Modelos de memória longa, GARCH e GARCH com memória longa para séries financeiras

Grazielle Yumi Soldá

Dissertação apresentada ao Instituto de Matemática e Estatística da Universidade de São Paulo Para obtenção do título de Mestre em Ciências

Área de Concentração: Estatística Orientadora: Profa. Dra. Chang Chiann

- São Paulo, maio de 2008 -

Modelos de memória longa, GARCH e GARCH com memória longa para séries financeiras

> Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Grazielle Yumi Soldá e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Profa. Dra. Chang Chiann (orientadora) IME-USP.
- Profa. Dra. Clélia Maria de Castro Toloi IME-USP.
- Profa. Dra. Silvia Regina Costa Lopes UFRGS.

Aos meus pais, irmão, namorado, amigos e professores.

Loucura? Sonho? Tudo é loucura ou sonho no começo. Nada do que o homem fez no mundo teve início de outra maneira - mas já tantos sonhos se realizaram que não temos o direito de duvidar de nenhum. (Monteiro Lobato)

Agradecimentos

Agradeço, antes de tudo, aos meus pais (Mauro e Lia) e ao meu irmão (Fabiano) pela educação, pela confiança e pelo amor de sempre.

Ao meu namorado e amigo, André Casado, obrigada por fazer os dias difíceis mais felizes. Obrigada pelo companheirismo e carinho. \heartsuit

Aos amigos da "família de São Paulo" (Karina, Mariana, Lauren, Augusto), do Mestrado (Ivan, Núbia, Joan, Fernando Cristiano, Fernando Mestre, Luiz Fernando, Luis Henrique, Josivon, Pedro, Joelmir, Susana, Kim, Betsabe, Cecília), aos amigos IMEanos, às meninas dos meus queridos times (principalmente, vôlei, basquete e hand do IME), às "irmãs" do ap51 (Mariana Pós e Marcia), aos amigos de Ribeirão Pires (em especial, Juliana e Roberta) e do Unibanco, muito obrigada pelo incentivo, compreensão e pelos bons momentos. Agradeço também ao amigo Josué que acreditou em mim desde o começo.

À minha querida orientadora, Chang Chiann, agradeço muito pela paciência e sabedoria. Às professoras da banca (Clélia Toloi e Silvia Lopes) pelas importantes observações e sugestões. Aos demais professores do IME, obrigada por todos estes anos de aprendizado.

Agradeço à professora Silvia Lopes pelas sugestões a respeito das simulações dos modelos GARCH com memória longa.

Agradeço a todos que, de alguma forma, me ajudaram e me incentivaram em mais esta longa caminhada de minha vida.

Agradecimentos

iv

Resumo

O objetivo deste trabalho é apresentar e comparar diferentes métodos de modelagem da volatilidade (variância condicional) de séries temporais financeiras. O modelo ARFIMA é empregado para capturar o comportamento de memória longa observado na volatilidade de séries financeiras. Por sua vez, o modelo GARCH é utilizado para modelar a volatilidade variando no tempo destas séries. Finalmente, o modelo FIGARCH é utilizado para modelar a dinâmica dos retornos de séries temporais financeiras juntamente com sua volatilidade. Serão apresentados alguns estimadores para os parâmetros dos modelos estudados. Foram realizadas simulações dos três tipos de modelos com o objetivo de comparar o comportamento dos estimadores para diferentes valores dos parâmetros. Por fim, serão apresentadas aplicações em séries reais.

Palavras-chave: memória longa, GARCH, FIGARCH, retornos, volatilidade, variância condicional.

Resumo

Abstract

The goal of this project is to present and compare differents methods of modeling volatility (conditional variance) in financial time series. ARFIMA model is applied to capture long memory behavior of volatility in financial time series. GARCH model is used to model the temporal variation in financial volatility. Finally, FIGARCH model is used to model dynamic of financial time series returns as well as its volatility behavior. We present some estimators for the studied models. Estimators behavior of the three types of models for different parameters is assessed through a simulation study. At last, applications to real data are presented.

Key words: long memory, GARCH, FIGARCH, asset returns, volatility, conditional variance.

Abstract

viii

Sumário

1	Introdução						
2	Pre	eliminares					
	2.1	Séries Temporais Estacionárias e Ergódicas	3				
	2.2	Processos Linares	5				
	2.3	Modelos da família ARIMA	5				
	2.4	Martingal e Seqüência de Diferenças de Martingal	10				
	2.5	Variância de Longo Prazo	11				
	2.6	Séries Temporais Não Estacionárias	13				
	2.7	Retornos	14				
3	Mo	odelos de Memória Longa					
	3.1	Introdução	17				
	3.2	Modelos ARFIMA	17				
	3.3	Testes Estatísticos para Memória Longa	21				
		3.3.1 Estatística R/S	22				
		3.3.2 Teste GPH	23				
	3.4	Estimação do Parâmetro de Memória Longa	24				
		3.4.1 Análise R/S	25				

SUMÁRIO

		3.4.2	Método do Periodograma	26
		3.4.3	Método de Whittle	28
		3.4.4	Método das Ondaletas	28
	3.5	Estim	ação de Modelos ARFIMA	31
		3.5.1	Modelos ARIMA Fracionários	31
		3.5.2	Método de Máxima Verossimilhança	35
	3.6	Simula	ações	36
4	Mo	delos (GARCH	49
	4.1	Introd	lução	49
	4.2	Model	los ARCH	49
		4.2.1	Testes para efeitos ARCH	52
	4.3	O Mo	delo GARCH e suas Propriedades	53
		4.3.1	Representação ARMA do Modelo GARCH	53
		4.3.2	Modelo GARCH e Fatos Estilizados	54
	4.4	Estim	ação de Modelos GARCH	56
		4.4.1	Estimador da Máxima Verossimilhança Condicional	56
		4.4.2	Estimador de Whittle	59
		4.4.3	Estimadores de Mínimos Desvios Absolutos	61
	4.5	Diagn	óstico de Modelos GARCH	64
	4.6	Exten	sões dos Modelos GARCH	66
		4.6.1	Efeitos de Alavancagem Assimétrica e Impactos de Informações Externas	66
		4.6.2	Modelo de Duas Componentes (<i>Two Components Model</i>)	72
		4.6.3	Modelo GARCH-in-the-mean	73
		4.6.4	Variáveis Exógenas na Equação Geral da Média Condicional	73

<i>SUMÁRIO</i> xi				
		4.6.5 Distribuição de Erros Não-Gaussianos	74	
	4.7	Seleção e Comparação de Modelos GARCH	75	
	4.8	Predição de Modelos GARCH	77	
	4.9	Simulações	78	
5	Mo	delos GARCH com Memória Longa	89	
	5.1	Introdução	89	
	5.2	Modelos GARCH com Memória Longa	90	
		5.2.1 Modelo $ARCH(\infty)$	90	
		5.2.2 Modelo LARCH	90	
		5.2.3 Modelo FIGARCH	91	
		5.2.4 Reparametrização do Modelo FIGARCH	95	
		5.2.5 Modelo FIEGARCH	96	
		5.2.6 Modelo LMGARCH	97	
	5.3	Estimação de Modelos GARCH com Memória Longa	98	
	5.4	4 Diagnóstico de Modelos GARCH com Memória Longa		
	5.5	Seleção e Comparação de Modelos GARCH e GARCH com Memória Longa 1	04	
	5.6	Simulações	06	
6	Apl	licação em Séries Reais	13	
	6.1	1 Aplicações		
		6.1.1 Preços das Ações do Unibanco	13	
		6.1.2 Preços das Ações do Itaú	45	
		6.1.3 Preços das Ações do Bradesco	60	
7	Cor	Conclusão e Trabalhos Futuros 175		

SUMÁRIO

Α	A Programas			
	A.1	Método de ondaletas para estimação do parâmetro de memória longa no modelo $\operatorname{ARFIMA}(p, d, q)$	177	
	A.2	Simulação do modelo FIGARCH $(1, d, 1)$	180	
Re	eferêi	ncias Bibliográficas	183	

Capítulo 1

Introdução

A volatilidade do mercado financeiro é uma medida importante na teoria e na prática financeira, como na precificação de ações, alocação de capital e gerenciamento de riscos, por exemplo. Apesar de muitos dos modelos conhecidos na literatura assumirem volatilidades e correlações constantes, é um fato conhecido que ambas variam conforme o tempo.

Modelos lineares do tipo ARIMA não são adequados para capturar a característica de evolução no tempo da variância condicional (volatilidade) de séries financeiras. Um modelo não-linear bastante utilizado para representar essa evolução dinâmica da volatilidade de séries temporais financeiras é o modelo GARCH ("Generalized Autoregressive Conditional Heretoscedasticity") de Engle (1982) e Bollerslev (1986).

Outra característica observada em estudos sobre a modelagem da volatilidade de séries financeiras, como retornos de ações, é o efeito persistente de alterações na variância condicional, ou seja, observa-se que as autocorrelações amostrais dos retornos ao quadrado tendem a decair lentamente em contraste com o decaimento exponencial observado em modelos do tipo GARCH, sugerindo existência de memória longa na variância condicional destas séries.

Dada a existência de uma série temporal que representa a volatilidade (como quadrados dos retornos ou dos retornos absolutos, por exemplo), os modelos ARFIMA (*"Autoregressivos Fracionários Integrados e de Médias móveis"*) podem ser usados com o intuito de modelar o comportamento de memória longa observado na volatilidade de retornos de ações financeiras.

Trabalhos recentes apresentam a alternativa de incorporar memória longa nos modelos GARCH. Os modelos FIGARCH (*"Fractionally Integrated Generalized Autoregressive Conditional Heretos-* *cedasticity*" ou *GARCH fracionário integrado*), introduzidos por Baillie et al. (1996), podem ser utilizados para modelar a dinâmica dos retornos juntamente com sua volatilidade.

Este trabalho tem como objetivo apresentar e comparar as diferentes abordagens da modelagem de volatilidade de séries financeiras, com ênfase em estimação de parâmetros em séries simuladas e em séries reais utilizando o módulo FinMetrics do software S-Plus.

No Capítulo 2 apresentamos alguns conceitos básicos utilizados no decorrer de todo o texto.

O Capítulo 3 aborda o conceito dos modelos de memória longa, apresentando testes para verificar a longa dependência e métodos de estimação do parâmetro de memória longa, além de simulações de modelos ARFIMA(0, d, 0), ARFIMA(1, d, 0) e ARFIMA(0, d, 1) para efeito de comparação de desempenho dos diferentes métodos de estimação apresentados.

Os modelos GARCH e suas extenções são estudados no Capítulo 4. Apresentamos testes para efeitos ARCH, métodos de estimação dos parâmetros dos modelos e algumas ferramentas de diagnóstico dos modelos ajustados. As simulações do modelo GARCH(1,1) foram realizadas com o intuito de comparar o comportamento de um dos estimadores apresentados (estimador de máxima verossimilhança condicional) para diferentes valores de parâmetros.

Os modelos GARCH com memória longa são descritos no Capítulo 5, com ênfase no estudo dos modelos FIGARCH e FIEGARCH (*EGARCH fracionário integrado*). Encontram-se neste capítulo a descrição do estimador de máxima verossimilhança condicional e ferramentas de diagnóstico dos ajustes dos modelos. Os resultados das simulações dos modelos FIGARCH(1, d, 0), FIGARCH(0, d, 1) e FIGARCH(1, d, 1) são apresentados.

Aplicações a dados reais podem ser encontradas no Capítulo 6. As séries utilizadas para as aplicações são séries de cotações de preços de ações de três grandes bancos brasileiros (Unibanco, Itaú e Bradesco) desde 1994 até 2008.

Finalmente, apresentamos algumas conclusões a respeito do trabalho realizado. Sugestões para trabalhos futuros também são apresentadas.

Capítulo 2

Preliminares

Neste capítulo serão apresentados alguns conceitos e definições que servirão de base para a desenvolvimento do restante dos capítulos. Mais detalhes podem ser encontrados em Morettin e Toloi (2004), Fuller (1996) e Hamilton (1994).

2.1 Séries Temporais Estacionárias e Ergódicas

Um processo estocástico $y = \{y(t, \omega), t \in T, \omega \in \Omega\}$ é uma família de variáveis aleatórias (v. a.) definidas num espaço de probabilidade (Ω, Λ, P) e indexadas pelos elementos de um conjunto de parâmetros T.

Na realidade, um processo estocástico é uma função de dois argumentos, $y(t, \omega), t \in T, \omega \in \Omega$. Para cada $t \in T$, a função $y(t, \cdot)$ é mensurável relativamente a Λ . Por outro lado, para cada $\omega \in \Omega$ fixado, obtemos uma função $y(\cdot, \omega)$ de $t : T \to \mathbb{R}$, que é chamada trajetória, realização, função amostral do processo ou série temporal.

Para facilitar a notação, representaremos $y(t, \omega)$ por y_t .

A série temporal y_t é estacionária de segunda ordem, fracamente estacionária, ou ainda, estacionária em sentido amplo se

- $E[y_t] = \mu$, constante, para todo t;
- $cov(y_t, y_{t-j}) = E[(y_t \mu)(y_{t-j} \mu)] = \gamma_j$, só depende da defasagem j, para todo t e qualquer j;
- $E[y_t^2] < \infty$, para qualquer t.

O valor γ_j é chamado de *autocovariância* de y_t de *j*-ésima ordem ou de lag *j*. A função de autocovariância, γ_j , calculada para todo *j*, satisfaz as seguintes propriedades:

- $\gamma_0 = var(y_t) > 0;$
- $\gamma_0 \ge |\gamma_j|;$
- $\gamma_{-j} = \gamma_j;$
- γ_j é não negativa definida.
- A função de autocorrelação (FAC) de y_t é definida por

$$\rho_j = \frac{cov(y_t, y_{t-j})}{\sqrt{var(y_t)var(y_{t-j})}} = \frac{\gamma_j}{\gamma_0}, \ j = 0, \pm 1, \cdots.$$

Intuitivamente, uma série temporal estacionária é definida por sua média, variância e função de autocorrelação.

Qualquer função de uma série temporal estacionária também é uma série temporal estacionária, isto é, se y_t é estacionária, então $z_t = g(y_t)$ também é estacionária, para qualquer função $g(\cdot)$.

A autocovariância amostral de lag j e a autocorrelação amostral de lag j são definidas por

$$\hat{\gamma}_j = \frac{1}{T} \sum_{t=j+1}^T (y_t - \bar{y})(y_{t-j} - \bar{y})$$
(2.1)

$$\hat{\rho}_j = \frac{\hat{\gamma}_j}{\hat{\gamma}_0},\tag{2.2}$$

onde $\bar{y} = \frac{1}{T} \sum_{t=1}^{T} y_t$ é a média amostral.

Dizemos que uma série temporal y_t é *ergódica* se os momentos amostrais convergirem em probabilidade para os momentos populacionais, ou seja, $\bar{y} \xrightarrow{p} \mu$, $\hat{\gamma}_j \xrightarrow{p} \gamma_j$ e $\hat{\rho}_j \xrightarrow{p} \rho_j$.

2.2. PROCESSOS LINARES

2.2 Processos Linares

Seja y_t uma série temporal estacionária. A série y_t pode ser representada por um *processo linear*, ou processo de médias móveis de ordem infinita, da seguinte forma

$$y_t = \mu + \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k},$$

$$\psi_0 = 1, \quad \sum_{k=0}^{\infty} \psi_k^2 < \infty,$$
(2.3)

 $\epsilon_t \sim RB(0,\sigma^2), \mbox{ onde } RB$ é ruído branco.

Para o processo linear (2.3), pode-se mostrar que

$$E[y_t] = \mu,$$

$$\gamma_0 = var(y_t) = \sigma^2 \sum_{k=0}^{\infty} \psi_k^2,$$

$$\gamma_j = cov(y_t, y_{t-1}) = \sigma^2 \sum_{k=0}^{\infty} \psi_k \psi_{k+j},$$

$$e \ \rho_j = \frac{\sum_{k=0}^{\infty} \psi_k \psi_{k+j}}{\sum_{k=0}^{\infty} \psi_k^2}.$$

Portanto, as autocorrelações de qualquer processo estacionário e ergódico y_t são determinadas pelos pesos $\{\psi_j\}$ da representação (2.3).

2.3 Modelos da família ARIMA

Modelos AR(p)

O modelo autor regressivo de ordem $p~(\mbox{AR}(p))$ é dado por

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \dots + \phi_p(y_{t-p} - \mu) + \epsilon_t,$$

ou, utilizando-se a notação de operadores,

$$\phi(L)(y_t - \mu) = \epsilon_t,$$

onde L é o operador de translação para o passado, definido como $L^i y_t = y_{t-i}$, e $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$ e $\mu = E[y_t]$.

Pode-se mostrar que um processo AR(p) é sempre invertível, e é estacionário e ergódico sob a condição de que as raízes da equação característica,

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0, \qquad (2.4)$$

estejam fora do círculo unitário. Uma condição necessária para estacionariedade muito útil na prática é que $|\phi_1 + \cdots + \phi_p| < 1$.

A função de autocovariância de ordem j de um processo AR(p) estacionário é dada por

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p},$$

 $\operatorname{com} \gamma_0 = \operatorname{var}(y_t).$

A função de autocorrelação de ordem j do processo é dada por

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_p \rho_{j-p}$$

Função de Autocorrelação Parcial

A função de autocorrelação parcial (FACP) é um importante instrumento para ajudar na identificação de modelos AR(p). A obtenção da FACP baseia-se na estimação da seqüência de modelos AR

$$z_{t} = \phi_{11}z_{t-1} + \epsilon_{1t}$$

$$z_{t} = \phi_{21}z_{t-1} + \phi_{22}z_{t-2} + \epsilon_{2t}$$

$$\vdots$$

$$z_t = \phi_{p1} z_{t-1} + \phi_{p2} z_{t-2} + \dots + \phi_{pp} z_{t-p} + \epsilon_{pt}$$

onde $z_t = y_t - \mu$. Os últimos coeficientes em cada AR(p) $(\phi_{jj}, \text{ para } j = 1, \dots, p)$ são chamados coeficientes de autocorrelação parcial.

Para um processo AR(p), o *p*-ésimo coeficiente de autocorrelação parcial é diferente de zero, e o restante é igual a zero para j > p.

Modelos MA(q)

O modelo de médias móveis de ordem q (MA(q)) tem a forma

$$y_t - \mu = \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q},$$

onde $\epsilon_t \sim RB(0, \sigma^2)$.

Em notação de operadores temos

$$(y_t - \mu) = \theta(L)\epsilon_t,$$

onde $\theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$.

O modelo MA(q) é estacionário e ergódico desde que $\theta_1, \dots, \theta_q$ sejam finitos, e é invertível se todas as raízes da equação característica,

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q = 0, \tag{2.5}$$

estiverem fora do círculo unitário.

A função de autocovariância de um processo MA(q) é dada por

$$\begin{aligned} \gamma_0 &= \sigma^2 (1 + \theta_1^2 + \dots + \theta_q^2); \\ \gamma_j &= \begin{cases} \sigma^2 (\theta_j + \theta_{j+1} \theta_1 + \theta_{j+2} \theta_2 + \dots + \theta_q \theta_{q-j}), \text{ para } j = 1, 2, \dots, q; \\ 0, \text{ para } j > q. \end{cases} \end{aligned}$$

Observa-se que a FAC de um processo MA(q) é diferente de zero para o lag q e igual a zero para

j > q. Dessa forma, a FAC pode ser usada como um importante instrumento na identificação deste tipo de processo.

E a função de autocorrelação do processo pode ser escrita como

$$\rho_j = \frac{\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \dots + \theta_q\theta_{q-j}}{1 + \theta_1^2 + \dots + \theta_q^2}.$$

A função de autocorrelação parcial de um processo MA(q) invertível apresenta decaimento exponencial.

Modelos $\operatorname{ARMA}(p,q)$

O modelo autorregressivo e de médias móveis (ARMA(p,q)) tem a forma

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \dots + \phi_p(y_{t-p} - \mu)$$

$$+ \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q},$$

$$\epsilon_t \sim RB(0, \sigma^2).$$
(2.6)

Pode ser escrito na forma da regressão

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q},$$

ou representado utilizando-se o operador L

$$\phi(L)(y_t - \mu) = \theta(L)\epsilon_t,$$

e é estacionário e ergódico se as raízes da equação característica $\phi(z) = 0$ estiverem fora do círculo unitário, e é invertível se as raízes da equação característica $\theta(z) = 0$ estiverem fora do círculo unitário. Assume-se a condição de que os polinômios $\phi(z) = 0$ e $\theta(z) = 0$ não tenham fatores em comum ou que se cancelam.

2.3. MODELOS DA FAMÍLIA ARIMA

A média de um processo estacionário e ergódico ARMA(p,q) é igual a

$$\mu = \frac{c}{1 - \phi_1 - \dots - \phi_p},$$

e suas autocovariâncias e autocorrelações safisfazem as seguintes relações recursivas

$$\gamma_{j} = \phi_{1}\gamma_{j-1} + \phi_{2}\gamma_{j-2} + \dots + \phi_{p}\gamma_{j-p}, \quad j > q,$$

$$\rho_{j} = \phi_{1}\rho_{j-1} + \phi_{2}\rho_{j-2} + \dots + \phi_{p}\rho_{j-p}, \quad j > q.$$
(2.7)

A forma geral da FAC de um processo ARMA(p,q) é complicada. Para detalhes, ver Hamilton (1994). Em geral, para o processo ARMA(p,q), a FAC tem o mesmo comportamento da FAC de um processo AR(p) para p > q, e a FACP tem comportamento parecido com a FACP de um processo MA(q) para q > p. Dessa forma, ambas FAC e FACP podem apresentar decaimento exponencial.

Na prática, modelos ARMA(p,q) de ordens altas são difíceis de serem estimados e, por este motivo, são raramente usados para analisar dados financeiros. Segundo Zivot e Wang (2003), modelos ARMA(p,q) de ordens baixas, com $p \in q$ menores do que três, são geralmente suficientes na análise de séries temporais financeiras.

Modelos $\operatorname{ARIMA}(p, d, q)$

A especificação do modelo ARMA(p,q) (2.6) assume que y_t é estacionário e ergódico. Se y_t possuir alguma tendência, como preços de ações e outras séries financeiras, então y_t deve sofrer alguma transformação com o objetivo de eliminar esta tendência.

O modelo ARIMA (*autorregressivo integrado e de médias móveis*) foi proposto para modelagem de processos nos quais pelo menos uma raiz da equação característica $\phi(z) = 0$ se encontra sobre o círculo unitário.

Se houver uma tendência linear em y_t , então a primeira diferença $\Delta y_t = y_t - y_{t-1}$ elimina tal tendência, onde $\Delta = 1 - L$ é o operador da diferença. Se houver tendência quadrática em y_t então haverá tendência linear em Δy_t , mas a segunda diferença $\Delta^2 y_t = (1 - 2L + L^2)y_t = y_t - 2y_{t-1} + y_{t-2}$ não apresentará tendência. O modelo ARIMA(p, d, q) é resultante de d diferenças para eliminar as tendências.

O modelo ARIMA(p, d, q) pode ser representado por

$$\phi(L)\Delta^d(y_t - \mu) = \theta(L)\epsilon_t,$$

onde $\Delta^d = (1-L)^d$.

2.4 Martingal e Sequência de Diferenças de Martingal

Seja $\{y_t\}$ uma seqüência de variáveis aleatórias e seja $I_t = \{y_t, y_{t-1}, \dots\}$ um conjunto de informações baseadas no passado histórico de $\{y_t\}$. A seqüência $\{y_t, I_t\}$ é chamada de um *martingal* se

- $I_{t-1} \subset I_t;$
- $E[|y_t|] < \infty;$
- $E[y_t|I_{t-1}] = y_{t-1}$ (propriedade martingal).

Um exemplo comum de um martingal é o modelo de passeio aleatório

$$y_t = y_{t-1} + \epsilon_t, \ \epsilon_t \sim RB(0, \sigma^2),$$

onde y_0 é um valor inicial fixado. Se $I_t = \{y_t, \dots, y_0\}$ então $E[y_t|I_{t-1}] = y_{t-1}$, dado que $E[\epsilon_t|I_{t-1}] = 0$.

Seja $\{\epsilon_t\}$ uma seqüência de variáveis aleatórias com um conjunto de informações associadas I_t . A seqüência $\{\epsilon_t, I_t\}$ é chamada de uma seqüência de diferenças de martingal (martingale difference sequence (MDS)) se

- $I_{t-1} \subset I_t;$
- $E[\epsilon_t|I_{t-1}] = 0$ (propriedade MDS).

Se $\{y_t, I_t\}$ é um martingal, um MDS $\{\epsilon_t, I_t\}$ pode ser construído definindo-se

$$\epsilon_t = y_t - E[y_t | I_{t-1}].$$

2.5. VARIÂNCIA DE LONGO PRAZO

Por construção, um MDS é um processo não correlacionado. Para k > 0 temos

$$E[\epsilon_t \epsilon_{t-k}] = E[E[\epsilon_t \epsilon_{t-k} | I_{t-1}]]$$

$$= E[\epsilon_{t-k} E[\epsilon_t | I_{t-1}]]$$

$$= 0.$$
(2.8)

Embora um processo MDS seja não correlacionado, ele não é necessariamente independente, isto é, pode existir dependência em momentos de maior ordem de ϵ_t .

Processos MDS possuem muitos resultados úteis de convergência (lei dos grandes números, teorema central do limite, etc.). White (1984), Hamilton (1994) e Hayashi (2000) apresentam muitos destes resultados para a análise de séries temporais financeiras.

2.5 Variância de Longo Prazo

Seja y_t uma série temporal estacionária e ergódica. Pelo Teorema Central do Limite

$$\sqrt{T}(\bar{y}-\mu) \xrightarrow{d} N(0, \sum_{j=-\infty}^{\infty} \gamma_j),$$

ou

$$\bar{y} \stackrel{A}{\sim} N\left(\mu, \frac{1}{T} \sum_{j=-\infty}^{\infty} \gamma_j\right),$$

onde T é o tamanho da amostra, \bar{y} é a média amostral de y_t e γ_j são as autocovariâncias.

A variância de longo prazo (vlp) de y_t é dada pela multiplicação de T pela variância assintótica (avar) da média amostral:

$$vlp(y_t) = T \cdot avar(\bar{y}) = \sum_{j=-\infty}^{\infty} \gamma_j.$$

Dado que $\gamma_{-j} = \gamma_j$, a variância de longo prazo de y_t pode ser escrita como

$$vlp(y_t) = \gamma_0 + 2\sum_{j=1}^{\infty} \gamma_j.$$

Estimando a Variância de Longo Prazo

Se y_t é um processo linear temos que

$$\sum_{j=-\infty}^{\infty} \gamma_j = \sigma^2 \left(\sum_{j=0}^{\infty} \psi_j \right)^2 = \sigma^2 \psi(1)^2,$$

e, portanto

$$vlp(y_t) = \sigma^2 \psi(1)^2.$$
 (2.9)

Se $y_t \sim \text{ARMA}(p,q)$ então

$$\psi(1) = \frac{1+\theta_1+\dots+\theta_q}{1-\phi_1-\dots-\phi_p} = \frac{\theta(1)}{\phi(1)}$$

e, portanto

$$vlp(y_t) = \frac{\sigma^2 \theta(1)^2}{\phi(1)^2}.$$
 (2.10)

Uma estimativa consistente para $vlp(y_t)$ pode ser calculada pela substituição de parâmetros estimados do modelo ARMA(p,q) apropriado em (2.10). De forma alternativa, pode-se aproximar o modelo ARMA(p,q) por um processo AR (p^*)

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_{p^*} y_{t-p^*} + \epsilon_t,$$

onde p^* é escolhido de tal forma que ϵ_t seja não correlacionado. Assim, a estimativa da variância de longo prazo autorregressiva fica dada por

$$vlp_{AR}(y_t) = \frac{\sigma^2}{\phi^*(1)^2}.$$
 (2.11)

Um estimador popularizado por Newey e West $(1987)^1$ é um estimador ponderado pelas autocovariâncias

$$\widehat{vlp}_{NW}(y_t) = \hat{\gamma}_0 + 2\sum_{j=1}^{M_T} w_{j,T} \cdot \hat{\gamma}_j,$$

 $^{^{1}}$ A estimativa de Newey-West para a variância de longo prazo será utilizada para o cálculo da Estimativa R/S. Ver subseção 3.3.1.

onde $w_{j,T}$ são pesos que somam 1 e M_T é um parâmetro de truncamento do lag que satisfaz $M_T = O(T^{1/3})$.

Para um processo $MA(q), \gamma_j = 0$ para j > q. New
ey e West (1987) sugerem o uso de pesos retangulares

$$w_{j,T} = \begin{cases} 1, \text{ para } j \le M_T = q; \\ 0, \text{ caso contrário.} \end{cases}$$
(2.12)

Para processos lineares em geral, Newey e West (1987) sugerem o uso dos pesos Bartlett, $w_{j,T} = 1 - \frac{j}{M_T + 1}$, com $M_T = [4(T/100)^{2/9}]$.²

2.6 Séries Temporais Não Estacionárias

Uma vez que um processo estacionário possui momentos invariantes, um processo *não estacionário* deve ter alguma dependência nos momentos. As formas mais comuns de não estacionariedade são causadas pela dependência temporal na média e na variância.

Processos Integrados

Chamamos y_t de um processo integrado de ordem 1 ($y_t \sim I(1)$) se possui a forma

$$y_t = y_{t-1} + u_t, \tag{2.13}$$

onde u_t é uma série temporal estacionária. Desta forma, tem-se que a primeira diferença de y_t é estacionária

$$\Delta y_t = u_t.$$

Devido a esta propriedade, o processo I(1) é chamado também de processo de diferença estacionária.

Uma vez que o processo estacionário u_t não precisa sofrer diferenças, é chamado de processo integrado de ordem zero $(u_t \sim I(0))$.

O processo $y_t \in I(d)$ se $\Delta^d y_t \sim I(0)$. Em séries temporais financeiras raramente são modelados processos I(d) com d > 2.

 $^{{}^{2}[\}cdot]$ denota a parte inteira de um número real.

2.7 Retornos

Para avaliar o risco de uma carteira de ativos no mercado financeiro é comum medir as variações de preços destes ativos. Na prática, é preferível trabalhar com retornos do que com preços pois os retornos são livres de escala e têm propriedades estatísticas interessantes (como estacionariedade e ergodicidade). Os modelos apresentados neste trabalho têm como objetivo modelar a volatilidade dos retornos de séries financeiras, como retornos de preços de ações, por exemplo.

Seja P_t o preço de um ativo no instante t. A variação de preços entre os instantes t-1 e t é dada por $\Delta P_t = P_t - P_{t-1}$ e a variação relativa de preços ou *retorno simples* deste ativo é definido por

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}}.$$
(2.14)

De (2.14) temos que $R_t = \frac{P_t}{P_{t-1}} - 1$. Chamamos $1 + R_t = \frac{P_t}{P_{t-1}}$ de retorno bruto simples ou taxa de retorno.

Definimos o retorno composto continuamente, ou log-retorno ou simplesmente retorno, como

$$r_t = \ln \frac{P_t}{P_{t-1}} = \ln(1+R_t) = \ln(P_t) - \ln(P_{t-1}) = p_t - p_{t-1},$$
(2.15)

onde $p_t = \ln(P_t)$.

Podemos definir também retornos multiperíodos. Dado um período k, o retorno simples entre os instantes $t - k \in t$ é dado por

$$R_t(k) = \frac{P_t - P_{t-k}}{P_{t-k}}.$$
(2.16)

Em termos de retornos de um mesmo período podemos escrever

$$1 + R_t(k) = (1 + R_t)(1 + R_{t-1})\cdots(1 + R_{t-k+1})$$

= $\frac{P_t}{P_{t-1}}\frac{P_{t-1}}{P_{t-2}}\cdots\frac{P_{t-k+1}}{P_{t-k}} = \frac{P_t}{P_{t-k}},$

portanto temos

$$R_t(k) = \frac{P_t}{P_{t-k}} - 1.$$
(2.17)

2.7. RETORNOS

O log-retorno de período k fica dado por

$$r_t(k) = \ln\left(\frac{P_t}{P_{t-k}}\right) = \ln(1 + R_t(k)) = p_t - p_{t-k}.$$
(2.18)

Se houver pagamento de dividendos D_t no período, retorno simples (R_t) e o log-retorno (r_t) são dados por

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1, \tag{2.19}$$

e

$$r_t = \ln(1 + R_t) = \ln(P_t + D_t) - \ln(P_{t-1}).$$
(2.20)

Fatos Estilizados sobre os Retornos

Séries financeiras apresentam algumas características que são comuns a outras séries temporais, como tendências, sazonalidade, pontos influentes (atípicos), heteroscedasticidade condicional e nãolinearidade. Os retornos financeiros, por outro lado, apresentam outras características peculiares que outras séries não apresentam. Retornos raramente apresentam tendências ou sazonalidades, com excessão eventualmente de retornos intra-diários, e séries de taxas de câmbio ou taxa de juros que podem apresentar tendências que variam com o tempo. Podemos resumir os principais *fatos estilizados* relativos a retornos financeiros:

- 1. retornos são em geral não autocorrelacionados;
- 2. os quadrados dos retornos são autocorrelacionados;
- 3. séries de retornos apresentam agrupamentos de volatilidades ao longo do tempo;
- 4. a distribuição (não-condicional) dos retornos apresenta caudas mais pesadas do que uma distribuição normal;
- 5. algumas séries de retornos são não-lineares (respondem de maneira diferente a choques grandes ou pequenos, ou a choques negativos ou positivos).

Capítulo 3

Modelos de Memória Longa

3.1 Introdução

Os modelos de memória longa em séries temporais ou espaciais já eram empiricamente reconhecidos nas áreas de hidrologia e climatologia na década de 50, embora sua aplicação em séries econômicas e financeiras tenha ocorrido apenas por volta de 1980.

