

Modelagem GARCH multivariada

Mauricio Alejandro Mazo Lopera

DISSERTAÇÃO
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
MESTRE EM CIÊNCIAS

Programa: Estatística

Orientadora: Profa. Dra. Chang Chiann

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro do CNPq

São Paulo, 6 de abril de 2011

Modelagem GARCH multivariada

Esta versão definitiva da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa realizada por Mauricio Alejandro Mazo Lopera em 24/03/2011.

Comissão Julgadora:

- Profa. Dra. Chang Chiann (Presidenta) - IME-USP
- Profa. Dra. Clélia Maria de Castro Toloi - IME-USP
- Prof. Dr. João Ricardo Sato - UFABC

Agradecimentos

À Professora Chang Chiann por ter me dado a honra de ser minha orientadora e por ter me ajudado sempre.

À minha família Genia, Libardo e Ema que, sem eles esse trabalho nem teria começado.

À Carolina Santafe pela paciência que teve nesses dois anos, estando sempre comigo.

Aos colegas e amigos do Departamento de Estatística do IME e ao CNPq, pelo apoio financeiro, por possibilitarem a realização desse trabalho.

Resumo

A maioria dos ativos financeiros nos mercados mundiais apresentam variância condicional evoluindo no tempo. Para modelar tal variância, Engle (1982) e posteriormente Bollerslev (1986), propuseram os modelos ARCH ("Autoregressive Conditional Heterocedasticity") e a generalização destes, conhecidos como os modelos GARCH ("Generalized Autoregressive Conditional Heterocedasticity"). Estes modelos conseguem ajustar individualmente a variância condicional de uma série temporal de forma autoregressiva, no entanto em algumas situações também é importante descrever as relações entre várias séries.

Para modelar estas relações, Engle e Kroner (1995) propuseram os modelos GARCH multivariados, os quais descrevem a matriz de covariâncias condicional de um grupo de séries de forma análoga ao caso GARCH univariado. Os modelos que propuseram foram os modelos Média Movel Exponencialmente Ponderada, VEC, Diagonal VEC e BEKK ("Baba-Engle-Kraft-Kroner"). Estes modelos ajustam de forma direta a matriz de covariâncias condicional usando matrizes de parâmetros de forma autoregressiva, sob certas condições que garantem as matrizes de covariâncias condicionais sejam positivas semidefinidas e o processo seja estacionário na covariância.

Nos modelos diretos o número de parâmetros a serem estimados é muito grande, e para resolver este problema foi necessário desenvolver outros métodos que ajustam a volatilidade condicional multivariada utilizando ajustes GARCH univariados, como por exemplo, os modelos de Correlação Condicional Constante e de Componentes Principais.

Algumas simulações para avaliar os principais modelos serão consideradas, além de aplicações com dados reais do mercado financeiro do Brasil, usando o software S-PLUS.

Palavras-chave: Modelagem GARCH, Análise Multivariada, Componentes Principais, Modelagem GARCH Multivariada.

Abstract

Most of financial assets in the world markets have conditional variance evolved over time. To model this conditional volatility, Engle (1982) and then Bollerslev (1986) proposed the ARCH ("Autoregressive Conditional Heterocedasticity") and the generalization of these, known as GARCH ("Generalized Autoregressive Conditional Heterocedasticity"). These models can individually adjust the conditional variance of a time series in an autoregressive form, but in some situations it is also important to describe the relationships between various series.

To model these relationships, Engle e Kroner (1995) proposed the multivariate GARCH models, which describe the conditional covariance matrix of a group of series analogous to univariate GARCH case. The models proposed were Exponentially Weighted Moving Average, VEC, BEKK and VEC Diagonal ("Baba-Engle-Kraft-Kroner"). These models fit directly the conditional covariance matrix using arrays of parameters in an autoregressive form, under certain conditions that guarantee the conditional covariance matrices to be positive semidefinite and the process to be covariance stationary.

In direct models the number of parameters to be estimated is very large, and to solve this problem it was necessary to develop other methods that adjust the multivariate conditional volatility using univariate GARCH models, such as the Constant Conditional Correlation Models and Principal Components.

Some simulations to evaluate the main models will be considered, as well as applications with real data from Brazil's financial market, using the software S-PLUS.

Keywords: GARCH Model, Multivariate Analysis, Principal Component, Multivariate GARCH Model.

Sumário

Notações	ix
1 Introdução	1
1.1 Objetivos	3
1.2 Organização do Trabalho	4
2 Conceitos Básicos	5
2.1 Traço da Matriz	5
2.2 Operador VEC	5
2.3 Formas Quadráticas e Matrizes Positivas Definidas	7
2.4 Decomposição de Matrizes Simétricas	9
2.5 Propriedades da Esperança Condicional	9
3 Modelos GARCH Multivariados	11
3.1 Modelo GARCH Univariado	12
3.2 Modelos GARCH Multivariados Diretos	12
3.2.1 Média Móvel Exponencialmente Ponderada	13
3.2.2 Modelo VEC	15
3.2.3 Modelo VEC Diagonal	19
3.2.4 Modelo Matriz-Diagonal	21
3.2.5 Modelo BEKK	24
3.2.6 Estacionariedade na Covariância	27
3.2.7 Estimação dos Modelos	29
3.2.8 Testes para os modelos GARCH multivariados	33
3.2.9 Previsão dos modelos GARCH multivariados	38
3.3 Modelos GARCH Multivariados Indiretos	41
3.3.1 Modelo de Correlação Condicional Constante	42
3.3.2 Modelo GARCH Ortogonal	44

4	Simulações	53
4.1	Modelo EWMA	53
4.2	Modelo DVEC(1,1)	77
4.3	Modelo BEKK(1,1,1)	92
4.4	Modelo BEKK(1,1,1) com distribuição t-student	104
5	Aplicações	111
5.1	Ações Materiais de Construção	112
5.2	Ações Construção Civil	123
6	Conclusões e Trabalhos Futuros	135
A	Provas dos Resultados	137
B	Comandos no S-PLUS	149
	Referências Bibliográficas	153

Notações

- Os vetores serão denotados com letras minúsculas em negrito: $\mathbf{a}, \mathbf{b}, \mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\eta}, \dots$
- As matrizes serão denotadas com letras maiúsculas em negrito: $\mathbf{A}, \mathbf{B}, \boldsymbol{\Sigma}, \mathbf{G}, \dots$
- Os escalares serão considerados com letras minúsculas: a, b, λ, \dots
- k : número de séries consideradas no vetor de retornos.
- $\mathbf{a}^{(i)}$: i -ésimo elemento do vetor \mathbf{a} .
- $\mathbf{A}^{(ij)}$: elemento na linha i e na coluna j da matriz \mathbf{A} .
- $tr(\cdot)$: traço da matriz.
- $\|\cdot\|$: norma do vetor.
- $|\cdot|$: norma da matriz.
- $vec(\cdot)$: operador vec .
- \otimes : produto Kronecker.
- \odot : produto Hadamard.
- \mathbf{A}^{-1} : inversa da matriz \mathbf{A} .
- \mathbf{A}' : transposta da matriz \mathbf{A} .
- \mathcal{F}_{t-1} : informações passadas até o tempo $t - 1$.
- $E(\cdot | \mathcal{F}_{t-1})$: esperança condicional dada as informações passadas.
- $E_{t-1}(\cdot) := E(\cdot | \mathcal{F}_{t-1})$.
- \mathbf{I}_k : matriz identidade ($k \times k$).
- $\mathbf{0}_{k \times l}$: matriz de zeros de dimensões ($k \times l$).

x NOTAÇÕES

- $\mathbf{N}_k(\cdot, \cdot)$: distribuição normal multivariada de um vetor ($k \times 1$).
- $N(\cdot, \cdot)$: distribuição normal univariada.

Capítulo 1

Introdução

Na análise de séries temporais financeiras a modelagem da volatilidade desempenha um papel fundamental na previsão da mesma e no cálculo do valor em risco ("VaR") de carteiras. A maioria dos ativos financeiros apresentam variância condicional evoluindo no tempo, tornando-se necessário ajustar modelos para tal variância, como por exemplo os modelos ARCH ("Autoregressive Conditional Heterocedasticity") ou a generalização destes, conhecidos como os modelos GARCH ("Generalized Autoregressive Conditional Heterocedasticity") (Morettin (2002) ou Zivot e Wang (2006)).

Além da volatilidade, em algumas situações é necessário modelar também as covariâncias condicionais num conjunto de ativos e para isto, foram desenvolvidos os modelos GARCH multivariados (Engle e Kroner (1995)).

A Figura 1.1 apresenta os log-retornos semanais do conjunto de ativos financeiros do mercado Brasileiro VALE, PETROBRAS e GERDAU de Janeiro de 1998 até Abril de 2010, num total de 570 observações.

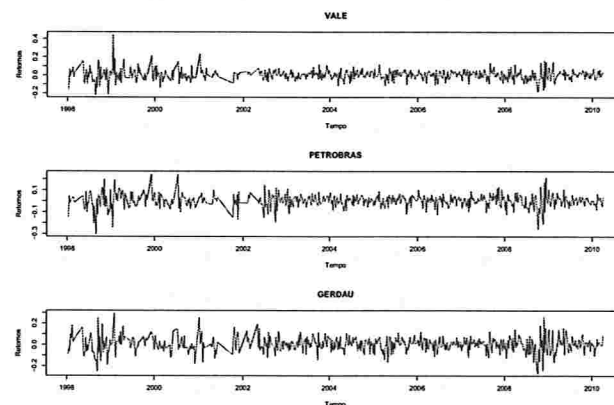


Figura 1.1: Retornos VALE, PETRO. e GERDAU.

Observa-se que ao longo do tempo existem conglomerados de volatilidade alta nas séries, sendo que os períodos que possuem a maior variância estão nos anos 1999 e 2009.

Modelos GARCH univariados podem ser ajustados separadamente para cada série de log-retornos, mas com isso não é possível identificar e descrever as correlações condicionais que existem entre elas ao longo do tempo. Portanto, para considerar estas correlações, o mais adequado é ajustar um modelo GARCH multivariado.

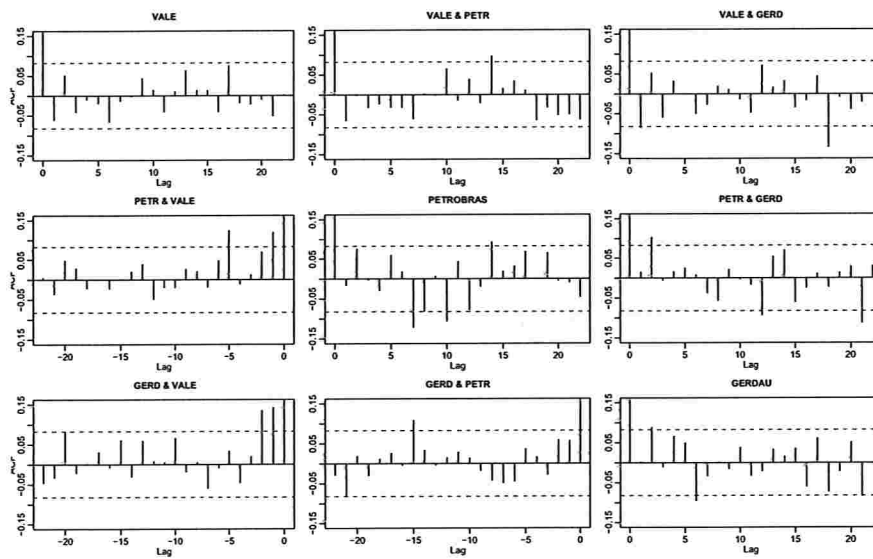


Figura 1.2: FAC e FAC cruzadas dos retornos.

A Figura 1.2 apresenta as FAC ("Funções de Autocorrelação") e FAC cruzada dos retornos. Nota-se que há poucas autocorrelações significativas até o lag 10. Já para o quadrado dos retornos, a Figura 1.3 apresenta as FAC e FAC cruzadas deles. Observa-se que, em geral, existem autocorrelações e autocorrelações cruzadas significativas entre eles, confirmando a existência de uma dependência quadrática nos log-retornos. Segundo Engle e Kroner (1995), os modelos GARCH multivariados conseguem descrever de forma adequada tal dependência.

Existem duas formas de generalizar os modelos GARCH univariados ao caso multivariado. A primeira consiste numa modelagem direta da matriz de covariâncias condicionais considerando matrizes de parâmetros. Neste grupo estão os modelos EWMA ("Exponentially Weighted Covariance Estimate"), VEC, DVEC ("Diagonal VEC") e BEKK ("Baba-Engle-Kraft-Kroner"). A segunda forma consiste numa modelagem indireta da matriz de covariâncias condicionais utilizando ajustes GARCH univariados. Os principais modelos neste grupo são os modelos de Correlação Condicional Constante ("CCC") e o modelo de Componentes

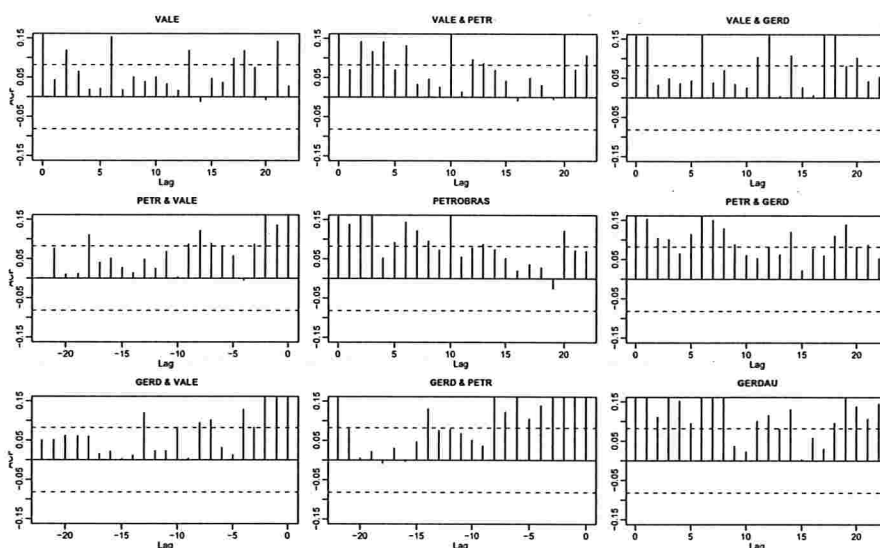


Figura 1.3: *FAC e FAC cruzadas dos retornos ao quadrado.*

Principais. As propriedades destes modelos dependem diretamente das matrizes de parâmetros e da distribuição dos vetores de erros, que nesta dissertação vão ser a distribuição normal e a distribuição t-student multivariadas.

1.1 Objetivos

O objetivo principal deste trabalho é apresentar os diferentes modelos GARCH multivariados para a matriz de covariâncias condicionais, e discutir suas vantagens e desvantagens ao longo do texto.

Os objetivos secundários são:

- Discutir a propriedade de estacionariedade na covariância e os métodos de estimação destes modelos como uma generalização dos modelos GARCH univariados.
- Apresentar os diferentes testes para a avaliação dos modelos ajustados, considerando principalmente o teste multivariado de Portmanteau.
- Avaliar os diferentes modelos a partir de simulações feitas no software S-PLUS.
- Comparar as generalizações diretas e indiretas do modelo GARCH univariado.
- Aplicar os diferentes modelos GARCH multivariados num conjunto de séries reais, com o intuito de comparar e obter o melhor modelo para previsões.

1.2 Organização do Trabalho

No Capítulo 2 apresentam-se conceitos básicos de matrizes, operadores vetoriais e algumas propriedades da esperança condicional. Os modelos GARCH multivariados, junto com as principais condições de estacionariedade, estimação e diagnóstico dos ajustes, são apresentados no Capítulo 3. O Capítulo 4 apresenta algumas simulações dos principais modelos usando o software S-PLUS. O Capítulo 5 apresenta uma aplicação com dados reais do mercado financeiro Brasileiro. As conclusões estão no Capítulo 6. No Apêndice A apresentam-se as provas dos principais resultados e os comandos e algoritmos usados no software S-PLUS estão no Apêndice B.

Capítulo 2

Conceitos Básicos

Neste capítulo apresentam-se alguns conceitos básicos que serão utilizados no decorrer desta dissertação. As provas das propriedades matriciais não serão consideradas, uma vez que essas podem ser encontradas nos livros de álgebra linear. Algumas referências são Wadsworth (1969) e Searle (1982).

2.1 Traço da Matriz

O **traço** de uma matriz quadrática ($m \times m$), $\mathbf{A} = [a_{ij}]$, é a soma dos elementos da diagonal,

$$\text{tr}(\mathbf{A}) := a_{11} + \dots + a_{mm}.$$

Por exemplo,

$$\text{tr} \left(\begin{bmatrix} 5 & 3 \\ 2 & 7 \end{bmatrix} \right) = 5 + 7 = 12.$$

Propriedade 2.1. *Sejam \mathbf{A} e \mathbf{B} matrizes de dimensão ($m \times m$), $\lambda_1, \dots, \lambda_m$ são os autovalores de \mathbf{A} , tem-se que*

1. $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$.
2. $\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}')$.
3. Se \mathbf{C} é ($m \times n$) e \mathbf{D} é ($n \times m$), $\text{tr}(\mathbf{CD}) = \text{tr}(\mathbf{DC})$.
4. $\text{tr}(\mathbf{A}) = \lambda_1 + \dots + \lambda_m$.

2.2 Operador VEC

Seja $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ uma matriz ($m \times n$), em que \mathbf{a}_i é o i -ésimo vetor coluna ($m \times 1$), para $i = 1, \dots, n$. O operador *vec* transforma \mathbf{A} num vetor ($mn \times 1$) empilhando as colunas,

isto é,

$$\text{vec}(\mathbf{A}) = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{bmatrix}.$$

Por exemplo, para

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & -2 \\ 8 & -1 & 4 \end{pmatrix},$$

tem-se que

$$\text{vec}(\mathbf{A}) = \begin{bmatrix} 1 \\ 8 \\ 3 \\ -1 \\ -2 \\ 4 \end{bmatrix}.$$

Definição 2.2.1. Sejam $\mathbf{A} = [a_{ij}]$ e $\mathbf{B} = [b_{kl}]$ matrizes $(m \times n)$ e $(p \times q)$, respectivamente. A matriz $(mp \times nq)$

$$\mathbf{A} \otimes \mathbf{B} := \begin{bmatrix} a_{11}\mathbf{B} & \dots & a_{1n}\mathbf{B} \\ \vdots & & \vdots \\ a_{m1}\mathbf{B} & \dots & a_{mn}\mathbf{B} \end{bmatrix},$$

é o **produto Kronecker** de \mathbf{A} e \mathbf{B} .

Propriedade 2.2. Sejam \mathbf{A} , \mathbf{B} e \mathbf{C} matrizes com as dimensões adequadas, tem-se que

1. $\mathbf{A} \otimes \mathbf{B} \neq \mathbf{B} \otimes \mathbf{A}$, em geral.
2. $(\mathbf{A} \otimes \mathbf{B})' = \mathbf{A}' \otimes \mathbf{B}'$.
3. $\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = \mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{C}$.
4. $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$.
5. Se \mathbf{A} e \mathbf{B} são invertíveis, então $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$.
6. $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr}(\mathbf{A})\text{tr}(\mathbf{B})$.
7. $\text{vec}(\mathbf{A} + \mathbf{B}) = \text{vec}(\mathbf{A}) + \text{vec}(\mathbf{B})$.
8. $\text{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\text{vec}(\mathbf{B})$.

$$9. \text{vec}(\mathbf{AB}) = (\mathbf{I} \otimes \mathbf{A})\text{vec}(\mathbf{B}) = (\mathbf{B}' \otimes \mathbf{I})\text{vec}(\mathbf{A}).$$

$$10. \text{vec}(\mathbf{ABC}) = (\mathbf{I} \otimes \mathbf{AB})\text{vec}(\mathbf{C}) = (\mathbf{C}'\mathbf{B}' \otimes \mathbf{I})\text{vec}(\mathbf{A}).$$

$$11. \text{vec}(\mathbf{B}')'\text{vec}(\mathbf{A}) = \text{tr}(\mathbf{BA}) = \text{tr}(\mathbf{AB}) = \text{vec}(\mathbf{A}')'\text{vec}(\mathbf{B}).$$

Definição 2.2.2. O *Produto Hadamard* entre duas matrizes $\mathbf{A} = [a_{ij}]$ e $\mathbf{B} = [b_{ij}]$ da mesma ordem ($m \times n$), foi definido pelo Francés Jacques Hadamard (1865-1963) como,

$$\mathbf{A} \odot \mathbf{B} := [a_{ij}] \odot [b_{ij}] = [a_{ij} \times b_{ij}],$$

isto é, o produto Hadamard define uma matriz de dimensão ($m \times n$), obtida multiplicando componente a componente as matrizes \mathbf{A} e \mathbf{B} .

Propriedade 2.3. Sejam \mathbf{A} e \mathbf{B} matrizes de dimensão ($m \times n$), tem-se a seguinte relação entre o operador vec e o produto Hadamard;

$$\text{vec}(\mathbf{A} \odot \mathbf{B}) = \text{diag}\{\text{vec}(\mathbf{A})\}\text{vec}(\mathbf{B}),$$

em que $\text{diag}\{\mathbf{a}\}$ denota a matriz diagonal.

Propriedade 2.4. Seja \mathbf{A} uma matriz de dimensão ($n \times n$) positiva semi-definida e \mathbf{b} um vetor ($n \times 1$) diferente de $\mathbf{0}$, então

$$\mathbf{A} \odot \mathbf{b}\mathbf{b}' = \text{diag}[\mathbf{b}] \mathbf{A} \text{diag}[\mathbf{b}],$$

em que $\text{diag}[\mathbf{b}]$ é uma matriz diagonal.

2.3 Formas Quadráticas e Matrizes Positivas Definidas

Seja \mathbf{A} uma matriz simétrica ($m \times m$) e \mathbf{x} um vetor ($m \times 1$). A função $\mathbf{x}'\mathbf{A}\mathbf{x}$ é chamada **forma quadrática** em \mathbf{x} . A matriz simétrica \mathbf{A} ou a correspondente forma quadrática é

- *positiva definida* se $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$, para todos os vetores m -dimensionais $\mathbf{x} \neq \mathbf{0}$;
- *positiva semidefinida* se $\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$, para todos os vetores m -dimensionais \mathbf{x} ;
- *negativa definida* se $\mathbf{x}'\mathbf{A}\mathbf{x} < 0$, para todos os vetores m -dimensionais $\mathbf{x} \neq \mathbf{0}$;
- *negativa semidefinida* se $\mathbf{x}'\mathbf{A}\mathbf{x} \leq 0$, para todos os vetores m -dimensionais \mathbf{x} ;

- *indefinida* se $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ para alguns \mathbf{x} , e $\mathbf{x}'\mathbf{A}\mathbf{x} < 0$ para outros \mathbf{x} .

Propriedade 2.5. *Seja \mathbf{A} uma matriz $(m \times m)$ simétrica, tem-se que*

1. *\mathbf{A} é negativa definida (semidefinida) se e somente se $-\mathbf{A}$ é positiva definida (semidefinida).*
2. *Todos os autovalores de uma matriz positiva (negativa) definida são maiores (menores) que zero.*
3. *Uma matriz diagonal é positiva (negativa) definida se e somente se os elementos na diagonal são positivos (negativos).*
4. *Se \mathbf{A} é positiva definida e \mathbf{B} é uma matriz $(m \times n)$, então $\mathbf{B}'\mathbf{A}\mathbf{B}$ é positiva semidefinida.*
5. *Se \mathbf{A} é positiva definida, então \mathbf{A}^{-1} é positiva definida.*

Também é possível definir propriedades para que a forma quadrática de uma matriz não simétrica \mathbf{A} seja positiva (negativa) definida.

Propriedade 2.6. *Seja \mathbf{A} uma matriz não simétrica. Uma condição suficiente para que a forma quadrática $\mathbf{x}'\mathbf{A}\mathbf{x}$ seja maior (menor) que zero é que $(\mathbf{A} + \mathbf{A}')/2$ seja positiva (negativa) definida. Além disso, tem-se que*

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{x}' \left(\frac{\mathbf{A} + \mathbf{A}'}{2} \right) \mathbf{x}.$$

Definição 2.3.1. Decomposição de Cholesky. *Se \mathbf{A} é uma matriz positiva definida $(m \times m)$ então existe uma matriz triangular inferior (superior), \mathbf{P} , com elementos positivos na diagonal principal, tal que*

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}'^{-1} = \mathbf{I} \quad \text{ou} \quad \mathbf{A} = \mathbf{P}\mathbf{P}'.$$

2.4 Decomposição de Matrizes Simétricas

Se \mathbf{A} é uma matriz simétrica ($m \times m$), então existe uma matriz ortogonal \mathbf{P} tal que

$$\mathbf{P}'\mathbf{A}\mathbf{P} = \mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{bmatrix} \quad \text{ou} \quad \mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}',$$

em que os λ_i 's são os autovalores de \mathbf{A} e as colunas de \mathbf{P} são os correspondentes autovetores.

Propriedade 2.7. *Seja \mathbf{A} uma matriz simétrica ($m \times m$) então,*

1. $\mathbf{A}^k = \mathbf{P}\mathbf{\Lambda}^k\mathbf{P}'$.
2. *Se os autovalores de \mathbf{A} são menores que 1, tem-se que quando $k \rightarrow \infty$,*
 - $\mathbf{A}^k \rightarrow \mathbf{0}$ e
 - $\mathbf{I} + \mathbf{A} + \dots + \mathbf{A}^k \rightarrow (\mathbf{I} - \mathbf{A})^{-1}$.

2.5 Propriedades da Esperança Condicional

Sejam (Ω, \mathcal{F}, P) um espaço de probabilidade, $\mathcal{G} \subset \mathcal{F}$ uma σ -álgebra e X e Y variáveis aleatórias integráveis, tem-se que

1. $E(E(X|\mathcal{G})) = E(X)$,
2. se X é independente de \mathcal{H} , então $E(X|\mathcal{H}) = E(X)$,
3. se \mathcal{H} é uma sub- σ -álgebra de \mathcal{G} , então

$$E(E(X|\mathcal{G})|\mathcal{H}) = E(X|\mathcal{H}),$$

4. $E(aX + bY|\mathcal{G}) = aE(X|\mathcal{G}) + bE(Y|\mathcal{G})$.

As provas destas propriedades podem ser encontradas em Williams (1997).

Capítulo 3

Modelos GARCH Multivariados

Neste capítulo serão apresentados os diferentes modelos GARCH multivariados, discutindo as vantagens e desvantagens de cada um deles em relação ao número de parâmetros a serem estimados e à dinâmica destes com respeito à dependência do passado.

Seja $\{\mathbf{r}_t\}$ a série de retornos¹ multivariada. Pode-se escrever \mathbf{r}_t ($k \times 1$), para cada instante t , como

$$\mathbf{r}_t = \mathbf{c} + \boldsymbol{\varepsilon}_t, \quad (3.1)$$

em que \mathbf{c} é o vetor médio constante ($k \times 1$), e $\boldsymbol{\varepsilon}_t$ é um vetor ($k \times 1$) ruído branco com média zero, isto é, $E(\boldsymbol{\varepsilon}_t) = \mathbf{0}$ e $E\{\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s'\} = \mathbf{0}_{k \times k}$, para todo $t \neq s$, em que $\mathbf{0}_{k \times k}$ denota uma matriz de zeros de dimensão ($k \times k$). Para um caso mais geral, a constante \mathbf{c} pode ser substituída pela esperança condicional de \mathbf{r}_t , $\boldsymbol{\mu}_t = E(\mathbf{r}_t | \mathcal{F}_{t-1})$, dada a sigma-álgebra, \mathcal{F}_{t-1} , gerada pelos valores passados $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_{t-1}\}$ e considerar um modelo ARMA vetorial para o processo $\{\mathbf{r}_t\}$.

A matriz de covariâncias condicional de $\boldsymbol{\varepsilon}_t$ dada \mathcal{F}_{t-1} , é uma matriz $\boldsymbol{\Sigma}_t$ positiva semidefinida de dimensão ($k \times k$) e é definida como $\boldsymbol{\Sigma}_t = Cov(\boldsymbol{\varepsilon}_t | \mathcal{F}_{t-1}) = E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' | \mathcal{F}_{t-1})$.

A modelagem da volatilidade multivariada é baseada na evolução no tempo de $\boldsymbol{\Sigma}_t$. Assim, o modelo para o processo $\{\boldsymbol{\Sigma}_t\}$ é o modelo para a volatilidade dos retornos $\{\mathbf{r}_t\}$. O objetivo é ajustar um modelo multivariado para a matriz de covariâncias condicionais analogamente à modelagem GARCH univariada para a volatilidade de uma série.

¹Como convenção considera-se como retorno ao log-retorno do preço de um ativo financeiro (ver Morettin (2002)).

3.1 Modelo GARCH Univariado

A parametrização da variância condicional utilizada por Engle (1982) para modelar os momentos não observáveis de segunda ordem permite que a variância condicional dependa dos elementos do conjunto de informações passadas numa forma autoregressiva. Seja \mathcal{F}_{t-1} a sigma álgebra gerada pelos valores passados de ε_t , o modelo ARCH univariado da ordem p , normalmente distribuído², pode ser escrito como

$$\varepsilon_t | \mathcal{F}_{t-1} \sim N(0, \sigma_t^2),$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2,$$

em que $\{\varepsilon_t\}$ são variáveis aleatórias independentes, $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, \dots, p-1$, e $\alpha_p > 0$. Bollerslev (1986) generalizou o modelo ARCH permitindo que as variâncias condicionais dependam também de seu passado. A variância condicional foi dada por

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2,$$

em que $\{\varepsilon_t\}$ são variáveis aleatórias independentes, $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, \dots, p-1$, $\alpha_p > 0$, $\beta_j \geq 0$, $j = 1, \dots, q-1$, e $\beta_q > 0$. Este modelo ARCH generalizado é conhecido como modelo GARCH(p,q) univariado. A estacionariedade deste processo é obtida quando $(\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j) < 1$. A prova desta propriedade e todas as demais características dos modelos GARCH univariados podem ser encontradas em Bollerslev (1986) e Morettin (2002).

3.2 Modelos GARCH Multivariados Diretos

Dadas as observações $\{\mathbf{r}_1, \dots, \mathbf{r}_{t-1}\}$, uma forma intuitiva de ajustar um modelo multivariado para o processo $\{\Sigma_t\}$ é considerar o estimador clássico para a matriz de covariâncias dado por

$$\widehat{\Sigma}_t = \frac{1}{t-1} \sum_{j=1}^{t-1} (\mathbf{r}_j - \mathbf{c})(\mathbf{r}_j - \mathbf{c})'. \quad (3.2)$$

Na equação (3.2) pode-se notar que todas as observações têm peso igual a $1/(t-1)$, que é uma desvantagem devido às observações mais recentes terem maior influência na estimação. Portanto, para distribuir pesos diferentes ao longo do tempo, foi considerado o seguinte modelo nomeado modelo de Média Móvel Exponencialmente Ponderada.

²Para outras distribuições, que descrevam melhor as caudas pesadas de séries financeiras, ver Morettin (2002).

3.2.1 Média Móvel Exponencialmente Ponderada

Para estimar uma matriz de covariâncias variando no tempo, além de dar maior peso às observações mais recentes, uma possibilidade é usar pesos exponenciais e estimar a matriz de covariâncias de $\mathbf{r}_t - \mathbf{c} = \boldsymbol{\varepsilon}_t$ por

$$\widehat{\boldsymbol{\Sigma}}_t = \frac{1 - \lambda}{1 - \lambda^{t-1}} \sum_{j=1}^{t-1} \lambda^{j-1} \boldsymbol{\varepsilon}_{t-j} \boldsymbol{\varepsilon}'_{t-j}, \quad (3.3)$$

em que $0 < \lambda < 1$, e a soma dos pesos $(1 - \lambda)\lambda^{j-1}/(1 - \lambda^{t-1})$ é igual a 1, uma vez que

$$\begin{aligned} \sum_{j=1}^{t-1} \frac{(1 - \lambda)\lambda^{j-1}}{(1 - \lambda^{t-1})} &= \frac{(1 - \lambda)}{(1 - \lambda^{t-1})} \sum_{j=1}^{t-1} \lambda^{j-1} \\ &= \frac{\cancel{(1 - \lambda)}}{\cancel{(1 - \lambda)}(1 + \lambda + \dots + \lambda^{t-2})} \sum_{j=1}^{t-1} \lambda^{j-1} \\ &= \frac{(1 + \lambda + \dots + \lambda^{t-2})}{(1 + \lambda + \dots + \lambda^{t-2})} \\ &= 1. \end{aligned}$$

Da equação (3.3) tem-se que

$$\widehat{\boldsymbol{\Sigma}}_t = \left(\frac{1 - \lambda}{1 - \lambda^{t-1}} \right) \boldsymbol{\varepsilon}_{t-1} \boldsymbol{\varepsilon}'_{t-1} + \left(\frac{1 - \lambda}{1 - \lambda^{t-1}} \right) \sum_{j=2}^{t-1} \lambda^{j-1} \boldsymbol{\varepsilon}_{t-j} \boldsymbol{\varepsilon}'_{t-j},$$

e fazendo $i = j - 1$ obtém-se que

$$\begin{aligned} \widehat{\boldsymbol{\Sigma}}_t &= \left(\frac{1 - \lambda}{1 - \lambda^{t-1}} \right) \boldsymbol{\varepsilon}_{t-1} \boldsymbol{\varepsilon}'_{t-1} + \left(\frac{1 - \lambda}{1 - \lambda^{t-1}} \right) \sum_{i=1}^{(t-1)-1} \lambda^i \boldsymbol{\varepsilon}_{(t-1)-i} \boldsymbol{\varepsilon}'_{(t-1)-i} \\ &= \left(\frac{1 - \lambda}{1 - \lambda^{t-1}} \right) \boldsymbol{\varepsilon}_{t-1} \boldsymbol{\varepsilon}'_{t-1} + \left(\frac{1 - \lambda^{t-2}}{1 - \lambda^{t-1}} \right) \lambda \widehat{\boldsymbol{\Sigma}}_{t-1}. \end{aligned}$$

Para um t suficientemente grande, $\lambda^{t-1} \approx 0$ e portanto $\widehat{\boldsymbol{\Sigma}}_t$ pode ser escrita como

$$\widehat{\boldsymbol{\Sigma}}_t = (1 - \lambda) \boldsymbol{\varepsilon}_{t-1} \boldsymbol{\varepsilon}'_{t-1} + \lambda \widehat{\boldsymbol{\Sigma}}_{t-1}. \quad (3.4)$$

A equação (3.4) é chamada modelo EWMA ("Média Móvel Exponencialmente Ponderada") da matriz de covariâncias condicional.

Da equação (3.4) tem-se que, dados os retornos multivariados $\{\mathbf{r}_1, \dots, \mathbf{r}_{t-1}\}$, um valor de λ e uma estimativa inicial $\widehat{\Sigma}_1$, a matriz $\widehat{\Sigma}_t$ pode ser calculada recursivamente. Na prática o valor de λ é escolhido de maneira conveniente considerando o nível de suavização que for necessário, mas se assume que $\boldsymbol{\varepsilon}_t = \mathbf{r}_t - \mathbf{c}$ segue uma distribuição normal multivariada com média zero e matriz de covariâncias Σ_t , o vetor média \mathbf{c} e o parâmetro λ podem ser estimados conjuntamente pelo método de máxima verossimilhança, baseados na função de log-verossimilhança

$$\log L(\mathbf{c}, \lambda) \propto -\frac{1}{2} \sum_{t=1}^T \log |\Sigma_t| - \frac{1}{2} \sum_{t=1}^T (\mathbf{r}_t - \mathbf{c})' \Sigma_t^{-1} (\mathbf{r}_t - \mathbf{c}).$$

Este modelo EWMA dado pela equação (3.4) é o modelo mais simples para ajustar o processo multivariado $\{\Sigma_t\}$, já que precisa estimar somente um parâmetro λ , embora ele tenha a desvantagem de não ser estacionário, uma vez que a matriz de covariâncias não condicional não está definida, similar ao modelo IGARCH univariado³. Para resolver este problema Bollerslev *et al.* (1988) e Engle e Kroner (1995) generalizaram o método EWMA propondo um novo modelo conhecido como modelo VEC.

Exemplo 3.1. Considera-se um ajuste GARCH multivariado EWMA normalmente distribuído, para o conjunto de retornos das três séries dadas no Capítulo 1, VALE, PETROBRAS e GERDAU.

A Tabela 3.1 apresenta as estimativas do vetor \mathbf{c} na equação (3.1) e o parâmetro estimado $\widehat{\lambda}$ do modelo (3.4). Nota-se que segundo os p -valores todos os parâmetros são significativos ao nível de significância de 5 %.

Tabela 3.1: *Parâmetros estimados EWMA*

Parâmetros Estimados EWMA:				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	0,008085	0,001689	4,786	<0,001
C(2)	0,005700	0,001771	3,219	0,001
C(3)	0,008521	0,002015	4,229	<0,001
$\widehat{\lambda}$	0,953920	0,003074	14,992	<0,001

Segundo a Tabela 3.1 o modelo estimado é dado por:

³Ver Morettin (2002)

$$\mathbf{r}_t = \begin{pmatrix} 0,008085 \\ 0,005700 \\ 0,008521 \end{pmatrix} + \boldsymbol{\varepsilon}_t \quad (3.5)$$

$$\boldsymbol{\Sigma}_t = 0,95392(\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}'_{t-1}) + 0,04608\boldsymbol{\Sigma}_{t-1},$$

em que $\boldsymbol{\varepsilon}_t \sim N_3(\mathbf{0}, \boldsymbol{\Sigma}_t)$. A Figura 3.1 apresenta as variâncias e covariâncias condicionais estimadas segundo o modelo (3.5). Nota-se que a maior variabilidade foi nos anos 1999 e finais de 2009.

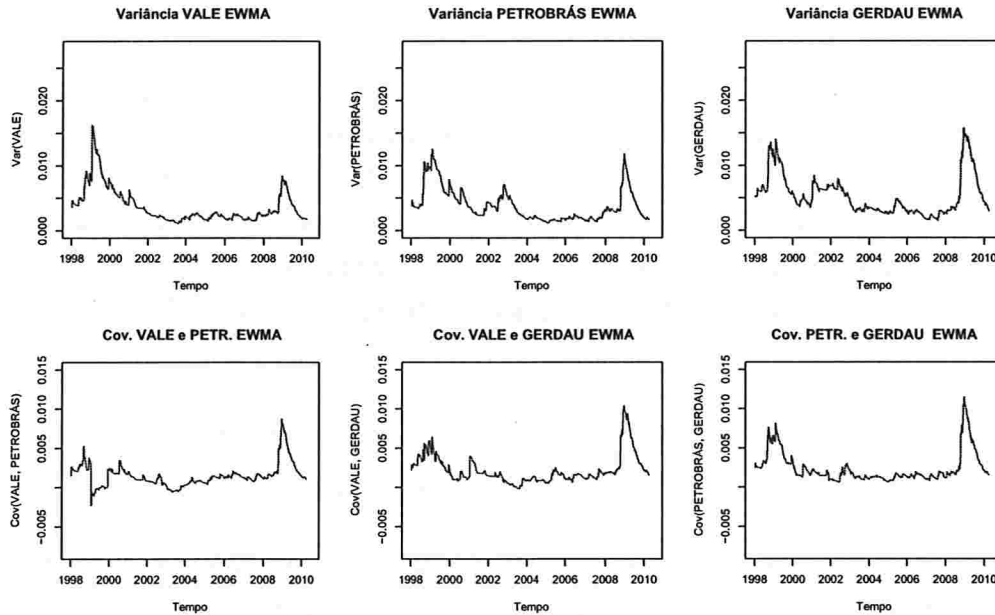


Figura 3.1: Variâncias e covariâncias condicionais estimadas segundo o modelo EWMA.

3.2.2 Modelo VEC

Calculando a esperança não condicional em ambos os lados do modelo EWMA dado em (3.4), tem-se que

$$E(\widehat{\boldsymbol{\Sigma}}_t) = (1 - \lambda)E(\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}'_{t-1}) + \lambda E(\widehat{\boldsymbol{\Sigma}}_{t-1}), \quad (3.6)$$

assim, do lado esquerdo da equação (3.6):

$$E(\widehat{\Sigma}_t) = \Sigma,$$

e do lado direito desta equação:

$$\begin{aligned} (1 - \lambda)E(\varepsilon_{t-1}\varepsilon'_{t-1}) + \lambda E(\widehat{\Sigma}_{t-1}) &= (1 - \lambda)\Sigma + \lambda\Sigma \\ &= \Sigma - \lambda\Sigma + \lambda\Sigma \\ &= \Sigma, \end{aligned}$$

portanto, pelos os dois lados $\Sigma = \Sigma$, logo a equação (3.6) tem infinitas soluções e assim, a matriz de covariâncias não condicional não está definida. Logo, o modelo EWMA é não estacionário na covariância.

Para preservar a intuição por trás do modelo EWMA, e no mesmo tempo permitindo um modelo flexível em que é possível considerar a estacionariedade na covariância, Bollerslev *et al.* (1988) e Engle e Kroner (1995) generalizaram o modelo EWMA num modelo VEC. Como o modelo GARCH univariado, o modelo VEC permite que a matriz de covariâncias Σ_t seja dependente das informações passadas \mathcal{F}_{t-1} , usando p lags passados dos quadrados e produtos cruzados de ε_t , e também de q lags passados de Σ_t . Assim, o modelo VEC é dado por

$$\begin{aligned} \text{vec}(\Sigma_t) &= \mathbf{c}_0 + \mathbf{C}_1 \text{vec}(\varepsilon_{t-1}\varepsilon'_{t-1}) + \dots + \mathbf{C}_p \text{vec}(\varepsilon_{t-p}\varepsilon'_{t-p}) \\ &\quad + \mathbf{G}_1 \text{vec}(\Sigma_{t-1}) + \dots + \mathbf{G}_q \text{vec}(\Sigma_{t-q}), \end{aligned} \tag{3.7}$$

em que $\text{vec}(\cdot)$ é o operador que permite empilhar uma matriz num vetor, escrevendo uma coluna embaixo da outra, \mathbf{c}_0 é um vetor de parâmetros ($k^2 \times 1$), e \mathbf{C}_i e \mathbf{G}_j são matrizes de parâmetros ($k^2 \times k^2$).

Em forma matricial este modelo pode ser escrito como

$$\begin{aligned}
 \mathbf{h}_t &= \begin{bmatrix} \mathbf{c}_0 & \vdots & \mathbf{C}_1 & \vdots & \dots & \vdots & \mathbf{C}_p & \vdots & \mathbf{G}_1 & \vdots & \dots & \vdots & \mathbf{G}_q \end{bmatrix} \begin{bmatrix} 1 \\ \boldsymbol{\eta}_{t-1} \\ \vdots \\ \boldsymbol{\eta}_{t-p} \\ \mathbf{h}_{t-1} \\ \vdots \\ \mathbf{h}_{t-q} \end{bmatrix} \\
 &= \mathbf{F}\mathbf{z}_t \\
 &= (\mathbf{z}'_t \otimes \mathbf{I}) \text{vec}(\mathbf{F}) \\
 &= \mathbf{Z}_t\boldsymbol{\alpha},
 \end{aligned} \tag{3.8}$$

em que

\mathbf{I} : matriz identidade,

\otimes : produto Kronecker,

$\mathbf{h}_t = \text{vec}(\boldsymbol{\Sigma}_t)$,

$\boldsymbol{\eta}_t = \text{vec}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}'_t)$,

$\mathbf{z}_t = (1, \boldsymbol{\eta}'_{t-1}, \dots, \boldsymbol{\eta}'_{t-p}, \mathbf{h}'_{t-1}, \dots, \mathbf{h}'_{t-q})$,

$\mathbf{F} = \begin{bmatrix} \mathbf{c}_0 & \vdots & \mathbf{C}_1 & \vdots & \dots & \vdots & \mathbf{C}_p & \vdots & \mathbf{G}_1 & \vdots & \dots & \vdots & \mathbf{G}_q \end{bmatrix}$,

$\boldsymbol{\alpha} = \text{vec}(\mathbf{F})$

e

$\mathbf{Z}_t = (\mathbf{z}'_t \otimes \mathbf{I})$.

A equação (3.8) define a parametrização da matriz de covariâncias condicional conhecida como representação VEC(p,q). Para ter uma ideia intuitiva deste modelo, apresenta-se o

modelo VEC(1,1) bivariado

$$\mathbf{h}_t = \begin{bmatrix} \mathbf{h}_t^{(11)} \\ \mathbf{h}_t^{(12)} \\ \mathbf{h}_t^{(22)} \end{bmatrix} = \begin{bmatrix} \mathbf{c}_0^{(1)} \\ \mathbf{c}_0^{(2)} \\ \mathbf{c}_0^{(3)} \end{bmatrix} + \begin{bmatrix} \mathbf{C}_1^{(11)} & \mathbf{C}_1^{(12)} & \mathbf{C}_1^{(13)} \\ \mathbf{C}_1^{(21)} & \mathbf{C}_1^{(22)} & \mathbf{C}_1^{(23)} \\ \mathbf{C}_1^{(31)} & \mathbf{C}_1^{(32)} & \mathbf{C}_1^{(33)} \end{bmatrix} \begin{bmatrix} \varepsilon_{t-1}^{(1)} \varepsilon_{t-1}^{(1)} \\ \varepsilon_{t-1}^{(1)} \varepsilon_{t-1}^{(2)} \\ \varepsilon_{t-1}^{(2)} \varepsilon_{t-1}^{(2)} \end{bmatrix} + \begin{bmatrix} \mathbf{G}_1^{(11)} & \mathbf{G}_1^{(12)} & \mathbf{G}_1^{(13)} \\ \mathbf{G}_1^{(21)} & \mathbf{G}_1^{(22)} & \mathbf{G}_1^{(23)} \\ \mathbf{G}_1^{(31)} & \mathbf{G}_1^{(32)} & \mathbf{G}_1^{(33)} \end{bmatrix} \begin{bmatrix} \mathbf{h}_{t-1}^{(11)} \\ \mathbf{h}_{t-1}^{(12)} \\ \mathbf{h}_{t-1}^{(22)} \end{bmatrix}.$$

Nota-se que foi omitido o elemento $\mathbf{h}_t^{(21)}$ e que não foram dados parâmetros para os elementos $\varepsilon_{t-1}^{(2)} \varepsilon_{t-1}^{(1)}$ e $\mathbf{h}_{t-1}^{(21)}$, já que estes são redundantes, gerando sete parâmetros a mais em cada uma das matrizes \mathbf{C}_1 e \mathbf{G}_1 . Redundâncias similares aparecem no modelo geral n -variado VEC(1,1). Em particular, todas as equações para as covariâncias aparecem duas vezes (i.e., existe uma equação para $\mathbf{h}_t^{(ij)}$ e outra para $\mathbf{h}_t^{(ji)}$) e todos os termos fora da diagonal aparecem também duas vezes (i.e., os dois termos $\varepsilon_{t-1}^{(i)} \varepsilon_{t-1}^{(j)}$ e $\varepsilon_{t-1}^{(j)} \varepsilon_{t-1}^{(i)}$ junto com $\mathbf{h}_{t-1}^{(ij)}$ e $\mathbf{h}_{t-1}^{(ji)}$ aparecem duas vezes em cada equação). Os termos redundantes podem ser eliminados sem afectar o modelo, levando num total de $(k(k+1)/2)^2$ parâmetros em cada uma das matrizes \mathbf{C}_i e \mathbf{G}_j . Numa formulação direta de (3.8) aparecem k^4 parâmetros em cada matriz, mas muitos deles são considerados duas vezes.

Para facilitar a implementação prática é desejável restringir a quantidade de parâmetros no modelo VEC(p,q). Uma restrição natural que foi usada no contexto ARCH por Engle *et al.* (1984) e no contexto GARCH por Bollerslev *et al.* (1988), é a *representação diagonal*, em que cada elemento da matriz de covariâncias condicional, $\mathbf{h}_t^{(ij)}$, depende só do seu passado e de valores passados de $\varepsilon_t^{(i)} \varepsilon_t^{(j)}$. Isto é, as variâncias dependem só dos seus próprios passados e do quadrado dos seus erros, e as covariâncias dependem dos seus passados e do produto cruzado dos erros. Como o modelo obtido com a representação diagonal é mais aplicável, ele será estudado separadamente com base na abordagem feita por Zivot e Wang (2006).

3.2.3 Modelo VEC Diagonal

A abordagem dada em Zivot e Wang (2006) para o modelo DVEC ("Diagonal VEC") é modelar a matriz de covariâncias condicional como

$$\Sigma_t = \mathbf{A}_0 + \sum_{i=1}^p \mathbf{A}_i \odot (\varepsilon_{t-i} \varepsilon'_{t-i}) + \sum_{j=1}^q \mathbf{B}_j \odot \Sigma_{t-j}, \quad (3.9)$$

em que p e q são inteiros não negativos, \mathbf{A}_i e \mathbf{B}_j matrizes simétricas de dimensão $(k \times k)$ e " \odot " denota o produto Hadamard, isto é, produto elemento por elemento. O modelo dado em (3.9) é conhecido como o modelo VEC diagonal e a notação usada é DVEC(p,q). Comparando este modelo com o modelo GARCH univariado nota-se que existe certa analogia, em que p representa o número de parâmetros na parte ARCH e q representa o número destes na parte GARCH.

Para ter uma ideia intuitiva por trás do modelo DVEC, considera-se o modelo bivariado DVEC(1,1):

$$\begin{pmatrix} \Sigma_t^{(11)} \\ \Sigma_t^{(21)} & \Sigma_t^{(22)} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_0^{(11)} \\ \mathbf{A}_0^{(21)} & \mathbf{A}_0^{(22)} \end{pmatrix} + \begin{pmatrix} \mathbf{A}_1^{(11)} \\ \mathbf{A}_1^{(21)} & \mathbf{A}_1^{(22)} \end{pmatrix} \odot \begin{pmatrix} \varepsilon_{t-1}^{(1)} \varepsilon_{t-1}'^{(1)} \\ \varepsilon_{t-1}^{(2)} \varepsilon_{t-1}'^{(1)} & \varepsilon_{t-1}^{(2)} \varepsilon_{t-1}'^{(2)} \end{pmatrix} \\ + \begin{pmatrix} \mathbf{B}_1^{(11)} \\ \mathbf{B}_1^{(21)} & \mathbf{B}_1^{(22)} \end{pmatrix} \odot \begin{pmatrix} \Sigma_{t-1}^{(11)} \\ \Sigma_{t-1}^{(21)} & \Sigma_{t-1}^{(22)} \end{pmatrix}, \quad (3.10)$$

em que só a parte inferior do sistema foi considerada devido à simetria das matrizes, com $\mathbf{X}^{(ij)}$ denotando o elemento (i, j) da matriz \mathbf{X} e $\varepsilon^{(i)}$ denota o i -ésimo elemento do vetor ε .

Considerando o operador vec , e aplicando este só à parte inferior das matrizes simétricas tem-se que, para Σ_t

$$vec \left[\begin{pmatrix} \Sigma_t^{(11)} \\ \Sigma_t^{(21)} & \Sigma_t^{(22)} \end{pmatrix} \right] = \begin{pmatrix} \Sigma_t^{(11)} \\ \Sigma_t^{(21)} \\ \Sigma_t^{(22)} \end{pmatrix},$$

e portanto o modelo (3.10) pode ser escrito como

$$vec(\Sigma_t) = vec(\mathbf{A}_0) + vec(\mathbf{A}_1 \odot \varepsilon_{t-1} \varepsilon'_{t-1}) + vec(\mathbf{B}_1 \odot \Sigma_{t-1}). \quad (3.11)$$

Considerando a relação entre o produto Hadamard e o operador vec , dada por $vec(\mathbf{A} \odot \mathbf{B}) = diag\{vec(\mathbf{A})\}vec(\mathbf{B})$, em que $diag\{vec(\mathbf{A})\}$ representa a matriz diagonal com o vetor $vec(\mathbf{A})$ na diagonal principal, a equação (3.11) fica

$$vec(\Sigma_t) = vec(\mathbf{A}_0) + diag\{vec(\mathbf{A}_1)\} vec(\varepsilon_{t-1}\varepsilon'_{t-1}) + diag\{vec(\mathbf{B}_1)\} vec(\Sigma_{t-1}).$$

Generalizando esta ideia, o modelo dado em (3.9) pode ser escrito como

$$\begin{aligned} vec(\Sigma_t) &= vec(\mathbf{A}_0) + \sum_{i=1}^p vec(\mathbf{A}_i \odot \varepsilon_{t-i}\varepsilon'_{t-i}) + \sum_{j=1}^q vec(\mathbf{B}_j \odot \Sigma_{t-j}) \\ &= vec(\mathbf{A}_0) + \sum_{i=1}^p diag\{vec(\mathbf{A}_i)\} vec(\varepsilon_{t-i}\varepsilon'_{t-i}) \\ &\quad + \sum_{j=1}^q diag\{vec(\mathbf{B}_j)\} vec(\Sigma_{t-j}). \end{aligned} \tag{3.12}$$

A forma dada em (3.12), em que o operador vec foi usado, dá uma ideia intuitiva em relação a razão do nome VEC Diagonal para este modelo, já que coincide com o modelo (3.7) quando $\mathbf{C}_1, \dots, \mathbf{C}_p, \mathbf{G}_1, \dots, \mathbf{G}_q$ são matrizes diagonais.

A forma matricial (3.10) pode ser escrita também como

$$\begin{aligned} \Sigma_t^{(11)} &= \mathbf{A}_0^{(11)} + \mathbf{A}_1^{(11)} \varepsilon_{t-1}^{(1)} \varepsilon_{t-1}^{(1)'} + \mathbf{B}_1^{(11)} \Sigma_{t-1}^{(11)}, \\ \Sigma_t^{(21)} &= \mathbf{A}_0^{(21)} + \mathbf{A}_1^{(21)} \varepsilon_{t-1}^{(2)} \varepsilon_{t-1}^{(1)'} + \mathbf{B}_1^{(21)} \Sigma_{t-1}^{(21)}, \\ \Sigma_t^{(22)} &= \mathbf{A}_0^{(22)} + \mathbf{A}_1^{(22)} \varepsilon_{t-1}^{(2)} \varepsilon_{t-1}^{(2)'} + \mathbf{B}_1^{(22)} \Sigma_{t-1}^{(22)}, \end{aligned} \tag{3.13}$$

então, o elemento (i, j) da matriz de covariâncias condicionais depende de seu próprio elemento um lag anterior e do correspondente produto cruzado dos erros. Também é possível notar que a volatilidade de cada série segue um processo GARCH e enquanto à covariância, pode ser também tratada como um modelo GARCH em relação ao momento cruzado dos erros.

