

UNICAMP

UNIVERSIDADE ESTADUAL DE
CAMPINAS

Instituto de Matemática, Estatística e
Computação Científica

EDUARDO GABRIEL PINHEIRO

**Estimação Robusta de Modelos cDCC em Alta
Dimensão**

Campinas

2022

Eduardo Gabriel Pinheiro

Estimação Robusta de Modelos cDCC em Alta Dimensão

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Estatística.

Orientador: Luiz Koodi Hotta

Este trabalho corresponde à versão final da Dissertação defendida pelo aluno Eduardo Gabriel Pinheiro e orientada pelo Prof. Dr. Luiz Koodi Hotta.

Campinas
2022

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

P655e Pinheiro, Eduardo Gabriel, 1996-
Pinheiro. – Estimação robusta de modelos cDCC em alta dimensão / Eduardo Gabriel Pinheiro. – Campinas, SP : [s.n.], 2022.

Orientador: Luiz Koodi Hotta.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Verossimilhança composta. 2. Modelo de volatilidade multivariada. 3. Portfólio de grandes dimensões. 4. Portfólio de variância mínima. I. Hotta, Luiz Koodi, 1952-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Robust estimation of high-dimensional cDCC model

Palavras-chave em inglês:

Composite likelihood

Multivariate volatility model

Large dimension portfolio

Minimum variance portfolio

Área de concentração: Estatística

Titulação: Mestre em Estatística

Banca examinadora:

Luiz Koodi Hotta [Orientador]

Carlos César Trucíos Maza

Mauricio Enrique Zevallos Herencia

Data de defesa: 29-04-2022

Programa de Pós-Graduação: Estatística

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0003-2279-5601>

- Curículo Lattes do autor: <http://lattes.cnpq.br/4368383901487026>

**Dissertação de Mestrado defendida em 29 de abril de 2022 e aprovada
pela banca examinadora composta pelos Profs. Drs.**

Prof(a). Dr(a). LUIZ KOODI HOTTA

Prof(a). Dr(a). CARLOS CÉSAR TRUCÍOS MAZA

Prof(a). Dr(a). MAURICIO ENRIQUE ZEVALLOS HERENCIA

A Ata da Defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós-Graduação do Instituto de Matemática, Estatística e Computação Científica.

À minha Família

Agradecimentos

Agradeço a Deus pela oportunidade, força e saúde.

Aos meus pais Francisco de Castro Pinheiro e Jonailda Gabriel Pinheiro que sempre tiveram como prioridade o bem dos filhos, estando presentes para compartilhar dos sucessos e também consolar nos fracassos.

Ao meu irmão Felipe Gabriel Pinheiro que sempre me ajudou em todas as minhas dificuldades, o primeiro e melhor amigo da minha vida.

Ao Professor Dr. Luiz K. Hotta pela orientação e paciência no decorrer do desenvolvimento desta dissertação.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

*“Quem controla o passado controla o futuro.
Quem controla o presente, controla o passado
(George Orwell, 1984)*

Resumo

Volatilidades de retornos de ativos são importantes para tomada de decisão em diferentes áreas de finanças econômétricas. Na literatura de modelagem estatística o modelo cDCC (*corrected dynamic conditional correlation*) é utilizado com frequência para aplicações desta natureza. Com o aumento do número de ativos, técnicas como a verossimilhança composta são utilizados para viabilizar a estimação em dados de alta dimensão. No entanto, na presença de *outliers* aditivos este estimador se torna severamente viciado. Esta dissertação apresenta um método robusto de alta dimensão resultante da combinação de diferentes metodologias presentes na literatura. Os resultados do estudo de simulação mostraram que o estimador proposto produz estimativas com menor viés e gera portfólios de mínima variância com menores variâncias. Isso foi confirmado na aplicação, que envolve retornos de S&P500, em que os estimadores foram comparados considerando rebalanceamentos diários e mensais. Em geral, o método robusto produziu portfólios de menor variância na maioria dos anos, inclusive nos de crise.

Palavras-chave: Verossimilhança composta; Modelo de volatilidade multivariada; Portfólio de grandes dimensões; Portfólio de variância mínima.

Abstract

Asset return volatility is important for decision making in different areas of econometric finance. In the statistical modeling literature, the cDCC (corrected dynamic conditional correlation) model is frequently used for applications of this nature. With the increase in the number of assets, techniques such as composite likelihood are used to enable to estimate high-dimensional data. However, in the presence of additive outliers this estimator becomes severely biased. This dissertation presents a robust high-dimensional method resulting from the combination of different methodologies present in the literature. The results of the simulation study showed that the proposed estimator produces less biased estimates and generates minimum variance portfolios with lower variances. This was confirmed in the application, which involves returns from S&P500, in which the estimators were compared considering daily and monthly rebalancing. In general, the robust method produced portfolios with lower variance in most years, including crisis years.

Keywords: Composite likelihood. Multivariate volatility model. Large dimension portfolio. Minimum variance portfolio.

Lista de ilustrações

Figura 1 – Estimativas de $\boldsymbol{\eta} = (0, 10, 0, 10, 0, 80)$ do modelo GARCH(1, 1) obtidas por $\hat{\boldsymbol{\eta}}_q$ e $\hat{\boldsymbol{\eta}}_r$. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\boldsymbol{\eta}}_q$ e $\hat{\boldsymbol{\eta}}_r$, respectivamente. Na figura superior estimativas do parâmetro α_1 e na inferior de β_1 . Linha tracejada indica real valor do respectivo parâmetro. $T = 2.000, \delta = 0,975$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.	56
Figura 2 – Distribuições das persistências estimadas $(\hat{\alpha}_i + \hat{\beta}_i)$ pelos estimadores tradicional e robusto. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\theta}_q$ e $\hat{\theta}_r$, respectivamente. Em (A) os resultados para $\boldsymbol{\eta}_1 = (0, 10, 0, 10, 0, 80)$ e (B) os resultados para $\boldsymbol{\eta}_2 = (0, 10, 0, 20, 0, 70)$. $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas.	57
Figura 3 – Viés e REQM (linhas) obtidos por $\hat{\eta}_q$ e $\hat{\eta}_r$ para parâmetros α_1 e β_1 (colunas). $T = 2.000, \delta = 0,975$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.	59
Figura 4 – Estimativas de $\phi = (0, 10, 0, 80)$ de modelo cDCC(1, 1) obtidas por $\hat{\phi}_q$ e $\hat{\phi}_r$. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\phi}_q$ e $\hat{\phi}_r$, respectivamente. No eixo vertical, estimativas para parâmetros a e b . Linha tracejada indica real valor do respectivo parâmetro. $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.	60
Figura 5 – Distribuições das persistências estimadas para $\phi = (0, 10, 0, 80)$ de modelo cDCC(1, 1) obtidas por $\hat{\phi}_q$ e $\hat{\phi}_r$. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\phi}_q$ e $\hat{\phi}_r$, respectivamente. Linha tracejada indica valor real de $(a+b)$. $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas.	61
Figura 6 – Viés, REQM e erro padrão (colunas) obtidos por $\hat{\phi}_q$ e $\hat{\phi}_r$ para os parâmetros a e b (linha), para $T = 2.000, \delta = 0,975$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.	62
Figura 7 – IPV por modelo. Resultados para saltos consecutivos e não consecutivos. $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.	64

Figura 8 – Razão entre as variâncias dos portfólios de variância miníma estimado e real. Resultados para saltos consecutivos e não consecutivos. $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.	64
Figura 9 – Norma de Frobenius. Resultados para saltos consecutivos e não consecutivos. $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.	65
Figura 10 – Estimativas para o modelo cDCC(1, 1) obtidas por $\hat{\phi}_q$ e $\hat{\phi}_r$. Contaminação com saltos isolados, $T = 1.250, \delta = 0,975, N = 1.000$ e 100 réplicas. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\phi}_q$ e $\hat{\phi}_r$, respectivamente. No eixo vertical, estimativas para parâmetros a e b . Em (A), (B), (C) e (D) resultados para $a = 0,10, a = 0,05, b = 0,80$ e $b = 0,93$, respectivamente. Linha tracejada indica real valor do respectivo parâmetro. Os saltos são de magnitude d vezes o desvio padrão condicional.	67
Figura 11 – Estimativas para o modelo cDCC(1, 1) obtidas por $\hat{\phi}_q$ e $\hat{\phi}_r$. Contaminação com saltos consecutivos, $T = 1.250, \delta = 0,975, N = 1.000$ e 100 réplicas. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\phi}_q$ e $\hat{\phi}_r$, respectivamente. No eixo vertical, estimativas para parâmetros a e b . Em (A), (B), (C) e (D) resultados para $a = 0,10, a = 0,05, b = 0,80$ e $b = 0,93$, respectivamente. Linha tracejada indica real valor do respectivo parâmetro. Os saltos são de magnitude d vezes o desvio padrão condicional.	68
Figura 12 – Distribuições das persistências estimadas para modelo cDCC(1, 1) obtidas por $\hat{\phi}_q$ e $\hat{\phi}_r$. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\phi}_q$ e $\hat{\phi}_r$, respectivamente. Linha tracejada indica valor real de $(a + b)$. Em (A) e (B) estão os resultados para $\phi = (0,10,0,80)'$ e $\phi = (0,05,0,93)'$, respectivamente. $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas.	71
Figura 13 – IPV por modelo. Resultados para saltos consecutivos e isolados. $T = 1.250, \delta = 0,975$ e 100 réplicas. Em (A) e (B) os resultados para $\phi = (0,10,0,80)$ e $(0,05,0,93)$, respectivamente. Os saltos são de magnitude d vezes o desvio padrão condicional.	73

Figura 14 – Razão entre as variâncias dos portfólios de variância mínima estimado e real. Resultados para saltos consecutivos e isolados. $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Em (A) e (B) os resultados para $\phi = (0, 10, 0, 80)$ e $(0, 05, 0, 93)$, respectivamente. Os saltos são de magnitude d vezes o desvio padrão condicional.	74
Figura 15 – Norma de Frobenius. Resultados para saltos consecutivos e isolados. $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Em (A) e (B) os resultados para $\phi = (0, 10, 0, 80)$ e $(0, 05, 0, 93)$, respectivamente. Os saltos são de magnitude d vezes o desvio padrão condicional.	74
Figura 16 – Ativos considerados na construção de portfólios com rebalanceamento diário ao longo dos anos. Em preto se o respectivo ativo foi considerado na produção do MVP na respectiva data e em branco caso contrário.	76
Figura 17 – Estimativas produzidas pelos métodos tradicional (Q) (figura superior) robusto (R) (figura inferior) para rebalanceamento diário. À esquerda estimativas para o parâmetro a e à direita para parâmetro b . Em destaque, anos de crise.	79
Figura 18 – Distribuição conjunta das estimativas de a e b obtidas pelos métodos tradicional (Q) e robusto (R) na atividade de estimação do MVP para séries de retornos do S&P500 considerando rebalanceamentos diárioss. Em destaque, anos de crise.	79
Figura 19 – Histogramas para estimativas de $\eta_1 = (0, 10, 0, 10, 0, 80)$ do modelo GARCH(1,1) obtidas por $\hat{\eta}_{q,1}$ (Q) e $\hat{\eta}_{r,1}$ (R). A linha tracejada indica valor verdadeiro do respectivo parâmetro. $T = 2.000$, $\delta = 0,975$ e 10.000 réplicas.	94
Figura 20 – Histogramas para estimativas de $\eta_2 = (0, 10, 0, 20, 0, 70)$ do modelo GARCH(1,1) obtidas por $\hat{\eta}_{q,2}$ (Q) e $\hat{\eta}_{r,2}$ (R). A Linha tracejada indica valor verdadeiro do respectivo parâmetro. Linha tracejada indica real valor do respectivo parâmetro. $T = 2.000$, $\delta = 0,975$ e 10.000 réplicas.	95
Figura 21 – Histogramas para estimativas de $\phi = (0, 10, 0, 80)$ de modelo cDCC(1, 1) obtidas por $\hat{\phi}_q$ e $\hat{\phi}_r$. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\phi}_q$ e $\hat{\phi}_r$, respectivamente. No eixo vertical, estimativas para parâmetros a e b . Linha tracejada indica real valor do respectivo parâmetro. $T = 2.000$, $\delta = 0,975$, $N = 2$ e 10.000 réplicas.	96
Figura 22 – Viés, REQM e erro padrão (colunas) obtidos por $\hat{\phi}_q$ e $\hat{\phi}_r$ para os parâmetros a e b (linha). Contaminação com saltos isolados, $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Resultados para $\phi = (0, 10, 0, 80)$. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.	97

- Figura 23 – Viés, REQM e erro padrão (colunas) obtidos por $\hat{\phi}_q$ e $\hat{\phi}_r$ para parâmetros a e b (linha). Contaminação com saltos consecutivos, $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Resultados para $\phi = (0, 10, 0, 80)$. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional. 98
- Figura 24 – Viés, REQM e erro padrão (colunas) obtidos por $\hat{\phi}_q$ e $\hat{\phi}_r$ para parâmetros a e b (linha). Contaminação com saltos isolados, $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Resultados para $\phi = (0, 05, 0, 93)$. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional. 99
- Figura 25 – Viés, REQM e erro padrão (colunas) obtidos por $\hat{\phi}_q$ e $\hat{\phi}_r$ para parâmetros a e b (linha). Contaminação com saltos consecutivos, $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Resultados para $\phi = (0, 05, 0, 93)$. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional. 100

Lista de tabelas

<p>Tabela 1 – Viés e REQM para estimadores de $\alpha_1 = 0, 10$ e $\beta_1 = 0, 8$ de modelo GARCH(1,1). $T = 2.000, \delta = 0, 975$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ. Os saltos são de magnitude d vezes o desvio padrão condicional.</p> <p>Tabela 2 – Métricas de desempenho para estimadores de $a = 0, 10$ e $b = 0, 80$ de modelo cDCC(1,1). $T = 2.000, \delta = 0, 975, N = 2$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ. Os saltos são de magnitude d vezes o desvio padrão condicional.</p> <p>Tabela 3 – Acurácia do filtro de volatilidade para etapa cDCC bidimensional. $T = 2.000, \delta = 0, 975, N = 2$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ. Os saltos são de magnitude d vezes o desvio padrão condicional.</p> <p>Tabela 4 – Métricas de desempenho para estimadores de a e b de modelo cDCC(1,1). Contaminação com saltos isolados, $T = 1.250, \delta = 0, 975, N = 1.000$ e 100 réplicas. Em negrito, para cenário de contaminação, a melhor performance em termos de viés e REQM. O percentual de tempos com contaminação é denotado por ϵ. Os saltos são de magnitude d vezes o desvio padrão condicional.</p> <p>Tabela 5 – Métricas de desempenho para estimadores de a e b de modelo cDCC(1,1). Contaminação com saltos consecutivos, $T = 1.250, \delta = 0, 975, N = 1.000$ e 100 réplicas. Em negrito, para cenário de contaminação, a melhor performance em termos de viés e REQM. O percentual de tempos com contaminação é denotado por ϵ. Os saltos são de magnitude d vezes o desvio padrão condicional.</p> <p>Tabela 6 – Desempenho do filtro de volatilidade na etapa cDCC de alta dimensão. $T = 1.250, N = 1.000, \delta = 0, 975$ e 100 réplicas. O percentual de tempos com contaminação é denotado por ϵ. Os saltos são de magnitude d vezes o desvio padrão condicional.</p>	<p>58</p> <p>62</p> <p>63</p> <p>69</p> <p>70</p> <p>72</p>
--	---

Tabela 7 – Medidas de desempenho dos portfólios selecionados pelos métodos tradicional (Q) e robusto (R) com rebalanceamento mensal. Em negrito, para cada ano, a melhor performance em termos de média (M), desvio padrão (DP) e quando a diferença entre os índices de informação (ΔIR) tradicional e robusto é positiva. M e DP são porcentagens anualizadas. Na coluna DP(R) (*), (**) e (***) indicam significância a um nível de 10%, 5% e 1%, respectivamente, para o teste de igualdade de variâncias por <i>bootstrap</i> entre os métodos robusto e tradicional. Linhas sombreadas indicam anos com crise econômica.	80
Tabela 8 – Medidas de desempenho dos portfólios selecionados pelo métodos tradicional (Q) e robusto (R) com rebalanceamento diário. Em negrito, para cada ano, a melhor performance em termos de média (M), desvio padrão (DP) e quando a diferença entre os índices de informação (ΔIR) tradicional e robusto é positiva. M e DP são porcentagens anualizadas. Na coluna DP(R) (*), (**) e (***) indicam significância a um nível de 10%, 5% e 1%, respectivamente, para o teste de igualdade de variâncias por <i>bootstrap</i> entre os métodos robusto e tradicional. Linhas sombreadas indicam anos com crise econômica. O (ΔIR^{net}) é apresentado para custos de transação de c pontos base.	81
Tabela 9 – Média, desvio padrão e quantis empíricos de IPV calculado para portfólios estimados sob método tradicional (Q) e robusto (R). $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas.	89
Tabela 10 – Média, desvio padrão e quantis empíricos de razão entre variância do portfólio de variância mínima estimado calculada sob método tradicional (Q) e robusto (R). $T = 2000, \delta = 0,975, N = 2$ e 10.000 réplicas.	90
Tabela 11 – Média, desvio padrão e quantis empíricos de norma de Frobenius calculada para os portfólios estimados sob os métodos tradicional (Q) e robusto (R). $T = 2000, \delta = 0,975, N = 2$ e 10.000 réplicas.	90
Tabela 12 – Média, desvio padrão e quantis empíricos de IPV calculado para portfólios estimados sob método tradicional (Q) e robusto (R). $T = 1250, \delta = 0,975, N = 1.000$ e 100 réplicas.	91
Tabela 13 – Média, desvio padrão e quantis empíricos da razão entre as variâncias dos portfólios de variância mínima estimados pelos métodos tradicional (Q) e robusto (R). $T = 1.250, \delta = 0,975, N = 1.000$ e 100 réplicas.	92
Tabela 14 – Média, desvio padrão e quantis empíricos da norma de Frobenius calculada para os portfólios estimados sob os métodos tradicional (Q) e robusto (R). $T = 1.250, \delta = 0,975, N = 1.000$ e 100 réplicas.	93

Sumário

1	INTRODUÇÃO	18
2	MODELO GARCH	20
2.1	Especificação do modelo	20
2.2	<i>Outliers</i>	22
2.3	Estimação dos parâmetros e volatilidade	23
2.4	Estimação robusta dos parâmetros e volatilidade	25
2.4.1	Estimador QMVT	25
2.4.2	Estimador-M para o modelo BIP-GARCH	26
2.4.3	Estimadores robustos para média e variância não condicionais	28
3	MODELOS MGARCH	29
3.1	Modelo DCC	30
3.2	Modelo cDCC	32
3.2.1	Condições de estacionaridade	33
3.3	Modelo cDCC contaminado	35
3.4	Estimação dos parâmetros e correlações condicionais	35
3.5	Estimação robusta dos parâmetros e correlações condicionais	37
3.5.1	Estimador de correlação não condicional reponderada	38
3.5.2	Estimador-M para o modelo BIP-cDCC	39
4	ESTIMAÇÃO EM ALTA DIMENSÃO	41
4.1	Definição e estimação dos modelos	41
4.2	Matriz de correlação condicional completa	44
4.2.1	Estimador Shrinkage	45
4.2.2	Estimador Wrapping	47
5	SIMULAÇÃO PARA MODELOS GARCH E CDCC BIVARIADO	49
5.1	Processos geradores de dados	49
5.2	Estimadores	50
5.3	Medidas de desempenho	52
5.3.1	Estimação dos parâmetros	52
5.3.2	Desempenho do filtro de volatilidade	53
5.3.3	Portfólio de variância globalmente mínima	53
5.4	Estimação para modelos GARCH	55
5.5	Estimação do modelo cDCC	59

5.5.1	Estimação dos parâmetros	59
5.5.2	Estimação da variância condicional	61
5.5.3	Desempenho do filtro de volatilidade	63
5.6	Conclusão	63
6	SIMULAÇÃO EM ALTA DIMENSÃO	66
6.1	Processo gerador de dados	66
6.2	Estimação dos parâmetros	67
6.3	Estimação da matriz de variância condicional	68
6.4	Desempenho do filtro de volatilidade	71
6.5	Conclusão	72
7	APLICAÇÃO	75
7.1	Resultados	77
7.2	Conclusão	78
8	CONSIDERAÇÕES FINAIS	82
	REFERÊNCIAS	83
	APÊNDICE A – TABELAS COMPLEMENTARES DOS CAPÍTULOS 5 E 6	89
	APÊNDICE B – GRÁFICOS COMPLEMENTARES DOS CAPÍTULOS 5, 6 E 7	94
	APÊNDICE C – TESTE DE IGUALDADE DE VARIÂNCIAS DE PORTFÓLIOS	101

1 Introdução

Em diversas aplicações de econometria financeira é importante compreender a variabilidade e estrutura de dependência existente em um conjunto de ativos, pois esse conhecimento poderá ser usado para auxiliar na tomada de decisão de diversas áreas como precificação de ativos, alocação de portfólio, gerenciamento de riscos, etc. Além disso, é comum em séries de retornos financeiros que variâncias e covariâncias condicionais sejam dinâmicas no tempo. Diante disso, um dos modelos mais populares para estimação e predição de variância condicionais, proposto por [Bollerslev \(1986\)](#), é o modelo de heteroscedasticidade autoregressiva condicional generalizada (GARCH, *generalized autoregressive conditional heteroskedasticity*). Já, para modelagem das covariâncias condicionais o modelo de correlação condicional dinâmica (DCC, *dynamic conditional correlation*) [Engle \(2002\)](#) ou DCC corrigido (cDCC) de [Aielli \(2013\)](#) estão entre os mais utilizados, os quais fazem parte da classe de modelos GARCH multivariados (MGARCH). Veja, [Bauwens, Laurent e Rombouts \(2006\)](#), [Almeida, Hotta e Ruiz \(2018\)](#) e [Boudt et al. \(2019\)](#) para uma revisão de modelos MGARCH.

Em séries de retornos financeiros existem eventos isolados que podem resultar em pontos atípicos (*outliers*). Em geral, espera-se que isto não impacte de forma significativa as volatilidades futuras e/ou a estrutura de dependências com outras séries. O mesmo vale para saltos conjuntos. No entanto, a presença desses saltos chegam a viciar os estimadores das volatilidades os preditores das correlações dos modelos GARCH e cDCC. Motivado por isso, [Muler e Yohai \(2008\)](#) propõem uma alternativa robusta em que o impacto das observações no modelo GARCH é limitado (BIP-GARCH). O termo BIP refere-se a *bounded innovation propagation*. [Boudt, Danielsson e Laurent \(2013\)](#) estendem essa solução para o caso multivariado, especificamente para o modelo cDCC (BIP-cDCC). Algumas alterações desses métodos tem sido desenvolvidas em [Trucíos, Hotta e Ruiz \(2018\)](#) e [Trucíos, Hotta e Ruiz \(2017\)](#).

Na prática é comum considerar centenas de ativos em alocação de portfólio e gerenciamento de riscos. Para grandes bases de dados estimar as correlações condicionais torna-se inviável por métodos tradicionais tanto por produzir previsões consideravelmente viciadas quanto no custo computacional atrelado; fenômeno conhecido como a maldição da dimensionalidade. [Pakel et al. \(2021\)](#) propõem um método, uma solução baseada no método da máxima verossimilhança composta considerando pares de séries consecutivas aplicável a modelos MGARCH, que resulta em estimativas menos viciadas para os parâmetros e sendo computacionalmente mais barato. Basicamente, a solução reduz um problema de dimensão N (número de ativos) a vários de menor dimensão. Considerando a estrutura do modelo cDCC esta técnica trata da etapa de estimação dos parâmetros, mas ainda

é necessário estimar a matriz de correlações não condicionais na dimensão original, um possível estimador é dado por sua respectiva quantidade amostral, mas esta tende a ser viesada em dimensões elevadas como investigado em [Ledoit e Péché \(2011\)](#). [Ledoit e Wolf \(2012\)](#) propõem o estimador *Shrinkage* não linear que aprimora a estimativa de matrizes de correlações em alta dimensão. [Engle, Ledoit e Wolf \(2019\)](#) propõem um estimador de alta dimensão para o modelo DCC resultante da combinação com diferentes metodologias já existentes na literatura: o método da máxima verossimilhança composta e *Shrinkage* não linear. O objetivo desta dissertação é robustecer o estimador de [Engle, Ledoit e Wolf \(2019\)](#) a *outliers* aditivos, mas considerando a versão corrigida de [Aielli \(2013\)](#).

Os métodos de simulação e estimativa, para GARCH, BIP-GARCH, cDCC e BIP-cDCC, utilizados nessa dissertação foram implementados utilizando o *software R* ([R Core Team, 2021](#)). Para reduzir o tempo computacional na simulação, estimativa dos parâmetros e da cálculo da matriz de correlação condicional as rotinas foram implementadas usando os pacotes *Rcpp* ([EDDELBUETTEL; BALAMUTA, 2018](#)) e *RcppArmadillo* ([EDDELBUETTEL; SANDERSON, 2014](#)) que integram R e C++. As estimativas para matriz de correlação não condicional por *Shrinkage* não linear foram obtidas por meio do pacote *nlshrink* ([RAMPRASAD, 2016](#)). As figuras foram produzidas usando o pacote *ggplot2* ([WICKHAM, 2016](#)).

O restante da dissertação é organizado da seguinte forma. No Capítulo 2 é apresentado o modelo GARCH; comenta-se sobre o método de estimativa por quase-máxima verossimilhança e como a contaminação por *outliers* aditivos pode impactar a estimativa dos parâmetros e consequentemente as volatilidades preditas. Posteriormente, apresenta-se um revisão sobre o modelo robusto BIP-GARCH, que utiliza um filtro de volatilidade na identificação de *outliers* além de possuir uma função de verossimilhança mais robusta. No Capítulo 3 são apresentados os modelos cDCC (tradicional) e BIP-cDCC (robusto), comentando sobre os respectivos métodos de estimativa e uma alternativa robusta para estimativa da matriz de variâncias e covariâncias não condicionais. O Capítulo 4 apresenta métodos específicos para modelagem de dados de alta dimensão: máxima verossimilhança composta, *Shrinkage* e *Wrapping*. Nos Capítulos 5 e 6 são apresentados resultados de estudos de simulação de Monte Carlo para séries bivariadas e de alta dimensão, respectivamente; os estimadores são comparados em diferentes graus de contaminação por *outliers* aditivos com respeito a estimativa dos parâmetros e da matriz de variâncias e covariâncias condicionais para construção do portfólio de variância mínima. Além disso, há uma revisão do desempenho do filtro de volatilidade na identificação dos saltos. No Capítulo 7 é apresentada uma aplicação considerando retornos de ativos que compõem o S&P500 negociados ao longo de 20 anos. Duas estratégias de rebalanceamento são consideradas: diárias e mensais. Por fim, no Capítulo 8 são apresentadas conclusões e considerações para trabalhos futuros.

2 Modelo GARCH

O desenvolvimento de modelos de volatilidade em tempo discreto foi acompanhado da necessidade de ajustar séries temporais financeiras que são caracterizadas por atributos como caudas pesadas, assimetria, aglomerados de volatilidade e dependência serial com correlação próxima, ou igual a zero, que não podem ser explicadas por modelos lineares tradicionais como ARIMA (BROCKWELL; DAVIS, 2016, p. 196). O modelo autoregressivo de heteroscedasticidade condicional (*autoregressive conditional heteroskedasticity*, ARCH) proposto por Engle (1982) foi o primeiro destinado a contemplar um modelo com estas propriedades. Uma explicação mais detalhada sobre a criação deste pode ser consultada em Engle (2004). Basicamente a motivação era encontrar um modelo para validar uma conjectura de Friedman (1977) de que a imprevisibilidade da inflação é causada principalmente pelos ciclos de negócios. A hipótese era de que incerteza sobre custos e preços futuros, e não a inflação, levariam empreendedores a não investir, resultando em recessão, que só seria plausível se a incerteza variasse com o tempo, conhecida também por heteroscedasticidade.

O modelo ARCH descreve a variância condicional futura um passo à frente como a média ponderada dos quadrados das observações presente e passadas. Posteriormente, surgiram diversas generalizações, uma das mais famosas é o modelo GARCH (*generalized autoregressive conditional heteroskedasticity*) proposto por Bollerslev (1986) que essencialmente incorpora a variância condicional atual (e passadas) no cálculo da variância condicional futura um passo à frente.

2.1 Especificação do modelo

A série de retornos r_1, \dots, r_T é gerada por um processo GARCH(p, q) centralizado se

$$r_t = \sqrt{h_t} \epsilon_t, \quad (2.1)$$

em que $\epsilon_1, \dots, \epsilon_T$ são variáveis aleatórias independentes e identicamente distribuídas (i.i.d) com função densidade de probabilidade (fpd) f com valor esperado e variância $E(\epsilon_t) = 0$ e $var(\epsilon_t) = 1$, respectivamente. Seja \mathcal{F}_{t-1} a informação disponível até o tempo $(t-1)$; é assumido que

$$h_t = var(r_t | \mathcal{F}_{t-1}) = \omega + \sum_{i=1}^p \alpha_i r_{t-i}^2 + \sum_{i=1}^q \beta_i h_{t-i}, \quad p \geq 0, \quad q > 0 \quad (2.2)$$

sendo $\alpha_i \geq 0$; $1 \leq i \leq p$; $\beta_i \geq 0$; $1 \leq j \leq q$ e $\omega > 0$. A variância condicional (ou volatilidade ao quadrado) de r_t , denotada por h_t , é expressa como a combinação linear do quadrado

das observações passadas e das variâncias condicionais passadas. Para $q = 0$ o processo é reduzido a um ARCH(p) e para $p = q = 0$ têm-se variância condicional constante, i.e., um ruído branco pois, como apresentado em (2.5) existe correlação serial nula. Uma condição necessária e suficiente para que o processo GARCH(p, q) definido em (2.2) seja estritamente estacionário para o momento de segunda ordem finito é

$$\sum_{i=1}^p \alpha_i + \sum_{i=1}^q \beta_i < 1. \quad (2.3)$$

veja [Bollerslev \(1986\)](#), [Nelson \(1990\)](#), [Bougerol e Picard \(1992\)](#). Dado isso

$$E(r_t) = 0, \quad (2.4)$$

$$\text{cov}(r_t, r_r) = 0, \quad \text{para } t \neq r \quad (2.5)$$

$$h = E(r_t^2) = \frac{\omega}{1 - (\sum_{i=1}^p \alpha_i + \sum_{i=1}^q \beta_i)}. \quad (2.6)$$

Sob (2.3), considerando que o processo iniciou-se no tempo $-\infty$ (na prática, em um passado suficientemente distante), a forma explícita para (2.2) é dada por

$$h_t = \frac{\omega}{1 - \sum_{j=1}^q \beta_j} + \sum_{i=1}^p \alpha_i r_{t-i}^2 + \sum_{i=1}^p \alpha_i \sum_{k=1}^{\infty} \sum_{j_1=1}^q \dots \sum_{j_k=1}^q \beta_{j_1} \dots \beta_{j_k} r_{t-i-j_1-\dots-j_k}^2. \quad (2.7)$$

Como α_i e β_j são não negativos, e $E(h_t) < \infty$ então (2.7) converge com probabilidade 1 ([HALL; YAO, 2003](#)).

O modelo mais frequentemente utilizado na prática é o processo GARCH(1,1), que será utilizado na dissertação, exceto quando mencionado explicitamente; o mesmo vale para a omissão dos índices em α_i e β_i . O modelo é definido pela equação de observação (2.1) e pela equação de volatilidade

$$h_t = \omega + \alpha r_{t-1}^2 + \beta h_{t-1}, \quad (2.8)$$

com $\omega > 0$, α e $\beta \geq 0$; $(\alpha + \beta) < 1$, que garante estacionariedade de segunda ordem do processo ([BOLLERSLEV, 1986](#)). Para uma revisão sobre aplicações envolvendo metodologias ARCH veja [Bollerslev, Chou e Kroner \(1992\)](#).

[Nelson \(1990\)](#) mostra que para $\omega > 0$ e $h_t < \infty$, $\{r_t, h_t\}$ é estritamente estacionário e ergótico se e somente se

$$E(\ln(\beta + \alpha \epsilon_t^2)) < 0. \quad (2.9)$$

O termo $(\alpha + \beta)$ quantifica a persistência de h_t , i.e., a velocidade com que as volatilidades futuras convergem para a variância não condicional. Para soma próxima a 1 as variâncias condicionais passadas permanecem relevantes por mais tempo no cálculo das previsões de volatilidades futuras; enquanto para soma igual a 1 obtém-se um processo

de raiz unitária que se comporta como um passeio aleatório; com soma maior do que 1 têm-se $E(r_t^2) = \infty$ (VERHOEVEN; MCALLEER, 2000).

Bollerslev (1986) apresenta uma condição suficiente para a existência do quarto momento de r_t que é dada por

$$k\alpha^2 + 2\alpha\beta + \beta^2 < 1,$$

sendo k o quarto momento condicional de ϵ_t .

O modelo GARCH pode ser escrito como um ARCH de ordem infinita, i.e

$$h_t = \omega + \sum_{i=1}^{\infty} \alpha_i r_{t-i}^2.$$

Essa relação pode ser útil por permitir aproveitar propriedades do modelo ARCH na prova, para o GARCH, da existência de momentos e memória longa; como apresentado em Giraitis, Kokoszka e Leipus (2000).

2.2 Outliers

Em análise de séries temporais existem vários tipos de *outliers*, sendo os aditivos e os de inovação os mais comuns. O primeiro afeta uma única observação como um erro de digitação ou coleta, por exemplo. O segundo contamina um conjunto de observações e tende a impactar, com efeito de decaimento gradual, a volatilidade devido a perturbações externas à série observada (HOTTA; TSAY, 2012). Segundo Franses e Ghysels (1999), negligenciar *outliers* aditivos pode resultar em previsões viciadas para as volatilidades.

Como apresentado em Boudt, Danielsson e Laurent (2013), considere que a série de retornos observada y_t centralizada, pode ser escrita em função do componente GARCH normal definido em (2.1) e um componente de salto a_t , isto é

$$y_t = r_t + a_t. \quad (2.10)$$

Quando não há pontos atípicos (ou seja, $a_t = 0 \quad \forall t$) em (2.10), y_t é diretamente gerada por um processo GARCH(1,1); sendo $\boldsymbol{\eta}$, o conjunto de todos os parâmetros desconhecidos, usualmente estimado por (quase) máxima verossimilhança (entraremos em detalhes sobre o processo de estimação na Seção 2.3). Na presença de *outliers*, $a_t \neq 0$ para algum $t \in \{1, \dots, T\}$ de modo que assumir um processo GARCH(1,1) para y_t implicaria que

$$h_{t+1} = \omega + \alpha(y_t + a_t)^2 + \beta h_t,$$

ou seja, a_t tem impacto com decaimento lento na previsão de h_{t+1} , efeito não esperado de um *outlier* aditivo. Além disso, com $E(a_t) \neq 0$ é esperado que as estimativas para $\boldsymbol{\eta}$ sejam

viciadas (BOUDT; DANIELSSON; LAURENT, 2013); podendo impactar de forma não linear as volatilidades estimadas (CARNERO; PEÑA; RUIZ, 2012). Por meio de estudos de Monte Carlo Mendes (2000) quantifica a falta de robustez do EQMVn, demonstrando que uma pequena quantidade de *outliers* pode resultar em estimativas altamente viesadas.

Na presença de *outliers* a série de retornos observada pode ser especificada por um modelo GARCH contaminado definido em Hotta e Tsay (2012) que é dado por

$$\begin{aligned} y_t &= r_t + \text{sign}(r_t)a_t I(t \in A) \\ r_t | \mathcal{F}_{t-1} &\sim \mathcal{N}(0, h_t) \\ h_t &= \omega + \alpha r_{t-1}^2 + \beta h_{t-1}, \end{aligned}$$

em que $t \in \{1, \dots, T\}$; a_t representa a magnitude do *outlier* no tempo t ; $I(\cdot)$ é a função indicadora e A o conjunto dos tempos em que há observações contaminadas, podendo ser fixo e desconhecido ou resultante de um processo estocástico.

Os modelos GARCH, assim como outros modelos de volatilidade, assumem que cada observação tem o mesmo impacto relativo na estimativa de h_t , independente da magnitude do retorno. Uma propriedade importante desta classe de modelos é a capacidade de se ajustar a dados com excesso de curtose. Apesar disso, frequentemente observa-se que os resíduos estimados apresentam excesso de curtose muito maior que o esperado, que pode ser ocasionado por *outliers* aditivos (FRANSES; GHIJSELS, 1999). A presença destes valores atípicos na série temporal observada pode resultar em problemas de estimação dos parâmetros ou induzir a erros de especificação do modelo (HOTTA; TRUCÍOS, 2018). Evidências sugerem que retornos atípicos tem efeito relativamente menor na volatilidade futura do que aqueles de menor magnitude (ANDERSEN et al., 2003).

2.3 Estimação dos parâmetros e volatilidade

Considere que $\boldsymbol{\eta} \in H$, em que H é um subespaço completo de um espaço Euclidiano. Assumindo que ϵ_t tem distribuição normal padrão, condicionado a h_1 e y_1 , o estimador de máxima verossimilhança de $\boldsymbol{\eta}$ é dado por

$$\hat{\boldsymbol{\eta}}^M = \underset{\boldsymbol{\eta} \in H}{\operatorname{argmax}} L(\boldsymbol{\eta}),$$

em que

$$L(\boldsymbol{\eta}) \propto -\frac{1}{T} \sum_{t=2}^T [\log(h_t) + y_t^2/h_t], \quad (2.11)$$

a estimativa de máxima verossimilhança (EMV) de $\boldsymbol{\eta}$ pode ser obtida utilizando técnicas de otimização numérica.

Na prática nem sempre é plausível supor que $\epsilon_1, \dots, \epsilon_T$ seja uma amostra de uma distribuição normal padrão. No entanto, é possível ainda assim tratá-las como

normalmente distribuídas, fazendo uso de sua fdp para estimação dos parâmetros. Neste caso, a EMV obtida utilizando (2.11) deve ser interpretada como uma estimativa de quasi-máxima verossimilhança (EQMVn) (VERHOEVEN; MCALLEER, 2000; CRYER; CHAN, 2008).