Um processo de memória longa é caracterizado pela dependência significante entre observações distantes (em tempo ou espaço). Nesse caso, existe persistência nas autocorrelações observadas, que decaem lentamente para zero.

A presença desse fenômeno foi primeiramente observada por Hurst (1951, 1957), Mandelbrot e Wallis (1968), Mandelbrot (1972), e McLeod e Hipel (1978), dentre outros.

Neste capítulo serão apresentados os conceitos de memória longa, assim como alguns testes para verificar a longa dependência. Detalharemos também alguns estimadores para estimar o parâmetro de memória longa, comparando suas eficiências via séries simuladas.

3.2 Modelos ARFIMA

Para ilustrar as características de memória longa em séries financeiras, consideremos a série de retornos de preços diários das ações da Petrobras, de 03 de janeiro de 1995 à 27 de dezembro de 2000, apresentada na Figura 3.1 (b) juntamente com a série original de preços, Figura 3.1 (a).¹

¹Alguns dos gráficos deste trabalho foram construídos a partir dos softwares S-Plus e Excel, que têm como padrão o "ponto" como símbolo para separar as casas decimais dos números.



Figura 3.1: (a) série original e (b) série de retornos de preços diários das ações da Petrobras

Uma vez que séries de retornos diários possuem média muito próxima de zero, os **retornos absolutos**, Figura 3.2 (b), podem ser usados como medida de volatilidade. Para a série de retornos dos preços das ações da Petrobras a média é igual a 0,0010938.

Pode-se perceber que a autocorrelação dos valores da série de retornos absolutos, apresentada na Figura 3.2 (b), possui forte persistência e permanece muito significante até os lags altos.

Ajustando um processo estacionário tradicional ARMA para a série de retornos absolutos, pode-se obter um número excessivo de parâmetros² e, além disso, tem-se que a função de autocorrelação teórica obtida descreve bem a autocorrelação da série real apenas nos lags iniciais, conforme apresentado na Figura 3.2.

Enquanto um processo ARMA tem memória curta, ou seja, sua função de autocorrelação ρ_j decresce exponencialmente para zero, de modo que

$$|\rho_k| \le Cr^{-k}, \, \mathbf{k} = 1, 2, \dots,$$
(3.1)

²Para a série de retornos absolutos de preços diários das ações da Petrobras, obtemos um AR de ordem 17.



Figura 3.2: (a) série de retornos absolutos de preços diários das ações da Petrobras e (b) FAC amostral

onde C > 0 e 0 < r < 0, 5, um processo estacionário de memória longa é definido conforme segue:

Definição 3.2.1. Seja y_t um processo estacionário. Se existir um número real $\alpha \in (0, 1)$ e uma constante C > 0 de modo que

$$\rho_k \to C k^{-\alpha} \quad quando \quad k \to \infty,$$
(3.2)

então y_t é chamado de processo estacionário de memória longa.

A função de autocorrelação de um processo de memória longa decai hiperbolicamente. Esse decaimento é muito lento e, desse modo, tem-se que as autocorrelações não são somáveis:

$$\sum_{k=-\infty}^{\infty} \rho_k = \infty. \tag{3.3}$$

Do ponto de vista estatístico, para um processo estacionário, a função de autocorrelação contém a mesma informação que sua densidade espectral. Em particular, a densidade espectral é definida como:

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho_k e^{ik\omega}, \qquad (3.4)$$

onde ω é a freqüência de Fourier. De (3.2) pode-se mostrar que a função espectral de um processo estacionário com memória longa é dada por

$$f(\omega) \to C_f \omega^{\alpha - 1}$$
 quando $\omega \to 0,$ (3.5)

sendo C_f uma constante positiva. Assim, para um processo de memória longa, sua função de densidade espectral tende para infinito na freqüência zero. Fazendo

$$H = 1 - \alpha/2 \in (0,5, 1),$$

temos o *coeficiente de Hurst* (Hurst, 1951), que é usado para mensurar a longa memória em y_t . Quanto maior for o valor de H, a função de autocorrelação do processo decai mais lentamente, ou seja, o processo possui memória mais longa.

Baseados na propriedade da função de autocorrelação (3.2) e no espectro do processo (3.5), Granger e Joyeux (1980) e Hosking (1981) mostraram, independentemente, que um processo de memória longa y_t pode ser modelado parametricamente pela extensão de um processo integrado para um *Processo Integrado Fracionário*. Consideremos a integração fracionária da série y_t conforme segue:

$$(1-L)^d (y_t - \mu) = u_t, (3.6)$$

onde L indica o operador de translação para o passado, d é o parâmetro de diferença fracionária, μ é a esperança de y_t e u_t é um ruído estacionário com média zero.

Na prática, quando uma série temporal tem longa dependência ou tem indícios de não ser estacionária, toma-se a primeira diferença (d = 1) da série para atingir estacionariedade. Entretanto, para algumas séries temporais econômicas e financeiras com memória longa, tomar uma diferença inteira pode ser demais, levando à "super-diferenciação" da série.

Para respeitar as características de uma série de memória longa e evitar a diferença inteira de y_t , toma-se d fracionário.

Definição 3.2.2. O operador de diferença fracionária, para qualquer número real d>-1, é dado por

$$(1-L)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-1)^k L^k,$$
(3.7)

com coeficientes binomiais:

$$\binom{d}{k} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}.$$

Nota-se que o filtro de diferença fracionária pode ser equivalentemente tratado como um filtro autorregressivo de ordem infinita ³. Em relação ao valor de d podemos mostrar que:

- se |d| > 1/2, então y_t é não estacionária;
- se 0 < d < 1/2, então y_t é estationária e possui memória longa;
- se -1/2 < d < 0, então y_t é estacionária e tem memória curta, e é identificada como *anti*persistente.

Quando uma série integrada fracionária y_t possui memória longa, pode-se mostrar também que

$$d = H - 1/2,$$

e, portanto, temos que tanto d quanto H podem ser usados como medidas de memória longa. Hosking (1981) mostrou que as propriedades (3.2) e (3.5) são satisfeitas quado 0 < d < 1/2.

Para ilustrar a FAC teórica de um processo integrado fracionário, apresentamos na Figura 3.3 os gráficos da FAC do processo para d = 0, 3 e d = -0, 3, com ruído u_t normal padrão.

3.3 Testes Estatísticos para Memória Longa

Existem vários testes para determinar a existência de memória longa em uma série temporal. Nessa seção vamos introduzir a estatística R/S e o teste GPH.

 $^{^{3}}$ No módulo S+Fin
Metrics do S-Plus a função FARIMA. d2ar pode ser usada para calcular a representação autor
regressiva do filtro de diferença fracionária.



Figura 3.3: Autocorrelação do Processo Integrado Fracionário

3.3.1 Estatística R/S

O teste mais conhecido para testar a existência de memória longa é a estatística R/S, também conhecida como *rescaled range* ou *range over standard deviation*, que foi originalmente proposto por Hurst (1951) e mais tarde refinado por Mandelbrot e seus co-autores.

Definição 3.3.1. Considere a série temporal y_t , para t=1,...,T. A estatística R/S é definida por:

$$Q_T = \frac{1}{s_T} \left[max_{1 \le k \le T} \sum_{j=1}^k (y_j - \bar{y}) - min_{1 \le k \le T} \sum_{j=1}^k (y_j - \bar{y}) \right],$$
(3.8)

onde $\bar{y} = 1/T \sum_{i=1}^{T} y_i$ é a média amostral e $s_T = \sqrt{1/T \sum_{i=1}^{T} (y_i - \bar{y})^2}$ é o desvio padrão amostral.

Pode-se mostrar que se y_t 's são variáveis aleatórias normais i.i.d , então

$$\frac{1}{\sqrt{T}}Q_t \Rightarrow V$$

onde \Rightarrow denota convergência fraca e V é uma variável aleatória que está no domínio de atração de
3.3. TESTES ESTATÍSTICOS PARA MEMÓRIA LONGA

uma ponte browniana no intervalo unitário.

Lo (1991) fornece os quantis de V e aponta que a estatística R/S não é robusta para dependência de memória curta. Em particular, se y_t é autocorrelacionado (possui memória curta) então a distribuição limite de $\frac{Q_T}{\sqrt{T}}$ é V com escala dada pela raiz quadrada da variância de longo prazo de y_t . Para permitir a presença da dependência de memória curta em y_t , Lo (1991) propõe a alteração da estatística R/S conforme segue:

$$\tilde{Q}_T = \frac{1}{\hat{\sigma}_T(q)} \left[max_{1 \le k \le T} \sum_{j=1}^k (y_j - \bar{y}) - min_{1 \le k \le T} \sum_{j=1}^k (y_j - \bar{y}) \right],$$
(3.9)

onde o desvio padrão amostral é substituído pela raiz quadrada da estimativa de Newey-West para a variância de longo prazo com largura de faixa q. Lo (1991) mostra que se existe memória curta mas não existe memória longa (hipótese nula H_0), \tilde{Q}_T também converge para V.

Aplicando este teste⁴ na série de retornos dos preços de ações da Petrobras, temos que a estatística do teste é igual a 2,0832 e, portanto, rejeitamos a hipótese nula (não existência de memória longa) a um nível de significância de 5%.

3.3.2 Teste GPH

Geweke e Porter-Hudak (1983) propuseram um teste semi-paramétrico para testar a presença de memória longa baseado na representação de uma série temporal como um processo integrado fracionário, conforme descrito em (3.6). Para um processo integrado fracionário y_t , temos que a função densidade espectral é dada por:

$$f(\omega) = \left[4sen^2\left(\frac{\omega}{2}\right)\right]^{-d} f_u(\omega) \tag{3.10}$$

onde ω é a frequência de Fourier e $f_u(\omega)$ é a função densidade espectral correspondente a u_t . O parâmetro de diferença fracionária d pode ser estimado pela seguinte regressão:

$$lnf(\omega_j) = \beta - dln \left[4sen^2 \left(\frac{\omega_j}{2} \right) \right] + e_j, \qquad (3.11)$$

⁴Função rosTest do módulo FinMetrics do sofware S-Plus.

para $j = 1, 2, ..., n_f(T)$. Geweke e Porter-Hudak (1983) mostraram que usando o *periodograma* como estimativa de $f(\omega_j)$, a estimativa \hat{d} de mínimos quadrados, obtida quando aplicada a regressão descrita acima, possui distribuição normal em grandes amostras se $n_f(T) = T^{\alpha} \mod 0 < \alpha < 1$:

$$\hat{d} \xrightarrow{D} N\left(d, \frac{\pi^2}{6\sum_{j=1}^{n_f(T)} (U_j - \bar{U})^2}\right)$$

onde

$$U_j = ln \left[4 sen^2 \left(\frac{\omega_j}{2} \right) \right]$$

e \overline{U} é a média amostral de U_j , $j=1,...,n_f(T)$. Sob a hipótese nula da não existência de memória longa (d=0), a estatística t

$$t_{d=0} = \hat{d} \cdot \left(\frac{\pi^2}{6\sum_{j=1}^{n_f(T)} (U_j - \bar{U})^2}\right)^{-1/2}$$
(3.12)

tem distribuição assintótica normal padrão.

Para a série de retornos dos preços de ações da Petrobras, este teste⁵ apresenta o valor da estatística de 4,5882 e, portanto, rejeitamos a hipótese nula (d = 0) a um nível de significância de 1%. Porém, o teste apresenta uma estimativa do valor de *d* como sendo 0,5565, o que indica não estacionariedade.

3.4 Estimação do Parâmetro de Memória Longa

Na seção anterior apresentamos dois testes para testar a presença de memória longa. Mas também é de interesse obter uma estimativa do parâmetro de memória longa, d ou H. Por meio do teste GPH obtemos automaticamente uma estimativa de d. Nessa seção vamos mostrar que a estatística R/S também pode ser usada para se obter uma estimativa do coeficiente de Hurst H. Serão introduzidos também dois métodos baseados no periodograma para se obter a estimativa do parâmetro de memória longa, o método do periodograma e o método de Whittle, e o método de Ondaletas. O método de Whittle foi implementado para estimar o parâmetro d somente em modelos ARFIMA(0, d, 0). Além disso, o parâmetro de diferença fracionária d pode ser estimado em um modelo ARFIMA(p, d, q)simultaneamente com os outros parâmetros ($\phi \in \theta$) pelo método de máxima verossimilhança, que

⁵Função gphTest do módulo FinMetrics do sofware S-Plus.

será introduzido na próxima seção. O desempenho desses e de outros estimadores para o parâmetro de memória longa foram comparadas por Taqqu, Teverovsky e Willinger (1995) e Taqqu e Teverovsky (1998).

3.4.1 Análise R/S

Quando um processo estacionário y_t possui memória longa, a estatística R/S converge para uma variável aleatória com taxa de convergência T^H , onde H é o coeficiente de Hurst, conforme mostrou Mandelbrot (1975). Baseando-se nesse resultado, temos que o gráfico (na escala log-log) da estatística R/S contra o tamanho da amostra utilizada deve oscilar em torno de uma linha reta com coeficiente de inclinação 1/2 para séries temporais com memória curta. Por outro lado, para séries com memória longa, o gráfico log-log deve oscilar em torno de uma linha reta com coeficiente de inclinação igual a H > 1/2, levando-se em conta que o tamanho da amostra é grande o suficiente.

Para estimar o parâmetro de memória longa H utilizando-se esse método, é preciso, primeiramente, calcular a estatística R/S usando k_1 observações consecutivas, onde k_1 é grande o suficiente. Depois, aumenta-se o número de observações usando-se um fator f. Assim, a estatística R/S será calculada com $k_i = fk_{i-1}$ observações consecutivas para i = 2, ..., s. Para se obter a estatística R/S com k_i observações consecutivas, deve-se dividir a amostra em $[T/k_i]$ blocos e obter $[T/k_i]$ valores diferentes. Quanto maior o valor k_i menor vai ser a quantidade $[T/k_i]$. A reta ajustada com todos esses valores da estatística R/S versus os valores de k_i , i = 1, ..., s, na escala log-log, produz uma estimativa do coeficiente de Hurst H.

O gráfico log-log da estatística R/S versus k_i para a série de retornos absolutos de preços diários das ações da Petrobras é apresentado na Figura 3.4. A linha sólida representa a reta ajustada, e a linha pontilhada representa o caso de ausência de memória longa. Pode-se observar que a linha sólida está afastada da linha pontilhada, o que evidencia a presença de memória longa na série apresentada. O valor da estimativa do parâmetro H para a série de retornos dos preços de ações da Petrobras é igual a 0, 6730694 e, portanto, a estimativa para o parâmetro d é igual a 0, 1730694.

Um ponto fraco desse procedimento é que para uma amostra em particular não é claro o quanto que podemos considerar o valor k_1 como "grande o suficiente". Além disso, para valores grandes de k_i , poucos valores da estatística R/S podem ser calculados, a não ser que o tamanho da amostra seja muito grande. Para minimizar esse problema, pode-se usar o método LAD⁶ (*least absolute deviation*)

⁶No software S-Plus usamos a função d.ros para calcular a estatística R/S. Para calcular a estatística R/S usando o



Log-Log R/S Plot

Figura 3.4: Gráfico Log-Log - estatística R/S

para o ajuste da reta, obtendo-se como resultado uma estimativa robusta do parâmetro de memória longa. Os valores das estimativas dos parâmetros de memória longa para os retornos da Petrobras, quando usamos o método LAD, são $\hat{H} = 0,6753404$ e $\hat{d} = 0,1753404$.

3.4.2 Método do Periodograma

Conforme descrito na seção (3.2), a densidade espectral de um processo de memória longa se aproxima de $C_f \omega^{1-2H}$ quando a freqüência ω é próxima a zero. Uma vez que a densidades espectral pode ser estimada pelo periodograma, o gráfico na escala log-log do periodograma versus a freqüência deve oscilar em torno de uma linha reta com coeficiente de inclinação igual a 1-2H para freqüência

método LAD, deve-se adicionar o argumento opcional method="l1".

próximas a zero. Esse método também pode ser usado para se obter uma estimativa do parâmetro de memória longa H, e é usualmente chamado de *método do periodograma*.

Assim como na estimativa de R/S para o parâmetro de memória longa, pode-se usar o método LAD^7 para o ajuste da reta, obtendo-se como resultado uma estimativa mais robusta do parâmetro de memória longa.

Aplicando este método à série de retornos dos preços de ações da Petrobras temos uma estimativa de 0, 4612033 para o parâmetro d. Utilizando o método LAD, obtém-se $\hat{d} = 0, 212811$. A Figura 3.5 apresenta os gráficos (método mínimos quadrados e LAD) na escala log-log do periodograma versus a freqüência para a série de retornos dos preços de ações da Petrobras.



Log-Log Plot - Periodograma

Figura 3.5: Gráfico Log-Log do Periodograma - (a) mínimos quadrados e (b) método LAD

⁷Função d.pgram do sofware S-Plus. Para calcular a estatística robusta usando o método LAD, deve-se adicionar o argumento opcional method="l1".

3.4.3 Método de Whittle

O método de Whittle para estimação do parâmetro de memória longa d é baseado na estimativa de máxima verossimilhança no domínio da frequência de um processo fracionário integrado (3.6). Pode-se mostrar que os parâmetros desconhecidos de (3.6) podem ser estimados se minimizarmos uma versão discreta de

$$Q(\theta) = \int_{-\pi}^{\pi} \frac{I(\omega)}{f(\theta;\omega)} d\omega$$

onde θ é o vetor de parâmetros desconhecidos, incluindo o parâmetro de diferença fracionária $d, I(\omega)$ é o periodograma de y_t , e $f(\theta; \omega)$ é a densidade espectral teórica de y_t .

A estimativa do parâmetro de memória longa resultante do método de Whittle para a série de retornos da Petrobras é de 0,2569083.⁸

3.4.4 Método das Ondaletas

Ondaletas são consideradas alternativas a outros sistemas de funções usados como base para representação de funções pertencentes a certos espaços, como senos e cossenos, polinômios ortogonais, funções de Haar, Walsh, etc. Um fato básico sobre as ondaletas é que elas são *localizadas* no tempo (ou espaço), ao contrário do que ocorre com as funções trigonométricas. Bases de Fourier são localizadas em freqüência, mas não no tempo. Pequenas mudanças em algumas observações podem provocar mudanças em todas as componentes de uma expansão de Fourier, o que não acontece com uma expansão em série de ondaletas.

As ondaletas são funções que satisfazem algumas propriedades. Considere o espaço $L^2(\Re)$ de todas as funções mensuráveis de quadrado integrável sobre \Re , onde as funções f(t) devem decair para zero quando $|t| \to \infty$. A idéia é obter dilatações (compressões) e translações de uma única função ψ , de modo a cobrir \Re . Assim, consideramos ondaletas como

$$\psi_{j,k}(t) = \psi(2^{j}t - k), j, k \in \mathbb{Z}.$$
(3.13)

As funções $\{\psi_{j,k}(t), j, k \in \mathbb{Z}\}$ formam uma base que não precisa ser necessariamente ortonormal. Como existem muitas vantagens em se trabalhar com bases ortogonais, consideremos uma base or-

⁸Função d.whittle do software S-Plus. Esta função está implementada levando-se em conta que u_t tem distribuição normal padrão e, portanto, é aplicado quando y_t é um processo ARFIMA(0, d, 0).

tonormal gerada por ψ :

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k), j, k \in \mathbb{Z}.$$
(3.14)

Existem vários tipos de ondaletas na literatura. Trabalharemos com a mais simples e mais antiga delas, a ondaleta Haar, que é uma função que assume os valores 1 e -1, em [0, 1/2) e [1/2, 1), respectivamente. Qualquer função contínua pode ser uniformemente aproximada pela ondaleta Haar (Vidakovic e Müller (1991)).

Definição 3.4.1. Seja x(t) uma função real de $L^2(\Re)$, e suponha que existam observações para x(t)em $t = 0, 1, ..., 2^p - 1$, onde $p \in \mathbb{Z}^+$. Definimos o produto escalar, $\langle \cdot \rangle$, da seguinte forma:

$$\langle x,g\rangle = \int x(t)g(t)dt,$$

onde $g \in L^2(\Re)$.

O coeficiente de ondaletas de x(t) é obtido pelo produto escalar:

$$w_{j,k} = \langle x, \psi_{j,k} \rangle = 2^{j/2} \int x(t)\psi(2^jt - k)dt.$$
 (3.15)

Se a função ψ for a ondaleta Haar, temos que $\psi_{j,k}$ é dado por:

$$\psi^{H}_{j,k}(t) = \begin{cases} 2^{j/2}, 2^{-j}k \leq t < 2^{-j}(k+1/2) \\ -2^{j/2}, 2^{-j}(k+1/2) \leq t < 2^{-j}(k+1) \\ 0, \text{caso contrário.} \end{cases}$$

Supondo que x(t) é um processo integrado fracionário, I(d), conforme definido em (3.6), temos que a função de autocovariância (Granger e Joyeux (1980); Hosking (1981); Brockwell e Davis (1991); Beran (1994)) desse processo é dada por:

$$R_{x}(t,s) = E[x(t)x(s)] = \frac{\sigma_{\epsilon}^{2}\Gamma(1-2d)\Gamma(|t-s|+d)}{\Gamma(d)\Gamma(1-d)\Gamma(|t-s|+1-d)}$$
(3.16)

$$\approx |t-s|^{2d-1} \text{ quando } |t-s| \to \infty.$$
(3.17)

Teorema 3.4.1. Seja x(t) um processo I(d) com média zero e |d| < 1/2. Os coeficientes de ondaletas,

 $w_{j,k}$, associados ao processo x(t) possuem distribuição $N(0, \sigma^2 2^{-2jd})$ quando $j \to \infty$, onde σ^2 é uma constante finita.

Demonstração. (ver Jensen (1999))

Pelo Teorema (3.4.1), os coeficientes de ondaletas de um processo I(d) possuem variância que é função do parâmetro de escala, j, mas que é independente do parâmetro de translação, k. Dessa forma, defina a variância do coeficiente de ondaleta na escala j como sendo $R(j) = \sigma^2 2^{-2jd}$. Tomando a transformação logarítmica de R(j), temos a relação:

$$\ln(R(j)) = \ln(\sigma^2) - d\ln(2^{2j}), \tag{3.18}$$

onde ln(R(j)) é linearmente relacionado com $\ln 2^{2j}$, ponderado pelo parâmetro de memória longa d. Portanto, o parâmetro d de uma série integrada fracionária pode ser estimado por meio do estimador de mínimos quadrados ordinários (\hat{d}) .

Para prosseguir com essa regressão de mínimos quadrados ordinários é necessário um estimador da variância populacional do coeficiente de ondaletas, R(j). Na escala j, defina a variância amostral do coeficiente de ondaletas da seguinte forma:

$$\bar{R}(j) = \frac{1}{2^j} \sum_{k=0}^{2^j - 1} w_{j,k}^2.$$
(3.19)

Se existe um número grande de coeficientes de ondaletas para a escala j, a variância amostral do coeficiente de ondaletas pode ser usada como um estimador consistente para a variância verdadeira (ou populacional), R(j).

Seja

$$y_j = \ln(2^{-2j}) - \frac{1}{p} \sum_{j=0}^{p-1} \ln(2^{-2j}),$$

onde p é tal que $T = 2^p$ e T é o número de observações do processo integrado fracionário x(t). Segundo Jensen (1999), o estimador de mínimos quadrados ordinários para o parâmetro de diferença

3.5. ESTIMAÇÃO DE MODELOS ARFIMA

fracionária é consistente e dado por:

$$\hat{d} = \left[\sum_{j=0}^{p-1} y_j^2\right]^{-1} \left[\sum_{j=0}^{p-1} y_j \ln(\bar{R}(j))\right].$$

3.5 Estimação de Modelos ARFIMA

As seções anteriores ilustraram como realizar testes de presença de memória longa em séries temporais e como estimar os parâmetros de memória longa H ou d. Essa seção introduz os modelos ARIMA fracionários (ARFIMA) que são capazes de modelar tanto a estrutura de correlação de ordens altas (memória longa) quanto a estrutura de correlação de ordens baixas. Será apresentado o método de estimação de máxima verossimilhança para estimação dos parâmetros d, $\phi \in \theta$ de um modelo ARFIMA(p, d, q).

3.5.1 Modelos ARIMA Fracionários

A abordagem tradicional pra modelar uma série temporal y_t do tipo I(d) é usar um modelo ARIMA:

$$\phi(L)(1-L)^d(y_t-\mu) = \theta(L)\epsilon_t \tag{3.20}$$

onde $\phi(L) \in \theta(L)$ são polinômios em L

$$\phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i$$
$$\theta(L) = 1 - \sum_{j=1}^{q} \theta_j L^j$$

com raízes fora do círculo unitário, e ϵ_t uma variável aleatória *iid* com média zero e variância σ^2 . Este é o chamado modelo ARIMA(p, d, q). Nos casos de o parâmetro d ser um número real, em vez de um número inteiro, o modelo ARIMA passa a ser um modelo autorregressivo fracionário integrado de médias móveis, ou ARFIMA(p, d, q)

Para o processo ARFIMA(p, d, q) dado por (3.20), Hosking (1981) demonstrou que:

• se trata de um processo estacionário caso $d < \frac{1}{2}$ e se todas as raízes de $\phi(L)$ estiverem fora do

círculo unitário;

• é um processo invertível se $d > -\frac{1}{2}$ e todas as raízes de $\theta(L)$ estiverem fora do círculo unitário.

A seguir detalharemos alguns casos particulares do modelo $\operatorname{ARFIMA}(p, d, q)$.

Exemplo 3.5.1. Seja y_t um ARFIMA(0, d, 0). Esse processo é chamado de ruído branco fracionário e é dado por:

$$(1-L)^d y_t = \epsilon_t. aga{3.21}$$

• Quando $d < \frac{1}{2}$, y_t é um processo fracionário estacionário representado por $y_t = \psi(L)\epsilon_t$, com pesos

$$\psi_k = \frac{d(1+d)\cdots(k-1+d)}{k!} = \frac{(k+d-1)!}{k!(d-1)!} = \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)},$$

onde

e

$$\psi_k \sim \frac{k^{(d-1)}}{(d-1)!} = C_1 k^{d-1}, k \to \infty,$$

sendo C_1 uma constante.

• Quando $d > -\frac{1}{2}$, o processo é invertível com representação na forma $\pi(L)y_t = \epsilon_t$, com pesos da forma

$$\pi_k = \frac{-d(1-d)\cdots(k-1-d)}{k!} = \frac{(k-d-1)!}{k!(-d-1)!} = \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)},$$
$$\pi_k \sim \frac{k^{-d-1}}{(-d-1)!} = C_2 k^{-d-1}, k \to \infty,$$

 $com C_2$ constante.

• Quando $-\frac{1}{2} < d < \frac{1}{2}$:

(a) A função de densidade espectral é dada por

$$f(\lambda) = \begin{cases} \frac{\sigma_{\epsilon}^{2}}{2\pi} (2sen(\frac{\lambda}{2}))^{-2d}, 0 < \lambda \le \pi, \\ \lambda^{-2d}, \lambda \to 0; \end{cases}$$
(3.22)

3.5. ESTIMAÇÃO DE MODELOS ARFIMA

(b) A função de autocorrelação pode ser escrita da seguinte forma

$$\rho_h = \frac{(-d)!(h+d-1)!}{(d-1)!(h-d)!} = \prod_{0 \le k \le h} \frac{k-1+d}{k-d}, h = 1, 2, \dots$$
(3.23)

Em particular,

$$\rho_1 = \frac{d}{1-d} \tag{3.24}$$

e

$$\rho_h \sim \frac{(-d)!h^{2d-1}}{(d-1)!} = C_3 h^{2d-1}, h \to \infty,$$

com C_3 constante;

(c) A função de autocorrelação parcial, por sua vez, é descrita como

$$\phi_{hh} = \frac{d}{h-d}, h = 1, 2, \dots;$$

(d) A variância γ_0 é dada por

$$\gamma_0 = \frac{(-2d)!}{(-d)!^2}.$$

Exemplo 3.5.2. Seja y_t um ARFIMA(1, d, 0). Este processo é estacionário e invertível se $-\frac{1}{2} < d < \frac{1}{2}$ e $|\phi| < 1$ e é dado por

$$(1-L)^d (1-\phi L)y_t = \epsilon_t.$$

(a) Os pesos $\psi_j e \pi_j$ das representações autoregressiva ($y_t = \psi(L)\epsilon_t$) e de médias móveis infinitas ($\pi(L)y_t = \epsilon_t$) podem ser escritos, respectivamente, como

$$\psi_j = \frac{(j+d-1)!}{j!(d-1)!} F(1;-j;1-d-j;\phi)^9 \sim \frac{j^{d-1}}{(1-\phi)(d-1)!}$$

e

$$\pi_j = \frac{(j-d-2)!}{(j-1)!(-d-1)!} \left\{ 1 - \phi - (1+d)/j \right\} \sim \frac{(1-\phi)}{(-d-1)!} j^{-d-1}, \text{ quando } j \to \infty;$$

 ${}^9F(a;b;c;d) = 1 + \frac{ab}{c}z + \frac{a(a+1)b(b+1)}{c(c+1)}z^2 + \cdots$ é a chamada função hipergeométrica

(b) A função densidade espectral desse processo é dada por

$$f(\lambda) = \begin{cases} \frac{\sigma_{\epsilon}^2}{2\pi} \frac{(2sen(\frac{\lambda}{2}))^{-2d}}{1+\phi^2 - 2\phi \cos\lambda}, 0 < \lambda \le \pi, \\ \frac{\lambda^{-2d}}{(1-\phi)^2}, \lambda \to 0; \end{cases}$$

(c) Em particular, a função de autocorrelação pode ser escrita da seguinte forma

$$\rho_1 = \frac{(1+\phi^2)F(1;d;1-d;\phi)-1}{\phi[2F(1;d;1-d;\phi)-1]}$$

e

$$\rho_h = \frac{(-d)!(1+\phi)h^{2d-1}}{(d-1)!(1-\phi)^2 F(1;1+d;1-d;\phi)}, h \to \infty;$$

(d) A variância desse processo é dada por

$$\gamma_0 = \frac{(-2d)F(1; 1+d; 1-d; \phi)!}{(1+\phi)[(-d)!]^2}.$$

Exemplo 3.5.3. Consideremos, agora, y_t um processo ARFIMA(0, d, 1) dado por

$$(1-L)^d y_t = (1-\theta L)\epsilon_t.$$

 $\label{eq:trata-se} \textit{ de um processo estacionário e invertível se } -\tfrac{1}{2} < d < \tfrac{1}{2} \ e \ |\theta \ | < 1.$

(a) Os pesos $\psi_j e \pi_j$ das representações auto-regressiva e de médias móveis infinitas são, respectivamente, dados por

$$\psi_j = \frac{(j-d-1)!}{j!(-d-1)!} F(1;-j;1+d-j;\theta) \sim \frac{j^{-d-1}}{(1-\theta)(-d-1)!}$$

e

$$\pi_j = \frac{(j+d-2)!}{(j-1)(d-1)!} \left[1 - \theta - \frac{(1+d)}{j} \right] \sim \frac{(1-\theta)}{(d-1)!} j^{d-1}, quando \ j \to \infty,$$

sendo $F(\cdot)$ a função hipergeométrica;

(b) A função de densidade espectral pode ser escrita como

$$f(\lambda) = \frac{\sigma_{\epsilon}^2}{2\pi} \left[1 + \theta^2 - 2\theta \cos \lambda \right] \left[2\sin\left(\frac{\lambda}{2}\right) \right]^{-2d} \sim \left[(1 - \theta^2 \lambda^{-2d}) \right] quando \ \lambda \to \infty;$$

3.5. ESTIMAÇÃO DE MODELOS ARFIMA

(c) Em particular, para a função de autocorrelação, temos

$$\rho_1 = \frac{(1+\theta^2)d(2-d) - 2\theta(1-d+d^2)}{(1-d)(2-d)1 + \theta^2 - 2\theta d/(1-d)}$$

e

$$\rho_h = \frac{(-d)!}{(d-1)!} ah^{2d-1}, h \to \infty,$$

sendo $a = \frac{(1-\theta^2)}{(1+\theta^2-2\theta d)/(1-d)}.$

3.5.2 Método de Máxima Verossimilhança

A função de verossimilhança de $\mathbf{Z} = (Z_1, \dots, Z_n)$ proveniente de um processo ARFIMA(p, d, q)pode ser expressa na forma

$$L(\eta, \sigma_a^2) = (2\pi\sigma_a^2)^{-n/2} (r_0 \cdots r_{n-1})^{-1/2} \exp\left[-\frac{1}{2\sigma_a^2} \sum_{j=1}^n (Z_j - \hat{Z}_j)^2 / r_{j-1}\right],$$
(3.25)

em que $\eta = (d, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q), \hat{Z}_j, j = 1, \dots, n$, são as previsões um passo à frente e $r_{j-1} = (\sigma_a^2)^{-1} E(Z_j - \hat{Z}_j)^2$.

