a_{t-k} a'_{t-j}] \neq 0$ apenas quando $j = k$.

$$\text{Cov}(v_t, a_t) = \mathbb{E}[v_t a'_t] = \mathbb{E} \left[\left(\sum_{i=0}^{\infty} (\Phi_1)^i a_{t-i} \right) a'_t \right] = \mathbb{E} \left[\sum_{i=0}^{\infty} (\Phi_1)^i (a_{t-i} a'_t) \right] = \mathbb{E}[a_t a'_t] = \Sigma \tag{2.4.8}$$

$$\text{Cov}(v_t, a_{t-k}) = \mathbb{E}[v_{t-k} a'_t] = \mathbb{E} \left[\left(\sum_{i=0}^{\infty} (\Phi_1)^i a_{t-i} \right) a'_t \right] = \mathbb{E} \left[\sum_{i=0}^{\infty} (\Phi_1)^i (a_{t-k-i} a'_t) \right] = 0 \tag{2.4.9}$$

2.4.1 Simulação

A simulação de séries temporais segundo o modelo VAR(p) pode ser realizada através da aplicação direta da equação que o define. Considere a série temporal v_t , definida na Equação (2.4.4), e a filtração gerada pela mesma, $\{\mathcal{F}_t\}_{t \geq 0}$.

Utilizando a equação do modelo, podemos calcular $v_{t+1} | \mathcal{F}_t$, que corresponde a uma variável aleatória \mathcal{F}_t -mensurável com distribuição normal.

$$v_{t+1} | \mathcal{F}_t = a_{t+1} + \sum_{i=1}^p \Phi_i v_{t+1-i} \tag{2.4.10}$$

Da mesma forma, podem ser simulados valores de $\{v_t\}_{t \geq 1}$ para os tempos $t + 2, \dots, t + h$

$$v_{t+2} | \mathcal{F}_t = a_{t+2} + \Phi_1 \{v_{t+1} | \mathcal{F}_t\} + \sum_{i=2}^p \Phi_i v_{t+2-i}, \quad (2.4.11)$$

$$v_{t+3} | \mathcal{F}_t = a_{t+3} + \Phi_1 \{v_{t+1} | \mathcal{F}_t\} + \Phi_2 \{v_{t+2} | \mathcal{F}_t\} + \sum_{i=3}^p \Phi_i v_{t+3-i}, \quad (2.4.12)$$

⋮

$$v_{t+h} | \mathcal{F}_t = a_{t+h} + \sum_{i=1}^n \Phi_i \{v_{t+h-i} | \mathcal{F}_t\} + \sum_{i=n+1}^p \Phi_i v_{t+h-i}, \quad (2.4.13)$$

onde $n = \min(h - 1, p)$ corresponde ao número de termos resultantes de simulação que são utilizados no cálculo de $v_{t+h} | \mathcal{F}_t$.

Conforme mostrado em [25, 52], segundo o critério de erro quadrático mínimo, a esperança $\mathbb{E}[v_{t+h} | \mathcal{F}_t]$ é a melhor previsão para o valor de v_{t+h} com base na informação disponível até o tempo t .

Seja $v_t(h) = \mathbb{E}[v_{t+h} | \mathcal{F}_t]$. Como ocorre na simulação de valores de v_{t+h} , as previsões $v_t(h)$ serão uma combinação de valores conhecidos e calculados, que podem ser calculados iterativamente com base na equação do modelo VAR(p). Assim,

$$v_t(1) = \mathbb{E}[v_{t+1} | \mathcal{F}_t] = \sum_{i=1}^p \Phi_i v_{t+1-i}, \quad (2.4.14)$$

$$v_t(2) = \mathbb{E}[v_{t+2} | \mathcal{F}_t] = \Phi_1 v_t(1) + \sum_{i=2}^p \Phi_i v_{t+2-i}, \quad (2.4.15)$$

$$v_t(3) = \mathbb{E}[v_{t+3} | \mathcal{F}_t] = \Phi_1 v_t(2) + \Phi_2 v_t(1) + \sum_{i=3}^p \Phi_i v_{t+3-i}, \quad (2.4.16)$$

⋮

$$v_t(h) = \mathbb{E}[v_{t+h} | \mathcal{F}_t] = \sum_{i=1}^n \Phi_i v_t(h-i) + \sum_{i=n+1}^p \Phi_i v_{t+h-i}. \quad (2.4.17)$$

Finalmente, voltando à Equação (2.4.10), pode-se reparar que o somatório na equação corresponde ao valor de $v_t(1)$. Assim, $v_{t+1} | \mathcal{F}_t \sim N(v_t(1), \Sigma)$, pois $a_{t+1} \sim N(0, \Sigma)$ e $v_t(1) \in \mathbb{R}$. Logo, simular valores para v_{t+1} consiste simplesmente em gerar amostras da distribuição de $v_{t+1} | \mathcal{F}_t$. O mesmo argumento é válido para v_{t+h} com $h \geq 1$, como pode ser comprovado por uma simples inspeção das Equações (2.4.13) e (2.4.17).

2.4.2 Ajuste do Modelo

Como demonstrado em [25], os estimadores de máxima verossimilhança (MLE) dos parâmetros de $\{\mu_t\}_{t \geq 1}$ ($\phi_0, \Phi_1, \dots, \Phi_p$ e Σ) podem ser calculados através da resolução de m problemas de regressão linear múltipla ou OLS (*Ordinary Least Squares*, veja [27]).

Assim, considere um conjunto $A = \{v_{-p+1}, v_{-p+2}, \dots, v_0, \dots, v_1, v_2, v_T\}$ que contenha $n = T + p$ observações de v_t . Além disso, sejam B e w_t , respectivamente, a matriz com todos os parâmetros de

$\{v_t\}_{t \geq 1}$ e o vetor que contém, no tempo t , as p últimas observações de v_t , ou seja,

$$B = [\phi_0 \ \Phi_1 \ \Phi_2 \ \dots \ \Phi_p] \quad (2.4.18)$$

e

$$w'_t = [1 \ v'_{t-1} \ v'_{t-2} \ \dots \ v'_{t-p}]. \quad (2.4.19)$$

Tomando $\beta_i = B'_i$, onde B_i é a i -ésima linha da matriz B , a Equação (2.4.1) pode ser escrita como um conjunto de m equações, correspondentes aos m problemas de OLS mencionados acima, ou seja,

$$v_{it} = a_{it} + w'_t \beta_i, \quad i = 1, \dots, m. \quad (2.4.20)$$

Desta forma, após calcular os estimadores OLS de β_i , denominados b_i (ver Equação (2.4.22)), o estimador da matriz B pode ser obtido ao concatenar os diversos b_i calculados. Assim,

$$\widehat{B} = [\widehat{\phi}_0 \ \widehat{\Phi}_1 \ \widehat{\Phi}_2 \ \dots \ \widehat{\Phi}_p] = [b_1 \ b_2 \ \dots \ b_m]', \quad (2.4.21)$$

onde b_i é dado por

$$b_i = \left(\sum_{t=1}^T w_t w'_t \right)^{-1} \left(\sum_{t=1}^T w_t v_{t,i} \right). \quad (2.4.22)$$

Uma vez calculados os estimadores OLS em \widehat{B} , é possível calcular o estimador MLE $\widehat{\Sigma}$ com base no resíduo \hat{a}_t das regressões lineares. Sendo estes dados por

$$\hat{a}_t = v_t - \widehat{B} w_t \quad (2.4.23)$$

e

$$\widehat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \hat{a}_t \hat{a}'_t. \quad (2.4.24)$$

Como alternativa, os parâmetros de um modelo VAR(p) também podem ser estimados através da função de máxima verossimilhança do modelo. A busca do conjunto de parâmetros que maximiza a verossimilhança pode ser realizada, por exemplo, com o método de Newton ou com métodos quasi-Newton como o BFGS (veja [6] para maiores detalhes).

Uma outra alternativa, ainda baseada na maximização da função de verossimilhança, consiste em escrever o modelo VAR(p) como um modelo no espaço de estados (ou *state space model*) [45, 25, 52] e utilizar o Filtro de Kalman para a avaliação da verossimilhança. Aliado ao método de Newton ou BFGS, esta abordagem é bastante conveniente devido à disponibilidade de pacotes numéricos para sua implementação.

2.5 Modelo GO-GARCH

Seja $\{x_t\}_{t \geq 1}$ o processo estocástico de dimensão m correspondente à equação de volatilidade do modelo de retornos (Equação (2.2.1)) e $\{\mathcal{F}_t\}_{t \geq 0}$ a filtração gerada por ele.

Segundo o modelo GO-GARCH, os retornos podem ser calculados como uma transformação linear não singular de $\{y_t\}_{t \geq 1}$, um processo vetorial não observável, ou seja, para Z uma matriz $m \times m$ não singular,

$$x_t = Z y_t. \quad (2.5.1)$$

Quanto ao processo $\{y_t\}_{t \geq 1}$, o modelo assume o seguinte conjunto de hipóteses

Hipótese 1. O processo estocástico $\{y_t\}_{t \geq 1}$ satisfaz:

- (i) É estritamente estacionário e ergódico,
- (ii) Tem curtose finita,
- (iii) Suas coordenadas (ou fatores) são condicionalmente não correlacionadas, têm esperança condicional nula e variância condicional h_{it} positiva e com esperança unitária (ver [16]), ou seja,

$$\mathbb{E}[y_t | \mathcal{F}_{t-1}] = 0, \quad (2.5.2)$$

$$\mathbb{E}[h_{it}] = 1, \quad h_{it} \geq 0 \quad (2.5.3)$$

e

$$\text{Cov}(y_{it}, y_{jt} | \mathcal{F}_{t-1}) = \begin{cases} h_{it}, & i = j \\ 0, & i \neq j \end{cases}. \quad (2.5.4)$$

Na primeira definição do modelo GO-GARCH [53], seus autores assumiram a hipótese de independência entre as coordenadas de y_t . Depois, como feito em [16], foi adotada a hipótese menos restritiva de componentes condicionalmente não correlacionadas.

O modelo GO-GARCH também define a estrutura de $\{y_t\}_{t \geq 1}$,

$$y_t = H_t^{1/2} \epsilon_t, \quad (2.5.5)$$

onde a matriz H_t , dada por

$$H_t = \text{diag}(h_{1t}, \dots, h_{mt}), \quad (2.5.6)$$

é a matriz de covariância condicional do modelo e $\{\epsilon_t\}_{t \geq 1}$ é uma sequência vetorial martingal de diferenças (ver Definição 4 e [25]) satisfazendo $\mathbb{E}[\epsilon_t | \mathcal{F}_{t-1}] = 0$ e $\text{Var}(\epsilon_t | \mathcal{F}_{t-1}) = 1$. Além disso, as coordenadas de $\{\epsilon_t\}_{t \geq 1}$ formam sequências de variáveis aleatórias iid (*independentes e identicamente distribuídas*).

Definição 4. Sequência Vetorial Martingal de Diferenças

Uma sequência martingal de diferenças é um sequência $\{X_t\}_{t=1}^{\infty}$ de vetores aleatórios $m \times 1$ que satisfaz para todo t :

$$\mathbb{E}[X_t] = 0, \quad (2.5.7)$$

e

$$\mathbb{E}[X_t | \mathcal{F}_{t-1}] = 0. \quad (2.5.8)$$

O modelo assume também que as variâncias condicionais h_{it} seguem modelos univariados do tipo GARCH (como os citados na Seção 2.2). Em [53] é adotado o modelo GARCH(1,1) para a dinâmica dos fatores. Já em [8], é utilizado o modelo APARCH. Na realidade, o modelo GO-GARCH permite que diferentes modelos univariados sejam utilizados para os fatores que compõem y_t . É possível, por exemplo, utilizar ao mesmo tempo o modelo GARCH(1,1) para alguns fatores e o modelo APARCH para outros. Sob certas condições, definidas em [44], é possível recorrer a modelos de volatilidade estocástica (ver [23]).

Como em [53], neste trabalho é adotado o modelo GARCH(1,1) para h_{it} , ou seja,

$$h_{it} = (1 - \alpha_i - \beta_i) + \alpha_i y_{i,t-1}^2 + \beta_i h_{i,t-1}, \quad (2.5.9)$$

com $\alpha_i, \beta_i \geq 0$ e $\alpha_i + \beta_i < 1$. Na Seção 2.3 são apresentadas maiores informações sobre o modelo GARCH(1,1).

As hipóteses de estacionariedade estrita e curtose finita adotadas para $\{y_t\}_{t \geq 1}$ implicam em sua estacionariedade fraca. Consequentemente, $\{x_t\}_{t \geq 1}$ também é fracamente estacionária, pois como demonstram as Equações (2.5.10) e (2.5.11), se $\mathbb{E}[y_t]$ e $\text{Cov}(y_t, y_{t-l})$ são invariantes no tempo, o mesmo valerá para $\{x_t\}_{t \geq 1}$, pois Z é uma matriz constante.

$$\mathbb{E}[x_t] = \mathbb{E}[Zy_t] = Z\mathbb{E}[y_t] = 0 \quad (2.5.10)$$

$$\text{Cov}(x_t, x_{t-l}) = \mathbb{E}[Zy_t y'_{t-l} Z'] = Z\mathbb{E}[y_t y'_{t-l}] Z' = Z\text{Cov}(y_t, y_{t-l})Z' \quad (2.5.11)$$

Sejam Σ_t e Σ as variâncias condicionais e não condicionais de $\{x_t\}_{t \geq 1}$. Ambas podem ser calculadas a partir da Equação (2.5.1), como é feito nas Equações (2.5.12) e (2.5.13).

$$\begin{aligned} \Sigma_t &= \text{Cov}(x_t | \mathcal{F}_{t-1}) \\ &= \mathbb{E}[x_t x'_t | \mathcal{F}_{t-1}] \\ &= \mathbb{E}[Zy_t y'_t Z' | \mathcal{F}_{t-1}] \\ &= Z \mathbb{E}[y_t y'_t | \mathcal{F}_{t-1}] Z' \\ &= ZH_t Z' \end{aligned} \quad (2.5.12)$$

$$\begin{aligned} \Sigma &= \text{Cov}(x_t) \\ &= \mathbb{E}[\text{Cov}(x_t | \mathcal{F}_{t-1})] \\ &= \mathbb{E}[ZH_t Z'] \\ &= Z \mathbb{E}[H_t] Z' \\ &= ZZ' \end{aligned} \quad (2.5.13)$$

Finalmente, considerando a distribuição condicional de x_t , pode-se escrever para uma distribuição \mathcal{D} com média zero e matriz de covariância Σ_t : $x_t | \mathcal{F}_{t-1} \sim \mathcal{D}(0, \Sigma_t)$. Quando as componentes de y_t têm distribuição normal, $x_t | \mathcal{F}_{t-1}$ terá distribuição normal multivariada, ou seja,

$$x_t | \mathcal{F}_{t-1} \sim N(0, \Sigma_t). \quad (2.5.14)$$

2.5.1 Simulação

No caso simples ilustrado pela Equação (2.5.14), a simulação de valores para x_{t+1} com base na informação disponível até o tempo t é realizada através da amostragem do vetor aleatório $x_{t+1} | \mathcal{F}_t$, que tem distribuição normal multivariada com média zero e matriz de covariância Σ_{t+1} . Vale ressaltar que Σ_{t+1} é determinística, pois H_{t+1} é \mathcal{F}_t -mensurável, ou seja, toda a informação necessária para seu cálculo é conhecida no tempo t .

Outra alternativa, com base na equação $x_t = Zy_t$, consiste em realizar a simulação dos fatores y_t e então combiná-los através da multiplicação pela matriz Z .

Sendo H_t uma matriz diagonal, a i -ésima coordenada de y_t pode ser escrita como o produto do termo aleatório ϵ_{it} pelo fator de escala $\sqrt{h_{it}}$, que determina a variância condicional a cada instante, ou seja,

$$y_{it} = \epsilon_{it} \sqrt{h_{it}}. \quad (2.5.15)$$

Desta forma, é possível simular os valores de $y_{i,t+h}$ aplicando a Equação (2.5.15) de forma iterativa. A repetição deste procedimento para $i = 1, \dots, m$ permite a obtenção dos valores de y_t e depois o cálculo de x_t . Assim, adotando o modelo GARCH(1,1), as volatilidades condicionais de cada fator

entre os tempos $t + 1$ e $t + h$ são dadas por

$$h_{i,t+1}|\mathcal{F}_t = (1 - \alpha_i - \beta_i) + \alpha_i y_{it}^2 + \beta_i h_{it}, \quad (2.5.16)$$

$$h_{i,t+2}|\mathcal{F}_t = (1 - \alpha_i - \beta_i) + \alpha_i \{y_{i,t+1}|\mathcal{F}_t\}^2 + \beta_i \{h_{i,t+1}|\mathcal{F}_t\}, \quad (2.5.17)$$

⋮

$$h_{i,t+h}|\mathcal{F}_t = (1 - \alpha_i - \beta_i) + \alpha_i \{y_{i,t+h-1}|\mathcal{F}_t\}^2 + \beta_i \{h_{i,t+h-1}|\mathcal{F}_t\}. \quad (2.5.18)$$

Logo, no mesmo período, a i -ésima coordenada de y_t assume os valores

$$y_{i,t+1}|\mathcal{F}_t = \epsilon_{i,t+1} \sqrt{h_{i,t+1}|\mathcal{F}_t}, \quad (2.5.19)$$

$$y_{i,t+2}|\mathcal{F}_t = \epsilon_{i,t+2} \sqrt{h_{i,t+2}|\mathcal{F}_t}, \quad (2.5.20)$$

⋮

$$y_{i,t+h}|\mathcal{F}_t = \epsilon_{i,t+h} \sqrt{h_{i,t+h}|\mathcal{F}_t}. \quad (2.5.21)$$

Como apresentado na Seção 2.5, as coordenadas de $\{\epsilon_t\}_{t \geq 1}$ formam sequências de variáveis aleatórias iid. Entretanto, a independência existente entre os valores destas sequências não implica na independência das coordenadas de y_t , pois a dependência entre estas pode estar presente no modelo adotado para as variâncias condicionais h_{it} . Isto ocorre, por exemplo, na extensão do modelo GARCH(1,1) que é adotada em [16].

A independência em $\{\epsilon_t\}_{t \geq 1}$, porém, implica na ausência de correlação entre os fatores. Como as Equações (2.5.22) e (2.5.23) demonstram.

$$\begin{aligned} \text{Cov}(y_t|\mathcal{F}_{t-1}) &= \mathbb{E}[y_t y_t'|\mathcal{F}_{t-1}] \\ &= \mathbb{E}\left[H_t^{1/2} \epsilon_t \epsilon_t' H_t^{1/2} \middle| \mathcal{F}_{t-1}\right] \\ &= H_t^{1/2} \mathbb{E}[\epsilon_t \epsilon_t'|\mathcal{F}_{t-1}] H_t^{1/2} \\ &= H_t^{1/2} I_m H_t^{1/2} \\ &= H_t \end{aligned} \quad (2.5.22)$$

$$\text{Cov}(y_t) = \mathbb{E}[\mathbb{E}[y_t y_t'|\mathcal{F}_{t-1}]] = \mathbb{E}[H_t] = I_m \quad (2.5.23)$$

Como feito para os modelos VAR(p) na Seção 2.4, a esperança $\mathbb{E}[x_{t+h}|\mathcal{F}_t]$ é adotada como a previsão para o retorno segundo o modelo GO-GARCH. Em seu cálculo, devido a não existência de uma fórmula fechada para a distribuição de $x_{t+h}|\mathcal{F}_t$, será necessário recorrer à integrações de Monte Carlo [50]. Desta forma, dado um conjunto A com N amostras de $x_{t+h}|\mathcal{F}_t$, a previsão de x_{t+h} é dada por

$$x_t(h) = Z \mathbb{E}[y_{t+h}|\mathcal{F}_t] = Z \left(\frac{1}{N} \sum_{k=1}^N A_k \right), \quad (2.5.24)$$

onde A_k é a k -ésima amostra de $x_{t+h}|\mathcal{F}_t$.

2.5.2 Ajuste do Modelo

Diversos métodos já foram publicados para o ajuste de modelos GO-GARCH. Os primeiros métodos, baseados na maximização de funções de verossimilhança (ver [53, 7, 16]) sofrem dos problemas

comuns a problemas de otimização com muitas variáveis, como a não convergência de métodos numéricos ou a existência de mínimos locais, conforme descrito em [44]. Entretanto, métodos mais recentes foram apresentados para lidar com estas dificuldades (ver [8, 44, 55]).

Entre estes, dois foram escolhidos para estudo: o método de momentos publicado em [44] e o método baseado na análise de componentes independentes, ou ICA para *Independent Component Analysis*, publicado em [8]. O primeiro método, aqui chamado MM2011, foi escolhido devido à sua robustez numérica e o segundo por apresentar os melhores resultados nos testes de estimação realizados em [55]. Além disso, em termos de implementação, os dois são bastante semelhantes.

Estimação em três etapas

O ajuste de um modelo GO-GARCH consiste em estimar Z e os parâmetros dos m modelos univariados de volatilidade que determinam as componentes de y_t .

Ambos os métodos selecionados realizam este ajuste em três etapas. Destas, as duas primeiras se referem ao cálculo de \hat{Z} e a última ao ajuste dos m modelos univariados.

O ajuste de Z tem como base sua decomposição polar, através da qual Z é descrita como o produto de uma matriz positiva definida simétrica S e de uma matriz ortogonal U , ou seja,

$$Z = SU. \quad (2.5.25)$$

Assim, o ajuste de um modelo GO-GARCH ocorre em três etapas:

1. Estimar S (usando a covariância não condicional de x_t)
2. Estimar U (usando o método MM2011 ou ICA)
3. Estimar os parâmetros dos modelos de volatilidade univariados

A matriz S pode ser calculada a partir de Σ , a variância não condicional de x_t , pois

$$\Sigma = ZZ' = (SU)(SU)' = SUU'S' = SS' = S^2. \quad (2.5.26)$$

Assim, recorrendo à decomposição espectral de Σ , S pode ser calculada facilmente, já que

$$\Sigma = PLP' \longrightarrow S = PL^{1/2}P'. \quad (2.5.27)$$

Explorando a relação entre S e Σ , \hat{S} pode ser calculada a partir da matriz de covariância amostral $\hat{\Sigma}$, que é um estimador de Σ calculado com base em uma amostra $A = \{x_1, \dots, x_N\}$, como mostra a Equação (2.5.28).

$$\hat{\Sigma} = \frac{1}{N} \sum_{t=1}^N x_t x_t' \quad (2.5.28)$$

Para completar o ajuste de Z , ainda é necessário estimar U . Esta, entretanto, não pode ser estimada através da variância de x_t , já que a Equação (2.5.26) é válida mesmo quando U é substituída por outras matrizes ortogonais.

Considerando que o estimador \hat{U} para U seja conhecido (a estimação de U será vista nas sessões a seguir), a matriz Z pode ser finalmente estimada:

$$\hat{Z} = \hat{S}\hat{U} \quad (2.5.29)$$

De acordo com as hipóteses do modelo, Z não é singular. Assim, é possível calcular sua inversa e, em seguida, calcular as estimativas para os fatores não observáveis y_t que controlam a dinâmica da

série de retornos $\{x_t\}_{t \geq 1}$, ou seja,

$$\hat{y}_t = \hat{Z}^{-1}x_t. \quad (2.5.30)$$

Finalmente, passa-se à terceira etapa do método de estimação, que consiste em estimar os m processos univariados de volatilidade do modelo GO-GARCH. Este tópico, porém, não é abordado neste trabalho. Veja [25, 52, 56, 24] para um excelente tratamento sobre modelos de volatilidade.

Nas próximas duas sessões, os dois métodos utilizados para a estimação de U são apresentados. Como mencionado anteriormente, esta estimação é a única diferença existente entre os métodos MM2011 e ICA.

2.5.3 Estimador ICA

Considere a série temporal $\{s_t\}_{t \geq 1}$ dos retornos padronizados, cujos elementos são dados por

$$s_t = \Sigma^{-1/2}x_t = S^{-1}x_t. \quad (2.5.31)$$

Como feito em [8], a matriz ortogonal U do modelo GO-GARCH pode ser estimada aplicando a *Análise de Componentes Independentes* (ou ICA) sobre a série dos retornos padronizados.

Assim, considere um vetor aleatório u , m -dimensional. A ICA é um método estatístico que permite descrever u como a combinação linear de um conjunto de fatores com variância unitária que são o mais independentes possível de acordo com alguma métrica. Assim, considerando que os fatores estejam agrupados em um vetor f com $\text{Cov}(f) = I_m$, a ICA determina a matriz B tal que

$$u = Bf. \quad (2.5.32)$$

Voltando à definição do modelo GO-GARCH (Equação (2.5.1)), é fácil ver que s_t também pode ser descrito como uma combinação linear dos fatores y_t , já que

$$s_t = S^{-1}x_t = S^{-1}(SU)y_t = Uy_t. \quad (2.5.33)$$

Desta forma, considerando que as amostras disponíveis de s_t são na verdade realizações de um único vetor aleatório, é possível estimar a matriz U através da ICA, já que tanto o modelo GO-GARCH quanto o modelo da ICA consideram que seus fatores têm variância unitária e são não correlacionados, ou seja, $\text{Cov}(f) = \text{Cov}(y_t) = I_m$.

Ao estimar a matriz U por este método, a estrutura temporal da série de retornos é ignorada. O mesmo, entretanto, não ocorre na estimação de U realizada pelo estimador MM2011, como pode ser visto na Seção 2.5.4.

Implementação do Estimador ICA

A implementação do estimador para U usando ICA não apresenta desafios, pois a implementação da ICA é simples (veja o Apêndice A) e, geralmente, já disponível em vários pacotes.

O Algoritmo 1 apresenta o pseudocódigo de uma função que implementa o estimador ICA para as matrizes do modelo GO-GARCH.

A entrada da função GOGARCH.ICA consiste na matriz X_t , cujas colunas $X_{t,i}$ correspondem às observações do processo x_t feitas entre o tempo $t = 1$ e $t = T$. Desta forma, X_t é uma matriz de tamanho $m \times T$.

Em primeiro lugar, a função calcula o estimador para a matriz S através da decomposição espectral da matriz Σ de covariância amostral (linhas 2 a 5). Em seguida, são calculados os retornos padronizados

como na Equação (2.5.31) (linha 6). Por último, então, são calculadas as estimativas para a matriz U e para a série $\{y_t\}_{t \geq 1}$ (linhas 7 e 8).

O cálculo da estimativa para U no Algoritmo 1 é realizado através da função FastICA, como em [8]. Este método realiza a ICA através de um processo iterativo, utilizando a entropia negativa como medida de independência dos fatores y_t .

No Apêndice A pode ser encontrada uma descrição do modelo ICA e do algoritmo FastICA, incluindo as justificativas teóricas e detalhes de implementação para o mesmo.

Algorithm 1 Algoritmo para estimar as matrizes do modelo GO-GARCH usando ICA

```

1: procedure GOGARCH_ICA( $X_t$ )
2:    $\Sigma := \frac{1}{n} \sum_{i=1}^T X_{t,i} X'_{t,i}$ 
3:    $P := \mathbf{autovetores}(\Sigma)$ 
4:    $L := \mathbf{autovalores}(\Sigma)$ 
5:    $S := PL^{1/2}P'$ 
6:    $s_t := (PL^{-1/2}P') X_t$ 
7:    $U := \mathbf{FastICA}(s_t, m)$ 
8:    $y_t := U' s_t$ 
9:   return  $S, U, y_t$ 
10: end procedure

```

2.5.4 Estimador MM2011

Considere novamente as séries $\{s_t\}_{t \geq 1}$ dos retornos padronizados e $\{y_t\}_{t \geq 1}$ dos fatores não observáveis do modelo GO-GARCH.

O método MM2011 realiza a estimação de U explorando a estrutura de autocorrelação dos processos matriciais $\{S_t\}_{t \geq 1}$ e $\{Y_t\}_{t \geq 1}$, definidos a partir de s_t e y_t por

$$S_t = s_t s'_t - I_m \quad (2.5.34)$$

e

$$Y_t = y_t y'_t - I_m. \quad (2.5.35)$$

Já que $\text{Cov}(y_t) = I_m$ (pela Hipótese 1) e U é ortogonal, ambos os processos matriciais têm média zero como mostram as Equações (2.5.36) e (2.5.37).

$$\mathbb{E}[Y_t] = \mathbb{E}[y_t y'_t - I_m] = \mathbb{E}[y_t y'_t] - I_m = 0 \quad (2.5.36)$$

$$\mathbb{E}[S_t] = \mathbb{E}[s_t s'_t - I_m] = \mathbb{E}[U y_t y'_t U'] - I_m = U \mathbb{E}[y_t y'_t] U' - I_m = 0 \quad (2.5.37)$$

A rigor, o estudo das autocovariâncias entre Y_t e Y_{t-k} deveria ser feito com base na matriz de $m^2 \times m^2$ elementos definida por

$$\text{Cov}(Y_t, Y_{t-k}) = \mathbb{E}[\text{vec}(Y_t) \text{vec}(Y_{t-k})']. \quad (2.5.38)$$

Entretanto, devido ao grande número de variáveis envolvido, as autocovariâncias são modeladas como a matriz $\Gamma_k(y)$, de tamanho $m \times m$, definida por

$$\Gamma_k(y) = \mathbb{E}[Y_t Y_{t-k}']. \quad (2.5.39)$$

As autocorrelações entre Y_t e Y_{t-k} , $\Phi_k(y)$ são definidas em função de suas autocovariâncias, ou seja,

$$\Phi_k(y) = \Gamma_0(y)^{-1/2} \Gamma_k(y) \Gamma_0(y)^{-1/2}. \quad (2.5.40)$$

Além disso, como $S_t = UY_tU'$, já que

$$S_t = s_t s_t' - I_m = U y_t y_t' U' - U U' = U Y_t U', \quad (2.5.41)$$

as autocovariâncias e autocorrelações da série de retornos padronizados também podem ser definidas em função de $\Gamma_k(y)$ e $\Phi_k(y)$, como mostram as Equações (2.5.42) e (2.5.43).

$$\begin{aligned} \Gamma_k(s) &= \mathbb{E}[S_t S_{t-k}] \\ &= \mathbb{E}[(U Y_t U') (U Y_{t-k} U')] \\ &= U \mathbb{E}[Y_t Y_{t-k}] U' \\ &= U \Gamma_k(y) U' \end{aligned} \quad (2.5.42)$$

$$\begin{aligned} \Phi_k(s) &= \Gamma_0(s)^{-1/2} \Gamma_k(s) \Gamma_0(s)^{-1/2} \\ &= \left(U \Gamma_0(y)^{-1/2} U' \right) \left(U \Gamma_k(y) U' \right) \left(U \Gamma_0(y)^{-1/2} U' \right) \\ &= U \Gamma_0(y)^{-1/2} \Gamma_k(y) \Gamma_0(y)^{-1/2} U' \\ &= U \Phi_k(y) U' \end{aligned} \quad (2.5.43)$$

As matrizes $\Phi_k(s)$ e $\Gamma_k(s)$ são de fundamental importância, pois fornecem as relações necessárias para estimar U , já que estas podem ser estimadas a partir dos retornos padronizados, que são observáveis. Uma hipótese adicional, entretanto, é feita:

Hipótese 2. *O processo estocástico $\{y_t\}_{t \geq 1}$ satisfaz:*

$$\min_{1 \leq i \leq m} \max_{1 \leq k \leq p} |\text{Corr}(y_{it}^2, y_{i,t-k}^2)| > 0 \quad (2.5.44)$$

$$\max_{\substack{1 \leq i < j \leq m \\ 1 \leq k \leq p}} |\text{Cov}(y_{it}^2, y_{i,t-k} y_{j,t-k})| = 0 \quad (2.5.45)$$

Dadas as Hipóteses 1 e 2, os únicos elementos não nulos de $\Gamma_k(y)$ são os de sua diagonal. Assim, $\Gamma_k(s)$, $\Phi_k(y)$ e $\Phi_k(s)$ também são diagonais.

Considerar que as matrizes de covariância são diagonais traz duas vantagens. A primeira vantagem consiste em reduzir ainda mais o número de variáveis a serem estimadas. A segunda, mais importante, é fornecer um método para a estimação de U , que pode ser vista como a matriz de autovetores da decomposição espectral de $\Gamma_k(s)$, já que $\Gamma_k(s) = U \Gamma_k(y) U'$ com U ortogonal e $\Gamma_k(y)$ diagonal (ver Equação (2.5.42)). Desta forma, a estimativa de U pode ser dada pelos autovetores de $\hat{\Gamma}_k(s)$ ou de $\hat{\Phi}_k(s)$, que são calculadas a partir de um conjunto de N amostras da série $\{s_t\}_{t \geq 1}$ de retornos padronizados, como mostram as Equações (2.5.46) e (2.5.47).

$$\hat{\Gamma}_k(s) = \frac{1}{N} \sum_{t=k+1}^N S_t S_{t-k} \quad (2.5.46)$$

$$\hat{\Phi}_k(s) = \hat{\Gamma}_0(s)^{-1/2} \hat{\Gamma}_k(s) \hat{\Gamma}_0(s)^{-1/2} \quad (2.5.47)$$

Experimentos de simulação realizados em [44], entretanto, indicam que um estimador mais eficiente para U é obtido ao adotar $\tilde{\Phi}_k$ (ver Equação (2.5.48)), que é uma versão simétrica de $\hat{\Phi}_k(s)$.

$$\tilde{\Phi}_k = \frac{1}{2} \left(\hat{\Phi}_k(s) + \hat{\Phi}_k(s)' \right) \quad (2.5.48)$$

Logo, o estimador para U pode ser definido da seguinte maneira:

Definição 5. Estimador \hat{U}_k

O estimador de lag k para U é a matriz ortogonal \hat{U}_k que diagonaliza a matriz simétrica de autocorrelação $\tilde{\Phi}_k$, calculada a partir de um conjunto de N amostras do processo $\{s_t\}_{t \geq 1}$ de retornos padronizados, ou seja, \hat{U}_k é a matriz de autovetores da decomposição espectral de $\tilde{\Phi}_k$, dada por

$$\tilde{\Phi}_k = \hat{U}_k \hat{D}_k \hat{U}_k' \quad (2.5.49)$$

Uma melhor estimativa para U pode ser obtida, como mostrado em [44], ao utilizar a *Análise de Componentes Principais Comuns* (CPCA) sobre um conjunto $C = \{\tilde{\Phi}_1, \dots, \tilde{\Phi}_p\}$ de matrizes de autocorrelação com diversos lags.

A CPCA é uma extensão da *Análise de Componentes Principais* (PCA) para lidar com um conjunto $A = (A_1, \dots, A_p)$ de p amostras, que procura identificar a estrutura comum a todas elas. Desta forma, sendo Σ_i a matriz de covariância da amostra A_i , a CPCA considera que todas as matrizes Σ_i compartilham um mesmo conjunto V de autovetores.

Como consequência deste compartilhamento de autovetores, todas as matrizes Σ_i são simultaneamente diagonalizáveis por matrizes cujas colunas são os autovetores em V . Logo, qualquer uma destas matrizes seria uma candidata natural a estimador de U , tendo em vista a Definição 5.

Assim, adota-se a hipótese de que, dado o conjunto C , todas as matrizes que este contém podem ser diagonalizadas simultaneamente por um mesmo conjunto de autovetores, que será utilizado como a nova estimativa \hat{U} .

Na prática, entretanto, dificilmente será possível encontrar uma matriz que diagonalize todos os elementos de C simultaneamente. Desta forma, é preciso contentar-se com uma solução aproximada e buscar a matriz que melhor diagonalize as matrizes de autocorrelação. Tal solução aproximada pode ser obtida resolvendo o problema

$$\min_{B \in \mathcal{O}_m} g(B) = \sum_{k=1}^p \sum_{\substack{i,j=1 \\ i \neq j}}^m \left(B \tilde{\Phi}_k B' \right)_{ij}^2, \quad (2.5.50)$$

onde \mathcal{O}_m é o conjunto de matrizes ortogonais de tamanho $m \times m$ e $(A)_{ij}$ é o elemento de sua i -ésima linha e j -ésima coluna da matriz A .

Como feito em [20, 4, 44], as soluções para o Problema (2.5.50) podem ser obtidas através do algoritmo FG, que é explicado em detalhes no Apêndice B.

Quanto às soluções do Problema (2.5.50), analisando a definição do problema, pode-se ver que são matrizes ortogonais que minimizam a soma do quadrado dos elementos fora das diagonais das matrizes $B \hat{\Phi}_k B'$, que deveriam ser nulos pela hipótese de diagonalização simultânea.

Assim, finalmente, o estimador utilizado para U pode ser definido:

Definição 6. Estimador \hat{U}

O estimador de U com base na informação de autocorrelações de lag $k = 1, \dots, p$ do processo vetorial $\{s_t\}_{t \geq 1}$ é a matriz ortogonal \hat{U} obtida ao resolver o Problema (2.5.50).

Implementação do Estimador MM2011

A implementação do estimador para a matriz ortogonal U pelo método de momentos é mais elaborada que a por ICA. Entretanto, também não apresenta desafios, já que o algoritmo FG tem implementação simples e é bastante robusto numericamente (ver Apêndice B).

Assim, considere o pseudocódigo no Algoritmo 2. Este contém a implementação da função GOGARCH_MM2011, que recebe como entrada uma matriz X_t de tamanho $m \times T$ cujas colunas são as observações do processo x_t até o tempo $t = T$.

De forma geral, três operações principais são realizadas na função. Em primeiro lugar, a função calcula a matriz S através da decomposição espectral da matriz Σ de covariância amostral (linhas 2 a 5). Em seguida, o conjunto $C = \{\tilde{\Phi}_1, \dots, \tilde{\Phi}_p\}$ é calculado (linhas 6 a 21). E, por último, a estimativa da matriz U é calculada aplicando o algoritmo FG às matrizes do conjunto C (linha 22).

Como pode ser visto no pseudocódigo de GOGARCH_MM2011, o cálculo das matrizes do conjunto C envolve vários passos intermediários. Sendo estes o cálculo dos retornos padronizados (linha 6), o cálculo do processo matricial S_t (linhas 7 a 9) e das matrizes de autocovariância $\hat{\Gamma}_k$ (linhas 10 a 12) e de autocorrelação $\hat{\Phi}_k$ (linhas 13 a 18).

Algorithm 2 Algoritmo GO-GARCH MM2011

```

1: procedure GOGARCH_MM2011( $X_t, p$ )
2:    $\Sigma := \frac{1}{n} \sum_{i=1}^T X_{t,i} X_{t,i}'$ 
3:    $P := \text{autovetores}(\Sigma)$ 
4:    $L := \text{autovalores}(\Sigma)$ 
5:    $S := PL^{1/2}P'$ 
6:    $s_t := (PL^{-1/2}P') X_t$ 
7:   for  $i := 1, \dots, T$  do
8:      $S_{t,i} := s_{t,i} s_{t,i}' - I_m$ 
9:   end for
10:  for  $k := 0, \dots, p$  do
11:     $\hat{\Gamma}_k := \frac{1}{T} \sum_{i=1}^T S_t S_{t-k}$ 
12:  end for
13:   $P := \text{autovetores}(\hat{\Gamma}_0)$ 
14:   $L := \text{autovalores}(\hat{\Gamma}_0)$ 
15:   $M := PL^{-1/2}P'$ 
16:  for  $k := 0, \dots, p$  do
17:     $\hat{\Phi}_k := M \hat{\Gamma}_k M$ 
18:  end for
19:  for  $k := 0, \dots, p$  do
20:     $\tilde{\Phi}_k := \frac{1}{2} (\hat{\Phi}_k + \hat{\Phi}_k')$ 
21:  end for
22:   $U := \mathbf{F}(\tilde{\Phi}_1, \dots, \tilde{\Phi}_p; T, \dots, T)$ 
23:   $y_t := U' s_t$ 
24:  return  $S, U, y_t$ 
25: end procedure

```

2.5.5 Resíduos Padronizados

Uma vez realizado o ajuste de um modelo GO-GARCH, é possível calcular a série temporal $\{z_t\}_{t \geq 1}$ dos seus resíduos padronizados através da equação

$$z_t = \Sigma_t^{-1/2} x_t, \quad (2.5.51)$$

onde Σ_t é matriz de covariância condicional do modelo, definida anteriormente.

De forma geral, a análise da série de resíduos permite determinar o quão adequado um modelo ajustado é. Os resíduos de um modelo bem ajustado têm média nula, componentes não correlacionadas e não apresentam autocorrelações no primeiro e no segundo momento. Veja [52, 15].

Neste aspecto, os modelos GO-GARCH não são exceção. É possível demonstrar, como feito a seguir, que:

$$(i) \mathbb{E}[z_t] = 0,$$

$$(ii) \text{Cov}(z_t) = I_m.$$

Em primeiro lugar, é mostrado que $\mathbb{E}[z_t | \mathcal{F}_{t-1}] = 0$. Pela Hipótese 1 do modelo GO-GARCH, $\mathbb{E}[y_t | \mathcal{F}_{t-1}] = 0$. Desta forma, $\mathbb{E}[x_t | \mathcal{F}_{t-1}] = 0$, pois

$$\mathbb{E}[x_t | \mathcal{F}_{t-1}] = \mathbb{E}[Z y_t | \mathcal{F}_{t-1}] = Z \mathbb{E}[y_t | \mathcal{F}_{t-1}] = 0. \quad (2.5.52)$$

Então, vale que $\mathbb{E}[z_t | \mathcal{F}_{t-1}] = 0$, já que

$$\mathbb{E}[z_t | \mathcal{F}_{t-1}] = \mathbb{E}\left[\Sigma_t^{-1/2} x_t \middle| \mathcal{F}_{t-1}\right] = \Sigma_t^{-1/2} \mathbb{E}[x_t | \mathcal{F}_{t-1}] = 0. \quad (2.5.53)$$

Com este resultado, a propriedade (i) está comprovada, pois

$$\mathbb{E}[z_t] = \mathbb{E}[\mathbb{E}[z_t | \mathcal{F}_{t-1}]] = \mathbb{E}[0] = 0. \quad (2.5.54)$$

A propriedade (ii) pode ser comprovada de forma semelhante. Em primeiro lugar, é comprovado que $\text{Cov}(z_t | \mathcal{F}_{t-1}) = I_m$. Veja a Equação (2.5.55).

$$\begin{aligned} \text{Cov}(z_t | \mathcal{F}_{t-1}) &= \mathbb{E}[z_t z_t' | \mathcal{F}_{t-1}] \\ &= \mathbb{E}\left[\Sigma_t^{-1/2} x_t x_t' \Sigma_t^{-1/2} \middle| \mathcal{F}_{t-1}\right] \\ &= \Sigma_t^{-1/2} \mathbb{E}[x_t x_t' | \mathcal{F}_{t-1}] \Sigma_t^{-1/2} \\ &= \Sigma_t^{-1/2} \Sigma_t \Sigma_t^{-1/2} \\ &= I_m \end{aligned} \quad (2.5.55)$$

Assim, a validade da propriedade (ii) pode ser provada por

$$\text{Cov}(z_t) = \mathbb{E}[z_t z_t'] = \mathbb{E}[\mathbb{E}[z_t z_t' | \mathcal{F}_{t-1}]] = \mathbb{E}[\text{Cov}(z_t | \mathcal{F}_{t-1})] = \mathbb{E}[I_m] = I_m. \quad (2.5.56)$$

As propriedades (i) e (ii) são importantes por fornecerem dois testes simples para a verificação do ajuste de um modelo GO-GARCH. Com este propósito, o teste de Hosking, apresentado na Seção 2.1, também pode ser utilizado para verificar a existência de autocorrelações no primeiro ou no segundo momento da série de resíduos.

2.5.6 Modelo O-GARCH

O modelo GO-GARCH, como seu nome indica, pode ser visto como uma generalização do modelo O-GARCH (*Orthogonal GARCH*). Ambos os modelos descrevem o processo $\{x_t\}_{t \geq 1}$ como uma combinação linear de processos não correlacionados $\{y_t\}_{t \geq 1}$, ou seja, ambos assumem que

$$x_t = Zy_t. \quad (2.5.57)$$

Entretanto, os modelos adotam características diferentes para a matriz Z . O modelo GO-GARCH, como visto na Seção 2.5, considera que Z é uma matriz não singular. Enquanto isso, o modelo O-GARCH adota uma característica mais restritiva ao considerar que Z é uma matriz ortogonal.

A maior diferença entre os modelos, porém, se encontra nos estimadores adotados para Z . No modelo O-GARCH, suas colunas são os autovetores ortonormais da matriz de covariância amostral, ou seja, a matriz Z corresponde à matriz usualmente adotada para realizar a análise de componentes principais. Desta forma, maiores problemas de identificação de parâmetros podem ser encontrados, pois a estimação dos mesmos é feita apenas com base em informação não condicional. Para maiores detalhes, consulte [53, 2].

Capítulo 3

Portfólios com Risco Mínimo

Este capítulo trata da avaliação objetiva do risco de perdas financeiras e da criação de portfólios de ativos que tenham risco mínimo.

