

UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA  
DEPARTAMENTO DE ESTATÍSTICA

# Desconvolução Não-Paramétrica Aplicada a Modelos de Volatilidade Estocástica

**Tatiana Andréa Benaglia**

**Orientador: Prof. Dr. Aluísio de Souza Pinheiro**

Dissertação apresentada junto ao Departamento de Estatística do Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas, para obtenção do Título de Mestre em Estatística.

Campinas - SP

2004

# Desconvolução Não-Paramétrica Aplicada a Modelos de Volatilidade Estocástica

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Tatiana Andréa Benaglia e aprovada pela comissão julgadora.

Campinas, 30 de abril de 2004.

---

Prof. Dr. Alúcio de Souza Pinheiro  
Orientador

Banca Examinadora:

1. Prof. Dr. Alúcio de Souza Pinheiro (Orientador) - IMECC/UNICAMP
2. Profa. Dra. Sílvia Regina Costa Lopes - UFRGS
3. Prof. Dr. Luiz Koodi Hotta - IMECC/UNICAMP

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para obtenção do Título de Mestre em Estatística.

Aos meus pais, Aduino e Alzira.

# Agradecimentos

À Deus, por estar sempre ao meu lado e me permitir alcançar mais esse objetivo.

Aos meus pais, por me apoiarem em todas as minhas decisões e me darem todo o suporte necessário, muitas vezes adiando a concretização de seus planos para que eu realizasse os meus. O amor que vocês me dão me dá forças para transpor qualquer obstáculo. Eu amo vocês demais, lhes serei eternamente grata e sinto um orgulho imenso em poder chamá-los de meus pais. Um sincero agradecimento também a Taís, minha amada irmã.

Ao Benilton Carvalho, por se fazer presente em todos os momentos mesmo naqueles em que estive distante. Obrigada por sorrir comigo nos momentos de alegria e por me animar nas horas difíceis. Suas contribuições acadêmicas, essenciais para o desenvolvimento desse trabalho, não têm preço. Seu incentivo constante só me faz olhar pra frente e seguir cada vez mais adiante. Duas palavras definem bem o que você é: companheiro e amigo, incondicional.

Ao Prof. Aluísio Pinheiro, meu orientador, por toda paciência, amizade e ajuda ao longo desses anos de trabalho. Suas correções, sempre bem humoradas, e seus ensinamentos foram fundamentais para a realização dessa dissertação e para o meu crescimento intelectual.

Aos Professores Sílvia Lopes e Luis Hotta, por terem aceitado fazer parte da banca examinadora e pelas inúmeras contribuições que só melhoraram a qualidade desse trabalho.

Ao Departamento de Estatística, de forma especial aos professores Hildete Pinheiro, Eliana Marques, Nancy Garcia, Ronaldo Dias, Filidor Labra, por sua amizade e orientações, fundamentais para minha boa formação pessoal e acadêmica. Claro que eu não deixaria de mencionar alguém a quem tenho um carinho imenso, prof. Mauro Marques. Obrigada por me passar um pouquinho da sua sabedoria e por me dar tanta atenção e carinho.

---

Aos meus amigos Artur Iuri, Clarissa Feliciano, Eliane Fusca, Laura Lupic nia, Leandro Sanchez, Leilaine Oliveira, Marianne Macedo, Patr cia Viana e Sandra Wakaguri, simplesmente por serem quem s o. Saibam que eu os considero parte da minha fam lia e voc es estar o sempre comigo, independente de onde eu estiver. Um obrigada especial ao Helder Palaro, por suas contribui es diretas a esse trabalho e por seu exemplo de disciplina e ao Ricardo Takeyama, por sua amizade, serenidade e companhia nas noites em fic vamos at  tarde no laborat rio terminando nossas respectivas disserta es.

Aos que participaram de forma mais constante no meu dia-a-dia, seja no laborat rio, em casa ou num momento de descontra o: Alexandre R besam, C zar Anselmo, Cl cio Ferreira, Edimilson Lucas, Neale El-Dash, Nelson Lopes, Paula Moura, Rodrigo Tsai, Rossana Lopez, Roberta e Anderson de Souza, Samara Kiihl, Lilian Hanamoto, Adriana e Emerson Louren o. Voc es certamente contrib iram para que essa fase fosse menos estressante. A conviv ncia com voc es j  deixa saudades.

  CAPES, pelo suporte financeiro, fundamental para o desenvolvimento desse projeto.

*“Statistics (...) the most important science in the whole world: for upon it depends the practical application of every other science and of every art: the one science essential to all political and social administration, all education, all organization based on experience, for it only gives results of our experience.”*

Florence Nightingale - Statistician

# Resumo

É de grande interesse o estudo da evolução de uma série financeira, ou seja, entender o comportamento dos preços do mercado financeiro. Dois dos principais objetivos no estudo dessas séries são interpretação e predição. Interpretar uma série financeira significa entender o comportamento dos preços de um modo geral. Conseguir identificar características do processo num ponto futuro caracteriza o problema de predição. Para isso, é necessário conhecer, ou estimar com precisão, o comportamento da volatilidade, pois essa quantidade está diretamente associada à variabilidade dos preços ou retornos.

Modelos bem conhecidos na literatura, como os de Volatilidade Estocástica, predizem a volatilidade considerando as observações passadas, conhecida como volatilidade condicional. Neste trabalho, o principal interesse recai sobre predições a médio e longo prazos e para isso utiliza-se a noção incondicional da volatilidade.

A densidade de volatilidade incondicional é estimada através de técnicas de desconvolução. Tal técnica é muito utilizada em modelos de regressão com erros nas variáveis e é baseada em métodos não-paramétricos para estimação de curvas como função-núcleo e ondaletas. Comte (2000) apresenta uma aplicação do estimador por desconvolução via função-núcleo em Modelos de Volatilidade Estocástica, vistos na forma de uma regressão com erros nas variáveis. Fan e Koo (2002) apresentam estimadores lineares e não-lineares por desconvolução via ondaletas e discutem suas propriedades. Nós utilizamos a idéia de Comte (2000), escrevendo o Modelo de Volatilidade Estocástica como uma regressão, e empregamos a proposta de Fan e Koo (2002) para estimar a densidade da volatilidade.

Estuda-se o comportamento dos estimadores por simulações e mede-se o desempenho de cada técnica através do Erro Quadrático Médio e de medidas de distância tais como a distância em  $\mathcal{L}^2$  e a distância de Hellinger. Na desconvolução via ondaletas, estudam-se também quão próximas da verdadeira densidade estão suas melhores aproximações por multi-resolução tanto para séries geradas a partir de um modelo AR(1)-SV quanto um GARCH(1,1). A metodologia é aplicada na série diária de taxa de câmbio Iene/Dólar no período de 02 de janeiro de 1990 a 05 de dezembro de 2003.

# Abstract

Series of financial returns are a topic of great interest nowadays. Two of the main concerns in their study are the characterization (estimation) and the future return evaluation (prediction). For both paradigms one is faced with the problem of using very precise estimation of volatility values.

Well-known models in the literature, such as Stochastic Volatility Models, attack the problem by predicting the volatility as a random variable partially dependent on the past observations. Therefore, they are considered conditional volatility models.

If one is interested in medium or long range predictions, it is wise to avoid the use of conditional volatility. We study in this work the unconditional volatility density from a deconvolution non-parametric point of view.

Comte (2000) presents a kernel deconvolution estimator for Stochastic Volatility Models on a regression model with errors in the variable setup. Fan e Koo (2002) proposed linear and non-linear deconvolution wavelets estimators.

Due to its implementation difficulties (numerical deconvolution and complex integrals), both papers present only the theoretical aspects of the techniques. We develop the ideas on simulational studies and on application to real data. Performance comparisons are made through IMSE and Hellinger distance as well as discussions on actual versus nominal rates of convergence and numerical feasibilities. The techniques are illustrated on the daily Yen/Dolar exchange rates from 01/02/1990 to 12/05/2003.

# Conteúdo

<b>Lista de Tabelas</b>	<b>xi</b>
<b>Lista de Figuras</b>	<b>xii</b>
<b>Introdução</b>	<b>1</b>
<b>1 Volatilidade em Séries Financeiras</b>	<b>4</b>
1.1 Séries Financeiras . . . . .	4
1.2 Fatos Estilizados . . . . .	5
1.3 Modelos de Volatilidade . . . . .	7
1.3.1 Modelo de Volatilidade Estocástica . . . . .	9
1.4 Predição . . . . .	10
1.4.1 Predição em Modelos ARMA . . . . .	11
<b>2 Análises de Fourier e de Ondaletas</b>	<b>15</b>
2.1 Transformada de Fourier . . . . .	15
2.1.1 Transformada Discreta de Fourier . . . . .	17
2.2 Análises de Multiresolução . . . . .	18
2.3 Espaços de Besov e Ondaletas . . . . .	19
2.3.1 Construção das Funções de Ondaletas . . . . .	21
2.3.2 Representação das Ondaletas por Filtros . . . . .	23
2.3.3 Regularidade das Ondaletas . . . . .	27

---

2.3.4	Bases de Ondaletas . . . . .	28
2.3.4.1	Base de Haar . . . . .	29
2.3.4.2	Ondaletas de Meyer . . . . .	31
2.3.4.3	Ondaletas de Daubechies . . . . .	33
<b>3</b>	<b>Estimadores por Desconvolução</b>	<b>37</b>
3.1	Classes de Erro . . . . .	37
3.2	Função-Núcleo . . . . .	38
3.2.1	Conceitos de Função-Núcleo . . . . .	39
3.2.1.1	Estimador de Nadaraya-Watson . . . . .	42
3.3	Desconvolução via Função-Núcleo . . . . .	43
3.4	Desconvolução via Ondaletas . . . . .	45
3.4.1	Estimadores Lineares . . . . .	45
3.4.2	Estimadores Não-lineares . . . . .	51
<b>4</b>	<b>Simulações</b>	<b>54</b>
4.1	Função-Núcleo . . . . .	54
4.2	Ondaletas . . . . .	60
4.2.1	Família Daubechies . . . . .	62
4.2.1.1	Erros Regulares Ordinários . . . . .	62
4.2.1.2	Erros Super-Regulares . . . . .	64
4.2.2	Ondaletas de Meyer . . . . .	66
4.2.2.1	Erros Regulares Ordinários . . . . .	66
4.2.2.2	Erros Super-Regulares . . . . .	67
4.2.3	Séries Simuladas - GARCH(1,1) e AR(1)-SV . . . . .	68
<b>5</b>	<b>Aplicação em Dados Reais</b>	<b>74</b>

Conteúdo	x
<b>Considerações Finais</b>	<b>79</b>
<b>Bibliografia</b>	<b>81</b>

# Lista de Tabelas

2.1	Funções $\nu$ para $\theta = 1, \dots, 5$ . . . . .	31
3.2	Exemplos de Funções-Núcleo . . . . .	40
4.3	Média dos EQM's para Diferentes Valores de $h$ . . . . .	58
4.4	Parâmetro de Alisamento Ótimo para os Modelos $L(a)$ , $S(a)$ e $FY(a)$ . . . . .	59
4.5	Erros por $DAUB8$ sem Limiar - Regulares Ordinários . . . . .	63
4.6	Erros por $DAUB8$ com Limiar - Regulares Ordinários . . . . .	63
4.7	Erros por $DAUB8$ sem Limiar - Super-Regulares . . . . .	64
4.8	Erros por $DAUB8$ com Limiar - Super-Regulares . . . . .	65
4.9	Erros por Meyer - Regulares Ordinários . . . . .	66
4.10	Erros por Meyer - Super-Regulares . . . . .	67
4.11	Desempenho dos Estimadores . . . . .	73

# Lista de Figuras

1.1	Cotação Diária Iene/Dólar - 02/01/1990 a 05/12/2003 . . . . .	6
1.2	Gráfico dos Log-Retornos e Retornos Quadrados da Série Iene/Dólar	6
1.3	Gráfico de Normalidade dos Retornos para a Série Iene/Dólar . . . . .	7
2.4	Bases de Haar: (a) Função Escala e (b) Função de Detalhe . . . . .	30
2.5	Ondaletas de Meyer no Domínio da Freqüência: (a) $\phi^*(\omega)$ e (b) $ \psi^*(\omega) $	32
2.6	Ondaletas de Meyer: (a) $\phi(x)$ e (b) $\psi(x)$ . . . . .	33
2.7	Ondaletas <i>DAUB8</i> : (a) $ \phi^*(\omega) $ e (b) $ \psi^*(\omega) $ . . . . .	35
2.8	Bases de Daubechies . . . . .	35
3.9	Exemplos de Funções-Núcleo . . . . .	41
4.10	Estimativas para Modelo $S(0,7)$ , $h = 0, 12; 0, 16; 0, 20; 0, 30$ . . . . .	58
4.11	Estimativas para Modelo $FY(0,5)$ , $h = 0, 12; 0, 16; 0, 20; 0, 30$ . . . . .	59
4.12	Estimativas de Densidade por <i>DAUB8</i> , $n=128, 1024$ e $4096$ . . . . .	62
4.13	Funções de Distribuição Acumuladas por <i>DAUB8</i> , $n=128, 1024$ e $4096$	62
4.14	Estimativas de Densidade por <i>DAUB8</i> com Limiar . . . . .	63
4.15	Funções de Distribuição Acumuladas por <i>DAUB8</i> com Limiar . . . . .	63
4.16	Estimativas de Densidade por <i>DAUB8</i> , $n=128, 1024$ e $4096$ . . . . .	64
4.17	Funções de Distribuição Acumuladas por <i>DAUB8</i> , $n=128, 1024$ e $4096$	64
4.18	Estimativas de Densidade por <i>DAUB8</i> com Limiar . . . . .	65
4.19	Funções de Distribuição Acumuladas por <i>DAUB8</i> com Limiar . . . . .	65
4.20	Estimativas de Densidade por Meyer, $n=128, 1024$ e $4096$ . . . . .	66
4.21	Funções de Distribuição Acumuladas por Meyer, $n=128, 1024$ e $4096$	66

---

4.22	Estimativas de Densidade por Meyer, n=128, 1024 e 4096 . . . . .	67
4.23	Funções de Distribuição Acumuladas por Meyer, n=128, 1024 e 4096	67
4.24	Retornos gerados por AR(1)-SV e GARCH(1,1) . . . . .	70
4.25	Volatilidade gerada por AR(1)-SV e GARCH(1,1) . . . . .	70
4.26	Estimativa da Densidade da Log-Volatilidade - AR(1)-SV . . . . .	71
4.27	Estimativa da Densidade da Volatilidade - AR(1)-SV . . . . .	71
4.28	Estimativa da Densidade da Log-Volatilidade - GARCH(1,1) . . . . .	72
4.29	Estimativa da Densidade da Volatilidade - GARCH(1,1) . . . . .	72
4.30	Desempenho do Estimadores . . . . .	73
5.31	Gráfico dos Log-Retornos e Retornos Quadrados da Série Iene/Dólar	74
5.32	Estimativas de Densidade para a Série Completa . . . . .	75
5.33	Estimativas de Densidade para as séries de tamanho 512 . . . . .	76
5.34	Estimativas de Densidade para as séries de tamanho 1024 . . . . .	76
5.35	Estimativas de Densidade para as séries de tamanho 2048 . . . . .	77
5.36	$D_{\mathcal{L}^2}$ e $D_H$ entre Estimativas de Densidades para n=512 . . . . .	77
5.37	$D_{\mathcal{L}^2}$ e $D_H$ entre Estimativas de Densidades para n=1024 . . . . .	78
5.38	$D_{\mathcal{L}^2}$ e $D_H$ entre Estimativas de Densidades para n=2048 . . . . .	78

# Introdução

Entender o comportamento dos *preços* de um produto do mercado financeiro e ter a capacidade de prever características do processo são dois dos objetivos principais no estudo de séries financeiras. Séries financeiras são seqüências de valores negociados, como, por exemplo, taxas de câmbio, taxas de juros, ativos e cotações da bolsa de valores.

Existem modelos muito difundidos na literatura para a análise de séries temporais, como os modelos ARMA vistos em Wei (1990), de fácil e eficiente implementação. No entanto, esses modelos não são capazes de reproduzir algumas características presentes nas séries, como por exemplo a não-linearidade. Dessa forma, costuma-se analisar uma função da variação do preço de dois dias consecutivos, conhecida como retorno. A série de retornos é estacionária e geralmente não-correlacionada e tem propriedades estatísticas mais interessantes do que a série de preços.

As séries de retornos geralmente apresentam características peculiares como, por exemplo, distribuição com caudas pesadas; medidas de variabilidade; estacionaridade; correlação entre os log-retornos; e não-linearidade. Essas características, ilustradas na Seção 1.2, são conhecidas como *Fatos Estilizados* e os modelos ARMA não são capazes de descrevê-las ou reproduzi-las.

Na implementação de um modelo para o preço justo de um produto, é fundamental estimar com precisão o valor de uma quantidade que é, em geral, inobservável: a volatilidade. A volatilidade está diretamente associada à variabilidade dos retornos e, por isso, é interpretada como uma medida do risco associado a cada produto. Os modelos que se propõem a estimar essa variância são conhecidos como Modelos de Volatilidade. Estes são modelos paramétricos e diferem quanto ao tratamento que dão à volatilidade: tratando-a como uma variável observável ou latente (inobservável).

Duas famílias de modelos bastante utilizadas são as do tipo ARCH (Autore-

gressivos de Variância Condicional Heteroscedástica) e suas extensões, tendo como referências Engle (1982), Bollerslev (1986) e Bollerslev, Engle e Nelson (1994); e os SV (Volatilidade Estocástica), primeiramente proposto por Taylor (1980) cujas características principais são detalhadas na Seção 1.3.1. Entre os modelos da família ARCH, o mais utilizado nas aplicações é o GARCH(1,1) e, entre os da família SV, destaca-se o AR(1)-SV. Ambos são capazes de reproduzir teoricamente as características observadas nos retornos das séries financeiras e de explicar características comuns às séries financeiras como: alta curtose, correlação no quadrado das perturbações, perturbações não-correlacionadas e conglomerados de volatilidade. Os modelos ARCH, consideram a volatilidade observável e recorrem ao procedimento de expressá-la como função dos retornos passados. Os modelos SV supõem que a volatilidade presente, considerada uma variável latente, depende parcialmente de seus valores passados, mas é independente dos retornos passados.

A metodologia apresentada nesse trabalho tem por finalidade auxiliar no problema de predição, ou seja, conseguir identificar características da distribuição do processo num ponto futuro. Pode-se considerar o problema de predição, dependendo do instante de tempo à frente cuja predição queremos fazer, como predição a curto, médio ou longo prazos.

Os modelos da família ARCH e SV utilizam a volatilidade condicional para fazer previsões. Um caso bem conhecido na literatura e utilizado na prática é a previsão um passo à frente. No entanto, quando o horizonte de predição é grande, seria mais plausível utilizar a volatilidade incondicional, ou seja, uma quantidade constante ao longo do tempo, como pode ser visto nos mais antigos modelos de séries temporais, como por exemplo, os modelos ARMA.

Embora o modelo ARMA seja simples demais para a análise de séries financeiras, utilizamo-no na Seção 1.4 para ilustração. Calculamos algumas quantidades e, pela facilidade de obtermos o erro de predição e sua variância, motivamos a necessidade do estudo incondicional da volatilidade. Mostramos que, para o curto prazo, o emprego da predição condicional é indicada; já para o caso de médio e longo prazos, a predição incondicional pode ser superior, tanto pela complexidade da estimação quanto pelo fato que, se olharmos uma série por um longo período de tempo, é comum ocorrerem mudanças bruscas, o que aumentaria significativamente o erro da predição condicional, sendo esta parametricamente baseada nas observações passadas. A predição incondicional estimaria todos os valores futuros como sendo a esperança da série, o que pode resultar em um erro muito menor do que se usarmos

um preditor condicional.