Os estimadores de máxima verossimilhança dos parâmetros são dados por

$$\hat{\sigma}_{MV}^2 = n^{-1} S(\hat{\eta}_{MV}), \qquad (3.26)$$

onde

$$S(\hat{\eta}_{MV}) = \sum_{j=1}^{n} (Z_j - \hat{Z}_j)^2 / r_{j-1}$$

e $\hat{\eta}_{MV}$ é o valor de η que minimiza

$$l(\eta) = \ln(S(\eta)|n) + n^{-1} \sum_{j=1}^{n} \ln r_{j-1}.$$

Entretanto, o cálculo de $l(\eta)$ é bastante lento. Para um procedimento alternativo, considera-se

uma aproximação para $l(\eta)$ dada por

$$l(\eta) \cong l_*(\eta) = \ln \frac{1}{n} \sum_j \frac{I(\omega_j)}{2\pi f(\eta; \omega_j)},$$
(3.27)

em que

$$I(\omega_j) = \frac{1}{n} \left| \sum_{t=1}^n Z_t e^{-it\omega_j} \right|^2$$

é o periodograma dos dados,

$$f(\eta;\omega_j) = \frac{\sigma_{\epsilon}^2}{2\pi} \frac{\left|1 - \theta_1 e^{-i\omega_j} - \dots - \theta_q e^{-iq\omega_j}\right|^2}{\left|1 - \phi_1 e^{-i\omega_j} - \dots - \phi_p e^{-ip\omega_j}\right|^2} \cdot \left|1 - e^{-i\omega_j}\right|^2$$

é a função de densidade espectral do processo Z_t e \sum_j é a soma sobre todas as freqüências de Fourier, $\omega_j = 2\pi j/n \in (-\pi, \pi].$

Hannan (1973) e Fox e Taqqu (1986) mostram que:

- (i) o estimador $\hat{\eta}_{MV}$ que minimiza (3.27) é consistente;
- (ii) se d > 0, então

$$\hat{\eta}_{MV} \xrightarrow{D} N(\eta, n^{-1}A^{-1}(\eta)), \qquad (3.28)$$

em que $A(\eta)$ é uma matriz de ordem $(p+q+1) \times (p+q+1)$ com (j,k)-ésimo elemento dado por

$$A_{jk}(\eta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial \ln f(\lambda;\eta)}{\partial \eta_j} \frac{\partial \ln f(\lambda;\eta)}{\partial \eta_k} d\lambda;$$

(iii) a variância σ_{ϵ}^2 é estimada por

$$\hat{\sigma}_{MV}^2 = \frac{1}{n} \sum_j \frac{I_n(w_j)}{2\pi f(w_j; \hat{\eta}_{MV})}.$$

3.6 Simulações

O objetivo desta seção é avaliar o desempenho e o comportamento dos estimadores do parâmetro de memória longa d apresentados na Seção 3.4. Para tanto, foram realizadas 500 simulações de

3.6. SIMULAÇÕES

amostras de tamanho 1024 dos modelos ARFIMA(0, d, 0), ARFIMA(1, d, 0) e ARFIMA(0, d, 1), com d tomando os valores 0,05, 0,1, 0,2, 0,25, 0,3, 0,4 e 0,45 em cada um dos diferentes processos. Para os modelos ARFIMA(1, d, 0) e ARFIMA(0, d, 1), os parâmetros autorregressivo (ϕ_1) e de média móveis (θ_1) assumem valores 0,3 e -0, 6, respectivamente.

Para gerar as séries deste capítulo, foi utilizada uma rotina do software S-Plus que simula modelos ARFIMA(p, d, q). Os estimadores dos métodos R/S, periodograma, Whittle e máxima verossimilhança foram obtidos por meio de funções já implementadas no S-Plus. Para o estimador de Ondaletas foi implementado um programa no próprio S-Plus que está apresentado no Apêndice A.

O estimador de Whittle será utilizado para estimar o parâmetro d somente das simulações do modelo ARFIMA(0, d, 0).

Para avaliar e comparar os estimadores do parâmetro d utilizamos:

(a) Média dos valores estimados

$$\bar{d_e} = \frac{\sum_{i=1}^{n=500} d_{ei}}{500}; \tag{3.29}$$

(b) Variância estimada de \hat{d}

$$\hat{Var}(\hat{d}_e) = \frac{\sum_{i=1}^{n=500} (\hat{d}_{ei} - \bar{d}_e)^2}{500};$$
(3.30)

(c) Erro quadrático médio (EQM) de \hat{d} estimado

$$E\hat{Q}M(\hat{d}_e) = \frac{\sum_{i=1}^{n=500} (\hat{d}_{ei} - d)^2}{500},$$
(3.31)

onde e se refere ao método utilizado para obtenção do estimador de d.

Nomenclatura dos estimadores:

Estimador	nomenclatura
estimador baseado na estatística R/S	\mathbf{rs}
estimador baseado na estatística $\mathrm{R/S}$ (método LAD)	rs2
estimador baseado no periodograma	р
estimador baseado no periodograma (robusto)	p2
método de Whittle	whi
método de ondaletas	wav
método de máxima verossimilhança	mv

Modelos ARFIMA(0, d, 0)

Na Tabela 3.1 estão apresentadas as estimativas da média, da variância e do erro quadrático médio (EQM) do parâmetro de memória longa dos 500 modelos ARFIMA(0, d, 0) simulados, com d tomando valores 0,05, 0,1, 0,2, 0,25, 0,3, 0,4 e 0,45. Na Figura 3.6 apresentamos os gráficos box-plot com 500 estimativas dos parâmetros de memória longa, e a Figura 3.7 apresenta os gráficos das variâncias estimadas e dos EQM's estimados. Podemos observar o comportamento dos diferentes tipos de estimadores utilizados na estimação dos parâmetros de memória longa d.

Verifica-se que os estimadores que possuem menor variância e menor EQM são os estimadores do método de Whittle e do método de máxima verossimilhança. Além disso, estes dois estimadores apresentam comportamento muito semelhante para todos os valores de d simulados.

Os estimadores baseados na estatística R/S possuem pequena variação, entretanto, o comportamento do EQM não mantém um padrão para os diferentes valores de d. Para d = 0, 2 e d = 0, 25 o EQM assume valores baixos, e para os outros valores de d o EQM tem comportamento crescente à medida que d se aproxima de zero ou de 0, 5.

As maiores variâncias são observadas para o estimador do método de ondaletas e para o estimador baseado no periodograma, apesar do estimador do periodograma resultar sempre em medianas muito próximas do verdadeiro valor de d. O EQM do estimador baseado no periodograma robusto possui o pior comportamento, uma vez que aumenta consideravelmente à medida que d se aproxima de 0, 5.

3.6. SIMULAÇÕES

		rs	rs2	р	p2	whi	wav	mv
d=0,05	Média	0,1368	0,1392	$0,\!0579$	0,0209	0,0447	-0,0268	0,0462
	Variância Amostral	0,0003	0,0003	0,0101	0,0026	0,0006	0,0099	0,0006
	EQM	0,0078	0,0083	0,0102	0,0035	0,0007	0,0158	0,0006
d=0,1	Média	$0,\!1566$	$0,\!1598$	0,1029	0,0445	0,0937	0,0269	0,0950
	Variância Amostral	0,0002	0,0003	0,0112	0,0028	0,0006	0,0090	0,0006
	\mathbf{EQM}	0,0034	0,0039	0,0112	0,0059	0,0007	0,0144	0,0006
d=0,2	Média	0,2003	0,2047	0,1948	0,0852	$0,\!1950$	$0,\!1137$	$0,\!1955$
	Variância Amostral	0,0003	0,0003	0,0108	0,0027	0,0008	0,0093	0,0006
	EQM	0,0003	0,0004	0,0108	0,0158	0,0008	0,0168	0,0007
-	Média	0,2203	0,2257	0,2429	0,1044	0,2434	$0,\!1568$	0,2438
d=0,25	Variância Amostral	0,0002	0,0003	0,0104	0,0024	0,0007	0,0095	0,0006
	\mathbf{EQM}	0,0011	0,0009	0,0104	0,0236	0,0008	0,0182	0,0007
d=0,3	Média	0,2402	0,2467	0,2994	$0,\!1292$	$0,\!2937$	$0,\!2057$	$0,\!2933$
	Variância Amostral	0,0002	0,0003	0,0111	0,0025	0,0008	0,0098	0,0007
	\mathbf{EQM}	0,0038	0,0032	0,0111	0,0316	0,0008	0,0187	0,0007
d=0,4	Média	0,2789	0,2872	0,4029	$0,\!1729$	0,3939	0,3020	$0,\!3917$
	Variância Amostral	0,0002	0,0003	0,0128	0,0029	0,0007	0,0110	0,0006
	EQM	0,0149	0,0130	0,0128	$0,\!0544$	0,0007	0,0206	0,0007
d=0,45	Média	0,2977	0,3077	0,4609	$0,\!1978$	0,4433	0,3629	0,4402
	Variância Amostral	0,0002	0,0003	0,0113	0,0025	0,0007	0,0090	0,0008
	EQM	0,0234	0,0206	0,0114	0,0661	0,0008	0,0165	0,0009

Tabela 3.1: Estimativas da Média, Variância e EQM do parâmetro dnas simulações do processo $\operatorname{ARFIMA}(0,d,0)$



Figura 3.6: Boxplot das 500 estimativas do parâmetro de memória longa das simulações do processo ARFIMA(0,d,0)

3.6. SIMULAÇÕES



Figura 3.7: Variância estimada e EQM estimado das estimativas do parâmetro de memória longa das simulações do processo ARFIMA(0, d, 0)

Modelos ARFIMA(1, d, 0)

O resultado das 500 simulações do modelo ARFIMA(1, d, 0) estão apresentados na Tabela 3.2, que apresenta as estimativas da média, da variância e do erro quadrático médio (EQM) do parâmetro d. As Figuras 3.8 e 3.9 apresentam os gráficos box-plot com as estimativas dos parâmetros de memória longa e os gráficos das variâncias estimadas e dos EQM's estimados, respectivamente.

Neste caso, não há um estimador que possui melhor performance para todos os valores de d. Observa-se que os estimadores que possuem pequena variação e valores menores do EQM são os estimadores baseados na estatística R/S quando $d \ge 0, 2$, atingindo o menor valor de EQM para d = 0, 3. Entretanto, os valores do EQM para valores mais próximos de zero aumentam consideravelmente.

O estimador de máxima verossimilhança, que apresentou bons resultados para os modelos AR-FIMA (0, d, 0), apresenta para esta situação valores menores de EQM apenas quando $d \leq 0, 1$, ou seja, nos casos em que os estimadores baseados na estatística R/S não apresentam bom desempenho.

Assim como no caso dos modelos ARFIMA(0, d, 0) simulados, o EQM do estimador baseado no periodograma robusto possui o pior comportamento, aumentando conforme d se aproxima de 0, 5.

Os resultados para o estimador baseado no periodograma mostram valores médios de d mais próximos aos valores verdadeiros, porém a variância observada é a maior de todos os casos.

/ / /							
		rs	rs2	р	p2	wav	mv
d=0,05	Média	0,2211	$0,\!2348$	0,0410	0,0190	$0,\!0335$	0,0152
	Variância Amostral	0,0002	0,0003	0,0111	0,0028	0,0101	0,0068
	EQM	0,0295	$0,\!0344$	0,0112	0,0038	0,0103	0,0080
	Média	0,2382	0,2529	0,1075	0,0431	0,0814	0,0696
d=0,1	Variância Amostral	0,0002	0,0003	0,0119	0,0026	0,0100	0,0069
	\mathbf{EQM}	0,0193	0,0236	0,0119	0,0058	0,0104	0,0078
d=0,2	Média	0,2723	0,2891	0,2174	0,0934	0,1828	0,1425
	Variância Amostral	0,0002	0,0003	0,0108	0,0026	0,0090	0,0051
	\mathbf{EQM}	0,0055	0,0082	0,0111	0,0140	0,0093	0,0059
d=0,25	Média	0,2907	0,3083	0,2585	0,1115	0,2240	0,2149
	Variância Amostral	0,0002	0,0003	0,0109	0,0024	0,0077	0,0058
	EQM	0,0019	0,0036	0,0110	0,0216	0,0083	0,0070
d=0,3	Média	0,3051	0,3235	0,3118	0,1358	0,2672	0,2581
	Variância Amostral	0,0002	0,0002	0,0097	0,0024	0,0093	0,0091
	EQM	0,0002	0,0008	0,0098	0,0294	0,0103	0,0109
d=0,4	Média	0,3362	0,3561	0,4173	0,1782	0,3653	0,3103
	Variância Amostral	0,0002	0,0002	0,0106	0,0027	0,0088	0,0458
	EQM	0,0043	0,0022	0,0109	$0,\!0519$	0,0100	0,0537
d=0,45	Média	0,3506	0,3719	0,4637	0,2005	0,4085	0,3693
	Variância Amostral	0,0002	0,0002	0,0106	0,0029	0,0093	0,0489
	EQM	0,0100	0,0063	0,0108	0,0652	0,0110	0,0553

Tabela 3.2: Estimativas da Média, Variância e EQM do parâmetro dnas simulações do processo $\operatorname{ARFIMA}(1,d,0)$

3.6. SIMULAÇÕES



Figura 3.8: Boxplot das 500 estimativas do parâmetro de memória longa das simulações do processo $\mathrm{ARFIMA}(1,d,0)$



Figura 3.9: Variância estimada e EQM estimado das estimativas do parâmetro de memória longa das simulações do processo ARFIMA(1, d, 0)

Modelos ARFIMA(0, d, 1)

A Tabela 3.3 e as Figuras 3.10 e 3.11 apresentam os resultados das simulações dos modelos ARFIMA(0, d, 1).

Observa-se que o estimador que possui menor EQM em quase todos os casos é o estimador de máxima verossimilhança, apresentando comportamento muito semelhante para todos os valores de d simulados.

Assim como para os modelos simulados ARFIMA(0, d, 0) e ARFIMA(1, d, 0) os estimadores baseados na estatística R/S possuem pequena variação, mas os valores do EQM não mantém um padrão para os diferentes valores de d. O EQM atinge o menor valor quando d = 0, 3 e maior valor quando d = 0, 05.

As maiores variâncias são observadas para o estimador baseado no periodograma. E o EQM do estimador baseado no periodograma robusto aumenta gradualmente à medida que d se aproxima de 0, 5, repetindo o comportamento observado para os outros dois tipos de modelos simulados.

3.6. SIMULAÇÕES

		rs	rs2	р	p2	wav	mv
d=0,05	Média	0,2233	0,2458	0,0608	0,0241	0,0368	$0,\!0451$
	Variância Amostral	0,0002	0,0003	0,0106	0,0026	0,0090	0,0009
	EQM	0,0302	0,0386	0,0107	0,0033	0,0092	0,0009
	Média	0,2387	0,2616	0,0956	0,0410	0,0785	0,0926
d=0,1	Variância Amostral	0,0002	0,0003	0,0097	0,0025	0,0101	0,0008
	EQM	0,0194	0,0264	0,0097	0,0059	0,0106	0,0008
d=0,2	Média	0,2727	0,2960	0,2051	0,0864	0,1675	0,1940
	Variância Amostral	0,0002	0,0003	0,0105	0,0024	0,0101	0,0008
	\mathbf{EQM}	0,0055	0,0095	0,0105	0,0153	0,0111	0,0009
	Média	0,2876	0,3119	0,2563	0,111	0,2188	0,2441
d=0,25	Variância Amostral	0,0002	0,0003	0,0114	0,0027	0,0091	0,0008
	\mathbf{EQM}	0,0016	0,0041	0,0114	0,0220	0,0101	0,0008
	Média	0,3019	0,3264	0,3053	0,1299	0,2614	0,2925
d=0,3	Variância Amostral	0,0002	0,0003	0,0111	0,0026	0,0103	0,0007
	\mathbf{EQM}	0,0002	0,0010	0,0111	0,0315	0,0117	0,0008
d=0,4	Média	0,3302	0,3557	0,4077	0,1731	0,3548	0,3884
	Variância Amostral	0,0002	0,0002	0,0116	0,0025	0,0109	0,0008
	EQM	0,0051	0,0022	0,0117	0,0539	0,0129	0,0009
d=0,45	Média	0,3443	0,3699	0,4643	0,1974	0,4117	$0,\!4395$
	Variância Amostral	0,0002	0,0002	0,0109	0,0027	0,0081	0,0010
	EQM	0,0114	0,0066	0,0111	0,0665	0,0096	0,0011

Tabela 3.3: Estimativas da Média, Variância e EQM do parâmetro dnas simulações do processo $\operatorname{ARFIMA}(0,d,1)$



Figura 3.10: Boxplot das 500 estimativas do parâmetro de memória longa das simulações do processo ARFIMA(0,d,1)

3.6. SIMULAÇÕES



Figura 3.11: Variância estimada e EQM estimado das estim
ativas do parâmetro de memória longa das simulações do processo
 $\operatorname{ARFIMA}(0,d,1)$

Capítulo 4

Modelos GARCH

4.1 Introdução

No mercado financeiro grandes mudanças têm tendência de serem seguidas por grandes mudanças, e pequenas mudanças são normalmente seguidas por pequenas mudanças. Podemos dizer, então, que o mercado financeiro é por vezes mais volátil, e por vezes menos ativo.

O comportamento volátil do mercado financeiro é usualmente referido como "volatilidade". A volatilidade se transformou num conceito muito importante em diferentes áreas da teoria e da prática financeira, como gerenciamento de risco, seleção de portfólio, derivação de preços, etc. Em termos estatísticos, a volatilidade é medida pela variância, ou pelo desvio padrão.

Nesse capítulo vamos introduzir a classe dos modelos GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*) desenvolvidos por Engle (1982), Bollerslev (1986), Nelson (1991), e outros, que são capazes de modelar a volatilidade variando no tempo e de capturar muitos dos fatos estilizados do comportamento da volatilidade observada em séries temporais financeiras.

4.2 Modelos ARCH

Na Figura 4.1 temos os retornos diários da série de preços de ações da Petrobras, no período de 03 de janeiro de 1995 à 27 de dezembro de 2000. Pode-se perceber que tanto as pequenas mudanças quanto as grandes mudanças parecem estar agrupadas. Essa tendência é típica de muitas séries temporais macroeconômicas e financeiras. Para confirmar este comportamento são apresentadas na Figura 4.2 as funções de autocorrelação dos retornos e dos retornos ao quadrado da série de preços da Petrobras.



Figura 4.1: Retornos da série de Preços diários das ações da Petrobras

Observando os gráficos das funções de autocorrelação percebemos que há autocorrelações significantes apenas para alguns lags iniciais da série de retornos. Por outro lado, a autocorrelação da série dos retornos ao quadrado se mostra significante pelo menos até o lag 30. Uma vez que os retornos ao quadrado medem o momento de segunda ordem da série original, este resultado indica que a variância dos retornos da série de preços da Petrobras condicionada ao passado histórico pode sofrer alterações ao longo do tempo, ou seja, a série de retornos pode mostrar a presença de *heteroscedasticidade* condicional variando com o tempo.

A heteroscedasticidade condicional da série dos retornos ao quadrado pode ser modelada usando um processo autorregressivo simples (AR) para resíduos quadráticos. Seja y_t uma série temporal estacionária tal como retornos financeiros, então y_t pode ser expresso como sua média acrescida de



Figura 4.2: (a) FAC dos retornos e (b) FAC do quadrado dos retornos da série de preços da Petrobras

um ruído branco, caso não exista autocorrelação significante em y_t :

$$y_t = c + \epsilon_t, \tag{4.1}$$

onde c é a média de y_t e ϵ_t é iid com média zero.

Para incluir o efeito da heteroscedasticidade condicional, seja $Var_{t-1}(\epsilon_t) = \sigma_t^2$, onde $Var_{t-1}(\cdot)$ é a variância condicional com informação até o tempo t-1, e

$$\sigma_t^2 = a_0 + a_1 \epsilon_{t-1}^2 + \dots + a_p \epsilon_{t-p}^2.$$
(4.2)

Se ϵ_t tiver média zero, $Var_{t-1}(\epsilon_t) = E_{t-1}(\epsilon_t^2) = \sigma_t^2$, a equação acima pode ser reescrita como:

$$\epsilon_t^2 = a_0 + a_1 \epsilon_{t-1}^2 + \dots + a_p \epsilon_{t-p}^2 + u_t, \tag{4.3}$$

onde $u_t = \epsilon_t^2 - E_{t-1}(\epsilon_t^2)$ é um processo diferença de martingal com média zero, mas com variância não constante.

A equação (4.3) representa um processo AR(p) para ϵ_t^2 , e o modelo em (4.1) e (4.2) é conhecido como o modelo *autorregressivo com heteroscedasticidade condicional* de Engle (1982), que é usualmente chamado de modelo ARCH(p).

Uma formulação alternativa para o modelo ARCH é

$$y_t = c + \epsilon_t$$

$$\epsilon_t = z_t \sigma_t$$

$$\sigma_t^2 = a_0 + a_1 \epsilon_{t-1}^2 + \dots + a_p \epsilon_{t-p}^2,$$

onde z_t é uma variável aleatória *iid* com uma distribuição especificada. No modelo ARCH básico a distribuição de z_t é a normal padrão *iid*.

4.2.1 Testes para efeitos ARCH

Antes de estimar um modelo ARCH para uma série temporal financeira, é interessante testar a presença de efeitos ARCH nos resíduos. Se não houver efeitos ARCH nos resíduos, então o modelo ARCH é desnecessário e mal especificado.

Uma vez que o modelo ARCH pode ser escrito como um modelo AR em termos de resíduos quadráticos como em (4.3), um teste simples de Multiplicador de Lagrange (LM) para efeitos ARCH pode ser construído baseado na regressão auxiliar (4.3). Sob a hipótese nula de que não há efeitos ARCH: $a_1 = a_2 = \cdots = a_p = 0$, temos o teste estatístico

$$LM = T \cdot R^2 \stackrel{A}{\sim} \chi^2(p),$$

onde T é o tamanho da amostra e R^2 é o quadrado do coeficiente de correlação múltipla da regressão (4.3) usando resíduos estimados.¹

Outro teste que pode ser utilizado para verificar se a série apresenta heteroscedasticidade condicional é o teste de Box-Pierce-Ljung para ϵ_t^2 .

¹ver Engle (1982) para detalhes.

4.3 O Modelo GARCH e suas Propriedades

Se o teste LM para efeitos ARCH é significante para a série temporal, pode-se estimar um modelo ARCH e obter estimativas da volatilidade σ_t baseadas no passado. No entanto, tem-se que na prática geralmente é preciso um número grande de lags p para se obter um bom ajuste no modelo, o que implica em um número grande de parâmetros. Um modelo mais parcimonioso foi proposto por Bollerslev (1986) e substitui o modelo AR em (4.2) pela formulação:

$$\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \epsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2, \qquad (4.4)$$

onde os coeficientes a_i $(i = 0, \dots, p)$ e b_j $(j = 1, \dots, q)$ são todos positivos para garantir que a variância condicional σ_t^2 seja sempre positiva.² O modelo (4.4) junto com (4.1) é conhecido como o modelo GARCH(p,q) ("generalized ARCH"). Quando q = 0, o modelo GARCH se reduz ao modelo ARCH.

No modelo GARCH(p,q) a variância condicional de ϵ_t , σ_t^2 , depende do quadrado dos resíduos nos *p* períodos anteriores, e da variância condicional nos *q* períodos anteriores. Normalmente um modelo GARCH(1,1) com apenas três parâmetros na equação da variância condicional é adequado para se obter um bom modelo ajustado para séries temporais financeiras.

4.3.1 Representação ARMA do Modelo GARCH

Assim como um modelo ARCH pode ser escrito como um modelo AR em termos de resíduos quadráticos, um modelo GARCH pode ser escrito na forma de um modelo ARMA de resíduos quadráticos. Considere o modelo GARCH(1,1):

$$\sigma_t^2 = a_0 + a_1 \epsilon_{t-1}^2 + b_1 \sigma_{t-1}^2. \tag{4.5}$$

Sendo $E_{t-1}(\epsilon_t^2) = \sigma_t^2$, a equação acima pode ser escrita como:

$$\epsilon_t^2 = a_0 + (a_1 + b_1)\epsilon_{t-1}^2 + u_t - b_1 u_{t-1}, \tag{4.6}$$

 $^{^{2}}$ Coeficientes positivos são condições suficientes mas não necessárias para garantir que a variância condicional seja positiva. Ver Nelson e Cao (1992).

que é um modelo ARMA(1,1) onde $u_t = \epsilon_t^2 - E_{t-1}(\epsilon_t^2)$ é ruído branco de média zero e variância não constante.

Dada a representação ARMA de um modelo GARCH, muitas propriedades do modelo GARCH seguem facilmente da correspondência de um processo ARMA para ϵ_t^2 . Por exemplo, para o modelo GARCH(1,1) ser estacionário é preciso que $a_1 + b_1 < 1$. Assumindo a estacionariedade do modelo GARCH(1,1), pode-se mostrar que a variância incondicional de ϵ_t é dada por $Var(\epsilon_t) = E(\epsilon_t^2) = a_0/(1-a_1-b_1)$, pois de (4.6) temos:

$$E(\epsilon_t^2) = a_0 + (a_1 + b_1)E(\epsilon_{t-1}^2)$$

e portanto

$$E(\epsilon_t^2) = a_0 + (a_1 + b_1)E(\epsilon_t^2),$$

baseado na suposição de que ϵ_t^2 é estacionário.

Para o modelo GARCH(p, q) (4.4), os resíduos quadráticos ϵ_t^2 têm comportamento de um processo ARMA(max(p, q),q). A estacionariedade requer que $\sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1$ e a variância incondicional de ϵ_t é dada por

$$\bar{\sigma}^2 = Var(\epsilon_t) = \frac{a_0}{1 - (\sum_{i=1}^p a_i + \sum_{j=1}^q b_j)}.$$
(4.7)

4.3.2 Modelo GARCH e Fatos Estilizados

Na prática, exitem alguns "fatos estilizados" sobre a volatilidade em séries temporais financeiras. Bollerslev, Engle e Nelson (1994) dão uma visão completa sobre esses fatos. Usando a representação ARMA dos modelos GARCH é possível mostrar que o modelo GARCH é capaz de explicar muitos desses fatos estilizados. Nessa seção serão detalhados três importantes fatos estilizados: agrupamento de volatilidades (volatility clustering), caudas pesadas e reversão à média da volatilidade (volatility mean reversion).

Agrupamento de Volatilidades

No modelo GARCH(1,1) dado em (4.5) o coeficiente b_1 encontrado é normalmente um número próximo de 0.9 para séries temporais financeiras diárias ou semanais. Dessa maneira, temos que valores grandes de σ_{t-1}^2 são seguidos por valores grandes de σ_t^2 , e valores pequenos de σ_{t-1}^2 , por sua vez, são seguidos por valores pequenos de σ_t^2 . Seguindo o mesmo raciocínio para a representação ARMA em (4.6), temos que grandes/pequenas mudanças em ϵ_{t-1}^2 são seguidas por grandes/pequenas mudanças em ϵ_t^2 .

Caudas pesadas

É fato conhecido que a distribuição de muitas séries temporais financeiras com alta freqüência geralmente possuem caudas mais pesadas que a distribuição normal. Ou seja, grandes mudanças ocorrem com mais freqüência do que em uma distribuição normal. Bollerslev (1986) descreve as condições para a existência do momento de quarta ordem de um processo GARCH(1, 1). Assumindo que o momento de quarta ordem existe, Bollerslev (1986) mostra que a curtose de um processo GARCH(1,1) é maior do que 3, que é a curtose de uma distribuição normal. He e Teräsvirta (1999a, 1999b) extenderam o resultado de Bollerslev para os modelos GARCH(p,q). Portanto, um modelo GARCH pode reproduzir as caudas pesadas observadas em séries temporais financeiras.

Reversão à média da volatilidade (Volatility Mean Reversion)

Os mercados financeiros podem passar por momentos de excessiva volatilidade, apesar disso, a tendência é que a volatilidade se estabilize em um nível de longo prazo. Conforme mostrado em (4.7), a variância de longo prazo de ϵ_t para o modelo estacionário GARCH(1,1) é igual a $a_0/(1 - a_1 - b_1)$. Nesse caso, reescrevendo a representação ARMA em (4.6), temos que a volatilidade é sempre atraída para seu nível de longo prazo:

$$\left(\epsilon_t^2 - \frac{a_0}{1 - a_1 - b_1}\right) = (a_1 + b_1)\left(\epsilon_{t-1}^2 - \frac{a_0}{1 - a_1 - b_1}\right) + u_t - b_1 u_{t-1}.$$

Na k-ésima iteração da equação acima, pode-se mostrar que

$$\left(\epsilon_{t+k}^2 - \frac{a_0}{1 - a_1 - b_1}\right) = (a_1 + b_1)^k \left(\epsilon_t^2 - \frac{a_0}{1 - a_1 - b_1}\right) + \eta_{t+k},$$

onde η_t é um processo de médias móveis. Uma vez que $a_1 + b_1 < 1$ para um modelo estacionário GARCH(1, 1), temos que $(a_1 + b_1)^k \to 0$ quando $k \to \infty$. Embora no tempo t deva haver um grande

desvio entre ϵ_t^2 e a variância de longo prazo, $\epsilon_{t+k}^2 - a_0/(1 - a_1 - b_1)$ se aproxima de zero "na média" conforme k se torna grande, ou seja, a volatilidade "reverte-se na média" para seu nível de longo prazo $a_0/(1 - a_1 - b_1)$. Por outro lado, se $a_1 + b_1 > 1$ e o modelo GARCH é não-estacionário, a volatilidade "explode" para infinito quando $k \to \infty$. Para o modelo GARCH(p,q) pode-se construir uma explicação similar.

4.4 Estimação de Modelos GARCH

O objetivo dessa seção é estimar um modelo GARCH. Conforme apresentado anteriormente, o modelo geral GARCH(p,q) é da forma

$$y_t = c + \epsilon_t$$

$$\epsilon_t = z_t \sigma_t$$

$$\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \epsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2,$$
(4.8)

para $t = 1, \dots, T$, onde $\sigma_t^2 = Var_{t-1}(\epsilon_t)$.

Serão introduzidos três tipos de estimadores para os parâmetros a_0 , $a_i e b_j$: estimador de máxima verossimilhança condicional, estimador de Whittle e estimadores de mínimos desvios absolutos. O primeiro é o estimador mais conhecido e é o que vem sendo bastante usado na indústria de operações bancárias. Os últimos são interessantes quando tratamos de erros com caudas pesadas. Nas simulações utilizaremos apenas o estimador de máxima verossimilhança condicional.

4.4.1 Estimador da Máxima Verossimilhança Condicional

Assim como na estimação dos modelos ARMA, os estimadores mais frequentemente usados para modelos ARCH/GARCH são aqueles derivados de uma função Gaussiana de máxima verossimilhança (condicional). Por exemplo, se z_t em (4.8) é normal e q = 0 (ou seja, um modelo ARCH puro), a logverossimilhança condicional negativa baseada nas observações $\epsilon_1, \dots, \epsilon_T$, ignorando-se as constantes, é igual a

$$-l(\sigma_t^2|\epsilon_1,\cdots,\epsilon_T) = \sum_{t=p+1}^T (log\sigma_t^2 + \epsilon_t^2/\sigma_t^2), \qquad (4.9)$$

4.4. ESTIMAÇÃO DE MODELOS GARCH

onde $\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \epsilon_{t-i}^2$. Os estimadores de máxima verossimilhança são obtidos minimizando-se a função em (4.9). Podemos perceber que esta função de verossimilhança é baseada na função de densidade de probabilidades condicional de $\epsilon_{p+1}, \dots, \epsilon_T$, dados $\epsilon_1, \dots, \epsilon_p$, uma vez que a função de densidade de probabilidades incondicional, que envolve a densidade conjunta de $\epsilon_1, \dots, \epsilon_p$, é difícil de ser obtida.

Para o modelo GARCH geral (ou seja, q > 0 no modelo (4.4)) a variância condicional σ_t^2 não pode ser expressa em termos de um número finito de observações passadas $\epsilon_{t-1}, \epsilon_{t-2}, \cdots$. Por indução, podemos escrever

$$\sigma_t^2 = \frac{a_0}{1 - \sum_{j=1}^q b_j} + \sum_{i=1}^p a_i \epsilon_{t-i}^2$$

$$+ \sum_{i=1}^p a_i \sum_{k=1}^\infty \sum_{j_1=1}^q \cdots \sum_{j_k=1}^q b_{j_1} \cdots b_{j_k} \epsilon_{t-i-j_1-\cdots-j_k}^2,$$
(4.10)

onde a soma múltipla desaparece quando q = 0. Nota-se que a soma múltipla acima converge com probabilidade 1 desde que cada $a_i \in b_j$ seja não negativo, e desde que o valor esperado da série múltipla seja finito. Na prática, a expressão (4.10) é substituída por uma versão truncada

Temos que, quando q = 0, $\tilde{\sigma}_t^2 = \sigma_t^2 = a_0 + \sum_{i=1}^p a_i \epsilon_{t-i}^2$. Seja $\mathbf{a} = (a_1, \dots, a_p)^{\tau}$ e $\mathbf{b} = (b_1, \dots, b_q)^{\tau}$. O estimador de máxima verossimilhança (condicional) $(\hat{a}_0, \hat{\mathbf{a}}, \hat{\mathbf{b}})$ é definido minimizando-se a expressão

$$l_{\nu}(a_0, \mathbf{a}, \mathbf{b}) = \sum_{t=\nu}^{T} (\log \tilde{\sigma}_t^2 + \epsilon_t^2 / \tilde{\sigma}_t^2), \qquad (4.12)$$

onde $\nu > p$ é um número inteiro.

Supondo que $f(\cdot)$ seja a função densidade de probabilidades de z_t conhecida, temos que os esti-

madores de máxima verossimilhança são obtidos minimizando-se

$$l_{\nu}(a_0, \mathbf{a}, \mathbf{b}) = \sum_{t=\nu}^{T} \{ log \tilde{\sigma}_t^2 - 2log f(\epsilon_t / \tilde{\sigma}_t) \}$$
(4.13)

em vez de minimizar a função (4.12).

Além da distribuição normal, algumas formas de $f(\cdot)$ frequentemente usadas são:

• distribuição t
 com ν graus de liberdade:

$$f(x) = \frac{\Gamma((\nu+1)/2)}{(\pi\nu)^{1/2}\Gamma(\nu/2)} \left(\frac{\nu}{\nu-2}\right)^{1/2} \left(1 + \frac{x^2}{\nu-2}\right)^{-\frac{(\nu+1)}{2}},$$

onde $\nu>2$ pode ser tratado como um parâmetro contínuo.