Uma das desvantagens deste modelo DVEC(p,q) é que ele não é dinâmico, por exemplo, as volatilidades não dependem da covariância entre as séries e além disso não é possível garantir de forma geral que a matriz Σ_t seja *positiva semidefinida* ("PSD") para cada instante t , que é uma condição fundamental das matrizes de covariâncias. Para resolver este

último problema foram propostas algumas extensões do modelo DVEC(p,q), conhecidos como modelos *matrizes diagonais*.

3.2.4 Modelo Matriz-Diagonal

Embora o modelo DVEC dado em (3.9) proporcione um bom modelo para a matriz de covariâncias condicional Σ_t , ele não garante que para cada instante t esta matriz seja PSD. Uma condição suficiente para que Σ_t seja PSD é que \mathbf{A}_i , $i = 0, 1, \dots, p$, e \mathbf{B}_j , $j = 1, \dots, q$, sejam todas PSD (Ver Apêndice A Proposição A.1). Segundo esta observação, Ding (1994) e Bollerslev *et al.* (1994) propuseram estimar os fatores de Cholesky das matrizes de parâmetros⁴

$$\Sigma_t = \mathbf{A}_0 \mathbf{A}'_0 + \sum_{i=1}^p (\mathbf{A}_i \mathbf{A}'_i) \odot (\varepsilon_{t-i} \varepsilon'_{t-i}) + \sum_{j=1}^q (\mathbf{B}_j \mathbf{B}'_j) \odot \Sigma_{t-j}, \quad (3.14)$$

em que \mathbf{A}_i , para $i = 0, 1, \dots, p$ e \mathbf{B}_j , para $j = 1, \dots, q$ são matrizes triangulares. O modelo (3.14) é conhecido como *Modelo Matriz-Diagonal*.

O modelo Matriz-Diagonal pode ser simplificado considerando vetores em lugar das matrizes \mathbf{A}_i e \mathbf{B}_j , o qual resulta em

$$\Sigma_t = \mathbf{A}_0 \mathbf{A}'_0 + \sum_{i=1}^p (\mathbf{a}_i \mathbf{a}'_i) \odot (\varepsilon_{t-i} \varepsilon'_{t-i}) + \sum_{j=1}^q (\mathbf{b}_j \mathbf{b}'_j) \odot \Sigma_{t-j}, \quad (3.15)$$

em que \mathbf{a}_i , $i = 1, \dots, p$, e \mathbf{b}_j , $j = 1, \dots, q$, são vetores de dimensão $(k \times 1)$. Ainda mais simples, é considerar

$$\Sigma_t = \mathbf{A}_0 \mathbf{A}'_0 + \sum_{i=1}^p a_i (\varepsilon_{t-i} \varepsilon'_{t-i}) + \sum_{j=1}^q b_j \Sigma_{t-j}, \quad (3.16)$$

com a_i e b_j escalares positivos, para $i = 1, \dots, p$ e $j = 1, \dots, q$, respectivamente.

O modelo (3.15) gera matrizes positivas semidefinidas devido às matrizes $\mathbf{a}_i \mathbf{a}'_i$ e $\mathbf{b}_j \mathbf{b}'_j$ serem PSD. Para ver isto considere um vetor qualquer $\mathbf{x} \neq \mathbf{0}$ de dimensão $(k \times 1)$, assim

$$\mathbf{x}'(\mathbf{a}_i \mathbf{a}'_i)\mathbf{x} = (\mathbf{x}'\mathbf{a}_i)(\mathbf{x}'\mathbf{a}_i)' = (\mathbf{x}'\mathbf{a}_i)^2 \geq 0,$$

já que $\mathbf{x}'\mathbf{a}_i$ tem dimensão (1×1) . Analogamente, obtém-se que $\mathbf{b}_j \mathbf{b}'_j$ é PSD. Já para o modelo

⁴Uma matriz \mathbf{A} é simétrica e PSD se e somente se esta matriz pode ser escrita na decomposição de Cholesky $\mathbf{A} = \mathbf{C}\mathbf{C}'$, em que \mathbf{C} é uma matriz triangular inferior.

(3.16) a condição de que a_i e b_j são escalares positivos é suficiente para garantir que Σ_t seja PSD.

Exemplo 3.2. Consideram-se novamente os retornos do conjunto de séries VALE, PETROBRAS e GERDAU. Um ajuste DVEC(1,1) normalmente distribuído foi considerado para estes retornos e os parâmetros estimados estão na Tabela 3.2. Todos os p -valores são menores que 5% e portanto todos eles são estatisticamente significantes.

Tabela 3.2: *Parâmetros estimados DVEC(1,1)*

Parâmetros Estimados DVEC(1,1):				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	0,00900368	0,00210235	4,283	<0,001
C(2)	0,00645814	0,00225033	2,870	0,002
C(3)	0,01021560	0,00255175	4,003	<0,001
A(1, 1)	0,00005617	0,00002217	2,534	0,006
A(2, 1)	0,00005282	0,00001999	2,643	0,004
A(3, 1)	0,00004929	0,00001629	3,026	0,001
A(2, 2)	0,00009062	0,00003191	2,840	0,002
A(3, 2)	0,00005069	0,00002068	2,451	0,007
A(3, 3)	0,00011122	0,00005125	2,170	0,010
ARCH(1; 1, 1)	0,06356726	0,01525376	4,167	<0,001
ARCH(1; 2, 1)	0,05082172	0,01110251	4,577	<0,001
ARCH(1; 3, 1)	0,04171108	0,00983030	4,243	<0,001
ARCH(1; 2, 2)	0,08618286	0,01434725	6,007	<0,001
ARCH(1; 3, 2)	0,04432857	0,01007950	4,398	<0,001
ARCH(1; 3, 3)	0,06323435	0,01403759	4,505	<0,001
GARCH(1; 1, 1)	0,92060709	0,01790613	51,413	<0,001
GARCH(1; 2, 1)	0,91920809	0,01772600	51,856	<0,001
GARCH(1; 3, 1)	0,93495529	0,01351536	69,177	<0,001
GARCH(1; 2, 2)	0,89271719	0,01697574	52,588	<0,001
GARCH(1; 3, 2)	0,93275149	0,01137124	82,027	<0,001
GARCH(1; 3, 3)	0,91420737	0,01880394	48,618	<0,001

Segundo a equação (3.9), o modelo estimado DVEC(1,1) é dado por:

$$\begin{aligned} \mathbf{\Gamma}_t &= \begin{pmatrix} 0,00900368 \\ 0,00645814 \\ 0,01021560 \end{pmatrix} + \varepsilon_t \\ \Sigma_t &= \begin{pmatrix} 0,00005617 & 0,00005282 & 0,00004929 \\ 0,00005282 & 0,00009062 & 0,00005069 \\ 0,00004929 & 0,00005069 & 0,00011122 \end{pmatrix} \\ &+ \begin{pmatrix} 0,06356726 & 0,05082172 & 0,04171108 \\ 0,05082172 & 0,89271719 & 0,04432857 \\ 0,04171108 & 0,04432857 & 0,06323435 \end{pmatrix} \odot (\varepsilon_{t-1}\varepsilon'_{t-1}) \\ &+ \begin{pmatrix} 0,92060709 & 0,91920809 & 0,93495529 \\ 0,91920809 & 0,89271719 & 0,93275149 \\ 0,93495529 & 0,93275149 & 0,91420737 \end{pmatrix} \odot \Sigma_{t-1}, \end{aligned} \tag{3.17}$$

em que $\varepsilon_t \sim N_3(0, \Sigma_t)$. A Figura 3.2 apresenta as variâncias e covariâncias condicionais dadas pelo modelo (3.17) e nota-se que têm um comportamento parecido com o modelo EWMA ajustado anteriormente, em que a variabilidade é maior nos anos 1999 e 2009.

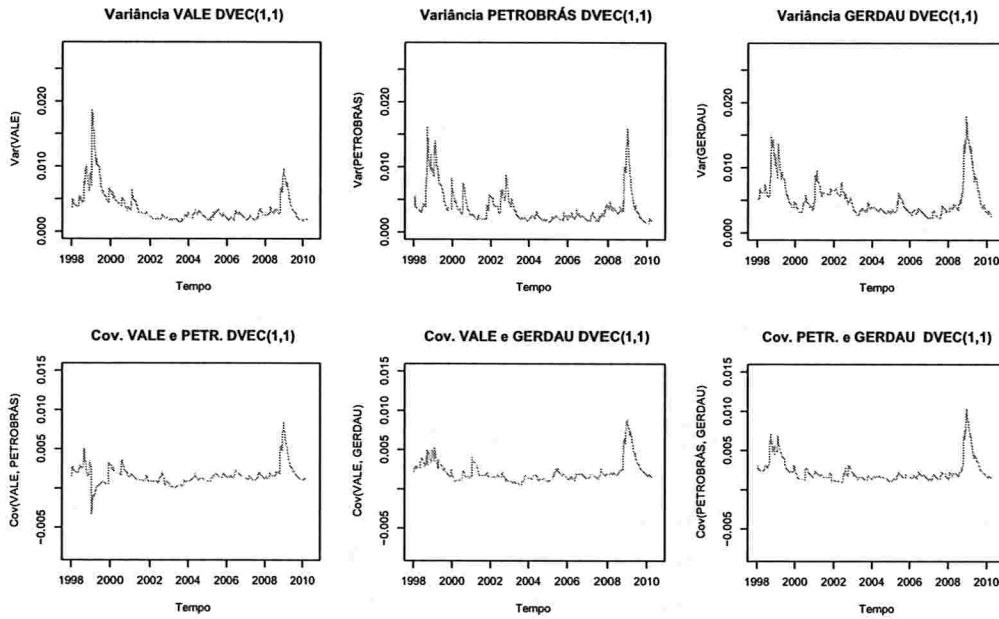


Figura 3.2: Variâncias e covariâncias condicionais estimadas segundo o modelo DVEC(1,1).

3.2.5 Modelo BEKK

O modelo DVEC pode ser modificado para garantir que a matriz Σ_t seja PSD para cada instante t , mas a dinâmica permitida na matriz de covariâncias condicional ainda continua, de alguma forma, restringida. Em particular, na equação (3.13), as variâncias e as covariâncias condicionais dependem só de seu próprio elemento um lag anterior e do correspondente produto cruzado dos erros.

Engle e Kroner (1995) propuseram o modelo Baba-Engle-Kraft-Kroner ("BEKK"), que oferece uma formulação alternativa para a matriz de covariâncias condicional,

$$\Sigma_t = \mathbf{A}_0 \mathbf{A}'_0 + \sum_{l=1}^r \sum_{i=1}^p \mathbf{A}_{il} (\varepsilon_{t-i} \varepsilon'_{t-i}) \mathbf{A}'_{il} + \sum_{l=1}^r \sum_{j=1}^q \mathbf{B}_{jl} \Sigma_{t-j} \mathbf{B}'_{jl}, \quad (3.18)$$

em que \mathbf{A}_0 é uma matriz triangular e \mathbf{A}_{il} , $i = 1, \dots, p$; $l = 1, \dots, r$, e \mathbf{B}'_{jl} , $j = 1, \dots, q$; $l = 1, \dots, r$, são matrizes quadráticas sem restrições. A constante r determina a generalidade do processo e na prática considera-se $r = 1$. Este modelo é denotado por BEKK(p,q,r) e a seguinte proposição mostra que ele gera matrizes PSD para cada instante t .

Proposição 3.1. *O modelo BEKK(p,q,r) dado em (3.18) gera matrizes, Σ_t , simétricas e PSD para todo t .*

A prova desta proposição está no Apêndice A.

A Proposição 3.1 garante que o processo $\{\Sigma_t\}$ gerado pelo modelo BEKK(p,q,r) gera matrizes PSD, a qual é uma propriedade fundamental para as matrizes de covariâncias condicionais. Além disso, este modelo permite uma maior dinâmica que o modelo DVEC, a qual pode ser ilustrada, por exemplo, considerando o elemento (2,2) de Σ_t do modelo BEKK(1,1,1),

$$\begin{aligned} \Sigma_t^{(22)} &= \mathbf{A}_0^{(22)} \mathbf{A}_0^{(22)} + [\mathbf{A}_{11}^{(21)} \varepsilon_{t-1}^{(1)} + \mathbf{A}_{11}^{(22)} \varepsilon_{t-1}^{(2)}]^2 + [\mathbf{B}_{11}^{(21)} \mathbf{B}_{11}^{(21)} \Sigma_{t-1}^{(11)} \\ &\quad + 2\mathbf{B}_{11}^{(21)} \mathbf{B}_{11}^{(22)} \Sigma_{t-1}^{(21)} + \mathbf{B}_{11}^{(22)} \mathbf{B}_{11}^{(22)} \Sigma_{t-1}^{(22)}], \end{aligned}$$

em que ambos $\varepsilon_{t-1}^{(1)}$ e $\varepsilon_{t-1}^{(2)}$ estão presentes nesta equação. Além disso, $\Sigma_{t-1}^{(11)}$, a volatilidade da primeira série do vetor de retornos, tem um impacto sobre $\Sigma_{t-1}^{(22)}$, a volatilidade da segunda série. O modelo BEKK(1,1,1) bivariado tem uma dinâmica maior devido a um número maior de parâmetros presentes no modelo, já que por exemplo, $\mathbf{A}_1^{(12)}$ e $\mathbf{B}_1^{(12)}$ são incluídos, em relação ao modelo DVEC(1,1) dado nas equações (3.13). Em geral, um modelo BEKK(p,q,r)

requer $k(k-1)(p+q)/2$ parâmetros a mais que o modelo DVEC da mesma ordem.

A representação do modelo BEKK(p,q,r) é suficientemente geral incluindo todas as representações diagonais PSD dadas pelo modelo DVEC(p,q) e muitas das representações PSD dadas pelo modelo VEC(p,q). Engle e Kroner (1995) mostraram algumas propriedades em relação ao valor que deve tomar a constante r , para que o modelo (3.18) seja suficientemente geral no sentido de ser equivalente à maior quantidade possível de representações VEC(p,q).

Existe uma relação entre os modelos VEC e BEKK dada matematicamente entre as matrizes de parâmetros pela propriedade $vec(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A}) vec(\mathbf{B})$, sendo \mathbf{A} , \mathbf{B} e \mathbf{C} matrizes com as dimensões adequadas. A Proposição 3.2 apresenta formalmente esta relação.

Proposição 3.2. *As parametrizações dos modelos VEC e BEKK dados em (3.7) e (3.18) respectivamente são equivalentes, se e somente se, existem matrizes \mathbf{A}_0 , \mathbf{A}_{il} e \mathbf{B}_{jl} , para $i = 1, \dots, p$ e $j = 1, \dots, q$, tais que*

$$\mathbf{c}_0 = (\mathbf{A}_0 \otimes \mathbf{A}_0) vec(\mathbf{I}_k), \quad \mathbf{C}_i = \sum_{l=1}^r (\mathbf{A}_{il} \otimes \mathbf{A}_{il}) \quad e \quad \mathbf{G}_j = \sum_{l=1}^r (\mathbf{B}_{jl} \otimes \mathbf{B}_{jl}).$$

A prova desta proposição está no Apêndice A.

As consequências da Proposição 3.2 são as seguintes:

- O modelo VEC obtido a partir do modelo BEKK é único, mas o recíproco não é verdadeiro. A transformação de um modelo VEC num modelo BEKK (quando esta existe) não é única uma vez que, para uma matriz \mathbf{C}_1 , a seleção de \mathbf{A}_{1l} não é única. Isto pode ser visto, considerando que $(\mathbf{A}_{1l} \otimes \mathbf{A}_{1l}) = (-\mathbf{A}_{1l} \otimes -\mathbf{A}_{1l})$, portanto embora $\mathbf{C}_i = \sum_{l=1}^r (\mathbf{A}_{il} \otimes \mathbf{A}_{il})$ seja única, a seleção de \mathbf{A}_{il} não é única.
- O modelo DVEC é obtido de um modelo BEKK, se e somente se, cada uma das matrizes \mathbf{A}_{il} e \mathbf{B}_{jl} são matrizes diagonais.
- As condições dadas na Proposição 3.2 dão uma caracterização em relação a quais modelos VEC têm representação BEKK e quais não têm, isto é, pode-se ver que os modelos VEC excluídos da parametrização geral BEKK são aqueles para os quais não existem \mathbf{A}_0 , \mathbf{A}_{il} e \mathbf{B}_{jl} satisfazendo as condições dadas na proposição. Como já foi comentado anteriormente, Engle e Kroner (1995) dão algumas condições para o valor da constante r , para que o modelo BEKK represente o maior número possível de modelos VEC.

Exemplo 3.3. Para os retornos do conjunto de séries VALE, PETROBRAS e GERDAU, considera-se um ajuste BEKK(1,1,1) normalmente distribuído e na Tabela 3.3 apresentam-se os parâmetros estimados.

Tabela 3.3: *Parâmetros estimados BEKK(1,1,1)*

Parâmetros Estimados BEKK(1,1,1):				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	0,00721590	0,002302	3,1342473	<0,001
C(2)	0,00507189	0,002343	2,1642734	0,015
C(3)	0,00878135	0,002913	3,0143479	0,001
A(1, 1)	0,01612854	0,002493	6,4685544	<0,001
A(2, 1)	0,01496409	0,002908	5,1459648	<0,001
A(3, 1)	0,02645911	0,003926	6,7398783	<0,001
A(2, 2)	0,00492814	0,001799	2,7389572	0,003
A(3, 2)	0,00680918	0,002912	2,3382028	0,009
A(3, 3)	0,00009193	0,172339	0,0005334	0,499
ARCH(1; 1, 1)	0,08315570	0,037110	2,2408181	0,013
ARCH(1; 2, 1)	0,00285317	0,041761	0,0683212	0,473
ARCH(1; 3, 1)	-0,14158466	0,054924	-2,5778357	0,005
ARCH(1; 1, 2)	-0,06735654	0,045736	-1,4727312	0,071
ARCH(1; 2, 2)	0,25131210	0,044582	5,6371246	<0,001
ARCH(1; 3, 2)	0,06365392	0,059364	1,0722658	0,142
ARCH(1; 1, 3)	0,23805990	0,042932	5,5450282	<0,001
ARCH(1; 2, 3)	0,10729915	0,044305	2,4218228	0,008
ARCH(1; 3, 3)	0,41437163	0,060689	6,8278443	<0,001
GARCH(1; 1, 1)	0,96075946	0,013575	70,7724793	<0,001
GARCH(1; 2, 1)	-0,00518259	0,012928	-0,4008658	0,344
GARCH(1; 3, 1)	-0,02860463	0,021592	-1,3247509	0,093
GARCH(1; 1, 2)	-0,00942665	0,017942	-0,5254069	0,300
GARCH(1; 2, 2)	0,95098311	0,016306	58,3223593	<0,001
GARCH(1; 3, 2)	-0,02456000	0,024660	-0,9960000	0,156
GARCH(1; 1, 3)	-0,07767000	0,022570	-3,4420000	<0,001
GARCH(1; 2, 3)	-0,07408000	0,021440	-3,4550000	<0,001
GARCH(1; 3, 3)	0,85888000	0,031210	27,5210000	<0,001

Da equação (3.18) e da Tabela 3.3, tem-se que o modelo BEKK(1,1,1) é dado por:

$$\begin{aligned} \mathbf{r}_t &= \begin{pmatrix} 0,007 \\ 0,005 \\ 0,009 \end{pmatrix} + \varepsilon_t \\ \Sigma_t &= \begin{pmatrix} 0,016 & 0,015 & 0,027 \\ 0,015 & 0,005 & 0,007 \\ 0,026 & 0,007 & 0,000 \end{pmatrix} + \begin{pmatrix} 0,083 & 0,000 & 0,238 \\ 0,000 & 0,251 & 0,107 \\ -0,142 & 0,000 & 0,414 \end{pmatrix} (\varepsilon_{t-1} \varepsilon'_{t-1}) \begin{pmatrix} 0,083 & 0,000 & -0,142 \\ 0,000 & 0,251 & 0,000 \\ 0,238 & 0,107 & 0,414 \end{pmatrix} \\ &+ \begin{pmatrix} 0,961 & 0,000 & -0,078 \\ 0,000 & 0,951 & -0,074 \\ -0,029 & 0,000 & 0,859 \end{pmatrix} \Sigma_{t-1} \begin{pmatrix} 0,961 & 0,000 & -0,029 \\ 0,000 & 0,951 & 0,000 \\ -0,078 & -0,074 & 0,859 \end{pmatrix}, \end{aligned} \quad (3.19)$$

em que $\varepsilon_t \sim N_3(\mathbf{0}, \Sigma_t)$. A Figura 3.3 apresenta as variâncias e covariâncias condicionais estimadas pelo modelo (3.19). O comportamento é parecido com os modelos EWMA e DVEC(1,1), ajustados anteriormente, mas nota-se que este modelo BEKK(1,1,1) é um pouco mais volátil.

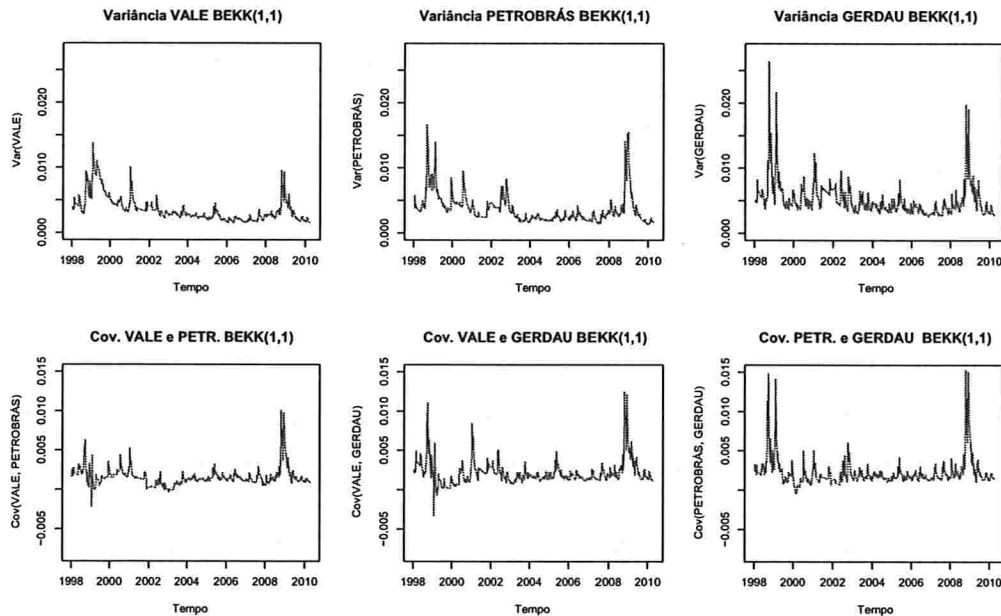


Figura 3.3: Variâncias e covariâncias condicionais estimadas segundo o modelo BEKK(1,1,1).

3.2.6 Estacionariedade na Covariância

Já foi provado anteriormente que o modelo EWMA, para a matriz de covariâncias condicional Σ_t , é não estacionário devido à matriz de covariâncias não condicional Σ não es-

tar definida. Foram propostos os modelos VEC(p,q) e BEKK(p,q,r) para considerar esta propriedade fundamental no contexto das séries temporais. Os seguintes resultados dão as condições sobre as matrizes de parâmetros, em cada modelo, para garantir a estacionariedade na covariância.

Seja o operador de diferenças L tal que $L^i w_t \equiv w_{t-i}$. Além disso, suponha que ε_t é uma sequência, $t = -\infty, \dots, 0, \dots, \infty$ e seja o processo geral para a volatilidade multivariada dado por

$$\mathbf{h}_t = \sum_{j=1}^{\infty} G(L)^{j-1} [\mathbf{c}_0 + A(L) \boldsymbol{\eta}_t], \quad (3.20)$$

em que $A(L)$ e $G(L)$ denotam polinômios em L , e os termos \mathbf{h}_t , \mathbf{c}_0 e $\boldsymbol{\eta}_t$ seguem a notação dada na Secção 3.2.2. A parametrização (3.20) contém os modelos VEC e BEKK, e para ver isto, note que (3.20) implica

$$\begin{aligned} \mathbf{h}_t &= \mathbf{c}_0 + A(L) \boldsymbol{\eta}_t + \sum_{j=2}^{\infty} G(L)^{j-1} [\mathbf{c}_0 + A(L) \boldsymbol{\eta}_t] \\ &= \mathbf{c}_0 + A(L) \boldsymbol{\eta}_t + G(L) \sum_{j=1}^{\infty} G(L)^{j-1} [\mathbf{c}_0 + A(L) \boldsymbol{\eta}_t] \\ &= \mathbf{c}_0 + A(L) \boldsymbol{\eta}_t + G(L) \mathbf{h}_t. \end{aligned} \quad (3.21)$$

Assim, definindo

$$A(L) = C_1 L + C_2 L^2 + \dots + C_p L^p$$

e

$$G(L) = G_1 L + G_2 L^2 + \dots + G_q L^q,$$

obtem-se o modelo VEC dado em (3.7). Considerando novamente a relação $\text{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\text{vec}(\mathbf{B})$ usada na Proposição (3.2), o modelo BEKK (3.18) é obtido definindo

$$A(L) = \sum_{l=1}^r (A_{1l} \otimes A_{1l}) L + \sum_{l=1}^r (A_{2l} \otimes A_{2l}) L^2 + \dots + \sum_{l=1}^r (A_{pl} \otimes A_{pl}) L^p$$

e

$$G(L) = \sum_{l=1}^r (B_{1l} \otimes B_{1l}) L + \sum_{l=1}^r (B_{2l} \otimes B_{2l}) L^2 + \dots + \sum_{l=1}^r (B_{ql} \otimes B_{ql}) L^q.$$

A seguinte proposição foi dada por Engle e Kroner (1995) e é uma generalização da condição de estacionariedade na variância nos modelos GARCH univariados.

Proposição 3.3. *Seja o processo $\{\varepsilon_t, t \in \mathcal{Z}\}$ e suponha que a equação (3.20) define um processo GARCH multivariado. Então, $\{\varepsilon_t\}$ é estacionário na covariância, se e somente se, todos os autovalores de $A(1) + G(1)$ são em módulo menores que 1.*

A prova desta proposição está no Apêndice A.

A proposição anterior implica que no modelo VEC, o processo $\{\varepsilon_t\}$ é estacionário na covariância se e somente se, os autovalores de $\sum_{i=1}^p C_i + \sum_{j=1}^q G_j$ são menores que 1 em módulo. Já para o modelo BEKK, o processo $\{\varepsilon_t\}$ é estacionário na covariância se e somente se os autovalores de $\sum_{l=1}^r \sum_{i=1}^p (A_{il} \otimes A_{il}) + \sum_{l=1}^r \sum_{j=1}^q (B_{jl} \otimes B_{jl})$ são menores que 1 em módulo. Na demonstração da Proposição 3.3, a matriz de covariâncias não condicional, quando esta existe, é dada por

$$E(\eta_t) = [I - A(1) - G(1)]^{-1} \mathbf{c}_0.$$

Isto implica que para o modelo VEC

$$E(\eta_t) = \left[I - \sum_{i=1}^p C_i - \sum_{j=1}^q G_j \right]^{-1} \mathbf{c}_0,$$

e para o modelo BEKK

$$E(\eta_t) = \left[I - \sum_{l=1}^r \sum_{i=1}^p (A_{il} \otimes A_{il}) - \sum_{l=1}^r \sum_{j=1}^q (B_{jl} \otimes B_{jl}) \right]^{-1} \text{vec}(\mathbf{A}_0 \mathbf{A}'_0).$$

3.2.7 Estimação dos Modelos

Seja $\Sigma_t(\theta)$ a matriz de covariâncias condicional do vetor \mathbf{r}_t , positiva semidefinida, de dimensão $(k \times k)$ e parametrizada pelo vetor θ . Por exemplo, para o modelo VEC dado em (3.7), o vetor de parâmetros é igual a $\theta' = [\mathbf{c}', \text{vec}(\mathbf{c}_0)', \text{vec}(\mathbf{C}_1)', \dots, \text{vec}(\mathbf{C}_p)', \text{vec}(\mathbf{G}_1)', \dots, \text{vec}(\mathbf{G}_q)']'$, em que \mathbf{c} é o vetor de médias dado em (3.1). Denotando $E_{t-1}[\cdot] = E[\cdot | \mathcal{F}_{t-1}]$, com \mathcal{F}_{t-1} o conjunto de informações passadas disponíveis até o tempo $t-1$, o modelo para a matriz de covariâncias condicional do processo $\{\varepsilon_t\}$, pode ser escrito como

$$E_{t-1}[\varepsilon_t] = \mathbf{0},$$

$$E_{t-1}[\varepsilon_t \varepsilon_t'] = \Sigma_t(\theta).$$

Em geral o vetor $\varepsilon_t = \mathbf{r}_t - \mathbf{c}$, segue uma distribuição normal multivariada ou uma distribuição t-student multivariada, cada uma com função densidade de probabilidade conjunta, dada

respectivamente por

$$f_N(\mathbf{r}_t) = \frac{1}{(2\pi)^{k/2} |\boldsymbol{\Sigma}_t|^{1/2}} \exp\{-[(\mathbf{r}_t - \mathbf{c})' \boldsymbol{\Sigma}_t^{-1} (\mathbf{r}_t - \mathbf{c})]/2\},$$

e

$$f_t(\mathbf{r}_t) = \frac{\Gamma[(\nu + k)/2]}{[\pi(\nu - 2)]^{k/2} \Gamma(\nu/2)} \frac{|\boldsymbol{\Sigma}_t|^{-1/2}}{[1 + (\mathbf{r}_t - \mathbf{c})' \boldsymbol{\Sigma}_t^{-1} (\mathbf{r}_t - \mathbf{c})/(\nu - 2)]^{(\nu+k)/2}},$$

em que ν denota o grau de liberdade. Assim, para uma amostra de tamanho T , tem-se que as funções de log-verossimilhança para as distribuições multivariadas normais e t-student são, respectivamente,

$$\log L_N(\boldsymbol{\theta}) = -\frac{kT}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log |\boldsymbol{\Sigma}_t| - \frac{1}{2} \sum_{t=1}^T (\mathbf{r}_t - \mathbf{c})' \boldsymbol{\Sigma}_t^{-1} (\mathbf{r}_t - \mathbf{c}),$$

e

$$\begin{aligned} \log L_t(\boldsymbol{\theta}) &= T \log \left[\Gamma \left(\frac{\nu + k}{2} \right) \right] - T \log \left[\Gamma \left(\frac{\nu}{2} \right) \right] - \frac{Tk}{2} \log[(\nu - 2)\pi] \\ &\quad - \sum_{t=1}^T \left\{ \frac{1}{2} \log |\boldsymbol{\Sigma}_t| \right\} - \frac{(\nu + k)}{2} \sum_{t=1}^T \left\{ \log \left[1 + \frac{(\mathbf{r}_t - \mathbf{c})' \boldsymbol{\Sigma}_t^{-1} (\mathbf{r}_t - \mathbf{c})}{(\nu - 2)} \right] \right\}. \end{aligned}$$

A ideia é encontrar o vetor $\boldsymbol{\theta}$ que maximize as funções de log-verossimilhança, mas não existe uma forma fechada e, portanto, é necessário usar um método iterativo como por exemplo o algoritmo BHHH (ver Berndt *et al.* (1974)), que também foi usado para estimar os modelos GARCH no caso univariado (ver, por exemplo, Bollerslev (1986)).

Uma forma alternativa para estimar os parâmetros quando a distribuição do vetor $\boldsymbol{\varepsilon}_t$ não é conhecida é usar o método de quase-verossimilhança (Hafner e Herwartz (2003)). Igual ao caso anterior, suponha que deseja-se estimar um vetor de parâmetros desconhecidos $\boldsymbol{\theta}_0$ baseado numa amostra dada de T observações. O método de quase máxima verossimilhança (QMV) estima $\boldsymbol{\theta}_0$ maximizando a função de log-verossimilhança $L(\boldsymbol{\theta}) = \sum_{t=1}^T l_t(\boldsymbol{\theta})$ com

$$l_t(\boldsymbol{\theta}) = -\frac{k}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Sigma}_t(\boldsymbol{\theta})| - \frac{1}{2} (\mathbf{r}_t - \mathbf{c})' \boldsymbol{\Sigma}_t(\boldsymbol{\theta})^{-1} (\mathbf{r}_t - \mathbf{c}). \quad (3.22)$$

Comte e Lieberman (2003) dão condições para que o estimador QMV, $\widehat{\boldsymbol{\theta}}$, seja consistente e com distribuição assintoticamente normal, embora as observações não sejam normais. A distribuição assintótica é dada por

$$\sqrt{T}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{D} \mathbf{N}(\mathbf{0}, \mathcal{J}^{-1} \mathcal{I} \mathcal{J}^{-1}),$$

e

$$\mathcal{I} = E \left[\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \middle| \boldsymbol{\theta}_0 \right], \quad \mathcal{J} = -E \left[\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \middle| \boldsymbol{\theta}_0 \right].$$

A matriz \mathcal{I} é conhecida como a matriz de informação e \mathcal{J} é a esperança negativa do Hessiano avaliada em $\boldsymbol{\theta}_0$. Se o processo $\{\mathbf{r}_t\}$ é condicionalmente Gaussiano, $\mathcal{I} = \mathcal{J}$ e a matriz de covariâncias assintótica é \mathcal{I}^{-1} , que é a cota inferior de Cramer-Rao.

Para fazer inferência em relação ao estimador $\hat{\boldsymbol{\theta}}$, é necessário calcular o vetor score e o Hessiano. Comte e Lieberman (2003) mostraram que

$$\frac{\partial l_t(\boldsymbol{\theta})}{\partial \theta_i} = \text{tr} \left[\dot{\Sigma}_{t,i} \Sigma_t^{-1} - \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Sigma_t^{-1} \dot{\Sigma}_{t,i} \Sigma_t^{-1} \right]$$

e

$$\begin{aligned} \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} = & \text{tr} \left[\ddot{\Sigma}_{t,i,j} \Sigma_t^{-1} - \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Sigma_t^{-1} \ddot{\Sigma}_{t,i,j} \Sigma_t^{-1} - \dot{\Sigma}_{t,i} \Sigma_t^{-1} \dot{\Sigma}_{t,j} \Sigma_t^{-1} \right. \\ & \left. + \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Sigma_t^{-1} \dot{\Sigma}_{t,j} \Sigma_t^{-1} \dot{\Sigma}_{t,i} \Sigma_t^{-1} + \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Sigma_t^{-1} \dot{\Sigma}_{t,i} \Sigma_t^{-1} \dot{\Sigma}_{t,j} \Sigma_t^{-1} \right], \end{aligned} \quad (3.23)$$

em que $\text{tr}(\cdot)$ denota o traço de uma matriz e

$$\dot{\Sigma}_{t,i} = \frac{\partial \Sigma_t}{\partial \theta_i}, \quad \ddot{\Sigma}_{t,i,j} = \frac{\partial^2 \Sigma_t}{\partial \theta_i \partial \theta_j},$$

todas Σ_t , $\dot{\Sigma}_{t,i}$ e $\ddot{\Sigma}_{t,i,j}$ são derivadas em $\boldsymbol{\theta}$.

Note que pela consistência do estimador QMV $\hat{\boldsymbol{\theta}}$, as matrizes \mathcal{I} e \mathcal{J} podem ser consistentemente estimadas por

$$\hat{\mathcal{I}} = \frac{1}{T} \sum_{t=1}^T \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \middle|_{\hat{\boldsymbol{\theta}}} \quad (3.24)$$

e

$$\hat{\mathcal{J}} = -\frac{1}{T} \sum_{t=1}^T \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \middle|_{\hat{\boldsymbol{\theta}}}. \quad (3.25)$$

Pela definição de Σ_t , $\Sigma_t(\boldsymbol{\theta}_0) = E_{t-1}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$ e aplicando a esperança condicional em (3.23), tem-se que

$$\begin{aligned}
M_{t,ij}(\boldsymbol{\theta}_0) &:= E \left[\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \Big|_{\boldsymbol{\theta}_0} \right] \\
&= \text{tr} \left[\ddot{\Sigma}_{t,i,j} \Sigma_t^{-1} - E_{t-1}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') \Sigma_t^{-1} \ddot{\Sigma}_{t,i,j} \Sigma_t^{-1} - \dot{\Sigma}_{t,i} \Sigma_t^{-1} \dot{\Sigma}_{t,j} \Sigma_t^{-1} \right. \\
&\quad \left. + E_{t-1}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') \Sigma_t^{-1} \dot{\Sigma}_{t,j} \Sigma_t^{-1} \dot{\Sigma}_{t,i} \Sigma_t^{-1} + E_{t-1}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') \Sigma_t^{-1} \dot{\Sigma}_{t,i} \Sigma_t^{-1} \dot{\Sigma}_{t,j} \Sigma_t^{-1} \right].
\end{aligned}$$

Portanto,

$$\begin{aligned}
M_{t,ij}(\boldsymbol{\theta}_0) &= \text{tr} \left[\ddot{\Sigma}_{t,i,j} \Sigma_t^{-1} - \Sigma_t \Sigma_t^{-1} \ddot{\Sigma}_{t,i,j} \Sigma_t^{-1} - \dot{\Sigma}_{t,i} \Sigma_t^{-1} \dot{\Sigma}_{t,j} \Sigma_t^{-1} \right. \\
&\quad \left. + \Sigma_t \Sigma_t^{-1} \dot{\Sigma}_{t,j} \Sigma_t^{-1} \dot{\Sigma}_{t,i} \Sigma_t^{-1} + \Sigma_t \Sigma_t^{-1} \dot{\Sigma}_{t,i} \Sigma_t^{-1} \dot{\Sigma}_{t,j} \Sigma_t^{-1} \right] \\
&= \text{tr} \left[\ddot{\Sigma}_{t,i,j} \Sigma_t^{-1} - \ddot{\Sigma}_{t,i,j} \Sigma_t^{-1} - \dot{\Sigma}_{t,i} \Sigma_t^{-1} \dot{\Sigma}_{t,j} \Sigma_t^{-1} \right. \\
&\quad \left. + \dot{\Sigma}_{t,j} \Sigma_t^{-1} \dot{\Sigma}_{t,i} \Sigma_t^{-1} + \dot{\Sigma}_{t,i} \Sigma_t^{-1} \dot{\Sigma}_{t,j} \Sigma_t^{-1} \right] \\
&= \text{tr} \left[\dot{\Sigma}_{t,j} \Sigma_t^{-1} \dot{\Sigma}_{t,i} \Sigma_t^{-1} \right] \\
&= \text{tr} \left[\dot{\Sigma}_{t,i} \Sigma_t^{-1} \dot{\Sigma}_{t,j} \Sigma_t^{-1} \right], \quad (\text{já que, } \text{tr}[\mathbf{AB}] = \text{tr}[\mathbf{BA}]).
\end{aligned}$$

Pela regra da esperança da esperança condicional tem-se que $\mathcal{J} = -E[M_t(\boldsymbol{\theta}_0)]$, portanto a estimação de \mathcal{J} é mais simples substituindo o parâmetro verdadeiro $\boldsymbol{\theta}_0$ em $M_t(\boldsymbol{\theta}_0)$ pelo estimador QMV, i.e.

$$\tilde{\mathcal{J}} = -\frac{1}{T} \sum_{t=1}^T M_t(\hat{\boldsymbol{\theta}}).$$

O estimador $\tilde{\mathcal{J}}$ evita o cálculo das derivadas de segunda ordem de Σ_t e portanto é mais fácil de implementar que o estimador $\hat{\mathcal{J}}$. Os dois estimadores são assintoticamente equivalentes, e por isso é de esperar que eles tenham um mesmo bom comportamento para grandes amostras⁵.

Na prática o estimador de máxima verossimilhança de $\boldsymbol{\theta}_0$ é calculado usando um algoritmo iterativo como

$$\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^{(i)} + \alpha_i \left(\sum_{t=1}^T \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right) \sum_{t=1}^T \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}},$$

⁵Para o comportamento destes estimadores em pequenas amostras ver Hafner e Herwartz (2003)

em que α_i é uma variável que depende do passo, escolhida para maximizar a função de verossimilhança numa direção dada. Este algoritmo é conhecido como o algoritmo BHHH e foi desenvolvido por Berndt *et al.* (1974).

3.2.8 Testes para os modelos GARCH multivariados

Para obter a representação VARMA (ARMA vetorial) do processo GARCH multivariado geral dado em (3.21) considera-se que $A(L) = \mathbb{A}_1L + \mathbb{A}_2L^2 + \dots + \mathbb{A}_pL^p$ e $G(L) = \mathbb{G}_1L + \mathbb{G}_2L^2 + \dots + \mathbb{G}_qL^q$, em que \mathbb{A}_i e \mathbb{G}_j são matrizes quadráticas de parâmetros, por exemplo, para gerar o modelo VEC dado em (3.7), tome $\mathbb{A}_i = \mathbf{C}_i$ e $\mathbb{G}_j = \mathbf{G}_j$, e para gerar o modelo BEKK dado em (3.18), tome $\mathbb{A}_i = \sum_{l=1}^r (A_{il} \otimes A_{il})$ e $\mathbb{G}_j = \sum_{l=1}^r (B_{jl} \otimes B_{jl})$. Assim, definindo $\mathbf{v}_t := \boldsymbol{\eta}_t - \mathbf{h}_t$ e substituindo $\boldsymbol{\eta}_t - \mathbf{v}_t$ por \mathbf{h}_t , o modelo (3.21) pode ser escrito como

$$\boldsymbol{\eta}_t = \mathbf{c}_0 + \sum_{i=1}^M (\mathbb{A}_i + \mathbb{G}_i) \boldsymbol{\eta}_{t-i} + \mathbf{v}_t - \sum_{j=1}^q \mathbb{G}_j \mathbf{v}_{t-j}, \quad (3.26)$$

com $M = \max\{p, q\}$, $\mathbb{A}_j = \mathbf{0}$ para $j > p$ e $\mathbb{G}_j = \mathbf{0}$ para $j > q$. Tem-se a representação VARMA de $\boldsymbol{\eta}_t = \text{vec}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$.

Teste Portmanteu

Os testes mais conhecidos para detectar efeitos ARCH são os testes *portmanteu* Box-Pierce/Ljung-Box. Dada a representação VARMA (3.26) dos modelos GARCH multivariados, o teste multivariado de portmanteu, dado por Hosking (1980), pode ser aplicado a $\boldsymbol{\eta}_t = \text{vec}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$ para testar se existem efeitos ARCH em $\boldsymbol{\varepsilon}_t$, ou seja, testar se o processo $\boldsymbol{\eta}_t$ tem correlação serial. As hipóteses a serem testadas são:

H_0 : a série multivariada $\boldsymbol{\eta}_t$ é ruído branco,

H_a : a série $\boldsymbol{\eta}_t$ não é ruído branco.

A estatística do teste é

$$Q(h) := T \sum_{i=1}^h \text{tr}(\widehat{\boldsymbol{\Gamma}}(i)' \widehat{\boldsymbol{\Gamma}}(0)^{-1} \widehat{\boldsymbol{\Gamma}}(i) \widehat{\boldsymbol{\Gamma}}(0)^{-1}),$$

em que T é o tamanho da amostra. A versão modificada é dada por

$$\overline{Q}(h) := T^2 \sum_{i=1}^h (T-i)^{-1} \text{tr}(\widehat{\boldsymbol{\Gamma}}(i)' \widehat{\boldsymbol{\Gamma}}(0)^{-1} \widehat{\boldsymbol{\Gamma}}(i) \widehat{\boldsymbol{\Gamma}}(0)^{-1}),$$

em que $\widehat{\Gamma}(i) = T^{-1} \sum_{t=i+1}^T (\boldsymbol{\eta}_t - \bar{\boldsymbol{\eta}})(\boldsymbol{\eta}_{t-i} - \bar{\boldsymbol{\eta}})'$ ($i = 0, 1, \dots, h$).

A distribuição assintótica destas estatísticas, sob a hipótese nula H_0 , é $\chi^2(k^2(h - q))$, em que k é o número de séries; para os detalhes ver Lütkepohl (2006). O valor h é o número de lags a serem testados e deve ser maior que q , para que o número de graus de liberdade seja positivo.

Multiplicadores de Lagrange

Um outro teste multivariado para verificar se existem efeitos ARCH nos resíduos é o *teste LM* (Multiplicadores de Lagrange). A ideia é considerar o modelo auxiliar

$$\boldsymbol{\eta}_t = \boldsymbol{\beta}_0 + \boldsymbol{\Phi}_1 \boldsymbol{\eta}_{t-1} + \dots + \boldsymbol{\Phi}_p \boldsymbol{\eta}_{t-p} + \text{erro}_t, \quad (3.27)$$

em que $\boldsymbol{\beta}_0$ é um vetor de dimensão $(k(k + 1)/2 \times 1)$, e $\boldsymbol{\Phi}_i$ são matrizes de parâmetros de dimensões $(k(k + 1)/2 \times k(k + 1)/2)$, com $i = 1, \dots, p$. Se todas as matrizes $\boldsymbol{\Phi}_i$ são nulas, então não existem efeitos ARCH nos resíduos. Portanto, as hipóteses são

$$H_0 : \boldsymbol{\Phi}_1 = \boldsymbol{\Phi}_2 = \dots = \boldsymbol{\Phi}_p = \mathbf{0} \quad \text{versus} \quad H_a : \boldsymbol{\Phi}_i \neq \mathbf{0} \quad \text{para pelo menos um valor de } i.$$

Denotando por $\widehat{\boldsymbol{\Sigma}}_{res}$ o estimador da matriz de covariâncias residual, baseado no modelo (3.27), e por $\widehat{\boldsymbol{\Sigma}}_0$ a matriz de covariâncias sob a hipótese nula, a estatística *LM* é da forma

$$LM(p) = \frac{1}{2} T k(k + 1) - T \text{tr}(\widehat{\boldsymbol{\Sigma}}_{res} \widehat{\boldsymbol{\Sigma}}_0^{-1}),$$

e tem, sob a hipótese nula, uma distribuição assintótica $\chi^2(pk^2(k + 1)^2/4)$ (para os detalhes ver Doornik e Hendry (1997), Sec. 10.9.2.4).

Na equação (3.27), cada uma das matrizes $\boldsymbol{\Phi}_i$ tem dimensões $(k(k + 1)/2 \times k(k + 1)/2)$ e portanto, o modelo auxiliar contém um número muito grande de parâmetros, mesmo o valor de p e o número de séries k sejam pequenos. Assim, o teste não funciona muito bem, a menos que o tamanho amostral seja grande. Porém, é possível aplicar o teste a cada uma das séries individualmente.

Métodos Alternativos

Existem métodos alternativos para testar se o número de lags p é adequado. Considera-se

o modelo auxiliar (3.27), para $p = 1, \dots, l$, isto é

$$\begin{aligned}\eta_t &= \beta_{1,0} + \Phi_{1,1}\eta_{t-1} + \mathbf{a}_{1,t}, \\ \eta_t &= \beta_{2,0} + \Phi_{2,1}\eta_{t-1} + \Phi_{2,2}\eta_{t-2} + \mathbf{a}_{2,t}, \\ &\vdots \\ \eta_t &= \beta_{l,0} + \Phi_{l,1}\eta_{t-1} + \dots + \Phi_{l,l}\eta_{t-l} + \mathbf{a}_{l,t},\end{aligned}$$

em que $\mathbf{a}_{l,t}$ denota o vetor de resíduos do modelo. Testamos

$$H_0 : \Phi_{l,l} = \mathbf{0} \quad \text{versus} \quad H_a : \Phi_{l,l} \neq \mathbf{0}, \quad l = 1, 2, \dots \quad (3.28)$$

O teste da razão de verossimilhanças é baseado nas estimativas das matrizes de covariâncias dos resíduos dos modelos ajustados. Para a l -ésima equação, considerando

$$\hat{\mathbf{a}}_{l,t} = \eta_t - \hat{\beta}_{l,0} - \hat{\Phi}_{l,1}\eta_{t-1} - \dots - \hat{\Phi}_{l,l}\eta_{t-l},$$

a matriz de covariâncias dos resíduos, é estimada por

$$\hat{\Sigma}_l = \frac{1}{T-l} \sum_{t=l+1}^T \hat{\mathbf{a}}_{l,t}(\hat{\mathbf{a}}_{l,t})', \quad l \geq 0,$$

em que define-se para $l = 0$, $\hat{\mathbf{a}}_{0,t} = \eta_t - \bar{\eta}$. A estatística da razão de verossimilhança para o teste (3.28) é dada por

$$RV(l) = (T-l) \ln \frac{|\hat{\Sigma}_{l-1}|}{|\hat{\Sigma}_l|},$$

que tem distribuição $\chi^2(k^2)$ sob H_0 .

Outra forma alternativa de identificar a ordem p do modelo GARCH multivariado é usar algum critério de informação, como

$$\begin{aligned}AIC(l) &= \ln(|\hat{\Sigma}_l|) + 2lk^2/T \quad (\text{Akaike}), \\ BIC(l) &= \ln(|\hat{\Sigma}_l|) + lk^2 \ln(T)/T \quad (\text{Schwarz}), \\ HQC(l) &= \ln(|\hat{\Sigma}_l|) + lk^2 \ln(\ln(T))/T \quad (\text{Hannan-Quinn}).\end{aligned}$$

Exemplo 3.4. AIC e BIC dos modelos EWMA, DVEC(1,1) e BEKK(1,1,1) aplicados ao conjunto de retornos das séries VALE, PETROBRAS e GERDAU, estão apresentados na

Tabela 3.4.

Tabela 3.4: *Critérios AIC e BIC.*

Modelo	AIC	BIC
EWMA	-5209,864	-5192,482
DVEC(1,1)	-5244,584	-5153,326
BEKK(1,1,1)	-5243,592	-5126,26

Pela Tabela 3.4, o modelo DVEC(1,1) é o mais adequado segundo o critério AIC e segundo o critério BIC o melhor modelo é o EWMA.

Os testes multivariados Portmanteau para o quadrado dos resíduos em cada ajuste estão apresentados na Tabela 3.5 e segundo os p -valores o modelo mais adequado é o DVEC(1,1).

Tabela 3.5: *Teste de portmanteau para o quadrado dos resíduos*

Teste Multivariado Portmanteau: Tipo Ljung-Box		
Hipótese Nula: não correlação serial		
Modelo	Test Stat	p.value
EWMA	193,4817	<0,001
DVEC(1,1)	125,8185	0,116
BEKK(1,1,1)	147,6508	0,007
Dist. sob Hip. Nula: qui-quadrado com 108 graus de liberdade		
Total Observ.: 570		

A Tabela 3.6 apresenta os testes de Normalidade, Ljung-Box univariados e Multiplicadores de Lagrange para os resíduos do modelo DVEC(1,1). Pode-se notar que segundo os p -valores do teste de Shapiro-Wilk as séries VALE e GERDAU seguem uma distribuição normal ao nível de significância de 5%. Os testes de Ljung-Box para os resíduos e os quadrados dos resíduos não rejeitam a hipótese de não autocorrelação serial para cada série de retornos, e este resultado é confirmado pelos testes Multiplicadores de Lagrange.

A Figura 3.4 apresenta as FAC e FAC cruzada dos quadrados dos resíduos do modelo ajustado DVEC(1,1), e observa-se que não existe autocorrelação e autocorrelação cruzada significativa entre eles antes do lag 10.

Tabela 3.6: Testes do ajuste DVEC(1,1)

Testes DVEC(1,1):						
Testes de Normalidade:						
	Jarque-Bera	P-valor	Shapiro-Wilk	P-valor		
VALE	94,793	0,0000	0,9855	0,5280		
PETROBRAS	132,181	0,0000	0,9789	0,0198		
GERDAU	4,589	0,1008	0,9899	0,9350		
Teste Ljung-Box para os resíduos padronizados:						
	Estatística	P-valor	$\chi^2 - d, f,$			
VALE	7,396	0,8304	12			
PETROBRAS	72,630	0,3965	12			
GERDAU	9,249	0,6815	12			
Teste Ljung-Box para o quadrado dos resíduos padronizados:						
	Estatística	P-valor	$\chi^2 - d, f,$			
VALE	5,630	0,9336	12			
PETROBRAS	4,226	0,9790	12			
GERDAU	10,319	0,5880	12			
Teste Multiplicadores de Lagrange:						
	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
VALE	-0,3127	0,0121	-0,57012	-0,5917	1,88010	0,5647
PETROBRAS	0,4578	0,5570	0,04308	0,6768	0,44552	-0,5148
GERDAU	-0,8488	-1,0658	1,08945	-0,5609	-0,03552	0,5952
	Lag 7	Lag 8	Lag 9	Lag 10	Lag 11	Lag 12
VALE	-0,5304	-0,2503	-0,02358	-0,5490	-0,91280	-0,17605
PETROBRAS	-0,5304	-0,2503	0,12863	-0,8433	0,09744	0,09767
GERDAU	0,3640	-1,1492	1,56776	-1,3190	-0,66656	-0,05155
	C					
VALE	0,0625					
PETROBRAS	-0,6511					
GERDAU	1,8960					
	TR^2	P-valor	F-estat	P-valor		
VALE	6,268	0,9020	0,5763	0,9374		
PETROBRAS	4,145	0,9807	0,3796	0,9973		
GERDAU	12,050	0,4417	1,1196	0,4531		

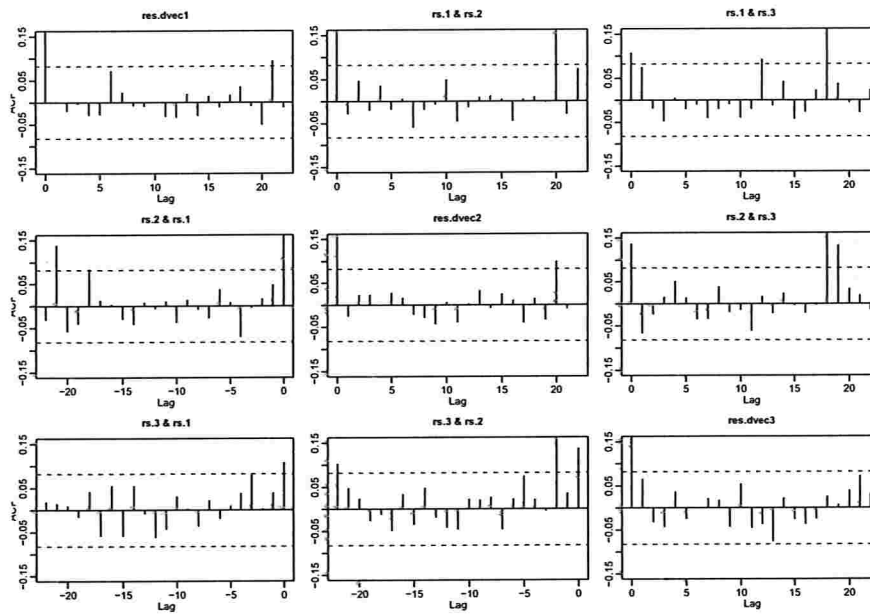


Figura 3.4: *FAC e FAC cruzadas dos resíduos do ajuste DVEC(1,1)*

3.2.9 Previsão dos modelos GARCH multivariados

Considera-se novamente o modelo GARCH multivariado geral dado em (3.21),

$$\mathbf{h}_t = \mathbf{c}_0 + A(L)\boldsymbol{\eta}_t + G(L)\mathbf{h}_t,$$

e define-se como na Secção 3.2.8 os polinômios em L , $A(L) = \mathbb{A}_1L + \mathbb{A}_2L^2 + \dots + \mathbb{A}_pL^p$ e $G(L) = \mathbb{G}_1L + \mathbb{G}_2L^2 + \dots + \mathbb{G}_qL^q$.