A densidade de volatilidade incondicional é estimada através de técnicas de desconvolução. Tal técnica é muito utilizada em modelos de regressão com erros nas variáveis e é baseada em métodos não-paramétricos para estimação de curvas como através de função-núcleo ou ondaletas. Detalhes sobre essas duas famílias de estimadores não-paramétricos são discutidos nas Seções 2.3 e 3.2. Comte (2000) apresenta uma aplicação do estimador por desconvolução via função-núcleo em Modelos de Volatilidade Estocástica, apresentado na Seção 3.4, visto na forma de uma regressão com erros nas variáveis. Fan e Koo (2002) apresentam estimadores lineares e não-lineares por desconvolução via ondaletas e discutem suas propriedades teóricas, como pode ser visto na Seção 3.4. Nós utilizamos a idéia de Comte (2000), escrevendo o Modelo de Volatilidade Estocástica como uma regressão, e empregamos a proposta de Fan e Koo (2002) para estimar a densidade da volatilidade. Pelas dificuldades numéricas (desconvolução numérica e integrais complexas), esta é a primeira em vez que os estimadores por desconvolução via ondaletas são implementados. Fan e Koo (2002) apresentam apenas aspectos teóricos da técnica.

É importante ressaltar aqui que essas técnicas serão utilizadas com propósitos diferentes. A desconvolução via função-núcleo é utilizada apenas como ilustração. Já a desconvolução via ondaletas é empregada com o objetivo de estimar a densidade de probabilidade da volatilidade. As propriedades dos estimadores são discutidas de acordo com a classe de erro do modelo. Duas classes de erro são utilizadas e suas definição e características são apresentadas na Seção 3.1.

No Capítulo 4 são apresentadas as simulações realizadas, tanto para o caso de desconvolução via função-núcleo quanto por ondaletas. Estuda-se o comportamento dos estimadores por simulações e mede-se o desempenho de cada técnica através do Erro Quadrático Médio Integrado, distância  $\mathcal{L}^2$  e distância de Hellinger. A função-núcleo será usada apenas para um caso específico de erros, pois devido à complexidade dos cálculos, não é possível utilizá-lo no caso geral. Faz-se um estudo mais detalhado para ondaletas. Séries de retornos são geradas, tanto por um modelo GARCH(1,1) quanto por um AR(1)-SV, para as quais a densidade da log-volatilidade ou da volatilidade do modelo, apresentado na Subseção 1.3.1, é estimada. No Capítulo 5, aplicamos a metodologia proposta para a série diária de taxa de câmbio Iene/Dólar e verificamos a consistência do estimador para a densidade da volatilidade.

# 1 Volatilidade em Séries Financeiras

## 1.1 Séries Financeiras

Existem vários modelos muito difundidos na literatura para a análise de séries temporais, como os modelos ARMA, que são de fácil e eficiente implementação. Embora de grande utilidade, esses modelos não são capazes de reproduzir algumas características, como por exemplo a não-linearidade presente nas séries. Dessa forma, para tornar a série estacionária, costuma-se analisar uma função da variação do preço de dois dias consecutivos. Essas quantidades são conhecidas como retornos, que são normalmente não-correlacionados. Os retornos têm propriedades estatísticas mais interessantes do que os preços. Uma das definições mais empregadas é a chamada de retornos compostos continuamente ou log-retornos. Considerando a série financeira  $\{P_t\}_{t \geq 1}$ , os log-retornos são definidos por:

$$R_t = \ln \left( \frac{P_t}{P_{t-1}} \right) = \ln(P_t) - \ln(P_{t-1}).$$

Nessa dissertação, trabalharemos com essa definição e chamaremos simplesmente de retorno.

As séries de retornos geralmente apresentam características peculiares como, por exemplo, distribuição com caudas pesadas; medidas de variabilidade; estacionaridade; correlação entre os log-retornos; e não-linearidade. Essas características

são conhecidas como *Fatos Estilizados* e os modelos ARMA não são capazes de descrevê-las ou reproduzi-las.

Na implementação de um modelo para o preço justo de um produto, é fundamental estimar com precisão o valor de uma quantidade que é, em geral, inobservável, a volatilidade. Informalmente, define-se a volatilidade como o grau de variabilidade que apresentam os retornos, o que é interpretada como uma medida do risco associado a cada produto. Os modelos que se propõem a estimar essa variância são conhecidos como Modelos de Volatilidade. Existem vários modelos paramétricos muito citados na literatura como os Modelos Autoregressivos de Variância Condicional Heteroscedástica (ARCH) e os Modelos de Volatilidade Estocástica (SV). Os modelos ARCH, vistos em Bollerslev, Engle e Nelson (1994), conseguem reproduzir algumas dessas características empíricas encontradas nas séries de retornos.

## 1.2 Fatos Estilizados

Como citado anteriormente, as séries de retornos possuem uma série de características empíricas comuns, chamadas de *Fatos Estilizados*. Nesta seção, são mostradas algumas dessas características, ilustradas por 3503 observações diárias da taxa de câmbio Iene/Dólar, no período de 02 de janeiro de 1990 a 05 de dezembro de 2003. Essa série também será utilizada para aplicação da metodologia discutida no trabalho.

Analisando-se a série original e os retornos nas Figuras 1.1 e 1.2, respectivamente, pode-se notar que a série é não estacionária, mas a série dos retornos tem nível constante.

Em geral, as séries de retornos não apresentam autocorrelação, mas as séries dos quadrados dos retornos a apresentam, de forma significativa, com decaimento lento para instantes superiores, ou seja, os retornos são serialmente dependentes. Essa dependência na distribuição dos retornos, em que existem períodos de alta (baixa) volatilidade, onde a variabilidade tende a permanecer alta (baixa), é conhecida como **conglomerados de volatilidade**.

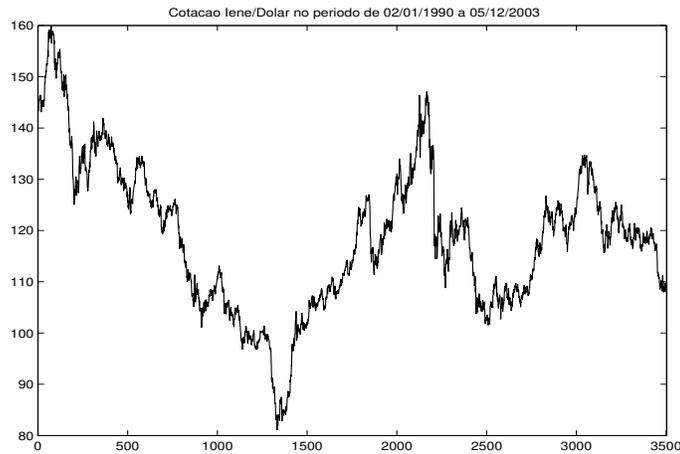


Figura 1.1: Cotação Diária Iene/Dólar - 02/01/1990 a 05/12/2003

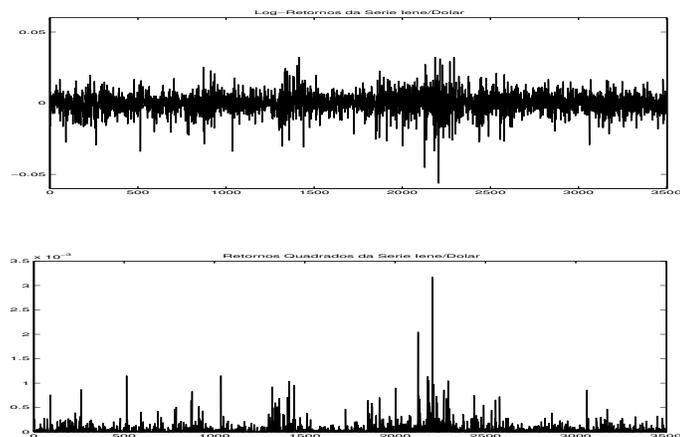


Figura 1.2: Gráfico dos Log-Retornos e Retornos Quadrados da Série Iene/Dólar

Outro fenômeno típico de séries financeiras é o **Efeito Alavancagem**, que faz as variações nos preços serem correlacionadas com a volatilidade, no sentido de que, quando os preços de um determinado produto financeiro estão em queda, sua volatilidade tende a ser maior do que em períodos de alta.

Em séries com alta e média frequência, como por exemplo de dados diários, observa-se uma alta persistência da volatilidade, tendendo a permanecer próxima à observada no instante anterior. Esse fato é conhecido como **memória longa**.

Outra característica dos retornos é a que eles apresentam uma **distribuição leptocúrtica**, ou seja, possuem uma distribuição com caudas mais pesadas do que

a da distribuição normal, o que se traduz em que valores longe da média ocorrem com uma frequência relativamente alta para um modelo ARAMA (Figura 1.3).

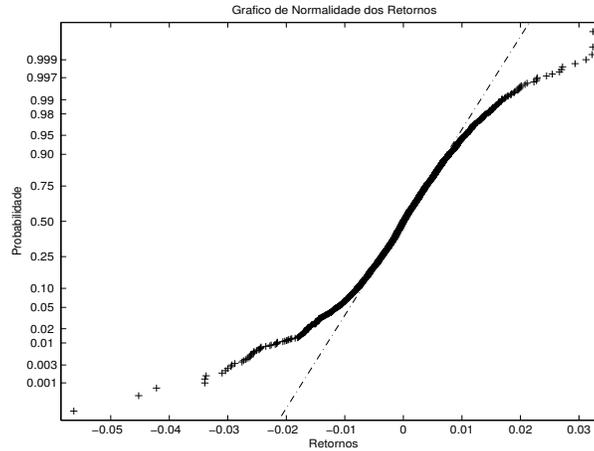


Figura 1.3: Gráfico de Normalidade dos Retornos para a Série Iene/Dólar

## 1.3 Modelos de Volatilidade

Considere a decomposição dos retornos em duas componentes seriais: uma de níveis funcionalmente determinísticos e a outra de variáveis aleatórias (v.a.'s), chamadas de erros ou perturbações aleatórias, isto é:

$$R_t = \mu_t + \epsilon_t,$$

em que  $\mu_t$  é o nível da série e as perturbações são tais que

$$\mathbb{E}(\epsilon_t) = 0 \text{ e } \mathbb{E}(\epsilon_t \mu_t) = 0, \quad \forall t.$$

As incertezas dos retornos estão diretamente ligadas à variabilidade dessa sequência de perturbações, ou seja, à variância em cada instante de tempo. A variância da perturbação num determinado instante de tempo  $t$ , condicionada à informação das observações até o instante  $t - 1$ , é chamada de volatilidade no tempo  $t$ .

Assim, define-se a volatilidade no instante  $t$  como

$$\sigma_t^2 = \text{Var}(R_t | \mathcal{F}_{t-1}), \quad (1.1)$$

em que  $\mathcal{F}_{t-1}$  é a informação até o tempo  $t - 1$ .

Os retornos têm média constante, ou seja,  $\mathbb{E}(R_t | \mathcal{F}_{t-1}) = 0$  e (1.1) se reduz a:

$$\sigma_t^2 = \mathbb{E}(R_t^2 | \mathcal{F}_{t-1}). \quad (1.2)$$

As regularidades apresentadas pelos retornos das séries financeiras podem ser classificadas de acordo com as características próprias das distribuições, como forma e momentos, ou quanto à evolução temporal retratada pela estrutura de autocorrelação e da natureza não-linear do processo gerador dos dados. Os *Fatos Estilizados* que caracterizam os retornos não podem ser explicados pelos modelos tradicionais de séries temporais lineares, já citados anteriormente, como o ARMA. O tratamento dessas séries requer a implementação de outros tipos de modelos que sejam capazes de reproduzir essas características empíricas.

Existem na literatura vários modelos propostos para o estudo da volatilidade, que são conhecidos como Modelos de Volatilidade. Estes são modelos paramétricos e diferem quanto ao tratamento que dão à volatilidade, ou seja, tratando-a como uma variável observável ou latente (inobservável).

Duas famílias de modelos bastante utilizadas são as do tipo ARCH (e suas extensões), e as do tipo SV. Dentro da família dos modelos do tipo ARCH, o mais utilizado nas aplicações é o GARCH(1,1) e, dentro da família dos SV, o AR(1)-SV. Ambos são capazes de reproduzir teoricamente as características observadas nos retornos das séries financeiras e de explicar características comuns às séries financeiras como alta curtose, correlação no quadrado das perturbações, perturbações não-correlacionadas e conglomerados de volatilidade. Nos modelos ARCH, a volatilidade é considerada observável e, nos modelos SV, a volatilidade é uma variável latente. Os modelos estatísticos que consideram a volatilidade como observável recorrem ao procedimento de expressá-la como função dos retornos.

Apresentaremos aqui com maiores detalhes apenas o modelo de Volatilidade Estocástica que será utilizado nesse trabalho. Para os modelos ARCH, ótimas referências são Engle (1982) e Bollerslev (1986).

### 1.3.1 Modelo de Volatilidade Estocástica

Os modelos de Volatilidade Estocástica, conhecidos como SV (Stochastic Volatility), primeiramente propostos por Taylor (1980), não supõem que a volatilidade condicional dependa dos retornos passados como nos modelos da família ARCH. Este modelo supõe que a volatilidade presente depende de seus valores passados, mas é independente dos retornos passados.

Uma série de retornos  $\{R_t\}_{t \geq 1}$  é dita seguir um modelo de volatilidade estocástica se:

$$R_t = \sigma_t \xi_t, \quad (1.3)$$

$$\sigma_t = \exp(X_t/2), \quad (1.4)$$

em que  $\{\xi_t\}_{t \geq 1}$  é uma seqüência estacionária, com média 0 e variância 1, e  $\{X_t\}_{t \geq 1}$  é uma seqüência que pode ser estacionária ou não, com uma densidade de probabilidade  $f_{X_t}$ .

A forma mais simples do modelo supõe que o logaritmo da volatilidade segue uma forma linear, isto é,

$$\ln(\sigma_t^2) = \alpha_0 + \alpha_1 \ln(\sigma_{t-1}^2) + \eta_t, \quad (1.5)$$

ou, em termos de  $X_t$ , como

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \eta_t, \quad (1.6)$$

em que  $\eta_t$  segue uma distribuição  $N(0, \sigma_\eta^2)$ . O parâmetro  $\alpha_1$  deve satisfazer  $|\alpha_1| \leq 1$ .

De forma geral, temos que

$$X_t = m(X_{t-1}) + \eta_t. \quad (1.7)$$

Comte (2000) mostra como estimar a função  $m$  através de desconvolução via função-núcleo. Empregando desconvolução via ondaletas, é possível estimar  $f_{X_t}$  e, conseqüentemente, estimaríamos a densidade de  $\sigma_t^2$ . Detalhes sobre esses métodos serão vistos no Capítulo 3.

Se assumirmos que  $\xi_t \sim N(0, 1)$ , então  $\ln(\xi_t^2)$  segue uma distribuição log-qui-

quadrado com  $\mathbb{E}(\ln(\xi_t^2)) \simeq -1,2704$  e  $Var(\ln(\xi_t^2)) = \pi^2/2$ .

De (1.3) e (1.4), obtemos:

$$\ln(R_t^2) = \ln(\sigma_t^2) + \ln(\xi_t^2), \quad (1.8)$$

$$X_t = \ln(\sigma_t^2). \quad (1.9)$$

Definindo-se  $w = \mathbb{E}(\ln(\xi_t^2))$ ,

$$\begin{aligned} Y_t &= \ln(R_t^2) - w \quad \text{e} \\ \varepsilon_t &= \ln(\xi_t^2) - w \simeq \ln(\xi_t^2) + 1,27, \end{aligned}$$

tem-se o modelo:

$$X_t = m(X_{t-1}) + \eta_t, \quad \eta_t \sim i.i.d. N(0, \sigma_\eta^2) \quad (1.10)$$

$$Y_t = X_t + \varepsilon_t, \quad \varepsilon_t \sim i.i.d. (0, \pi^2/2), \quad (1.11)$$

com  $\{\eta\}_{t \geq 1}$  e  $\{\varepsilon_s\}_{s \geq 1}$  processos estocásticos independentes.

Note que  $\mathbb{E}(\varepsilon_t) = 0$  e  $Var(\varepsilon) = \pi^2/2$ . Denotaremos  $f_\eta$  e  $f_\varepsilon$  as densidades de probabilidade de  $\eta$  e  $\varepsilon$ , respectivamente.

Os únicos valores observáveis no modelo acima são os  $Y_t$ 's, já que são função dos retornos. Dessa forma, o modelo (1.10)-(1.11) pode ser considerado como um caso particular de um modelo de regressão com erro nas variáveis:

$$Z_t = m(X_t) + \eta_t \quad (1.12)$$

$$Y_t = X_t + \varepsilon_t, \quad (1.13)$$

sob as mesmas suposições sobre os ruídos e os dois processos  $(Y, Z)$  sendo observados para  $t = 1, \dots, n$ . No nosso caso, no entanto, apenas um dos processos é observado.

## 1.4 Predição

Como citado na Seção 1.1, em análise de séries financeiras, um dos principais objetivos é conseguir identificar características da distribuição do processo num ponto futuro, sendo esse procedimento conhecido como predição. Para isso, é fundamental

conhecer, ou estimar com precisão, a volatilidade, pois essa variável está diretamente associada à variabilidade dos preços.

Suponha que tenhamos uma série de preços diários até o tempo  $n$  e desejemos fazer previsão para o tempo  $n+l$ , conhecida como previsão  $l$  passos à frente. Quando esse valor  $l$  é pequeno, é interessante utilizar a volatilidade condicional, ou seja, a variância dos retornos condicionada às informações do passado. Um caso bem conhecido na literatura e utilizado na prática é a previsão um passo à frente, ou seja, utiliza-se a volatilidade no tempo  $n$  para estimar a volatilidade no tempo  $n+1$ . Os modelos da família ARCH e SV utilizam a volatilidade condicional para fazer previsões.

No entanto, quando o valor de  $l$  é grande, seria mais plausível utilizar a volatilidade incondicional, ou seja, uma quantidade constante ao longo do tempo, como pode ser visto nos mais clássicos modelos de séries temporais, como por exemplo, os modelos ARMA. Sabemos que, em toda previsão, existe um erro associado e ocorre a propagação desse erro à medida em que o tempo dos valores a serem previstos se distancia do tempo do último valor observado.

### 1.4.1 Predição em Modelos ARMA

Para ilustrar o que ocorre em previsões a curto, médio e longo prazo, iremos utilizar um dos modelos mais comuns de séries financeiras, o modelo ARMA(p,q), ou seja, o processo  $Z_t$  é escrito na forma:

$$\phi_p(B)(Z_t - \mu) = \theta_q(B)a_t, \quad (1.14)$$

em que  $a_t$  é um ruído branco com  $\mathbb{E}(a_t) = 0$  e  $Var(a_t) = \sigma_a^2$ . O operador  $B$  é tal que  $B^j X_t = X_{t-j}$  e os polinômios  $\phi_p$  e  $\theta_q$ , conhecidos como polinômios de médias móveis e autoregressivo, respectivamente, são definidos por:

$$\begin{aligned} \phi_p(B) &= 1 - \phi_1 B - \dots - \phi_p B^p; & e \\ \theta_q(B) &= 1 - \theta_1 B - \dots - \theta_q B^q. \end{aligned}$$

Uma função  $\psi(B)$  representa a série  $\sum_{j=1}^{\infty} \psi_j B^j$ . Assumindo-se que o modelo

seja estacionário, podemos reescrevê-lo na representação média móvel:

$$\begin{aligned} Z_t &= \mu + \frac{\theta_q(B)}{\phi_p(B)} a_t = \mu + \psi(B)a_t \\ &= \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \end{aligned} \quad (1.15)$$

Suponha que conheçamos o processo até o tempo  $n$ , ou seja, temos as observações  $Z_1, Z_2, \dots, Z_n$  e queiramos saber o valor de  $Z_{n+l}$ . Esse valor pode ser estimado de duas formas: a primeira consiste em considerar as observações passadas, isto é, a predição condicional; e a segunda, é a chamada predição incondicional, em que não consideraremos observações anteriores. A seguir, iremos descrever esses dois métodos de estimar o valor de uma observação num tempo futuro, mostrando as vantagens e desvantagens de cada um.