• distribuição Gaussiana generalizada:

$$\lambda = \nu \{ \lambda 2^{1+1/\nu} \Gamma(1/\nu) \}^{-1} exp\{ -\frac{1}{2} \left| \frac{x}{\lambda} \right|^{\nu} \},$$

onde
$$\lambda = \{2^{-\frac{2}{\nu}} \Gamma(\frac{1}{\nu}) / \Gamma(\frac{3}{\nu})\}^{1/2} \in 0 < \nu < 2.$$

Quando $\nu = 1$, a distribuição Gaussiana generalizada se reduz para a função $f(x) = exp\{-\sqrt{2}|x|\}/\sqrt{2}$.

Todas as distribuições acima foram normalizadas para ter média 0 e variância 1, e todas elas possuem caudas mais pesadas do que a distribuição normal.

Para comparar dois ou mais modelos ajustados, usamos os valores AIC e BIC para cada um dos modelos ajustados, que são calculados da seguinte maneira:

$$AIC = l_{\nu}(\hat{a_0}, \hat{\mathbf{a}}, \hat{\mathbf{b}}) + 2(p+q+1), \qquad (4.14)$$

 \mathbf{e}

$$BIC = l_{\nu}(\hat{a_0}, \hat{\mathbf{a}}, \hat{\mathbf{b}}) + (p+q+1)\log(T-\nu+1), \qquad (4.15)$$

onde $l_{\nu}(\cdot)$ segue de (4.13).
4.4. ESTIMAÇÃO DE MODELOS GARCH

4.4.2 Estimador de Whittle

Para o modelo GARCH(p, q) definido em (4.4), a variância condicional pode ser escrita como

$$\sigma_t^2 = a_0 / \left(1 - \sum_{j=1}^q b_j \right) + \sum_{j=1}^\infty d_j \epsilon_{t-j}^2,$$

onde $d_j \ge 0$ e $\sum_{j=1}^{\infty} d_j = \sum_{i=1}^{p} a_i / (1 - \sum_{j=1}^{q} b_j)^3$. Supondo que ϵ_t é estacionário de quarta ordem no sentido que seus primeiros quatro momentos são todos invariantes no tempo. Seja $Y_t = \epsilon_t^2$. Então Y_t é um processo AR(∞) estacionário satisfazendo

$$Y_t = a_0 / \left(1 - \sum_{j=1}^q b_j \right) + \sum_{j=1}^\infty d_j Y_{t-j} + e_t,$$
(4.16)

.

onde e_t é uma diferença de martingal

$$e_t = (z_t^2 - 1) \left\{ a_0 / \left(1 - \sum_{j=1}^q b_j \right) + \sum_{j=1}^\infty d_j Y_{t-j} \right\}$$

com $\sigma_e^2 \equiv Var(e_t) < \infty.$ Além disso, a densidade espectral do processo Y_t é dado por

$$g(\omega) = \frac{\sigma_e^2}{2\pi} \left| 1 - \sum_{j=1}^{\infty} d_j e^{i\omega j} \right|^{-2}$$

Baseados em (4.16), Giraitis e Robinson (2001) propuseram os estimadores de Whittle para a_i e b_j minimizando-se a expressão

$$\sum_{j=1}^{T-1} I_T(\omega_j) / g(\omega_j),$$

onde $I_T(\cdot)$ é o periodograma de Y_t , e $\omega_j = 2\pi j/T$ é a freqüência de Fourier. Giraitis e Robinson (2001) estabeleceram também a normalidade assintótica dos estimadores.

Para um processo AR, o estimador de Whittle é assintoticamente equivalente ao estimador de

³Para o caso geral, tem-se a equação $\sum_{j=1}^{\infty} d_j k^j = \sum_{i=1}^p a_i k^i / (1 - \sum_{j=1}^q b_j k^j)$

Yule-Walker (ver Box, Jenkins e Reinsel (1994)). Portanto, suas propriedades assintóticas dependem apenas de um número finito de autocorrelações amostrais. Dessa forma, o estimador de Whittle baseado na estrutura ARMA de um processo geral GARCH(p,q) depende de um crescente número de autocorrelações amostrais.

Mikosch e Straumann (2002) estudaram o estimador de Whittle para o modelo com caudas pesadas GARCH(1, 1) e o compararam com o estimador de máxima verossimilhança condicional. Uma vantagem do estimador de Whittle apontada em Mikosch e Straumann (2002), em contraste com o estimador de máxima verossimilhança condicional, é que o estimador de Whittle depende apenas dos valores $\epsilon_1^2, \dots, \epsilon_T^2$, ou seja, os valores não observados de σ_t^2 não precisam ser calculados. Porém, os resultados de algumas simulações detalhados a seguir mostram a superioridade do estimador de máxima verossimilhança condicional em relação ao estimador de Whittle.

Na Tabela 4.1 estão apresentados os parâmetros dos quatro modelos GARCH(1,1) considerados por Mikosch e Straumann (2002) em suas simulações. O parâmetro k representa a curtose das distribuições. Para cada um dos quatro modelos foram simuladas séries temporais de tamanho n = 250, 1000, 5000, 10000 e 20000 com 2000 réplicas independentes. Os resultados obtidos por Mikosch e Straumann (2002) estão apresentados nas Tabelas 4.2 e 4.3.

A comparação dos desvios padrão dos dois estimadores mostra valores menores para as estimativas do estimador de máxima verossimilhança condicional em todos os modelos simulados e, além disso, observa-se que os desvios padrão tendem a diminuir conforme aumentam os tamanhos das amostras. Temos ainda que as simulações apresentadas em Mikosch e Straumann (2002) indicam que o estimador de máxima verossimilhança condicional melhora sua performance à medida que os erros apresentam caudas menos pesadas, ou seja, conforme o valor de k diminui. Observa-se também que ambos os estimadores de ϕ_1 e b_1 tendem a ser viesados para a esquerda e possuem assimetria negativa. Entretanto, o viés e a assimetria diminuem quando o tamanho da amostra aumenta.

Para mais detalhes, ver Giraitis e Robinson (2001) e Mikosch e Straumann (2002).

Modelo	a_0	a_1	b_1	$\phi_1 = a_1 + b_1$	k
1	$8,58 \times 10^{-6}$	0,072	0,9	0,972	10
2	$8,58 \times 10^{-6}$	0,072	0,91	0,982	7,5
3	$8,58 imes 10^{-6}$	0,072	0,92	0,992	5
4	$8,58 \times 10^{-6}$	0,072	0,925	0,997	3,2

Tabela 4.1: Parâmetros dos modelos GARCH(1,1) das simulações de Mikosch e Straumann (2002)

Modelo		Estimador de Whittle		Estir	nador de máxi	ma verossimilhanç	a condicional		
	n	Média	Mediana	Desvio Padrão	Assimetria	Médi	a Mediana	Desvio Padrão	Assimetria
	250	0,888	0,943	0,153	-3,068	0,90	3 0,951	0,142	-3,387
	1000	0,952	0,960	0,038	-5,121	0,96	0,968	0,023	-2,049
1	5000	0,967	0,968	0,010	-1,162	0,97	0,971	0,007	-0,723
	10000	0,969	0,970	0,007	-0,690	0,97	0,972	0,005	-0,573
	20000	0,970	0,970	0,005	-0,313	0,97	2 0,972	0,003	-0,302
	250	0.895	0.947	0.153	-3.315	0.92	2 0.962	0.132	-4.190
	1000	0.963	0.971	0.040	-8.248	0.97	1 0.978	0.018	-3.146
2	5000	0.977	0,978	0.010	-7.334	0,98	0.981	0,005	-0,669
	10000	0,979	0,979	0,006	-3,533	0,98	0,982	0,004	-0,555
	20000	0,980	0,980	0,004	-0,197	0,98	2 0,982	0,003	-0,287
	250	0,911	0,954	$0,\!137$	-3,848	0,93	7 0,970	0,111	-4,346
	1000	0,975	0,980	0,025	-9,854	0,98	5 0,988	0,012	-2,631
3	5000	0,986	0,987	0,006	-1,046	0,99	0,991	0,004	-0,755
	10000	0,988	0,989	0,005	-2,679	0,99	0,992	0,002	-0,670
	20000	0,989	0,989	0,003	-0,785	0,99	2 0,992	0,002	-0,323
	250	0.918	0.959	0.127	-3.992	0,94	6 0.978	0.114	-4.801
	1000	0.980	0.985	0.024	-11.414	0.99	0.993	0.009	-1.884
4	5000	0.990	0,991	0,009	-28,023	0,99	6 0,996	0,003	-0,534
	10000	0,992	0,992	0.004	-6,173	0,99	6 0.997	0,002	-0,549
	20000	0,993	0,993	0,005	-18,779	0,99	7 0,997	0,001	-0,531

Tabela 4.2: Estimativas de ϕ_1 nos processos GARCH(1,1) via estimadores Whittle e Máxima Verossimilhança Condicional

4.4.3 Estimadores de Mínimos Desvios Absolutos

Os estimadores apresentados anteriormente são derivados da maximização da verossimilhança Gaussiana ou de uma aproximação da verossimilhança Gaussiana. Dessa forma, estes são L_2 estimadores. Sabe-se que os L_1 -estimadores são mais robustos do que os L_2 -estimadores, no que diz respeito às distribuições com caudas pesadas. Algumas evidências empíricas sugerem que este método é mais apropriado para séries financeiras pois estas apresentam caudas mais pesadas que a distribuição normal. Para detalhes ver Mandelbrot (1963), Fama (1965), Mittnik, Rachev and Paolella (1998), e Mittnik e Rachev (2000).

A idéia deste estimador implica numa reparametrização do modelo (4.4), onde $E(z_t) = 0$ e a mediana de z_t^2 , ao invés da variância de z_t , é igual a 1. Sob essa nova formulação, os parâmetros a_0 e a_i 's diferem daqueles da formulação original por um fator constante, enquanto que os parâmetros b_j 's não se alteram.

Modelo		Estimador de Whittle			Estimador de máxima verossimilhança condicional				
	n	Média	Mediana	Desvio Padrão	Assimetria	Média	Mediana	Desvio Padrão	Assimetria
	250	0,821	0,879	0,195	-2,585	0,825	0,874	0,168	-2,915
	1000	0,882	0,893	0,059	-4,586	0,889	0,894	0,034	-1,122
1	5000	0,896	0,899	0,020	-1,635	0,898	0,898	0,012	-0,246
	10000	0,898	0,900	0,014	-1,172	0,899	0,899	0,009	-0,213
	20000	0,899	0,899	0,011	-0,971	0,899	0,899	0,006	-0,049
	250		0.00 7				0.00 -	0.1 50	
	250	0,826	0,885	$0,\!194$	-2,761	0,847	0,885	0,152	-3,521
	1000	0,894	0,906	0,061	-6,097	0,902	0,905	0,028	-1,268
2	5000	0,906	0,909	0,022	-6,273	0,908	0,909	0,010	-0,153
	10000	0,909	0,910	0,016	-7,316	0,909	0,909	0,007	-0,095
	20000	0,909	0,910	0,010	-1,031	0,910	0,910	0,005	-0,140
	250	0.843	0.890	0.175	-3.126	0.864	0.894	0.131	-3,677
	1000	0.907	0.915	0,049	-8,267	0.912	0.914	0.021	-0,913
3	5000	0,918	0,920	0,019	-1,404	0,918	0,918	0,008	-0,048
	10000	0,920	0,922	0,016	-2,450	0,919	0,919	0,006	-0,115
	20000	0,920	0,921	0,013	-1,682	0,920	0,920	0,004	-0,082
	250	0.850	0.806	0 166	2 296	0.879	0.001	0 129	4 105
	200	0,650	0,890	0,100	-3,320	0,872	0,901	0,152	-4,105
	1000	0,912	0,921	0,046	-6,504	0,917	0,918	0,019	-0,443
4	5000	0,923	0,927	0,029	-17,006	0,923	0,923	0,007	-0,029
	10000	0,926	0,928	0,017	-3,526	0,924	0,924	0,005	-0,092
	20000	0,925	0,927	0,021	-9,406	0,925	0,925	0,003	0,011

Tabela 4.3: Estimativas de b_1 nos processos GARCH(1,1) via estimadores Whittle e Máxima Verossimilhança Condicional

Seja

$$\epsilon_t^2 / \sigma_t^2 = 1 + e_{t,1}, \tag{4.17}$$

onde $e_{t,1} = (z_t^2 - 1)$, que possui mediana igual a zero. Assim, temos que o estimador de mínimos desvios absolutos $\hat{\theta}_1$, que é um L_1 -estimador baseado na regressão (4.17), pode ser obtido minimizando-se

$$\sum_{t=\nu}^{T} \left| \epsilon_t^2 / \tilde{\sigma}_t^2 - 1 \right|, \tag{4.18}$$

onde $\tilde{\sigma}_t^2$ é definido em (4.11) e $\nu = p + 1$ se q = 0 e $\nu > p + 1$ se q > 0. Peng e Yao (2003) mostraram que este estimador é viesado. Para resolver este problema, Peng e Yao (2003) definiram uma forma modificada para o estimador de mínimos desvios absolutos (θ_2), que resulta da minimização da expressão:

$$\sum_{t=\nu}^{T} \left| log(\epsilon_t^2) - log(\tilde{\sigma}_t^2) \right|, \qquad (4.19)$$

4.4. ESTIMAÇÃO DE MODELOS GARCH

motivados pelo modelo de regressão

$$log(\epsilon_t^2) = log(\sigma_t^2) + e_{t,2}, \tag{4.20}$$

onde os erros $e_{t,2} = log(z_t^2)$ são i.i.d com mediana zero.

Peng e Yao (2003) demonstraram que sob condições muito amenas, os estimadores de mínimos desvios absolutos são assintoticamente normais com razão de convergência padrão $T^{1/2}$, independentemente do fato de a distribuição de z_t possuir caudas pesadas ou não. Esta é uma diferença marcante entre estes estimadores e os estimadores de máxima verossimilhança condicional derivados de (4.12), que podem apresentar convergência lenta quando z_t possui caudas pesadas.

Outro estimador proposto por Peng e Yao (2003) foi motivado pela equação de regressão

$$\epsilon_t^2 = \sigma_t^2 + e_{t,3},\tag{4.21}$$

onde $e_{t,3} = \sigma_t^2(z_t^2 - 1)$ possui mediana igual a zero. Assim, o estimador $\hat{\theta}_3$ é obtido minimizando-se:

$$\sum_{t=\nu}^{T} \left| \epsilon_t^2 - \tilde{\sigma}_t^2 \right|. \tag{4.22}$$

Intuitivamente, é preferível usar o estimador θ_2 ao θ_3 pois os termos de erros $e_{t,2}$ do modelo de regressão (4.20) são independentes e identicamente distribuídos enquanto que os erros $e_{t,3}$ do modelo (4.21) não são independentes.

Em Peng e Yao (2003) foram realizadas comparações entre os três estimadores de mínimos desvios absolutos com o estimador de máxima verossimilhança Gaussiana, $\hat{\theta}_{ML}$, para modelos simulados ARCH(2) e GARCH(1,1), tomando os erros z_t com distribuição normal padrão ou t de Student padronizada com d = 3 ou d = 4 graus de liberdade. Foram realizadas 500 réplicas de tamanho 300, com parâmetros: constante = 0, 3, $a_1 = 0, 5, a_2 = b_1 = 0, 4$ e $\nu = 20$. A Figura 4.3 apresenta os boxplots dos erros absolutos médios destas simulações. Pode-se perceber que, para modelos com erros de caudas muito pesadas, ou seja $u_t \sim t_3$, o estimador de mínimos desvios absolutos $\hat{\theta}_2$ obteve a melhor performance. Para $u_t \sim t_4$ podemos perceber que os estimadores $\hat{\theta}_2$ e $\hat{\theta}_{ML}$ possuem comportamento parecido, e são melhores que os estimadores $\hat{\theta}_1$ e $\hat{\theta}_3$. E no caso onde u_t é normal, temos que $\hat{\theta}_{ML}$ possui a melhor performance. De fato, a performance do estimador de máxima verossimilhança Gaussiana piora conforme as caudas da distribuição dos erros tendem a ser mais pesadas. Entretanto, esse comportamento não ocorre para os estimadores de mínimos desvios absolutos uma vez que estes apresentam-se mais robustos para caudas pesadas.

Para mais detalhes, ver Peng e Yao (2003).



Figura 4.3: Boxplots dos erros absolutos médios das estimativas de máxima verossimilhança, MLE, e das estimativas de mínimos desvios absolutos, $\hat{\theta}_1$, $\hat{\theta}_2 \in \hat{\theta}_3$. As distribuições dos erros indicadas por t(3), t(4) e Norm são, respectivamente, as distribuições t de Student com 3 e 4 graus de liberdade e a distribuição normal.

4.5 Diagnóstico de Modelos GARCH

As seções anteriores apresentaram métodos para estimação dos modelos GARCH. Para diagnosticar os modelos ajustados são usados basicamente dois métodos: a análise gráfica e a análise das estatísticas resumo.

Na análise das estatísticas resumo, podemos observar os erros padrão e os níveis descritivos (p valores) das estatísticas para avaliar se os coeficientes do modelo são estatisticamente iguais a zero.

4.5. DIAGNÓSTICO DE MODELOS GARCH

Além disso, existem testes variados para os resíduos padronizados $\hat{\epsilon}_t/\hat{\sigma}_t$. Para a série de retornos dos preços de ações da Petrobras, as Tabelas 4.4 e 4.5 apresentam os coeficientes ajustados por meio do método de máxima verossimilhança condicional e alguns testes para resíduos padronizados do modelo GARCH(1, 1): testes de normalidade Jarque-Bera e Shapiro-Wilks, teste de Ljung-Box para resíduos padronizados e para o quadrado dos resíduos padronizados e teste do multiplicador de Lagrange para efeitos ARCH nos resíduos padronizados.

No teste Ljung-Box para os resíduos padronizados, rejeitamos a hipótese nula H_0 da não existência de autocorrelação. Por outro lado, no teste Ljung-Box para o quadrado dos resíduos padronizados, não rejeitamos a hipótese da não existência de autocorrelação. Portanto, o modelo capturou com sucesso a estrutura de correlação serial na variância condicional mas não obtivemos um resultado satisfatório para a média condicional.

Além dos testes para autocorrelação nos resíduos padronizados, pode-se aplicar o teste do multiplicador de Lagrange para efeitos ARCH nos resíduos padronizados para verificar se ainda existe algum efeito ARCH restante. Com p-valor de 0, 8075, não rejeitamos a hipótese nula da não existência de efeitos ARCH.

Tabela 4.4. Coefficientes Estimados do modelo entren(1,1)						
Coeficiente	Valor Estimado	Erro Padrão	Valor t	Pr(> t)		
С	0,0025344	0,0006610	3,834	0,0000657		
Α	0,0000325	0,0000069	4,706	0,0000014		
ARCH(1)	0,1415773	0,0102600	13,792	0,0000000		
GARCH(1)	0,8340078	0,0127700	$65,\!289$	0,0000000		

Tabela 4.4: Coeficientes Estimados do modelo GARCH(1,1)

O modelo GARCH básico assume distribuição normal para os erros ϵ_t . Se o modelo está corretamente especificado então os resíduos padronizados ϵ_t/σ_t devem se comportar como um variável aleatória normal padrão. Os testes de normalidade Jarque-Bera e Shapiro-Wilks da Tabela 4.5 levam a conclusões opostas (p-valores 0 e 0, 7088). Para se obter uma conclusão decisiva sobre a suposição da normalidade podemos usar a ferramenta gráfica, construindo o gráfico QQ-plot⁴ (quantis da distribuição gaussiana x quantis dos resíduos padronizados). Para o modelo GARCH(1, 1) ajustado para a série de retornos dos preços das ações da Petrobras temos como resultado o gráfico QQ-plot apresentado na Figura 4.4. Podemos perceber que nas caudas há um significante desvio da linha QQ

 $^{{}^{4}}$ QQ-plot de y_t é o gráfico de dispersão dos quantis padronizados da distribuição empírica de y_t versus os quantis da distribuição normal padrão. Se y_t seguir uma distribuição normal, então os pontos do gráfico serão marcados sobre a linha de inclinação de 45 graus.

	Testes de Normalidade:		e:	
	Jarque-Bera	P-valo	or	
	772,4	0		
	Shapiro-Wilk	P-valo	r	
	0,9886	0,7088	8	
Teste Lju	ng-Box para os re	esíduos j	padronizados:	
Estatístic	a P-vale	or	$Chi^2 - d.f.$	
43,77	0,00001	672	12	
Teste Ljung-Box	z para o quadrado	o dos res	síduos padroniz	ados:
Estatística	P-valor		$Chi^2 - d.f.$	
8,515	0,7437		12	
Teste do multiplica	dor de Lagrange	para os	resíduos padro	nizados:
Estatística	P-valo	or	$Chi^2 - d.f$	
7,7085	0,807	5	12	

Tabela 4.5: Testes para diagnóstico do modelo $\operatorname{GARCH}(1,1)$ ajustado

normal, e portanto a suposição de normalidade dos resíduos não é apropriada.

Outros gráficos também podem ser utilizados para visualizar o ajuste do modelo. Por exemplo, na Figura 4.5 apresentamos a FAC dos retornos padronizados e do quadrado dos retornos padronizados. Observa-se que existe ainda pequena autocorrelação nos resíduos padronizados.

4.6 Extensões dos Modelos GARCH

Em muitos casos, o modelo GARCH básico (4.4) fornece um modelo razoavelmente bom para analisar séries temporais financeiras e para estimar a volatilidade condicional. No entanto, existem alguns aspectos do modelo que podem ser melhorados para que, assim, ocorra uma melhor captura das características e da dinâmica de uma série temporal particular. Esta seção introduz uma série de extensões do modelo GARCH básico que fazem da modelagem GARCH mais flexível.

4.6.1 Efeitos de Alavancagem Assimétrica e Impactos de Informações Externas

No modelo GARCH básico (4.4), uma vez que apenas os resíduos ao quadrado ϵ_{t-i}^2 entram na equação, os sinais dos resíduos ou choques não têm efeitos na volatilidade condicional. Entretanto,



QQ-Plot dos Resíduos Padronizados

Figura 4.4: Gráfico QQ-plot dos resíduos padronizados - model
o $\mathrm{GARCH}(1,1)$ da série de retornos Petrobras

um fato estilizado da volatilidade financeira é que más notícias (choques negativos) tendem a oferecer maior impacto na volatilidade do que boas notícias (choques positivos). Black (1976) atribui esse efeito ao fato de que más notícias tendem a abaixar os preços das ações, e portanto aumentam a alavancagem da ação causando maior volatilidade nessa ação. Baseado nessa conjectura, o impacto assimétrico das notícias, das informações externas, é normalmente chamado de *efeito de alavancagem* ou *efeito alavanca (leverage effect)*. Nesta subseção serão apresentados os modelos EGARCH, TGARCH e PGARCH que são capazes de incorporar este efeito.



Figura 4.5: FAC dos resíduos padronizados e do quadrado dos resíduos padronizados - modelo GARCH(1,1) da série de retornos Petrobras

Modelo EGARCH

Nelson (1991) propôs o seguinte modelo *GARCH exponencial* (EGARCH) para capturar os efeitos de alavancagem:

$$h_t = a_0 + \sum_{j=1}^q b_j \frac{|\epsilon_{t-j}| + \gamma_j \epsilon_{t-j}}{\sigma_{t-j}} + \sum_{i=1}^p a_i h_{t-i}, \qquad (4.23)$$

onde $h_t = \log(\sigma_t^2)$ ou $\sigma_t^2 = e^{h_t}$. Note que quando ϵ_{t-j} é positivo ou se houve "boas notícias", o efeito total de ϵ_{t-j} é $(1 + \gamma_j) |\epsilon_{t-j}|$; por outro lado, quando ϵ_{t-j} é negativo ou se houve "más notícias", o efeito total de ϵ_{t-j} é $(1 - \gamma_j) |\epsilon_{t-j}|$. Más notícias podem ter um impacto maior na volatilidade, e é esperado que o valor de γ_j seja negativo.

Outra vantagem do modelo EGARCH sobre o modelo GARCH básico é a garantia de que a variância condicional σ_t^2 é positiva, independentemente dos valores dos coeficientes de (4.23), porque é modelado o logaritmo de σ_t^2 em vez de se modelar σ_t^2 .

4.6. EXTENSÕES DOS MODELOS GARCH

Modelo TGARCH

Outra variação do modelo GARCH que é capaz de modelar o efeito alavanca é o modelo TGARCH (*threshold* GARCH), que possui a seguinte forma:

$$\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \epsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i S_{t-i} \epsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2, \qquad (4.24)$$

onde

$$S_{t-i} = \begin{cases} 1, \text{ se } \epsilon_{t-i} < 0 \text{ ``más notícias'',} \\ 0, \text{ se } \epsilon_{t-i} \ge 0 \text{ ``boas notícias''.} \end{cases}$$

Ou seja, ϵ_{t-i}^2 apresenta diferentes efeitos sobre a varância condicional σ_t^2 : quando ϵ_{t-i} é positivo, os efeitos totais são dados por $a_i \epsilon_{t-i}^2$; quando ϵ_{t-i} é negativo, os efeitos totais são dados por $(a_i + \gamma_i) \epsilon_{t-i}^2$. Dessa forma, espera-se um valor positivo para γ_i no caso de más notícias para que haja impactos maiores.

Este modelo também é conhecido como o modelo GJR, pois Glosten, Jagannathan e Runkle (1993) propuseram essencialmente o mesmo modelo.

Modelo PGARCH

O modelo GARCH básico também pode ser extendido para permitir a presença do efeito alavanca. Isto é possível se o modelo GARCH básico for tratado como um caso especial do modelo *power* GARCH (PGARCH) proposto por Ding, Granger e Engle (1993):

$$\sigma_t^d = a_0 + \sum_{i=1}^p a_i \left(|\epsilon_{t-i}| + \gamma_i \epsilon_{t-i} \right)^d + \sum_{j=1}^q b_j \sigma_{t-j}^d,$$
(4.25)

onde d é positivo, e γ_i denota o coeficiente do efeito alavanca. Nota-se que para d = 2, o modelo (4.25) se reduz ao modelo GARCH básico com efeito alavanca.

O expoente d do modelo PGARCH também pode ser fixado com algum valor diferente de 2. Por exemplo, uma escolha usual é fixar d = 1 que faz com que o modelo GARCH seja robusto para outliers. Curva de Impacto das Informações Externas

As subseções anteriores mostraram que os modelos EGARCH, TGARCH e PGARCH são todos capazes de modelar o efeito alavanca. A escolha por um modelo em particular pode ser feita por meio de um critério de seleção de modelos tal como o critério de BIC (Bayesian information criterion). Como uma alternativa, Engle e Ng (1993) propuseram que a curva de impacto das informações *externas* também pode ser usada para comparar os diferentes modelos.

Definição 4.6.1. A curva de impacto é a relação entre a variância condicional no tempo t e o termo de choque (erro) no tempo t-1, mantendo-se constantes as informações com tempo t-2 ou anterior, e com todas as variâncias condicionais passadas estimadas no mesmo nível da variância incondicional.

Para facilitar a comparação das curvas de impacto das notícias para os diferentes modelos GARCH, a Tabela 4.6 resume a variância incondicional, $\bar{\sigma}^2$, das variações do modelo GARCH e a Tabela 4.7 apresenta as curvas de impacto para os modelos quando p = 1 e q = 1.

Tabela 4.6: Variância Incondicional dos Processos GARCH					
GARCH(p,q)	$\bar{\sigma}^2 = a_0 / \left[1 - \sum_{i=1}^p a_i (1 + \gamma_i^2) - \sum_{j=1}^q b_j \right]$				
$\mathrm{TGARCH}(\mathbf{p},\mathbf{q})$	$\bar{\sigma}^2 = a_0 / \left[1 - \sum_{i=1}^p a_i + \gamma_i / 2) - \sum_{j=1}^q b_j \right]$				
PGARCH(p,q,1)	$ar{\sigma}^2 = a_0^2 / \left[1 - \sum_{i=1}^p a_i \sqrt{2/\pi} - \sum_{j=1}^q b_j \right]^2$				
EGARCH(p,q)	$\bar{\sigma}^2 = exp\left\{ \left(a_0 + \sum_{i=1}^p a_i \sqrt{2/\pi} / (1 - \sum_{j=1}^q b_j) \right\} \right\}$				

Tabela 4.7: Curvas de Impacto das Informações Externas dos Processos GARCH

GARCH(1,1)	$\sigma_t^2 = A + a_1 \left(\epsilon_{t-1} + \gamma_1 \epsilon_{t-1} \right)^2$
	$A = a_0 + b_1 \bar{\sigma}^2$
TGARCH(1,1)	$\sigma_t^2 = A + \left(a_1 + \gamma_1 S_{t-1} \epsilon_{t-1}^2\right)$
	$A = a_0 + b_1 \bar{\sigma}^2$
PGARCH(1,1)	$\sigma_t^2 = A + 2\sqrt{Aa_1} \left(\epsilon_{t-1} + \gamma_1 \epsilon_{t-1} \right) + a_1^2 \left(\epsilon_{t-1} + \gamma_1 \epsilon_{t-1} \right)^2$
	$A = (a_0 + b_1 \bar{\sigma})^2$
EGARCH(1,1)	$\sigma_t^2 = Aexp\left\{a_1\left(\epsilon_{t-1} + \gamma_1\epsilon_{t-1}\right)/\bar{\sigma}\right\}$
	$A = \bar{\sigma}^{2b_1} exp\left\{a_0\right\}$

Por exemplo, para comparar as curvas de impacto dos modelos EGARCH(1,1), TGARCH(1,1), PGARCH (1,1) e GARCH(1,1) com efeito alavanca dos retornos da série de preços das ações da

4.6. EXTENSÕES DOS MODELOS GARCH

Petrobras, apresentamos a Figura 4.6. Nesse gráfico, o intervalo de ϵ_t é determinado pelos resíduos dos modelos ajustados. O gráfico mostra que todas as curvas de impacto são assimétricas porque os efeitos alavanca estão sendo capturados pelos quatro modelos, e pode-se perceber também que choques negativos ou más notícias implicam em maiores impactos na volatilidade.

O modelo GARCH sugere maiores impactos na volatilidade do que o modelo TGARCH, que por sua vez apresenta maiores impactos do que o modelo EGARCH, independente dos valores dos choques. Além disso, observa-se que impactos na volatilidade dos valores grandes de choques ($\epsilon_t < -0, 7$ ou $\epsilon_t > 0, 14$) no modelo PGARCH são menores em comparação com os modelos EGARCH, GARCH e TGARCH. Para valores de ϵ_t próximos de zero (entre $-0, 7 \in 0, 14$), o modelo que apresenta menores impactos na volatilidade é o modelo EGARCH.



Figura 4.6: Comparação das Curvas de Impacto das Informações Externas

4.6.2 Modelo de Duas Componentes (Two Components Model)

Na seção (4.3.2) foi introduzido o fato de que o modelo GARCH pode ser usado para modelar a reversão à média na volatilidade condicional, ou seja, a volatilidade condicional sempre "reverte-se na média" para seu nível de longo prazo se o modelo GARCH é estacionário. Retomando a forma de reversão à média do modelo básico GARCH(1, 1) temos:

$$(\epsilon_t^2 - \bar{\sigma}^2) = (a_1 + b_1)(\epsilon_t^2 - \bar{\sigma}^2) + u_t - b_1 u_{t-1}.$$

onde $\bar{\sigma}^2 = a_0/(1-a_1-b_1)$ é o nível de longo prazo incondicional da volatilidade.

Engle e Lee (1999) sugeriram que a alta persistência está associada à volatilidade de longo prazo variando no tempo. Em particular, eles sugeriram a decomposição da variância condicional em duas componentes:

$$\sigma_t^2 = q_t + s_t \tag{4.26}$$

onde q_t é uma componente de longo prazo com alta persistência, e s_t é uma componente transitória de curto prazo.

A forma geral do modelo de duas componentes é baseada em uma versão modificada de Ding e Granger (1996):

$$\sigma_t^d = q_t^d + s_t^d \tag{4.27}$$

$$q_t^d = \alpha_1 \, |\epsilon_{t-1}|^d + \beta_1 q_{t-1}^d \tag{4.28}$$

$$s_t^d = a_0 + \alpha_2 |\epsilon_{t-1}|^d + \beta_2 s_{t-1}^d.$$
(4.29)

Ou seja, a componente de longo prazo q_t segue um modelo PGARCH(1,1) de alta persistência, e a componente transitória s_t segue outro modelo PGARCH(1,1). Expressando estes dois modelos PGARCH usando a notação de operador de translação para o passado

$$q_t^d = (1 - \beta_1 \mathbf{L})^{-1} \alpha_1 |\epsilon_{t-1}|^d$$
$$s_t^d = a_0 + (1 - \beta_2 \mathbf{L})^{-1} \alpha_2 |\epsilon_{t-1}|^d$$

e substituindo estas expressões em (4.27), pode-se mostrar que a forma reduzida do modelo de duas

4.6. EXTENSÕES DOS MODELOS GARCH

componentes é dada por:

$$\sigma_t^d = a_0 + (\alpha_1 + \alpha_2) |\epsilon_{t-1}|^d - (\alpha_1 \beta_2 + \alpha_2 \beta_1) |\epsilon_{t-2}|^d + (\beta_1 + \beta_2) \sigma_{t-1}^d - \beta_1 \beta_2 \sigma_{t-2}^d,$$

que é a forma de um modelo PGARCH(2, 2) limitado. Entretanto, o modelo de duas componentes não é completamente equivalente ao modelo PGARCH(2, 2) porque nem todo modelo PGARCH(2, 2)possui uma estrutura de componentes. De fato, uma vez que o modelo de componentes é uma versão limitada do modelo PGARCH(2, 2), a estimação do modelo de duas componentes é numericamente mais estável que a estimação de um modelo PGARCH(2, 2).