A previsão um passo à frente para \mathbf{h}_t , $\widehat{\mathbf{h}}_t(1)$, é dada por

$$\begin{aligned}
\widehat{\mathbf{h}}_t(1) &= E [\mathbf{h}_{t+1} | \mathcal{F}_t] \\
&= E [\mathbf{c}_0 + A(L)\boldsymbol{\eta}_{t+1} + G(L)\mathbf{h}_{t+1} | \mathcal{F}_t] \\
&= \mathbf{c}_0 + \mathbb{A}_1 E [\boldsymbol{\eta}_t | \mathcal{F}_t] + \mathbb{A}_2 E [\boldsymbol{\eta}_{t-1} | \mathcal{F}_t] + \dots + \mathbb{A}_p E [\boldsymbol{\eta}_{t-(p-1)} | \mathcal{F}_t] \\
&\quad + \mathbb{G}_1 E [\mathbf{h}_t | \mathcal{F}_t] + \mathbb{G}_2 E [\mathbf{h}_{t-1} | \mathcal{F}_t] + \dots + \mathbb{G}_q E [\mathbf{h}_{t-(q-1)} | \mathcal{F}_t] \\
&= \mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_t + \mathbb{A}_2 \boldsymbol{\eta}_{t-1} + \dots + \mathbb{A}_p \boldsymbol{\eta}_{t-p+1} + \mathbb{G}_1 \mathbf{h}_t + \mathbb{G}_2 \mathbf{h}_{t-1} + \dots + \mathbb{G}_q \mathbf{h}_{t-q+1}.
\end{aligned}$$

A previsão dois passos à frente é dada por:

$$\begin{aligned}
\widehat{\mathbf{h}}_t(2) &= E [\mathbf{h}_{t+2} | \mathcal{F}_t] \\
&= E [\mathbf{c}_0 + A(L)\boldsymbol{\eta}_{t+2} + G(L)\mathbf{h}_{t+2} | \mathcal{F}_t] \\
&= \mathbf{c}_0 + \mathbb{A}_1 E [\boldsymbol{\eta}_{t+1} | \mathcal{F}_t] + \mathbb{A}_2 E [\boldsymbol{\eta}_t | \mathcal{F}_t] + \dots + \mathbb{A}_p E [\boldsymbol{\eta}_{t-(p-2)} | \mathcal{F}_t] \\
&\quad + \mathbb{G}_1 E [\mathbf{h}_{t+1} | \mathcal{F}_t] + \mathbb{G}_2 E [\mathbf{h}_t | \mathcal{F}_t] + \dots + \mathbb{G}_q E [\mathbf{h}_{t-(q-2)} | \mathcal{F}_t] \\
&= \mathbf{c}_0 + \mathbb{A}_1 \widehat{\mathbf{h}}_t(1) + \mathbb{A}_2 \boldsymbol{\eta}_t + \dots + \mathbb{A}_p \boldsymbol{\eta}_{t-p+2} + \mathbb{G}_1 \widehat{\mathbf{h}}_t(1) + \mathbb{G}_2 \mathbf{h}_t + \dots + \mathbb{G}_q \mathbf{h}_{t-q+2} \\
&= \mathbf{c}_0 + (\mathbb{A}_1 + \mathbb{G}_1) \widehat{\mathbf{h}}_t(1) + \mathbb{A}_2 \boldsymbol{\eta}_t + \dots + \mathbb{A}_p \boldsymbol{\eta}_{t-p+2} + \mathbb{G}_2 \mathbf{h}_t + \dots + \mathbb{G}_q \mathbf{h}_{t-q+2}.
\end{aligned}$$

Seja $M = \max\{p, q\}$, a previsão l passos à frente, para $l \leq M$ é dada por

$$\begin{aligned}\widehat{\mathbf{h}}_t(l) &= E [\mathbf{h}_{t+l} | \mathcal{F}_t] \\ &= E [\mathbf{c}_0 + A(L)\boldsymbol{\eta}_{t+l} + G(L)\mathbf{h}_{t+l} | \mathcal{F}_t] \\ &= \mathbf{c}_0 + (\mathbb{A}_1 + \mathbb{G}_1)\widehat{\mathbf{h}}_t(1) + (\mathbb{A}_2 + \mathbb{G}_2)\widehat{\mathbf{h}}_t(2) + \dots + (\mathbb{A}_{l-1} + \mathbb{G}_{l-1})\widehat{\mathbf{h}}_t(l-1) \\ &\quad + \mathbb{A}_l\boldsymbol{\eta}_t + \dots + \mathbb{A}_M\boldsymbol{\eta}_{t-M+l} + \mathbb{G}_l\mathbf{h}_t + \dots + \mathbb{G}_M\mathbf{h}_{t-M+l},\end{aligned}$$

em que $\mathbb{A}_i = \mathbf{0}_{k \times k}$, para $p < i \leq M$ e $\mathbb{G}_j = \mathbf{0}_{k \times k}$, para $q < j \leq M$.

Agora, para $M < l$, tem-se que

$$\begin{aligned}\widehat{\mathbf{h}}_t(l) &= E [\mathbf{h}_{t+l} | \mathcal{F}_t] \\ &= E [\mathbf{c}_0 + A(L)\boldsymbol{\eta}_{t+l} + G(L)\mathbf{h}_{t+l} | \mathcal{F}_t] \\ &= \mathbf{c}_0 + (\mathbb{A}_1 + \mathbb{G}_1)\widehat{\mathbf{h}}_t(1) + (\mathbb{A}_2 + \mathbb{G}_2)\widehat{\mathbf{h}}_t(2) + \dots + (\mathbb{A}_{l-1} + \mathbb{G}_{l-1})\widehat{\mathbf{h}}_t(l-1),\end{aligned}$$

em que $\mathbb{A}_i = \mathbf{0}_{k \times k}$ para $p < i \leq M$, e $\mathbb{G}_j = \mathbf{0}_{k \times k}$ para $q < j \leq M$.

Com $\mathbb{A}_i = \mathbf{C}_i$ ($i = 1, \dots, p$) e $\mathbb{G}_j = \mathbf{G}_j$ ($j = 1, \dots, q$) obtêm-se a previsão para o modelo VEC, e com $\mathbb{A}_i = \sum_{l=1}^r \mathbf{A}_{il} \otimes \mathbf{A}_{il}$, $i = 1, \dots, p$, e $\mathbb{G}_j = \sum_{l=1}^r \mathbf{B}_{jl} \otimes \mathbf{B}_{jl}$, para $j = 1, \dots, q$, obtêm-se a previsão para o modelo BEKK.

Exemplo 3.5. Apresentam-se na Figura 3.5 as previsões dos desvios padrão e das correlações 10 semanas à frente do modelo ajustado DVEC(1,1) dado por (3.17), para os retornos do conjunto de séries VALE, PETROBRAS e GERDAU.

As previsões estão depois da linha vertical pontilhada e nota-se que os desvios padrão condicionais estimados têm uma tendência crescente e as correlações condicionais estimadas uma tendência decrescente.

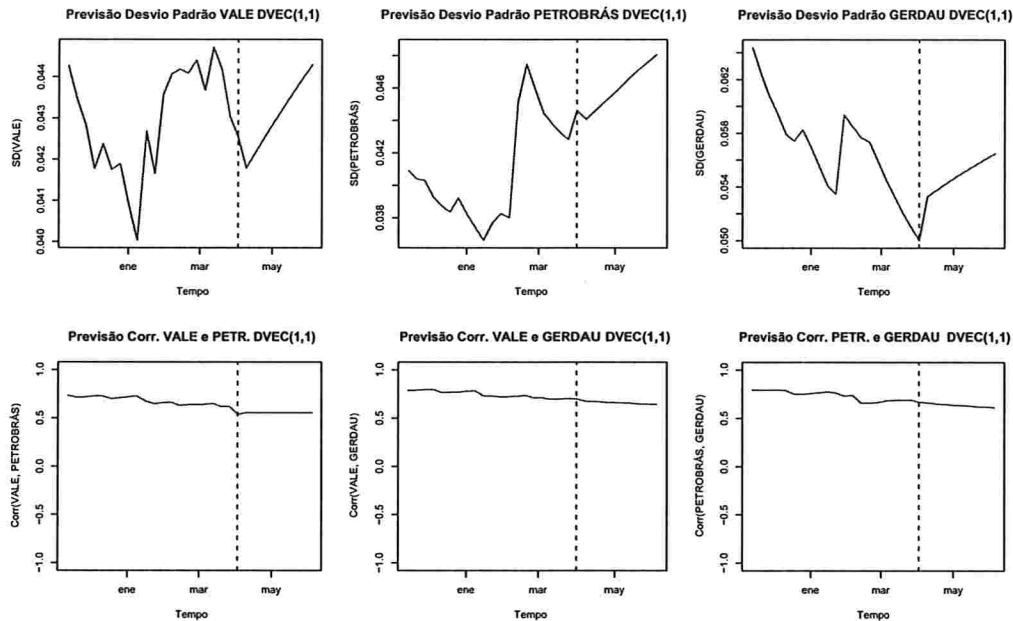


Figura 3.5: *Previsão 10 semanas à frente DVEC(1,1)*

3.3 Modelos GARCH Multivariados Indiretos

Até agora foram apresentados modelos para a matriz de covariâncias condicional, Σ_t , considerando uma abordagem direta. Uma das desvantagens destes modelos é a quantidade de parâmetros a serem estimados simultaneamente e que aumenta consideravelmente com o número de séries. Computacionalmente existem muitos problemas em relação à convergência dos algoritmos usados para estimar os modelos GARCH diretos, além da impossibilidade de suportar uma quantidade muito grande de séries. Para tentar resolver este problema foram desenvolvidos alguns modelos como por exemplo o modelo de Correlação Condicional Constante e o modelo GARCH Ortogonal ou de Componentes Principais, que serão estudados nesta seção.

Estes modelos estão baseados principalmente em modelos GARCH univariados e portanto os ajustes não são muito complicados, mesmo que o número de séries seja muito grande. A estacionariedade, a estimação dos parâmetros, os testes para avaliar a qualidade de ajuste e a previsão, estão baseadas nos modelos GARCH univariados⁶.

⁶Para as propriedades dos modelos GARCH univariados consultar Bollerslev (1986).

3.3.1 Modelo de Correlação Condicional Constante

Em geral, uma matriz de covariâncias Σ de dimensões $(k \times k)$, pode ser decomposta da seguinte forma

$$\Sigma = \Delta \mathbf{R} \Delta,$$

em que \mathbf{R} é a matriz de correlação, Δ é uma matriz diagonal com os elementos $\sigma_1, \dots, \sigma_k$ na diagonal, e σ_i é o desvio padrão da i -ésima série. Bollerslev (1990), com o intuito de estudar as matrizes de covariâncias condicionais do dólar americano contra o marco alemão, o marco francês e a lira italiana, introduziu o modelo conhecido como o modelo de Correlação Condicional Constante ("CCC"), assumindo que a matriz de correlações condicionais é constante no tempo. Assim, o modelo que sugeriu para a matriz Σ_t foi

$$\Sigma_t = \Delta_t \mathbf{R} \Delta_t,$$

em que \mathbf{R} é a matriz de correlação condicional constante, e Δ_t é a seguinte matriz diagonal

$$\Delta_t = \begin{bmatrix} \sigma_{1t} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \sigma_{kt} \end{bmatrix}, \quad (3.29)$$

em que σ_{it}^2 segue um processo GARCH univariado, para $i = 1, \dots, k$. Nesta situação o processo de estimação pode ser dividido em duas fases: na primeira estimam-se modelos da família GARCH univariados para cada série de retornos individualmente e obtêm-se estimativas das variâncias condicionais; na segunda, \mathbf{R} pode ser estimada considerando a matriz de correlação não condicional das séries padronizadas.

Considerando a suposição de correlação condicional constante, o número de parâmetros a serem estimados é muito menor que para os modelos GARCH multivariados diretos. Por exemplo, se cada variância condicional segue um processo GARCH(1,1) univariado, isto é,

$$\sigma_{it}^2 = \omega_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2 \quad i = 1, \dots, k,$$

este modelo contém $k(k+1)/2 - k + 3k = k(k+5)/2$ parâmetros (já que \mathbf{R} contém $k(k+1)/2 - k$ parâmetros e os modelos GARCH(1,1) contêm $3k$ parâmetros), comparado com os $k(5k+1)/2$ parâmetros do modelo BEKK(1,1,1) a serem estimados simultaneamente. A matriz Σ_t é positiva definida se e somente se \mathbf{R} é positiva definida.

Embora o modelo CCC seja fácil de ajustar, em muitas situações reais não é possível

considerar que a matriz de correlações seja constante. A generalização do modelo CCC é chamada modelo de Correlação Condicional Dinâmica ("DCC"), em que a matriz \mathbf{R} varia com o tempo. Este último modelo não será considerado neste trabalho, mas pode ser consultado, por exemplo, em Christodoulakis e Satchell (2002), Engle (2002) e Tse e Tsui (2002). Um outro modelo baseado em modelos GARCH univariados e que não tem uma restrição tão forte como é que a matriz de correlação condicional seja constante, é o modelo GARCH Ortogonal que será apresentado a seguir.

Exemplo 3.6. Considera-se um ajuste de Correlação Condicional Constante normalmente distribuído para os retornos do grupo de séries VALE, PETROBRAS e GERDAU, fazendo para cada um dos elementos da diagonal da matriz dada por (3.29) um ajuste GARCH(1,1) univariado.

A Tabela 3.7 apresenta os parâmetros estimados para o modelo CCC-GARCH(1,1). Todos os p -valores são significantes o nível de significância de 5%. Também a matriz de correlação constante é apresentada junto com seus desvios padrão.

Os critérios AIC e BIC são, respectivamente -5195,121 e -5142,974 e o p -valor do teste multivariado de Portmanteau para o quadrados dos resíduos foi 0,1392, não rejeitando a hipótese nula de autocorrelação e autocorrelação não significativa.

O modelo CCC-GARCH(1,1) ajustado para o vetor de retornos \mathbf{r}_t (3×1), é dado por

$$\mathbf{r}_t = \begin{pmatrix} 0,00886708 \\ 0,00659223 \\ 0,01033374 \end{pmatrix} + \boldsymbol{\varepsilon}_t, \quad (3.30)$$

$$\boldsymbol{\Sigma}_t = \boldsymbol{\Delta}_t \mathbf{R} \boldsymbol{\Delta}_t,$$

com

$$\boldsymbol{\Delta}_t = \begin{pmatrix} \sigma_{1t} & 0 & 0 \\ 0 & \sigma_{2t} & 0 \\ 0 & 0 & \sigma_{3t} \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} 1,0000 & 0,4796 & 0,5166 \\ 0,4796 & 1,0000 & 0,5512 \\ 0,5166 & 0,5512 & 1,0000 \end{pmatrix},$$

em que σ_{it} , para $i = 1, 2, 3$, é o desvio padrão condicional da i -ésima série de retornos e além disso, $\boldsymbol{\varepsilon}_t \sim N_3(\mathbf{0}, \boldsymbol{\Sigma}_t)$. Portanto, segundo a Tabela 3.7 tem-se que as variâncias seguem os

Tabela 3.7: *Resumo ajuste CCC-GARCH(1,1).*

Parâmetros Estimados CCC-GARCH(1,1):				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	0,00886708	0,00211837	4,186	<0,001
C(2)	0,00659223	0,00228996	2,879	0,002
C(3)	0,01033374	0,00261828	3,947	<0,001
A(1, 1)	0,00005714	0,00002620	2,181	0,015
A(2, 2)	0,00006497	0,00003096	2,099	0,018
A(3, 3)	0,00015919	0,00007332	2,171	0,010
ARCH(1; 1, 1)	0,05965359	0,01728203	3,452	<0,001
ARCH(1; 2, 2)	0,07962065	0,01338374	5,949	<0,001
ARCH(1; 3, 3)	0,07962689	0,01909987	4,169	<0,001
GARCH(1; 1, 1)	0,92186900	0,02105657	43,781	<0,001
GARCH(1; 2, 2)	0,90433462	0,01483176	60,973	<0,001
GARCH(1; 3, 3)	0,88583800	0,02828570	31,318	<0,001
Matriz de Correlação Condicional Constante:				
	VALE	PETROBRAS	GERDAU	
VALE	1,0000	0,4796	0,5166	
PETROBRAS	0,4796	1,0000	0,5512	
GERDAU	0,5166	0,5512	1,0000	
Desvios Padrão:				
	VALE	PETROBRAS	GERDAU	
VALE	NA	0,02593	0,02613	
PETROBRAS	0,02593	NA	0,02831	
GERDAU	0,02613	0,02831	NA	

seguintes modelos estimados:

$$\begin{aligned}
 \sigma_{1t}^2 &= 0,00005714 + 0,05965359\varepsilon_{1(t-1)}^2 + 0,92186900\sigma_{1(t-1)}^2 \\
 \sigma_{2t}^2 &= 0,00006497 + 0,07962065\varepsilon_{2(t-1)}^2 + 0,90433462\sigma_{2(t-1)}^2 \\
 \sigma_{3t}^2 &= 0,00015919 + 0,07962689\varepsilon_{3(t-1)}^2 + 0,88583800\sigma_{3(t-1)}^2.
 \end{aligned} \tag{3.31}$$

A Figura 3.6 apresenta as variâncias e covariâncias condicionais estimadas pelas equações (3.30) e (3.31). Observa-se que em contraste com os modelos EWMA, DVEC(1,1) e BEKK(1,1,1) o comportamento é similar, mostrando maior variabilidade nos anos 1999 e 2009.

3.3.2 Modelo GARCH Ortogonal

Considera-se um conjunto de T observações num conjunto de k séries temporais, resumidas numa matriz \mathbf{Y} de dimensões $(T \times k)$. A análise de componentes principais dá como resultado k variáveis estacionárias não correlacionadas, chamadas componentes principais

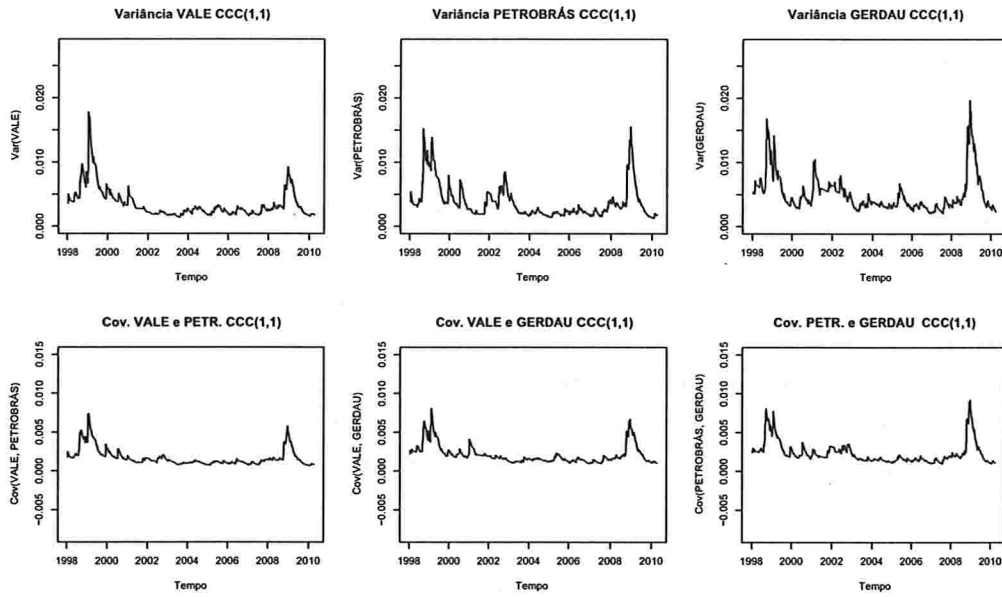


Figura 3.6: Variâncias e covariâncias condicionais estimadas segundo o ajuste CCC-GARCH(1,1).

de \mathbf{Y} , em que cada uma é uma combinação linear das séries originais. Ao mesmo tempo é estabelecido quanto da variação total no sistema original das séries é explicado por cada componente principal, e as componentes são ordenadas de acordo com a quantidade de variação que elas explicam.

Os passos para fazer uma análise de componentes principais são:

1. Padronizar os dados numa matriz \mathbf{X} de dimensões $(T \times k)$, que representa as variáveis de \mathbf{Y} , com a condição de que cada coluna de \mathbf{X} seja da forma $\mathbf{x}_i = (\mathbf{y}_i - \mu_i)/\sigma_i$, em que μ_i e σ_i são a média e o desvio padrão de \mathbf{y}_i , para $i = 1, \dots, k$.
2. Calcular a matriz de autovetores \mathbf{W} da matriz de correlação $\mathbf{X}'\mathbf{X}/T$, e a matriz de autovalores associada $\mathbf{\Lambda}$ ordenando os valores em ordem de magnitude decrescente. Assim, $(\mathbf{X}'\mathbf{X}/T)\mathbf{W} = \mathbf{W}\mathbf{\Lambda}$, onde \mathbf{W} é uma matriz ortogonal de autovetores e $\mathbf{\Lambda}$ é uma matriz diagonal de autovalores.
3. As componentes principais de \mathbf{Y} são obtidas como

$$\mathbf{P} = \mathbf{X}\mathbf{W}. \quad (3.32)$$

Então, foi obtida uma transformação linear dos dados originais de tal forma que as novas séries transformadas sejam ortogonais, e portanto elas têm correlação zero. Para

isto, considera-se $\mathbf{P}'\mathbf{P} = \mathbf{W}'\mathbf{X}'\mathbf{X}\mathbf{W} = \mathbf{W}'\mathbf{W}\mathbf{\Lambda}T$, mas \mathbf{W} é uma matriz ortogonal, logo $\mathbf{P}'\mathbf{P}/T = \mathbf{\Lambda}$, que é uma matriz diagonal. A variância não condicional da i -ésima componente principal é portanto o i -ésimo autovalor, e a correlação não condicional entre as componentes principais é zero.

Se \mathbf{W} for invertível, a equação (3.32) é equivalente a $\mathbf{X} = \mathbf{P}\mathbf{W}'$, já que \mathbf{W} é ortogonal, assim

$$\mathbf{x}_i = w_{1i}\mathbf{P}_1 + w_{2i}\mathbf{P}_2 + \dots + w_{ki}\mathbf{P}_k, \quad (3.33)$$

e a matriz \mathbf{W} é chamada a matriz de "pesos fatoriais". Considerando o conceito de inércia do sistema ("inertia of the system"), I_s , dado em componetes principais (Peña (2002)) e que é definido como o traço da matriz $\mathbf{X}'\mathbf{X}/T$, a variabilidade do sistema original pode ser estimada por

$$\begin{aligned} I_s &:= \text{tr}(\mathbf{X}'\mathbf{X}/T) \\ &= \text{tr}(\mathbf{W}\mathbf{P}'\mathbf{P}\mathbf{W}'/T) \\ &= \text{tr}(\mathbf{W}\mathbf{\Lambda}\mathbf{W}') \\ &= \text{tr}(\mathbf{W}'\mathbf{W}\mathbf{\Lambda}) \\ &= \text{tr}(\mathbf{\Lambda}) \\ &= \sum_{i=1}^k \lambda_i, \end{aligned}$$

em que λ_i representa o i -ésimo elemento na diagonal da matriz $\mathbf{\Lambda}$. Logo, a proporção da variação total em \mathbf{X} que é explicada pela i -ésima componente principal é igual a λ_i/I_s . Os autovalores estão ordenados na diagonal da matriz $\mathbf{\Lambda}$ de tal forma que $\lambda_1 > \lambda_2 > \dots > \lambda_k$, e a i -ésima componente é escolhida na matriz \mathbf{W} de acordo com o i -ésimo autovalor correspondente. Considere m primeiras componentes que representam uma porcentagem mínima da variabilidade τ , isto é,

$$100 \sum_{i=1}^m \lambda_i/I_s \geq \tau.$$

O sistema (3.33) pode se escrito em termos das variáveis originais \mathbf{Y} como

$$\mathbf{y}_i = \mu_i + \omega_{1i}\mathbf{P}_1 + \omega_{2i}\mathbf{P}_2 + \dots + \omega_{mi}\mathbf{P}_m + \epsilon_i, \quad (3.34)$$

onde $\omega_{ij} = w_{ij}\sigma_i$ e o termo ϵ_i representa o erro na aproximação usando somente as primeiras m componentes principais. Estas m componentes principais são os fatores "chave" no sistema e a variação restante é considerada um "ruído".

Para obter a matriz de covariâncias, calcula-se a variância não condicional no sistema (3.34) e considerando que as componentes principais são não correlacionadas, tem-se que

$$\Sigma = ADA' + V_\epsilon, \quad (3.35)$$

onde $A = [\omega_{ij}]$ é a matriz ($k \times m$) de pesos fatoriais reescalados,

$$D = \text{diag}\{\text{var}(P_1), \dots, \text{var}(P_m)\}$$

é a matriz diagonal de variâncias das m primeiras componentes principais e V_ϵ é a matriz de covariâncias dos erros. Ignorando este último termo obtem-se a aproximação

$$\Sigma \approx ADA', \quad (3.36)$$

com uma precisão que é controlada escolhendo mais ou menos componentes para representar o sistema. Isto mostra como a matriz de covariâncias ($k \times k$) das séries originais Σ pode ser obtida através de umas poucas componentes principais. Quando $k = m$ a matriz V_ϵ é nula e (3.36) fica como:

$$\Sigma = ADA'.$$

Nota-se que Σ obtida através de (3.36) é positiva semidefinida, mas não é estritamente positiva definida a menos que $k = m$. Embora D seja positiva definida (uma matriz diagonal de elementos positivos), não é possível garantir que ADA' seja positiva definida quando $m < k$. Para mostrar isto, considere um vetor z de dimensões ($k \times 1$), e

$$z'ADA'z = v'Dv,$$

com $v = A'z$. Como $m < k$ então na equação anterior existe um sistema linear com menos equações que variáveis z , e portanto v pode ser zero para algum z diferente de zero, assim que $z'ADA'z$ não é estritamente positivo para todo z diferente de zero. Assim que ADA' é somente positiva semidefinida.

O modelo *GARCH Ortogonal*⁷ considera que a matriz D varia no tempo, isto é, o modelo GARCH Ortogonal considera que

$$\Sigma_t = AD_tA', \quad (3.37)$$

em que A é a matriz ($k \times m$) de pesos fatoriais reescalados definida anteriormente, e D_t é uma matriz diagonal ($m \times m$), com as variâncias das primeiras m componentes princi-

⁷Em alguns textos, como por exemplo Zivot e Wang (2006), o modelo GARCH Ortogonal também é conhecido como o modelo GARCH de Componentes Principais.

país dado o passado \mathcal{F}_{t-1} , na diagonal. É possível ajustar para cada uma das componentes principais um modelo GARCH univariado de maneira independente, e para cada tempo t calcular a matriz de covariâncias condicional das séries originais Σ_t . Definidas desta forma estas matrizes são PSD para cada instante t .

As componentes principais são não correlacionadas, mas quando são consideradas as informações passadas pode existir correlação entre elas, isto é, a matriz de covariâncias dado \mathcal{F}_{t-1} , \mathbf{D}_t , não é necessariamente diagonal. O modelo de componentes principais está baseado na suposição de que a matriz \mathbf{D}_t é diagonal em cada instante t .

O procedimento para ajustar um modelo GARCH Ortogonal é o seguinte:

1. Para o conjunto de vetores de séries $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_T$, considere a matriz $\mathbf{Y} = [\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_T]$ de dimensões $(k \times T)$.
2. Padronizar a matriz \mathbf{Y} , subtraindo a média e dividindo pelo desvio padrão de cada coluna, obtendo a matriz \mathbf{X} .
3. Obter as matrizes de autovalores Λ , e de autovetores \mathbf{W} , da matriz $\mathbf{X}'\mathbf{X}/T$.
4. Escolher o número de componentes principais m , que representam bem o sistema original, baseados nos autovalores.
5. Obter as matriz de componentes principais, \mathbf{P} de dimensões $(T \times m)$, considerando a relação (3.32).
6. Para cada componente ajustar um modelo GARCH univariado para a volatilidade, e gerar as matrizes diagonais \mathbf{D}_t , $t = 1, \dots, T$, através destes ajustes.
7. Obter a matriz de pesos fatoriais reescalados \mathbf{A} , definida em (3.35) e para cada instante t considerar a matriz de covariâncias condicionais $\Sigma_t = \mathbf{A}\mathbf{D}_t\mathbf{A}'$.

Exemplo 3.7. Considera-se agora um ajuste de Componentes Principais normalmente distribuído para os retornos das séries VALE, PETROBRAS e GERDAU, fazendo para cada uma das componentes um ajuste GARCH(1,1) univariado.

A Tabela 3.8 apresenta os parâmetros estimados para o modelo PCOMP-GARCH(1,1). Todos os p -valores são significativos ao nível de significância de 5%. Também são apresentados os autovetores com seus respectivos autovalores.

Tabela 3.8: *Resumo do ajuste PCOMP-GARCH(1,1).*

Parâmetros Estimados CCC-GARCH(1,1):				
	Value	Std.Error	t value	Pr(> t)
A(1, 1)	0,00038871	1,514e-004	2,567	0,005
A(2, 2)	0,00001316	6,923e-006	1,901	0,029
A(3, 3)	0,00001491	1,053e-005	1,416	0,079
ARCH(1; 1, 1)	0,08762347	2,098e-002	4,177	< 0,001
ARCH(1; 2, 2)	0,08194596	1,702e-002	4,816	< 0,001
ARCH(1; 3, 3)	0,06318318	1,734e-002	3,644	< 0,001
GARCH(1; 1, 1)	0,86414292	3,159e-002	27,359	< 0,001
GARCH(1; 2, 2)	0,91533191	1,638e-002	55,866	< 0,001
GARCH(1; 3, 3)	0,92809756	1,731e-002	53,624	< 0,001
Autovetores: (matriz da transformação ortogonal)				
	VALE	PETROBRAS	GERDAU	
VALE	0,4825	-0,78628	0,3859	
PETROBRAS	0,5230	0,61208	0,5931	
GERDAU	0,7026	0,08435	-0,7066	
Autovalores:				
0,008529	0,002156	0,001851		

Neste ajuste os critérios **AIC** e **BIC** são respectivamente **-5268,673** e **-5216,525**. Já para o teste de Portmanteau para o quadrado dos resíduos, o p -valor foi **0,4803**, não rejeitando a hipótese nula.

O modelo PCOMP-GARCH(1,1) ajustado é dado por

$$\Sigma_t = \begin{pmatrix} 0,48 & -0,79 & 0,39 \\ 0,52 & 0,61 & 0,59 \\ 0,71 & 0,08 & -0,71 \end{pmatrix} D_t \begin{pmatrix} 0,48 & 0,52 & 0,71 \\ -0,79 & 0,61 & 0,08 \\ 0,39 & 0,59 & -0,71 \end{pmatrix} \quad (3.38)$$

com

$$D_t = \begin{pmatrix} \sigma_{1t}^2 & 0 & 0 \\ 0 & \sigma_{2t}^2 & 0 \\ 0 & 0 & \sigma_{3t}^2 \end{pmatrix},$$

em que σ_{it}^2 , para $i = 1, 2, 3$, é a variância da i -ésima componente principal. Portanto, segundo a Tabela 3.8, as variâncias seguem os seguintes modelos estimados:

$$\begin{aligned} \sigma_{1t}^2 &= 0,00038871 + 0,08762347p_{1(t-1)}^2 + 0,86414292\sigma_{1(t-1)}^2 \\ \sigma_{2t}^2 &= 0,00001316 + 0,08194596p_{2(t-1)}^2 + 0,91533191\sigma_{2(t-1)}^2 \\ \sigma_{3t}^2 &= 0,06318318p_{3(t-1)}^2 + 0,92809756\sigma_{3(t-1)}^2, \end{aligned} \quad (3.39)$$

em que $\mathbf{p}_t = (p_{1(t-1)} \ p_{2(t-1)} \ p_{3(t-1)})'$ é o vetor de componentes principais no instante t , e $\mathbf{p}_t \sim N_3(\mathbf{0}, \mathbf{D}_t)$.

A Figura 3.7 apresenta as variâncias e covariâncias condicionais estimadas pelas equações (3.38) e (3.39). Observa-se que o comportamento é parecido com os modelos EWMA, DVEC(1,1), BEKK(1,1,1) e CCC-GARCH(1,1) ajustados até agora, mas o gráfico da covariância condicional dos retornos das ações VALE e PETROBRAS tem valores menores no ano 1999, comparando com os outros modelos.

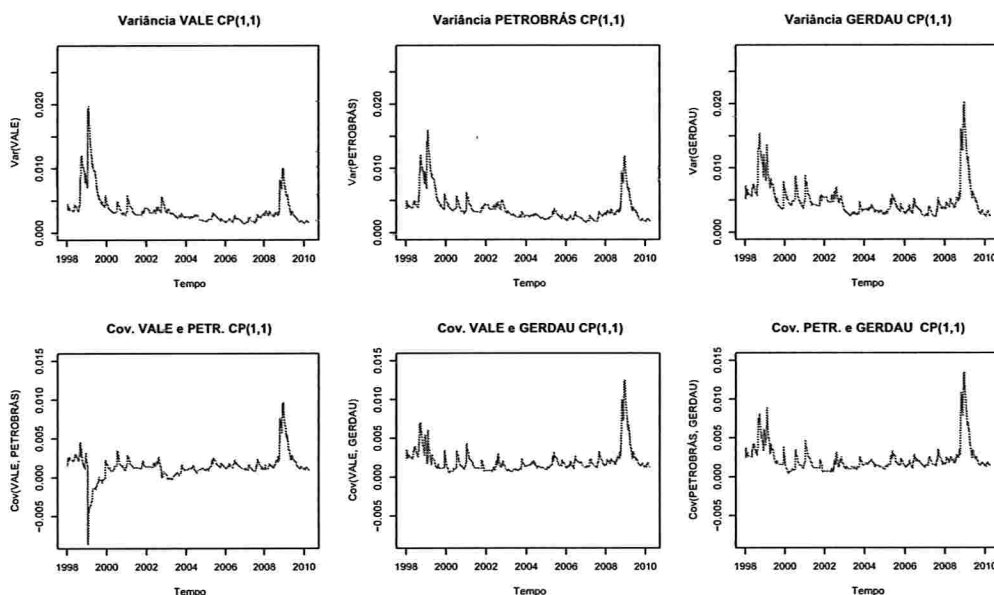


Figura 3.7: Variâncias e covariâncias estimadas condicionais segundo o ajuste PCOMP-GARCH(1,1).

Finalmente, apresentam-se na Figura 3.8 as variâncias e covariâncias condicionais de todos os modelos considerados. Nota-se que em geral o comportamento é similar em todos os casos, mostrando uma variabilidade maior nos anos 1999 e 2009 e um comportamento mais volátil nos modelos BEKK(1,1,1) e PCOMP-GARCH(1,1), ao contrário dos modelos EWMA e CCC-GARCH(1,1), em que o comportamento foi mais suave em cada variância e cada covariância.

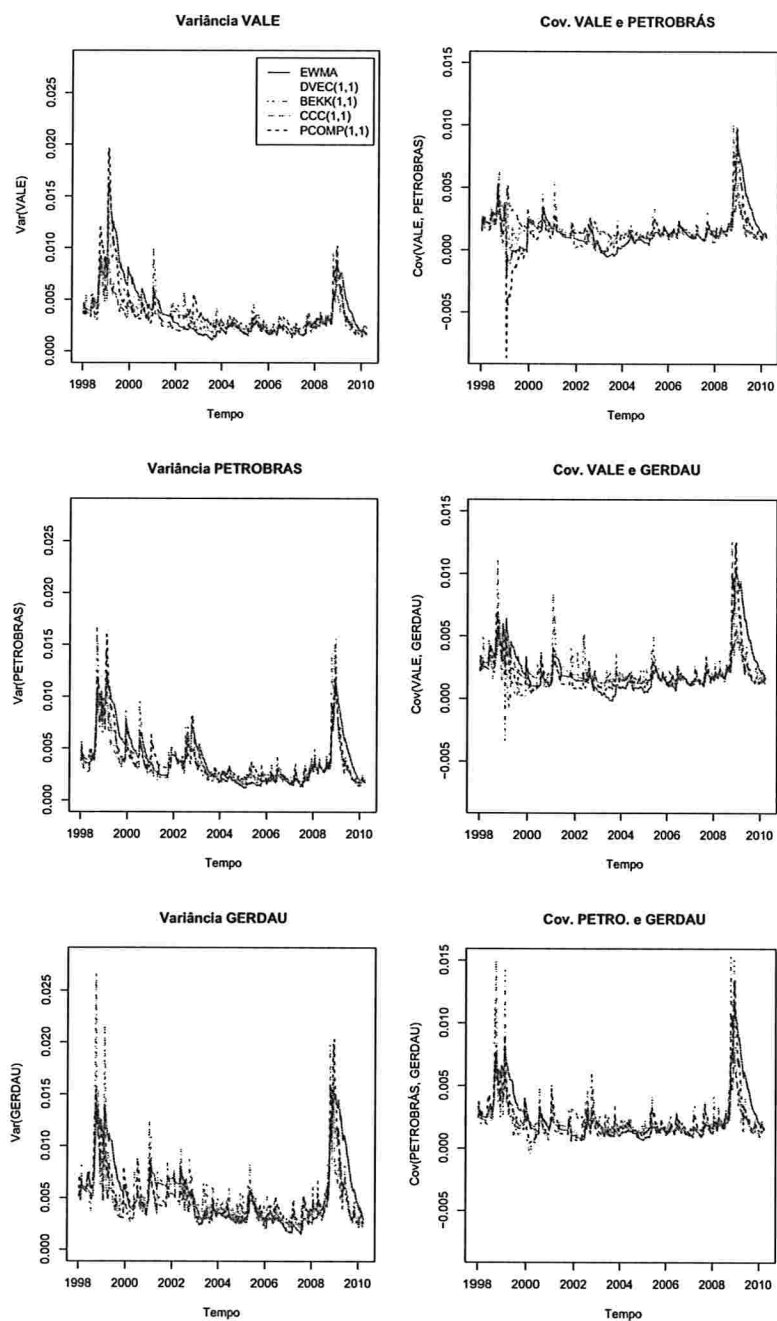


Figura 3.8: Variâncias e covariâncias condicionais estimadas em todos os ajustes.

Capítulo 4

Simulações

Neste capítulo serão apresentadas algumas simulações dos modelos considerados no Capítulo 3, com o intuito de comparar as estimativas de variâncias e covariâncias condicionais, utilizando o software S-PLUS. O Apêndice B apresenta os códigos e algoritmos utilizados nas simulações.

4.1 Modelo EWMA

Considere o modelo dado em (3.4) junto com a equação (3.1), com $\mathbf{c} = \mathbf{0}$,

$$\mathbf{r}_t = \boldsymbol{\varepsilon}_t, \quad (4.1)$$

$$\boldsymbol{\Sigma}_t = (1 - \lambda)\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' + \lambda\boldsymbol{\Sigma}_{t-1}, \quad (4.2)$$

em que $\boldsymbol{\varepsilon}_t$ é um vetor de dimensão (4×1) normal multivariado, $\mathbf{N}_4(\mathbf{0}, \boldsymbol{\Sigma}_t)$, e $t = 1, \dots, 500$.

A matriz $\boldsymbol{\Sigma}_0$ inicial foi obtida de tal forma que seja PSD, e está dada por

$$\boldsymbol{\Sigma}_0 = \begin{bmatrix} 0,16 & 0,12 & 0,04 & 0,32 \\ 0,12 & 0,73 & -0,29 & 0,40 \\ 0,04 & -0,29 & 0,98 & -0,63 \\ 0,32 & 0,40 & -0,63 & 1,42 \end{bmatrix}.$$

Considerando que na equação (4.2) o valor de λ determina o peso da matriz $\boldsymbol{\Sigma}_{t-1}$, os valores escolhidos para simulação foram 0,90, 0,92, 0,94, 0,96, 0,98 e 0,99.

As Figuras 4.1, 4.2 e 4.3 apresentam as 4 séries simuladas para cada valor de λ , segundo o modelo dado por (4.1) e (4.2). Observa-se que existem conglomerados ao longo do tempo, em que a variabilidade é maior para cada valor de λ . Além disso, nota-se que a variabilidade converge rapidamente para zero quando $\lambda \rightarrow 0,90$. Algumas simulações que não foram apresentadas nesta dissertação foram feitas para valores menores que 0,90, como por exemplo 0,88 e 0,89, no entanto as matrizes de covariâncias condicionais convergiram muito rápido para zero, obtendo matrizes nulas que o software S-PLUS não reconhece como matriz PSD, gerando problemas nas simulações.

As Figuras 4.4, 4.5 e 4.6 apresentam as funções de autocorrelação e autocorrelação cruzada das séries simuladas ao quadrado. Nota-se que existe autocorrelação e autocorrelação cruzada entre elas, sendo mais significativa para os valores de $\lambda \in \{0,90, 0,92, 0,94, 0,96\}$ e menos significativa para $\lambda \in \{0,98, 0,99\}$, o qual pode ser explicado ao considerar que segundo a equação (4.2), quando $\lambda \rightarrow 1$, a matriz de covariâncias condicional depende menos de $\varepsilon_t \varepsilon_t'$ que determina o efeito autoregressivo no modelo EWMA.

Considerando que existe autocorrelação e autocorrelação cruzada entre as séries simuladas ao quadrado, foram feitos para cada valor de λ ajustes GARCH multivariados EWMA. Os parâmetros estimados para cada um dos valores de λ são apresentados na Tabela 4.1. As estimativas de λ foram muito próximas dos valores reais. Além disso, nota-se que para $\lambda \in \{0,92, 0,94, 0,96, 0,98, 0,99\}$, os valores $C(1)$, $C(2)$, $C(3)$ e $C(4)$, que representam as estimativas do vetor c na equação (3.1), são estatisticamente não significantes ao nível de 5%. No caso $\lambda = 0,90$ as estimativas $C(1)$ e $C(2)$ foram significantes.

Para o quadrado dos resíduos de cada um dos modelos foi feito o teste multivariado Portmanteau com o intuito de testar se eles têm autocorrelação cruzada. Os p -valores para os testes são apresentados na Tabela 4.2 e ao nível de significância de 5%, a hipótese nula de não autocorrelação cruzada é rejeitada para $\lambda \in \{0,90, 0,02, 0,94\}$. Para tentar explicar isto, considera-se o modelo EWMA dado na equação (3.4), em que o efeito ARCH depende do termo

$$(1 - \lambda)\varepsilon_{t-1}\varepsilon'_{t-1},$$

e quando $\lambda \rightarrow 0,90$ o valor $(1 - \lambda)$ é maior, portanto o efeito ARCH tem um maior peso que o ajuste EWMA não consegue descrever bem. O contrário acontece quando $\lambda \rightarrow 1$, em que os p -valores do teste de Portmanteau são significantes.

As variâncias condicionais são apresentadas nas Figuras 4.7 a 4.12, junto com os desvios padrão. As variâncias simuladas, com linha contínua, e as variâncias estimadas pelo modelo

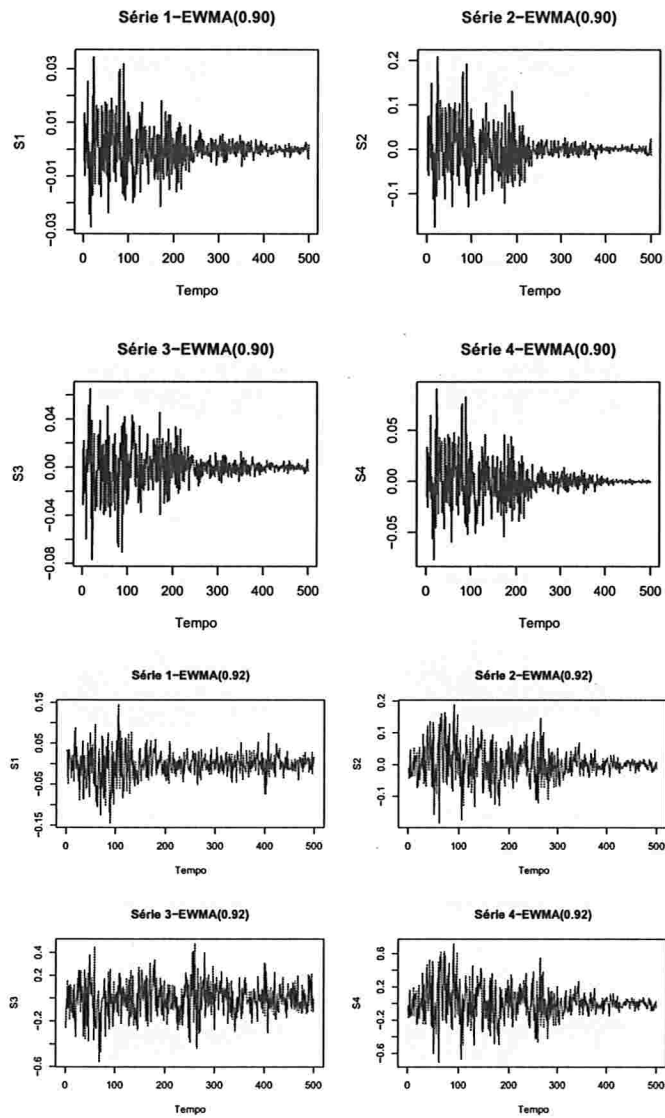


Figura 4.1: Séries da simulação EWMA para $\lambda = 0,90$ e $0,92$.

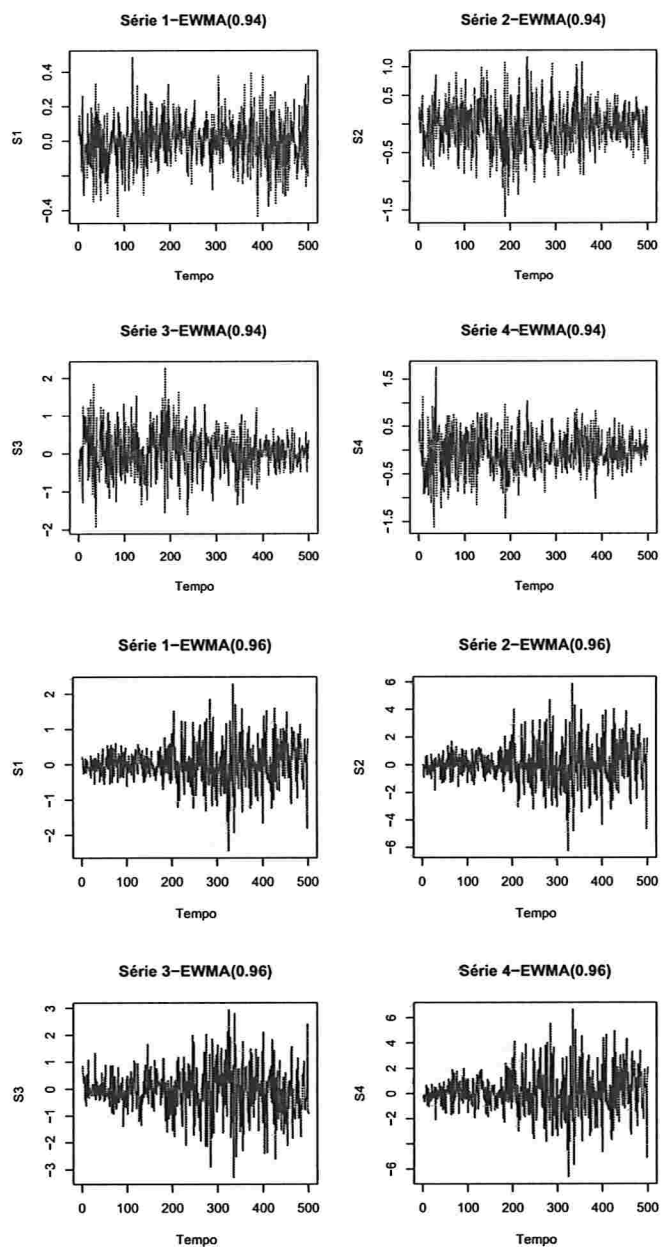


Figura 4.2: Séries da simulação EWMA para $\lambda = 0,94$ e $0,96$.

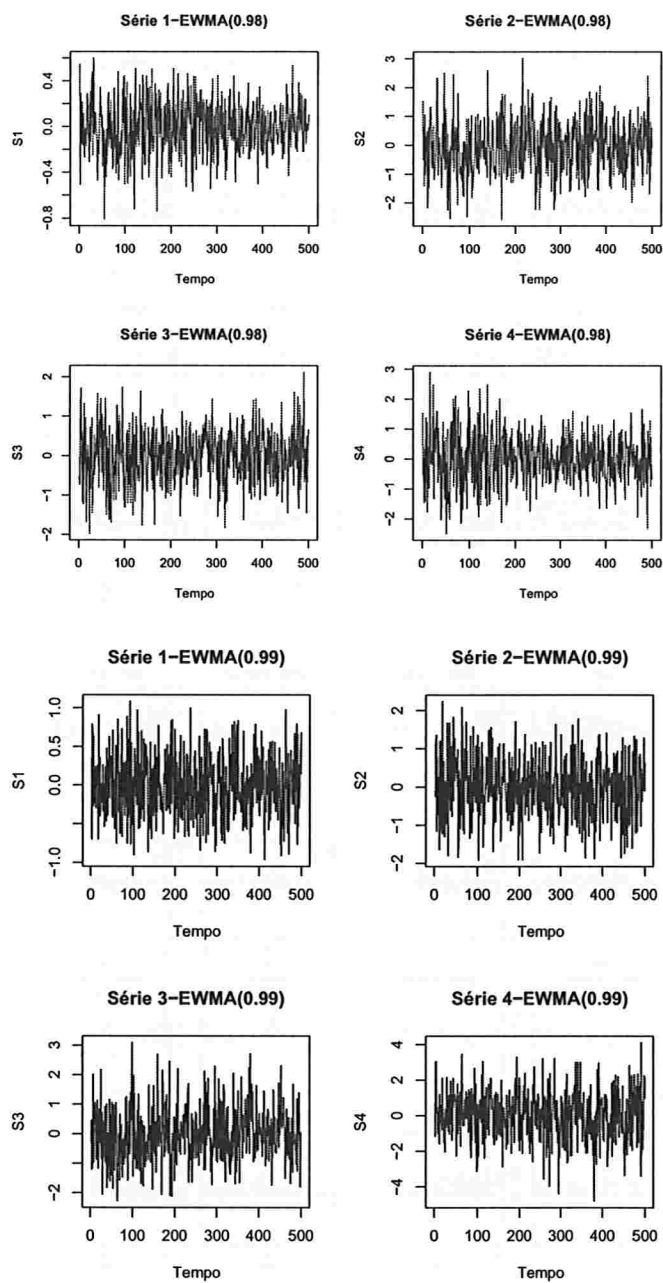


Figura 4.3: Séries da simulação EWMA para $\lambda = 0,98$ e $0,99$.