O valor teórico de  $Z_{n+l}$  é:

$$Z_{n+l} = \mu + a_{n+l} + \psi_1 a_{n+l-1} + \psi_2 a_{n+l-2} + \dots + \psi_l a_n + \psi_{l+1} a_{n-1} + \dots \quad (1.16)$$

Um estimador de  $Z_{n+l}$  é dado por:

$$\widehat{Z}_n(l) = \mu + \widehat{\psi}_l a_n + \widehat{\psi}_{l+1} a_{n-1} + \widehat{\psi}_{l+2} a_{n-2} + \dots, \quad (1.17)$$

em que os  $\widehat{\psi}$ 's devem ser determinados a partir da amostra.

Um critério para a escolha desses parâmetros, visto em Wei (1990), é aquele que minimiza o Erro Quadrático Médio (EQM), ou seja, o EQM da predição é

$$\begin{aligned} \mathbb{E}[Z_{n+l} - \widehat{Z}_n(l)]^2 &= \mathbb{E} \left[ \sum_{j=0}^{l-1} a_{n+l-j} \psi_j + \sum_{j=l}^{\infty} a_{n+l-j} (\psi_j - \widehat{\psi}_j) \right]^2 \\ &= \mathbb{E} \left[ \sum_{j=0}^{l-1} a_{n+l-j} \psi_j \right]^2 + \mathbb{E} \left[ \sum_{j=l}^{\infty} a_{n+l-j} (\psi_j - \widehat{\psi}_j) \right]^2 \\ &= \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 + \sigma_a^2 \sum_{j=l}^{\infty} (\psi_j - \widehat{\psi}_j)^2 \\ &= \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 + \sigma_a^2 \sum_{j=0}^{\infty} (\psi_{j+l} - \widehat{\psi}_{j+l})^2. \end{aligned} \quad (1.18)$$

Nota-se que, se os parâmetros  $\psi_j, j = 0, \dots, l-1$  são conhecidos, o EQM é minimizado quando  $\hat{\psi}_{j+l} = \psi_{j+l}$ . Então, o melhor estimador sob o critério de EQM de predição, é dado por

$$\hat{Z}_n(l) = \mu + \psi_l a_n + \psi_{l+1} a_{n-1} + \psi_{l+2} a_{n-2} + \dots \quad . \quad (1.19)$$

Utilizando-se a equação (1.16) e o fato que

$$\mathbb{E}(a_{n+j}|Z_n, Z_{n-1}, \dots, Z_1) = \begin{cases} 0 & \text{se } j > 0, \\ a_{n+j} & \text{se } j \leq 0, \end{cases} \quad (1.20)$$

temos que

$$\mathbb{E}(Z_{n+l}|Z_n, Z_{n-1}, \dots, Z_1) = \mu + \psi_l a_n + \psi_{l+1} a_{n-1} + \psi_{l+2} a_{n-2} + \dots \quad . \quad (1.21)$$

Note que a esperança condicional (1.21) é exatamente o valor de  $\hat{Z}_n(l)$ . Dessa forma, a predição incondicional,  $\hat{Z}_n^*(l)$ , é definida apenas como a esperança de  $Z_{n+l}$ :

$$\hat{Z}_n^*(l) = \mathbb{E}(Z_{n+l}) = \mu. \quad (1.22)$$

Também, pode-se considerar a predição incondicional como o limite da predição condicional quando o número de passos a frente vai para o infinito, ou seja,  $\hat{Z}_n^*(l) = \lim_{l \rightarrow \infty} \hat{Z}_n(l) = \mu$ .

O erro de predição condicional é dado por

$$e_n(l) = Z_{n+l} - \hat{Z}_n(l) = \sum_{j=0}^{l-1} \psi_j a_{n+l-j}, \quad (1.23)$$

enquanto o erro de predição incondicional é dado por

$$e_n^*(l) = Z_{n+l} - \hat{Z}_n^*(l) = \sum_{j=0}^{\infty} \psi_j a_{n+l-j}. \quad (1.24)$$

Como  $\mathbb{E}(e_n(l)|Z_t, t \leq n) = 0$  e  $\mathbb{E}(e_n^*(l)) = 0$ , tanto a predição condicional, quanto

a incondicional, são não-viciadas e têm variâncias, respectivamente, dadas por:

$$\text{Var}(e_n(l)) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2, \quad (1.25)$$

e

$$\text{Var}(e_n^*(l)) = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j^2 = \text{Var}(Z_t). \quad (1.26)$$

Pelos erros de predição e suas respectivas variâncias, vemos que, teoricamente, a predição condicional é sempre melhor. No entanto, na prática, podemos encontrar algumas justificativas para o uso da predição incondicional. Por exemplo, na representação média móvel do modelo ARMA(p,q), é necessário estimar o valor dos parâmetros  $\psi$ 's. Dessa forma, só podemos conhecer o valor estimado de (1.25) e (1.26). No caso condicional, quanto maior o valor de  $l$ , maior seria o número de parâmetros a ser estimado, pois a quantidade de  $\psi$ 's é igual ao número de passos a frente para a predição. Já no caso incondicional, embora tenhamos que estimar infinitos  $\psi$ 's para conhecer o valor da série, existe a alternativa de estimação não-paramétrica, por exemplo, que nos permite estimar  $\text{Var}(e_n^*(l))$  como um todo, já que essa quantidade é igual à  $\text{Var}(Z_t)$ , o que independe do valor de  $l$ .

Com isso, podemos dizer que: para o curto prazo, utilizar a predição condicional é indicada, uma vez que poucos parâmetros seriam estimados e, com isso, o erro propagado não teria grandes efeitos; já para o caso de médio e longo prazos, a predição incondicional deve ser superior, tanto pela complexidade da estimação quanto pelo fato que, se olharmos uma série por um longo período de tempo, é comum ocorrerem mudanças bruscas, o que aumentaria significativamente o erro da predição condicional, sendo esta parametricamente baseada nas observações passadas. A predição incondicional estimaria todos os valores futuros como sendo a esperança da série, o que pode resultar em um erro muito menor do que se usarmos um preditor condicional.

Embora o modelo ARMA seja simples demais para a análise de séries financeiras, utilizamo-los aqui apenas para ilustrar algumas quantidades e pela facilidade de obtermos o erro de predição e sua variância. Espera-se com isso, motivar a necessidade do estudo incondicional da volatilidade pois essas diferenças de comportamento, de acordo com o horizonte de predição, são possíveis em modelos mais complexos, mesmo que sua não-linearidade nos impeça de quantificá-las.

## 2 Análises de Fourier e de Ondaletas

A Análise de Fourier é uma ferramenta muito utilizada em fenômenos que trabalham no domínio do tempo e da frequência, como por exemplo, séries temporais e processamento de sinais. Apesar de suas inúmeras vantagens teóricas, essa técnica apresenta problemas para representar características locais de funções, pela natureza oscilatória das exponenciais.

Ondaletas (também conhecidas na literatura como *wavelets*) são ferramentas matemáticas com uma grande variedade de aplicações. As ondaletas são utilizadas para representar funções através de uma série de funções ortogonais. Além de serem funções consideradas razoavelmente simples, são bem localizadas, usando em geral uma quantidade pequena de coeficientes para representar sinais e imagens, se comparadas com outros sistemas ortogonais.

Uma das vantagens das ondaletas em relação à Análise de Fourier, por exemplo, é que as ondaletas são funções locais em escala (dilatações) e em tempo (translações).

### 2.1 Transformada de Fourier

Nessa seção, apresentaremos a definição de Transformada de Fourier, assim como a Transformada Inversa, que serão utilizadas ao longo desse trabalho.

Considere o espaço  $\mathcal{L}^2(\mathbb{R})$ , ou seja, o espaço que contém as funções quadrado-integráveis. A norma de uma função  $f$  em  $\mathcal{L}^2(\mathbb{R})$  é dada por  $\|f\|_2 = \sqrt{\langle f, f \rangle}$ .

**Definição 2.1.1.** A transformada de Fourier de uma função  $f \in \mathcal{L}^2(\mathbb{R})$  é definida por:

$$f^*(\omega) = \mathcal{F}[f(x)] = \int_{-\infty}^{+\infty} e^{i\omega x} f(x) dx. \quad (2.1)$$

Se  $f^*$  é a Transformada de Fourier de  $f$ , a Transformada Inversa é definida, para todo  $x$  ponto de continuidade de  $f$ , por:

$$f(x) = \mathcal{F}^{-1}[f^*(\omega)] = \frac{1}{2\pi} (f^*)^*(-x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\omega x} f^*(\omega) d\omega. \quad (2.2)$$

Note que, se a função  $f$  for uma densidade de probabilidade de uma v.a.  $X$ , a Transformada de Fourier dessa função é chamada de **função-característica** e é obtida por  $f^*(\omega) = \mathbb{E}(e^{i\omega X})$ . Essa ferramenta é muito utilizada em Estatística, principalmente quando se trabalha com operações de v.a.'s, por exemplo, soma de v.a.'s independentes, uma vez que os cálculos se tornam muito mais simples. Além disso, pode-se encontrar momentos de várias ordens de uma v.a., o que poderia não ser uma operação tão simples se feita da maneira tradicional.

Algumas das principais propriedades das Transformadas de Fourier são descritas a seguir.

P1. **Linearidade:**  $(\alpha f + \beta f)^*(\omega) = \alpha f^*(\omega) + \beta f^*(\omega);$

P2. **Localização e Escala:**  $f^*(a\omega - b) = \frac{1}{|a|} e^{\frac{i\omega b}{a}} f^*\left(\frac{\omega}{a}\right);$

P3. **Simetria:**  $(f^*)^*(x) = 2\pi f(-x);$

P4. **Derivativa:**  $f^{(n)*}(\omega) = (i\omega)^n f^*(\omega);$

P5. **Convolução:**  $(f * g)^*(\omega) = f^*(\omega)g^*(\omega).$

A convolução de  $f$  e  $g$  é definida como  $(f * g)(x) = \int_{\mathbb{R}} f(x - u)g(u)du;$

P6. **Identidade de Plancherel:**  $\langle f, g \rangle = \frac{1}{2\pi} \langle f^*, g^* \rangle$  e  $\|f\|_2^2 = \frac{1}{2\pi} \|f^*\|_2^2;$  e

P7. **Momentos:**

$$\int_{-\infty}^{+\infty} x^n f(x) dx = i^n \left. \frac{d^n f^*(\omega)}{d\omega^n} \right|_{\omega=0}.$$

### 2.1.1 Transformada Discreta de Fourier

A Transformada de Fourier, por ser contínua, impõe dificuldades computacionais a sua aplicação, o que motiva o emprego da Transformada Discreta de Fourier. O teorema de amostragem de Shannon (Teorema 2.1.1), mostra que é possível obter, para  $t \in \mathbb{R}$ , o valor de uma função de  $h \in \mathcal{L}^2(\mathbb{R})$  apenas a partir dos seus valores amostrados em frequências adequadas.

Seja  $\Delta$  o intervalo entre duas amostras consecutivas tal que a seqüência de valores amostrados no tempo seja, por exemplo,

$$h(t_n) = h(n\Delta), \quad n \in \mathbb{Z}.$$

Para qualquer intervalo amostral  $\Delta$ , existe uma frequência especial  $f_c$ , chamada de *frequência crítica de Nyquist*. Essa frequência é a menor frequência possível de amostragem que possibilita recuperar a função de forma única. Note que a frequência,  $\omega$ , considerada na definição da Transformada de Fourier (2.1) é a frequência angular. Sendo assim, a frequência angular de Nyquist é dada por  $\omega_c = 2\pi f_c$ .

**Teorema 2.1.1 (Teorema de Amostragem de Shannon).** *Seja  $h \in \mathcal{L}^2(\mathbb{R})$  e  $h^*(\omega) = 0$  para  $|\omega| > 2\pi f_c$ . Então,  $h$  pode ser reconstruída a partir de seus valores amostrados nos tempos  $t_n = n\Delta$ ,  $n \in \mathbb{Z}$ , pela seguinte expressão:*

$$h(t) = \Delta \sum_{n \in \mathbb{Z}} \frac{\text{sen}[2\pi f_c(t - n\Delta)]}{\pi(t - n\Delta)} h(t_n). \quad (2.3)$$

Conhecendo-se apenas um número finito de valores amostrados de uma função  $h$ , temos interesse em calcular a Transformada de Fourier dessa função utilizando tal amostra. Essa transformada é conhecida como Transformada Discreta de Fourier.

Suponha  $N$  consecutivos valores amostrais,  $h_n = h(t_n)$ , sendo  $t_n = n\Delta$ , para  $n = 0, 1, \dots, N - 1$ . Em vez de tentar estimar a Transformada de Fourier  $h^*(\omega)$  para todo valor de  $\omega$  no intervalo  $[-\omega_c, \omega_c]$ , estima-se a transformada apenas nas frequências:

$$w_k = \frac{2\pi k}{N\Delta}, \quad k = \frac{-N}{2}, \dots, \frac{N}{2}. \quad (2.4)$$

Note que os extremos de  $k$  correspondem exatamente aos extremos do intervalo

$[-\omega_c, \omega_c]$  e, assim, definem-se a Transformada Discreta e a Transformada Discreta Inversa de Fourier.

**Definição 2.1.2.** A Transformada Discreta de Fourier da seqüência  $\{h(t_n) : n = 0, 1, \dots, N - 1\}$  é:

$$h^*(\omega_k) = \Delta \sum_{n=0}^{N-1} h(t_n) e^{i\omega_k t_n}. \quad (2.5)$$

**Definição 2.1.3.** Para a seqüência  $\{h^*(\omega_k) : k = 0, 1, \dots, N - 1\}$ , obtemos a Transformada Discreta Inversa de Fourier através de:

$$h(t_k) = \frac{1}{2\pi N \Delta} \sum_{k=0}^{N-1} h^*(\omega_k) e^{-i\omega_k t_k}. \quad (2.6)$$

## 2.2 Análises de Multiresolução

Uma Análise de Multiresolução (AMR) é ferramenta indispensável na construção de bases de ondaletas ortonormais, processo que, antes de sua formulação, era extremamente complicado. AMR é gerada a partir de uma função  $\phi$ , chamada de função-escala. A seguir, vemos o conceito de ortonormalidade seguido por uma definição de uma AMR que pode ser encontrada em Chui (1992).

**Definição 2.2.1.** Uma seqüência  $\{f_n\}_{n=0}^{\infty}$  de funções reais (ou complexas) definidas em  $\mathcal{L}^2(\mathbb{R})$  é dita ortogonal se:

$$\langle f_n, f_m \rangle = \int_{-\infty}^{+\infty} f_n(x) \overline{f_m(x)} dx = 0 \quad n \neq m; \quad n, m = 0, 1, 2, \dots$$

Se, além disso, tivermos  $\langle f_n, f_n \rangle = 1$ ,  $n = 0, 1, 2, \dots$ , essa seqüência é dita ortonormal.

**Definição 2.2.2.** Uma Análise de Multiresolução em  $\mathcal{L}^2(\mathbb{R})$  é uma seqüência aninhada de subespaços fechados  $\{V_j\}_{j \in \mathbb{Z}}$  de modo que:

$$(i) \quad \dots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \dots \subset \mathcal{L}^2(\mathbb{R});$$

- (ii) Os espaços hierárquicos têm uma intersecção trivial e união densa em  $\mathcal{L}^2(\mathbb{R})$ , isto é,

$$\bigcap_{j \in \mathbb{Z}} V_j = \{0\} \quad \text{e} \quad \overline{\bigcup_{j \in \mathbb{Z}} V_j} = \mathcal{L}^2(\mathbb{R});$$

- (iii) Os espaços  $V_j$  e  $V_{j+1}$  são similares, ou seja,  $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$ ,  $j \in \mathbb{Z}$ .

- (iv) Existe uma função  $\phi \in V_0$  tal que a seqüência  $\{\phi(\bullet - k)\}_{k \in \mathbb{Z}}$  é uma base ortonormal do espaço de referência  $V_0$ , ou seja, para uma constante  $c_k$ ,

$$V_0 = \left\{ f \in \mathcal{L}^2(\mathbb{R}) : f(x) = \sum_{k \in \mathbb{Z}} c_k \phi(x - k) \right\}$$

De (iii)-(iv), segue que:

$$f(x - k) \in V_0 \Leftrightarrow f(x) \in V_0, \quad \forall k \in \mathbb{Z} \quad \text{e} \quad (2.7)$$

$$f(x) \in V_j \Leftrightarrow f(x - 2^j k) \in V_j, \quad \forall k, j \in \mathbb{Z}. \quad (2.8)$$

A condição (iii) se refere à resolução. É intuitivo dobrar ou reduzir pela metade a resolução em que se está trabalhando. Quando se passa de  $V_j$  para  $V_{j+1}$ , a resolução é dobrada, ou seja,  $f(2x)$  é uma versão comprimida de  $f(x)$ .

## 2.3 Espaços de Besov e Ondaletas

Iniciaremos essa seção com uma definição de espaços de Besov. Um espaço de Besov é natural para as ondaletas pois existe uma conexão entre normas de Besov e a expansão de uma função em coeficientes de ondaletas. Mais precisamente, espaços de Besov podem ser caracterizados em termos dos coeficientes de ondaletas. Härdle (1998) mostra alguns resultados sobre quão bem adaptados são os espaços de Besov para AMR. Fan e Koo (2002) mostram que em espaços de Besov com parâmetro  $p < 2$ , os estimadores por desconvolução não-lineares via ondaletas podem alcançar taxa ótima de convergência, com erros regulares ordinários.

**Definição 2.3.1.** Seja  $E_j$  o operador da projeção ortogonal associado a  $V_j$  e  $D_j = E_{j+1} - E_j$ . Espaços de Besov dependem de três parâmetros ( $\sigma > 0$ ,  $1 \leq p \leq \infty$  e  $1 \leq q \leq \infty$ ) e são denotados por  $B_{\sigma pq}$ . Uma função  $f \in B_{\sigma pq}$  se e somente se a norma

$$J_{\sigma pq}(f) = \|E_0 f\|_{\mathcal{L}^p(\mathbb{R})} + \left\{ \sum_{j \geq 0} (2^{j\sigma} \|D_j f\|_{\mathcal{L}^p(\mathbb{R})})^q \right\}^{1/q} < \infty, \quad (2.9)$$

com usual modificação para  $q = \infty$ .

Usando as decomposições:

$$E_0 f = \sum_{k \in \mathbb{Z}} \alpha_{0,k} \phi_{0,k} \quad \text{e} \quad D_j f = \sum_{k \in \mathbb{Z}} \beta_{j,k} \psi_{j,k},$$

o espaço de Besov  $B_{\sigma pq}$  pode ser definido pela norma equivalente:

$$\|f\|_{\sigma pq} = \|\alpha_0\|_p + \left\{ \sum_{j \geq 0} (2^{j\sigma'} \|\beta_j\|_p)^q \right\}^{1/q} < \infty, \quad (2.10)$$

sendo  $\sigma' = \sigma + \frac{1}{2} - \frac{1}{p}$ .

Temos

$$\|\beta_j\|_p = \left\{ \sum_{k \in \mathbb{Z}} |\beta_{j,k}|^p \right\}^{1/p}$$

e a mesma notação aplicada a seqüência  $(\alpha_{0,k})$ . Abusando da notação, também escreve-se  $\|\beta\|_{\sigma pq}$  para a norma da seqüência acima aplicada aos coeficientes de ondaletas.

Os espaços de Besov tem as seguintes relações:

$$B_{\sigma_1 p_1 q_1} \subset B_{\sigma p q}, \quad \text{para } \sigma_1 \geq \sigma \text{ e } q_1 \leq q$$

e

$$B_{\sigma p q} \subset B_{\sigma_1 p_1 q}, \quad \text{para } p_1 > p \text{ e } \sigma_1 = \sigma - \frac{1}{p} + \frac{1}{p_1}.$$

Estão incluídos nesses espaços os espaços de Sobolev  $B_{\sigma 22}$  e as funções limitadas  $\sigma$ -Lipschitz  $B_{\sigma \infty \infty}$  para  $\sigma \in (0, 1)$ .