Na Seção 3.1 são apresentadas as características comuns às boas medidas de risco e também as medidas *Value at Risk* (VaR) e *Conditional Value at Risk* (CVaR). Em seguida, na Seção 3.2, é apresentado um método robusto para a estimação do CVaR de um portfólio. Por último, na Seção 3.3, a estratégia usada para a estimação do CVaR é adaptada para a criação de portfólios de ativos com risco mínimo.

3.1 Medidas de Risco

No trabalho [3], que se tornou referência, são encontradas as definições necessárias para representar matematicamente o risco associado a um ativo financeiro e também para a definição de medidas objetivas de risco.

Sejam o espaço de probabilidade definido por $(\Omega, \mathcal{F}, \mathbb{P})$ e uma variável aleatória $X : \Omega \rightarrow \mathbb{R}$, tal que para qualquer $\omega \in \Omega$, $X(\omega)$ é o valor futuro descontado de um ativo (ou de um portfólio) no tempo T , ou seja, o *risco* associado ao ativo (ou portfólio) no período definido entre as datas t e $t + T$ é modelado como uma variável aleatória no espaço de probabilidade acima. Além disso, seja \mathcal{G} o conjunto de todas as variáveis aleatórias de Ω em \mathbb{R} que modelam o valor futuro de algum ativo financeiro.

Neste contexto, uma *medida de risco* é uma função $\rho : \mathcal{G} \rightarrow \mathbb{R}$, que indica o nível de risco de um ativo utilizando apenas um número real.

O valor $\rho(X)$ é interpretado de duas maneiras distintas. Quando $\rho(X) > 0$, $\rho(X)$ indica a quantidade de capital extra que deve ser investida de uma forma livre de risco para atingir níveis aceitáveis de risco. Entretanto, quando $\rho(X) < 0$, $\rho(X)$ indica a quantidade de capital que pode ser removida da posição atual, mantendo o risco em níveis aceitáveis.

Segundo os autores deste modelo, a representação do risco através de um único valor real acarreta em uma grande perda de informação, porém, simplifica a tomada de decisões ao indicar se um determinado nível de risco é aceitável ou não, ou seja, esta representação é adequada para que $\rho(X)$ desempenhe o papel esperado de uma medida de risco.

Finalmente, na Definição 7, são apresentadas as propriedades do que poderiam ser chamadas de "boas medidas de risco" e, em seguida, a interpretação de algumas delas.

Definição 7. *Medida Coerente de Risco*

Uma medida de risco ρ é dita coerente quando esta satisfaz as seguintes propriedades:

- (i) *Invariância a Translação: Para todo $X \in \mathcal{G}$ e todo $\alpha \in \mathbb{R}$, vale que $\rho(X + \alpha) = \rho(X) - \alpha$.*

(ii) *Subaditividade:* Para todos $X, Y \in \mathcal{G}$, $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

(iii) *Homogeneidade Positiva:* Para todo $X \in \mathcal{G}$ e todo $\lambda \in \mathbb{R}$, temos que $\rho(\lambda X) = \lambda \rho(X)$.

(iv) *Monotonicidade:* Para todo $X, Y \in \mathcal{G}$, com $Y \geq X$, $\rho(Y) \leq \rho(X)$.

Algumas das propriedades podem parecer pouco intuitivas, assim, interpretá-las pode ser de grande valia para auxiliar no seu entendimento. A invariância a translações indica que ao adicionarmos o capital α ao portfólio X , o risco do mesmo diminui em uma quantidade igual a α . Como $\rho(X)$ mede a quantidade de capital que deve ser adicionada ao portfólio, após a adição de tal capital é natural que $\rho(X)$ diminua em α .

Já a subaditividade indica que o risco do portfólio formado por X e Y não pode ser maior que a soma dos riscos individuais destes ativos. Essa propriedade ajuda a modelar o princípio da diversificação, que diz que ao investir em uma quantidade maior de ativos, o risco do portfólio diminui. Ver [14].

Finalmente, a monotonicidade nos mostra que para dois ativos X e Y , se o valor de Y é sempre maior que o valor de X , então seu risco deve ser sempre menor que o risco de X .

Com isso, está concluída a introdução sobre a modelagem de medidas de risco e sobre as propriedades exigidas das mesmas. Assim, nas seções 3.1.1 e 3.1.2, são apresentadas as medidas de risco selecionadas para este trabalho: VaR e CVaR.

3.1.1 Value at Risk (VaR)

O *Value at Risk*, ou simplesmente VaR, não é uma medida coerente de risco pois não satisfaz à propriedade de subaditividade. Veja [33, 14, 1].

Entretanto, a medida VaR é a medida de risco mais difundida no mercado, devido às suas características positivas, como a aplicabilidade para diferentes tipos de ativos financeiros e as facilidades de cálculo, de uso e de verificação (*backtesting*).

A definição encontrada em [33] diz que a medida VaR indica a pior perda em um horizonte de tempo pré-especificado com um nível de confiança β . Como exemplo, considere que o VaR diário (com horizonte de um dia) e nível de confiança $\beta = 99\%$ para um determinado ativo é igual a R\$ 10 milhões. Isto indica que para este ativo existe 1 chance em 100 da perda de um dia para o outro ser maior que R\$ 10 milhões.

Como feito em [49], o VaR é definido formalmente utilizando a *função de perda* associada a um *portfólio de ativos*. Nas definições 8 e 9 estes dois conceitos são formalizados.

Definição 8. Portfólio de Ativos

Seja $A = \{A_1, \dots, A_m\}$ um conjunto formado por m ativos financeiros e $x = (x_1, \dots, x_m)$ um vetor m -dimensional que indica a quantidade de cada um dos ativos A_i em A . Neste contexto, um portfólio Π é definido como o par

$$\Pi := (A, x). \quad (3.1.1)$$

Definição 9. Função de Perda

Seja t a data corrente, $\Pi = (A, x)$ um portfólio como o da Definição 8 e $Y = (Y_1, \dots, Y_m)$ um vetor cujas coordenadas indicam o valor descontado de cada um dos ativos em A no tempo $t + T$. A função de perda $f(x, Y)$ indica a perda no valor total do portfólio Π no intervalo de tempo definido pelas datas t e $t + T$.

Considere agora a probabilidade $\Psi(x, \alpha)$, que é dada por

$$\Psi(x, \alpha) = \mathbb{P}(f(x, Y) \leq \alpha), \quad (3.1.2)$$

e que corresponde à probabilidade da perda ser menor que α . Como uma função de α , para x fixo, $\Psi(x, \alpha)$ é a função de distribuição acumulada para a perda associada a x , ou seja, para x fixo, $f(x, Y)$ é uma variável aleatória que toma valores em \mathbb{R} e que tem sua distribuição induzida pela distribuição do vetor Y .

Finalmente, com base na probabilidade $\Psi(x, \alpha)$, o VaR é definido formalmente na Definição 10.

Definição 10. *Value at Risk (VaR)*

Seja $\Pi = (A, x)$ um portfólio como o da Definição 8. O Value at Risk com nível de confiança β para este portfólio, $\alpha_\beta(x)$, é definido por

$$\alpha_\beta(x) = \min\{\alpha \in \mathbb{R} : \Psi(x, \alpha) \geq \beta\}. \quad (3.1.3)$$

Pela Definição 10, fica claro que $\alpha_\beta(x)$ é simplesmente um quantil de nível β da função de distribuição de perdas $\Psi(x, \alpha)$ do portfólio Π . Como exemplo, considere a Figura 3.1, onde está indicado que a perda do portfólio será menor que $\alpha_{95\%}(x)$ em 95% dos casos.

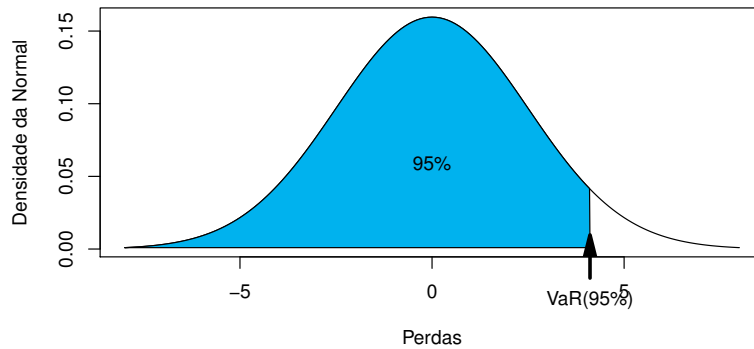


Figura 3.1: VaR como um quantil da distribuição das perdas.

A seguir, como exemplos, são apresentados os valores de VaR para as distribuições Normal e t de Student.

VaR para as Distribuições Normal e t de Student

Considere um ativo A , com valor futuro descontado indicado pela variável aleatória $Y \in \mathcal{G}$, e com perdas $f(Y) \sim N(\mu, \sigma^2)$. Seja α_β o VaR com nível de confiança β para o ativo A . Desta forma, de acordo com a Definição 3.1.3,

$$\Psi(x, \alpha_\beta) = \mathbb{P}(f(Y) \leq \alpha_\beta) = \beta. \quad (3.1.4)$$

Logo,

$$\mathbb{P}\left(\frac{f(Y) - \mu}{\sigma} \leq \frac{\alpha_\beta - \mu}{\sigma}\right) = \beta \quad (3.1.5)$$

e, como $f(Y)$ tem distribuição normal,

$$\frac{\alpha_\beta - \mu}{\sigma} = \Phi^{-1}(\beta). \quad (3.1.6)$$

Finalmente,

$$\alpha_\beta = \sigma \Phi^{-1}(\beta) + \mu, \quad (3.1.7)$$

onde Φ é a função de distribuição acumulada de uma variável aleatória normal padrão.

Mais uma vez, o resultado obtido na Equação (3.1.7) mostra que o VaR é apenas um quantil da distribuição de perdas do ativo A .

Repare que o resultado acima também é válido para portfólios contendo mais de um ativo. Quando Y tem distribuição normal multivariada, $f(x, Y)$ terá, geralmente, uma distribuição normal univariada, pois a soma ponderada de variáveis aleatórias normais é uma variável aleatória normal.

Séries temporais de retornos financeiros têm, em geral, caudas pesadas. Veja [52]. Desta forma, por também apresentar esta característica, a distribuição t de Student é bastante utilizada na modelagem de retornos financeiros.

O cálculo do VaR para a distribuição t é similar ao realizado para a distribuição Normal. Como anteriormente, o VaR é dado pelo quantil de nível β , ou seja,

$$\alpha_\beta = t_\nu^{-1}(\beta), \quad (3.1.8)$$

onde t_ν é a densidade de probabilidade da distribuição t com ν graus de liberdade.

3.1.2 Conditional Value at Risk (CVaR)

Nesta seção é apresentada a medida *Conditional Value at Risk* (ou CVaR), também conhecida como *Expected Shortfall*. Veja [1].

Dada a ampla utilização da medida VaR, a comparação com a mesma é inevitável, principalmente pelo fato da medida CVaR corrigir dois problemas apresentados pela primeira.

O primeiro destes problemas é de caráter teórico. Ao contrário da medida VaR, CVaR é uma medida coerente de risco. Desta forma, é mais apropriada para gerenciar o risco financeiro.

O segundo problema se refere à informação fornecida por cada uma das medidas. Considere o VaR α_β . Este indica que a probabilidade de uma perda ser menor α_β é igual a β . Entretanto, nenhuma informação é dada sobre o valor da perda quando a mesma é maior que α_β , o que ocorre com probabilidade $1 - \beta$. Esta característica torna a medida perigosa para gerenciar portfólios contendo ativos com distribuições que apresentam caudas pesadas ou ativos com dependência de cauda.

A medida CVaR complementa a informação da medida VaR informando o valor esperado da perda que ocorre com a frequência $1 - \beta$, ou seja, informando o valor esperado da perda, uma vez que esta é maior que o VaR α_β .

De certa forma, como dito em [14], é possível dizer que a medida CVaR leva em consideração o formato da cauda da distribuição das perdas, enquanto VaR não o faz. Como ilustração desta afirmação, considere as figuras 3.1 e 3.2. Comparando as duas figuras fica claro que as medidas utilizam intervalos de valores de perda diferentes.

Finalmente, considere a Definição 11, onde a medida CVaR é definida.

Definição 11. *Conditional Value at Risk (CVaR)*

Seja $\Pi = (A, x)$ um portfólio de ativos financeiros. O valor da medida CVaR com nível de confiança β para o portfólio Π , $\phi_\beta(x)$, é dado por

$$\phi_\beta(x) = \mathbb{E}[f(x, Y) | f(x, Y) \geq \alpha_\beta(x)], \quad (3.1.9)$$

ou, de forma equivalente, pela integral

$$\phi_\beta(x) = \frac{1}{1 - \beta} \left(\int_{\Omega} f(x, Y) \mathbb{1}_{[f(x, Y) \geq \alpha_\beta(x)]} d\mathbb{P} \right), \quad (3.1.10)$$

onde $\alpha_\beta(x)$ é o VaR com nível de confiança β , como na Definição 3.1.3.

Como a definição acima deixa evidente, uma das dificuldades no uso de CVaR é sua dependência em VaR. Esta dependência torna estimativas de CVaR menos precisas, pois existem duas possíveis fontes de erro: a estimação de $\alpha_\beta(x)$ e também a estimação de $\phi_\beta(x)$. Veja [14].

Outra dificuldade associada à medida CVaR é sua falta de robustez. Em [12], foi demonstrado que medidas de risco que têm a propriedade de subaditividade não podem ser robustas. Em particular, foram demonstradas a falta de robustez de CVaR e a robustez de VaR.

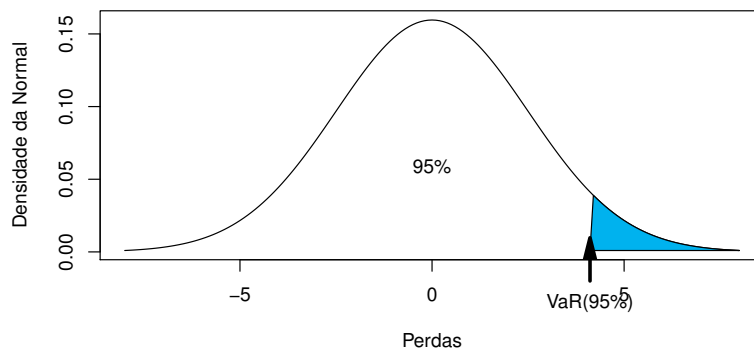


Figura 3.2: Valores de perda usado no cálculo de CVaR.

Além disso, estimações de CVaR são difíceis pois estas são baseadas em amostras da cauda da distribuição de $f(x, Y)$, como indicado na Figura 3.2. Já que o número de amostras na cauda da distribuição geralmente não é grande, é improvável que as estimativas feitas com estas amostras sejam precisas.

Assim, enquanto VaR é frequentemente calculado usando dados de mercado, para CVaR geralmente é necessário adotar um modelo para a distribuição das caudas de $f(x, Y)$ e, então, ajustar o mesmo aos dados disponíveis. Ver [14].

Apesar desta dificuldade, conforme indicado em [49], CVaR tem uma propriedade fundamental para a otimização de portfólios: é uma função convexa. A importância deste fato fica clara ao considerar a quantidade de métodos existentes na literatura para a otimização de funções convexas.

CVaR para a Distribuição Normal e t de Student

Como feito na Seção 3.1.1, considere um ativo A ao qual está associada uma função de perdas $f(Y)$ com distribuição normal de média μ e variância σ^2 . Partindo da Equação (3.1.10) e utilizando uma simples mudança de variáveis é possível chegar, sem grandes desafios, à fórmula para o CVaR normal publicada em [37]:

$$\phi_\beta(\alpha_\beta) = \mu + \left[\frac{\frac{1}{\sigma} \phi\left(\frac{\alpha_\beta - \mu}{\sigma}\right)}{1 - \Phi\left(\frac{\alpha_\beta - \mu}{\sigma}\right)} \right] \sigma^2 \quad (3.1.11)$$

onde ϕ e Φ são, respectivamente, a densidade e a função de distribuição acumulada de uma variável aleatória normal padrão.

Para a distribuição normal padrão, a Equação (3.1.11) se resume a

$$\phi_{\beta}(\alpha_{\beta}) = \frac{\phi(\alpha_{\beta})}{1 - \Phi(\alpha_{\beta})}. \quad (3.1.12)$$

Em [39], são apresentadas fórmulas fechadas para calcular CVaR para as distribuições de probabilidade mais utilizadas. Entre estas, encontra-se a distribuição t de Student, cuja fórmula para o cálculo de CVaR é dada por

$$\phi_{\beta}(\alpha_{\beta}) = \frac{\nu + \alpha_{\beta}^2}{\nu + 1} \frac{t_{\nu}(\alpha_{\beta})}{1 - T_{\nu}(\alpha_{\beta})}, \quad (3.1.13)$$

onde t_{ν} é a densidade de probabilidade da distribuição t com ν graus de liberdade e T_{ν} é sua função de probabilidade acumulada.

3.2 Estimação do CVaR

Na Definição 11, o CVaR de um portfólio é definido em função de seu VaR. Desta forma, é natural considerar que a estimação do primeiro deve ser realizada em duas etapas. Entretanto, a seguir é apresentado um estimador para o CVaR que independe de estimativas prévias para o VaR.

Ao eliminar a primeira etapa, na qual o VaR do portfólio seria estimado, obtém-se um procedimento de estimação mais simples e, além disso, também elimina-se uma das possíveis fontes de erro.

Considere um portfólio Π , definido de acordo com a Definição 8. Como mostrado em [49], tanto o VaR quanto o CVaR podem ser definidos a partir da função $F_{\beta}(x, \alpha)$, dada por

$$F_{\beta}(x, \alpha) = \alpha + \frac{1}{1 - \beta} \int_{\Omega} [f(x, Y) - \alpha]^+ d\mathbb{P}. \quad (3.2.1)$$

Mais precisamente, o VaR e o CVaR de Π são definidos em função dos pontos de mínimo de F_{β} , ou seja, em função dos pontos do conjunto A_{β} que é definido como

$$A_{\beta} = \underset{\alpha}{\operatorname{argmin}} F_{\beta}(x, \alpha). \quad (3.2.2)$$

Assim, como será provado a seguir, o VaR é dado por

$$\alpha_{\beta}(x) = \min \{A_{\beta}\} \quad (3.2.3)$$

e o CVaR por

$$\phi_{\beta}(x) = F_{\beta}(x, \alpha_{\beta}(x)). \quad (3.2.4)$$

De acordo com [49], F_{β} é convexa e continuamente diferenciável em relação ao parâmetro α . Desta forma, os elementos de A_{β} correspondem aos pontos em que a derivada de F_{β} em relação a α é nula. Conforme comprovado na Equação (3.2.8), esta derivada é dada por

$$\frac{\partial F_{\beta}}{\partial \alpha} = \frac{\Psi(x, \alpha) - \beta}{1 - \beta}. \quad (3.2.5)$$

Logo, o conjunto A_{β} é dado por

$$A_{\beta} = \{\alpha : \Psi(x, \alpha) = \beta\}, \quad (3.2.6)$$

já que

$$\frac{\partial F_{\beta}}{\partial \alpha} = 0 \longrightarrow \Psi(x, \alpha) = \beta, \quad (3.2.7)$$

ou seja, os elementos de A_β são os quantis de nível β da distribuição da função de perdas $f(x, Y)$.

De forma geral, espera-se que exista apenas um elemento em A_β . Porém, pode haver mais elementos, caso existam conjuntos de valores da função de perda com medida nula. Por esta razão, o VaR é definido como o menor elemento em A_β . Com isso, está comprovada a Equação (3.2.3).

$$\begin{aligned}
\frac{\partial F_\beta}{\partial \alpha} &= \frac{\partial}{\partial \alpha} \left(\alpha + (1 - \beta)^{-1} \int_{\Omega} [f(x, Y) - \alpha]^+ d\mathbb{P} \right) \\
&= 1 + (1 - \beta)^{-1} \frac{\partial}{\partial \alpha} \int_{\Omega} [f(x, Y) - \alpha]^+ d\mathbb{P} \\
&= 1 + (1 - \beta)^{-1} (\Psi(x, \alpha) - 1) \\
&= \frac{\Psi(x, \alpha) - \beta}{1 - \beta}
\end{aligned} \tag{3.2.8}$$

Na Equação (3.2.8), foi utilizado o fato da função

$$G(\alpha) = \int_{\Omega} [f(x, Y) - \alpha]^+ d\mathbb{P}, \tag{3.2.9}$$

que também é convexa e continuamente diferenciável, ter derivada

$$G'(\alpha) = \Psi(x, \alpha) - 1, \tag{3.2.10}$$

como demonstrado na Equação (3.2.11).

$$\begin{aligned}
\frac{dG}{d\alpha} &= \frac{d}{d\alpha} \left(\int_{\Omega} [f(x, Y) - \alpha]^+ d\mathbb{P} \right) \\
&= \frac{d}{d\alpha} \left(\int_{\Omega} [f(x, Y) - \alpha] \mathbb{1}_{[f(x, Y) \geq \alpha]} d\mathbb{P} \right) \\
&= \frac{d}{d\alpha} \left(\int_{\Omega} f(x, Y) \mathbb{1}_{[f(x, Y) \geq \alpha]} d\mathbb{P} \right) - \frac{d}{d\alpha} \left(\alpha \int_{\Omega} \mathbb{1}_{[f(x, Y) \geq \alpha]} d\mathbb{P} \right) \\
&= 0 - \int_{\Omega} \mathbb{1}_{[f(x, Y) \geq \alpha]} d\mathbb{P} \\
&= -\mathbb{P}[f(x, Y) \geq \alpha] \\
&= -1 + \mathbb{P}[f(x, Y) \leq \alpha] \\
&= \Psi(x, \alpha) - 1
\end{aligned} \tag{3.2.11}$$

Agora, seja α_β um elemento de A_β . Como demonstrado na Equação (3.2.12), $F_\beta(x, \alpha_\beta) = \phi_\beta(x)$. Este resultado comprova a Equação (3.2.4) e também indica como estimar o CVaR de forma

independente do VaR.

$$\begin{aligned}
F_\beta(x, \alpha_\beta) &= \alpha_\beta + \frac{1}{1-\beta} \int_{\Omega} [f(x, Y) - \alpha_\beta]^+ d\mathbb{P} \\
&= \alpha_\beta + \frac{1}{1-\beta} \int_{\Omega} [f(x, Y) - \alpha_\beta] \mathbb{1}_{[f(x, Y) \geq \alpha_\beta]} d\mathbb{P} \\
&= \alpha_\beta + \frac{1}{1-\beta} \left(\int_{\Omega} f(x, Y) \mathbb{1}_{[f(x, Y) \geq \alpha_\beta]} d\mathbb{P} \right) - \frac{\alpha_\beta}{1-\beta} \left(\int_{\Omega} \mathbb{1}_{[f(x, Y) \geq \alpha_\beta]} d\mathbb{P} \right) \\
&= \alpha_\beta + \phi_\beta(x) - \alpha_\beta \\
&= \phi_\beta(x).
\end{aligned} \tag{3.2.12}$$

Como o valor do CVaR é o mesmo para todos os elementos α_β de A_β , o mesmo pode ser calculado diretamente através da minimização da função F_β , ou seja,

$$\phi_\beta(x) = \min_{\alpha} F_\beta(x, \alpha). \tag{3.2.13}$$

O Problema (3.2.13), entretanto, não fornece um estimador adequado para o CVaR, já que, na prática, a distribuição da função de perdas $f(x, Y)$ pode ser desconhecida ou de difícil avaliação. Nestes casos, o cálculo da esperança em F_β pode não ser viável. Como alternativa, F_β pode ser aproximada por uma integração de Monte Carlo.

Considere, então, q amostras y_1, \dots, y_q do vetor $Y \in \mathcal{G}^m$. A aproximação por integração de Monte Carlo para F_β é dada por

$$\tilde{F}_\beta(x, \alpha) = \alpha + \frac{1}{q(1-\beta)} \sum_{k=1}^q [f(x, y_k) - \alpha]^+, \tag{3.2.14}$$

onde, como de costume, a integral da esperança é substituída por um somatório sobre o conjunto de amostras.

Assim, utilizando a aproximação \tilde{F}_β , é obtido o estimador para o CVaR do portfólio Π , dado por

$$\hat{\phi}_\beta(x) = \min_{\alpha} \tilde{F}_\beta(x, \alpha). \tag{3.2.15}$$

Repare que ao adotar \tilde{F}_β , não é necessário fazer hipóteses sobre a distribuição de Y . Entretanto, é preciso que o número de amostras q seja grande para que as estimativas sejam confiáveis. Além disso, o CVaR também pode ser estimado a partir de um conjunto de amostras da função de perda $f(x, Y)$ e não, necessariamente, a partir de amostras de Y .

Como ocorre para a expressão original, a aproximação \tilde{F}_β também é convexa com respeito a α . Além disso, é linear por partes. Desta forma, seus pontos de mínimo podem ser encontrados com certa facilidade ao recorrer a métodos numéricos.

Como mostrado em [49], é possível calcular as estimativas $\hat{\phi}_\beta$ por programação linear. Para isso, o somatório existente na expressão de \tilde{F}_β ,

$$\sum_{k=1}^q [f(x, y_k) - \alpha]^+, \tag{3.2.16}$$

é substituído por um conjunto $U = \{u_1, \dots, u_q\}$ de variáveis de folga e de restrições. Desta forma, é

obtido o Problema (3.2.17), cujas soluções fornecem as estimativas de CVaR utilizadas neste trabalho.

$$\left\{ \begin{array}{l} \min_{\alpha} \alpha + \frac{1}{q(1-\beta)} \sum_{k=1}^q u_k \\ \text{Sujeito a:} \\ u_k \geq 0 \\ u_k + \alpha - f(x, y_k) \geq 0 \end{array} \right. \quad (3.2.17)$$

3.3 Portfólios com CVaR Mínimo

Nesta seção é abordado o problema de *alocação de capital* ou de *seleção de portfólio*, que consiste em determinar a quantidade de capital a ser investida em cada ativo de um portfólio. Veja [41, 42].

Assim, considere um portfólio $\Pi = (A, x)$, definido por um conjunto $A = \{A_1, \dots, A_m\}$ de ativos e pelo vetor real $x = (x_1, \dots, x_m)$, que indica a quantidade de cada ativo em Π . Neste contexto, o problema de alocação de capital consiste em determinar o vetor x de forma a satisfazer certos critérios.

Como em [49], neste trabalho a alocação de capital é feita de forma a minimizar o risco do portfólio Π , avaliado pela medida CVaR, e de forma a satisfazer uma expectativa de rendimento mínimo dada pela função $\mu(x)$.

De forma geral, o problema de alocação de capital pode ser definido como o Problema (3.3.1).

$$\left\{ \begin{array}{l} \min_{(x, \alpha)} F_{\beta}(x, \alpha) \\ \text{Sujeito a:} \\ \sum_{j=1}^m x_j = 1, \text{ onde } x_j \geq 0 \\ \mu(x) \geq R \end{array} \right. \quad (3.3.1)$$

Repare que a primeira restrição do Problema (3.3.1) indica que a soma dos elementos de x tem a unidade como resultado, ou seja, cada componente x_i representa a fração do capital total que deve ser investida no ativo A_i . Além disso, repare que a expectativa de rendimento mínimo deve ser superior ao valor real R , a ser definido pelo investidor.

Considere agora que o conjunto A é formado apenas por ações. Neste caso, tanto a função de perdas $f(x, Y)$ quanto o rendimento esperado $\mu(x)$ podem ser calculados a partir de variáveis aleatórias que modelam os retornos das ações em A . Assim, as perdas do portfólio Π são dadas por

$$f(x, Y) = - \sum_{j=1}^m x_j r_j \quad (3.3.2)$$

e o rendimento esperado por

$$\mu(x) = \sum_{j=1}^m x_j \mathbb{E}[r_j], \quad (3.3.3)$$

onde r_j é a variável aleatória que modela o retorno da ação A_j .

Vale ressaltar que as equações (3.3.2) e (3.3.3) têm como base a aproximação

$$r_{\Pi} \approx \sum_{j=1}^m x_j r_j \quad (3.3.4)$$

para o retorno r_{Π} do portfólio Π . Esta aproximação, segundo [52], é bastante utilizada no estudo dos retornos de portfólios.

Quanto à distribuição do vetor aleatório $r = (r_1, \dots, r_m)$, esta pode ser dada por qualquer modelo multivariado para retornos. Como exemplos, é possível citar os modelos VAR e GO-GARCH, apresentados no Capítulo 2.

Finalmente, como ocorre para a estimação do CVaR, muitas vezes é preciso recorrer à função \tilde{F}_{β} , apresentada na Seção 3.2, para mensurar o risco do portfólio. Nestes casos, como feito para o estimador $\hat{\phi}_{\beta}$, é possível transformar o Problema (3.3.5) em um problema de programação linear. Para isto, basta definir um conjunto apropriado de variáveis de folga e de restrições, como feito para o Problema (3.2.17). O resultado desta transformação é o Problema (3.3.5).

$$\left\{ \begin{array}{l} \min_{(x, \alpha)} \alpha + \frac{1}{q(1-\beta)} \sum_{k=1}^q u_k \\ \text{Sujeito a:} \\ \sum_{j=1}^m x_j = 1, \text{ onde } x_j \geq 0 \\ \mu(x) \geq R \\ u_k \geq 0 \\ u_k + \alpha - f(x, y_k) \geq 0 \end{array} \right. \quad (3.3.5)$$

Parte II

Aplicações Práticas

Capítulo 4

Implementação e Avaliação dos Estimadores

Neste capítulo são apresentados os exemplos calculados para validar a implementação dos estimadores e também para compreender as limitações das técnicas utilizadas. Além disso, também são apresentados alguns detalhes sobre a implementação computacional dos estimadores.

A seguir, na Seção 4.1 é analisado o estimador de CVaR. Depois, na Seção 4.2, os estimadores de modelos GO-GARCH.

4.1 Estimação de VaR e CVaR

O estimador para CVaR foi implementado na linguagem R, com o auxílio da biblioteca de otimização *MOSEK* e do pacote *R-to-MOSEK*, que faz a ligação entre os dois. Veja [47, 22, 43].

Como as estimativas de VaR e CVaR são encontradas através da resolução do Problema (3.2.17), a implementação do estimador é simples e consiste, basicamente, em construir as matrizes que definem a instância do problema de programação linear. A solução do problema, propriamente dita, fica a cargo da biblioteca de otimização. No caso, a cargo da biblioteca *MOSEK*.

Para validar a implementação e avaliar o método de estimação foram escolhidas duas distribuições bem conhecidas: a normal padrão e a t de Student. Como essas distribuições possuem fórmulas fechadas para o cálculo de VaR e CVaR, é possível comparar os valores corretos dos mesmos com os valores estimados a partir de uma amostra.

Os valores corretos de VaR e de CVaR foram calculados através das fórmulas apresentadas nas seções 3.1.1 e 3.1.2. Estes valores podem ser encontrados na Tabela 4.1.

Considere o conjunto $A = \{A_1, \dots, A_N\}$ formado por N amostras de mesmo tamanho e o conjunto $X = \{\hat{x}_1, \dots, \hat{x}_N\}$ formado pelas estimativas calculadas sobre cada amostra A_i . Sendo x o valor correto a ser estimado, o erro absoluto de estimação pode ser medido através da raiz do erro quadrático médio (ou RMSE), que é dado por

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{x}_i - x)^2}. \quad (4.1.1)$$

Muitas vezes o erro relativo também é de interesse. Este, no caso do erro RMSE é dado por

$$RRMSE = 100 \left(\frac{RMSE}{x} \right). \quad (4.1.2)$$

Considere agora os gráficos das figuras 4.1 e 4.2. Estes mostram, respectivamente, os valores dos erros RMSE e RRMSE calculados para as duas distribuições escolhidas e para diferentes valores β

do nível de VaR. Nos gráficos, o tamanho das amostras foi variado, entretanto, o número total de 500 amostras foi mantido constante em todos os testes realizados.

	dist	beta	VaR	CVaR
1	Std. Normal	90%	1.2816	1.7550
2	Std. Normal	95%	1.6449	2.0627
3	Std. Normal	99%	2.3263	2.6652
4	t(5)	90%	1.4759	2.3022
5	t(5)	95%	2.0150	2.8901
6	t(5)	99%	3.3649	4.4524

Tabela 4.1: VaR e CVaR para as distribuições normal padrão e t_5

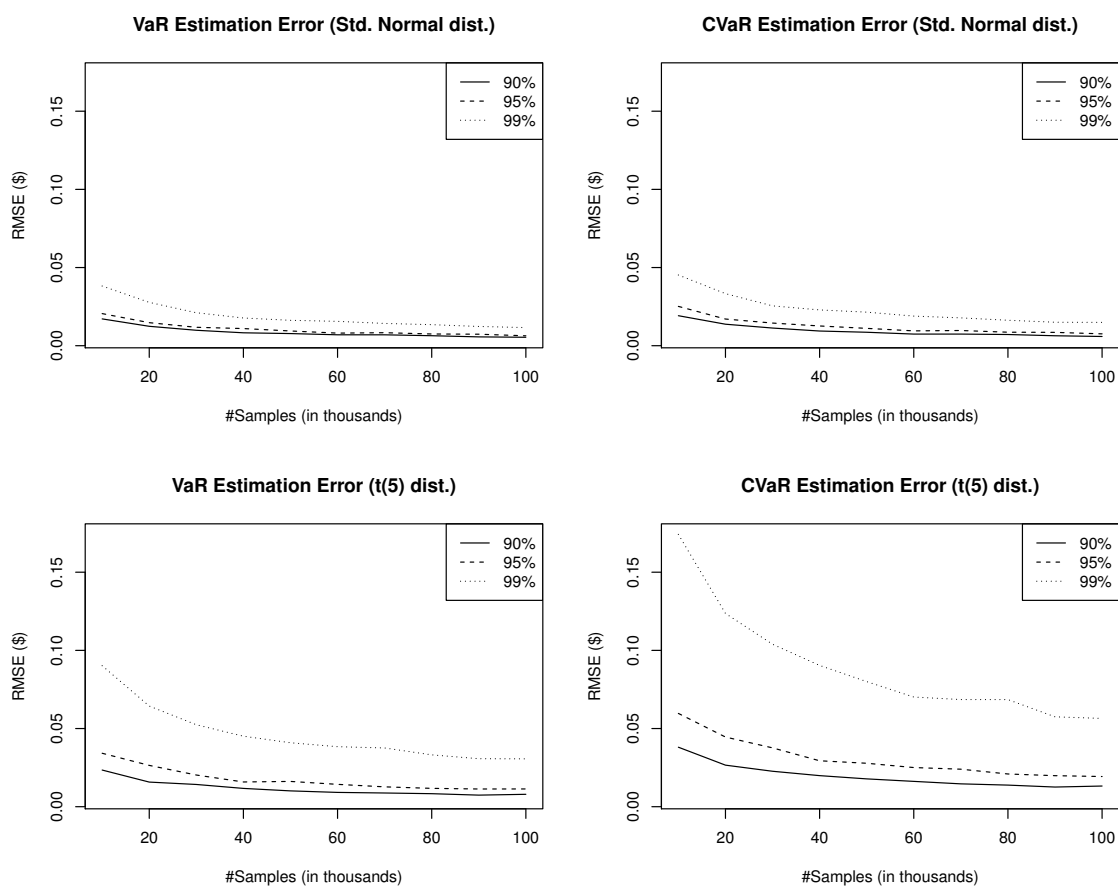


Figura 4.1: Teste de estimação para VaR e CVaR - Erro absoluto (RMSE)

Observando os resultados obtidos nos gráficos calculados, é possível observar que:

- (i) Os erros diminuem à medida que o tamanho da amostra aumenta,
- (ii) Os erros aumentam à medida que o nível β de VaR aumenta,
- (iii) Estimativas de CVaR têm erro maior que as de VaR,
- (iv) Estimativas da distribuição t de Student apresentam erros maiores que as da distribuição normal.

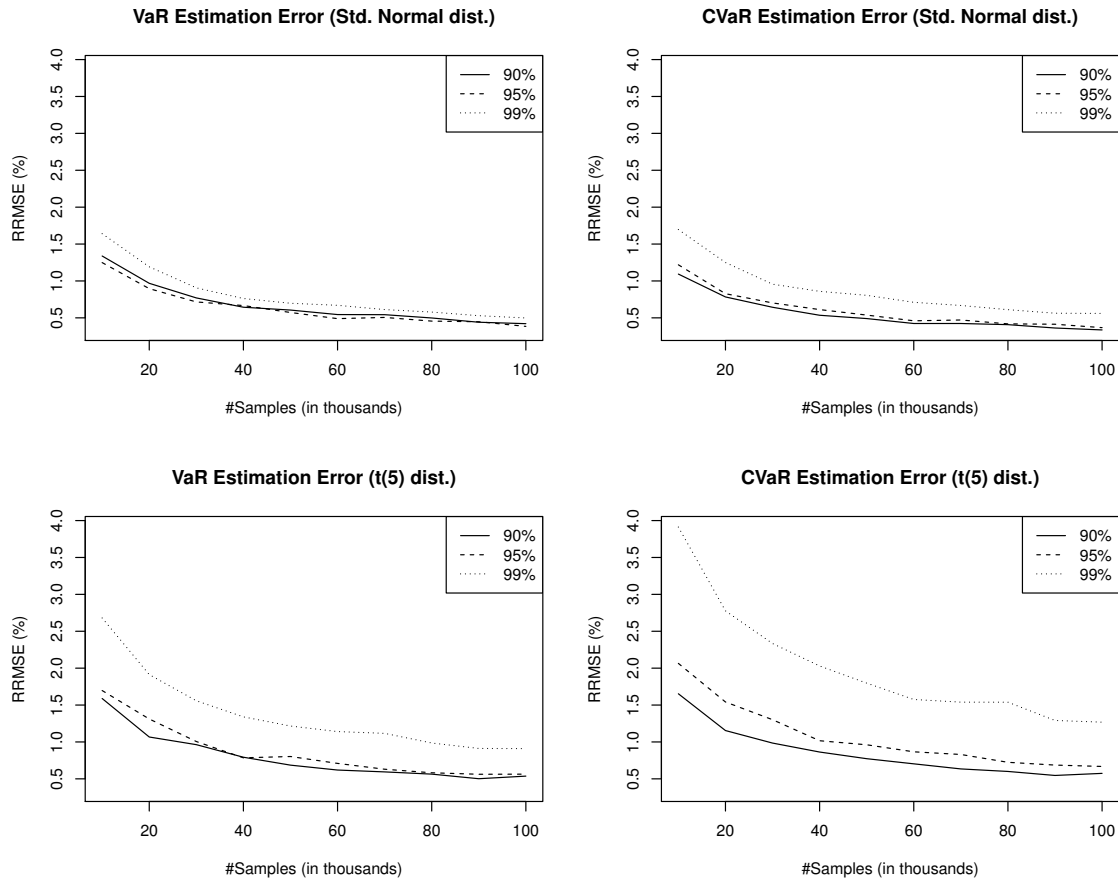


Figura 4.2: Teste de estimação para VaR e CVaR - Erro relativo (RRMSE)

De forma geral, quanto maior o tamanho de uma amostra, melhores as estimativas calculadas sobre a mesma. Com isso o resultado do item (i) corresponde ao esperado.

Naturalmente, à medida que o nível β de VaR aumenta, é esperada uma menor quantidade de valores para calcular as estimativas de CVaR, já que estas são calculadas com valores na cauda de uma distribuição de probabilidade. Com isso, é natural que os erros aumentem com valores de β maiores, como indica a observação (ii).

Quanto à observação (iii), já que CVaR é o valor esperado de uma perda, suas estimativas dependem do formato da cauda da distribuição de perdas. Como as amostras de uma distribuição apresentam poucos valores na cauda da mesma, é natural esta determinação apresentar maiores erros, já que a esperança é aproximada por uma média aritmética calculada sobre um número reduzido de valores.

Finalmente, considere a Figura 4.3, onde estão representadas as densidades de probabilidade das duas distribuições. Como a figura indica, a cauda da distribuição t é mais longa que a da distribuição normal, ou seja, o intervalo de valores para os quais a densidade de probabilidade é significativa é maior para a distribuição t . Assim, é natural que sejam necessárias amostras maiores para representar esta cauda de forma apropriada. Ou, colocando de outra forma, é natural que para amostras de mesmo tamanho, a distribuição t apresente erros maiores, como indicado na observação (iv).

Assim, com base nos resultados obtidos e nas observações acima, é possível concluir que a implementação computacional do estimador para VaR e CVaR é correta.

Quanto às limitações do método de estimação, foi possível verificar que os erros de estimação são bastante sensíveis ao tamanho das amostras utilizadas. Além disso, também é possível concluir que, de

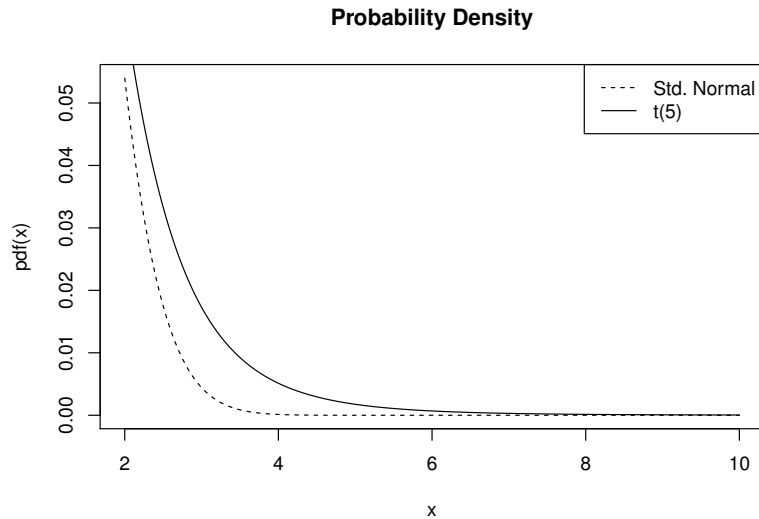


Figura 4.3: Densidade de probabilidade na cauda das distribuições

forma geral, os erros na estimação do CVaR devem aumentar com a curtose da distribuição.

4.2 Estimação de Modelos GO-GARCH

Como o estimador de CVaR, os dois estimadores para modelos GO-GARCH também foram implementados em R, seguindo de perto os algoritmos 1 e 2, apresentados no Capítulo 2. Na implementação do estimador ICA, foi utilizado o pacote *fastICA*, que implementa o algoritmo FastICA para a análise de componentes independentes. Veja [47, 40] e o Apêndice A para maiores detalhes.

A exemplo dos trabalhos [44, 55, 16], os testes de validação são baseados em um processo de simulação e estimação, ou seja, dado um modelo GO-GARCH com um conjunto Θ de parâmetros conhecidos, um conjunto de amostras $A = \{A_1, \dots, A_N\}$ é gerado a partir de simulações. Em seguida, o conjunto $P = \{\hat{\Theta}_1, \dots, \hat{\Theta}_N\}$ de estimativas para Θ é obtido ao aplicar um dos estimadores a cada amostra A_i . Finalmente, a capacidade do estimador de recuperar os parâmetros originais é avaliada através de estatísticas das distâncias entre Θ e as estimativas $\hat{\Theta}_i$, o que depende de uma função de distância apropriada.

Nas seções seguintes, o procedimento de teste descrito acima é detalhado. Em primeiro lugar, a estratégia de simulação é definida. Em seguida, a métrica do erro de estimação e, por último, são descritos os testes realizados e os resultados obtidos.

4.2.1 Estratégia de Simulação

Todos os conjuntos de amostras utilizados nos testes foram gerados de acordo com a estratégia de simulação apresentada na Seção 2.5.1, ou seja, cada amostra é obtida através da simulação de d processos independentes que formam o vetor y_t de fatores não correlacionados e que são combinados através da multiplicação do mesmo por uma matriz inversível Z . Vale lembrar que o modelo GO-GARCH tem a forma $x_t = Zy_t$.

Para permitir uma comparação direta com resultados já publicados, nas simulações foram utilizados os mesmos parâmetros definidos em [55]. Estes podem ser encontrados na Figura 4.4, onde as matrizes

Z_d são usadas para combinar os d processos GARCH(1,1) definidos pelos vetores α_d e β_d , como na Equação (2.5.9).

4.2.2 Métrica do Erro de Estimação

Geralmente, os processos associados às componentes do vetor y_t do modelo GO-GARCH são estimados através de procedimentos bem conhecidos, como os utilizados no ajuste de modelos GARCH univariados. Desta forma, a análise dos estimadores para modelos GO-GARCH se concentra nos erros de estimação da matriz Z .