Seja $\lambda_t = \log(y_t^2)$ uma variável aleatória com fdp g_0 e considere f como a fdp de $\epsilon_1, \dots, \epsilon_T$. Por Muler e Yohai (2008) o estimador de quasi-máxima verossimilhança pode ser obtido utilizando uma forma equivalente de (2.11) dada por

$$\hat{\boldsymbol{\eta}}_q = \underset{\boldsymbol{\eta} \in H}{\operatorname{argmax}} M_0(\boldsymbol{\eta}),$$

em particular, considerando que f corresponde a uma distribuição normal padrão têm-se que

$$\begin{aligned} M_0(\boldsymbol{\eta}) &\propto \frac{1}{T-p} \sum_{t=p+1}^T \rho_0(\ln(y_t^2/h_t)), \\ \rho_0 &= -\ln(g_0), \\ g_0(\lambda_t) &= \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2}(e^{\lambda_t} - \lambda_t) \right\}. \end{aligned}$$

O subscrito “q” indica estimativas de quasi-máxima verossimilhança (tradicionais) não robustas.

O estimador de quasi-máxima verossimilhança para o modelo GARCH(p, q) correspondente a qualquer distribuição simétrica f^* , não necessariamente coincidente com f , é obtido minimizando

$$M(\boldsymbol{\eta}) \propto \frac{1}{T-p} \sum_{t=p+1}^T \rho^*(\ln(y_t^2/h_t)),$$

condicionado a h_1 e y_1 , em que

$$\rho^* = -\log(g^*), \quad (2.12)$$

$$g^*(\lambda) = f^*(e^{\lambda/2})e^{\lambda/2}. \quad (2.13)$$

No modelo GARCH(1,1) a variância condicional prevista 1 passo à frente é dada por

$$\hat{h}_{q,t+1} = \hat{\omega}_q + \hat{\alpha}_{q,1}y_t^2 + \hat{\beta}_q\hat{h}_{q,t}. \quad (2.14)$$

Weiss (1986), Bollerslev e Wooldridge (1992), Berkes, Horváth e Kokoszka (2003), Straumann e Mikosch (2006) afirmam que o estimador de QMVn é consistente e tem distribuição assintoticamente normal, mesmo na ausência de normalidade para ϵ_t , se as condições para o segundo e quarto momentos forem satisfeitas. Porém, a precisão das estimativas e predições serão comprometidas caso f seja assimétrica. Quando f é uma distribuição com caudas pesadas Hall e Yao (2003) alertam sobre dificuldades no processo de estimação utilizando QMVn, em que a distribuição assintótica dos estimadores pode não ser normal.

2.4 Estimação robusta dos parâmetros e volatilidade

O modelo GARCH definido nas Equações (2.1)-(2.8) não possui mecanismos para detecção e tratamento de *outliers* e isso pode afetar significativamente a estimação de η (HOTTA; TRUCÍOS, 2018). Na literatura existem diversas alternativas para estimação robusta. O estimador QMVt, proposto por Bollerslev (1987), é baseado na maximização da log-verossimilhança da distribuição t de Student, sendo uma alternativa eficiente para modelagem de dados com excesso de curtose. Outra alternativa é o estimador BIP-GARCH proposto por Muler e Yohai (2008), que considera o uso de uma função limitada ρ^* como definido em (2.12) e (2.13) a fim de robustecer a estimação dos parâmetros diante de qualquer tipo de *outlier*, como apresentado em Yohai (1987) para estimativas-M de modelos de regressão considerando ρ limitada. Adicionalmente, é utilizado um filtro robusto que restringe a propagação do efeito de *outliers* na estimação dos h_t 's. Um terceiro estimador é o BQMVt (CARNERO; PEÑA; RUIZ, 2012) baseado na combinação entre QMVt e o mecanismo limitador utilizado em BIP-GARCH, além de utilizar um filtro robusto alternativo. O estimador BIP-GARCH (BOUDT; DANIELSSON; LAURENT, 2013), também baseado no BIP-GARCH, apresenta uma modificação na equação de atualização da volatilidade e na estimação da variância marginal.

Segundo Huber e Ronchetti (2009, Seção 1.2, pág. 5), um estimador robusto precisa satisfazer duas propriedades:

- Ser eficiente quando todas as observações seguem o modelo proposto.
- Substituir uma pequena parcela da amostra por *outliers* não deve impactar significativamente as estimativas.

De modo geral, quando comparado a métodos robustos, o EQMVn pode produzir estimativas de menor viés e EQM. Por outro lado, quando há *outliers*, o desempenho deste estimador é afetado. Neste cenário QMVt, BIP-GARCH, BQMVt e BIP-GARCH apresentam melhores resultados. Para séries com *outliers* isolados ou com dois consecutivos o método BIP-GARCH apresenta a melhor performance (HOTTA; TRUCÍOS, 2018). Por meio de estudos de simulação, Boudt, Danielsson e Laurent (2013) perceberam que o desempenho do estimador robusto proposto não é significativamente inferior ao EQMVn em cenários sem contaminação, além de apresentar melhor performance em séries contaminadas.

2.4.1 Estimador QMVt

Substituir (2.11) pela função de log-verossimilhança da distribuição t de Student pode reduzir a influência de outliers por esta ser uma distribuição com caudas mais pesadas

que da distribuição Gaussiana (CARNERO; PEÑA; RUIZ, 2012; HOTTA; TRUCÍOS, 2018). O logaritmo da função de quase-verossimilhança t de Student é dada por,

$$L_t(\boldsymbol{\eta}) \propto -\sum_{t=1}^T \left(\log \left(t_v \left(\frac{y_t}{\sqrt{h_t}} \right) \right) - \frac{1}{2} \log(h_t) \right), \quad (2.15)$$

sendo $\boldsymbol{\eta}$ o vetor de parâmetros desconhecidos incluindo os graus de liberdade e $t_v(\cdot)$ a densidade da distribuição t de Student com v graus de liberdade padronizada para ter variância unitária. Condicionado a h_1 e y_1 o estimador QMvt de $\boldsymbol{\eta}$ é dado por

$$\hat{\boldsymbol{\eta}}^t = \underset{\boldsymbol{\eta} \in H}{\operatorname{argmax}} L_t(\boldsymbol{\eta}),$$

2.4.2 Estimador-M para o modelo BIP-GARCH

O impacto de a_t no processo de estimação pode ser reduzido utilizando uma modificação do modelo GARCH que limita o efeito dos *outliers* passados na variância condicional por meio de um filtro robusto na expressão que define h_t responsável por identificar um retorno atípico. Visto que a volatilidade é variante no tempo, a identificação dos *outliers* é feita utilizando o quadrado dos retornos desvolatilizados, $y_{t-1}/\sqrt{h_{t-1}}$. Assim, a Equação (2.8) é substituída por

$$h_t = \omega + \alpha w_{\delta,1} \left(\frac{y_{t-1}^2}{h_{t-1}} \right) y_{t-1}^2 + \beta h_{t-1}. \quad (2.16)$$

em que $w_{\delta,N}(\cdot)$ é uma função peso dada por

$$w_{\delta,N}(u) = \min(u, k_{\delta,N}) u^{-1} \times c_{\delta,N}, \quad (2.17)$$

$$c_{\delta,N} = \frac{E(U)}{E(\min(U, k_{\delta,N}))}, \quad (2.18)$$

em que $U \sim \chi_1^2$ e, para o modelo univariado GARCH, $N = 1$ em (2.17) e (2.18). No caso gaussiano, y_{t-1}^2/h_{t-1} segue uma distribuição qui-quadrado com parâmetro de não centralidade a_{t-1}^2 e 1 grau de liberdade. Caso não haja salto então $a_{t-1}^2 = 0$ e $y_{t-1}^2/h_{t-1} \sim \chi_1^2$. Portanto, uma maneira de identificar se o quadrado do retorno desvolatilizado y_{t-1}^2/h_{t-1} é um *outlier* pode ser feita mediante a comparação do mesmo com $k_{\delta,1}$, o quantil de ordem δ de uma distribuição qui-quadrado com 1 grau de liberdade.

A função peso definida na Equação (2.17) foi proposta por Boudt, Danielsson e Laurent (2013) como uma modificação da função peso adotada em Muler e Yohai (2008), caracterizada pela inclusão da constante $c_{\delta,1}$ que garante que a esperança condicional dos quadrado dos choques ponderados é a variância condicional, mesmo na ausência de *outliers*.

A estimativa de $\boldsymbol{\eta}$ por QML descrita na Seção 2.3 não é robusta a *outliers* mesmo substituindo (2.8) por (2.16). Um dos motivos é que $\rho(z) = z$ é não limitada, de

modo que um retorno atípico pode ter um efeito desproporcional em $M(\boldsymbol{\eta})$. Dado isso, diferentes escolhas para $\rho(\cdot)$ podem ser adotadas a fim de robustecer o procedimento de estimação. Muler e Yohai (2008) classificam as estimativas-M em dois tipos: (1) estimativas-M com ρ' (derivada de ρ) limitada e ρ não limitada (2) estimativas-M com $\rho(\cdot)$ e $\rho(\cdot)'$ limitadas. Em (1) as estimativas são robustas quando f é uma distribuição de cauda pesada, no entanto as estimativas ainda podem ser afetadas por *outliers* aditivos. Devido a isso, considera-se (2) para aumentar o grau de robustez. Por robustez, a função $\rho(\cdot)$ precisa minimizar o efeito de pontos extremos. A escolha da função ρ rege uma troca entre robustez e eficiência. Algumas candidatas foram investigadas em Boudt e Croux (2010). Portanto, a estimativa para η no modelo BIP-GARCH pode ser feita utilizando um estimador-M que minimiza o valor médio de uma função objetivo $\rho(\cdot)$ avaliada no logaritmo dos retornos desvolatilizados ao quadrado condicionada a h_1 e y_1 , dado por

$$\hat{\boldsymbol{\eta}}_r = \underset{\boldsymbol{\eta} \in H}{\operatorname{argmin}} \frac{1}{T} \sum_{t=1}^T \rho \left(\log \frac{y_t^2}{h_t} \right). \quad (2.19)$$

O subscrito “r” indica estimativas de quasi-máxima verossimilhança robustas.

Muler e Yohai (2008) recomendam uma função- ρ dada por $\rho_1(z) = 0,8m(g_0(z)/0,8)$, em que

$$m(z) = \begin{cases} z, & \text{se } z \leq 4, \\ P(z), & \text{se } 4 < z < 4,3, \\ 4,15, & \text{se } z > 4,3 \end{cases}$$

e

$$\begin{aligned} P(z) &= \frac{2}{(x_1 - x_0)^3} \left(\frac{1}{4}(x^4 - x_0^4) - \frac{1}{3}(2x_0 + x_1)(x^3 - x_0^3) + \frac{1}{2}(x_0^2 + 2x_0x_1)(x^2 - x_0^2) \right) \\ &\quad - \frac{2x_0^2x_1}{(x_1 - x_0)}(x - x_0) - \frac{1}{3(x_1 - x_0)^2}(x - x_0)^3 + x, \end{aligned}$$

sendo $x_0 = 4$ e $x_1 = 4,3$.

No entanto, $\rho_1(\cdot)$ não pode ser estendida facilmente para o caso multivariado, assim Boudt, Danielsson e Laurent (2013) recomendam, por meio de estudos de simulação comparando diversos candidatas, $\rho_2(\cdot)$ que é baseada na fdp de uma distribuição t de Student com v graus de liberdade, dada por

$$\begin{aligned} \rho_2(z) &= -z + \sigma_{1,v} \rho_{t_{1,v}}(\exp(z)), \\ \rho_{t_{1,v}}(u) &= (1+v) \log \left(1 + \frac{u}{v-2} \right), \\ \sigma_{1,v} &= \frac{1}{E[\rho'_{t_{1,v}}(U)U]}, \\ U &\sim \chi_1^2, \end{aligned}$$

com $v = 4$ e sendo χ_1^2 uma distribuição qui-quadrado com 1 graus de liberdade. Deste modo, a estimação de ω , α e β é realizada conjuntamente. No entanto, dificuldades surgem quando α e β somam aproximadamente 1, implicando que ω será muito pequeno mas obrigatoriamente positivo, como discutido em [Mezrich e Engle \(1996\)](#). Alternativamente, é possível utilizar o procedimento de *variance targeting* em que ω é reparametrizado como $\hat{h}(1 - \alpha - \beta)$, sendo \hat{h} um estimador da variância não condicional.

No modelo BIP-GARCH(1,1) a variância condicional prevista 1 passo à frente é dada por

$$\hat{h}_{r,t+1} = \hat{\omega}_r + \hat{\alpha}_r w_{\delta,1} \left(\frac{y_t^2}{\hat{h}_{r,t}} \right) y_t^2 + \hat{\beta}_r \hat{h}_{r,t}. \quad (2.20)$$

[Trucíos, Hotta e Ruiz \(2017\)](#) e [Trucíos, Hotta e Ruiz \(2018\)](#) propõem algumas modificações ao BIP a fim de obter melhores previsões de volatilidades/correlações.

2.4.3 Estimadores robustos para média e variância não condicionais

As definições apresentadas consideram que y_t tem média ($\mu = 0$), quando isso não ocorre é necessário estimá-la. Estimadores naturais para μ e h são a média e variância amostrais. No entanto, esses estimadores são muito sensíveis a *outliers*, se fazendo necessário a utilização de um estimador robusto, mais especificamente, os estimadores robustos reponderados de média e variância propostos por [Boudt, Croux e Laurent \(2011\)](#), dados por

$$\hat{\mu} = \frac{\sum_{t=1}^T y_t^* I_t}{\sum_{t=1}^T I_t} \quad \text{e} \quad \hat{h} = b_{0,95,1} \cdot \frac{\sum_{t=1}^T (y_t^* - \hat{\mu})^2 J_t}{\sum_{t=1}^T J_t},$$

em que

$$I_t = I \left[\frac{(y_t^* - \text{mediana}_t(y_t^*))^2}{mad_t^2} \leq \chi_1^2(0, 95) \right] \quad \text{e} \quad J_t = I \left[\frac{(y_t^* - \hat{\mu})^2}{mad_t^2} \leq \chi_1^2(0, 95) \right],$$

sendo $\chi_1(\delta)$ um quantil de ordem δ de uma distribuição qui-quadrado com 1 grau de liberdade, mad_t o valor absoluto dos desvios dos retornos em relação a mediana em uma janela local em torno da observação y_t^* .

3 Modelos MGARCH

Os modelos MGARCH foram desenvolvidos a partir do fim da década de 1980 como extensão dos processos GARCH univariados para o domínio multivariado (BOLLERSLEV; ENGLE; WOOLDRIDGE, 1988). Motivada pelo conceito de que volatilidades financeiras evoluem conjuntamente ao longo do tempo, a modelagem de séries multivariadas leva a modelos empíricos mais relevantes que os univariados. Eles são úteis para estudar a relação entre volatilidades e co-volatilidades de diferentes mercados, verificar e quantificar se choques em um mercado impactam a volatilidade de outro, avaliar se correlações entre retorno de ativos mudam no decorrer do tempo, se são maiores em períodos de alta volatilidade (como em períodos de crise financeira) entre outras questões (BAUWENS; LAURENT; ROMBOUTS, 2006).

Um bom modelo MGARCH deve ser capaz de representar variâncias e covariâncias; ter especificação parcimoniosa que facilite a estimação e interpretação dos parâmetros do modelos, mas que ainda seja capaz de descrever dinâmicas relevantes na estrutura da matriz de covariância condicional; seja flexível e viável na modelagem de séries de dimensão considerável; que a estrutura imposta para matriz covariâncias resulte em estimativas positivas definidas (SILVENNOINEN; TERÄSVIRTA, 2009; BOUDT et al., 2019).

A generalização dos modelos GARCH univariados para o caso multivariado acontece pela substituição de h_t , a variância condicional, por \mathbf{H}_t , a matriz de covariâncias condicionais. Considere um vetor de processo estocástico $\{\mathbf{r}_t\}$ de dimensão $N \times 1$, tal que

$$\mathbf{r}_t = \mathbf{H}_t^{1/2} \boldsymbol{\epsilon}_t,$$

A matriz de covariância, condicionada a \mathcal{F}_{t-1} , é dada por $\mathbf{H}_t = \text{Var}(\mathbf{r}_t | \mathcal{F}_{t-1})$, uma matriz positiva definida de dimensão $(N \times N)$. O vetor de variáveis aleatórias independentes $\boldsymbol{\epsilon}_t$ tem dimensão $(N \times 1)$, $E(\boldsymbol{\epsilon}_t) = \mathbf{0}$ e $\text{Var}(\boldsymbol{\epsilon}_t) = \mathbf{I}_N$ sendo $\mathbf{0}$ um vetor de zeros de dimensão $(N \times 1)$ e \mathbf{I}_N a matriz identidade de ordem N .

Uma estrutura para \mathbf{H}_t ainda precisa ser apresentada; Silvennoinen e Teräsvirta (2009) dividem a literatura de especificação para matriz de covariância condicional em quatro categorias: (1) modelagem de \mathbf{H}_t diretamente, que inclui, dentre outros, os modelos VEC de Bollerslev, Engle e Wooldridge (1988) e Baba-Engle-Kraft-Kroner (BEKK) de Engle e Kroner (1995); (2) modelos de fatores, em que se assume que \mathbf{r}_t é gerado por um pequeno número de fatores heteroscedásticos não observáveis; Engle, Ng e Rothschild (1990) introduzem o primeiro modelo de fatores GARCH; (3) especificação das variâncias e correlações condicionais, englobando modelos como o de correlação condicional constante (*Constant Conditional Correlation*, CCC) de Bollerslev (1990), correlação condicional

dinâmica (*dynamic conditional correlation*, DCC) de Engle (2002) e a versão de Tse e Tsui (2002) e o modelo DCC de Engle (2002) corrigido (*corrected DCC*, cDCC) de Aielli (2013); (4) abordagem semi e não paramétrica destinada a situações em que há perda de eficiência dos estimadores paramétricos devido especificação incorreta da matriz de covariância condicional, veja Linton (2008).

O supracitado modelo VECH foi um dos primeiros MGARCH, sendo uma generalização direta do modelo GARCH univariado. A generalidade deste modelo proporciona muita flexibilidade porém trás muitas exigências, devido a restrições impostas nos parâmetros para que \mathbf{H}_t seja positiva definida, e custo computacional para estimação de um elevado número de parâmetros. Dado isto, diferentes alternativas mais parcimoniosas surgiram, acomodadas em duas linhas de pesquisa. A primeira foca em impor restrições nos parâmetros do modelo VEC, como os modelos BEKK e fatorial. A segunda baseia-se na modelagem da covariância condicional pelas variâncias e correlações condicionais. Estes modelos tem como vantagem a facilidade relativa na etapa de estimação e que seus parâmetros possuem interpretação, tornando-se muito populares (SILVENNOINEN; TERÄSVIRTA, 2009).

Para modelar séries de altas dimensões é natural considerar inicialmente modelos mais simples. Neste sentido, será considerada a família de modelos de correlação condicional, mais especificamente os modelos DCC e cDCC. Estes modelos comportam uma estrutura dinâmica para a matriz de correlação condicional, tornando-os mais realistas em comparação a especificação constante adotada no modelo CCC.

3.1 Modelo DCC

O modelo de DCC proposto por Engle (2002) baseia na decomposição da matriz de covariância condicional \mathbf{H}_t em desvios padrão e matriz de correlação condicional, i.e

$$\mathbf{H}_t = \mathbf{D}_t^{1/2} \mathbf{R}_t \mathbf{D}_t^{1/2}, \quad (3.1)$$

sendo \mathbf{R}_t a matriz de correlação condicional e $\mathbf{D}_t \equiv \text{diag}(h_{1,t}, \dots, h_{N,t})$ uma matriz diagonal com $h_{i,t} = \text{var}(r_{i,t} | \mathcal{F}_{t-1})$, ambas de dimensão $(N \times N)$ e avaliadas no tempo t . A variância condicional é tipicamente tratada como um processo GARCH univariado de ordem (p, q) . A estrutura para correlação dinâmica, caracterizando o modelo DCC de ordem (U, V) é dada em função de $\tilde{\mathbf{r}}_t \equiv (r_{t,1}, \dots, r_{t,N})' = \mathbf{D}_t^{-1} \mathbf{r}_t$ por

$$\mathbf{Q}_t = (1 - \sum_{u=1}^U a_u - \sum_{v=1}^V b_v) \mathbf{S} + \sum_{u=1}^U a_u (\tilde{\mathbf{r}}_{t-u} \tilde{\mathbf{r}}'_{t-u}) + \sum_{v=1}^V b_v \mathbf{Q}_{t-v}, \quad (3.2)$$

$$\mathbf{R}_t = \mathbf{Q}_t^{*-1/2} \mathbf{Q}_t \mathbf{Q}_t^{*-1/2}, \quad (3.3)$$

em que $\mathbf{S} := [s_{i,j}]$ é a matriz de correlação não condicional de $\tilde{\mathbf{r}}_t$, tal que $s_{i,i} = 1$ e a_u e b_v são escalares. \mathbf{Q}_t^* é uma matriz diagonal composta pelos elementos da diagonal de

$\mathbf{Q}_t := [q_{i,j,t}]$, i.e

$$\mathbf{Q}_t^* = \begin{bmatrix} q_{1,1,t} & 0 & \dots & 0 \\ 0 & q_{2,2,t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & q_{N,N,t} \end{bmatrix},$$

e R_t tem forma explícita dada por

$$\mathbf{R}_t = \begin{bmatrix} 1 & \frac{q_{1,2,t}}{\sqrt{q_{1,1,t}q_{2,2,t}}} & \dots & \frac{q_{1,N,t}}{\sqrt{q_{1,1,t}q_{N,N,t}}} \\ \frac{q_{2,1,t}}{\sqrt{q_{1,1,t}q_{2,2,t}}} & 1 & \dots & \frac{q_{2,N,t}}{\sqrt{q_{2,2,t}q_{N,N,t}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{q_{N,1,t}}{\sqrt{q_{N,N,t}q_{1,1,t}}} & \frac{q_{N,2,t}}{\sqrt{q_{N,N,t}q_{2,2,t}}} & \dots & 1 \end{bmatrix}.$$

Uma das complicações do modelo é que \mathbf{R}_t precisa ser positiva definida para todo t . Um resultado de álgebra linear útil para verificar esta condição é: uma matriz quadrada \mathbf{A} é positiva definida se e somente se $\mathbf{B} = \mathbf{A}^{*-1}\mathbf{A}\mathbf{A}^{*-1}$ é positiva definida; em que \mathbf{A} e \mathbf{A}^* são análogas a \mathbf{Q}_t e \mathbf{Q}_t^* . Portanto, para que \mathbf{R}_t seja positiva definida, \mathbf{Q}_t também deve ser. Para verificar isso, [Engle e Sheppard \(2001\)](#) apresentam um conjunto de condições suficientes, para as partes GARCH univariada e DCC, para que \mathbf{H}_t seja uniformemente positiva definida.

Condições suficientes para \mathbf{H}_t ser positiva definida: Seja $\boldsymbol{\eta}_i = (\omega_i, \alpha_i, \beta_i)'$ o vetor de parâmetros do modelo GARCH(1,1) associado a i -ésima série unidimensional satisfazendo $(\alpha_i + \beta_i) < 1$ para $i \in [1, \dots, N]$:

- a. $\omega_i, h_{i,0} > 0$;
- b. α_i e β_i são tais que $h_{it} > 0$ com probabilidade 1;
- c. As raízes de $(1 - \alpha_i Z + \beta_i Z)$ estão fora do círculo unitário;
- d. Os parâmetros do modelo DCC satisfazem (3.2);
- d. $a_u, b_v \geq 0$ para $u \in [1, \dots, U]$ e $v \in [1, \dots, V]$;
- e. $\sum_{u=1}^U a_u + \sum_{v=1}^V b_v < 1$;
- f. O menor autovalor de $\mathbf{S} > \delta > 0$;

se todas as condições forem satisfeitas então \mathbf{H}_t é positiva definida para todo t . \mathbf{Q}_t será positiva definida para todo t por ser a média ponderada de uma matriz positiva definida

\mathbf{S} , uma matriz positiva semi definida ($\tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t$) e uma matriz positiva definida \mathbf{Q}_{t-1} ; \mathbf{Q}_0 é positiva definida pela condição f .

O processo DCC(1,1)-GARCH(1,1) é uma versão mais simples e parcimoniosa e é usualmente muito utilizada na prática, este será o padrão para o restante da dissertação, exceto quando mencionado explicitamente. A estrutura imposta para \mathbf{Q}_t passa a ser

$$\mathbf{Q}_t = (1 - a - b)\mathbf{S} + a\tilde{\mathbf{r}}_{t-1}\tilde{\mathbf{r}}'_{t-1} + b\mathbf{Q}_{t-1}. \quad (3.4)$$

3.2 Modelo cDCC

No modelo DCC a matriz \mathbf{S} é tratada como o segundo momento de $\tilde{\mathbf{r}}_t$, mas [Aielli \(2013\)](#) alerta sobre a invalidade desta igualdade. Considere que $(a + b) < 1$ e $E(\mathbf{Q}_t)$ e $E(\tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t)$ são independentes de t . Então, pela Equação (3.4), têm-se que

$$\begin{aligned} E(\mathbf{S}) &= \frac{1}{1-a-b}(E(\mathbf{Q}_t) - bE(\mathbf{Q}_{t-1}) - aE(\tilde{\mathbf{r}}_{t-1}\tilde{\mathbf{r}}'_{t-1})) \\ \mathbf{S} &= \frac{1}{1-a-b}(E(\mathbf{Q}_t) - bE(\mathbf{Q}_t) - aE(\tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t)) \quad (\text{independente de } t) \end{aligned}$$

usando o fato de que $E(\mathbf{Q}_t) = E(\mathbf{Q}_t^{*1/2} \mathbf{R}_t \mathbf{Q}_t^{*1/2})$

$$\mathbf{S} = \frac{1-b}{1-a-b}E(\mathbf{Q}_t^{*1/2} \tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t \mathbf{Q}_t^{*1/2}) - \frac{a}{1-a-b}E(\tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t), \quad (3.5)$$

com $\mathbf{S} = E(\tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t)$, se e somente se, $E(\tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t) = E(\mathbf{Q}_t)$. Para o caso de correlação condicional constante ($a = b = 0$) têm-se que $\mathbf{S} = E(\tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t)$. Como comentado em [Aielli \(2013\)](#), este parece ser o único caso em que a igualdade ocorre. Note pela Equação (3.4) que

$$\begin{aligned} E(\tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t) &= E(E(\tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t | \mathcal{F}_{t-1})) \\ &= E(\mathbf{R}_t) \\ &= E(\mathbf{Q}_t^{*-1/2} \mathbf{Q}_t \mathbf{Q}_t^{*-1/2}), \end{aligned}$$

com exceção do caso de correlação condicional constante, segue que $\mathbf{Q}_t^{*-1/2} \mathbf{Q}_t \mathbf{Q}_t^{*-1/2} \neq \mathbf{Q}_t$. A tratabilidade do modelo DCC pode ser aprimorada aplicando-se um correção na Equação (3.4). Para o modelo cDCC (*corrected DCC*) de ordem (1, 1), segue que

$$\mathbf{Q}_t = (1 - a - b)\mathbf{S} + a \mathbf{Q}_{t-1}^{*1/2} \tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t \mathbf{Q}_{t-1}^{*1/2} + b \mathbf{Q}_{t-1}. \quad (3.6)$$

Pré e pós multiplicar (3.6) por uma matriz diagonal \mathbf{Z} resulta em uma fórmula análoga, tal que, $\mathbf{S}, \mathbf{Q}_t, \mathbf{Q}_t^*$ passam a ser $\bar{\mathbf{S}} = \mathbf{Z} \mathbf{S} \mathbf{Z}$, $\bar{\mathbf{Q}}_t = \mathbf{Z} \mathbf{Q}_t \mathbf{Z}$, e $\bar{\mathbf{Q}}_t^* = \mathbf{Z} \mathbf{Q}_t^* \mathbf{Z}$, respectivamente. Como $\mathbf{R}_t = \mathbf{Q}_t^{*-1/2} \mathbf{Q}_t \mathbf{Q}_t^{*-1/2} = \bar{\mathbf{Q}}_t^{*-1/2} \bar{\mathbf{Q}}_t \bar{\mathbf{Q}}_t^{*-1/2}$, os parâmetros (\mathbf{S}, a, b) e $(\bar{\mathbf{S}}, a, b)$ são indistinguíveis de $\tilde{\mathbf{r}}_t$. Portanto, a identificabilidade de (\mathbf{S}, a, b) é garantida assumindo que \mathbf{S} é diagonal unitária.. As condições para a positividade e estacionariedade de \mathbf{H}_t são as mesmas do modelo DCC.

Sob a especificação na Equação (3.6) e ainda considerando $(a + b) < 1$ e que $E(\mathbf{Q}_t)$ e $E(\tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t)$ são independentes de t , a matriz \mathbf{S} passa a definida por

$$\mathbf{S} = E(\mathbf{Q}_t^{*1/2} \tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_t \mathbf{Q}_t^{*1/2}), \quad (3.7)$$

assim, no modelo cDCC, \mathbf{S} é o segundo momento de $\mathbf{Q}_t^{*1/2} \tilde{\mathbf{r}}_t$.

3.2.1 Condições de estacionaridade

Para verificar as condições estacionariedade, Aielli (2013) reescreve o modelo cDCC como um caso particular do modelo BEKK a fim de evitar a estrutura não linear característica de sua forma original (FERMANIAN; MALONGO, 2017). Dessa forma, no contexto de um processo cDCC de ordem (U, V) com representação BEKK para $K > 1$ têm-se que \mathbf{Q}_t passa a ser representada por

$$\mathbf{Q}_t = \mathbf{C} + \sum_{u=1}^U \sum_{k=1}^K \bar{A}_{u,k} \left\{ \mathbf{Q}_{t-u}^* \tilde{\mathbf{r}}_{t-u} \tilde{\mathbf{r}}_{t-u}^\top \mathbf{Q}_{t-u}^{*1/2} \right\} \bar{A}_{u,k}^\top + \sum_{v=1}^V \sum_{k=1}^K \bar{B}_{v,k} \mathbf{Q}_{t-v} \bar{B}_{v,k}^\top$$

para

$$\mathbf{C} \equiv \mathbf{S} - \sum_{u=1}^U \sum_{k=1}^K \bar{A}_{u,k} \mathbf{S} \bar{A}_{u,k}^\top - \sum_{v=1}^V \sum_{k=1}^K \bar{B}_{v,k} \mathbf{S} \bar{B}_{v,k}^\top,$$

tal que para $l = 1, \dots, K$, têm-se

$$\begin{aligned} \bar{A}_{i,l} &= \begin{bmatrix} a_{11,l} & a_{12,l} & \dots & a_{1N,l} \\ a_{21,l} & a_{22,l} & \dots & a_{2N,l} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1,l} & a_{N2,l} & \dots & a_{NN,l} \end{bmatrix}, \quad i = 1, \dots, U \\ \bar{B}_{j,l} &= \begin{bmatrix} b_{11,l} & b_{12,l} & \dots & b_{1N,l} \\ b_{21,l} & b_{22,l} & \dots & b_{2N,l} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N1,l} & b_{N2,l} & \dots & b_{NN,l} \end{bmatrix}, \quad j = 1, \dots, V \end{aligned}$$

em que \mathbf{Q}_t é positiva definida se \mathbf{C} também for. Este modelo é denotado por BEKK cDCC.

Dado que os processos de variância são estacionários sob a suposição de estacionaridade univariada de inovações padronizadas, e que os processos de correlação são estacionários, qualquer forma funcional para o processo de variância é permitida, sem que prejudique a estacionaridade do processo da covariância. Dado isso, o princípio da estacionaridade para o modelo cDCC resulta da verificação da estacionaridade de $(\mathbf{R}_t, \tilde{\mathbf{r}}_t)$ e de $diag(\mathbf{D}_t)$ para garantir estacionaridade de $(\mathbf{H}_t, \mathbf{r}_t)$. Para o modelo BEKK cDCC

considere

$$\mathbf{A}_q \equiv \mathbf{U} \left\{ \sum_{k=1}^K A_{q,k} \otimes A_{q,k} \right\} \mathbf{W}^\top, \quad q = 1, 2, \dots, \bar{Q},$$

$$\mathbf{B}_p \equiv \mathbf{U} \left\{ \sum_{k=1}^K B_{p,k} \otimes B_{p,k} \right\} \mathbf{W}^\top, \quad p = 1, 2, \dots, \bar{P},$$

em que \mathbf{U} e \mathbf{W} são matrizes únicas de dimensão $N(N+1)/2 \times N^2$, tal que para uma matriz N -dimensional e simétrica \mathbf{Z} , $vech(\mathbf{Z}) = \mathbf{U}vec(\mathbf{Z})$, $vec(\mathbf{Z}) = \mathbf{W}^\top vech(\mathbf{Z})$, e $\mathbf{U}\mathbf{W}^\top = I_{N(N+1)/2}$, em que vec e $vech$ denotam, respectivamente, os operadores de vetorização de colunas e parte triangular inferior da matriz argumento. Considere $\boldsymbol{\theta} = (\eta'_1, \dots, \eta'_N)'$ e $\mathbf{D}_t(\boldsymbol{\theta}, \tilde{\mathbf{r}}_{t-1})$ como a representação de todas as variâncias condicionais em função de $\tilde{\mathbf{r}}_{t-1}$. Assumindo que $\tilde{\mathbf{r}}_t = \mathbf{R}_t^{1/2} \kappa_t$, tal que \mathbf{R}_t é positiva definida e $\mathbf{R}^{1/2} \mathbf{R}^{1/2} = \mathbf{R}_t$ e κ_t é iid com $E[\kappa_t] = 0$ e $E[\kappa_t \kappa_t^\top] = I_N$. Dado isso, um processo cDCC é estacionário se:

H1. A função densidade de probabilidade de κ_t é absolutamente contínua com respeito a medida de Lebesgue e positiva em uma vizinhança da origem;

H2. \mathbf{C} é positiva definida e o maior autovalor de $\sum_{u=1}^U \mathbf{A}_u + \sum_{v=1}^V \mathbf{B}_v$ é menor que um em módulo;

H3. Os elementos em $diag(\mathbf{D}_t(\boldsymbol{\theta}, \tilde{\mathbf{r}}_{t-1}))$ são mensuráveis;

H4. $E[r_{i,t}^2] < \infty$ para $i = 1, \dots, N$.

Supondo *H1* e *H2* o processo $[vech(\mathbf{R}_t)', \tilde{\mathbf{r}}_t']'$ admite estacionaridade fraca, estrita e não antecipativa, além de solução ergótica. Por *H1*, *H2* e *H3* o processo $[vech(\mathbf{H}_t)', \mathbf{r}_t', vech(\mathbf{R}_t)', \tilde{\mathbf{r}}_t']'$ admite estacionaridade estrita e não antecipativa bem como solução ergótica. Em adição a *H1* – *H3*, *H4* garante que o processo \mathbf{r}_t admite solução fracamente estacionaria. As demonstrações para validade de cada uma das suposições de correlação citadas podem ser consultadas em Aielli (2013). Basicamente, *H1* é tipicamente satisfeita em aplicações; *H2* é análoga as suposições em (3.9)-(3.10) para o modelo Vech cDCC Diagonal expresso em (3.10). Para *H3*, os elementos em $diag(\mathbf{D}_t(\boldsymbol{\theta}, \tilde{\mathbf{r}}_{t-1}))$ são mensuráveis sob estacionaridade estrita de $\tilde{\mathbf{r}}_t$. Para inovações GARCH (1, 1) isso é verificado pela condição em (2.9) e $E[\max(0, \log(\alpha_i \epsilon_{i,t}^2 + \beta_i))] < \infty$. Por fim, para *H4* ainda no contexto de inovações GARCH (1, 2), $E[r_{i,t}^2] < \infty$ se $\alpha_i + \beta_i < 1$.

$$\sum_{u=1}^U a_{ii,u} + \sum_{v=1}^V b_{ii,v} < 1, \quad i = 1, \dots, N \tag{3.8}$$

$$\mathbf{C}^* = \left(\mathbf{u}^\top - \sum_{u=1}^U \mathbf{A}_u - \sum_{v=1}^V \mathbf{B}_v \right) \odot \mathbf{S} \quad \text{é pd} \tag{3.9}$$

$$\mathbf{Q}_t = \mathbf{C}^* + \sum_{u=1}^U A_q \odot \left\{ \mathbf{Q}_t^{*1/2} \tilde{\mathbf{r}}_{t-u} \tilde{\mathbf{r}}_{t-u}^\top \mathbf{Q}_t^{*1/2} \right\} + \sum_{v=1}^V B_v \odot Q_{t-v} \tag{3.10}$$

3.3 Modelo cDCC contaminado

Considere o vetor de retornos observados $\{\mathbf{y}_t^*\}$ de dimensão $N \times 1$ com vetor de médias $\boldsymbol{\mu}$ dado \mathcal{F}_{t-1} e seja $\{\mathbf{y}_t\}$ o vetor de retornos centralizados têm-se que a versão contaminada do modelo cDCC, proposta por [Boudt, Danielsson e Laurent \(2013\)](#), é dada por

$$\mathbf{y}_t = \mathbf{y}_t^* - \boldsymbol{\mu} = \mathbf{r}_t + \mathbf{d}_t I(t \in B), \quad (3.11)$$

$$\mathbf{r}_t = \mathbf{H}_t^{1/2} \boldsymbol{\epsilon}_t \text{ em que } \boldsymbol{\epsilon}_t \stackrel{i.i.d.}{\sim} (0, I_N), \quad (3.12)$$

em que \mathbf{y}_t é composto por dois componentes não-observáveis, um processo GARCH(1,1)-cDCC denotado por $\mathbf{r}_t = (r_{1,t}, \dots, r_{N,t})'$ e um processo aditivo N-dimensional \mathbf{d}_t . B denota o conjunto dos tempos em que há observações contaminadas. Os componentes de \mathbf{d}_t são definidos por

$$d_{i,t} = \text{sinal}(y_{i,t}) d I(i \in \mathcal{I}_t),$$

sendo d o tamanho do *outlier* para a i-ésima série no tempo t e \mathcal{I}_t o conjunto de índices associado às séries com contaminação no tempo t .