4.6.3 Modelo GARCH-in-the-mean

Em investimentos financeiros, espera-se que alto risco resulte em altos retornos. Apesar da teoria moderna de precificação de ações não implicar em uma relação simples, ela sugere que existem algumas interações entre retornos esperados e risco que são medidas pela volatilidade. Engle, Lilien e Robins (1987) propuseram extender o modelo GARGH básico de modo que a volatilidade condicional pudesse gerar um prêmio de risco (*risk premium*), que é parte dos retornos esperados. Este modelo GARCH extendido é usualmente chamado de modelo *GARCH-in-the-mean*, ou modelo GARCH-M.

O modelo GARCH-M extende a equação da média condicional (4.1) conforme segue:

$$y_t = c + \alpha g(\sigma_t) + \epsilon_t \tag{4.30}$$

onde $g(\cdot)$ pode ser uma função arbitrária da volatilidade σ_t . Por exemplo, algumas possíveis funções $g(\sigma_t)$ podem ser σ_t , σ_t^2 ou $\log(\sigma_t^2)$.

4.6.4 Variáveis Exógenas na Equação Geral da Média Condicional

Até o momento a equação da média condicional ficou restrita a uma constante nos modelos GARCH, com excessão do modelo GARCH-M onde a volatilidade leva em conta a equação da média como uma variável explicativa.

A forma geral da média condicional é dada por

$$y_{t} = c + \sum_{i=1}^{r} \phi_{i} y_{t-i} + \sum_{j=1}^{s} \theta_{j} \epsilon_{t-j} + \sum_{l=1}^{L} \beta_{l}' \mathbf{x}_{t-l} + \epsilon_{t}$$
(4.31)

onde \mathbf{x}_t é um vetor $k \times 1$ de variáveis exógenas, e β_t é o vetor $k \times 1$ de coeficientes.

4.6.5 Distribuição de Erros Não-Gaussianos

Até o momento, foi utilizada a suposição da distribuição normal para os erros. No entanto, é conhecido o fato de que séries temporais financeiras possuem caudas pesadas, portanto é de interesse usar distribuições que possuem caudas mais pesadas do que as da distribuição normal. Descreveremos aqui três possíveis distribuições de erros com caudas pesadas para o ajuste de modelos GARCH: a *distribuição t de Student*, a *distribuição exponencial dupla* e a *distribuição do erro generalizado*.

Distribuição t de Student

Se uma variável aleatória u_t tem distribuição t
 de Student com ν graus de liberdade e com parâmetro de escala s_t , a função densidade de probabilidade (fdp) de u_t é dada por

$$f(u_t) = \frac{\Gamma\left[\left(\nu+1\right)/2\right]}{\left(\pi\nu\right)^{1/2}\Gamma\left(\nu/2\right)} \frac{s_t^{-1/2}}{\left[1+u_t^2/\left(s_t\nu\right)\right]^{(\nu+1)/2}},$$

onde $\Gamma(\cdot)$ é a função gama. A variância de u_t é dada por:

$$Var(u_t) = \frac{s_t\nu}{\nu - 2}, \ \nu > 2.$$

Se o termo de erro ϵ_t em um modelo GARCH segue uma distribuição t
 de Student com ν graus de liberdade e $Var_{t-1}(\epsilon_t) = \sigma_t^2$, o parâmetro de escala s_t pode ser escrito como

$$s_t = \frac{\sigma_t^2 \left(\nu - 2\right)}{\nu}.$$

Portanto, a função de log verossimilhança de um modelo GARCH com distribuição t de Student para os erros pode ser facilmente construída baseada na fdp acima.

Distribuição do Erro Generalizado

Nelson (1991) propôs o uso da distribuição do erro generalizado (GED) para capturar as caudas

pesadas geralmente observadas na distribuição de séries temporais financeiras. Se a variável aleatória u_t é uma GED com média zero e variância unitária, a fdp de u_t é dada por:

$$f(u_t) = \frac{\nu \exp\left[-(1/2) |u_t/\lambda|^{\nu}\right]}{\lambda \cdot 2^{(\nu+1)/\nu} \Gamma(1/\nu)}$$

onde

$$\lambda = \left[\frac{2^{-2/\nu}\Gamma(1/\nu)}{\Gamma(3/\nu)}\right]^{1/2}$$

e ν é um parâmetro positivo que determina o comportamento das caudas dessa distribuição. Quando $\nu = 2$, a fdp acima se reduz a fdp normal padrão; quando $\nu < 2$, a densidade possui caudas mais densas do que da distribuição normal; quando $\nu > 2$, a densidade possui caudas mais leves do que da densidade normal.

Quando $\nu = 1$, a fdp de GED se reduz à fdp da distribuição exponencial dupla:

$$f(u_t) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|u_t|}$$

Baseado na fdp acima, a função de log verossimilhança do modelo GARCH com GED ou de erros com distribuição exponencial dupla pode ser facilmente construída.

4.7 Seleção e Comparação de Modelos GARCH

As seções anteriores ilustraram extensões dos modelos GARCH. Selecionar o melhor modelo para determinado conjunto de dados pode ser uma tarefa muito complicada. O diagnóstico de modelos baseado nos resíduos padronizados e curvas de impacto para os efeitos de alavancagem pode ser usado para comparar a eficácia de diferentes aspectos dos modelos GARCH. Além disso, uma vez que os modelos GARCH podem ser tratados como modelos ARMA para resíduos quadráticos, os critérios tradicionais de seleção de modelos como o critério da informação de Akaike (AIC) e o critério da informação Bayesiana (BIC) também podem ser usados para selecionar os melhores modelos.

Considerando o ajuste de um modelo GARCH(1, 1) simples e um modelo com distribuição t de Student para a série de retornos dos preços das ações da Petrobras, temos que o BIC do modelo com distribuição t de Student é menor do que o do modelo com distribuição normal, o que sugere que a distribuição t de Student seja melhor que a distribuição normal. Os valores de BIC, AIC e log da verossimilhança dos modelos ajustados são apresentados na Tabela 4.8.

Tabela 4.8: AIC, BIC e log da verossimilhança dos modelos GARCH ajustados para a série de retornos da Petrobras

	distribuição normal	distribuição t de Student
AIC	-6348	-6443
BIC	-6326	-6416
log verossimilhança	3178	3226

Podemos comparar graficamente os ajustes dos modelos. Na Figura 4.7 são apresentadas as FAC do quadrado dos resíduos padronizados dos modelos ajustados. Este gráfico sugere que os dois modelos são apropriados para modelar a volatilidade condicional. Podemos comparar também os gráficos QQ-plots dos resíduos padronizados, que estão apresentados na Figura 4.8. A análise gráfica reforça a conclusão obtida na análise dos critérios tradicionais de seleção de modelos, sugerindo também que a distribuição t de Student é melhor que a distribuição normal.



Figura 4.7: Comparação das FACs do Quadrado dos Resíduos Padronizados - (a) Distribuição Normal e (b) Distribuição t de Student



Figura 4.8: Comparação dos QQ-plots dos Resíduos Padronizados - (a) Distribuição Normal e (b) Distribuição t de Student

4.8 Predição de Modelos GARCH

Uma tarefa importante na modelagem da volatilidade condicional é a geração de previsões precisas tanto para valores futuros de uma série temporal financeira quanto para sua volatilidade condicional. Uma vez que a média condicional do modelo geral GARCH dado em (4.31) assume a forma de um ARMA tradicional, a previsão de valores futuros de uma série temporal pode ser obtida seguindo-se a abordagem tradicional para predição de modelos ARMA. Portanto, levando-se em conta também a variância condicional variando com o tempo, modelos GARCH podem gerar previsões adequadas para os valores futuros, especialmente sobre horizontes pequenos. Esta seção ilustra como prever a volatilidade usando modelos GARCH.

Para simplificar, consideremos o modelo básico GARCH(1,1):

$$\sigma_t^2 = a_0 + a_1 \epsilon_{t-1}^2 + b_1 \sigma_{t-1}^2,$$

com $t = 1, 2, \dots, T$. Para obter $E_T[\sigma_{T+k}^2]$, que é a previsão da volatilidade futura σ_{T+k}^2 , para k > 0,

dada a informação no tempo T, a partir da equação anterior temos:

$$E_T[\sigma_{T+1}^2] = a_0 + a_1 E_T[\epsilon_T^2] + b_1 E_T[\sigma_T^2] = a_0 + a_1 \epsilon_T^2 + b_1 \sigma_T^2,$$

onde ϵ_T^2 e σ_T^2 são os valores obtidos depois da estimação.
5 Para T+2 temos:

$$E_T[\sigma_{T+2}^2] = a_0 + a_1 E_T[\epsilon_{T+1}^2] + b_1 E_T[\sigma_{T+1}^2]$$

= $a_0 + (a_1 + b_1) E_T[\sigma_{T+1}^2],$

onde $E_T[\epsilon_{T+1}^2] = E_T[\sigma_{T+1}^2]$. Seguindo o mesmo raciocínio, pode-se obter a equação de previsão da volatilidade condicional:

$$E_T[\sigma_{T+k}^2] = a_0 \sum_{i=1}^{k-2} (a_1 + b_1)^i + (a_1 + b_1)^{k-1} E_T[\sigma_{T+1}^2], \qquad (4.32)$$

para $k \ge 2$. Para $k \to \infty$, a previsão da volatilidade em (4.32) se aproxima da variância incondicional $a_0/(1 - a_1 - b_1)$ se o processo GARCH for estacionário (ou seja, se $a_1 + b_1 < 1$).

O algoritmo de predição (4.32) produz uma previsão para a variância condicional σ_{T+k}^2 . A previsão para a volatilidade condicional, σ_{T+k} , é definida pela raiz quadrada da previsão de σ_{T+k}^2 .

A volatilidade prevista pode ser usada juntamente com valores previstos da série para gerar intervalos de confiança para os valores previstos da série.

4.9 Simulações

Nesta seção serão apresentados os resultados das simulações realizadas do modelo GARCH (1, 1). Todas as rotinas para simulação e estimação foram implementadas para o software S-Plus. Foram realizadas 500 simulações com tamanho de séries 1000, com constantes $a_0 = 0, 2$ e c = 0, e com os coeficientes a_1 e b_1 variando de 0,05 a 0,85. O objetivo destas simulações é avaliar o comportamento do estimador de máxima verossimilhança para os diferentes valores dos parâmetros a_1 e b_1 e, além disso, avaliarmos também o comportamento da soma destes dois parâmetros.

Para avaliar e comparar o comportamento do estimador de máxima verossimilhança para os dife-

 $^{{}^{5}}A$ notação utilizada se refere aos valores ajustados em vez de representar os valores "reais" não observados.

4.9. SIMULAÇÕES

rentes parâmetros utilizamos a média dos valores estimados, a variância estimada e o erro quadrático médio (EQM) dos parâmetros, definidos em (3.29), (3.30) e (3.31), respectivamente.

Os resultados das simulações estão apresentados nas Tabelas 4.9 e 4.10.

0,1-0,0007 0,40,00120,6-0,0014 b_1 0.20.3 0.50.70.8 0.85-0,0014 -0,0005 -0,0010 0,0026 -0,0021 -0.0010 -0.0004 С 0,1560 0,17520,1822 0,2025 0,2115 0,2306 0,2591 0,3069 0,3387 0,3391 a_0 Média a_1 0,0419 0,0449 0,0449 0,0477 0,0485 0,0494 0,0516 0,0524 0,0520 0,0545 b_1 0.25750.20470.27000.29380.36940.42900.49440.56070.6909 0.77240,2497 0,3149 0,8269 0,2994 0,3414 0,4179 0,4784 0,5460 0,6131 0,7429 $a_1 + b$ 0.00020.0002 0.0002 0.0003 0.0004 0.0006 0.0008 0.0013 0.0019 0,0176 a_0 0.01260.01530,0211 0.02760.03210.0487 0.0817 0,15600,1444 $a_1 = 0,05$ Variância 0,0013 0,0013 0,0015 0,0013 0,0012 0,0011 0,0012 0,0010 0,0008 0,0007 a_1 Amostral b_1 0.2645 $0,2905 \\ 0,2829$ 0,26210.2303 0.21460.1716 0,1553 0.12820.0942 0.0413 0.25090,24970,2238 0,16740,15190,12790,0947 0.0411 $a_1 + b$ 0,2088 0,0002 0,0002 0,0002 0,0003 0,0003 0,00040,0006 0,0008 0,0013 0,0019 0.0145 0.0159 0.0179 0.0211 0.0277 0.0330 0.0521 0.0929 0 1750 0 1635 a_0 0,0010 EQM 0,0013 0,0015 0,0012 0,0011 0,0012 0,0008 0,0007 0.0014 0.0013 a_1 0,1662 b_1 0,3070 0,3009 0,2665 0,2299 0,17630,14730,1059 0,0472 0,21510 2902 0 2923 0 2534 0 2234 0 2094 0 1722 0 1624 0 1464 0 1060 0.0463 *a* 1 + b0,0002 0,0000 -0,0003 0,0010 0,0002 0,0003 -0,0012 -0,0022 -0,0002 0,0019 C 0,1877 0,1860 0,2017 0,2032 0,2137 0,2175 0,2288 0,2346 0,2314 0,2436 a_0 Média a_1 0.0945 0.0915 0.09520.09780.0985 0,1019 0,1013 0,1017 0,1014 0,1026 0,1060 0,1640 0,1973 0,2906 0,3646 0,6621 0,7814 0,4614 0,55310,8351 b_1 0,2006 0,25550,00030,29250,00030,38850,00030,65440,00070,76380,00090,8828 0,0018 0,56330,9376 0,4631 $a_1 + b_1$ 0.0004 0.0005 0.0033 0,0072 0,0081 0,0096 0,0094 0,0131 0,0135 0,0157 0,0246 0,0109 0,0106 a_0 Variância 0,0019 0,0018 0,0019 0,0018 0,0017 0,0015 0,0013 0,0011 0,0007 0,0006 $a_1 = 0, 1$ a_1 0.0946 0.0929 0.0606 0.0049 0.0019 Amostral b_1 0.13560.14130.12780.0420 0.0306 0,1303 0,0361 0,1278 0,1220 0,0871 0,0863 0,0554 0,0268 0,0030 0,0008 $a_1 + b$ 0.00020.00030.00030.00030.00040.00050.00070.00090.0018 0.00330.0096 0.0073 0.0083 0.0094 0.0133 0.0137 0.01650.0258 0.0118 0.0125 a_0 0,0007 EQM0,0019 0,0019 0,0019 0,0018 0,0017 0,0016 0,0013 0,0011 0,0006 a_1 b_1 0.13850.14510,12750.09450.0940 0.06200,0441 0.03200,0052 0.0021 0,1218 0.0870 0,0875 0,0567 0,1007 0.0280 0,0033 0.0010 0.13010.1332 $a_1 + b_1$ -0,0011 -0,0004 -0,0005 -0,0029 0,0008 0,0004 -0,0008 0,0030 -0.0052 0.2000 0,1996 0,19780,2069 0,2047 0,2114 0,21580,21710.2509 a_0 Média a_1 0.1947 0.19370.19520.1926 0.1987 0.19850.1991 0.2001 0.2027 b_1 0,1049 0,2070 0,2858 0,3900 0,5835 0,6877 0,7910 0,0486 0,4804 0.24330.29860 4023 0 4784 0.58880.6789 0 7825 0.8878 0.9938 $a_1 + b$ 0,0002 0,0003 0,0003 0,0006 0,0008 0,0003 0,0004 0,0014 0,0071 C 0,0022 0,0027 0,0028 0,0038 0,0040 0,0042 0,0045 0,0042 0,0095 a_0 $a_1 = 0, 2$ Variância 0.0023 0.0026 0.0024 0.0026 0.0023 0.0021 0.0017 0.0014 0.0010 $a_1 \\ b_1$ 0,0345 0,0380 0,0296 0,0295 0,0221 0,0145 0,0080 0,0034 0,0008 Amostral $a_1 +$ 0,0327 0,0328 0,0249 0,0248 0,0169 0,0104 0,0052 0,0016 0,0003 0.00710.00020.00030.0003 0.0003 0.0004 0.00060.0008 0.00140,0022 0,0027 0,0027 0,0038 0,0040 0,0043 0,0047 0,0045 0,0121 a_0 EQM 0,0023 0,0026 0,0024 0,0026 0,0023 0,0021 0,0017 0,0014 0,0010 a_1 0,0296 b_1 0.03450.0379 0.0297 0,0221 0.0148 0,0083 0,0035 0,0009 0.0327 0,0327 0,0248 0,0252 0.0170 0.0108 0,0055 0.0017 0.0003 $a_1 + b$ 0,0001 -0,0001 0,0001 -0,0013 0,0012 -0,0006 -0,0020 -0,0022 a_0 0.20370.20150.20660.20310.20360.20770.20680.2198Média 0,2946 0,2966 0,2918 0,2960 0,2961 0,2962 0,2974 0,3038 a_1 b_1 0,0412 0,0981 0,1841 0,29410,3960 0,49250,5965 0,6921 0.5901 $a_1 + b$ 0.33580.39470.47590.69220.78880.89390.99590,0002 0,0003 0,0003 0,0003 0,0005 0,0007 0,0011 0,0026 0.0011 0.0015 0,0017 0,0019 0,0019 0.0021 0.0024 0.0038 a_0 a_1 0.0031 0.0032 0.0025 $a_1 = 0, 3$ Variância 0.0030 0.0028 0.0028 0.00220.00150,0029 Amostral b_1 0,0124 0,0155 0,0142 0,0122 0,0082 0,0050 0,0010 $a_1 +$ 0.01250.0146 0.0132 0.0087 0.0049 0.00320.0013 0.0004 0,0003 0,0003 0,0007 0,0026 0,0002 0,0003 0,0005 0,0011 C 0,0011 0,0015 0,0018 0,0020 0,0019 0,0022 0,0024 0,0042 a_0 EQM $a_1 \\ b_1$ 0.00310.0030 0.0033 0.0028 0.0029 0.00250.0022 0.00150,0145 0,0051 0,0124 0,0155 0,0122 0,0082 0,0029 0,0011 0,0126 0,0146 0,0137 0,0088 0,0050 0.0033 0,0013 0,0005 $a_1 +$ 0.0001 0.00130.00020.0006 -0.00160.0020-0.00170,2009 0,2000 0,2058 0,2047 0,2047 0,2059 0,2160 a_{\cap} Média 0,3945 0,3975 0,3951 0,3936 0,3962 0.3934 0,4032 a_1 b_1 0.04710,10260.18740.29300.3956 0.49920.59060,4405 0,4971 0,5849 0,6881 0,7891 0,8954 0,9938 $a_1 + b_1$ 0.0003 0.0003 0.0003 0.0004 0.0006 0.0008 0.0016 0.0007 0.0011 0.0017 0.0008 0.0012 0.0014 0.0026 a_0 0,0033 0,0040 0,0035 0,0034 0,0030 0,0030 0,0024 $a_1 = 0, 4$ Variância a_1 b_1 0,0056 0,0063 $0,0064 \\ 0,0062$ 0,0060 0,0039 0.0026 0,0013 Amostral 0.00710.0066 0.0047 0.0029 0.0016 0.0009 *a* 1 + b0,0003 0,0003 0,0003 0,0004 0,0006 0,0016 С 0.0007 0.0008 0.0011 0.0013 0.0014 0.0017 0.0029 a_0 EQM $a_1 \\ b_1$ 0.0033 0,0040 0,0035 0,0034 0,0030 0,0030 0,0024 0,0056 0,0066 0,0039 0,0026 0,0063 0,0060 0,0014 $a_1 + b_1$ 0.00720.0066 0.0065 0.0048 0.0030 0.0016 0.0009

Tabela 4.9: Estimativas da Média, Variância e EQM do parâmetro nas simulações do processo GARCH(1,1)

4.9. SIMULAÇÕES

 b_1 0,05-0,0015 0,20,00100,30,00100,4-0,0004 0.1 -0,0010 -0,0016 С 0,2035 0,2022 0,2048 0,2008 0,2024 0,2103 a_0 Média a_1 0,4997 0,4937 0,4958 0,4988 0,4982 0,5053 0.0391 0.0983 0.19510.29190.39760.4891 b_1 0,5921 0,7907 0,8958 0,9944 0,5389 0,6910 $a_1 + b_1$ 0.0002 0.0003 0.0005 0.0007 0.0011 0,0007 0,0008 0,0010 0,0012 0,0018 a_0 0,0006 $a_1 = 0, 5$ Variância 0,0046 0,0044 0,0042 0,0038 0,0035 0,0033 a_1 Amostral b_1 0,0033 0.00450.0039 $0,0037 \\ 0,0032$ $0,0029 \\ 0,0021$ 0.0016 0,0062 0,0053 0,0041 0,0015 $a_1 + b$ 0,0002 0,0003 0,0003 0,00050,0007 0,0011 a_0 0.0006 0.0007 0.0008 0.0011 0.00120.0019 EQM 0,0046 0,0044 0,0042 0,0038 0,0035 0,0033 a_1 b_1 0,0034 0,0045 0,0039 0,0038 0,0029 0,0017 0 0021 $a_1 + b_1$ 0.0063 0.0054 0.0041 0.0033 0.00150,0000 0,0001 -0,0005 0,0012 -0,0009 с 0,2013 0,2002 0,2001 0,2026 0,2053 a_0 Média a_1 $0,5966 \\ 0,0474$ 0,5937 0,5905 0.5961 0,5976 0,1012 0,2031 0,2967 0,3968 b_1 0,64400,00030,69490,00030,8928 0,99440,00080,7937 $a_1 + b_2$ 0.0004 0,0005 0,0006 0,0007 0,0013 0,0010 a_0 $a_1 = 0, 6$ Variância 0,0045 0,0044 0,0047 0,0048 0,0039 a_1 0.0022 0.0024 0.0030 0.0023 0.0017 Amostral b_1 0,0050 0,0045 0,0037 0,0029 0,0020 $a_1 + b$ 0.00030.00030.00040.00050,0008 0,0013 0,0010 0.0006 0.0007 0.0005 a_0 EQM 0,0045 0,0044 0,0048 0,0048 0,0039 a_1 $0,0022 \\ 0,0050$ b_1 0.0024 0.0030 0,0023 0.0017 0,0045 0,0037 0,0029 0,0020 $a_1 + b_1$ 0,0007 0,0003 -0,0019 0,0021 с $0,2011 \\ 0,6970$ $0,2011 \\ 0,6948$ 0,20120,69860,20370,6977 a_0 Média a_1 b_1 0,0476 0,0980 0,1963 0,2957 0,74460,00030,7928 $a_1 + b_2$ 0.89490.99340,00040,0006 0,0004 0,0006 0,0007 0,0010 a_0 $a_1 = 0, 7$ Variância 0.0050 0.0053 0.0050 0.0048 a_1 0,0014 0,0018 0,0018 b_1 0,0021 Amostral $a_1 +$ b 0,0049 0,0045 0,0035 0,0031 0.00060.00030.0004 0,0010 0,0004 0,0006 0,0007 a_0 $0,0053 \\ 0,0018$ $0,0050 \\ 0,0021$ $0,0048 \\ 0,0018$ EQM 0,0050 a_1 0,0014 b_1 0,0050 0,0045 0,0035 0,0031 $a_1 + b_1$ 0,0006 0,0000 -0,0013 0.2018 0.2011 a_0 0,2013Média 0,7943 0,7928 0,7957 a_1 b_1 0,04740,0978 0,1980 0.9936 $a_1 + b_2$ 0.84170.8906 0,0003 0,0003 0,0004 0,0005 0,0005 0,0008 a_0 $a_1 = 0.8$ 0.0058 Variância 0.0057 0.0061 a_1 Amostral b_1 0,0010 0,0013 0,0015 $a_1 + b$ 0.0056 0.0051 0.0041 0,0003 0,0003 0,0004 с 0,0005 0,0005 0,0008 a_0 EQM $0,0057 \\ 0,0010$ $a_1 \\ b_1$ 0.0061 0,0058 0,0013 0,0015 0,0056 0,0052 0,0042 $a_1 + b_1$ 0,0005 -0.00130,2014 0,2011 a_0 Média 0,8434 0,8391 a_1 $0,0479 \\ 0,8913$ 0.0982 b_1 0,9373 $a_1 + b$ 0.0003 0.0004 0.0005 $a_1 = 0,85$ Variância a_0 0,0054 0,0058 a_1 Amostral b_1 $0,0008 \\ 0,0053$ $0,0012 \\ 0,0047$ $a_1 + b_2$ 0,0003 0,0003 с 0,00050.0005 a_0 0,0059 EQM 0,0055 $b_1^{a_1}$ 0,0008 0,0012 $a_1 + b_1$ 0,0053 0.0048

Tabela 4.10: Estimativas da Média, Variância e EQM do parâmetro nas simulações do processo GARCH(1,1)

Na Figura 4.9 temos os gráficos box-plot das estimativas do parâmetro a_1 e seu comportamento conforme aumenta o valor do parâmetro b_1 . Cada quadro desta figura apresenta as estimativas para um determinado valor de a_1 , e dentro de cada quadro observa-se as variações possíveis de b_1 , de modo que a soma $a_1 + b_1$ seja no máximo 1. Ainda a respeito do parâmetro a_1 , temos a Figura 4.10 que apresenta a variância estimada e o EQM estimado das estimativas do parâmetro a_1 . Cada linha representa um valor de a_1 e o eixo das abscissas apresenta os valores possíveis de b_1 . Podemos observar um leve decaimento da variância e do EQM do parâmetro a_1 à medida que o parâmetro b_1 aumenta. Nota-se ainda que, independentemente do valor de b_1 , a variância e o EQM das estimativas são menores para valores menores do parâmetro a_1 . O estimador apresentou melhores resultados em relação ao parâmetro a_1 quando $a_1 = 0, 1$ e $b_1 = 0, 85$ ou $a_1 = 0, 05$ e $b_1 = 0, 85$.

4.9. SIMULAÇÕES



Figura 4.9: Boxplot das estimativas do parâmetro a_1 das simulações do processo GARCH(1, 1), com a_1 fixo e b_1 variando de 0,05 a 0,85



Figura 4.10: Variância estimada e EQM estimado das estimativas do parâmetro a_1 das simulações do processo GARCH(1, 1), com a_1 fixo e b_1 variando de 0,05 a 0,85

A Figura 4.11 apresenta os gráficos box-plot das estimativas do parâmetro b_1 . Analogamente à Figura 4.9, temos que cada quadro apresenta um valor fixo de b_1 . Em cada quadro observa-se as variações possíveis de a_1 até que $a_1 + b_1 = 1$. Além dos box-plot, podemos observar a Figura 4.12 para analisar o comportamento das estimativas de b_1 . Cada linha representa um valor de b_1 e o eixo das abscissas apresenta os valores possíveis de a_1 .

Observa-se que o estimador de b_1 possui melhor performance para valores maiores do parâmetro a_1 . Os box-plots da Figura 4.11 mostram que, quanto maior o valor de a_1 , as estimativas de b_1 são mais próximas do valor real e possuem menor variabilidade. Fixado um valor de b_1 , podemos perceber forte tendência decrescente da variância e do EQM deste parâmetro, apresentados na Figura 4.12, conforme aumenta-se o valor de a_1 .

O melhor dos casos de estimação do parâmetro b_1 ocorre quando $a_1 = 0,85$ e $b_1 = 0,05$, confirmando a melhor performance para o maior valor de a_1 possível

4.9. SIMULAÇÕES



Figura 4.11: Boxplot das estimativas do parâmetro b_1 das simulações do processo GARCH(1, 1), com b_1 fixo e a_1 variando de 0,05 a 0,85



Figura 4.12: Variância estimada e EQM estimado das estimativas do parâmetro b_1 das simulações do processo GARCH(1, 1), com a_1 fixo e b_1 variando de 0,05 a 0,85

Finalmente, temos as figuras que representam o comportamento da soma dos parâmetros $a_1 e b_1$. A Figura 4.13 apresenta os box-plot e a Figura 4.14 apresenta a variância e o EQM das estimativas da soma dos dois parâmetros. No geral, fixado um valor de b_1 , observa-se melhores resultados para valores maiores de a_1 . E, se fixado um valor de a_1 , tem-se melhores resultados à medida que b_1 aumenta. Ou seja, quanto maior a soma, melhor a performance do estimador. Os melhores resultados são obtidos quando o valor de a_1 é pequeno e o valor de b_1 é grande, de modo que a soma se aproxime de 1 ($a_1 = 0, 2 e b_1 = 0, 8$ ou $a_1 = 0, 1 e b_1 = 0, 85$).

4.9. SIMULAÇÕES



Figura 4.13: Boxplot das estimativas da soma dos parâmetros $a_1 e b_1$ das simulações do processo GARCH(1, 1), com a_1 fixo e b_1 variando de 0,05 a 0,85



Figura 4.14: Variância estimada e EQM estimado das estimativas da soma dos parâmetros a_1 e b_1 das simulações do processo GARCH(1, 1), com a_1 fixo e b_1 variando de 0,05 a 0,85

Capítulo 5

Modelos GARCH com Memória Longa

5.1 Introdução

É um fato já conhecido que os retornos de ações determinados pelo mercado financeiro especulativo são aproximadamente não correlacionados mas que não são independentes ao longo do tempo, uma vez que a maioria das séries de retornos apresenta grandes variações na volatilidade. Da mesma forma que existem estudos visando modelar adequadamente a dependência de longo prazo na média condicional de séries temporais econômicas e financeiras, questões sobre a modelagem da variância condicional destas séries tornaram-se relevantes.

Estudos recentes realizados por de Lima, Breidt e Crato (1994), Dacorogna et al. (1993), Ding, Granger e Engle (1993), e Harvey (1993) registram a presença de memória longa nas autocorrelações dos quadrados dos retornos ou dos retornos absolutos de variadas séries de preços de ações financeiras.

O Capítulo 3 mostrou como o modelo ARFIMA pode ser usado diretamente para modelar o comportamento de memória longa observado na volatilidade de retornos de ações financeiras, dada a existência da série temporal que representa a volatilidade. Entretanto, pode ser difícil encontrar uma estimativa confiável para a volatilidade, ou ainda, pode ser de interesse modelar a dinâmica dos retornos juntamente com sua volatilidade.

Esta seção mostra como os modelos GARCH podem ser estendidos para capturar diretamente o comportamento de memória longa e a alta persistência na volatilidade. Serão introduzidos os processos FIGARCH (*Fractionally Integrated Generalized Autoregressive Conditionally Heteroskedastic*) para a variância condicional que são capazes de explicar e representar a dependência temporal observada na volatilidade do mercado financeiro.

5.2 Modelos GARCH com Memória Longa

5.2.1 Modelo ARCH(∞)

Definiu-se a classe de modelos ARCH(p) em (4.1) e (4.2) na seção 4.2 do Capítulo 4. Os processos $ARCH(\infty)$, por sua vez, são definidos como:

$$y_t = c + \epsilon_t$$

$$\epsilon_t = z_t \sigma_t$$

$$\sigma_t^2 = a_0 + \sum_{j=1}^{\infty} a_j \epsilon_{t-j}^2,$$
(5.1)

onde

$$a_0 > 0, \ a_j > 0 \ (j \ge 1), \ \sum_{j=1}^{\infty} a_j < \infty,$$
 (5.2)

e z_t é uma variável aleatória *iid*.

Os processos $ARCH(\infty)$, que são extensões do processo ARCH(p) em (4.1) e (4.2) e do processo GARCH(p,q) em (4.1) e (4.4), foram considerados por Robinson (1991) como uma classe de alternativas paramétricas em testes para independência serial de y_t .

Para capturar o comportamento de memória longa observado nas autocorrelações dos quadrados dos retornos de séries temporais financeiras, Robinson (1991) considerou memória longa na formulação (5.1) de ϵ_t^2 , tomando $a_0 = 0$ (contrariando as condições especificadas em 5.2) e a_j como sendo os pesos autorregressivos de um processo integrado fracionário, com a imposição $\sum_{j=1}^{\infty} a_j = 1$.

Sob a condição $a_0 > 0$, Giraitis, Kokoszka e Leipus (2000) propuseram que $\sum_{j=1}^{\infty} a_j < 1$ para que a_j 's sejam consistentes com a estacionariedade de ϵ_t .

Para detalhes, ver Robinson (1991), Robinson e Zaffaroni (2006), Granger e Ding (1995) e Ding e Granger (1996).

5.2.2 Modelo LARCH

As condições estabelecidas por Giraitis, Kokoszka e Leipus (2000) para o modelo ARCH(∞) a fim de garantir estacionariedade não consideram memória longa nas autocorrelações de ϵ_t^2 . Desta forma,

5.2. MODELOS GARCH COM MEMÓRIA LONGA

a captura do potencial de memória longa do processo $ARCH(\infty)$ definido em (5.1) é perdida.

Podemos considerar o modelo alternativo LARCH (*Linear ARCH*), dado por:

$$\epsilon_t = z_t \sigma_t$$

$$\sigma_t = a_0 + \sum_{j=1}^{\infty} a_j \epsilon_{t-j}, \ t \in \mathbb{Z} .$$
(5.3)

Temos, então, que a variância condicional σ_t^2 é dada por:

$$\sigma_t^2 = (a_0 + \sum_{j=1}^{\infty} a_j \epsilon_{t-j})^2, \ t \in \mathbb{Z} .$$
(5.4)

Robinson (1991) comparou as implicações no comportamento do terceiro momento de ϵ_t sob as condições (5.1) e (5.4).

Observa-se também que (5.4), em contraste com (5.1), não possui a restrição da não negatividade, desta forma σ_t não é considerado um desvio padrão e perde-se um pouco da intepretação usual de volatilidade.