Seja $Z = SU$ a decomposição espectral de Z . Como detalhado na Seção 2.5.2, a estimação da matriz Z é feita através das matrizes S e U . Assim, já que a estimativa \hat{S} é igual para os dois estimadores, os erros de estimação são dados pela distância $d(U, \hat{U}_i)$ entre a matriz U e suas estimativas \hat{U}_i . Como em [16], a distância $d(U, \hat{U}_i)$ entre estas duas matrizes ortogonais é definida por

$$d(\hat{U}_i, U) = \sqrt{\frac{1}{2} [D(U, \hat{U}_i) + D(\hat{U}_i, U)]}, \quad (4.2.1)$$

onde $D(\hat{U}, U)$ é dada por

$$D(\hat{U}, U) = 1 - \frac{1}{m} \sum_{i=1}^m \max_{1 \leq j \leq m} |u'_i \hat{u}_j|. \quad (4.2.2)$$

Assim, para um conjunto de amostras A , o erro de estimação é medido pela raiz da distância quadrática média (RMSD), ou seja,

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N d(U, \hat{U}_i)^2}. \quad (4.2.3)$$

4.2.3 Testes e Resultados

Diversos gráficos foram criados para analisar os erros de estimação da matriz U , sendo estes medidos em função do tamanho das amostras e da dimensão d do modelo. Como feito na Seção 4.1, foi utilizado um total de 500 amostras por experimento, sendo o tamanho das mesmas alterado a cada experimento.

No caso do estimador MM2011, um parâmetro adicional é necessário. Este corresponde ao número total de matrizes de autocorrelação que são consideradas pelo algoritmo de estimação. Em todos os testes realizados, este número foi fixado em 100.

Os resultados obtidos podem ser vistos na Figura 4.5. Observando a mesma é possível fazer as seguintes observações:

- (i) Os erros de estimação diminuem à medida que o tamanho das amostras aumenta,
- (ii) Modelos de maior dimensão apresentam maior erro de estimação,
- (iii) O estimador ICA apresentou melhor desempenho que o estimador MM2011.

Todas as observações feitas acima confirmam os resultados obtidos na literatura. As duas primeiras são bem naturais, considerando que os erros de estimação são menores em amostras maiores. No caso dos processos de maior dimensão, é natural que amostras maiores sejam necessárias para que o erro seja menor, já que modelos de maior dimensão contêm um número maior de parâmetros a estimar.

Considerando a estratégia de simulação adotada, é fácil ver que as coordenadas do vetor y_t são realmente independentes. Desta maneira, não é surpreendente que o estimador ICA apresente um

melhor desempenho que o estimador MM2011. Afinal, o estimador ICA calcula \hat{Z} de forma que as componentes de y_t sejam o mais independentes possível. Desta forma, conceitualmente, o estimador ICA parece ser o mais apropriado.

Finalmente, os resultados publicados em [55] são muito semelhantes aos obtidos neste trabalho. Com isso, é possível considerar que a implementação dos modelos GO-GARCH foi bem sucedida.

$$\begin{aligned}
 \alpha_3 &= \begin{bmatrix} 0.03 \\ 0.08 \\ 0.12 \end{bmatrix} & \beta_3 &= \begin{bmatrix} 0.94 \\ 0.89 \\ 0.85 \end{bmatrix} & Z_3 &= \begin{bmatrix} -0.6300 & 3.2030 & -4.4890 \\ 0.3470 & 4.6790 & -1.4470 \\ 2.3780 & 1.4940 & 1.7820 \end{bmatrix} \\
 \alpha_4 &= \begin{bmatrix} 0.03 \\ 0.08 \\ 0.12 \\ 0.15 \end{bmatrix} & \beta_4 &= \begin{bmatrix} 0.94 \\ 0.89 \\ 0.85 \\ 0.82 \end{bmatrix} & Z_4 &= \begin{bmatrix} -0.9860 & -3.5920 & 4.7760 & -6.7020 \\ -1.6110 & 0.9870 & -0.6420 & -4.6880 \\ 0.7470 & 0.0700 & -3.1620 & -4.3350 \\ -1.1560 & 4.5710 & -0.4010 & -2.5800 \end{bmatrix} \\
 \alpha_6 &= \begin{bmatrix} 0.03 \\ 0.08 \\ 0.11 \\ 0.14 \\ 0.16 \\ 0.18 \end{bmatrix} & \beta_6 &= \begin{bmatrix} 0.95 \\ 0.90 \\ 0.87 \\ 0.84 \\ 0.82 \\ 0.80 \end{bmatrix} & Z_6 &= \begin{bmatrix} 1.8680 & -2.4810 & 2.1760 & -5.2500 & -4.1570 & 1.4740 \\ 0.6240 & 5.0450 & 4.4930 & 3.8410 & -4.0520 & -1.1120 \\ 4.7850 & 2.0860 & 0.2510 & -2.8270 & 8.0870 & -3.6840 \\ -0.1670 & 0.9070 & 0.1060 & -2.2990 & 1.7230 & 3.1190 \\ -4.0590 & -1.6580 & -2.0360 & -11.4510 & -4.3180 & -1.5050 \\ 6.3010 & -4.8140 & 7.1500 & 3.8200 & -2.8200 & 2.4580 \end{bmatrix} \\
 \alpha_8 &= \begin{bmatrix} 0.02 \\ 0.06 \\ 0.09 \\ 0.11 \\ 0.13 \\ 0.15 \\ 0.17 \\ 0.19 \end{bmatrix} & \beta_8 &= \begin{bmatrix} 0.96 \\ 0.92 \\ 0.89 \\ 0.87 \\ 0.85 \\ 0.83 \\ 0.81 \\ 0.79 \end{bmatrix} & Z_8 &= \begin{bmatrix} -8.4960 & 4.5640 & 1.5260 & -2.8840 & -0.3660 & 1.4390 & -5.7900 & 3.0480 \\ -1.3430 & -5.4180 & 6.4020 & -4.2590 & 0.9180 & -10.7550 & -1.1990 & -1.9340 \\ -2.4820 & -5.7960 & -3.5670 & -6.3800 & 1.8280 & -8.1560 & -3.3280 & 7.5840 \\ -4.6440 & -5.6890 & 0.2640 & -0.6390 & 5.7650 & 1.8350 & 8.4050 & -0.4790 \\ 3.3600 & -3.2280 & -3.0830 & -2.1630 & 13.2690 & 3.4750 & -3.9330 & 5.2420 \\ -1.1210 & 1.6190 & -0.8550 & 5.2890 & -2.2670 & -6.5510 & -3.2760 & -1.3080 \\ 2.6160 & 6.6740 & 3.4010 & -2.3740 & 4.1780 & 7.9270 & 2.9450 & -4.7090 \\ 4.8540 & -2.1450 & -3.2710 & 3.9950 & -4.8070 & -1.9060 & -5.5120 & 5.4640 \end{bmatrix}
 \end{aligned}$$

Figura 4.4: Parâmetros usados na simulação de processos GO-GARCH

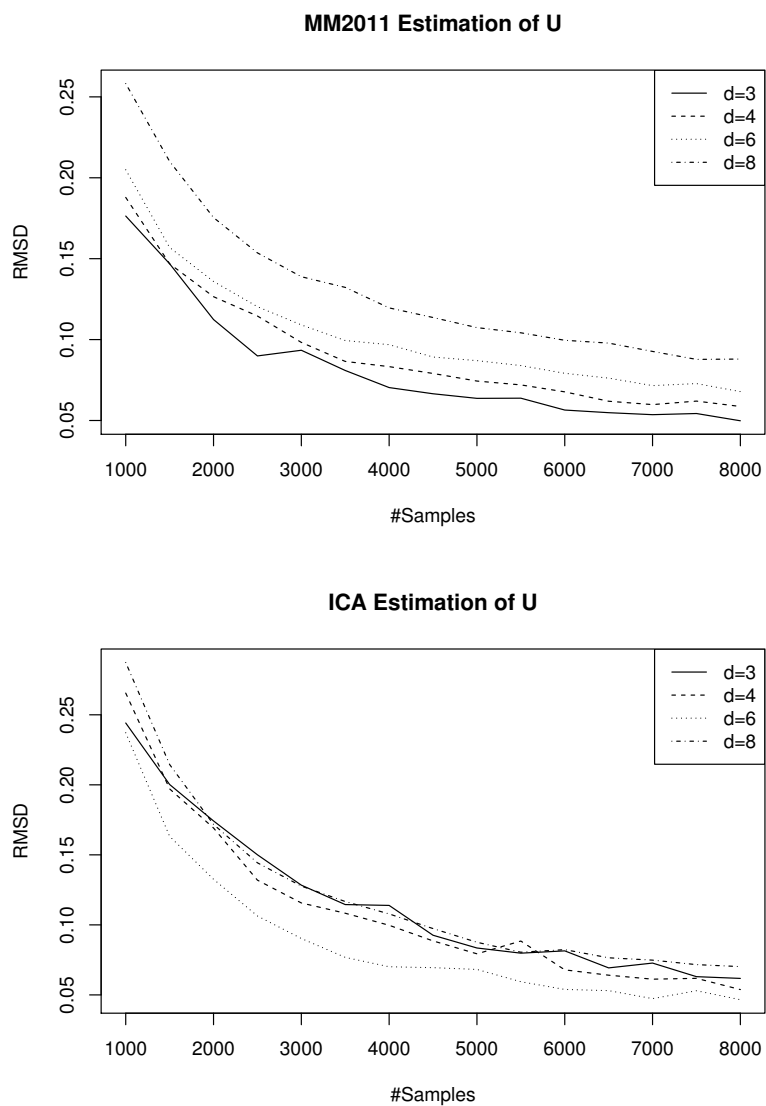


Figura 4.5: Erros de estimação da matriz U usando os estimadores MM2011 e ICA

Capítulo 5

Seleção de Portfólios

Neste capítulo são apresentados os resultados dos testes referentes à criação de portfólios com risco mínimo. Em primeiro lugar, na Seção 5.1 são vistos alguns detalhes de implementação. Em seguida, na Seção 5.2 são descritos os testes e apresentados os resultados obtidos.

5.1 Implementação

As formulações apresentadas na Seção 3.3 para a seleção de portfólios de risco mínimo são bastante gerais. Entretanto, neste trabalho é tratado apenas o caso mais simples dos portfólios de ações.

A simplicidade no tratamento dos portfólios de ações é devida à facilidade com que as suas métricas de risco, as suas perdas e os seus retornos esperados podem ser calculados, como mostrado a seguir.

Considere um portfólio $\Pi = (A, x)$ definido pelo conjunto $A = \{A_1, \dots, A_m\}$ de m ações e pelo vetor x , cujas coordenadas x_i indicam a fração do capital a ser investida nas ações A_i . Sendo r o vetor aleatório m -dimensional que modela os retornos das ações em A , o retorno do portfólio Π é dado por

$$\mu(x) = x' \mathbb{E}[r] \quad (5.1.1)$$

e as perdas do mesmo dadas por

$$f(x, Y) = -x'r, \quad (5.1.2)$$

já que uma perda corresponde ao negativo de um retorno.

Assim, as expressões obtidas para o retorno esperado $\mu(x)$ e para as perdas $f(x, Y)$ são bastante simples. Estas expressões, porém, são definidas em função do vetor aleatório r , que geralmente tem distribuição de probabilidade desconhecida.

Desta forma, evitando assumir hipóteses que poderiam restringir o uso do algoritmo implementado, as perdas e os retornos dos portfólios são estimados com base em uma amostra de retornos. Seja $r_t = \{r_{t1}, \dots, r_{tN}\}$ uma amostra de tamanho N do vetor de retornos r . Com base em r_t , a amostra Q de perdas do portfólio Π é dada por

$$Q = \{-x'r_1, \dots, -x'r_N\} \quad (5.1.3)$$

e a estimativa $\hat{\mu}(x)$ do seu retorno esperado por

$$\hat{\mu}(x) = x'\hat{\mu}_A, \quad (5.1.4)$$

onde $\hat{\mu}_A$ é a estimativa de $\mathbb{E}[r]$, calculada através de uma média amostral, ou seja,

$$\hat{\mu}_A = \frac{1}{N} \sum_{i=1}^N r_i. \quad (5.1.5)$$

Considere agora as métricas de risco, VaR e CVaR, do portfólio Π . Na Seção 3.2, foi mostrado que estas podem ser estimadas a partir da função \tilde{F}_β que, utilizando a amostra de perdas Q , tem a seguinte expressão:

$$\tilde{F}_\beta(x, \alpha) = \alpha + \frac{1}{N(1-\beta)} \sum_{k=1}^N [-x' r_k - \alpha]^+. \quad (5.1.6)$$

Finalmente, substituindo as equações (5.1.4) e (5.1.6) no Problema (3.3.5), é obtido o programa linear que foi adotado no algoritmo para a seleção de portfólios:

$$\left\{ \begin{array}{l} \min_{(x, \alpha)} \alpha + \frac{1}{N(1-\beta)} \sum_{k=1}^N u_k \\ \text{Sujeito a:} \\ \sum_{j=1}^m x_j = 1, \text{ onde } x_j \geq 0 \\ \hat{\mu}(x) \geq R \\ u_k \geq 0 \\ u_k + \alpha + x' r_k \geq 0 \end{array} \right. \quad (5.1.7)$$

Já que o Problema (5.1.7) pode ser resolvido por programação linear, a implementação computacional do algoritmo para a seleção de portfólios é bem simples. Esta consiste, basicamente, em construir as matrizes que definem a instância do problema a ser resolvido, obedecendo ao formato definido pela biblioteca de otimização escolhida. Novamente, foram utilizados a biblioteca *MOSEK* e o pacote *R-to-MOSEK*, pois o algoritmo foi implementado em R. Veja [47, 22, 43].

5.2 Resultados

Nesta seção são apresentados os resultados dos testes do algoritmo que foi implementado para a seleção de portfólios (ou alocação de capital). Em primeiro lugar, são descritos os valores adotados para os parâmetros de entrada do algoritmo. Depois, um dos portfólios selecionados é analisado em detalhes. Por último, os portfólios criados pelo algoritmo são analisados quanto às suas medidas de risco e quanto às suas composições.

5.2.1 Parâmetros de Entrada

O primeiro passo para realizar o teste de qualquer algoritmo consiste na determinação de um conjunto adequado de parâmetros de entrada. No caso aqui tratado, observando o programa linear descrito no Problema (5.1.7), é possível ver que a seleção de um portfólio depende de quatro parâmetros: o conjunto de ativos A , a amostra de retornos r_t , o nível β de VaR e o retorno esperado mínimo R . A seguir, cada um destes parâmetros é visto em maiores detalhes.

O conjunto de ativos A foi o primeiro parâmetro a ser determinado. Vinte ações com boa liquidez que são negociadas na bolsa BM&FBovespa foram selecionadas como seus elementos. A lista destas vinte ações pode ser vista na Tabela C.1, presente no Apêndice C.

Uma vez determinado o conjunto A das ações, a amostra r_t de retornos diários pôde ser construída com base nas séries históricas de preço de fechamento que são disponibilizadas pela BM&FBovespa. Vale ressaltar que todas as séries utilizadas foram devidamente ajustadas, como descrito no Apêndice C.

Determinados A e r_t , restam os principais parâmetros para a análise do algoritmo: β e R . A escolha de valores para o nível β de VaR não apresenta dificuldades. Nos testes foram adotados os valores típicos (90%, 95% e 99%). Valores para R , entretanto, devem ser escolhidos com cuidado para garantir que o Problema (5.1.7) sempre tenha solução. Dado que o retorno esperado de um portfólio de ações é uma combinação convexa das componentes do vetor $\hat{\mu}_A$, os valores válidos para R estão contidos no intervalo fechado determinado pelos valores $\min(\hat{\mu}_A)$ e $\max(\hat{\mu}_A)$.

5.2.2 Análise de um Portfólio

O primeiro teste do algoritmo consistiu em analisar um dos portfólios calculados em relação a três características: seu retorno esperado, sua composição e suas medidas de risco. Assim, considere o portfólio $\Pi_1 = (A, x_1)$, que foi selecionado pelo algoritmo para o nível de VaR $\beta = 95\%$ e retorno esperado mínimo $R = 0.6 \max(\hat{\mu}_A)$.

Em primeiro lugar, foi verificado o retorno esperado de Π_1 . A aplicação da Equação (5.1.4) comprovou que $\hat{\mu}(x_1)$ é realmente igual a $0.6 \max(\hat{\mu}_A)$. Desta forma, a restrição do retorno esperado mínimo foi satisfeita, já que $\hat{\mu}(x_1) = R$.

Em seguida, Π_1 foi analisado em relação à sua composição. Observando a representação do vetor x_1 na Figura 5.1 e as métricas de risco das ações em A na Figura 5.2, fica evidente que as ações selecionadas pelo algoritmo são as de menor CVaR, com exceção de CPLE6.

A ação CPLE6 parece uma boa candidata para reduzir o risco de Π_1 , já que tem CVaR menor que outras ações presentes no portfólio Π_1 (CCRO3, DASA3, EMBR3, KLBN4 e NATU3). Entretanto, ao substituir estas ações por CPLE6, foi possível verificar que os valores de CVaR dos novos portfólios foram sempre maiores que o CVaR de Π_1 , como pode ser verificado na Tabela 5.1.

Portfólio	VaR	CVaR
Π_1	0.022517	0.033406
Π_{CCRO3}	0.022122	0.033658
Π_{NATU3}	0.023599	0.034832
Π_{KLBN4}	0.022736	0.033460
Π_{EMBR3}	0.022907	0.034356
Π_{DASA3}	0.022865	0.033727

Tabela 5.1: Comparação do risco de portfólios

Então, por último, os valores de VaR e CVaR do portfólio Π_1 foram analisados. A comparação destes com os valores obtidos para cada ação deixa claro o benefício da diversificação de um investimento. Os valores de VaR e CVaR do portfólio Π_1 , assinalados pelas duas linhas horizontais na Figura 5.2, são significativamente menores que os valores encontrados para as ações.

Desta forma, com base nas observações acima e no fato da biblioteca de otimização indicar ter encontrado a solução ótima para a instância do Problema (5.1.7), é possível concluir que o portfólio Π_1 realmente apresenta CVaR mínimo.

5.2.3 Análise de Gráficos

No teste apresentado na Seção 5.2.2 foi feita a análise detalhada de um único portfólio. Neste segundo teste, buscou-se uma visão mais ampla do comportamento do algoritmo. Para isso, foi feita a análise

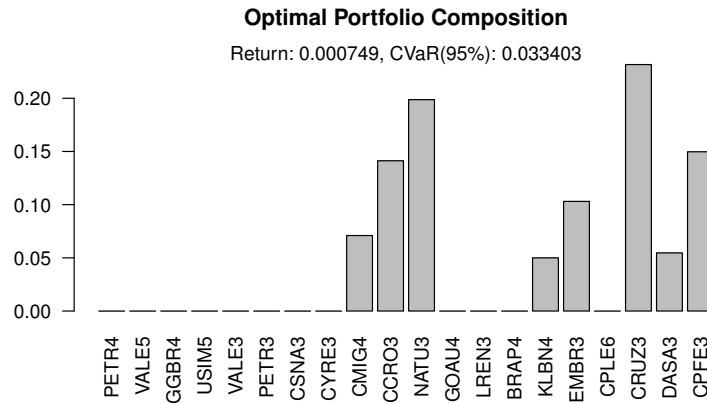


Figura 5.1: Composição x_1 do portfólio Π_1

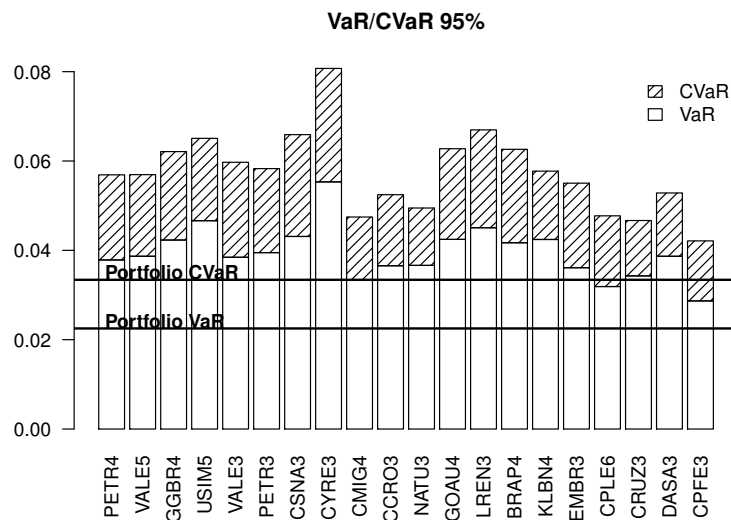


Figura 5.2: Comparação de VaR/CVaR: portfólio Π_1 versus ações individuais

das métricas de risco e da composição dos portfólios selecionados em função dos parâmetros β e R .

Desta forma, em primeiro lugar, os portfólios definidos pelo algoritmo foram analisados quanto às suas métricas de risco. Através dos gráficos presentes na Figura 5.3, que mostram os valores de VaR e CVaR em função do retorno esperado mínimo R , foi possível fazer as seguintes observações:

- (i) Os valores de VaR e CVaR aumentam à medida que β aumenta,
- (ii) O CVaR é uma curva convexa,
- (iii) O CVaR é não decrescente em relação a R ,
- (iv) Uma redução em R não implica, necessariamente, em uma redução do CVaR.

Em sequência, os portfólios usados na construção dos gráficos da Figura 5.3 foram analisados quanto à sua composição. Para isto, a composição de cada portfólio foi representada por um gráfico

vertical de barras e as barras obtidas dispostas lado a lado, como pode ser visto na Figura 5.4. Desta maneira, é possível analisar a composição dos portfólios em função do parâmetro R .

Assim, com base nos gráficos da Figura 5.4, foi verificado que:

- (v) Muitas ações nunca foram escolhidos pelo algoritmo,
- (vi) Para diversos valores de R , não houve alteração na composição dos portfólios,
- (vii) À medida que R se aproxima do seu valor máximo, aumenta a fração do portfólio atribuída à ação CRUZ3, que tem o maior retorno esperado entre todas as ações no conjunto A . Veja a Figura C.2, localizada no Apêndice C.

Com base nas observações feitas acima, é possível concluir que o algoritmo para a seleção de portfólios apresenta o comportamento esperado em relação aos parâmetros β e R , como justificado a seguir.

As duas primeiras observações estão de acordo com a definição das duas medidas de risco e com suas propriedades, como pode ser visto na Seção 3.1.2.

As observações (iii), (iv) e (vi) estão claramente relacionadas. A princípio, espera-se que o risco associado a um portfólio aumente à medida que seu retorno esperado aumente. A terceira observação parece contrariar este princípio da Teoria Moderna de Portfólios. Veja [41]. Entretanto, a observação da Figura 5.4 mostra que não há alteração na composição dos portfólios para um intervalo significativo de valores de R . Logo, neste mesmo intervalo, não deve haver alteração nos valores de VaR e de CVaR. Na região em que há alteração na composição dos portfólios, pode-se verificar que um aumento do retorno R implica em um aumento do risco associado e vice-versa, como esperado. Assim, as observações (iii), (iv) e (vi) estão de acordo com o esperado.

Finalmente, a concentração do capital no ativo de maior retorno quando R aumenta é bastante natural, dado que o retorno esperado do portfólio é um combinação convexa dos retornos das ações. Desta forma, a observação (vii) também corresponde ao esperado.

5.3 Conclusão

Os resultados obtidos na análise detalhada de um portfólio e na análise do comportamento do algoritmo em função do nível de VaR e do retorno esperado mínimo indicam que o algoritmo implementado para a seleção de portfólios de ações funciona como o esperado.

Além disso, os resultados obtidos na Seção 5.2.2 também reforçam a importância da utilização de modelos multivariados para o retorno de um conjunto de ações. As tentativas manuais sem sucesso de melhorar o portfólio Π_1 substituindo as ações da carteira por uma ação de menor CVaR refletem, de certa forma, as estratégias que devem ser adotadas por algoritmos baseados nas avaliações de risco individuais de cada ação. Assim, é possível que em muitos casos os modelos multivariados levem a soluções superiores.

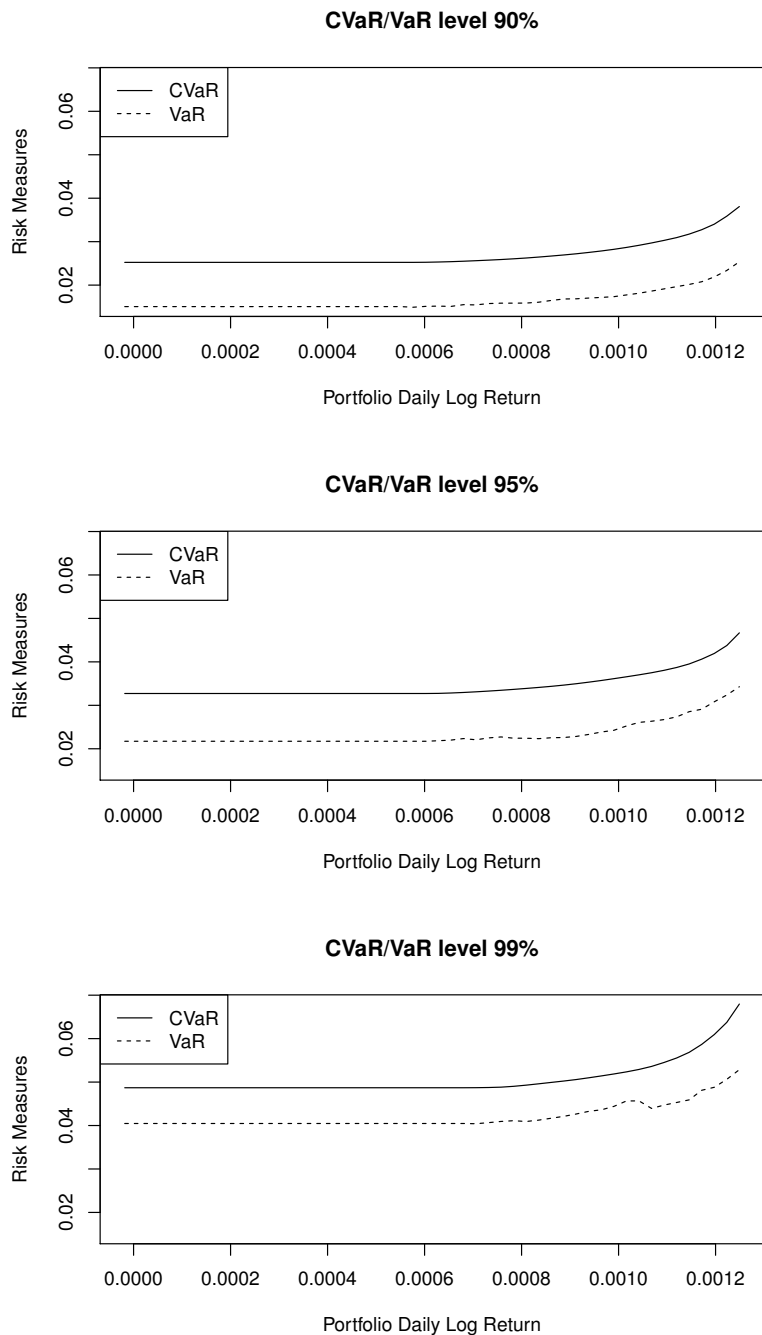


Figura 5.3: VaR/CVaR do portfólio em função do retorno médio

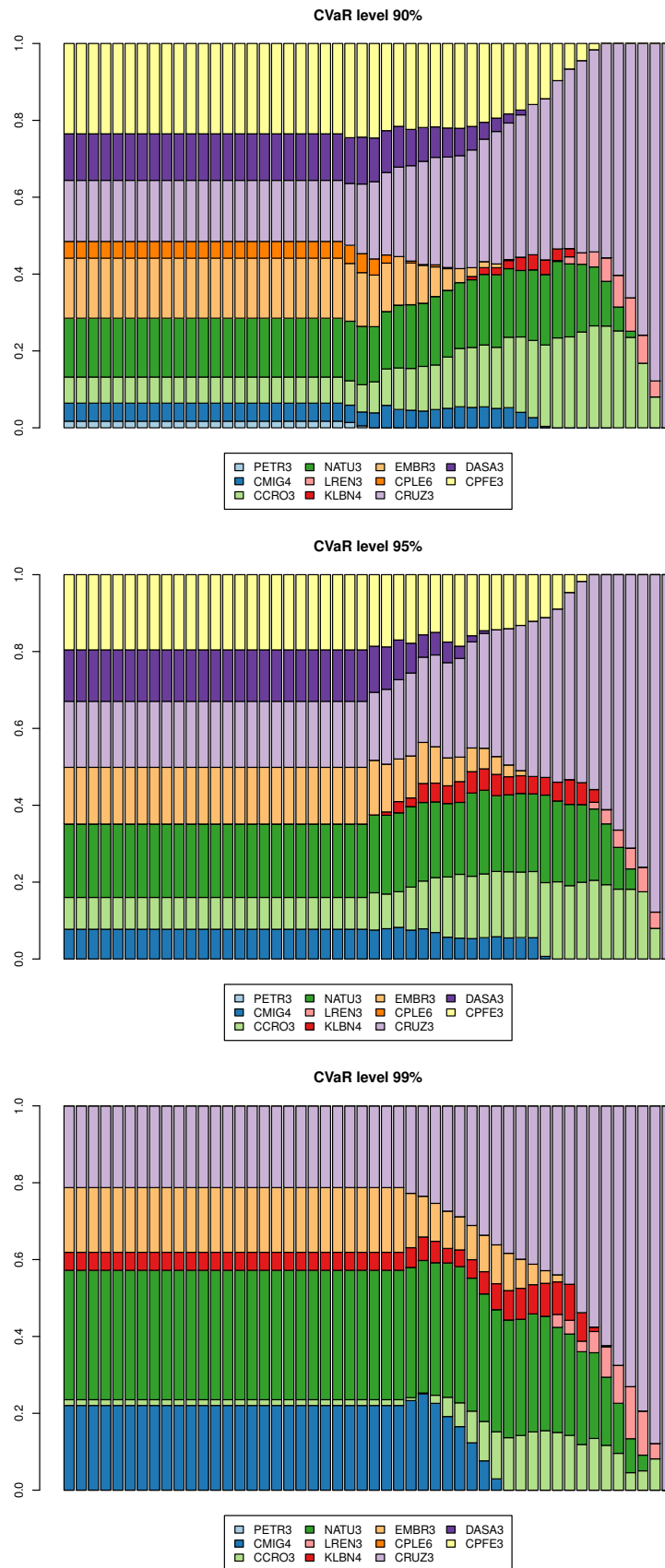


Figura 5.4: Composição do portfólio em função do retorno médio

Capítulo 6

Ajuste do Modelo VAR+GO-GARCH

Neste capítulo é apresentado o algoritmo que foi implementado para o ajuste a dados históricos de um modelo VAR+GO-GARCH. Na Seção 6.1 é feita uma descrição do algoritmo e apresentado o seu pseudocódigo. Em seguida, na Seção 6.2, são apresentados os resultados do ajuste realizado às séries de retornos históricos de cinco ações da bolsa BM&FBovespa. Por último, na Seção 6.3, são apresentadas as conclusões deste capítulo.

6.1 Implementação

Uma série temporal $\{r_t\}_{t \geq 1}$ que segue o modelo VAR+GO-GARCH é composta pela soma de duas outras séries temporais, correspondentes à sua média e à sua volatilidade. A primeira destas séries, $\{\mu_t\}_{t \geq 1}$, segue o modelo VAR e a segunda série, $\{x_t\}_{t \geq 1}$, o modelo GO-GARCH, como mostrado no Capítulo 2.

Desta forma, a primeira etapa do ajuste de um modelo VAR+GO-GARCH à uma amostra de retornos $R_t = \{r_1, \dots, r_T\}$ deve ser o ajuste do modelo VAR, pois o modelo GO-GARCH e seus estimadores assumem a hipótese de que os retornos têm média nula.

Este primeiro ajuste fornece o conjunto $M_t = \{\hat{\mu}_1, \dots, \hat{\mu}_T\}$ dos valores de $\{\mu_t\}_{t \geq 1}$ ajustados a R_t . Com estes, os resíduos $X_t = \{\hat{x}_1, \dots, \hat{x}_T\}$ podem ser calculados pela expressão

$$\hat{x}_t = r_t - \hat{\mu}_t \quad (6.1.1)$$

e sobre X_t pode ser feito, então, o ajuste do modelo GO-GARCH.

Recorrendo novamente ao Capítulo 2, Seção 2.5.2, é possível ver que o ajuste de um modelo GO-GARCH ocorre em três etapas:

- (i) Estimação da matriz S,
- (ii) Estimação da matriz U,
- (iii) Estimação dos m processos GARCH(1,1) que regem $\{x_t\}_{t \geq 1}$.

As etapas (i) e (ii) podem ser completadas utilizando o estimador ICA ou o estimador MM2011, que foram apresentados nas seções 2.5.3 e 2.5.4. Em seguida, a série y_t de fatores pode ser calculada através da equação

$$\hat{y}_t = Z^{-1} \hat{x}_t, \quad (6.1.2)$$

onde $Z = SU$. Finalmente, o ajuste do modelo GO-GARCH é completado após os m processos GARCH(1,1) serem ajustados às coordenadas de y_t .

Todo o processo de ajuste de um modelo VAR+GO-GARCH pode ser visto no Algoritmo 3, onde está descrita a função *VAR_GOGARCH*. Esta foi implementada em R com o auxílio do estimador ICA e dos pacotes *vars* e *rugarch*, que foram usados para realizar os ajustes do modelo VAR e do modelo GARCH(1,1). Veja [46, 24].

Algorithm 3 Algoritmo para o ajuste de um modelo VAR+GO-GARCH

```

1: procedure VAR_GOGARCH( $R_t, p$ )
2:    $\{M_t, B\} := \mathbf{VAR}(R_t, p)$ 
3:    $X_t := R_t - M_t$ 
4:    $\{Y_t, S, U\} := \mathbf{GOGARCH\_ICA}(X_t)$ 
5:    $Z := S U$ 
6:   for  $k := 1, \dots, m$  do
7:      $\{\sigma_k^2, g_k\} := \mathbf{GARCH}(Y'_{t,k}, 1, 1)$ 
8:   end for
9:    $G := [g_1 \ g_2 \ \dots \ g_m]$ 
10:   $H_t := [\sigma_1^2 \ \sigma_2^2 \ \dots \ \sigma_m^2]'$ 
11:  return  $B, M_t, X_t, Z, Y_t, G, H_t$ 
12: end procedure

```

Seguindo o pseudocódigo do Algoritmo 3, é possível ver que a primeira operação da função *VAR_GOGARCH* é o ajuste do modelo VAR de ordem p através da chamada à função *VAR*. Esta recebe como parâmetros a ordem do modelo e a matriz R_t de tamanho $m \times T$ que contém a amostra da série de retornos. Em seguida, são calculados os resíduos X_t do modelo VAR (linha 3) e realizado o ajuste do modelo GO-GARCH com a chamada à função *GOGARCH_ICA*, que foi definida na Seção 2.5.3. Finalmente, é realizado o ajuste dos modelos GARCH(1,1) às coordenadas da série de fatores y_t (linhas 6 a 8). Por último, então, são criadas a matriz G de parâmetros GARCH(1,1) e a a matriz H_t com as volatilidades univariadas ajustadas.

Vale ressaltar que o ajuste do k -ésimo processo univariado fornece o vetor σ_k^2 de dimensão $1 \times T$ e o vetor coluna g_k . O primeiro destes contém os valores estimados para a volatidade da k -ésima coordenada de y_t e o segundo os parâmetros do modelo GARCH(1,1) ajustado.

6.2 Resultados

Nesta seção são apresentados os resultados do ajuste realizado e para avaliar a implementação do Algoritmo 3 e para determinar a adequação do modelo VAR+GO-GARCH em relação aos retornos de um conjunto de ações. Desta forma, na Seção 6.2.1 é feita uma descrição dos parâmetros do ajuste realizado. Em seguida, na Seção 6.2.2 é feito o diagnóstico do ajuste. Por último, na Seção 6.2.3, é feita a análise das volatilidades e covariâncias condicionais calculadas pelo modelo VAR+GO-GARCH.

6.2.1 Descrição do Ajuste

Sendo o ajuste do modelo VAR+GO-GARCH feito através da função *VAR_GOGARCH*, é preciso especificar três características para a realização do teste: a amostra R_t dos retornos, a ordem p do modelo VAR e a distribuição dos choques ϵ_t dos processos GARCH(1,1) que são usados para modelar os fatores. A seguir, cada uma destas características é descrita.

A amostra R_t escolhida para o teste é formada pelos retornos de cinco ações que foram calculados com base em 600 preços de fechamento observados entre 04/01/2006 e 12/06/2008. As cinco ações escolhidas, todas negociadas na BM&FBovespa, foram: PETR4, VALE5, GGBR4, USIM5 e CSNA3.

A ordem p do modelo VAR foi escolhida como $p = 5$, após a realização do ajuste de diversos modelos VAR, sendo cada um deles seguido da análise das autocorrelações dos resíduos.

Finalmente, dado que os retornos financeiros apresentam distribuições assimétricas e com caudas pesadas, foi selecionada a distribuição t de Student assimétrica, definida em [17] e implementada no pacote *rugarch*, como a distribuição dos choques dos processos GARCH(1,1).

6.2.2 Diagnóstico do Ajuste

Na ausência de testes específicos para o diagnóstico de modelos GO-GARCH na literatura, recorreu-se à análise de gráficos, a testes gerais que podem ser aplicados a qualquer modelo GARCH multivariado e à verificação das hipóteses e restrições do modelo para analisar a qualidade do ajuste obtido através da função *VAR_GOGARCH*.

Análise das Correlações

O diagnóstico do ajuste do modelo VAR+GO-GARCH foi iniciado com a análise das matrizes de correlação associadas às séries de resíduos obtidas no ajuste do modelo: $\{x_t\}_{t \geq 1}$ e $\{z_t\}_{t \geq 1}$. A primeira destas, $\{x_t\}_{t \geq 1}$, corresponde aos resíduos do ajuste VAR. A segunda série, $\{z_t\}_{t \geq 1}$, corresponde aos resíduos padronizados do modelo GO-GARCH.

Considere a Figura 6.1, onde a matrizes de correlação analisadas podem ser encontradas. As matrizes de correlação de r_t e r_t^2 indicam a presença de correlações significativas entre as componentes dos retornos r_t no primeiro e no segundo momento.

O mesmo pode ser observado nas matrizes referentes a x_t , que são muito semelhantes às obtidas para os retornos. Aparentemente, o modelo VAR não capturou totalmente a dinâmica da série $\{r_t\}_{t \geq 1}$.

Entretanto, ao considerar as matrizes referentes a z_t , é possível notar uma grande redução nas correlações encontradas no primeiro e no segundo momento. Assim, é possível que o modelo GO-GARCH ajustado tenha sido capaz de capturar a dinâmica dos retornos de forma apropriada. Além disso, a dinâmica da série de retornos parece ser explicada, principalmente, pelo segundo momento.

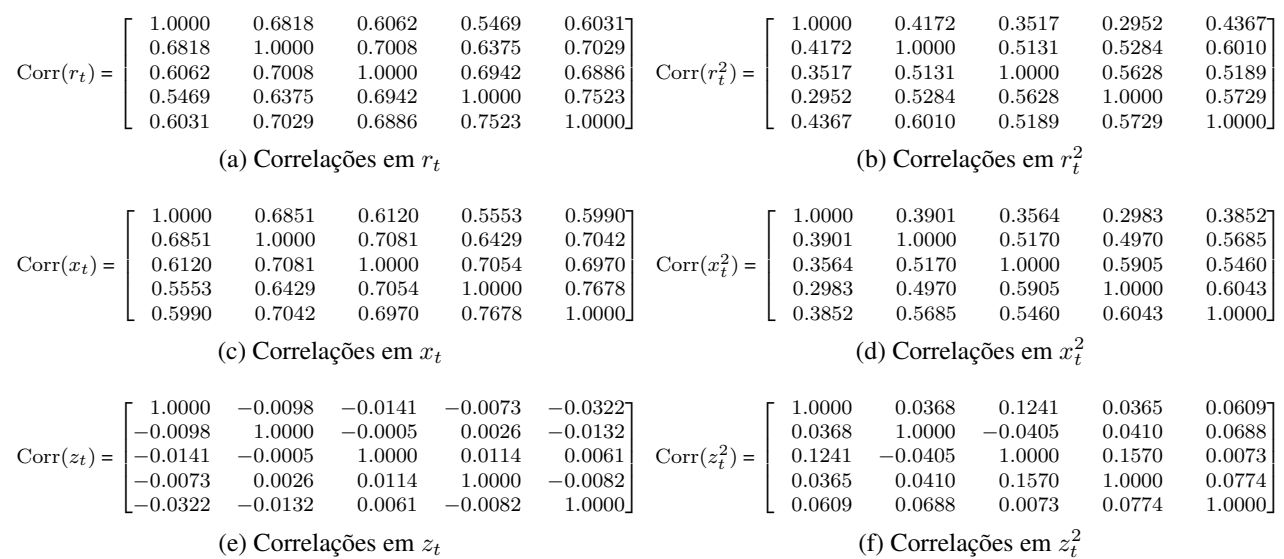


Figura 6.1: Matrizes de correlação: r_t , x_t , z_t e seus quadrados

Testes de Autocorrelação

Após a análise das matrizes de correlação, foi testada a existência de autocorrelações nas séries de retornos e de resíduos. Assim, o teste de Hosking, apresentado na Seção 2.1, foi aplicado às amostras das séries $\{r_t\}_{t \geq 1}$, $\{x_t\}_{t \geq 1}$, $\{z_t\}_{t \geq 1}$ e seus quadrados. Os resultados obtidos podem ser encontrados na Figura 6.2.

Analisando as tabelas presentes na Figura 6.2, é possível ver que os retornos r_t apresentam autocorrelações em seus dois primeiros momentos, pois os p-valores fornecidos pelo teste de Hosking indicam a rejeição da hipótese nula. Esta, no teste de Hosking, é de não existência de autocorrelações. Repare que a rejeição da hipótese nula é assinalada pelos p-valores em vermelho nas tabelas (a) e (b).

Lags	Statistic	df	p-value
6	192.3920	150	0.0111
10	309.9618	250	0.0058
15	427.7906	375	0.0309
20	578.5314	500	0.0085
25	749.3717	625	0.0004
30	943.8149	750	0

(a) Hosking em r_t

Lags	Statistic	df	p-value
6	354.7336	150	0
10	464.2409	250	0
15	645.1586	375	0
20	784.1894	500	0
25	973.4747	625	0
30	1172.6119	750	0

(b) Hosking em r_t^2

Lags	Statistic	df	p-value
6	22.0671	25	0.6319
10	140.5794	125	0.1614
15	245.1438	250	0.5748
20	403.0684	375	0.1528
25	545.5175	500	0.0780
30	727.1244	625	0.0029

(c) Hosking em x_t

Lags	Statistic	df	p-value
5	279.7014	125	0
10	398.5589	250	0
15	562.6849	375	0
20	680.7541	500	0
25	877.0728	625	0
30	1060.1254	750	0

(d) Hosking em x_t^2

Lags	Statistic	df	p-value
5	13.7861	125	1.0000
10	150.8289	250	1.0000
15	247.7314	375	1.0000
20	393.7500	500	0.9998
25	531.5097	625	0.9972
30	693.9615	750	0.9288

(e) Hosking em z_t

Lags	Statistic	df	p-value
5	118.6954	125	0.6418
10	257.9895	250	0.3507
15	388.3685	375	0.3062
20	540.3178	500	0.1033
25	767.2099	625	0.0001
30	995.0267	750	0

(f) Hosking em z_t^2

Figura 6.2: Testes de autocorrelação: r_t , x_t , z_t e seus quadrados

Em seguida, analisando os p-valores obtidos nos testes feitos sobre $\{x_t\}_{t \geq 1}$, é possível ver que as autocorrelações existentes em $\{r_t\}_{t \geq 1}$ foram removidas até o lag 25 pelo ajuste do modelo VAR. As autocorrelações do segundo momento, como esperado, ainda estão presentes na série $\{x_t^2\}_{t \geq 1}$, como indicam os p-valores da tabela (d).

Finalmente, analisando os resultados obtidos para $\{z_t^2\}_{t \geq 1}$, é possível ver que o ajuste do modelo GO-GARCH foi capaz de remover as autocorrelações existentes até o vigésimo lag. Além disso, este ajuste também parece ter tido um efeito positivo sobre a série de resíduos $\{x_t\}_{t \geq 1}$, pois as autocorrelações que nesta foram acusadas, não mais ocorrem em $\{z_t\}_{t \geq 1}$.

Análise dos Resíduos VAR

Completados os testes de autocorrelação, passou-se, então, à análise de outras propriedades dos resíduos. Assim, considere o ajuste do modelo VAR. Para que este seja considerado bem sucedido, é necessário que o valor esperado de x_t seja nulo e que não existam autocorrelações na série $\{x_t\}_{t \geq 1}$. A última dessas condições foi comprovada nos testes de autocorrelação (veja a Figura 6.2). Assim, ainda deve ser verificada a primeira condição.

Para avaliar se a esperança $\mathbb{E}[x_t]$ é nula, foi utilizado o teste t de Student para a média de uma amostra A . Este teste indica se o valor esperado dos elementos em A corresponde ao valor μ_0 , especificado previamente. Veja a Definição 12 e [54].