3.4 Estimação dos parâmetros e correlações condicionais

Seja $\boldsymbol{\phi} = (a, b)'$ o vetor de parâmetros do modelo cDCC de ordem (1, 1) definido no espaço paramétrico Φ . Assumindo que os modelos de volatilidade para todas as séries de retorno são processos GARCH de ordem (1, 1), definidos como nas Equações (2.1) e (2.8), tal que $\boldsymbol{\eta}_i = (\omega_i, \alpha_i, \beta_i)'$ para a i-ésima série. Considerando $\boldsymbol{\theta} = (\boldsymbol{\eta}'_1, \boldsymbol{\eta}'_2, \dots, \boldsymbol{\eta}'_N)'$ e $(\mathbf{S}, \boldsymbol{\theta})$ como parâmetros de perturbação, o processo de estimação para \mathbf{H}_t é dado por:

1. com $\mathbf{y}_t = (y_{1,t}^* - \hat{\mu}_{1,t}, \dots, y_{N,t}^* - \hat{\mu}_{N,t})'$, obter $\hat{\boldsymbol{\theta}}_q$, o EQMVn para $\boldsymbol{\theta}$, tal que cada componente $\boldsymbol{\eta}_i$ é estimada utilizando o método apresentado na Seção (2.3) e especificação GARCH(1, 1) dada pela Equação (2.8). O subscrito “q” indica estimativas de quase-verossimilhança (tradicional) não robustas.
2. calcular matriz de volatilidades estimadas $\hat{\mathbf{D}}_{q,t} \equiv \text{diag}(\hat{h}_{q,1,t}, \dots, \hat{h}_{q,N,t})$ em que cada componente $\hat{h}_{q,i,t}$ é calculado usando a Equação (2.14);
3. calcular retornos desvolatilizados $\tilde{\mathbf{y}}_{q,t} = (\tilde{y}_{q,1,t}, \dots, \tilde{y}_{q,N,t})' = (\hat{\mathbf{D}}_{q,t})^{-1} \mathbf{y}_t$;
4. calcular $\hat{\boldsymbol{\phi}}_q = \underset{\boldsymbol{\phi} \in \Phi}{\text{argmax}} L_T(\hat{\boldsymbol{\theta}}_q, \hat{\mathbf{S}}_q(\hat{\boldsymbol{\theta}}_q, \boldsymbol{\phi}), \boldsymbol{\phi})$;
5. calcular $\hat{\mathbf{H}}_{q,t}$ dado $(\hat{\mathbf{D}}_{q,t}, \hat{\mathbf{S}}_q(\hat{\boldsymbol{\theta}}_q, \hat{\boldsymbol{\phi}}_q), \hat{\boldsymbol{\phi}}_q)$.

No passo (4) o estimador do parâmetro de perturbação, condicionado ao parâmetro de interesse ϕ , é denotado por $(\hat{\theta}_q, \hat{\mathbf{S}}_q(\hat{\theta}_q, \phi))$. Dado \mathbf{D}_1 e $\tilde{\mathbf{y}}_{q,1}$, ϕ pode ser estimado por meio da maximização da função de quase verossimilhança normal dada por

$$L_T(\hat{\theta}_q, \mathbf{S}(\hat{\theta}_q, \phi), \phi) \propto -\frac{1}{2} \sum_{t=1}^T \left[\log(\det \mathbf{R}_t(\hat{\theta}_q, \phi)) + \tilde{\mathbf{y}}'_{q,t} \mathbf{R}_t^{-1}(\hat{\theta}_q, \phi) \tilde{\mathbf{y}}_{q,t} \right], \quad (3.13)$$

em que $\mathbf{S} = E \left[\mathbf{Q}_t^{*1/2} \tilde{\mathbf{y}}_t \tilde{\mathbf{y}}'_t \mathbf{Q}_t^{*1/2} \right]$ é a matriz de correlação não condicional de $\mathbf{Q}_t^{*1/2} \tilde{\mathbf{y}}_t$.

\mathbf{S} e ϕ são estimados simultaneamente, com $\hat{\mathbf{S}}_q(\hat{\theta}, \phi)$, a matriz de correlação amostral de $\mathbf{Q}_t^{*1/2} \tilde{\mathbf{y}}_t$, como o estimador de \mathbf{S} dado ϕ . Por definição, condicionado a (θ, ϕ) os elementos de \mathbf{Q}_t^* são dados por

$$q_{i,i,t} = (1 - a - b)s_{i,i} + aq_{i,i,t-1}\tilde{y}_{q,i,t-1}^2 + bq_{i,i,t-1}, \quad i = 1, \dots, N, \quad (3.14)$$

$$q_{i,j,t} = 0, \quad i \neq j, \quad (3.15)$$

sendo $s_{i,i} = 1 \forall_i$.

Assim, os elementos de \mathbf{Q}_t^* independem de \mathbf{S} . Com isso, é possível calcular $\hat{\mathbf{S}}_q(\hat{\theta}, \phi)$ dado ϕ ; \mathbf{Q}_t dado $\hat{\mathbf{S}}_q(\hat{\theta}, \phi)$, \mathbf{R}_t dado \mathbf{Q}_t e por fim o valor da log-verossimilhança no tempo t . Com isso, utilizando métodos iterativos para maximização de (3.13) é possível estimar ϕ por $\hat{\phi}$. A estimativa de \mathbf{S} é dada pelo valor de $\hat{\mathbf{S}}_q(\hat{\theta}, \hat{\phi})$ na última etapa do processo iterativo.

Análogo ao que fora comentado no caso univariado, as Equações (3.13) e (3.14) não consideram a existência de *outliers*; na presença de saltos aditivos é provável que as estimativas sejam viesadas, mesmo que procedimentos robustos sejam adotados no passo (1), persistindo as complicações na estimação e predição das correlações. Por exemplo, se em apenas um ativo for observado um salto grande nos preços, a estimativa da correlação com os demais é atraída para zero (BOUDT; DANIELSSON; LAURENT, 2013).

O estimador obtido pela minimização de (3.13) pertence a uma classe de estimadores-M cuja definição, apresentada em Boudt e Croux (2010), é dada por

$$\hat{\phi}^* = \underset{\phi \in \Phi}{\operatorname{argmin}} M_T(\phi, \rho) \equiv \left(\frac{1}{T} \sum_{t=1}^T [\log(\det \mathbf{R}_t) + \sigma^* \rho^* (\tilde{\mathbf{y}}'_t \mathbf{R}_t^{-1} \tilde{\mathbf{y}}_t)] \right). \quad (3.16)$$

em que ρ^* é uma função perda escalar contínua, positiva e não decrescente associada ao estimador-M definida nos reais positivos. O fator de correção σ^* é necessário para garantir a consistência do estimador-M sob distribuição de $f(\cdot)$, a verdadeira fdp de ϵ_t , e assegurar que este estima a matriz de covariâncias condicionais dos retornos e não um múltiplo da mesma. A escolha de σ^* depende da verdadeira fdp $f(\cdot)$ e da função perda escolhida. Caso $f(\cdot)$ não seja especificada, o fator de correção precisa ser estimado, para mais detalhes ver Boudt e Croux (2010). O EQMVn é um caso particular para $\sigma^* = 1$ e $\rho^*(z) = z$.

Uma vez calculado $\hat{\phi}$ é possível estimar \mathbf{S} pela matriz de correlação amostral das inovações revolatilizadas $\mathbf{Q}_t^{*1/2} \tilde{\mathbf{y}}'_{q,t}$, dada por

$$\hat{\mathbf{S}}_q(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}) = \frac{1}{T} \sum_{t=1}^N \mathbf{Q}_{j,t}^{*1/2}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}) \tilde{\mathbf{y}}_{q,t} \tilde{\mathbf{y}}_{q,t}^\top \mathbf{Q}_{j,t}^{*1/2}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}), \quad (3.17)$$

em que os retornos desvolatilizados podem ser obtidos utilizando o mesmo artifício apresentado na Equação (3.14). Dado $(\hat{\mathbf{S}}_q, \hat{\boldsymbol{\theta}}_q, \hat{\boldsymbol{\phi}}_q)$, o estimador para \mathbf{H}_t é dado por

$$\hat{\mathbf{H}}_{q,t} = \hat{\mathbf{D}}_{q,t}^{1/2} \hat{\mathbf{R}}_{q,t} \hat{\mathbf{D}}_{q,t}^{1/2},$$

3.5 Estimação robusta dos parâmetros e correlações condicionais

Um procedimento robusto para estimação da matriz de correlações condicionais em casos de séries contaminadas pode ser pensado usando a especificação e procedimento de estimação abordado em [Boudt, Danielsson e Laurent \(2013\)](#), definido pelas seguintes etapas

1. com $\mathbf{y}_t = (y_{1,t}^* - \hat{\mu}_{1,t}, \dots, y_{N,t}^* - \hat{\mu}_{N,t})$, obter $\hat{\boldsymbol{\theta}}_r$ utilizando o estimador-M com função de perda t de Student com 4 graus de liberdade definido na Seção 2.4.2 e especificação BIP-GARCH dada na Equação (2.16). O subscrito “r” indica estimativas do método robusto proposto;
2. calcular matriz de volatilidades estimadas $\hat{\mathbf{D}}_{r,t} \equiv \text{diag}(\hat{h}_{r,1,t}, \dots, \hat{h}_{r,N,t})$ em que cada componente $\hat{h}_{r,i,t}$ é calculado usando a Equação (2.20).
3. calcular retornos desvolatilizados $\tilde{\mathbf{y}}_{r,t} = (\hat{\mathbf{D}}_{r,t})^{-1} \mathbf{y}_t$;
4. obter a estimativa de $\boldsymbol{\phi}$, do modelo cDCC, através de $\hat{\boldsymbol{\phi}}_r = \underset{\boldsymbol{\phi} \in \Phi}{\text{argmax}} L_T(\hat{\boldsymbol{\theta}}_r, \mathbf{S}_r(\hat{\boldsymbol{\theta}}_r, \boldsymbol{\phi}), \boldsymbol{\phi})$, sendo \mathbf{S}_r o estimador robusto da correlação reponderada descrito na Seção 3.5.1;
5. estimar $\hat{\mathbf{H}}_{r,t}$ dado $(\hat{\mathbf{D}}_{r,t}, \hat{\mathbf{S}}_r, \hat{\boldsymbol{\phi}}_r)$.

O modelo BIP-cDCC é uma extensão do modelo BIP-GARCH para o caso multivariado, em que o efeito de $\tilde{\mathbf{y}}_{t-1} \tilde{\mathbf{y}}'_{t-1}$ em \mathbf{Q}_t é limitado; apenas retornos mais extremos são restringidos e, mesmo assim, a matriz de covariância resultante é positiva semidefinida. No modelo BIP-cDCC, a especificação de \mathbf{Q}_t é dada por

$$\mathbf{Q}_t = (1 - a - b)\mathbf{S} + a w_{\delta,N}(d_{t-1}) \mathbf{Q}_{t-1}^{*1/2} \tilde{\mathbf{y}}_{t-1} \tilde{\mathbf{y}}'_{t-1} \mathbf{Q}_{t-1}^{*1/2} + b \mathbf{Q}_{t-1}, \quad (3.18)$$

sendo a função $w_{\delta,N}(\cdot)$ definida, pelas Equações (2.17) e (2.18), dadas na Seção 2.4.2 na estimação robusta dos modelos GARCH, e d_{t-1} é uma medida escalar da discrepância de $\tilde{\mathbf{y}}_{t-1}$ com relação à sua distribuição condicional. Para esta medida é necessário conhecer

sua distribuição. Por exemplo, supondo que a distribuição condicional de \mathbf{r}_{t-1} é normal com vetor de médias zero e matriz de covariância \mathbf{R}_{t-1} , a distância de Mahalanobis pode ser utilizada para identificar extremos em dados com distribuição normal multivariada (HEALY, 1968), assim

$$d_{t-1} = \tilde{\mathbf{y}}'_{t-1} \mathbf{R}_{t-1}^{-1} \tilde{\mathbf{y}}_{t-1}, \quad \text{em que } d_{t-1} \stackrel{i.i.d.}{\sim} \chi_N^2,$$

se d_{t-1} for maior que $k_{\delta, N}$, o quantil de ordem δ da distribuição χ_N^2 , é provável que \mathbf{y}_{t-1} seja um *outlier* e deve ter seu peso reduzido. Como no caso univariado, a escolha de δ regula uma troca entre eficiência e robustez. Para séries não contaminadas, as estimativas obtidas para os parâmetros do modelo BIP-cDCC são viesadas e com alto erro quadrático médio; valores altos para δ aproximam o BIP-cDCC do cDCC, reduzindo o viés. Porém, δ elevado aumenta o efeito de retornos extremos nas correlações estimadas.

3.5.1 Estimador de correlação não condicional reponderada

No modelo cDCC, assim como no modelo BIP-cDCC, a matriz \mathbf{Q}_t depende de \mathbf{S} , que por sua vez é a matriz de correlação de $\mathbf{v}_t = \mathbf{Q}_t^{*1/2} \tilde{\mathbf{y}}_t$. Na modelagem de séries contaminadas a estimativa de \mathbf{S} é potencialmente viesada, sendo necessário um estimador robusto para este componente. Boudt, Danielsson e Laurent (2013) propõem o uso de um estimador robusto da correlação reponderada que é proporcional a correlação amostral de uma janela local de observações no intervalo $[t - K/2, t + K/2]$ em torno de \mathbf{v}_t . Para $t \leq K$ e $t > T - K/2$ os intervalos são dados por $[1, K + 1]$ e $[T - K, T]$, respectivamente. K é um parâmetro de ajuste do tamanho da janela local.

O processo de obtenção do estimador da correlação reponderada é dado em duas etapas. Primeiro, para cada janela, calcula-se a matriz de correlação de Spearman de \mathbf{v}_t , denotada por \mathbf{C}_t . O fator de correção a seguir é aplicado para garantir consistência

$$\mathbf{SC}_t = 2 \operatorname{sen} \left(\frac{1}{6} \pi \mathbf{C}_t \right).$$

Posteriormente, calcula-se o estimador da correlação quasi-reponderada

$$\mathbf{RC} = \frac{c_{0,95,N}}{\sum_{t=1}^T L_t} \sum_{t=1}^T \tilde{\mathbf{y}}_t \tilde{\mathbf{y}}'_t L_t,$$

em que $L_t = I[\tilde{\mathbf{y}}'_t S C_t^{-1} \tilde{\mathbf{y}}_t \leq \chi_N^2(0, 95)]$ decorrente da suposição de normalidade multivariada imposta; e $c_{0,95,N}$ definida como em (2.18). O estimador da matriz de correlação não condicional reponderada é dado por

$$\hat{\mathbf{S}}_r(\boldsymbol{\theta}, \boldsymbol{\phi}) = \operatorname{diag}(RC_{11}^{-1/2}, \dots, RC_{NN}^{-1/2}) RC \operatorname{diag}(RC_{11}^{-1/2}, \dots, RC_{NN}^{-1/2}). \quad (3.19)$$

3.5.2 Estimador-M para o modelo BIP-cDCC

O estimador-M robusto é dado por

$$\hat{\phi}_r = \underset{\phi \in \Phi}{\operatorname{argmin}} \left(\frac{1}{T} \sum_{t=1}^T \rho_3(\mathbf{R}_t, \tilde{\mathbf{y}}_t) \right), \quad (3.20)$$

em que o subscrito “r” indica estimativas de quase-verossimilhança robustas e

$$\rho_3(\mathbf{R}, \mathbf{y}) = \log(\det \mathbf{R}) + \sigma_{N,4} \rho_{t_{N,4}}(\mathbf{y}' \mathbf{R}^{-1} \mathbf{y}), \quad (3.21)$$

$$\rho_{t_{N,v}}(u) = (N+v) \log \left(1 + \frac{u}{v-2} \right), \quad (3.22)$$

$$\sigma_{N,v} = \frac{N}{E[\rho'_{t_{N,v}}(U)U]}, \quad (3.23)$$

$$U \sim \chi_N^2, \quad (3.24)$$

com $v = 4$ e sendo χ_N^2 uma distribuição qui-quadrado com N graus de liberdade.

A motivação para escolha de $\sigma_{N,v}$ em (3.23) vem da suposição de que $f(\cdot)$, a verdadeira fdp de ϵ_t , pertence à família de distribuições t de Student com v graus de liberdade. A robustez da função de perda $\rho_{t_{N,v}}$ está atrelada a capacidade da mesma de reduzir a influência de *outliers* nas estimativas-M. Para obter $\hat{\phi}_r$, a solução para o problema de minimização em (3.20), a condição de primeira ordem precisa ser satisfeita, i.e

$$\frac{\partial M_T(\phi, \rho)}{\partial \phi_i} = \frac{1}{T} \sum_{t=1}^T \text{Tr} \left[I_N - \sigma^* \rho^{**}(\tilde{\mathbf{y}}_t \mathbf{R}_t^{-1} \tilde{\mathbf{y}}_t) \tilde{\mathbf{y}}_t \tilde{\mathbf{y}}_t' \mathbf{R}_t^{-1} \right] \frac{\partial \mathbf{H}_t}{\partial \phi_i} \mathbf{H}_t^{-1} = 0,$$

sendo $\text{Tr}(\cdot)$ o operador para o traço de uma matriz. Se a t -ésima observação é afetada por um *outlier* então o estimador-M será robusto a este se $\rho^{**}(\tilde{\mathbf{y}}_t' \mathbf{R}_t^{-1} \tilde{\mathbf{y}}_t)$ limitar suficientemente $\tilde{\mathbf{y}}_t \tilde{\mathbf{y}}_t'$; e isto ocorre para $\rho^*(z)$ com derivada decrescente para zero a uma taxa de pelo menos $1/z$. Para o QMVn, em que $\rho^*(z) = 1$, o estimador-M atribui peso igual a todas as observações sem distinguir dados contaminados de não contaminados, por isso é muito sensível a *outliers*. Adotando $\rho^*(z) = \rho_{t_{N,v}}(z)$ definido em (3.22) têm-se $\rho^*(z) = (N+v)/(v+z-2)$ que decresce para zero a um taxa de $1/z$, além disso quanto menor v for, mais robusto será o estimador-M com função perda t_v .

Assim como no cDCC, a estimação de $\mathbf{S}(\theta, \phi)$ e ϕ é feita de forma simultânea, iniciando com a estimação de $\mathbf{S}(\theta, \phi)$ como a matriz de correlação reponderada de $\mathbf{Q}_1^{*1/2} \tilde{\mathbf{y}}_1, \dots, \mathbf{Q}_t^{*1/2} \tilde{\mathbf{y}}_t$; cujos elementos diagonais de \mathbf{Q}_t podem obtidos de forma similar a (3.14), mas utilizando a formulação BIP dada por

$$q_{i,i,t} = (1 - a - b) + a q_{i,i,t-1} \omega_{\delta,1}(\tilde{y}_{i,t-1}^2) \tilde{y}_{i,t-1}^2 + b q_{i,i,t-1}, \quad i = 1, \dots, N. \quad (3.25)$$

O procedimento itera até a convergência através da estimação de $\mathbf{S}(\theta, \phi)$ e estimação-M Student t_4 de ϕ usando a especificação BIP-cDCC. Assim como o processo

de estimação no cDCC(1,1) a expressão (3.25) torna possível a definição de $\mathbf{S}(\boldsymbol{\theta}, \boldsymbol{\phi})$ que é exigida na Equação recursiva definidora do processo em (3.18). Mas note que diferentemente da especificação não robusta, os valores obtidos utilizando (3.25) não são exatamente iguais aos elementos diagonais em \mathbf{Q}_t , justamente pela diferença do filtros de volatilidade utilizados em cada equação. Dado isso, um vez calculada $\mathbf{S}(\boldsymbol{\theta}, \boldsymbol{\phi})$, os elementos diagonais de \mathbf{Q}_t são recalculados em cada tempo. O estimador robusto proposto de \mathbf{H}_t , condicionado a $(\mathbf{S}(\hat{\boldsymbol{\theta}}_r, \hat{\boldsymbol{\phi}}_r), \hat{\boldsymbol{\theta}}_r, \hat{\boldsymbol{\phi}}_r)$, é dado por

$$\hat{\mathbf{H}}_{r,t} = \hat{\mathbf{D}}_{r,t}^{1/2} \hat{\mathbf{R}}_{r,t} \hat{\mathbf{D}}_{r,t}^{1/2}.$$

4 Estimação em Alta Dimensão

Os métodos de estimação apresentados nas Seções 3.4-3.5 são viáveis em aplicações com número moderado de séries. Porém, computacionalmente intensivos para conjuntos de dados de alta dimensão. O principal motivo é o cálculo de H_t^{-1} , em todos os tempos, para avaliação da Equação (3.13) em cada etapa do processo iterativo até a convergência numérica, com custo computacional de $O(N^3)$. Além disso, as estimativas obtidas são viesadas. A redução significativa do viés requer uma amostra maior do que disponível em muitas aplicações (PAKEL et al., 2021). Como alternativa, Pakel et al. (2021) propõem utilizar o método da verossimilhança composta (*composite likelihood*, CL) (LINDSAY, 1988; COX; REID, 2004; VARIN; VIDONI, 2005; VARIN; REID; FIRTH, 2011), que aproxima a verossimilhança completa utilizando combinações lineares de densidades marginais de menor dimensão. Por exemplo, uma log-CL básica considerando variáveis aleatórias independentes, é dada por

$$\frac{1}{N} \sum_{i=1}^N l_i T(\cdot), \quad (4.1)$$

sendo $l_i T(\cdot)$ a log-verossimilhança marginal associada à i -ésima série de retornos. Pakel et al. (2021) propõem construir a função log-CL a partir de densidades bivariadas. Seja L a quantidade de pares distintos em N séries de retornos e $\mathbf{y}_{i,t} = (y_{i_1,t}, y_{i_2,t})'$ o vetor de retornos observados no i -ésimo par e no tempo t , tal que $i = 1, \dots, L$, $(i_1, i_2) \in \{1, \dots, N\}^2$ e $i_1 \neq i_2$; o método de estimação envolvendo a Equação (3.13) pode ser feito utilizando a função log-CL definida na Equação (4.1) e

$$l_{i,T}(\boldsymbol{\theta}_i, \boldsymbol{\phi}) = \frac{1}{T} \sum_{t=1}^T [\log |\mathbf{H}_{i,t}| + \mathbf{y}_{i,t} \mathbf{H}_{i,t}^{-1} \mathbf{y}'_{i,t}], \quad (4.2)$$

sendo $\mathbf{H}_{i,t}$ a matriz de covariância condicional e $l_{i,T}(\boldsymbol{\theta}_i, \boldsymbol{\phi})$ a log-verossimilhança do i -ésimo par.

No decorrer desta seção serão apresentados algumas técnicas que são comuns a ambos os métodos de estimação apresentados até o momento. Portanto, nesses casos, optou-se pela omissão dos subscritos “q” e “r”.

4.1 Definição e estimação dos modelos

Supondo que $y_{i_1,t}$ e $y_{i_2,t}$ são gerados por processos GARCH com parâmetros $\boldsymbol{\eta}_{i_1}$ e $\boldsymbol{\eta}_{i_2}$, respectivamente, e que $\boldsymbol{\theta}_i = (\boldsymbol{\eta}'_{i_1}, \boldsymbol{\eta}'_{i_2})'$. Os retornos desvolatilizados do i -ésimo par são dados por

$$\tilde{\mathbf{y}}_{i,t}(\boldsymbol{\theta}_i) = (\tilde{y}_{i_1,t}, \tilde{y}_{i_2,t})' = \hat{\mathbf{D}}_{i,t}^{-1/2} \mathbf{y}_{i,t}, \quad (4.3)$$

em que

$$\hat{\mathbf{D}}_{i,t} = \begin{bmatrix} \hat{h}_{i_1,t} & 0 \\ 0 & \hat{h}_{i_2,t} \end{bmatrix}, \quad (4.4)$$

sendo $\hat{h}_{i,t}$ a volatilidade estimada no tempo t para a i_1 -ésima série.

Para cada $\mathbf{y}_{i,t}$ serão ajustados modelo GARCH(1,1)-cDCC(1,1) e BIP-GARCH(1,1)-BIP-cDCC(1,1) com estrutura e processo de estimação descritos nas seguintes etapas:

1. Equações para volatilidades:

GARCH(1,1)

$$\begin{aligned} h_{i_1,t} &= \omega_{i_1} + \alpha_{i_1} y_{i_1,t-1}^2 + \beta_{i_1} h_{i_1,t-1}, \\ h_{i_2,t} &= \omega_{i_2} + \alpha_{i_2} y_{i_2,t-1}^2 + \beta_{i_2} h_{i_2,t-1} \end{aligned}$$

BIP-GARCH(1,1)

$$\begin{aligned} h_{i_1,t} &= \omega_{i_1} + \alpha_{i_1} w_{\delta,1} \left(\frac{y_{i_1,t-1}^2}{h_{i_1,t-1}} \right) y_{i_1,t-1}^2 + \beta_{i_1} h_{i_1,t-1}, \\ h_{i_2,t} &= \omega_{i_2} + \alpha_{i_2} w_{\delta,1} \left(\frac{y_{i_2,t-1}^2}{h_{i_2,t-1}} \right) y_{i_2,t-1}^2 + \beta_{i_2} h_{i_2,t-1} \end{aligned}$$

com $\boldsymbol{\eta}_{i_1} = (\omega_{i_1}, \alpha_{i_1}, \beta_{i_1})'$ e $\boldsymbol{\eta}_{i_2} = (\omega_{i_2}, \alpha_{i_2}, \beta_{i_2})'$.

2. Equação para correlações:

cDCC(1,1)

$$\begin{aligned} \mathbf{H}_{i,t}(\boldsymbol{\theta}_i, \mathbf{S}_i(\boldsymbol{\theta}_i, \boldsymbol{\phi}), \boldsymbol{\phi}) &= \mathbf{D}_{i,t}^{1/2}(\boldsymbol{\theta}_i) \mathbf{R}_{i,t}(\boldsymbol{\theta}_i, \mathbf{S}_i(\boldsymbol{\theta}_i, \boldsymbol{\phi}), \boldsymbol{\phi}) \mathbf{D}_{i,t}^{1/2}(\boldsymbol{\theta}_i) \\ \mathbf{R}_{i,t}(\boldsymbol{\theta}_i, \mathbf{S}_i(\boldsymbol{\theta}_i, \boldsymbol{\phi}), \boldsymbol{\phi}) &= \mathbf{Q}_{i,t}^{*1/2} \mathbf{Q}_{i,t}(\boldsymbol{\theta}_i, \mathbf{S}_i(\boldsymbol{\theta}_i, \boldsymbol{\phi}), \boldsymbol{\phi}) \mathbf{Q}_{i,t}^{*1/2}, \\ \mathbf{Q}_{i,t}(\boldsymbol{\theta}_i, \mathbf{S}_i(\boldsymbol{\theta}_i, \boldsymbol{\phi}), \boldsymbol{\phi}) &= (1 - a - b) \mathbf{S}_i(\boldsymbol{\theta}_i, \boldsymbol{\phi}) \\ &\quad + a (\mathbf{Q}_{i,t-1}^{*1/2}(\boldsymbol{\theta}_i, \boldsymbol{\phi}) \tilde{\mathbf{y}}_{i,t-1}'(\boldsymbol{\theta}_i) \tilde{\mathbf{y}}_{i,t-1}(\boldsymbol{\theta}_i) \mathbf{Q}_{i,t-1}^{*1/2}(\boldsymbol{\theta}_i, \boldsymbol{\phi})) \\ &\quad + b \mathbf{Q}_{i,t-1}(\boldsymbol{\theta}_i, \boldsymbol{\phi}), \\ \mathbf{S}_i(\boldsymbol{\theta}_i, \boldsymbol{\phi}) &= E \left[\mathbf{Q}_{i,t}^{*1/2}(\boldsymbol{\theta}_i, \boldsymbol{\phi}) \tilde{\mathbf{y}}_{i,t}(\boldsymbol{\theta}_i) \tilde{\mathbf{y}}_{i,t}'(\boldsymbol{\theta}_i) \mathbf{Q}_{i,t}^{*1/2}(\boldsymbol{\theta}_i, \boldsymbol{\phi}) \right]. \end{aligned}$$

BIP-cDCC(1,1)

Para o modelo BIP-cDCC(1,1) apenas $\mathbf{Q}_{i,t}$ sofre modificação, passando a incorporar a distância de Mahalanobis associada ao i -ésimo par

$$\begin{aligned} d_{i,t-1} &= \tilde{\mathbf{y}}_{i,t}'(\boldsymbol{\theta}_i) \mathbf{R}_{i,t}^{-1}(\boldsymbol{\theta}_i, \mathbf{S}_i(\boldsymbol{\theta}_i, \boldsymbol{\phi}), \boldsymbol{\phi}) \tilde{\mathbf{y}}_{i,t}(\boldsymbol{\theta}_i) \\ \mathbf{Q}_{i,t}(\boldsymbol{\theta}_i, \mathbf{S}_i(\boldsymbol{\theta}_i, \boldsymbol{\phi}), \boldsymbol{\phi}) &= (1 - a - b) \mathbf{S}_i(\boldsymbol{\theta}_i, \boldsymbol{\phi}) \\ &\quad + aw_{\delta,2}(d_{i,t-1})(\mathbf{Q}_{i,t-1}^{*1/2}(\boldsymbol{\theta}_i, \boldsymbol{\phi}) \tilde{\mathbf{y}}_{i,t-1}'(\boldsymbol{\theta}_i) \tilde{\mathbf{y}}_{i,t-1}(\boldsymbol{\theta}_i) \mathbf{Q}_{i,t-1}^{*1/2}(\boldsymbol{\theta}_i, \boldsymbol{\phi})) \\ &\quad + b \mathbf{Q}_{i,t-1}(\boldsymbol{\theta}_i, \boldsymbol{\phi}). \end{aligned}$$

Consequentemente os elementos diagonais de $\mathbf{Q}_{i,t}$ também diferem dos do modelo cDCC(1,1) pela incorporação de formulação BIP, que para cada par é dada por

$$q_{i,j,j,t} = (1 - a - b) + aq_{i,j,j,t-1}\omega_{\delta,1}(\tilde{y}_{j,t-1}^2)\tilde{y}_{j,t-1}^2 + bq_{i,j,j,t-1}, \quad j = 1, 2.$$

3. Estimação dos parâmetros:

A estimação por CL segue a estrutura dos procedimentos apresentados nas Seções 3.4 e 3.5 análogos a cada par, com uma modificação no cálculo da verossimilhança conjunta completa, i.e

- a) Estimar $\boldsymbol{\theta}_i$ por QMVn ou estimação robusta para BIP-GARCH descrito na Subseção 2.4.2; posteriormente calcular os retornos desvolatilizados estimados $\tilde{\mathbf{y}}_{i,t}(\hat{\boldsymbol{\theta}}_i) = (\tilde{y}_{i_1,t}(\hat{\boldsymbol{\eta}}_{i_1}), \tilde{y}_{i_2,t}(\hat{\boldsymbol{\eta}}_{i_2}))'$, $\tilde{y}_{i_1,t}(\hat{\boldsymbol{\eta}}_{i_1}) = y_{i_1,t}h_{i_1,t}^{-1/2}(\hat{\boldsymbol{\eta}}_{i_1})$.
- b) Em seguida $\boldsymbol{\phi}$ é estimado utilizando as Equações (4.1), (4.2) e o estimador por CL dado por

$$\hat{\boldsymbol{\phi}} = \operatorname{argmin} \frac{1}{L} \sum_{j=1}^L l_{iT}(\hat{\boldsymbol{\theta}}_i, \hat{\mathbf{S}}_j(\hat{\boldsymbol{\theta}}_j, \boldsymbol{\phi}), \boldsymbol{\phi}), \quad (4.5)$$

com $\hat{\boldsymbol{\phi}}$ a matriz de correlação não condicional de $\mathbf{Q}_{i,t}^{*1/2}\tilde{\mathbf{y}}_t$ pode ser estimada, para cada par, utilizando

$$\mathbf{S}_i(\hat{\boldsymbol{\theta}}_i, \boldsymbol{\phi}) = \frac{1}{T} \sum_{t=1}^T \mathbf{Q}_{j,t}^{*1/2}(\hat{\boldsymbol{\theta}}_j, \boldsymbol{\phi}) \tilde{\mathbf{y}}_{i,t}(\hat{\boldsymbol{\theta}}_i) \tilde{\mathbf{y}}'_{i,t}(\hat{\boldsymbol{\theta}}_i) \mathbf{Q}_{j,t}^{*1/2}(\hat{\boldsymbol{\theta}}_j, \boldsymbol{\phi}), \quad (4.6)$$

O estimador-M para $\boldsymbol{\phi}$ por CL pode ser obtido utilizando a função de verossimilhança expressa em (3.20) análoga a cada par, i.e,

$$l_{iT} = \frac{1}{T} \sum_{t=1}^T [\log |\mathbf{R}_{i,t}| + \sigma_{2,4}\rho(d_{i,t})]$$

A matriz de correlação não condicional $\mathbf{S}_i(\boldsymbol{\theta}_i, \boldsymbol{\phi})$ também pode ser estimada utilizando o estimador robusto apresentado na Seção 3.5.

4. Configuração dos Pares:

Existem algumas configurações que podem ser adotadas na definição dos pares da CL, Pakel et al. (2021) realizam experimentos de Monte Carlo comparando viés, desvio padrão e outras estatísticas obtidas usando estimadores de $\boldsymbol{\phi}$ obtidos sob verossimilhança completa (2MLE), CL para todos os pares (2MCLE) e CL com pares consecutivos (2MSCLE). O que se observou foi o deterioramento do viés para 2MLE com o aumento de N ; por exemplo, para $N = 200$ o viés de estimação para o parâmetro a é de cerca de 30% mesmo com $T = 2000$. Em contra partida 2MCLE e 2MSCLE não resultaram em estimativas significativamente viesadas. Mais do que isso,

ambos apresentam pouca diferença em termos de EQM. Dado isso, neste trabalho optou-se por utilizar 2MSCLE tal que $X_{1t} = (y_{1t}, y_{2t})'$, $X_{2t} = (y_{2t}, y_{3t})'$, ..., $X_{Lt} = (y_{N-1,t}, y_{N,t})'$, por ser uma opção mais atrativa em termos de eficiência na etapa de estimação e menor custo computacional.

4.2 Matriz de correlação condicional completa

Note que por 2MSCLE em nenhum momento do processo de estimação de ϕ trabalhou-se com matrizes de dimensão de \mathbf{r}_t , i.e., $N \times N$. A matriz completa de correlações condicionais \mathbf{H}_t é estimada dado $(\hat{\phi}, \hat{\theta})$, estimadores quaisquer para (ϕ, θ) . Mas antes se faz necessário estimar $\mathbf{S}(\theta, \phi)$ a matriz completa de correlações não condicionais dos retornos revolatilizado $\mathbf{Q}_t^{*1/2}\tilde{\mathbf{y}}_t$, que por sua vez é função de \mathbf{Q}_t^* . Resumindo, condicionado a $(\hat{\phi}, \hat{\theta})$ todos os passos no processo de estimação precisam ser refeitos, mas na dimensão original do problema. Portanto, inicialmente calcula-se $\mathbf{Q}_t^{*1/2}$ explorando o fato de $\hat{\mathbf{S}}$ ser diagonal unitária, assim

$$\mathbf{Q}_t^{*1/2} = (1 - \hat{a} - \hat{b}) + \hat{a}\tilde{\mathbf{y}}_{t-1}'\tilde{\mathbf{y}}_{t-1} + \hat{b}\mathbf{Q}_{t-1}^{*1/2}. \quad (4.7)$$

Utilizando estes valores, estima-se $\mathbf{S}(\hat{\theta}, \hat{\phi})$. Por fim, para cada tempo t , o estimador para \mathbf{H}_t é dada por

$$\begin{aligned} \hat{\mathbf{H}}_t(\hat{\theta}, \hat{\mathbf{S}}(\hat{\theta}, \hat{\phi}), \hat{\phi}) &= \hat{\mathbf{D}}_t^{1/2}(\hat{\theta})\hat{\mathbf{R}}_t(\hat{\theta}, \hat{\mathbf{S}}(\hat{\theta}, \hat{\phi}), \hat{\phi})\hat{\mathbf{D}}_t^{1/2}(\hat{\theta}) \\ \hat{\mathbf{R}}_t(\hat{\theta}, \hat{\mathbf{S}}(\hat{\theta}, \hat{\phi}), \hat{\phi}) &= \mathbf{Q}_t^{*1/2}\mathbf{Q}_t(\hat{\theta}, \hat{\mathbf{S}}(\hat{\theta}, \hat{\phi}), \hat{\phi})\mathbf{Q}_t^{*1/2}, \\ \mathbf{Q}_t(\hat{\theta}, \hat{\mathbf{S}}(\hat{\theta}, \hat{\phi}), \hat{\phi}) &= (1 - \hat{a} - \hat{b})\hat{\mathbf{S}}(\hat{\theta}, \hat{\phi}) \\ &\quad + \hat{a}(\mathbf{Q}_{t-1}^{*1/2}(\hat{\theta}, \hat{\phi})\tilde{\mathbf{y}}_{t-1}(\hat{\theta})\tilde{\mathbf{y}}_{t-1}'(\hat{\theta})\mathbf{Q}_{t-1}^{*1/2}(\hat{\theta}, \hat{\phi})) \\ &\quad + \hat{b}\mathbf{Q}_{t-1}(\hat{\theta}, \hat{\phi}), \\ \mathbf{S}(\theta, \phi) &= E\left[\mathbf{Q}_t^{*1/2}(\theta, \phi)\tilde{\mathbf{y}}_t(\theta)\tilde{\mathbf{y}}_t'(\theta)\mathbf{Q}_t^{*1/2}(\theta, \phi)\right]. \end{aligned}$$

Em que $\mathbf{S}(\theta, \phi)$ pode ser estimada utilizando a matriz de variância e covariâncias amostral dada por

$$\hat{\mathbf{S}}(\hat{\theta}, \hat{\phi}) = T^{-1} \sum_{t=1}^T \mathbf{Q}_t^{*1/2}(\hat{\theta}, \hat{\phi})\tilde{\mathbf{y}}_t(\hat{\theta})\tilde{\mathbf{y}}_t'(\hat{\theta})\mathbf{Q}_t^{*1/2}(\hat{\theta}, \hat{\phi}),$$

mas complicações também aparecem nesta etapa quando se trabalha com um grande número de séries. Mas como destacado em [Engle, Ledoit e Wolf \(2019\)](#) a matriz de correlação amostral apresenta sobreajuste com o comprometimento do desempenho fora da amostra, que é geralmente o cenário de investimento e administração de portfólio.

Basicamente, a relação entre o número de séries (dimensão do problema) e de tempos rege a eficácia do estimador de correlação amostral. Para $O(N^2)$ observações não é possível estimar os $N(N - 1)/2$ parâmetros, que é um cenário conhecido como a maldição

da dimensionalidade. Pela teoria de matrizes aleatórias, $T/N < 0,01$ a quantidade amostral é uma boa aproximação da real assintoticamente. Em outros casos, métodos especiais precisam ser adotados para sanar problemas de sobreajuste. Em casos mais extremos, quando $T/N > 1$ a matriz de correlação é singular.

Dado isso, a matriz de correlação amostral apresenta viés justificado no campo dos autovalores. Basicamente, os menores (maiores) autovalores amostrais são muito pequenos (grandes). Uma solução se dá ao aumentar os menores autovalores e diminuir os maiores reduzindo assim a dispersão da distribuição dos autovalores. Este procedimento é conhecido como *shrinkage*.