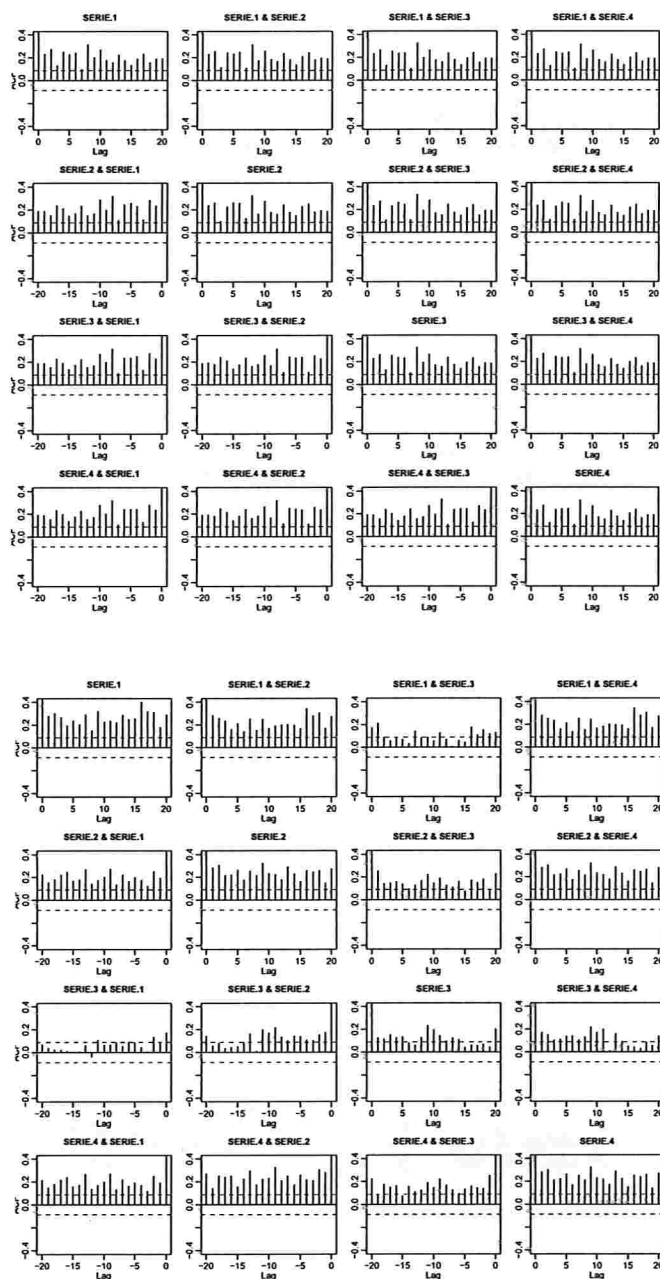


Figura 4.4: *FAC* e *FAC* cruzadas das séries simuladas ao quadrado para $\lambda = 0,90$ (Acima) e $\lambda = 0,92$ (Embaixo).

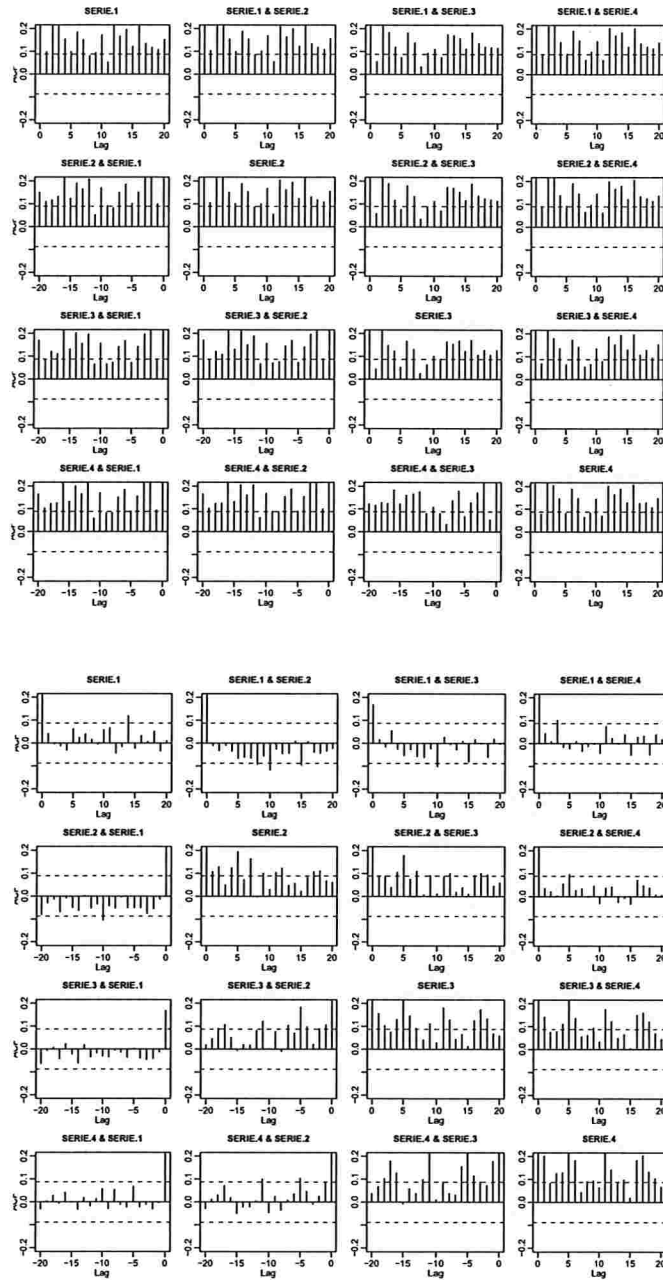


Figura 4.5: *FAC e FAC cruzadas das séries simuladas ao quadrado para $\lambda = 0,94$ (Acima) e $\lambda = 0,96$ (Embaixo).*

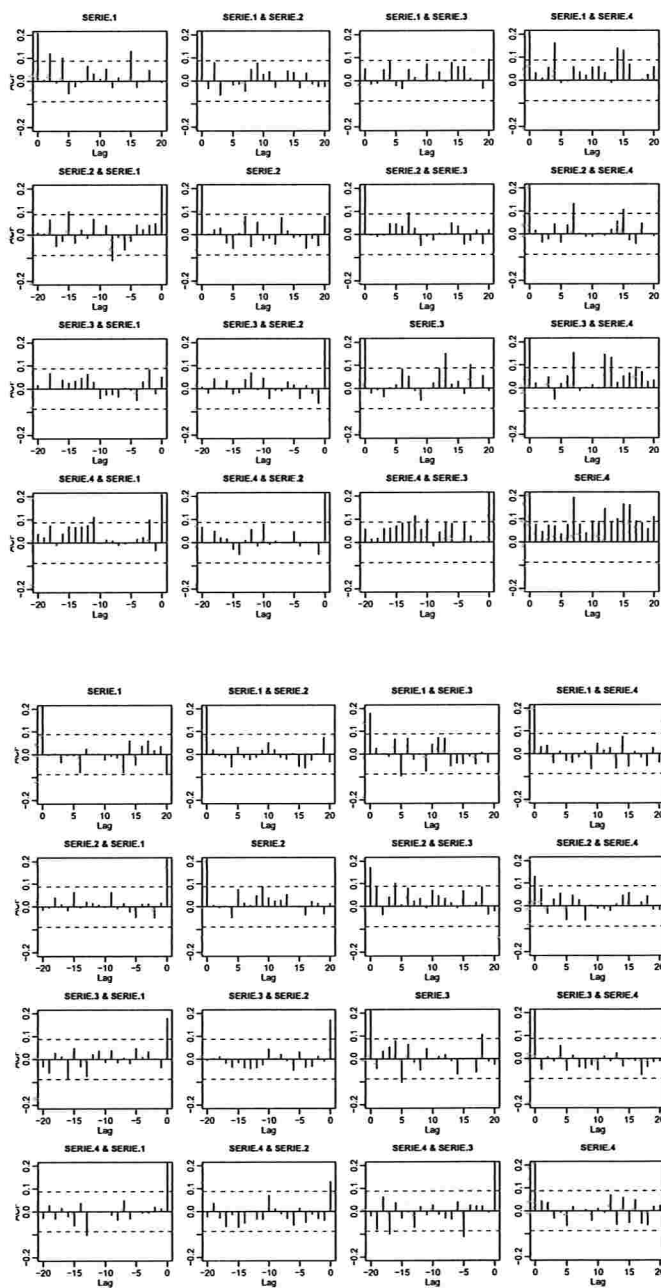


Figura 4.6: *FAC e FAC cruzadas das séries ao quadrado para $\lambda = 0,98$ (Acima) e $\lambda = 0,99$ (Embaixo).*

Tabela 4.1: Parâmetros e valores de λ estimados.

Parâmetros estimados para $\lambda = 0,90$				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	-2,311e-005	8,633e-006	-2,677	0,0077
C(2)	2,027e-004	1,679e-004	1,207	0,2279
C(3)	4,292e-005	1,441e-005	2,979	0,0030
C(4)	5,152e-006	2,020e-005	0,255	0,7988
$\hat{\lambda}$	0,90067	0,0045	22,316	< 0,001
Parâmetros estimados para $\lambda = 0,92$				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	-8,191e-006	0,0005281	-0,01551	0,9876
C(2)	-2,469e-004	0,0007062	-0,34963	0,7268
C(3)	1,102e-003	0,0026765	0,41190	0,6806
C(4)	-8,979e-004	0,0027006	-0,33247	0,7397
$\hat{\lambda}$	0,92064	0,0041350	19,19215	< 0,001
Parâmetros estimados para $\lambda = 0,94$				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	0,00045085	0,003768	0,119642	0,9048
C(2)	-0,00008615	0,011638	-0,007402	0,9941
C(3)	-0,00087670	0,015119	-0,057989	0,9538
C(4)	0,00014815	0,010959	0,013519	0,9892
$\hat{\lambda}$	0,94015157	0,004109	14,565704	< 0,001
Parâmetros estimados para $\lambda = 0,96$				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	0,01586	0,015042	1,054	0,2923
C(2)	0,04079	0,038923	1,048	0,2952
C(3)	-0,04900	0,022257	-2,202	0,0815
C(4)	0,07085	0,043328	1,635	0,1026
$\hat{\lambda}$	0,95768	0,003421	12,371	< 0,001
Parâmetros estimados para $\lambda = 0,98$				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	0,010514	0,010572	0,99445	0,3205
C(2)	0,019738	0,018050	1,09354	0,2747
C(3)	0,001507	0,018506	0,08145	0,9351
C(4)	0,001488	0,036807	0,04043	0,9678
$\hat{\lambda}$	0,978366	0,002664	8,12045	< 0,001
Parâmetros estimados para $\lambda = 0,99$				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	-0,004528	0,01717	-0,2636	0,7922
C(2)	0,001004	0,03476	0,0289	0,97696
C(3)	0,009070	0,03991	0,2273	0,8203
C(4)	0,010914	0,05633	0,1938	0,8464
$\hat{\lambda}$	0,989895	0,00232	4,3565	< 0,001

Tabela 4.2: Testes multivariados Portmanteau para o quadrado dos resíduos.

Teste Multivariado Portmanteau: Tipo Ljung-Box		
Hipótese nula: não correlação serial		
λ	Estat. do Teste	p.valor
0,90	545,2814	<0,001
0,92	250,2524	0,0086
0,94	275,3182	0,0051
0,96	222,2164	0,0654
0,98	210,3124	0,0916
0,99	185,1671	0,5434
Dist. sob Hipót. Nula: qui-quadrado com 192 graus de liberdade		
Total Observ.: 500		

ajustado, com linha pontilhada, são apresentadas no mesmo gráfico para cada valor de λ e para cada série. Em todos os gráficos existe tendência, o qual confirma que o processo dado por (4.1) e (4.2) é não estacionário. Para valores pequenos de λ , 0,90 e 0,92, a volatilidade e o desvio padrão decrescem de forma rápida e para os outros valores a tendência parece ser mais suave. Observa-se também que para os valores de λ mais grandes, a variância estimada demora mais para convergir à variância simulada, mas em geral em todos os gráficos o modelo ajustado descreve corretamente a variância simulada.

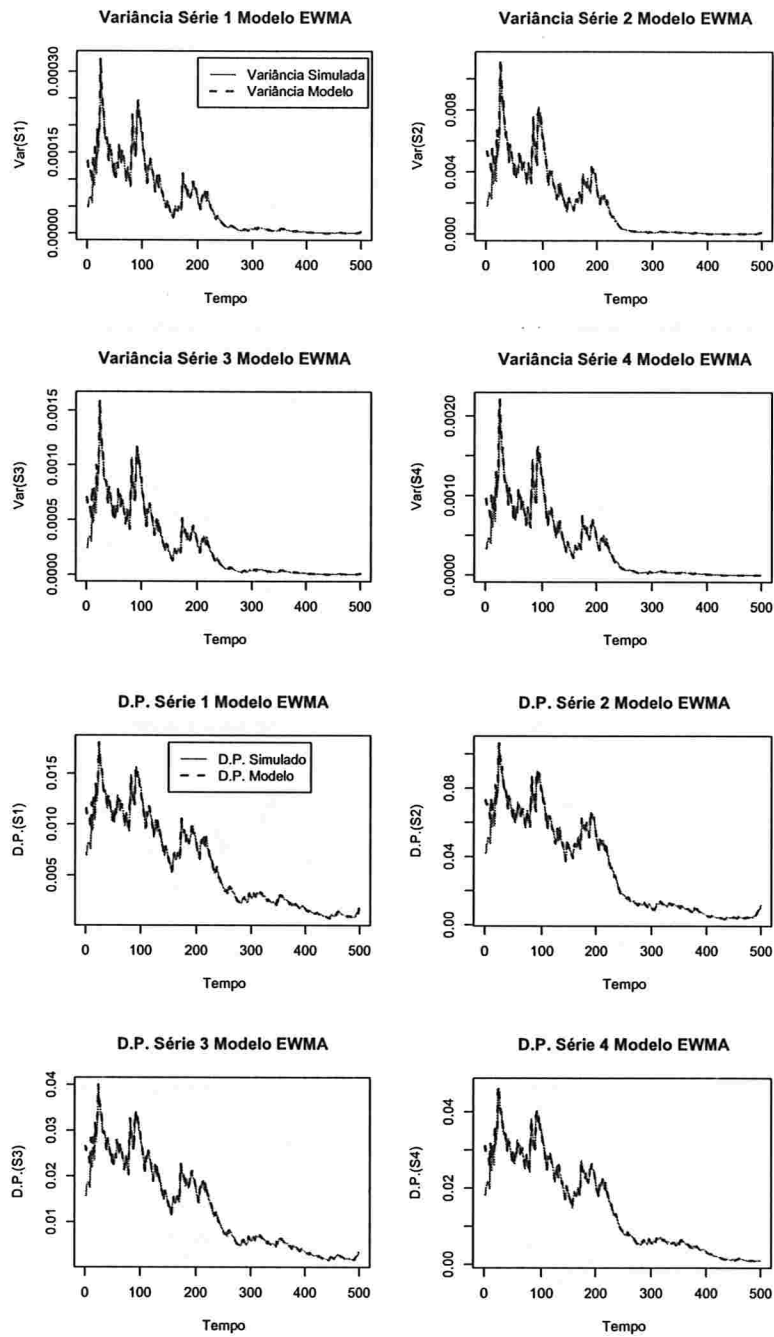


Figura 4.7: Variâncias e desvios padrão condicionais da simulação EWMA para $\lambda = 0,90$.

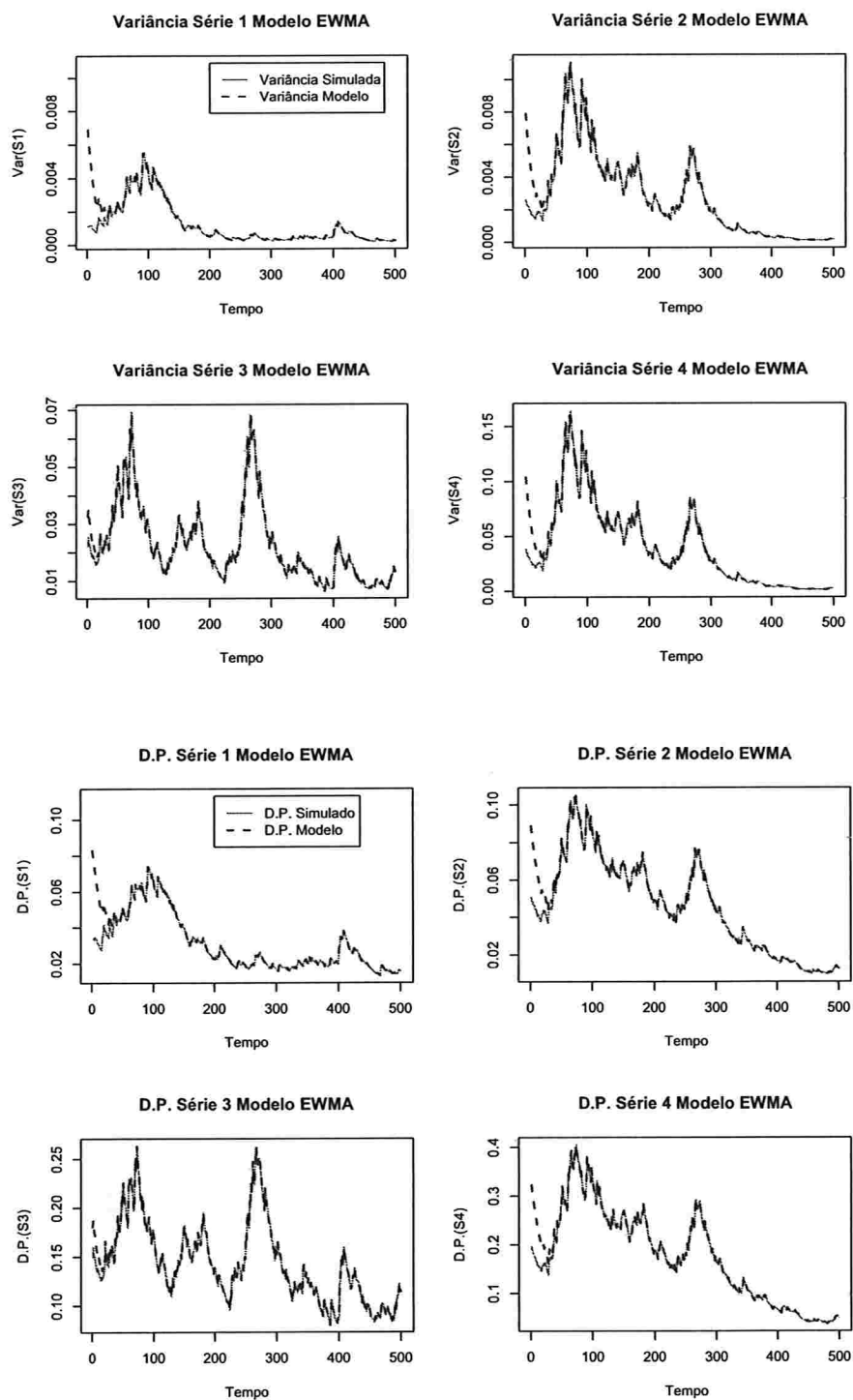


Figura 4.8: Variâncias e desvios padrão condicionais da simulação EWMA para $\lambda = 0,92$.

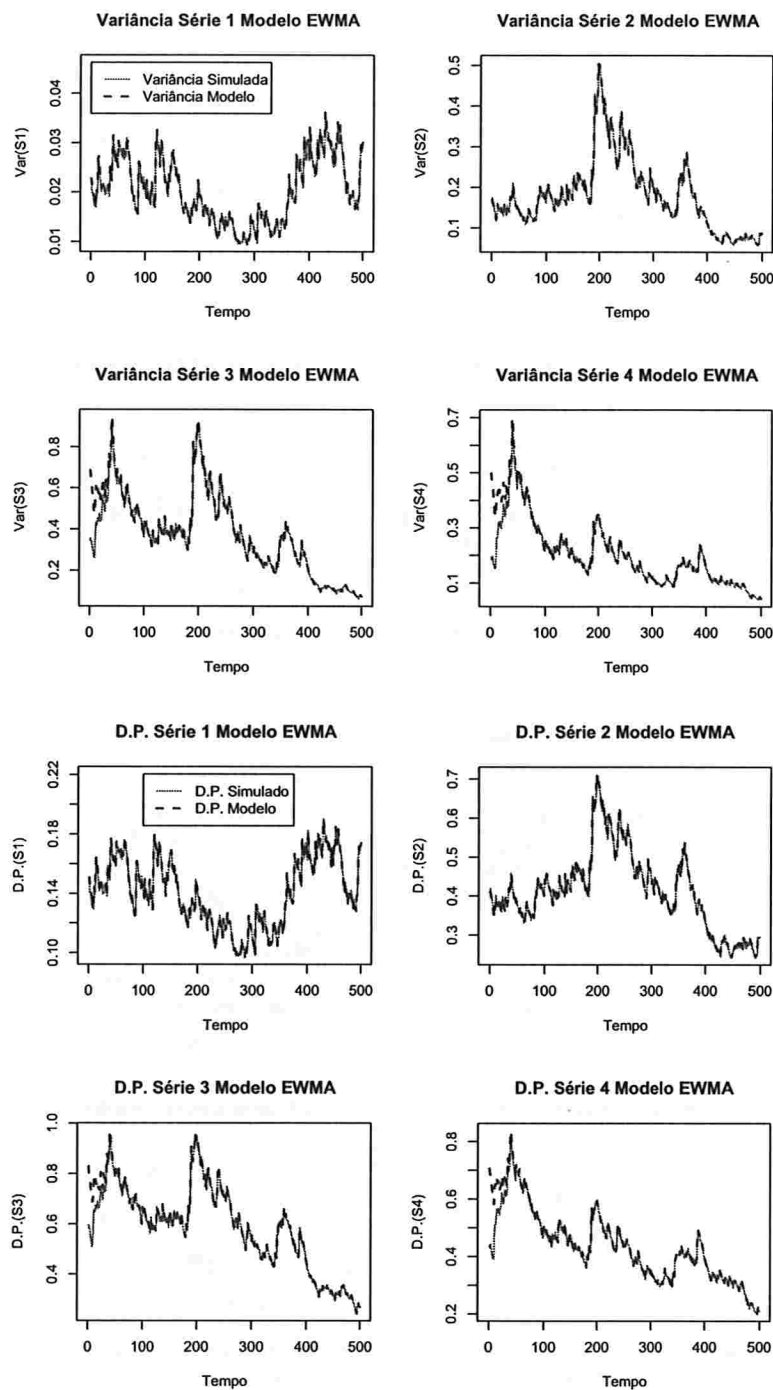


Figura 4.9: Variâncias e desvios padrão da simulação EWMA para $\lambda = 0,94$.

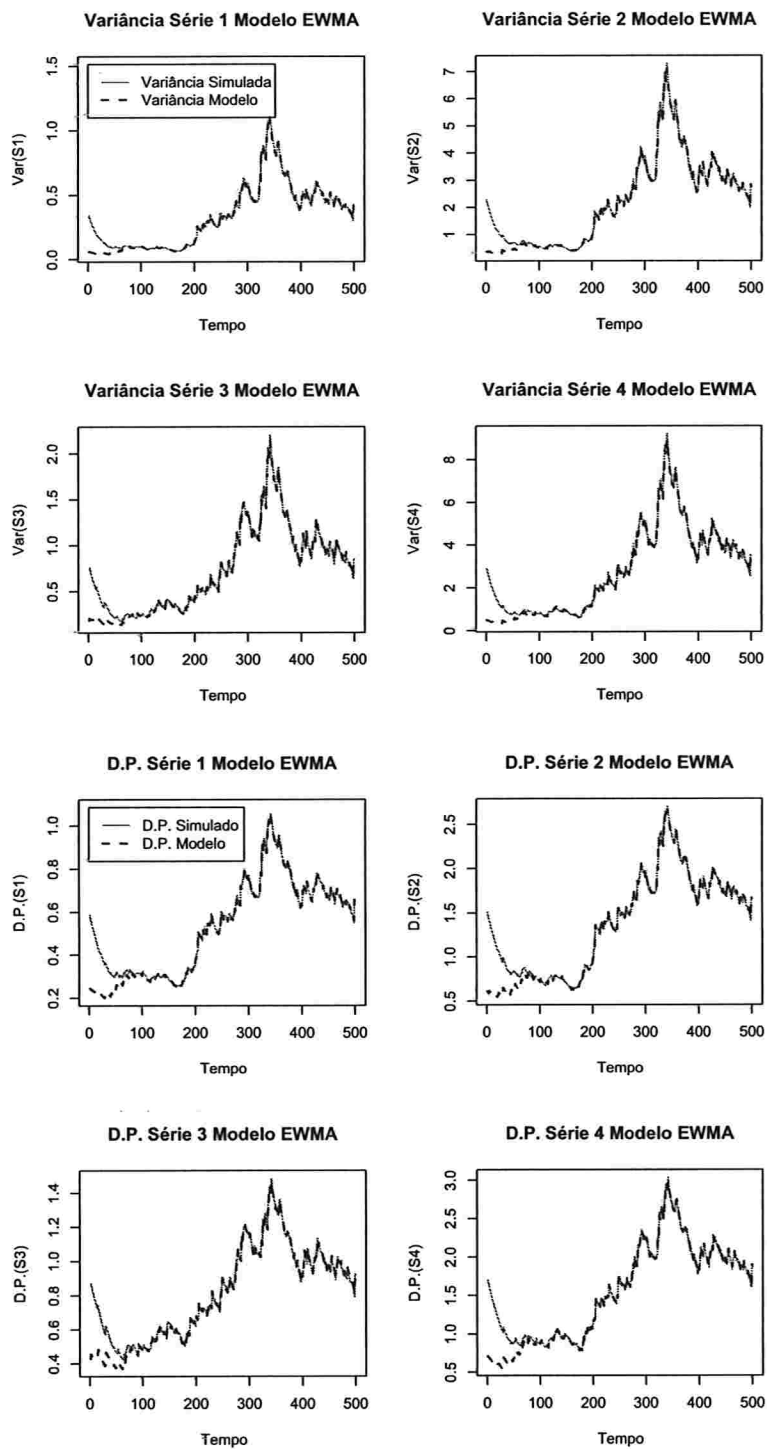


Figura 4.10: Variâncias e desvios padrão condicionais da simulação EWMA para $\lambda = 0,96$.

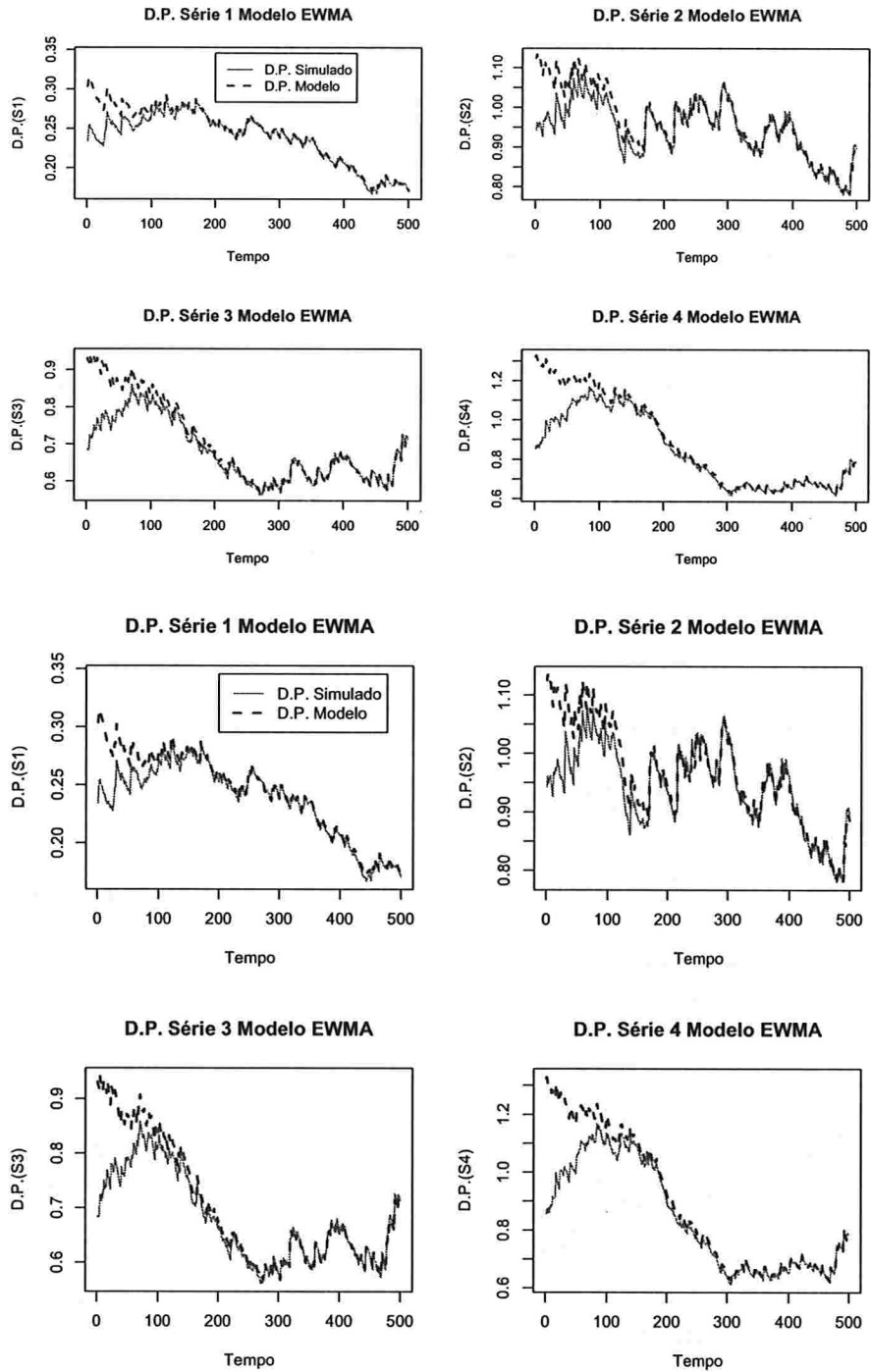


Figura 4.11: Variâncias e desvios padrão condicionais da simulação EWMA para $\lambda = 0,98$.

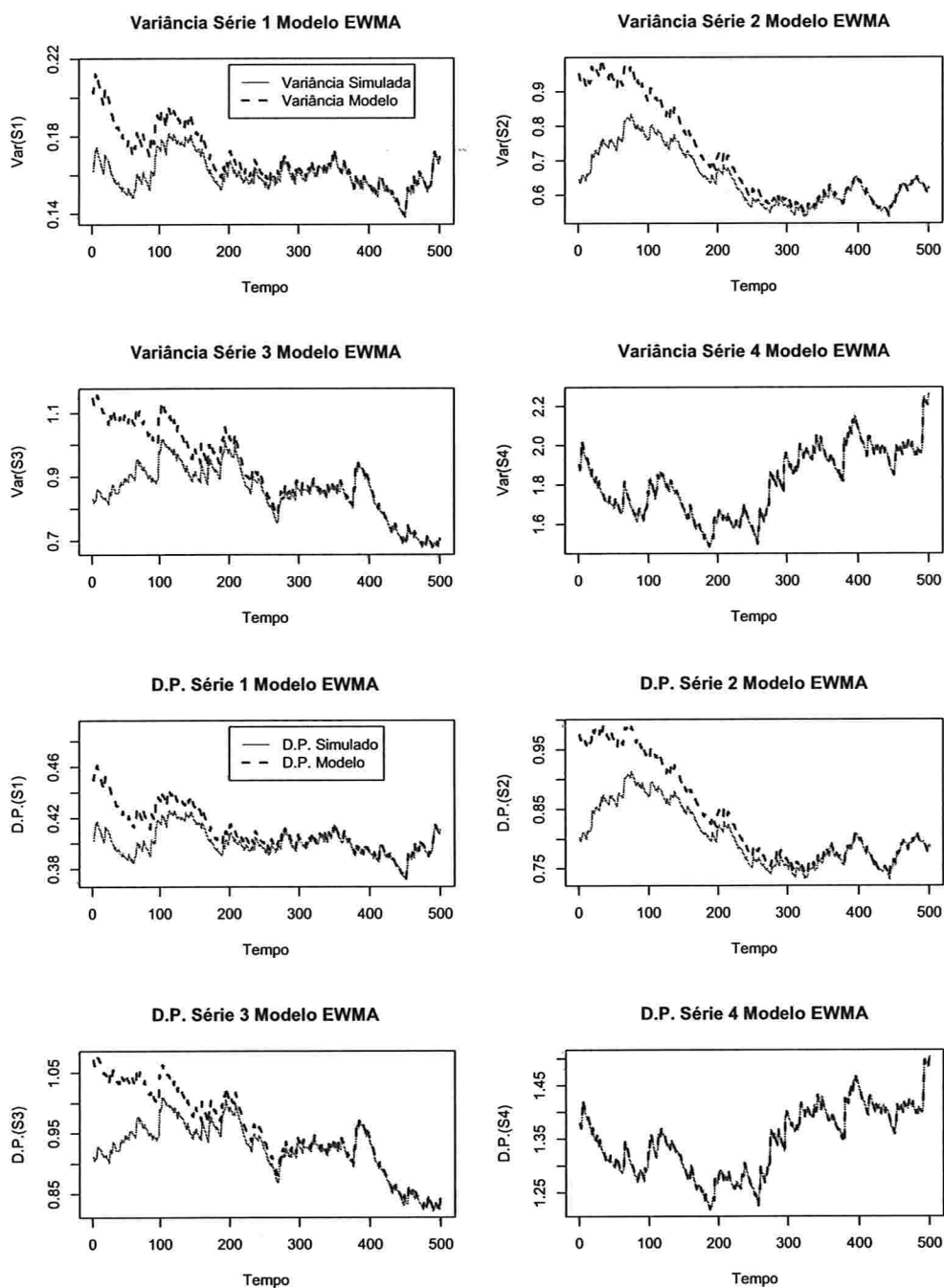


Figura 4.12: Variâncias e desvios padrão condicionais da simulação EWMA para $\lambda = 0,99$.

As Figuras 4.13 a 4.18 apresentam as covariâncias condicionais simuladas (linha contínua) e as covariâncias condicionais ajustadas (linha pontilhada), junto com as correlações condicionais. Os comportamentos são similares aos das variâncias condicionais. Observa-se que existe em todos os casos tendência ao longo do tempo. Nota-se também que as covariâncias para $\lambda = 0,90$ e $0,92$ convergem para zero rapidamente. Já para as correlações, a tendência parece ser mais suave para $\lambda \rightarrow 1$.

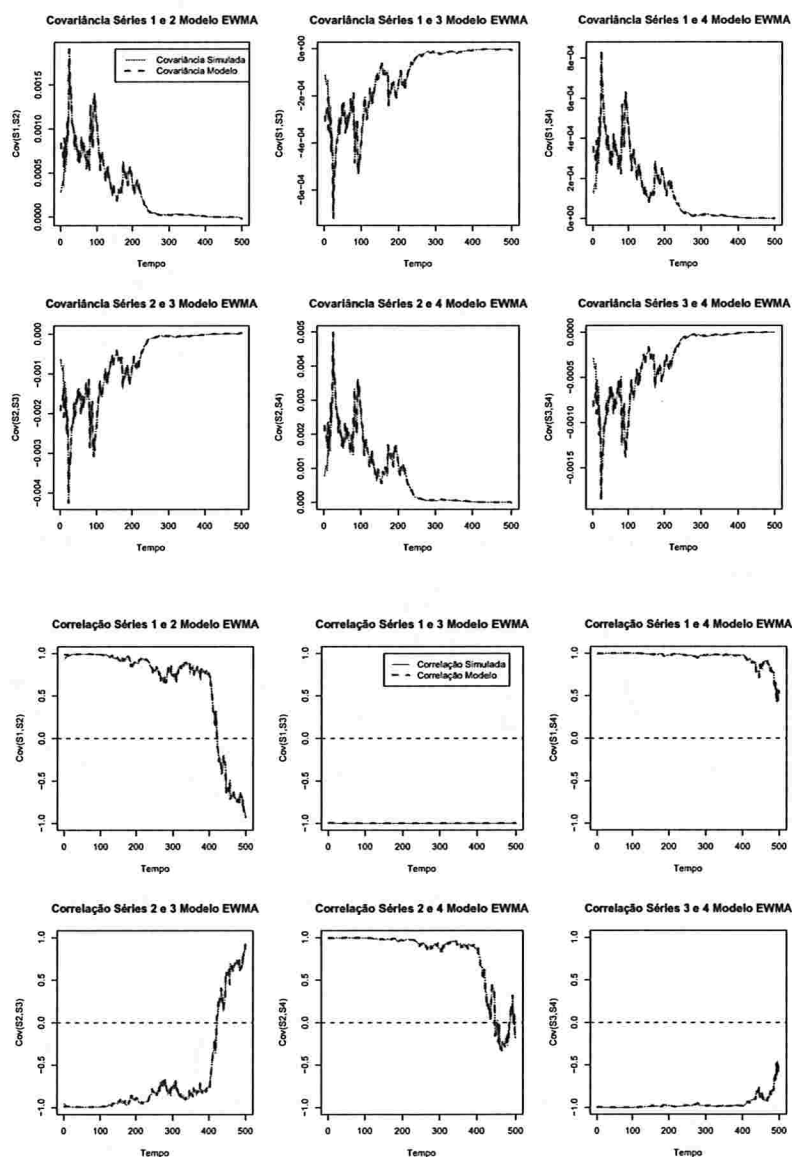


Figura 4.13: Covariâncias e correlações condicionais da simulação EWMA para $\lambda = 0,90$.

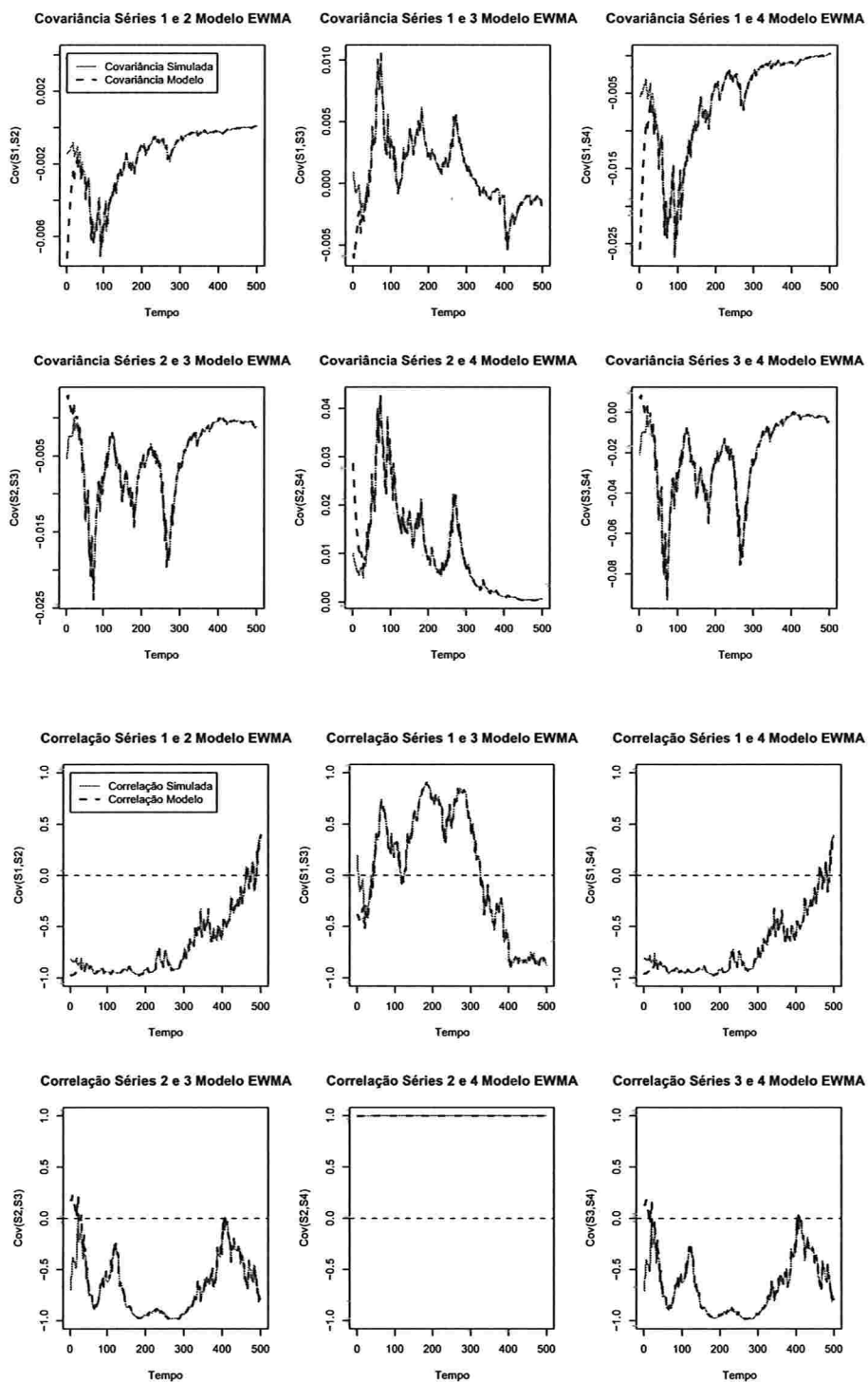


Figura 4.14: Covariâncias e correlações condicionais da simulação EWMA para $\lambda = 0,92$.

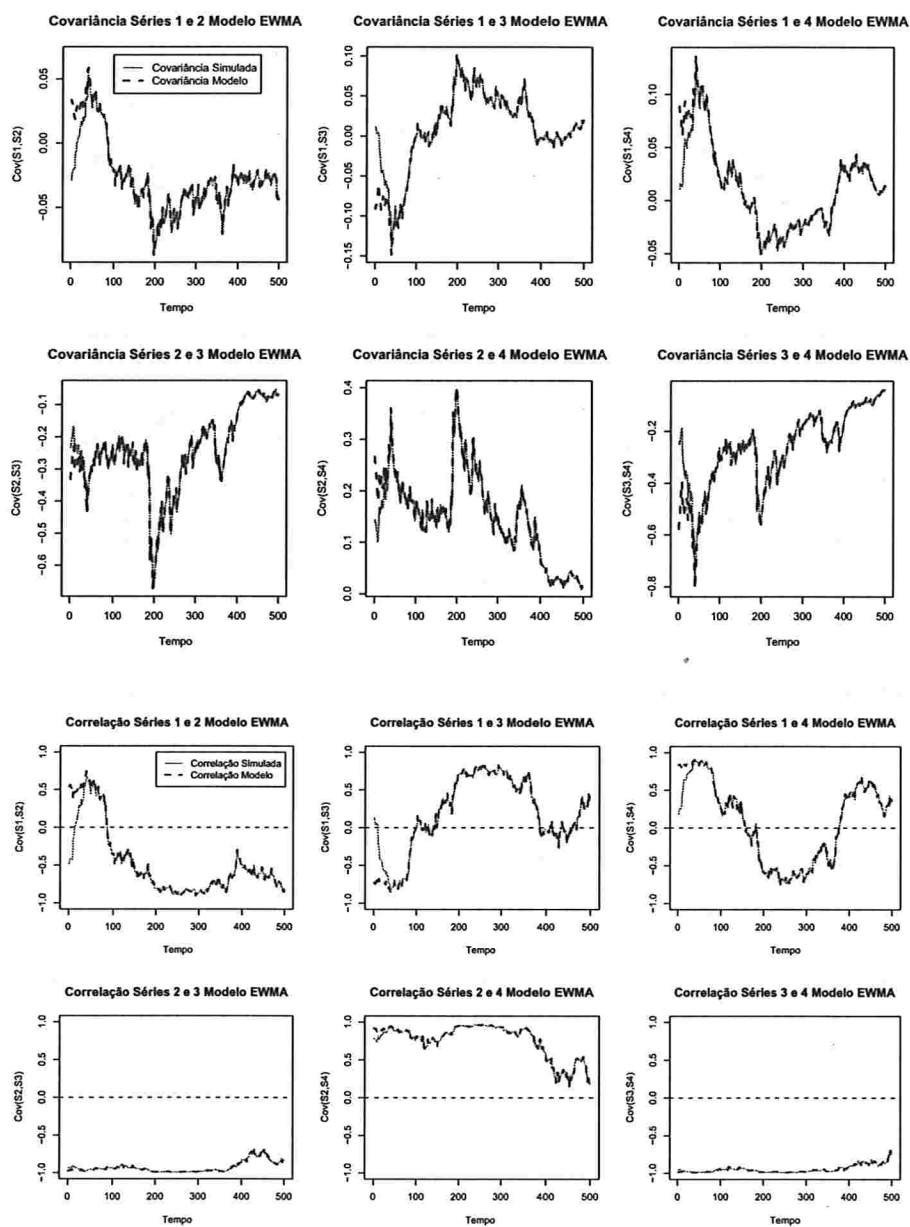


Figura 4.15: Covariâncias e correlações condicionais da simulação EWMA para $\lambda = 0,94$.

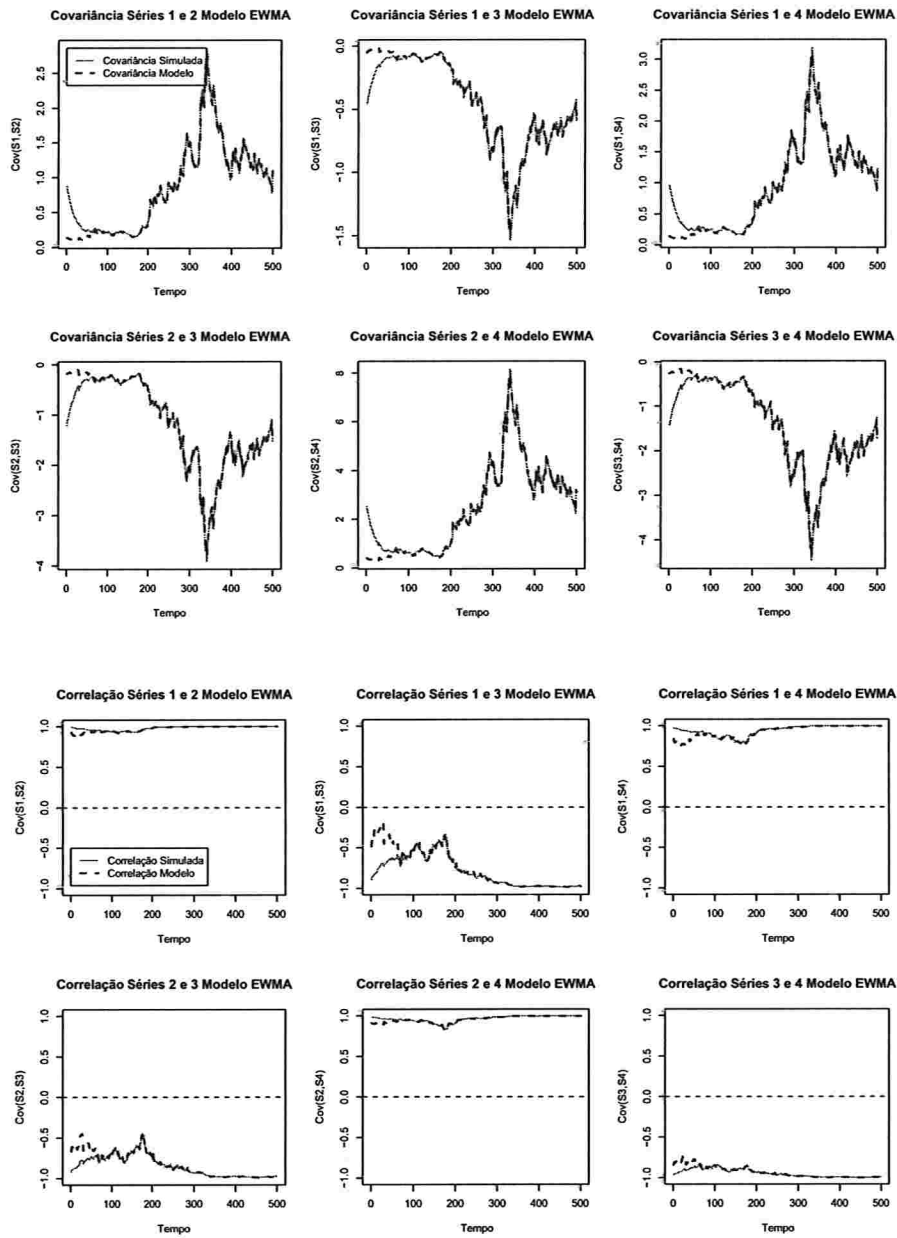


Figura 4.16: Covariâncias e correlações condicionais da simulação EWMA para $\lambda = 0,96$.

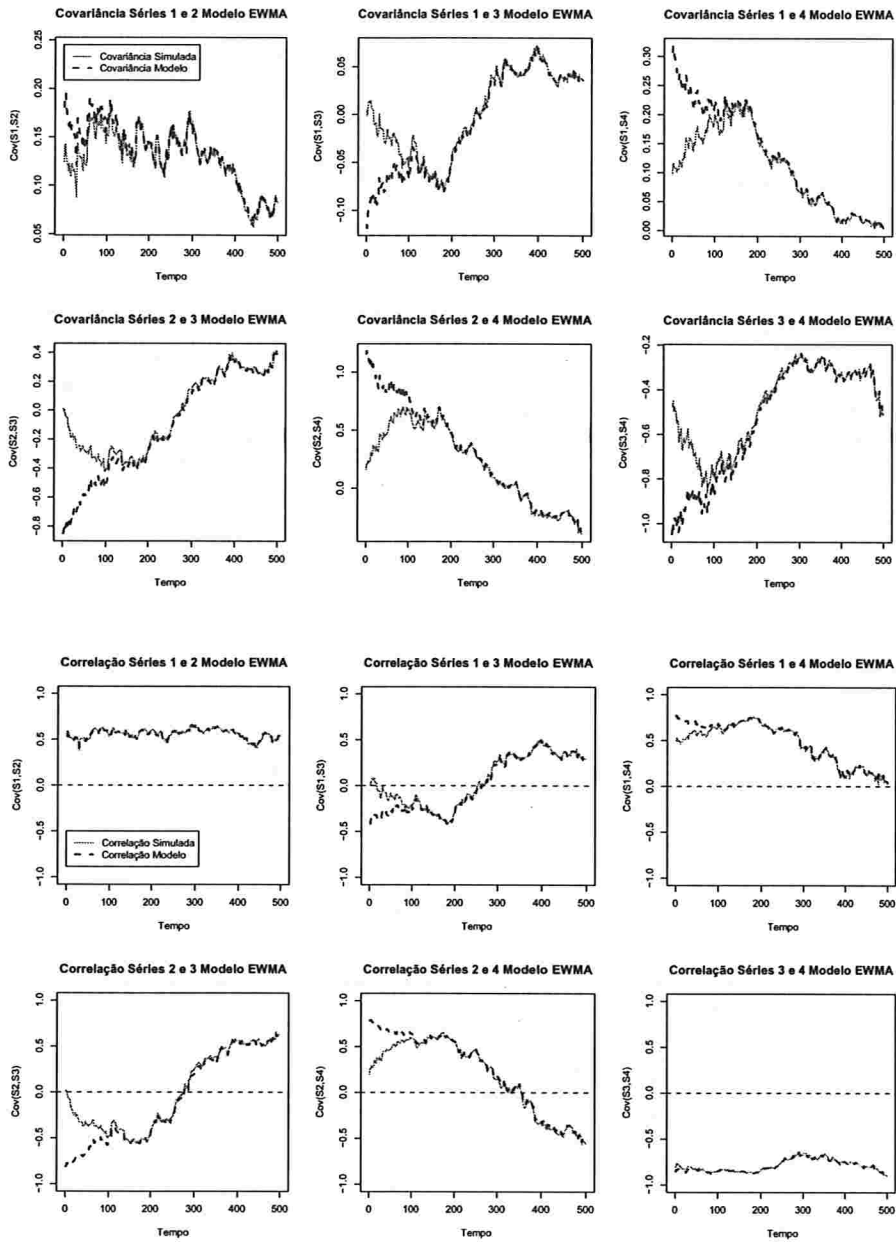


Figura 4.17: Covariâncias e correlações condicionais da simulação EWMA para $\lambda = 0,98$.

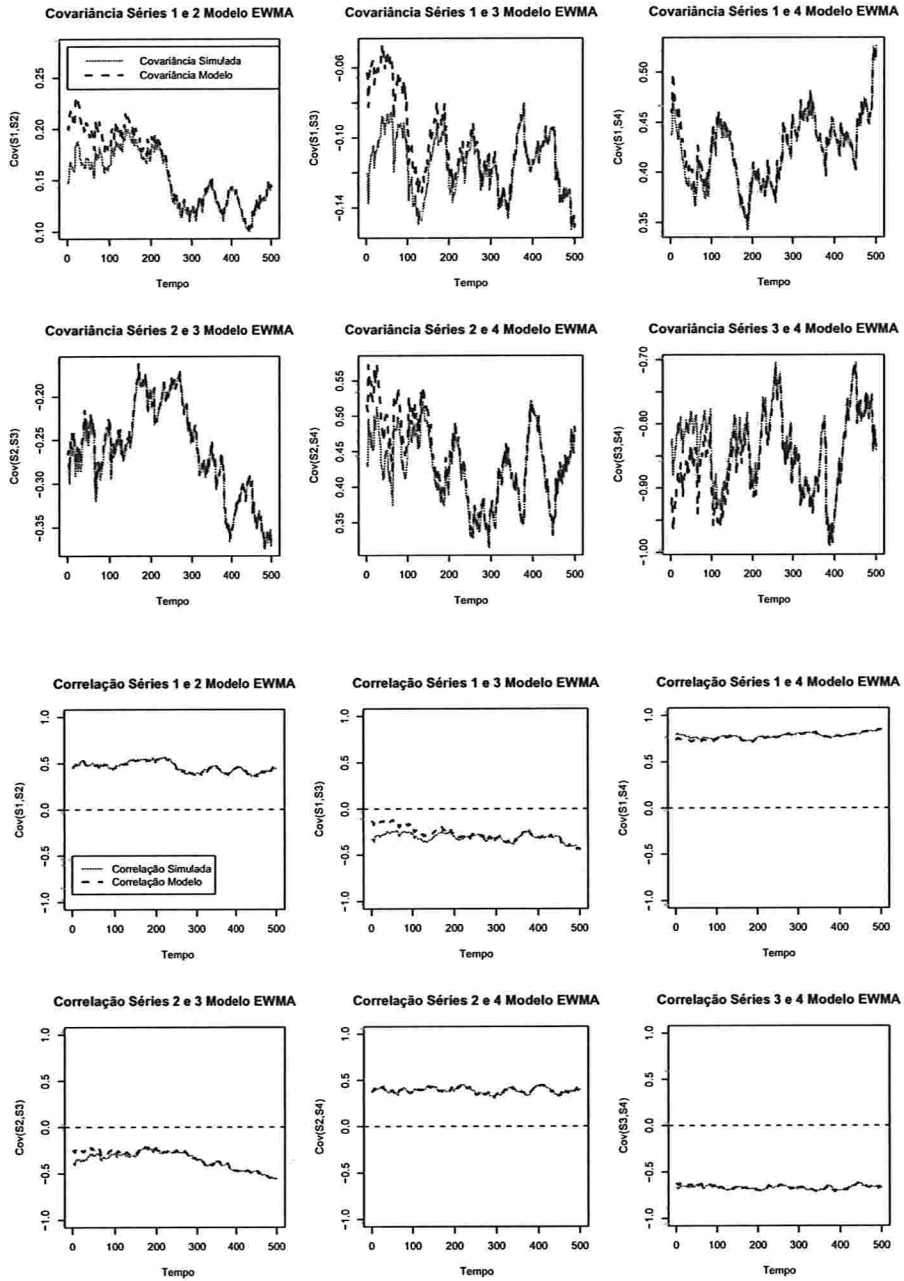


Figura 4.18: Covariâncias e correlações condicionais da simulação EWMA para $\lambda = 0,99$.