Definição 12. *Teste t de Student para a Média de uma Amostra A*

$$H_0: \mathbb{E}[A] = \mu_0$$

$$H_a: \mathbb{E}[A] \neq \mu_0$$

$$\text{Estatística: } t = \frac{(\hat{\mu} - \mu_0) \sqrt{n}}{\hat{s}}$$

$$\text{Regra: } H_0 \text{ é rejeitada em favor de } H_a \text{ quando } |t| > t_{n-1}(\alpha/2),$$

onde $\hat{\mu}$ e \hat{s} são, respectivamente, a média e o desvio padrão amostrais de A ; $t_{n-1}(\alpha/2)$ é o quantil de ordem $\alpha/2$ da distribuição t de Student com $n - 1$ graus de liberdade e n é o tamanho da amostra A .

Desta forma, os valores esperados das coordenadas de x_t foram testados contra $\mu_0 = 0$. Como os p -valores obtidos nos testes indicam, todas as coordenadas de x_t têm média nula. Veja a Tabela 6.1.

	Value	t-Statistic	p-value
$\mathbb{E}[x_{1t}]$	0	0	1.0000
$\mathbb{E}[x_{2t}]$	0	0	1.0000
$\mathbb{E}[x_{3t}]$	0	0	1.0000
$\mathbb{E}[x_{4t}]$	0	0	1.0000
$\mathbb{E}[x_{5t}]$	0	0	1.0000

Tabela 6.1: Teste do valor esperado dos resíduos x_t do ajuste VAR

Assim, comprovada a ausência de autocorrelações em $\{x_t\}_{t \geq 1}$ e que $\mathbb{E}[x_t] = 0$, é possível concluir que o ajuste do modelo VAR foi bem sucedido.

Análise dos Resíduos VAR+GO-GARCH

Considere agora, o ajuste do modelo GO-GARCH e a série $\{z_t\}_{t \geq 1}$ de resíduos padronizados. Na Seção 2.5.5, são descritas propriedades dos resíduos padronizados z_t que podem ser exploradas para avaliar o ajuste de um modelo GO-GARCH. Estes resíduos, quando o ajuste é bem sucedido, devem apresentar:

- (i) Ausência de autocorrelações nos dois primeiros momentos,
- (ii) $\mathbb{E}[z_t] = 0$,
- (iii) $\text{Cov}(z_t) = I_m$.

A validade da primeira propriedade foi verificada anteriormente, durante a realização dos testes de autocorrelação. Nestes, ao aplicar o teste de Hosking nas séries $\{z_t\}_{t \geq 1}$ e $\{z_t^2\}_{t \geq 1}$, foi possível comprovar que a série de resíduos GO-GARCH não apresenta autocorrelações em seus dois primeiros momentos. Veja os resultados da Figura 6.2.

Para avaliar a segunda propriedade, referente ao valor esperado de z_t , recorreu-se novamente ao teste t de Student para a média de uma amostra. Conforme os resultados obtidos, presentes na Tabela 6.2, todas as componentes de z_t têm média zero.

	Value	t-Statistic	p-value
$\mathbb{E}[z_{1t}]$	0.0027	0.0645	0.9486
$\mathbb{E}[z_{2t}]$	0.0032	0.0768	0.9388
$\mathbb{E}[z_{3t}]$	-0.0093	-0.2269	0.8206
$\mathbb{E}[z_{4t}]$	-0.0065	-0.1558	0.8763
$\mathbb{E}[z_{5t}]$	-0.0010	-0.0241	0.9808

Tabela 6.2: Teste do valor esperado dos resíduos padronizados z_t

Em seguida, para avaliar se a matriz de covariância de z_t é igual à identidade, recorreu-se novamente ao teste t de Student.

Uma observação sobre este teste é necessária. Repare que o elemento (i, j) da matriz de covariância corresponde à esperança $\mathbb{E}[z_{it}z_{jt}]$, para a qual o valor μ_0 a testar é conhecido (elementos na diagonal têm $\mu_0 = 1$ e os demais têm $\mu_0 = 0$). Desta forma, é possível aplicar o teste t a cada uma das entradas da matriz $\text{Cov}(z_t)$, o que permite uma análise mais aprofundada sobre z_t através dos p-valores obtidos. Estes, por conveniência, foram agrupados na matriz $\text{PV}(z_t)$, que pode ser encontrada na Figura 6.3.

$$\text{Cov}(z_t) = \begin{bmatrix} 1.0412 & -0.0101 & -0.0140 & -0.0072 & -0.0332 \\ -0.0101 & 1.0028 & -0.0004 & 0.0027 & -0.0134 \\ -0.0140 & -0.0004 & 1.0074 & 0.0105 & 0.0066 \\ -0.0072 & 0.0027 & 0.0105 & 1.0189 & -0.0080 \\ -0.0332 & -0.0134 & 0.0066 & -0.0080 & 1.0119 \end{bmatrix} \quad \text{PV}(z_t) = \begin{bmatrix} 0.7444 & 0.8247 & 0.7879 & 0.8755 & 0.4890 \\ 0.8247 & 0.9655 & 0.9925 & 0.9499 & 0.7660 \\ 0.7879 & 0.9925 & 0.9012 & 0.8279 & 0.8754 \\ 0.8755 & 0.9499 & 0.8279 & 0.7715 & 0.8610 \\ 0.4890 & 0.7660 & 0.8754 & 0.8610 & 0.8620 \end{bmatrix}$$

Figura 6.3: Teste da matriz de covariância dos retornos padronizados z_t

Conforme os p-valores de $\text{PV}(z_t)$, a matriz $\text{Cov}(z_t)$ é realmente igual à identidade. Desta forma, comprovadas todas as três propriedades, é possível concluir que o ajuste do modelo GO-GARCH também foi bem sucedido.

Análise da Distribuição dos Choques ϵ_t

Após a verificação das propriedades dos resíduos, foi feita a análise da distribuição do vetor aleatório ϵ_t , que corresponde aos choques dos processos GARCH(1,1) usados para modelar o vetor de fatores y_t . Esta análise tem por objetivo determinar se a distribuição do vetor ϵ_t foi especificada corretamente.

Quando a dinâmica dos retornos é capturada corretamente, os resíduos z_t obtidos têm distribuição compatível com a distribuição de ϵ_t , ou seja, os resíduos podem ser considerados realizações do vetor aleatório ϵ_t .

Sendo as coordenadas de ϵ_t independentes entre si, a comparação com os resíduos deve ser feita coordenada a coordenada, ou seja, é preciso determinar se ϵ_{kt} , a k -ésima coordenada de ϵ_t é compatível com o conjunto $Z_k = \{z_{k1}, \dots, z_{kT}\}$ das k -ésimas coordenadas de z_t .

Felizmente, esta comparação pode ser feita de forma bem conveniente. Basta contrapor Z_k a uma amostra E_k que tenha sido gerada segundo a distribuição de ϵ_{kt} . Esta contraposição pode ser

feita através do teste de Kolmogorov-Smirnov (ou teste KS) e da criação de QQ-plots, que permitem determinar se duas amostras pertencem ou não à uma mesma distribuição. Veja [52, 50].

Assim, foram geradas as amostras E_1, \dots, E_5 segundo a distribuição das coordenadas de ϵ_t , cujos parâmetros foram determinados após o ajuste dos processos GARCH(1,1). Estas foram, então, comparadas com as amostras Z_1, \dots, Z_5 através de QQ-plots, que podem ser encontrados na Figura 6.4.

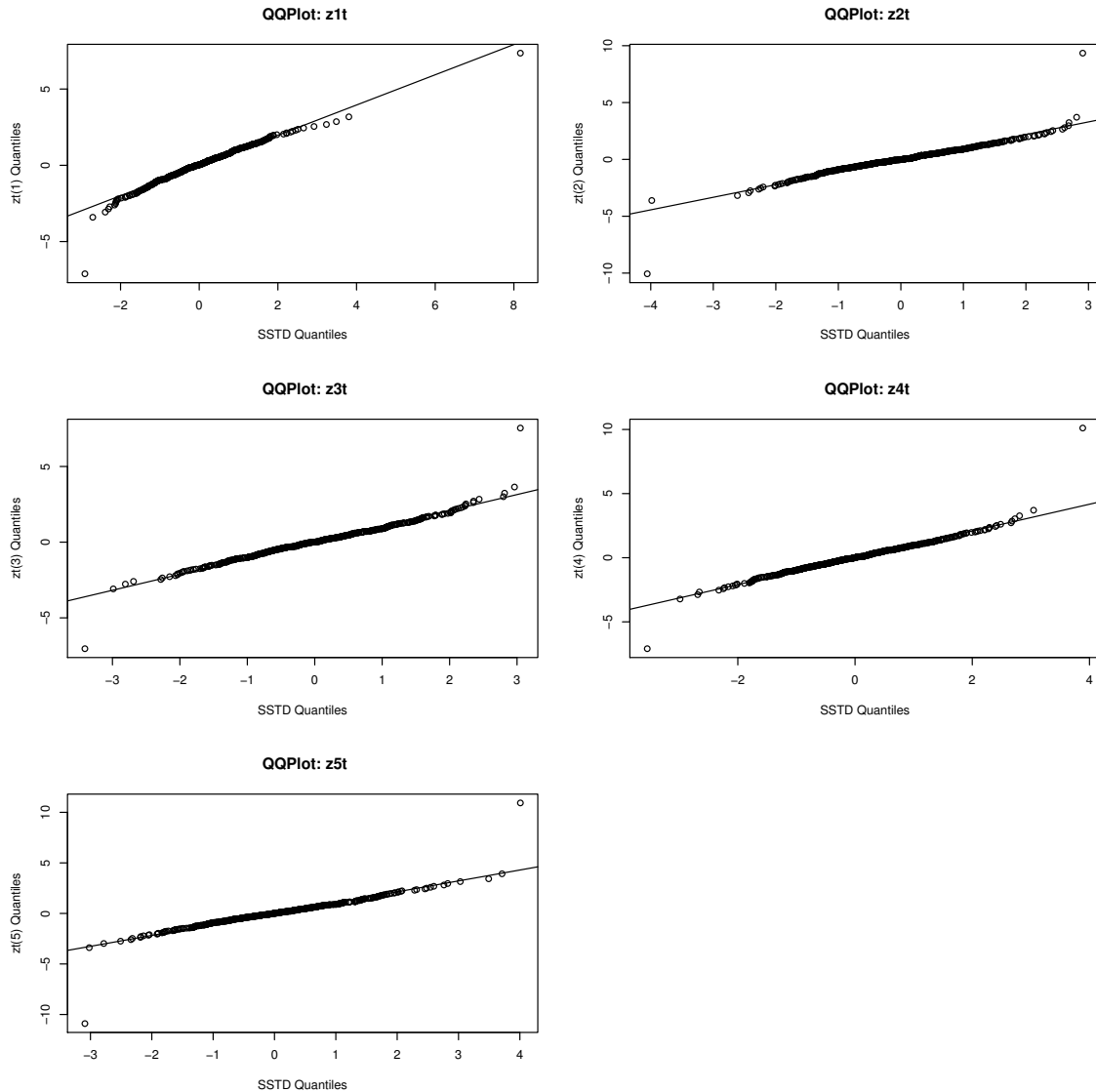


Figura 6.4: Q-Q Plot dos resíduos padronizados do modelo VAR+GO-GARCH

Dada a disposição dos elementos das amostras nos gráficos da Figura 6.4 (dispostos sobre uma mesma reta), é possível concluir que estas podem realmente ter sido geradas segundo uma mesma distribuição de probabilidade.

Em seguida, após o resultado positivo na análise dos QQ-plots, foi aplicado o teste KS aos pares de amostras E_k e Z_k . Os resultados obtidos podem ser encontrados na Tabela 6.3.

Analisando os p-valores obtidos presentes na Tabela 6.3, chega-se novamente à conclusão de que E_k e Z_k são amostras de uma mesma variável aleatória, pois os alto valores encontrados não permitem rejeitar a hipótese nula do teste KS. Esta, logicamente, é de que as amostras foram geradas segundo uma mesma distribuição.

	Statistic	p-value
z_{1t} versus ϵ_{1t}	0.0381	0.3563
z_{2t} versus ϵ_{2t}	0.0359	0.4280
z_{3t} versus ϵ_{3t}	0.0338	0.5067
z_{4t} versus ϵ_{4t}	0.0228	0.9180
z_{5t} versus ϵ_{5t}	0.0316	0.5936

Tabela 6.3: Teste de Kolmogorov-Smirnov: distribuição de z_t versus a distribuição de ϵ_t

Desta forma, dados os resultados positivos dos testes KS e dos QQ-plots, é possível concluir que a distribuição t de Student assimétrica é realmente uma escolha apropriada para ϵ_t .

Matrizes do Modelo GO-GARCH

Os testes realizados até o momento indicam que o modelo escolhido para os retornos r_t foi capaz de capturar a dinâmica dos mesmos de forma apropriada. Assim, agora são consideradas as hipóteses adotadas pelo modelo GO-GARCH.

O primeiro teste realizado neste sentido se refere às matrizes do modelo. Como mostrado na Seção 2.5, a matriz $Z = SU$ deve ser inversível, o que pôde ser comprovado facilmente, calculando sua inversa.

As matrizes obtidas no ajuste e a matriz Z correspondente podem ser vistas na Figura 6.5.

$$S = \begin{bmatrix} 0.0190 & 0.0064 & 0.0052 & 0.0048 & 0.0052 \\ 0.0064 & 0.0187 & 0.0066 & 0.0060 & 0.0069 \\ 0.0052 & 0.0066 & 0.0192 & 0.0074 & 0.0067 \\ 0.0048 & 0.0060 & 0.0074 & 0.0225 & 0.0092 \\ 0.0052 & 0.0069 & 0.0067 & 0.0092 & 0.0207 \end{bmatrix} \quad U = \begin{bmatrix} 0.0545 & 0.1949 & -0.3556 & 0.1251 & 0.9038 \\ -0.5039 & -0.6071 & -0.5991 & 0.1032 & -0.0888 \\ 0.4303 & 0.0961 & -0.2801 & 0.8093 & -0.2689 \\ 0.4696 & 0.1619 & -0.6215 & -0.5604 & -0.2301 \\ -0.5809 & 0.7470 & -0.2234 & 0.0686 & -0.2235 \end{bmatrix}$$

$$Z = \begin{bmatrix} -0.0008 & 0.0050 & -0.0162 & 0.0049 & 0.0129 \\ -0.0074 & -0.0034 & -0.0206 & 0.0052 & -0.0005 \\ 0.0048 & 0.0050 & -0.0173 & 0.0132 & -0.0043 \\ 0.0057 & 0.0085 & -0.0234 & -0.0048 & -0.0055 \\ -0.0080 & 0.0145 & -0.0182 & 0.0030 & -0.0045 \end{bmatrix}$$

Figura 6.5: Matrizes do modelo GO-GARCH ajustado

Análise dos Fatores y_t

Voltando à definição do modelo GO-GARCH, no Capítulo 2, é possível ver que a série $\{y_t\}_{t \geq 1}$ dos fatores segue, por hipótese, a estrutura dada pela equação

$$y_t = H_t^{1/2} \epsilon_t, \quad (6.2.1)$$

onde $H_t = \text{diag}(h_{1t}, \dots, h_{mt})$ é a matriz de variância condicional, h_{it} é a volatilidade condicional dada pelo i -ésimo processo GARCH(1,1) do modelo e $\{\epsilon_t\}_{t \geq 1}$ é a sequência vetorial martingal de diferenças que corresponde aos choques dos processos GARCH(1,1). Veja a Definição 4.

Além desta, outras hipóteses também são feitas sobre a série de fatores:

- (i) A série $\{y_{it}\}_{t \geq 1}$ formada pela i -ésima coordenada de $\{y_t\}_{t \geq 1}$ é estacionária,
- (ii) $\mathbb{E}[h_{it}] = 1$ e $h_{it} \geq 0$,
- (iii) $\mathbb{E}[y_t | \mathcal{F}_{t-1}] = 0$,
- (iv) $\text{Cov}(y_t | \mathcal{F}_{t-1}) = \mathbb{E}[y_t y_t' | \mathcal{F}_{t-1}] = H_t$.

Infelizmente, é impossível analisar as esperanças condicionais estatisticamente, pois existe apenas uma realização de $y_t | \mathcal{F}_{t-1}$ para cada tempo t . Entretanto, como $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X | \mathcal{F}_t]]$ para uma variável aleatória X , é possível recorrer à média amostral das realizações de $y_t | \mathcal{F}_{t-1}$ para a realização de testes.

Desta forma, obtém-se o seguinte conjunto de propriedades que podem ser testadas:

- (i) As séries $\{y_{it}\}_{t \geq 1}$ são estacionárias,
- (ii) $\mathbb{E}[h_{it}] = 1$ e $h_{it} \geq 0$,
- (iii) $\mathbb{E}[y_{it}] = 0$,
- (iv) $\mathbb{E}[\epsilon_{it}] = 0$,
- (v) $\text{Cov}(\epsilon_t) = I_m$,
- (vi) $\text{Cov}(y_t) = I_m$.

Sobre estas propriedades, é necessário fazer alguns comentários. As propriedades (iv) e (v) são provenientes da definição de uma sequência vetorial martingal de diferenças (Definição 4), sendo assim facilmente comprovadas. A propriedade (vi), entretanto, requer um pouco mais de atenção. Esta pode ser justificada por

$$\text{Cov}(y_t) = \mathbb{E}[y_t y_t'] = \mathbb{E}[\mathbb{E}[y_t y_t' | \mathcal{F}_{t-1}]] = \mathbb{E}[H_t] = I_m. \quad (6.2.2)$$

Assim, uma vez que todos os testes realizados sobre as séries de fatores foram propriamente definidos e justificados, é possível passar à análise de seus resultados.

Em primeiro lugar, foi testada a estacionariedade dos fatores $\{y_{it}\}_{t \geq 1}$. Para isto, foi empregado o teste KPSS, que tem como hipótese nula a estacionariedade da série sendo analisada. Veja [36].

De acordo com os p-valores obtidos, veja a Tabela 6.4, é possível concluir que todas as séries $\{y_{it}\}_{t \geq 1}$ são realmente estacionárias, como esperado.

	Statistic	p-value
KPSS Test em y_{1t}	0.0827	0.1000
KPSS Test em y_{2t}	0.1170	0.1000
KPSS Test em y_{3t}	0.2908	0.1000
KPSS Test em y_{4t}	0.1128	0.1000
KPSS Test em y_{5t}	0.2642	0.1000

Tabela 6.4: Teste de estacionariedade dos fatores y_{it} do modelo GO-GARCH ajustado

Em seguida, o teste t de Student foi utilizado para validar as hipóteses feitas sobre os valores esperados de h_{it} , y_{it} e ϵ_{it} . Como os p-valores presentes nas tabelas da Figura 6.6 indicam, todas

	Value	t-Statistic	p-value		Value	t-Statistic	p-value
$\mathbb{E}[y_{1t}]$	-0.2182	-0.3151	0.7684	$\mathbb{E}[\varepsilon_{1t}]$	-0.0078	-0.1883	0.8507
$\mathbb{E}[y_{2t}]$	0.1583	0.3973	0.7115	$\mathbb{E}[\varepsilon_{2t}]$	-0.0044	-0.1065	0.9152
$\mathbb{E}[y_{3t}]$	0.0343	0.1447	0.8920	$\mathbb{E}[\varepsilon_{3t}]$	0.0031	0.0767	0.9389
$\mathbb{E}[y_{4t}]$	0.0783	0.2079	0.8455	$\mathbb{E}[\varepsilon_{4t}]$	-0.0043	-0.1043	0.9170
$\mathbb{E}[y_{5t}]$	-0.4187	-1.5045	0.2069	$\mathbb{E}[\varepsilon_{5t}]$	0.0061	0.1450	0.8847

(a) Teste t para $\mathbb{E}[y_{it}] = 0$

	Value	t-Statistic	p-value
$\mathbb{E}[h_{1t}]$	0.9998	-0.0141	0.9888
$\mathbb{E}[h_{2t}]$	0.9903	-1.3757	0.1694
$\mathbb{E}[h_{3t}]$	1.0038	0.3239	0.7461
$\mathbb{E}[h_{4t}]$	1.0032	0.2026	0.8395
$\mathbb{E}[h_{5t}]$	0.9427	-3.6000	0.0003

(b) Teste t para $\mathbb{E}[\varepsilon_{it}] = 0$

(c) Teste t para $\mathbb{E}[h_{it}] = 1$

Figura 6.6: Teste dos valores esperados de y_t , ε_t e h_t

as hipóteses do modelo foram confirmadas, com exceção da esperança $\mathbb{E}[h_{5t}]$. Esta se mostrou estatisticamente diferente da unidade, porém com valor próximo da mesma.

Dando continuidade aos testes, a exemplo do que foi feito na avaliação da matriz de covariância $\text{Cov}(z_t)$, o teste t de Student também foi aplicado às matrizes de covariância de y_t e de ε_t para determinar se estas eram iguais à identidade. De acordo com os p-valores obtidos, que podem ser encontrados nas matrizes $\text{PV}(y_t)$ e $\text{PV}(\varepsilon_t)$ da Figura 6.7, isto é verdade, ou seja, $\text{Cov}(y_t) = \text{Cov}(\varepsilon_t) = I_m$.

$\text{Cov}(y_t) = \begin{bmatrix} 1.0004 & 0 & 0.0007 & -0.0005 & 0.0010 \\ 0 & 1.0017 & 0 & 0 & 0 \\ 0.0007 & 0 & 1.0013 & 0.0003 & -0.0005 \\ -0.0005 & 0 & 0.0003 & 1.0015 & 0.0004 \\ 0.0010 & 0 & -0.0005 & 0.0004 & 1.0009 \end{bmatrix}$	$\text{PV}(y_t) = \begin{bmatrix} 0.9957 & 0.9999 & 0.9878 & 0.9912 & 0.9802 \\ 0.9999 & 0.9809 & 0.9999 & 0.9999 & 0.9999 \\ 0.9878 & 0.9999 & 0.9855 & 0.9946 & 0.9897 \\ 0.9912 & 0.9999 & 0.9946 & 0.9836 & 0.9918 \\ 0.9802 & 0.9999 & 0.9897 & 0.9918 & 0.9952 \end{bmatrix}$
$\text{Cov}(\varepsilon_t) = \begin{bmatrix} 1.0072 & 0.0033 & -0.0215 & -0.0018 & 0.0098 \\ 0.0033 & 1.0156 & 0.0098 & 0.0033 & -0.0118 \\ -0.0215 & 0.0098 & 0.9987 & 0.0006 & 0.0020 \\ -0.0018 & 0.0033 & 0.0006 & 1.0009 & -0.0040 \\ 0.0098 & -0.0118 & 0.0020 & -0.0040 & 1.0599 \end{bmatrix}$	$\text{PV}(\varepsilon_t) = \begin{bmatrix} 0.9141 & 0.9419 & 0.6244 & 0.9639 & 0.8088 \\ 0.9419 & 0.8297 & 0.8044 & 0.9392 & 0.7725 \\ 0.6244 & 0.8044 & 0.9858 & 0.9867 & 0.9625 \\ 0.9639 & 0.9392 & 0.9867 & 0.9895 & 0.9142 \\ 0.8088 & 0.7725 & 0.9625 & 0.9142 & 0.6804 \end{bmatrix}$

Figura 6.7: Teste das matrizes de covariância de y_t e ε_t

Finalmente, foram analisados os gráficos dos fatores y_{it} , que podem ser encontrados na Figura 6.8. A semelhança destes com os gráficos de retornos financeiros é aparente e natural, já que os retornos são modelados como uma combinação linear dos fatores.

6.2.3 Análise da Covariância Condicional

As estimativas da matriz Σ_t de covariância condicional, de certa forma, definem a distribuição de probabilidade dos retornos r_t . Estas têm, então, efeito direto sobre a seleção de um portfólio de risco mínimo e também sobre o cálculo de avaliações de risco. Assim, estas estimativas têm grande importância em relação aos objetivos deste trabalho.

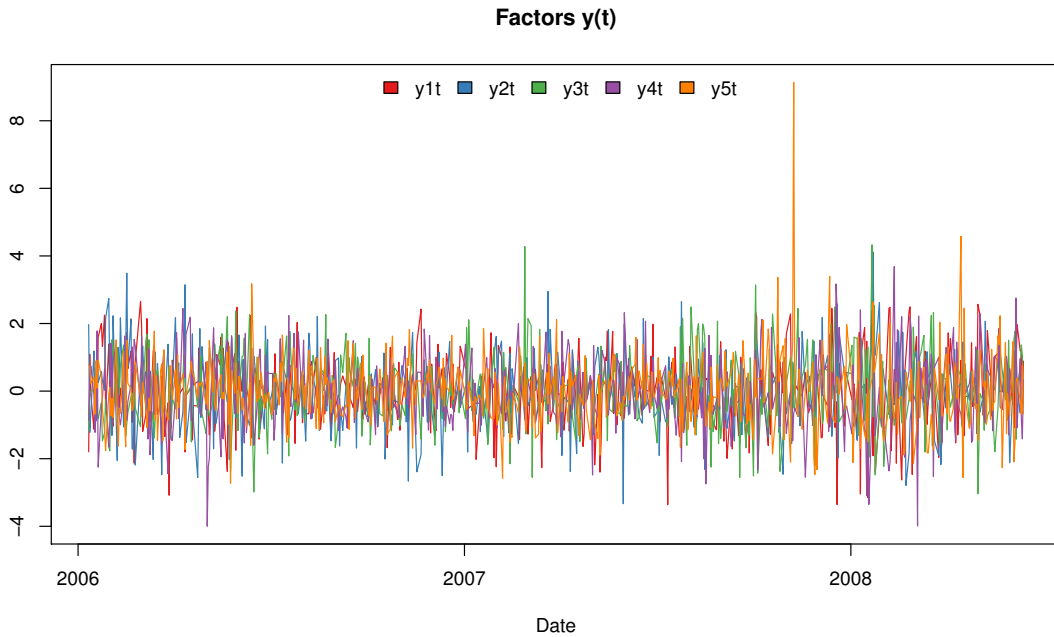


Figura 6.8: Fatores do modelo GO-GARCH ajustado

Desta forma, nesta seção é apresentada uma análise qualitativa das estimativas de Σ_t , dadas por

$$\Sigma_t = Z H_t Z', \quad (6.2.3)$$

buscando determinar a validade das mesmas. Para isso, foram criados diversos gráficos mostrando a evolução das entradas de Σ_t ao longo do tempo. Além disso, foi feita uma comparação das volatilidades condicionais calculadas, tendo o modelo IGARCH(1,1) como referência.

Em primeiro lugar, considere as figuras 6.9 e 6.10. Na primeira destas, são mostrados os elementos da diagonal de Σ_t ao longo do tempo. Estes correspondem às variâncias condicionais de $\{x_t\}_{t \geq 1}$. Na segunda figura, cada gráfico corresponde a uma das colunas de Σ_t , ou seja, cada gráfico mostra as covariâncias condicionais de $\{x_t\}_{t \geq 1}$ em função do tempo t .

Observando os gráficos das figuras 6.9 e 6.10, é possível notar que as séries de variâncias e covariâncias apresentam um mesmo comportamento. Isto é, suas subidas e descidas são sincronizadas. Uma vez que as variâncias e as covariâncias condicionais são calculadas com base no mesmo conjunto de fatores, esta observação não chega a ser surpreendente. De acordo com os resultados, o que muda de série para série é a intensidade com que estas respondem às mudanças nos fatores.

A sincronia existente entre as séries das figuras 6.9 e 6.10 também parece de acordo com as correlações identificadas na Tabela C.2 do Apêndice C. Consultando esta tabela, é possível verificar que as ações aqui utilizadas apresentam correlação positiva bastante alta.

Considere agora a Figura 6.11, onde estão representados os retornos das ações juntamente com os desvios padrão condicionais calculados pelos modelos VAR+GO-GARCH e IGARCH.

Pelos gráficos, é possível ver que os dois desvios padrão calculados parecem aumentar na proximidade de um cluster de volatilidade, como esperado. Os desvios padrão IGARCH, entretanto, parecem oscilar mais que os calculados pelo modelo VAR+GO-GARCH. Estes, como identificado em [53], aparentemente não ultrapassam um patamar mínimo. Esta maior oscilação dos desvios padrão IGARCH pode ser vista com maior clareza na Figura 6.12, que compara as volatilidades dos dois modelos.

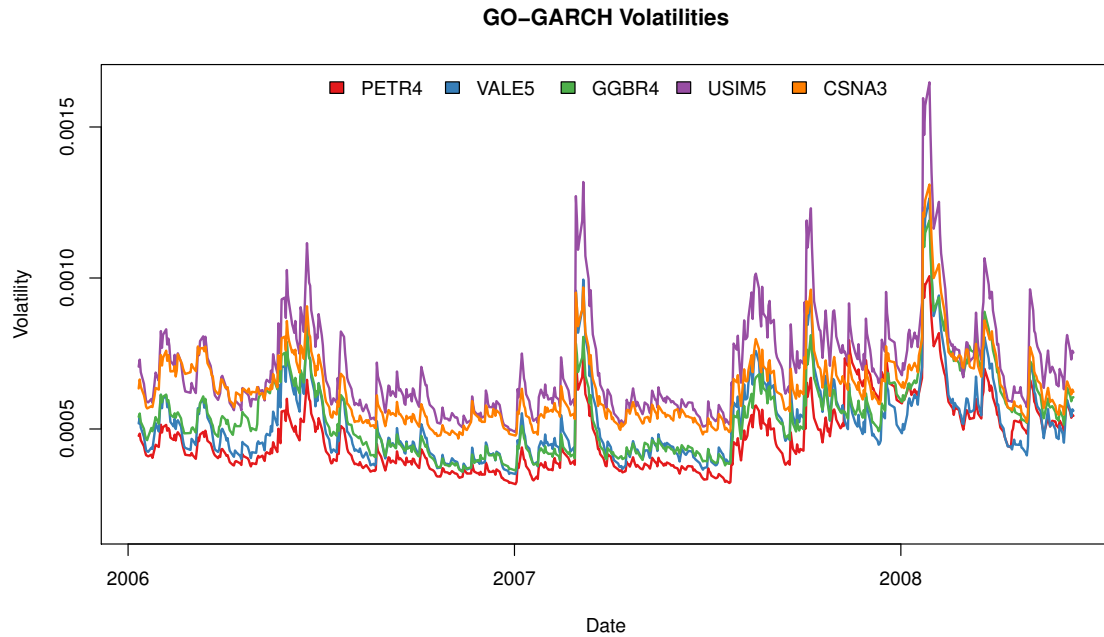


Figura 6.9: Volatilidades condicionais do modelo GO-GARCH

Na Figura 6.12 é possível ver que em diversos momentos a volatilidade estimada pelo modelo IGARCH se encontra bem acima ou bem abaixo da estimada pelo modelo VAR+GO-GARCH.

6.3 Conclusão

Neste capítulo foi apresentado o algoritmo para ajuste de um modelo VAR+GO-GARCH a dados históricos e os resultados do ajuste às séries históricas de cinco ações da BM&FBovespa: PETR4, VALE5, GGBR4, USIM5 e CSNA3.

O diagnóstico do ajuste demonstrou que o modelo VAR+GO-GARCH foi bem sucedido em capturar a dinâmica dos retornos das cinco ações selecionadas. Isto foi comprovado, principalmente, pela ausência de autocorrelações nos dois momentos das séries de resíduos VAR e GO-GARCH. Frente a estes, o fato da esperança $\mathbb{E}[h_{5t}]$ ser menor que a unidade não é tão relevante. Além disso, tendo sido este o único desvio encontrado das hipóteses do modelo, é possível considerar que o ajuste do modelo foi bem sucedido.

Finalmente, a comparação das volatilidades calculadas pelo modelo VAR+GO-GARCH com as calculadas pelo modelo IGARCH(1,1) demonstrou que em relação aos retornos r_t ambas apresentam o comportamento esperado. Entretanto, a volatilidade IGARCH se mostrou menos estável e com maiores oscilações.

Como explicado na Seção 2.2, o modelo IGARCH(1,1) foi escolhido após uma extensiva bateria de testes de modelos univariados. Desta forma, caso as volatilidades VAR+GO-GARCH sejam adequadas para aplicações de avaliação de risco, é possível que exista uma grande vantagem no seu uso, dado que as volatilidades calculadas não são tão extremas quanto às encontradas com o modelo IGARCH.

Os relatórios contendo o diagnóstico dos cinco ajustes IGARCH realizados pode ser encontrado no Apêndice D.

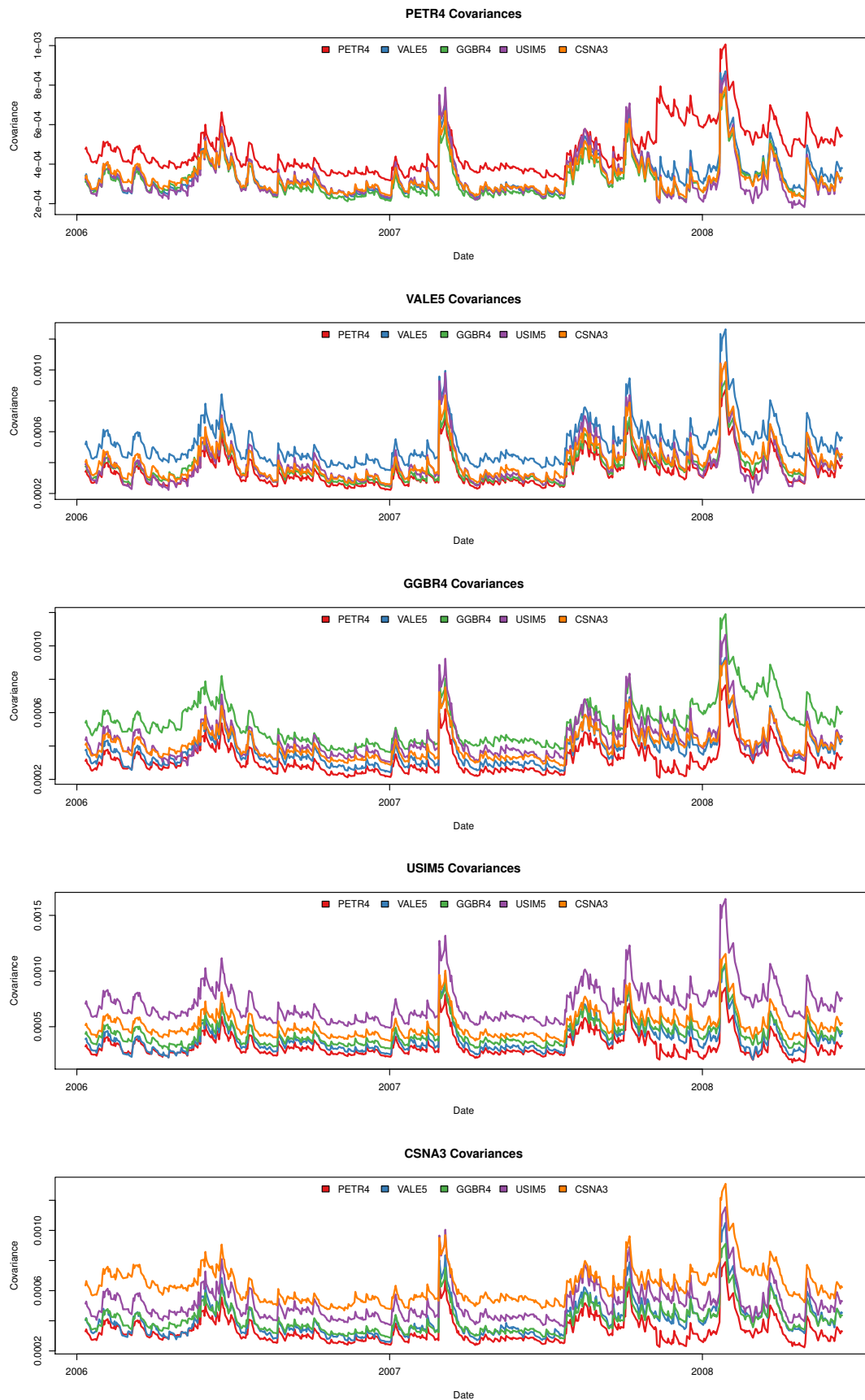


Figura 6.10: Covariâncias condicionais do modelo GO-GARCH ajustado

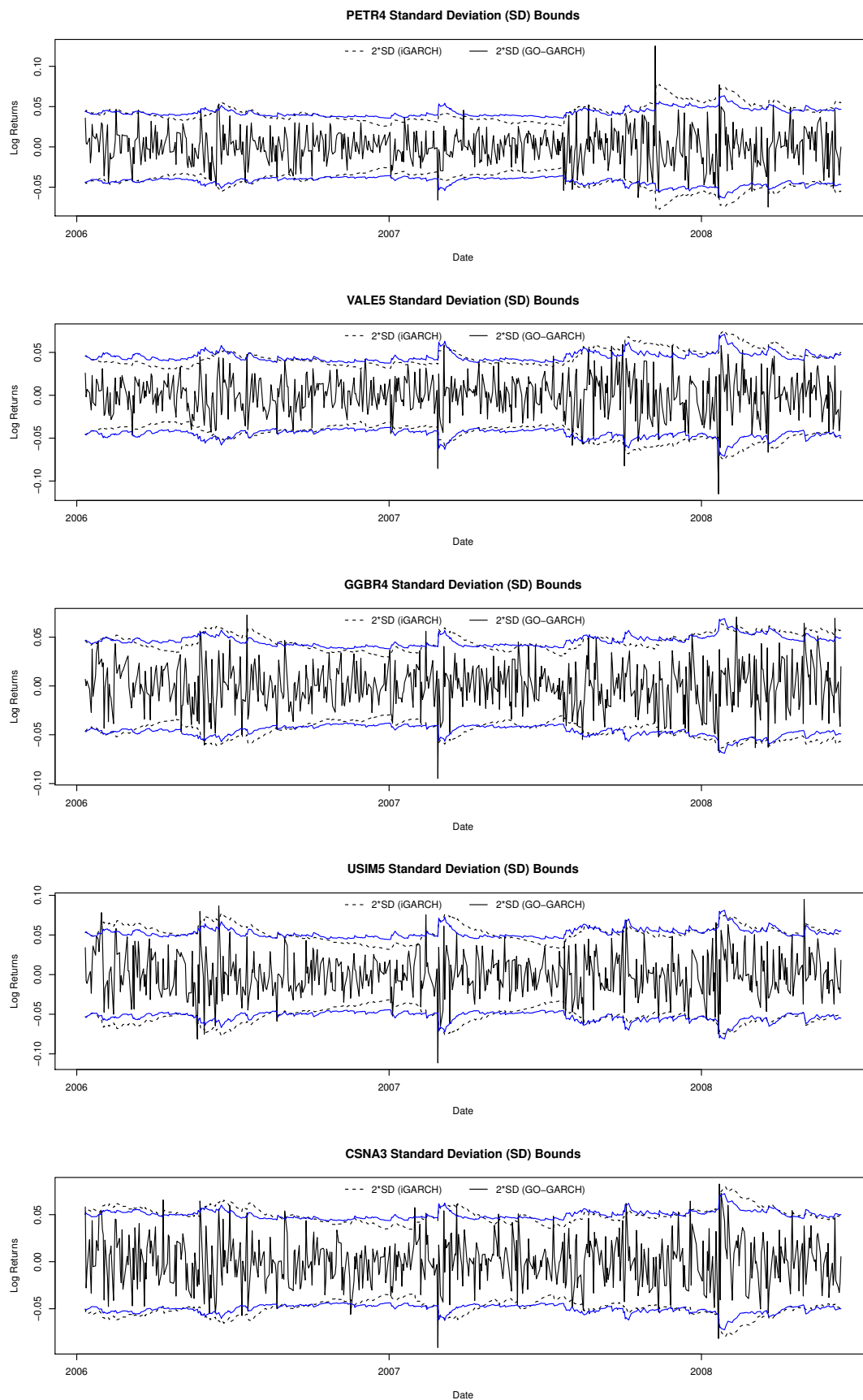


Figura 6.11: Comparação de desvios padrão: GO-GARCH versus IGARCH

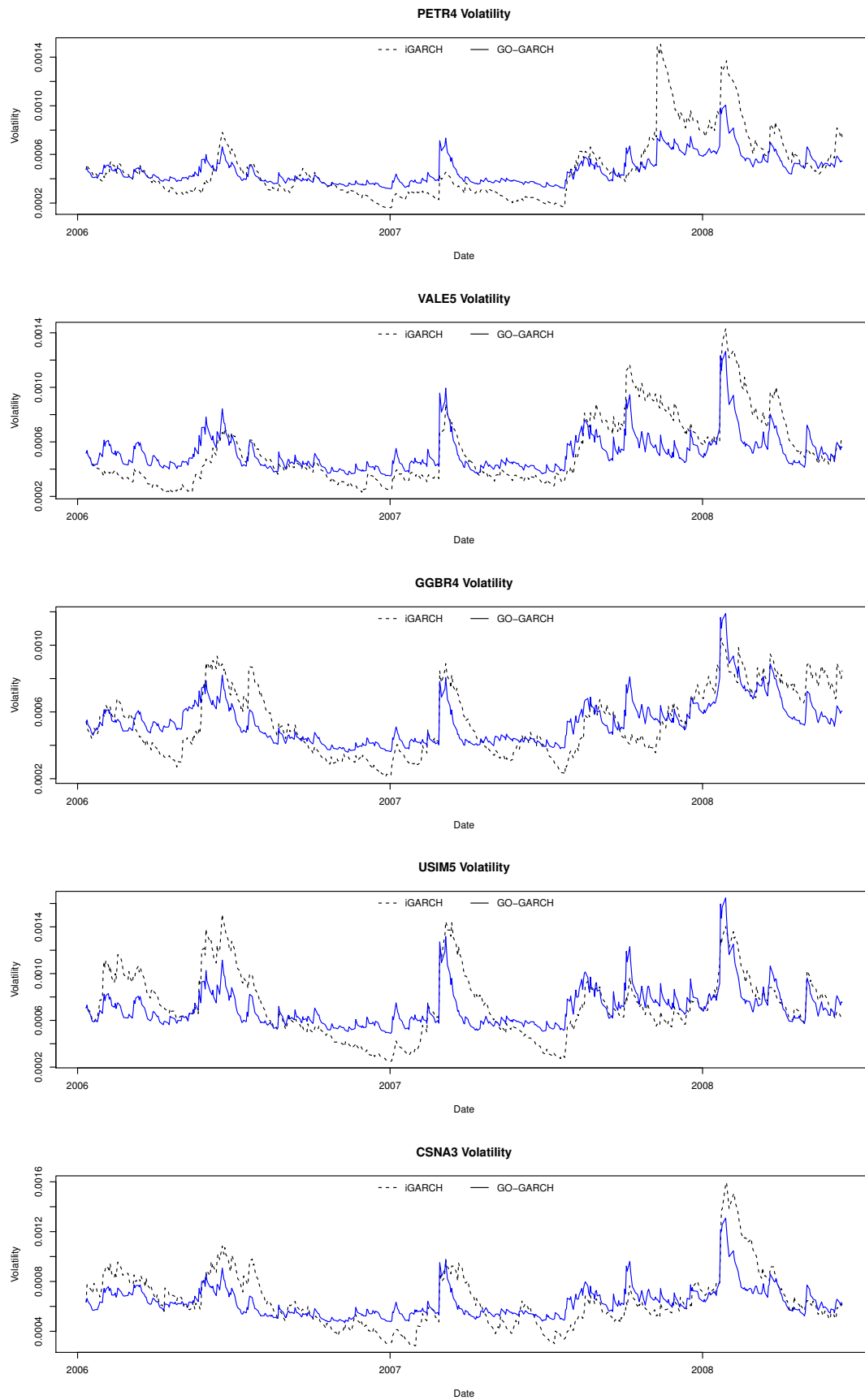


Figura 6.12: Comparação de volatilidades: GO-GARCH versus IGARCH

Capítulo 7

Avaliação do Risco de um Portfólio

Previsões de medidas de risco são fundamentais para o bom gerenciamento de qualquer portfólio, pois permitem que o investidor tome ações de forma a se prevenir contra grandes perdas financeiras. Com o objetivo de auxiliar na tomada de decisões de um investidor, foi implementado um algoritmo para calcular previsões de VaR e CVaR com o horizonte de um dia.

Na Seção 7.1, são apresentados detalhes do algoritmo implementado e também a metodologia de cálculo das medidas de risco. Em seguida, na Seção 7.2, é apresentado o procedimento de *backtesting* realizado para avaliar o algoritmo. Depois, na Seção 7.3, são apresentados os resultados obtidos. Finalizando o capítulo, na Seção 7.4, são apresentadas as conclusões.

7.1 Implementação

A metodologia adotada para o cálculo das previsões de risco é bem conhecida. Em primeiro lugar, escolhe-se um modelo para o retorno dos ativos que formam o portfólio em avaliação. Tendo este modelo e as informações colhidas até uma data t como referência, é feita a simulação de diversos cenários para a data futura $t + 1$. Então, as medidas de risco podem ser estimadas com base nos diversos cenários calculados. Veja [52, 14].