4.2.1 Estimador Shrinkage

Seja $\mathbf{C} := [\epsilon_{i,t}]$ uma matriz de dimensão $N \times T$ tal que $E[\epsilon_{i,t}] = 0$ e $E[\epsilon_{i,t}^2] = 1$ para $i = 1, \dots, N$ e $t = 1, \dots, N$. A matriz de correlação amostral é dada por

$$\hat{\mathbf{S}} = \frac{1}{T} \mathbf{C} \mathbf{C}', \quad (4.8)$$

que pode ser decomposta em autovalores $(\lambda_1, \dots, \lambda_N)$, tal que $\lambda_1 > \lambda_2 > \dots, \lambda_N$, e autovetores correspondentes (u_1, u_2, \dots, u_N) , de modo que de (4.8) pode ser substituída por

$$\hat{\mathbf{S}} := \sum_{i=1}^N \lambda_i u_i u_i', \quad (4.9)$$

em que $\lambda_i = u_i' \hat{\mathbf{S}} u_i$ pode ser interpretada como a variância (dentro da amostra) de um portfólio com pesos dados pelo vetor u_i . O estimador *shrinkage* busca substituir λ_i por $u_i' \mathbf{S} u_i$ ou a variância (fora da amostra) do portfólio. Com isso o desafio se resume a estimar os N autovalores ao invés de \mathbf{S} que tem dimensão $N \times N$. Dado isso, os autovalores de \mathbf{S} , denotados por $\boldsymbol{\tau} = (\tau_1, \dots, \tau_N)'$, podem ser estimados utilizando uma função multivariada não aleatória denominada QuEST (*Quantized Eigenvalues Sampling Transform*) proposta em [Ledoit e Wolf \(2015\)](#) e denotada por $Q_{N,T}$ que é definida por

$$\begin{aligned} Q_{N,T} : [0, \infty)^N &\longrightarrow [0, \infty)^N \\ \mathbf{t} := (t_1, \dots, t_N)' &\longrightarrow Q_{N,T}(\mathbf{t}) := (q_{N,T}^1(\mathbf{t}), \dots, q_{N,T}^N(\mathbf{t})), \end{aligned}$$

em que

$$\begin{aligned} q_{N,T}^i(\mathbf{t}) &:= N \int_{(i-1)/N}^{i/N} (F_{N,T}^{\mathbf{t}})^{-1}(u) du \\ (F_{N,T}^{\mathbf{t}})^{-1}(u) &:= \sup \{x \in \mathbb{R} : F_{N,T}^{\mathbf{t}}(x) \leq u\}, \end{aligned}$$

sendo $F_{N,T}^{\mathbf{t}}$ a função de distribuição limite para os autovalores amostrais correspondentes a função de distribuição espectral $1/p \sum_{i=1}^N I[t_i, +\infty]$. A função QuEST recebe como argumento

um conjunto de autovalores populacionais $\mathbf{t} := (t_1, \dots, T_N)$ e retorna um equivalente determinístico dos autovalores amostrais $Q_{N,T}(\mathbf{t}) = (q_{N,T}^1(\mathbf{t}), \dots, q_{N,T}^N(\mathbf{t}))$. Dessa forma, os autovalores populacionais podem ser estimados invertendo numericamente a função QuEST

$$\hat{\boldsymbol{\tau}} := \underset{\mathbf{t} \in [0, \infty)}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^N [q_{N,T}^i(\mathbf{t}) - \lambda_i]^2.$$

Dado os autovalores populacionais estimados $\hat{\boldsymbol{\tau}}$, é possível utilizar o Teorema 4 de [Ledoit e Péché \(2011\)](#) para calcular uma formula *shrinkage* não linear que é assintoticamente ótima quando N e T tendem a infinito. Os autovalores *shrunk* são $\hat{\boldsymbol{\lambda}} := (\hat{\lambda}_1(\boldsymbol{\tau}), \dots, \hat{\lambda}_1(\boldsymbol{\tau}))'$. Pelo Teorema 2.1 em [Ledoit e Wolf \(2015\)](#) têm-se que

$$\frac{1}{N} \sum_{i=1}^N [\hat{\tau}_i - \tau_i]^2 \xrightarrow{a.s.} 0, \quad (4.10)$$

indicando que os autovalores amostrais convergem individualmente para seus análogos não aleatórios na função QuEST. A ideia de convergência é estabelecida utilizando a distância Euclidiana entre $\boldsymbol{\lambda}$ e $Q_{N,T}(\boldsymbol{\tau})$ normalizado por N . Dessa forma, os estimadores, por (4.10), convergem individualmente para os autovalores populacionais. Portanto, o estimador *shrinkage* para \mathbf{S} é dado por

$$\hat{\mathbf{S}}_s := \sum_{i=1}^N \hat{\lambda}_i(\hat{\boldsymbol{\tau}}) u_i u_i'. \quad (4.11)$$

Dois atributos importante deste estimador são sua capacidade de lidar com aplicações em que $T/N < 1$ e não exigir normalidade.

Uma alternativa mais simples é utilizar o estimador *skrinkage* linear ([LEDOIT; WOLF, 2004](#)), dado por

$$\hat{\mathbf{S}}_{sl} := \sum_{i=1}^N [\rho \bar{\lambda} + (1 - \rho) \lambda_i] u_i u_i', \quad (4.12)$$

sendo $\bar{\lambda}$ a média dos autovalores, i.e

$$\bar{\lambda} := \frac{1}{N} \sum_{i=1}^N \lambda_i.$$

O estimador consistente ρ representa a intensidade do encolhimento (*shrinkage*), sendo uma quantidade entre zero e um que controla o quanto os autovalores são aproximados de $\bar{\lambda}$. Apesar de popular, o estimador *skrinkage* linear é ótimo apenas em dimensão dois. Portanto, assintoticamente, quando N e T crescem conjuntamente o estimador não linear deve apresentar melhores resultados ([ENGLE; LEDOIT; WOLF, 2019](#)). A justificativa se dá por ρ ser restrito a um mesmo valor para todos os autovalores, enquanto no não linear este é individualmente ajustados para cada λ_i .

4.2.2 Estimador Wrapping

Estimar a matriz intercepto \mathbf{S} por *Shrinkage* auxilia no processo de estimação da mesma quando a dimensão da aplicação é elevada; mas este procedimento não é robusto a presença de *outliers*. E novamente, a importância de estimar este componente corretamente vem desde as etapas de estimação dos parâmetros dos modelo; lembrando da matriz de correlação não condicional reponderada discutido em [Boudt, Danielsson e Laurent \(2013\)](#) e apresentada neste trabalho na Seção 3.5.1. Esta ferramente poderia ser utilizada agora, o problema reside na componente K que regula o tamanho da janela local; para uma aplicação com $N = 1.000$ e $T = 1.250$ estimar \mathbf{S} completa torna-se inviável mesmo escolhendo um K mais elevado. Portanto, se faz necessário utilizar uma abordagem um pouco diferente. Relembre que em determinada etapa do cálculo da matriz de correlação não condicional reponderada utilizou-se a matriz de correlação de Spearman de $\mathbf{Q}_t^{*1/2}\tilde{\mathbf{y}}_t$; que na verdade é a matriz de correlação de Pearson de uma transformação do próprio ($\mathbf{Q}_t^{*1/2}\tilde{\mathbf{y}}_t$). A motivação portanto seria a escolha de uma transformação g adequada que não seja tão restrita quanto os postos de Spearman. Transformações g aplicada a variáveis individualmente são denotadas por produto- g . Versões mais antigas de produtos- g robustos são funções da correlação por postos como em [Spearman \(1904\)](#). Outro tipo, já comentando na Seção 2.4.3 decorre da contribuição de [Huber e Ronchetti \(2009, Seção 8.3\)](#) baseado em estimação-M, com a transformada de Huber para a i -ésima observação de uma série de retornos r_1, \dots, r_T sendo definida por

$$g(r_i) = \psi((r_i - \hat{\mu})/\hat{\sigma}). \quad (4.13)$$

Dado isso, [Raymaekers e Rousseeuw \(2021\)](#) propõem uma função ψ adequada para correlações; eliminando o efeito de *outliers* a c desvios padrão robustos de distância. A função ψ é dada por

$$\psi_{b,c}(z) = \begin{cases} z, & \text{se } 0 \neq |z| \leq b, \\ q_1 \tanh(q_2(c - |z|))\text{sign}(z), & \text{se } b \leq |z| \leq c, \\ 0, & \text{se } c \leq |z|, \end{cases}$$

para $q_1 = \sqrt{A(k-1)}$ e $q_2 = (B/2)\sqrt{(k-1)/A}$; sendo A, B e k constantes dadas por

$$\begin{aligned} A &= \int_{-c}^c \psi(x)^2 d\Phi(x), \\ B &= \int_{-c}^c \psi'(x)^2 d\Phi(x), \\ \kappa^*(\psi) &= k, \end{aligned}$$

com escolhas padrões de $b = 1,5$ e $c = 4$ ([RAYMAEKERS; ROUSSEUW, 2021](#)); em particular $c = 4$ elimina observações a mais de 4 desvios padrão. Sob normalidade, $\psi_{b,c}$

mantém inalterado 86,6% das observações, dado $b = 1,5$; com isso conservando grande parte da informação das distribuições das variáveis e o tipo de associação entre as mesmas (ex: linear). Para o intervalo (b, c) são ponderados em direção ao centro, tornando assim *outliers* intermediários ainda influentes no cálculo do coeficiente. Por esta razão, $\psi_{b,c}$ é conhecida como função *wrapping*; envolvendo os dados em torno do intervalo $[-b, b]$.

Portanto, \mathbf{S} pode ser estimada simplesmente calculando-se a matriz de correlação de Pearson na amostra transformada definida em (4.13). Dessa forma o estimador *wrapping* de \mathbf{S} é dado por

$$\hat{\mathbf{S}}_w = T^{-1} \sum_{t=1}^T \mathbf{x}_t^* \mathbf{x}_t^{*\prime}, \quad (4.14)$$

em que $\mathbf{x}_t^* = [x_{1,t}^*, \dots, x_{N,t}^*]$ tal que

$$\begin{aligned} x_{i,t}^* &= g(x_{i,t}), \\ x_{i,t} &= q_{i,i,t} * \tilde{y}_{i,t}, \quad i = 1, \dots, N, \end{aligned}$$

sendo $q_{i,i,t}$ o i -ésimo elemento na diagonal da matriz \mathbf{Q}_t^* .

Alguns trabalhos utilizaram este estimador em contextos relacionados como Trucíos et al. (2021).

5 Simulação para Modelos GARCH e cDCC Bivariado

Embora o objetivo desta dissertação seja a previsão de \mathbf{H}_t para N elevado; neste capítulo optou-se por realizar um estudo inicial dos estimadores tradicional e robusto apresentados nos Capítulos 2 e 3 para $N = 2$. Desta forma, será realizado um exercício simulação que considera as partes de estimação dos parâmetros (univariadas e multivarada) e previsão de \mathbf{H}_t . Os resultados dos estudos de simulação de Monte Carlo destinados a comparar o desempenho dos estimadores apresentados nas Seções 2.3 e 2.4.2 para modelos GARCH, e nas Seções 3.4 e 3.5 para modelos cDCC.

Na Seção 5.1 são apresentados os processos geradores de dados que dividem o estudo em duas partes: na primeira é estudada a estimação dos parâmetros e na segunda a estimação das volatilidades e correlações condicionais no caso bivariado. Na Seção 5.2 são apresentados os estimadores, enquanto na Seção 5.3 as medidas de desempenho que avaliam diferentes atributos do processo de estimação como qualidade das estimativas, desempenho do filtro de volatilidade no caso robusto e performance global na estimação da matriz \mathbf{H}_t . Na Seção 5.4 são apresentados os resultados de estimação dos parâmetros da parte GARCH, enquanto na Seção 5.5 são apresentados todos os resultados associados a parte cDCC: estimação de ϕ , estimação de \mathbf{H}_t e desempenho do filtro de volatilidade.

5.1 Processos geradores de dados

Os modelos adotados foram: (1) GARCH(1,1) para variâncias condicionais e (2) cDCC(1,1) para correlações condicionais. O plano de contaminação difere com a dimensão do problema. A seguir são apresentados os esquemas para o estudo de Monte Carlo considerando séries bivariadas.

Seguindo as configurações adotadas no estudo de simulação em Boudt, Danielsson e Laurent (2011), a série de retornos \mathbf{r}_t é gerada como a soma entre um processo GARCH(1,1)-cDCC(1,1) e um processo \mathbf{a}_t , seguindo a especificação apresentada na Seção 3.3. Foram consideradas 10.000 réplicas de amostras de tamanho $T = 2.000$. Os retornos gerados seguem a estrutura apresentada nas Equações (3.11)-(3.12) com $\boldsymbol{\mu}$, \mathbf{h} , $\boldsymbol{\theta}$ e ϕ dados

por

$$\begin{aligned}\boldsymbol{\mu} &= (\mu_1 \ \mu_2)' = (0, 05 \ - 0, 05)' \\ \mathbf{h} &= (h_1^* \ h_2^*)' = (1 \ 1)' \\ \boldsymbol{\theta} &= \begin{bmatrix} 0,10 & 0,10 & 0,80 \\ 0,10 & 0,20 & 0,70 \end{bmatrix} \\ \boldsymbol{\phi} &= (0,10 \ 0,80)',\end{aligned}$$

em que h_i^* é a variância não condicional para a i -ésima série.

Para $\mathbf{S} = [s_{ij}]$ adotou-se $s_{1,2} = s_{2,1} = 0,4$. Os elementos em B , conjunto de tempos com *outliers*, foram selecionados de forma igualmente espaçada, com percentual de tempos com contaminação $\epsilon = 0\%, 1\%, 5\%$ e 10% . Em 40% desses tempos os saltos são conjuntos (selecionados aleatoriamente) e nos 60% restantes seleciona-se uma série (com probabilidade $1/2$) para contaminação. Dessa forma, $\epsilon = 1\%$ corresponde, em média, a $0,7\%$ de tempos com contaminação para cada série. A magnitude do salto é igual ao desvio padrão condicional correspondente ao respectivo elemento de \mathbf{r}_t vezes d para a primeira série e d negativo para a segunda, com d valendo 3 ou 4.

5.2 Estimadores

Na estimação do parâmetro $\boldsymbol{\eta}_i = (\omega_i, \alpha_i, \beta_i)'$ de um processo GARCH de ordem $(1, 1)$ serão comparados os estimadores $\hat{\boldsymbol{\eta}}_{q,i}$ e $\hat{\boldsymbol{\eta}}_{r,i}$ que para efeito de notação estão associados a i -ésima série de retornos. Novamente, o estimador de quase-máxima verossimilhança ($\hat{\boldsymbol{\eta}}_q$) para $\boldsymbol{\eta}$ de uma série qualquer $\{y_t, t = 1, \dots, N\}$ é definido por:

$$\hat{\boldsymbol{\eta}}_q = \underset{\boldsymbol{\eta} \in \mathbf{H}}{\operatorname{argmax}} M_0(\boldsymbol{\eta}),$$

em que

$$h_t = \omega + \alpha y_{t-1}^2 + \beta h_{t-1},$$

tal que

$$\begin{aligned}M_0(\boldsymbol{\eta}) &\propto \frac{1}{T-p} \sum_{t=p+1}^T \rho_0(\ln(y_t^2) - \ln h_t), \\ \rho_0 &= -\ln(g_0), \\ g_0(\lambda_t) &= \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2}(e^{\lambda_t} - \lambda_t) \right\}.\end{aligned}$$

O estimador robusto $\hat{\boldsymbol{\eta}}_r$, definido na Seção 2.4.2, é dado por:

$$\hat{\boldsymbol{\eta}}_r = \underset{\boldsymbol{\eta} \in \mathbf{H}}{\operatorname{argmin}} \frac{1}{T} \sum_{t=1}^T \rho_2 \left(\log \frac{y_t^2}{h_t} \right),$$

tal que

$$\begin{aligned} h_t &= \omega + \alpha w_{\delta,1} \left(\frac{y_{t-1}^2}{h_{t-1}} \right) y_{t-1}^2 + \beta h_{t-1}, \\ w_{\delta,N}(u) &= \min(u, k_{\delta,N}) u^{-1} \times c_{\delta,N}, \\ c_{\delta,N} &= \frac{E(U)}{E(\min(U, k_{\delta,N}))}, \end{aligned} \quad (5.1)$$

com

$$\begin{aligned} \rho_2(z) &= -z + \sigma_{1,v} \rho_{t_{1,v}}(\exp(z)), \\ \rho_{t_{N,v}}(u) &= (N+v) \log \left(1 + \frac{u}{v-2} \right), \\ \sigma_{N,v} &= \frac{N}{E[\rho'_{t_{1,v}}(U)U]}, \\ U &\sim \chi_N^2, \end{aligned}$$

considerando $v = 4$, $\delta = 0,975$, $\sigma_{1,4} = 0,8260$ e $c_{0,975;1} = 1,0465$. Valores de $c_{\delta,N}$ e $\sigma_{N,v}$, para diferentes escolhas de N e δ podem ser consultados em [Boudt, Danielsson e Laurent \(2013, Tabela 1\)](#).

Para o processo cDCC de ordem (1, 1) exibido na Seção 3.2 com parâmetro $\phi = (a, b)'$ serão comparados os estimadores $\hat{\phi}_q$ e $\hat{\phi}_r$, em que $\hat{\phi}_q$ é o estimador de quasi-máxima verossimilhança de ϕ apresentado na Seção 3.4, dado por

$$\hat{\phi}_q = \underset{\phi \in \Phi}{\text{argmax}} -\frac{1}{2} \sum_{t=1}^T [\log(\det \mathbf{R}_t) + \tilde{\mathbf{y}}'_{q,t} \mathbf{R}_t^{-1} \tilde{\mathbf{y}}_{q,t}] .$$

O estimador robusto $\hat{\phi}_r$ apresentado na Seção 3.5 é dada por:

$$\hat{\phi}_r = \underset{\phi \in \Phi}{\text{argmin}} \left(\frac{1}{T} \sum_{t=1}^T \rho_3(\mathbf{R}_t, \tilde{\mathbf{y}}_t) \right),$$

em que

$$\begin{aligned} \mathbf{R}_t &= \mathbf{Q}_t^{*1/2} \mathbf{Q}_t \mathbf{Q}_t^{*1/2}, \\ \mathbf{Q}_t &= (1-a-b)\mathbf{S} + aw_{\delta,2}(d_{t-1}) \mathbf{Q}_{t-1}^{*1/2} \tilde{\mathbf{y}}'_{t-1} \tilde{\mathbf{y}}_{t-1} \mathbf{Q}_{t-1}^{*1/2} + b \mathbf{Q}_{t-1}, \end{aligned}$$

com

$$\rho_3(\mathbf{R}, \mathbf{y}) = \log(\det \mathbf{R}) + \sigma_{N,4} \rho_{t_{N,4}}(\mathbf{y}' \mathbf{R}^{-1} \mathbf{y}),$$

considerando $v = 4$, $N = 2$, $\sigma_{2,4} = 0,8258$ e $c_{0,975;2} = 1,0257$.

Os resultados do estudo de Monte Carlo para $\hat{\eta}_q$ e $\hat{\eta}_r$ são apresentados na Seção 5.4, e para $\hat{\phi}_q$ e $\hat{\phi}_r$ na Seção 5.5.

5.3 Medidas de desempenho

A comparação dos modelos foi feita utilizando uma seleção de métricas destinadas a captar diferentes competências. Neste trabalho, elas podem ser classificadas em três classes principais: (1) desempenho na estimação dos parâmetros, (2) desempenho do filtro de volatilidade e (3) desempenho na estimação de portfólio de variância mínima. Para o (1) procurou-se investigar o desempenho dos EQMVn ($\boldsymbol{\theta}_q, \boldsymbol{\phi}_q$) e robusto ($\boldsymbol{\theta}_r, \boldsymbol{\phi}_r$) precisamente na etapa de estimação, sendo avaliados considerando viés, erro quadrático médio, desvio padrão etc. As métricas em (2) buscam quantificar o desempenho do filtro de volatilidade, verificando a frequência em que este é utilizado para limitar um possível *outlier*, quantas vezes foi utilizado corretamente para ponderar (sensibilidade) e não ponderar (especificidade) uma observação. Por fim, em (3) há o interesse em observar a competência de cada método na estimação da matriz \mathbf{H}_t .

5.3.1 Estimação dos parâmetros

Seja $\hat{\tau}$ o estimador do parâmetro τ e $\hat{\tau}_j$ a estimativa obtida para a j -ésima réplica de Monte Carlo. Considerando um total de M replicações as medidas utilizadas para a avaliar os estimadores são descritas a seguir.

- **Média:**

$$\text{Média}(\hat{\tau}) = \frac{1}{M} \sum_{j=1}^M \hat{\tau}_j = \bar{\hat{\tau}};$$

- **Viés:**

$$\text{Viés}(\tau, \hat{\tau}) = \frac{1}{M} \sum_{j=1}^M (\hat{\tau}_j - \tau) = \bar{\hat{\tau}} - \tau;$$

- **Raiz do Erro Quadrático Médio (REQM):**

$$REQM(\tau, \hat{\tau}) = \left(\frac{1}{M} \sum_{j=1}^M (\hat{\tau}_j - \tau)^2 \right)^{1/2};$$

- **Desvio Padrão (DP):**

$$DP(\hat{\tau}) = \left(\frac{1}{M} \sum_{j=1}^M (\hat{\tau}_j - \bar{\hat{\tau}})^2 \right)^{1/2};$$

- **Quantis Empíricos ($100 \times \delta$) % (QE):**

Sejam $\hat{\tau}_1, \dots, \hat{\tau}_M$ as estimativas de Monte Carlo para τ . Dada a distribuição dessas estimativas é possível construir estimativas intervalares utilizando os quantis empíricos $\hat{\tau}_{(1-\delta)/2}$ e $\hat{\tau}_{(1+\delta)/2}$ de ordem $(1 - \delta)$ e $(1 + \delta)/2$, com $\delta \in (0, 1)$. Dessa forma os intervalos empíricos são dados por

$$QE[\tau, 100 \times \delta\%] = [\hat{\tau}_{(1-\delta)/2}; \hat{\tau}_{(1+\delta)/2}].$$

5.3.2 Desempenho do filtro de volatilidade

A qualidade do filtro de volatilidade é dada pela capacidade do filtro de identificar corretamente um *outlier*. Para o conjunto de tempos contaminados B apresentado na Seção 3.3, têm-se que $\#B$ é a cardinalidade do conjunto, i.e., a quantidade de tempos contaminados, e $\#B^c$ a quantidade de tempos não contaminados de modo que $\#B + \#B^c = T$. Dado isso, as métricas são dadas por:

- **Especificidade:** A especificidade quantifica o acerto por parte do filtro de volatilidade, que ocorre quando o mesmo não pondera uma observação sem *outlier*. Considerando o estudo de Monte Carlo, a especificidade estimada considerando as M réplicas é dada por

$$\text{especificidade}_M(w_{\delta,N}) = \frac{1}{\#B^c M} \sum_{m=1}^M \sum_{t=1}^T I[w_{\delta,N}(d_{m,t}) = 1, \|\mathbf{a}_t\| = 0].$$

- **Sensibilidade:** A medida de sensibilidade quantifica o desempenho do filtro de volatilidade em outro tipo de acerto, i.e., detectar outlier nos casos com contaminação ($a_t \neq 0$, ou $\|\mathbf{a}_t\| \neq 0$ no caso multivariado). Para M replicações a sensibilidade geral pode ser calculada por

$$\text{sensibilidade}_M(w_{\delta,N}) = \frac{1}{\#B M} \sum_{m=1}^M \sum_{t=1}^T I[w_{\delta,N}(d_{m,t}) < 1, \|\mathbf{a}_{m,t}\| \neq 0],$$

sendo $r_{m,t}$, $d_{m,t}$ e $\mathbf{a}_{m,t}$ a observação, distância de Mahalanobis e processo aditivo N-dimensional no tempo t para m -ésima replicação, respectivamente.

- **Acurácia:** Por fim, a acurácia contabiliza o proporção de acertos totais, i.e., ponderar (não ponderar) uma observação quando for (não for) um *outliers*. Para M replicações a acurácia pode ser calculada por

$$\text{acurácia}_M(w_{\delta,N}) = \frac{\#B^c \text{especificidade}_M(w_{\delta,N}) + \#B \text{sensibilidade}_M(w_{\delta,N})}{T}.$$

5.3.3 Portfólio de variância globalmente mínima

O portfólio de variância globalmente mínima (PVGM) é um específico portfólio ótimo cuja variância é a menor dentre todos os portfólios na fronteira eficiente sem levar em consideração o retorno esperado (BODNAR; MAZUR; OKHRIN, 2017). O PVGM é a solução para a seguinte formulação

$$\min_{\omega} \quad \omega'_t \mathbf{H}_t \omega_t \tag{5.2a}$$

$$\text{s.a} \quad \omega'_t \mathbf{1} = \mathbf{1}, \tag{5.2b}$$

sendo $\mathbf{1}$ um vetor de uns de dimensão $N \times 1$ e $\boldsymbol{\omega}_t$ o vetor de pesos para o portfólio. Dado \mathbf{H}_t a solução analítica é dada por

$$\boldsymbol{\omega}_t = \frac{\mathbf{H}_t^{-1}\mathbf{1}}{\mathbf{1}\mathbf{H}_t^{-1}\mathbf{1}},$$

em que \mathbf{H}_t precisa ser estimada por $\hat{\mathbf{H}}_t$ resultando no portfólio estimado

$$\hat{\boldsymbol{\omega}}_t = \frac{\hat{\mathbf{H}}_t^{-1}\mathbf{1}}{\mathbf{1}\hat{\mathbf{H}}_t^{-1}\mathbf{1}}.$$

A tarefa de estimar o PVGM permite avaliar exclusivamente o desempenho do estimador da matriz de variâncias e covariâncias condicionais ao omitir a etapa de estimação do vetor de retornos esperados.

Conforme apresentado em [Engle e Colacito \(2006\)](#) a variância do portfólio ótimo estimado considerando um estimador genérico de \mathbf{H}_t e vetor de médias $\boldsymbol{\mu}$ é dada por

$$\mathcal{L}_V(\hat{\mathbf{H}}_t, \mathbf{H}_t, \boldsymbol{\mu}) = \frac{\boldsymbol{\mu}'\mathbf{H}_t\hat{\mathbf{H}}_t\boldsymbol{\mu}}{\boldsymbol{\mu}'\hat{\mathbf{H}}\boldsymbol{\mu}}.$$

Por meio desta, é possível mensurar o desempenho do estimador da matriz de variâncias e covariâncias condicional, sendo útil para problemas de tomada de decisão baseados na minimização do risco. Note que \mathcal{L}_V será mínima quando $\hat{\mathbf{H}}_t = \mathbf{H}_t$. Para um função perda não negativa, [Engle, Ledoit e Wolf \(2019\)](#) apresentam a formulação para o excesso de variância resultante do erro de estimação

$$\mathcal{L}_E(\hat{\mathbf{H}}_t, \mathbf{H}_t, \boldsymbol{\mu}) = \frac{\boldsymbol{\mu}'\mathbf{H}_t\hat{\mathbf{H}}_t\boldsymbol{\mu}}{\boldsymbol{\mu}'\hat{\mathbf{H}}\boldsymbol{\mu}} - \frac{1}{\boldsymbol{\mu}'\mathbf{H}_t^{-1}\boldsymbol{\mu}}. \quad (5.3)$$

O vetor de médias $\boldsymbol{\mu}$ não precisa ser igual ao real, na verdade [Engle, Ledoit e Wolf \(2019\)](#) recomendam integrar \mathcal{L}_E em um domínio factível de vetores de retornos esperados. Na verdade, os componentes na Equação (5.3) não dependem do comprimento de $\boldsymbol{\mu}$, mas sim do ângulo. Por isso, [Engle e Colacito \(2006\)](#) consideraram para 2 dimensões $\boldsymbol{\mu} = [\cos(\theta), \sin(\theta)]$ para algum ângulo θ . Assim, o problema a função perda dependeria somente da matrizes de covariância

$$\mathcal{L}_I(\hat{\mathbf{H}}_t, \mathbf{H}_t) = \int_{\mathcal{U}_2} \mathcal{L}_E(\hat{\mathbf{H}}_t, \mathbf{H}_t, \boldsymbol{\mu}) d\boldsymbol{\mu},$$

em que para dimensão N , \mathcal{U}_N é um esfera unitária N -dimensional, i.e., $\mathcal{U}_N = \{\mathbf{x} \in \mathbb{R}^N : \|\mathbf{x}\| = 1\}$.

A integral em \mathcal{L}_I não é de simples resolução analítica. Como solução pode-se recorrer ao seguinte lema fundamental na Teoria de Matrizes Aleatórias

$$\mathbf{x}'\mathbf{A}\mathbf{x} \approx \frac{\text{Tr}(\mathbf{A})\|\mathbf{x}\|^2}{N},$$

sendo $\text{Tr}(\cdot)$ o traço da matriz, \mathbf{A} uma matriz aleatória simétrica de dimensão N e \mathbf{x} um vetor N -dimensional independente de \mathbf{A} e sua distribuição é invariante à rotação. Assim, uma alternativa factível para \mathcal{L}_I é dada por

$$\mathcal{L}(\hat{\mathbf{H}}_t, \mathbf{H}_t, \boldsymbol{\mu}) = \frac{\text{Tr}(\hat{\mathbf{H}}_t^{-1} \mathbf{H}_t \hat{\mathbf{H}}_t^{-1})/N}{\left[\text{Tr}(\hat{\mathbf{H}}_t^{-1})/N\right]^2} - \frac{1}{\text{Tr}(\mathbf{H}_t^{-1})/N}. \quad (5.4)$$

Por padronização, também é possível considerar (5.4) em termos do incremento percentual na variância do portfólio, dado por

$$\mathcal{L}_n(\hat{\mathbf{H}}_t, \mathbf{H}_t, \boldsymbol{\mu}) = \log(v_t/u_t), \quad (5.5)$$

em que

$$\begin{aligned} v_t &= \frac{\text{Tr}(\hat{\mathbf{H}}_t^{-1} \mathbf{H}_t \hat{\mathbf{H}}_t^{-1})/N}{\left[\text{Tr}(\hat{\mathbf{H}}_t^{-1})/N\right]^2}, \\ u_t &= \frac{1}{\text{Tr}(\mathbf{H}_t^{-1})/N}. \end{aligned}$$

As funções de perda $\mathcal{L}_I(\hat{\mathbf{H}}_t, \mathbf{H}_t)$ e $\mathcal{L}(\hat{\mathbf{H}}_t, \mathbf{H}_t, \boldsymbol{\mu})$ convergem quase certamente para o mesmo limite não estocástico quando $T \rightarrow \infty$ com $T/N \rightarrow c \in (0, 1)$ pelo Teorema 3.1 de [Ledoit e Wolf \(2017\)](#).

No estudo de Monte Carlo considera-se a função de perda avaliada no período fora da amostra, referenciada neste trabalho como **incremento percentual de variância** (IPV). Para a m -ésima réplica o IPV é dado por

$$\hat{L}_m = \mathcal{L}_{n,m}(\hat{\mathbf{H}}_{T+1,m}, \mathbf{H}_{T+1,m}), \quad (5.6)$$

sendo T o número de tempos, $\mathcal{L}_{n,m}$ a função perda apresentada na Equação (5.5), $\mathbf{H}_{T+1,m}$ e $\hat{\mathbf{H}}_{T+1,m}$ matrizes de variâncias e covariâncias condicionais real e estimada para m -ésima réplica fora da amostra, respectivamente. Para cada percentual de contaminação, magnitude do salto, réplica e ambos os modelos, será calculada a perda média obtida. Em seguida uma análise da distribuição dos erros médios será feita a fim de comparar o desempenho dos modelos na estimativa do portfólio de variância mínima.

No decorrer da seção, $\mathcal{L}_{n,q}$ irá denotar o IPV do portfólio cuja matriz \mathbf{H}_t é estimada utilizando $\hat{\phi}_q$, $\hat{\eta}_q$ e especificação de \mathbf{Q}_t dada em (3.6). Para a abordagem robusta, $\mathcal{L}_{n,r}$ denota o IPV do portfólio com matriz \mathbf{H}_t estimada utilizando $\hat{\phi}_r$, $\hat{\eta}_r$ e \mathbf{Q}_t com especificação dada em (3.18).

5.4 Estimação para modelos GARCH

As distribuições das estimativas dadas por modelo, percentual de contaminação e magnitude do salto, são ilustradas na Figura 1. Sem contaminação, as estimativas

robustas e tradicionais para α_1 estão distribuídas em torno de 0,10; com $\hat{\alpha}_{q,1}$ tendo mediana mais próxima. Quando há *outliers* as estimativas robustas ficam relativamente mais próximas de α_1 , mas em cenários de maior contaminação os desvios são consideráveis, por exemplo, com 5% de contaminação, 75% dos valores são menores que 0,10. Com relação a β_1 os resultados são semelhantes, com uma particularidade para 5% de contaminação, no qual as estimativas tradicionais apresentam grande variância. Para 10% mais de 95% das estimativas são inferiores (α_1) ou superiores (β_1) ao real valor do parâmetro. As distribuições das persistências estimadas para $(\alpha_i + \beta_i)$ das duas séries podem ser conferidas na Figura 2, o estimador robusto foi capaz de produzir estimativas mais próximas do real em grande parte dos cenários.

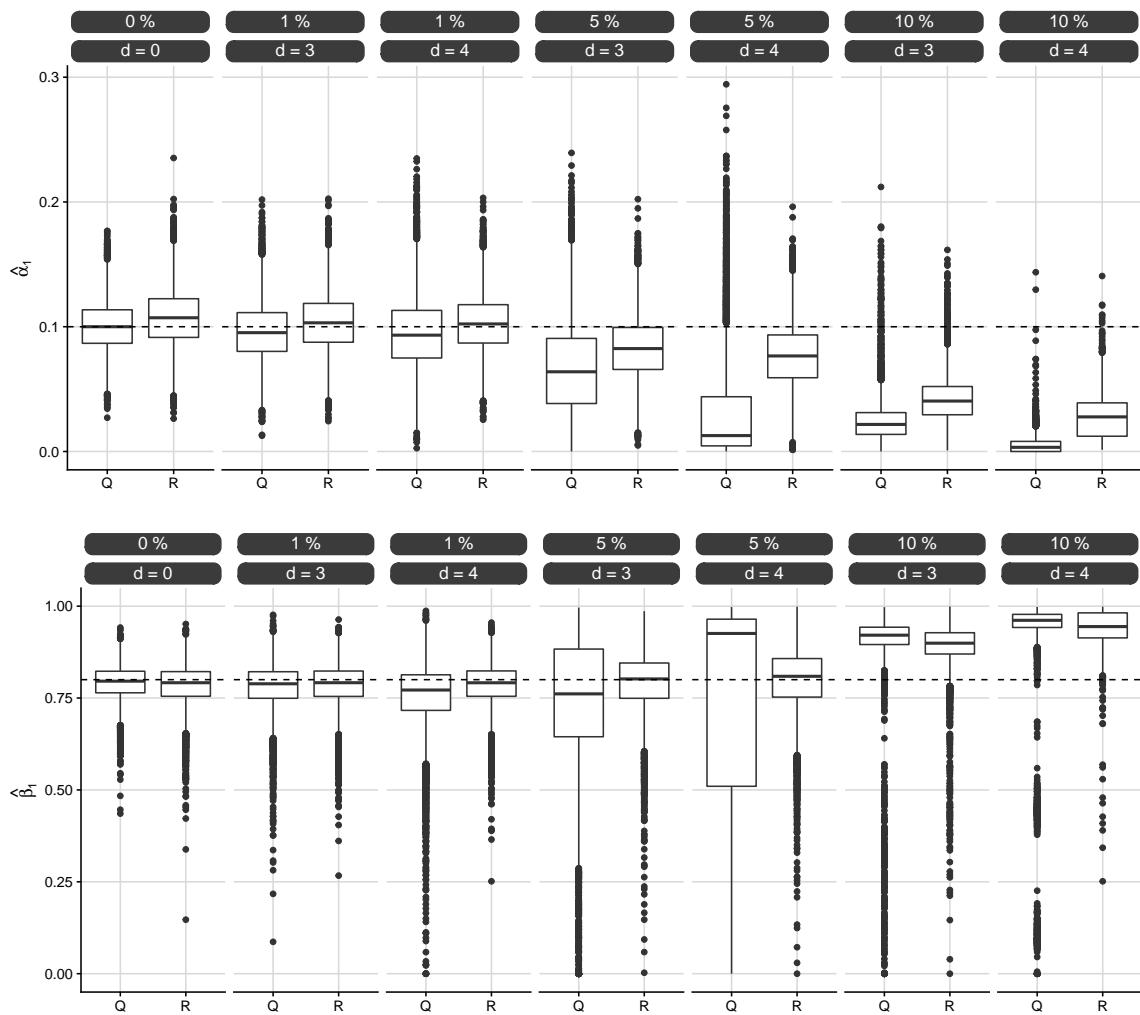


Figura 1 – Estimativas de $\eta = (0, 10, 0, 10, 0, 80)$ do modelo GARCH(1, 1) obtidas por $\hat{\eta}_q$ e $\hat{\eta}_r$. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\eta}_q$ e $\hat{\eta}_r$, respectivamente. Na figura superior estimativas do parâmetro α_1 e na inferior de β_1 . Linha tracejada indica real valor do respectivo parâmetro. $T = 2.000$, $\delta = 0,975$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

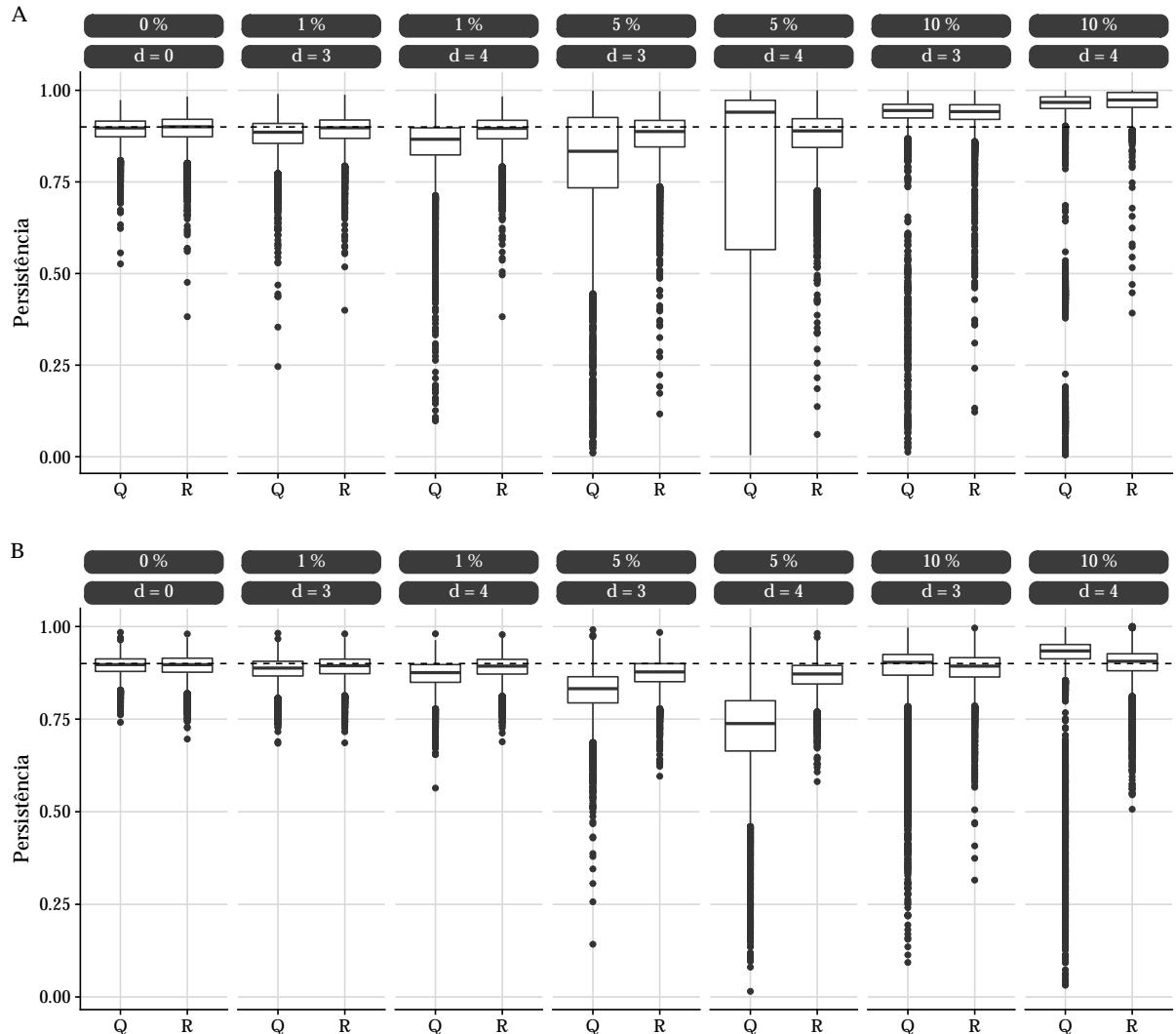


Figura 2 – Distribuições das persistências estimadas ($\hat{\alpha}_i + \hat{\beta}_i$) pelos estimadores tradicional e robusto. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\theta}_q$ e $\hat{\theta}_r$, respectivamente. Em (A) os resultados para $\eta_1 = (0, 10, 0, 10, 0, 80)$ e (B) os resultados para $\eta_2 = (0, 10, 0, 20, 0, 70)$. $T = 2.000$, $\delta = 0, 975$, $N = 2$ e 10.000 réplicas.