Foram feitas simulações de Monte Carlo do modelo EWMA para valores de $\lambda=0,90, 0,92, 0,94, 0,96, 0,98$ e $0,99$.

Apresentam-se na Tabela 4.3 as porcentagens de aceitação dos testes de Portmanteau para 10 , 100 e 1000 réplicas e observa-se que estas aumentam quando $\lambda \rightarrow 1$. Isto confirma o fato que quando $\lambda \rightarrow 1$ os ajustes EWMA conseguem descrever melhor o efeito ARCH, dado na equação (3.4) pelo termo

$$(1 - \lambda)\varepsilon_{t-1}\varepsilon'_{t-1}.$$

Tabela 4.3: Porcentagens de aceitação dos testes de Portmanteau Monte Carlo do modelo EWMA

λ	10 Réplicas	100 Réplicas	1000 Réplicas	% Apróximada
0.90	0	3	10	1,333
0,92	0	5	34	2,800
0,94	1	19	247	17,900
0,96	4	33	415	38,167
0,98	5	49	670	55,333
0,99	8	85	869	83,967

A Tabela 4.4 apresenta os parâmetros verdadeiros e estimados para 10, 100 e 1000 réplicas, junto com os seus respectivos erros quadráticos médios (EQM). Nota-se que já para 10 réplicas os parâmetros estimados convergiram aos parâmetros verdadeiros, mostrando que com poucas réplicas há convergência.

Tabela 4.4: Parâmetros estimados Monte Carlo Modelo EWMA.

$\lambda = 0,90$		10 Réplicas		100 Réplicas		1000 Réplicas	
	Verd.	Sim.	EQM	Sim.	EQM	Sim.	EQM
C(1)	0,000	0,000	0,000	0,001	0,000	0,000	0,000
C(2)	0,000	0,000	0,000	0,002	0,001	0,000	0,000
C(3)	0,000	0,007	0,000	0,001	0,000	0,001	0,000
C(4)	0,000	0,010	0,000	0,003	0,004	0,000	0,001
λ	0,900	0,918	0,010	0,911	0,005	0,909	0,001
$\lambda = 0,92$							
	Verd.	Sim.	EQM	Sim.	EQM	Sim.	EQM
C(1)	0,000	0,001	0,000	0,000	0,000	0,002	0,001
C(2)	0,000	0,020	0,005	-0,003	0,001	0,000	0,000
C(3)	0,000	0,002	0,000	-0,001	0,001	0,001	0,000
C(4)	0,000	0,008	0,000	0,001	0,001	0,000	0,000
λ	0,920	0,933	0,012	0,926	0,000	0,921	0,000
$\lambda = 0,94$							
	Verd.	Sim.	EQM	Sim.	EQM	Sim.	EQM
C(1)	0,000	-0,002	0,000	0,001	0,000	0,000	0,000
C(2)	0,000	0,000	0,000	0,000	0,000	0,000	0,000
C(3)	0,000	0,000	0,000	-0,002	0,001	0,001	0,000
C(4)	0,000	-0,0040	0,000	0,002	0,003	0,000	0,000
λ	0,940	0,942	0,000	0,941	0,000	0,940	0,000
$\lambda = 0,96$							
	Verd.	Sim.	EQM	Sim.	EQM	Sim.	EQM
C(1)	0,000	0,005	0,000	-0,001	0,001	0,001	0,000
C(2)	0,000	0,002	0,003	0,002	0,007	0,002	0,001
C(3)	0,000	0,001	0,002	-0,001	0,005	-0,003	0,004
C(4)	0,000	0,001	0,001	0,010	0,005	0,001	0,001
λ	0,960	0,960	0,000	0,961	0,000	0,960	0,000
$\lambda = 0,98$							
	Verd.	Sim.	EQM	Sim.	EQM	Sim.	EQM
C(1)	0,000	0,004	0,000	0,001	0,000	0,000	0,000
C(2)	0,000	0,001	0,001	-0,001	0,001	0,000	0,000
C(3)	0,000	-0,003	0,008	0,004	0,003	0,000	0,000
C(4)	0,000	0,003	0,009	0,000	0,000	0,001	0,000
λ	0,980	0,980	0,000	0,981	0,000	0,980	0,000
$\lambda = 0,99$							
	Verd.	Sim.	EQM	Sim.	EQM	Sim.	EQM
C(1)	0,000	-0,001	0,003	0,000	0,000	0,000	0,000
C(2)	0,000	-0,004	0,007	0,008	0,005	0,000	0,000
C(3)	0,000	0,009	0,007	0,001	0,006	0,005	0,001
C(4)	0,000	-0,001	0,005	0,000	0,001	-0,002	0,004
λ	0,990	0,993	0,000	0,992	0,000	0,990	0,000

4.2 Modelo DVEC(1,1)

Para simular o modelo DVEC(1,1) considere o modelo dado em (3.9), junto com a equação (3.1), com $\mathbf{c} = \mathbf{0}$, isto é,

$$\mathbf{r}_t = \boldsymbol{\varepsilon}_t, \quad (4.3)$$

$$\boldsymbol{\Sigma}_t = \mathbf{A}_0 \mathbf{A}'_0 + \mathbf{A}_1 \mathbf{A}'_1 \odot (\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}'_t) + \mathbf{B}_1 \mathbf{B}'_1 \odot \boldsymbol{\Sigma}_{t-1}, \quad (4.4)$$

em que $\boldsymbol{\varepsilon}_t$ é um vetor (4×1) normal multivariado, $N_4(\mathbf{0}, \boldsymbol{\Sigma}_t)$ e $t = 1, \dots, 500$, isto é, considera-se um tamanho amostral $T = 500$. As matrizes de parâmetros e a matriz inicial $\boldsymbol{\Sigma}_0$ foram escolhidas de tal forma que as matrizes de covariâncias condicionais sejam PSD, e são dadas por

$$\mathbf{A}_0 \mathbf{A}'_0 = \begin{bmatrix} 0,040 & -0,0460 & 0,05200 & 0,046000 \\ -0,046 & 0,1154 & -0,12730 & 0,019600 \\ 0,052 & -0,1273 & 0,18890 & 0,028360 \\ 0,046 & 0,0196 & 0,02836 & 0,249969 \end{bmatrix},$$

$$\mathbf{A}_1 \mathbf{A}'_1 = \begin{bmatrix} 0,11560 & 0,106760 & 0,112200 & 0,065620 \\ 0,10676 & 0,210821 & 0,213165 & 0,167467 \\ 0,11220 & 0,213165 & 0,332793 & 0,285309 \\ 0,06562 & 0,167467 & 0,285309 & 0,398035 \end{bmatrix},$$

$$\mathbf{B}_1 \mathbf{B}'_1 = \begin{bmatrix} 0,02250 & 0,018150 & 0,029100 & 0,018150 \\ 0,01815 & 0,028330 & 0,038918 & 0,030436 \\ 0,02910 & 0,038918 & 0,072221 & 0,058455 \\ 0,01815 & 0,030436 & 0,058455 & 0,084996 \end{bmatrix}, \quad \boldsymbol{\Sigma}_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

As matrizes \mathbf{A}_1 e \mathbf{B}_1 foram escolhidas segundo a equação (3.12) de tal forma que os autovalores de

$$\text{diag}\{\text{vec}(\mathbf{A}_1 \mathbf{A}'_1)\} + \text{diag}\{\text{vec}(\mathbf{B}_1 \mathbf{B}'_1)\} \quad (4.5)$$

sejam menores que um, garantindo assim que o processo seja estacionário¹ (neste caso o maior autovalor foi 0,483031). Para obter as matrizes \mathbf{A}_0 , \mathbf{A}_1 e \mathbf{B}_1 aplica-se a decomposição de Cholesky às matrizes $\mathbf{A}_0 \mathbf{A}'_0$, $\mathbf{A}_1 \mathbf{A}'_1$ e $\mathbf{B}_1 \mathbf{B}'_1$ respectivamente. A matriz $\boldsymbol{\Sigma}_0$ foi considerada como a matriz identidade, para controlar a variabilidade no início da simulação.

¹Ver a Proposição 3.3.

As séries simuladas estão apresentadas na Figura 4.19. Nota-se que não existe tendência em nenhuma das séries e que existem alguns conglomerados, ao longo do tempo, em que a variabilidade é maior.

A Figura 4.20 apresenta as FAC e FAC cruzadas para as séries ao quadrado e pode-se observar que existem correlações cruzadas significativas, sugerindo uma dependência quadrática nas séries. Ajusta-se um modelo DVEC(1,1) usando o software S-PLUS e são apresentados na Tabela 4.5 os parâmetros estimados.

Como o número de parâmetros no modelo DVEC(1,1) é 34 (muito maior que para os modelos EWMA), e o tamanho da amostra é 500, considera-se um nível de significância de 10%. Nota-se que os parâmetros estimados relacionados com o vetor constante \mathbf{c} e o valor $A(4, 2)$ da matriz constante $\mathbf{A}_0\mathbf{A}'_0$ são estatisticamente zero. Na parte ARCH, matriz $\mathbf{A}_1\mathbf{A}'_1$, todos os parâmetros estimados são significativos. Já para a parte GARCH, matriz $\mathbf{B}_1\mathbf{B}'_1$, os parâmetros estimados GARCH(1; 1, 1), GARCH(1; 2, 1), GARCH(1; 3, 1), GARCH(1; 4, 1) e GARCH(1; 2, 2) são não significativos.

Observando a Figura 4.20, nota-se que as autocorrelações cruzadas da Série 1 com as outras séries não é muito forte e isto pode estar relacionado com o fato que os parâmetros da parte GARCH relacionados com a Série 1 sejam não significativos.

A Tabela 4.6 apresenta os testes do ajuste. Todos são testes univariados e segundo os p -valores, não rejeitam-se as hipóteses nulas de normalidade (testes de Jarque-Bera e Shapiro-Wilk) e de não autocorrelação serial para os resíduos e o quadrado deles, considerando as séries individualmente (testes de Ljung-Box e Multiplicadores de Lagrange), ao nível de significância de 10%.

As Figuras 4.21 e 4.22 apresentam as FAC e FAC cruzadas do quadrado dos resíduos e o QQ-plot deles, respectivamente, e observa-se que não existe autocorrelação e autocorrelação cruzada entre eles e que seguem uma distribuição normal.

Para verificar se existe autocorrelação e autocorrelação cruzada entre os resíduos ao quadrado do modelo ajustado, considera-se o teste multivariado de Portmanteau, apresentado na Tabela 4.7. O p -valor do teste foi 0,7057, não rejeitando a hipótese nula de não autocorrelação e autocorrelação cruzada nos resíduos ao quadrado, indicando que o modelo é adequado.

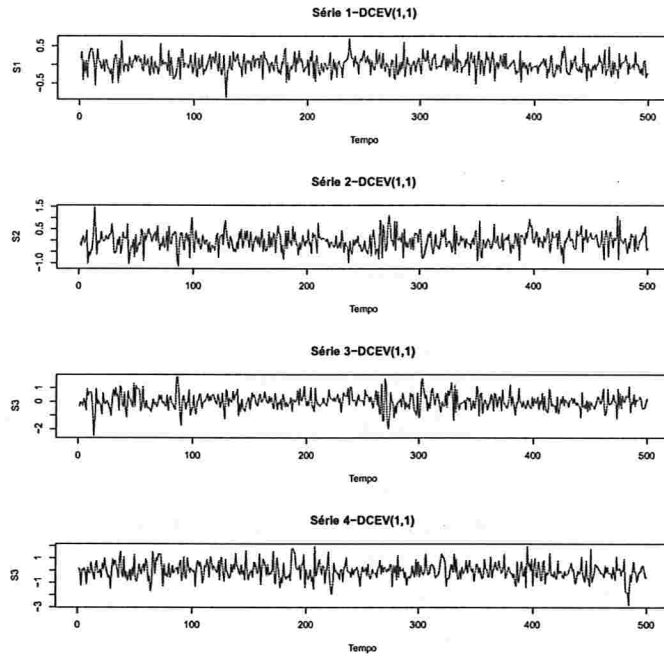


Figura 4.19: Séries simuladas DVEC(1,1).

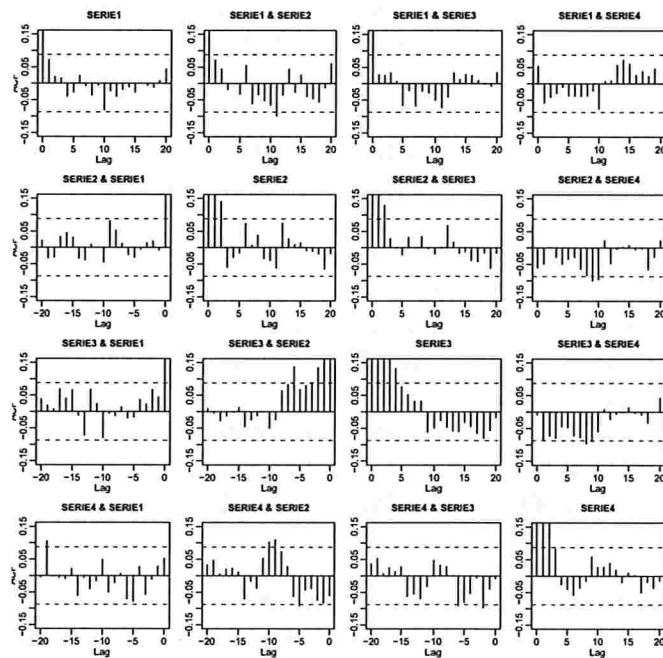


Figura 4.20: FAC e FAC cruzadas das séries simuladas ao quadrado DVEC(1,1).

Tabela 4.5: Parâmetros estimados DVEC(1,1).

Parâmetros Estimados DVEC(1,1):				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	0,004282	0,009030	0,4742	0,635
C(2)	-0,008126	0,013915	-0,5840	0,559
C(3)	0,008128	0,020510	0,3963	0,692
C(4)	-0,008185	0,027087	-0,3022	0,763
A(1, 1)	0,035296	0,007186	4,9117	<0,001
A(2, 1)	-0,036955	0,006998	-5,2811	<0,001
A(3, 1)	0,041752	0,010404	4,0131	<0,001
A(4, 1)	0,031032	0,016702	1,8580	0,064
A(2, 2)	0,095054	0,011623	8,1780	<0,001
A(3, 2)	-0,093142	0,013348	-6,9778	<0,001
A(4, 2)	0,007810	0,006437	1,2133	0,223
A(3, 3)	0,157265	0,025285	6,2196	<0,001
A(4, 3)	0,033141	0,011499	2,8821	0,004
A(4, 4)	0,220885	0,045897	4,8126	<0,001
ARCH(1; 1, 1)	0,186177	0,042504	4,3802	<0,001
ARCH(1; 2, 1)	0,144184	0,036646	3,9344	<0,001
ARCH(1; 3, 1)	0,137166	0,037539	3,6539	<0,001
ARCH(1; 4, 1)	0,064144	0,038344	1,6728	0,095
ARCH(1; 2, 2)	0,271642	0,043058	6,3088	<0,001
ARCH(1; 3, 2)	0,242065	0,042313	5,7208	<0,001
ARCH(1; 4, 2)	0,119629	0,038991	3,0681	0,002
ARCH(1; 3, 3)	0,316531	0,055493	5,7040	<0,001
ARCH(1; 4, 3)	0,228152	0,042003	5,4318	<0,001
ARCH(1; 4, 4)	0,32023	0,05864	5,4607	<0,001
GARCH(1; 1, 1)	0,06084	0,15090	0,4032	0,687
GARCH(1; 2, 1)	0,15158	0,12281	1,2343	0,218
GARCH(1; 3, 1)	0,19711	0,16132	1,2219	0,222
GARCH(1; 4, 1)	0,34924	0,32684	1,0685	0,286
GARCH(1; 2, 2)	0,06243	0,06717	0,9296	0,353
GARCH(1; 3, 2)	0,16057	0,08447	1,9009	0,058
GARCH(1; 4, 2)	0,43639	0,13965	3,1249	0,002
GARCH(1; 3, 3)	0,16067	0,09599	1,6739	0,095
GARCH(1; 4, 3)	0,32247	0,09513	3,3898	0,001
GARCH(1; 4, 4)	0,23741	0,11111	2,1367	0,033

A Figura 4.23 apresenta as variâncias e os desvios padrão condicionais do modelo simulado (linha contínua) e do modelo ajustado (linha pontilhada). Observa-se que o comportamento é similar nas variâncias simuladas e ajustadas.

Tabela 4.6: Testes DVEC(1,1).

Testes da Simulação DVEC(1,1):						
AIC(34) = 1068,597						
BIC(34) = 1211,894						
Testes de Normalidade:						
	Jarque-Bera	P-valor	Shapiro-Wilk	P-valor		
serie1.DV11	0,1424	0,9313	0,9881	0,8289		
serie2.DV11	0,1421	0,9314	0,9854	0,5435		
serie3.DV11	0,2474	0,8836	0,9874	0,7633		
serie4.DV11	0,6406	0,7259	0,9871	0,7368		
Teste de Ljung-Box para os resíduos padronizados:						
	Estatística	P-valor	$\chi^2 - d, f,$			
serie1.DVEC11	8,778	0,7217	12			
serie2.DVEC11	9,622	0,6491	12			
serie3.DVEC11	9,236	0,6826	12			
serie4.DVEC11	7,409	0,8295	12			
Teste de Ljung-Box para o quadrado dos resíduos padronizados:						
	Estatística	P-valor	$\chi^2 - d, f,$			
serie1.BEKK11,s	8,719	0,7268	12			
serie2.BEKK11,s	7,440	0,8272	12			
serie3.BEKK11,s	10,028	0,6135	12			
serie3.BEKK11,s	6,150	0,9083	12			
Teste multiplicadores de Lagrange:						
	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
serie1.DV11	-0,004019	0,4535	-0,9134	-0,3643	0,4883	-0,06215
serie2.DV11	1,254669	0,3366	-0,9868	0,1900	0,5217	0,48808
serie3.DV11	-1,482287	-2,2147	-1,2507	0,5318	-0,2048	-0,24011
serie4.DV11	-0,348562	1,1291	0,4352	0,1346	-0,1831	-0,69402
	Lag 7	Lag 8	Lag 9	Lag 10	Lag 11	Lag 12
serie1.DV11	-1,0836	-0,1282	-1,8567	-0,5462	-1,0661	-0,28188
serie2.DV11	0,3652	1,8958	-0,1170	0,5439	0,5529	-0,03941
serie3.DV11	0,2221	-0,8547	0,2938	0,6066	0,4193	0,31935
serie4.DV11	-0,1208	-0,6611	-1,2244	0,9127	0,6273	0,01217
	C					
serie1.DV11	-1,3785					
serie2.DV11	0,1543					
serie3.DV11	-1,1401					
serie4.DV11	-0,4099					
	TR^2	P-valor	F-estat.	P-valor		
serie1.DV11	9,300	0,6771	0,8619	0,6876		
serie2.DV11	7,996	0,7854	0,7390	0,8080		
serie3.DV11	11,215	0,5106	1,0436	0,5161		
serie4.DV11	5,626	0,9337	0,5174	0,9666		

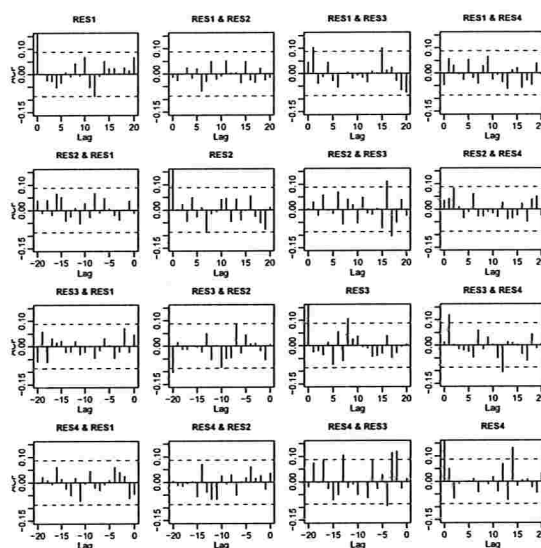


Figura 4.21: *FAC e FAC cruzadas do quadrado dos resíduos do ajuste DVEC(1,1).*

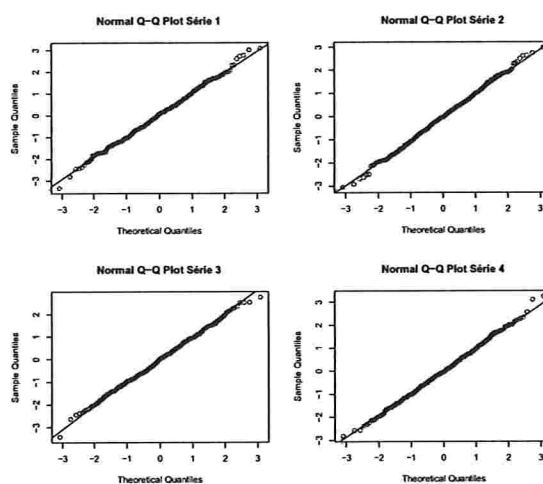


Figura 4.22: *QQ-Plot dos resíduos do ajuste DVEC(1,1).*

As covariâncias condicionais ajustadas e simuladas estão na Figura 4.24 e comparando estas, nota-se que o comportamento é similar em cada caso. Em comparação com o modelo EMWA, não existe uma tendência nas correlações condicionais no modelo DVEC(1,1), já que este é estacionário.

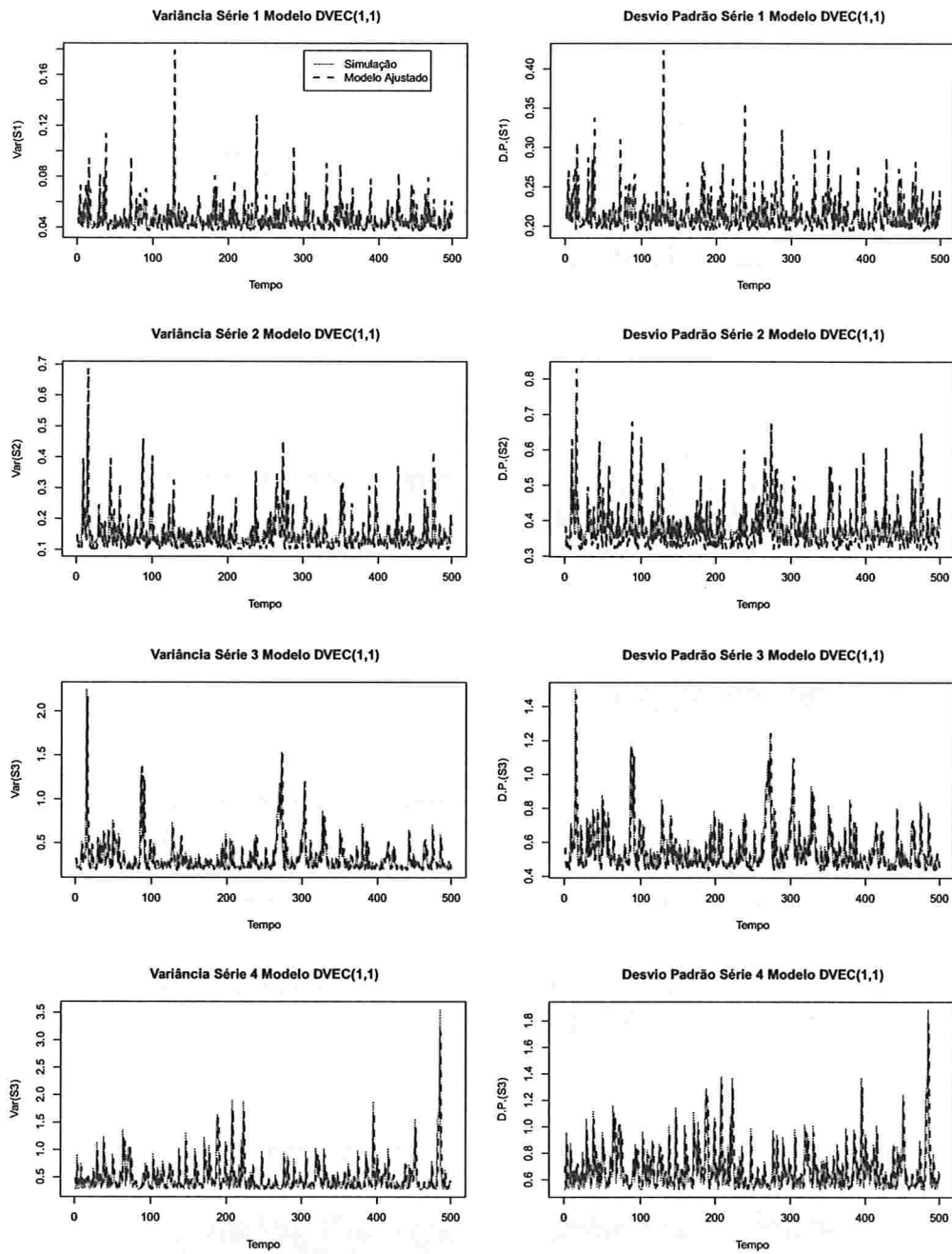


Figura 4.23: Variâncias e desvios padrão condicionais DVEC(1,1).

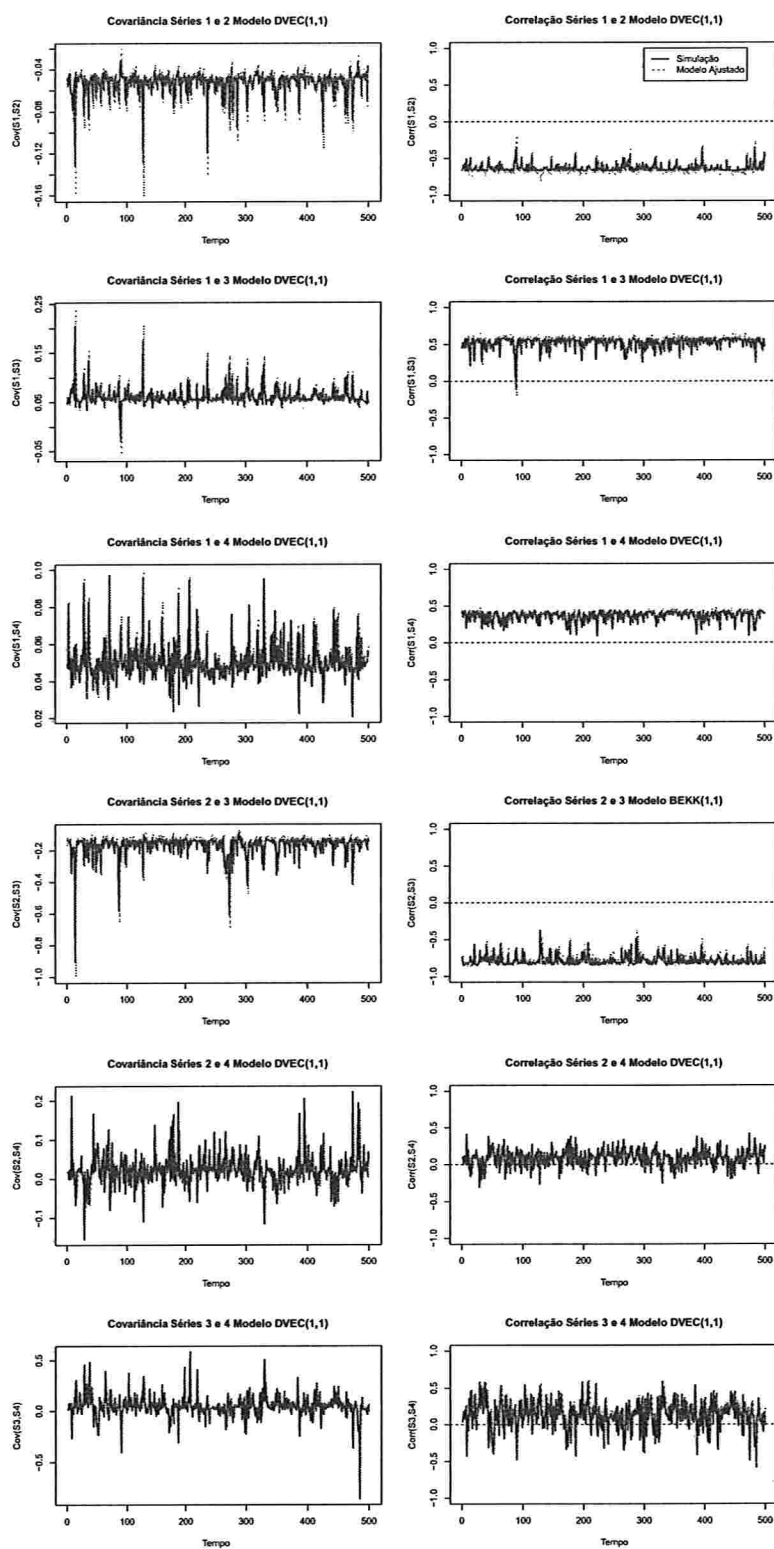


Figura 4.24: Covariâncias e correlações condicionais do ajuste $DVEC(1,1)$.

Tabela 4.7: Teste multivariado de Portmanteau para o quadrado dos resíduos do ajuste DVEC(1,1).

Teste Multivariado Portmanteau: Tipo Ljung-Box	
Hipótese Nula: não correlação serial	
Estat. do teste	p.valor
180,9527	0,7057
Dist. sob Hipótese Nula: qui-quadrado com 192 graus de liberdade	
Total Observ.: 500	

Segundo a Tabela 4.5 o modelo ajustado DVEC(1,1) é dado por:

$$\begin{aligned}
 \mathbf{r}_t &= \begin{pmatrix} 0,004282 \\ -0,008126 \\ 0,008128 \\ -0,008185 \end{pmatrix} + \boldsymbol{\varepsilon}_t \\
 \boldsymbol{\Sigma}_t &= \begin{pmatrix} 0,035296 & -0,036955 & 0,041752 & 0,031032 \\ -0,036955 & 0,095054 & -0,093142 & 0,007810 \\ 0,041752 & -0,093142 & 0,157265 & 0,033141 \\ 0,031032 & 0,007810 & 0,033141 & 0,220885 \end{pmatrix} \\
 &+ \begin{pmatrix} 0,186177 & 0,144184 & 0,137166 & 0,064144 \\ 0,144184 & 0,271642 & 0,242065 & 0,119629 \\ 0,137166 & 0,242065 & 0,316531 & 0,228152 \\ 0,064144 & 0,119629 & 0,228152 & 0,320230 \end{pmatrix} \odot (\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}'_{t-1}) \\
 &+ \begin{pmatrix} 0,06084 & 0,15158 & 0,19711 & 0,34924 \\ 0,15158 & 0,06243 & 0,16057 & 0,43639 \\ 0,19711 & 0,16057 & 0,16067 & 0,32247 \\ 0,34924 & 0,43639 & 0,32247 & 0,23741 \end{pmatrix} \odot \boldsymbol{\Sigma}_{t-1},
 \end{aligned}$$

em que $\boldsymbol{\varepsilon}_t \sim \mathbf{N}_4(\mathbf{0}, \boldsymbol{\Sigma}_t)$.

Para analisar o comportamento dos parâmetros, foram feitas simulações Monte Carlo com 10, 100 e 1000 réplicas do modelo DVEC(1,1) dado por (4.3) e (4.4), considerando tamanhos amostrais $T = 500$ e $T = 5000$. Em cada caso foi considerada a média dos parâmetros nas réplicas e os p -valores do teste multivariado de Portmanteau.

A Tabela 4.8 apresenta os parâmetros verdadeiros (" v ") e estimados (" s ") com 10, 100 e 1000 réplicas, além do erro quadrático médio ("EQM") para cada caso, considerando um tamanho amostral $T = 500$. Os parâmetros com um EQM maior foram os correspondentes à parte GARCH, $B_1 B_1'$. Na Figura 4.25, apresentam-se os gráficos das estimativas (" s ") e os valores dos parâmetros verdadeiros (" v "), para 10, 100 e 1000 réplicas. Nota-se que graficamente os parâmetros da parte GARCH, estão sendo sobrestimados, o qual pode ser causado pelo algoritmo que usa o software S-PLUS na estimação destes. Além disso, com o número de réplicas aumentando as estimativas dos parâmetros GARCH não foram melhores, já que para 100 e 1000 réplicas a convergência foi similar que com 10 réplicas. Por outro lado, aumentando o tamanho da série para $T = 5000$, tem-se que os valores obtidos para os parâmetros na parte GARCH também não foram melhores, conforme os resultados apresentados na Tabela 4.9 e a Figura 4.26.

Observando os gráficos dos p -valores do teste Portmanteau para tamanhos amostrais $T = 500$ e $T = 5000$, nas Figuras 4.25 e 4.26, nota-se que a maioria destes p -valores ficaram acima do valor 0,05, com linha pontilhada, em cada caso. No caso em que $T = 500$, para 10 réplicas os 10 p -valores foram significantes, para 100 réplicas 93 deles foram significantes e para 1000 réplicas 934 foram significantes. No caso em que $T = 5000$, para 10 réplicas 9 p -valores foram significantes, para 100 réplicas 95 deles foram significantes e para 1000 réplicas 953 foram significantes, indicando que aproximadamente num 94.45% das vezes o modelo ajustado pelo S-PLUS consegue ajustar bem o modelo DVEC(1,1) dado por (4.3) e (4.4), com as matrizes escolhidas neste caso. A Tabela 4.10 apresenta o resumo destas porcentagens.

Tabela 4.8: Parâmetros estimados Monte Carlo DVEC(1,1) com tamanho amostral $T = 500$.

	Verd. ("v")	10 Réplicas		100 Réplicas		1000 Réplicas	
		Sim. ("s")	EQM	Sim. ("s")	EQM	Sim. ("s")	EQM
c	0,0000	0,0019	0,0001	-0,0001	0,0001	0,0005	0,0001
	0,0000	-0,0021	0,0003	-0,0010	0,0002	-0,0004	0,0002
	0,0000	-0,0016	0,0004	0,0015	0,0003	-0,0003	0,0004
	0,0000	0,0039	0,0002	-0,0019	0,0005	0,0005	0,0006
A₀A'₀	0,0400	0,0352	0,0004	0,0376	0,0002	0,0349	0,0002
	-0,0460	-0,0373	0,0005	-0,0421	0,0002	-0,0398	0,0002
	0,0520	0,0420	0,0003	0,0466	0,0003	0,0459	0,0003
	0,0460	0,0402	0,0006	0,0416	0,0007	0,0370	0,0006
	0,1154	0,1165	0,0005	0,1109	0,0004	0,1064	0,0005
	-0,1273	-0,1241	0,0006	-0,1216	0,0005	-0,1178	0,0007
	0,0196	0,0237	0,0001	0,0169	0,0001	0,0180	0,0001
	0,1889	0,1907	0,0010	0,1862	0,0009	0,1798	0,0010
	0,0284	0,0224	0,0005	0,0277	0,0001	0,0272	0,0002
	0,2500	0,2520	0,0006	0,2457	0,0018	0,2428	0,0016
A₁A'₁	0,1156	0,0796	0,0027	0,1007	0,0022	0,1024	0,0019
	0,1068	0,0895	0,0026	0,0980	0,0016	0,0950	0,0013
	0,1122	0,0940	0,0028	0,1074	0,0020	0,1012	0,0016
	0,0656	0,0483	0,0021	0,0545	0,0017	0,0565	0,0016
	0,2108	0,2006	0,0033	0,1956	0,0021	0,1966	0,0021
	0,2132	0,2079	0,0023	0,1998	0,0021	0,1988	0,0022
	0,1675	0,1756	0,0034	0,1497	0,0026	0,1529	0,0022
	0,3328	0,3088	0,0030	0,3123	0,0043	0,3118	0,0047
	0,2853	0,2753	0,0039	0,2609	0,0041	0,2653	0,0037
	0,3980	0,3914	0,0063	0,3536	0,0084	0,3707	0,0076
B₁B'₁	0,0225	0,1507	0,1881	0,0915	0,1185	0,1385	0,1106
	0,0182	0,2087	0,1781	0,1038	0,0777	0,1394	0,0910
	0,0291	0,1930	0,1185	0,1232	0,0774	0,1332	0,0933
	0,0182	0,1227	0,1587	0,1129	0,2667	0,1995	0,2375
	0,0283	0,0383	0,0160	0,0606	0,0180	0,0919	0,0250
	0,0389	0,0539	0,0217	0,0731	0,0181	0,0993	0,0259
	0,0304	0,0483	0,0299	0,1248	0,0510	0,1068	0,0454
	0,0722	0,0648	0,0141	0,0864	0,0102	0,1110	0,0139
	0,0585	0,0413	0,0083	0,0908	0,0138	0,0984	0,0199
	0,0850	0,0806	0,0018	0,1111	0,0101	0,1127	0,0116

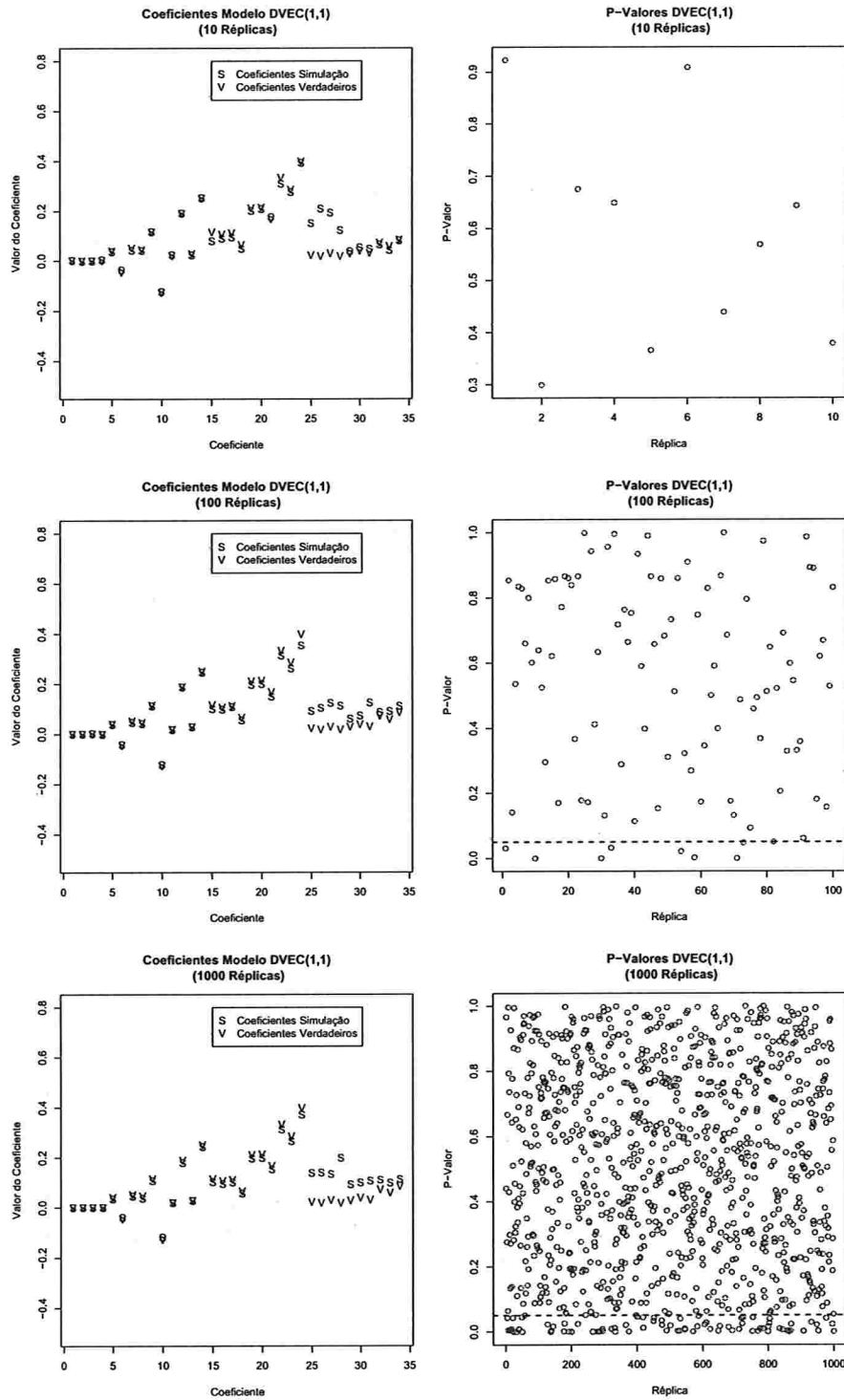


Figura 4.25: Parâmetros simulados e estimados e teste Portmanteau DVEC(1,1) com tamanho amostral $T = 500$.

Tabela 4.9: Parâmetros estimados Monte Carlo DVEC(1,1) com tamanho amostral $T = 5000$.

	Verd. ("v")	10 Réplicas		100 Réplicas		1000 Réplicas	
		Sim. ("s")	EQM	Sim. ("s")	EQM	Sim. ("s")	EQM
c	0,0000	-0,0009	0,0002	0,0001	0,0001	0,0001	0,0004
	0,0000	0,0006	0,0001	-0,0003	0,0005	-0,0002	0,0002
	0,0000	-0,0016	0,0003	0,0001	0,0006	0,0001	0,0003
	0,0000	-0,0030	0,0001	-0,0007	0,0001	-0,0002	0,0004
A₀A'₀	0,0400	0,0386	0,0004	0,0401	0,0004	0,0401	0,0007
	-0,0460	-0,0447	0,0006	-0,0448	0,0005	-0,0461	0,0003
	0,0520	0,0502	0,0007	0,0511	0,0003	0,0521	0,0001
	0,0460	0,0441	0,0002	0,0455	0,0003	0,0468	0,0002
	0,1154	0,1140	0,0003	0,1131	0,0002	0,1151	0,0001
	-0,1273	-0,1270	0,0006	-0,1249	0,0005	-0,1257	0,0004
	0,0196	0,0188	0,0002	0,0196	0,0006	0,0193	0,0005
	0,1889	0,1897	0,0001	0,1872	0,0002	0,1882	0,0002
	0,0284	0,0286	0,0005	0,0283	0,0003	0,0281	0,0003
	0,2500	0,2465	0,0002	0,2487	0,0004	0,2497	0,0007
A₁A'₁	0,1156	0,1159	0,0000	0,1107	0,0001	0,1150	0,0002
	0,1068	0,1064	0,0001	0,1037	0,0003	0,1057	0,0004
	0,1122	0,1137	0,0004	0,1105	0,0001	0,1123	0,0007
	0,0656	0,0656	0,0001	0,0651	0,0001	0,0650	0,0003
	0,2108	0,2108	0,0001	0,2057	0,0007	0,2154	0,0002
	0,2132	0,2166	0,0002	0,2093	0,0001	0,2179	0,0001
	0,1675	0,1679	0,0003	0,1567	0,0009	0,1638	0,0007
	0,3328	0,3420	0,0002	0,3245	0,0002	0,3348	0,0003
	0,2853	0,2898	0,0001	0,2930	0,0006	0,2867	0,0005
	0,3980	0,4047	0,0003	0,3806	0,0010	0,3886	0,0020
B₁B'₁	0,0225	0,1549	0,1783	0,1090	0,1064	0,1223	0,0900
	0,0182	0,1835	0,1583	0,0939	0,0165	0,1284	0,0865
	0,0291	0,1847	0,1342	0,0403	0,1575	0,0903	0,1924
	0,0182	0,1279	0,1240	0,1101	0,1431	0,1979	0,1961
	0,0283	0,0319	0,0342	0,0593	0,0147	0,0900	0,0133
	0,0389	0,0533	0,0980	0,0714	0,0889	0,0938	0,0776
	0,0304	0,0404	0,0342	0,1133	0,0299	0,1038	0,0184
	0,0722	0,0636	0,0283	0,0782	0,0190	0,0918	0,0124
	0,0585	0,0657	0,0928	0,0801	0,0873	0,0856	0,0753
	0,0850	0,0893	0,0015	0,1163	0,0132	0,1086	0,0232

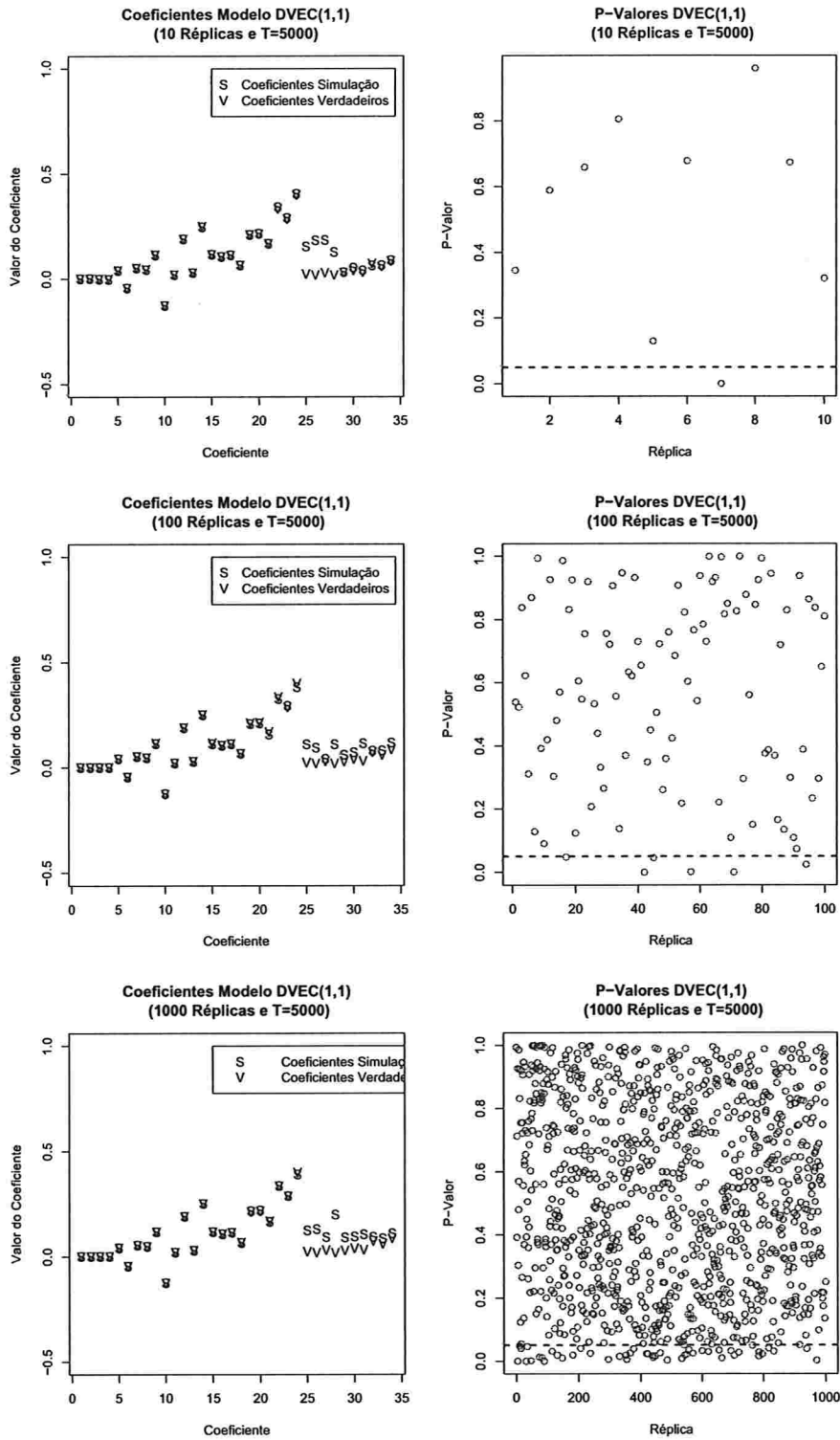


Figura 4.26: Parâmetros simulados e estimados e teste Pormanteau DVEC(1,1) com tamanho amostral $T = 5000$.

Tabela 4.10: *Porcentagens de aceitação do teste Portmanteau usando o método Monte Carlo DVEC(1,1).*

	Tamanho amostral $T = 500$			Tamanho amostral $T = 5000$		
Número de réplicas	10	100	1000	10	100	1000
Porcentagem	100%	93%	93,4%	90%	95%	95,3%

Conforme apresentado anteriormente, pode-se concluir que a parte GARCH do modelo simulado DVEC(1,1) não consegue convergir aos parâmetros reais, mas a parte constante e a parte ARCH conseguem ser bem ajustadas. Considerando outras matrizes de covariâncias iniciais, as estimativas foram iguais, descartando uma possível convergência em mínimos locais da função de log-verossimilhança.

4.3 Modelo BEKK(1,1,1)

Considere o modelo BEKK(1,1,1) dado em (3.18), junto com a equação (3.1), com $\mathbf{c} = \mathbf{0}$, isto é,

$$\mathbf{r}_t = \boldsymbol{\varepsilon}_t, \quad (4.6)$$

$$\boldsymbol{\Sigma}_t = \mathbf{C}_0 \mathbf{C}'_0 + \mathbf{C}_1 (\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}'_t) \mathbf{C}'_1 + \mathbf{D}_1 \boldsymbol{\Sigma}_{t-1} \mathbf{D}'_1, \quad (4.7)$$

em que $\boldsymbol{\varepsilon}_t$ é um vetor (3×1) normal multivariado, $\mathbf{N}_3(\mathbf{0}, \boldsymbol{\Sigma}_t)$ e $t = 1, \dots, 500$, isto é, o tamanho amostral é $T = 500$. As matrizes de parâmetros e a matriz $\boldsymbol{\Sigma}_0$ inicial são

$$\mathbf{C}_0 \mathbf{C}'_0 = \begin{bmatrix} 1,690 & 0,3340 & 0,0810 \\ 0,334 & 0,8356 & 0,2430 \\ 0,081 & 0,2430 & 0,0729 \end{bmatrix}, \quad \mathbf{C}_1 = \begin{bmatrix} 0,30 & 0,05 & -0,05 \\ -0,05 & 0,30 & -0,07 \\ 0,04 & 0,07 & 0,30 \end{bmatrix},$$

$$\mathbf{D}_1 = \begin{bmatrix} 0,90 & -0,02 & 0,04 \\ 0,05 & 0,90 & -0,05 \\ -0,04 & 0,02 & 0,90 \end{bmatrix}, \quad \boldsymbol{\Sigma}_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

A matriz $\boldsymbol{\Sigma}_0$ foi considerada como a matriz identidade para que a variabilidade no início seja mais estável. Como o modelo BEKK é muito volátil, as matrizes de parâmetros \mathbf{C}_1 e \mathbf{D}_1 possuem valores grandes nas diagonais e para que o processo seja estacionário² os parâmetros foram escolhidos de tal forma que os autovalores da matriz

$$\mathbf{C}_1 \otimes \mathbf{C}_1 + \mathbf{D}_1 \otimes \mathbf{D}_1 \quad (4.8)$$

sejam menores que um (neste caso o maior autovalor foi 0,91).

Estão na Figura 4.27 as 3 séries simuladas e observa-se que não têm tendência ao longo do tempo e que existem conglomerados em que a variabilidade é maior. Apresentam-se as FAC e FAC cruzadas das séries ao quadrado na Figura 4.28, em que existem lags em cada gráfico nos quais as correlações são significativas, confirmando a existência de uma dependência quadrática nas séries.

A Tabela 4.11 apresenta os parâmetros ajustados. Observa-se que os valores significativos são aqueles que estão nas diagonais das matrizes de parâmetros das partes ARCH

²Ver a Proposição 3.3.

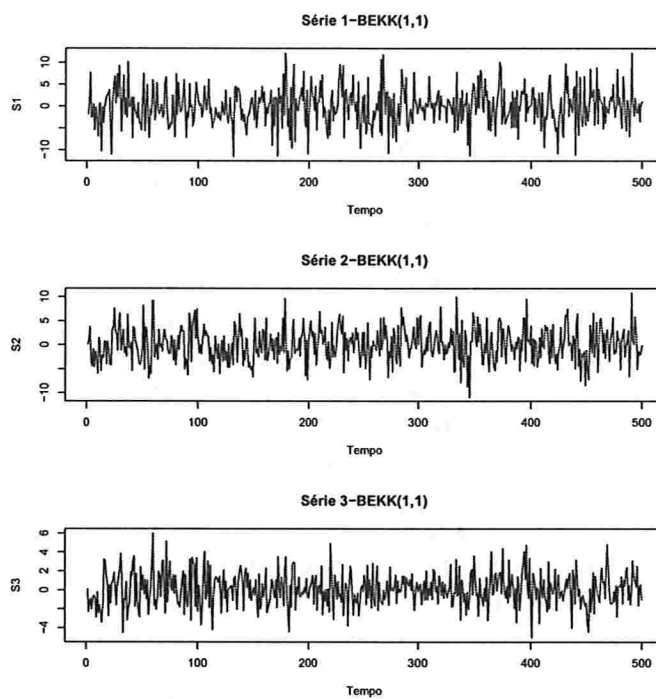


Figura 4.27: Séries simuladas BEKK(1,1,1).