A metodologia descrita acima foi implementada utilizando os modelos GO-GARCH, O-GARCH e também o estimador de CVaR baseado na função \tilde{F}_β , definido na Seção 3.2. A escolha destes é justificada a seguir.

Em primeiro lugar, como pôde ser verificado na Seção 5.2.2, o uso de modelos multivariados pode resultar em melhores estimativas para o VaR e para o CVaR de um portfólio. Desta forma, dado que o ajuste dos modelos VAR e GO-GARCH aos dados históricos apresentou bons resultados no Capítulo 6, a escolha destes seria bastante natural. O modelo O-GARCH foi adotado para avaliar o desempenho do modelo GO-GARCH, que tem estimadores mais complexos. Desta forma, espera-se que o modelo GO-GARCH apresente melhores resultados.

Também seria natural a escolha pela estimação das medidas de risco através de amostras de retornos. Em muitos casos, a distribuição dos retornos simulados é bastante complexa e pode não ter fórmula fechada. Como exemplo, os próprios modelos GO-GARCH podem ser citados, pois estes favorecem a obtenção de distribuições complexas ao permitir a combinação de fatores regidos por diferentes modelos. Nestes casos, a utilização do estimador baseado na função \tilde{F}_β se apresenta como a única opção viável.

Considere, então, o Algoritmo 4. Neste, o procedimento completo para o cálculo das previsões de risco foi implementado na função *RISK_VAR_GOGARCH*, que recebe cinco parâmetros: a matriz R_t de retornos históricos multivariados, a composição w do portfólio, a ordem p do modelo VAR, o nível β de VaR e o número N de amostras a serem geradas nas simulações dos fatores.

Acompanhando o pseudocódigo da função *RISK_VAR_GOGARCH*, é possível ver que a primeira operação realizada consiste no ajuste dos modelos VAR e GO-GARCH à série de retornos multivariados que está armazenada na matriz R_t . Este ajuste é realizado através da chamada à função *VAR_GOGARCH*, que foi definida na Seção 6.1.

Em seguida, a função *GARCH_SIM* é utilizada para simular os m processos GARCH(1,1) usados na modelagem dos fatores do modelo GO-GARCH (linhas 3 a 5). Terminadas as simulações, o conjunto r_{t+1} dos possíveis retornos para o tempo $t + 1$ é, então, calculado (linhas 6 a 9). Neste cálculo, os valores obtidos para os fatores são combinados pela matriz Z do modelo GO-GARCH e depois somados à previsão do retorno médio m_{t+1} que é calculada pela função *VAR_PREVISAO*, como indicado na Seção 2.4.1.

Então, a função *CVaR* é utilizada para calcular as estimativas das medidas VaR e CVaR. Esta função recebe como parâmetros a composição w do portfólio, os retornos simulados para o tempo $t + 1$ e o nível β de VaR. Vale ressaltar que a função *CVaR* calcula as estimativas de risco por programação linear, através da resolução do Problema (3.2.17).

Algorithm 4 Algoritmo para estimar o risco de um portfólio com modelos VAR e GO-GARCH

```

1: procedure RISK_VAR_GOGARCH( $R_t, w, p, \beta, N$ )
2:    $\{B, M_t, X_t, Z, Y_t, G, H_t\} := \mathbf{VAR\_GOGARCH}(R_t, p)$ 
3:   for  $k := 1, \dots, m$  do
4:      $y_k := \mathbf{GARCH\_SIM}(V'_{t,k}, G_k, N)$ 
5:   end for
6:    $y_{t+1} := [y_1 \ y_2 \ \dots \ y_m]'$ 
7:    $m_{t+1} := \mathbf{VAR\_PREVISAO}(M_t, B)$ 
8:    $x_{t+1} := Z \ y_{t+1}$ 
9:    $r_{t+1} := x_{t+1} + m_{t+1}$ 
10:   $\{VaR_\beta, CVaR_\beta\} := \mathbf{CVaR}(w, r_{t+1}, \beta)$ 
11:  return  $VaR_\beta, CVaR_\beta$ 
12: end procedure

```

Finalmente, quanto à sua implementação computacional, o algoritmo foi implementado em R, utilizando as funções previamente criadas para o ajuste de modelos GO-GARCH e para o cálculo das estimativas de risco. Assim, o algoritmo depende, indiretamente, da biblioteca de otimização *MOSEK* e dos pacotes *R-to-MOSEK* e *fastICA*. Veja [47, 43, 22, 40].

7.2 Backtesting

A avaliação das previsões de VaR e CVaR geralmente é feita através de um procedimento de *backtesting*, ou seja, através da comparação das perdas observadas com as previsões feitas com base em dados históricos. Veja [18].

Assim, recorreu-se a este procedimento a fim de avaliar as previsões calculadas pelo Algoritmo 4. Para isto, foi necessário definir: o portfólio a ser testado, o modelo dos retornos financeiros, os parâmetros de simulação, o conjunto de dados históricos e o critério de avaliação das previsões. Cada um destes tópicos é abordado a seguir.

Primeiramente, dados os bons resultados obtidos no Capítulo 6, optou-se por utilizar o mesmo conjunto A de ações e o mesmo modelo para os retornos. Assim, foram escolhidos o portfólio Π_5 , cuja composição pode ser vista na Tabela 7.1, e o modelo VAR de ordem $p = 5$ em conjunto com o modelo GO-GARCH, sendo este último estimado por ICA.

A_i	x_i
PETR4	0.25
VALE5	0.20
GGBR4	0.10
USIM5	0.15
CSNA3	0.30

Tabela 7.1: Composição do portfólio Π_5

Para o cálculo das estimativas, foi utilizado um total de $N = 50000$ simulações. Com base nos testes da Seção 4.1, este valor pode ser considerado um bom compromisso entre o tempo computacional e a precisão das estimativas calculadas.

Em seguida, foi definido o conjunto de dados históricos. Naturalmente, foram utilizadas todas as informações históricas disponíveis, ou seja, foram utilizados os preços de fechamento entre 01/01/2006 e 14/01/2013. Entretanto, como são necessárias diversas previsões das medidas de risco, é necessário limitar a quantidade de informação com que cada uma destas é calculada. Assim, optou-se por calculá-las com base em uma janela deslizante de 600 dias. Isto é, a previsão para uma data $t + 1$ foi feita com base nos preços observados entre a data t e $t - 600$. Desta forma, foi possível calcular estimativas para todas as datas entre 13/06/2008 e 14/01/2013.

Por último, então, foram definidos os critérios para avaliar as previsões feitas. Dada a disponibilidade de funções implementadas em R para o teste de previsões de VaR¹, a avaliação do Algoritmo 4 se limitou a esta medida. Foram utilizados três testes estatísticos que são baseados em violações do nível de VaR: o teste de Kupiec, o teste de independência serial das violações (IT98) que foi publicado em [10] e o teste de independência baseado em *durations* (DBIT), que foi publicado em [9].

O teste de Kupiec indica se o número de violações de VaR corresponde à porcentagem esperada do número total de dias observados. Esta porcentagem deve ser igual a $1 - \beta$. Veja [35].

Os testes de independência medem se as violações são realmente independentes entre si. Quando estas são independentes, o modelo adotado para calculá-las se prova capaz de adaptar suas previsões corretamente, após a ocorrência de um choque. Esta característica é desejável dada a presença dos *clusters de volatilidade* em séries de retornos financeiros. Veja [52, 18].

Conforme o esperado, os dois testes de independência diferem na forma como a medem. O primeiro teste, IT98, utiliza cadeias de Markov para determinar se a ocorrência de uma violação em uma data t provê alguma informação sobre a ocorrência de uma nova violação na data $t + 1$. Por considerar apenas dias consecutivos, este é um teste de primeira ordem. Já o segundo teste, DBIT, utiliza a série de intervalos de tempo entre violações consecutivas para avaliar a independência das mesmas. Assim, este é um teste de ordem superior.

Detalhes sobre estes métodos estão fora do escopo deste trabalho. Aqui, limita-se a avaliar o resultado dos mesmos através dos p-valores por eles fornecidos. Para maiores detalhes sobre estes métodos, é recomendada a leitura de [18].

Um segundo procedimento de *backtesting* foi realizado para avaliar o efeito dos diferentes modelos de retorno sobre as previsões de VaR. Neste, foi utilizada uma janela deslizante de 1000 dias e um total de $N = 300000$ simulações, buscando reduzir ao máximo possível os erros de estimação das medidas de risco. Foram utilizados os dois estimadores disponíveis para modelos GO-GARCH e também o modelo O-GARCH.

Na próxima seção são apresentadas as estimativas de risco calculadas nos dois procedimentos de *backtesting* e o resultado da avaliação das mesmas.

7.3 Resultados

Finalmente, considere a Tabela 7.2, que contém o resultado dos testes estatísticos utilizados para avaliar a qualidade das previsões de VaR realizadas. As colunas desta tabela apresentam: o período de teste (colunas *from* e *to*), o teste aplicado, o nível de VaR sendo testado, o p-valor resultante, a estatística de teste, o número de amostras, o número de violações observadas e a porcentagem de não violações, que corresponde ao nível β de VaR.

from	to	test	VaR.level	p.value	statistic	samples	violations	non.violations.ratio
2008-06-13	2013-01-14	Kupiec	90	0.0610	3.5095	1136	133	88.2923
2008-06-13	2013-01-14	IT	90	0.2199	1.5052	1136	133	88.2923
2008-06-13	2013-01-14	DBIT	90	1.0000	-5.3555	1136	133	88.2923
2008-06-13	2013-01-14	Kupiec	95	0.1387	2.1927	1136	68	94.0141
2008-06-13	2013-01-14	IT	95	0.9688	0.0015	1136	68	94.0141
2008-06-13	2013-01-14	DBIT	95	0.3644	0.8228	1136	68	94.0141
2008-06-13	2013-01-14	Kupiec	99	0.1924	1.6989	1136	16	98.5915
2008-06-13	2013-01-14	IT	99	0.4988	0.4576	1136	16	98.5915
2008-06-13	2013-01-14	DBIT	99	0.0012	10.4741	1136	16	98.5915

(a) Jun/2008 a Jan/2013

from	to	test	VaR.level	p.value	statistic	samples	violations	non.violations.ratio
2010-01-04	2013-01-14	Kupiec	90	0.2372	1.3969	751	85	88.6818
2010-01-04	2013-01-14	IT	90	0.0659	3.3831	751	85	88.6818
2010-01-04	2013-01-14	DBIT	90	1.0000	-0.4017	751	85	88.6818
2010-01-04	2013-01-14	Kupiec	95	0.3720	0.7970	751	43	94.2743
2010-01-04	2013-01-14	IT	95	0.7258	0.1230	751	43	94.2743
2010-01-04	2013-01-14	DBIT	95	1.0000	-0.9812	751	43	94.2743
2010-01-04	2013-01-14	Kupiec	99	0.8589	0.0316	751	8	98.9348
2010-01-04	2013-01-14	IT	99	0.6779	0.1725	751	8	98.9348
2010-01-04	2013-01-14	DBIT	99	0.2515	1.3152	751	8	98.9348

(b) Jan/2010 a Jan/2013 (pós-crise)

from	to	test	VaR.level	p.value	statistic	samples	violations	non.violations.ratio
2008-06-13	2009-12-30	Kupiec	90	0.1187	2.4350	385	48	87.5325
2008-06-13	2009-12-30	IT	90	0.6334	0.2274	385	48	87.5325
2008-06-13	2009-12-30	DBIT	90	1.0000	-4.4678	385	48	87.5325
2008-06-13	2009-12-30	Kupiec	95	0.1977	1.6591	385	25	93.5065
2008-06-13	2009-12-30	IT	95	0.5743	0.3156	385	25	93.5065
2008-06-13	2009-12-30	DBIT	95	0.1944	1.6839	385	25	93.5065
2008-06-13	2009-12-30	Kupiec	99	0.0634	3.4472	385	8	97.9221
2008-06-13	2009-12-30	IT	99	0.5596	0.3405	385	8	97.9221
2008-06-13	2009-12-30	DBIT	99	0.0516	3.7892	385	8	97.9221

(c) Jun/2008 a Dez/2009 (crise)

Tabela 7.2: Avaliação das previsões de VaR feitas com modelo GO-GARCH (ICA)

Como pode ser observado na Tabela 7.2, a informação sobre os testes foi dividida em três blocos. O primeiro destes contém as informações referentes a todas as previsões realizadas. O segundo, as relativas ao período entre 2010 e 2013. E o terceiro bloco, as relativas às datas anteriores a 2010, ou seja, relativas ao período que contém a crise de 2008.

Seguindo a prática, as hipóteses nulas dos testes são rejeitadas apenas quando os seus p-valores são menores que 5%. No caso do teste de Kupiec, a rejeição da hipótese nula indica que o número de violações de VaR não corresponde ao esperado. Para os testes de independência, esta indica a existência de algum tipo de dependência entre as violações de VaR.

Voltando à Tabela 7.2, nota-se que a única hipótese nula rejeitada, assinalada em vermelho, é a do teste de independência DBIT para $\beta = 99\%$, no primeiro bloco de testes. Além deste, alguns outros

¹Os dois testes de independência utilizados foram implementados pela autora do trabalho [18], a quem agradecemos por tê-las tornado disponíveis para este trabalho.

testes apresentaram p-valores baixos, mas não o suficiente para terem suas hipóteses nulas rejeitadas.

Foi a rejeição do teste DBIT que motivou a divisão das previsões em dois blocos, para permitir a análise dos períodos durante e após a crise de 2008. Analisando cada um dos blocos separadamente, não é possível rejeitar a hipótese nula de nenhum dos testes realizados. Entretanto, o teste DBIT para $\beta = 99\%$ ainda apresenta um p-valor baixo para o período da crise.

De forma geral, no período pós crise (segundo bloco), as previsões de VaR parecem ser bastante adequadas. Apenas um dos testes teve p-valor abaixo de 20%. Entretanto, para os valores mais críticos de β (95% e 99%), bons resultados foram obtidos em todos os testes, ou seja, os p-valores destes são grandes o suficiente para que suas hipóteses nulas não sejam rejeitadas com confiança.

Durante o período da crise, os testes não apresentaram resultados tão bons quanto os do segundo bloco. Porém, considerando que a crise de 2008 foi marcada por um período de grandes oscilações, os resultados obtidos podem ser considerados animadores. Neste período, os únicos p-valores abaixo de 10% foram encontrados nos testes de Kupiec e DBIT para $\beta = 99\%$.

Considere agora a Tabela 7.3, que contém os resultados obtidos no segundo procedimento de *backtesting*. Nesta, pode-se ver que as previsões de VaR calculadas com o modelo O-GARCH não apresentaram violações com as frequências esperadas, como indicam os testes de Kupiec rejeitados. Além disso, pode-se ver que, de forma geral, os resultados obtidos para o modelo GO-GARCH com o estimador ICA foram melhores que os obtidos com o estimador MM2011. Os testes estatísticos realizados para estes dois casos, entretanto, apresentaram resultados positivos.

from	to	test	VaR.level	p.value	statistic	samples	violations	non.violations.ratio
2010-01-26	2013-01-14	Kupiec	90	0.9412	0.0054	736	73	90.0815
2010-01-26	2013-01-14	IT	90	0.9174	0.0108	736	73	90.0815
2010-01-26	2013-01-14	DBIT	90	1.0000	-6.8079	736	73	90.0815
2010-01-26	2013-01-14	Kupiec	95	0.2354	1.4079	736	30	95.9239
2010-01-26	2013-01-14	IT	95	0.1100	2.5540	736	30	95.9239
2010-01-26	2013-01-14	DBIT	95	1.0000	-3.8412	736	30	95.9239
2010-01-26	2013-01-14	Kupiec	99	0.8930	0.0181	736	7	99.0489
2010-01-26	2013-01-14	IT	99	0.7137	0.1346	736	7	99.0489
2010-01-26	2013-01-14	DBIT	99	0.2361	1.4040	736	7	99.0489

(a) GO-GARCH (ICA)

from	to	test	VaR.level	p.value	statistic	samples	violations	non.violations.ratio
2010-01-26	2013-01-14	Kupiec	90	0.2820	1.1577	736	65	91.1685
2010-01-26	2013-01-14	IT	90	0.1629	1.9473	736	65	91.1685
2010-01-26	2013-01-14	DBIT	90	1.0000	-4.6313	736	65	91.1685
2010-01-26	2013-01-14	Kupiec	95	0.0544	3.7011	736	26	96.4674
2010-01-26	2013-01-14	IT	95	0.1673	1.9073	736	26	96.4674
2010-01-26	2013-01-14	DBIT	95	1.0000	-1.5076	736	26	96.4674
2010-01-26	2013-01-14	Kupiec	99	0.8930	0.0181	736	7	99.0489
2010-01-26	2013-01-14	IT	99	0.7137	0.1346	736	7	99.0489
2010-01-26	2013-01-14	DBIT	99	0.3924	0.7315	736	7	99.0489

(b) GO-GARCH (MM2011)

from	to	test	VaR.level	p.value	statistic	samples	violations	non.violations.ratio
2010-01-26	2013-01-14	Kupiec	90	0.0003	13.0940	736	46	93.7500
2010-01-26	2013-01-14	IT	90	0.5612	0.3375	736	46	93.7500
2010-01-26	2013-01-14	DBIT	90	1.0000	-5.4747	736	46	93.7500
2010-01-26	2013-01-14	Kupiec	95	0.0038	8.3934	736	21	97.1467
2010-01-26	2013-01-14	IT	95	0.2663	1.2355	736	21	97.1467
2010-01-26	2013-01-14	DBIT	95	0.3513	0.8689	736	21	97.1467
2010-01-26	2013-01-14	Kupiec	99	0.1729	1.8573	736	4	99.4565
2010-01-26	2013-01-14	IT	99	0.8343	0.0438	736	4	99.4565
2010-01-26	2013-01-14	DBIT	99	1.0000	-1.3467	736	4	99.4565

(c) O-GARCH

Tabela 7.3: Avaliação das previsões de VaR feitas com diversos modelos de retornos

Além dos testes estatísticos, foram criados diversos gráficos para auxiliar na avaliação do algoritmo implementado. Na Figura 7.1, é possível ver os gráficos das estimativas de risco utilizadas nos testes da Tabela 7.2 junto com os retornos diários do portfólio. Na Figura 7.2 foi feita uma versão suavizada dos mesmos gráficos, para permitir uma melhor visualização dos diferentes níveis das medidas de VaR e CVaR.

Também foram criados gráficos sobrepondo as violações de VaR observadas, representadas por pontos azuis, e as estimativas de risco para um mesmo valor de β . Estes podem ser vistos nas figuras 7.3, 7.4 e 7.5. Nestas figuras, também foram incluídas as versões suaves das curvas de VaR e CVaR.

Analisando os gráficos da Figura 7.2, é possível ver que os aumentos e reduções nos valores das medidas de risco acompanham a maior ou menor volatilidade encontrada nas séries de retornos diários. Como exemplos, é possível notar que as métricas de risco têm valores bastante elevados na proximidade do grande cluster de volatilidade encontrado em 2008 e no outro cluster encontrado em meados de 2010.

Além disso, conforme o esperado, para quaisquer dois valores β_1 e β_2 de nível de VaR, tais que $\beta_1 > \beta_2$, foi possível verificar que $VaR_{\beta_1} > VaR_{\beta_2}$, que $CVaR_{\beta_1} > CVaR_{\beta_2}$ e que $CVaR_{\beta_1} > VaR_{\beta_1}$. Apesar dos gráficos disponíveis indicarem estes resultados visualmente, os mesmos também foram verificados numericamente.

Finalmente, considere os gráficos da Figura 7.5, calculados para $\beta = 99\%$. O número reduzido de violações permite ver que as mesmas acontecem, em geral, em momentos de grande variação dos retornos, de acordo com o esperado.

7.4 Conclusão

Neste capítulo foi apresentado um algoritmo para a avaliação do risco de um portfólio através do cálculo de previsões das medidas VaR e CVaR que são baseadas em simulações de um modelo para os retornos diários do portfólio. No algoritmo foram adotados o modelo VAR+GO-GARCH (ver Capítulo 6) e o estimador de CVaR baseado na função \tilde{F}_β , definido na Seção 3.2.

Para avaliar a qualidade das previsões calculadas pelo algoritmo, foram realizados diversos testes estatísticos e também foram elaborados gráficos que permitem visualizar as medidas de risco calculadas e as violações de VaR observadas. De forma geral, os testes de independência e de Kupiec realizados para avaliar as previsões não acusaram grandes problemas.

Entretanto, a rejeição apontada no primeiro bloco da Tabela 7.2 para o teste DBIT com $\beta = 99\%$, merece atenção. Como mostrado na Seção 7.3, ao dividir as previsões em dois blocos, a rejeição apontada anteriormente não mais ocorre. Assim, a rejeição pode ter sido causada pela presença de dois períodos econômicos bastante diferentes no primeiro bloco de testes. Algo semelhante foi reportado em [14], onde foram analisadas previsões de VaR para o índice S&P 500.

Finalmente, dado que a análise visual dos gráficos criados não foi capaz de apontar nada fora do esperado e dados os resultados positivos dos testes estatísticos, é possível concluir que o algoritmo cumpre o seu propósito de forma adequada.

Os resultados do segundo procedimento de *backtesting* mostram que, entre os modelos adotados, o modelo O-GARCH foi o menos apropriado para o cálculo das previsões de VaR. Este modelo, dado o menor número de violações, parece superestimar o risco de perdas financeiras nos casos estudados.

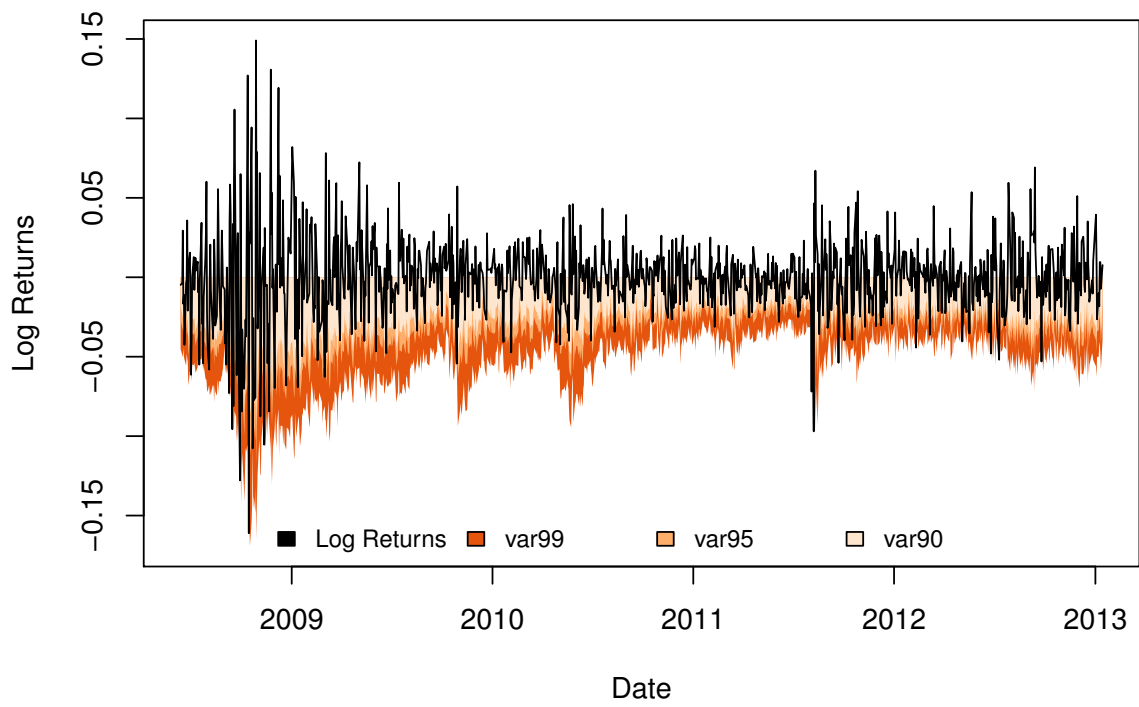
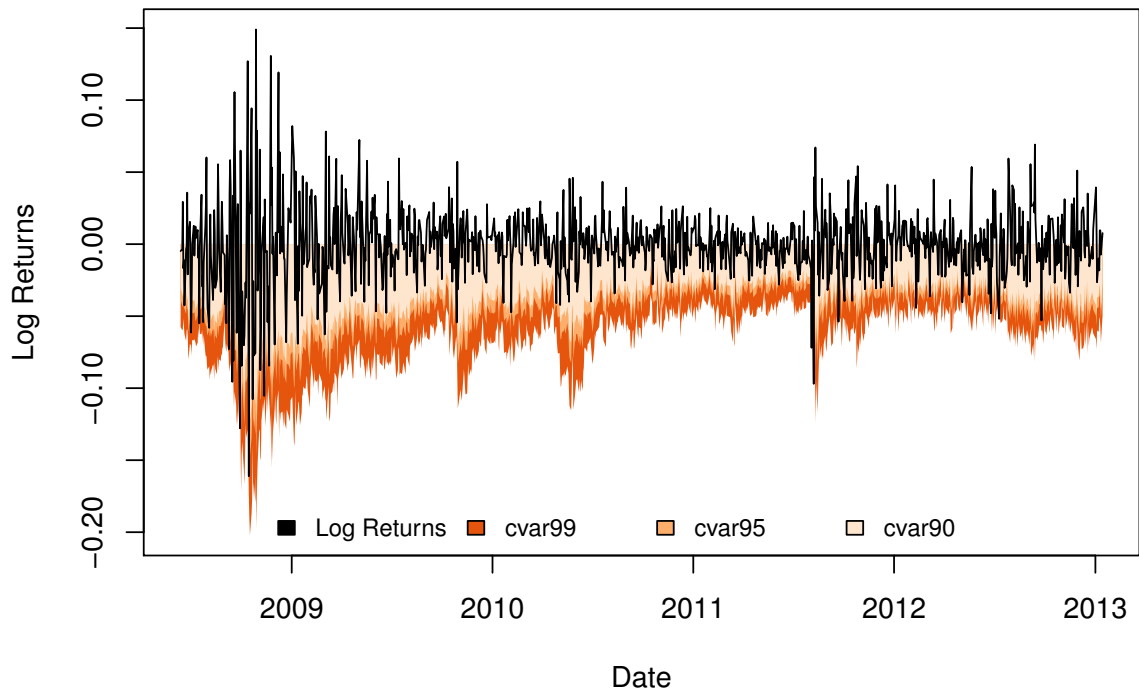


Figura 7.1: Previsões de VaR e CVaR

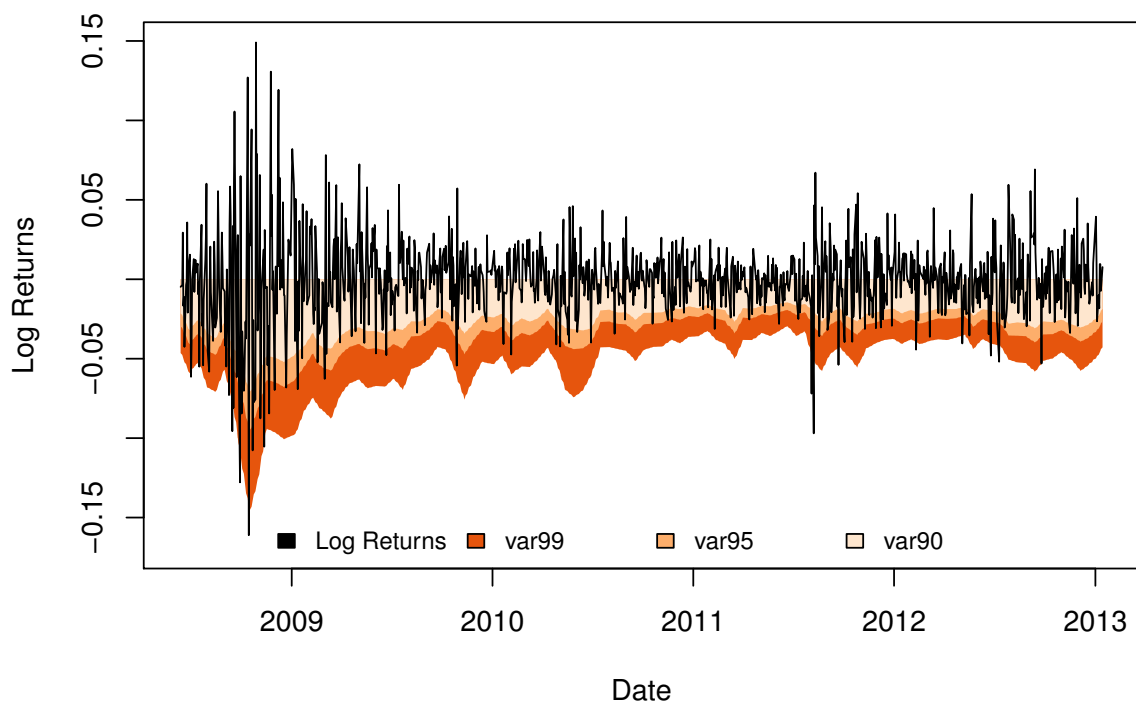
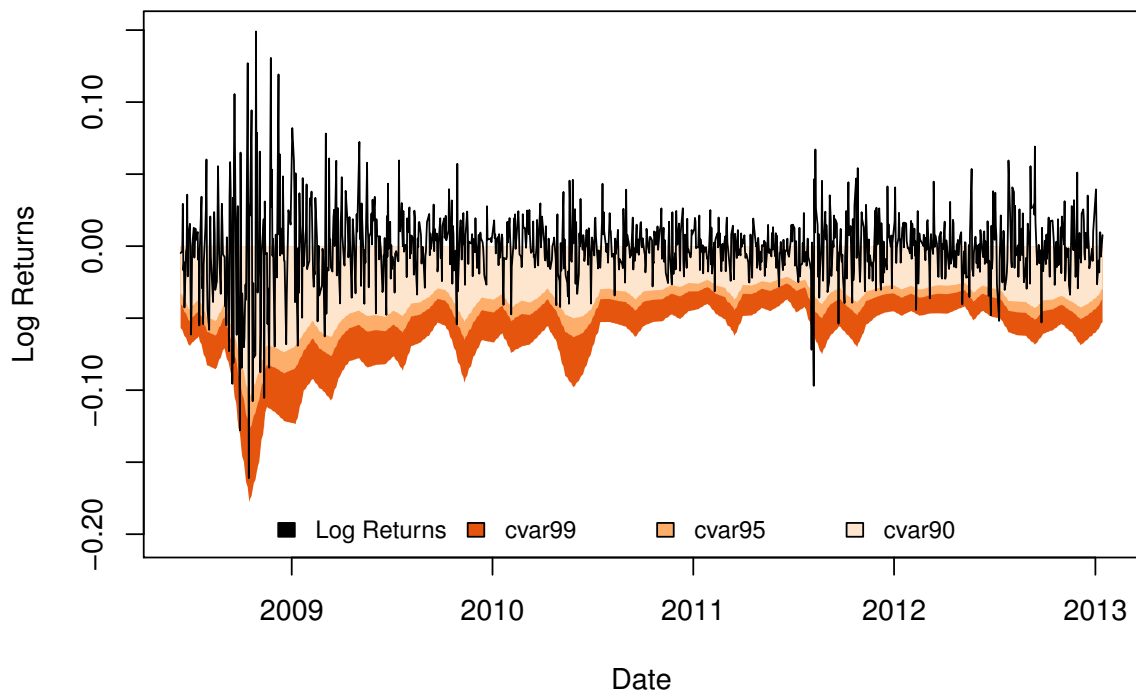


Figura 7.2: Previsões de VaR e CVaR (suavizadas)

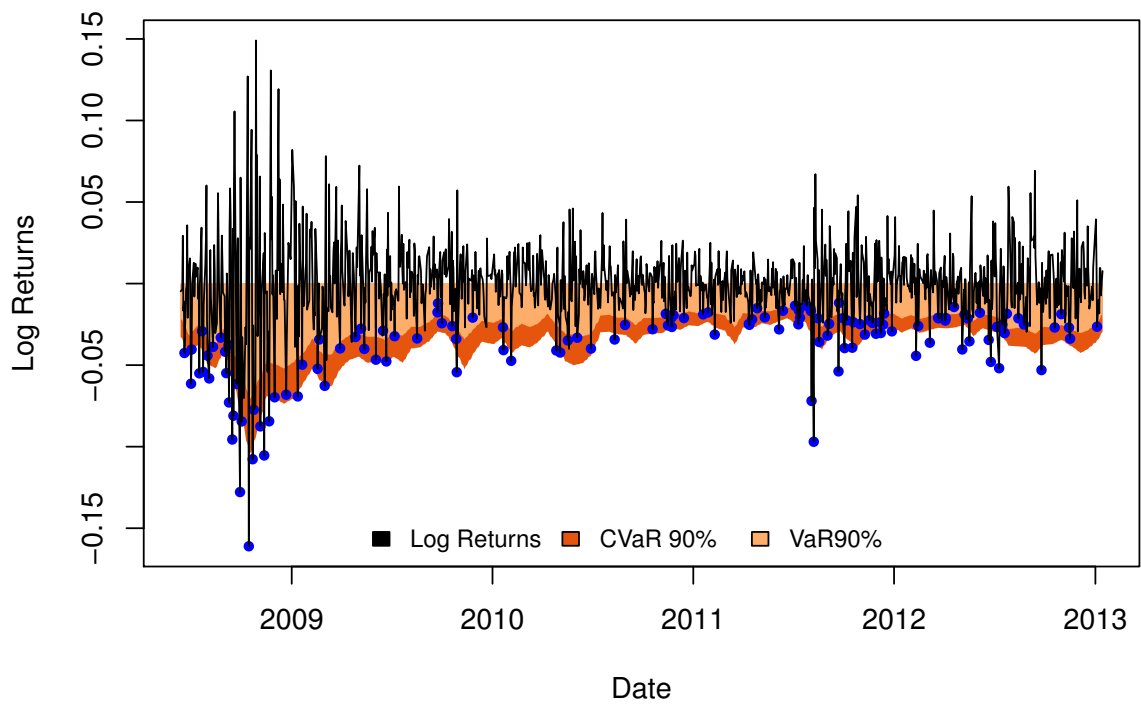
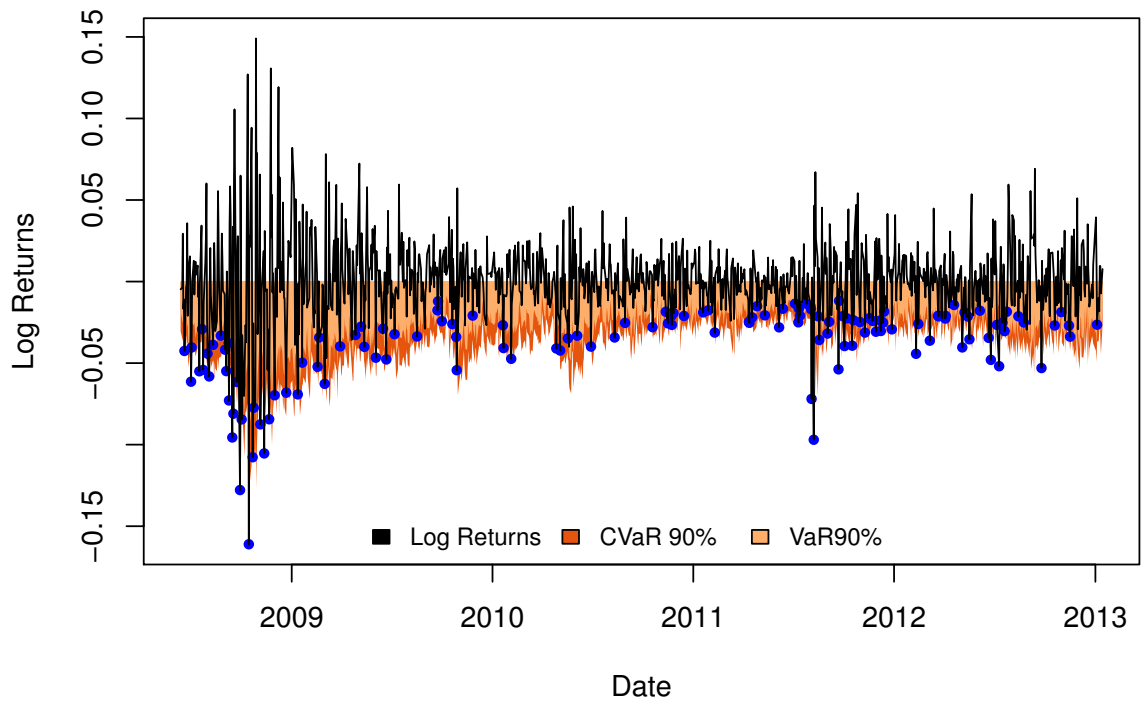


Figura 7.3: Previsões de VaR e CVaR, com violações de VaR, para $\beta = 90\%$

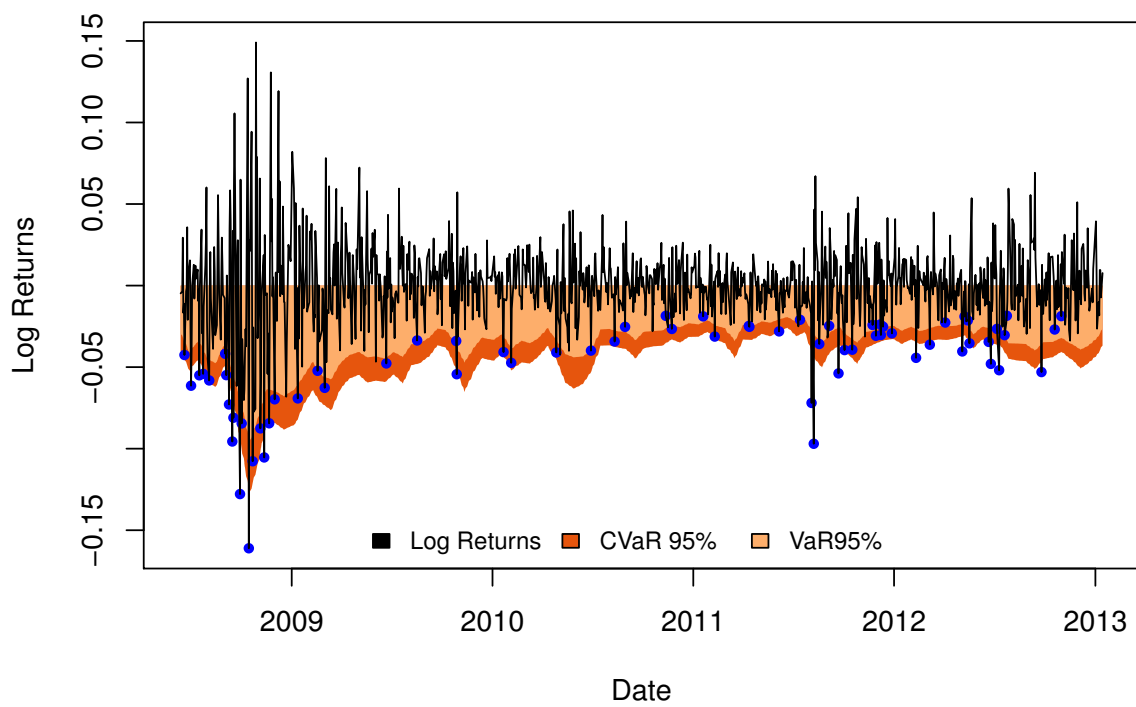
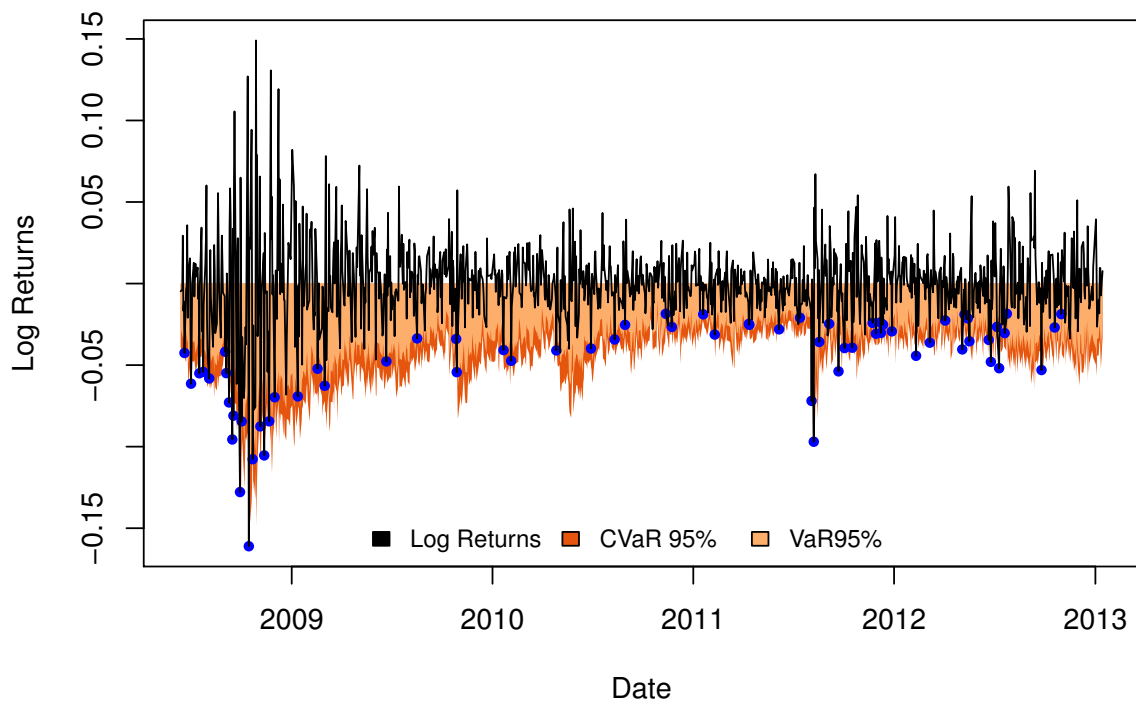


Figura 7.4: Previsões de VaR e CVaR, com violações de VaR, para $\beta = 95\%$

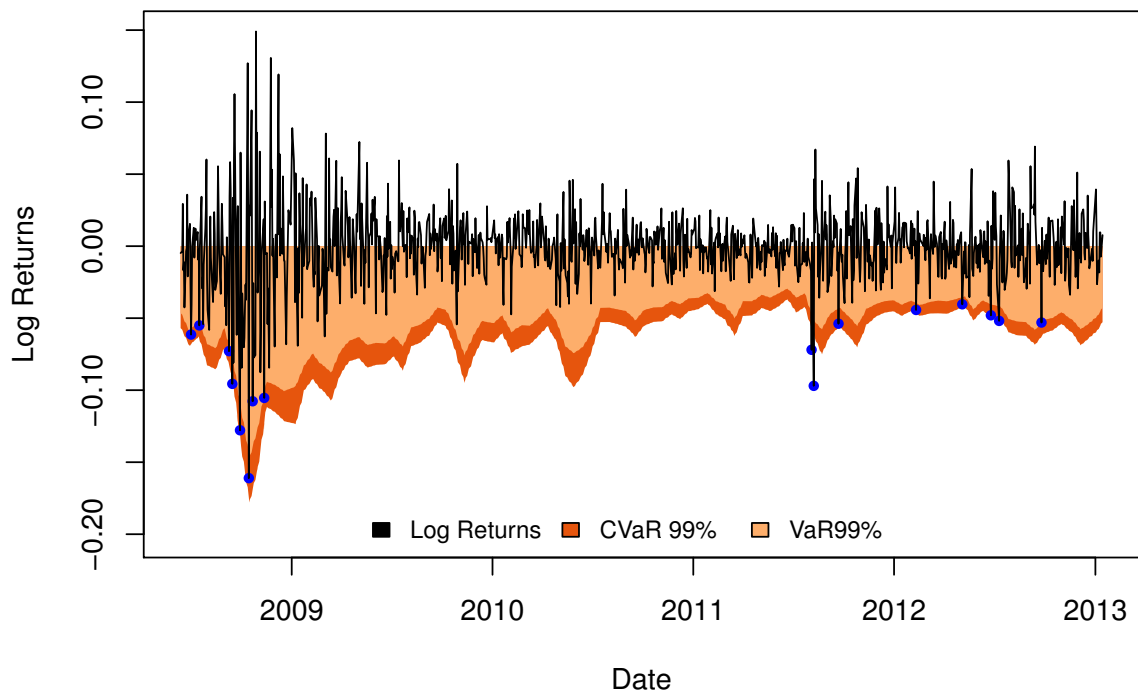
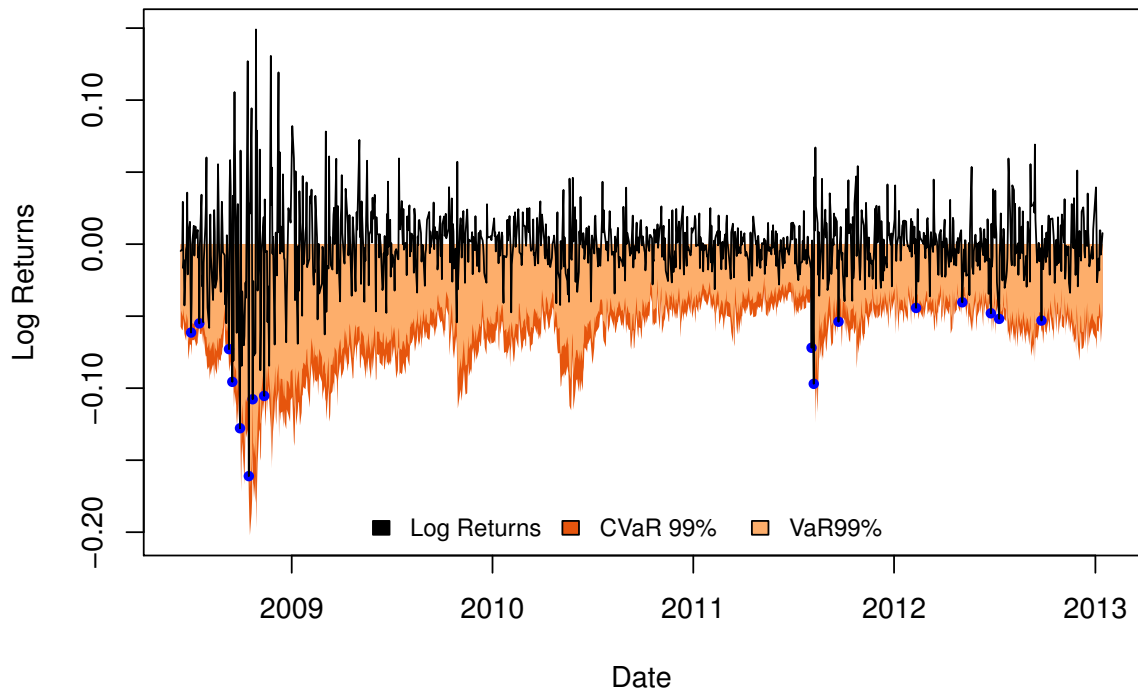


Figura 7.5: Previsões de VaR e CVaR, com violações de VaR, para $\beta = 99\%$

Parte III

Conclusão e Complementos

Capítulo 8

Conclusão

Este trabalho teve como objetivo o desenvolvimento de algoritmos para a avaliação de risco de um portfólio e também para a seleção de portfólios de risco mínimo.