Na Tabela 1 os estimadores tradicional e robusto são comparados considerando algumas medidas de qualidade de ajuste; os resultados estão ilustrados na Figura 3. Quando não há *outliers* aditivos $\hat{\eta}_{q,1}$ pode ser preferível, porém em cenários com contaminação o estimador robusto tem menor REQM, principalmente quando ϵ, d ou ambos aumentam. Em particular, para β_1 com 10% de contaminação, as estimativas tem maior (menor) viés (REQM), provavelmente por se concentrarem no limite superior do intervalo $[0, 1]$. Outro ponto a se destacar é o efeito que a magnitude dos saltos tem em um mesmo percentual de contaminação; para $\hat{\eta}_{q,1}$ o aumento de 3 para 4 desvios padrão condicional tem grande impacto, tanto em termos de viés quanto de REQM, enquanto que $\hat{\eta}_{r,1}$ é menos afetado (ver Figura 3).

Tabela 1 – Viés e REQM para estimadores de $\alpha_1 = 0,10$ e $\beta_1 = 0,8$ de modelo GARCH(1,1). $T = 2.000, \delta = 0,975$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

	d		μ_1	ω	α_1	β_1	h_1^*
$\epsilon = 0\%$	0	Viés	Q	0,0000	0,0081	0,0005	-0,0091
	0		R	0,0000	0,0085	0,0074	-0,0145
$\epsilon = 1\%$	3		Q	0,0208	0,0301	-0,0037	-0,0193
	3		R	0,0044	0,0134	0,0037	-0,0143
	4		Q	0,0277	0,0650	-0,0048	-0,0445
	4		R	0,0026	0,0143	0,0028	-0,0145
$\epsilon = 5\%$	3		Q	0,1037	0,1528	-0,0331	-0,0618
	3		R	0,0221	0,0348	-0,0169	-0,0089
	4		Q	0,1383	0,2695	-0,0686	-0,0727
	4		R	0,0130	0,0319	-0,0231	-0,0012
$\epsilon = 10\%$	3		Q	0,2076	0,0165	-0,0759	0,1022
	3		R	0,0513	-0,0253	-0,0584	0,0942
	4		Q	0,2768	0,0318	-0,0947	0,1292
	4		R	0,0278	-0,0650	-0,0728	0,1423
$\epsilon = 0\%$	0	REQM	Q	0,0221	0,0357	0,0200	0,0486
	0		R	0,0258	0,0409	0,0243	0,0570
$\epsilon = 1\%$	3		Q	0,0304	0,0585	0,0237	0,0654
	3		R	0,0263	0,0448	0,0236	0,0590
	4		Q	0,0356	0,1037	0,0297	0,1019
	4		R	0,0260	0,0456	0,0235	0,0596
$\epsilon = 5\%$	3		Q	0,1062	0,2514	0,0490	0,1909
	3		R	0,0345	0,0771	0,0301	0,0822
	4		Q	0,1402	0,5128	0,0800	0,3215
	4		R	0,0292	0,0796	0,0345	0,0879
$\epsilon = 10\%$	3		Q	0,2090	0,1624	0,0777	0,1535
	3		R	0,0582	0,0622	0,0612	0,1133
	4		Q	0,2780	0,2929	0,0950	0,1962
	4		R	0,0385	0,0736	0,0748	0,1492

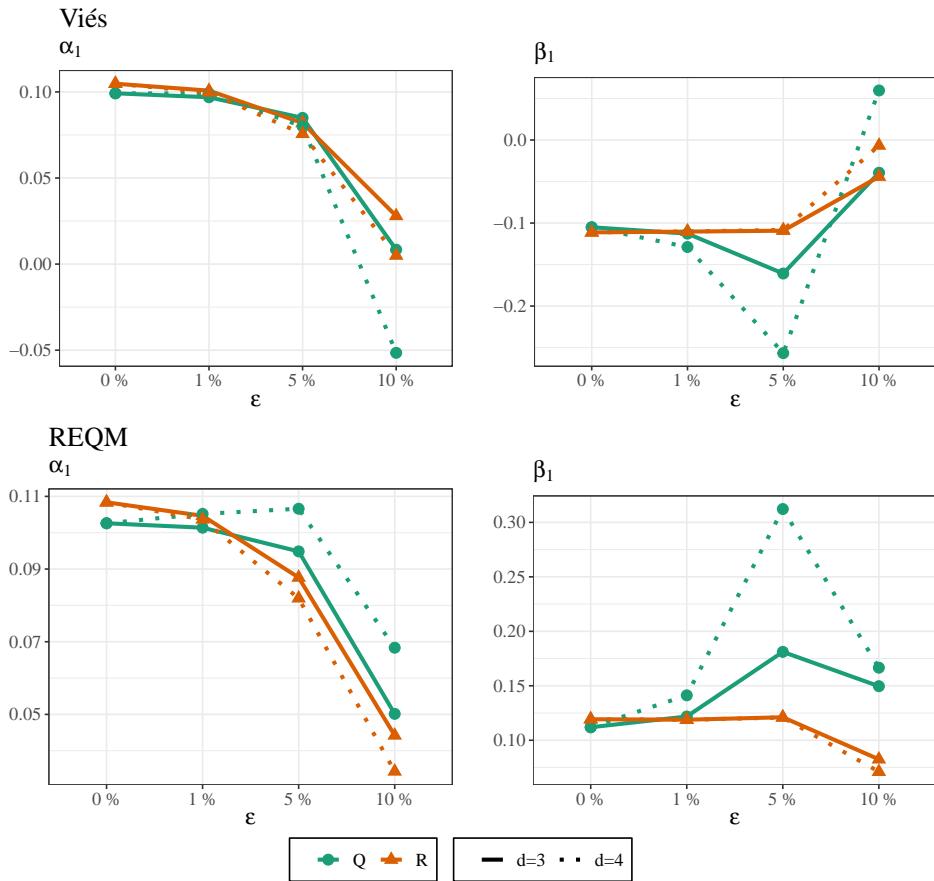


Figura 3 – Viés e REQM (linhas) obtidos por $\hat{\eta}_q$ e $\hat{\eta}_r$ para parâmetros α_1 e β_1 (colunas). $T = 2.000$, $\delta = 0,975$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

5.5 Estimação do modelo cDCC

Nesta seção serão apresentados os resultados relativos a estimação do modelo cDCC bivariado. Os resultados sobre a estimação de $\phi = (a, b)'$ são apresentados na Subseção 5.5.1, enquanto a Subseção 5.5.2 apresenta os resultados sobre a estimação da matriz de covariância condicional fora da amostra $H_{2.001}$ e a Subseção 5.5.3 a análise dos filtros de volatilidade utilizados.

5.5.1 Estimação dos parâmetros

As distribuições das estimativas de ϕ e das persistências estimadas estão ilustradas nas Figuras 4 e 5, respectivamente. Algumas medidas de qualidade de ajuste podem ser consultadas na Tabela 2, alguns destes resultados são ilustrados na Figura 6.

Na parte cDCC, os estimadores tradicional e robusto produzem estimativas similares quando não há saltos, mas a contaminação por *outliers* aditivos destaca a robustez de $\hat{\phi}_r$, que produz estimativas mais próximas dos valores reais (ver Figura 4). Análogo ao observado na parte GARCH, o aumento no número de *outliers* (especificamente, 5% e 10%)

desloca a distribuições para os limites do intervalo $[0, 1]$, aumentando viés e reduzindo a variância, conforme apresentado na Tabela 2 que contém, além destas, outras medidas de qualidade de ajuste. Os valores estimados para $(a + b)$ por $\hat{\phi}_r$ estão mais próximos da persistência real ($=0,90$) na maioria dos cenários com exceção de 10%, em grande parte por \hat{b}_q produzir valores próximos a 0,80 (ver Figura 5).

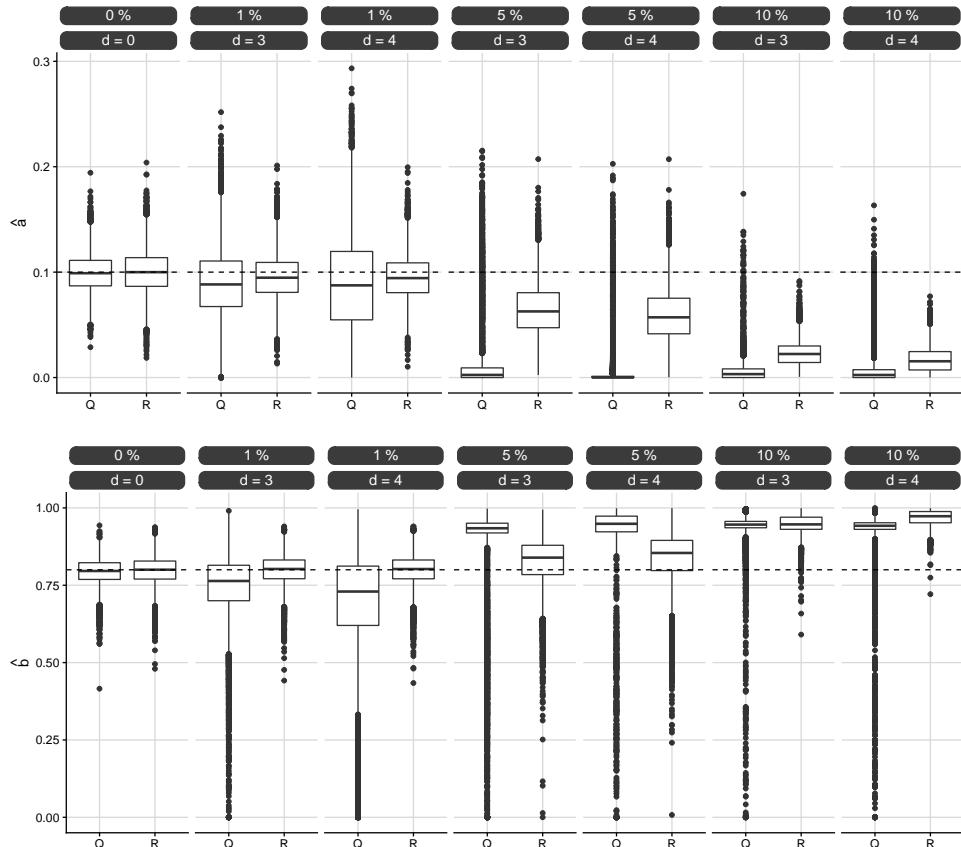


Figura 4 – Estimativas de $\phi = (0, 10, 0, 80)$ de modelo cDCC(1, 1) obtidas por $\hat{\phi}_q$ e $\hat{\phi}_r$. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\phi}_q$ e $\hat{\phi}_r$, respectivamente. No eixo vertical, estimativas para parâmetros a e b . Linha tracejada indica real valor do respectivo parâmetro. $T = 2.000$, $\delta = 0, 975$, $N = 2$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

Pela Figura 6 é possível visualizar o efeito do aumento do percentual de contaminação no viés, desvio padrão e REQM; com 1% o estimador robusto obtém REQM menor que o tradicional; com 5% essa diferença aumenta (diminui) para o parâmetro a (b); com 10% os REQMs do dois estimadores são similares com pequena vantagem para o robusto, com exceção do parâmetro b para saltos de 4 desvios padrão condicional. Além disso, algo a se observar é que o aumento do viés para o estimador tradicional de b não ocorre de forma linear, entre 0% e 1% o viés passa a ser positivo enquanto que entre 1% e 5%, negativo. Por fim, 1% de contaminação é o cenário em que o estimador tradicional

apresenta maior variância, em específico para saltos de 4 desvios padrão condicional. Com isso, têm-se que o estimador robusto consegue ser melhor (em termos de REQM) até 5% de contaminação; para 10% o estimador robusto não difere muito do tradicional, cabe a discussão deste ponto em trabalhos futuros.

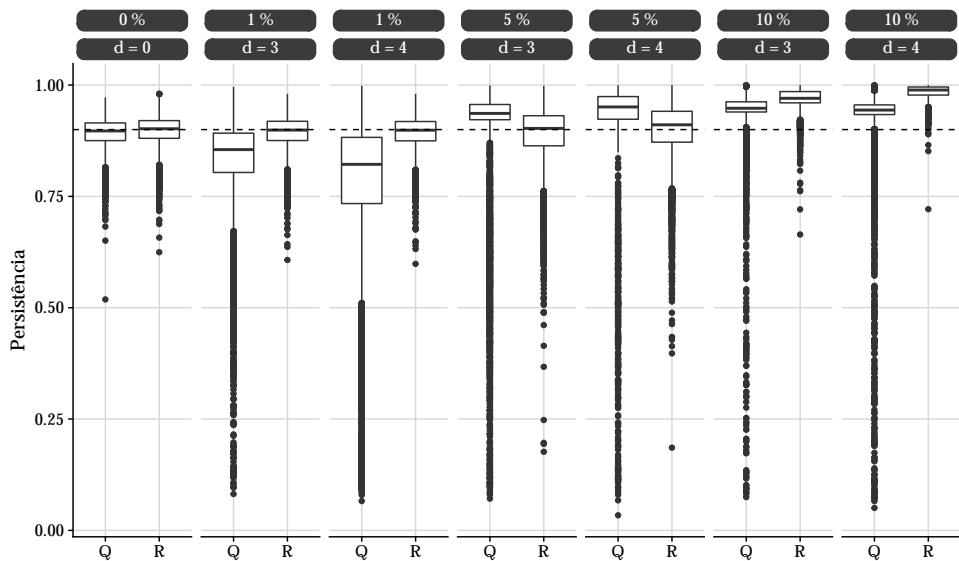


Figura 5 – Distribuições das persistências estimadas para $\phi = (0, 10, 0, 80)$ de modelo cDCC(1, 1) obtidas por $\hat{\phi}_q$ e $\hat{\phi}_r$. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\phi}_q$ e $\hat{\phi}_r$, respectivamente. Linha tracejada indica valor real de $(a + b)$. $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas.

5.5.2 Estimação da variância condicional

Nesta seção serão apresentados os resultados para estimação de $\mathbf{H}_{2.001}$. O desempenho é mensurado usando três métricas: (1) incremento percentual de variância (IPV); (2) razão entre as variâncias dos portfólios de variância mínima estimado e real, e (3) norma de Frobenius para $\hat{\mathbf{H}}_{2.001}$ e $\mathbf{H}_{2.001}$, apresentados nas Figuras 7, 8 e 9, respectivamente. Em específico, avalia-se o desempenho de predição das correlações condicionais fora da amostra. Nas três métricas é possível perceber como a estimação de $\mathbf{H}_{2.001}$ por EQMVn passa a ficar cada vez mais comprometida com o aumento do grau de contaminação, com o estimador robusto gerando resultados relativamente melhores e sofrendo menos impacto em cenários com mais *outliers*. Com 0%, o método tradicional resultou em portfólios com variância próxima ao PVGM; em média o excesso de variância percentual é cerca de 0,19%, enquanto o robusto teve desempenho um pouco pior com 0,28%. Quantificando o que já foi observado na Figura 9, em um cenário de 1% de contaminação o método robusto apresentou aumento percentual de 0,30%, menos da metade do obtido por $\hat{\mathbf{H}}_{q,2.001}$. Além disso, note a alta variabilidade das estimativas, com desvio padrão quase sempre superior a respectiva média.

Tabela 2 – Métricas de desempenho para estimadores de $a = 0, 10$ e $b = 0, 80$ de modelo cDCC(1,1). $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

	d	média		viés		REQM		DP		QE(95%)				
		a	b	a	b	a	b	a	b	a	b			
$\epsilon = 0\%$	0	Q	0,0994	0,7934	-0,0006	-0,0066	0,0182	0,0437	0,0182	0,0432	0,0650	0,1369	0,6998	0,8683
	0	R	0,1005	0,7975	0,0005	-0,0025	0,0208	0,0459	0,0208	0,0459	0,0610	0,1432	0,6993	0,8791
$\epsilon = 1\%$	3	Q	0,0897	0,7420	-0,0103	-0,0580	0,0338	0,1314	0,0322	0,1179	0,0304	0,1558	0,4366	0,8978
	3	R	0,0955	0,7989	-0,0045	-0,0011	0,0218	0,0491	0,0213	0,0491	0,0557	0,1391	0,6903	0,8835
	4	Q	0,0870	0,6961	-0,0130	-0,1039	0,0504	0,2121	0,0487	0,1849	0,0000	0,1829	0,1730	0,9589
	4	R	0,0951	0,7986	-0,0049	-0,0014	0,0219	0,0496	0,0213	0,0496	0,0556	0,1394	0,6902	0,8837
$\epsilon = 5\%$	3	Q	0,0143	0,8839	-0,0857	0,0839	0,0912	0,1899	0,0311	0,1704	0,0000	0,1183	0,3263	0,9893
	3	R	0,0651	0,8244	-0,0349	0,0244	0,0428	0,0854	0,0248	0,0818	0,0237	0,1208	0,6302	0,9393
	4	Q	0,0041	0,9312	-0,0959	0,1312	0,0974	0,1704	0,0174	0,1088	0,0000	0,0640	0,5748	0,9911
	4	R	0,0600	0,8373	-0,0400	0,0373	0,0473	0,0919	0,0252	0,0840	0,0195	0,1178	0,6359	0,9501
$\epsilon = 10\%$	3	Q	0,0060	0,9359	-0,0940	0,1359	0,0946	0,1563	0,0105	0,0771	0,0000	0,0259	0,8237	0,9893
	3	R	0,0227	0,9473	-0,0773	0,1473	0,0782	0,1506	0,0121	0,0312	0,0021	0,0493	0,8781	0,9971
	4	Q	0,0059	0,9254	-0,0941	0,1254	0,0948	0,1603	0,0116	0,0998	0,0000	0,0311	0,7132	0,9893
	4	R	0,0168	0,9682	-0,0832	0,1682	0,0839	0,1699	0,0112	0,0238	0,0018	0,0414	0,9196	0,9978

Algumas estatísticas calculadas para os IPVs, razão entre as variâncias dos portfólios de variância mínima estimado e real e norma de Frobenius são apresentadas no Apêndice A, nas Tabelas 9, 10 e 11, respectivamente. Os resultados corroboram os comentários anteriores.

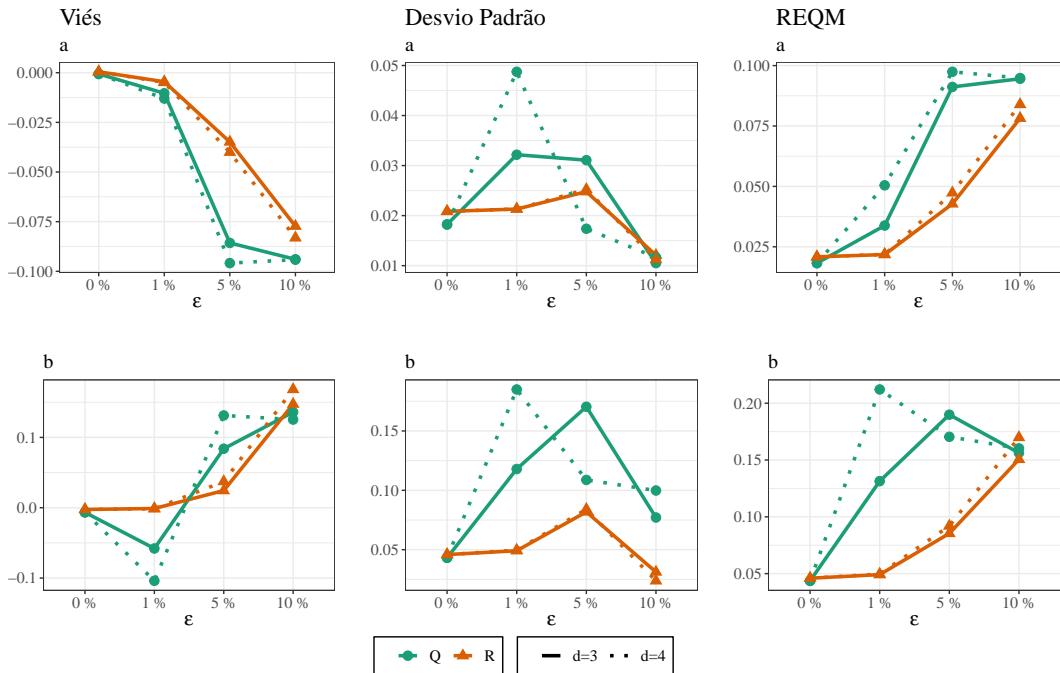


Figura 6 – Viés, REQM e erro padrão (colunas) obtidos por $\hat{\phi}_q$ e $\hat{\phi}_r$ para os parâmetros a e b (linha), para $T = 2.000, \delta = 0,975$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

5.5.3 Desempenho do filtro de volatilidade

Os resultados para análise do desempenho do filtro de volatilidade na etapa cDCC são apresentados na Tabela 3. Note que a sensibilidade tende a diminuir com o aumento de tempos contaminados, por outro lado, o filtro é mais assertivo quando os saltos são de maior magnitude. A especificidade se mantém relativamente constante para os diferentes cenários de contaminação. Por fim, a acurácia próxima a 1 resume o bom desempenho geral do filtro de volatilidade.

Tabela 3 – Acurácia do filtro de volatilidade para etapa cDCC bidimensional. $T = 2.000$, $\delta = 0,975$, $N = 2$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

d	ϵ	Especificidade	Sensibilidade	Acurácia
0	0%	0,98	-	0,98
3	1%	0,98	0,86	0,98
3	5%	0,98	0,83	0,98
3	10%	0,99	0,76	0,97
4	1%	0,98	0,98	0,98
4	5%	0,98	0,97	0,98
4	10%	0,99	0,93	0,98

5.6 Conclusão

Nesta seção foi realizado um estudo de Monte Carlo para avaliar a performance de dois estimadores: tradicional e robusto, aplicados tanto na etapa univariada para estimação das variâncias condicionais (parte GARCH) quanto na bivariada para correlações condicionais (parte cDCC). Esses estimadores foram avaliados em cenários com diferentes graus de contaminação a fim de validar possíveis vantagens que o estimador robusto pode trazer na modelagem de séries com *outliers* aditivos. O estudo foi dividido em 3 partes: (1) desempenho na estimação dos parâmetros, (2) eficácia na estimação de portfólio de variância mínima e (3) desempenho do filtro de volatilidade. Em geral, o estimador robusto apresentou melhores resultados. Embora que em termos de estimação o viés para 10% de séries contaminadas ainda é consideravelmente alto, este é um ponto para discussão em trabalhos futuros. Em termos práticos, o estimador robusto trás ganhos consideráveis e até mesmo com 10% de contaminação, os resultados são relativamente bons, com IPV mediano inferior a 0,10. Por fim, o desempenho do filtro de volatilidade parece estar diretamente ligado à magnitude dos saltos.

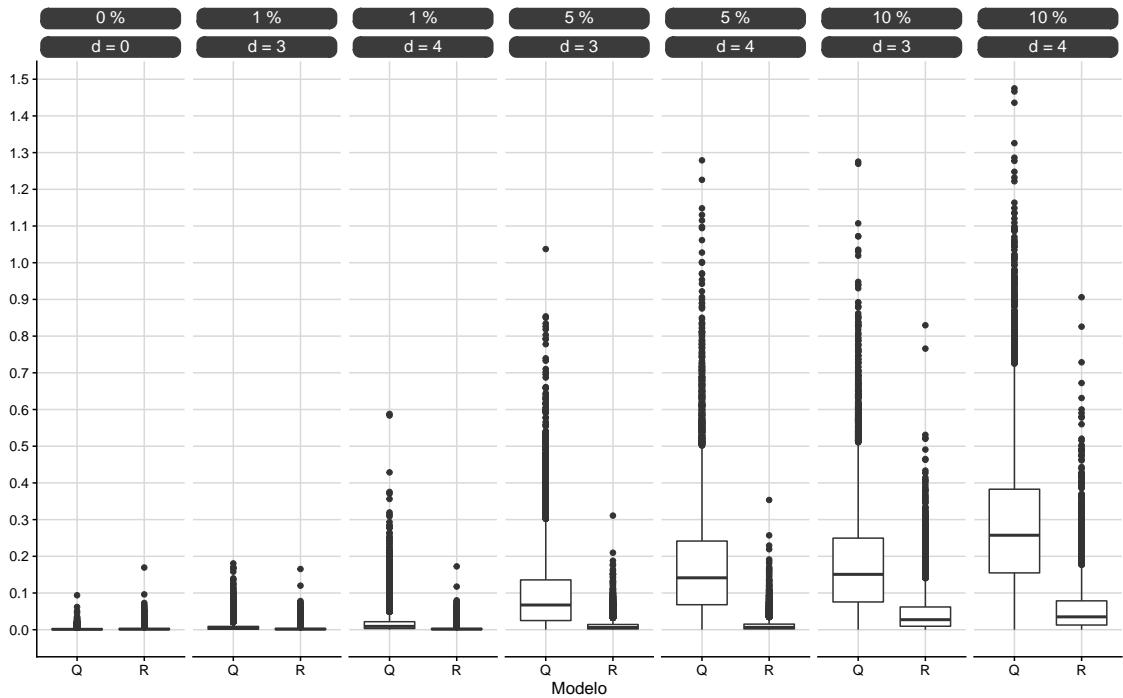


Figura 7 – IPV por modelo. Resultados para saltos consecutivos e não consecutivos. $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

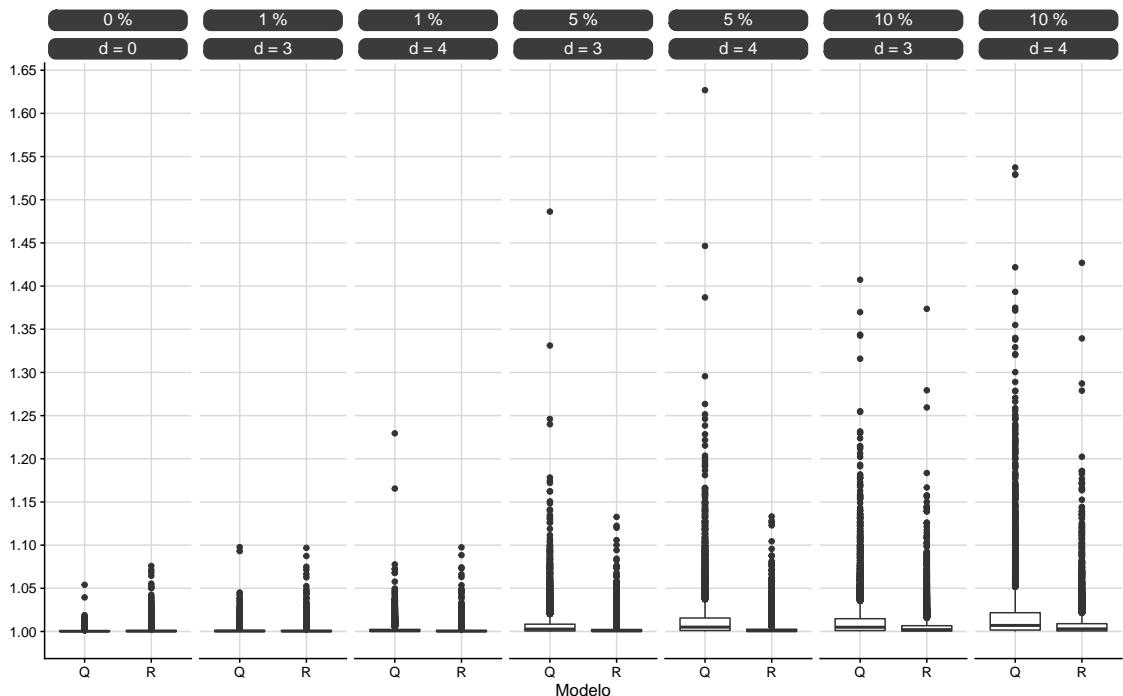


Figura 8 – Razão entre as variâncias dos portfólios de variância mínima estimado e real. Resultados para saltos consecutivos e não consecutivos. $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

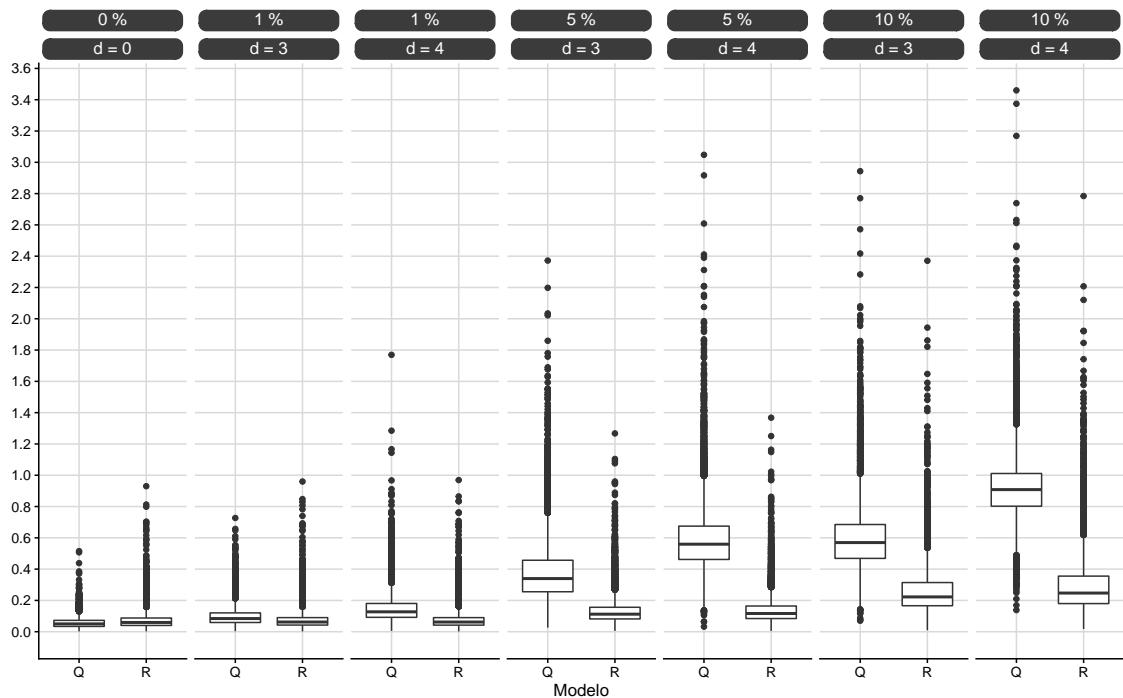


Figura 9 – Norma de Frobenius. Resultados para saltos consecutivos e não consecutivos. $T = 2.000, \delta = 0,975, N = 2$ e 10.000 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

6 Simulação em Alta Dimensão

Neste capítulo serão apresentados resultados do estudo de simulação de Monte Carlo a modelagem multivariada com dimensão $N = 1.000$. Os estimadores tradicional e robusto serão comparados em termos de qualidade de ajuste tanto no quesito de estimação dos parâmetros quanto na construção do portfólio de variância mínima para dados com diferentes graus de contaminação. O método de verosimilhança composta e estimadores *Shrinkage* e *Wrapping* apresentados no Capítulo 4 serão comuns aos dois estimadores, o objetivo é verificar a robustez de $\hat{\phi}_r$ combinado a estes métodos. Medidas de desempenho definidas no capítulo anterior serão utilizadas.

Na Seção 6.1 será apresentado o processo gerador de dados; seguido da exposição dos resultados divididos em três seções: (1) análise das estimativas para ϕ obtidas pelos estimadores tradicional e robusto na Seção 6.2, (2) na Seção 6.3 é avaliado o desempenho da estimativa de \mathbf{H}_t com as mesmas métricas apresentadas na Seção 5.3.2 e (3) na Seção 6.4 o desempenho do filtro de volatilidade utilizando as medidas apresentadas na Seção 5.3.2.

6.1 Processo gerador de dados

Seguindo o plano de contaminação adotado em Trucíos, Hotta e Pereira (2019) foram geradas 100 réplicas de tamanho $T = 1.250$. Os tempos em B são divididos em dois casos: (1) isolados e (2) consecutivos. Em (1) os tempos $t = 625$ e $t = 1249$ são contaminados por $a_{i,t} = d\sqrt{h_{i,t}}$ para $i \in \mathcal{I}$ e d igual a 5 ou 10. Em (2) as contaminações ocorrem em $t = 625, 626$ e $t = 1249$. Por fim, ϵ passa a representar o percentual de séries contaminadas com $\epsilon = 0\%, 25\%$ e 50% . Assim como em Pakel et al. (2021), o foco nesta parte está na estimativa de ϕ em problemas de alta dimensão, por isso a etapa de estimação das volatilidades foi omitida neste estudo, fixando-se $h_{i,t} = 1$ para todos os tempos e séries, ou seja, 1º estágio estimado sem erro. Foram consideradas duas estruturas de dependência temporal para volatilidade multivariadas: $\phi = (0, 10, 0, 80)$ e $(0, 05, 0, 93)$. A matriz de correlação não condicional \mathbf{S} é dada por:

$$\mathbf{S}_{i,j} = \begin{cases} \pi_i \pi_j, & i \neq j, \\ 1, & i = j \end{cases}$$

em que π_i e π_j , $i, j = 1, \dots, N$ são independentes e seguem uma distribuição normal truncada com média 0,5, desvio padrão 0,1 e truncamento em ± 4 desvios padrão, de modo que $\pi \in (0, 1, 0, 9)$ e a correlação média é de 0,25, produzindo assim retornos positivamente correlacionados.

6.2 Estimação dos parâmetros

As distribuições das estimativas considerando saltos isolados e consecutivos estão ilustradas nas Figuras 10 e 11, respectivamente. As persistências estimadas por sua vez são ilustradas na Figura 12. Algumas medidas de qualidade de ajuste podem ser consultados nas Tabelas 4 e 5 para saltos isolados e consecutivos, respectivamente.

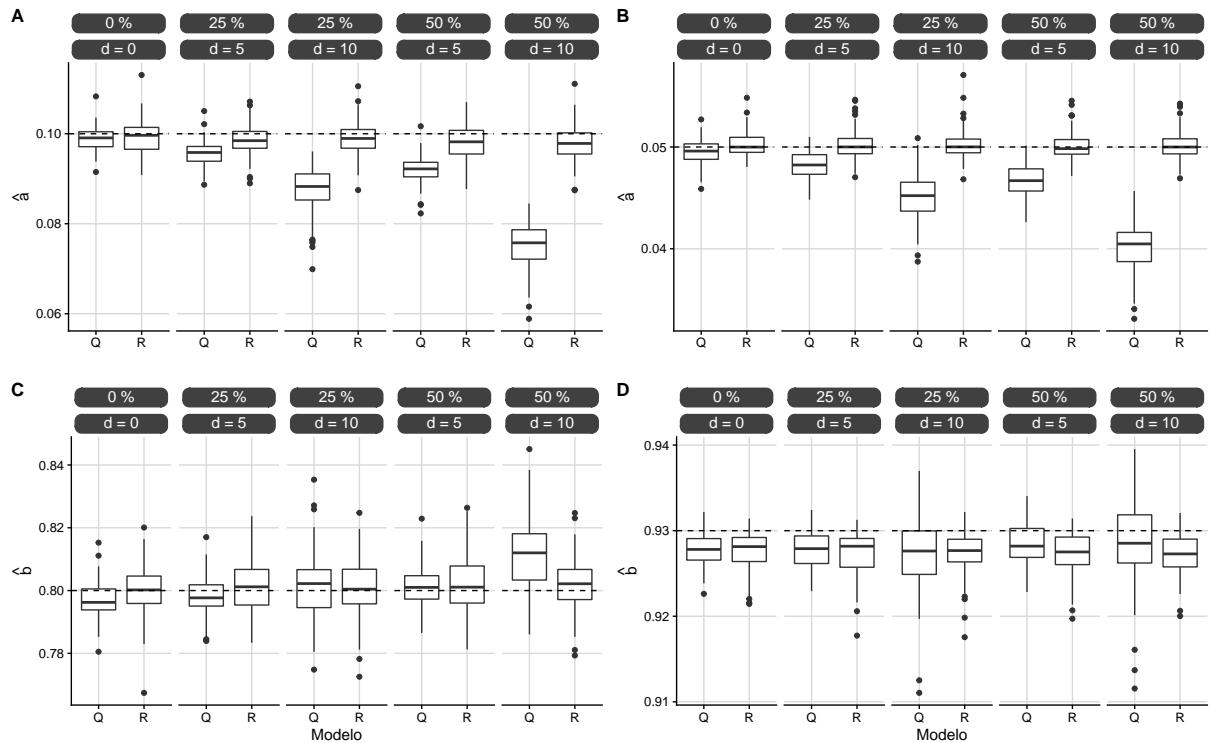


Figura 10 – Estimativas para 0 modelo cDCC(1, 1) obtidas por $\hat{\phi}_q$ e $\hat{\phi}_r$. Contaminação com saltos isolados, $T = 1.250$, $\delta = 0,975$, $N = 1.000$ e 100 réplicas. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\phi}_q$ e $\hat{\phi}_r$, respectivamente. No eixo vertical, estimativas para parâmetros a e b . Em (A), (B), (C) e (D) resultados para $a = 0, 10$, $a = 0, 05$, $b = 0, 80$ e $b = 0, 93$, respectivamente. Linha tracejada indica real valor do respectivo parâmetro. Os saltos são de magnitude d vezes o desvio padrão condicional.