Giraitis, Robinson e Surgailis (2000) exploraram a estrutura de σ_t e seu possível comportamento de memória longa e apresentam condições para alta persistência na autocorrelação de ϵ_t^2 .

5.2.3 Modelo FIGARCH

Na seção 4.3 do Capítulo 4 mostrou-se que um modelo GARCH(1, 1) pode ser escrito como um modelo ARMA(1, 1) em termos dos resíduos quadráticos. Da mesma forma, o modelo GARCH(p, q)

$$\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \epsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2,$$

pode ser reescrito como:

$$\phi(L)\epsilon_t^2 = a_0 + b(L)u_t, \tag{5.5}$$

onde

$$u_t = \epsilon_t^2 - \sigma_t^2,$$

CAPÍTULO 5. MODELOS GARCH COM MEMÓRIA LONGA

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_m L^m,$$

$$b(L) = 1 - b_1 L - b_2 L^2 - \dots - b_a L^q,$$

com $m = \max(p, q)$ e $\phi_i = a_i + b_i$. A equação (5.5) representa um processo ARMA(m, q) em termos dos resíduos quadráticos ϵ_t^2 onde u_t possui média zero e é não correlacionado serialmente. Portanto, o processo u_t pode ser interpretado como "inovações" para a variância condicional.

A alta persistência nos modelos GARCH sugere que o polinômio $\phi(z) = 0$ tenha raízes unitárias, e nesse caso o modelo GARCH se torna o modelo GARCH integrado (IGARCH). Para levar em conta alta persistência e memória longa na variância condicional, extende-se o processo ARMA(m, q) em (5.5) para um processo ARFIMA(m, d, q) conforme segue:

$$\phi(L)(1-L)^{d}\epsilon_{t}^{2} = a_{0} + b(L)u_{t}, \qquad (5.6)$$

onde todas as raízes de $\phi(z) = 0$ e b(z) = 0 estão fora do círculo unitário. Quando d = 0, este processo se reduz ao modelo GARCH usual; quando d = 1, torna-se o modelo IGARCH; quando 0 < d < 1, os resíduos quadráticos de diferença fracionária, $(1 - L)^d \epsilon_t^2$, seguem um processo ARMA(m, q). O processo (5.6) para ϵ_t^2 pode ser reescrito em termos da variância condicional σ_t^2 :

$$b(L)\sigma_t^2 = a_0 + [b(L) - \phi(L)(1-L)^d]\epsilon_t^2.$$
(5.7)

Baillie, Bollerslev e Mikkelsen (1996) se referem ao modelo (5.7) como o modelo GARCH integrado fracionário (FIGARCH(m, d, q)). Quando 0 < d < 1, os coeficientes em $\phi(L)$ e b(L) capturam a dinâmica de curto prazo da volatilidade, enquanto que o parâmetro de diferença fracionária d modela as características de longo prazo da volatilidade.

A variância condicional de ϵ_t pode ser escrita como:

$$\sigma_t^2 = a_0 b(L)^{-1} + \left\{ 1 - b(L)^{-1} \phi(L) (1-L)^d \right\} \epsilon_t^2$$

$$\equiv a_0 b(L)^{-1} + \lambda(L) \epsilon_t^2, \qquad (5.8)$$

onde $\lambda(L) = \lambda_1 L + \lambda_2 L^2 + \cdots$.

Para que o processo FIGARCH(m, d, q) em (5.6) seja bem definido e para que a variância condicional seja positiva para todo t, todos os coeficientes de (5.8), que é a representação ARCH do

92

5.2. MODELOS GARCH COM MEMÓRIA LONGA

processo FIGARCH, devem ser não negativos, ou seja, $\lambda_k \geq 0$ para $k=1,2,\cdots.$

De (5.8) temos que

$$\lambda(L) = 1 - b(L)^{-1} \phi(L) (1 - L)^d.$$
(5.9)

O operador de diferença fracionária, $(1-L)^d$, definido em (3.2.2), pode ser expandido da seguinte forma

$$(1-L)^{d} = \sum_{k=0}^{\infty} {\binom{d}{k}} (-1)^{k} L^{k}$$

= $1 - dL + \frac{1}{2!} d(d-1)L^{2} - \frac{1}{3!} d(d-1)(d-2)L^{3} + \cdots$
= $1 - \sum_{k=1}^{\infty} \delta_{d,k} L^{k}$
= $1 - \delta_{d}(L),$

onde

$$\delta_{d,0} = -1$$
$$\delta_{d,1} = d,$$

 \mathbf{e}

$$\delta_{d,k} = \delta_{d,k-1} \frac{k-1-d}{k}, \text{ para } k \ge 2.$$
 (5.10)

Os pesos λ_k podem ser obtidos a partir da igualdade (5.9). Multiplicando ambos os lados por b(L), substituindo $(1-L)^d$ por $1-\delta_d(L)$ e fazendo $\phi(L) = b(L)+a(L)$, onde $a(L) = -a_1L-a_2L^2-\cdots-a_pL^p$, temos

$$\lambda(L)b(L) = b(L) - \phi(L)(1-L)^{d}$$

= $b(L) - (b(L) + a(L))(1 - \delta_{d}(L))$
= $-a(L) + b(L)\delta_{d}(L) + a(L)\delta_{d}(L).$ (5.11)

Seja(I)o lado esquerdo da igualdade (5.11)
e(II)o lado direito. Para o processo FIGARCH
(1, d, 1)temos

$$(I) = \lambda_1 - \lambda_1 b_1 L^2 + \lambda_2 L^2 - \lambda_2 b_1 L^3 - \cdots,$$

$$(II) = a_1L + (1 - b_1L)\delta_d(L) - a_1L\delta_d(L)$$

= $a_1L + \delta_d(L) - b_1L\delta_d(L) - a_1L\delta_d(L)$
= $(a_1 + \delta_{d,1})L + (\delta_{d,2} - b_1\delta_{d,1} - a_1\delta_{d,1})L^2 + (\delta_{d,3} - b_1\delta_{d,2} - a_1\delta_{d,2})L^3 + \cdots$

Portanto, os pesos λ_k do processo FIGARCH(1, d, 1) são dados por

$$\lambda_1 = a_1 + \delta_{d,1},$$

$$\lambda_k = b_1 \lambda_{k-1} + \delta_{d,k} - (a_1 + b_1) \delta_{d,k-1}, \text{ para } k \ge 2.$$

Analogamente, os coeficientes λ_k do modelo FIGARCH(1, d, 0) são dados por

$$\lambda_1 = a_1 + \delta_{d,1},$$

 $\lambda_k = \delta_{d,k} - (a_1)\delta_{d,k-1}, \text{ para } k \ge 2.$

E, para o modelo FIGARCH(0, d, 1) temos

 $\lambda_1 = \delta_{d,1},$

$$\lambda_k = b_1 \lambda_{k-1} + \delta_{d,k} - (b_1) \delta_{d,k-1}, \text{ para } k \ge 2.$$

Considerando que $a_0 > 0$, as condições

$$b_1 - d \le \phi_1 \le \frac{(2-d)}{3},\tag{5.12}$$

e

$$d\left[\phi_1 - \frac{(1-d)}{2}\right] \le b_1(\phi_1 - b_1 + d), \tag{5.13}$$

são suficientes para garantir que a variância condicional do modelo FIGARCH(1, d, 1) seja positiva para todo t.

Em Baillie et. al (1996) são apresentadas as condições $a_0 > 0$ e

$$0 \le b_1 < d \le 1, \tag{5.14}$$

94
como necessárias e suficientes para garantir que a variância condicional no modelo FIGARCH(1, d, 0) seja positiva.

Para mais detalhes a respeito dos pesos λ_k no modelo geral FIGARCH(m, d, q), ver Lopes e Mendes (2006).

5.2.4 Reparametrização do Modelo FIGARCH

Tomando a especificação do modelo FIGARCH para ϵ_t^2 em (5.6), temos que o termo constante a_0 possui estrutura diferente do termo constante μ do modelo ARFIMA (3.20). O operador de diferença fracionária $(1 - L)^d$ é aplicado à μ mas não é aplicado à a_0 . Para evitar esta discrepância na parametrização, Chung (1999) sugeriu reescrever o modelo GARCH(p, q) (5.5) da seguinte maneira:

$$\phi(L)(\epsilon_t^2 - \bar{\sigma}^2) = b(L)u_t, \tag{5.15}$$

e então redefinir o modelo FIGARCH como

$$\phi(L)(1-L)^d(\epsilon_t^2 - \bar{\sigma}^2) = b(L)u_t, \tag{5.16}$$

onde $\bar{\sigma}^2$ é a variância incondicional de ϵ_t descrita em (4.7).

Breidt, Crato e de Lima (1998) e Teyssière (1996) também consideraram esta parametrização particular.

A relação entre o parâmetro a_0 de (5.6) e o parâmetro $\bar{\sigma}^2$ fica dada por:

$$a_0 = \phi(L)(1-L)^d \bar{\sigma}^2. \tag{5.17}$$

Dada esta nova parametrização, a variância condicional derivada do modelo FIGARCH (5.16) pode ser expressa conforme segue:

$$b(L)\sigma_t^2 = b(L)\epsilon_t^2 - \phi(L)(1-L)^d(\epsilon_t^2 - \bar{\sigma}^2).$$
(5.18)

Segundo Chung (1999), há ainda três diferenças estruturais entre a definição do processo ARFIMA (3.20) para a média condicional e a definição do modelo FIGARCH (5.16) para a variância condicional:

- 1. O modelo ARFIMA(p, 0, q) se reduz ao modelo ARMA(p, q) enquanto que o modelo FIGARCH (p, 0, q) não é exatamente o modelo GARCH(p, q).
- 2. Para garantir que o modelo FIGARCH seja estritamente estacionário o parâmetro de diferença fracionária deve ser tal que $0 \le d < 1$. Portanto, o intervalo de variação dos parâmetros de diferença fracionária para os dois modelos são diferentes, fazendo com que o grau de persistência permitida na variância condicional seja maior que no caso da média condicional.
- 3. Enquanto não há restrições em relação ao sinal no modelo ARFIMA para a média condicional, os parâmetros do modelo FIGARCH necessitam de restrições adicionais para garantir que as variâncias condicionais resultantes sejam todas não negativas.

A simulações realizadas por Chung (1999) demonstram propriedades para pequenas amostras da nova especificação para o processo FIGARCH. As estimações de máxima verossimilhança dos parâmetros parecem seguir a teoria assintótica padrão, com exceção do parâmtro $\bar{\sigma}^2$. Entretanto, os resultados das simulações indicam que o estimador de $\bar{\sigma}^2$ no modelo FIGARCH apresenta razão de convergência lenta, fenômeno também conhecido para o estimador da média condicional no modelo ARFIMA.

Para mais detalhes ver Chung (1999).

5.2.5 Modelo FIEGARCH

O modelo FIGARCH extende diretamente a representação ARMA do modelo GARCH em termos dos resíduos quadráticos para um modelo fracionário integrado. Entretanto, é complicado garantir que um modelo geral FIGARCH seja estacionário e que a variância condicional σ_t^2 seja sempre positiva, sendo necessária a imposição de restrições difíceis de serem estabelecidas para os coeficientes do modelo. Para detalhes, ver Baillie, Bollerslev e Mikkelsen (1996) ou Bollerslev e Mikkelsen (1996).

Observando que o modelo EGARCH pode ser representado por um processo ARMA em termos do logaritmo da variância condicional, garantindo que a variância condicional seja sempre positiva, Bollerslev e Mikkelsen (1996) propuseram o modelo *EGARCH fracionário integrado* (FIEGARCH):

$$\phi(L)(1-L)^d \ln \sigma_t^2 = a_0 + \sum_{j=1}^q (b_j |x_{t-j}| + \gamma_j x_{t-j}),$$
(5.19)

5.2. MODELOS GARCH COM MEMÓRIA LONGA

onde $\phi(L)$ é o mesmo definido para o modelo FIGARCH, $\gamma_j \neq 0$ leva em conta os efeitos de alavancagem e x_t é o resíduo padronizado:

$$x_t = \frac{\epsilon_t}{\sigma_t}.\tag{5.20}$$

Quando d = 0, este processo se reduz ao modelo EGARCH usual e ao modelo IGARCH quando d = 1. Analogamente à classe de modelos ARFIMA para a média condicional, $ln(\sigma_t^2)$ é estacionária e invertível para d entre $-0, 5 \in 0, 5$.

É importante observar que os parâmetros do modelo FIEGARCH não precisam satisfazer a condição de não-negatividade para que o modelo seja bem definido, ao contrário da formulação do processo FIGARCH apresentada anteriormente.

5.2.6 Modelo LMGARCH

A classe de processos LMGARCH (*long-memory GARCH*) pertencente à família de processos com heteroscedasticidade condicional também é capaz de modelar o comportamento de memória longa de séries temporais financeiras. Estes processos são muito relacionados com os processos FI-GARCH descritos anteriormente e compartilham de algumas características dos processos ARFIMA. Em particular, oscilações na variância condicional de um processo LMGARCH tendem a zero com uma taxa hiperbólica lenta, ao invés de decair rapidamente a uma taxa geométrica como ocorre com os processos GARCH fracamente estacionários.

O processo GARCH pode ser generalizado se ϵ_t^2 for expresso na forma da equação:

$$\epsilon_t^2 = \omega + \Omega(L)u_t, \tag{5.21}$$

para $\omega \in (0,\infty), u_t = \epsilon_t^2 - \sigma_t^2, \ \Omega(L) = \sum_{j=0}^{\infty} \omega_j L^j, \ 0 < \sum_{j=0}^{\infty} \omega_j^2 < \infty.$

Uma importante parametrização finita de $\Omega(L)$ para considerar memória longa neste processo é

$$\Omega(L) = \frac{b(L)}{a(L)(1-L)^d},$$
(5.22)

para $d \in (0, 1/2), a(L) = 1 - a_1L - a_2L^2 - \cdots - a_qL^p \in b(L)$ conforme definido anteriormente.

Um processo ϵ_t que atenda as especificações (5.21) e (5.22) é chamado de LMGARCH(p, d, q).

Karanasos, Psaradakis e Sola (2001) apresentam algumas propriedades para a função de autocor-

relação do processo LMGARCH.

5.3 Estimação de Modelos GARCH com Memória Longa

Na seção anterior apresentamos alguns modelos para capturar o comportamento de memória longa na volatilidade de séries temporais financeiras. Nesta seção serão introduzidos métodos de estimação dos parâmetros para os processos FIGARCH e FIEGARCH.

Dadas as formulações iterativas da variância condicional apresentadas em (5.7) e (5.19), as estimativas dos coeficientes dos modelos FIGARCH e FIEGARCH podem ser obtidas por estimação de máxima verossimilhança se os resíduos seguirem uma distribuição condicional normal.

Para o processo FIGARCH(p, d, q), com amostra $\epsilon_1, \epsilon_2, \dots, \epsilon_T$, as estimativas de máxima verossimilhança podem ser obtidas maximizando-se a expressão:

$$\log L(\theta; \epsilon_1, \epsilon_2, \cdots, \epsilon_T | I_0) = -0.5T \log(2\pi) - 0.5 \sum_{t=1}^T [\log(\sigma_1^2) + \epsilon_1^2 \sigma_t^{-2}],$$
(5.23)

onde $\theta' \equiv (a_0, d, a_1, \dots, a_p, b_1, \dots, b_q)$ e I_0 são as condições iniciais usadas para maximizar a função de variância condicional.

Em muitas aplicações com dados financeiros de alta freqüência a condição de inovações com distribuição normal padrão, $z_t = \epsilon_t \sigma_t^{-1}$, não é verdadeira. Entretanto, segundo Weiss (1986) e Bollerslev e Wooldridge (1992), na obtenção de $\hat{\theta}$, que é a estimativa de Quasi Máxima Verossimilhança resultante de (5.23), válidas inferências assintóticas podem ser baseadas em versões robustas dos testes estatísticos padrões. Em particular, uma matriz de covariância robusta para as estimativas dos parâmetros é consistentemente estimada por $A(\hat{\theta})^{-1}B(\hat{\theta})A(\hat{\theta})^{-1}$, onde $A(\hat{\theta}) \in B(\hat{\theta})$ são a matriz Hessiana e o produto externo dos gradientes, respectivamente, calculados para $\hat{\theta}$.

Outra complicação existente na estimação de modelos do tipo ARCH diz respeito ao tratamento adequado das condições iniciais, I_0 . Uma abordagem para os modelos fracionários integrados pode ser baseada na representação ARCH do modelo FIGARCH(p, d, q) em (5.8), com os valores pré-amostrais de ϵ_t^2 para $t = 0, -1, -2, \cdots$, e com σ_t^2 fixado igual à variância amostral incondicional. Além disso, esta abordagem necessita de um truncamento do polinômio infinito (5.8). Uma vez que o operador de diferença fracionária tem a função de capturar o comportamento de memória longa do processo, truncar a um lag baixo pode destruir dependências de longo prazo importantes. Baillie, Bollerslev e Mikkelsen (1996) ou Bollerslev e Mikkelsen (1996) utilizaram lag de truncamento igual a 1000 em suas simulações.

5.4 Diagnóstico de Modelos GARCH com Memória Longa

Para diagnosticar os modelos FIGARCH e FIEGARCH ajustados, utilizam-se as mesmas ferramentas de análise dos modelos GARCH apresentadas na seção 4.5: análise gráfica e análise das estatísticas resumo.

Para a série de retornos dos preços de ações da Petrobras, as Tabelas 5.1 e 5.2 apresentam os coeficientes ajustados por meio do método de máxima verossimilhança condicional e alguns testes para resíduos padronizados do modelo FIGARCH(1, d, 1): testes de normalidade Jarque-Bera e Shapiro-Wilks, teste de Ljung-Box para resíduos padronizados e para o quadrado dos resíduos padronizados e teste do multiplicador de Lagrange para efeitos ARCH nos resíduos padronizados. Os resultados para o modelo FIEGARCH(1, d, 1) estão apresentados nas Tabelas 5.3 e 5.4.

Observamos que, para ambos os modelos (FIGARCH e FIEGARCH), nos testes Ljung-Box para os resíduos padronizados, rejeitamos a hipótese nula H_0 da não existência de autocorrelação nos resíduos padronizados. Por outro lado, nos testes Ljung-Box para o quadrado dos resíduos padronizados, não rejeitamos a hipótese nula. Portanto, os modelos capturam com sucesso a estrutura de correlação serial nas variâncias condicionais, porém não capturam a estrutura de correlação serial nas médias condicionais.

Além disso, nos testes do multiplicador de Lagrange para efeitos ARCH nos resíduos padronizados, não rejeitamos a hipótese nula com p-valores de 0,9973 e 0,9834 para os modelos FIGARCH e FIEGARCH, respectivamente. Assim, não restaram efeitos ARCH nos resíduos padronizados.

Tabela 5.1: Coeficientes Estimados (FIGARCH $(1, d, 1)$)							
Coeficiente	Valor Estimado	Erro Padrão	Valor t	Pr(> t)			
C	0,00235754	0,00066423	$3,\!5493$	1,991e-004			
А	0,00006311	0,00001796	$3,\!5134$	2,278e-004			
GARCH(1)	$0,\!17916869$	$0,\!15252055$	$1,\!1747$	1,201e-001			
ARCH(1)	-0,04169004	$0,\!11869763$	-0,3512	$3,\!627\mathrm{e}{-}001$			
d	$0,\!39679840$	$0,\!05885819$	6,7416	$1,\!116e-011$			

	:	Testes de Norr	nalidade	<u> </u>	
		Jarque-Bera	P-valor	 •	
		1042	0		
		Shapiro-Wilk	P-valor	•	
		$0,\!9873$	$0,\!4351$		
	:			=	
	Teste Ljung	g-Box para os re	esíduos p	adronizados:	
	Estatística	P-valo	or ($Chi^2 - d.f.$	
	42,46	0,00002	783	12	
Te	ste Ljung-Box j	para o quadrado	o dos resí	duos padroniz	ados:
Es	tatística	P-valor		$Chi^2 - d.f.$	
	2,9	0,9962		12	
Teste	do multiplicad	or de Lagrange	para os i	resíduos padro	nizados:
Estat	ística	P-valo	or	$Chi^2 - d.f$	
2,6	699	0,997	3	12	

Tabela 5.2: Testes para diagnóstico do modelo FIGARCH ajustado

Tabela 5.3: Coeficientes Estimados (FIEGARCH(1, d, 1))

Coeficiente	Valor Estimado	Erro Padrão	Valor t	Pr(> t)
С	0,002866	0,0005837	4,910	5,046e-007
А	-0,269708	0,0259481	-10,394	0,000e+000
GARCH(1)	0,388895	$0,\!1496727$	$2,\!598$	4,730e-003
ARCH(1)	0,319076	0,0350552	9,102	0,000e+000
d	0,614067	0,0605971	$10,\!134$	0,000e+000

	P	0			
		Testes de Norr	:		
		Jarque-Bera	P-valor	 	
		$869,\!8$	0		
		Shapiro-Wilk	P-valor	• -	
		$0,\!9887$	0,7387		
	:			=	
	Teste Ljung	g-Box para os re	esíduos p	adronizados:	
	Estatística	P-valo	or	$Chi^2 - d.f.$	
	40,99	0,00004	914	12	
Teste	e Ljung-Box j	para o quadrado	o dos resi	íduos padroniz	ados:
Estat	tística	P-valor		$Chi^2 - d.f.$	
4,4	411	0,9748		12	
Teste de	o multiplicad	or de Lagrange	para os i	resíduos padro	nizados:
Estatíst	ica	P-valo	or	$Chi^2 - d.f$	•
4,004	:	0,9834	4	12	

Tabela 5.4:	Testes par	a diag	nóstico	do	modelo	FIEGARC	'H a	aiustado
100010 011		a arang	11000100	~~~	1110 01010			

Graficamente, para se obter uma análise adicional sobre a suposição da normalidade nos resíduos podemos construir o gráfico QQ-plot (quantis da distribuição gaussiana x resíduos padronizados). Nos gráficos QQ-plot apresentados na Figura 5.1, pode-se perceber que nas caudas há um significante desvio da linha QQ normal, e portanto a suposição de normalidade dos resíduos não é apropriada. Além destes gráficos, apresentamos na Figura 5.2 a FAC dos resíduos padronizados e do quadrado dos resíduos padronizados para os dois modelos ajustados às séries de retornos dos preços de ações da Petrobras. Observa-se que existe ainda pequena autocorrelação nos resíduos padronizados.



Figura 5.1: Gráfico QQ-plot dos resíduos padronizados - modelos (a) FIGARCH(1, d, 1) e (b)

FIEGARCH(1, d, 1)

QQ-Plot dos resíduos padronizados



Figura 5.2: FAC dos resíduos padronizados e do quadrado dos resíduos padronizados - modelos (a) FIGARCH(1, d, 1) e (b) FIEGARCH(1, d, 1)

5.5 Seleção e Comparação de Modelos GARCH e GARCH com Memória Longa

Na seção anterior apresentamos algumas ferramentas para diagnóstico de modelos FIGARCH e FIEGARCH que podem ser usadas também para selecionar o melhor modelo para um determinado conjunto de dados. Assim como exemplificado no Capítulo 4, os critérios tradicionais de seleção de modelos como o critério da informação de Akaike (AIC) e o critério da informação Bayesiana (BIC) também podem ser usados na seleção dos melhores modelos.

No capítulo anterior comparamos os modelos GARCH(1, 1) simples e com distribuição t de Student ajustados para a série de retornos dos preços das ações da Petrobras. Vamos comparar os ajustes destes dois modelos com os modelos EGARCH, FIGARCH e FIEGARCH. Os valores BIC, AIC e log da verossimilhança dos modelos ajustados são apresentados na Tabela 5.5.

Tabela 5.5: AIC, BIC e log da verossimilhança dos modelos GARCH e GARCH com memória longa ajustados para a série de retornos da Petrobras

	GARCH (dist. normal)	GARCH (dist. t de Student)	EGARCH	FIGARCH	FIEGARCH
AIC	-6348	-6443	-6350	-6350	-6341
BIC	-6326	-6416	-6328	-6324	-6314
log veross.	3178	3226	3179	3180	3176

Analisando os valores BIC dos modelos ajustados, temos que o modelo GARCH com distribuição t de Student oferece o melhor ajuste. Em seguida, temos os modelos EGARCH e FIGARCH, que apresentam resultados muito semelhantes.

Para auxiliar na decisão da escolha do melhor modelo, podemos analisar os gráficos das funções de autocorrelação (FAC) do quadrado dos resíduos padronizados dos modelos ajustados (Figura 5.3) e os gráficos QQ-plots dos resíduos padronizados (Figura 5.4).

Aparentemente, todos os modelos são apropriados para modelar a volatilidade condicional. Porém, os QQ-Plots confirmam que o modelo GARCH com distribuição t de Student apresenta o melhor ajuste.

5.5. SELEÇÃO E COMPARAÇÃO DE MODELOS GARCH E GARCH COM MEMÓRIA LONGA105



Figura 5.3: Comparação das FACs do Quadrado dos Resíduos Padronizados - FIGARCH, FIE-GARCH, GARCH com distribuição Normal, GARCH com distribuição t de Student e EGARCH



Figura 5.4: Comparação dos QQ-plots dos Resíduos Padronizados - FIGARCH, FIEGARCH, GARCH com distribuição Normal, GARCH com distribuição t de Student e EGARCH

5.6 Simulações

Nesta seção apresentamos os resultados das simulações dos modelos FIGARCH (1, d, 0), FI-GARCH (0, d, 1) e FIGARCH (1, d, 1). Os programas para estas simulações foram implementados no software S-Plus e estão apresentados no Apêndice A. Ainda não há uma rotina pronta para este tipo de simulação neste software.

5.6. SIMULAÇÕES

Foram realizadas 100 simulações de tamanho de séries 2000. Porém, descartamos os 200 primeiros valores simulados para tirar o efeito da escolha do valor inicial de σ_t^2 ($\sigma_1^2 = 0, 01$).

O parâmetro de memória longa, d, recebeu dois valores: 0,45 e 0,75. Os demais parâmetros foram escolhidos respeitando-se as condições apresentadas em (5.12), (5.13) e (5.14), resultando nos modelos:

Modelo	d	a_1	b_1
	$0,\!45$	0,2	0
$\mathrm{FIGARCH}(1, d, 0)$	0,45	-0,2	0
	0,75	0,1	0
	0,75	-0,1	0
	$0,\!45$	0	0,4
$\mathrm{FIGARCH}(0, d, 1)$	$0,\!45$	0	$0,\!02$
	0,75	0	0,7
	0,75	0	$0,\!02$
	$0,\!45$	-0,2	0,4
$\mathrm{FIGARCH}(1, d, 1)$	$0,\!45$	-0,2	0,7
	0,75	-0,3	0,4
	0,75	-0,3	0,7

Tabela 5.6: Modelos FIGARCH simulados

O objetivo destas simulações é avaliar o comportamento do estimador de máxima verossimilhança do parâmetro de memória longa nos diferentes modelos simulados. Para tanto, utilizamos a média dos valores estimados, a variância estimada e o erro quadrático médio (EQM) do parâmetro, definidos em (3.29), (3.30) e (3.31), respectivamente.

Os resultados das simulações do modelo FIGARCH(1, d, 0) estão apresentados nas Tabelas 5.7 e 5.8. Para o modelo FIGARCH(0, d, 1), temos as Tabelas 5.9 e 5.10. E, finalmente, apresentamos os resultados das simulações do modelo FIGARCH(1, d, 1) nas Tabelas 5.11 e 5.12.

Os gráficos box-plot das estimativas do parâmetro de memória longa nos modelos simulados FIGARCH(1, d, 0), FIGARCH(0, d, 1) e FIGARCH(1, d, 1) podem ser observados nas Figuras 5.5, 5.6 e 5.7, respectivamente.

Analisando-se as tabelas e os gráficos box-plot, pode-se perceber que a estimativa do parâmetro dpouco se altera quando modificamos os demais parâmetros, $a_1 \in b_1$. Para o modelo FIGARCH(1, d, 0), notamos que o EQM do parâmetro d é um pouco menor para valores negativos do parâmetro a_1 (-0, 1 e -0, 2). No modelo FIGARCH(0, d, 1), por outro lado, não observamos este padrão, pois para d = 0, 45 o menor EQM ocorre quando $b_1 = 0, 02$, mas para d = 0, 75 o menor EQM ocorre quando $b_1 = 0, 75$ o menor EQM ocorre quando $b_1 = 0, 75$. E, para o modelo FIGARCH(1, d, 1), temos o melhor resultado se $a_1 = -0, 2$ e $b_1 = 0, 4$, para d = 0, 45, e $a_1 = -0, 3$ e $b_1 = 0, 7$, para d = 0, 75.

Apesar de não ser o objetivo aqui, vale observar o comportamento das estimativas dos parâmetros a_1 e b_1 dos modelos simulados. No modelo FIGARCH(1, d, 0), as estimativas do parâmetro a_1 são bem próximas do real, para qualquer valor do parâmetro d. As estimativas do parâmetro b_1 dos modelos simulados FIGARCH(0, d, 1) não apresentam bons resultados. Finalmente, para o modelo FIGARCH(1, d, 1), observamos boas estimativas para o parâmetro b_1 em todos os modelos simulados. Porém, as estimativas do parâmetro a_1 não podem ser consideradas satisfatórias.

Tabela 5.7: Estimativas da Média, Variância e EQM dos parâmetros nas simulações do processo FIGARCH(1, d, 0) (d=0,45)

			$a_1 = 0.2$	$a_1 = (-0, 2)$
	mádia	~	$\frac{a_1 - 0, 2}{0.2042}$	$\frac{u_1 - (-0, 2)}{0.2002}$
	media	a_1	0,2045	-0,2005
		d	0,4358	0,4480
d=0,45	variância	a_1	0,0052	0,0026
	amostral	d	0,0090	0,0027
	EQM	a_1	$0,\!0051$	0,0026
		d	0,0091	0,0027

Tabela 5.8: Estimativas da Média, Variância e EQM dos parâmetros nas simulações do processo FIGARCH(1, d, 0) (d=0,75)

			$a_1 = 0, 1$	$a_1 = (-0, 1)$
	média	a_1	0,0995	-0,0986
		d	0,7498	0,7462
d=0,75	variância	a_1	0,0016	0,0018
	amostral	d	0,0032	0,0018
	EQM	a_1	0,0016	0,0018
		d	0,0032	0,0018

Tabela 5.9: Estimativas da Média, Variância e EQM dos parâmetros nas simulações do processo FIGARCH(0, d, 1) (d=0,45)_____

			$b_1 = 0, 4$	$b_1 = 0,02$
	média	b_1	-0,0031	-0,0075
		d	$0,\!4399$	$0,\!4505$
d=0,45	variância	b_1	0,0033	0,0028
	$\operatorname{amostral}$	d	0,0048	0,0037
	EQM	b_1	$0,\!1657$	0,0035
		d	0,0048	0,0036

Tabela 5.10: Estimativas da Média, Variância e EQM dos parâmetros nas simulações do processo FIGARCH(0, d, 1) (d=0,75)

			$b_1 = 0, 7$	$b_1 = 0,02$
	média	b_1	-0,0019	0,0076
		d	0,7503	0,7625
d=0,75	variância	b_1	0,0020	0,0023
	amostral	d	0,0020	0,0022
	EQM	b_1	$0,\!4946$	0,0025
		d	0,0020	0,0024

Tabela 5.11: Estimativas da Média, Variância e EQM dos parâmetros nas simulações do processo FIGARCH(1, d, 1) (d=0,45)

			$a_1 = (-0, 2), b_1 = 0, 4$	$a_1 = (-0, 2), b_1 = 0, 7$
		b_1	0,3839	$0,\!6792$
	média	a_1	0,1969	0,4778
		d	$0,\!4329$	$0,\!4465$
d=0,45	variância	b_1	0,0575	0,0118
	amostral	a_1	0,0607	0,0105
		d	0,0092	0,0123
		b_1	0,0572	0,0121
	\mathbf{EQM}	a_1	0,2176	0,4698
		d	0,0094	0,0122

/ (/	,			
			$a_1 = (-0, 3), b_1 = 0, 4$	$a_1 = (-0, 3), b_1 = 0, 7$
		b_1	$0,\!4459$	$0,\!6748$
	média	a_1	0,2199	0,3881
		d	$0,\!6751$	0,7354
d=0,75	variância	b_1	0,0525	0,0108
	amostral	a_1	$0,\!1197$	0,0079
		d	0,0319	0,0168
		b_1	0,0536	0,0114
	\mathbf{EQM}	a_1	0,3877	$0,\!4813$
		d	0,0369	0,0169

Tabela 5.12: Estimativas da Média, Variância e EQM dos parâmetros nas simulações do processo FIGARCH(1, d, 1) (d=0,75)



Figura 5.5: Boxplot das estimativas do parâmetro d das simulações do processo FIGARCH(1, d, 0)



Figura 5.6: Boxplot das estimativas do parâmetro d das simulações do processo FIGARCH(0, d, 1)



Figura 5.7: Boxplot das estimativas do parâmetro d das simulações do processo FIGARCH(1, d, 1)

Capítulo 6

Aplicação em Séries Reais

6.1 Aplicações

Neste capítulo serão apresentadas algumas aplicações a dados reais dos testes e estimadores apresentados nos capítulos anteriores. As aplicações foram realizadas em séries de retornos de preços de ações de três grandes bancos brasileiros, Unibanco, Itaú e Bradesco, obtidos no site da BOVESPA (Bolsa de Valores de São Paulo)¹. O software utilizado para as aplicações foi o S-Plus e apresentaremos os códigos utilizados ao longo da aplicação. Muitas das rotinas utilizadas são provenientes do módulo FinMetrics deste software. Ver Zivot (2003) para mais detalhes sobre uso deste software no estudo de séries financeiras.