Todo o desenvolvimento teve como base a medida de risco CVaR e os modelos VAR e GO-GARCH para os retornos financeiros. A medida CVaR, como explicado na Seção 1.1, foi escolhida por ser apropriada para mensurar as perdas financeiras (pois é uma medida coerente de risco) e por ser uma função convexa, o que permite construir algoritmos de otimização com maior facilidade. Quanto aos modelos VAR e GO-GARCH, estes foram escolhidos por fornecer uma representação de retornos mais realista, ou seja, retornos não gaussianos, multivariados e explicados principalmente pela sua volatilidade.

Ao longo do trabalho foram implementados algoritmos para as seguintes operações:

- (i) Estimação das medidas de risco VaR e CVaR a partir de uma amostra de retornos,
- (ii) Seleção de um portfólio de CVaR mínimo em função do retorno esperado mínimo e de uma amostra de retornos,
- (iii) Ajuste de modelos VAR e GO-GARCH a séries históricas de retorno,
- (iv) Avaliação diária do risco de um portfólio.

Estes algoritmos foram desenvolvidos em R com o auxílio da biblioteca de otimização *MOSEK* e dos pacotes *vars*, *rugarch*, *R-to-MOSEK* e *fastICA*. Também foram utilizados os testes de independência de violações implementados no trabalho [18], que foram gentilmente cedidos pela autora do mesmo, para testar os valores de VaR calculados no Capítulo 7. Foram feitas implementações próprias dos estimadores ICA e MM2011 para modelos GO-GARCH e dos estimadores das medidas de risco utilizadas.

Quanto ao uso dos modelos VAR e GO-GARCH, foi possível verificar que estes são adequados para descrever os retornos financeiros. De forma geral, os ajustes realizados foram capazes de capturar a dinâmica dos retornos, o que pôde ser verificado através da ausência de autocorrelações de primeiro e segundo momento no vetor dos resíduos, por estes apresentarem coordenadas não correlacionadas entre si e também pela distribuição destas ser compatível com a distribuição escolhida para modelar os choques dos fatores que controlam a dinâmica do modelo GO-GARCH. Além disso, foi possível verificar que o ajuste dos modelos a dados históricos não apresentou problemas de convergência. Entretanto, o ajuste do modelo por análise de componentes independentes (ICA) se mostrou sensível a parâmetros iniciais.

Em relação aos portfólios de risco mínimo, foi possível confirmar que o CVaR destes aumenta à medida que o rendimento esperado mínimo aumenta e também à medida que o nível β de VaR aumenta.

Além disso, o risco dos portfólios selecionados era menor que o risco individual de qualquer um dos ativos disponíveis, o que demonstra o benefício da diversificação de um investimento.

De acordo com o esperado, o estimador de CVaR é bastante sensível ao tamanho da amostra utilizada em seu cálculo. Testes de simulação indicaram que o erro relativo médio, ao variar o tamanho da amostra entre dez mil e cem mil, fica entre 0.5% e 4.0%. Assim, o estimador pode ser utilizado sem grandes problemas nos casos onde é possível controlar o tamanho das amostras utilizadas. Em outros casos, é possível que os erros de estimação sejam proibitivos.

O estimador de CVaR e os modelos VAR e GO-GARCH foram utilizados em conjunto para gerenciar o risco de um portfólio, calculando estimativas de VaR e CVaR sobre uma amostra de cenários gerados para o dia seguinte. A análise das violações de VaR indicou que esta estratégia é adequada, uma vez que as violações ocorreram com a frequência esperada e também eram independentes entre si. No teste destas violações foram utilizados o teste de Kupiec e também os testes de independência utilizados em [18]. Em um segundo exemplo, foi utilizado o modelo O-GARCH ao invés do modelo GO-GARCH. Neste caso, os resultados obtidos não foram satisfatórios, pois as violações de nível de VaR não ocorreram com a frequência esperada.

Havendo disponibilidade de rotinas para a resolução de problemas de programação linear e para o ajuste de modelos GARCH univariados, todos os algoritmos têm implementação simples. Estes são sequenciais, não havendo uso de recursão, paralelismo ou estruturas de dados complexas. Desta forma, dados os bons resultados obtidos no gerenciamento do risco de um portfólio com o estimador de CVaR e com os modelos VAR e GO-GARCH, é possível que esta seja uma alternativa mais apropriada que a adoção de modelos baseados na distribuição normal. Afinal, os retornos representados pelos modelos VAR e GO-GARCH são mais próximos aos encontrados na prática.

8.1 Trabalhos Futuros

Este trabalho ainda pode ser estendido de diversas formas. O modelo adotado para os retornos financeiros, por exemplo, utiliza fatores que seguem o modelo GARCH(1,1). Apesar dos bons resultados obtidos, seria interessante testar outros modelos para estes fatores, já que isto é permitido pelo modelo GO-GARCH.

O algoritmo para a seleção de portfólios também pode sofrer modificações. Atualmente, o seu usuário fornece apenas o retorno esperado mínimo. Entretanto, permitir uma maior influência do usuário na escolha do portfólio pode ser bastante útil. Este poderia fornecer requisitos mínimos de liquidez ou uma quantidade máxima para cada ativo. Outra possível modificação é a incorporação de outros ativos financeiros, como opções ou *bonds*.

Em relação à avaliação do risco de um portfólio, seria interessante comparar as previsões de VaR e CVaR calculadas com as obtidas por outras metodologias, como a RiskMetrics. Além disso, os testes realizados até o momento foram concentrados na medida VaR, devido à disponibilidade dos testes para tal medida. Assim, ainda é preciso realizar testes sobre as estimativas de CVaR calculadas pelo algoritmo. Os testes para CVaR estudados em [18] podem ser uma boa opção.

Finalmente, também seria interessante tratar a falta de robustez da medida CVaR, como indicado em [12].

Apêndice A

Análise de Componentes Independentes

Neste apêndice é feita uma breve apresentação da análise de componentes independentes (ou ICA) de um vetor aleatório. Para um tratamento completo deste tópico, recomenda-se a consulta a [11, 29, 32, 31, 30].

Esta apresentação tem três objetivos principais. Em primeiro lugar, definir formalmente a ICA, diferenciando-a da análise de componentes principais (PCA), que é uma técnica bastante semelhante. Em segundo lugar, apresentar os conceitos teóricos que justificam formalmente o método de estimação das componentes independentes. Finalmente, apresentar o algoritmo FastICA [31], que foi utilizado na estimação de modelos GO-GARCH.

A.1 Definição

Tanto a PCA quanto a ICA são baseadas na hipótese de que as coordenadas de um vetor aleatório x podem ser calculadas a partir de combinações lineares de um outro vetor aleatório y que, em geral, tem menor dimensão e algumas propriedades especiais. Desta forma, vale a expressão

$$x = Fy. \quad (\text{A.1.1})$$

Os nomes *componentes principais* ou *componentes independentes* são atribuídos às coordenadas de y , dependendo do tipo de análise realizado.

A Definição 13, adaptada de [11], define formalmente a PCA de um vetor aleatório x .

Definição 13. *Análise de Componentes Principais (PCA)*

Seja x um vetor aleatório m -dimensional com média zero e matriz de covariância Σ_x , associado a um vetor aleatório y de componentes principais, n -dimensional. A PCA de x consiste no par de matrizes $\{F, \Delta\}$ satisfazendo:

- (i) F é uma matriz $m \times n$ de posto completo satisfazendo $F'F = I_n$,
- (ii) Δ é uma matriz diagonal $n \times n$, real e positiva,
- (iii) $\Sigma_x = F\Delta^2F'$.

Pela Equação (A.1.1) e pela restrição $F'F = I_n$, as componentes principais (ou independentes) podem ser obtidas através da expressão

$$y = F'x \quad (\text{A.1.2})$$

Considerando a Equação (A.1.1) e a Definição 13, a PCA do vetor x pode ser vista como sua decomposição em variáveis não correlacionadas, já que a matriz Σ_y (matriz de covariâncias de y) é diagonal, como demonstra a Equação (A.1.3).

$$\Sigma_y = \mathbb{E}[yy'] = \mathbb{E}[F'x(F'x)'] = F'\Sigma_x F = (F'F)\Delta^2(F'F) = \Delta^2 \quad (\text{A.1.3})$$

Finalmente, pela Definição 14, a ICA pode ser vista como uma extensão da PCA. Sendo que a diferença entre as duas análises reside na troca do requisito de não existência de correlações entre as componentes de y pelo de independência entre as mesmas.

Vale lembrar que a independência entre variáveis aleatórias implica na ausência de correlação entre as mesmas. O contrário, porém, não é verdade. Assim, fica claro que a ICA e a PCA são semelhantes, porém conceitualmente diferentes.

Definição 14. *Análise de Componentes Independentes (ICA)*

Seja x um vetor aleatório m -dimensional com média zero e matriz de covariância Σ_x , associado a um vetor y de componentes independentes, n -dimensional. A ICA de x consiste no par de matrizes $\{F, \Delta\}$ satisfazendo:

- (i) F é uma matriz $m \times n$ de posto completo satisfazendo $F'F = I_n$,
- (ii) Δ é uma matriz diagonal $n \times n$, real e positiva,
- (iii) $\Sigma_x = F\Delta^2F'$,
- (iv) y tem componentes o mais independentes possível, de acordo com uma função de contraste a ser definida.

Assim como para as componentes principais, a Definição 14 e a Equação (A.1.1) permitem interpretar a ICA como uma decomposição do vetor x em variáveis aleatórias. Sendo estas, entretanto, independentes.

A.2 Estimação

De acordo com a Definição 14, a estimação das componentes independentes se resume a determinar o par de matrizes $\{F, \Delta\}$ que maximiza uma *função de contraste* $I(y)$, cujo propósito é medir o grau de independência existente entre as componentes de y .

O problema de otimização a ser resolvido pode ser simplificado utilizando duas hipóteses adotadas no algoritmo FastICA.

Na primeira hipótese, considera-se que as componentes independentes têm variância unitária, ou seja, que $\Sigma_y = I_n$.

Como mostrado em [30], já que $x = Fy$ e tanto F quanto y são não observáveis, é impossível determinar a variância das componentes independentes, pois alterações em y podem ser compensadas por alterações em F .

Desta forma, não há perda de generalidade ao considerar que as componentes independentes têm variância unitária. Na verdade, a estimação das mesmas se torna mais robusta.

A segunda hipótese, que aparentemente resulta em perda de generalidade, considera que o vetor x é *branco* (ver Definição 15).

Definição 15. *Vetor Aleatório Branco*

Um vetor aleatório x de média zero, m -dimensional, é dito *branco* quando suas coordenadas são não correlacionadas e têm variância unitária, ou seja, $\text{Cov}(x) = I_m$.

Felizmente, como mostrado na Seção A.3, a operação de branqueamento permite obter uma representação de x que corresponde a um vetor aleatório branco. Desta forma, a estimação de componentes independentes não fica restrita apenas a vetores aleatórios que sejam originalmente brancos.

Com as duas hipóteses adicionadas, a restrição $\Sigma_x = F\Delta^2F'$ equivale a $FF' = I_m$. Desta forma, duas restrições são obtidas para o problema de estimação: $FF' = I_m$ e $F'F = I_n$.

Assim, estimar as componentes independentes de x se resume a resolver o problema de otimização a seguir.

$$\left\{ \begin{array}{l} \max_F I(F'x) \\ \text{Sujeito a:} \\ F'F = I_n \\ FF' = I_m \end{array} \right. \quad (\text{A.2.1})$$

Daqui em diante, considera-se o caso em que as dimensões m e n são iguais. Sob esta condição, as restrições sobre F implicam na sua ortogonalidade, já que $FF' = F'F = I_m$.

A ortogonalidade de F , como mostrado em [30], pode ser explorada para melhorar o processo de estimação. Quando F é ortogonal, apenas $m(m-1)/2$ parâmetros precisam ser estimados contra m^2 parâmetros para uma matriz qualquer. Desta forma, o branqueamento do vetor a ser analisado se apresenta como uma técnica obrigatória, principalmente para problemas de maior dimensão.

A.3 Centralização e Branqueamento

Seja z um vetor aleatório com média μ_z não nula e com matriz de covariância Σ_z . Através das operações de centralização e branqueamento, como mostrado em [30], é possível obter um vetor branco z_b , que pode ser usado para estimar as componentes independentes de z através da solução do Problema (A.2.1).

A operação de centralização consiste em simplesmente subtrair de z sua média, para obter um vetor z_0 de média nula, ou seja,

$$z_0 = z - \mu_z. \quad (\text{A.3.1})$$

Após a centralização, o vetor branco z_b é obtido multiplicando z_0 pela matriz $V = PL^{-1/2}P'$, onde PLP' é a decomposição espectral de Σ_z .

$$z_b = Vz_0. \quad (\text{A.3.2})$$

Como mostram as Equações (A.3.3) e (A.3.4), z_b tem as propriedades de um vetor branco, como desejado.

$$\mu_{z_b} = \mathbb{E}[z_b] = V\mathbb{E}[z_0] = 0 \quad (\text{A.3.3})$$

$$\text{Cov}(z_b) = \mathbb{E}[z_b z_b'] = V\mathbb{E}[z_0 z_0']V' = (PL^{-1/2}P')(PLP')(PL^{-1/2}P') = I_m \quad (\text{A.3.4})$$

Finalmente, seja F_{z_b} a matriz ortogonal, estimada pela resolução da Problema (A.2.1), para o vetor z_b . A partir das Equações (A.3.1) e (A.3.2), a decomposição de z em componentes independentes pode ser escrita como

$$z = (V^{-1}F_{z_b})y + \mu_z. \quad (\text{A.3.5})$$

A.4 Função de Contraste

A fim de completar a definição do Problema (A.2.1), é necessário definir formalmente a função de contraste $I(x)$, responsável por medir o grau de independência entre as coordenadas do vetor x .

As componentes de um vetor aleatório são independentes entre si quando a densidade de probabilidade conjunta do vetor é igual ao produto de suas densidades marginais. Assim, sendo f_x a densidade de probabilidade conjunta de x e g_x o produto de suas densidades marginais (ver Equação (A.4.1)), é natural utilizar a diferença entre f_x e g_x como métrica de independência.

$$g_x(u) = \prod_{i=1}^m f_{x_i}(u_i) \quad (\text{A.4.1})$$

Desta forma, a função de contraste $I(x)$ é definida como a *Divergência de Kullback-Leibler* (KLD) entre f_x e g_x , que fornece uma medida não simétrica da diferença entre as duas densidades de probabilidade (ver a Definição 16 e [11, 26, 34] para maiores informações), ou seja,

$$I(x) = \delta(f_x, g_x). \quad (\text{A.4.2})$$

Definição 16. Divergência de Kullback-Leibler (KLD)

Sejam f_x e g_x densidades de probabilidade. A KLD entre f_x e g_x , $\delta(f_x, g_x)$, é dada por

$$\delta(f_x, g_x) = \int_{\mathbb{R}^m} f_x(u) \log \frac{f_x(u)}{g_x(u)} du. \quad (\text{A.4.3})$$

Como a KLD de qualquer par de densidades de probabilidade é não negativa (veja [11]), vale que $I(x) = \delta(f_x, g_x) \geq 0$ com igualdade apenas quando f_x e g_x diferem em conjuntos de medida nula.

Além da formulação baseada na KLD, a função de contraste $I(x)$ possui formulações derivadas de conceitos da Teoria da Informação.

Utilizando a *Entropia Diferencial* (veja a Definição 17), é possível escrever $I(x)$ como

$$I(x) = \sum_{i=1}^m H(x_i) - H(x). \quad (\text{A.4.4})$$

Definição 17. Entropia Diferencial

Seja x um vetor aleatório com densidade conjunta de probabilidade f_x . A Entropia Diferencial de x é dada por

$$H(x) = - \int_{\mathbb{R}^m} f_x(u) \log f_x(u) du. \quad (\text{A.4.5})$$

A entropia diferencial é uma medida da quantidade de informação que pode ser obtida a partir de observações de uma variável aleatória (ver [30]). Quanto maior sua entropia (ou aleatoriedade), menor a quantidade de informação que pode ser obtida sobre o valor real de suas realizações.

Desta maneira, $I(x)$, chamada de *Informação Mútua* na Teoria da Informação, mede a quantidade de informação redundante que as componentes x_i possuem. Assim, quando a informação mútua é nula, não existe informação redundante nas componentes de x , sendo todas necessárias para atingir o nível de informação $H(x)$.

Conforme esperado, as formulações para $I(x)$ usando a entropia diferencial e a KLD são equivalentes, como mostra a Equação (A.4.6).

$$\begin{aligned}
I(x) &= \delta(f_x, g_x) \\
&= \int_{\mathbb{R}^m} f_x(u) \log \frac{f_x(u)}{g_x(u)} du \\
&= \int_{\mathbb{R}^m} f_x(u) \log f_x(u) du - \int_{\mathbb{R}^m} f_x(u) \log g_x(u) du \\
&= -H(x) - \int_{\mathbb{R}^m} f_x(u) \log \prod_{i=1}^m f_{x_i}(u_i) du \\
&= -H(x) - \int_{\mathbb{R}^m} \sum_{i=1}^m f_x(u) \log f_{x_i}(u_i) du \\
&= -H(x) - \sum_{i=1}^m \int_{\mathbb{R}^m} f_x(u) \log f_{x_i}(u_i) du \\
&= -H(x) - \sum_{i=1}^m \int_{\mathbb{R}} f_{x_i}(u) \log f_{x_i}(u_i) du_i \\
&= -H(x) + \sum_{i=1}^m H(x_i)
\end{aligned} \tag{A.4.6}$$

Um outro conceito de Teoria da Informação, o de *Entropia Negativa* (ver Definição 18), ainda é necessário para a definição final da informação mútua $I(x)$.

A distribuição Normal multivariada tem a maior entropia entre todas as distribuições de mesma variância (ver [13]). Desta forma, sua entropia é utilizada para definir uma medida não nula da quantidade de informação de um vetor aleatório, que tem importância na prática. Esta medida é a entropia negativa.

Definição 18. *Entropia Negativa*

Seja x um vetor aleatório e x_g um vetor normal multivariado, ambos com matriz de covariância Σ_x . A Entropia Negativa de x é dada por

$$J(x) = H(x_g) - H(x). \tag{A.4.7}$$

A exemplo da entropia diferencial, a entropia negativa também pode ser escrita a partir da KLD, como mostra a Equação (A.4.8). No caso da entropia negativa, entretanto, g_x é substituída por ϕ_x , densidade de uma Normal multivariada com mesma matriz de covariância do vetor x .

$$\begin{aligned}
J(x) &= H(x_g) - H(x) \\
&= H(x_g) - \int_{\mathbb{R}^m} f_x(u) \log \phi_x(u) du + \int_{\mathbb{R}^m} f_x(u) \log \phi_x(u) du - H(x) \\
&= H(x_g) - \int_{\mathbb{R}^m} f_x(u) \log \phi_x(u) du + \int_{\mathbb{R}^m} f_x(u) \log \frac{f_x(u)}{\phi_x(u)} du \\
&= 0 + \int_{\mathbb{R}^m} f_x(u) \log \frac{f_x(u)}{\phi_x(u)} du \\
&= \delta(f_x, \phi_x)
\end{aligned} \tag{A.4.8}$$

Sendo uma KLD, a entropia negativa herda suas propriedades. Assim, $J(x) \geq 0$. Mais uma vez, valendo a igualdade apenas quando f_x e ϕ_x diferem apenas em conjuntos de medida nula.

Finalmente, tendo sido apresentados os conceitos necessários, é possível definir a informação mútua para uma transformação linear $y = Wx$, caso de interesse para resolução do Problema (A.2.1). Em primeiro lugar, considere a entropia de y . Como mostrado em [13],

$$H(y) = H(x) + \log |\det W| = H(x), \quad (\text{A.4.9})$$

já que W é uma matriz ortogonal no contexto do Problema (A.2.1).

Em seguida, considere as variáveis aleatórias z , de variância unitária, e a variável $z_g \sim N(0, 1)$. Aplicando a Definição 17, chega-se a entropia de z_g , dada pela constante $H(z_g) = 1/2(1 + \log 2\pi)$. Assim, pela Definição 18, é fácil ver que as entropias diferencial e negativa de z diferem apenas pela constante $H(z_g)$, ou seja,

$$H(z) = H(z_g) - J(z) = 1/2(1 + \log 2\pi) - J(z). \quad (\text{A.4.10})$$

Logo, pelas Equações (A.4.9) e (A.4.10), a informação mútua $I(y)$ pode ser escrita em função das entropias negativas $J(y_i)$ e de uma constante K_x , cujo valor é dado pelas propriedades do vetor x , ou seja,

$$\begin{aligned} I(y) &= \sum_{i=1}^m H(y_i) - H(y) \\ &= \sum_{i=1}^m \{1/2(1 + \log 2\pi) - J(y_i)\} - H(x) \\ &= \left(\sum_{i=1}^m 1/2(1 + \log 2\pi) - H(x) \right) - \sum_{i=1}^m J(y_i) \\ &= K_x - \sum_{i=1}^m J(y_i) \end{aligned} \quad (\text{A.4.11})$$

Com a derivação da Equação (A.4.11), a justificativa teórica para os algoritmos de otimização é concluída, já que esta relaciona a função objetivo $I(y)$ às variáveis de otimização $J(y_i)$.

Assim, considere novamente o Problema (A.2.1). Seu objetivo final é a estimação das componentes independentes e não a obtenção do valor de $I(y)$ propriamente dito. Logo, não há perdas em trocar a minimização da informação mútua $I(y)$ pela maximização da soma das entropias negativas $J(y_i)$. A equivalência entre estas duas alternativas fica evidente pela Equação (A.4.11).

Como mostrado em [32, 31, 30] é possível aproximar as entropias negativas $J(y_i)$ utilizando funções não lineares. Estas aproximações, que seguem a forma

$$J(y) \propto (\mathbb{E}[G(y)] - \mathbb{E}[G(\nu)])^2, \quad (\text{A.4.12})$$

têm implementação computacional simples e permitem avaliar $I(y)$ de forma bastante eficiente.

Entretanto, como $J(y)$ tem valor máximo quando a esperança $\mathbb{E}[G(y)]$ é máxima, a última pode ser usada como a função objetivo do problema de estimação, simplificando-o ainda mais.

Desta forma, o problema de otimização original pode ser reescrito como o Problema (A.4.13).

$$\left\{ \begin{array}{l} \max_F \mathbb{E}[G(F'x)] \\ \text{Sujeito a:} \\ F'F = I_n \\ FF' = I_m \end{array} \right. \quad (\text{A.4.13})$$

Quanto à função não linear G , diversos exemplos são apresentados em [30], onde é descrito com detalhes o processo para obtenção das mesmas. Como exemplos, as funções G_1 e G_2 podem ser usadas no lugar de G (ver Equações (A.4.14) e (A.4.15)).

$$G_1(y) = \frac{1}{a_1} \log \cosh(a_1 y) \quad (\text{A.4.14})$$

$$G_2(y) = -\frac{1}{a_2} \exp\left(-\frac{a_2 y^2}{2}\right) \quad (\text{A.4.15})$$

A.5 Algoritmo FastICA

A resolução do Problema (A.4.13) é feita com base em uma modificação do Método de Newton, que dá origem a um algoritmo eficiente e simples para a estimação de componentes independentes.

Seguindo [29, 31, 30], a apresentação do algoritmo FastICA é realizada em duas partes. Inicialmente, é apresentado o algoritmo para estimação de uma única componente independente. Em seguida é apresentado o algoritmo completo, obtido com uma simples modificação do algoritmo inicial.

Aqui, a apresentação se limita à descrição dos algoritmos utilizados na ICA, para a prova de convergência dos mesmos, veja [29, 31, 30].

A.5.1 Estimação de uma componente independente

Ao considerar apenas uma componente, o Problema (A.4.13) se reduz à determinação do vetor unitário w , para o qual a esperança $\mathbb{E}[G(w'x)]$ é máxima, como indica o Problema (A.5.1).

$$\begin{cases} \max_w \mathbb{E}[G(w'x)] \\ \text{Sujeito a:} \\ w'w = 1 \end{cases} \quad (\text{A.5.1})$$

Em geral, como a distribuição de $w'x$ é desconhecida, a esperança $\mathbb{E}[G(w'x)]$ é avaliada como a média amostral dos valores $G(w'x_i)$, para uma amostra $A = \{x_1, \dots, x_N\}$.

A fim de resolver o Problema (A.5.1) utiliza-se o Método de Newton e a técnica de Multiplicadores de Lagrange (ver [5]). Assim, a função Lagrangiana \mathcal{L} é dada por

$$\mathcal{L}(w) = \mathbb{E}[G(w'x)] - \beta(1 - w'w), \quad (\text{A.5.2})$$

e suas derivadas pelas Equações (A.5.3) e (A.5.4).

$$\frac{\partial \mathcal{L}}{\partial w} = \mathbb{E}[xG'(w'x)] + \beta w \quad (\text{A.5.3})$$

$$\frac{\partial^2 \mathcal{L}}{\partial w^2} = \mathbb{E}[xx'G''(w'x)] + \beta I_m \quad (\text{A.5.4})$$

As iterações do Método de Newton são dadas por

$$w_{k+1} = w_k - \left[\frac{\partial^2 \mathcal{L}}{\partial w^2}(w_k) \right]^{-1} \frac{\partial \mathcal{L}}{\partial w}(w_k), \quad (\text{A.5.5})$$

o que exige a inversão da hessiana de \mathcal{L} .

Entretanto, ao invés de inverter a hessiana original de \mathcal{L} , é adotada uma aproximação que a torna uma matriz diagonal, como mostra a Equação (A.5.6). Segundo [30], tal aproximação é possível por x ser um vetor branco (veja a Definição 15).

$$\mathbb{E}[xx'G''(w'x)] \approx \mathbb{E}[xx']\mathbb{E}[G''(w'x)] = I_m\mathbb{E}[G''(w'x)] \quad (\text{A.5.6})$$

Assim, a hessiana aproximada de \mathcal{L} é dada por

$$\frac{\partial^2 \mathcal{L}}{\partial w^2} \approx \mathbb{E}[G''(w'x)]I_m + \beta I_m \quad (\text{A.5.7})$$

e as iterações aproximadas do Método de Newton por

$$w_{k+1} = w_k - \frac{\mathbb{E}[xG'(w'_k x)] + \beta w}{\mathbb{E}[G''(w'_k x)] + \beta}. \quad (\text{A.5.8})$$

Seja $\alpha = -\beta - \mathbb{E}[G''(w'_k x)]$. Multiplicando os dois lados da Equação (A.5.8) por α , é possível simplificá-la. Assim, obtém-se

$$\begin{aligned} w_{k+1}\alpha &= \left(w_k - \frac{\mathbb{E}[xG'(w'_k x)] + \beta w}{\mathbb{E}[G''(w'_k x)] + \beta} \right) \alpha \\ &= w_k\alpha + \mathbb{E}[xG'(w'_k x)] + \beta w_k \\ &= -\beta w_k - \mathbb{E}[G''(w'_k x)] w_k + \mathbb{E}[xG'(w'_k x)] + \beta w_k \\ &= \mathbb{E}[xG'(w'_k x)] - \mathbb{E}[G''(w'_k x)] w_k, \end{aligned} \quad (\text{A.5.9})$$

que leva à Equação (A.5.10) para a iteração do Método de Newton aproximado.

$$w_{k+1} = 1/\alpha (\mathbb{E}[xG'(w'_k x)] - \mathbb{E}[G''(w'_k x)] w_k) \quad (\text{A.5.10})$$

Para garantir o respeito às restrições do Problema (A.5.1), a cada iteração do algoritmo as soluções encontradas pela Equação (A.5.10) são projetadas sobre o espaço das restrições.

No caso da estimação de uma única componente independente, a projeção no espaço das restrições consiste em normalizar o vetor w . Assim, dada a normalização obrigatória, o termo $1/\alpha$ pode ser descartado da Equação (A.5.10). Desta forma, a cada iteração, o algoritmo realiza as operações

$$\begin{cases} w_{k+1} = \mathbb{E}[xG'(w'_k x)] - \mathbb{E}[G''(w'_k x)] w_k, \\ w_{k+1} = \frac{w_{k+1}}{\|w_{k+1}\|}. \end{cases} \quad (\text{A.5.11})$$

O Algoritmo 5 apresenta o pseudocódigo para implementação da estimação de uma única componente independente.

Vale destacar o objetivo das funções *branqueamento* e *rand*. A primeira função, *branqueamento*, aplica à sua entrada as operações necessárias para transformá-la em um vetor branco, como feito na Seção A.3. A segunda, *rand*, cria um vetor aleatório que é usado para iniciar o algoritmo.

A.5.2 FastICA

O algoritmo para a estimação simultânea de n componentes independentes é obtido com modificações simples no Algoritmo 5. A estrutura geral deste, entretanto, permanece inalterada.

A primeira modificação consiste em calcular atualizações para todos os vetores w_1, \dots, w_n a cada iteração.

Algorithm 5 IC - Estimaco de uma nica componente independente

```

1: procedure IC( $x$ )
2:    $x_b := \mathbf{branqueamento}(x)$ 
3:    $w_k := \mathbf{rand}()$ 
4:    $w_k := w_k / \|w_k\|$ 
5:    $w_{k+1} := 0$ 
6:   while  $|w'_k w_{k+1}| \neq 1$  do
7:      $w_k := w_{k+1}$ 
8:      $w_{k+1} := \mathbb{E}[x_b G'(w'_k x_b)] - \mathbb{E}[G''(w'_k x_b)] w_k$ 
9:      $w_{k+1} := w_{k+1} / \|w_{k+1}\|$ 
10:  end while
11:  return  $w_k$ 
12: end procedure

```

Como no Algoritmo 5, as atualizaes so calculadas pela Equaco (A.5.11), sendo estas independentes entre si. Observe que, sendo estas independentes entre si,  possvel calcul-las em paralelo quando houver necessidade de um algoritmo computacionalmente mais eficiente.

A segunda modificaco, correspondente  projeo de solues sobre o espao de restries, consiste em ortogonalizar a matriz $W = [w_1 \ \dots \ w_n]$ aps a atualizao dos vetores w_1, \dots, w_n .

Esta ortogonalizao pode ser facilmente justificada considerando a covarincia das componentes independentes y . Como mostrado na Seo A.2, $\text{Cov}(y) = I_n$ por hiptese. Logo, as colunas de W , que so os vetores w_1, \dots, w_n , devem ser ortogonais entre si, pois

$$I_n = \text{Cov}(y) = \mathbb{E}[(Wx)(Wx)'] = W\mathbb{E}[xx']W' = WW'. \quad (\text{A.5.12})$$

Quanto ao procedimento para ortogonalizao, em [30]  recomendada a ortogonalizao simultnea, descrita a seguir.

Seja PLP' a decomposio espectral de W , com P ortogonal e L diagonal. Para $V = (W'W)^{-1/2}$, a matriz $W_o = VW$  ortogonal, pois

$$V = (W'W)^{-1/2} = ((PLP')(PLP'))^{-1/2} = (PL^2P')^{-1/2} = PL^{-1}P', \quad (\text{A.5.13})$$

e, conseqentemente,

$$W'_o W_o = (VW)'(VW) = (PLP')(PL^{-1}P')(PL^{-1}P')(PLP') = I_n. \quad (\text{A.5.14})$$

Desta forma, est comprovada a ortogonalidade de W_o .

Finalmente, incorporando estas duas modificaces ao pseudocdigo do Algoritmo 5, obtm-se o algoritmo FastICA, cujo pseudocdigo pode ser visto no Algoritmo 6.

Em [31], diversas propriedades do algoritmo FastICA so apresentadas. Aqui, so destacadas as seguintes:

- (i) A convergncia do algoritmo FastICA  cbica (ou pelo menos quadrtica).
- (ii) Por no requerer a definio de um tamanho de passo, o algoritmo tem fcil utilizao.
- (iii) A escolha das funes no lineares G no  determinada por hipteses feitas sobre a distribuo do dado analisado.

- (iv) O algoritmo é passível de configuração, sendo possível melhorar seu desempenho através de escolhas apropriadas da não linearidade G .

Algorithm 6 FastICA - Estimação de n componentes independentes

```

1: procedure FASTICA( $x, n$ )
2:    $x_b := \text{branqueamento}(x)$ 
3:   for  $i := 1, \dots, n$  do
4:      $w_{ik} := \text{rand}()$ 
5:      $w_{ik} := w_{ik}/\|w_{ik}\|$ 
6:      $w_{ik+1} := 0$ 
7:   end for
8:   repeat
9:     for  $i := 1, \dots, n$  do
10:       $w_{ik+1} := \mathbb{E}[x_b G'(w'_{ik} x_b)] - \mathbb{E}[G''(w'_{ik} x_b)] w_{ik}$ 
11:       $w_{ik+1} := w_{ik+1}/\|w_{ik+1}\|$ 
12:    end for
13:     $W := (w_{1k}, \dots, w_{mk})'$ 
14:     $V := (WW')^{-1/2}$ 
15:    for  $i := 1, \dots, n$  do
16:       $w_{ik+1} := V w_{ik+1}$ 
17:    end for
18:     $\text{convergencia} := \text{TRUE}$ 
19:    for  $i := 1, \dots, n$  do
20:      if  $|w'_{ik} w_{ik+1}| \neq 1$  then
21:         $\text{convergencia} := \text{FALSE}$ 
22:      end if
23:    end for
24:    for  $i := 1, \dots, n$  do
25:       $w_{ik} = w_{ik+1}$ 
26:    end for
27:  until  $\text{convergencia}$ 
28:  return  $(w_{1k}, \dots, w_{mk})'$ 
29: end procedure

```

Apêndice B

Análise de Componentes Principais Comuns

Neste apêndice é feita uma breve revisão sobre a análise de componentes principais comuns (ou CPCA), tendo em vista três objetivos principais. Em primeiro lugar, definir formalmente a CPCA, esclarecendo a relação desta com as análises de componentes independentes (ICA) e principais (PCA). Em seguida, definir o problema de otimização usado para estimar as componentes principais comuns (ou CPCs) e, por último, apresentar a adaptação do algoritmo FG para a estimação das mesmas.

Para um melhor tratamento sobre CPCs, recomenda-se a consulta a [20, 4]. Além disso, como nesta revisão são omitidas as provas de corretude e convergência do algoritmo FG, recomenda-se a consulta a [21, 19] para um tratamento mais completo. Aqui, a apresentação se limita à descrição do algoritmo e à apresentação de seu pseudocódigo.

B.1 Definição

A PCA e a ICA são métodos estatísticos de amostra única, ou seja, estudam as propriedades de apenas uma amostra. Em [20], a CPCA é apresentada como uma generalização da PCA para lidar com um conjunto de amostras, buscando identificar a estrutura comum a todas elas.

Desta forma, por se tratar de uma extensão, as principais características da PCA também podem ser encontradas na CPCA. Entre estas, por exemplo, estão a diagonalização da matriz de covariância e o cálculo das componentes principais através da multiplicação do dado original por uma matriz ortogonal.

Conceitualmente, a CPCA é baseada na *hipótese de componentes principais comuns*, que é caracterizada pela existência de uma matriz capaz de diagonalizar simultaneamente todo o conjunto de matrizes de covariância das amostras em estudo.

Assim, considere um conjunto de amostras $A = \{A_1, \dots, A_p\}$, onde cada amostra A_i tem matriz de covariância amostral S_i e é obtida através de observações de um vetor aleatório x_i . Neste contexto, diz-se que as amostras em A satisfazem a hipótese de componentes principais comuns se existe uma matriz ortogonal B_S que as diagonaliza simultaneamente, ou seja,

$$S_i = B_S \Delta_i^2 B_S', \quad 1 \leq i \leq p, \quad (\text{B.1.1})$$

onde as matrizes Δ_i^2 são diagonais.

Como todas as matrizes S_i são diagonalizáveis por uma mesma matriz ortogonal B_S , a estrutura compartilhada pelas amostras A_i consiste em nada mais que um conjunto de autovetores. No caso, os autovetores das matrizes S_i , que são os vetores coluna da matriz B_S .

Considere agora a matriz S_i associada ao vetor aleatório x_i . Pela Equação (B.1.1) e pela ortogona-

lidade de B_S , vale que

$$\Delta_i^2 = B_S' S_i B_S = B_S' \mathbb{E} [x_i x_i'] B_S = \mathbb{E} [(B_S' x_i) (x_i' B_S)]. \quad (\text{B.1.2})$$

Logo, tomando

$$y_i = B_S' x_i, \quad (\text{B.1.3})$$

obtem-se

$$\Delta_i^2 = \mathbb{E} [y_i y_i'] = \text{Cov} (y_i). \quad (\text{B.1.4})$$

Assim, como na PCA, a Equação (B.1.3) permite obter um vetor y_i com coordenadas não correlacionadas entre si. Neste contexto, entretanto, as coordenadas de y_i são denominadas *componentes principais comuns* de x_i .

Finalmente, pela ortogonalidade de B_S e pela Equação (B.1.3), os vetores x_i podem ser escritos como uma combinação linear de suas componentes principais comuns y_i . Isto é,

$$x_i = B_S y_i. \quad (\text{B.1.5})$$

Desta maneira, pelas Equações (B.1.1), (B.1.3) e (B.1.5), os vetores x_i têm as principais propriedades encontradas na PCA (veja o Apêndice A), o que demonstra a grande semelhança existente entre a CPCA e a PCA.

Até este ponto, considerou-se que a hipótese de componentes principais comuns é satisfeita. Na prática, entretanto, dificilmente é possível encontrar a matriz B_S que diagonalize todas as matrizes de covariância simultaneamente.

Assim, a definição formal da CPCA (veja a Definição 19) abandona este requisito em favor de um menos restritivo, exigindo apenas que as matrizes Δ_i sejam simultaneamente o mais diagonais possível.

Além disso, a Definição 19 considera o caso em que a matriz B_S tem tamanho $m \times n$ para permitir uma melhor comparação com as definições da PCA e da ICA, feitas no Apêndice A. No restante deste apêndice, entretanto, considera-se o caso mais simples em que B_S é uma matriz ortogonal.

Definição 19. *Análise de Componentes Principais Comuns (CPCA)*

Seja $X = \{x_1, \dots, x_p\}$ um conjunto de p vetores aleatórios m -dimensionais com médias zero e matrizes de covariância $\Sigma_1, \dots, \Sigma_p$. Além disso, cada vetor x_i está associado a um vetor n -dimensional y_i , que contém as chamadas componentes principais comuns (ou CPCs) de x_i . Neste contexto, a CPCA de X consiste no conjunto de matrizes $\{B_S, \Delta_1, \dots, \Delta_p\}$ satisfazendo:

- (i) B_S é uma matriz $m \times n$ de posto completo satisfazendo $B_S' B_S = I_n$,
- (ii) $\Delta_1, \dots, \Delta_p$ são matrizes reais $n \times n$, positivas e, simultaneamente, o mais diagonais possível
- (iii) $\Sigma_i = B_S \Delta_i^2 B_S'$, $1 \leq i \leq p$.

B.2 Estimação

Uma vez disponível a estimativa \hat{B} para a matriz ortogonal B_S , que extrai as componentes principais comuns de um conjunto de amostras, as estimativas para a CPCA e para as CPCs podem ser calculadas com facilidade. Desta forma a estimação de B_S é de fundamental importância no contexto da CPCA.

As CPCs do vetor aleatório x_i , por exemplo, podem ser calculadas pela Equação (B.1.3), ou seja, as CPCs y_i são estimadas como

$$\hat{y}_i = \hat{B}' x_i. \quad (\text{B.2.1})$$

No caso da CPCA, pela Definição 19 e pela Equação (B.1.1), a estimativa é dada pelo conjunto

$$\hat{B} \cup \left\{ \Delta_1 = \hat{B}' S_1 \hat{B}, \dots, \Delta_p = \hat{B}' S_p \hat{B} \right\}. \quad (\text{B.2.2})$$

Assumindo que os vetores aleatórios x_i têm distribuição Normal multivariada, é possível obter um estimador para B_S baseado na maximização de verossimilhança, como feito em [20]. A hipótese de normalidade, entretanto, limita a aplicação do estimador obtido.

Desta forma, aqui é adotado o estimador desenvolvido em [4], onde a estimativa para B_S é calculada por mínimos quadrados, explorando diretamente a propriedade de diagonalização simultânea das matrizes de covariância. Mais especificamente, a estimativa \hat{B} é dada pela matriz ortogonal que minimiza o erro total de diagonalização das p matrizes S_1, \dots, S_p , ou seja,

$$\hat{B} = \operatorname{argmin}_{F \in \mathcal{O}_m} \sum_{k=1}^p g(F, S_k), \quad (\text{B.2.3})$$

onde \mathcal{O}_m é o conjunto de matrizes ortogonais de tamanho $m \times m$ e $g(F, S_k)$ a função que mede o erro de diagonalização da matriz S_k pela matriz F .

Quanto ao erro de diagonalização da matriz S_k pela matriz F , este é dado pela soma do quadrado dos elementos fora da diagonal da matriz $F' S_k F$, que são nulos quando S_k é diagonalizável por F . Assim, o erro $g(F, S_k)$ é dado por

$$g(F, S_k) = \|F' S_k F - \operatorname{diag}(F' S_k F)\|_F^2, \quad (\text{B.2.4})$$

onde $\|\cdot\|_F$ é a norma de Frobenius.