Uma análise dos resultados mostram que o estimador tradicional do parâmetro a é altamente afetado pela presença de *outliers*. As estimativas, que já subestimavam a na ausência de saltos, se afastam ainda mais com o aumento do grau de contaminação. Para a , o estimador robusto apresenta menor viés (absoluto) e REQM em quase todos os cenários com exceção para $a = 0, 10$ e séries não contaminadas, em que existe praticamente igualdade de REQMs (ver Tabelas 4 e 5). Com o aumento do grau de contaminação existe uma pequena diminuição do vício e um ligeiro aumento na variância resultando em REQMs um pouco maiores, isso indica que este estimador parece não ser muito afetado por *outliers*.

Já para b as conclusões diferem, dependendo inclusive de ϕ . Quando $b = 0, 8$ a

presença de *outliers* deslocam as estimativas produzidas pelo estimador não robusto para cima, o deslocamento sendo maior quanto maior for o tamanho dos *outliers*, proporção de séries afetadas e principalmente com saltos consecutivos. O estimador robusto é melhor em termos de viés em quase todos os casos mas só é preferível sob REQM quando há saltos de tamanho 10; se $b = 0,93$ apenas a vantagem para REQM persiste. No entanto, quando o estimador não robusto é melhor, sua vantagem não é grande (ver Tabelas 4 e 5).

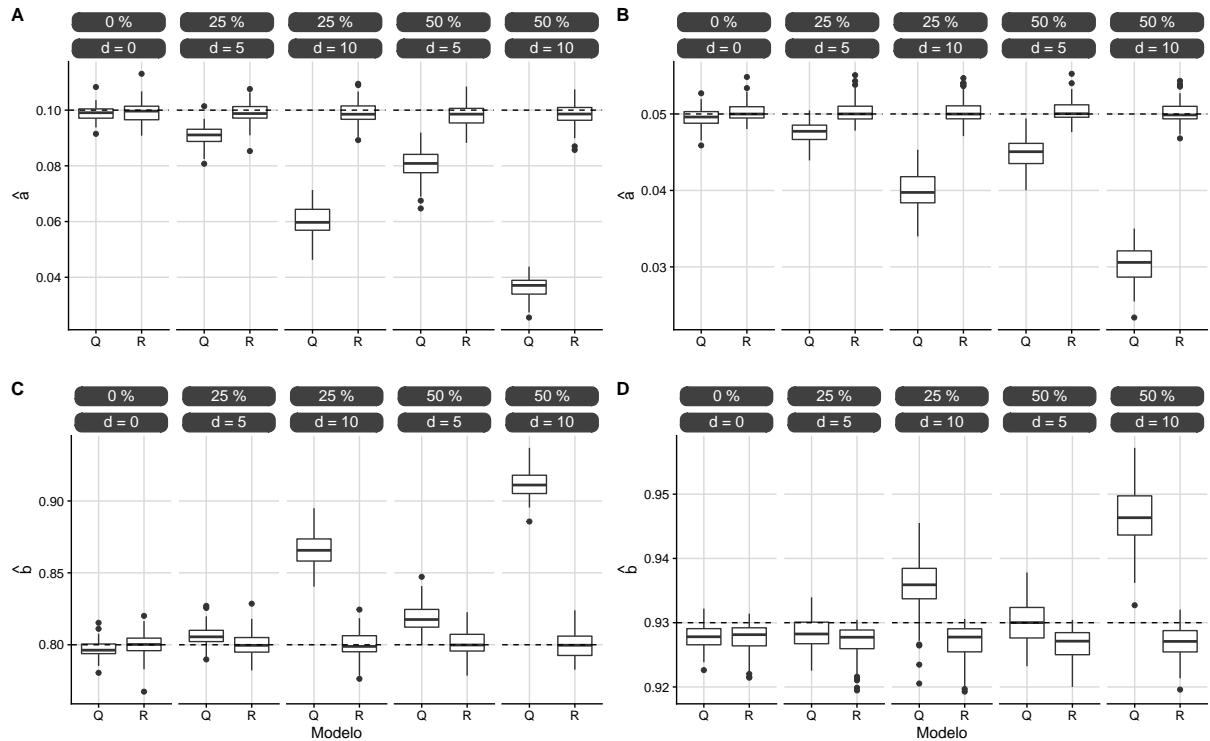


Figura 11 – Estimativas para o modelo cDCC(1, 1) obtidas por $\hat{\phi}_q$ e $\hat{\phi}_r$. Contaminação com saltos consecutivos, $T = 1.250$, $\delta = 0,975$, $N = 1.000$ e 100 réplicas. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\phi}_q$ e $\hat{\phi}_r$, respectivamente. No eixo vertical, estimativas para parâmetros a e b . Em (A), (B), (C) e (D) resultados para $a = 0,10$, $a = 0,05$, $b = 0,80$ e $b = 0,93$, respectivamente. Linha tracejada indica real valor do respectivo parâmetro. Os saltos são de magnitude d vezes o desvio padrão condicional.

Note ainda que mesmo na ausência de saltos as estimativas para $\phi = (0,05, 0,93)'$ (tradicional e robustas) tendem a subestimar b , e isso impacta diretamente nas persistências estimadas apresentadas na Figura 12. Para $\phi = (0,10, 0,80)'$ as estimativas robustas estão mais próximas do valor real em todos os cenários de contaminação; o mesmo não ocorre para $\phi = (0,05, 0,93)'$ em que a maioria é inferior a $a + b = 0,98$.

6.3 Estimação da matriz de variância condicional

Nesta seção os estimadores serão avaliados com relação a eficiência da estimação da matriz $\mathbf{H}_{1.251}$. Para isso, diferentes métricas serão utilizadas: (1) incremento percentual

de variância (IPV); (2) variância do portfólio de variância mínima estimado e (3) norma de Frobenius para $\hat{\mathbf{H}}_{1.251}$ e $\mathbf{H}_{1.251}$. Os passos nesta etapa do estudo de Monte Carlo são dados por:

1. Para a i -ésima réplica de Monte Carlo, utiliza-se as estimativas obtidas por QMVn e o modelo robusto, denotadas por $\hat{\phi}_{i,q}$ e $\hat{\phi}_{i,r}$, para calcular $\mathbf{Q}_t^{*1/2}\tilde{\mathbf{y}}_t$;
2. Transformar os respectivos retornos revolatilizados utilizando a função *wrapping*. O retorno desta transformação será denotado por **retornos revolatilizados embrulhados**.
3. Estimar a matriz intercepto \mathbf{S} por *Shrinkage* aplicada aos retornos revolatilizados embrulhados. $\hat{\mathbf{S}}_r$ e $\hat{\mathbf{S}}_q$ denota as estimativas utilizando $\hat{\phi}_r$ e $\hat{\phi}_q$, respectivamente.
4. Usar retornos até $t = 1.250$ para estimar $\mathbf{H}_{1.251}$ e calcular as métricas supracitadas.

Tabela 4 – Métricas de desempenho para estimadores de a e b de modelo cDCC(1,1). Contaminação com saltos isolados, $T = 1.250$, $\delta = 0,975$, $N = 1.000$ e 100 réplicas. Em negrito, para cenário de contaminação, a melhor performance em termos de viés e REQM. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

d	média		viés		REQM		DP		QE(95%)					
	a	b	a	b	a	b	a	b	a	b	a	b		
			$a = 0,10, b = 0,80$											
$\epsilon = 0\%$	0	Q	0,0990	0,7968	-0,0010	-0,0032	0,0027	0,0064	0,0025	0,0056	0,0951	0,1033	0,7856	0,8070
	0	R	0,0992	0,8000	-0,0008	0,0000	0,0035	0,0076	0,0034	0,0076	0,0938	0,1054	0,7858	0,8129
$\epsilon = 25\%$	5	Q	0,0958	0,7983	-0,0042	-0,0017	0,0050	0,0060	0,0027	0,0058	0,0904	0,1003	0,7859	0,8077
	5	R	0,0986	0,8012	-0,0014	0,0012	0,0036	0,0074	0,0034	0,0073	0,0912	0,1054	0,7890	0,8146
	10	Q	0,0877	0,8014	-0,0123	0,0014	0,0133	0,0104	0,0049	0,0103	0,0760	0,0946	0,7829	0,8232
	10	R	0,0988	0,8004	-0,0012	0,0004	0,0040	0,0091	0,0038	0,0091	0,0910	0,1062	0,7816	0,8182
$\epsilon = 50\%$	5	Q	0,0919	0,8011	-0,0081	0,0011	0,0086	0,0064	0,0031	0,0063	0,0855	0,0968	0,7883	0,8143
	5	R	0,0979	0,8018	-0,0021	0,0018	0,0045	0,0092	0,0040	0,0090	0,0908	0,1058	0,7866	0,8199
	10	Q	0,0751	0,8111	-0,0249	0,0111	0,0254	0,0157	0,0051	0,0111	0,0640	0,0838	0,7904	0,8353
	10	R	0,0979	0,8018	-0,0021	0,0018	0,0043	0,0086	0,0038	0,0084	0,0910	0,1052	0,7854	0,8176
	$a = 0,05, b = 0,93$													
$\epsilon = 0\%$	0	Q	0,0495	0,9278	-0,0005	-0,0022	0,0013	0,0029	0,0013	0,0020	0,0468	0,0518	0,9240	0,9314
	0	R	0,0502	0,9275	0,0002	-0,0025	0,0012	0,0033	0,0012	0,0022	0,0483	0,0527	0,9221	0,9301
$\epsilon = 25\%$	5	Q	0,0483	0,9279	-0,0017	-0,0021	0,0022	0,0030	0,0014	0,0022	0,0458	0,0507	0,9242	0,9320
	5	R	0,0502	0,9273	0,0002	-0,0027	0,0014	0,0036	0,0014	0,0024	0,0478	0,0537	0,9216	0,9301
	10	Q	0,0451	0,9272	-0,0049	-0,0028	0,0054	0,0050	0,0023	0,0042	0,0405	0,0491	0,9204	0,9338
	10	R	0,0502	0,9273	0,0002	-0,0027	0,0014	0,0035	0,0014	0,0023	0,0482	0,0531	0,9221	0,9300
$\epsilon = 50\%$	5	Q	0,0467	0,9284	-0,0033	-0,0016	0,0036	0,0029	0,0015	0,0024	0,0440	0,0496	0,9239	0,9333
	5	R	0,0501	0,9274	0,0001	-0,0026	0,0014	0,0035	0,0014	0,0024	0,0475	0,0531	0,9214	0,9312
	10	Q	0,0402	0,9284	-0,0098	-0,0016	0,0101	0,0052	0,0025	0,0050	0,0348	0,0448	0,9180	0,9367
	10	R	0,0502	0,9272	0,0002	-0,0028	0,0014	0,0037	0,0014	0,0024	0,0478	0,0536	0,9226	0,9310

Os resultados são ilustrados nas Figuras 13, 14 e 15. Sob as três métricas de desempenho nota-se que os portfólios produzidos pelo estimador robusto apresentaram melhores resultados, seja por terem variância próximas ao do PVGM ou em termos de estimação dos pesos (ω_{1251}). Analisando a métrica de IPV percebe-se que os portfólios

estimados tem excesso de variância muito maior que os observados no estudo de séries bivariadas, naturalmente relacionando ao aumento na dimensão do problema. Dito isto, percebe-se que o estimador não robusto é altamente afetado pela presença de *outliers*. Por exemplo, para séries não contaminadas e considerando $\phi = (0, 10, 0, 80)$ o estimador tradicional apresentou, em média, excesso de variância percentual de 50% enquanto o robusto de 48%, o desvio padrão é similar. Com *outliers* isolados o método robusto reduz consideravelmente a variabilidade do portfólio estimado pelo método tradicional, no cenário mais parelho há uma redução de 78% para 54%. Com *outliers* consecutivos o excesso de variância percentual para o método tradicional chega a passar de 100% em média no pior cenário, enquanto que para o robusto o maior é de 56%. Em termos de desvio padrão, o estimador robusto apresentou variabilidade de resultados maior em todos os cenários; a maior diferença ocorre para 25% de séries contaminadas com saltos consecutivos de tamanho 10 com desvio padrão igual a 0,20 para o estimador robusto contra 0,13 para o tradicional. Para $\phi = (0, 05, 0, 93)$ as conclusões são similares, diferindo que as distribuições são mais assimétricas e os percentuais são em geral maiores; com 0% de contaminação a mediana de ambos é próxima a 0,80.

Tabela 5 – Métricas de desempenho para estimadores de a e b de modelo cDCC(1,1). Contaminação com saltos consecutivos, $T = 1.250$, $\delta = 0,975$, $N = 1.000$ e 100 réplicas. Em negrito, para cenário de contaminação, a melhor performance em termos de viés e REQM. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

	d	média		viés		REQM		DP		QE(95%)				
		a	b	a	b	a	b	a	b	a	b			
<i>a = 0,10, b = 0,80</i>														
$\epsilon = 0\%$	0	Q	0,0990	0,7968	-0,0010	-0,0032	0,0027	0,0064	0,0025	0,0056	0,0951	0,1033	0,7856	0,8070
	0	R	0,0992	0,8000	-0,0008	0,0000	0,0035	0,0076	0,0034	0,0076	0,0938	0,1054	0,7858	0,8129
$\epsilon = 25\%$	5	Q	0,0909	0,8062	-0,0091	0,0062	0,0098	0,0092	0,0035	0,0068	0,0831	0,0965	0,7935	0,8190
	5	R	0,0987	0,8003	-0,0013	0,0003	0,0039	0,0080	0,0037	0,0080	0,0912	0,1056	0,7850	0,8171
	10	Q	0,0603	0,8664	-0,0397	0,0664	0,0401	0,0672	0,0054	0,0105	0,0496	0,0697	0,8500	0,8872
	10	R	0,0989	0,7997	-0,0011	-0,0003	0,0039	0,0087	0,0038	0,0087	0,0913	0,1063	0,7811	0,8160
$\epsilon = 50\%$	5	Q	0,0804	0,8192	-0,0196	0,0192	0,0203	0,0212	0,0050	0,0091	0,0691	0,0873	0,8054	0,8387
	5	R	0,0980	0,8005	-0,0020	0,0005	0,0045	0,0087	0,0040	0,0087	0,0900	0,1051	0,7838	0,8177
	10	Q	0,0364	0,9121	-0,0636	0,1121	0,0637	0,1125	0,0040	0,0094	0,0281	0,0431	0,8976	0,9311
	10	R	0,0984	0,7999	-0,0016	-0,0001	0,0042	0,0091	0,0039	0,0091	0,0904	0,1052	0,7850	0,8193
<i>a = 0,05, b = 0,93</i>														
$\epsilon = 0\%$	0	Q	0,0495	0,9278	-0,0005	-0,0022	0,0013	0,0029	0,0013	0,0020	0,0468	0,0518	0,9240	0,9314
	0	R	0,0502	0,9275	0,0002	-0,0025	0,0012	0,0033	0,0012	0,0022	0,0483	0,0527	0,9221	0,9301
$\epsilon = 25\%$	5	Q	0,0476	0,9283	-0,0024	-0,0017	0,0028	0,0029	0,0015	0,0024	0,0441	0,0503	0,9235	0,9328
	5	R	0,0503	0,9270	0,0003	-0,0030	0,0014	0,0038	0,0014	0,0025	0,0482	0,0536	0,9212	0,9300
	10	Q	0,0399	0,9357	-0,0101	0,0057	0,0104	0,0070	0,0024	0,0041	0,0350	0,0443	0,9265	0,9425
	10	R	0,0503	0,9270	0,0003	-0,0030	0,0040	0,0016	0,0026	0,0476	0,0539	0,9202	0,9303	
$\epsilon = 50\%$	5	Q	0,0448	0,9300	-0,0052	-0,0000	0,0055	0,0028	0,0019	0,0029	0,0404	0,0480	0,9247	0,9354
	5	R	0,0504	0,9266	0,0004	-0,0034	0,0014	0,0043	0,0013	0,0026	0,0481	0,0532	0,9207	0,9300
	10	Q	0,0304	0,9464	-0,0196	0,0164	0,0198	0,0169	0,0024	0,0042	0,0260	0,0344	0,9381	0,9530
	10	R	0,0502	0,9268	0,0002	-0,0032	0,0014	0,0041	0,0014	0,0025	0,0480	0,0532	0,9220	0,9304

Um resumo das estatísticas de desempenho calculadas para os IPVs, razão entre as variâncias dos portfólios de variância mínima estimado e real e norma de Frobenius

são apresentadas no Apêndice A, nas Tabelas 12, 13 e 14, respectivamente. Os resultados corroboram os comentários anteriores.

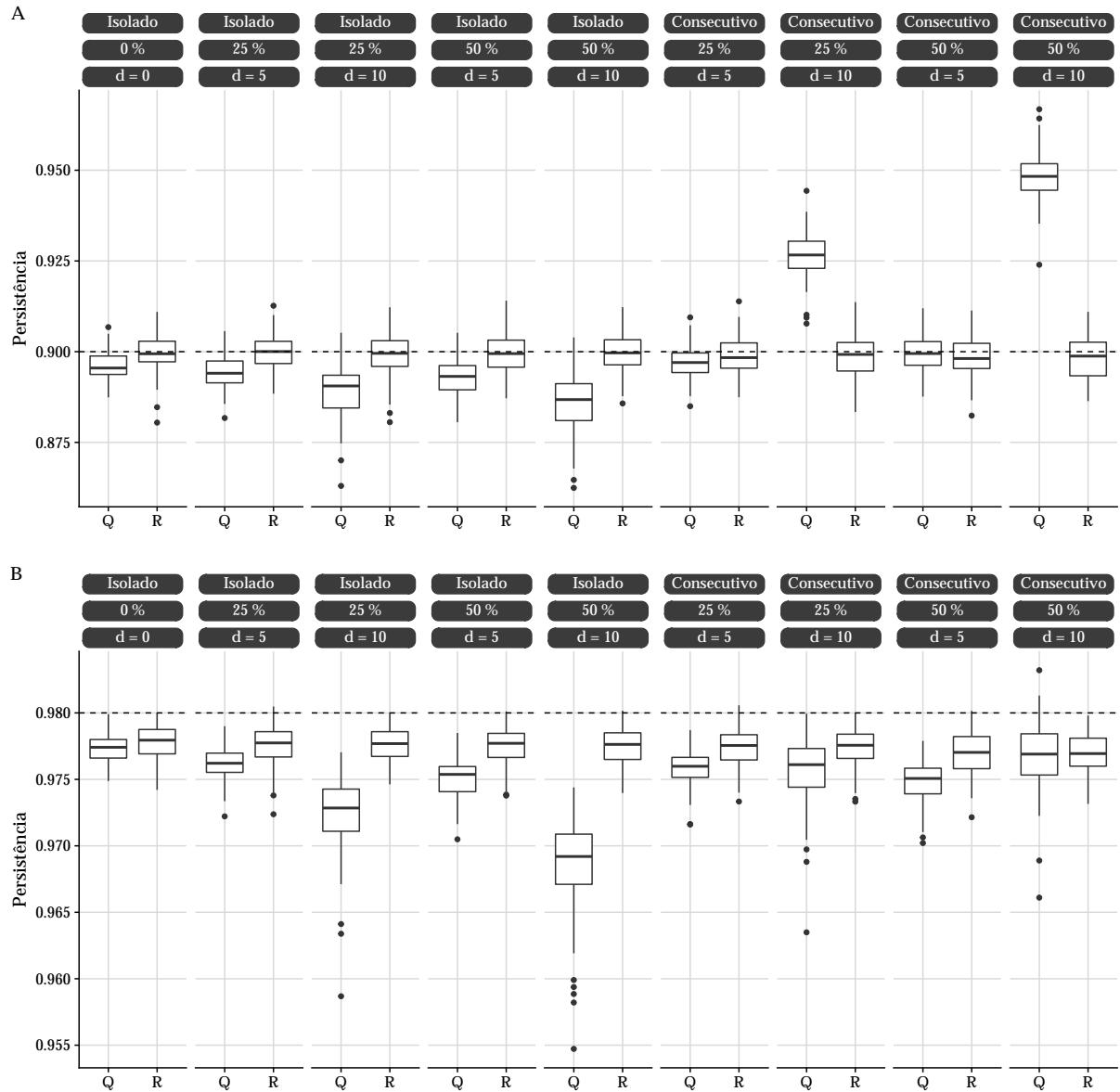


Figura 12 – Distribuições das persistências estimadas para modelo cDCC(1, 1) obtidas por $\hat{\phi}_q$ e $\hat{\phi}_r$. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\phi}_q$ e $\hat{\phi}_r$, respectivamente. Linha tracejada indica valor real de $(a + b)$. Em (A) e (B) estão os resultados para $\phi = (0, 10, 0, 80)'$ e $\phi = (0, 05, 0, 93)'$, respectivamente. $T = 2.000$, $\delta = 0,975$, $N = 2$ e 10.000 réplicas.

6.4 Desempenho do filtro de volatilidade

Os resultados para análise do desempenho do filtro de volatilidade na dimensão original ($N = 1.000$) são apresentados na Tabela 6. Diferente do observado para séries

bivariadas a acurácia do filtro de volatilidade é baixa, principalmente para $\phi = (0, 05, 0, 93)$. Quando há contaminação a acurácia (e especificidade) média sofre uma leve redução. Aparentemente em cenários de maior persistência, de ($a + b = 0, 90$) para ($a + b = 0, 98$), há redução da especificidade. Lembrando que neste cenário o vetor de retornos no tempo t está sendo comparado ao quantil δ de uma distribuição $\chi^2_{1,000}$, o filtro está sendo avaliado na etapa de estimação de \mathbf{H}_t e não dos parâmetros, que é feita considerando múltiplos pares consecutivos de séries. Os resultados indicam que para $\phi = (0, 05, 0, 93)$ é necessário utilizar um valor maior para δ de forma a tornar o filtro de volatilidade apresentado na Equação (5.2) mais “rígido” na classificação de *outliers*. Um estudo neste sentido é bem-vindo.

Tabela 6 – Desempenho do filtro de volatilidade na etapa cDCC de alta dimensão. $T = 1.250$, $N = 1.000$, $\delta = 0,975$ e 100 réplicas. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

d	ϵ	Especificidade	Sensibilidade	Acurácia
$a = 0, 10, b = 0, 80$				
0	0%	0,79	-	0,79
5	25%	0,78	1,00	0,78
5	50%	0,80	1,00	0,80
10	25%	0,78	1,00	0,78
10	50%	0,80	1,00	0,80
$a = 0, 05, b = 0, 93$				
0	0%	0,15	-	0,15
5	25%	0,16	1,00	0,16
5	50%	0,16	1,00	0,16
10	25%	0,16	1,00	0,16
10	50%	0,16	1,00	0,16

6.5 Conclusão

Nesta seção realizou-se um estudo de Monte Carlo para avaliar a performance de dois estimadores: tradicional e robusto, em conjunto com o método de máxima verossimilhança composta. O interesse foi avaliar exclusivamente os métodos na etapa cDCC, portanto considerou-se $h_{i,t} = 1$ para todas as séries e tempos. Esses estimadores foram avaliados em cenários com diferentes graus de contaminação a fim de validar possíveis vantagens que o estimador robusto pode trazer na modelagem de séries com *outliers* aditivos. O estudo foi dividido em 3 partes: (1) desempenho na estimação dos parâmetros, (2) eficácia na estimativa de portfólio de variância mínima e (3) desempenho do filtro de volatilidade. Em geral, considerando viés e REQM, o estimador robusto foi melhor para a .

Em b sua preferência depende do valor do parâmetro e da quantidade/tamanho dos saltos. Para $b = 0,80$ o estimador robusto obtém geralmente menor viés mas a preferência em termos de REQM só ocorre quando há saltos de tamanho 10. Quando $b = 0,93$, apenas o critério de REQM é favorável. Na prática, os portfólios de variância mínima obtidos com $\hat{\mathbf{H}}_r$, em média, foram melhores do que os obtidos pelo método tradicional, considerando as três métricas de desempenho e para diferentes valores de ϕ . Por fim, o filtro de volatilidade tem desempenho inferior se comparado ao estudo com séries bivariadas. Em cenários de maior persistência ($\phi = (0,05,0,93)$) a acurácia é ainda mais afetada, muitos retornos erroneamente classificados como *outliers*, uma alternativa é considerar uma valor maior de δ de forma a deixar o filtro mais rígido na detecção de retornos com saltos.

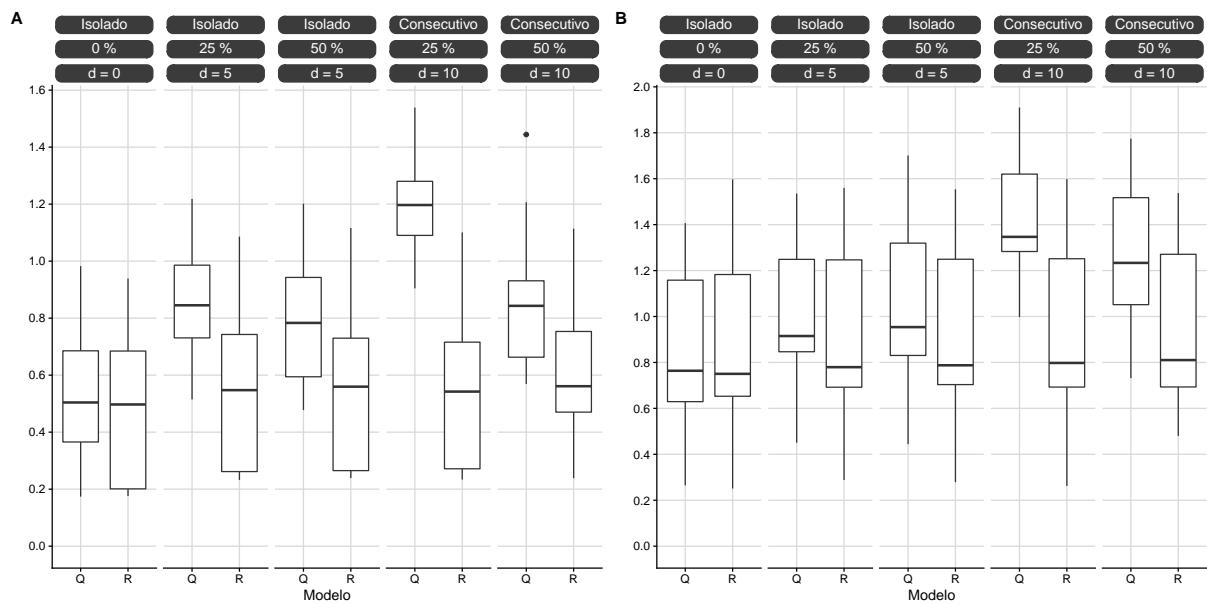


Figura 13 – IPV por modelo. Resultados para saltos consecutivos e isolados. $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Em (A) e (B) os resultados para $\phi = (0,10,0,80)$ e $(0,05,0,93)$, respectivamente. Os saltos são de magnitude d vezes o desvio padrão condicional.

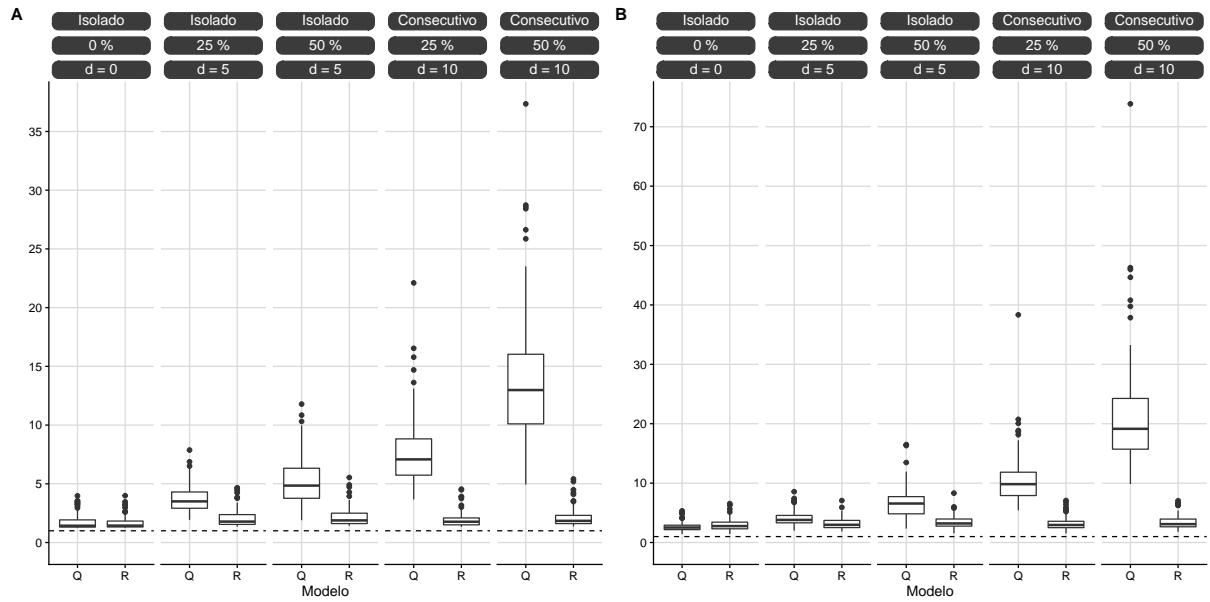


Figura 14 – Razão entre as variâncias dos portfólios de variância mínima estimado e real. Resultados para saltos consecutivos e isolados. $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Em (A) e (B) os resultados para $\phi = (0, 10, 0, 80)$ e $(0, 05, 0, 93)$, respectivamente. Os saltos são de magnitude d vezes o desvio padrão condicional.

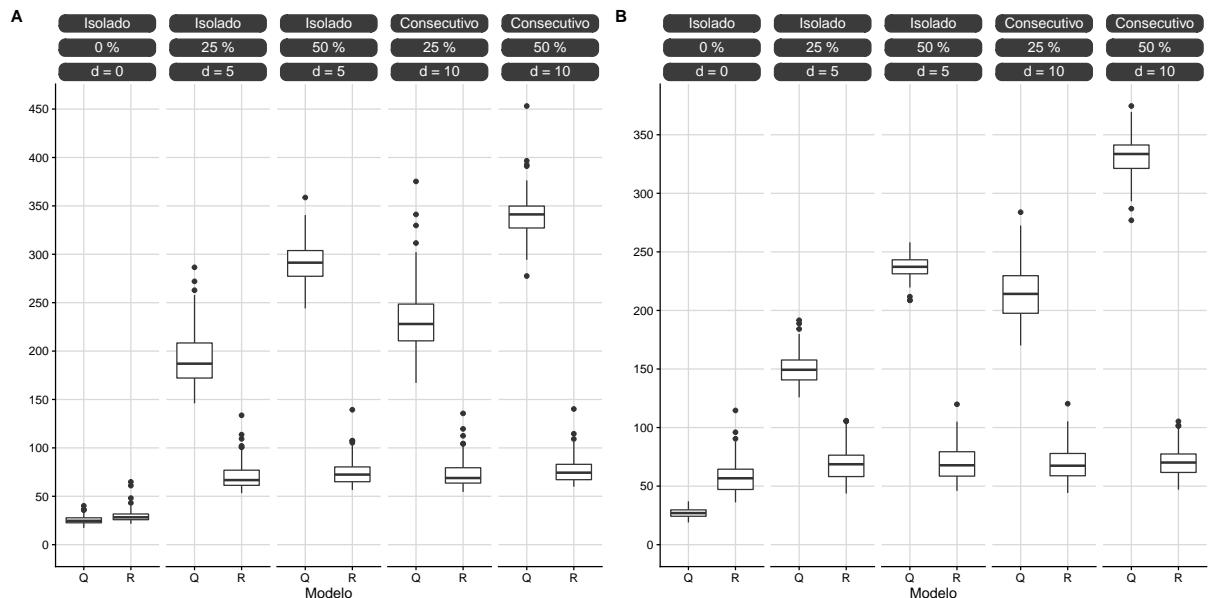


Figura 15 – Norma de Frobenius. Resultados para saltos consecutivos e isolados. $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Em (A) e (B) os resultados para $\phi = (0, 10, 0, 80)$ e $(0, 05, 0, 93)$, respectivamente. Os saltos são de magnitude d vezes o desvio padrão condicional.

7 Aplicação

Neste capítulo será avaliado o desempenho da predição 1 passo à frente dos métodos robusto e não robusto na seleção do MVP (*minimum variance portfolio*, portfólio de variância mínima). Os dados utilizados são retornos de ativos que compõem o S&P500, negociados de 13 de Janeiro de 1999 a 9 de Setembro de 2021. Os dados disponíveis em *finance.yahoo.com* foram coletados com auxílio do pacote *BatchGetSymbols* (PERLIN, 2020) disponível para o *software R* (R Core Team, 2021).

Como nem todos os ativos foram negociados durante todo o período do estudo, a dimensão do portfólio é variável com o mínimo de $N = 352$ a um máximo de $N = 485$. Seguindo Engle, Ledoit e Wolf (2019), será considerado um sistema de janela móvel ($T = 1.250$) para estimativa da matriz de covariâncias condicionais. Foram utilizadas duas estratégias de rebalanceamento: (1) mensal, em que os pesos são fixos durante um “mês”, mas os retornos do portfólio são calculados diariamente e (2) diário. Por “mês” considera-se 21 dias úteis consecutivos, uma prática comum. Na primeira estratégia desconsidera-se os custos de transação. Para (2) serão apresentados resultados com custos de 0, 1 e 10 pontos base, custos considerados em Boudt, Danielsson e Laurent (2013). Desta forma, no caso de rebalanceamento mensal (diário) foram considerados apenas os ativos sem falta de informação nas 1.250 observações da janela de estimativa e na 21 (1) observações seguintes, dessa forma a quantidade de ativos nos períodos dentro e fora da amostra são iguais. Além disso, para pares de ativos com correlação maior que 0,95, excluiu-se aquele que tinha menor volume de transações. O período fora da amostra vai de 2 de Janeiro de 2004 a 9 de Setembro de 2021 totalizando 4.453 retornos para rebalanceamentos diários, enquanto para rebalanceamentos mensais o MVP foi selecionado 213 vezes. Em ambos os casos os pesos são não negativos e menor do que 10%. Os anos foram classificados em períodos calmos e de crise. Os períodos de crise incluem a crise financeira (2007-2009) e a crise de Covid-19 (2020-2021). Os demais anos foram classificados como período sem crise.

A Figura 16 apresenta os ativos considerados na construção dos portfólios com rebalanceamento ao longo dos anos. Em preto se o respectivo ativo foi considerado na atividade de produção do MVP na respectiva data e em branco caso contrário. A dimensão do portfólio aumenta de forma constante. Até meados de 2009, nenhuma série que entrou saiu. Depois, poucas séries entraram e saíram, com maior frequência após 2016.

Devido ao sistema de janelas móveis e condições para inclusão/exclusão dos ativos, o cálculo para os custos de transação leva em conta àqueles que foram considerados na janela passada e atual simultaneamente. Portanto, os resultados envolvendo custos de transação devem ser considerados com um pouco mais de cuidado, visto que o conjunto de

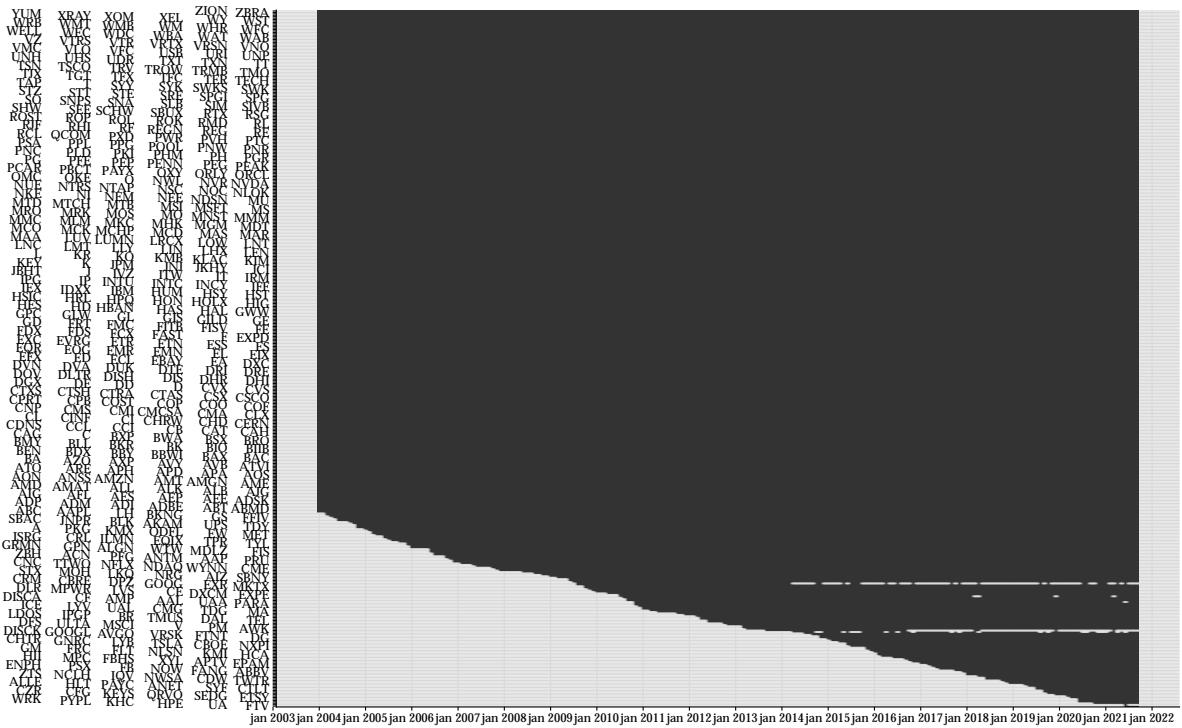


Figura 16 – Ativos considerados na construção de portfólios com rebalanceamento diário ao longo dos anos. Em preto se o respectivo ativo foi considerado na produção do MVP na respectiva data e em branco caso contrário.

ações consideradas entre janelas consecutivas pode variar.