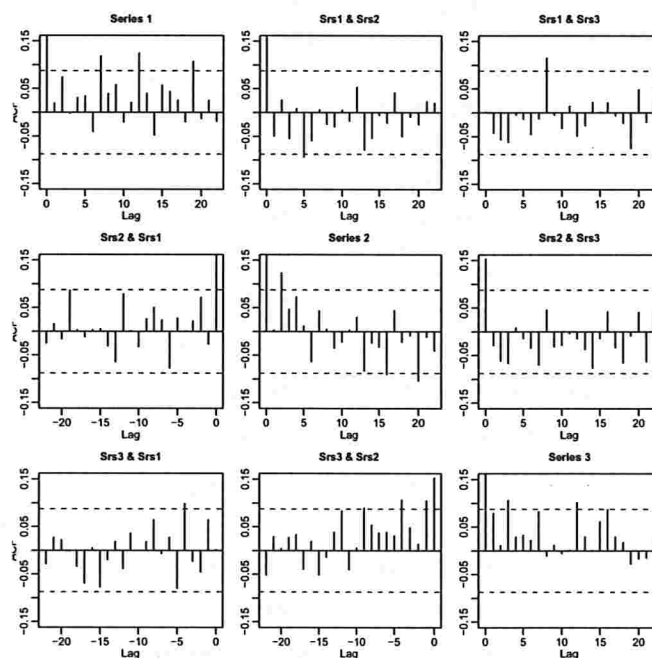


Figura 4.28: FAC e FAC cruzadas das séries simuladas ao quadrado BEKK(1,1,1).

Tabela 4.11: *Parâmetros estimados BEKK(1,1,1)*,

Parâmetros Estimados BEKK(1,1,1):				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	-0,006851	0,18880	-0,0362897	0,971
C(2)	0,120765	0,14871	0,8120701	0,417
C(3)	0,051355	0,07438	0,6904797	0,490
A(1, 1)	1,678158	0,53499	3,1368294	0,002
A(2, 1)	0,992676	0,60729	1,6346057	0,103
A(3, 1)	0,106956	0,23374	0,4575751	0,648
A(2, 2)	1,150550	0,37749	3,0479221	0,002
A(3, 2)	0,120997	0,23947	0,5052655	0,613
A(3, 3)	0,003864	25,29160	0,0001528	0,999
ARCH(1; 1, 1)	0,256540	0,09023	2,8430882	0,005
ARCH(1; 2, 1)	-0,046466	0,07559	-0,6147126	0,539
ARCH(1; 3, 1)	0,042193	0,03086	1,3670685	0,172
ARCH(1; 1, 2)	0,034211	0,11534	0,2965984	0,767
ARCH(1; 2, 2)	0,370333	0,09963	3,7170210	<0,001
ARCH(1; 3, 2)	0,078831	0,04053	1,9449509	0,052
ARCH(1; 1, 3)	-0,035204	0,16567	-0,2124929	0,832
ARCH(1; 2, 3)	-0,092381	0,15172	-0,6089025	0,543
ARCH(1; 3, 3)	0,231540	0,06301	3,6745831	<0,001
GARCH(1; 1, 1)	0,933486	0,07481	12,4784606	<0,001
GARCH(1; 2, 1)	0,062702	0,06415	0,9774720	0,329
GARCH(1; 3, 1)	-0,029965	0,02561	-1,1701318	0,242
GARCH(1; 1, 2)	-0,113179	0,12009	-0,9424418	0,346
GARCH(1; 2, 2)	0,790187	0,12096	6,5325560	<0,001
GARCH(1; 3, 2)	-0,00871	0,04244	-0,2052	0,837
GARCH(1; 1, 3)	0,08712	0,12450	0,6997	0,484
GARCH(1; 2, 3)	0,03437	0,12895	0,2666	0,789
GARCH(1; 3, 3)	0,93814	0,04434	21,1585	<0,001

e GARCH, além do valor ARCH(1; 3, 2). Na matriz constante, $C_0 C_0'$, somente os valores A(1, 1) e A(2, 2) são significativos. O modelo BEKK é muito volátil e a escolha dos parâmetros a serem simulados é complicada, neste caso por exemplo, os parâmetros foram escolhidos dando valores maiores nas diagonais das matrizes.

A Tabela 4.12 apresenta os testes do ajuste. Todos os testes são univariados e pode-se observar que ao nível de significância de 5% as hipóteses de normalidade, para cada série, e de não correlação serial, entre os resíduos e o quadrado destes, não são rejeitadas.

Tabela 4.12: Testes BEKK(1,1)

Testes da Simulação BEKK(1,1):						
AIC(27) = 7205,93						
BIC(27) = 7319,725						
Testes de Normalidade:						
		Jarque-Bera	P-valor	Shapiro-Wilk	P-valor	
serie1.BEKK11.s		0,4386	0,8031	0,9827	0,24171	
serie2.BEKK11.s		0,7493	0,6875	0,9801	0,07243	
serie3.BEKK11.s		1,3315	0,5139	0,9863	0,64834	
Teste de Ljung-Box para os resíduos padronizados:						
			Estatística	P-valor	$\chi^2 - d, f,$	
serie1.BEKK11.s			10,079	0,6090	12	
serie2.BEKK11.s			7,186	0,8451	12	
serie3.BEKK11.s			7,704	0,8078	12	
Teste de Ljung-Box para o quadrado dos resíduos padronizados:						
			Estatística	P-valor	$\chi^2 - d, f,$	
serie1.BEKK11.s			16,996	0,1497	12	
serie2.BEKK11.s			8,291	0,7620	12	
serie3.BEKK11.s			5,808	0,9255	12	
Teste Multiplicadores de Lagrange:						
	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
serie1.BEKK11.s	0,3588	-1,2497	-0,5163	-0,2110	-1,3918	1,3162
serie2.BEKK11.s	-0,2023	1,7306	0,4122	1,0813	-0,5451	0,3673
serie3.BEKK11.s	-0,7901	-0,3041	-0,4090	0,5306	0,3189	0,9971
	Lag 7	Lag 8	Lag 9	Lag 10	Lag 11	Lag 12
serie1.BEKK11.s	0,05898	0,8172	-0,7113	0,1907	2,5607	0,08312
serie2.BEKK11.s	-1,31492	-0,7252	-0,2069	1,2776	0,9648	0,22132
serie3.BEKK11.s	0,72370	-0,4544	-1,0263	-0,9304	-0,6035	-0,01678
	C					
serie1.BEKK11.s	-0,99390					
serie2.BEKK11.s	-0,92075					
serie3.BEKK11.s	0,06852					
			TR^2	P-valor	F-estat.	P-valor
serie1.BEKK11.s			15,440	0,2183	1,4495	0,2514
serie2.BEKK11.s			9,276	0,6792	0,8596	0,6898
serie3.BEKK11.s			5,706	0,9302	0,5248	0,9635

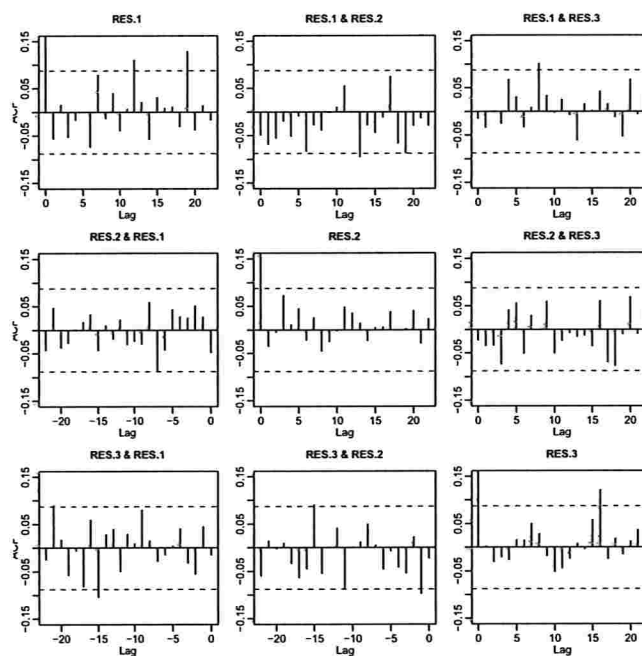


Figura 4.29: *FAC e FAC cruzadas do quadrado dos resíduos do ajuste BEKK(1,1,1).*

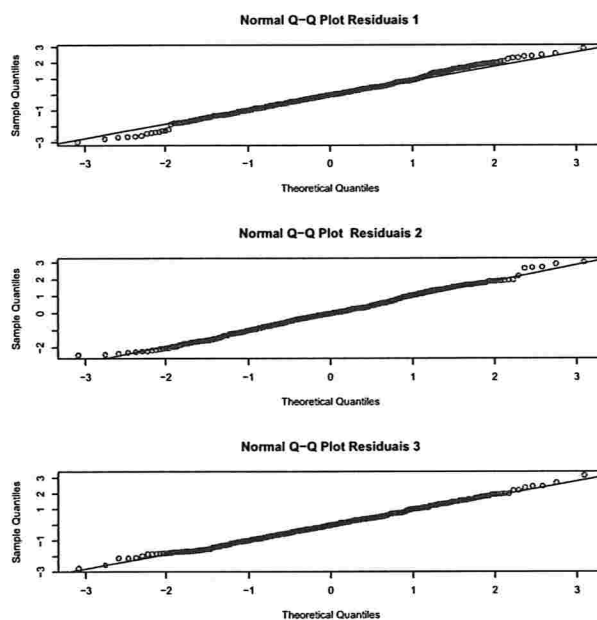


Figura 4.30: *QQ-Plot dos resíduos do ajuste BEKK(1,1,1).*

As Figuras 4.29 e 4.30 apresentam as FAC e FAC cruzada do quadrado dos resíduos e o QQ-plot deles, respectivamente, e nota-se que não existe autocorrelação e autocorrelação cruzada nos resíduos ao quadrado e além disso, que eles seguem uma distribuição normal.

Segundo o p -valor do teste de Portmanteau dado na Tabela 4.13 não existe autocorrelação e autocorrelação cruzada entre os resíduos ao quadrado, já que o valor obtido foi 0,6614.

Tabela 4.13: Teste multivariado de Portmanteau para o quadrado dos resíduos do ajuste BEKK(1,1,1).

Teste Multivariado Portmanteau: Tipo Ljung-Box	
Hipótese Nula: não correlação serial	
Estat. do Teste	p.valor
101,3538	0,6614
Dist. sob Hipótese Nula: qui-quadrado com 108 graus de liberdade	
Total Observ.: 500	

As variâncias e covariâncias condicionais simuladas e ajustadas, além dos desvios padrão e as correlações, estão apresentadas nas Figuras 4.31 e 4.32, respectivamente. Ao longo do tempo, as volatilidades das séries e as covariâncias condicionais destas não são constantes, mostrando alguns picos significativos em alguns períodos do tempo e além disso, o comportamento é similar em cada caso.

Segundo a Tabela 4.11 o modelo BEKK(1,1,1) ajustado é dado por

$$\mathbf{r}_t = \begin{pmatrix} -0,006851 \\ 0,120765 \\ 0,051355 \end{pmatrix} + \varepsilon_t$$

$$\Sigma_t = \begin{pmatrix} 1,678158 & 0,992676 & 0,106956 \\ 0,992676 & 1,150550 & 0,120997 \\ 0,106956 & 0,120997 & 0,003864 \end{pmatrix}$$

$$+ \begin{pmatrix} 0,256540 & 0,034211 & -0,035204 \\ -0,046466 & 0,370333 & -0,092381 \\ 0,042193 & 0,078831 & 0,231540 \end{pmatrix} (\varepsilon_{t-1} \varepsilon'_{t-1}) \begin{pmatrix} 0,256540 & 0,034211 & -0,035204 \\ -0,046466 & 0,370333 & -0,092381 \\ 0,042193 & 0,078831 & 0,231540 \end{pmatrix}$$

$$+ \begin{pmatrix} 0,933486 & -0,113179 & 0,08712 \\ 0,062702 & 0,790187 & 0,03437 \\ -0,029965 & -0,00871 & 0,93814 \end{pmatrix} \Sigma_{t-1} \begin{pmatrix} 0,933486 & -0,113179 & 0,08712 \\ 0,062702 & 0,790187 & 0,03437 \\ -0,029965 & -0,00871 & 0,93814 \end{pmatrix},$$

em que $\varepsilon_t \sim N_3(0, \Sigma_t)$.

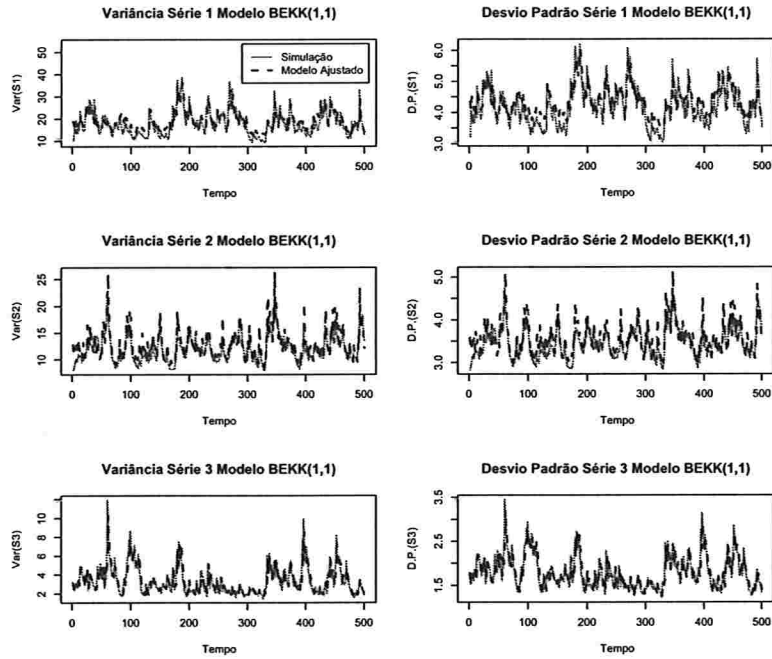


Figura 4.31: Variâncias e desvios padrão BEKK(1,1,1).

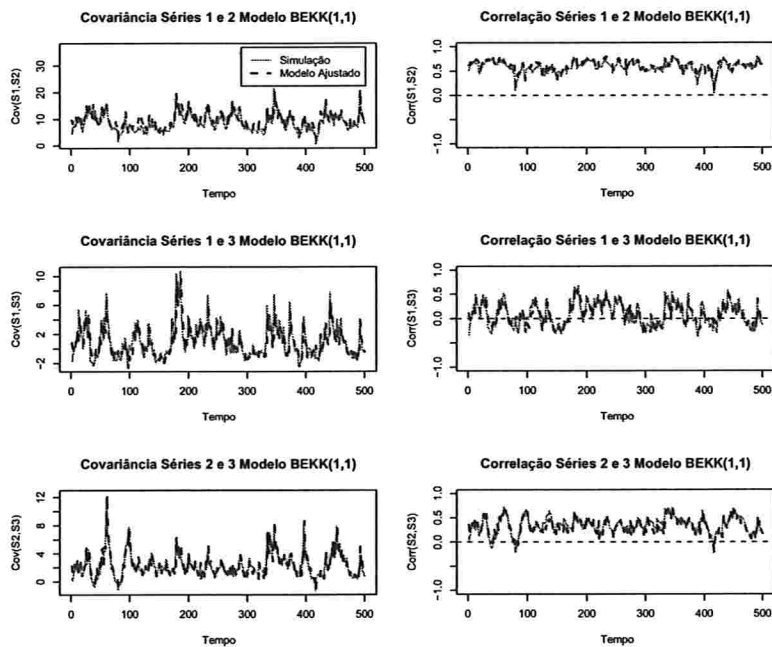


Figura 4.32: Covariâncias e correlações BEKK(1,1,1).

Foram feitas simulações Monte Carlo com 10, 100 e 1000 réplicas do modelo BEKK(1,1,1) dado por (4.6) e (4.7), considerando tamanhos amostrais $T = 500$ e $T = 5000$. Em cada caso, foi considerada a média dos parâmetros nas réplicas e os p -valores do teste multivariado de Portmanteau para os resíduos ao quadrado em cada ajuste.

Estão na Tabela 4.14 os parâmetros verdadeiros ("v") e estimados ("s") com 10, 100 e 1000 réplicas, além do erro quadrático médio (EQM) para cada caso, considerando um tamanho amostral $T = 500$. Em geral os parâmetros estimados conseguiram convergir aos parâmetros verdadeiros e os valores EQM foram pequenos. A Figura 4.33, apresenta as estimativas ("s") e os valores dos parâmetros verdadeiros ("v"), para 10, 100 e 1000 réplicas. Gráficamente os parâmetros estimados conseguiram convergir aos parâmetros verdadeiros e só no caso do elemento na linha 1 e na coluna 1 da matriz $C_0 C_0'$ não convergiu ao valor verdadeiro. Por outro lado, aumentando o tamanho da série para $T = 5000$, os resultados obtidos não mudaram muito, conforme a Tabela 4.15 e a Figura 4.34.

Nos gráficos dos p -valores do teste Portmanteau nas Figuras 4.33 e 4.34 para tamanhos amostrais $T = 500$ e $T = 5000$, respectivamente, observa-se que a maioria destes p -valores ficaram acima do valor 0,05, com linha pontilhada, em cada caso. No caso $T = 500$, para 10 réplicas 8 p -valores foram significativos, para 100 réplicas 94 deles foram significativos e para 1000 réplicas 957 foram significativos. Para o caso $T = 5000$, para 10 réplicas 9 p -valores foram significativos, para 100 réplicas 96 deles foram significativos e para 1000 réplicas 959 foram significativos, indicando que aproximadamente num 92% das vezes, o modelo ajustado pelo S-PLUS, consegue ajustar bem o modelo BEKK(1,1,1) dado por (4.6) e (4.7), com as matrizes escolhidas neste caso. A Tabela 4.16 apresenta o resumo destas porcentagens.

Tabela 4.14: Parâmetros estimados Monte Carlo BEKK(1,1,1) com tamanho amostral $T = 500$.

	Verd. ("v")	10 Réplicas		100 Réplicas		1000 Réplicas	
		Sim. ("s")	EQM	Sim. ("s")	EQM	Sim. ("s")	EQM
c	0,0000	0,0254	0,0205	0,0227	0,0282	0,0071	0,0270
	0,0000	0,0350	0,0214	0,0190	0,0180	0,0074	0,0197
	0,0000	0,0039	0,0042	0,0093	0,0040	0,0026	0,0047
C₀C'₀	1,6900	1,3898	0,2093	1,2939	0,2476	1,3377	0,2463
	0,3340	0,3072	0,0774	0,3564	0,01778	0,3479	0,0176
	0,0810	0,0243	0,0464	0,0551	0,0351	0,0318	0,0372
	0,8356	0,7389	0,0598	0,8648	0,0828	0,8542	0,0902
	0,2430	0,2124	0,0239	0,1780	0,0547	0,1991	0,0585
	0,0729	0,0851	0,0317	0,0505	0,0116	0,0442	0,0140
C₁	0,3000	0,2793	0,0029	0,2892	0,0032	0,2923	0,0035
	-0,0500	-0,0372	0,0020	-0,0439	0,0019	-0,0439	0,0021
	0,0400	0,0329	0,0003	0,0279	0,0007	0,0338	0,0006
	0,0500	0,0758	0,0054	0,0489	0,0049	0,0505	0,0054
	0,3000	0,2967	0,0027	0,3031	0,0029	0,2983	0,0035
	0,0700	0,0643	0,0014	0,0686	0,0009	0,0648	0,0011
	-0,0500	-0,1229	0,0106	-0,0581	0,0084	-0,0482	0,0126
	-0,0700	-0,0851	0,0051	-0,0845	0,0089	-0,0685	0,0089
0,3000	0,2931	0,0011	0,2810	0,0030	0,2862	0,0024	
D₁	0,9000	0,8790	0,0052	0,8949	0,0030	0,8878	0,0027
	0,0500	0,0340	0,0011	0,0429	0,0019	0,0413	0,0020
	-0,0400	-0,0372	0,0002	-0,0327	0,0006	-0,0360	0,0005
	-0,0200	-0,0210	0,0056	-0,0218	0,0055	-0,0197	0,0056
	0,9000	0,9060	0,0017	0,8816	0,0040	0,8860	0,0036
	0,0200	0,0178	0,0003	0,0155	0,0010	0,0204	0,0010
	0,0400	0,0978	0,0090	0,0424	0,0082	0,0433	0,0078
	-0,0500	-0,0299	0,0038	-0,0330	0,0048	-0,0453	0,0051
	0,9000	0,9000	0,0011	0,8995	0,0011	0,8964	0,0012

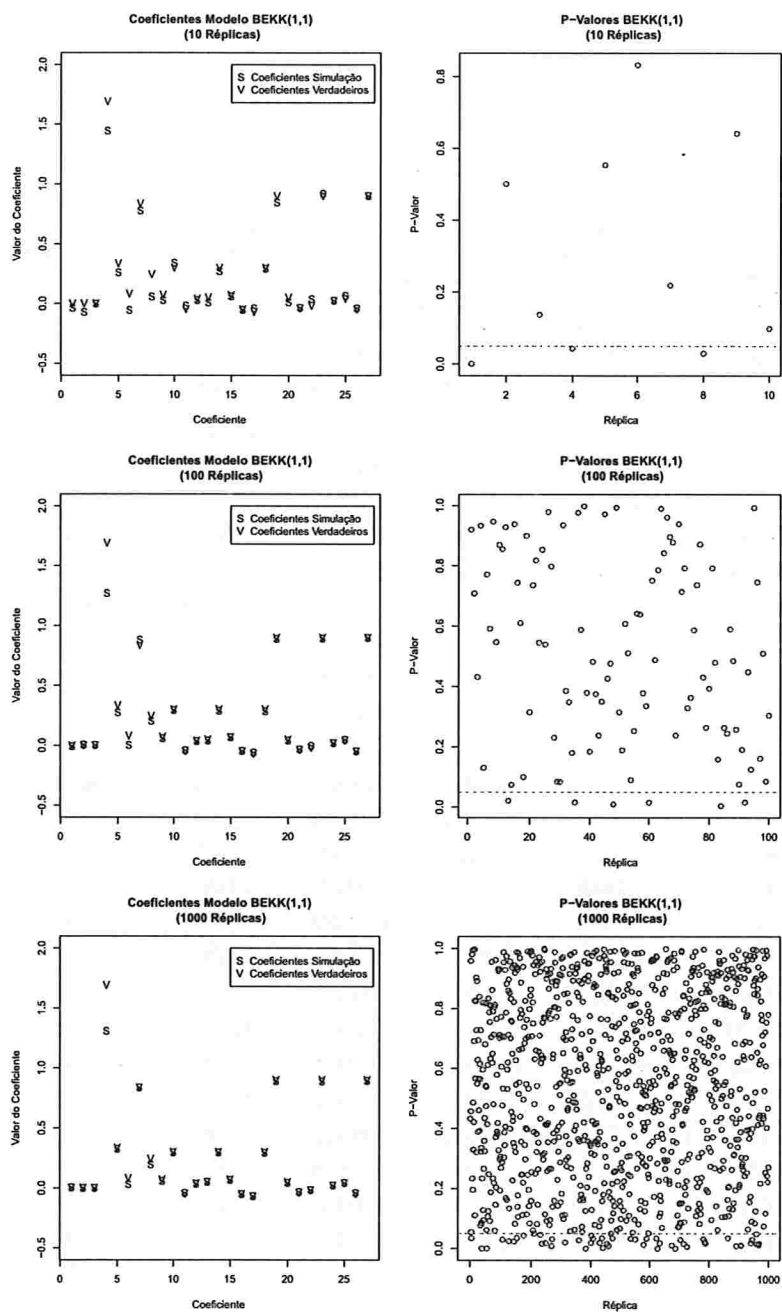


Figura 4.33: Parâmetros simulados e estimados e teste Pormanteau BEKK(1,1,1) com tamanho amostral $T = 500$.

Tabela 4.15: Parâmetros estimados Monte Carlo BEKK(1,1,1) com tamanho amostral $T = 5000$.

	Verd. ("v")	10 Réplicas		100 Réplicas		1000 Réplicas	
		Sim. ("s")	EQM	Sim. ("s")	EQM	Sim. ("s")	EQM
c	0,0000	0,0001	0,0031	0,0039	0,0023	0,0004	0,0028
	0,0000	-0,0003	0,0015	0,0018	0,0020	-0,0014	0,0019
	0,0000	-0,0041	0,0007	-0,0016	0,0005	0,0003	0,0005
C₀C'₀	1,6900	1,3048	0,1982	1,2603	0,2430	1,3687	0,2126
	0,3340	0,2585	0,0187	0,2985	0,0185	0,3315	0,0002
	0,0810	0,0883	0,0020	0,0709	0,0030	0,0610	0,0027
	0,8356	0,9277	0,0131	0,8572	0,0105	0,8363	0,0012
	0,2430	0,2537	0,0026	0,2469	0,0041	0,2582	0,0037
	0,0729	0,0891	0,0066	0,0687	0,0077	0,0662	0,0073
C₁	0,3000	0,2945	0,0005	0,2965	0,0004	0,2994	0,0003
	-0,0500	-0,0536	0,0001	-0,0531	0,0002	-0,0505	0,0002
	0,0400	0,0000	0,0000	0,0388	0,0000	0,0397	0,0000
	0,0500	0,0508	0,0010	0,0548	0,0005	0,0500	0,0005
	0,3000	0,3073	0,0002	0,3037	0,0004	0,3018	0,0004
	0,0700	0,0705	0,0001	0,0694	0,0001	0,0700	0,0001
	-0,0500	-0,0376	0,0011	-0,0524	0,0010	-0,0483	0,0012
	-0,0700	-0,0660	0,0008	-0,0715	0,0007	-0,0703	0,0007
0,3000	0,3044	0,0003	0,2982	0,0003	0,2987	0,0002	
D₁	0,9000	0,9041	0,0001	0,8997	0,0002	0,8977	0,0002
	0,0500	0,0514	0,0001	0,0517	0,0001	0,0508	0,0001
	-0,0400	-0,0401	0,0000	-0,0405	0,0000	-0,0398	0,0000
	-0,0200	-0,0186	0,0003	-0,0198	0,0003	-0,0195	0,0003
	0,9000	0,8915	0,0002	0,8952	0,0002	0,8958	0,0003
	0,0200	0,0189	0,0001	0,0207	0,0001	0,0192	0,0001
	0,0400	0,0279	0,0006	0,0368	0,0005	0,0398	0,0004
	-0,0500	-0,0509	0,0002	-0,0494	0,0003	-0,0500	0,0003
	0,9000	0,8976	0,0001	0,9001	0,0001	0,8999	0,0001

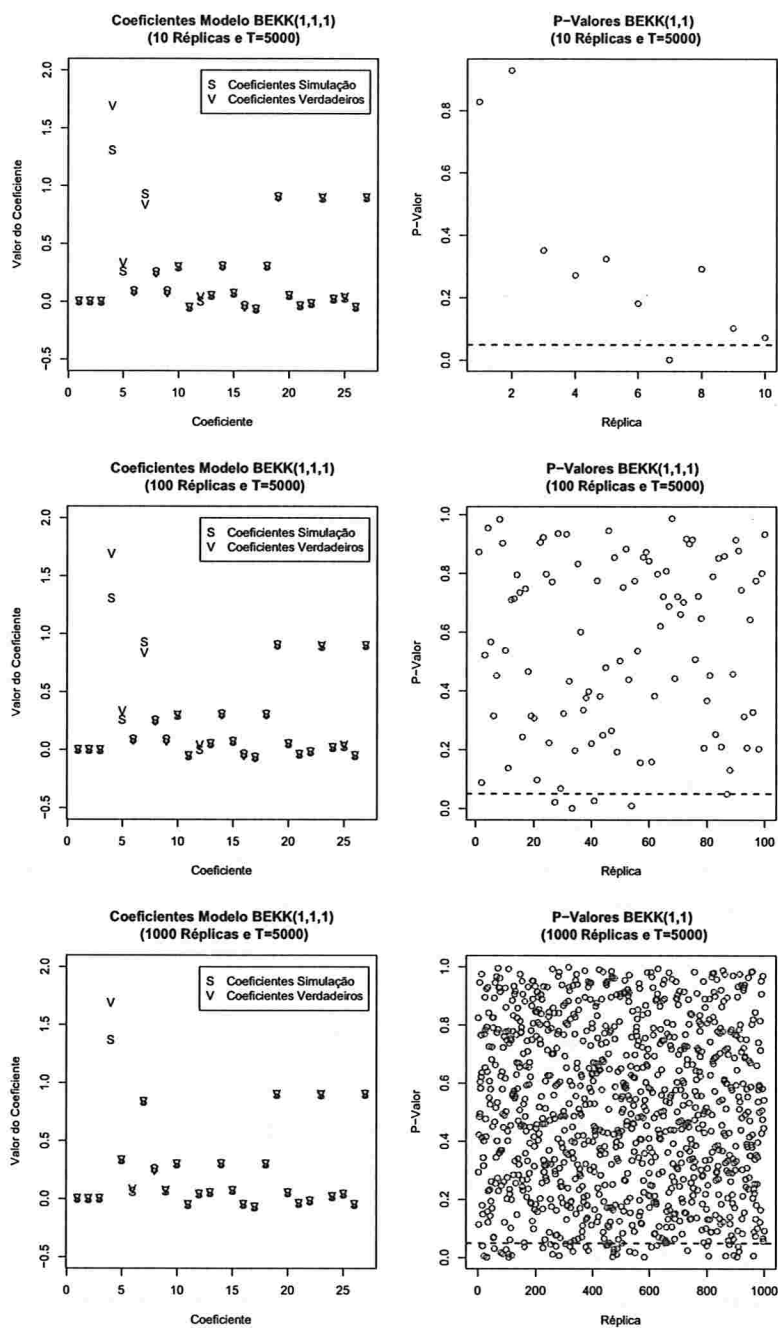


Figura 4.34: Parâmetros simulados e estimados e teste Pormanteau BEKK(1,1,1) com tamanho amostral $T = 5000$.

Tabela 4.16: *Porcentagens de aceitação do teste Portmanteau usando o método Monte Carlo BEKK(1,1,1).*

	Tamanho amostral $T = 500$			Tamanho amostral $T = 5000$		
Número de réplicas	10	100	1000	10	100	1000
Porcentagem	80%	94%	95,7%	90%	96%	95,9%

4.4 Modelo BEKK(1,1,1) com distribuição t-student

Foram simuladas 3 séries de tamanho 500 com as mesmas matrizes do modelo BEKK(1,1,1) normalmente distribuído, considerando agora a distribuição t-student com 8 graus de liberdade. Este modelo também segue as equações (4.6) e (4.7).

A Figura 4.35 apresenta as 3 séries simuladas e observa-se que as séries não têm tendência ao longo do tempo e que existem conglomerados em que a volatilidade é maior. Apresentam-se as FAC e as FAC cruzadas das séries ao quadrado na Figura 4.36. Nota-se que em todos os gráficos existe autocorrelação significativa antes do lag 10, confirmando a existência de uma dependência quadrática nas séries.

A Tabela 4.17 apresenta os parâmetros estimados pelo S-PLUS usando uma distribuição t-student e na Tabela 4.18 são apresentados os testes do ajuste BEKK(1,1,1) com distribuição t-student. Nota-se que segundo estes testes univariados a hipótese de normalidade do teste de Jarque Bera foi rejeitada ao nível de significância de 5%, mas os testes de Ljung-Box não rejeitam a hipótese de não correlação entre os resíduos e o quadrado destes, também ao nível de 5%.

As Figuras 4.37 e 4.38 apresentam as FAC e FAC cruzadas dos resíduos ao quadrado e o QQ-plot deles, respectivamente, e observa-se que não existe autocorrelação e autocorrelação cruzada entre os resíduos ao quadrado e além disso, no QQ-plot nota-se que as caudas são mais pesadas que na distribuição normal, já que a distribuição escolhida foi a t-student.

A Tabela 4.19 apresenta o p -valor do teste de Portmanteau para o quadrado dos resíduos e segundo ele, não existe autocorrelação e autocorrelação cruzada nos resíduos ao quadrado, já que o valor obtido foi 0,9997, indicando a adequação do modelo ajustado.

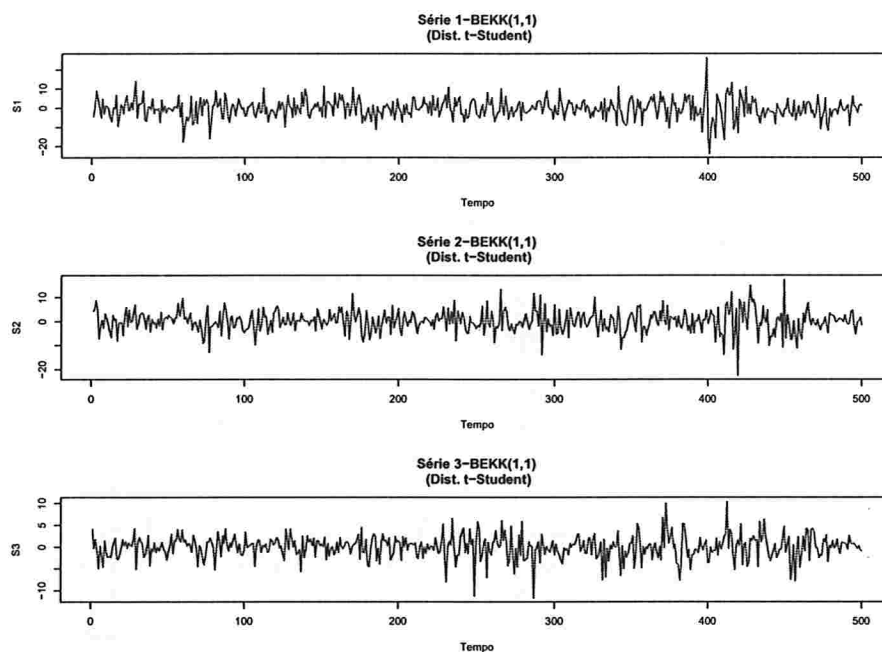


Figura 4.35: Séries simuladas $BEKK(1,1,1)$ t -student.

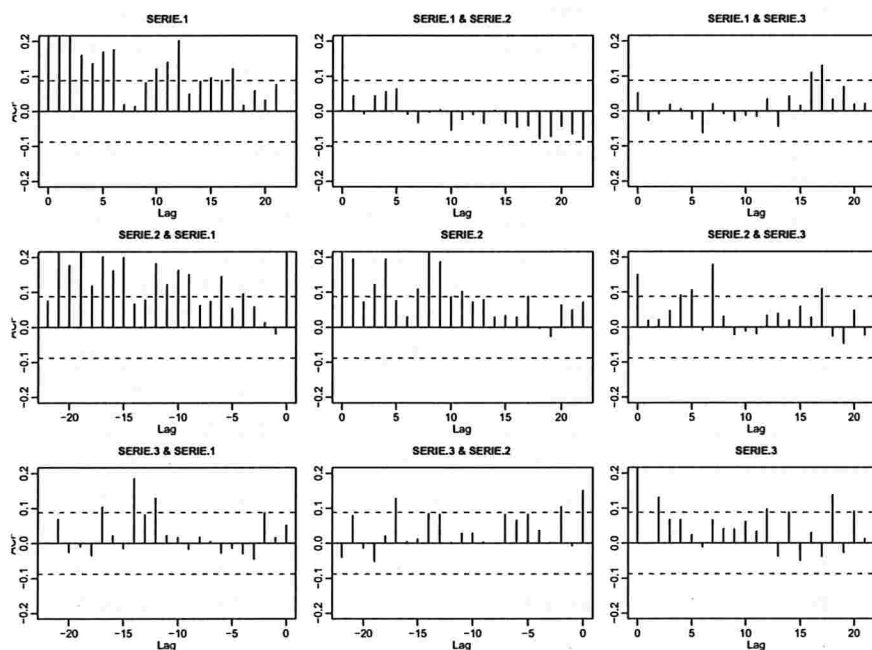


Figura 4.36: FAC e FAC cruzadas do quadrado das séries simuladas modelo $BEKK(1,1,1)$ t -student.

Tabela 4.17: Parâmetros estimados BEKK(1,1,1) t-student.

Graus de liberdade estimado:				
Graus	D.P.			
6,874039	1,143702			
Parâmetros Estimados BEKK(1,1,1) t-student:				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	-0,127750	0,18591	-0,6871786	0,246
C(2)	0,053293	0,17114	0,3113970	0,378
C(3)	0,042259	0,10253	0,4121641	0,340
A(1, 1)	1,434207	0,35389	4,0526396	<0,001
A(2, 1)	-0,114439	0,44566	-0,2567848	0,399
A(3, 1)	0,193722	0,25082	0,7723390	0,220
A(2, 2)	1,003571	0,36110	2,7792325	0,003
A(3, 2)	0,642404	0,32331	1,9869332	0,024
A(3, 3)	0,005790	37,90081	0,0001528	0,499
ARCH(1; 1, 1)	0,283314	0,05246	5,4007276	<0,001
ARCH(1; 2, 1)	-0,084315	0,04941	-1,7064300	0,044
ARCH(1; 3, 1)	0,037941	0,02736	1,3867351	0,083
ARCH(1; 1, 2)	0,048270	0,06705	0,7199207	0,236
ARCH(1; 2, 2)	0,336069	0,06168	5,4486892	<0,001
ARCH(1; 3, 2)	0,059842	0,03584	1,6697176	0,048
ARCH(1; 1, 3)	-0,025554	0,10172	-0,2512120	0,401
ARCH(1; 2, 3)	0,033244	0,09312	0,3569921	0,360
ARCH(1; 3, 3)	0,311863	0,05646	5,5232554	<0,001
GARCH(1; 1, 1)	0,906589	0,04223	21,4682029	<0,001
GARCH(1; 2, 1)	0,074751	0,03741	1,9983478	0,023
GARCH(1; 3, 1)	-0,034793	0,02104	-1,6535660	0,049
GARCH(1; 1, 2)	0,001225	0,04200	0,0291605	0,488
GARCH(1; 2, 2)	0,898423	0,03547	25,3321494	<0,001
GARCH(1; 3, 2)	0,010737	0,02188	0,49064	0,312
GARCH(1; 1, 3)	0,005051	0,05911	0,08546	0,466
GARCH(1; 2, 3)	-0,084204	0,06097	-1,38098	0,084
GARCH(1; 3, 3)	0,895657	0,03088	29,00741	<0,001

Tabela 4.18: Testes BEKK(1,1,1) t-student.

Testes da Simulação BEKK(1,1,1) t-student:						
AIC(27) = 7884,429						
BIC(27) = 8002,438						
Testes de Normalidade:						
	Jarque-Bera		P-valor	Shapiro-Wilk	P-valor	
serie1.BEKK11.s	28,54		<0,001	0,9953	0,99959	
serie2.BEKK11.s	186,54		<0,001	0,9800	0,07025	
serie3.BEKK11.s	176,00		<0,001	0,9797	0,05929	
Teste de Ljung-Box para os resíduos padronizados:						
	Estatística		P-valor	$\chi^2 - d.f.$		
serie1.BEKK11.s	18,012		0,1153	12		
serie2.BEKK11.s	4,929		0,9603	12		
serie3.BEKK11.s	10,768		0,5489	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados:						
	Estatística		P-valor	$\chi^2 - d.f.$		
serie1.BEKK11.s	11,833		0,4592	12		
serie2.BEKK11.s	10,403		0,5807	12		
serie3.BEKK11.s	3,636		0,9891	12		
Teste Multiplicadores de Lagrange:						
	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
serie1.BEKK11.s	0,5417	1,4439	-0,5394	0,2645	0,6199	-1,0743
serie2.BEKK11.s	-0,3407	0,1960	0,7683	-0,3484	-1,5270	-1,1382
serie3.BEKK11.s	0,8139	-0,5311	0,4653	-0,8123	-0,8589	0,4897
	Lag 7	Lag 8	Lag 9	Lag 10	Lag 11	Lag 12
serie1.BEKK11.s	-0,6395	-0,5037	-0,08166	0,8625	2,0986	0,05731
serie2.BEKK11.s	1,7220	-0,6946	-0,04728	-0,5327	-0,2717	0,01876
serie3.BEKK11.s	-0,2245	0,4080	0,03825	0,0261	0,1139	0,04190
	C					
serie1.BEKK11.s	0,7695					
serie2.BEKK11.s	1,0614					
serie3.BEKK11.s	-0,4720					
	TR^2		P-valor	F-stat	P-valor	
serie1.BEKK11.s	11,157		0,5155	1,0380	0,5209	
serie2.BEKK11.s	9,656		0,6461	0,8955	0,6543	
serie3.BEKK11.s	3,427		0,9917	0,3137	0,9997	

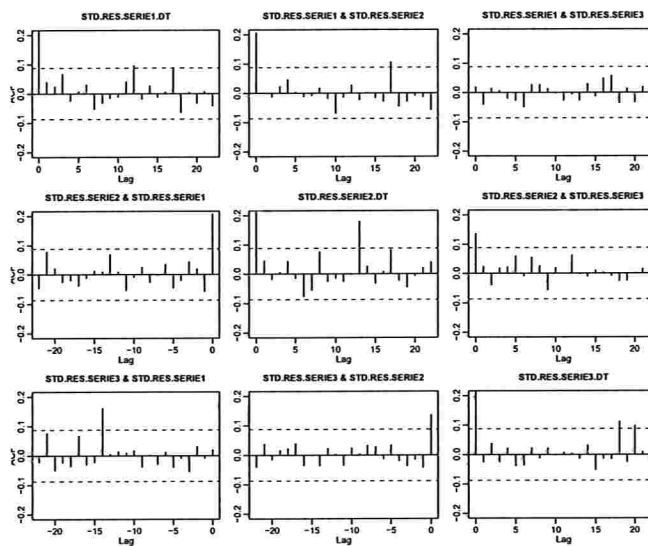


Figura 4.37: FAC e FAC cruzadas do quadrado dos resíduos do ajuste $BEKK(1,1,1)$ t -student.

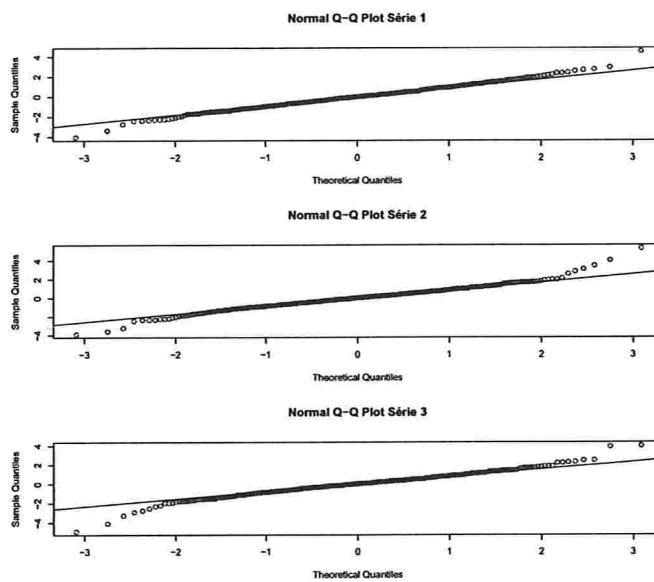
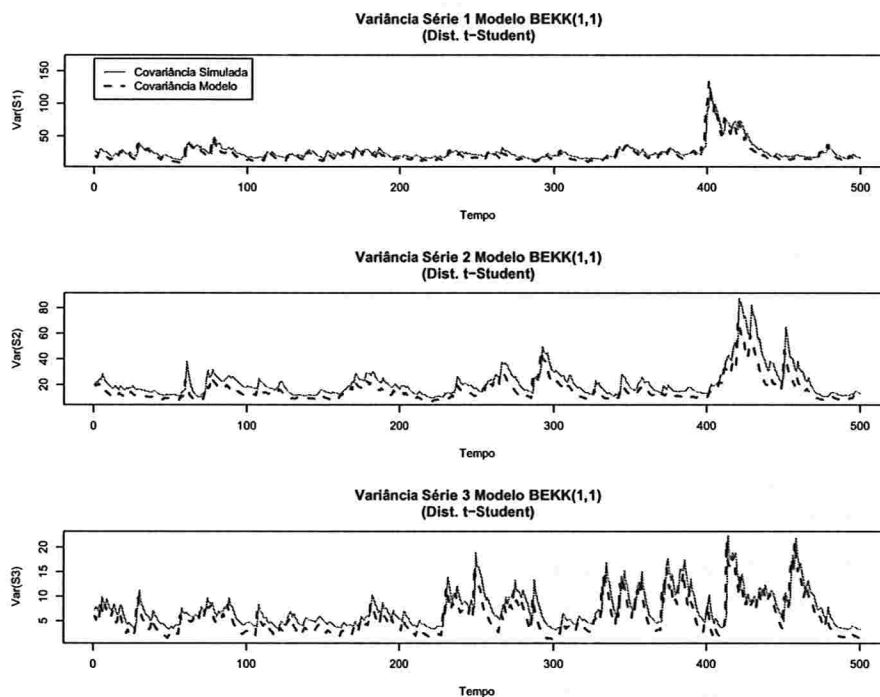


Figura 4.38: QQ-Plot dos resíduos do ajuste $BEKK(1,1,1)$ t -student.

Tabela 4.19: Teste de portmanteau para o quadrado dos resíduos do ajuste $BEKK(1,1,1)$ t -student.

Teste Multivariado Portmanteau: Tipo Ljung-Box	
Hipótese Nula: não correlação serial	
Estat. do Teste	p.valor
64,6114	0,9997
Dist. sob Hipótese Nula: qui-quadrado com 108 graus de liberdade	
Total Observ.: 500	

O comportamento das variâncias e covariâncias condicionais simuladas e ajustadas pode ser visto nas Figuras 4.39 e 4.40, respectivamente. Observa-se que nos dois casos é similar e além disso, existem alguns períodos no tempo que apresentam picos altos.

Figura 4.39: Variâncias juntas $BEKK(1,1,1)$ t -student.

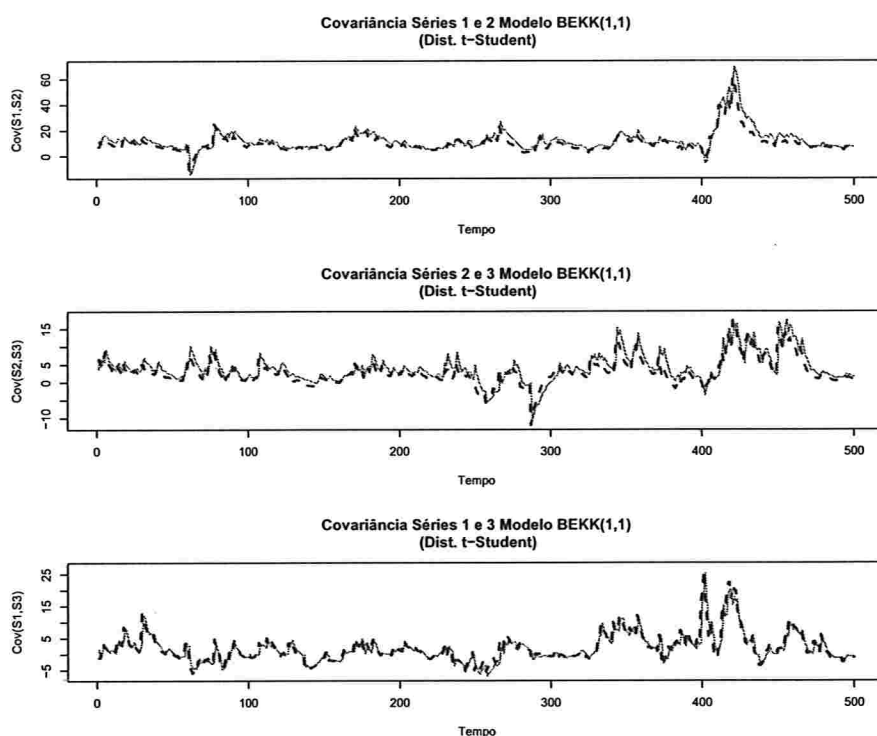


Figura 4.40: Covariâncias juntas BEKK(1,1,1) t-student.

Finalmente, o modelo ajustado é dado por:

$$\mathbf{r}_t = \begin{pmatrix} -0,127750 \\ 0,053293 \\ 0,042259 \end{pmatrix} + \varepsilon_t$$

$$\Sigma_t = \begin{pmatrix} 1,434207 & -0,114439 & 0,193722 \\ -0,114439 & 1,003571 & 0,642404 \\ 0,193722 & 0,642404 & 0,005790 \end{pmatrix}$$

$$+ \begin{pmatrix} 0,283314 & 0,048270 & -0,025554 \\ -0,084315 & 0,336069 & 0,033244 \\ 0,037941 & 0,059842 & 0,311863 \end{pmatrix} (\varepsilon_{t-1} \varepsilon'_{t-1}) \begin{pmatrix} 0,283314 & 0,048270 & -0,025554 \\ -0,084315 & 0,336069 & 0,033244 \\ 0,037941 & 0,059842 & 0,311863 \end{pmatrix}$$

$$+ \begin{pmatrix} 0,906589 & 0,001225 & 0,005051 \\ 0,074751 & 0,898423 & -0,084204 \\ -0,034793 & 0,010737 & 0,895657 \end{pmatrix} \Sigma_{t-1} \begin{pmatrix} 0,906589 & 0,001225 & 0,005051 \\ 0,074751 & 0,898423 & -0,084204 \\ -0,034793 & 0,010737 & 0,895657 \end{pmatrix},$$

em que ε_t segue uma distribuição t-student multivariada com aproximadamente 7 graus de liberdade e matriz de covariâncias condicional Σ_t .

Capítulo 5

Aplicações

Neste capítulo apresentam-se duas aplicações usando dados reais do mercado financeiro Brasileiro, a primeira com quatro ações que fazem parte do mercado dos Materiais de Construção e depois consideram-se quatro ações que fazem parte do mercado da Construção Civil. Estes dois mercados estão altamente correlacionados e fazem parte importante do índice Ibovespa. Os dados são os log-retornos obtidos com os preços de fechamento diários entre 2 de Janeiro de 2009 e 16 de Agosto de 2010, num total de 400 observações.

A Tabela 5.1 apresenta as ações consideradas com seus respectivos códigos e o setor no qual pertencem.

Tabela 5.1: Ações da aplicação.

Setor	Ação	Código
Materiais de Construção	Vale	VALE
	Gerdau	GGBR
	Usiminas	USIM
	Aços Vill	AVIL
Construção Civil	Cyrela Realt	CYRE
	Inpar S/A	INPR
	Tecnisa	TCSA
	Gafisa	GFSA

Pode-se considerar um ajuste GARCH multivariado para o conjunto das ações da Tabela 5.1, mas o número de parâmetros a serem estimados é consideravelmente alto, e portanto é razoável reagrupar as séries em grupos menores para depois aplicar os modelos GARCH multivariados em cada grupo. Na prática o conjunto de ativos financeiros da Tabela 5.1 estão fortemente correlacionados dentro de cada setor e é conveniente fazer um ajuste GARCH multivariado em cada grupo para reduzir o número de parâmetros. Com o intuito de verificar

que no conjunto de 8 ativos existem de fato dois grupos fortemente correlacionados entre eles, aplica-se o método das componentes principais (ver Johnson e Wichern (1998)). A Figura 5.1 apresenta o biplot das duas primeiras componentes do conjunto de ações e observa-se que existe uma separação da variabilidade em dois grupos, considerados na segunda componente. Os grupos obtidos são os mesmos que na Tabela 5.1, dois setores Materiais de Construção e Construção Civil.

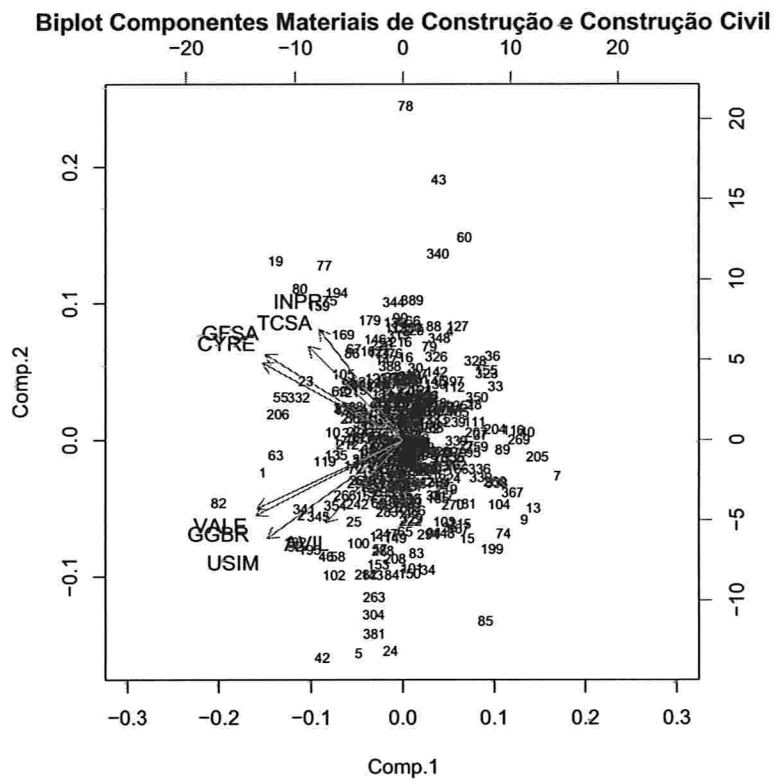


Figura 5.1: Biplot das duas primeiras componentes principais das ações.

5.1 Ações Materiais de Construção

Os modelos dados no Capítulo 3 foram aplicados aos log-retornos do conjunto de ações do setor Materiais de Construção.

A Figura 5.2 apresenta os log-retornos das ações VALE, USIMINAS, GERDAU e AÇOS VILL no período de 2 de Janeiro de 2009 até 16 de Agosto de 2010. Nota-se que não há uma tendência definida e que existem, ao longo do tempo, períodos de maior variabilidade

em cada uma das séries.