A Equação (B.2.4), entretanto, pode ser escrita de forma mais conveniente. Tomando $Q = F' S_k F$ e $D = \operatorname{diag}(F' S_k F)$, obtém-se

$$\begin{aligned} g(F, S_k) &= \|Q - D\|_F^2 \\ &= \operatorname{traço} [(Q - D)' (Q - D)] \\ &= \operatorname{traço} [Q' Q] - \operatorname{traço} [Q' D] - \operatorname{traço} [D' Q] + \operatorname{traço} [D^2] \\ &= \operatorname{traço} [Q' Q] - 2 \operatorname{traço} [Q' D] + \operatorname{traço} [D^2]. \end{aligned} \quad (\text{B.2.5})$$

Pela ortogonalidade de F e pela invariância do traço de uma matriz quanto a transformações de similaridade, $Q' Q$ e $S'_k S_k$ têm o mesmo traço, como demonstra a Equação (B.2.6).

$$\operatorname{traço} [Q' Q] = \operatorname{traço} [(F' S_k F)' (F' S_k F)] = \operatorname{traço} [F' S'_k S_k F] = \operatorname{traço} [S'_k S_k] \quad (\text{B.2.6})$$

Além disso, é válido que

$$\operatorname{traço} [Q' D] = \operatorname{traço} [D^2] = \sum_{j=1}^m (F'_j S_k F_j)^2. \quad (\text{B.2.7})$$

Assim, pelas Equações (B.2.5), (B.2.6) e (B.2.7), o erro $g(F, S_k)$ pode ser escrito como

$$g(F, S_k) = \operatorname{traço} [S'_k S_k] - \sum_{j=1}^m (F'_j S_k F_j)^2. \quad (\text{B.2.8})$$

O erro total de diagonalização é, então, dado por

$$\sum_{k=1}^p g(F, S_k) = \sum_{k=1}^p \text{traço} [S'_k S_k] - \sum_{k=1}^p \sum_{j=1}^m (F'_j S_k F_j)^2 = K - \sum_{k=1}^p \sum_{j=1}^m (F'_j S_k F_j)^2. \quad (\text{B.2.9})$$

Como o termo K na Equação (B.2.9) tem valor independente de F , a minimização do erro total de diagonalização pode ser escrita como um problema, mais simples, de maximização. Desta forma, a estimativa de B_S pode ser calculada como

$$\hat{B} = \operatorname{argmax}_{F \in \mathcal{O}_m} \sum_{k=1}^p \sum_{j=1}^m (F'_j S_k F_j)^2. \quad (\text{B.2.10})$$

Em [4], o cálculo de \hat{B} é feito atribuindo diferentes pesos às matrizes S_k , sendo o peso de cada matriz determinado pelo número n_k de amostras utilizado em sua estimação. Adotando esta estratégia, o valor de \hat{B} passa ser dado por

$$\hat{B} = \operatorname{argmax}_{F \in \mathcal{O}_m} \sum_{k=1}^p \sum_{j=1}^m 1/2 (n_k - 1)^2 (F'_j S_k F_j)^2. \quad (\text{B.2.11})$$

A fim de encontrar as soluções do Problema (B.2.11), é preciso calcular sua função lagrangiana. Sendo as restrições de ortogonalidade descritas pelas Equações (B.2.12),

$$\begin{cases} F'_i F_i = 1, & 1 \leq i \leq m \\ F'_i F_j = 0, & 1 \leq i \leq m, 1 \leq j \leq m, i \neq j \end{cases} \quad (\text{B.2.12})$$

a lagrangiana em questão é dada por

$$\mathcal{L}(F) = \sum_{k=1}^p \sum_{j=1}^m 1/2 (n_k - 1)^2 [F'_j S_k F_j]^2 - 2 \sum_{i=1}^m \sum_{j=i+1}^m \lambda_{ij} F'_j F_i - \sum_{i=1}^m \lambda_i (1 - F'_i F_i). \quad (\text{B.2.13})$$

Como as restrições $F'_i F_j$ e $F'_j F_i$ são iguais, considera-se que λ_{ij} e λ_{ji} , os multiplicadores de Lagrange associados a elas, também o são. Com isso, na Equação (B.2.13), cada multiplicador λ_{ij} é contabilizado duas vezes.

Finalmente, o sistema de equações que permite calcular \hat{B} é obtido através dos valores, nos pontos de máximo da lagrangiana \mathcal{L} , dos multiplicadores λ_{ij} . Estes, por sua vez, podem ser calculados multiplicando as condições de primeira ordem, dadas por

$$\frac{\partial \mathcal{L}}{\partial F_i} = \sum_{k=1}^p 2 (n_k - 1)^2 (F'_i S_k F_i) S_k F_i - 2 \sum_{j=1, j \neq i}^m \lambda_{ij} F_j - 2 \lambda_i F_i = 0, \quad (\text{B.2.14})$$

à esquerda pelo vetor coluna F_q , ou seja, a partir da equação

$$F'_q \left(\frac{\partial \mathcal{L}}{\partial F_i} \right) = \sum_{k=1}^p (n_k - 1)^2 (F'_i S_k F_i) F'_q S_k F_i - \sum_{j=1, j \neq i}^m \lambda_{ij} F'_q F_j - \lambda_i F'_q F_i = 0. \quad (\text{B.2.15})$$

Tomando $q \neq i$ na Equação (B.2.15), obtém-se que

$$\lambda_{iq} = \sum_{k=1}^p (n_k - 1)^2 (F'_i S_k F_i) (F'_q S_k F_i), \quad (\text{B.2.16})$$

já que $F'_q F_j = 1$ e $F'_q F_i = 0$, pela ortogonalidade de F .

Como as matrizes de covariância S_k são simétricas, $F'_i S_k F_i$ também o são. Assim, vale que

$$F'_q S_k F_i = F'_i S_k F_q. \quad (\text{B.2.17})$$

Desta forma, pelas Equações (B.2.16) e (B.2.17), o multiplicador λ_{iq} também é dado pela expressão

$$\lambda_{iq} = \sum_{k=1}^p (n_k - 1)^2 (F'_i S_k F_i) (F'_i S_k F_q). \quad (\text{B.2.18})$$

Já que λ_{iq} é igual a λ_{qi} , a diferença destes é nula e dada por

$$\begin{aligned} \lambda_{qi} - \lambda_{iq} &= \sum_{k=1}^p (n_k - 1)^2 (F'_q S_k F_q) (F'_q S_k F_i) - \sum_{k=1}^p (n_k - 1)^2 (F'_i S_k F_i) (F'_q S_k F_i) \\ &= \sum_{k=1}^p (n_k - 1)^2 [F'_q S_k F_q - F'_i S_k F_i] (F'_q S_k F_i) \\ &= F'_q \left(\sum_{k=1}^p (n_k - 1)^2 [F'_q S_k F_q - F'_i S_k F_i] S_k \right) F_i \\ &= 0 \end{aligned} \quad (\text{B.2.19})$$

Assim, aplicando a Equação (B.2.19) a todos os pares de colunas de F , obtém-se o conjunto de equações que, se resolvido, fornece a estimativa \hat{B} . Lembrando que as soluções do sistema devem ser matrizes ortogonais, a estimação de B_S pode ser escrita, então, como o Problema (B.2.20), onde são incluídas as restrições de ortogonalidade.

$$\left\{ \begin{array}{l} \hat{B}'_j \left(\sum_{k=1}^p (n_k - 1)^2 [\hat{B}'_j S_k \hat{B}_j - \hat{B}'_i S_k \hat{B}_i] S_k \right) \hat{B}_i = 0, \quad 1 \leq i < j \leq m \\ \text{Sujeito a:} \\ \hat{B}' \hat{B} = I_m \end{array} \right. \quad (\text{B.2.20})$$

Quanto às soluções do Problema (B.2.20), estas podem ser encontradas através de uma adaptação do algoritmo FG (ver [4, 21, 19]), que é apresentada na Seção B.3.

A estimação por máxima verossimilhança, que considera apenas vetores x_i com distribuição Normal multivariada, é feita através da resolução do Problema (B.2.21) (ver [20]). Comparando-o com o problema obtido pela técnica de mínimos quadrados (Problema (B.2.20)), percebe-se a que a diferença entre os dois está na atribuição dos pesos dados às matrizes S_k .

$$\left\{ \begin{array}{l} \hat{B}'_j \left(\sum_{k=1}^p (n_k - 1) \left[\frac{\hat{B}'_j S_k \hat{B}_j - \hat{B}'_i S_k \hat{B}_i}{\hat{B}'_j S_k \hat{B}_j \quad \hat{B}'_i S_k \hat{B}_i} \right] S_k \right) \hat{B}_i = 0, \quad 1 \leq i < j \leq m \\ \text{Sujeito a:} \\ \hat{B}' \hat{B} = I_m \end{array} \right. \quad (\text{B.2.21})$$

B.3 Adaptação do algoritmo FG

O algoritmo FG foi originalmente criado para encontrar as soluções do Problema (B.2.21), ou seja, para fazer a estimação da matriz B_S no caso de vetores aleatórios com distribuição Normal (veja [21]). Em [4], sem fornecer detalhes sobre como adaptar o algoritmo FG, o mesmo é utilizado para resolver o Problema (B.2.20). Assim, nesta seção o algoritmo FG é apresentado em maiores detalhes, com foco na estimação de B_S por mínimos quadrados.

O algoritmo FG é uma boa alternativa para a resolução dos Problemas (B.2.20) e (B.2.21) devido à simplicidade de sua implementação computacional e robustez numérica. De acordo com [44], por depender principalmente do cálculo de autovetores de matrizes 2×2 , o algoritmo FG evita os transtornos tipicamente encontrados em problemas de otimização com muitas variáveis.

Assim, seja $S = \{S_1, \dots, S_p\}$ um conjunto de p matrizes de covariância amostral que devem ser diagonalizadas simultaneamente e $\Phi(F, S)$ o erro total de diagonalização dos elementos de S por uma matriz ortogonal F .

O algoritmo FG encontra as soluções para o Problema (B.2.20) iterativamente, produzindo uma sequência $\{\hat{B}_k\}_{k \geq 0}$ de matrizes ortogonais onde o erro total de diagonalização é reduzido progressivamente até que um mínimo local da função de erro $\Phi(\hat{B}_k, S)$ seja atingido, ou seja, para duas matrizes desta sequência, \hat{B}_j e \hat{B}_k com $j < k$, vale que $\Phi(\hat{B}_j, S) \geq \Phi(\hat{B}_k, S)$.

Considere agora a k -ésima iteração do algoritmo FG. Nesta iteração, o algoritmo calcula a matriz \hat{B}_{k+1} , aplicando uma rotação específica a cada um dos $m(m+1)/2$ pares de colunas da matriz \hat{B}_k , obtida na iteração passada.

Cada uma das rotações aplicadas é escolhida de forma com que o par de colunas rotacionado satisfaça à equação do Problema (B.2.20) a ele correspondente, ou seja, ao par de colunas $\hat{B}_{k,j}$ e $\hat{B}_{k,i}$ da matriz \hat{B}_k é aplicada uma rotação Q_{ji} de forma que as colunas rotacionadas, R_j e R_i , satisfaçam à equação

$$R'_j \left(\sum_{k=1}^p (n_k - 1)^2 [R'_j S_k R_j - R'_i S_k R_i] S_k \right) R_i = 0. \quad (\text{B.3.1})$$

Desta maneira, ao fim de cada iteração, os pares de colunas da matriz \hat{B}_{k+1} são soluções viáveis para o Problema (B.2.20), pois obedecem às restrições do mesmo. Além disso, como mostrado em [21], a cada iteração o erro total de diagonalização diminui ou não sofre alteração.

Finalmente, para completar a visão geral do algoritmo FG, é preciso definir a função do erro total de diagonalização. Adotando a métrica de erro definida na Equação (B.2.4), $\Phi(F, S)$ é dado por

$$\Phi(F, S) = \sum_{k=1}^p \|F' S_k F - \text{diag}(F' S_k F)\|_F^2. \quad (\text{B.3.2})$$

Outras métricas para o erro de diagonalização também podem ser adotadas. Em [21], por exemplo, o

erro de diagonalização total é dado por

$$\Phi(F, S, n) = \prod_{i=1}^k \left[\frac{\det(\text{diag}(F' S_i F))}{\det(F' S_i F)} \right]^{n_i}, \quad (\text{B.3.3})$$

onde n_i corresponde ao número de amostras usadas no cálculo da matriz de covariância amostral S_i .

No Algoritmo 7 é apresentado o pseudocódigo da função F , ponto de entrada do algoritmo FG. Como pode ser visto no mesmo, o algoritmo realiza um número indeterminado de iterações do *loop* mais externo (linhas 3 a 12) até que ocorra a convergência do erro total de diagonalização (na linha 12). Além disso, a cada iteração do *loop* mais interno (linhas 7 a 9), a rotação Q é aplicada à matriz formada pelas colunas B_j e B_i , resultando na matriz R de colunas rotacionadas. Então, as colunas B_j e B_i são atualizadas com as colunas de R (na linha 9).

Algorithm 7 Algoritmo F

```

1: procedure F( $A_1, \dots, A_p; n_1, \dots, n_p$ )
2:    $B := I_m$ 
3:   repeat
4:      $B_{old} := B$ 
5:     for  $j := 1, \dots, m$  do
6:       for  $i := l + 1, \dots, m$  do
7:          $Q := \mathbf{G}(B_j; B_i; A_1, \dots, A_p; n_1, \dots, n_p)$ 
8:          $R := [B_j \ B_i] Q$ 
9:          $B := [B_1 \ \dots \ B_{j-1} \ R_1 \ B_{j+1} \ \dots \ B_{i-1} \ R_2 \ B_{i+1} \ \dots \ B_m]$ 
10:      end for
11:    end for
12:    until  $|\Phi(B) - \Phi(B_{old})| \approx 0$ 
13:    return  $B$ 
14: end procedure

```

A prova de convergência da função F pode ser encontrada em [21]. Aqui, a apresentação se limita à sua descrição e à apresentação do seu pseudocódigo.

Para completar a descrição do algoritmo FG, é preciso ainda definir o algoritmo para o cálculo das matrizes de rotação Q . Assim, considere a matriz H , formada pelo par de colunas \hat{B}_j e \hat{B}_i , ou seja,

$$H = \begin{bmatrix} \hat{B}_j & \hat{B}_i \end{bmatrix}. \quad (\text{B.3.4})$$

Ao aplicar a rotação Q à matriz H , é obtida a matriz R com as colunas rotacionadas. Assim,

$$R = [R_j \ R_i] = HQ. \quad (\text{B.3.5})$$

Como as colunas rotacionadas R devem obedecer as equações do Problema (B.2.20), a equação correspondente a estas pode ser escrita como

$$R'_j \left(\sum_{k=1}^p w_k S_k \right) R_i = 0, \quad (\text{B.3.6})$$

onde w_k é dado por

$$w_k = (n_k - 1)^2 [R'_j S_k R_j - R'_i S_k R_i]. \quad (\text{B.3.7})$$

Considere agora a matriz simétrica M , dada por

$$M = \sum_{k=1}^p w_k [R' S_k R]. \quad (\text{B.3.8})$$

Como a matriz $R' S_k R$, de tamanho 2×2 , é dada por

$$R' S_k R = \begin{bmatrix} R'_j S_k R_j & R'_j S_k R_i \\ R'_i S_k R_j & R'_i S_k R_i \end{bmatrix}, \quad (\text{B.3.9})$$

os termos fora da diagonal de M são iguais a

$$M_{12} = M_{21} = \sum_{k=1}^p w_k [R'_j S_k R_i] = R'_j \left(\sum_{k=1}^p w_k S_k \right) R_i. \quad (\text{B.3.10})$$

Logo, a Equação (B.3.6) é satisfeita se, e somente se, R_j e R_i forem escolhidos de forma que M seja uma matriz diagonal.

Desta maneira, escrevendo a matriz M em função de Q , esta pode ser escolhida de forma que M seja diagonal e, conseqüentemente, de forma que a Equação (B.3.6) seja satisfeita.

Assim, considere as seguintes definições:

$$T_k := H' S_k H = \begin{bmatrix} \hat{B}'_j S_k \hat{B}_j & \hat{B}'_j S_k \hat{B}_i \\ \hat{B}'_i S_k \hat{B}_j & \hat{B}'_i S_k \hat{B}_i \end{bmatrix} \quad (\text{B.3.11})$$

e

$$v_k := (n_k - 1)^2 [Q'_1 T_k Q_1 - Q'_2 T_k Q_2]. \quad (\text{B.3.12})$$

Pela Equação (B.3.5), vale que

$$R' S_k R = (HQ)' S_k (HQ) = Q' (H' S_k H) Q = Q' T_k Q. \quad (\text{B.3.13})$$

Quanto a w_k , também pela Equação (B.3.5) vale que

$$\begin{aligned} w_k &= (n_k - 1)^2 [R'_j S_k R_j - R'_i S_k R_i] \\ &= (n_k - 1)^2 [Q'_1 T_k Q_1 - Q'_2 T_k Q_2] \\ &= v_k. \end{aligned} \quad (\text{B.3.14})$$

Logo, a matriz M pode ser escrita em função de Q como

$$\begin{aligned} M &= \sum_{k=1}^p v_k [Q' T_k Q] \\ &= Q' \left(\sum_{k=1}^p v_k T_k \right) Q \\ &= Q' W Q, \end{aligned} \quad (\text{B.3.15})$$

onde a matriz W é dada por

$$W = \sum_{k=1}^p v_k T_k. \quad (\text{B.3.16})$$

Finalmente, pela Equação (B.3.15), para que M seja diagonal, basta que a matriz de rotação Q seja composta pelos autovetores da matriz W . Entretanto, como o valor da matriz W depende da rotação Q através do termo v_k (veja Equação (B.3.12)), é necessário um algoritmo iterativo para calcular o valor de Q (ver [21]).

Assim, considere o Algoritmo 8, onde a função G , responsável pelo cálculo da rotação Q , é descrita.

Analisando o pseudocódigo de G , é possível ver que a mesma realiza um número indeterminado de iterações do *loop* principal (linhas 6 a 15) até que a matriz de rotação Q não apresente alterações significativas. Na linha 15, $\|\cdot\|$ representa uma norma matricial.

Além disso, percebe-se que, a cada iteração, a função G calcula a matriz W utilizando o valor mais recente da rotação Q . Em seguida, Q é atualizada com os autovetores da nova matriz W . Vale ressaltar que a matriz W tem tamanho 2×2 . Assim, o cálculo de seus autovetores é trivial.

Como no caso do algoritmo F , a prova de convergência do algoritmo G é omitida aqui. A mesma, entretanto, pode ser encontrada em [21].

Algorithm 8 Algoritmo G

```

1: procedure G( $B_j; B_i; A_1, \dots, A_p; n_1, \dots, n_p$ )
2:    $Q := I_2$ 
3:   for  $k := 1, \dots, p$  do
4:      $T_k := \begin{bmatrix} B_j' A_k B_j & B_j' A_k B_i \\ B_i' A_k B_j & B_i' A_k B_i \end{bmatrix}$ 
5:   end for
6:   repeat
7:      $W := 0$ 
8:     for  $k := 1, \dots, p$  do
9:        $d_1 := Q_1' T_k Q_1$ 
10:       $d_2 := Q_2' T_k Q_2$ 
11:       $W := W + (n_k - 1)^2 (d_1 - d_2) T_k$ 
12:    end for
13:     $Q_{old} := Q$ 
14:     $Q := \text{autovetores}(W)$ 
15:  until  $\|Q - Q_{old}\| \approx 0$ 
16:  return  $Q$ 
17: end procedure

```

Apêndice C

Dados Históricos

As séries de retornos utilizadas nos testes foram calculadas a partir de séries históricas de preços de fechamento de ações negociadas na BM&FBovespa, que podem ser obtidas gratuitamente no *web site* www.bmfbovespa.com.br.

Entre todas as séries disponíveis, vinte foram selecionadas por terem sido negociadas diariamente no período de estudo (de 01/01/2006 a 14/01/2013) e por apresentarem um número médio de negócios diários superior a 1000, o que demonstra uma boa liquidez.

Na Tabela C.1, são apresentados os códigos das ações escolhidas e algumas estatísticas diárias. Entre estas, destacam-se as mais importantes para a seleção das ações: o número médio de negócios (*Mean #Trades*), o volume médio de capital (*Mean Volume*) e a quantidade média de ações negociadas (*Mean Quantity*).

CODNEG	First Date	Last Date	#Dates	Mean #Trades	Mean Volume(R\$)	Mean Quantity	Name
PETR4	20060102	20130114	1737	17853	57864136121	18230315	PETROBRAS
VALE5	20060102	20130114	1737	15875	56916669690	13657950	VALE R DOCE
GGBR4	20060102	20130114	1737	7089	10615071368	4709302	GERDAU
USIM5	20060102	20130114	1737	5829	11335472607	3751863	USIMINAS
VALE3	20060102	20130114	1737	5330	14321899258	3096947	VALE R DOCE
PETR3	20060102	20130114	1737	5324	13272619724	3904257	PETROBRAS
CSNA3	20060102	20130114	1737	4672	9118076299	2643182	SID NACIONAL
CYRE3	20060102	20130114	1737	4521	5116288706	2798380	CYRELA REALT
CMIG4	20060102	20130114	1737	3349	5356616957	89385966	CEMIG
CCRO3	20060102	20130114	1737	2894	3485731420	1390844	CCR RODOVIAS
NATU3	20060102	20130114	1737	2454	3172934570	967617	NATURA
GOAU4	20060102	20130114	1737	2350	2707958185	886626	GERDAU MET
LREN3	20060102	20130114	1737	2306	3920239688	940654	LOJAS RENNEN
BRAP4	20060102	20130114	1737	2256	3943556971	972704	BRADSPAR
KLBN4	20060102	20130114	1737	2161	1241988401	2080341	KLABIN S/A
EMBR3	20060102	20130114	1737	2097	2015113006	1545376	EMBRAER
CPL6	20060102	20130114	1737	1617	2212459408	181023534	COPEL
CRUZ3	20060102	20130114	1737	1607	1840707132	512419	SOUZA CRUZ
DASA3	20060102	20130114	1737	1462	1612560645	813305	DASA
CPFE3	20060102	20130114	1737	1439	1624844233	520024	CPFL ENERGIA

Tabela C.1: Ações da BM&FBovespa selecionadas para estudo

Antes de calcular as séries de retorno diário, os preços históricos das vinte ações selecionadas foram devidamente ajustados em relação aos seus fatores de cotação e aos seus eventos corporativos.

O ajuste quanto ao fator de cotação é essencial pois cada preço fornecido pela bolsa está associado a uma determinada quantidade de ações. Esta quantidade pode ser diferente de ação para ação e também pode sofrer alterações ao longo do tempo, como ocorreu para a ação CMIG4. Assim, todos os preços históricos foram ajustados de forma a corresponder a uma única ação.

Como exemplos de eventos corporativos, é possível citar o pagamento de dividendos, de juros sobre o capital próprio, o grupamento (ou desdobramento) das ações e as aquisições e fusões com

outras empresas. A importância desses eventos reside no efeito direto que têm sobre o preço das ações. Além disso, como estas alterações no preço são fruto de decisões estratégicas e administrativas de cada empresa, estas geralmente não são incorporadas nos modelos de retornos. Assim, devem ser descontadas de maneira apropriada, permitindo a análise com base apenas nas flutuações de mercado do valor de cada ação.

O ajuste devido a um evento corporativo é bem simples. Considere uma ação que tem um evento corporativo E em uma data d . Como todos os eventos corporativos, E pode ser descrito como um fator multiplicativo real α que é aplicado ao preço da ação. Calculado α , o ajuste em relação a E consiste em multiplicar todos os preços anteriores à data d por α .

Considere agora a Figura C.1. Esta apresenta a série de preços de fechamento de duas ações (CMIG4 e EMBR3) antes e depois da realização do ajuste de preços. Como fica evidente, as séries ajustadas apresentam diferenças bastante significativas em relação às originais.

Finalmente, na Tabela C.2 é apresentada a matriz de correlação dos log-retornos e nas figuras C.2 e C.3 são apresentados dois gráficos que permitem comparar os retornos diários das ações selecionadas. O primeiro gráfico é um gráfico de barras que mostra o valor do log-retorno médio para cada uma das ações. O segundo gráfico é um *boxplot* dos retornos, onde é possível ver os quartis da distribuição empírica de cada ação junto com alguns valores extremos de retorno.

Entre os retornos médios, destaca-se o da ação EMBR3 que é o único negativo entre os selecionados. Através da Figura C.1, é possível ver que este valor negativo é razoável.

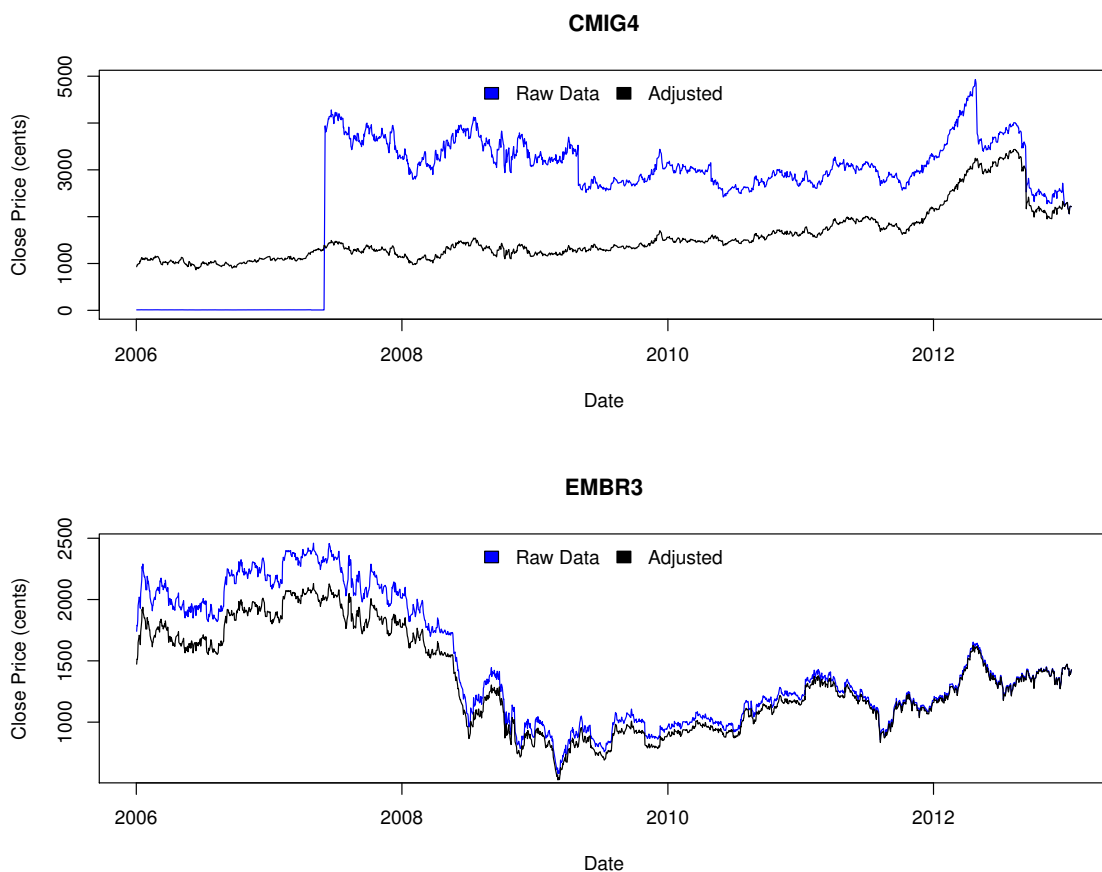


Figura C.1: Séries de preços de fechamento CMIG4 e EMBR3 antes e depois do ajuste

	PETRA4	VALE5	GGBR4	USIM5	VALE3	PETR3	CSNA3	CYRE3	CMIG4	CCRO3	NATU3	GOAU4	LREN3	BRAP4	KLBN4	EMBR3	CPL6	CRUZ3	DASA3	CPFE3
PETRA4	1.00	0.73	0.69	0.58	0.72	0.97	0.66	0.48	0.37	0.35	0.31	0.66	0.40	0.69	0.49	0.35	0.44	0.36	0.32	0.41
VALE5	0.73	1.00	0.75	0.66	0.97	0.72	0.76	0.53	0.43	0.39	0.33	0.73	0.48	0.90	0.54	0.41	0.47	0.38	0.39	0.46
GGBR4	0.69	0.75	1.00	0.74	0.75	0.68	0.78	0.56	0.43	0.40	0.33	0.95	0.48	0.74	0.53	0.42	0.49	0.40	0.40	0.47
USIM5	0.58	0.66	0.74	1.00	0.65	0.57	0.77	0.51	0.39	0.34	0.31	0.73	0.47	0.67	0.46	0.35	0.45	0.35	0.38	0.43
VALE3	0.72	0.97	0.75	0.65	1.00	0.72	0.75	0.52	0.41	0.37	0.33	0.73	0.47	0.91	0.52	0.41	0.45	0.37	0.39	0.44
PETR3	0.97	0.72	0.68	0.57	0.72	1.00	0.66	0.48	0.36	0.34	0.30	0.66	0.38	0.68	0.49	0.36	0.42	0.35	0.31	0.39
CSNA3	0.66	0.76	0.78	0.77	0.75	0.66	1.00	0.53	0.41	0.35	0.31	0.75	0.47	0.74	0.50	0.42	0.48	0.35	0.40	0.46
CYRE3	0.48	0.53	0.56	0.51	0.52	0.48	0.53	1.00	0.37	0.37	0.33	0.57	0.51	0.54	0.45	0.37	0.41	0.31	0.33	0.40
CMIG4	0.37	0.43	0.43	0.39	0.41	0.36	0.41	0.37	1.00	0.35	0.30	0.43	0.38	0.42	0.38	0.28	0.62	0.32	0.28	0.53
CCRO3	0.35	0.39	0.40	0.34	0.37	0.34	0.35	0.37	0.35	1.00	0.27	0.41	0.32	0.39	0.37	0.28	0.36	0.35	0.27	0.36
NATU3	0.31	0.33	0.33	0.31	0.33	0.30	0.31	0.33	0.30	0.27	1.00	0.34	0.37	0.34	0.30	0.21	0.32	0.33	0.22	0.36
GOAU4	0.66	0.73	0.95	0.73	0.73	0.66	0.75	0.57	0.43	0.41	0.34	1.00	0.47	0.73	0.52	0.40	0.48	0.40	0.39	0.47
LREN3	0.40	0.48	0.48	0.47	0.47	0.38	0.47	0.51	0.38	0.32	0.37	0.47	1.00	0.48	0.40	0.30	0.41	0.29	0.33	0.43
BRAP4	0.69	0.90	0.74	0.67	0.91	0.68	0.74	0.54	0.42	0.39	0.34	0.73	0.48	1.00	0.54	0.40	0.48	0.39	0.39	0.46
KLBN4	0.49	0.54	0.53	0.46	0.52	0.49	0.50	0.45	0.38	0.37	0.30	0.52	0.40	0.54	1.00	0.38	0.41	0.38	0.32	0.40
EMBR3	0.35	0.41	0.42	0.35	0.41	0.36	0.42	0.37	0.28	0.28	0.21	0.40	0.30	0.40	0.38	1.00	0.29	0.21	0.25	0.34
CPL6	0.44	0.47	0.49	0.45	0.45	0.42	0.48	0.41	0.62	0.36	0.32	0.48	0.41	0.48	0.41	0.29	1.00	0.36	0.32	0.56
CRUZ3	0.36	0.38	0.40	0.35	0.37	0.35	0.35	0.31	0.32	0.35	0.33	0.40	0.29	0.39	0.38	0.21	0.36	1.00	0.26	0.38
DASA3	0.32	0.39	0.40	0.38	0.39	0.31	0.40	0.33	0.28	0.27	0.22	0.39	0.33	0.39	0.32	0.25	0.32	0.26	1.00	0.31
CPFE3	0.41	0.46	0.47	0.43	0.44	0.39	0.46	0.40	0.53	0.36	0.36	0.47	0.43	0.46	0.40	0.34	0.56	0.38	0.31	1.00

Tabela C.2: Matriz de correlação dos log-retornos diários

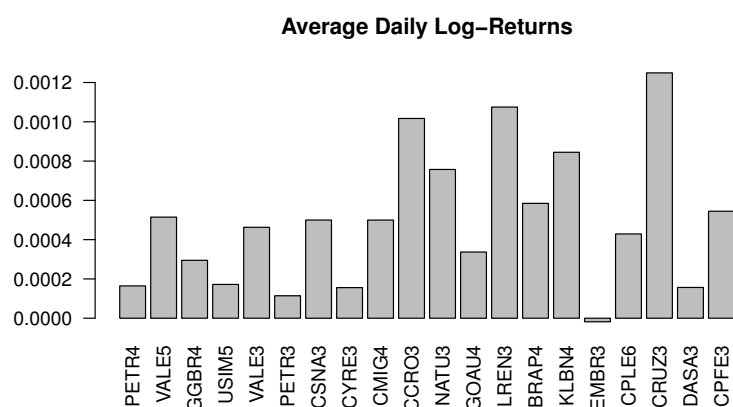


Figura C.2: Média dos log-retornos diários

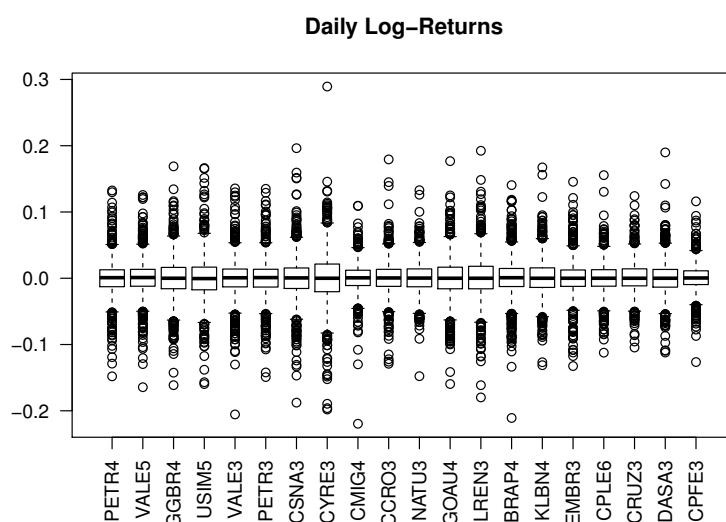


Figura C.3: Boxplot dos log-retornos diários

Apêndice D

Diagnóstico Modelos IGARCH

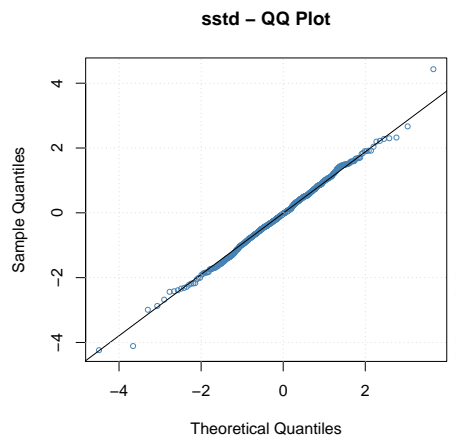
PETR4 – Diagnóstico do Modelo IGARCH(1,1)

Test	Null.Hypothesis	pvalue	Rejected.0.01	Rejected.0.05
1 Kurtosis Test	Kurtosis == 3	0.0000	TRUE	TRUE
2 Skewness Test	Skewness == 0	0.0017	TRUE	TRUE
3 Jarque Bera Test	Normality	0.0000	TRUE	TRUE
4 LM Test for ARCH effects (lags=12)	No ARCH effects	0.7685	FALSE	FALSE
5 KPSS Test	Level stationarity	0.1000	FALSE	FALSE
6 Ljung-Box (lags=1)	No autocorrelation	0.8056	FALSE	FALSE
7 Ljung-Box (lags=2)	No autocorrelation	0.8856	FALSE	FALSE
8 Ljung-Box (lags=3)	No autocorrelation	0.9414	FALSE	FALSE
9 Ljung-Box (lags=4)	No autocorrelation	0.9750	FALSE	FALSE
10 Ljung-Box (lags=5)	No autocorrelation	0.6860	FALSE	FALSE
11 Ljung-Box (lags=6)	No autocorrelation	0.7911	FALSE	FALSE
12 Ljung-Box (lags=7)	No autocorrelation	0.8712	FALSE	FALSE

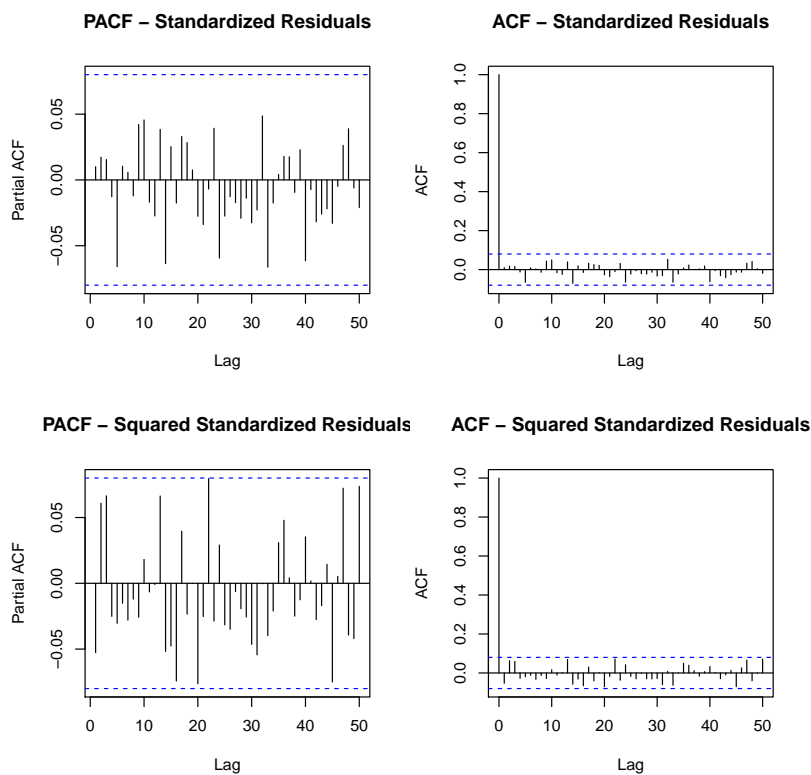
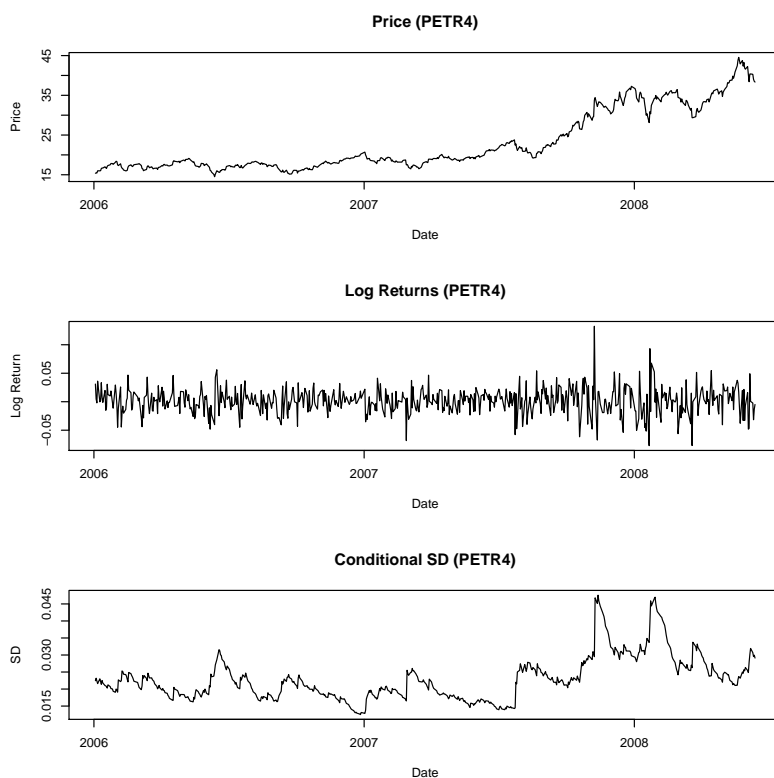
Tabela D.1: Testes para os resíduos padronizados PETR4

Test	Null.Hypothesis	pvalue	Rejected.0.01	Rejected.0.05
1 Kurtosis Test	Kurtosis == 3	0.0000	TRUE	TRUE
2 Skewness Test	Skewness == 0	0.0000	TRUE	TRUE
3 Jarque Bera Test	Normality	0.0000	TRUE	TRUE
4 LM Test for ARCH effects (lags=12)	No ARCH effects	1.0000	FALSE	FALSE
5 KPSS Test	Level stationarity	0.1000	FALSE	FALSE
6 Ljung-Box (lags=1)	No autocorrelation	0.1958	FALSE	FALSE
7 Ljung-Box (lags=2)	No autocorrelation	0.1283	FALSE	FALSE
8 Ljung-Box (lags=3)	No autocorrelation	0.0994	FALSE	FALSE
9 Ljung-Box (lags=4)	No autocorrelation	0.1514	FALSE	FALSE
10 Ljung-Box (lags=5)	No autocorrelation	0.2246	FALSE	FALSE
11 Ljung-Box (lags=6)	No autocorrelation	0.3178	FALSE	FALSE
12 Ljung-Box (lags=7)	No autocorrelation	0.3606	FALSE	FALSE

Tabela D.2: Testes para o quadrado dos resíduos padronizados PETR4



PETR4 – Diagnóstico do Modelo IGARCH(1,1) – Gráficos



PETR4 – Resultados do Ajuste do Modelo IGARCH(1,1)

```

*-----*
*          GARCH Model Fit          *
*-----*

Conditional Variance Dynamics
-----
GARCH Model : iGARCH(1,1)
Mean Model  : ARFIMA(0,0,0)
Distribution : sstd

Optimal Parameters
-----
      Estimate  Std. Error  t value  Pr(>|t|)
mu      0.001665   0.000838   1.9867  0.046953
omega   0.000005   0.000004   1.1334  0.257059
alpha   0.080920   0.036645   2.2082  0.027232
beta1   0.919080           NA           NA           NA
skew    0.879619   0.054739   16.0693  0.000000
shape   7.516615   2.311313   3.2521  0.001146

Robust Standard Errors:
      Estimate  Std. Error  t value  Pr(>|t|)
mu      0.001665   0.000934   1.78193  0.074760
omega   0.000005   0.000007   0.77574  0.437904
alpha   0.080920   0.059568   1.35845  0.174322
beta1   0.919080           NA           NA           NA
skew    0.879619   0.057650   15.25783  0.000000
shape   7.516615   2.619332   2.86967  0.004109

LogLikelihood : 1463.434

Information Criteria
-----
Akaike      -4.8614
Bayes       -4.8248
Shibata     -4.8616
Hannan-Quinn -4.8472

Q-Statistics on Standardized Residuals
-----
      statistic  p-value
Lag[1]          0.06056  0.8056
Lag[p+q+1][1]  0.06056  0.8056
Lag[p+q+5][5]  3.09096  0.6860
d.o.f=0
H0 : No serial correlation

Q-Statistics on Standardized Squared Residuals
-----
      statistic  p-value
Lag[1]          1.674  0.19578
Lag[p+q+1][3]   6.264  0.01232
Lag[p+q+5][7]   7.691  0.17410
d.o.f=2

ARCH LM Tests
-----
      Statistic  DoF  P-Value
ARCH Lag[2]     3.85   2  0.1458
ARCH Lag[5]     7.29   5  0.2000
ARCH Lag[10]    8.45  10  0.5850

Nyblom stability test
-----
Joint Statistic: 0.9309
Individual Statistics:
mu      0.15350
omega   0.32981
alpha   0.04183
skew    0.17349
shape   0.02816

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic: 1.28 1.47 1.88
Individual Statistic: 0.35 0.47 0.75

Sign Bias Test
-----
      t-value   prob sig
Sign Bias      2.2455  0.025104  **
Negative Sign Bias 0.5888  0.556226
Positive Sign Bias 2.7149  0.006821  ***
Joint Effect    8.3370  0.039536  **

Adjusted Pearson Goodness-of-Fit Test:
-----
      group  statistic  p-value (g-1)
1      20      14.00      0.7837
2      30      28.90      0.4703
3      40      26.13      0.9429
4      50      41.33      0.7736

```

VALE5 – Diagnóstico do Modelo IGARCH(1,1)

Test	Null.Hypothesis	pvalue	Rejected.0.01	Rejected.0.05
1 Kurtosis Test	Kurtosis == 3	0.0000	TRUE	TRUE
2 Skewness Test	Skewness == 0	0.0001	TRUE	TRUE
3 Jarque Bera Test	Normality	0.0000	TRUE	TRUE
4 LM Test for ARCH effects (lags=12)	No ARCH effects	0.9805	FALSE	FALSE
5 KPSS Test	Level stationarity	0.1000	FALSE	FALSE
6 Ljung-Box (lags=1)	No autocorrelation	0.9577	FALSE	FALSE
7 Ljung-Box (lags=2)	No autocorrelation	0.9639	FALSE	FALSE
8 Ljung-Box (lags=3)	No autocorrelation	0.8123	FALSE	FALSE
9 Ljung-Box (lags=4)	No autocorrelation	0.9156	FALSE	FALSE
10 Ljung-Box (lags=5)	No autocorrelation	0.7561	FALSE	FALSE
11 Ljung-Box (lags=6)	No autocorrelation	0.8439	FALSE	FALSE
12 Ljung-Box (lags=7)	No autocorrelation	0.8567	FALSE	FALSE