A avaliação é feita considerando um conjunto de métricas de desempenho calculadas para k retornos de portfólio fora da amostra denotados por $R_k = \{r_{p,1}, \dots, r_{p,k}\}$. Como as matrizes de variância e covariâncias reais fora da amostra $\mathbf{H}_{T+1}, \dots, \mathbf{H}_{T+k}$ são desconhecidas, mensurar a qualidade das previsões usando (5.6) não é mais viável. Visto isso, as métricas de desempenho consideradas nesta aplicação são:

1. O retorno médio dos retornos anualizados (M):

$$M = 252 \frac{\sum_{i=1}^k r_{p,i}}{k};$$

2. O desvio padrão dos retornos anualizados (DP):

$$DP = \sqrt{252} \left(\frac{\sum_{i=1}^k (r_{p,i} - M/252)^2}{k} \right)^{1/2};$$

3. A Razão de informação (IR - *Information ratio*):

$$IR = M/DP;$$

4. A diferença entre JB dos portfólio selecionado pelos métodos robusto e não robusto:

5. O retorno médio após descontar os custos de transação de 1 e 10 pontos base como em [Boudt, Danielsson e Laurent \(2013\)](#);
6. IR após descontar os custos de transação.
7. Diferença entre IR dos portfólio selecionado pelos métodos robusto e não robusto considerando custos de transação.

Para testar a hipótese nula de igualdade das variâncias dos portfólios estimados pelos métodos não robusto e robusto foi utilizado o teste *bootstrap* de [Ledoit e Wolf \(2011\)](#), que é apresentado no Apêndice [C](#).

7.1 Resultados

O procedimento de estimação de \mathbf{H}_t é similar ao apresentado na Seção [6.3](#), com adição da etapa de estimação das volatilidades. Os resultados da estimação de ϕ para as janelas de rebalanceamento diário são apresentados na Figura [17](#). Para os dois estimadores é notável que as estimativas exibem mudanças de nível no decorrer do tempo, justificando a utilização de janelas móveis. Adicionalmente, percebe-se que em determinados períodos há mais variabilidade, destacando-se dois: 2006-2009 e 2016-2020. Isso pode ser parcialmente explicado pela inclusão e/ou exclusão de ativos ao longo do tempo além da inclusão e exclusão de dias. Com respeito a a , não é tão claro estabelecer qual estimador produziu estimativas de menor variabilidade. Já para b a vantagem do estimador robusto parece ser mais evidente principalmente para os períodos de alta variabilidade supracitados, em que o estimador tradicional exibe mudanças de nível muito mais evidentes. Portanto, em geral, o estimador robusto tem menor variabilidade, principalmente para b .

Na Figura [18](#) está ilustrada a distribuição conjunta das estimativas robustas e tradicionais, os pontos no gráfico de dispersão estão mapeados por linhas a fim de conservar a estrutura temporal das estimativas produzidas. Note que em nenhum dos dois estimadores há casos de linhas conectando pontos distantes, que reflete certa consistência das estimativas geradas. Nos dois métodos os pontos seguem a relação de $a + b < 1$, além disso, o robusto apresenta menor variabilidade da persistência $a + b$.

Os resultados para a estratégia envolvendo rebalanceamento mensal estão presentes na Tabela [7](#). Visto que o objetivo é obter o MVP, a análise dos resultados prioriza o desvio padrão como critério principal para escolha do melhor modelo. Em grande parte dos anos os retornos dos portfólios construídos utilizando estimador robusto apresentaram menor variância (11 em 18 anos), em grande parte sendo significativamente menores em comparação ao estimador tradicional a 1% de significância (7), tendo menor variância em todos os anos de crise. Considerando todo o período, o desvio padrão para os retornos dos portfólios obtidos usando o método robusto foi de 12,78; um redução

significativa, comparado ao valor 12,89 encontrado nos portfólios estimados pelo método tradicional (considerando 1% de significância). Em termos de IR o número de anos com predominância do robusto é maior (14). No geral, o método robusto tem melhor desempenho.

Para rebalanceamento diários as conclusões são similares, como pode ser consultado na Tabela 8. Em comparação aos resultados de rebalanceamento mensais a quantidade de anos em que o método robusto proporcionou redução na variâncias dos portfólios aumenta (13); com 10 significativas à 1% de significância. Novamente, com melhor desempenho em todos os anos de crise. Com relação ao índice de informação, em 6 anos houve vantagem do estimador robusto para todos os custos de transação, em 2 anos para os dois primeiros (0 e 1 pontos base) e em 1 para o último custo de transação (10 pontos base). Com custos de 0 e 1 pontos base o estimador robusto obteve IR's maiores em 8 anos, com ampla vantagem em 2020-2021. Para 10 pontos base acabou sendo melhor em 7 anos sendo 3 destes de crise. Para o histórico completo, o estimador robusto foi melhor com 0 e 1 pontos base.

7.2 Conclusão

Nesta seção foram realizados o comparativo entre os estimadores tradicional e robusto apresentados na Seção 6.3, com adição da etapa de estimação das volatilidades, na atividade de construção do portfólio de variância mínima. Os dados utilizados são retornos de ativos que compõem o S&P500, negociados entre 13 de Janeiro de 1999 e 9 de Setembro de 2021. O período fora da amostra inicia em 2 de Janeiro de 2004. Considerou-se um sistema de janela móvel em que as últimas 1.250 observações são consideradas para estimação da matriz de covariâncias condicionais. Rebalanceamentos diários e mensais também foram considerados. Em geral, as estimativas geradas pelo método robusto exibiram menor variabilidade, principalmente para b . Além disso, \mathbf{H}_t prevista resultou em retornos de portfólios com menor variabilidade em 11 (13) dos 18 anos avaliados considerando rebalanceamentos mensais (diários). Para os custos de transação de 0 e 1 pontos base o estimador robusto obteve maiores IR's em 8 anos, conseguindo ser amplamente vantajoso durante os anos da crise de Covid. Com 10 pontos base foi superior em 7 anos sendo 3 destes de crise. Para o histórico completo, o estimador robusto foi melhor com 0 e 1 pontos base.

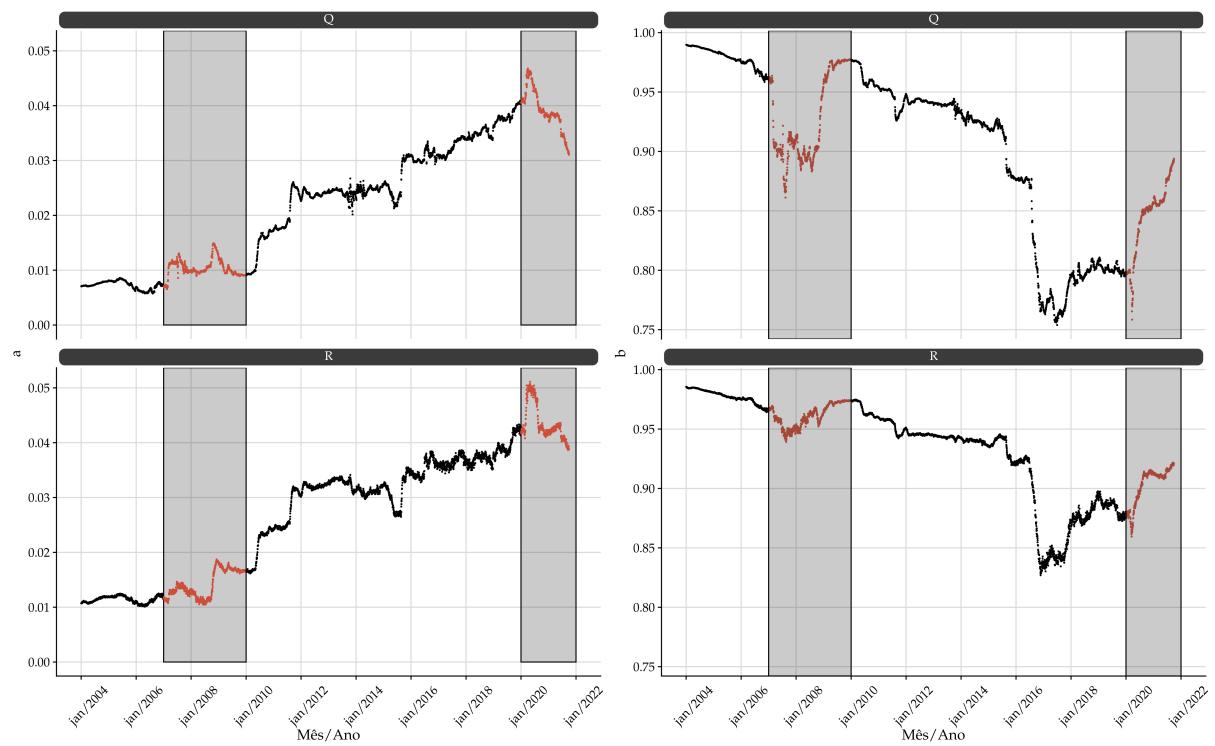


Figura 17 – Estimativas produzidas pelos métodos tradicional (Q) (figura superior) robusto (R) (figura inferior) para rebalanceamento diário. À esquerda estimativas para o parâmetro a e à direita para parâmetro b . Em destaque, anos de crise.

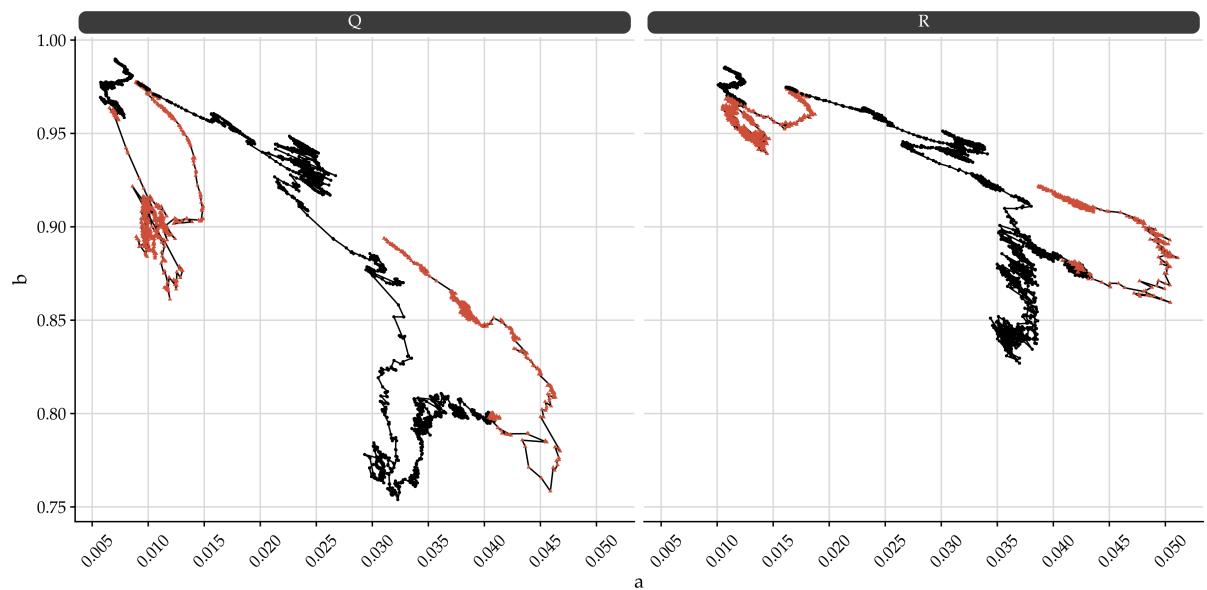


Figura 18 – Distribuição conjunta das estimativas de a e b obtidas pelos métodos tradicional (Q) e robusto (R) na atividade de estimativa do MVP para séries de retornos do S&P500 considerando rebalanceamentos diários. Em destaque, anos de crise.

Tabela 7 – Medidas de desempenho dos portfólios selecionados pelos métodos tradicional (Q) e robusto (R) com rebalanceamento mensal. Em negrito, para cada ano, a melhor performance em termos de média (M), desvio padrão (DP) e quando a diferença entre os índices de informação (ΔIR) tradicional e robusto é positiva. M e DP são porcentagens anualizadas. Na coluna DP(R) (*), (**) e (***) indicam significância a um nível de 10%, 5% e 1%, respectivamente, para o teste de igualdade de variâncias por *bootstrap* entre os métodos robusto e tradicional. Linhas sombreadas indicam anos com crise econômica.

Ano	M		DP		ΔIR
	Q	R	Q	R	
2004	15,78	17,81	9,12	9,11	0,22
2005	2,42	3,52	8,65	8,51 ***	0,13
2006	21,01	18,19	7,24	7,17 ***	-0,36
2007	1,45	1,78	10,89	10,35 ***	0,04
2008	-19,30	-23,42	24,25	23,77 ***	-0,19
2009	14,61	15,33	14,99	14,49 ***	0,08
2010	8,52	9,78	10,74	11,01 ***	0,10
2011	13,14	13,19	13,10	13,15	< 0,01
2012	6,45	8,19	7,63	7,60 **	0,23
2013	10,55	13,96	8,82	8,82	0,39
2014	14,94	14,14	9,03	8,99 **	-0,08
2015	-0,14	6,23	12,65	12,43 ***	0,51
2016	14,25	17,51	9,63	9,70 **	0,33
2017	12,54	13,12	5,27	5,31 ***	0,09
2018	-10,74	-10,48	11,74	12,17 ***	0,05
2019	17,37	21,96	8,40	8,46 **	0,53
2020	12,61	19,78	27,00	26,91	0,27
2021	20,82	22,63	9,93	9,43 ***	0,30
Total	8,42	9,91	12,89	12,78 ***	0,12

Tabela 8 – Medidas de desempenho dos portfólios selecionados pelo métodos tradicional (Q) e robusto (R) com rebalanceamento diário. Em negrito, para cada ano, a melhor performance em termos de média (M), desvio padrão (DP) e quando a diferença entre os índices de informação (ΔIR) tradicional e robusto é positiva. M e DP são porcentagens anualizadas. Na coluna DP(R) (*), (**) e (***) indicam significância a um nível de 10%, 5% e 1%, respectivamente, para o teste de igualdade de variâncias por *bootstrap* entre os métodos robusto e tradicional. Linhas sombreadas indicam anos com crise econômica. O (ΔIR^{net}) é apresentado para custos de transação de c pontos base.

Ano	M		DP		ΔIR^{net}		
	Q	R	Q	R	0	10^{-4}	10^{-3}
2004	16,09	16,47	8,43	8,45	0,04	0,01	-0,27
2005	1,50	1,33	8,26	7,97 ***	-0,01	-0,04	-0,25
2006	22,00	18,89	7,11	7,04 ***	-0,41	-0,43	-0,60
2007	5,59	4,87	10,46	10,22 *	-0,06	-0,06	-0,11
2008	-24,01	-23,10	24,87	22,87 ***	-0,05	-0,04	0,02
2009	11,99	11,05	14,73	14,55 ***	-0,05	-0,06	-0,07
2010	14,70	17,76	10,44	10,36 **	0,30	0,29	0,19
2011	15,18	13,62	12,76	12,68 ***	-0,12	-0,12	-0,20
2012	6,98	8,38	7,69	7,56 ***	0,20	0,18	0,01
2013	15,20	16,27	8,83	9,07 ***	0,07	0,05	-0,12
2014	20,67	19,26	8,71	8,54 ***	-0,12	-0,12	-0,17
2015	4,80	4,96	12,58	12,19 ***	0,03	0,03	0,06
2016	16,84	17,27	9,39	9,07 ***	0,11	0,11	0,09
2017	14,58	12,89	5,24	5,35 ***	-0,38	-0,41	-0,68
2018	1,91	-4,35	11,50	11,68 ***	-0,54	-0,53	-0,47
2019	21,00	19,70	8,75	8,86 ***	-0,18	-0,18	-0,17
2020	14,22	18,98	25,00	24,64 **	0,20	0,21	0,25
2021	19,89	22,85	10,45	10,00 ***	0,38	0,37	0,25
Total	10,90	10,73	12,59	12,26 ***	0,01	(0; 0,01)	(-0,01; 0)

8 Considerações Finais

A dissertação teve por objetivo robustecer o estimador DCC de alta dimensão proposto por [Engle, Ledoit e Wolf \(2019\)](#) à *outliers* aditivos.

No Capítulo 5, no estudo de simulação de Monte Carlo com séries bivariadas, o estimador robusto obteve melhores resultados tanto na parte GARCH quanto na cDCC, em termos de viés e REQM das estimativas dos parâmetros. Embora que em cenários com um alto percentual de contaminação (10%) as estimativas obtidas tenham viés consideravelmente alto. Além disso, o BIP-GARCH-BIP-cDCC trouxe ganhos consideráveis na escolha do portfólio de variância mínima.

No Capítulo 6 o estudo de simulação é estendido para alta dimensão com $N = 1.000$. Esta etapa avaliou exclusivamente a estimação das correlações condicionais, mantendo-se $h_{i,t} = 1$ para todo i e t , considerando diferentes estruturas de dependência temporal para as volatilidades multivariadas: $\phi = (0, 10, 0, 80)$ e $(0, 05, 0, 93)$. Em termos de estimação de parâmetros, o estimador robusto apresentou menor vício e REQM para a em quase todos os cenários, para b sua preferência depende do valor do parâmetro e da quantidade/tamanho dos saltos, em específico tendo melhor desempenho quando $d = 10$. Na escolha do portfólio de variância mínima $\hat{\mathbf{H}}_r$ obteve melhores resultados sob as três métricas de desempenho consideradas. Por fim, na dimensão original o filtro de volatilidade é tem baixa acurácia principalmente com $\phi = (0, 05, 0, 93)$.

No Capítulo 7, os estimadores foram avaliados em uma aplicação envolvendo séries de retornos de ativos que compõem o S&P500. Os comparativos foram feitos considerando duas estratégias de rebalanceamento: (1) mensal e (2) diário. No quesito de estimação dos parâmetros o estimador robusto produziu, em geral, estimativas com menor variabilidade principalmente para b ; além disso \mathbf{H}_t predita produziu retornos de portfólios com menor variabilidade na maioria dos anos inclusive nos de crise, tanto para rebalanceamentos diários quanto mensais. Com rebalanceamentos mensais, o IR pelo método robusto foi maior na maioria dos anos. Já para rebalanceamentos diários a preferência por $\hat{\phi}_r$ persiste apenas para 0 e 1 pontos base.

Neste trabalho, considerou-se apenas um estimador robusto a *outliers* aditivos, o que de certa forma é uma limitação. Trabalhos futuros devem considerar outros tipos de *outliers* (veja, por exemplo [Tsay, Pena e Pankratz \(2000\)](#)). Além disso, cabe investigar alternativas para melhor desempenho do estimador quando o percentual de tempos com contaminação é elevado (10%, por exemplo). Por fim, é importante avaliar os possíveis ganhos de se utilizar um valor maior para δ de forma a tornar o filtro mais rígido na detecção de *outliers*.

Referências

- AIELLI, G. P. Dynamic conditional correlation: on properties and estimation. *Journal of Business & Economic Statistics*, v. 31, n. 3, p. 282–299, 2013. Citado 6 vezes nas páginas 18, 19, 30, 32, 33 e 34.
- ALMEIDA, D. de; HOTTA, L. K.; RUIZ, E. MGARCH models: Trade-off between feasibility and flexibility. *International Journal of Forecasting*, v. 34, n. 1, p. 45–63, 2018. Citado na página 18.
- ANDERSEN, T. G.; BOLLERSLEV, T.; DIEBOLD, F. X.; LABYS, P. Modeling and forecasting realized volatility. *Econometrica*, v. 71, n. 2, p. 579–625, 2003. Citado na página 23.
- ANDREWS, D. W. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica: Journal of the Econometric Society*, p. 817–858, 1991. Citado na página 102.
- BAUWENS, L.; LAURENT, S.; ROMBOUTS, J. V. Multivariate GARCH models: a survey. *Journal of Applied Econometrics*, v. 21, n. 1, p. 79–109, 2006. Citado 2 vezes nas páginas 18 e 29.
- BERKES, I.; HORVÁTH, L.; KOKOSZKA, P. GARCH processes: structure and estimation. *Bernoulli*, v. 9, n. 2, p. 201–227, 2003. Citado na página 24.
- BODNAR, T.; MAZUR, S.; OKHRIN, Y. Bayesian estimation of the global minimum variance portfolio. *European Journal of Operational Research*, v. 256, n. 1, p. 292–307, 2017. Citado na página 53.
- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, v. 31, n. 3, p. 307–327, 1986. Citado 4 vezes nas páginas 18, 20, 21 e 22.
- _____. A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, The MIT Press, v. 69, n. 3, p. 542–547, 1987. Citado na página 25.
- _____. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *The Review of Economics and Statistics*, p. 498–505, 1990. Citado na página 29.
- BOLLERSLEV, T.; CHOU, R. Y.; KRONER, K. F. ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, v. 52, n. 1-2, p. 5–59, 1992. Citado na página 21.
- BOLLERSLEV, T.; ENGLE, R. F.; WOOLDRIDGE, J. M. A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, v. 96, n. 1, p. 116–131, 1988. Citado na página 29.
- BOLLERSLEV, T.; WOOLDRIDGE, J. M. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, v. 11, n. 2, p. 143–172, 1992. Citado na página 24.

- BOUDT, K.; CROUX, C. Robust M-estimation of multivariate GARCH models. *Computational Statistics & Data Analysis*, v. 54, n. 11, p. 2459–2469, 2010. Citado 2 vezes nas páginas 27 e 36.
- BOUDT, K.; CROUX, C.; LAURENT, S. Outlyingness weighted covariation. *Journal of Financial Econometrics*, v. 9, n. 4, p. 657–684, 2011. Citado na página 28.
- BOUDT, K.; DANIELSSON, J.; LAURENT, S. Robust forecasting of dynamic conditional correlation GARCH models. 2011. Citado na página 49.
- BOUDT, K.; DANIELSSON, J.; LAURENT, S. Robust forecasting of dynamic conditional correlation GARCH models. *International Journal of Forecasting*, v. 29, n. 2, p. 244–257, 2013. Citado 14 vezes nas páginas 18, 22, 23, 25, 26, 27, 35, 36, 37, 38, 47, 51, 75 e 77.
- BOUDT, K.; GALANOS, A.; PAYSEUR, S.; ZIVOT, E. Multivariate GARCH models for large-scale applications: A survey. In: *Handbook of Statistics*. [S.l.]: Elsevier, 2019. v. 41, p. 193–242. Citado 2 vezes nas páginas 18 e 29.
- BOUGEROL, P.; PICARD, N. Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics*, v. 52, n. 1-2, p. 115–127, 1992. Citado na página 21.
- BROCKWELL, P. J.; DAVIS, R. A. *Introduction to Time Series and Forecasting*. 3. ed. [S.l.]: Springer Cham, 2016. (Springer Texts in Statistics). Citado na página 20.
- CARNERO, M. A.; PEÑA, D.; RUIZ, E. Estimating GARCH volatility in the presence of outliers. *Economics Letters*, v. 114, n. 1, p. 86–90, 2012. Citado 3 vezes nas páginas 23, 25 e 26.
- COX, D. R.; REID, N. A note on pseudolikelihood constructed from marginal densities. *Biometrika*, v. 91, n. 3, p. 729–737, 2004. Citado na página 41.
- CRYER, J. D.; CHAN, K.-S. *Time Series Analysis: with Applications in R*. [S.l.]: Springer Science & Business Media, 2008. Citado na página 24.
- EDDELBUETTEL, D.; BALAMUTA, J. J. Extending extitR with extitC++: A Brief Introduction to extitRcpp. *The American Statistician*, v. 72, n. 1, p. 28–36, 2018. Disponível em: <<https://doi.org/10.1080/00031305.2017.1375990>>. Citado na página 19.
- EDDELBUETTEL, D.; SANDERSON, C. Rcpparmadillo: Accelerating R with high-performance C++ linear algebra. *Computational Statistics and Data Analysis*, v. 71, p. 1054–1063, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.csda.2013.02.005>>. Citado na página 19.
- EFRON, B.; TIBSHIRANI, R. J. *An Introduction to the Bootstrap*. 29 West 35th Street New York, NY 10001-2299: Chapman & Hall, 1993. Citado na página 102.
- ENGLE, R. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, v. 20, n. 3, p. 339–350, 2002. Citado 2 vezes nas páginas 18 e 30.
- _____. Risk and volatility: Econometric models and financial practice. *American Economic Review*, v. 94, n. 3, p. 405–420, 2004. Citado na página 20.

- ENGLE, R.; COLACITO, R. Testing and valuing dynamic correlations for asset allocation. *Journal of Business & Economic Statistics*, v. 24, n. 2, p. 238–253, 2006. Citado na página 54.
- ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, p. 987–1007, 1982. Citado na página 20.
- ENGLE, R. F.; KRONER, K. F. Multivariate simultaneous generalized ARCH. *Econometric Theory*, p. 122–150, 1995. Citado na página 29.
- ENGLE, R. F.; LEDOIT, O.; WOLF, M. Large dynamic covariance matrices. *Journal of Business & Economic Statistics*, v. 37, n. 2, p. 363–375, 2019. Citado 6 vezes nas páginas 19, 44, 46, 54, 75 e 82.
- ENGLE, R. F.; NG, V. K.; ROTHSCHILD, M. Asset pricing with a factor-ARCH covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics*, v. 45, n. 1-2, p. 213–237, 1990. Citado na página 29.
- ENGLE, R. F.; SHEPPARD, K. *Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH*. [S.l.], 2001. (Working Paper Series, 8554). Citado na página 31.
- FERMANIAN, J.-D.; MALONGO, H. On the stationarity of dynamic conditional correlation models. *Econometric Theory*, v. 33, n. 3, p. 636–663, 2017. Citado na página 33.
- FRANSES, P. H.; GHIJSELS, H. Additive outliers, GARCH and forecasting volatility. *International Journal of Forecasting*, v. 15, n. 1, p. 1–9, 1999. Citado 2 vezes nas páginas 22 e 23.
- FRIEDMAN, M. Nobel lecture: inflation and unemployment. *Journal of Political Economy*, v. 85, n. 3, p. 451–472, 1977. Citado na página 20.
- GIRAITIS, L.; KOKOSZKA, P.; LEIPUS, R. Stationary ARCH models: dependence structure and central limit theorem. *Econometric Theory*, p. 3–22, 2000. Citado na página 22.
- HALL, P.; YAO, Q. Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica*, v. 71, n. 1, p. 285–317, 2003. Citado 2 vezes nas páginas 21 e 24.
- HEALY, M. Multivariate normal plotting. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, v. 17, n. 2, p. 157–161, 1968. Citado na página 38.
- HOTTA, L. K.; TRUCÍOS, C. Inference in (M)GARCH models in the presence of additive outliers: Specification, estimation, and prediction. In: LAVOR, C.; GOMES, F. A. M. (Ed.). Cham: Springer International Publishing, 2018. p. 179–202. ISBN 978-3-319-94015-1. Citado 3 vezes nas páginas 23, 25 e 26.
- HOTTA, L. K.; TSAY, R. S. Outliers in GARCH processes. In: BELL, W. R.; HOLAN, S. H.; MCELROY, T. S. (Ed.). *Economic Time Series: Modeling and Seasonality*. [S.l.: s.n.], 2012. p. 337–358. Citado 2 vezes nas páginas 22 e 23.
- HUBER, P. J.; RONCHETTI. *Robust Statistics*. Segunda ed. Hoboken, NJ: John Wiley & Sons, 2009. Citado 2 vezes nas páginas 25 e 47.

- LEDOIT, O.; PÉCHÉ, S. Eigenvectors of some large sample covariance matrix ensembles. *Probability Theory and Related Fields*, v. 151, n. 1, p. 233–264, 2011. Citado 2 vezes nas páginas 19 e 46.
- LEDOIT, O.; WOLF, M. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, v. 88, n. 2, p. 365–411, 2004. Citado na página 46.
- _____. Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, v. 15, n. 5, p. 850–859, 2008. Citado na página 103.
- _____. Robust performances hypothesis testing with the variance. *Wilmott*, v. 2011, n. 55, p. 86–89, 2011. Citado 3 vezes nas páginas 77, 101 e 102.
- _____. Nonlinear shrinkage estimation of large-dimensional covariance matrices. *The Annals of Statistics*, v. 40, n. 2, p. 1024–1060, 2012. Citado na página 19.
- _____. Spectrum estimation: A unified framework for covariance matrix estimation and PCA in large dimensions. *Journal of Multivariate Analysis*, v. 139, p. 360–384, 2015. Citado 2 vezes nas páginas 45 e 46.
- _____. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies*, v. 30, n. 12, p. 4349–4388, 2017. Citado na página 55.
- LINDSAY, B. G. Composite likelihood methods. *Contemporary Mathematics*, v. 80, n. 1, p. 221–239, 1988. Citado na página 41.
- LINTON, O. B. Semiparametric and nonparametric ARCH modelling. In: ANDERSEN, T. G.; DAVIS, R. A.; KREIß, J.-P.; MIKOSCH, T. (Ed.). *Handbook of Financial Time Series*. 1. ed. [S.l.]: Springer Berlin, Heidelberg, 2008. Citado na página 30.
- MENDES, B. V. D. M. Assessing the bias of maximum likelihood estimates of contaminated GARCH models. *Journal of Statistical Computation and Simulation*, v. 67, n. 4, p. 359–376, 2000. Citado na página 23.
- MEZRICH, J.; ENGLE, R. GARCH for groups. *Risk*, v. 9, n. 8, p. 36–40, 1996. Citado na página 28.
- MULER, N.; YOHAI, V. J. Robust estimates for GARCH models. *Journal of Statistical Planning and Inference*, v. 138, n. 10, p. 2918–2940, 2008. Citado 5 vezes nas páginas 18, 24, 25, 26 e 27.
- NELSON, D. B. Stationarity and persistence in the GARCH (1, 1) model. *Econometric Theory*, v. 6, n. 3, p. 318–334, 1990. Citado na página 21.
- PAKEL, C.; SHEPHARD, N.; SHEPPARD, K.; ENGLE, R. F. Fitting vast dimensional time-varying covariance models. *Journal of Business & Economic Statistics*, v. 39, n. 3, p. 652–668, 2021. Disponível em: <<https://doi.org/10.1080/07350015.2020.1713795>>. Citado 4 vezes nas páginas 18, 41, 43 e 66.
- PERLIN, M. *BatchGetSymbols: Downloads and Organizes Financial Data for Multiple Tickers*. [S.l.], 2020. R package version 2.6.1. Disponível em: <<https://CRAN.R-project.org/package=BatchGetSymbols>>. Citado na página 75.

- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>. Citado 2 vezes nas páginas 19 e 75.
- RAMPRASAD, P. *nlshrink: Non-Linear Shrinkage Estimation of Population Eigenvalues and Covariance Matrices*. [S.l.], 2016. R package version 1.0.1. Disponível em: <<https://CRAN.R-project.org/package=nlshrink>>. Citado na página 19.
- RAYMAEKERS, J.; ROUSSEEUW, P. J. Fast robust correlation for high-dimensional data. *Technometrics*, v. 63, n. 2, p. 184–198, 2021. Citado na página 47.
- SILVENNOINEN, A.; TERÄSVIRTA, T. Multivariate GARCH models. In: ANDERSEN, T. G.; DAVIS, R. A.; KREIß, J.-P.; MIKOSCH, T. (Ed.). *Handbook of Financial Time Series*. [S.l.]: Springer, 2009. p. 201–229. Citado 2 vezes nas páginas 29 e 30.
- SPEARMAN, C. "General Intelligence" objectively determined and measured. *The American Journal of Psychology*, v. 15, n. 2, p. 201–292, 1904. Citado na página 47.
- STRAUMANN, D.; MIKOSCH, T. Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *The Annals of Statistics*, v. 34, n. 5, p. 2449–2495, 2006. Citado na página 24.
- TRUCÍOS, C.; HOTTA, L. K.; PEREIRA, P. L. V. On the robustness of the principal volatility components. *Journal of Empirical Finance*, v. 52, p. 201–219, 2019. Citado na página 66.
- TRUCÍOS, C.; HOTTA, L. K.; RUIZ, E. Robust bootstrap forecast densities for GARCH returns and volatilities. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 87, n. 16, p. 3152–3174, 2017. Citado 2 vezes nas páginas 18 e 28.
- _____. Robust bootstrap densities for dynamic conditional correlations: implications for portfolio selection and value-at-risk. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 88, n. 10, p. 1976–2000, 2018. Citado 2 vezes nas páginas 18 e 28.
- TRUCÍOS, C.; MAZZEU, J. H.; HOTTA, L. K.; PEREIRA, P. L. V.; HALLIN, M. Robustness and the general dynamic factor model with infinite-dimensional space: identification, estimation, and forecasting. *International Journal of Forecasting*, Elsevier, v. 37, n. 4, p. 1520–1534, 2021. Citado na página 48.
- TSAY, R. S.; PENA, D.; PANKRATZ, A. E. Outliers in multivariate time series. *Biometrika*, v. 87, n. 4, p. 789–804, 2000. Citado na página 82.
- TSE, Y. K.; TSUI, A. K. C. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics*, v. 20, n. 3, p. 351–362, 2002. Citado na página 30.
- VARIN, C.; REID, N.; FIRTH, D. An overview of composite likelihood methods. *Statistica Sinica*, v. 21, n. 1, p. 5–42, 2011. Citado na página 41.
- VARIN, C.; VIDONI, P. A note on composite likelihood inference and model selection. *Biometrika*, v. 92, n. 3, p. 519–528, 2005. Citado na página 41.

VERHOEVEN, P.; MCALEER, M. Modeling outliers and extreme observations for ARMA-GARCH processes. In: ECONOMETRIC SOCIETY. *Econometric Society World Congress 2000 Contributed Papers No. 1922*. [S.l.], 2000. Citado 2 vezes nas páginas 22 e 24.

WEISS, A. A. Asymptotic theory for ARCH models: estimation and testing. *Econometric Theory*, v. 2, n. 1, p. 107–131, 1986. Citado na página 24.

WICKHAM, H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN 978-3-319-24277-4. Disponível em: <<https://ggplot2.tidyverse.org>>. Citado na página 19.

YOHAI, V. J. High breakdown-point and high efficiency robust estimates for regression. *The Annals of Statistics*, v. 15, n. 20, p. 642–656, 1987. Citado na página 25.

APÊNDICE A – Tabelas Complementares dos Capítulos 5 e 6

Neste apêndice são apresentadas algumas tabelas com resultados das simulações. As Tabelas 9, 10 e 11 apresentam alguns resultados da Seção 5.5.2; mais precisamente, relativos a IPVs, razão entre as variâncias dos portfólios de variância mínima estimado e real e norma de Frobenius são apresentadas nas, respectivamente. Estas tabelas são relativas a modelos cDCC bivariados. As Tabelas 12, 13 e 14 apresentam um resumo do IPV, razão entre variâncias dos portfólios de variância mínima e norma de Frobenius para as simulações para modelos cDCC em altas dimensões apresentadas na Seção 6.3.

Tabela 9 – Média, desvio padrão e quantis empíricos de IPV calculado para portfólios estimados sob método tradicional (Q) e robusto (R). $T = 2.000$, $\delta = 0,975$, $N = 2$ e 10.000 réplicas.

			Média	Desvio Padrão	2,5%	97,5%
$\epsilon = 0\%$	d = 0	Q	0,0019	0,0028	0,0000	0,0086
	d = 0	R	0,0028	0,0049	0,0001	0,0142
$\epsilon = 1\%$	d = 3	Q	0,0075	0,0119	0,0001	0,0378
	d = 3	R	0,0030	0,0054	0,0001	0,0153
	d = 4	Q	0,0183	0,0291	0,0002	0,0928
	d = 4	R	0,0030	0,0054	0,0001	0,0155
$\epsilon = 5\%$	d = 3	Q	0,0976	0,1020	0,0016	0,3770
	d = 3	R	0,0112	0,0153	0,0002	0,0527
	d = 4	Q	0,1746	0,1455	0,0067	0,5512
	d = 4	R	0,0122	0,0171	0,0002	0,0576
$\epsilon = 10\%$	d = 3	Q	0,1808	0,1432	0,0065	0,5462
	d = 3	R	0,0455	0,0541	0,0007	0,1901
	d = 4	Q	0,2862	0,1821	0,0258	0,7300
	d = 4	R	0,0578	0,0678	0,0010	0,2407

Tabela 10 – Média, desvio padrão e quantis empíricos de razão entre variância do portfólio de variância mínima estimado calculada sob método tradicional (Q) e robusto (R). $T = 2000$, $\delta = 0,975$, $N = 2$ e 10.000 réplicas.

			Média	Desvio Padrão	2,5%	97,5%
$\epsilon = 0\%$	d = 0	Q	1,0006	0,0013	1,0000	1,0033
	d = 0	R	1,0012	0,0031	1,0000	1,0070
$\epsilon = 1\%$	d = 3	Q	1,0012	0,0026	1,0000	1,0067
	d = 3	R	1,0012	0,0034	1,0000	1,0076
	d = 4	Q	1,0022	0,0050	1,0000	1,0124
	d = 4	R	1,0012	0,0034	1,0000	1,0076
$\epsilon = 5\%$	d = 3	Q	1,0074	0,0144	1,0000	1,0415
	d = 3	R	1,0024	0,0058	1,0000	1,0140
	d = 4	Q	1,0130	0,0229	1,0000	1,0713
	d = 4	R	1,0026	0,0063	1,0000	1,0162
$\epsilon = 10\%$	d = 3	Q	1,0123	0,0217	1,0000	1,0664
	d = 3	R	1,0061	0,0129	1,0000	1,0361
	d = 4	Q	1,0182	0,0321	1,0000	1,1012
	d = 4	R	1,0081	0,0161	1,0000	1,0449

Tabela 11 – Média, desvio padrão e quantis empíricos de norma de Frobenius calculada para os portfólios estimados sob os métodos tradicional (Q) e robusto (R). $T = 2000$, $\delta = 0,975$, $N = 2$ e 10.000 réplicas.