Os espaços de densidade considerados por Fan e Koo (2002) denotados por

$D_{\sigma pq}(M)$  e  $D_{\sigma pq}(M, S)$ , são ambos formados por funções  $f$ , de forma que,  $\int_{\mathbb{R}} f = 1$  e  $f \geq 0$ , cuja norma dada em (2.10) é limitada por uma dada constante  $M$ . O segundo deles também satisfaz a condição de as funções  $f$  terem suporte compacto, isto é:

$$D_{\sigma pq}(M) = \left\{ f : \int_{\mathbb{R}} f = 1, f \geq 0, \|f\|_{\sigma pq} \leq M \right\}$$

e

$$D_{\sigma pq}(M, S) = \{f \in D_{\sigma pq}(M), \text{suporte } f \in [-S, S]\},$$

para  $S$  constante.

### 2.3.1 Construção das Funções de Ondaletas

A função  $\phi(x)$ , definida na propriedade (iv) da Definição 2.2.2, é conhecida como função-escala. Fazendo-se translações e dilatações diádicas da função-escala, obtem-se a seguinte função:

$$\phi_{j,k}(x) = 2^{\frac{j}{2}} \phi(2^j x - k), \quad j \in \mathbb{Z}, \quad k \in \mathbb{Z}. \quad (2.11)$$

Vimos anteriormente que  $\{\phi_{0,j}(x), k \in \mathbb{Z}\}$  forma uma base ortonormal do espaço  $V_0$ . Agora, utilizando a condição (iii) da Definição 2.2.2 e o fato de que  $\|\phi(x)\| = \|\phi_{j,k}(x)\|$ , para  $\forall j, k \in \mathbb{Z}$ , temos que  $\{\phi_{j,k}(x), k \in \mathbb{Z}\}$  é uma base ortonormal de  $V_j$ , ou seja,

$$V_j = \left\{ f \in \mathcal{L}^2(\mathbb{R}) : f(x) = \sum_{k \in \mathbb{Z}} \alpha_{j,k} \phi_{j,k}(x) \right\}, \quad \forall j \in \mathbb{Z}. \quad (2.12)$$

Dessa forma, considerando também a estrutura hierárquica dos espaços  $V_j$  e a propriedade (ii) da Definição 2.2.2, uma função  $f$  em  $\mathcal{L}^2(\mathbb{R})$  pode ser representada, em norma, por

$$f(x) = \lim_{j \rightarrow +\infty} \sum_{k \in \mathbb{Z}} \alpha_{j,k} \phi_{j,k}(x) = \lim_{j \rightarrow +\infty} P_j f(x), \quad (2.13)$$

sendo  $\alpha_{j,k} = \langle f, \phi_{j,k} \rangle = \int_{\mathbb{R}} f(x) \phi_{j,k}(x) dx$  e  $P_j f(x)$  a projeção ortogonal de  $f(x)$  em  $V_j$ , ou seja, é a aproximação da função  $f$  na resolução  $j$ . Note também que  $\lim_{j \rightarrow -\infty} P_j f(x) = 0$ , pois seria como se olhássemos  $f$  com resoluções cada vez menores, tendendo assim à constante 0.

A diferença de informação entre as aproximações da função  $f$  nas resoluções  $j$  e  $j+1$  é chamada de detalhe na resolução  $j$ , que é dado por  $Q_j f$ , a projeção ortogonal de  $f$  sobre o complemento ortogonal,  $W_j$ , de  $V_j$  em  $V_{j+1}$ , ou seja,

$$P_{j+1}f(x) = P_j f(x) + Q_j f(x), \quad j \in \mathbb{Z}. \quad (2.14)$$

Assim, analogamente a  $V_j$ , a base de  $W_j$  pode ser construída por translações e dilatações de uma função  $\psi$ , conhecida como ondaleta-mãe. Ou seja, temos a função

$$\psi_{j,k}(x) = 2^{\frac{j}{2}} \psi(2^j x - k), \quad j \in \mathbb{Z}, \quad k \in \mathbb{Z}, \quad (2.15)$$

e o conjunto  $\{\psi_{j,k}(x), k \in \mathbb{Z}\}$  forma uma base ortonormal para o espaço do complemento ortogonal  $W_j$ , isto é,

$$W_j = \left\{ f \in \mathcal{L}^2(\mathbb{R}) : f(x) = \sum_{k \in \mathbb{Z}} \beta_{j,k} \psi_{j,k}(x) \right\}, \quad \forall j \in \mathbb{Z}. \quad (2.16)$$

Considere a notação  $V_{j+1} = V_j \oplus W_j$ . Sendo  $V_j \subset V_{j+1}$  e, por construção,  $W_j \perp V_j$  e  $W_{j+1} \perp V_{j+1}$ ,  $\forall j$ , temos que  $W_j \perp W_l$ ,  $\forall j, l \in \mathbb{Z}$  e  $j \neq l$ .

Levando-se ainda em consideração a união densa da AMR, chega-se que

$$V_j = \bigoplus_{l \leq j} W_l, \quad \forall j, l \in \mathbb{Z} \quad (2.17)$$

e

$$\mathcal{L}^2(\mathbb{R}) = \overline{\bigoplus_{j \in \mathbb{Z}} W_j}. \quad (2.18)$$

O espaço  $W_j$  herda muitas das propriedades de  $V_j$ . Fixando-se um nível de resolução  $j_0 \in \mathbb{Z}$ , temos que  $\{\phi_{j_0,k}(x), \psi_{j,k}(x), j \geq j_0, j, k \in \mathbb{Z}\}$  forma uma base ortonormal de  $\mathcal{L}^2(\mathbb{R})$  e, a partir disso, qualquer função  $f \in \mathcal{L}^2(\mathbb{R})$  pode ser escrita como

$$f(x) = \sum_{k \in \mathbb{Z}} \alpha_{j_0,k} \phi_{j_0,k}(x) + \sum_{j \geq j_0} \sum_{k \in \mathbb{Z}} \beta_{j,k} \psi_{j,k}(x), \quad (2.19)$$

sendo

$$\alpha_{j_0,k} = \langle f, \phi_{j_0,k} \rangle = \int_{\mathbb{R}} f(x) \phi_{j_0,k}(x) dx \quad (2.20)$$

e

$$\beta_{j,k} = \langle f, \psi_{j,k} \rangle = \int_{\mathbb{R}} f(x) \psi_{j,k}(x) dx. \quad (2.21)$$

Os coeficientes  $\alpha_{j_0,k}$  e  $\beta_{j,k}$  são conhecidos como coeficientes de ondaletas de aproximação e de detalhes, respectivamente.

A equação (2.19) é uma identidade em norma de  $\mathcal{L}^2(\mathbb{R})$ . Seu primeiro termo é a projeção de  $f(x)$  no espaço  $V_{j_0}$  e, para cada  $j \geq j_0$ , o segundo é a projeção de  $f(x)$  no espaço  $W_j$ . O índice  $j$  está associado à resolução e  $j_0$  é a escala de resolução mais baixa utilizada na AMR.

### 2.3.2 Representação das Ondaletas por Filtros

A representação das ondaletas através de filtros é bastante utilizada, pois possibilita a implementação computacional de esquemas de cálculos mais eficientes e rápidos.

Considerando-se que  $V_0 \subset V_1$ , a função  $\phi(x) \in V_0$  pode ser representada como uma combinação linear das funções de  $V_1$ , ou seja,

$$\phi(x) = \sum_{k \in \mathbb{Z}} h_k \phi_{1,k}(x) = \sum_{k \in \mathbb{Z}} h_k \sqrt{2} \phi(2x - k), \quad (2.22)$$

para alguns coeficientes  $h_k$ ,  $k \in \mathbb{Z}$ . Os coeficientes  $h_k$  são bastante empregados para relacionar a AMR à teoria de processamento de sinais. A seqüência  $\{h_k\}_{k \in \mathbb{Z}}$  é conhecida como *filtros da função-escala*.

Considere a função:

$$m_0(\omega) = \frac{1}{\sqrt{2}} \sum_{k \in \mathbb{Z}} h_k e^{ik\omega} = \frac{1}{\sqrt{2}} H(\omega). \quad (2.23)$$

Essa função é conhecida como *função de transferência* e descreve o comportamento dos filtros associados  $\{h_k\}_{k \in \mathbb{Z}}$ , no domínio da frequência. Note que essa função é periódica com período de  $2\pi$ . Como visto em Vidakovic (1999), a Transformada de Fourier da função-escala é dada pela relação

$$\begin{aligned}
\phi^*(\omega) &= \int_{-\infty}^{+\infty} e^{i\omega x} \phi(x) dx = \sum_k \sqrt{2} h_k \int_{-\infty}^{+\infty} e^{i\omega x} \phi(2x - k) dx \\
&= \sum_k \frac{h_k}{\sqrt{2}} e^{i\omega k/2} \int_{-\infty}^{+\infty} e^{i\omega x/2} \phi(x) dx = \sum_k \frac{h_k}{\sqrt{2}} e^{i\omega k/2} \phi^*\left(\frac{\omega}{2}\right) \\
&= m_0\left(\frac{\omega}{2}\right) \phi^*\left(\frac{\omega}{2}\right).
\end{aligned} \tag{2.24}$$

Por recorrência do resultado (2.24), temos

$$\phi^*(\omega) = \prod_{n=1}^{\infty} m_0\left(\frac{\omega}{2^n}\right), \tag{2.25}$$

que é convergente sob suaves condições (da taxa de decaimento da função-escala  $\phi$ ). Existe uma série de condições suficiente para a convergência do produto em (2.25). Por exemplo, a convergência uniforme sobre conjuntos compactos é assegurada se

- (i)  $m_0(\omega) = 1$  e;
- (ii)  $|m_0(\omega) - 1| < C|\omega|^\varepsilon$ , para algum  $C$  e  $\varepsilon$  positivos.

Os filtros da função-escala apresentam duas importantes propriedades associadas a uma AMR ortogonal: normalização; e ortogonalidade.

Assumindo-se que  $\int \phi(x) dx \neq 0$ ,

$$\int \phi(x) dx = \sqrt{2} \sum_k h_k \int \phi(2x - k) dx = \frac{\sqrt{2}}{2} \sum_k h_k \int \phi(x) dx.$$

Portanto,

$$\sum_k h_k = \sqrt{2}. \tag{2.26}$$

Para qualquer  $l \in \mathbb{Z}$ ,

$$\sum_{k \in \mathbb{Z}} h_k h_{k-2l} = \delta_l = \begin{cases} 0 & \text{se } l \neq 0 \\ 1 & \text{se } l = 0 \end{cases}. \tag{2.27}$$

Um caso particular da ortogonalidade é que, para  $l = 0$ ,  $\sum_{k \in \mathbb{Z}} h_k^2 = 1$ .

Tanto a Transformada de Fourier da função-escala quanto a função de transferência representam as propriedades da função-escala no domínio da frequência. Sendo assim, a ortonormalidade do conjunto  $\{\phi(\bullet - k), k \in \mathbb{Z}\}$  no domínio da frequência é equivalente às relações (2.28)-(2.29).

$$\sum_{l=-\infty}^{\infty} |\phi^*(\omega) + 2\pi l|^2 = 1, \quad (2.28)$$

ou

$$|m_0(\omega)|^2 + |m_0(\omega + \pi)|^2 = 1. \quad (2.29)$$

As duas condições de ortonormalidade dadas acima não são equivalentes. A primeira delas é uma condição suficiente e a segunda necessária.

Agora, vamos mostrar a construção da ondaleta-mãe a partir da função-escala, utilizando filtros.

Como o espaço  $W_0 \subset V_1$ ,  $\psi(x) \in V_1$  e, assim, pode ser representado como:

$$\psi(x) = \sum_{k \in \mathbb{Z}} g_k \sqrt{2} \phi_{1,k}(x) = \sum_{k \in \mathbb{Z}} g_k \sqrt{2} \phi(2x - k), \quad (2.30)$$

para alguns coeficientes  $g_k, k \in \mathbb{Z}$ , sendo a seqüência  $\{g_k\}_{k \in \mathbb{Z}}$  conhecida como filtro da ondaleta-mãe.

Assim como no caso dos filtros da função-escala, temos a função de transferência da seqüência  $\{g_k\}_{k \in \mathbb{Z}}$ :

$$m_1(\omega) = \frac{1}{\sqrt{2}} \sum_k g_k e^{ik\omega}. \quad (2.31)$$

De forma análoga a (2.24), a Transformada de Fourier da ondaleta-mãe é dada pela relação:

$$\psi^*(\omega) = m_1\left(\frac{\omega}{2}\right) \phi^*\left(\frac{\omega}{2}\right). \quad (2.32)$$

Como a função de transferência representa a função  $\psi$  no domínio da frequência, ela também é utilizada para representar as propriedades da ondaleta-mãe e, principalmente, sua relação com a função-escala  $\phi$  nesse domínio.

Por construção, tem-se que  $V_j \perp W_j, \forall j \in \mathbb{Z}$ . Em particular, para  $j = 0$ , tem-se

$$\int_{\mathbb{R}} \psi(x - l) \overline{\phi(x - k)} dx = 0, \quad \forall l, k \in \mathbb{Z}.$$

Pela ortogonalidade e pelas definições de  $m_0$  e  $m_1$ , essa condição é equivalente, no domínio da frequência, a

$$m_1(\omega)\overline{m_0(\omega)} + m_1(\omega + \pi)\overline{m_0(\omega + \pi)} = 0. \quad (2.33)$$

Da relação acima, Vidakovic (1999) conclui que existe uma função  $\lambda(\omega)$  tal que:

$$(m_1(\omega), m_1(\omega + \pi)) = \lambda(\omega) \left( \overline{m_0(\omega + \pi)}, -\overline{m_0(\omega)} \right) \quad (2.34)$$

Existem inúmeras possibilidades para a função  $\lambda(\omega)$ , desde que se satisfaçam as seguintes condições:

- (i)  $\lambda(\omega)$  é periódica com período de  $2\pi$ ;
- (ii)  $\lambda(\omega) = -\lambda(\omega + \pi)$ ; e
- (iii)  $|\lambda(\omega)|^2 = 1$ .

Escolhas padrão para  $\lambda(\omega)$  são  $-e^{i\omega}$ ,  $e^{i\omega}$ ,  $-e^{-i\omega}$  e  $e^{-i\omega}$ . Escolhendo-se a primeira delas, tem-se a definição de  $m_1(\omega)$ :

$$m_1(\omega) = -e^{i\omega}\overline{m_0(\omega + \pi)}. \quad (2.35)$$

Como  $|m_1(\omega)| = |\overline{m_0(\omega + \pi)}|$ , a condição de ortonormalidade (2.29) pode ser reescrita como:

$$|m_0(\omega)|^2 + |m_1(\omega)|^2 = 1. \quad (2.36)$$

Desenvolvendo-se a expressão (2.35), tem-se

$$\begin{aligned} m_1(\omega) &= -e^{i\omega} \frac{1}{\sqrt{2}} \sum_k h_k e^{-i(\omega+\pi)k} = -\frac{1}{\sqrt{2}} \sum_k h_k e^{i\omega(1-k)} e^{-i\pi k} \\ &= \frac{1}{\sqrt{2}} \sum_k h_k e^{i\omega(1-k)} (-1)^{1+k} = \frac{1}{\sqrt{2}} \sum_n (-1)^n h_{1-n} e^{i\omega n}. \end{aligned} \quad (2.37)$$

Comparando-se o resultado (2.37) com a expressão (2.31), temos uma relação entre os filtros  $g_k$  e  $h_k$ , ou seja,  $g_k = (-1)^k h_{1-k}$ . Essa relação é conhecida como *relação de quadratura espelhada*, uma vez que os filtros  $\{g_k\}_{k \in \mathbb{Z}}$  são reflexões dos filtros  $\{h_k\}_{k \in \mathbb{Z}}$ . Os filtros  $\{g_k\}_{k \in \mathbb{Z}}$  e  $\{h_k\}_{k \in \mathbb{Z}}$  são conhecidos como *filtros de quadratura*

*espelhada.*

### 2.3.3 Regularidade das Ondaletas

A regularidade das ondaletas está diretamente relacionada à taxa de decaimento da função-escala e ao número de momentos finitos que a função-escala e a ondaleta-mãe possuem.

O Teorema 2.3.1 é importante para que possamos relacionar a regularidade das ondaletas, o número de momentos nulos e a forma da função de transferência  $m_0(\omega)$ .

Definem-se por

$$\mathcal{M}_k = \int x^k \phi(x) dx \quad \text{e} \quad \mathcal{N}_k = \int x^k \psi(x) dx, \quad (2.38)$$

o  $k$ -ésimo momento da função-escala e da ondaleta-mãe, respectivamente.

**Teorema 2.3.1.** *Seja  $\{\psi_{j,k}(x), j, k \in \mathbb{Z}\}$  um sistema ortonormal de funções em  $\mathcal{L}^2(\mathbb{R})$ ,*

$$|\psi(x)| \leq \frac{C_1}{(1 + |x|)^\alpha}, \quad \alpha > N, \quad (2.39)$$

*e  $\psi \in \mathcal{C}^{N-1}$ , em que as derivadas  $\psi^{(k)}(x)$  são limitadas para  $k \leq N - 1$ . Então,  $\psi$  tem  $N$  momentos nulos,*

$$\mathcal{N}_k = 0, \quad 0 \leq k \leq N - 1.$$

*Se*

$$|\phi(x)| \leq \frac{C_2}{(1 + |x|)^\alpha}, \quad \alpha > N, \quad (2.40)$$

*então, a função de transferência  $m_0(\omega)$  associada a  $\phi(x)$  é necessariamente da forma*

$$m_0(\omega) = \left( \frac{1 + e^{-i\omega}}{2} \right)^N \mathcal{L}(\omega), \quad (2.41)$$

*sendo a função  $\mathcal{L}(\omega)$  periódica com período  $2\pi$  e pertencente a classe  $\mathcal{C}^{N-1}$ .*

**Definição 2.3.2.** Uma AMR (ou a função-escala) é dita ser  $r$ -regular, se para

qualquer  $\alpha \in \mathbb{Z}$ ,

$$|\phi^{(k)}(x)| \leq \frac{C}{(1+|x|)^\alpha}, \quad \text{para } k = 0, 1, \dots, r. \quad (2.42)$$

A restrição de que  $\psi$  tenha  $N$  momentos finitos pode ser expressa em termos de  $\psi^*$ ,  $m_0$  ou pelos filtros  $\{h_k\}$ .

Assumindo-se que a ondaleta-mãe  $\psi(x)$  tenha  $N$  momentos nulos, isto é,  $\mathcal{N}_k = 0, k = 0, 1, \dots, N-1$ , então, pela propriedade dos momentos da Transformada de Fourier, essa suposição é equivalente a

$$\left. \frac{d^k \psi^*(\omega)}{d\omega^k} \right|_{\omega=0} = 0, \quad k = 0, 1, \dots, N-1,$$

o que implica em que

$$m_1^{(k)}(\omega)|_{\omega=0} = m_1^{(k)}(0) = 0, \quad k = 0, \dots, N-1. \quad (2.43)$$

Usando a definição de  $m_1(\omega)$ , dada pela equação (2.31), e a relação entre os filtros  $\{h_k\}_{k \in \mathbb{Z}}$  e  $\{g_k\}_{k \in \mathbb{Z}}$ , temos a seguinte restrição para que  $\psi$  possua  $N$  momentos nulos:

$$\sum_{n \in \mathbb{Z}} n^k g_n = \sum_{n \in \mathbb{Z}} (-1)^n n^k h_n, \quad k = 0, \dots, N-1. \quad (2.44)$$

Se não fizermos restrições sobre as bases de ondaletas, os filtros podem ser séries infinitas, dificultando assim a aplicação dessa análise. No entanto, fazendo uso de certas restrições, como o número de momentos nulos da ondaleta-mãe ou exigindo-se que a mesma tenha suporte compacto, reduz-se o número de coeficientes dos filtros  $\{h_k\}_{k \in \mathbb{Z}}$  e  $\{g_k\}_{k \in \mathbb{Z}}$ .

### 2.3.4 Bases de Ondaletas

Uma importante característica na análise de ondaletas é a grande diversidade de bases existentes. Pode-se utilizar bases simétricas, bases com suporte compacto, bases com boa regularidade ou bases ortogonais. No entanto, não existe uma base que reúna todas essas características. Por exemplo, não existe uma ondaleta simétrica

com suporte compacto, assim como não é possível construir bases que sejam ortogonais e simétricas.