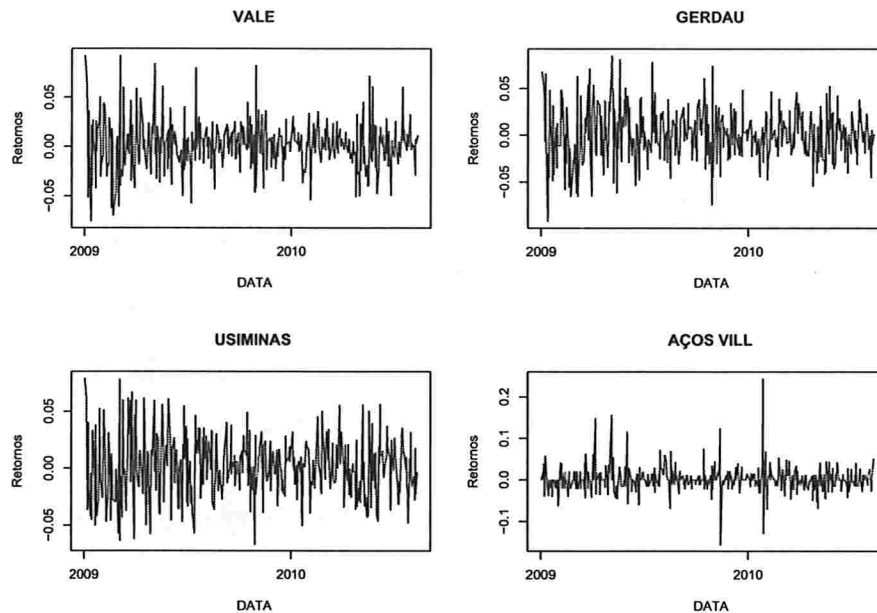


Figura 5.2: Ações materiais de construção, 2 de Janeiro de 2009 até 19 de Agosto de 2010.

Estão apresentadas na Figura 5.3 as FAC e FAC cruzadas dos log-retornos e observa-se que em alguns gráficos existe autocorrelação e autocorrelação cruzada significativa antes do lag 10, indicando que há dependência linear, não muito forte, entre eles.

Na Figura 5.4 apresentam-se as FAC e FAC cruzadas dos quadrados dos log-retornos no grupo materiais de construção. Nota-se que existe autocorrelação e autocorrelação cruzada significativa, mostrando que existe dependência quadrática nos log-retornos.

Os ajustes GARCH multivariados que foram considerados são EWMA, DVEC(1,1), BEKK(1,1,1), CCC com ajustes GARCH(1,1) univariados para cada um dos elementos na diagonal da matriz dada por (3.29) e Componentes Principais com ajustes GARCH(1,1) para cada uma das componentes. Alguns outros ajustes, como por exemplo DVEC(2,1), DVEC(1,2),..., foram feitos, mas os resultados obtidos em cada um deles não foram melhores. Também foram considerados ajustes ARMA(1,1) para os log-retornos multivariados.

A Tabela 5.2 apresenta os valores de AIC e BIC para cada um dos modelos. O modelo que apresentou o menor valor nos dois critérios foi o CCC-GARCH(1,1).

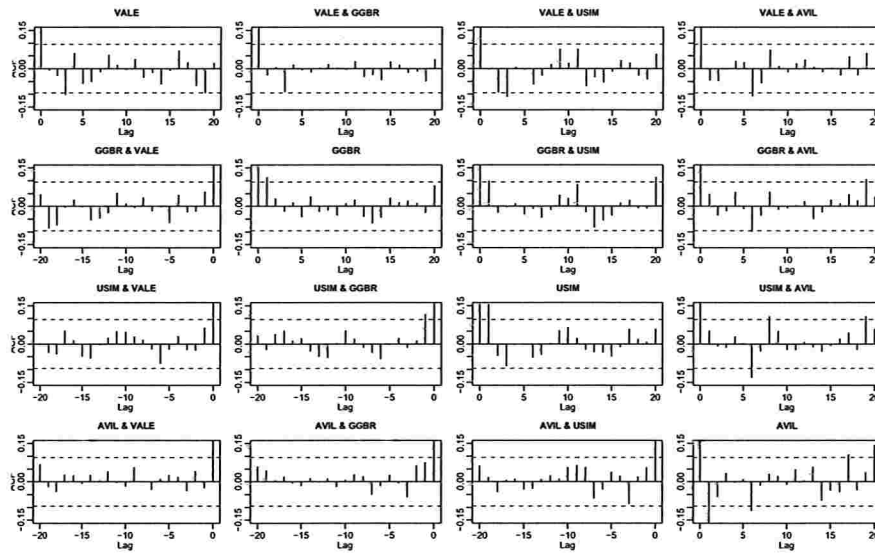


Figura 5.3: *FAC e FAC cruzadas dos log-retornos, materiais de construção.*

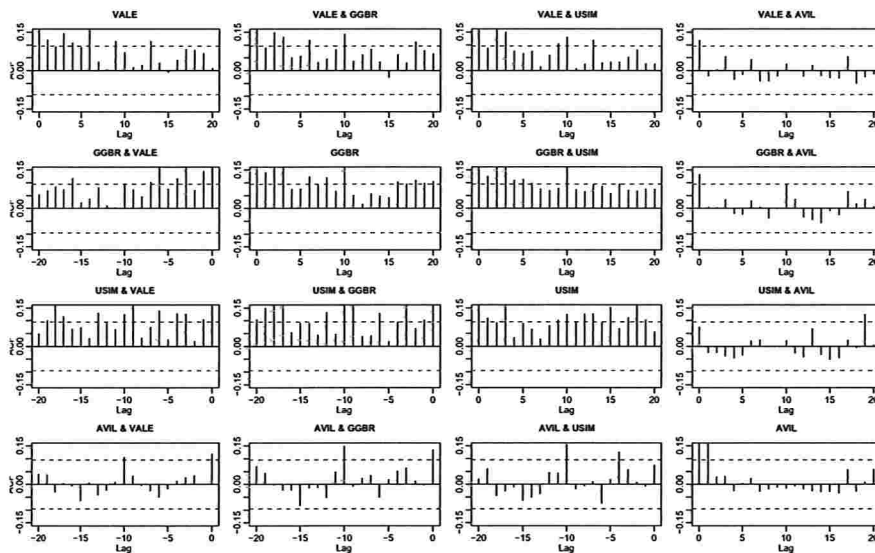


Figura 5.4: *FAC e FAC cruzadas dos log-retornos ao quadrado, materiais de construção.*

Tabela 5.2: Critérios de seleção AIC e BIC, materiais de construção.

Modelo	AIC	BIC
EWMA	-7954,073	-7898,193
DVEC(1,1)	-8010,344	-7798,796
BEKK(1,1)	-7969,844	-7750,313
CCC-GARCH(1,1)	-8044,968	-7945,181
PCOMP-GARCH(1,1)	-8002,156	-7902,37

Apresentam-se na Tabela 5.3 os testes multivariados Portmanteau para o quadrado dos resíduos em cada ajuste. Todos os p -valores são maiores que o nível de significância de 5%. Segundo os critérios AIC e BIC, o modelo mais adequado é o CCC-GARCH(1,1).

Tabela 5.3: Testes de Portmanteau para o quadrado dos resíduos dos ajustes, materiais de construção.

Teste Multivariado Portmanteau: Tipo Ljung-Box		
Hipótese Nula: não correlação serial		
Modelo	Estat. do teste	p.valor
EWMA	213,9478	0,1327
DVEC(1,1)	165,8898	0,9137
BEKK(1,1)	166,4166	0,9089
CCC-GARCH(1,1)	165,9967	0,9127
PCOMP-GARCH(1,1)	188,7550	0,5527

A Tabela 5.4 apresenta o resumo do ajuste CCC-GARCH(1,1). Nota-se que as estimativas $C(1)$, $C(2)$, $C(3)$ e $C(4)$ do vetor constante c , são não significativas ao nível de 5%. Na parte ARMA, unicamente os parâmetros estimados $AR(1; 4, 4)$ e $MA(1; 4, 4)$ foram significativos. Nas estimativas GARCH univariadas, os parâmetros estimados $A(2, 2)$ e $A(3, 3)$ são não significativos ao nível de significância de 5%. A matriz de correlação constante junto com os seus desvios padrão também são apresentados na Tabela 5.4. Para os graus de liberdade o valor estimado foi 8,21468 com um desvio padrão de 1,13567.

Na Tabela 5.5 encontram-se os testes do ajuste CCC-GARCH(1,1). Segundo os testes de normalidade, as séries GERDAU e USIMINAS seguem uma distribuição normal e as séries VALE e AÇOS VILL não se ajustam com esta distribuição. Já para os resíduos e o quadrado dos resíduos, todos os p -valores dos testes Ljung-Box e Lagrange são maiores que 5%, não

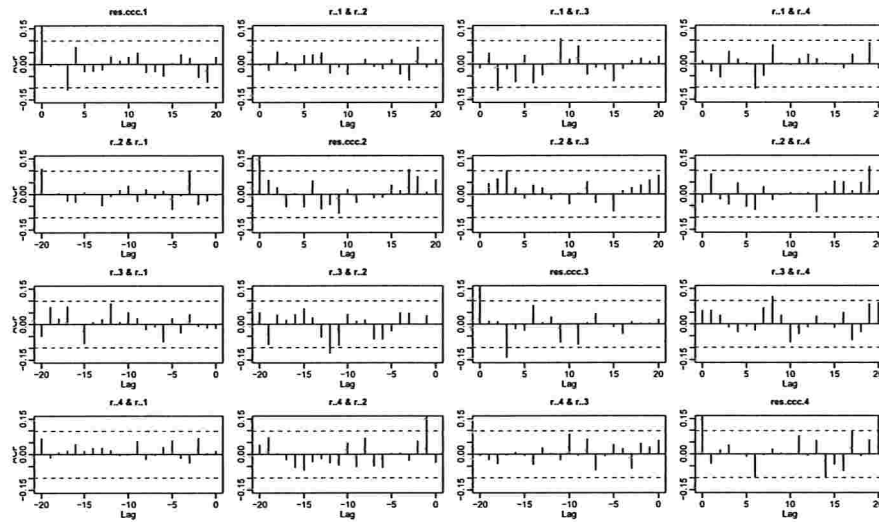


Figura 5.5: FAC e FAC cruzada dos resíduos do ajuste CCC-GARCH(1,1), materiais de construção.

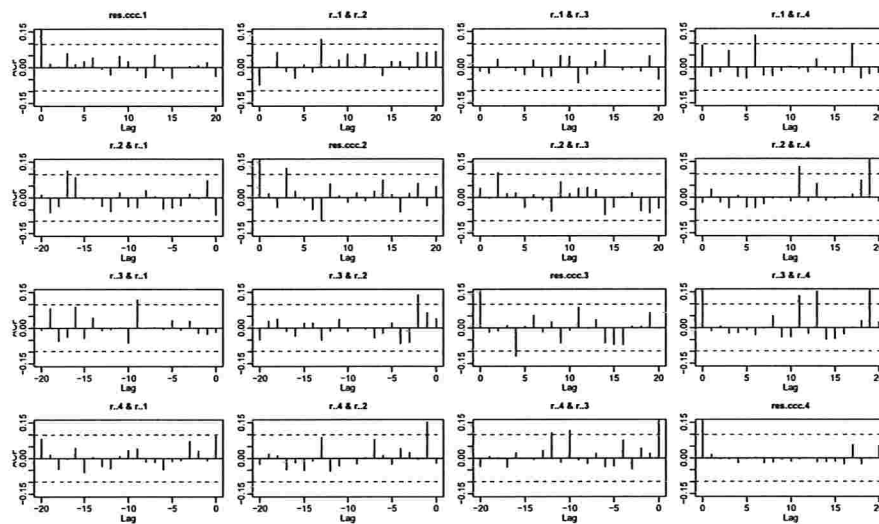


Figura 5.6: FAC e FAC cruzada do quadrado dos resíduos do ajuste CCC-GARCH(1,1), materiais de construção.

Tabela 5.4: Resumo do ajuste CCC-GARCH(1,1), materiais de construção.

Graus de liberdade estimado:				
Graus	D.P.			
8,21468	1,13567			
Parâmetros Estimados CCC-GARCH(1,1):				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	1,279e-003	1,447e-003	0,88392	0,188
C(2)	1,340e-004	1,062e-003	0,12615	0,450
C(3)	4,211e-004	1,174e-003	0,35871	0,360
C(4)	2,781e-004	3,872e-004	0,71819	0,237
AR(1; 1, 1)	-4,670e-002	8,851e-001	-0,05277	0,479
AR(1; 2, 2)	2,440e-001	2,049e-001	1,19090	0,117
AR(1; 3, 3)	1,293e-001	2,338e-001	0,55319	0,290
AR(1; 4, 4)	6,040e-001	2,204e-001	2,74006	0,003
MA(1; 1, 1)	8,635e-002	8,820e-001	0,09790	0,461
MA(1; 2, 2)	-9,498e-002	2,134e-001	-0,44507	0,328
MA(1; 3, 3)	3,479e-002	2,390e-001	0,14558	0,442
MA(1; 4, 4)	-6,981e-001	1,908e-001	-3,65971	< 0,001
A(1, 1)	2,168e-005	1,066e-005	2,03336	0,021
A(2, 2)	1,487e-005	1,096e-005	1,35730	0,088
A(3, 3)	8,786e-006	8,346e-006	1,05273	0,147
A(4, 4)	3,004e-004	8,229e-005	3,65002	< 0,001
ARCH(1; 1, 1)	5,728e-002	1,993e-002	2,87486	0,002
ARCH(1; 2, 2)	4,365e-002	1,815e-002	2,40561	0,008
ARCH(1; 3, 3)	3,535e-002	1,895e-002	1,86545	0,031
ARCH(1; 4, 4)	2,107e-001	6,626e-002	3,17951	< 0,001
GARCH(1; 1, 1)	8,987e-001	3,379e-002	26,59587	< 0,001
GARCH(1; 2, 2)	9,354e-001	2,760e-002	33,89061	< 0,001
GARCH(1; 3, 3)	9,526e-001	2,634e-002	36,16683	< 0,001
GARCH(1; 4, 4)	0,3514e-000	0,136e-000	2,57500	0,005
Matriz de Correlação Condicional Constante:				
	VALE	GGBR	USIM	AVIL
VALE	1,0000	0,7677	0,6919	0,4122
GGBR	0,7677	1,0000	0,7245	0,4106
USIM	0,6919	0,7245	1,0000	0,3604
AVIL	0,4122	0,4106	0,3604	1,0000
Desvios Padrão:				
	VALE	GGBR	USIM	AVIL
VALE	NA	0,02277	0,03037	0,04780
GGBR	0,02277	NA	0,02748	0,04972
USIM	0,03037	0,02748	NA	0,05233
AVIL	0,04780	0,04972	0,05233	NA

Tabela 5.5: Testes do ajuste CCC-GARCH(1,1), materiais de construção.

Testes Ajuste CCC-GARCH(1,1):						
Testes de Normalidade:						
	Jarque-Bera	P-valor	Shapiro-Wilk	P-valor		
VALE	42,8713	0,000	0,9728	0,001587		
GGBR	6,7319	0,034	0,9848	0,531655		
USIM	0,6048	0,739	0,9904	0,951286		
AVIL	6479,5896	0,000	0,8527	<0,001		
Teste Ljung-Box para os resíduos padronizados:						
	Estatística	P-valor	$\chi^2 - d.f.$			
VALE	10,277	0,5916	12			
GGBR	11,122	0,5185	12			
USIM	17,027	0,1486	12			
AVIL	8,045	0,7816	12			
Teste Ljung-Box para o quadrado dos resíduos padronizados:						
	Estatística	P-valor	$\chi^2 - d.f.$			
VALE	8,045	0,7816	12			
GGBR	14,1982	0,2882	12			
USIM	12,4123	0,4132	12			
AVIL	0,7764	1,0000	12			
Teste Multiplicadores de Lagrange:						
	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
VALE	0,1034	1,00295	0,3463	0,57829	0,03884	-0,1132
GGBR	-0,8100	2,63738	0,8414	-0,37270	-1,16713	-2,1207
USIM	-0,1355	0,10665	-2,3138	-0,15219	0,94383	-0,1590
AVIL	-0,1028	-0,09859	-0,4189	-0,06366	0,06842	-0,4240
	Lag 7	Lag 8	Lag 9	Lag 10	Lag 11	Lag 12
VALE	-0,8244	1,2146	0,838043	-0,03283	-1,01385	-0,423498
GGBR	1,1970	0,3165	0,239423	0,25881	-0,56896	-0,121132
USIM	0,1673	-1,2542	-0,004793	1,56990	0,02780	0,163066
AVIL	-0,3561	-0,1345	-0,300554	-0,12453	-0,01906	0,009367
	C					
VALE	-0,1574					
GGBR	0,4936					
USIM	-0,3278					
AVIL	0,2875					
	TR^2	P-valor	F-estat.	P-valor		
VALE	5,283	0,9478	0,4869	0,9772		
GGBR	15,205	0,2304	1,4386	0,2565		
USIM	11,279	0,5052	1,0561	0,5050		
AVIL	0,740	1,0000	0,0674	1,0000		

rejeitando a hipótese nula de não autocorrelação serial.

As Figuras 5.5 e 5.6 apresentam as FAC e FAC cruzadas dos resíduos e do quadrado dos resíduos do ajuste CCC-GARCH(1,1). Observa-se que existem poucos lags nos quais a autocorrelação e autocorrelação cruzada é significativa nos dois casos. A Tabela 5.6 apresenta os testes multivariados de Portmanteau para os resíduos e o quadrado dos resíduos do ajuste CCC-GARCH(1,1). Os p -valores dos testes foram 0,6830 e 0,9127, respectivamente, não rejeitando, em ambos os casos, a hipótese nula, indicando que o modelo ajustado é adequado para modelar a matriz de covariâncias condicional.

Os QQ-plot dos resíduos padronizados do ajuste CCC-GARCH(1,1) estão na Figura 5.7 e eles indicam que a suposição de normalidade não é apropriada. Portanto, é razoável considerar uma distribuição t-student multivariada.

Tabela 5.6: Testes de Portmanteau para os resíduos e o quadrado dos resíduos do ajuste CCC-GARCH(1,1), materiais de construção.

Teste Multivariado Portmanteau: Tipo Ljung-Box		
Hipótese Nula: não correlação serial		
	Estat. do teste	p.valor
Resíduos	182.1743	0,6830
Resíduos ao quadrado	165,9967	0,9127
Dist. sob Hipótese Nula: qui-quadrado com 192 graus de liberdade		
Total Observ.: 400		

Segundo a Seção 3.3.1, o modelo CCC-GARCH(1,1) ajustado para o vetor de log-retornos \mathbf{r}_t (4×1), é dado por

$$\mathbf{r}_t = \begin{pmatrix} 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,604 \end{pmatrix} \mathbf{r}_{t-1} - \begin{pmatrix} 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,6981 \end{pmatrix} \varepsilon_{t-1} + \varepsilon_t \quad (5.1)$$

$$\Sigma_t = \Delta_t \mathbf{R} \Delta_t,$$

com

$$\Delta_t = \begin{pmatrix} \sigma_{1t} & 0 & 0 & 0 \\ 0 & \sigma_{2t} & 0 & 0 \\ 0 & 0 & \sigma_{3t} & 0 \\ 0 & 0 & 0 & \sigma_{4t} \end{pmatrix}, \quad \widehat{\mathbf{R}} = \begin{pmatrix} 1,0000 & 0,7677 & 0,6919 & 0,4122 \\ 0,7677 & 1,0000 & 0,7245 & 0,4106 \\ 0,6919 & 0,7245 & 1,0000 & 0,3604 \\ 0,4122 & 0,4106 & 0,3604 & 1,0000 \end{pmatrix},$$

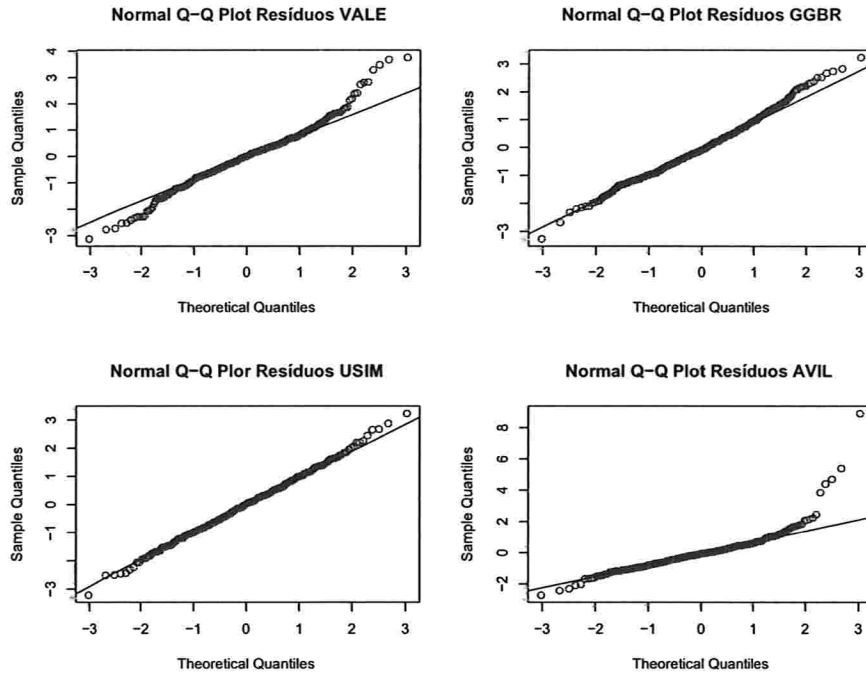


Figura 5.7: *QQ-plot dos resíduos padronizados, materiais de construção.*

em que σ_{it} , para $i = 1, 2, 3, 4$, é o desvio padrão da i -ésima série de log-retornos e além disso, ε_t segue uma distribuição t-student multivariada com aproximadamente 8 graus de liberdade e matriz de covariâncias Σ_t .

As variâncias condicionais seguem os seguintes modelos estimados:

$$\begin{aligned}
 \sigma_{1t}^2 &= 0,00002168 + 0,05728000\varepsilon_{1(t-1)}^2 + 0,89870000\sigma_{1(t-1)}^2 \\
 \sigma_{2t}^2 &= 0,04365000\varepsilon_{2(t-1)}^2 + 0,93540000\sigma_{2(t-1)}^2 \\
 \sigma_{3t}^2 &= 0,03535000\varepsilon_{3(t-1)}^2 + 0,95260000\sigma_{3(t-1)}^2 \\
 \sigma_{4t}^2 &= 0,00030040 + 0,21070000\varepsilon_{4(t-1)}^2 + 0,35140000\sigma_{4(t-1)}^2.
 \end{aligned} \tag{5.2}$$

As Figuras 5.8 e 5.9 apresentam as variâncias e covariâncias condicionais estimadas dos log-retornos. Nota-se que nas ações VALE, GERDAU e USIMINAS a variabilidade é maior no ano 2009 decrescendo até estabilizar-se um pouco mais no ano 2010. A ação AÇOS VILL tem alguns picos na variabilidade ao longo do tempo.

As covariâncias condicionais têm um comportamento similar às variâncias condicionais. Entre as ações VALE, GERDAU e USIMINAS estas foram maiores no ano 2009 e menores

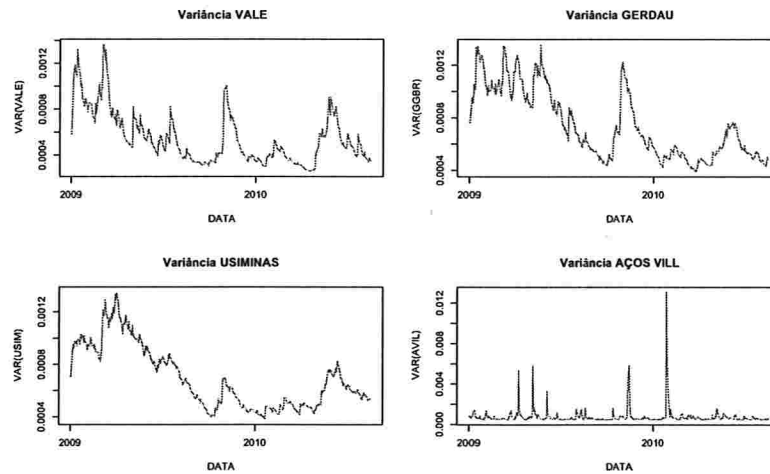


Figura 5.8: Variâncias estimadas pelo ajuste $CCC-GARCH(1,1)$, materiais de construção.

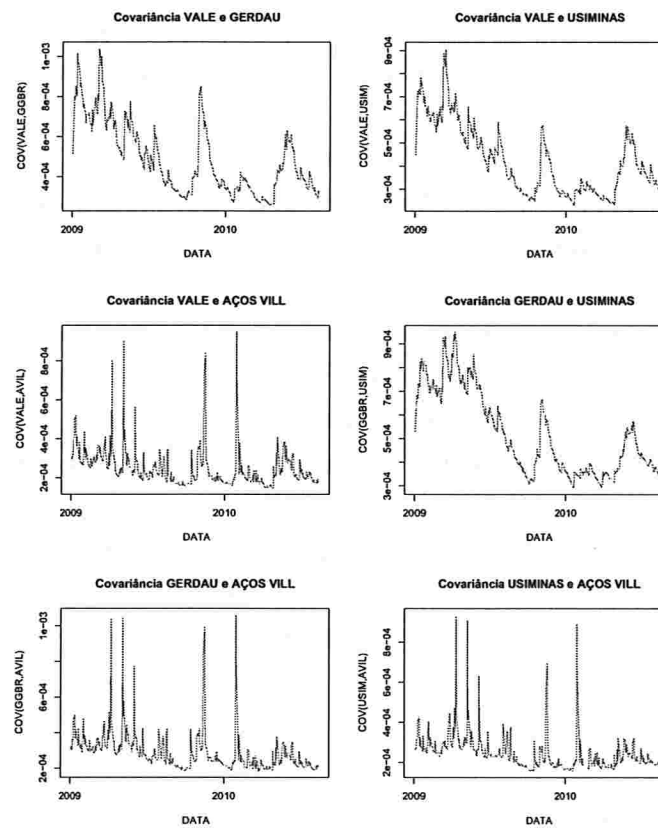


Figura 5.9: Covariâncias estimadas pelo ajuste $CCC-GARCH(1,1)$, materiais de construção.

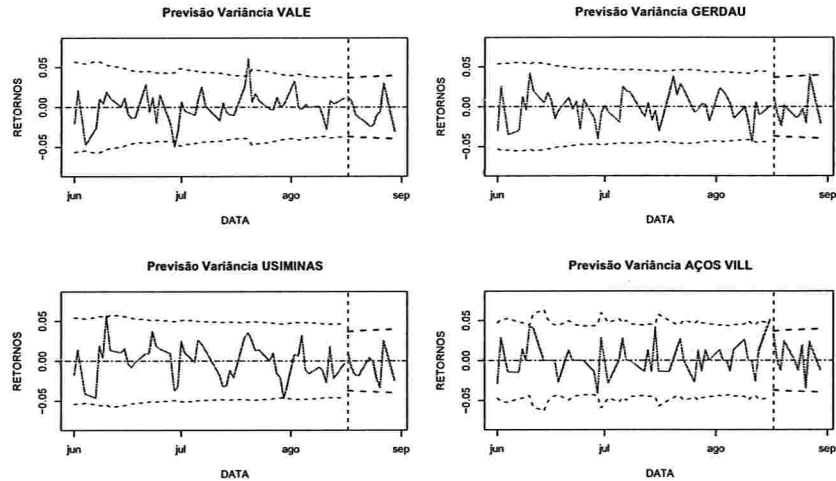


Figura 5.10: *Previsão e intervalos de confiança para os log-retornos 10 dias à frente das ações de materiais de construção, ajuste CCC-GARCH(1,1).*

no ano 2010. Já para a ação AÇOS VILL, todas as covariâncias entre ela e as outras ações têm um comportamento parecido, apresentando vários picos em alguns períodos.

Apresentam-se na Figura 5.10 os log-retornos das ações Materiais de Construção, no período 1 de Junho de 2010 até 30 de Agosto de 2010, as previsões 10 dias à frente e os respectivos intervalos de confiança. Antes da linha vertical pontilhada estão os log-retornos das ações e o intervalo de confiança estimado usando as variâncias estimadas pelo ajuste CCC-GARCH(1,1) dado em (5.1). Após desta linha vertical, apresentam-se com linha pontilhada grossa o intervalo de confiança estimado dos log-retornos, 10 dias à frente, usando as variâncias previstas. Nota-se que os intervalos previstos contêm os log-retornos, mostrando que o ajuste CCC-GARCH(1,1) é adequado para fazer previsões neste conjunto de ações.

A Figura 5.11, apresenta as covariâncias condicionais, no período 1 de Junho de 2010 até 16 de Agosto de 2010, junto com as previsões 10 dias à frente depois da linha vertical pontilhada e nota-se que todas as covariâncias condicionais previstas seguem um comportamento suave.

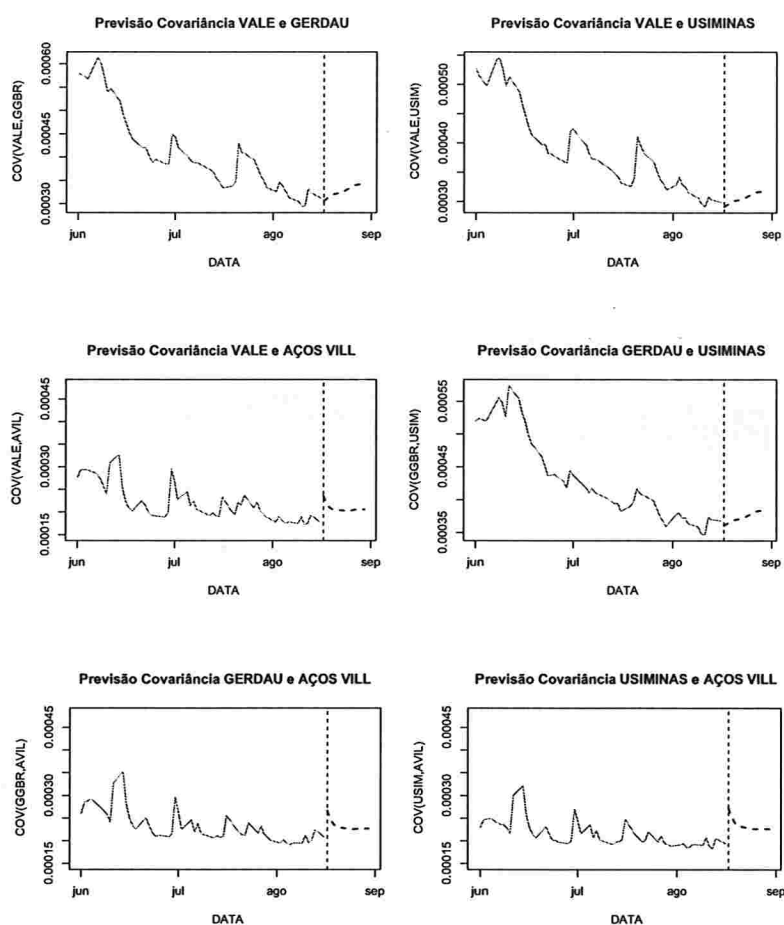


Figura 5.11: Previsão 10 dias à frente das covariâncias condicionais das ações materiais de construção, ajuste CCC-GARCH(1,1).

5.2 Ações Construção Civil

Considera-se o conjunto das ações que fazem parte do setor Construção Civil.

Na Figura 5.12 apresentam-se os log-retornos das ações CYRELA REALT, INPAR S/A, TECNISA e GAFISA no período 2 de Janeiro de 2009 até 16 de Agosto de 2010. Nenhuma das ações apresenta tendência ao longo do tempo mas existem alguns conglomerados de variabilidade alta em cada série.

Na Figura 5.13 encontram-se as FAC e FAC cruzadas dos log-retornos e observa-se que em alguns gráficos existe autocorrelação e autocorrelação cruzada significativa antes do lag 10. Já para o quadrado dos log-retornos, a Figura 5.14 apresenta as FAC e FAC cruzadas deles. Nota-se que existe autocorrelação e autocorrelação cruzada forte na maioria dos grá-

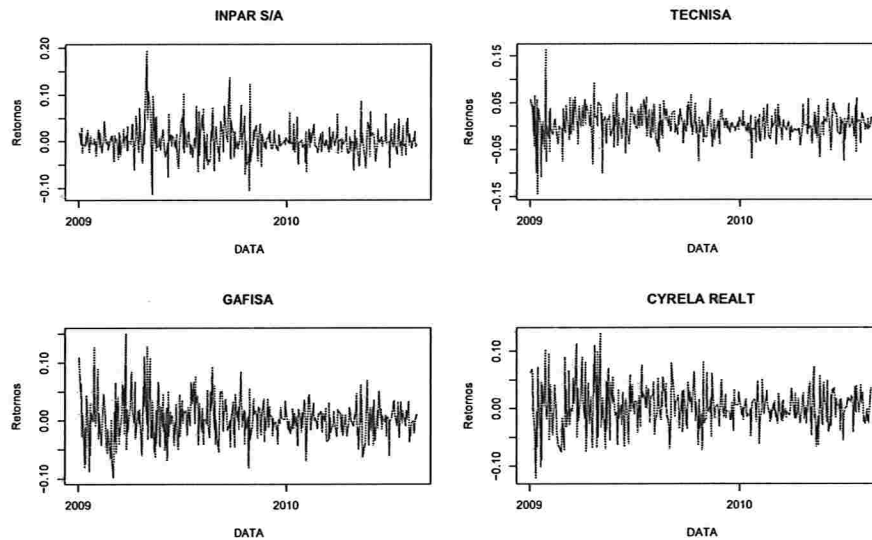


Figura 5.12: Ações construção civil, 2 de Janeiro de 2009 até 19 de Agosto de 2010.

ficos, mostrando a existência de uma dependência quadrática nos log-retornos.

Os ajustes GARCH multivariados considerados foram EWMA, DVEC(1,1), BEKK(1,1,1), CCC com ajustes GARCH(1,1) univariados para cada um dos elementos na diagonal da matriz dada por (3.29) e Componentes Principais com ajustes GARCH(1,1) para cada uma das componentes. Alguns outros ajustes, como por exemplo BEKK(2,1), BEKK(1,2), DVEC(1,2),..., foram feitos, mas os resultados obtidos não foram melhores. Além dos modelos GARCH multivariados, foi considerado um ajuste ARMA(1,1) para os log-retornos.

Os valores de AIC e BIC para cada um dos modelos estão apresentados na Tabela 5.7. O modelo que apresentou o menor valor nos dois critérios foi o modelo CCC-GARCH(1,1).

Tabela 5.7: Critérios de seleção AIC e BIC, construção civil.

Modelo	AIC	BIC
EWMA	-7083,857	-7027,976
DVEC(1,1)	-7102,707	-6931,074
BEKK(1,1)	-7050,526	-6830,996
CCC-GARCH(1,1)	-7135,983	-7036,197
PCOMP-GARCH(1,1)	-7125,3	-7025,513

Apresentam-se na Tabela 5.8 os testes multivariados Portmanteau para o quadrado dos resíduos de cada ajuste. O único modelo ajustado que tem um p -valor significativo ao nível

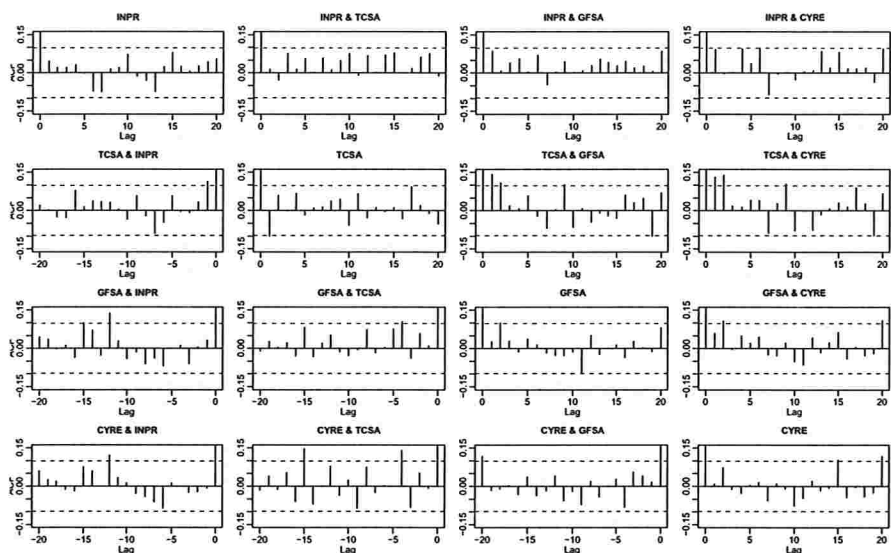


Figura 5.13: FAC e FAC cruzadas dos log-retornos, construção civil.

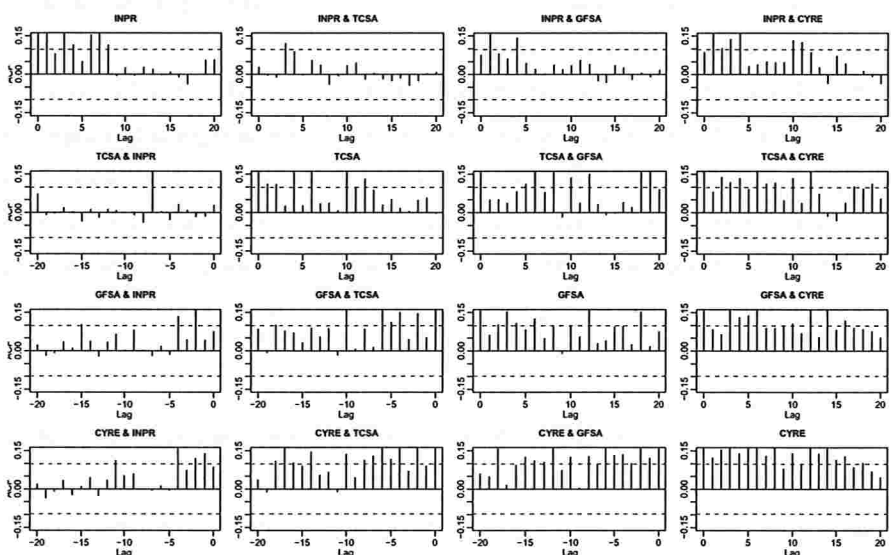


Figura 5.14: FAC e FAC cruzadas dos log-retornos ao quadrado, construção civil.

de significância de 5% é o modelo CCC-GARCH(1,1).

Tabela 5.8: Testes de Portmanteau para o quadrado dos resíduos dos ajustes, construção civil.

Teste Multivariado Portmanteau: Tipo Ljung-Box		
Hipótese Nula: não correlação serial		
Modelo	Estat. do teste	p.valor
EWMA	333,0461	<0,001
DVEC(1,1)	227,6082	0,040
BEKK(1,1)	344,881	<0,001
CCC-GARCH(1,1)	218,6206	0,091
PCOMP-GARCH(1,1)	262,8459	<0,001

Segundo estes critérios o modelo mais adequado é o modelo CCC-GARCH(1,1), igual que no grupo das ações Materiais de Construção.

A Tabela 5.9 apresenta o resumo do ajuste CCC-GARCH(1,1). Os valores estimados da constante na média, $C(1)$, $C(2)$, $C(3)$ e $C(4)$, foram todos não significativos, ao nível de 5%. Em cada um dos ajustes GARCH univariados, os parâmetros estimados $A(2, 2)$, $A(3, 3)$ e $A(4, 4)$ foram não significativos ao nível de significância de 5%. Na parte ARMA, os únicos parâmetros significativos foram $AR(1; 3, 3)$ e $MA(1; 3, 3)$. O número de graus de liberdade estimado é 11,27724 com um desvio padrão de 3,204989.

Na Tabela 5.10 apresentam-se os testes do ajuste CCC-GARCH(1,1) para as ações do setor Construção Civil. Os testes de normalidade de Jaque-Bera rejeitam a hipótese de normalidade das ações INPAR S/A, TECNISA e GAFISA, e o teste de Shapiro-Wilk rejeita a hipótese de normalidade unicamente da ação INPAR S/A, com um nível de significância de 5%. Em relação aos resíduos do ajuste e ao quadrado destes, os testes de Ljung-Box não rejeitam a hipótese de não correlação serial entre eles, também ao nível 5%, o qual é confirmado com os testes Multiplicadores de Lagrange.

As Figuras 5.15 e 5.16 apresentam as FAC e FAC cruzadas dos resíduos e do quadrado dos resíduos do ajuste. Observa-se que aparentemente existem alguns lags em que a autocorrelação e a autocorrelação cruzada são significativas em ambos os casos, mas para ter certeza foram feitos os testes multivariados de Portmanteu e são apresentados na Tabela 5.11. Os p -valores dos testes foram 0,0975 e 0,0911, respectivamente, não rejeitando nos dois casos a hipótese nula com um nível de significância de 5%, indicando a adequação do modelo ajustado.

Tabela 5.9: Resumo do ajuste CCC-GARCH(1,1), construção civil.

Graus de liberdade estimado:				
Graus	Std.Error			
11,27724	3,204989			
Parâmetros Estimados CCC-GARCH(1,1):				
	Estimativa	D.P.	t-valor	Pr(> t)
C(1)	-0,00060813	0,00177862	-0,3419	0,3663
C(2)	0,00232917	0,00207217	1,1240	0,3080
C(3)	0,00132261	0,00233443	0,5666	0,2857
C(4)	0,00171696	0,00190495	0,9013	0,1845
AR(1; 1, 1)	-0,28140269	0,83417215	-0,3373	0,3680
AR(1; 2, 2)	-0,30709794	0,62437552	-0,4918	0,3116
AR(1; 3, 3)	-0,65082643	0,23947557	-2,7177	0,0034
AR(1; 4, 4)	-0,31823720	0,35001105	-0,9092	0,1819
MA(1; 1, 1)	0,22338190	0,84541741	0,2642	0,3959
MA(1; 2, 2)	0,23784022	0,63712858	0,3733	0,3546
MA(1; 3, 3)	0,57559432	0,25623201	2,2464	0,0126
MA(1; 4, 4)	0,23285179	0,35315582	0,6593	0,2550
A(1, 1)	0,00019691	0,00006248	3,1518	< 0,001
A(2, 2)	0,00001918	0,00001208	1,5877	0,0566
A(3, 3)	0,00002324	0,00001497	1,5522	0,0607
A(4, 4)	0,00002354	0,00001588	1,4824	0,6951
ARCH(1; 1, 1)	0,25168568	0,06107458	4,1210	< 0,001
ARCH(1; 2, 2)	0,06153823	0,02351847	2,6166	0,0046
ARCH(1; 3, 3)	0,05823585	0,02110176	2,7598	0,0031
ARCH(1; 4, 4)	0,05796852	0,02163816	2,6790	0,0038
GARCH(1; 1, 1)	0,53260990	0,10763895	4,9481	< 0,001
GARCH(1; 2, 2)	0,91537872	0,03034441	30,1663	< 0,001
GARCH(1; 3, 3)	0,91717719	0,02884130	31,8008	< 0,001
GARCH(1; 4, 4)	0,92000000	0,03022000	30,4500	< 0,001
Matriz de Correlação Condicional Constante:				
	INPR	TCSA	GFSA	CYRE
INPR	1,0000	0,2543	0,3678	0,3681
TCSA	0,2543	1,0000	0,3841	0,3849
GFSA	0,3678	0,3841	1,0000	0,8086
CYRE	0,3681	0,3849	0,8086	1,0000
Desvios Padrão:				
	INPR	TCSA	GFSA	CYRE
INPR	NA	0,05464	0,04579	0,04894
TCSA	0,05464	NA	0,04711	0,04848
GFSA	0,04579	0,04711	NA	0,01973
CYRE	0,04894	0,04848	0,01973	NA

Tabela 5.10: Testes do ajuste CCC-GARCH(1,1), construção civil.

Testes Ajuste CCC-GARCH(1,1):						
Testes de Normalidade:						
	Jarque-Bera	P-valor	Shapiro-Wilk	P-valor		
INPR	122,733	<0,001	0,9682	<0,001		
TCSA	8,395	0,015	0,9884	0,866		
GFSA	6,173	0,046	0,9859	0,6486		
CYRE	1,573	0,456	0,9810	0,177		
Teste de Ljung-Box para os resíduos padronizados:						
	Estatística	P-valor	$\chi^2 - d.f.$			
INPR	9,379	0,6703	12			
TCSA	3,048	0,9952	12			
GFSA	14,003	0,3005	12			
CYRE	17,445	0,1336	12			
Teste de Ljung-Box para o quadrado dos resíduos padronizados:						
	Estatística	P-valor	$\chi^2 - d.f.$			
INPR	9,917	0,6233	12			
TCSA	15,044	0,2391	12			
GFSA	11,294	0,5039	12			
CYRE	13,469	0,3359	12			
Teste Multiplicadores de Lagrange:						
	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
INPR	-1,3068	-0,2498	-0,2862	0,19041	0,8790	2,1537
TCSA	-0,9591	0,4325	1,3755	-0,88909	0,8232	-0,7679
GFSA	-0,2828	1,3914	0,7240	0,07722	0,3246	-0,6978
CYRE	-0,3900	0,3200	0,4707	0,11818	1,1044	-0,1471
	Lag 7	Lag 8	Lag 9	Lag 10	Lag 11	Lag 12
INPR	-0,25328	0,23839	0,009396	0,49852	-0,08722	0,1416
TCSA	-1,14646	-0,03317	2,808409	0,03087	1,89935	-0,7035
GFSA	0,29644	-0,58990	1,540178	-1,54728	1,98226	-0,1073
CYRE	0,03938	-1,84373	-0,288315	2,26875	-1,34023	-0,2671
C						
INPR	0,9929					
TCSA	0,1060					
GFSA	1,6606					
CYRE	-0,4415					
	TR^2	P-valor	F-estat.	P-valor		
INPR	9,057	0,6981	0,843	0,7055		
TCSA	18,203	0,1097	1,736	0,1515		
GFSA	13,276	0,3493	1,250	0,3600		
CYRE	13,151	0,3582	1,237	0,3680		

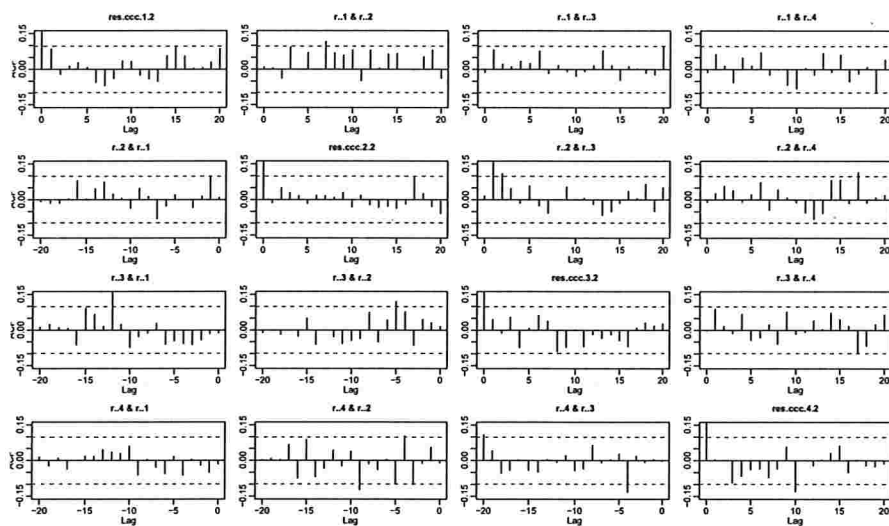


Figura 5.15: *FAC* resíduos do ajuste *CCC-GARCH(1,1)* construção civil.

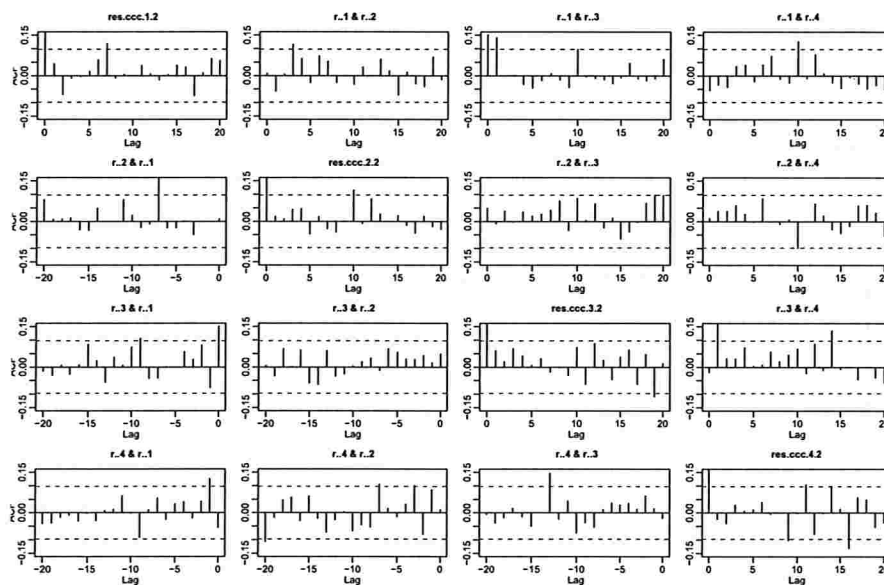


Figura 5.16: *FAC* resíduos ao quadrado do ajuste *CCC-GARCH(1,1)* construção civil.

Tabela 5.11: Testes de Portmanteau para os resíduos e resíduos ao quadrado do ajuste CCC-GARCH(1,1), construção civil.

Teste Multivariado Portmanteau: Tipo Ljung-Box		
Hipótese Nula: não correlação serial		
	Estat. do teste	p.valor
Resíduos	217,8122	0,0975
Resíduos ao quadrado	218,6206	0,0911
Dist. sob Hipótese Nula: qui-quadrado com 192 graus de liberdade		
Total Observ.: 400		

Os QQ-plot dos resíduos padronizados do ajuste CCC-GARCH(1,1) para as ações do setor construção civil estão na Figura 5.17 e eles indicam que a suposição de normalidade não é apropriada. Portanto, é razoável considerar uma distribuição t-student multivariada.

O modelo CCC-GARCH(1,1) ajustado para o vetor de log-retornos \mathbf{r}_t (4×1) do setor Construção Civil é dado por

$$\mathbf{r}_t = - \begin{pmatrix} 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,651 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 \end{pmatrix} \mathbf{r}_{t-1} + \begin{pmatrix} 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,576 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 \end{pmatrix} \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\varepsilon}_t \quad (5.3)$$

$$\boldsymbol{\Sigma}_t = \boldsymbol{\Delta}_t \mathbf{R} \boldsymbol{\Delta}_t,$$

com

$$\boldsymbol{\Delta}_t = \begin{pmatrix} \sigma_{1t} & 0 & 0 & 0 \\ 0 & \sigma_{2t} & 0 & 0 \\ 0 & 0 & \sigma_{3t} & 0 \\ 0 & 0 & 0 & \sigma_{4t} \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} 1,0000 & 0,2543 & 0,3678 & 0,3681 \\ 0,2543 & 1,0000 & 0,3841 & 0,3849 \\ 0,3678 & 0,3841 & 1,0000 & 0,8086 \\ 0,3681 & 0,3849 & 0,8086 & 1,0000 \end{pmatrix},$$

em que σ_{it} , para $i = 1, 2, 3, 4$, é o desvio padrão da i -ésima série de log-retornos e além disso, $\boldsymbol{\varepsilon}_t$ segue uma distribuição t-student multivariada com aproximadamente 11 graus de liberdade e matriz de covariâncias $\boldsymbol{\Sigma}_t$.

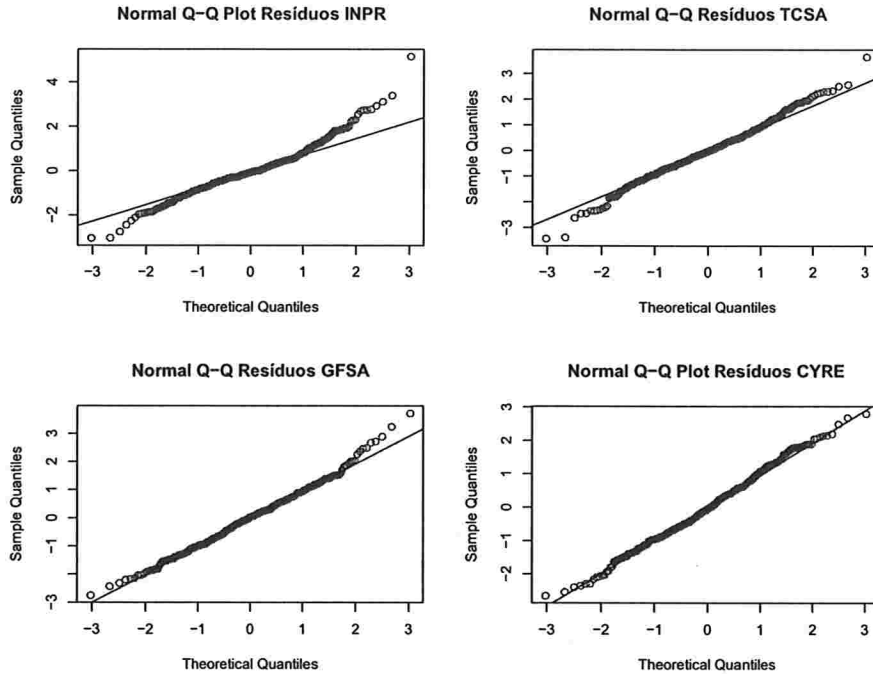


Figura 5.17: *QQ-plot dos resíduos padronizados, construção civil.*

As variâncias condicionais seguem os seguintes modelos:

$$\begin{aligned}
 \sigma_{1t}^2 &= 0,00019691 + 0,25168568\varepsilon_{1(t-1)}^2 + 0,53260990\sigma_{1(t-1)}^2 \\
 \sigma_{2t}^2 &= 0,06153823\varepsilon_{2(t-1)}^2 + 0,91537872\sigma_{2(t-1)}^2 \\
 \sigma_{3t}^2 &= 0,05823585\varepsilon_{3(t-1)}^2 + 0,91717719\sigma_{3(t-1)}^2 \\
 \sigma_{4t}^2 &= 0,05796852\varepsilon_{4(t-1)}^2 + 0,92000000\sigma_{4(t-1)}^2.
 \end{aligned} \tag{5.4}$$

As Figuras 5.18 e 5.19 apresentam as variâncias e covariâncias condicionais dos log-retornos, respectivamente. Todas as ações deste setor têm maior variabilidade no primeiro semestre do ano 2009 e depois começa a estabilizar-se. As covariâncias têm um comportamento similar às variâncias sendo também maiores no começo do ano 2009 decrescendo até estabilizarse um pouco mais no ano 2010.