			Média	Desvio Padrão	2,5%	97,5%
$\epsilon = 0\%$	d = 0	Q	0,0576	0,0357	0,0134	0,1458
	d = 0	R	0,0730	0,0577	0,0165	0,2129
$\epsilon = 1\%$	d = 3	Q	0,0992	0,0629	0,0254	0,2647
	d = 3	R	0,0758	0,0597	0,0169	0,2240
	d = 4	Q	0,1514	0,0973	0,0423	0,4111
	d = 4	R	0,0759	0,0603	0,0169	0,2234
$\epsilon = 5\%$	d = 3	Q	0,3831	0,1942	0,1413	0,8923
	d = 3	R	0,1325	0,0859	0,0395	0,3519
	d = 4	Q	0,5923	0,2179	0,2740	1,1366
	d = 4	R	0,1392	0,0931	0,0408	0,3822
$\epsilon = 10\%$	d = 3	Q	0,5986	0,2141	0,2775	1,1392
	d = 3	R	0,2614	0,1534	0,0875	0,6497
	d = 4	Q	0,9185	0,2150	0,5391	1,4086
	d = 4	R	0,2940	0,1812	0,0922	0,7694

Tabela 12 – Média, desvio padrão e quantis empíricos de IPV calculado para portfólios estimados sob método tradicional (Q) e robusto (R). $T = 1250$, $\delta = 0,975$, $N = 1.000$ e 100 réplicas.

				Média	Desvio Padrão	2,5%	97,5%	
$a = 0,10, b = 0,80$								
		$\epsilon = 0\%$	d = 0	Q	0.5039	0.2200	0.1765	0.8869
			d = 0	R	0.4853	0.2216	0.1779	0.8965
Isolado	$\epsilon = 25\%$	$\epsilon = 25\%$	d = 5	Q	0.8448	0.1702	0.5376	1.1746
			d = 5	R	0.5438	0.2160	0.2349	0.9044
	$\epsilon = 50\%$	$\epsilon = 50\%$	d = 5	Q	0.7876	0.1880	0.5050	1.1344
			d = 5	R	0.5420	0.2301	0.2418	0.9794
Consecutivo	$\epsilon = 25\%$	$\epsilon = 25\%$	d = 10	Q	1.1875	0.1311	0.9445	1.4176
			d = 10	R	0.5339	0.2066	0.2363	0.8901
	$\epsilon = 50\%$	$\epsilon = 50\%$	d = 10	Q	0.8350	0.1701	0.6030	1.1582
			d = 10	R	0.5672	0.2175	0.2451	0.9362
$a = 0,05, b = 0,93$								
		$\epsilon = 0\%$	d = 0	Q	0.8743	0.2884	0.4333	1.3445
			d = 0	R	0.8847	0.3217	0.4345	1.5050
Isolado	$\epsilon = 25\%$	$\epsilon = 25\%$	d = 5	Q	1.0206	0.2522	0.6141	1.4978
			d = 5	R	0.9316	0.3049	0.5103	1.4375
	$\epsilon = 50\%$	$\epsilon = 50\%$	d = 5	Q	1.0420	0.2859	0.5155	1.5066
			d = 5	R	0.9398	0.3058	0.4764	1.4660
Consecutivo	$\epsilon = 25\%$	$\epsilon = 25\%$	d = 10	Q	1.4312	0.2026	1.1805	1.8057
			d = 10	R	0.9185	0.3130	0.2941	1.4299
	$\epsilon = 50\%$	$\epsilon = 50\%$	d = 10	Q	1.2789	0.2607	0.8717	1.7066
			d = 10	R	0.9414	0.2935	0.6314	1.4453

Tabela 13 – Média, desvio padrão e quantis empíricos da razão entre as variâncias dos portfólios de variância mínima estimados pelos métodos tradicional (Q) e robusto (R). $T = 1.250$, $\delta = 0,975$, $N = 1.000$ e 100 réplicas.

				Média	Desvio Padrão	2,5%	97,5%
$a = 0,10, b = 0,80$							
	$\epsilon = 0\%$	d = 0	Q	1,7555	0,6446	1,2578	3,4756
		d = 0	R	1,6797	0,5390	1,2749	3,2099
Isolado	$\epsilon = 25\%$	d = 5	Q	3,7742	1,1594	2,3678	6,4296
		d = 5	R	2,0614	0,7471	1,3753	4,3410
	$\epsilon = 50\%$	d = 5	Q	5,2552	2,0887	2,2361	10,1500
		d = 5	R	2,1843	0,8027	1,4687	4,5148
Consecutivo	$\epsilon = 25\%$	d = 10	Q	7,6751	2,9962	3,8060	15,2671
		d = 10	R	1,9627	0,6632	1,3678	3,9072
	$\epsilon = 50\%$	d = 10	Q	13,8320	5,6137	6,5173	28,5110
		d = 10	R	2,1352	0,8110	1,4167	4,3704
$a = 0,05, b = 0,93$							
	$\epsilon = 0\%$	d = 0	Q	2.7037	0.7629	1.6416	4.8927
		d = 0	R	2.9878	1.0299	1.4716	5.9956
Isolado	$\epsilon = 25\%$	d = 5	Q	4.0713	1.1849	2.3199	7.0263
		d = 5	R	3.2088	0.9544	1.8724	5.2947
	$\epsilon = 50\%$	d = 5	Q	6.6878	2.6357	2.8456	12.7411
		d = 5	R	3.4104	1.0462	1.8758	5.8756
Consecutivo	$\epsilon = 25\%$	d = 10	Q	10.6843	4.2935	6.0352	19.4948
		d = 10	R	3.2320	1.1201	1.6920	6.1804
	$\epsilon = 50\%$	d = 10	Q	21.3514	9.1704	11.3019	45.3480
		d = 10	R	3.4174	1.0675	2.2402	6.4445

Tabela 14 – Média, desvio padrão e quantis empíricos da norma de Frobenius calculada para os portfólios estimados sob os métodos tradicional (Q) e robusto (R). $T = 1.250$, $\delta = 0,975$, $N = 1.000$ e 100 réplicas.

			Média	Desvio Padrão	2,5%	97,5%
$a = 0,10, b = 0,80$						
	$\epsilon = 0 \%$	d = 0	Q	25,5503	4,1434	20,1399
		d = 0	R	29,7969	6,8050	22,1953
Isolado	$\epsilon = 25 \%$	d = 5	Q	192,3067	28,3318	153,7176
		d = 5	R	71,2673	14,6703	54,7275
	$\epsilon = 50 \%$	d = 5	Q	292,0612	20,4955	255,9371
		d = 5	R	75,1561	13,7624	59,1474
Consecutivo	$\epsilon = 25 \%$	d = 10	Q	232,8704	35,9451	183,6651
		d = 10	R	73,2122	15,0830	55,9533
	$\epsilon = 50 \%$	d = 10	Q	341,3959	23,5282	301,2739
		d = 10	R	77,0796	14,0436	60,1294
$a = 0,05, b = 0,93$						
	$\epsilon = 0 \%$	d = 0	Q	27.0031	3.5924	21.0805
		d = 0	R	57.5881	13.3127	40.6165
Isolado	$\epsilon = 25 \%$	d = 5	Q	151.1682	13.8883	132.1502
		d = 5	R	69.3169	13.9644	49.2013
	$\epsilon = 50 \%$	d = 5	Q	236.8191	10.1846	215.4764
		d = 5	R	70.0970	14.5255	50.5516
Consecutivo	$\epsilon = 25 \%$	d = 10	Q	216.2888	24.3256	179.2365
		d = 10	R	69.6269	14.5984	50.2742
	$\epsilon = 50 \%$	d = 10	Q	331.3525	17.0167	294.0820
		d = 10	R	70.8404	13.0686	50.5001

APÊNDICE B – Gráficos Complementares dos Capítulos 5, 6 e 7

Neste apêndice são apresentadas algumas figuras com resultados das simulações e da aplicação. Nas Figuras 19, 20 e 21 alguns resultados das Seções 5.4 e 5.5, em específico viés e REQM dos estimadores nas etapas GARCH e cDCC. Na Figuras 22, 23, 24 e 25 alguns resultados da Seção 6.2.

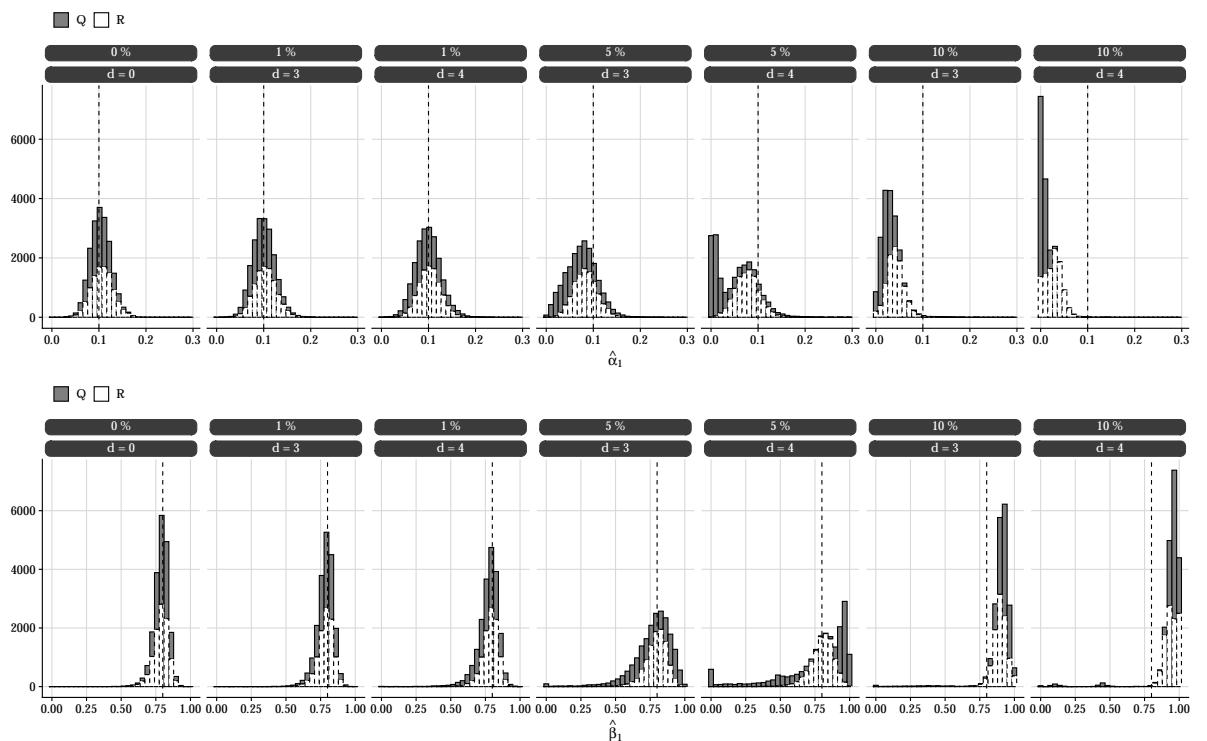


Figura 19 – Histogramas para estimativas de $\eta_1 = (0, 10, 0, 10, 0, 80)$ do modelo GARCH(1,1) obtidas por $\hat{\eta}_{q,1}$ (Q) e $\hat{\eta}_{r,1}$ (R). A linha tracejada indica valor verdadeiro do respectivo parâmetro. $T = 2.000$, $\delta = 0,975$ e 10.000 réplicas.

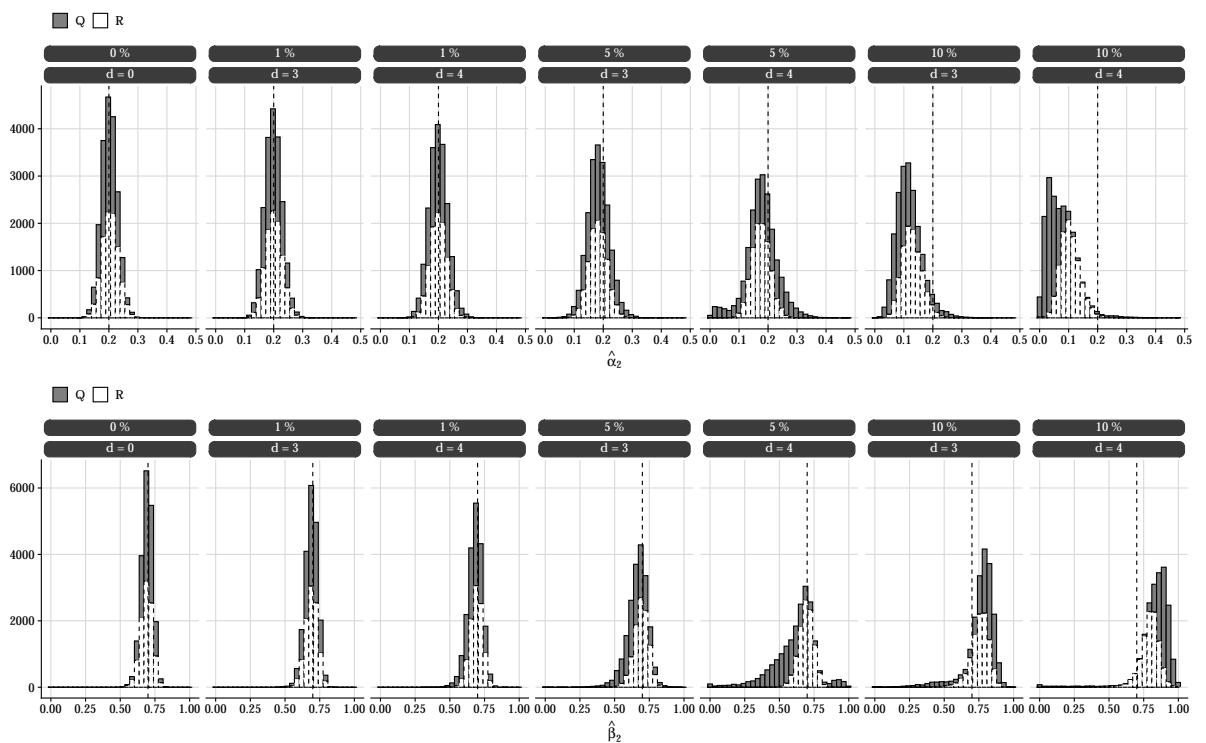


Figura 20 – Histogramas para estimativas de $\boldsymbol{\eta}_2 = (0, 10, 0, 20, 0, 70)$ do modelo GARCH(1,1) obtidas por $\hat{\boldsymbol{\eta}}_{q,2}$ (Q) e $\hat{\boldsymbol{\eta}}_{r,2}$ (R). A Linha tracejada indica valor verdadeiro do respectivo parâmetro. Linha tracejada indica real valor do respectivo parâmetro. $T = 2.000$, $\delta = 0,975$ e 10.000 réplicas.

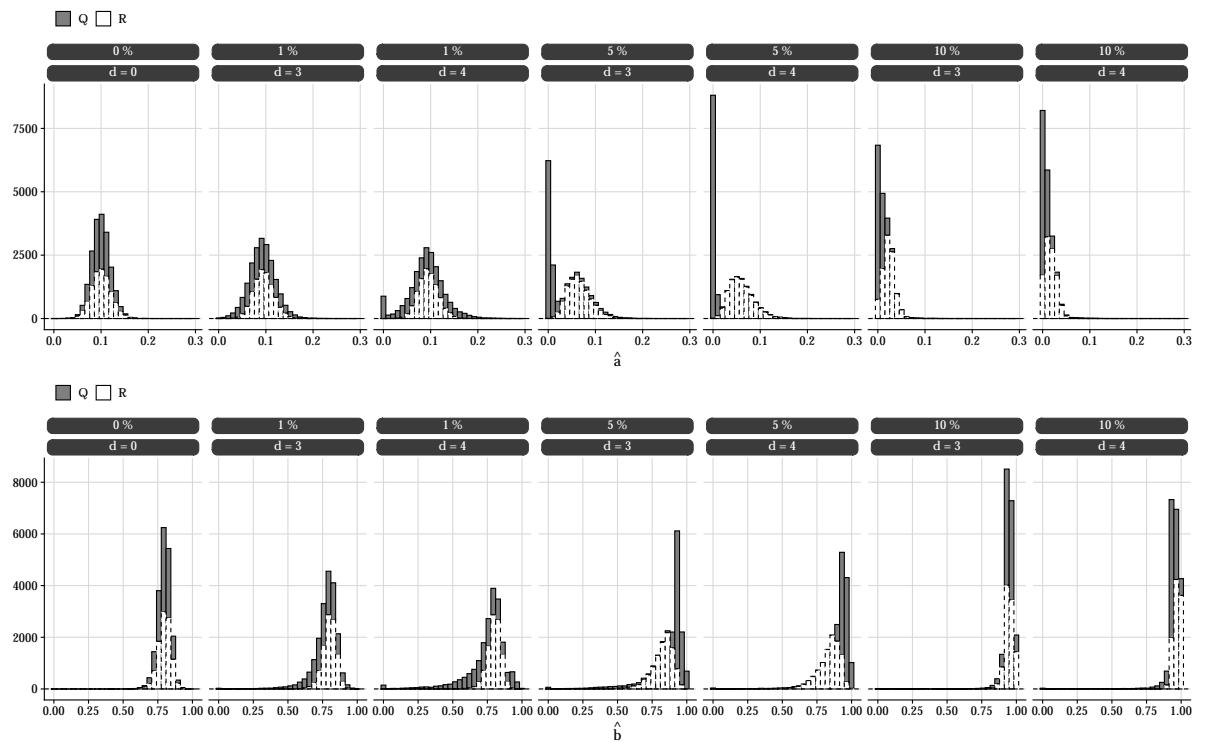


Figura 21 – Histogramas para estimativas de $\phi = (0, 10, 0, 80)$ de modelo cDCC(1, 1) obtidas por $\hat{\phi}_q$ e $\hat{\phi}_r$. No eixo horizontal, (Q) e (R) são legendas para estimativas obtidas utilizando $\hat{\phi}_q$ e $\hat{\phi}_r$, respectivamente. No eixo vertical, estimativas para parâmetros a e b . Linha tracejada indica real valor do respectivo parâmetro. $T = 2.000$, $\delta = 0, 975$, $N = 2$ e 10.000 réplicas.

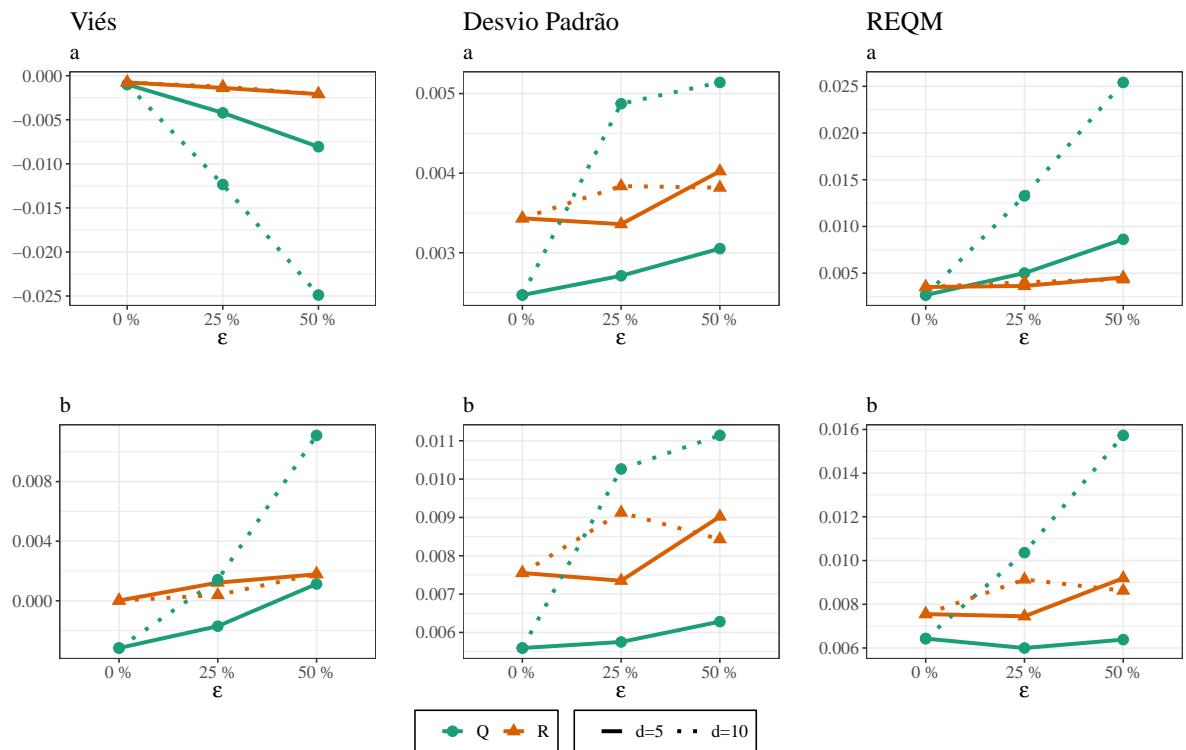


Figura 22 – Viés, REQM e erro padrão (colunas) obtidos por $\hat{\phi}_q$ e $\hat{\phi}_r$ para os parâmetros a e b (linha). Contaminação com saltos isolados, $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Resultados para $\phi = (0, 10, 0, 80)$. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

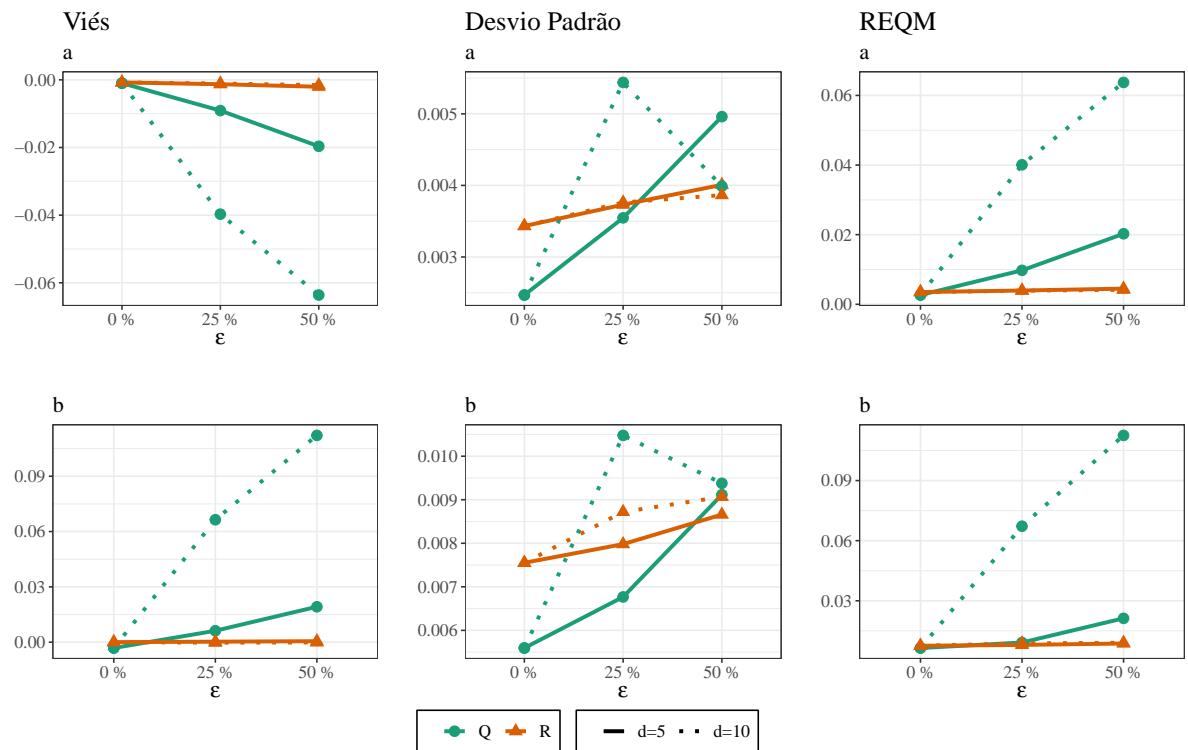


Figura 23 – Viés, REQM e erro padrão (colunas) obtidos por $\hat{\phi}_q$ e $\hat{\phi}_r$ para parâmetros a e b (linha). Contaminação com saltos consecutivos, $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Resultados para $\phi = (0, 10, 0, 80)$. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

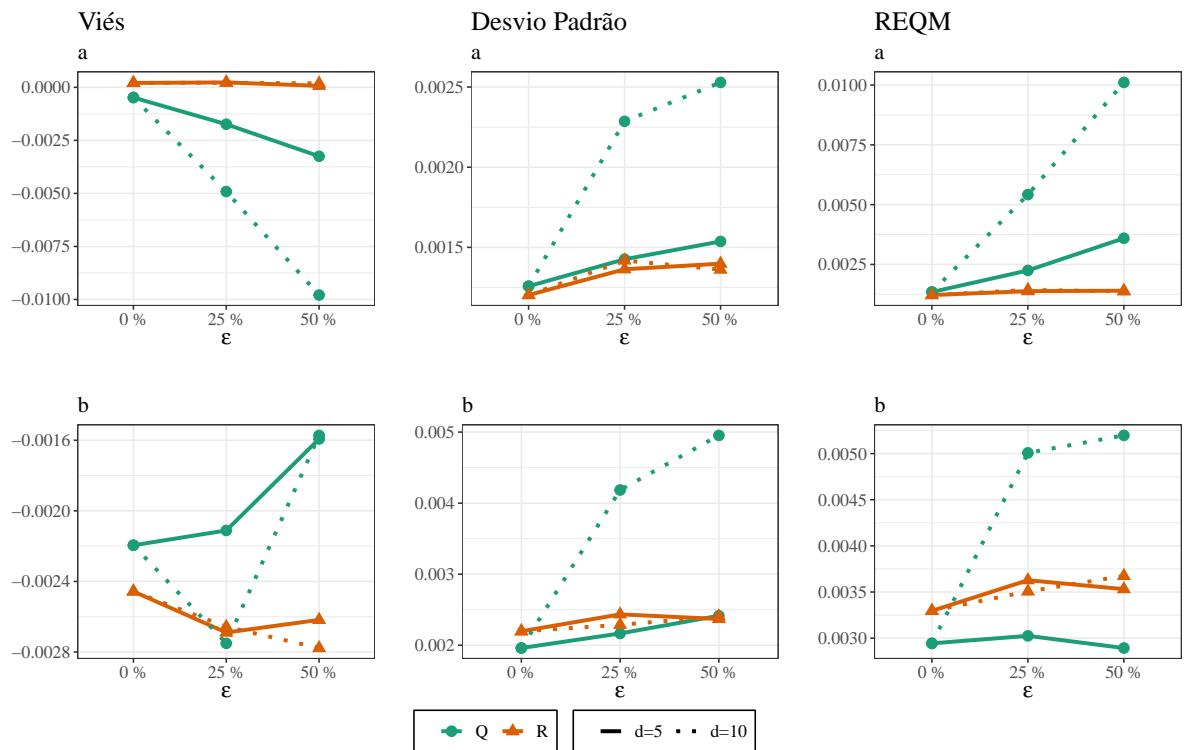


Figura 24 – Viés, REQM e erro padrão (colunas) obtidos por $\hat{\phi}_q$ e $\hat{\phi}_r$ para parâmetros a e b (linha). Contaminação com saltos isolados, $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Resultados para $\phi = (0,05,0,93)$. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

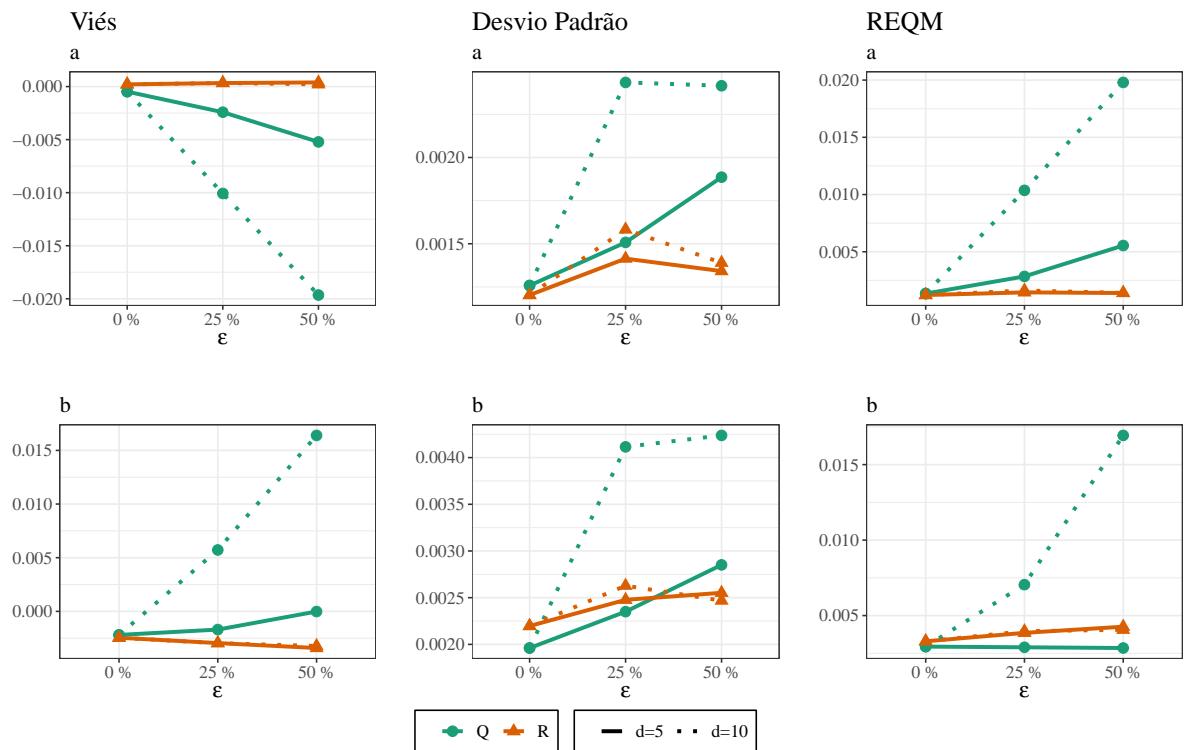


Figura 25 – Viés, REQM e erro padrão (colunas) obtidos por $\hat{\phi}_q$ e $\hat{\phi}_r$ para parâmetros a e b (linha). Contaminação com saltos consecutivos, $T = 1.250$, $\delta = 0,975$ e 100 réplicas. Resultados para $\phi = (0,05,0,93)$. O percentual de tempos com contaminação é denotado por ϵ . Os saltos são de magnitude d vezes o desvio padrão condicional.

APÊNDICE C – Teste de Igualdade de Variâncias de Portfólios

Neste apêndice são apresentados os teste de igualdade de variância dos retornos de dois métodos de seleção de portfólios sugeridos por (LEDOIT; WOLF, 2011). O primeiro é chamado de teste HAC e segundo um teste *bootstrap*. Embora os resultados da aplicação do primeiro não tenha sido apresentado no Capítulo 7, ele é apresentado no apêndice porque poderiam ser utilizados quando o número de observações é grande. Considere $(r_{1,1}, r_{2,1})', \dots, (r_{1,k}, r_{2,k})'$ como um total de k pares de retornos de portfólio. Considerando que esta série bivariada é estacionária, sua distribuição não muda com o tempo. Seu vetor de média e matriz de variâncias e covariâncias são dadas por:

$$(r_{1,i}, r_{2,i})' \sim (\mu, \Sigma), \quad (\text{C.1})$$

com

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad \text{e} \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{bmatrix}, \quad (\text{C.2})$$

Usualmente, para comparação de variâncias utiliza-se o clássico teste-F cujas hipóteses são dadas por:

$$H_0 : \vartheta = 1 \quad \text{ou} \quad H_1 : \vartheta \neq 1,$$

em que

$$\vartheta = \sigma_1^2 / \sigma_2^2$$

com a seguinte estatística de teste

$$F = \hat{\sigma}_1^2 / \hat{\sigma}_2^2,$$

sendo $\hat{\sigma}_1^2$ e $\hat{\sigma}_2^2$ as respectivas quantidades amostrais.

O teste F rejeitará a hipótese nula a um nível de significância α se e somente se $F < F_{\alpha/2, T-1, T-1}$ ou $F > F_{1-\alpha/2, T-1, T-1}$, considerando ainda que os dados são oriundos de uma distribuição normal bivariada com $\sigma_{1,2}^2 = 0$ e independente ao longo do tempo. Esta condição em conjunto com retornos com distribuição de caudas mais pesadas que a normal e/ou dependentes ao longo do tempo inviabilizam a utilização do teste F. Naturalmente, pelo menos um desses três impedimentos estão presentes em séries financeiras. Motivado

por isso, [Ledoit e Wolf \(2011\)](#) propõem uma transformação logarítmica nas hipóteses. As hipóteses passam a ser

$$H_0 : \Delta = 0 \quad \text{ou} \quad H_1 : \Delta \neq 0,$$

para

$$\Delta = \log(\vartheta) = \log(\sigma_1^2) - \log(\sigma_2^2).$$

em que é usada $\log(\vartheta)$ para estabilizar a variância, veja [Efron e Tibshirani \(1993\)](#).

Defina o vetor aleatório $w_j = (r_{1j}, r_{2j}, r_{1j}^2, r_{2j}^2)'$, que tem como esperança $v = E(w_j) = (\mu_1, \mu_2, \gamma_1, \gamma_2)'$. Portanto, considerando que a variância dos retornos são finitas, temos que os estimadores;

$$\hat{\mu}_i = n^{-1} \sum_{j=1}^n r_{1j} \tag{C.3}$$

$$\hat{\gamma}_i = n^{-1} \sum_{j=1}^n r_{1j}^2, \tag{C.4}$$

para $i = 1, 2$ são estimadores não viciados de μ_i e γ_i , respectivamente. Sob condições não muito fortes, por exemplo, que $E(r_{1,1}^{4+\delta}), E(r_{1,2}^{4+\delta}) < \infty$, para um δ pequeno constante positivo e condições de mistura (*mixing*) apropriadas (veja, por exemplo, [Andrews \(1991\)](#)), temos que:

$$\Delta = f(v) \quad \text{e} \quad \hat{\Delta} = f(\hat{v})$$

com

$$\begin{aligned} f(\mu_1, \mu_2, \gamma_1, \gamma_2) &= \log(\gamma_1 - \mu_1^2) - \log(\gamma_2 - \mu_2^2), \\ v = (\mu_1, \mu_2, \gamma_1, \gamma_2)' &\quad \text{e} \quad \hat{v} = (\hat{\mu}_1, \hat{\mu}_2, \hat{\gamma}_1, \hat{\gamma}_2)', \end{aligned}$$

assumindo que

$$\sqrt{T}(\hat{v} - v) \xrightarrow{d} \mathcal{N}(0, \Psi)$$

sendo Ψ uma matriz positiva semi-definida. Pelo método delta

$$\sqrt{T}(\Delta - \hat{\Delta}) \xrightarrow{d} \mathcal{N}(0; \nabla' f(v) \Psi \nabla' f(v)) \tag{C.5}$$

com

$$\nabla' f(\mu_1, \mu_2, \gamma_1, \gamma_2) = \left(-\frac{2\mu_1}{\gamma_1 - \mu_1^2}, \frac{2\mu_1}{\gamma_2 - \mu_2^2}, \frac{1}{\gamma_1 - \mu_1^2}, -\frac{1}{\gamma_2 - \mu_2^2} \right),$$

é possível obter um estimador consistente para Ψ utilizando métodos *kernel* heteroscedásticos e auto correlacionados consistentes (*heteroskedasticity and autocorrelation robust*, HAC), para mais detalhes veja [Ledoit e Wolf \(2008\)](#).

Dado um estimador consistente para Ψ ($\hat{\Psi}$), o erro padrão para $\hat{\Delta}$ é dado por

$$s(\hat{\Delta}) = \sqrt{\frac{\nabla' f(\hat{v}) \hat{\Psi} \nabla f(\hat{v})}{T}} \quad (\text{C.6})$$

Dessa forma, com $s(\hat{\Delta})$ combinado com a normalidade assintótica em [C.5](#) é possível calcular o valor-p para hipóteses nula

$$\hat{p} = 2\Phi\left(-\frac{\hat{\Delta}}{s(\hat{\Delta})}\right), \quad (\text{C.7})$$

em que $\Phi(\cdot)$ denota a função de distribuição acumulada de uma distribuição normal padrão.

No entanto, uma consideração acerca deste teste é: HAC é muito liberal o que implica em testes de hipóteses que rejeitam a hipótese nula com frequência alta para o nível de significância estabelecido. Devido a isso, [Ledoit e Wolf \(2008\)](#) propõem um teste baseado na inversão de intervalos de confiança *bootstrap* como uma alternativa mais conservadora se comparada a métodos HAC descritos anteriormente. Basicamente, dada a estrutura de dependência que existe entre os pares, as réplicas são geradas utilizando *bootstrap* em blocos circulares. Considerando M replicações, para cada uma estima-se $\hat{\Psi}^*$ dado por

$$\hat{\Psi}_m^* = \frac{1}{l} \sum_{j=1}^l \zeta_{j,m} \zeta'_{j,m}, \quad m = 1, \dots, M \quad (\text{C.8})$$

em que

$$\begin{aligned} \zeta_{j,m} &= \frac{1}{\sqrt{b}} \sum_{t=1}^b s_{(j-1)b+t,m}^* \quad t = 1, \dots, l, \\ s_{t,m}^* &= (r_{1,t,m}^* - \hat{\mu}_{1,m}^*, r_{2,t,m}^* - \hat{\mu}_{2,m}^*, r_{1,t,m}^{*2} - \gamma_{1,m}^*, r_{2,t,m}^{*2} - \gamma_{2,m}^*), \end{aligned}$$

sendo $l = [T/b]$; b o tamanho do bloco e $s_{t,m}^*$ calculado a partir da m -ésima replica *bootstrap*. Mais detalhes sobre o procedimento proposto para a escolha de b podem ser consultados em [Ledoit e Wolf \(2008\)](#). Com isso, a estatística de teste studentizada centralizada é dada por

$$\tilde{d}^{*,m} = \frac{|\hat{\Delta}^{*,m} - \hat{\Delta}|}{s(\hat{\Delta}^{*,m})}$$

com o erro padrão $s(\hat{\Delta}^{*,m})$ calculado usando [\(C.6\)](#) com a respectiva réplica. O valor-p é dado por

$$PV = \frac{\#\{\tilde{d}^{*,m} \geq d\} + 1}{M + 1}. \quad (\text{C.9})$$

em que d é a estatística studentizada original dada por $d = |\hat{\Delta}|/s(\hat{\Delta})$.