Nessa dissertação, utilizaremos as ondaletas de Meyer e as de Daubechies, pois, como visto em Fan e Koo (2002), o uso dessas bases, de acordo com o tipo de erro do modelo, faz que os estimadores atinjam taxas de convergência ótimas. A seguir, mostraremos algumas características dessas bases.

Primeiramente, apresentaremos as bases de Haar. Essas bases têm valor prático restrito; no entanto, são de grande utilidade didática. O estudo da base de Haar facilita a demonstração das propriedades vistas anteriormente e é a única base ortogonal com suporte compacto para qual pode-se obter, de forma explícita, as funções  $\phi$  e  $\psi$ . Embora bem localizadas no domínio do tempo, as ondaletas de Haar possuem taxa de decaimento muito baixa no domínio da frequência. Sua utilidade prática fica prejudicada pelo fato de que não é muito interessante aproximar funções contínuas por funções descontínuas.

#### 2.3.4.1 Base de Haar

Considere a função-escala  $\phi(x) = \mathbb{1}_{[0,1]}(x)$ . Pela definição do filtro  $\{h_k\}_{k \in \mathbb{Z}}$  em (2.22), tem-se que:

$$\phi(x) = \mathbb{1}_{[0,1]}(x) = \sum_{k \in \mathbb{Z}} \sqrt{2} h_k \mathbb{1}_{[0,1]}(2x - k) = \sum_{k=0}^1 \sqrt{2} h_k \mathbb{1}_{\left[\frac{k}{2}, \frac{k+1}{2}\right]}(x). \quad (2.45)$$

Note que, pela definição inicial da função-escala, os únicos valores possíveis são  $k = 0$  ou  $k = 1$ . Dessa forma, temos dois filtros da função-escala e para que a equação acima seja verdadeira é necessário que:

$$h_0 = h_1 = \frac{1}{\sqrt{2}}. \quad (2.46)$$

Pela relação de quadratura espelhada, os filtros da ondaleta-mãe são:

$$g_0 = \frac{1}{\sqrt{2}} \quad \text{e} \quad g_1 = -\frac{1}{\sqrt{2}}. \quad (2.47)$$

Para as ondaletas de Haar, a função de transferência vista em (2.23) é dada por:

$$m_0(\omega) = \sum_{k \in \mathbb{Z}} \frac{h_k}{\sqrt{2}} e^{i\omega k} = \frac{1 + e^{i\omega}}{2}. \quad (2.48)$$

E, conseqüentemente,

$$m_1(\omega) = -e^{i\omega} \overline{m_0(\omega + \pi)} = \frac{1 - e^{i\omega}}{2}. \quad (2.49)$$

Calculando-se a Transformada Inversa de Fourier da função  $\psi^*$  dada em (2.32), obtemos a ondaleta-mãe da base de Haar:

$$\begin{aligned} \psi(x) &= \frac{1}{2\pi} \int_{\mathbb{R}} \left( \frac{1 - e^{i\omega}}{2} \right) \phi^* \left( \frac{\omega}{2} \right) e^{-i\omega x} d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1}{2} \phi^* \left( \frac{\omega}{2} \right) e^{-i\omega x} d\omega - \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1}{2} \phi^* \left( \frac{\omega}{2} \right) e^{-i\omega(x-1/2)} d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \phi^*(\omega) e^{-i\omega 2x} d\omega - \frac{1}{2\pi} \int_{\mathbb{R}} \phi^*(\omega) e^{-i\omega(2x-1)} d\omega \\ &= \phi(2x) - \phi(2x - 1) \\ &= \mathbb{I}_{[0, \frac{1}{2}]}(x) - \mathbb{I}_{[\frac{1}{2}, 1]}(x). \end{aligned} \quad (2.50)$$

Seria muito mais simples obter a ondaleta-mãe de Haar através dos filtros  $g_0$  e  $g_1$ , mas utilizamos a Transformada de Fourier para demonstrar as propriedades vistas anteriormente.

A Figura (2.4) mostra a função-escala  $\phi$  (2.45) e a ondaleta-mãe  $\psi$  (2.50).

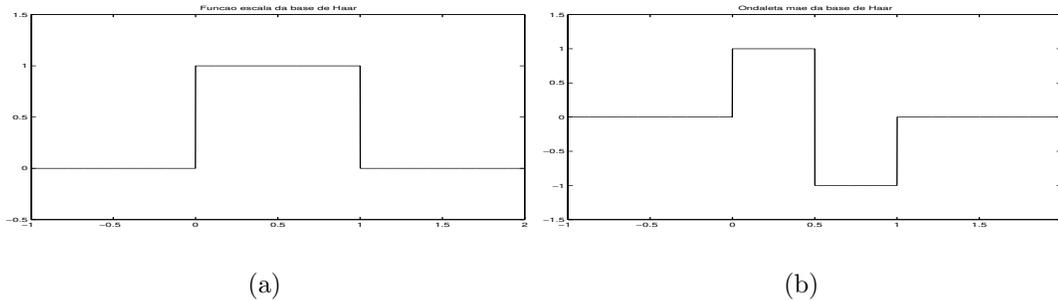


Figura 2.4: Bases de Haar: (a) Função Escala e (b) Função de Detalhe

### 2.3.4.2 Ondaletas de Meyer

Nessa seção iremos definir as ondaletas de Meyer. Fan e Koo (2002) mostram que os estimadores por desconvolução via ondaletas de Meyer atingem taxas de convergência ótimas quando empregadas no caso de modelos cuja distribuição dos erros pertença a classe de erros super-regulares.

A função-escala no domínio de Fourier é dada pelo seguinte equação:

$$\phi^*(\omega) = \begin{cases} 1, & \text{se } |\omega| \leq \frac{2\pi}{3} \\ \cos \left[ \frac{\pi}{2} \nu \left( \frac{3}{2\pi} |\omega| - 1 \right) \right], & \text{se } \frac{2\pi}{3} \leq |\omega| \leq \frac{4\pi}{3} \\ 0, & \text{caso contrário} \end{cases} \quad (2.51)$$

em que a função  $\nu$  satisfaz:

$$\begin{aligned} \nu(x) + \nu(1-x) &= 1 \\ \nu(x) &= 0, \quad x \leq 0 \\ \nu(x) &= 1, \quad x \geq 1. \end{aligned}$$

Muitas funções são possíveis para  $\nu$ , mas as mais comuns são as funções polinomiais. A Tabela 2.1 apresenta as funções  $\nu$  em função do parâmetro  $\theta$ . De acordo com esse parâmetro, a transformada correspondente  $\phi^*(\omega)$  tem  $(\theta - 1)$  derivadas contínuas. A função  $\nu$  utilizada nesse trabalho é o polinômio com parâmetro  $\theta = 4$ .

Tabela 2.1: Funções  $\nu$  para  $\theta = 1, \dots, 5$

$\theta$	$\nu_\theta(x)$
1	$x$
2	$x^2(2 - 3x)$
3	$x^3(10 - 15x + 6x^2)$
4	$x^4(35 - 84x + 70x^2 - 20x^3)$
5	$x^5(126 - 420x + 540x^2 - 315x^3 + 70x^4)$

Pelo fato de que  $\phi^*(\frac{\omega}{2}) = 1$  no domínio da frequência,  $m_0(\frac{\omega}{2}) = \phi^*(\omega)$ ,  $-\pi \leq \omega \leq \pi$ . E, como  $m_0(\omega)$  é uma função periódica em  $2\pi$  por construção, tem-se que:

$$m_0(\omega) = \sum_k \phi^*[2(\omega + 2k\pi)]. \quad (2.52)$$

Utilizando a propriedade (2.35), encontra-se uma expressão para a Transformada de Fourier da ondaleta-mãe:

$$\begin{aligned}\psi^*(\omega) &= -e^{i\omega/2} \overline{m_0\left(\frac{\omega}{2} + \pi\right)} \phi^*\left(\frac{\omega}{2}\right) \\ &= -e^{i\omega/2} \sum_l \phi^*(\omega + 2\pi(2l + 1)) \phi^*\left(\frac{\omega}{2}\right) \\ &= -e^{i\omega/2} [\phi^*(\omega - 2\pi) + \phi^*(\omega + 2\pi)] \phi^*\left(\frac{\omega}{2}\right),\end{aligned}\quad (2.53)$$

desde que para  $l \neq 0; -1$  as funções  $\phi^*(\omega + 2\pi(2l + 1))$  e  $\phi^*(\omega/2)$  não se sobreponham.

Substituindo-se (2.53) por (2.51), temos:

$$\psi^*(\omega) = \begin{cases} -e^{i\omega/2} \sin\left[\frac{\pi}{2}\nu\left(\frac{3}{2\pi}|\omega| - 1\right)\right], & \text{se } \frac{2\pi}{3} \leq |\omega| \leq \frac{4\pi}{3} \\ -e^{i\omega/2} \cos\left[\frac{\pi}{2}\nu\left(\frac{3}{4\pi}|\omega| - 1\right)\right], & \text{se } \frac{4\pi}{3} \leq |\omega| \leq \frac{8\pi}{3} \\ 0, & \text{caso contrário.} \end{cases}\quad (2.54)$$

A Figura 2.5 apresenta as ondaletas de Meyer no domínio da frequência.

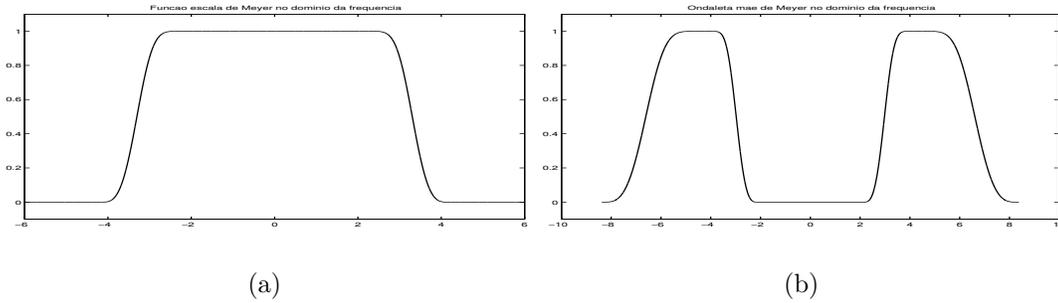


Figura 2.5: Ondaletas de Meyer no Domínio da Frequência: (a)  $\phi^*(\omega)$  e (b)  $|\psi^*(\omega)|$

Aplicando-se as Transformadas Inversas de Fourier nas funções (2.51) e (2.54), tem-se a função-escala  $\phi$  e a ondaleta-mãe  $\psi$ , respectivamente. Como  $\phi^*(\omega)$  é uma função par, tem-se que:

$$\phi(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\omega x} \phi^*(\omega) d\omega = \frac{1}{\pi} \int_0^{4\pi/3} \phi^*(\omega) \cos(\omega x) d\omega.\quad (2.55)$$

e

$$\psi(x) = -\frac{1}{\pi} \int_{2\pi/3}^{4\pi/3} \phi^*\left(\frac{\omega}{2}\right) \phi^*(\omega - 2\pi) \cos\left[\omega\left(x - \frac{1}{2}\right)\right] d\omega. \quad (2.56)$$

Os gráficos dessas funções são mostrados na Figura 2.6.

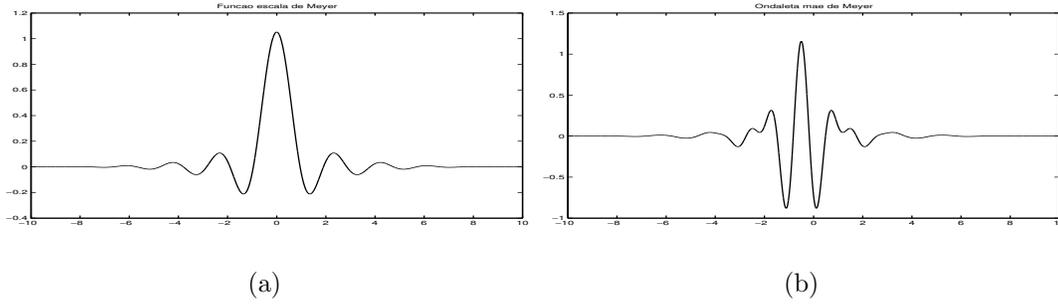


Figura 2.6: Ondaletas de Meyer: (a)  $\phi(x)$  e (b)  $\psi(x)$

Os filtros das ondaletas de Meyer, exceto para o caso em que  $\nu(x) = x$ , são obtidos por integração numérica da seguinte expressão:

$$h_n = \frac{\sqrt{2}}{\pi} \left\{ \int_0^{\pi/2} \cos(n\omega) d\omega + \int_{\pi/3}^{2\pi/3} \cos\left[\frac{\pi}{2}\nu\left(\frac{3}{\pi}\omega - 1\right)\right] d\omega \right\}. \quad (2.57)$$

### 2.3.4.3 Ondaletas de Daubechies

As bases de Daubechies foram as primeiras bases ortogonais de suporte compacto com um nível de suavidade previamente definido. Essas bases não possuem forma explícita, exceto para o caso da Daubechies de ordem 1, que é a base de Haar, vista anteriormente. Assim, as ondaletas de Daubechies são construídas de forma numérica através dos filtros  $\{h_k\}_{k \in \mathbb{R}}$ . Considerando a taxa de convergência dos estimadores por desconvolução, Fan e Koo (2002) mostram que essas bases devem ser empregadas no caso de erros regulares ordinários.

Suponha que  $\psi$  tenha  $N(\geq 2)$  momentos nulos, isto é,  $\mathcal{N}_k = 0, k = 0, 1, \dots, N - 1$ . Então pelo Teorema 2.3.1,  $m_0(\omega)$  tem a forma:

$$m_0(\omega) = \left(\frac{1 + e^{-i\omega}}{2}\right)^N \mathcal{L}(\omega), \quad (2.58)$$

em que a função  $\mathcal{L}(\omega)$  é um polinômio trigonométrico. Não entraremos em detalhes sobre esse polinômio nem sobre como os filtros são construídos. Essa seção tem o intuito apenas de ilustrar a construção das ondaletas de Daubechies.

A base de Daubechies possui filtro  $\{h_k\}_{k \in \mathbb{R}}$  de tamanho  $2N$  e tem suporte compacto  $[0, 2N - 1]$ . Como exemplo, temos os filtros das ondaletas *DAUB2*:

$$\begin{aligned} h_0 &= \frac{1 + \sqrt{3}}{4\sqrt{2}} = 0,483, & h_1 &= \frac{3 + \sqrt{3}}{4\sqrt{2}} = 0,837, \\ h_2 &= \frac{3 - \sqrt{3}}{4\sqrt{2}} = 0,224, & h_3 &= \frac{1 - \sqrt{3}}{4\sqrt{2}} = -0,129. \end{aligned} \quad (2.59)$$

Como as bases de Daubechies não possuem uma forma explícita, faz-se uso do algoritmo de Daubechies-Lagarias, que permite utilizar os filtros  $\{h_k\}_{k \in \mathbb{R}}$  para avaliar as funções  $\phi$  e  $\psi$  em qualquer ponto  $x \in \mathbb{R}$ , com um nível de precisão pré-determinado. Esse algoritmo pode ser utilizado com outras bases também, mas será ilustrado aqui para as bases de Daubechies.

**Teorema 2.3.2. (Daubechies-Lagarias)**

Seja  $\phi$  a função-escala da base *DAUBN* e  $\{h_0, h_1, \dots, h_{2N-1}\}$  a sua respectiva seqüência de filtros. O suporte de  $\phi$  é  $[0, 2N - 1]$ . Para  $x \in (0, 1)$  e  $d(x) = \{d_1, d_2, \dots, d_n, \dots\}$  o conjunto de dígitos 0 e 1 da representação diádica de  $x$ , definida por  $x = \sum_{j=1}^{\infty} d_j 2^{-j}$  e, definindo as matrizes  $T_0$  e  $T_1$  como:

$$T_0 = \left( \sqrt{2}h_{2i-j-1} \right)_{1 \leq i, j \leq 2N-1} \quad T_1 = \left( \sqrt{2}h_{2i-j} \right)_{1 \leq i, j \leq 2N-1}, \quad (2.60)$$

temos que, para um subgrupo  $\{d_1, d_2, \dots, d_n\}$ :

$$\begin{aligned} & \lim_{n \rightarrow \infty} T_{d_1} \cdot T_{d_2} \cdots T_{d_n} \\ &= \begin{bmatrix} \phi(x) & \phi(x) & \dots & \phi(x) \\ \phi(x+1) & \phi(x+1) & \dots & \phi(x+1) \\ \vdots & \vdots & \dots & \vdots \\ \phi(x+2N-2) & \phi(x+2N-2) & \dots & \phi(x+2N-2) \end{bmatrix}. \end{aligned} \quad (2.61)$$

Devido a relação entre  $\phi(x)$  e  $\psi(x)$  vista em (2.30), pode-se encontrar  $\psi(x)$  para todo  $x \in \mathbb{R}$  de forma numérica.

A Figura 2.7 apresenta, no domínio da freqüência, a base de *DAUB8*, enquanto

a Figura 2.8 apresenta as bases de Daubechies de ordem 2, 5 e 8, respectivamente, no domínio do tempo.

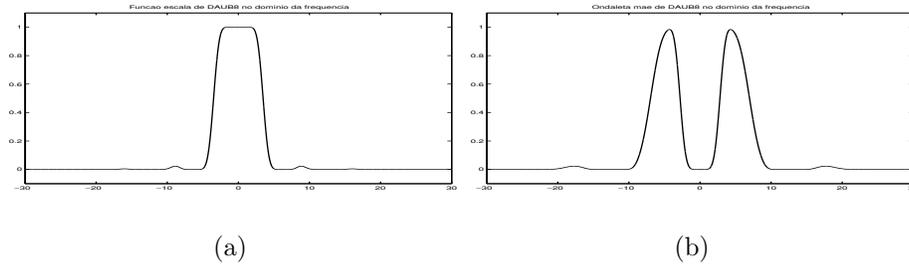


Figura 2.7: Ondaletas *DAUB8*: (a)  $|\phi^*(\omega)|$  e (b)  $|\psi^*(\omega)|$

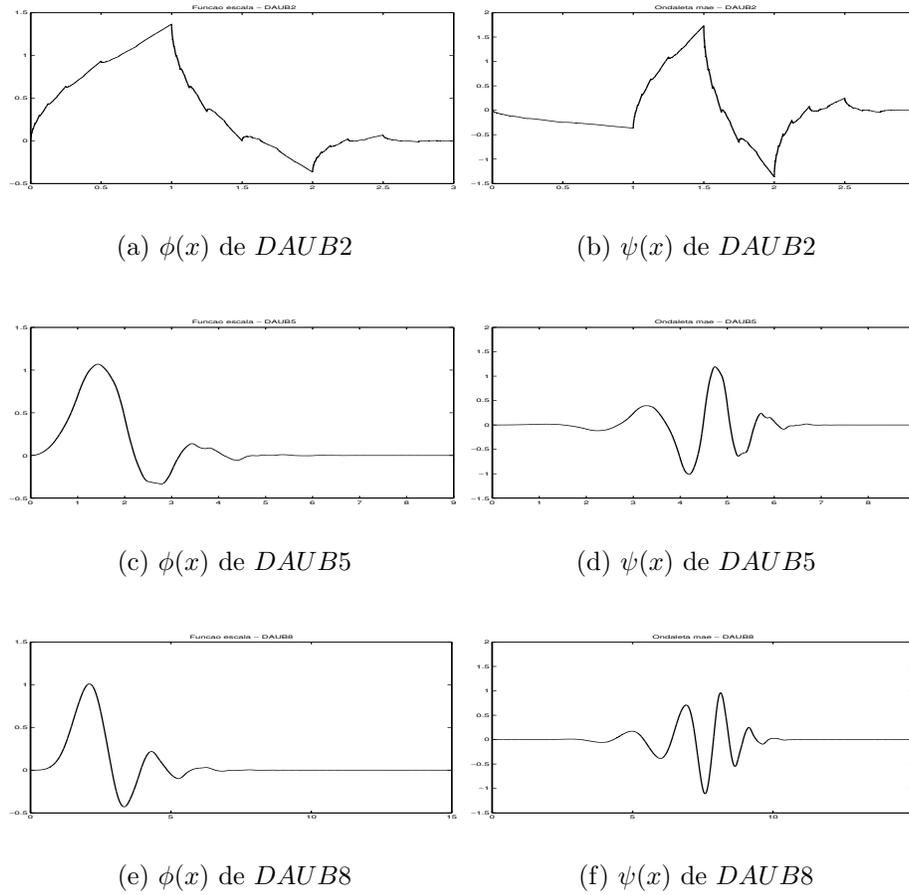


Figura 2.8: Bases de Daubechies

---

Note que, à medida que o valor de  $N$  aumenta, as funções apresentam uma oscilação maior. É importante observar que qualquer aparente descontinuidade das funções  $\phi$  e  $\psi$ , principalmente para  $N = 2$ , é apenas um artifício numérico, devido a aproximação.