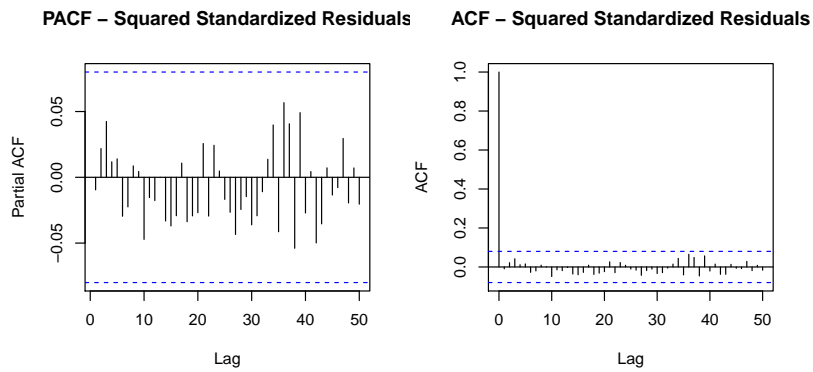
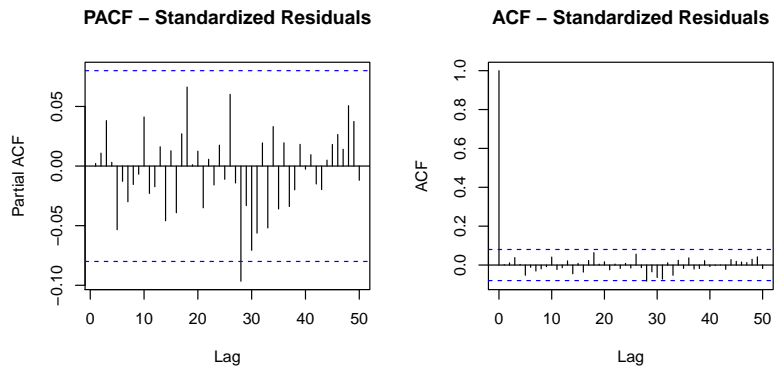
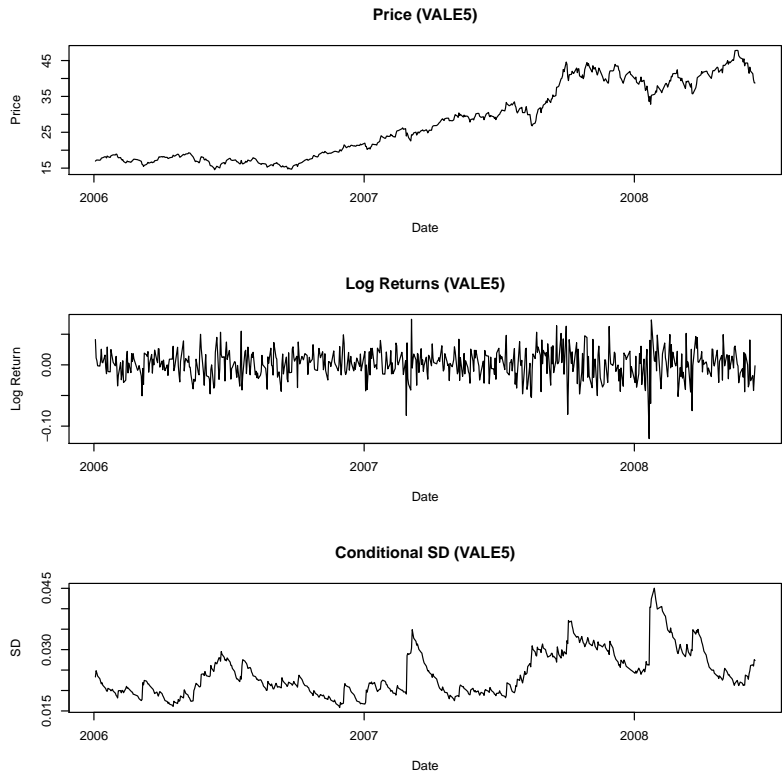
Tabela D.3: Testes para os resíduos padronizados VALE5

Test	Null.Hypothesis	pvalue	Rejected.0.01	Rejected.0.05
1 Kurtosis Test	Kurtosis == 3	0.0000	TRUE	TRUE
2 Skewness Test	Skewness == 0	0.0000	TRUE	TRUE
3 Jarque Bera Test	Normality	0.0000	TRUE	TRUE
4 LM Test for ARCH effects (lags=12)	No ARCH effects	1.0000	FALSE	FALSE
5 KPSS Test	Level stationarity	0.1000	FALSE	FALSE
6 Ljung-Box (lags=1)	No autocorrelation	0.8149	FALSE	FALSE
7 Ljung-Box (lags=2)	No autocorrelation	0.8415	FALSE	FALSE
8 Ljung-Box (lags=3)	No autocorrelation	0.7026	FALSE	FALSE
9 Ljung-Box (lags=4)	No autocorrelation	0.8282	FALSE	FALSE
10 Ljung-Box (lags=5)	No autocorrelation	0.8965	FALSE	FALSE
11 Ljung-Box (lags=6)	No autocorrelation	0.9103	FALSE	FALSE
12 Ljung-Box (lags=7)	No autocorrelation	0.9381	FALSE	FALSE

Tabela D.4: Testes para o quadrado dos resíduos padronizados VALE5



VALE5 – Diagnóstico do Modelo IGARCH(1,1) – Gráficos



VALE5 – Resultados do Ajuste do Modelo IGARCH(1,1)

```

*-----*
*          GARCH Model Fit          *
*-----*

Conditional Variance Dynamics
-----
GARCH Model : iGARCH(1,1)
Mean Model  : ARFIMA(0,0,0)
Distribution : sstd

Optimal Parameters
-----
      Estimate  Std. Error  t value  Pr(>|t|)
mu      0.001482   0.000904   1.6403  0.100936
omega   0.000005   0.000004   1.1635  0.244640
alpha   0.066321   0.029468   2.2507  0.024407
beta1   0.933679           NA           NA           NA
skew    0.906078   0.056444   16.0527 0.000000
shape   9.849934   3.951114   2.4930  0.012669

Robust Standard Errors:
      Estimate  Std. Error  t value  Pr(>|t|)
mu      0.001482   0.000892   1.66131 0.096652
omega   0.000005   0.000005   0.97741 0.328368
alpha   0.066321   0.038271   1.73293 0.083109
beta1   0.933679           NA           NA           NA
skew    0.906078   0.054357   16.66894 0.000000
shape   9.849934   4.341336   2.26887 0.023276

LogLikelihood : 1425.29

Information Criteria
-----
Akaike      -4.7343
Bayes      -4.6977
Shibata    -4.7344
Hannan-Quinn -4.7200

Q-Statistics on Standardized Residuals
-----
      statistic  p-value
Lag[1]          0.00281  0.9577
Lag[p+q+1][1]  0.00281  0.9577
Lag[p+q+5][5]  2.63436  0.7561
d.o.f=0
H0 : No serial correlation

Q-Statistics on Standardized Squared Residuals
-----
      statistic  p-value
Lag[1]          0.05482  0.8149
Lag[p+q+1][3]  1.41267  0.2346
Lag[p+q+5][7]  2.34776  0.7992
d.o.f=2

ARCH LM Tests
-----
      Statistic  DoF  P-Value
ARCH Lag[2]    0.3308  2  0.8475
ARCH Lag[5]    1.6798  5  0.8914
ARCH Lag[10]   3.8433 10  0.9541

Nyblom stability test
-----
Joint Statistic: 1.3021
Individual Statistics:
mu      0.10540
omega   0.41125
alpha   0.02078
skew    0.15800
shape   0.05681

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:      1.28 1.47 1.88
Individual Statistic: 0.35 0.47 0.75

Sign Bias Test
-----
      t-value  prob sig
Sign Bias      0.3324  0.7397
Negative Sign Bias 0.1637  0.8700
Positive Sign Bias 0.4709  0.6379
Joint Effect    0.8547  0.8363

Adjusted Pearson Goodness-of-Fit Test:
-----
      group  statistic  p-value (g-1)
1         20      28.07      0.08215
2         30      38.80      0.10555
3         40      56.40      0.03524
4         50      66.17      0.05148

```

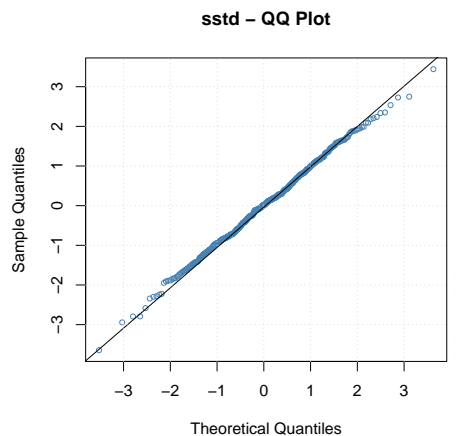
USIM5 – Diagnóstico do Modelo IGARCH(1,1)

Test	Null.Hypothesis	pvalue	Rejected.0.01	Rejected.0.05
1 Kurtosis Test	Kurtosis == 3	0.0232	FALSE	TRUE
2 Skewness Test	Skewness == 0	0.4946	FALSE	FALSE
3 Jarque Bera Test	Normality	0.1297	FALSE	FALSE
4 LM Test for ARCH effects (lags=12)	No ARCH effects	0.5088	FALSE	FALSE
5 KPSS Test	Level stationarity	0.1000	FALSE	FALSE
6 Ljung-Box (lags=1)	No autocorrelation	0.9780	FALSE	FALSE
7 Ljung-Box (lags=2)	No autocorrelation	0.9752	FALSE	FALSE
8 Ljung-Box (lags=3)	No autocorrelation	0.6217	FALSE	FALSE
9 Ljung-Box (lags=4)	No autocorrelation	0.5364	FALSE	FALSE
10 Ljung-Box (lags=5)	No autocorrelation	0.1588	FALSE	FALSE
11 Ljung-Box (lags=6)	No autocorrelation	0.2328	FALSE	FALSE
12 Ljung-Box (lags=7)	No autocorrelation	0.2114	FALSE	FALSE

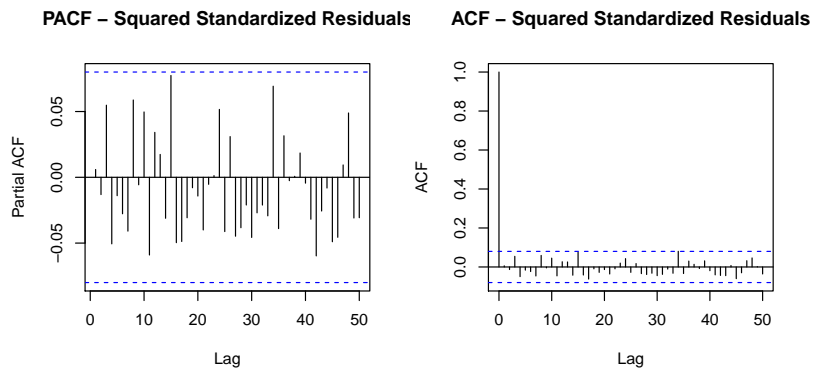
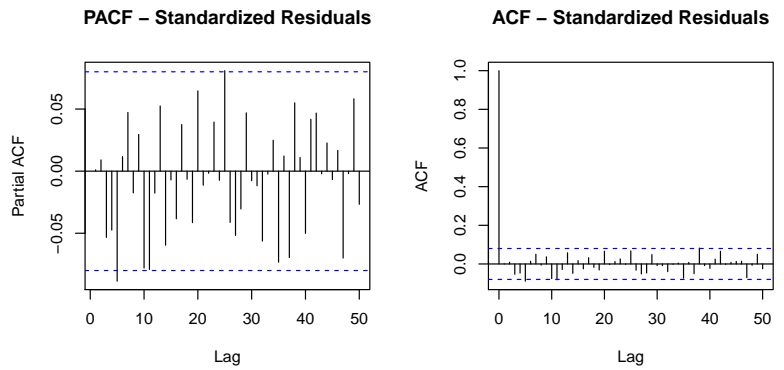
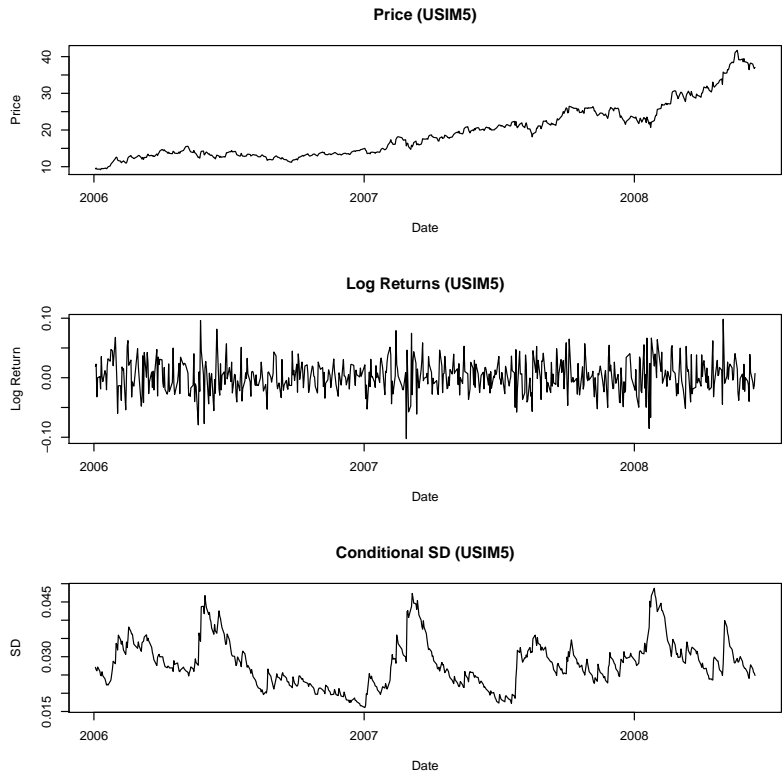
Tabela D.5: Testes para os resíduos padronizados USIM5

Test	Null.Hypothesis	pvalue	Rejected.0.01	Rejected.0.05
1 Kurtosis Test	Kurtosis == 3	0.0000	TRUE	TRUE
2 Skewness Test	Skewness == 0	0.0000	TRUE	TRUE
3 Jarque Bera Test	Normality	0.0000	TRUE	TRUE
4 LM Test for ARCH effects (lags=12)	No ARCH effects	0.6600	FALSE	FALSE
5 KPSS Test	Level stationarity	0.1000	FALSE	FALSE
6 Ljung-Box (lags=1)	No autocorrelation	0.8842	FALSE	FALSE
7 Ljung-Box (lags=2)	No autocorrelation	0.9395	FALSE	FALSE
8 Ljung-Box (lags=3)	No autocorrelation	0.5867	FALSE	FALSE
9 Ljung-Box (lags=4)	No autocorrelation	0.4901	FALSE	FALSE
10 Ljung-Box (lags=5)	No autocorrelation	0.6115	FALSE	FALSE
11 Ljung-Box (lags=6)	No autocorrelation	0.6884	FALSE	FALSE
12 Ljung-Box (lags=7)	No autocorrelation	0.6343	FALSE	FALSE

Tabela D.6: Testes para o quadrado dos resíduos padronizados USIM5



USIM5 – Diagnóstico do Modelo IGARCH(1,1) – Gráficos



USIM5 – Resultados do Ajuste do Modelo IGARCH(1,1)

```

-----*
*          GARCH Model Fit          *
*-----*

Conditional Variance Dynamics
-----
GARCH Model : iGARCH(1,1)
Mean Model  : ARFIMA(0,0,0)
Distribution : sstd

Optimal Parameters
-----
      Estimate  Std. Error  t value  Pr(>|t|)
mu      0.002165   0.001015   2.1328  0.032938
omega   0.000009   0.000006   1.4982  0.134094
alpha   0.095245   0.028850   3.3014  0.000962
beta1   0.904755           NA           NA           NA
skew    1.020804   0.061375  16.6322  0.000000
shape   14.449460   8.635908   1.6732  0.094291

Robust Standard Errors:
      Estimate  Std. Error  t value  Pr(>|t|)
mu      0.002165   0.001006   2.1523  0.031377
omega   0.000009   0.000006   1.3936  0.163445
alpha   0.095245   0.029982   3.1767  0.001489
beta1   0.904755           NA           NA           NA
skew    1.020804   0.058186  17.5438  0.000000
shape   14.449460   8.366013   1.7272  0.084139

LogLikelihood : 1326.954

Information Criteria
-----
Akaike      -4.4065
Bayes      -4.3699
Shibata    -4.4067
Hannan-Quinn -4.3923

Q-Statistics on Standardized Residuals
-----
      statistic  p-value
Lag[1]      0.000762  0.9780
Lag[p+q+1][1] 0.000762  0.9780
Lag[p+q+5][5] 7.953682  0.1588
d.o.f=0
H0 : No serial correlation

Q-Statistics on Standardized Squared Residuals
-----
      statistic  p-value
Lag[1]      0.02122  0.8842
Lag[p+q+1][3] 1.93181  0.1646
Lag[p+q+5][7] 5.21008  0.3908
d.o.f=2

ARCH LM Tests
-----
      Statistic  DoF  P-Value
ARCH Lag[2]    0.1233   2  0.9402
ARCH Lag[5]    3.5261   5  0.6194
ARCH Lag[10]   8.5710  10  0.5732

Nyblom stability test
-----
Joint Statistic: 0.3903
Individual Statistics:
mu      0.03442
omega   0.03863
alpha   0.05125
skew    0.07488
shape   0.12395

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic: 1.28 1.47 1.88
Individual Statistic: 0.35 0.47 0.75

Sign Bias Test
-----
      t-value  prob sig
Sign Bias      0.9066  0.3650
Negative Sign Bias 0.7357  0.4622
Positive Sign Bias 1.2350  0.2173
Joint Effect    2.1285  0.5462

Adjusted Pearson Goodness-of-Fit Test:
-----
      group  statistic  p-value (g-1)
1      20      30.53      0.04539
2      30      35.00      0.20454
3      40      54.40      0.05163
4      50      50.33      0.42045

```

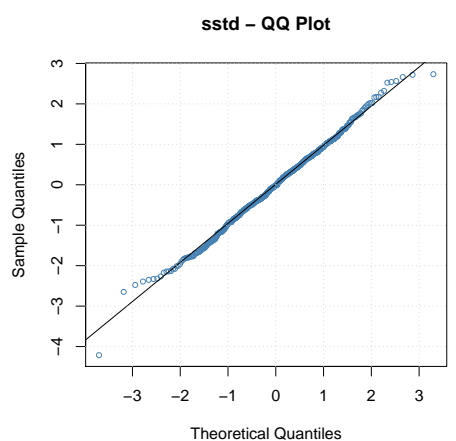
GGBR4 – Diagnóstico do Modelo IGARCH(1,1)

Test	Null.Hypothesis	pvalue	Rejected.0.01	Rejected.0.05
1 Kurtosis Test	Kurtosis == 3	0.0293	FALSE	TRUE
2 Skewness Test	Skewness == 0	0.0696	FALSE	FALSE
3 Jarque Bera Test	Normality	0.0530	FALSE	FALSE
4 LM Test for ARCH effects (lags=12)	No ARCH effects	0.6862	FALSE	FALSE
5 KPSS Test	Level stationarity	0.1000	FALSE	FALSE
6 Ljung-Box (lags=1)	No autocorrelation	0.7890	FALSE	FALSE
7 Ljung-Box (lags=2)	No autocorrelation	0.4966	FALSE	FALSE
8 Ljung-Box (lags=3)	No autocorrelation	0.6562	FALSE	FALSE
9 Ljung-Box (lags=4)	No autocorrelation	0.3587	FALSE	FALSE
10 Ljung-Box (lags=5)	No autocorrelation	0.1437	FALSE	FALSE
11 Ljung-Box (lags=6)	No autocorrelation	0.2117	FALSE	FALSE
12 Ljung-Box (lags=7)	No autocorrelation	0.2848	FALSE	FALSE

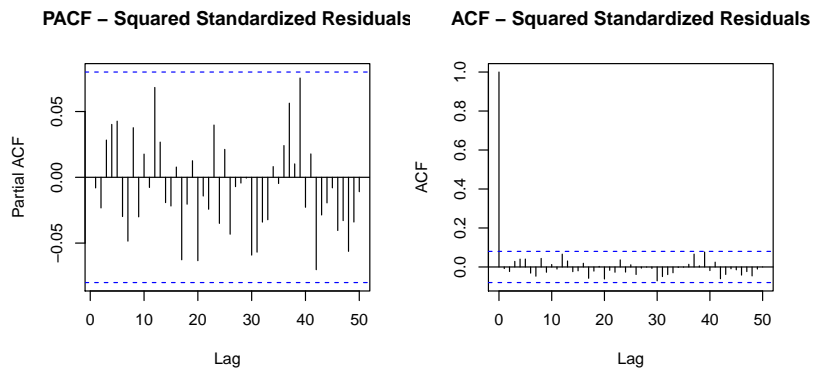
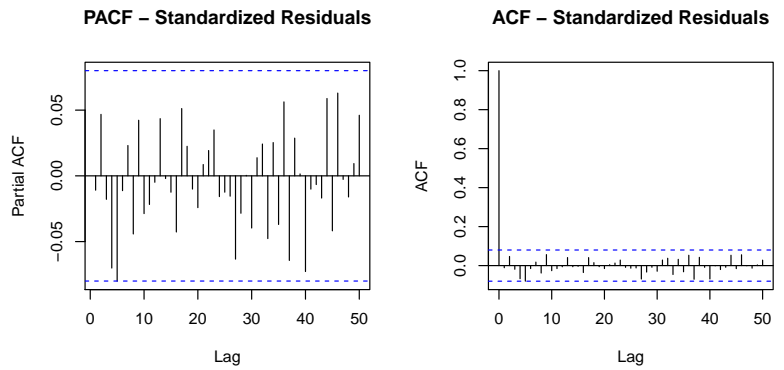
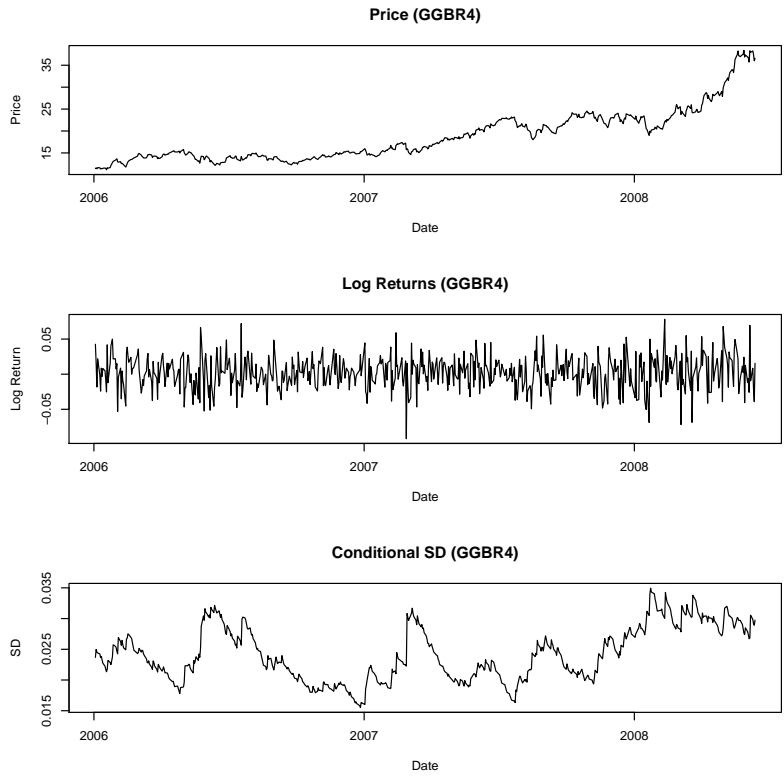
Tabela D.7: Testes para os resíduos padronizados GGBR4

Test	Null.Hypothesis	pvalue	Rejected.0.01	Rejected.0.05
1 Kurtosis Test	Kurtosis == 3	0.0000	TRUE	TRUE
2 Skewness Test	Skewness == 0	0.0000	TRUE	TRUE
3 Jarque Bera Test	Normality	0.0000	TRUE	TRUE
4 LM Test for ARCH effects (lags=12)	No ARCH effects	0.7016	FALSE	FALSE
5 KPSS Test	Level stationarity	0.1000	FALSE	FALSE
6 Ljung-Box (lags=1)	No autocorrelation	0.8434	FALSE	FALSE
7 Ljung-Box (lags=2)	No autocorrelation	0.8335	FALSE	FALSE
8 Ljung-Box (lags=3)	No autocorrelation	0.8346	FALSE	FALSE
9 Ljung-Box (lags=4)	No autocorrelation	0.7645	FALSE	FALSE
10 Ljung-Box (lags=5)	No autocorrelation	0.7243	FALSE	FALSE
11 Ljung-Box (lags=6)	No autocorrelation	0.7513	FALSE	FALSE
12 Ljung-Box (lags=7)	No autocorrelation	0.6840	FALSE	FALSE

Tabela D.8: Testes para o quadrado dos resíduos padronizados GGBR4



GGBR4 – Diagnóstico do Modelo IGARCH(1,1) – Gráficos



GGBR4 – Resultados do Ajuste do Modelo IGARCH(1,1)

```

*-----*
*          GARCH Model Fit          *
*-----*

Conditional Variance Dynamics
-----
GARCH Model : iGARCH(1,1)
Mean Model  : ARFIMA(0,0,0)
Distribution : sstd

Optimal Parameters
-----
      Estimate  Std. Error  t value  Pr(>|t|)
mu      0.001898   0.000918   2.0663  0.038801
omega   0.000003   0.000003   1.0000  0.317308
alpha   0.054451   0.027684   1.9668  0.049201
beta1   0.945549           NA           NA           NA
skew    0.918426   0.056683  16.2027  0.000000
shape  16.971533  11.515214   1.4738  0.140526

Robust Standard Errors:
      Estimate  Std. Error  t value  Pr(>|t|)
mu      0.001898   0.000944   2.00920  0.044516
omega   0.000003   0.000004   0.75525  0.450099
alpha   0.054451   0.040254   1.35267  0.176161
beta1   0.945549           NA           NA           NA
skew    0.918426   0.057619  15.93957  0.000000
shape  16.971533  11.878878   1.42871  0.153086

LogLikelihood : 1405.398

Information Criteria
-----
Akaike      -4.6680
Bayes      -4.6314
Shibata    -4.6681
Hannan-Quinn -4.6537

Q-Statistics on Standardized Residuals
-----
      statistic  p-value
Lag[1]          0.07162  0.7890
Lag[p+q+1][1]  0.07162  0.7890
Lag[p+q+5][5]  8.23655  0.1437
d.o.f=0
H0 : No serial correlation

Q-Statistics on Standardized Squared Residuals
-----
      statistic  p-value
Lag[1]          0.03904  0.8434
Lag[p+q+1][3]  0.86179  0.3532
Lag[p+q+5][7]  4.80276  0.4404
d.o.f=2

ARCH LM Tests
-----
      Statistic  DoF  P-Value
ARCH Lag[2]    0.3551  2  0.8373
ARCH Lag[5]    2.9608  5  0.7060
ARCH Lag[10]   6.3514  10 0.7849

Nyblom stability test
-----
Joint Statistic: 1.8642
Individual Statistics:
mu      0.0787
omega   1.1866
alpha   0.1504
skew    0.2116
shape   0.1327

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic: 1.28 1.47 1.88
Individual Statistic: 0.35 0.47 0.75

Sign Bias Test
-----
      t-value  prob sig
Sign Bias      0.7630  0.4458
Negative Sign Bias 0.0477  0.9620
Positive Sign Bias 0.2887  0.7729
Joint Effect    2.2017  0.5316

Adjusted Pearson Goodness-of-Fit Test:
-----
      group  statistic  p-value (g-1)
1      20      11.73      0.8967
2      30      20.90      0.8628
3      40      21.73      0.9885
4      50      35.00      0.9342

```

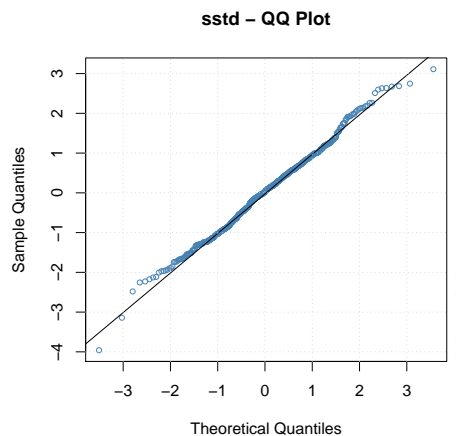
CSNA3 – Diagnóstico do Modelo IGARCH(1,1)

Test	Null.Hypothesis	pvalue	Rejected.0.01	Rejected.0.05
1 Kurtosis Test	Kurtosis == 3	0.0171	FALSE	TRUE
2 Skewness Test	Skewness == 0	0.3109	FALSE	FALSE
3 Jarque Bera Test	Normality	0.0886	FALSE	FALSE
4 LM Test for ARCH effects (lags=12)	No ARCH effects	0.9883	FALSE	FALSE
5 KPSS Test	Level stationarity	0.1000	FALSE	FALSE
6 Ljung-Box (lags=1)	No autocorrelation	0.3811	FALSE	FALSE
7 Ljung-Box (lags=2)	No autocorrelation	0.6093	FALSE	FALSE
8 Ljung-Box (lags=3)	No autocorrelation	0.6502	FALSE	FALSE
9 Ljung-Box (lags=4)	No autocorrelation	0.4348	FALSE	FALSE
10 Ljung-Box (lags=5)	No autocorrelation	0.3783	FALSE	FALSE
11 Ljung-Box (lags=6)	No autocorrelation	0.4835	FALSE	FALSE
12 Ljung-Box (lags=7)	No autocorrelation	0.5860	FALSE	FALSE

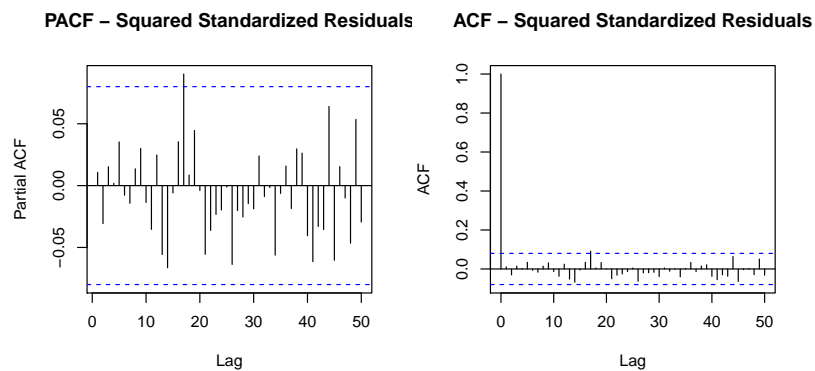
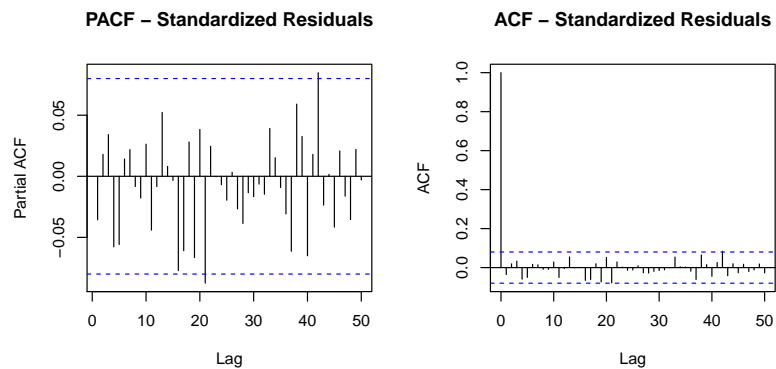
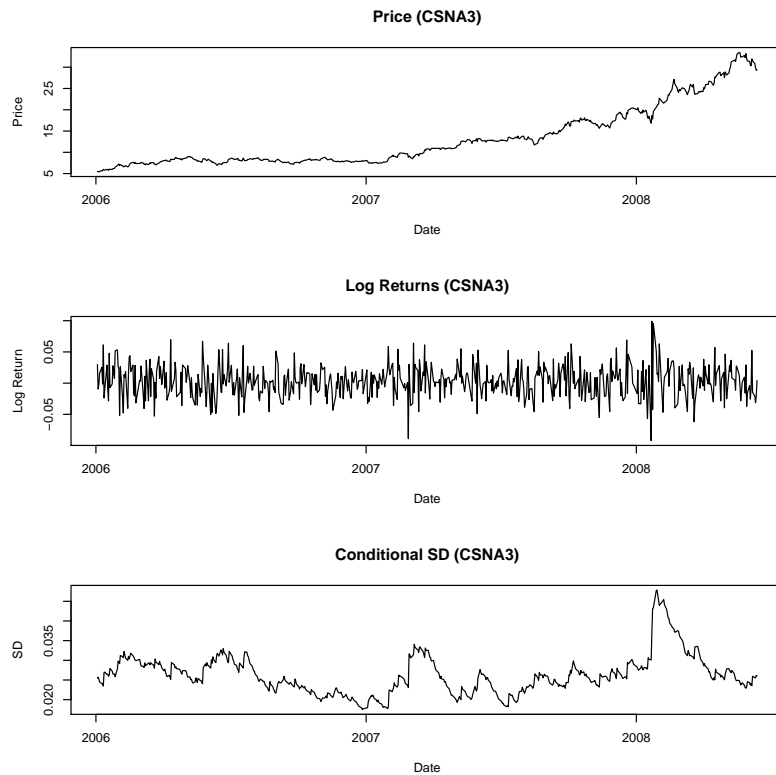
Tabela D.9: Testes para os resíduos padronizados CSNA3

Test	Null.Hypothesis	pvalue	Rejected.0.01	Rejected.0.05
1 Kurtosis Test	Kurtosis == 3	0.0000	TRUE	TRUE
2 Skewness Test	Skewness == 0	0.0000	TRUE	TRUE
3 Jarque Bera Test	Normality	0.0000	TRUE	TRUE
4 LM Test for ARCH effects (lags=12)	No ARCH effects	0.9998	FALSE	FALSE
5 KPSS Test	Level stationarity	0.1000	FALSE	FALSE
6 Ljung-Box (lags=1)	No autocorrelation	0.7932	FALSE	FALSE
7 Ljung-Box (lags=2)	No autocorrelation	0.7285	FALSE	FALSE
8 Ljung-Box (lags=3)	No autocorrelation	0.8584	FALSE	FALSE
9 Ljung-Box (lags=4)	No autocorrelation	0.9426	FALSE	FALSE
10 Ljung-Box (lags=5)	No autocorrelation	0.9147	FALSE	FALSE
11 Ljung-Box (lags=6)	No autocorrelation	0.9585	FALSE	FALSE
12 Ljung-Box (lags=7)	No autocorrelation	0.9754	FALSE	FALSE

Tabela D.10: Testes para o quadrado dos resíduos padronizados CSNA3



CSNA3 – Diagnóstico do Modelo IGARCH(1,1) – Gráficos



CSNA3 – Resultados do Ajuste do Modelo IGARCH(1,1)

```

*-----*
*          GARCH Model Fit          *
*-----*

Conditional Variance Dynamics
-----
GARCH Model : iGARCH(1,1)
Mean Model  : ARFIMA(0,0,0)
Distribution : sstd

Optimal Parameters
-----
      Estimate  Std. Error  t value  Pr(>|t|)
mu      0.002383   0.001003   2.37683  0.017462
omega   0.000004   0.000004   0.98712  0.323585
alpha   0.059489   0.025824   2.30361  0.021245
beta1   0.940511           NA           NA           NA
skew    1.010768   0.062901  16.06927  0.000000
shape   15.728459   9.660301   1.62815  0.103492

Robust Standard Errors:
      Estimate  Std. Error  t value  Pr(>|t|)
mu      0.002383   0.001017   2.34358  0.01910
omega   0.000004   0.000005   0.76511  0.44421
alpha   0.059489   0.037116   1.60276  0.10899
beta1   0.940511           NA           NA           NA
skew    1.010768   0.066917  15.10485  0.00000
shape   15.728459  10.372258   1.51640  0.12942

LogLikelihood : 1357.725

Information Criteria
-----
Akaike      -4.5091
Bayes      -4.4724
Shibata    -4.5092
Hannan-Quinn -4.4948

Q-Statistics on Standardized Residuals
-----
      statistic  p-value
Lag[1]          0.7672  0.3811
Lag[p+q+1][1]  0.7672  0.3811
Lag[p+q+5][5]  5.3183  0.3783
d.o.f=0
H0 : No serial correlation

Q-Statistics on Standardized Squared Residuals
-----
      statistic  p-value
Lag[1]          0.06869  0.7932
Lag[p+q+1][3]  0.76254  0.3825
Lag[p+q+5][7]  1.68077  0.8913
d.o.f=2

ARCH LM Tests
-----
      Statistic  DoF  P-Value
ARCH Lag[2]    0.6336   2  0.7285
ARCH Lag[5]    1.5123   5  0.9116
ARCH Lag[10]   2.6213  10  0.9890

Nyblom stability test
-----
Joint Statistic: 1.3801
Individual Statistics:
mu      0.09650
omega   0.41901
alpha   0.05696
skew    0.10930
shape   0.12326

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:      1.28 1.47 1.88
Individual Statistic: 0.35 0.47 0.75

Sign Bias Test
-----
      t-value  prob sig
Sign Bias      0.9688  0.3330
Negative Sign Bias 0.9993  0.3181
Positive Sign Bias 1.1635  0.2451
Joint Effect    2.4836  0.4783

Adjusted Pearson Goodness-of-Fit Test:
-----
      group  statistic  p-value (g-1)
1      20      28.07      0.08215
2      30      34.50      0.22143
3      40      42.67      0.31638
4      50      47.00      0.55457

```

Referências Bibliográficas

- [1] C. Acerbi and D. Tasche, *Expected Shortfall: a natural coherent alternative to Value at Risk*, May 2001.
- [2] Carol Alexander, *Principal component models for generating large GARCH covariance matrices*, *Economic Notes* **31** (2002), no. 2, 337–359.
- [3] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath, *Coherent Measures of Risk*, 1998.
- [4] Michael Beaghen, *Canonical Variate Analysis and Related Methods with Longitudinal Data*, Ph.D. thesis, Virginia Polytechnic Institute and State University, 1997.
- [5] D.P. Bertsekas, *Nonlinear programming*, Athena Scientific optimization and computation series, Athena Scientific, 1999.
- [6] J Frédéric Bonnans, *Numerical optimization: theoretical and practical aspects: with 26 figures*, Springer-Verlag New York Incorporated, 2003.
- [7] Herman Peter Boswijk and Roy Van der Weide, *Wake me up before you GO-GARCH*, (2006).
- [8] Simon A Broda and Marc S Paoletta, *CHICAGO: A fast and accurate method for portfolio risk calculation*, *Journal of Financial Econometrics* **7** (2009), no. 4, 412–436.
- [9] Peter Christoffersen and Denis Pelletier, *Backtesting value-at-risk: A duration-based approach*, *Journal of Financial Econometrics* **2** (2004), no. 1, 84–108.
- [10] Peter F Christoffersen, *Evaluating interval forecasts*, *International economic review* (1998), 841–862.
- [11] P. Comon, *Independent component analysis – a new concept?*, *Signal Processing* **36** (1994), no. 3, 287–314.
- [12] Rama Cont, Romain Deguest, and Giacomo Scandolo, *Robustness and sensitivity analysis of risk measurement procedures*, *Quantitative Finance* **10** (2010), no. 6, 593–606.
- [13] Thomas M Cover and Joy A Thomas, *Elements of information theory*, Wiley-interscience, 2012.
- [14] J. Danielsson, *Financial Risk Forecasting: The Theory and Practice of Forecasting Market Risk with Implementation in R and Matlab*, The Wiley Finance Series, Wiley, 2011.
- [15] Zhuanxin Ding and Robert Engle, *Large scale conditional covariance matrix modeling, estimation and testing*, (2001).

- [16] Jianqing Fan, Mingjin Wang, and Qiwei Yao, *Modelling multivariate volatilities via conditionally uncorrelated components*, Journal of the Royal Statistical Society: series B (statistical methodology) **70** (2008), no. 4, 679–702.
- [17] Carmen Fernández and Mark FJ Steel, *On Bayesian modeling of fat tails and skewness*, Journal of the American Statistical Association **93** (1998), no. 441, 359–371.
- [18] Cristiane Azevedo Ferreira, *Avaliação de Modelos de Risco através de Backtesting*, Master's thesis, Instituto Nacional de Matemática Pura e Aplicada, July 2013.
- [19] Bernard N Flury and Gregory Constantine, *Algorithm AS 211: The FG diagonalization algorithm*, Journal of the Royal Statistical Society. Series C (Applied Statistics) **34** (1985), no. 2, 177–183.
- [20] Bernhard N Flury, *Common principal components in k groups*, Journal of the American Statistical Association **79** (1984), no. 388, 892–898.
- [21] Bernhard N Flury and Walter Gautschi, *An algorithm for simultaneous orthogonal transformation of several positive definite symmetric matrices to nearly diagonal form*, SIAM J. Sci. Stat. Comput. **7** (1986), no. 1, 169–184.
- [22] Henrik Alsing Friberg, *Rmosek: The R-to-MOSEK Optimization Interface*, 2012, R package version 1.2.4.
- [23] Jim Gatheral, *The volatility surface: a practitioner's guide*, vol. 357, Wiley, 2006.
- [24] Alexios Ghalanos, *Introduction to the rugarch package*, January 2013.
- [25] J.D.A. Hamilton, *The Time Series Analysis*, Princeton Univers. Press, 1994.
- [26] Trevor Hastie, *A closer look at the deviance*, The American Statistician **41** (1987), no. 1, 16–20.
- [27] F. Hayashi, *Econometrics*, Princeton University Press, 2011.
- [28] John RM Hosking, *The multivariate portmanteau statistic*, Journal of the American Statistical Association **75** (1980), no. 371, 602–608.
- [29] A. Hyvärinen, *Fast and robust fixed-point algorithm for independent component analysis*, **10** (1999), no. 3, 626–634.
- [30] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, Adaptive and Learning Systems for Signal Processing, Communications and Control Series, Wiley, 2004.
- [31] A Hyvärinen and E Oja, *Independent component analysis: algorithms and applications*, Neural Netw **13** (2000), no. 4-5, 411–430.
- [32] Aapo Hyvärinen, *Survey on Independent Component Analysis*, 1999.
- [33] P. Jorion, *Value at risk: the new benchmark for managing financial risk*, vol. 2, McGraw-Hill New York, 2007.
- [34] Solomon Kullback and Richard A Leibler, *On information and sufficiency*, The Annals of Mathematical Statistics **22** (1951), no. 1, 79–86.
- [35] Paul Kupiec, *Techniques for verifying the accuracy of risk measurement models*, THE J. OF DERIVATIVES **3** (1995), no. 2.

- [36] Denis Kwiatkowski, Peter CB Phillips, Peter Schmidt, and Yongcheol Shin, *Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?*, *Journal of econometrics* **54** (1992), no. 1, 159–178.
- [37] Z.M. Landsman and E.A. Valdez, *Tail conditional expectations for elliptical distributions*, *North American Actuarial Journal* **7** (2003), no. 4, 55–71.
- [38] Greta M Ljung and George EP Box, *On a measure of lack of fit in time series models*, *Biometrika* **65** (1978), no. 2, 297–303.
- [39] Andriy Andreev-Antti Kanto-Pekka Malo, *On Closed-form calculation of CVAR*, (2005).
- [40] J L Marchini, C Heaton, and B D Ripley, *fastICA: FastICA Algorithms to perform ICA and Projection Pursuit*, 2012, R package version 1.1-16.
- [41] Harry Markowitz, *Portfolio selection**, *The journal of finance* **7** (1952), no. 1, 77–91.
- [42] A. Meucci, *Risk and asset allocation*, Springer finance, Springer-Verlag Berlin and Heidelberg GmbH & Company KG, 2005.
- [43] MOSEK, *The MOSEK Optimization Tools Manual*, (2012).
- [44] H Peter Boswijk and Roy van der Weide, *Method of moments estimation of GO-GARCH models*, *Journal of Econometrics* **163** (2011), no. 1, 118–126.
- [45] Giovanni Petris, Patrizia Campagnoli, and Sonia Petrone, *Dynamic linear models with R*, Springer, 2009.
- [46] Bernhard Pfaff, *VAR, SVAR and SVEC Models: Implementation Within R Package vars*, *Journal of Statistical Software* **27** (2008), no. 4.
- [47] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012, ISBN 3-900051-07-0.
- [48] TM Riskmetrics, *JP Morgan Technical Document*, 1996.
- [49] R. Rockafellar and S. Uryasev, *Optimization of Conditional Value-at-Risk*, *The Journal of Risk* (2000).
- [50] S.M. Ross, *Simulation*, Statistical Modeling and Decision Science, Elsevier Science, 2006.
- [51] Gideon Schwarz, *Estimating the dimension of a model*, *The annals of statistics* **6** (1978), no. 2, 461–464.
- [52] R.S. Tsay, *Analysis of Financial Time Series*, Wiley Series in Probability and Statistics, John Wiley & Sons, 2010.
- [53] Roy Van der Weide, *GO-GARCH: a multivariate generalized orthogonal GARCH model*, *Journal of Applied Econometrics* **17** (2002), no. 5, 549–564.
- [54] Dennis D Wackerly, William Mendenhall, and Richard L Scheaffer, *Mathematical statistics with applications*, CengageBrain. com, 2008.
- [55] Lingyu Zheng and William WS Wei, *Weighted scatter estimation method of the GO-GARCH models*, *Journal of Time Series Analysis* **33** (2011), no. 1, 81–95.

- [56] E. Zivot and J. Wang, *Modeling Financial Time Series With S-Plus*, International Federation for Information Processing, Springer, 2003.