6.1.1 Preços das Ações do Unibanco

Os dados desta aplicação são provenientes das cotações diárias dos preços das ações do Unibanco (série UBB), de 06 de setembro de 1997 a 11 de janeiro de 2008, totalizando 2456 observações. Os dados são apresentados em reais e a cotação estudada foi o preço de fechamento diário da ação.

Primeiramente, apresentamos na Figura 6.1 a série UBB (a), a série de retornos UBB (b), o histograma dos retornos UBB com densidade ajustada (c) e o gráfico QQ-plot dos retornos UBB (d).

¹http://www.bovespa.com.br/Principal.asp



Figura 6.1: (a) Série UBB, (b) Série de retornos UBB, (c) Histograma dos retornos UBB, (d) QQ-plot dos retornos UBB

Nota-se que a série UBB apresenta uma queda muito grande por volta da observação de número 1600, provavelmente devido a alguma ação de desdobramento de ações. Como o estudo não prevê esse tipo de alterações, seguiremos as análises com as primeiras 1500 observações. Chamaremos esta série reduzida de UBB1500. Na Figura 6.2 temos a série UBB1500, a série de retornos UBB1500, o histograma dos retornos UBB1500 com densidade ajustada e o gráfico QQ-plot dos retornos UBB1500.



Figura 6.2: (a) Série UBB1500, (b) Série de retornos UBB1500, (c) Histograma dos retornos UBB1500, (d) QQ-plot dos retornos UBB1500

(a) Modelo ARFIMA

O objetivo desta análise é verificar se a volatilidade condicional da série UBB1500 possui comportamento de memória longa.

Conforme descrito no Capítulo 3, a série de retornos absolutos pode ser utilizada para representar a volatilidade de séries financeiras, considerando que a média dos retornos é muito próxima de zero. Abaixo temos algumas medidas resumo da série de retornos UBB1500, onde observamos que a média é igual a 0,0003656.²

 $^{^{2}}$ Como o software S-Plus tem como padrão o "ponto" como símbolo para separar as casas decimais dos números, os resultados deste capítulo provenientes deste software serão apresentados com essa notação.

>	<pre>> summary(retornos.ubb1500)</pre>								
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.			
-	-0.4418328	-0.0180325	0.000000	0.0003656	0.0188685	0.3830858			

A Figura 6.3 apresenta a série de retornos absolutos UBB1500 e sua função de autocorrelação amostral. Observamos que as autocorrelações permanecem significantes até lags altos, o que pode indicar presença de memória longa. No código abaixo, utilizamos a função **ar** do S-Plus para ajustar um modelo AR para a série retornos absolutos UBB1500. O comando **acf**.FARIMA compara a função de autocorrelação teórica do modelo ajustado com as autocorrelações amostrais, conforme podemos conferir na Figura 6.3 (b). Como as autocorrelações amostrais têm um decaimento lento, processos tradicionais ARMA resultam normalmente em modelos com um número excessivo de parâmetros. No caso dos retornos absolutos UBB1500, obtivemos um AR(12), o que significa que foram necessários 12 coeficientes autorregressivos para capturar a dependência nos dados.

> retornos.ubb1500abs = abs(retornos.ubb1500) > par(mfrow = c(1, 2)) > tsplot(retornos.ubb1500abs) > acf(retornos.ubb1500abs, lag = 50) > retornos.ubb1500abs.ar = ar(retornos.ubb1500abs) > retornos.ubb1500abs.ar\$order [1] 12 > temp = list(ar = as.vector(retornos.ubb1500abs.ar\$ar), sigma2 = 1, d = 0) > ar.acf = acf.FARIMA(temp, lag.max = 50) > lines(ar.acf\$lags, ar.acf\$acf/ar.acf\$acf[1])



Figura 6.3: (a) Retornos absolutos UBB1500 e (b) Função de autocorrelação amostral da série de retornos absolutos UBB1500

Apresentamos no Capítulo 3 os testes da estatística R/S e GPH para verificar a presença do comportamento de memória longa em séries temporais. Os comandos e os resultados dos testes para os retornos absolutos UBB1500 estão apresentados a seguir.

Estatística R/S:

```
> rosTest(retornos.ubb1500abs)
Test for Long Memory: Modified R/S Test
Null Hypothesis: no long-term dependence
Test Statistics:
    1.9665*
    * : significant at 5% level
** : significant at 1% level
Total Observ.: 1499
Bandwidth : 7
```

Estatística R/S modificada:

```
> rosTest(retornos.ubb1500abs, bandwidth = 0)
Test for Long Memory: R/S Test
Null Hypothesis: no long-term dependence
Test Statistics:
    3.2924**
    * : significant at 5% level
** : significant at 1% level
Total Observ.: 1499
```

Teste GPH:

```
> gph.ubb = gphTest(retornos.ubb1500abs)
> gph.ubb
Test for Long Memory: GPH Test
Null Hypothesis: d = 0
Test Statistics:
    d 0.3288
stat 3.3715**
    * : significant at 5% level
** : significant at 1% level
Total Observ.: 1499
Number of Freq: 55
```

Erro padrão do parâmetro d:

> gph.ubb\$std.err
[1] 0.09752148

A estatística R/S (1,9665) é significante a um nível de 5%, a estatística R/S modificada (3,2924) e a estatística do teste GPH (3,3715) são significantes a um nível de significância de 1%. Portanto, podemos concluir que existe presença de memória longa na série retornos absolutos UBB1500.

O teste GPH apresenta também uma estimativa para o parâmetro d, $\hat{d} = 0,3288$, o que também sugere memória longa. O intervalo de confiança de 95% para d baseado no erro padrão assintótico, [0,23127852,0,42632148], indica que a série é estacionária.

Uma vez que concluímos que a série retornos absolutos UBB1500 tem comportamento de memória longa, tentamos ajustar um modelo ARFIMA para esses dados. Os resultados seguem:

```
> farima.retornos.ubb1500abs = FARIMA(retornos.ubb1500abs,
  + p.range = c(0, 2), q.range = c(0, 2), mmax = 0)
> summary(farima.retornos.ubb1500abs)
Call:
FARIMA(x = retornos.ubb1500abs, p.range = c(0, 2), q.range = c(0, 2), mmax = 0)
Coefficients:
         Value Std. Error t value Pr(>|t|)
   d
        0.3284
                 0.0498
                            6.5910
                                     0.0000
AR(1) -0.1209
                 0.0225
                           -5.3667
                                     0.0000
AR(2) -0.7166
                                     0.0000
                0.0196
                          -36.6358
MA(1) -0.0402
                 0.0290
                           -1.3843
                                     0.1665
MA(2) -0.8020
                 0.0206
                                     0.0000
                          -38.9230
Information Criteria:
log-likelihood
                      BIC
  2779.909
                -5523.261
Residual scale estimate: 0.0378
                   total residual
Degree of freedom: 1497
                             1491
BIC of all models estimated:
          q=0
                    q=1
                              q=2
p=0 -5503.511 -5496.105 -5503.384
p=1 -5501.648 -5497.851 -5509.445
p=2 -5511.303 -5507.420 -5523.261
```

Pelo critério BIC, o melhor ajuste é o do modelo ARFIMA(2, d, 2). Para diagnosticar este modelo, podemos utilizar ferramentas gráficas. Na Figura 6.4 apresentamos o QQ-plot dos resíduos padronizados e as autocorrelações dos resíduos. Observamos que não há normalidade nos dados e que o modelo ajustado é capaz de capturar o comportamento de memória longa. Para construir estes gráficos utilizamos o comando abaixo.

> plot(farima.retornos.ubb1500abs2)

120



Figura 6.4: (a) QQ-plot dos resíduos padronizados e (b) Autocorrelações dos resíduos

(b) Modelos GARCH

Por enquanto, assumimos que a série retornos absolutos UBB1500 representa a volatilidade dos retornos UBB1500 e tentamos capturar o comportamento de memória longa na volatilidade. Agora, vamos utilizar os modelos da família GARCH para modelar diretamente a volatilidade da série de retornos UBB1500.

Primeiro, apresentamos na Figura 6.5 os gráficos das autocorrelações dos retornos UBB1500 e as autocorrelações do quadrado dos retornos UBB1500.



Figura 6.5: (a) Autocorrelações dos retornos UBB1500 e (b) Autocorrelações do quadrado dos retornos UBB1500

Observamos que a série de quadrado dos retornos UBB1500 apresenta autocorrelação significante pelo menos até o lag 13. Conforme descrevemos no Capítulo 4, este comportamento indica a existência de heteroscedasticidade condicional nos dados.

Para comprovar a existência de efeitos ARCH nos resíduos, conforme descrito em 4.2.1, utilizamos os comandos:

```
> archTest(retornos.ubb1500)
Test for ARCH Effects: LM Test
Null Hypothesis: no ARCH effects
Test Statistics:
Test Stat 442.2017
  p.value  0.0000
Dist. under Null: chi-square with 31 degrees of freedom
  Total Observ.: 1499
```

Como a estatística do teste é igual a 442,2017, rejeitamos a hipótese nula da não existência de efeitos ARCH, a um nível de significância de 5%.

Podemos, então, ajustar modelos da família GARCH para a série de retornos UBB1500. Ajustando um modelo GARCH(1, 1), supondo normalidade nos erros, temos:

```
> retornos.ubb1500.garch = garch(retornos.ubb1500 ~ 1, ~ garch(1, 1))
> summary(retornos.ubb1500.garch)
Call:
garch(formula.mean = retornos.ubb1500 ~ 1, formula.var = ~ garch(1, 1))
Mean Equation: retornos.ubb1500 ~ 1
Conditional Variance Equation: ~ garch(1, 1)
Conditional Distribution: gaussian
_____
Estimated Coefficients:
_____
         Value Std.Error t value Pr(>|t|)
    C 0.0033010 0.00081401 4.055 0.00002633
     A 0.0001886 0.00001209 15.607 0.00000000
ARCH(1) 0.3684906 0.02310150 15.951 0.00000000
GARCH(1) 0.6246309 0.01098741 56.850 0.00000000
 _____
AIC(4) = -5381.255
BIC(4) = -5360.005
Normality Test:
_____
Jarque-Bera P-value Shapiro-Wilk P-value
     6079 0 0.9435 0
Ljung-Box test for standardized residuals:
Statistic P-value Chi<sup>2</sup>-d.f.
   14.12 0.2929
                  12
Ljung-Box test for squared standardized residuals:
_____
Statistic P-value Chi<sup>2</sup>-d.f.
   17.37 0.1362 12
```

```
Lagrange multiplier test:
```

Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10 0.8692 -0.1633 -1.363 -0.9026 1.383 -0.5963 -0.719 -0.2076 3.546 -0.04466 Lag 11 Lag 12 C -0.2725 0.07226 -0.0167 TR^2 P-value F-stat P-value 17.94 0.1175 1.651 0.1747

Na análise das estatísticas resumo, podemos observar que os coeficientes do modelo ajustado são todos significantes.

Os testes de normalidade Jarque-Bera (6079, com p-valor=0) e Shapiro-Wilks (0,9435, com p-valor=0) para os resíduos padronizados ϵ_t/σ_t indicam não normalidade. No teste Ljung-Box para os resíduos padronizados e para o quadrado dos resíduos padronizados, não rejeitamos a hipótese nula H_0 da não existência de autocorrelação. Portanto, o modelo capturou com sucesso as estruturas de correlação serial tanto na variância condicional quanto na média condicional.

Além dos testes para autocorrelação nos resíduos padronizados, temos o teste do multiplicador de Lagrange para efeitos ARCH nos resíduos padronizados para verificar se ainda existe algum efeito ARCH restante. Com p-valor de 0, 1175, não rejeitamos a hipótese nula da não existência de efeitos ARCH.

A Figura 6.6 apresenta o QQ-plot dos resíduos padronizados e o gráfico das autocorrelações do quadrado dos resíduos padronizados. Observamos no gráfico QQ-plot que a suposição de normalidade não é apropriada, confirmando os resultados dos testes apresentados anteriormente.



Figura 6.6: (a) QQ-plot dos resíduos padronizados e (b) Autocorrelações do quadrado dos resíduos padronizados

Agora, vamos ajustar um modelo GARCH(1,1) supondo distribuição t de Student para os erros. Além disso, ajustaremos as extensões do modelo GARCH apresentadas em (4.6) (EGARCH, TGARCH e PGARCH). Uma ferramenta do S-Plus que nos ajuda a comparar os resultados dos ajustes é a função compare.mgarch. Abaixo temos os resultados dos ajustes para a série de retornos UBB1500, e a comparação dos critérios AIC e BIC. Observa-se, pelos critérios AIC e BIC, que o melhor modelo GARCH ajustado é aquele que supõe erros com distribuição t de Student. Confirmamos este resultado pela Figura 6.7. A Figura 6.8 apresenta as autocorrelações do quadrado dos resíduos padronizados, indicando que todos os modelos são adequados para modelar a volatilidade condicional.

GARCH com distribuição t:

```
> retornos.ubb1500.garcht = garch(retornos.ubb1500 ~ 1, ~ garch(1, 1),
cond.dist = "t")
> summary(retornos.ubb1500.garcht)
Call:
garch(formula.mean = retornos.ubb1500 ~ 1, formula.var = ~ garch(1, 1),
cond.dist = "t")
Mean Equation: retornos.ubb1500 ~ 1
Conditional Variance Equation: ~ garch(1, 1)
Conditional Distribution: t
with estimated parameter 3.122728 and standard error 0.2996047
       _____
Estimated Coefficients:
_____
                                     ------
          Value Std.Error t value Pr(>|t|)
     C 0.0014458 0.00070683 2.045 2.049e-002
     A 0.0002328 0.00005326 4.372 6.595e-006
ARCH(1) 0.3846816 0.07850717 4.900 5.314e-007
GARCH(1) 0.6217564 0.03869152 16.070 0.000e+000
AIC(5) = -5725.653
BIC(5) = -5699.09
Normality Test:
_____
          _____
Jarque-Bera P-value Shapiro-Wilk P-value
      6919
               0 0.9405
                                 0
```

```
Ljung-Box test for standardized residuals:
_____
Statistic P-value Chi<sup>2</sup>-d.f.
   14.19 0.2888 12
Ljung-Box test for squared standardized residuals:
_____
Statistic P-value Chi<sup>2</sup>-d.f.
   19.53 0.07655 12
Lagrange multiplier test:
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10
0.9636 -0.08449 -1.357 -0.8535 1.628 -0.5114 -0.7055 -0.1573 3.817 0.03362
Lag 11 Lag 12
              С
-0.2637 0.06991 0.03735
 TR<sup>2</sup> P-value F-stat P-value
20.28 0.06195 1.869 0.1195
```

EGARCH:

```
> retornos.ubb1500.egarch = garch(retornos.ubb1500 ~ 1, ~ egarch(1, 1),
leverage = T)
> summary(retornos.ubb1500.egarch)
Call:
garch(formula.mean = retornos.ubb1500 ~ 1, formula.var = ~ egarch(1, 1),
leverage = T)
Mean Equation: retornos.ubb1500 ~ 1
Conditional Variance Equation: ~ egarch(1, 1)
Conditional Distribution: gaussian
_____
Estimated Coefficients:
_____
         Value Std.Error t value Pr(>|t|)
     C 0.001165 0.000872 1.336 9.084e-002
     A -0.778409 0.030020 -25.930 0.000e+000
ARCH(1) 0.412271 0.015178 27.163 0.000e+000
GARCH(1) 0.922724 0.004439 207.859 0.000e+000
 LEV(1) -0.209196 0.032734 -6.391 1.100e-010
     _____
AIC(5) = -5393.344
BIC(5) = -5366.781
Normality Test:
_____
Jarque-Bera P-value Shapiro-Wilk P-value
      5890
              0
                    0.9401
                              0
```

```
Ljung-Box test for standardized residuals:
          _____
Statistic P-value Chi<sup>2</sup>-d.f.
   15.09 0.2368
                    12
Ljung-Box test for squared standardized residuals:
_____
Statistic P-value Chi<sup>2</sup>-d.f.
   36.48 0.0002711
                     12
Lagrange multiplier test:
_____
 Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10
0.7143 0.2453 -1.324 -0.9288 1.609 -0.5148 -0.8064 -0.06193 5.616 -0.2487
 Lag 11 Lag 12
                С
-0.2503 0.06908 0.8566
 TR^2
     P-value F-stat P-value
37.11 0.0002146 3.46 0.01196
```

No teste Ljung-Box (p-valor igual a 0,0002711) para o quadrado dos resíduos padronizados, rejeitamos a hipótese nula H_0 da não existência de autocorrelação. Portanto, o modelo não captura com sucesso a estrutura de correlação serial variância condicional.
TGARCH:

```
> retornos.ubb1500.tgarch = garch(retornos.ubb1500 ~ 1, ~ tgarch(1, 1), trace
= F)
> summary(retornos.ubb1500.tgarch)
Call:
garch(formula.mean = retornos.ubb1500 ~ 1, formula.var = ~ tgarch(1, 1),
trace = F)
Mean Equation: retornos.ubb1500 ~ 1
Conditional Variance Equation: ~ tgarch(1, 1)
Conditional Distribution: gaussian
_____
Estimated Coefficients:
_____
         Value Std.Error t value Pr(>|t|)
     C 0.0015801 0.00093444 1.691 4.553e-002
     A 0.0001599 0.00001175 13.614 0.000e+000
ARCH(1) 0.2158605 0.02701008 7.992 1.332e-015
GARCH(1) 0.6601752 0.01278423 51.640 0.000e+000
GAMMA(1) 0.2172474 0.03793976 5.726 6.201e-009
         _____
AIC(5) = -5396.454
BIC(5) = -5369.891
Normality Test:
_____
Jarque-Bera P-value Shapiro-Wilk P-value
      5083
              0
                    0.9461
                              0
```

```
Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi<sup>2</sup>-d.f.
   14.63 0.262 12
Ljung-Box test for squared standardized residuals:
_____
Statistic P-value Chi<sup>2</sup>-d.f.
   14.91 0.2465 12
Lagrange multiplier test:
_____
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10
0.6815 -0.292 -1.45 -0.9687 1.409 -0.6466 -0.7293 -0.1441 3.098 -0.1609
Lag 11 Lag 12
              С
-0.2386 0.07708 0.1904
 TR<sup>2</sup> P-value F-stat P-value
15.43 0.2185 1.418 0.2655
```

PGARCH:

```
> retornos.ubb1500.pgarch = garch(retornos.ubb1500 ~ 1, ~ pgarch(1, 1), trace
= F)
> summary(retornos.ubb1500.pgarch)
Call:
garch(formula.mean = retornos.ubb1500 ~ 1, formula.var = ~ pgarch(1, 1),
trace = F)
Mean Equation: retornos.ubb1500 ~ 1
Conditional Variance Equation: ~ pgarch(1, 1)
Conditional Distribution: gaussian
_____
Estimated Coefficients:
_____
          Value Std.Error t value Pr(>|t|)
     \texttt{C} \ \texttt{0.0029164} \ \texttt{0.0008233} \qquad \texttt{3.542} \ \texttt{0.0002045}
     A 0.0003495 0.0001054 3.315 0.0004689
ARCH(1) 0.3261021 0.0202585 16.097 0.0000000
GARCH(1) 0.6707024 0.0167764 39.979 0.0000000
  POWER 1.7695038 0.1031289 17.158 0.0000000
    _____
AIC(5) = -5380.297
BIC(5) = -5353.735
Normality Test:
_____
Jarque-Bera P-value Shapiro-Wilk P-value
      6437
               0
                      0.9426
                                0
```

```
Ljung-Box test for standardized residuals:
_____
Statistic P-value Chi<sup>2</sup>-d.f.
   13.64 0.3241 12
Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi<sup>2</sup>-d.f.
   16.81 0.1568 12
Lagrange multiplier test:
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10
 1.15 -0.0726 -1.353 -0.9104 1.386 -0.6042 -0.7677 -0.2297 3.419 -0.02578
Lag 11 Lag 12
              С
-0.3419 0.07096 0.2192
 TR<sup>2</sup> P-value F-stat P-value
17.56 0.1298 1.615 0.1861
```

Comparação dos modelos ajustados:

<pre>> compara.ubb = compare.mgarch(retornos.ubb1500.garch, retornos.ubb1500.garcht,</pre>						
retornos.ubb1500.egarch, retornos.ubb1500.tgarch,						
retornos.uk	bb1500.pgarch)					
> compara.u	ıbb					
	retornos.ubb1500.garch i	retornos.ubb1500.garch	t			
AIC	-5381	-572	6			
BIC	-5360	-569	9			
Likelihood	2695	286	8			
	retornos.ubb1500.egarch	retornos.ubb1500.tgar	ch			
AIC	-5393	-53	96			
BIC	-5367	-53	70			
Likelihood	2702	27	03			
retornos.ubb1500.pgarch						
AIC	-5380					
BIC	-5354					
Likelihood	2695					

Gráficos de diagnóstico:

- > plot(compara.ubb)
- > plot(compara.ubb, qq = T)



Figura 6.7: QQ-plot dos resíduos padronizados - modelos TGARCH, PGARCH, GARCH, GARCH com distribuição t e EGARCH



Figura 6.8: FAC do quadrado dos resíduos padronizados - modelos TGARCH, PGARCH, GARCH, GARCH, GARCH com distribuição t e EGARCH

(c) Modelos GARCH com memória longa

Finalmente, com o objetivo de capturar diretamente o comportamento de memória longa e a alta persistência na volatilidade, vamos ajustar modelos FIGARCH e FIEGARCH na série de retornos UBB1500.

```
FIGARCH
```

```
> retornos.ubb1500.figarch = fgarch(retornos.ubb1500 ~ 1, ~ figarch(1, 1))
> summary(retornos.ubb1500.figarch)
Call:
fgarch(formula.mean = retornos.ubb1500 ~ 1, formula.var = ~ figarch(1, 1))
Mean Equation: retornos.ubb1500 ~ 1
Conditional Variance Equation: ~ figarch(1, 1)
    _____
Estimated Coefficients:
_____
        Value Std.Error t value Pr(>|t|)
     C 0.003459 0.00065093 5.314 6.181e-008
     A 0.000148 0.00001446 10.231 0.000e+000
GARCH(1) 0.672739 0.03200063 21.023 0.000e+000
ARCH(1) 0.150692 0.03380121 4.458 4.442e-006
fraction 1.000000 0.04998750 20.005 0.000e+000
_____
AIC(5) = -5388.206
BIC(5) = -5361.643
Normality Test:
_____
Jarque-Bera P-value Shapiro-Wilk P-value
     6112
              0
                    0.9441
                             0
```

```
Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi<sup>2</sup>-d.f.
   13.76 0.3163 12
Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi<sup>2</sup>-d.f.
   11.18 0.5138 12
Lagrange multiplier test:
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10
1.537 0.1625 -1.283 -0.8765 1.317 -0.5961 -0.8547 -0.2932 2.305 -0.07007
Lag 11 Lag 12
              С
-0.3423 0.07353 -0.04106
TR<sup>2</sup> P-value F-stat P-value
  12 0.4454 1.1 0.4693
```

```
FIEGARCH
```

```
> retornos.ubb1500.fiegarch = fgarch(retornos.ubb1500 ~ 1, ~ fiegarch(1, 1),
leverage = T)
> summary(retornos.ubb1500.fiegarch)
Call:
fgarch(formula.mean = retornos.ubb1500 ~ 1, formula.var = ~ fiegarch(1, 1),
leverage = T)
Mean Equation: retornos.ubb1500 ~ 1
Conditional Variance Equation: ~ fiegarch(1, 1)
      _____
Estimated Coefficients:
_____
          Value Std.Error t value Pr(>|t|)
     C 0.001355 0.0008375 1.618 5.289e-002
     A -0.339235 0.0155445 -21.824 0.000e+000
GARCH(1) 0.341826 0.0786678 4.345 7.428e-006
ARCH(1) 0.428973 0.0236208 18.161 0.000e+000
 LEV(1) -0.118341 0.0149878 -7.896 2.776e-015
fraction 0.567583 0.0304559 18.636 0.000e+000
_____
                                 _____
AIC(6) = -5390.52
BIC(6) = -5358.645
Normality Test:
_____
Jarque-Bera P-value Shapiro-Wilk P-value
      5320
              0
                     0.935
                               0
```

```
Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi<sup>2</sup>-d.f.
   17.55 0.1301
                    12
Ljung-Box test for squared standardized residuals:
_____
Statistic
          P-value Chi<sup>2</sup>-d.f.
   62.54 7.731e-009
                      12
Lagrange multiplier test:
_____
 Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10
0.4188 0.4792 -1.326 -0.7549 2.418 -0.228 -0.8587 0.3081 7.53 -0.5036
 Lag 11 Lag 12
                С
-0.4334 0.06799 0.7794
 TR^2
       P-value F-stat P-value
63.16 5.966e-009 5.996 0.001013
```

No teste Ljung-Box para o quadrado dos resíduos padronizados, rejeitamos a hipótese nula H_0 da não existência de autocorrelação. Portanto, o modelo não captura com sucesso a estrutura de correlação serial variância condicional.

Comparação dos modelos FIGARCH e FIEGARCH com modelo GARCH com distribuição t

<pre>> compara2.ubb = compare.mgarch(retornos.ubb1500.garcht,</pre>					
retornos.ubb1500.figarch, retornos.ubb1500.fiegarch)					
> compara2.ubb					
retornos	s.ubb1500.garcht retornos.u	bb1500.figarch			
AIC	-5726	-5388			
BIC	-5699	-5362			
Likelihood	2868	2699			
retornos	s.ubb1500.fiegarch				
AIC	-5391				
BIC	-5359				
Likelihood	2701				
> plot(compara2.ubb)					
> plot(compara2.ubb, qq = T)					



Figura 6.9: FAC do quadrado dos resíduos padronizados - modelos GARCH com distribuição t, FIGARCH e FIEGARCH





(d) Conclusão

Comparando os modelos ajustados nos itens (a), (b) e (c), concluímos que o modelo GARCH com distribuição t de Student apresenta o melhor ajuste segundo os valores dos critérios AIC e BIC, que

podem ser observados na Tabela 6.1. O segundo melhor modelo obtido, segundo os mesmos critérios, foi o modelo ARFIMA ajustado à série de retornos absolutos.

Modelo	AIC	BIC
ARFIMA	-	-5523
GARCH t	-5726	-5699
FIGARCH	-5388	-5362
FIEGARCH	-5391	-5359

Tabela 6.1: AIC e BIC dos modelos ajustados para a volatilidade da série UBB

6.1.2 Preços das Ações do Itaú

Os dados desta aplicação são provenientes das cotações diárias dos preços das ações do Itaú (série ITAU), de 04 de julho de 1994 a 11 de janeiro de 2008, totalizando 3348 observações. Os dados são apresentados em reais e a cotação estudada foi o preço de fechamento diário da ação.

Primeiramente, apresentamos na Figura 6.11 a série ITAU (a), a série de retornos ITAU (b), o histograma dos retornos ITAU com densidade ajustada (c) e o gráfico QQ-plot dos retornos ITAU (d).



Figura 6.11: (a) Série ITAU, (b) Série de retornos ITAU, (c) Histograma dos retornos ITAU, (d) QQ-plot dos retornos ITAU

Nota-se que a série ITAU apresenta quedas grandes, sendo que a primeira ocorre por volta da observação de número 1300, provavelmente devido a alguma ação de desdobramento de ações. Como o estudo não prevê esse tipo de alterações, seguiremos as análises com as primeiras 1200 observações. Chamaremos esta série reduzida de ITAU1200. Na Figura 6.12 temos a série ITAU1200, a série de retornos ITAU1200, o histograma dos retornos ITAU1200 com densidade ajustada e o gráfico QQ-plot dos retornos ITAU1200.



Figura 6.12: (a) Série ITAU1200, (b) Série de retornos ITAU1200, (c) Histograma dos retornos ITAU1200, (d) QQ-plot dos retornos ITAU1200

(a) Modelo ARFIMA

O objetivo desta análise é verificar se a volatilidade condicional da série ITAU1200 possui com-

portamento de memória longa.

Abaixo temos algumas medidas resumo da série de retornos UBB1500. A média dos retornos é igual a 0,0007471, que é próxima de zero. Ajustaremos o modelo ARFIMA para valores absolutos da série ITAU1200.

> summary(retornos.BBD1500)
 Min. 1st Qu. Median Mean 3rd Qu. Max.
-0.4228569 -0.0142099 0.0000000 0.0007471 0.0151710 0.2876821

A Figura 6.13 apresenta a série de retornos absolutos ITAU1200 e sua função de autocorrelação amostral. Observamos que as autocorrelações permanecem significantes até lags altos, o que pode indicar presença de memória longa. O processo tradicional ARMA ajustado resultou em um modelo com um número excessivo de parâmetros, AR(9), o que significa que foram necessários 9 coeficientes autorregressivos para capturar a dependência nos dados.



Figura 6.13: (a) Retornos absolutos ITAU1200 e (b) Função de autocorrelação amostral da série de retornos absolutos ITAU1200

A estatística R/S (2,8555), a estatística R/S modificada (4,5453) e a estatística do teste GPH

(2,7238) são significantes a um nível de significância de 1%. Portanto, podemos concluir que existe presença de memória longa na série retornos absolutos ITAU1200.

O teste GPH apresenta também uma estimativa para o parâmetro d, $\hat{d} = 0,2315$, o que também sugere memória longa. O intervalo de confiança de 95% para d baseado no erro padrão assintótico, [0, 1271637, 0, 3358363], indica que a série é estacionária.

Ajustando um modelo ARFIMA para esses dados temos:

```
> farima.retornos.itau1200abs = FARIMA(retornos.itau1200abs, p.range = c(0,
2), q.range = c(0, 2), mmax = 0)
> summary(farima.retornos.itau1200abs)
Call:
FARIMA(x = retornos.itau1200abs, p.range = c(0, 2), q.range = c(0, 2), mmax =
0)
Coefficients:
   Value Std. Error t value Pr(>|t|)
d 0.2341 0.0227
                     10.3350 0.0000
Information Criteria:
log-likelihood
                      BIC
 2989.928
                -5972.768
Residual scale estimate: 0.0199
                   total residual
Degree of freedom: 1197
                             1195
BIC of all models estimated:
          q=0
                    q=1
                              q=2
p=0 -5972.768 -5964.562 -5956.942
p=1 -5969.898 -5964.875 -5962.208
p=2 -5966.150 -5963.152 -5956.169
```

Pelo critério BIC, o melhor ajuste é o do modelo ARFIMA(0, d, 0). Para diagnosticar este modelo, podemos utilizar ferramentas gráficas. Na Figura 6.14 apresentamos o QQ-plot dos resíduos padronizados e as autocorrelações dos resíduos. Observamos que não há normalidade nos dados e que o modelo ajustado é capaz de capturar o comportamento de memória longa.



Figura 6.14: (a) QQ-plot dos resíduos padronizados e (b) Autocorrelações dos resíduos

(b) Modelos GARCH

Por enquanto, assumimos que a série retornos absolutos ITAU1200 representa a volatilidade dos retornos ITAU1200 e tentamos capturar o comportamento de memória longa na volatilidade. Agora, vamos utilizar os modelos da família GARCH para modelar diretamente a volatilidade da série de retornos ITAU1200.

Primeiro, apresentamos na Figura 6.15 os gráficos das autocorrelações dos retornos ITAU1200 e as autocorrelações do quadrado dos retornos ITAU1200.



Figura 6.15: (a) Autocorrelações dos retornos ITAU1200 e (b) Autocorrelações do quadrado dos retornos ITAU1200

Observamos que a série de quadrado dos retornos ITAU1200 apresenta autocorrelação significante pelo menos até o lag 14. Conforme descrevemos no Capítulo 4, este comportamento indica a existência de heteroscedasticidade condicional nos dados.

Para comprovar a existência de efeitos ARCH nos resíduos, realizamos o teste:

```
> archTest(retornos.itau1200)
Test for ARCH Effects: LM Test
Null Hypothesis: no ARCH effects
Test Statistics:
Test Stat 196.0797
  p.value  0.0000
Dist. under Null: chi-square with 30 degrees of freedom
  Total Observ.: 1199
```

Como a estatística do teste é igual a 196,0797, rejeitamos a hipótese nula da não existência de efeitos ARCH, a um nível de significância de 5%.

Podemos, então, ajustar modelos da família GARCH para a série de retornos ITAU1200. Ajustando um modelo GARCH(1,1), supondo normalidade nos erros, temos que, na análise das estatísticas resumo, os coeficientes do modelo ajustado são todos significantes:

Estimated Coefficients:					
	Value	Std.Error	t value	Pr(> t)	
C	0.00244736	0.00072734	3.365	3.952e-004	
А	0.00006948	0.00001155	6.016	1.188e-009	
ARCH(1)	0.21231705	0.02430928	8.734	0.000e+000	
GARCH(1)	0.71588897	0.02994936	23.903	0.000e+000	

Os testes de normalidade Jarque-Bera (234, com p-valor=0) e Shapiro-Wilks (0,9707, com pvalor=0) para os resíduos padronizados ϵ_t/σ_t indicam não normalidade (H_0 =normalidade). No teste Ljung-Box para os resíduos padronizados rejeitamos a hipótese nula H_0 da não existência de autocorrelação, mas para o quadrado dos resíduos padronizados, não rejeitamos a hipótese nula H_0 . Portanto, o modelo capturou com sucesso apenas a estrutura de correlação serial na variância condicional.

```
Normality Test:
          _____
Jarque-Bera P-value Shapiro-Wilk P-value
      234
                  0.9707
             0
                        0
Ljung-Box test for standardized residuals:
_____
         P-value Chi^2-d.f.
Statistic
   44.29 0.00001363
                    12
Ljung-Box test for squared standardized residuals:
_____
Statistic P-value Chi<sup>2</sup>-d.f.
   12.24 0.4265
                  12
```

No teste do multiplicador de Lagrange para efeitos ARCH nos resíduos padronizados não rejeitamos a hipótese nula da não existência de efeitos ARCH, com p-valor de 0,3387.

```
Lagrange multiplier test:

Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10

-0.6212 0.4992 -1.507 1.448 -0.5712 -0.6972 -0.8587 0.3251 -1.688 1.48

Lag 11 Lag 12 C

1.438 -0.1593 0.2425

TR^2 P-value F-stat P-value

13.43 0.3387 1.235 0.3697
```

A Figura 6.16 apresenta o QQ-plot dos resíduos padronizados e o gráfico das autocorrelações do quadrado dos resíduos padronizados. Observamos no gráfico QQ-plot que a suposição de normalidade não é apropriada, confirmando os resultados dos testes apresentados anteriormente.