## 3 Estimadores por Desconvolução

Nesse capítulo iremos apresentar os estimadores por desconvolução para os modelos de volatilidade estocástica. Duas técnicas não-paramétricas serão utilizadas para construir esses estimadores: funções-núcleo e ondaletas. No entanto, é importante ressaltar aqui que essas técnicas serão utilizadas com propósitos diferentes. A desconvolução via função-núcleo é empregada para estimar a função  $m$  do modelo (1.10). Já a desconvolução via ondaletas é empregada com o objetivo de estimar a densidade de probabilidade da variável  $X_t$  no modelo (1.11), uma vez que a volatilidade é uma função dessa variável. Sendo assim, é possível obter uma estimativa para a densidade da volatilidade.

As propriedades dos estimadores são discutidas de acordo com o tipo de erro do modelo. Esses erros são divididos em duas classes, apresentadas na Seção 3.1. Em seguida, na Seção 3.2 apresentamos conceitos de funções-núcleo, necessários para o entendimento da técnica e construção do estimador visto na Seção 3.3. Por fim, temos os estimadores por desconvolução lineares e não-lineares via ondaletas expostos na Seção 3.4.

### 3.1 Classes de Erro

Duas classes de erros  $\varepsilon$  serão consideradas, de acordo com a taxa de decaimento da cauda da função-característica de  $\varepsilon$ ,  $f_\varepsilon^*$ . Quanto mais rápido é esse decaimento da cauda, mais suave é a função-característica do  $\varepsilon$  correspondente. Essas duas classes

são:

1. Super-Regulares de ordem  $s$ :

$$\gamma_0 |t|^{s_0} e^{-|t|^{s/\lambda}} \leq |f_\varepsilon^*(t)| \leq \gamma_1 |t|^{s_1} e^{-|t|^{s/\lambda}} \text{ quando } t \rightarrow +\infty,$$

com  $s, \lambda > 0$  e alguma constante  $s_0$  e  $s_1$ .

2. Regulares Ordinárias de ordem  $s$ :

$$\gamma_0 |t|^{-s} \leq |f_\varepsilon^*(t)| \leq \gamma_1 |t|^{-s} \text{ quando } t \rightarrow +\infty,$$

com  $s, \lambda > 0$ .

Exemplos dessas classes são: distribuições Normal e Cauchy para as classes super-regulares de ordens  $s = 2$  e  $s = 1$ , respectivamente; e a distribuição exponencial dupla para a classe regular ordinária de ordem  $s = 2$ .

A classe a que os erros pertencem é fundamental no desempenho dos estimadores, tanto para a desconvolução via função-núcleo quanto por ondaletas.

## 3.2 Função-Núcleo

A técnica de estimação não-paramétrica de curvas por função-núcleo é uma das mais conhecidas e utilizadas entre as várias existentes na literatura. Detalhes podem ser encontrados em Härdle (1990, 1991).

A estimação não-paramétrica da curva de regressão têm quatro propósitos principais: fornecer um método versátil de explorar a relação geral entre duas variáveis; dar predições de observações sem referência a um modelo paramétrico; servir como ferramenta para encontrar observações espúrias pelo estudo de pontos isolados; e constituir um método flexível de substituição para valores faltantes ou interpolação entre valores. Em nosso caso, utilizaremos a técnica para predição.

### 3.2.1 Conceitos de Função-Núcleo

Considere  $n$  observações bidimensionais de um vetor aleatório  $(X, Y)$ . Uma questão de interesse é modelar a relação entre  $X$  e  $Y$ . Essa relação é comumente modelada através da esperança condicional  $\mathbb{E}(Y|X)$ . Seja  $m(x) = \mathbb{E}(Y|X = x)$ . Escreve-se:

$$Y_i = m(X_i) + \varepsilon_i, \quad i = 1, \dots, n, \quad (3.1)$$

sendo  $\varepsilon_i$  uma v. a., que denota a variação de  $Y$  em torno de  $m(X)$ , com  $\mathbb{E}(\varepsilon_i) = 0$  e  $Var(\varepsilon_i) = \sigma^2$ .

Assumindo-se que  $(X, Y)$  tenha densidade conjunta, pode escrever-se:

$$m(x) = \mathbb{E}(Y|X = x) = \int_{-\infty}^{+\infty} y \frac{f_{X,Y}(x, y)}{f_X(x)} dy, \quad (3.2)$$

em que  $f_{X,Y}(x, y)$  é a densidade conjunta de  $(X, Y)$  e  $f_X(x)$  a densidade marginal de  $X$ .

A idéia do ajuste da curva de regressão não-paramétrica por função-núcleo é, em cada ponto  $x$ , ponderar-se a variável resposta  $Y$  por suas observações em uma certa vizinhança desse ponto. Os pesos de cada observação  $y_i$  dependem das respectivas distâncias de  $x_i$  a  $x$ .

Dessa forma, o estimador de regressão por função-núcleo (Härdle (1991)), é dado por:

$$\hat{m}_h(x) = \frac{1}{n} \sum_{i=1}^n W_h(x; X_1, \dots, X_n) Y_i,$$

em que  $W_h$  define a função-peso, dependendo de  $h$ , chamado de parâmetro de alisamento, e da amostra  $X_1, \dots, X_n$  e é denotado por  $W_{hi}(x)$ .

Tem-se, então, a forma geral do alisamento por regressão não-paramétrica:

$$\hat{m}_h(x) = \frac{1}{n} \sum_{i=1}^n W_{hi}(x) Y_i. \quad (3.3)$$

Quase todas as técnicas de regressão não-paramétrica são médias ponderadas

das variáveis respostas  $Y_i$ , onde os pesos  $W_{hi}(x)$  dependem da técnica e também da distância entre  $x$  e  $X_i$ . A forma dos pesos depende da função-núcleo  $K$  escolhida e o *tamanho* dos pesos dado a cada observação é determinado pelo parâmetro  $h$ .

Considere a função-núcleo  $K$  tal que a seguinte relação é satisfeita:

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right).$$

Para que os estimadores de densidade por função-núcleo sejam também densidades, é suficiente que  $K$  seja uma densidade. Simetria de  $K$  é uma condição razoável por dar pesos iguais a observações à mesma distância do ponto de interesse. Portanto, consideram-se funções-núcleo tais que

$$K(x) \geq 0, \forall x \in \mathbb{R}, \quad \int_{-\infty}^{+\infty} K(x)dx = 1 \quad \text{e} \quad \int_{-\infty}^{+\infty} xK(x)dx = 0.$$

Alguns exemplos de funções-núcleo são apresentados na Tabela 3.2 e na Figura 3.2.1.

Tabela 3.2: Exemplos de Funções-Núcleo

Função-núcleo	$K(u)$
Uniforme	$\mathbb{I}_{( u  \leq 1/2)}$
Retangular	$0.5 \mathbb{I}_{( u  \leq 1)}$
Triangular	$(1 -  u ) \mathbb{I}_{( u  \leq 1)}$
Epanechnikov	$0.75(1 - u^2) \mathbb{I}_{( u  \leq 1)}$
Gaussiana	$\frac{1}{\sqrt{2\pi}} e^{-u^2/2}$

Calculando-se a média dos valores da função-núcleo em cada observação, chega-se à definição do estimador de densidade por função-núcleo:

$$\hat{f}_h(x) = \frac{1}{n} \sum_{j=1}^n K_h(x - X_j) = \frac{1}{nh} \sum_{j=1}^n K\left(\frac{x - X_j}{h}\right). \quad (3.4)$$

É importante considerar a escolha da largura de banda  $h$  pois, quando o valor de  $h$  é muito pequeno, a estimativa da densidade se torna excessivamente detalhada

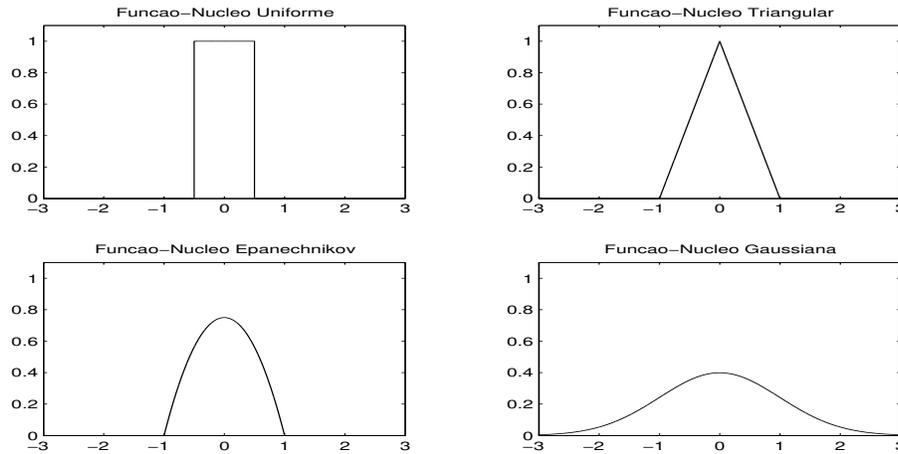


Figura 3.9: Exemplos de Funções-Núcleo

e cheia de ruídos. Por outro lado, se o valor de  $h$  for grande, a estimativa de densidade se torna alisada em demasiado. Intuitivamente, quando  $h \rightarrow \infty$ , a forma da estimativa de densidade se aproxima da forma da função-núcleo escolhida e, quando  $h \rightarrow 0$ , a estimativa resulta numa soma de *funções- $\delta$ -Dirac* centradas nas observações, isto é, a *pseudo-derivada* da função de distribuição empírica.

Algumas características de estimadores definidos por (3.4) são:

- As propriedades da função-núcleo são herdadas por  $\hat{f}_h(x)$ . Por exemplo, se  $K$  tem  $n$  derivadas contínuas,  $\hat{f}_h(x)$  tem  $n$  derivadas contínuas;
- A função de densidade estimada para um determinado conjunto de dados, correspondente a  $h$  e  $K$ , é única; e
- O estimador de densidade por função-núcleo é assintoticamente não-viciado. O vício desse estimador é quadrático em  $h$ . Assim, quanto menor for o valor de  $h$ , menor o vício. No entanto, a variância é proporcional a  $((nh)^{-1})$ , ou seja, para diminuir a variância, devemos aumentar o valor de  $h$ . Portanto, para minimizar o Erro Quadrado Médio (EQM), é necessário equilibrar o vício e variância, o que implica em que  $h$  não pode ser nem *pequeno* nem *grande demais*.

### 3.2.1.1 Estimador de Nadaraya-Watson

Suponha que se tenham  $n$  observações de v.a.'s independentes e identicamente distribuídas,  $\{(X_i, Y_i)\}_{i=1}^n$ . Deseja-se encontrar uma estimativa para a esperança condicional  $m(x) = \mathbb{E}(Y|X = x)$ .

Substituindo-se, em (3.2), as densidades populacionais por suas versões empíricas, temos

$$\hat{m}(x) = \mathbb{E}(\widehat{Y|X} = x) = \int_{-\infty}^{+\infty} y \frac{\widehat{f}_{X,Y}(x, y)}{\widehat{f}_X(x)} dy, \quad (3.5)$$

em que  $\widehat{f}_{X,Y}$  é o análogo bidimensional do estimador  $\widehat{f}_X$ , dado por (3.4), isto é,

$$\widehat{f}_{X,Y}(x, y) = \frac{1}{n} \sum_{i=1}^n K_h(x - X_i) K_h(y - Y_i). \quad (3.6)$$

Das equações (3.4) e (3.6), obtém-se o estimador de regressão por função-núcleo, conhecido como Estimador de Nadaraya-Watson:

$$\hat{m}_h(x) = \int_{-\infty}^{+\infty} y \frac{\frac{1}{n} \sum_{i=1}^n K_h(x - X_i) K_h(y - Y_i)}{\frac{1}{n} \sum_{j=1}^n K_h(x - X_j)} dy = \frac{\sum_{i=1}^n K_h(x - X_i) Y_i}{\sum_{j=1}^n K_h(x - X_j)}. \quad (3.7)$$

Desse modo, os pesos têm a forma  $W_{hi}(x) = K_h(x - X_i) / \widehat{f}_h(x)$ .

Note que:

- Em geral, os pesos  $W_{hi}(x)$  dependem de toda a amostra  $\{X_j\}_{j=1}^n$  através da estimativa de densidade por função-núcleo  $\widehat{f}_h(x)$ ;
- As observações obtêm maiores pesos nas regiões onde as densidades empíricas correspondentes aos  $X_i$ 's são menores;
- Quando  $h \rightarrow 0$ ,  $W_{hi}(x) \rightarrow n$ , se  $x = X_i$ . Então, a estimativa em  $X_i$  converge para  $Y_i$  e produz uma *interpolação* dos dados;

- Quando  $h \rightarrow \infty$ ,  $W_{hi}(x) \rightarrow 1$ ,  $\forall x \in \mathbb{R}$ . Então, a estimativa de  $m(x)$  converge para  $\bar{Y}$ ; e
- O estimador  $\hat{m}_h(x)$  é consistente para a curva de regressão  $m(x)$ , com escolhas adequadas de  $h$ .

### 3.3 Desconvolução via Função-Núcleo

Nessa seção, apresentaremos o estimador por desconvolução via função-núcleo definido em Comte (2000), considerando-se  $(X_t, Y_t)$  e o modelo (1.10)-(1.11), isto é,

$$\begin{aligned} X_t &= m(X_{t-1}) + \eta_t, \\ Y_t &= X_t + \varepsilon_t. \end{aligned}$$

A função  $m(x)$  é dada por

$$\mathbb{E}(Y_{j+1}|X_j = x) = m(x) = \int_{-\infty}^{+\infty} y \frac{f_{X,Y}(x, y)}{f_X(x)} dy,$$

em que  $f_{X,Y}$  denota a densidade conjunta de  $(X_1, Y_1)$  e  $f_X$  a densidade de  $X_1$ .

Se tanto  $X_j$ 's e  $Y_j$ 's fossem observáveis, poderíamos usar Nadaraya-Watson:

$$\hat{m}(x) = \frac{\sum_{j=1}^n Y_{j+1} K\left(\frac{x - X_j}{h}\right)}{\sum_{j=1}^n K\left(\frac{x - X_j}{h}\right)}. \quad (3.8)$$

Em nosso caso, os  $X_j$ 's são não observáveis. Portanto, uma estratégia alternativa será empregada.

Observe que  $f_Y(y) = (f_X * f_\varepsilon)(y)$ , ou seja, a convolução entre as funções de densidade de  $X$  e  $\varepsilon$ .

Usando a Transformada Inversa de Fourier, (2.2), tem-se

$$f_X(x) = \frac{1}{2\pi} \left( \frac{f_Y^*}{f_\varepsilon^*} \right)^* (-x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \frac{f_Y^*(t)}{f_\varepsilon^*(t)} dt. \quad (3.9)$$

Isso motiva o estimador sugerido por Stefanski e Carroll (1990):

$$\hat{f}_n(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} K^*(th) \frac{\hat{\Phi}_n(t)}{f_\varepsilon^*(t)} dt, \quad (3.10)$$

em que  $\hat{\Phi}_n(t)$  é a função-característica empírica de  $Y$ , dada por:

$$\hat{\Phi}_n(t) = \frac{1}{n} \sum_{j=1}^n e^{itY_j}. \quad (3.11)$$

Definindo-se

$$K_n(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \frac{K^*(t)}{f_\varepsilon^*(t/h)} dt, \quad (3.12)$$

pode-se reescrever  $\hat{f}_n(x)$  como:

$$\begin{aligned} \hat{f}_n(x) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} K^*(th) \frac{\hat{\Phi}_n(t)}{f_\varepsilon^*(t)} dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \frac{K^*(th)}{f_\varepsilon^*(t)} \left[ \frac{1}{n} \sum_{j=1}^n e^{itY_j} \right] dt. \end{aligned}$$

Substituindo-se  $z = th$ , temos

$$\begin{aligned} \hat{f}_n(x) &= \frac{1}{nh2\pi} \int_{-\infty}^{+\infty} e^{-i\frac{zx}{h}} \frac{K^*(z)}{f_\varepsilon^*(z/h)} \left[ \sum_{j=1}^n e^{i\frac{zY_j}{h}} \right] dz \\ &= \frac{1}{nh2\pi} \int_{-\infty}^{+\infty} \sum_{j=1}^n e^{-iz\left(\frac{x-Y_j}{h}\right)} \frac{K^*(z)}{f_\varepsilon^*(z/h)} dz \\ &= \frac{1}{nh} \sum_{j=1}^n \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iz\left(\frac{x-Y_j}{h}\right)} \frac{K^*(z)}{f_\varepsilon^*(z/h)} dz \\ &= \frac{1}{nh} \sum_{j=1}^n K_n \left( \frac{x - Y_j}{h} \right). \end{aligned} \quad (3.13)$$

No numerador de (3.8), os  $X_j$ 's desconhecidos foram substituídos pelos  $Y_j$ 's e a função-núcleo  $K$ , pela função-núcleo de desconvolução  $K_n$ . Por último, segundo Fan e Truong (1993), e por analogia a (3.8), considera-se o estimador de  $m$ , baseado nas observações  $Y_1, \dots, Y_{n+1}$ :

$$\hat{m}_n(x) = \frac{\sum_{j=1}^n K_n\left(\frac{x-Y_j}{h}\right) Y_{j+1}}{\sum_{j=1}^n K_n\left(\frac{x-Y_j}{h}\right)} = \frac{\frac{1}{nh} \sum_{j=1}^n K_n\left(\frac{x-Y_j}{h}\right) Y_{j+1}}{\hat{f}_n(x)}, \quad (3.14)$$

em que  $K_n$  é dado por (3.12) e  $\hat{f}_n(x)$  por (3.13).

É importante enfatizar que não é necessário conhecer a distribuição da inovação  $\eta$  para construir  $\hat{m}_n(x)$ , mas é necessário conhecer completamente a distribuição de  $\varepsilon$ . Sob essas suposições,  $\hat{m}_n$  e  $\hat{f}_n$  podem ser usados para estimar  $\sigma_\eta^2$ :

$$\sigma_\eta^2 = \int (x^2 - \hat{m}_n^2(x)) \hat{f}_n(x) dx,$$

sendo  $\sigma_\eta^2 = Var(X) - Var(m(X)) = \mathbb{E}(X^2) - \mathbb{E}(m^2(X))$ . Como  $\mathbb{E}(Y^2) = \mathbb{E}(X^2) + \sigma_\varepsilon^2$ , temos

$$\hat{\sigma}_\eta^2 = \frac{1}{n} \sum_{j=1}^n Y_j^2 - \sigma_\varepsilon^2 - \int \hat{m}_n^2(x) \hat{f}_n(x) dx,$$

como um estimador de  $\sigma_\eta^2$  para as observações  $Y_1, \dots, Y_{n+1}$ .

## 3.4 Desconvolução via Ondaletas

O conceito de ondaletas já foi discutido na Seção 2.3. Mostraremos aqui, os estimadores por desconvolução lineares e não-lineares via ondaletas. Algumas condições, apresentadas na Subseção 3.4.1, sobre as ondaletas utilizadas e sobre os erros, são necessárias para que esses estimadores atinjam taxas de convergência ótimas.