Apresentam-se na Figura 5.20 os log-retornos das ações Construção Civil, no período 1 de Junho de 2010 até 30 de Agosto de 2010, e as previsões e respectivos intervalos de confiança para os log-retornos. Antes da linha vertical pontilhada estão os log-retornos das ações e o intervalo estimado usando as variâncias condicionais estimadas pelo modelo CCC-GARCH(1,1). Após desta linha vertical, apresentam-se com linha pontilhada grossa, o inter-

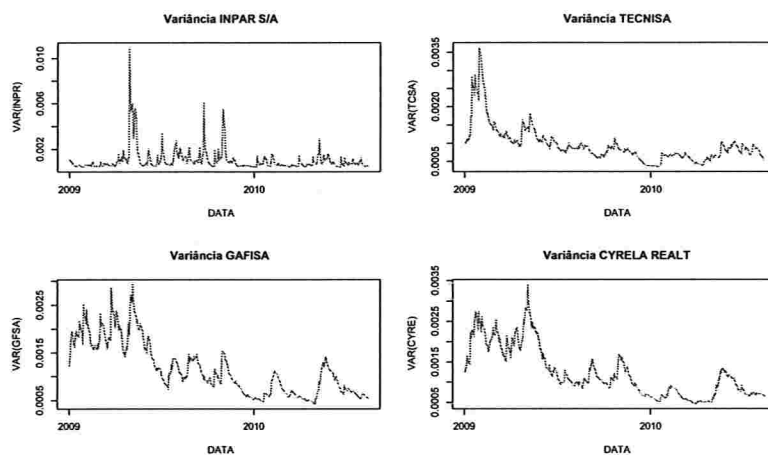


Figura 5.18: Variâncias condicionais estimadas das ações construção civil ajuste CCC-GARCH(1,1).

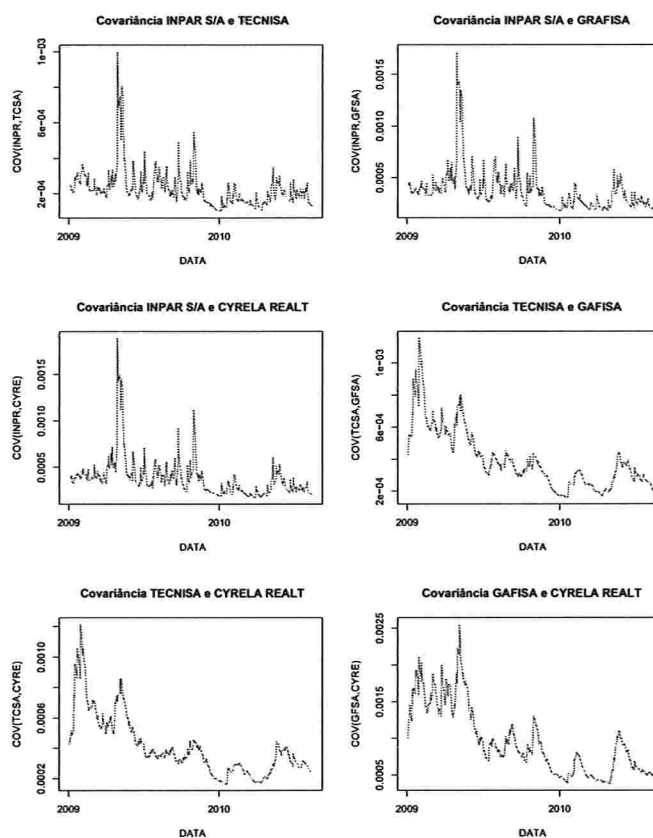


Figura 5.19: Covariâncias condicionais estimadas das ações construção civil ajuste CCC-GARCH(1,1).

valo de confiança estimado dos log-retornos 10 dias à frente, usando as variâncias condicionais previstas. Nota-se que os intervalos previstos contêm os log-retornos reais, mostrando que o ajuste CCC-GARCH(1,1) funciona bem para este conjunto de ações neste período de tempo.

Na Figura 5.21 apresentam-se as covariâncias condicionais no período 1 de Junho de 2010 até 30 de Agosto de 2010. Após da linha vertical pontilhada estão as previsões 10 dias à frente e nota-se que todas as covariâncias previstas têm uma tendência crescente.

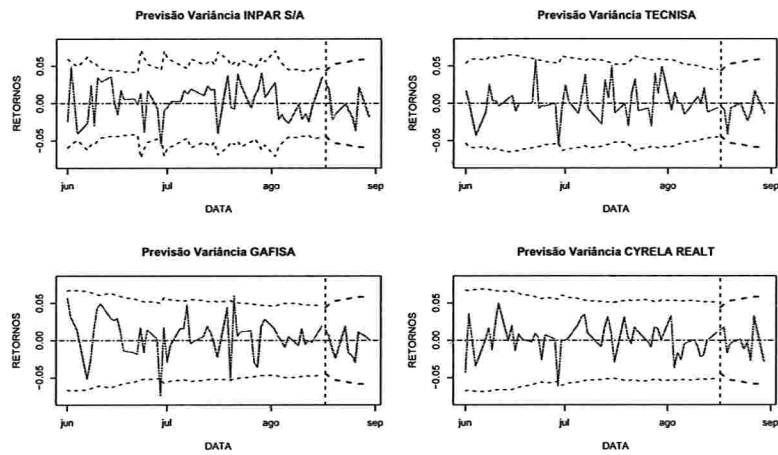


Figura 5.20: *Previsão e intervalos de confiança para os log-retornos 10 dias à frente das ações construção civil ajuste, CCC-GARCH(1,1).*

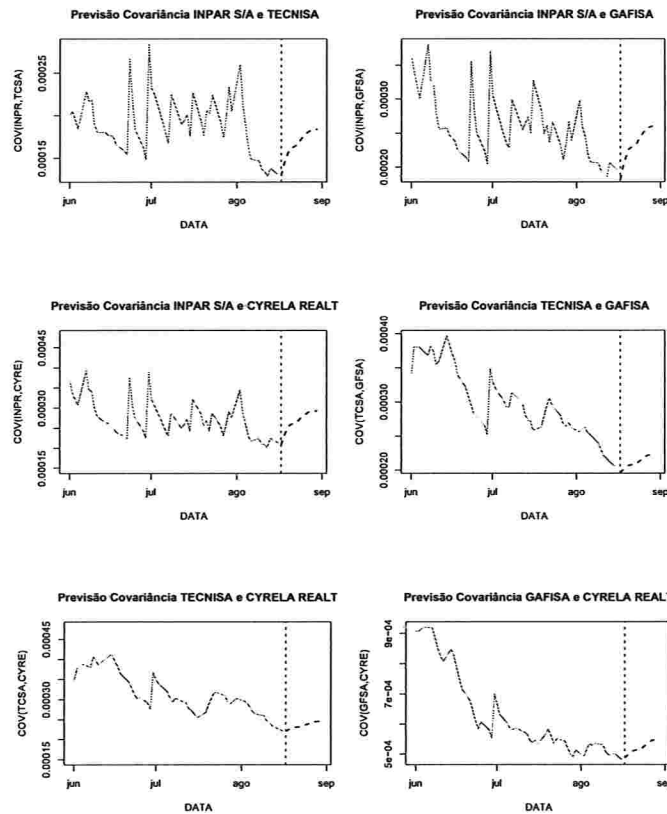


Figura 5.21: *Previsão 10 dias à frente das covariâncias condicionais das ações construção civil, ajuste CCC-GARCH(1,1).*

Capítulo 6

Conclusões e Trabalhos Futuros

Neste trabalho foram apresentados os modelos GARCH multivariados em dois grupos. O primeiro foi o grupo dos modelos que ajustam de forma direta a matriz de covariâncias condicional, usando matrizes de parâmetros análoga aos modelos GARCH univariados. As principais complicações destes modelos foram as condições para gerar matrizes positivas semidefinidas, a estacionariedade na covariância, a estimação devido ao grande número de parâmetros e finalmente os testes de avaliação dos modelos ajustados.

Cada um destes problemas foi abordado ao longo do Capítulo 3 chegando em soluções baseadas em restrições para as matrizes de parâmetros. Por exemplo, a Proposição 3.3 dá uma condição suficiente sobre estas matrizes para que o modelo seja estacionário.

O problema da estimação foi evidenciado nas simulações, em que alguns dos parâmetros estimados pelo método Monte Carlo nos modelos DVEC(1,1) e BEKK(1,1,1), não conseguiram convergir aos verdadeiros parâmetros. Embora isto aconteça, as variâncias e covariâncias condicionais simuladas e estimadas coincidiram em todos os casos, o qual é muito importante já que o objetivo final da modelagem GARCH multivariada é prever as volatilidades e correlações cruzadas.

Dos modelos diretos considerados, o modelo BEKK envolve uma variabilidade maior uma vez que ele gera uma estrutura mais dinâmica na matriz de covariâncias.

O maior problema com a modelagem direta foi que o número de parâmetros é muito alto. Para tentar resolver este problema, foram propostos alguns métodos com base em modelos GARCH univariados, chamados modelos indiretos. Os dois principais modelos são o modelo de Correlação Condicional Constante que supõe a matriz de correlação é constante, que é uma restrição muito forte, e o modelo de Componentes Principais que ajusta modelos GARCH univariados para cada componente principal, mas supõe que elas são condicional-

mente não correlacionadas, e isto não necessariamente é verdade.

Tanto para os modelos diretos como indiretos, o método de escolha do melhor modelo foi baseado primeiro nos critérios AIC e BIC, além dos Multiplicadores de Lagrange e os testes de Ljung-Box univariados para as séries ao quadrado. No contexto multivariado o teste principal foi o teste de Portmanteau, o qual avalia a hipótese nula de não autocorrelação e autocorrelação cruzada entre os resíduos ao quadrado do modelo ajustado.

Como trabalhos futuros poderia-se considerar modelos em que as matrizes de parâmetros variem com o tempo. Também seria interessante aprofundar na generalização do modelo GARCH multivariado de Componentes Principais em que as componentes são condicionalmente correlacionadas, modelo conhecido como *GARCH Ortogonal Generalizado* dado por Weide (2002) . Além do anterior, uma generalização do modelo CCC, como por exemplo o modelo de Correlação Condicional Dinâmica dado por Engle (2000), poderia ser considerada em estudos posteriores.

Apêndice A

Provas dos Resultados

Neste apêndice apresentam-se as provas das propriedades e proposições dadas ao longo deste trabalho.

A condição suficiente para garantir que a matriz Σ_t dada no modelo DVEC(p,q) da equação (3.9) seja PSD, é que as matrizes de parâmetros sejam todas PSD. Para demonstrar isto, serão considerados os seguintes lemas.

Lema A.1. *Seja \mathbf{A} uma matriz simétrica de dimensão $(n \times n)$ positiva semi-definida e \mathbf{b} um vetor $(n \times 1)$ diferente de $\mathbf{0}$, então $\mathbf{A} \odot \mathbf{b}\mathbf{b}'$ é também PSD.*

Prova.

Pela Propriedade 2.4 tem-se que

$$\mathbf{A} \odot \mathbf{b}\mathbf{b}' = \mathit{diag}[\mathbf{b}] \mathbf{A} \mathit{diag}[\mathbf{b}], \quad (\text{A.1})$$

em que $\mathit{diag}[\mathbf{b}]$ é uma matriz diagonal com $\mathbf{b}^{(i)}$ o i -ésimo elemento na diagonal. Como \mathbf{A} é PSD, pela decomposição de Cholesky ela pode ser escrita como $\mathbf{A} = \mathbf{L}\mathbf{L}'$, onde \mathbf{L} é uma matriz triangular inferior. Assim, da equação (A.1) tem-se que

$$\begin{aligned} \mathbf{A} \odot \mathbf{b}\mathbf{b}' &= \mathit{diag}[\mathbf{b}] \mathbf{L}\mathbf{L}' \mathit{diag}[\mathbf{b}] \\ &= (\mathit{diag}[\mathbf{b}] \mathbf{L}) (\mathbf{L}' \mathit{diag}[\mathbf{b}]) \\ &= (\mathit{diag}[\mathbf{b}] \mathbf{L}) (\mathit{diag}[\mathbf{b}] \mathbf{L})'. \end{aligned} \quad (\text{A.2})$$

Seja \mathbf{x} um vetor diferente de $\mathbf{0}$, da equação (A.2) tem-se que

$$\begin{aligned} \mathbf{x}'(\mathbf{A} \odot \mathbf{b}\mathbf{b}')\mathbf{x} &= \mathbf{x}'(\text{diag}[\mathbf{b}] \mathbf{L})(\text{diag}[\mathbf{b}] \mathbf{L})'\mathbf{x} \\ &= (\mathbf{x}'\text{diag}[\mathbf{b}] \mathbf{L})(\mathbf{x}'\text{diag}[\mathbf{b}] \mathbf{L})' \\ &= \|\mathbf{x}'\text{diag}[\mathbf{b}] \mathbf{L}\|^2 \geq 0, \end{aligned}$$

em que $\|\cdot\|^2$ denota a magnitude ao quadrado de um vetor. Do resultado anterior pode-se concluir que a matriz $\mathbf{A} \odot \mathbf{b}\mathbf{b}'$ é PSD. ■

Lema A.2. *Sejam $\mathbf{A}_1, \dots, \mathbf{A}_p$ matrizes simétricas de dimensão $(n \times n)$ positivas semi-definidas, então a matriz $\mathbf{S}_p = \mathbf{A}_1 + \dots + \mathbf{A}_p$ é PSD.*

Prova.

Seja \mathbf{x} um vetor de dimensão $(n \times 1)$ diferente de $\mathbf{0}$, logo como $\mathbf{A}_1, \dots, \mathbf{A}_p$ são todas PSD, então

$$\begin{aligned} \mathbf{x}'\mathbf{A}_1\mathbf{x} &\geq 0, \\ \mathbf{x}'\mathbf{A}_2\mathbf{x} &\geq 0, \\ &\vdots \\ \mathbf{x}'\mathbf{A}_p\mathbf{x} &\geq 0, \end{aligned}$$

e somando em ambos os lados destas desigualdades, tem-se que

$$\begin{aligned} \mathbf{x}'\mathbf{A}_1\mathbf{x} + \mathbf{x}'\mathbf{A}_2\mathbf{x} + \dots + \mathbf{x}'\mathbf{A}_p\mathbf{x} &\geq 0, \\ \mathbf{x}'(\mathbf{A}_1 + \mathbf{A}_2 + \dots + \mathbf{A}_p)\mathbf{x} &\geq 0, \\ \mathbf{x}'\mathbf{S}_p\mathbf{x} &\geq 0. \end{aligned}$$

Portanto, \mathbf{S}_p é PSD. ■

Lema A.3. *Se \mathbf{A} e \mathbf{B} são matrizes PSD de dimensão $(n \times n)$, então $\mathbf{A} \odot \mathbf{B}$ é também uma matriz PSD.*

Prova.

Como \mathbf{B} é PSD, da decomposição espectral tem-se que

$$\mathbf{B} = \sum_{i=1}^n \lambda_i \mathbf{b}_i \mathbf{b}_i',$$

em que os $\lambda_i \geq 0$ para todo i , representam os autovalores de \mathbf{A} e os \mathbf{b}_i representam os correspondentes autovetores. Logo,

$$\begin{aligned} \mathbf{A} \odot \mathbf{B} &= \mathbf{A} \odot \sum_{i=1}^n \lambda_i \mathbf{b}_i \mathbf{b}_i' \\ &= \sum_{i=1}^n \lambda_i (\mathbf{A} \odot \mathbf{b}_i \mathbf{b}_i'). \end{aligned}$$

Pelo Lema A.1, $\mathbf{A} \odot \mathbf{b}_i \mathbf{b}_i'$ é PSD, já que \mathbf{A} é PSD e como $\lambda_i \geq 0$, $i = 1, \dots, n$, então $\lambda_i (\mathbf{A} \odot \mathbf{b}_i \mathbf{b}_i')$ é também PSD. Assim, $\mathbf{A} \odot \mathbf{B}$ é uma soma de n matrizes PSD e pelo Lema A.2 obtem-se que $\mathbf{A} \odot \mathbf{B}$ é PSD. ■

Proposição A.1. *Se as matrizes de parâmetros no modelo DVEC(p, q) dado em (3.9) são PSD, então as matrizes geradas Σ_t são PSD para todo t .*

Prova.

O modelo (3.9) é dado por

$$\Sigma_t = \mathbf{A}_0 + \sum_{i=1}^p \mathbf{A}_i \odot (\boldsymbol{\varepsilon}_{t-i} \boldsymbol{\varepsilon}_{t-i}') + \sum_{j=1}^q \mathbf{B}_j \odot \Sigma_{t-j}.$$

Seja $M = \max\{p, q\}$ e $\Sigma_1, \dots, \Sigma_{M-1}$ matrizes iniciais geradas de tal forma que sejam PSD. Provando por indução, suponha que para $t \geq M$, as matrizes $\Sigma_1, \dots, \Sigma_{t-1}$ são PSD e será demonstrado que Σ_t é também PSD.

Assim, como \mathbf{A}_i é PSD e $\boldsymbol{\varepsilon}_{t-i} \neq \mathbf{0}$, $i = 1, \dots, p$, então pelo Lema A.1, $\mathbf{A}_i \odot (\boldsymbol{\varepsilon}_{t-i} \boldsymbol{\varepsilon}_{t-i}')$ é PSD, e do Lema A.2,

$$\mathbf{S}_1 = \sum_{i=1}^p \mathbf{A}_i \odot (\boldsymbol{\varepsilon}_{t-i} \boldsymbol{\varepsilon}_{t-i}'),$$

é PSD. Pelo Lema A.3, $\mathbf{B}_j \odot \boldsymbol{\Sigma}_{t-j}$ é PSD, já que \mathbf{B}_j e $\boldsymbol{\Sigma}_{t-j}$ são PSD, $j = 1, \dots, q$.

Do Lema A.2

$$\mathbf{S}_2 = \sum_{j=1}^q \mathbf{B}_j \odot \boldsymbol{\Sigma}_{t-j},$$

é PSD. Finalmente, como $\boldsymbol{\Sigma}_t$ é a soma de \mathbf{A}_0 (PSD por hipótese), \mathbf{S}_1 e \mathbf{S}_2 , então pelo Lema A.2, esta matriz é também PSD para todo t . ■

Agora, apresentam-se as provas de alguns resultados do Capítulo 3.

Prova da Proposição 3.1. Por construção $\boldsymbol{\Sigma}_t$ é simétrica para todo t . Já para a segunda parte, considera-se que $M = \max\{p, q\}$ e que as matrizes $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \dots, \boldsymbol{\Sigma}_{M-1}$ são matrizes iniciais, dadas de tal forma que sejam PSD. Mostra-se que $\boldsymbol{\Sigma}_t$ é PSD para todo $t \geq M$, usando o método de indução, isto é, supondo que as matrizes $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \dots, \boldsymbol{\Sigma}_{t-1}$ são PSD¹, e então demonstrar que $\boldsymbol{\Sigma}_t$ é também PSD. De fato, seja \mathbf{x} um vetor ($n \times 1$) diferente de $\mathbf{0}$, logo

$$\begin{aligned} \mathbf{x}'(\mathbf{A}_0\mathbf{A}_0')\mathbf{x} &= (\mathbf{x}'\mathbf{A}_0)(\mathbf{x}'\mathbf{A}_0)' \\ &= \|\mathbf{x}'\mathbf{A}_0\|^2 \geq 0, \end{aligned} \tag{A.3}$$

em que $\|\cdot\|^2$ denota a magnitude ao quadrado de um vetor. Por outra parte

$$\begin{aligned} \mathbf{x}'[\mathbf{A}_{il}(\boldsymbol{\varepsilon}_{t-i}\boldsymbol{\varepsilon}_{t-i}')\mathbf{A}_{il}']\mathbf{x} &= (\mathbf{x}'\mathbf{A}_{il}\boldsymbol{\varepsilon}_{t-i})(\mathbf{x}'\mathbf{A}_{il}\boldsymbol{\varepsilon}_{t-i})' \\ &= \|\mathbf{x}'\mathbf{A}_{il}\boldsymbol{\varepsilon}_{t-i}\|^2 \geq 0, \end{aligned} \tag{A.4}$$

para todo $i = 1, \dots, p$ e $l = 1, \dots, r$.

Agora, como $\boldsymbol{\Sigma}_{t-j}$, $j = 1, \dots, q$, é por hipótese PSD, pela decomposição de Cholesky,

¹A suposição inicial já garante que $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \dots, \boldsymbol{\Sigma}_{M-1}$ sejam matrizes PSD, mas a hipótese de indução supõe que $\boldsymbol{\Sigma}_M, \boldsymbol{\Sigma}_{M+1}, \dots, \boldsymbol{\Sigma}_{t-1}$ são também PSD.

$\Sigma_{t-j} = \mathbf{L}_{t-j} \mathbf{L}'_{t-j}$, onde \mathbf{L}_{t-j} é uma matriz triangular inferior. Assim,

$$\begin{aligned} \mathbf{x}'(\mathbf{B}_{jl} \Sigma_{t-j} \mathbf{B}'_{jl}) \mathbf{x} &= \mathbf{x}'(\mathbf{B}_{jl} \mathbf{L}_{t-j} \mathbf{L}'_{t-j} \mathbf{B}'_{jl}) \mathbf{x} \\ &= (\mathbf{x}' \mathbf{B}_{jl} \mathbf{L}_{t-j})(\mathbf{x}' \mathbf{B}_{jl} \mathbf{L}_{t-j})' \\ &= \|\mathbf{x}' \mathbf{B}_{jl} \mathbf{L}_{t-j}\|^2 \geq 0, \end{aligned} \quad (\text{A.5})$$

para todo $j = 1, \dots, q$ e $l = 1, \dots, r$. De (A.3), (A.4) e (A.5), tem-se que

$$\begin{aligned} \mathbf{x}' \mathbf{A}_0 \mathbf{A}'_0 \mathbf{x} + \sum_{l=1}^r \sum_{i=1}^p \mathbf{x}' [\mathbf{A}_{il} (\boldsymbol{\varepsilon}_{t-i} \boldsymbol{\varepsilon}'_{t-i}) \mathbf{A}'_{il}] \mathbf{x} + \sum_{l=1}^r \sum_{j=1}^q \mathbf{x}' (\mathbf{B}_{jl} \Sigma_{t-j} \mathbf{B}'_{jl}) \mathbf{x} &\geq 0, \\ \mathbf{x}' (\mathbf{A}_0 \mathbf{A}'_0 + \sum_{l=1}^r \sum_{i=1}^p \mathbf{A}_{il} (\boldsymbol{\varepsilon}_{t-i} \boldsymbol{\varepsilon}'_{t-i}) \mathbf{A}'_{il} + \sum_{l=1}^r \sum_{j=1}^q \mathbf{B}_{jl} \Sigma_{t-j} \mathbf{B}'_{jl}) \mathbf{x} &\geq 0, \\ \mathbf{x}' \Sigma_t \mathbf{x} &\geq 0. \end{aligned}$$

Portanto, pode-se concluir que Σ_t é PSD. ■

Prova da Proposição 3.2. Suponha que os modelos VEC e BEKK dados por (3.7) e (3.18) respectivamente, são equivalentes, logo para cada sequência $\{\boldsymbol{\varepsilon}_t\}$ estas parametrizações geram a mesma sequência $\{\Sigma_t\}$. Assim, aplicando o operador *vec* em ambos os lados da equação do modelo BEKK (3.18), tem-se que

$$\begin{aligned} \text{vec}(\Sigma_t) &= \text{vec}(\mathbf{A}_0 \mathbf{A}'_0) + \sum_{l=1}^r \sum_{i=1}^p \text{vec}(\mathbf{A}_{il} (\boldsymbol{\varepsilon}_{t-i} \boldsymbol{\varepsilon}'_{t-i}) \mathbf{A}'_{il}) + \sum_{l=1}^r \sum_{j=1}^q \text{vec}(\mathbf{B}_{jl} \Sigma_{t-j} \mathbf{B}'_{jl}) \\ &= (\mathbf{A}_0 \otimes \mathbf{A}_0) \text{vec}(\mathbf{I}_k) + \sum_{i=1}^p \sum_{l=1}^r (\mathbf{A}_{il} \otimes \mathbf{A}_{il}) \text{vec}(\boldsymbol{\varepsilon}_{t-i} \boldsymbol{\varepsilon}'_{t-i}) \\ &\quad + \sum_{j=1}^q \sum_{l=1}^r (\mathbf{B}_{jl} \otimes \mathbf{B}_{jl}) \text{vec}(\Sigma_{t-j}), \end{aligned} \quad (\text{A.6})$$

mas o modelo VEC dado em (3.7) tem a forma

$$\begin{aligned} \text{vec}(\Sigma_t) &= \mathbf{c}_0 + \mathbf{C}_1 \text{vec}(\boldsymbol{\varepsilon}_{t-1} \boldsymbol{\varepsilon}'_{t-1}) + \dots + \mathbf{C}_p \text{vec}(\boldsymbol{\varepsilon}_{t-p} \boldsymbol{\varepsilon}'_{t-p}) \\ &\quad + \mathbf{G}_1 \text{vec}(\Sigma_{t-1}) + \dots + \mathbf{G}_q \text{vec}(\Sigma_{t-q}), \end{aligned} \quad (\text{A.7})$$

e como isto é verdade para toda sequência $\{\boldsymbol{\varepsilon}_t\}$, então igualando as duas equações anteriores

e comparando termo a termo, tem-se que

$$\mathbf{c}_0 = (\mathbf{A}_0 \otimes \mathbf{A}_0) \text{vec}(\mathbf{I}_k), \quad \mathbf{C}_i = \sum_{l=1}^r (\mathbf{A}_{il} \otimes \mathbf{A}_{il}) \quad \text{e} \quad \mathbf{G}_i = \sum_{l=1}^r (\mathbf{B}_{jl} \otimes \mathbf{B}_{jl}).$$

Agora, para mostrar que estes modelos são equivalentes, suponha que as identidades dadas pela proposição se cumprem. Seja $\{\varepsilon_t\}$ uma sequência ruído branco multivariada e considera-se que as matrizes iniciais $\Sigma_1, \dots, \Sigma_M$, são iguais nos dois modelos, sendo $M = \max\{p, q\}$. Denotemos por $\Sigma_{V,t}$ e $\Sigma_{B,t}$ as matrizes de covariância geradas pelos processos VEC e BEKK no instante t , respectivamente e usando o método de indução, isto é, suponha que para $t \geq M$

$$\begin{aligned} \Sigma_{B,t} &= \Sigma_{V,t}, \\ \Sigma_{B,t-1} &= \Sigma_{V,t-1}, \\ &\vdots \\ \Sigma_{B,1} &= \Sigma_{V,1}, \end{aligned}$$

e será demonstrado que

$$\Sigma_{B,t+1} = \Sigma_{V,t+1}.$$

De fato, como $\{\varepsilon_t\}$ é a mesma sequência para ambos modelos, então pela hipótese de indução e devido às identidades da proposição se cumprem, tem-se que

$$(\mathbf{A}_0 \otimes \mathbf{A}_0) \text{vec}(\mathbf{I}_k) = \mathbf{c}_0,$$

$$\sum_{i=1}^p \sum_{l=1}^r (\mathbf{A}_{il} \otimes \mathbf{A}_{il}) \text{vec}(\varepsilon_{(t+1)-i} \varepsilon'_{(t+1)-i}) = \sum_{i=1}^p \mathbf{C}_i \text{vec}(\varepsilon_{(t+1)-i} \varepsilon'_{(t+1)-i}),$$

e

$$\sum_{j=1}^q \sum_{l=1}^r (\mathbf{B}_{jl} \otimes \mathbf{B}_{jl}) \text{vec}(\Sigma_{(t+1)-j}) = \sum_{j=1}^q \mathbf{G}_j \text{vec}(\Sigma_{(t+1)-j})$$

e somando em ambos os lados destas igualdades, tem-se que por definição

$$\Sigma_{B,t+1} = \Sigma_{V,t+1},$$

logo, para todo t as matrizes de covariância nos dois modelos VEC e BEKK são iguais e portanto, eles são equivalentes. ■

Prova da Proposição 3.3. Considera-se primeiro o caso GARCH(1,1), isto é, considerar que $G(L) = \mathbb{G}_1 L$ e $A(L) = \mathbb{A}_1 L$, em que \mathbb{G}_1 e \mathbb{A}_1 são matrizes quadráticas² e depois, numa segunda parte, considera-se o caso geral GARCH(p,q).

Seja E_{t-1} o operador esperança condicional dado o conjunto de informação passada \mathcal{F}_{t-1} , logo

$$\begin{aligned}
 E_{t-1}(\boldsymbol{\eta}_t) &= \mathbf{h}_t \\
 &= \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} L^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 L \boldsymbol{\eta}_t] \\
 &= \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 L^{i-1} L \boldsymbol{\eta}_t] \\
 &= \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 L^i \boldsymbol{\eta}_t] \\
 &= \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_{t-i}],
 \end{aligned}$$

em que $\mathbf{h}_t = \text{vec}(\boldsymbol{\Sigma}_{t-1})$ e $\boldsymbol{\eta}_t = \text{vec}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$. Por outra parte,

$$\begin{aligned}
 E_{t-2}(\boldsymbol{\eta}_t) &= E_{t-2}[E_{t-1}(\boldsymbol{\eta}_t)], \quad (\text{Ver Propriedade 3 da Secção 2.5}) \\
 &= E[E(\boldsymbol{\eta}_t | \mathcal{F}_{t-1}) | \mathcal{F}_{t-2}] \\
 &= E \left[\sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_{t-i}] \middle| \mathcal{F}_{t-2} \right] \\
 &= \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 E_{t-2}(\boldsymbol{\eta}_{t-i})],
 \end{aligned}$$

² Para o caso do modelo VEC dado em (3.7) tem-se que $\mathbb{G}_1 = \mathbf{G}_1$ e $\mathbb{A}_1 = \mathbf{C}_1$, e para o modelo BEKK dado em (3.18) tem-se que $\mathbb{G}_1 = \sum_{l=1}^r (\mathbf{A}_{1l} \otimes \mathbf{A}_{1l})$ e $\mathbb{A}_1 = \sum_{l=1}^r (\mathbf{B}_{1l} \otimes \mathbf{B}_{1l})$

logo,

$$\begin{aligned}
E_{t-2}(\boldsymbol{\eta}_t) &= \mathbf{c}_0 + \mathbb{A}_1 E_{t-2}(\boldsymbol{\eta}_{t-1}) + \sum_{i=2}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 E_{t-2}(\boldsymbol{\eta}_{t-i})] \\
&= \mathbf{c}_0 + \mathbb{A}_1 \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_{t-i-1}] + \sum_{i=2}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_{t-i}] \\
&= \mathbf{c}_0 + \mathbb{A}_1 \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1(L) \boldsymbol{\eta}_{t-i-1}] + \mathbb{G}_1 \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_{t-i-1}] \\
&= \mathbf{c}_0 + [\mathbb{A}_1 + \mathbb{G}_1] \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_{t-i-1}].
\end{aligned}$$

Agora,

$$\begin{aligned}
E_{t-3}(\boldsymbol{\eta}_t) &= E_{t-3}[E_{t-2}(\boldsymbol{\eta}_t)] \\
&= \mathbf{c}_0 + [\mathbb{A}_1 + \mathbb{G}_1] \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 E_{t-3}(\boldsymbol{\eta}_{t-i-1})] \\
&= \mathbf{c}_0 + [\mathbb{A}_1 + \mathbb{G}_1] [\mathbf{c}_0 + \mathbb{A}_1 E_{t-3}(\boldsymbol{\eta}_{t-2})] \\
&\quad + [\mathbb{A}_1 + \mathbb{G}_1] \sum_{i=2}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_{t-i-1}] \\
&= \mathbf{c}_0 + \mathbf{c}_0 [\mathbb{A}_1 + \mathbb{G}_1] + [\mathbb{A}_1 + \mathbb{G}_1] \mathbb{A}_1 \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_{t-i-2}] \\
&\quad + [\mathbb{A}_1 + \mathbb{G}_1] \sum_{i=2}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_{t-i-1}] \\
&= \mathbf{c}_0 + \mathbf{c}_0 [\mathbb{A}_1 + \mathbb{G}_1] + [\mathbb{A}_1 + \mathbb{G}_1] \mathbb{A}_1 \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_{t-i-2}] \\
&\quad + [\mathbb{A}_1 + \mathbb{G}_1] \mathbb{G}_1 \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_{t-i-1}] \\
&= \mathbf{c}_0 + \mathbf{c}_0 [\mathbb{A}_1 + \mathbb{G}_1] + [\mathbb{A}_1 + \mathbb{G}_1]^2 \sum_{i=1}^{\infty} \mathbb{G}_1^{i-1} [\mathbf{c}_0 + \mathbb{A}_1 \boldsymbol{\eta}_{t-i-2}].
\end{aligned}$$

Continuando recursivamente, tem-se que

$$\begin{aligned}
 E_{t-\tau}(\boldsymbol{\eta}_t) &= [\mathbf{I} + (\mathbf{A}_1 + \mathbf{G}_1) + \dots + (\mathbf{A}_1 + \mathbf{G}_1)^{\tau-2}] \mathbf{c}_0 \\
 &\quad + [\mathbf{A}_1 + \mathbf{G}_1]^{\tau-1} \sum_{i=1}^{\infty} \mathbf{G}_1^{i-1} [\mathbf{c}_0 + \mathbf{A}_1 \boldsymbol{\eta}_{t-i-\tau+1}] \\
 &= [\mathbf{I} + (\mathbf{A}_1 + \mathbf{G}_1) + \dots + (\mathbf{A}_1 + \mathbf{G}_1)^{\tau-2}] \mathbf{c}_0 \\
 &\quad + [\mathbf{A}_1 + \mathbf{G}_1]^{\tau-1} E_{t-\tau}(\boldsymbol{\eta}_{t-\tau+1}).
 \end{aligned}$$

Considerando o fato que $E(\boldsymbol{\eta}_t) = E[E_{t-\tau}(\boldsymbol{\eta}_t)]$ para todo τ e t inteiros, tem-se que

$$\begin{aligned}
 E(\boldsymbol{\eta}_t) &= [\mathbf{I} + (\mathbf{A}_1 + \mathbf{G}_1) + \dots + (\mathbf{A}_1 + \mathbf{G}_1)^{\tau-2}] \mathbf{c}_0 \\
 &\quad + [\mathbf{A}_1 + \mathbf{G}_1]^{\tau-1} E(\boldsymbol{\eta}_{t-\tau+1}).
 \end{aligned}$$

Fazendo $s = t - \tau + 1$, tem-se que $\tau = t - s + 1$, logo

$$\begin{aligned}
 E(\boldsymbol{\eta}_t) &= [\mathbf{I} + (\mathbf{A}_1 + \mathbf{G}_1) + \dots + (\mathbf{A}_1 + \mathbf{G}_1)^{t-s+1-2}] \mathbf{c}_0 \\
 &\quad + [\mathbf{A}_1 + \mathbf{G}_1]^{t-s+1-1} E(\boldsymbol{\eta}_s) \\
 &= [\mathbf{I} + (\mathbf{A}_1 + \mathbf{G}_1) + \dots + (\mathbf{A}_1 + \mathbf{G}_1)^{t-s-1}] \mathbf{c}_0 \\
 &\quad + [\mathbf{A}_1 + \mathbf{G}_1]^{t-s} E(\boldsymbol{\eta}_s),
 \end{aligned}$$

e quando $t \rightarrow \infty$, tem-se que $E(\boldsymbol{\eta}_t)$ converge para $[\mathbf{I} - \mathbf{A}_1 - \mathbf{G}_1]^{-1} \mathbf{c}_0$, se e somente se o módulo dos autovalores de $(\mathbf{A}_1 + \mathbf{G}_1)$ é menor que um, já que $\mathbf{Z}^t \rightarrow \mathbf{0}$ e $(\mathbf{I} + \mathbf{Z} + \dots + \mathbf{Z}^t) \rightarrow (\mathbf{I} - \mathbf{Z})^{-1}$ quando $t \rightarrow \infty$ se e somente se o módulo dos autovalores de \mathbf{Z} é menor que um.

Para generalizar o resultado ao caso GARCH(p,q), suponha que o módulo dos autovalores de $A(1) + G(1)$ é menor que um e será demonstrado que o processo tem matriz de covariâncias não condicional constante. Sejam $M = \max\{p, q\}$ e $k^* = k^2$, em que k é o número de séries

do modelo. Além disso, sejam

$$\tilde{\mathbf{h}}_t = \begin{pmatrix} \mathbf{h}_t \\ \mathbf{h}_{t-1} \\ \vdots \\ \mathbf{h}_{t-(M-1)} \end{pmatrix} \quad \tilde{\mathbf{c}}_0 = \begin{pmatrix} \mathbf{c}_0 \\ \mathbf{0}_{k^* \times 1} \\ \vdots \\ \mathbf{0}_{k^* \times 1} \end{pmatrix} \quad \tilde{\mathbf{A}} = \begin{pmatrix} \mathbb{A}_1 & \mathbb{A}_2 & \dots & \mathbb{A}_M \\ \mathbf{0}_{k^* \times k^*} & \mathbf{0}_{k^* \times k^*} & \dots & \mathbf{0}_{k^* \times k^*} \\ \vdots & \vdots & & \vdots \\ \mathbf{0}_{k^* \times k^*} & \mathbf{0}_{k^* \times k^*} & \dots & \mathbf{0}_{k^* \times k^*} \end{pmatrix}$$

$$\tilde{\mathbf{G}} = \begin{pmatrix} \mathbb{G}_1 & \mathbb{G}_2 & \dots & \dots & \mathbb{G}_M \\ \mathbf{I}_{k^*} & \mathbf{0}_{k^* \times k^*} & \dots & \dots & \mathbf{0}_{k^* \times k^*} \\ \mathbf{0}_{k^* \times k^*} & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0}_{k^* \times k^*} & \dots & \mathbf{0}_{k^* \times k^*} & \mathbf{I}_{k^*} & \mathbf{0}_{k^* \times k^*} \end{pmatrix} \quad \tilde{\boldsymbol{\eta}}_t = \begin{pmatrix} \boldsymbol{\eta}_t \\ \vdots \\ \boldsymbol{\eta}_{t-(p-1)} \\ \mathbf{0}_{k^* \times 1} \\ \vdots \\ \mathbf{0}_{k^* \times 1} \end{pmatrix}$$

em que $\mathbf{0}_{k^* \times 1}$ e $\mathbf{0}_{k^* \times k^*}$ denotam o vetor e a matriz de zeros de dimensões $(k^* \times 1)$ e $(k^* \times k^*)$, respectivamente. Também define-se $\mathbb{G}_i = \mathbf{0}$, para $q < i \leq M$, e $\mathbb{A}_i = \mathbf{0}$, para $p < i \leq M$.

O modelo GARCH geral dado em (3.21)

$$\mathbf{h}_t = \mathbf{c}_0 + A(L)\boldsymbol{\eta}_t + G(L)\mathbf{h}_t,$$

pode ser escrito como

$$\tilde{\mathbf{h}}_t = \tilde{\mathbf{c}}_0 + \tilde{\mathbf{A}}\tilde{\boldsymbol{\eta}}_{t-1} + \tilde{\mathbf{G}}\tilde{\mathbf{h}}_{t-1},$$

isto é, um modelo GARCH(1,1) em $\tilde{\mathbf{h}}_t$ e $\tilde{\boldsymbol{\eta}}_t$, mas como já foi provado anteriormente este modelo tem matriz de covariâncias não condicional constante se e somente se os autovalores da matriz

$$\tilde{\mathbf{A}} + \tilde{\mathbf{G}} = \begin{pmatrix} \mathbf{A}_1 + \mathbf{G}_1 & \mathbf{A}_2 + \mathbf{G}_2 & \dots & \dots & \mathbf{A}_M + \mathbf{G}_M \\ & \mathbf{I}_{k^*} & \mathbf{0}_{k^* \times k^*} & \dots & \dots & \mathbf{0}_{k^* \times k^*} \\ \mathbf{0}_{k^* \times k^*} & \ddots & \ddots & & & \vdots \\ & \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0}_{k^* \times k^*} & \dots & \mathbf{0}_{k^* \times k^*} & \mathbf{I}_{k^*} & \mathbf{0}_{k^* \times k^*} \end{pmatrix}, \quad (\text{A.8})$$

têm módulo menor que 1. Comte e Lieberman (2003) (Proposition A.1.) mostraram que se os autovalores da matriz $\mathbf{A}_1 + \dots + \mathbf{A}_p + \mathbf{G}_1 + \dots + \mathbf{G}_q$ têm módulo menor que 1, então os autovalores da matriz dada em (A.8) também têm módulo menor que 1. Portanto, o modelo geral GARCH multivariado dado em (3.21) tem matriz de covariâncias não condicional constante.

Por outra parte, Fuchs (2009) (Theorem 3.3.9. (ii)) mostrou que se o processo $\{\varepsilon_t\}$ é estacionário na covariância então os autovalores da matriz $\mathbf{A}_1 + \dots + \mathbf{A}_p + \mathbf{G}_1 + \dots + \mathbf{G}_q$ têm módulo menor que 1.

Finalmente prova-se que a função de autocovariância não depende do instante t . De fato, para $\gamma \neq 0$ tem-se que

$$\begin{aligned} E(\varepsilon_t \varepsilon'_{t+\gamma}) &= E[E(\varepsilon_t \varepsilon'_{t+\gamma} | \mathcal{F}_{t+\gamma-1})] \\ &= E[\varepsilon_t \underbrace{E(\varepsilon'_{t+\gamma} | \mathcal{F}_{t+\gamma-1})}_{\mathbf{0}'}] \\ &= \mathbf{0}_{k \times K}, \end{aligned}$$

em que $\mathbf{0}'$ denota o vetor de zeros transposto e $\mathbf{0}_{k \times k}$ denota a matriz nula de dimensão $(k \times k)$. Do anterior pode-se concluir que $\{\varepsilon_t\}$ é estacionário na covariância, se e somente se, o todos os autovalores de $A(1) + G(1)$ têm módulo menor que 1. ■

Apêndice B

Comandos no S-PLUS

O pacote FinMetrics do software S-PLUS faz o ajuste GARCH multivariado usando a função `mgarch`. A Tabela B.1 apresenta os comandos para os diferentes ajustes.

Modelo	Comando
EWMA	<code>mgarch(DADOS~1, ~ewmal, trace=F)</code>
DVEC(1,1)	<code>mgarch(DADOS~1, ~dvec.mat(1,1), trace=F)</code>
BEKK(1,1)	<code>mgarch(DADOS~1, ~bekk(1,1), trace=F)</code>
CCC-GARCH(1,1)	<code>mgarch(DADOS~1, ~ccc.garch(1,1), trace=F)</code>
PCOMP-GARCH(1,1)	<code>mgarch(DADOS~1, ~prcomp.garch(1,1), trace=F)</code>

Tabela B.1: *Comandos no software S-PLUS.*

Os comandos usados no S-PLUS para obter os diferentes ajustes do Capítulo 5 junto com os testes de Portmanteau são dados por:

```
module(finmetrics) # Comando usado para chamar o módulo FinMetrics

EWMA.AJUSTE<-mgarch(DADOS~1, ~ewmal, trace=F)
autocorTest(residuals(EWMA.AJUSTE, standardize = T)^2, lag = 12, bycol = F)

PCOMP.AJUSTE<-mgarch(DADOS~1, ~prcomp.garch(1,1), trace=F)
autocorTest(residuals(PCOMP.AJUSTE, standardize = T)^2, lag = 12, bycol = F)

BEKK.AJUSTE<-mgarch(DADOS~1, ~bekk(1,1), trace=F)
autocorTest(residuals(BEKK.AJUSTE, standardize = T)^2, lag = 12, bycol = F)

DVEC.AJUSTE<-mgarch(DADOS~1, ~dvec.mat(1,1), trace=F)
```

```
autocorTest(residuals(DVEC.AJUSTE, standardize = T)^2, lag = 12, bycol = F)
```

```
CCC.AJUSTE<-mgarch(DADOS~1,~ccc.garch(1,1), trace=F)
```

```
autocorTest(residuals(CCC.AJUSTE, standardize = T)^2, lag = 12, bycol = F)
```

Os algoritmos para simular os modelos do Capítulo 4 são os seguintes,

```
#=====
#MODELO EWMA:
#=====
SERIES.EWMA<-function(n,m,lambda){
# n: comprimento da série
# m: número de séries
n<-n
m<-m
l<-lambda
S<-matrix(ncol=n,nrow=m)
SIGMA<-matrix(ncol=n,nrow=(m^2))
sigma<-sigma0
for (i in 1:n){
epsilon<-t(rmnorm(1,rep(0,m),sigma))
S[,i]<-epsilon
SIGMA[,i]<-c(sigma)
E<-epsilon%*%t(epsilon)
sigma<-(1-l)*E+l*sigma}
return(S,SIGMA)}
#=====
#MODELO DVEC(1,1)
#=====
a0.DV11.m<-c(0.2 , -0.23 , 0.26, 0.23, 0, 0.25,
-0.27, 0.29, 0, 0, 0.22, 0.213, 0, 0, 0, 0.26)
A0.DV11.m<-matrix(a0.DV11.m, 4)
a1.DV11.m<-c(0.324 , 0.214 , 0.28, 0.1293, 0,0.25,
0.20, 0.277, 0, 0, 0.4242, 0.243, 0, 0, 0, 0.376)
A1.DV11.m<-matrix(a1.DV11.m, 4)
b1.DV11.m<-c(0.3150 , 0.21 , 0.194, 0.121, 0, 0.3117,
0.132, 0.135, 0, 0, 0.3131, 0.31, 0, 0, 0, 0.3187)
B1.DV11.m<-matrix(b1.DV11.m, 4)

SERIES.DV11.m<-function(n){
#n: Comprimento das séries
n<-n
S.DV11<-matrix(ncol=n,nrow=4)
VC.DV11<-matrix(ncol=n,nrow=10)
sigma.DV11<-sigma0.DV11.m
for (i in 1:n){
```

```

epsilon.DV11<-t(rmvnorm(1,rep(0,4),sigma.DV11))
S.DV11[,i]<-epsilon.DV11
E.DV11<-(epsilon.DV11**t(epsilon.DV11))
sigma1.DV11<-((A0.DV11.m**t(A0.DV11.m))+
(A1.DV11.m**t(A1.DV11.m))*E.DV11)+
(B1.DV11.m**t(B1.DV11.m))*sigma.DV11))
sigma.DV11<-(sigma1.DV11+t(sigma1.DV11))/2
VC.DV11[,i]<-c(sigma.DV11[1,1],sigma.DV11[2,1],
sigma.DV11[3,1],sigma.DV11[4,1],sigma.DV11[2,2],
sigma.DV11[3,2],sigma.DV11[4,2],sigma.DV11[3,3],
sigma.DV11[4,3],sigma.DV11[4,4])
return(S.DV11,VC.DV11)
#=====
#MODELO BEKK(1,1)
#=====
a01.s<-c(1.2,0,0,0.4,0.16,0,0.3,0.9,0.27)
A0.s<-matrix(a01.s,3)
a12.s<-c(0.3,-0.05,0.04,0.05,0.3,0.07,-0.05,-0.07,0.3)
A1.s<-matrix(a12.s,3)
b12.s<-c(0.9,0.05,-0.04,-0.02,0.9,0.02,0.04,-0.05,0.9)
B1.s<-matrix(b12.s,3)

SERIE.BEKK<-function(n){
#n: Comprimento das séries
n<-n
S<-matrix(ncol=n,nrow=3)
VC<-matrix(ncol=n,nrow=6)
sigma<-sigma0.s
for(i in 1:n){
epsilon<-t(rmvnorm(1,rep(0,3),sigma))
S[,i]<-epsilon
E<-(epsilon**t(epsilon))
sigma1<-((A0.s**t(A0.s))+(A1.s**E**t(A1.s))
+(B1.s**sigma**t(B1.s)))
sigma<-(sigma1+t(sigma))/2
VC[,i]<-c(sigma[1,1],sigma[2,1],sigma[3,1],
sigma[2,2],sigma[3,2],sigma[3,3])
return(S,VC)
#=====

```


Referências Bibliográficas

- Alexander(2000)** Carol Alexander. A primer on the orthogonal garch model. *Available in www.ismacentre.reading.ac.uk*. Citado na pág.
- Bauwens et al.(2006)** Luc Bauwens, Sebastian Laurent, e Jeroen V. K. Rombouts. Multivariate garch models: A survey. *Journal Of Applied Econometrics*, 21:79–109. Citado na pág.
- Berndt et al.(1974)** E.K. Berndt, B.H. Hall, R.E. Hall, e J.A. Hausman. Estimation and inference in nonlinear structural models. *Annals of Economic and Social Measurement*, 3:653–665. Citado na pág. 30, 33
- Bollerslev(1990)** T. Bollerslev. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *Review of Economics and Statistics*, 72:498–505. Citado na pág. 42
- Bollerslev(1986)** T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327. Citado na pág. iii, v, 12, 30, 41
- Bollerslev et al.(1988)** T. Bollerslev, R.F. Engle, e J.M. Wooldridge. A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, 96:116–131. Citado na pág. 14, 16, 18
- Bollerslev et al.(1994)** T. Bollerslev, R.F. Engle, e D.B. Nelson. Arch models. *Handbook of Econometrics*, 4. Citado na pág. 21
- Brooks et al.(2003)** C. Brooks, S. Burke, e G. Persaud. Multivariate garch models: Software choice and estimation issues. *Journal of Applied Econometrics*, 18:725–734. Citado na pág. 91
- Christodoulakis e Satchell(2002)** George A. Christodoulakis e Stephen E. Satchell. Correlated arch (corrarch): Modelling the time-varying conditional correlation between financial asset returns. *European Journal of Operational Research, Elsevier*, 139:351–370. Citado na pág. 43
- Comte e Lieberman(2003)** F. Comte e O. Lieberman. Asymptotic theory for multivariate garch process. *Journal of Multivariate Analysis*, 84:61–84. Citado na pág. 30, 31, 147
- Ding(1994)** Z. Ding. *Time Series Analysis of Speculative Returns*. Tese de Doutorado, Department of Economics, University of California, San Diego. Citado na pág. 21

- Ding e Engle(2001)** Zhuanxin Ding e Robert Engle. Large scale conditional covariance matrix modeling, estimation and testing. *Academia Economic Papers*, 29:157–184. Citado na pág. 34
- Doornik e Hendry(1997)** J.A. Doornik e D.F. Hendry. *Modelling Dynamic Systems Using PcFilm 9.0 for Windows*. International Thomson Business Press, London. Citado na pág. 34
- Engle(2000)** Robert Engle. Dynamic conditional correlation- a simple class of multivariate garch models. Citado na pág. 136
- Engle(2002)** Robert F. Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics, American Statistical Association*, 20:339–350. Citado na pág. 43
- Engle(1982)** Robert F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50:987–1007. Citado na pág. iii, v, 12
- Engle e Kroner(1995)** Robert F. Engle e Kenneth F. Kroner. Multivariate simultaneous generalized arch. *Econometric Theory*, 11:122–150. Citado na pág. iii, v, 1, 2, 14, 16, 24, 25, 28
- Engle et al.(1984)** Robert F. Engle, C.W.J. Granger, e D. Kraft. Combining competing forecastas of inflation using a bivariate arch model. *Journal of Econometric Dynamics and Control*, 8:151–165. Citado na pág. 18
- Fuchs(2009)** Florian Fuchs. Probabilistic analysis of multivariate garch models. *Project thesis, TU München*. Citado na pág. 147
- Hafner e Herwartz(2003)** Christian M. Hafner e Helmut Herwartz. Analytical quasi maximum likelihood inference in multivariate volatility models. *Econometrics Institute Report 21. Erasmus University Rotterdam*. Citado na pág. 30, 32
- Hosking(1980)** J.R.M. Hosking. The multivariate portmanteau statistic. *Journal of the American Statistical Association*, 75:602–608. Citado na pág. 33
- Johnson e Wichern(1998)** Richard A. Johnson e Dean W. Wichern. *Applied Multivariate Statistical Analysis*. Pentice Hall, Upper Saddle River, New Jersey, quarta edição. Citado na pág. 112
- Lütkepohl(2006)** Helmut Lütkepohl. *New Introduction to Multiple Time Series Analysis*. Springer-Verlag Berlin Heidelberg 2006, segunda edição. Citado na pág. 34
- Morettin(2002)** Pedro A. Morettin. *Econometria Financeira*. São Paulo: Instituto de Matemática e Ciências Afins, IMCA, and Universidad Católica do Perú, primeira edição. Citado na pág. 1, 11, 12, 14
- Peña(2002)** Daniel Peña. *Análisis de Datos Multivariantes*. Mc. Graw Hill, primeira edição. Citado na pág. 46
- Searle(1982)** S. Searle. *Matrix Algebra Useful for Statistics*. John Wiley. Citado na pág. 5

- Silvennoinen e Teräsvirta(2007)** Annastiina Silvennoinen e Timo Teräsvirta. Multivariate garch models. *Working Paper Series in Economics and Finance*, 669. Citado na pág.
- Tsay(2005)** R.S. Tsay. *Analysis of Financial Time Series*. Wiley, New Jersey. Citado na pág.
- Tse e Tsui(2002)** K.Y. Tse e A.K. Tsui. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business and Economic Statistics, American Statistical Association*, 20:351-362. Citado na pág. 43
- Wadsworth(1969)** Belmont Wadsworth. *Introduction to Matrices with Applications in Statistics*. Citado na pág. 5
- Weide(2002)** Roy Van Der Weide. Go-garch: A multivariate generalized orthogonal garch model. *Journal Of Applied Econometrics*, 17:549-564. Citado na pág. 136
- Williams(1997)** D. Williams. *Probability with Martingales*. Cambridge Mathematical Textbooks. Citado na pág. 9
- Zivot e Wang(2006)** Eric Zivot e Jiahui Wang. *Modelling Financial Time Series with S-PLUS*. Springer-Verlag. Citado na pág. 1, 18, 19, 47