Figura 6.16: (a) QQ-plot dos resíduos padronizados e (b) Autocorrelações do quadrado dos resíduos padronizados

Agora, vamos ajustar um modelo GARCH(1,1) supondo distribuição t de Student para os erros e as extensões do modelo GARCH apresentadas em (4.6) (EGARCH, TGARCH e PGARCH). Observase, pelos critérios AIC e BIC, que o melhor modelo GARCH ajustado é aquele que supõe erros com distribuição t de Student. Confirmamos este resultado pela Figura 6.17. A Figura 6.18 apresenta as autocorrelações do quadrado dos resíduos padronizados, indicando que todos os modelos são adequados para modelar a volatilidade condicional.

Comparação dos modelos ajustados:

<pre>> compara.itau = compare.mgarch(retornos.itau1200.garch,</pre>						
retornos.itau1200.garcht, retornos.itau1200.egarch,						
retornos.it	tau1200.tgarch,	retornos.itau12	200.pgarch)			
> compara.i	itau					
	retornos.itau1	200.garch retorr	os.itau1200.garcht			
AIC		-5315	-5421			
BIC		-5295	-5395			
Likelihood		2662	2715			
	retornos.itau1	200.egarch retor	nos.itau1200.tgarch			
AIC		-5340	-5336			
BIC		-5315	-5311			
Likelihood		2675	2673			
	retornos.itau1	200.pgarch				
AIC		-5320				
BIC		-5294				
Likelihood		2665				

Gráficos de diagnóstico:



Figura 6.17: QQ-plot dos resíduos padronizados - modelos TGARCH, PGARCH, GARCH, GARCH com distribuição t e EGARCH



Figura 6.18: FAC do quadrado dos resíduos padronizados - modelos TGARCH, PGARCH, GARCH, GARCH, GARCH com distribuição t e EGARCH

(c) Modelos GARCH com memória longa

Finalmente, com o objetivo de capturar diretamente o comportamento de memória longa e a alta persistência na volatilidade, vamos ajustar modelos FIGARCH e FIEGARCH na série de retornos ITAU1200.

Comparação dos modelos FIGARCH e FIEGARCH com modelo GARCH com distribuição t

<pre>> compara2.itau = compare.mgarch(retornos.itau1200.garcht,</pre>							
retornos.itau1200.figarch, retornos.itau1200.fiegarch)							
> compara2.	.itau						
	retornos.itau1200.garcht retornos.itau	1200.figarch					
AIC -5421 -5319							
BIC -5395 -529							
Likelihood	2715	2665					
retornos.itau1200.fiegarch							
AIC	-5335						
BIC -5305							
Likelihood 2674							



Figura 6.19: FAC do quadrado dos resíduos padronizados - modelos GARCH com distribuição t
, FIGARCH e FIEGARCH

158



Figura 6.20: QQ-plot dos resíduos padronizados - modelos FIEGARCH, GARCH com distribuição t e FIGARCH

(d) Conclusão

Comparando os modelos ajustados nos itens (a), (b) e (c), concluímos que o modelo ARFIMA (ajustado para a série de retornos absolutos) apresenta o melhor ajuste segundo o critérios BIC,

conforme podemos observar na tabela a seguir.

Tabela 6.2: AIC e BIC dos modelos ajustados para a volatilidade da série ITAU

Modelo	AIC	BIC
ARFIMA	-	-5973
GARCH t	-5421	-5395
FIGARCH	-5319	-5294
FIEGARCH	-5335	-5305

6.1.3 Preços das Ações do Bradesco

Os dados desta aplicação são provenientes das cotações diárias dos preços das ações do Bradesco (série BDC), de 04 de julho de 1994 a 11 de janeiro de 2008, totalizando 3348 observações. Os dados são apresentados em reais e a cotação estudada foi o preço de fechamento diário da ação.

Primeiramente, apresentamos na Figura 6.21 a série BDC (a), a série de retornos BDC (b), o histograma dos retornos BDC com densidade ajustada (c) e o gráfico QQ-plot dos retornos BDC (d).



Figura 6.21: (a) Série BDC, (b) Série de retornos BDC, (c) Histograma dos retornos BDC, (d) QQ-plot dos retornos BDC

Nota-se que a série BDC apresenta oscilação grande por volta da observação 2500, provavelmente devido a alguma ação de grupamento de ações. Como o estudo não prevê esse tipo de alterações, seguiremos as análises com as primeiras 1500 observações. Chamaremos esta série reduzida de BDC1500. Na Figura 6.22 temos a série BDC1500, a série de retornos BDC1500, o histograma dos retornos BDC1500 com densidade ajustada e o gráfico QQ-plot dos retornos BDC1500.



Figura 6.22: (a) Série BDC1500, (b) Série de retornos BDC1500, (c) Histograma dos retornos BDC1500, (d) QQ-plot dos retornos BDC1500

(a) Modelo ARFIMA

O objetivo desta análise é verificar se a volatilidade condicional da série BDC1500 possui comportamento de memória longa.

Abaixo temos algumas medidas resumo da série de retornos UBB1500. A média dos retornos é igual a 0,0007471, que é próxima de zero. Ajustaremos o modelo ARFIMA para valores absolutos da série de retornos BDC1500.

> ຣະ	> summary(retornos.BBD1500)					
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-0.	4228569	-0.0142099	0.0000000	0.0007471	0.0151710	0.2876821

A Figura 6.23 apresenta a série de retornos absolutos BDC1500 e sua função de autocorrelação amostral. Observamos que as autocorrelações permanecem significantes até lags altos, o que pode indicar presença de memória longa. O processo tradicional ARMA ajustado resultou em um modelo com um número excessivo de parâmetros, AR(12), o que significa que foram necessários 12 coeficientes autorregressivos para capturar a dependência nos dados.



Figura 6.23: (a) Retornos absolutos BDC1500 e (b) Função de autocorrelação amostral da série de retornos absolutos BDC1500

A estatística R/S (2,5218), a estatística R/S modificada (4,5453) e a estatística do teste GPH (3,7205) são significantes a um nível de significância de 1%. Portanto, podemos concluir que existe presença de memória longa na série retornos absolutos BDC1500.

O teste GPH apresenta também uma estimativa para o parâmetro d, $\hat{d} = 0,3164$, o que também sugere memória longa. O intervalo de confiança de 95% para d baseado no erro padrão assintótico,

[0, 2189, 0, 4139], indica que a série é estacionária.

Ajustando um modelo ARFIMA para esses dados temos:

```
> farima.retornos.BBD1500abs = FARIMA(retornos.BBD1500abs, p.range = c(0, 2),
q.range = c(0, 2), mmax = 0)
> summary(farima.retornos.BBD1500abs)
Call:
FARIMA(x = retornos.BBD1500abs, p.range = c(0, 2), q.range = c(0, 2), mmax = 0
)
Coefficients:
   Value Std. Error t value Pr(>|t|)
d 0.2094 0.0203
                    10.3381 0.0000
Information Criteria:
log-likelihood
                      BIC
 3459.829
               -6912.348
Residual scale estimate: 0.024
                   total residual
Degree of freedom: 1497
                             1495
BIC of all models estimated:
          q=0
                    q=1
                              q=2
p=0 -6912.348 -6902.740 -6896.887
p=1 -6908.388 -6901.158 -6894.586
p=2 -6906.592 -6899.647 -6897.497
```

Pelo critério BIC, o melhor ajuste é o do modelo ARFIMA(0, d, 0). Para diagnosticar este modelo, podemos utilizar ferramentas gráficas. Na Figura 6.24 apresentamos o QQ-plot dos resíduos padronizados (a) e as autocorrelações dos resíduos (b). Observamos que não há normalidade nos dados e que o modelo ajustado é capaz de capturar o comportamento de memória longa.



Figura 6.24: (a) QQ-plot dos resíduos padronizados e (b) Autocorrelações dos resíduos

(b) Modelos GARCH

Por enquanto, assumimos que a série retornos absolutos BDC1500 representa a volatilidade dos retornos BDC1500 e tentamos capturar o comportamento de memória longa na volatilidade. Agora, vamos utilizar os modelos da família GARCH para modelar diretamente a volatilidade da série de retornos BDC1500.

Primeiro, apresentamos na Figura 6.25 os gráficos das autocorrelações dos retornos BDC1500 e as autocorrelações do quadrado dos retornos BDC1500.



Figura 6.25: (a) Autocorrelações dos retornos BDC1500 e (b) Autocorrelações do quadrado dos retornos BDC1500

Para testar a existência de efeitos ARCH nos resíduos, realizamos o teste:

> archTest(retornos.BBD1500)
Test for ARCH Effects: LM Test
Null Hypothesis: no ARCH effects
Test Statistics:
Test Stat 18.3942
 p.value 0.9643
Dist. under Null: chi-square with 31 degrees of freedom
 Total Observ.: 1499

Como a estatística do teste é igual a 18,3942, não rejeitamos a hipótese nula da não existência de efeitos ARCH.
6.1. APLICAÇÕES

Ajustando um modelo GARCH(1,1), supondo normalidade nos erros, temos que, na análise das estatísticas resumo, os coeficientes do modelo ajustado são todos significantes:

Estimated Coefficients:					
	Value	Std.Error	t value	Pr(> t)	
C	0.0017654	0.00084997	2.077	1.898e-002	
А	0.0003319	0.00003285	10.105	0.000e+000	
ARCH(1)	0.3034684	0.03652567	8.308	1.110e-016	
GARCH(1)	0.4124113	0.05270524	7.825	4.774e-015	

Os testes de normalidade Jarque-Bera (168870, com p-valor=0) e Shapiro-Wilks (0,893, com p-valor=0) para os resíduos padronizados ϵ_t/σ_t indicam não normalidade (H_0 =normalidade).

No teste Ljung-Box para os resíduos padronizados rejeitamos a hipótese nula H_0 da não existência de autocorrelação, mas para o quadrado dos resíduos padronizados, não rejeitamos a hipótese nula H_0 .

```
Normality Test:
              _____
Jarque-Bera P-value Shapiro-Wilk P-value
    168870
             0
                   0.893
                            0
Ljung-Box test for standardized residuals:
_____
Statistic P-value Chi<sup>2</sup>-d.f.
   24.82 0.01571
                   12
Ljung-Box test for squared standardized residuals:
_____
Statistic P-value Chi<sup>2</sup>-d.f.
  0.2431
            1
                   12
```

No teste do multiplicador de Lagrange para efeitos ARCH nos resíduos padronizados não rejeitamos a hipótese nula da não existência de efeitos ARCH.

```
Lagrange multiplier test:
    Lag 1
             Lag 2
                      Lag 3
                               Lag 4
                                       Lag 5
                                                 Lag 6
                                                          Lag 7 Lag 8
                                                                         Lag 9
 -0.07948 -0.02453 -0.2404 -0.1199 -0.1685 -0.08155 -0.1863 0.206 -0.1447
  Lag 10 Lag 11
                    Lag 12
                                  С
 -0.0147 0.03677 -0.01459 -0.1683
   TR<sup>^</sup>2 P-value F-stat P-value
0.2425
               1 0.02205
                                1
```

A Figura 6.26 apresenta o QQ-plot dos resíduos padronizados e o gráfico das autocorrelações do quadrado dos resíduos padronizados. Observamos no gráfico QQ-plot que a suposição de normalidade não é apropriada, confirmando os resultados dos testes apresentados anteriormente.



Figura 6.26: (a) QQ-plot dos resíduos padronizados e (b) Autocorrelações do quadrado dos resíduos padronizados

Agora, vamos ajustar um modelo GARCH(1, 1) supondo distribuição t de Student para os erros e as extensões do modelo GARCH apresentadas em (4.6) (EGARCH, TGARCH e PGARCH). Observase, pelos critérios AIC e BIC, que o melhor modelo GARCH ajustado é aquele que supõe erros com

6.1. APLICAÇÕES

distribuição t de Student. Confirmamos este resultado pela Figura 6.27. A Figura 6.28 apresenta as autocorrelações do quadrado dos resíduos padronizados, indicando que todos os modelos são adequados para modelar a volatilidade condicional.

Comparação dos modelos ajustados:

> compara.BBD					
retornos.BBD1500.garch retornos.BBD1500.garcht					
AIC	-6211	-6628			
BIC	-6189	-6601			
Likelihood	3109	3319			
	retornos.BBD1500.egarch	retornos.BBD1500.tgarch			
AIC	-6253	-6249			
BIC	-6227	-6223			
Likelihood	3132	3130			
retornos.BBD1500.pgarch					
AIC	-6216				
BIC	-6189				
Likelihood	3113				

Gráficos de diagnóstico:



Figura 6.27: QQ-plot dos resíduos padronizados - modelos TGARCH, PGARCH, GARCH, GARCH com distribuição t ${\rm e}$ EGARCH



Figura 6.28: FAC do quadrado dos resíduos padronizados - modelos TGARCH, PGARCH, GARCH, GARCH, GARCH com distribuição t e EGARCH

(c) Modelos GARCH com memória longa

Finalmente, com o objetivo de capturar diretamente o comportamento de memória longa e a alta persistência na volatilidade, vamos ajustar modelos FIGARCH e FIEGARCH na série de retornos BDC1500.

Comparação dos modelos FIGARCH e FIEGARCH com modelo GARCH com distribuição t

> compara2.BBD = compare.mgarch(retornos.BBD1500.garcht,					
retornos.BBD1500.figarch, retornos.BBD1500.fiegarch)					
> compara2.	BBD				
	retornos.BBD150	00.garcht retor	nos.BBD1500.figarch		
AIC -6628 -6195					
BIC		-6601	-6168		
Likelihood		3319	3102		
retornos.BBD1500.fiegarch					
AIC		-6257			
BIC		-6225			
Likelihood		3134			



Figura 6.29: FAC do quadrado dos resíduos padronizados - modelos GARCH com distribuição t
, FIGARCH e FIEGARCH

6.1. APLICAÇÕES





Figura 6.30: QQ-plot dos resíduos padronizados - modelos FIEGARCH, GARCH com distribuição t e FIGARCH

(d) Conclusão

Comparando os modelos ajustados nos itens (a), (b) e (c), concluímos que o modelo ARFIMA (ajustado para a série de retornos absolutos) apresenta o melhor ajuste segundo o critérios BIC,

conforme mostra a Tabela 6.3.

	Tabela 6.3:	AIC e BI	C dos modele	s ajustados	para a	volatilidade	da série	BDC
--	-------------	----------	--------------	-------------	--------	--------------	----------	-----

Modelo	AIC	BIC
ARFIMA	-	-6912
GARCH t	-6628	-6601
FIGARCH	-6195	-6168
FIEGARCH	-6257	-6225

Capítulo 7

Conclusão e Trabalhos Futuros

Neste trabalho apresentamos os modelos ARFIMA, GARCH e GARCH com memória longa para modelar a volatilidade de séries temporais financeiras.

Para o modelo ARFIMA, apresentamos 7 estimadores para o parâmetro de memória longa e mostramos, via simulações dos modelos ARFIMA(1, d, 0), ARFIMA(0, d, 1) e ARFIMA(1, d, 1), que a melhor performance ocorre para o estimador de máxima verossimilhança.

No caso do modelo GARCH(1, 1), avaliamos o comportamento do estimador de máxima verossimilhança nas simulações, para diferentes pares dos parâmetros $a_1 e b_1$, e concluímos que quanto maior o valor de um parâmetro, menor é o EQM das estimativas do outro parâmetro, ou seja, quanto maior b_1 melhor será a estimativa de a_1 e quanto maior for o parâmetro a_1 , melhor será a estimativa de b_1 . Analisando os dois parâmetros em conjunto, observamos que as melhores estimativas da soma dos parâmetros ocorre quando o valor da soma dos parâmetro se aproxima de 1.

Para o modelo FIGARCH, avaliamos o comportamento do estimador de máxima verossimilhança para o parâmetro d nos modelos FIGARCH(1, d, 0), FIGARCH(0, d, 1) e FIGARCH(1, d, 1), e mostramos por meio de simulações que, para todos os modelos simulados, o estimador apresenta bom desempenho independentemente dos valores dos parâmetros de a_1 e b_1 .

Na aplicação a dados reais observamos que o modelo GARCH de erros com distribuição t de Student apresenta melhor ajuste para a série UBB, em termos dos critérios AIC e BIC, mesmo não levando em conta a característica de memória longa na variância condicional. Provavelmente esse resultado ocorre por conta do fato de que os retornos apresentam caudas mais pesadas que a distribuição normal e os modelos FIGARCH e FIEGARCH assumem normalidade dos resíduos. Por outro lado, para as séries ITAU e BDC observamos melhor resultado, segundo critério BIC, para o modelo ajustado ARFIMA.

Como trabalhos futuros pode-se desenvolver os algoritmos para os estimadores de Whittle e de Mínimos Desvios Absolutos para modelos GARCH apresentados no Capítulo 4, e estudar métodos de previsão para os modelos ajustados GARCH e FIGARCH. Além disso, pode-se trabalhar com os programas de simulação de modelos FIGARCH apresentados no Apêndice A para melhorar sua performance e diminuir o tempo de cada simulação.

Apêndice A

Programas

Neste anexo, apresentamos alguns dos programas utilizados na simulação e estimação dos parâmetros dos modelos estudados.

O software utilizado foi o S-Plus 6.1. Muitas das funções utilizadas são provenientes do módulo FinMetrics deste software.

A.1 Método de ondaletas para estimação do parâmetro de memória longa no modelo $\operatorname{ARFIMA}(p, d, q)$

```
###T é o número de observações #jota <- 10, 9, 8
TA <- 1024
jota <- 10
indice.t0 <- seq(0, (TA-1))
###divide por T para ficar entre 0 e 1
indice.t <- indice.t0/TA</pre>
###cria matriz de coeficientes da ondaleta
matriz.coef.wav <- matrix(0, TA, 2<sup>jota - 1</sup>)
contador <- 0
vetor.psi <- numeric()</pre>
for (j in 0:(jota-1)) {
for (k in 0:(2^j-1)) {
indice.j <- j</pre>
indice.k <- k
vetor.psi <- apply(as.array(indice.t), 1, matriz.coef.psi)</pre>
contador <- contador + 1</pre>
matriz.coef.wav[,contador] <- vetor.psi }}</pre>
###cria vetor y
vetor.y1 <- numeric()</pre>
contadory1 <- 0
###construindo o vetor y
for (j in 0:(jota-1)) {
contadory1 <- contadory1 + 1</pre>
vetor.y1[contadory1] <- log((2)^(-2*j))}</pre>
vetor.y <- numeric()</pre>
contadory <- 0
for (j in 0:(jota-1)) {
contadory <- contadory + 1</pre>
vetor.y[contadory] <- vetor.y1[contadory] - mean (vetor.y1)}</pre>
```

178

```
####calcula d.chapeu
####series.005 -> 500 séries geradas com parâmetro d=0.05 (rotina simulate.FARIMA)
d.wav.005 <- matrix(0, 1, 500)
contador <- 1
for (i in 1 : 500) {
####construindo o vetor de estimativas dos coeficientes da ondaleta
vetor.d <- series.005[,contador] %*% matriz.coef.wav</pre>
contador0 <- 1
vetor.d2 <- numeric()</pre>
for (j in 0:(jota-1)) {
for (k in 0:(2^j-1)) {
vetor.d2[contador0] <- vetor.d[contador0]</pre>
contador0 <- contador0 + 1 } }</pre>
contador1 <- 0
contador2 <- 0
vetor.r <- numeric()</pre>
for (j in 0:(jota-1)) {
contador1 <- contador1 + 1</pre>
valor.r <- 0
for (k in 0:(2^j-1)) {
contador2 <- contador2 + 1</pre>
valor.r <- vetor.d2[contador2]^2 + valor.r }</pre>
vetor.r[contador1] <- valor.r/(2^j) }</pre>
####criando log(vetor.r)
logvetor.r <- log(vetor.r)</pre>
####estimando o parâmetro de memória longa d
d.chapeu <- ((vetor.y %*% logvetor.r) / (sum(vetor.y^2)))</pre>
d.wav.005[,contador] <- d.chapeu
contador <- contador + 1}</pre>
```

A.2 Simulação do modelo FIGARCH(1, d, 1)

module(FinMetrics)

```
#####
                                          #####
                 FIGARCH(1,d,1)
#
    PARÂMETROS
              #
d<-0.45
b1<-0.4
a0<-0.1
a1<-(-0.2)
n<-2000
#
     DELTA
              #
vetor.delta<-numeric()</pre>
vetor.delta0 <- d*gamma(0-d)/(gamma(0+1)*gamma(1-d))</pre>
vetor.delta[1]<-d*gamma(1-d)/(gamma(1+1)*gamma(1-d))
for (k in 2:n){
vetor.delta[k]<-vetor.delta[k-1]*((k-1-d)/k) }</pre>
#
     LAMBDA
              #
vetor.lambda<-numeric()</pre>
vetor.lambda[1]<- a1 + vetor.delta[1]</pre>
```

```
for (i in 2:n){
vetor.lambda[i]<- vetor.delta[i] - a1*vetor.delta[i-1] }</pre>
epsilon.menos0204504.200=matrix(0,n-200,100)
sigma2.menos0204504.200=matrix(0,n-200,100)
param.figarch.menos0204504.200=matrix(0,100,3)
for (s in 1 : 100){
#
        z(t)
                      #
z<-rnorm(n+1000,0,1)
z <- z[z<3 & z>(-3)]
z<- z[1:n]
#
        EPSILON
                      #
epsilon = numeric()
epsilon2 = numeric()
sigma2 = numeric()
sigma2[1]=0.001
cte = ((1-b1)^{(-1)})*a0
epsilon[1]=z[1]*sqrt(sigma2[1])
epsilon2[1]=epsilon[1]^2
for (t in 2 : n){
      vmultiplica <-0
      for (k in 1 : (t-1)){
             vmultiplica = vmultiplica+vetor.lambda[k]*(epsilon2[t-k])
                                                                 }
      sigma2[t] = cte + vmultiplica
      epsilon[t]=z[t]*sqrt(sigma2[t])
      epsilon2[t]=(epsilon[t])^2}
```

```
epsilon.menos0204504.200[,s]<-epsilon[201:2000]
```

sigma2.menos0204504.200[,s]<-sigma2[201:2000]
param.figarch.temp = fgarch(epsilon.menos0204504.200[,s]~-1, ~figarch(1,0))
param.figarch.menos0204504.200[s,]= t(param.figarch.temp\$coef) }</pre>

Referências Bibliográficas

- Aparicio, F. M., Estrada, J. (2001). Empirical distributions of stock returns: European securities markets, 1990-1995, European Journal of Finance, 7, 1-21.
- [2] Baillie, R. T., Bollerslev, T., Mikkelsen, H. O. (1996). Fractionally Integrated Generalized Autorregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 74, 3-30.
- [3] Beran, J. (1994). Statistics for Long Memory Processes. Chapman and Hall, New York.
- [4] Black, F. (1976). Studies in stock price volatility changes, Proceedings of the 1976 Business Meeting of the Business and Economics Statistics Section, American Statistical Association, 177-181.
- [5] Blattberg, R. C., Gonedes, N. J. (1974). Comparison of the stable and student distributions as statistical models for stock prices, *Journal of Business*, 47, 244-280.
- [6] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 31, 307-327.
- [7] Bollerslev, T., Engle, R. F., Nelson, D. B. (1994). ARCH Models, in R. F. Engle and D. L. McFadden (eds.) Handbook of Econometrics, Vol. 4, North Holland Press, Amsterdan.
- [8] Bollerslev, T., Mikkelsen, H. O. (1996). Modeling and Pricing Long Memory in Stock Market Volatility, Journal of Econometrics, 73, 151-184.
- [9] Bollerslev, T., Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with timevarying covariances, *Econometric Reviews*, 11, 143-172.
- [10] Box, G.E.P., Jenkins, G.M., Reinsel, G.C. (1994). Time Series Analysis, Forecasting and Control, 3rd edn. Prentice Hall, New Jersey.
- [11] Breidt, F. J., Crato, N., de Lima. P. (1998). The Detection and Estimation of Long Memory in Stochastic Volatility, *Journal of Econometrics*, 83, 325-348.
- [12] Brockwell, P. J., Davis, R. A. (1991). Time Series: Theory and Methods, 2nd edn. Springer-Verlag, New York.
- [13] Chung, C.-F. (1999). Estimating the Fractionally Integrated GARCH Model, Working paper, National Taiwan University.

- [14] Dacorogna, M. M., Muller, U. A., Nagler, R. J., Olsen, R. B., Pictet, O. V. (1993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market, *Journal of International Money* and Finance, 12, 413-438.
- [15] de Lima, P., Breidt, F. J., Crato, N. (1994). Modeling long-memory stochastic volatility, Working Papers in Economics 323, Johns Hopkins University, Baltimore, MD.
- [16] Ding, Z., Granger, C. W. J. (1996). Modeling volatility persistence of speculative returns: A new approach, Journal of Econometrics, 73, 185-215.
- [17] Ding, Z., Granger, C. W. J., Engle, R. F. (1993). A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance*, 1, 83-106.
- [18] Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50 (4), 987-1007.
- [19] Engle, R. F., Lee, G. J. (1999). A Long-Run and Short-Run Component Model of Stock Return Volatility, in R. F. Engle and H.White (eds.), *Cointegration, Causality, and Forecasting*, Oxford University Press.
- [20] Engle, R. F., Lilien, D.M., Robins, R. P. (1987). Estimating Time Varying Risk Premia in the Term-Structure: the ARCH-M Model, *Econometrica*, 55 (2), 391-407.
- [21] Engle, R. F., Ng, V. (1993). Measuring and Testing the Impact of News on Volatility, Journal of Finance, 48 (5), 1749-1778.
- [22] Fama, E. F. (1965). The behavior of stock market prices, Journal of Business, 38, 34-105.
- [23] Fan, J., Yao, Q. (2005). Nonlinear Time Series: Nonparametric and Parametric Methods, Springer.
- [24] Fox, A. J., Taqqu, M. S. (1986). Large Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series, *The Annals of Statistics*, 14, 517-532.
- [25] Fuller, W. A. (1996). Introduction to Statistical Time Series, Second Edition. John Wiley & Sons, New York.
- [26] Geweke, J., Porter-Hudak, S. (1983). The Estimation and Application of Long Memory Time Series Models, *Journal of Time Series Analysis*, 4, 221-237.
- [27] Giraitis L., Kokoszka, P., Leipus, R. (2000). Stationary ARCH Models: Dependence Structure and Central Limit Theorem, *Econometric Theory*, 16, 3-22.
- [28] Giraitis, L., Robinson, P. M. (2001). Whittle Estimation of ARCH Models, *Econometric Theory*, 17, 608-631.
- [29] Giraitis, L., Robinson, P. M., Surgailis D. (2000). A Model for long memory conditional heteroscedasticity, Annals of Applied Probability, 10, 1002-1024.
- [30] Glosten, L. R., Jagannathan, R., Runkle, D. E. (1993). Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance*, 48, 1779-1801.

REFERÊNCIAS BIBLIOGRÁFICAS

- [31] Granger, C. W. J., Ding, Z. (1995). Some Properties of Absolute Returns, an Alternative Measure of Risk, Annual Econom. Statist., 40, 67-91.
- [32] Granger, C. W. J., Joyeux, R. (1980). An Introduction to Long-Memory Time Series Models and Fractional Differencing, *Journal of Time Series Analysis*, 1, 15-29.
- [33] Hamilton, J. D. (1994). Time Series Analysis. Princeton University Press, New Jersey.
- [34] Hannan, E. J. (1973). The Asymptotic Theory of Linear Time Series Models, Journal of Applied Probability, 10, 130-145.
- [35] Harvey, A. C. (1993). Long memory in stochastic volatility, Manuscript (London School of Economics, London).
- [36] Hayashi, F. (2000). *Econometrics*. Princeton University Press, New Jersey.
- [37] He, C., Teräsvirta, T. (1999a). Properties of Moments of a Family of GARCH Processes, Journal of Econometrics, 92, 173-192.
- [38] He, C., Teräsvirta, T. (1999b). Fourth Moment Structure of the GARCH(p,q) Process, *Econometric Theory*, 15, 824-846.
- [39] Hosking, J. R. M. (1981). Fractional Differencing, *Biometrika*, 68, 165-176.
- [40] Hurst, H. E. (1951). Long Term Storage Capacity of Reservoirs, Transactions of the American Society of Civil Engineers, 116, 770-799.
- [41] Hurst, H. E. (1957). A suggested statistical model of time series that occur in nature, Nature, 180, 494.
- [42] Jensen, M. J. (1999). Using Wavelets to Obtain a Consistent Ordinary Least Squares Estimator of the Long-Memory Parameter, *Journal of Forecasting*, 18, 17-32.
- [43] Karanasos M., Psaradakis Z., Sola M. (2001). On the Autocorrelation Properties of Long-Memory GARCH Processes, *Journal of Time Series Analysis*, 25 (2), 265-281.
- [44] Lo, A. W. (1991). Long Term Memory in Stock Market Prices, *Econometrica*, 59, 1279-1313.
- [45] Lopes, S. R. C., Mendes, B. V. M. (2006). Bandwidth selection in classical and robust estimation of long memory, *International Journal Of Statistics And Systems*, 1 (2), 177 - 200.
- [46] Mandelbrot, B. B. (1963). The variation of certain speculative prices, *Journal of Business*, 36, 394-419.
- [47] Mandelbrot, B. B. (1972). Statistical methodology for non-periodic cycles: from the covariance to R/S analysis, Annals of Economic and Social Measurement, 1(3), 259 - 290.
- [48] Mandelbrot, B. B. (1975). Limit Theorems on the Self-Normalized Range for Weakly and Strongly Dependent Processes, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 31, 271-285.
- [49] Mandelbrot, B. B., Wallis, J. (1968). Noah, Joseph and operational hydrology, Water Resources Research, 4, 909-918.

- [50] McLeod, A. I., Hipel, K. W. (1978). Preservation of the rescaled adjusted range, 1: A reassessment of the Hurst phenomenon, *Water Resources Research*, 14, 491-508.
- [51] Mikosch, T., Straumann, D. (2002). Whittle estimation in a heavy-tailed GARCH(1,1) model, Stochastic Processes and their Applications, 100, 187-222.
- [52] Mittnik, S., Rachev, S. T. (2000). Stable Paretian Models in Finance. Wiley, New York.
- [53] Mittnik, S., Rachev, S. T., Paolella, M. S. (1998). Stable Paretian modeling in finance: Some empirical and theoretical aspects, in R. J. Adler, R. E. Feldman and M. S. Taqqu (eds.): A Practical Guide to Heavy Tails: Statistical Techniques and Applications, 79-110, Birkhaüser, Boston.
- [54] Morettin, P. A., Toloi, C. M. C. (2004). Análise de Séries Temporais. Edgard Blücher, São Paulo.
- [55] Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: a New Approach, Econometrica, 59 (2), 347-370.
- [56] Nelson, D. B., Cao, C. Q. (1992). Inequality Constraints in the Univariate GARCH Model, Journal of Business and Economic Statistics, 10 (2), 229-235.
- [57] Newey, W. K., West, K. D. (1987). A Simple Positive Semidefinite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica*, 55, 703-708.
- [58] Peng, L., Yao, Q. (2003). Least absolute deviation estimation for ARCH and GARCH models, *Biometrika*, 90 (4), 967-975.
- [59] Robinson, P. M. (1991). Testing for Strong Serial Correlation and Dynamic Conditional Heteroskedasticity in Multiple Regression, *Journal of Econometrics*, 47, 67-84.
- [60] Robinson, P. M., Zaffaroni, P. (2006). Pseudo-Maximum Likelihood Estimation of ARCH(∞) Models, The Annals of Statistics, 34 (3), 1049-1074.
- [61] Taqqu, M. S., Teverovsky, V. (1998). On Estimating the Intensity of Long-Range Dependence in Finite and Infinite Variance Time Series, in R. J. Adler, R. E. Feldman and M. S. Taqqu (eds.): A Practical Guide to Heavy Tails: Statistical Techniques and Applications, Birkhaüser, Boston.
- [62] Taqqu, M. S., Teverovsky, V., Willinger, W. (1995). Estimators for Long Range Dependence: An Empirical Study, *Fractals*, 3(4), 785-798.
- [63] Teyssière, G. (1996). Double Long-Memory Financial Time Series, QMW working paper 348, Department of Economics, University of London, London.
- [64] Tsay, R. S. (2002). Analysis of Financial Time Series. Wiley, New York.
- [65] Vidakovic, B., Müller, P. (1991). Wavelets for Kids: A Tutorial Introduction, Institute of Statistics and Decision Sciences, Duke University.
- [66] Weiss, A. A., (1986). Asymptotic theory for ARCH models: Estimation and testing, *Econometric Theory*, 2, 107-131.
- [67] White, H. (1984). Asymptotic Theory for Econometricians. Academic Press, San Diego.

REFERÊNCIAS BIBLIOGRÁFICAS

- [68] Zaffaroni, P. (2000). Stationarity and Memory of $ARCH(\infty)$ Models, Unpublished Manuscript, Bank of Italy.
- [69] Zivot, E., Wang, J. (2003). Modelling Financial Time Series with S-PLUS. Springer, New York.