### 3.4.1 Estimadores Lineares

Nessa seção, desenvolvemos os estimadores discutidos em Fan e Koo (2002). É

estabelecida uma taxa minimax para o melhor estimador linear sob o critério da perda quadrática integrada. Para atingir tal taxa mínima, inicialmente é estabelecido um limite inferior minimax para a classe dos estimadores por desconvolução lineares e, depois, são propostos estimadores lineares por desconvolução via ondaletas que atinjam esse limite. Fan e Koo (2002) mostram que os estimadores lineares podem alcançar a taxa ótima somente quando a classe de funções pertence a um espaço de Sobolev (um caso especial do espaço de Besov,  $B_{\sigma,2,q}$ ). Para os espaços de Besov gerais  $B_{\sigma pq}$ , com  $p < 2$ , nenhum estimador linear pode alcançar a taxa de convergência ótima.

A classe  $\mathcal{C}_L$  dos estimadores lineares é definida pela representação

$$\widehat{f}_L(Y_1, \dots, Y_n, x) = \sum_{j=1}^n T_j(Y_j, x), \quad (3.15)$$

em que  $\{T_j(\cdot, \cdot)\}$  são funções arbitrárias mensuráveis. Os estimadores por desconvolução induzidos por um método de séries ortogonais são da forma linear.

Inicialmente, serão apresentadas algumas condições necessárias. Essas condições não são conceitualmente independentes e podem até mesmo ser mutuamente exclusivas. Certamente, algumas delas são mais fortes que outras.

Abaixo, apresentam-se condições necessárias para as funções-escala e ondaletas-mãe. Usa-se a notação  $h^{(m)}$  para indicar a  $m$ -ésima derivada de uma função  $h$  e  $\mathcal{C}^r$  denota a espaço das funções  $h$  com todas as derivadas  $h^{(m)}$  contínuas, para  $m \leq r$ .

- A1. As funções  $\phi$  e  $\psi$  têm suporte compacto;
- A2. As funções  $\phi, \psi \in \mathcal{C}^r$ , para  $r \geq s + 2$  e  $s \geq 0$ ;
- A3.  $s \geq 2$ ;
- A4. Os suportes de  $\phi$  e  $\psi$  são  $\{t : |t| \leq 4\pi/3\}$  e  $\{t : 2\pi/3 \leq |t| \leq 8\pi/3\}$ , respectivamente; e
- A5. Para todo inteiro  $N \geq 1$  e  $0 \leq q \leq r$ , existe uma constante  $d_N > 0$  tal que

$$|\phi^{(q)}(x)| \leq d_N(1 + |x|)^{-N} \quad \text{e} \quad |\psi^{(q)}(x)| \leq d_N(1 + |x|)^{-N}.$$

Em particular, as ondaletas de Daubechies satisfazem A1-A3 e são indicadas

quando os erros pertencem à classe de erros regulares ordinários. Já as ondaletas de Meyer satisfazem A4 e A5 e têm bom desempenho no caso de erros super-regulares.

As condições necessárias para erros regulares ordinários são expostas em B1-B3.

$$\text{B1. } |f_\varepsilon^*(t)| \leq C(1 + |t|)^{-s};$$

$$\text{B2. } |f_\varepsilon^*(t)| \geq C(1 + |t|)^{-s}; \text{ e}$$

$$\text{B3. } |f_\varepsilon^{*(m)}(t)| \leq C(1 + |t|)^{-s-m}, \text{ para } m = 0, 1, 2.$$

As distribuições Exponencial Dupla e Gama são exemplos de erros regulares ordinários. Note que a condição B2 implica que  $f_\varepsilon^*(t) \neq 0$  para qualquer  $t$ .

Quando a distribuição dos erros é super-regular, as condições a seguir são requeridas:

$$\text{C1. } |f_\varepsilon^*(t)| \leq C(1 + |t|)^{s_1} e^{-|t|^s/\lambda}, \text{ com } s, \lambda \text{ e } s_1 \text{ constantes sendo } s, \lambda > 0;$$

$$\text{C2. } \mathbf{P}(|\varepsilon - x| \leq |x|^{\bar{\alpha}}) = O(|x|^{-(a-\bar{\alpha})}), \text{ quando } x \rightarrow \pm\infty \text{ para algum } 0 < \bar{\alpha} < 1, \\ a > 1 + \bar{\alpha};$$

$$\text{C3. } |f_\varepsilon^*(t)| \geq C(1 + |t|)^{s_0} e^{-|t|^s/\lambda}, \text{ com } s, \lambda \text{ e } s_0 \text{ constantes sendo } s, \lambda > 0; \text{ e}$$

$$\text{C4. Para } m = 0, 1, 2,$$

$$|f_\varepsilon^{*(m)}(t)| \leq C(1 + |t|)^{s_2} e^{-|t|^s/\lambda},$$

com  $s, \lambda > 0$  e  $s_2$  uma constante.

Se  $\varepsilon$  tem distribuição Normal ou Cauchy, então C1-C3 são satisfeitas. A distribuição Normal dos erros também satisfaz a condição C4 com  $s_2 = 2$ . Observe que a condição C3 implica em que  $f_\varepsilon^*(t) \neq 0$  para qualquer  $t$ .

Tendo em vista essas condições, Fan e Koo (2002) propõem estimadores não viciados dos coeficientes de ondaletas  $\alpha_{j_0, k}$  da função de densidade desconhecida  $f_X$ .

Desde que a densidade desconhecida  $f_X$  seja elemento de  $\mathcal{L}^2(\mathbb{R})$ , ela pode ser decomposta em coeficientes de ondaletas como em (2.19). Sendo assim, os coeficientes

de ondaletas no nível  $j_0$ , pela identidade de Plancherel, podem ser escritos como:

$$\begin{aligned}\alpha_{j_0,k} &= \int_{-\infty}^{+\infty} \phi_{j_0,k}(x) f_X(x) dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi_{j_0,k}^*(-t) \frac{f_Y^*(t)}{f_\varepsilon^*(t)} dt.\end{aligned}\quad (3.16)$$

A função-característica desconhecida  $f_Y^*$  pode ser estimada pela função-característica empírica da v.a.  $Y$ , vista em (3.11). Dessa forma, define-se o estimador para  $\alpha_{j_0,k}$ :

$$\hat{\alpha}_{j_0,k} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi_{j_0,k}^*(-t) \frac{\hat{\Phi}_n(t)}{f_\varepsilon^*(t)} dt, \quad (3.17)$$

em que  $\phi_{j_0,k}^*(-t) = 2^{\frac{j_0}{2}} \phi^*(-2^{j_0}t - k) = 2^{-\frac{j_0}{2}} e^{-itk} 2^{-j_0} \phi^*(-t/2^{j_0})$ , pelas propriedades de linearidade, escala e localização das Transformadas de Fourier.

Desenvolvendo-se a equação (3.17) e fazendo-se  $w = 2^{-j_0}t$ , obtém-se a seguinte expressão:

$$\begin{aligned}\hat{\alpha}_{j_0,k} &= \frac{1}{n} \sum_{j=1}^n \frac{1}{2\pi} \int_{-\infty}^{+\infty} 2^{-j_0/2} e^{-itk} 2^{-j_0} \frac{\phi^*(-t/2^{j_0})}{f_\varepsilon^*(t)} e^{itY_j} dt \\ &= \frac{1}{n} \sum_{j=1}^n \frac{1}{2\pi} \int_{-\infty}^{+\infty} 2^{j_0/2} e^{-iwk} \frac{\phi^*(-w)}{f_\varepsilon^*(2^{j_0}w)} e^{i2^{j_0}wY_j} dw \\ &= \frac{1}{n} \sum_{j=1}^n \frac{1}{2\pi} \int_{-\infty}^{+\infty} 2^{j_0/2} e^{iw(2^{j_0}Y_j - k)} \frac{\phi^*(-w)}{f_\varepsilon^*(2^{j_0}w)} dw \\ &= \frac{1}{n} \sum_{j=1}^n 2^{j_0/2} (\mathcal{K}_{j_0}^- \phi)(2^{j_0}Y_j - k) \\ &= \frac{1}{n} \sum_{j=1}^n (\mathcal{K}_{j_0}^- \phi)_{j_0,k}(Y_j).\end{aligned}\quad (3.18)$$

O operador  $\mathcal{K}_j^-$  aplicado em um função  $g$  é formalmente definido como:

$$(\mathcal{K}_j^- g)(y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ity} \frac{g^*(-t)}{f_\varepsilon^*(2^j t)} dt. \quad (3.19)$$

Sob a condição A2, tem-se que

$$|\phi^{*(m)}(t)| \leq C(1 + |t|)^{-r} \quad \text{e} \quad |\psi^{*(m)}(t)| \leq C(1 + |t|)^{-r}, \quad (3.20)$$

para  $m \geq 0$ . Isso implica em que  $\mathcal{K}_{j_0}^- \phi$  e  $\mathcal{K}_j^- \psi$  são bem definidos para distribuições de erros regulares ordinárias de ordem  $s$ . Por outro lado, o uso das ondaletas de Meyer, cuja transformada de Fourier tem suporte compacto, implica em que  $\mathcal{K}_{j_0}^- \phi$  e  $\mathcal{K}_j^- \psi$  são também bem definidos.

Sendo assim, tem-se o estimador linear por ondaletas definido por:

$$\hat{f}_{j_0}(x) = \sum_k \hat{\alpha}_{j_0,k} \phi_{j_0,k}(x), \quad x \in \mathbb{R}. \quad (3.21)$$

Através da escolha de um  $j_0$  apropriado, o estimador de Fan e Koo (2002), definido por (3.17)-(3.21), alcança uma taxa de convergência ótima entre os estimadores lineares.

As ondaletas de Daubechies de suporte compacto, satisfazendo as condições A1-A3, devem ser empregadas no problema de desconvolução para a classe de erros regulares ordinários. O teorema 3.4.1 mostra a melhor taxa possível para a classe de estimadores lineares, quando a distribuição dos erros é regular ordinária.

**Teorema 3.4.1.** *Suponha que  $1 \leq q \leq \infty$ ,  $1 \leq p \leq 2$  e  $\sigma \geq 1/p$ . Então, sob as condições A1, A2 e B3, tem-se que*

$$\inf_{\hat{f}_L \in \mathcal{C}_L} \sup_{f_X \in D_{\sigma pq}(M)} \left[ \mathbb{E}_{f_X} \|\hat{f}_L - f_X\|_2^2 \right]^{1/2} \geq C n^{-\sigma'/(2\sigma'+2s+1)}. \quad (3.22)$$

Como  $\phi$  tem suporte compacto, a somatória em  $k$  do estimador linear de desconvolução via ondaletas (3.21) é finito para cada  $x \in \mathbb{R}$ . O teorema 3.4.2 mostra que o limite inferior sobre os estimadores lineares é alcançado pelos estimadores de ondaletas lineares, quando o erro é regular ordinário.

Sejam duas seqüências positivas  $\{a_n\}$  e  $\{b_n\}$ . Então a relação  $a_n \asymp b_n$  significa que, para  $C$  uma constante,  $C^{-1} \leq a_n/b_n < C$ .

**Teorema 3.4.2.** *Suponha que  $1 \leq q \leq \infty$ ,  $1 \leq p \leq 2$  e  $\sigma \geq 1/p$ . Então, sob as condições A1, A2 e B2, tem-se que*

$$\sup_{f_X \in D_{\sigma pq}(M)} \left[ \mathbb{E}_{f_X} \|\hat{f}_{j_0} - f_X\|_2^2 \right]^{1/2} \leq C n^{-\sigma'/(2\sigma'+2s+1)}, \quad (3.23)$$

resultando que  $2^{j_0} \asymp n^{1/(2\sigma'+2s+1)}$ .

O resultado do teorema acima também é válido para a classe  $f_X \in D_{\sigma pq}(M)$ , sob as condições adicionais A3 e B3.

Para erros super-regulares, ondaletas de Meyer satisfazendo A4 e A5 têm desempenho ótimo, pelo teorema 3.4.3.

**Teorema 3.4.3.** *Se  $1 \leq p \leq 2$ ,  $q \leq \infty$  e  $\sigma \geq 1/p$ , então, sob as condições A4, A5, C1 e C2, tem-se que*

$$\inf_{\hat{f}_L \in \mathcal{C}_L} \sup_{f_X \in D_{\sigma pq}(M)} \left[ \mathbb{E}_{f_X} \|\hat{f}_L - f_X\|_2^2 \right]^{1/2} \geq C(\log n)^{-\sigma'/s}. \quad (3.24)$$

**Lema 3.4.1.** *Suponha que  $g(y)$  é uma função sobre  $\mathbb{R}$ , tal que*

$$|g(y)| \leq C(1 + |y|)^{-2}. \quad (3.25)$$

Então, para qualquer seqüência de escalares  $\{\alpha_k\}_{k \in \mathbb{Z}}$ , tem-se que, para  $1 \leq p \leq \infty$

$$\left\| \sum_{k \in \mathbb{Z}} \alpha_k h_{j,k} \right\|_p \leq C \|\alpha\|_p. \quad (3.26)$$

O Lema 3.4.1 implica em que, para todo  $x \in \mathbb{R}$ ,

$$\left\| \sum_k \hat{\alpha}_{j_0,k} \phi_{j_0,k}(x) \right\|_\infty \leq C \sup_k |\hat{\alpha}_{j_0,k}| \leq C, \quad (3.27)$$

sob as condições C1-C3. Sendo assim, o estimador linear por ondaletas (3.21) é também bem definido para o caso super-regular. De acordo com o teorema 3.4.4, o limite inferior entre os estimadores lineares é atingido pelos estimadores lineares por ondaletas quando os erros são super-regulares.

**Teorema 3.4.4.** *Suponha que  $1 \leq q \leq \infty$ ,  $1 \leq p \leq 2$  e  $\sigma \geq 1/p$ . Sob as condições A4, A5 e C1-C3, tem-se que*

$$\sup_{f_X \in D_{\sigma pq}(M,S)} \left[ \mathbb{E}_{f_X} \|\hat{f}_{j_0} - f_X\|_2^2 \right]^{1/2} \leq C(\log n)^{-\sigma'/s}, \quad (3.28)$$

sendo

$$j_0 = \left\lceil \log_2 \left( \frac{3}{4\pi} \left( \frac{\lambda}{4} \right)^{1/s} (\log n)^{1/s} \right) \right\rceil. \quad (3.29)$$

O resultado acima também é válido para a classe  $f_X \in D_{\sigma pq(M)}$  sob a condição adicional C4.

Considere  $\alpha = \sigma/(2\sigma + 2s + 1)$ , e  $\mathcal{A}$  a classe dos estimadores de  $f_X$  baseados em  $Y_1, \dots, Y_n$ , definidos por  $\hat{f}_n^*$ . Tem-se o seguinte limite inferior minimax para as distribuições dos erros regulares ordinárias.

**Teorema 3.4.5.** *Suponha que  $1 \leq q \leq \infty$  e  $\sigma > 1/p$ . Sob as condições A1-A3 e B3, tem-se que*

$$\inf_{\hat{f}_n^* \in \mathcal{A}} \sup_{f_X \in D_{\sigma pq(M)}} \left[ \mathbb{E}_{f_X} \|\hat{f}_n^* - f_X\|_2^2 \right]^{1/2} \geq Cn^{-\alpha}. \quad (3.30)$$

Para erros super-regulares, os estimadores lineares via ondaletas tem uma taxa de convergência muito baixa; no entanto, são ótimos dentro da ordem logarítmica, como pode ser visto pelo teorema 3.4.6

**Teorema 3.4.6.** *Suponha que  $1 \leq q \leq \infty$  e  $\sigma > 1/p$ . Sob as condições A4, A5, C1 e C2,*

$$\inf_{\hat{f}_n^* \in \mathcal{A}} \sup_{f_X \in D_{\sigma pq(M)}} \left[ \mathbb{E}_{f_X} \|\hat{f}_n^* - f_X\|_2^2 \right]^{1/2} \geq C(\log n)^{-\alpha/s}. \quad (3.31)$$

### 3.4.2 Estimadores Não-lineares

Nessa seção, iremos mostrar a construção dos estimadores não-lineares por desconvolução via ondaletas, assim como as taxas minimax de convergência para a estimação da densidade  $f_X$  em  $B_{\sigma pq}$ . Para erros da classe regular ordinária, Fan e Koo (2002) mostram que estimadores por desconvolução via ondaletas com limiar são ótimos dentro da ordem logarítmica. No entanto, para o caso super-regular, o desempenho dos estimadores por desconvolução não-lineares via ondaletas com limiar é semelhante ao dos estimadores lineares por ondaletas.

Analogamente ao que foi feito para encontrar o estimador de  $\alpha_{j_0, k}$ , escreve-se

$\beta_{j,k}$  como:

$$\begin{aligned}\beta_{j,k} &= \int_{-\infty}^{+\infty} \psi_{j,k}(x) f_X(x) dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \psi_{j,k}^*(-t) \frac{f_Y^*(t)}{f_\varepsilon^*(t)} dt.\end{aligned}\quad (3.32)$$

Substituindo-se  $f_Y^*$  por sua função-característica empírica, tem-se o estimador:

$$\hat{\beta}_{j,k} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \psi_{j,k}^*(-t) \frac{\hat{\Phi}_n(t)}{f_\varepsilon^*(t)} dt.\quad (3.33)$$

Desenvolvendo-se a expressão (3.33) como em (3.18) e empregando-se o operador  $\mathcal{K}_j^-$  na função  $\psi$ , temos:

$$\hat{\beta}_{j,k} = \frac{1}{n} \sum_{j=1}^n (\mathcal{K}_j^- \psi)_{j,k}(Y_j).\quad (3.34)$$

Pelas propriedades da AMR, o estimador visto em (3.21) pode também ser escrito como:

$$\hat{f}_{j_1, j_0}(x) = \sum_k \hat{\alpha}_{j_0, k} \phi_{j_0, k}(x) + \sum_{j_0}^{j_1} \sum_k \hat{\beta}_{j, k} \psi_{j, k}(x), \quad x \in \mathbb{R},\quad (3.35)$$

em que  $j_0$  e  $j_1$  são, respectivamente, os níveis mínimo e máximo de resolução utilizados.

É comum em análise de ondaletas empregar um limiar nos coeficientes. Essa técnica, também conhecida como *thresholding*, se faz necessária, para que possamos utilizar apenas os coeficientes que realmente são importantes para a reconstrução do sinal e desprezar coeficientes que expressam apenas o ruído amostral. Existem dois tipos de limiar:

- Limiar Suave: procedimento que encolhe os coeficientes; e
- Limiar Duro: procedimento que anula os coeficientes considerados muito pequenos.

Fan e Koo (2002) sugerem a seguinte regra de limiar duro:

$$\tilde{\beta}_{j,k} = \begin{cases} \hat{\beta}_{j,k} & \text{se } |\hat{\beta}_{j,k}| > Tc_j n^{-1/2} \\ 0 & \text{se } |\hat{\beta}_{j,k}| \leq Tc_j n^{-1/2} \end{cases}, \quad (3.36)$$

sendo  $c_j = 2^{sj} \sqrt{j}$  e a constante  $T$  determinada de acordo com o Teorema 3.4.7.

O estimador que emprega os coeficientes de detalhes da decomposição de ondaletas encolhidos é conhecido como estimador por desconvolução não-linear. Esse estimador, associado com os parâmetros  $j_0(n)$ ,  $j_1(n)$  e  $T$ , é dado por:

$$\hat{f}_T(x) = \sum_k \hat{\alpha}_{j_0,k} \phi_{j_0,k}(x) + \sum_{j_0}^{j_1} \sum_k \tilde{\beta}_{j,k} \psi_{j,k}(x), \quad x \in \mathbb{R}. \quad (3.37)$$

Os níveis de resolução  $j_0$  e  $j_1$  são escolhidos de forma que:

$$2^{j_0(n)} \asymp (n(\log n)^{(2-p)/p})^{1/(2\sigma+2s+1)} \quad (3.38)$$

e

$$2^{j_1(n)} \asymp n^{\alpha/\sigma'}. \quad (3.39)$$

**Teorema 3.4.7.** *Suponha que  $1 \leq q \leq \infty$ ,  $1 \leq p \leq 2$ ,  $\sigma > 1/p$  e  $s < \{p/(2-p)\}\sigma'$ . Então, sob as condições A1-A3 e B2, existem constantes  $C = C(\sigma, p, q, M)$  e  $T_0 = T_0(\sigma, p, M)$  tal que, para  $T > T_0$ ,*

$$\sup_{f_X \in D_{\sigma pq}(M, S)} \left[ \mathbb{E}_{f_X} \|\hat{f}_T - f_X\|_2^2 \right]^{1/2} \geq C n^{-\alpha} (\log n)^{(1-\delta/\sigma p)\alpha}, \quad (3.40)$$

em que  $\delta = \sigma p - \frac{2-p}{2} - s(2-p)$ .

## 4 Simulações

Nesse capítulo, serão apresentadas as simulações realizadas, tanto para o caso de desconvolução via função-núcleo quanto por ondaletas. Os programas foram desenvolvidos no aplicativo *Matlab*.

A função-núcleo será usada apenas para um caso específico de erros regulares ordinários. Devido à complexidade dos cálculos, não é possível utilizar seu estimador para o caso geral.

Faz-se um estudo mais detalhado para ondaletas. Inicialmente, essa técnica é utilizada para as duas classes de erro (super-regulares e regulares ordinários) num caso mais simples do modelo (1.11), para ilustrar o comportamento dos estimadores propostos. Depois, séries de retornos são geradas, tanto por um modelo GARCH(1,1) quanto por um AR(1)-SV, para as quais a densidade da log-volatilidade ou da volatilidade do modelo apresentado na Subseção 1.3.1 é estimada.

### 4.1 Função-Núcleo

A técnica de desconvolução é empregada em modelos de volatilidade estocástica com o objetivo de estimar a função  $m$ , dada em (1.7). Normalmente, assume-se que essa função seja linear. Nesse trabalho, simulamos casos em que a função  $m$  assume três formas distintas, sugeridas por Comte (2000), que serão detalhadas mais adiante.

Erros da classe regular ordinária são empregados, mais especificamente a distribuição Exponencial Dupla.

Assumindo-se a seguinte densidade para  $\xi$ ,

$$f_{\xi}(u) = \frac{1}{4} \mathbb{I}_{[-1,1]}(u) + \frac{1}{4u^2} \mathbb{I}_{(1,+\infty)}(|u|), \quad (4.1)$$

obtém-se uma distribuição exponencial dupla para  $\varepsilon$ , ou seja, um elemento da classe de erros regular ordinária de ordem  $s = 2$ .

Assim,

$$\begin{aligned} w &= \mathbb{E}(\ln(\xi^2)) = \frac{1}{4} \int_{-1}^1 \ln(x^2) dx + \frac{1}{4} \int_{|x| \geq 1} \frac{\ln(x^2)}{x^2} dx \\ &= [x \ln(x) - x]_0^1 + \int_1^{+\infty} u e^{-2u} e^u = -1 + \Gamma(2) = 0. \end{aligned}$$

A variância de  $\varepsilon = \ln(\xi^2)$  é dada por:

$$\begin{aligned} Var(\varepsilon) &= Var(\ln(\xi^2) - w) = (\mathbb{E}(\ln(\xi^2))^2) \\ &= 2 \left[ \int_0^{+\infty} u^2 e^{-u} du + \int_0^{+\infty} \frac{u^2 e^u}{e^{2u}} du \right] \\ &= 2 [\Gamma(3) + \Gamma(3)] = 8. \end{aligned}$$

Como  $\varepsilon = \ln(\xi^2) - w = \ln(\xi^2)$ ,  $\varepsilon$  segue uma distribuição exponencial dupla e sua densidade é dada por:

$$\begin{aligned} f_{\varepsilon}(v) &= \left| \frac{1}{2} e^{\frac{v}{2}} \right| f_{\xi}(e^{\frac{v}{2}}) + \left| \frac{-1}{2} e^{\frac{v}{2}} \right| f_{\xi}(-e^{\frac{v}{2}}) \\ &= \frac{1}{4} e^{\frac{-|v|}{2}} \mathbb{I}_{(-\infty, +\infty)}(v). \end{aligned}$$

De forma análoga ao cálculo de  $Var(\varepsilon)$ , temos a transformada de Fourier de  $f_{\varepsilon}$ :

$$\begin{aligned} f_{\varepsilon}^*(t) &= \mathbb{E}(e^{it \ln(\xi^2)}) = \frac{1}{4} \left[ \int_{-1}^1 e^{it \ln(x^2)} dx + \int_{|x| \geq 1} \frac{e^{it \ln(x^2)}}{x^2} dx \right] \\ &= \frac{1}{2} \left[ \frac{1}{1 + 2it} + \frac{1}{1 - 2it} \right] = \frac{1}{1 + 4t^2}. \end{aligned} \quad (4.2)$$

O estimador de  $m$ ,

$$\hat{m}_n(x) = \frac{\frac{1}{nh} \sum_{j=1}^n K_n \left( \frac{x-Y_j}{h} \right) Y_{j+1}}{\hat{f}_n(x)},$$

será obtido calculando-se separadamente as estimativas do numerador e denominador. O denominador é dado por (3.13) e o numerador é estimado da seguinte forma:

$$\begin{aligned} \frac{1}{nh} \sum_{j=1}^n K_n \left( \frac{x-Y_j}{h} \right) Y_{j+1} &= \frac{1}{nh} \sum_{j=1}^n \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-it \left( \frac{x-Y_j}{h} \right)} \frac{K^*(t)}{f_\varepsilon^*(t/h)} dt Y_{j+1} \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i \frac{tx}{h}} \frac{K^*(t)}{f_\varepsilon^*(t/h)} \left[ \frac{1}{nh} \sum_{j=1}^n e^{i \frac{tY_j}{h}} Y_{j+1} \right] dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \frac{K^*(th)}{f_\varepsilon^*(t)} \left[ \frac{1}{n} \sum_{j=1}^n e^{itY_j} Y_{j+1} \right] dt. \end{aligned}$$

O problema está no cálculo da função-núcleo  $K_n$  diretamente, uma vez que tem-se que trabalhar com números complexos. Uma das alternativas encontradas para resolver essa questão é utilizar a técnica de quadratura, como pode ser visto em Press (1992), para aproximar  $K_n$ . Essa dificuldade no cálculo de  $K_n$  impede a utilização do estimador para o caso geral, assumindo-se qualquer distribuição para os erros.

Os  $\varepsilon$ 's utilizados aqui serão como descritos anteriormente juntamente com a função-núcleo gaussiana,  $K(x) = e^{-x^2/2}/\sqrt{2\pi}$ , cuja transformada de Fourier é dada por  $K^*(t) = e^{-t^2/2}$ .

Considerando-se a definição de  $K_n$ , dada pela equação (3.12), e  $f_\varepsilon^*$ , dada por (4.2), será utilizado o seguinte  $K_n$ :

$$\begin{aligned} K_n(x) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} e^{-t^2/2} \left[ 1 + 4 \left( \frac{t}{h} \right)^2 \right] dt \\ &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \int_{-\infty}^{+\infty} \left[ 1 + \frac{4}{h^2} t^2 \right] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t+ix)^2} dt. \end{aligned}$$

A última integral da expressão acima pode ser escrita como a esperança de uma

função de  $T$ , sendo  $T \sim N(-ix, 1)$ . Assim temos,

$$\begin{aligned} K_n(x) &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \mathbb{E} \left( 1 + \frac{4}{h^2} T^2 \right) \\ &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \left( 1 + \frac{4(1+x^2)}{h^2} \right). \end{aligned} \quad (4.3)$$

Foram utilizados três modelos para  $m(x)$ , sugeridos por Comte (2000):

1. Modelo Linear (L(a)):  $m(x) = ax + 1$ ;
2. Modelo Senoidal (S(a)):  $m(x) = \text{sen}(ax) + 2$ ; e
3. Modelo de Fan e Yao (1998) (FY(a)):  $m(x) = a(x + 5e^{-x^2})$ .

A escolha do parâmetro  $a$  é restrita à condição de manter a estabilidade da parte autoregressiva dos modelos, ou seja, é necessário que  $|a| < 1$ . Dessa forma, foram realizadas 500 replicações de cada um dos modelos citados acima para os valores de  $a$ : 0,3; 0,5 e 0,7. Além de obterem-se as estimativas empregando-se a função-núcleo obtida por desconvolução, também foram calculadas as estimativas usando-se a função-núcleo gaussiana.

Existem várias medidas de qualidade das estimativas. A medida utilizada para esse caso é o Erro Quadrático Médio Integrado (EQMI) que, considerando-se uma malha de  $p$  pontos, no nosso caso 100, num intervalo igualmente espaçado entre o mínimo e o máximo da amostra  $Y_1, Y_2, \dots, Y_{500}$ , gerada a partir dos modelos (1.10) e (1.11), é definido por:

$$\text{EQM} = \mathbb{E}(\hat{m}_n(x) - m(x))^2 = \frac{1}{p} \sum_{i=1}^p (\hat{m}_n(x_i) - m(x_i))^2.$$

A Tabela 4.3 apresenta a média dos EQM's para os modelos L(a), S(a) e FY(a), respectivamente, de acordo com o valor de  $a$  e o valor de  $h$ .

As Figuras 4.10 e 4.11 apresentam as estimativas pela função-núcleo  $K_n$  e pela gaussiana para os modelos S(0,7) e FY(0,5), respectivamente, para valores de  $h$ : 0,12; 0,16; 0,20 e 0,30.

Tabela 4.3: Média dos EQM's para Diferentes Valores de  $h$

$h$	Função Núcleo	L(a)			S(a)			FY(a)		
		0,3	0,5	0,7	0,3	0,5	0,7	0,3	0,5	0,7
0,12	$K_n$	0,141	0,152	0,165	0,129	0,144	0,147	0,144	0,231	0,312
	Normal	0,176	0,182	0,202	0,160	0,179	0,177	0,174	0,262	0,346
0,14	$K_n$	0,103	0,131	0,161	0,105	0,116	0,138	0,149	0,232	0,239
	Normal	0,129	0,161	0,191	0,131	0,147	0,169	0,177	0,260	0,265
0,16	$K_n$	0,105	0,140	0,130	0,106	0,115	0,126	0,145	0,212	0,272
	Normal	0,136	0,166	0,157	0,131	0,144	0,150	0,170	0,228	0,291
0,18	$K_n$	0,108	0,091	0,142	0,086	0,105	0,154	0,133	0,179	0,215
	Normal	0,138	0,114	0,165	0,112	0,129	0,182	0,156	0,198	0,219
0,20	$K_n$	0,107	0,131	0,114	0,073	0,104	0,127	0,127	0,184	0,251
	Normal	0,130	0,157	0,135	0,096	0,126	0,152	0,143	0,197	0,252
0,22	$K_n$	0,094	0,109	0,109	0,058	0,082	0,101	0,120	0,185	0,245
	Normal	0,117	0,127	0,131	0,075	0,102	0,116	0,134	0,187	0,239
0,24	$K_n$	0,075	0,093	0,106	0,063	0,079	0,093	0,127	0,195	0,274
	Normal	0,093	0,111	0,123	0,083	0,094	0,107	0,139	0,189	0,259
0,26	$K_n$	0,084	0,081	0,101	0,062	0,078	0,106	0,098	0,195	0,281
	Normal	0,103	0,095	0,116	0,079	0,092	0,120	0,103	0,183	0,252
0,28	$K_n$	0,064	0,081	0,119	0,060	0,063	0,100	0,097	0,214	0,303
	Normal	0,077	0,093	0,132	0,074	0,073	0,109	0,101	0,201	0,267
0,30	$K_n$	0,068	0,076	0,116	0,046	0,067	0,079	0,102	0,221	0,301
	Normal	0,082	0,086	0,125	0,058	0,078	0,085	0,103	0,200	0,258

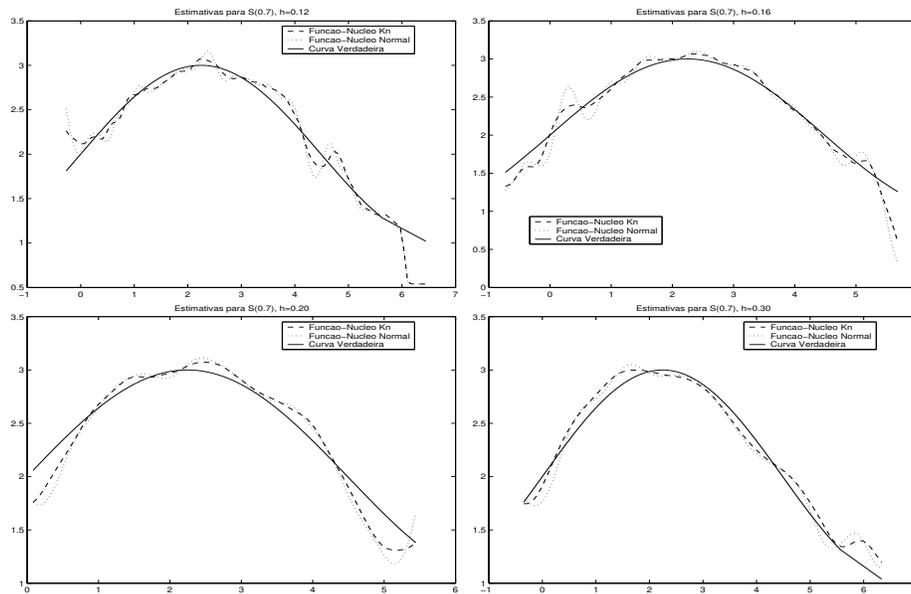


Figura 4.10: Estimativas para Modelo  $S(0,7)$ ,  $h = 0,12; 0,16; 0,20; 0,30$

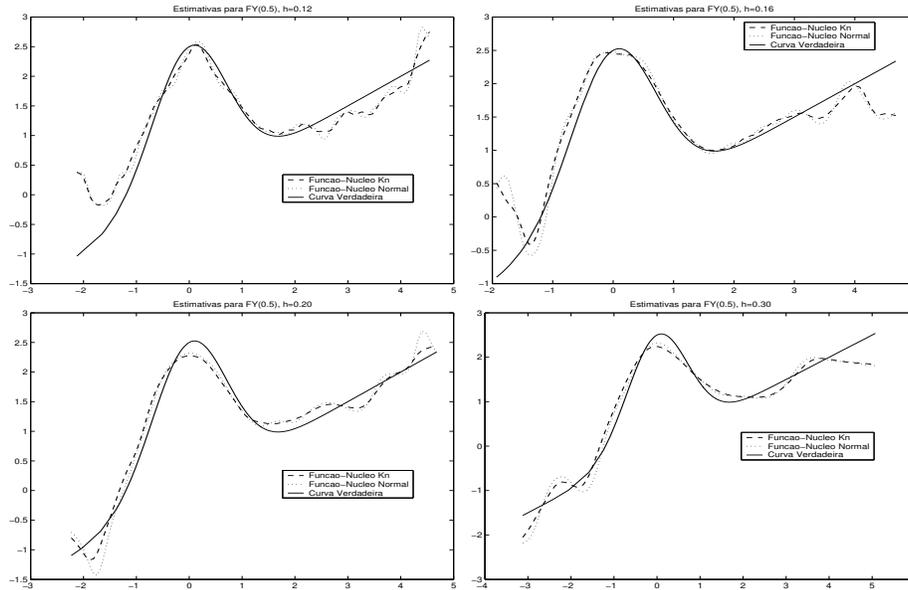


Figura 4.11: Estimativas para Modelo  $FY(0,5)$ ,  $h = 0, 12; 0, 16; 0, 20; 0, 30$

As estimativas obtidas se comportam razoavelmente bem no intervalo escolhido de  $h$ . Note que, se a variância dos erros for muito pequena, a variável  $Y$  é praticamente igual à variável  $X$ , o que resulta em estimativas por  $K_n$  e pela função-núcleo gaussiana praticamente iguais. No entanto, é importante ressaltar que a estimativa mais realista é aquela que utiliza  $K_n$ , pela justificativa teórica para sua construção.

Pelos valores dos EQM's, é possível determinar o  $h$  ótimo empírico para cada modelo e para cada tipo de função-núcleo utilizada, como visto na Tabela 4.4

Tabela 4.4: Parâmetro de Alisamento Ótimo para os Modelos  $L(a)$ ,  $S(a)$  e  $FY(a)$

Função Núcleo	L(a)			S(a)			FY(a)		
	0,3	0,5	0,7	0,3	0,5	0,7	0,3	0,5	0,7
$K_n$	0,28	0,30	0,26	0,30	0,28	0,30	0,28	0,18	0,18
Normal	0,30	0,30	0,26	0,30	0,28	0,30	0,28	0,26	0,18

As estimativas conseguem reproduzir bem as forma das curvas, principalmente na região central da amostra. Nota-se que nos valores extremos a qualidade das estimativas não é muito satisfatória.

Seria interessante poder aplicar o estimador para o caso geral, em que fosse possível escolher qualquer distribuição para  $\varepsilon$ . No entanto, devido às dificuldades

computacionais de trabalhar-se com integração complexa e o custo computacional, foi-nos possível mostrar apenas um caso específico.

## 4.2 Ondaletas

No caso do estimador por desconvolução via ondaletas, inicialmente obtêm-se as estimativas dos coeficientes  $\hat{\alpha}_{j_0,k}$  e  $\hat{\beta}_{j,k}$  vistos em (3.17) e (3.33), respectivamente, e, a partir desses coeficientes, é possível reconstruir a estimativa  $\hat{f}_{j_1,j_0}$  da densidade  $f_X$ , dada em (3.35) ou (3.37). Ambos os estimadores, linear e não-linear, foram utilizados. Não houve diferença nos desempenhos desses estimadores quando se empregam as ondaletas de Meyer. Dessa forma, mostraremos as estimativas lineares e não-lineares apenas para as bases de Daubechies.

Para estudar as propriedades de cada estimador proposto, será inicialmente simulado um caso bem simples no qual a densidade a ser estimada é a curva normal-padrão.

Considere o modelo:

$$Y = X + \varepsilon,$$

em que  $X$  e  $\varepsilon$  são v.a.'s independentes. Nosso interesse é estimar a densidade da variável aleatória  $X$ ,  $f_X$ , sendo que se observam apenas os valores de  $Y$ . Iremos supor que  $X$  siga uma distribuição  $N(0, 1)$ . O que muda de uma simulação para a outra é a distribuição associada ao erro. Erros com distribuição normal padrão são utilizados como exemplos da classe super-regulares e erros com distribuição exponencial dupla como exemplo da classe regular ordinária.

Teoricamente, uma função pode ser escrita como uma combinação linear de bases de ondaletas, mas o que se tem na realidade é uma aproximação para essa função, já que se utiliza um número finito de coeficientes. Sendo assim, além da curva verdadeira, será considerada também a densidade construída a partir dos coeficientes teóricos dados por (3.16) e (3.32), chamada de  $\tilde{f}_{j_1,j_0}$ , ou densidade aproximada.

Dessa forma, as estimativas obtidas serão comparadas com a verdadeira densidade e com a densidade aproximada através do EQM. Este último é chamado de Erro de Estimção. Para cada caso, será indicado qual o Erro de Aproximação.

Muitas dificuldades foram encontradas para obter as estimativas por desconvolução, utilizando-se as bases de ondaletas, pela complexidade computacional dos cálculos. Note que, na reconstrução das estimativas a partir dos coeficientes de ondaletas, é necessário conhecer as funções  $\phi_{j_0,k}$  e  $\psi_{j,k}$  numa malha de pontos pré-determinada. Esse fato não é obtido de forma automática. Sendo assim, utilizamos métodos de interpolação para obter as funções  $\phi_{j_0,k}$  e  $\psi_{j,k}$  nos pontos desejados. Apesar de ser um método bem utilizado e eficiente, sabe-se que existe um erro associado à aproximação de uma função  $g(x)$  por  $\hat{g}(x)$  através de interpolação.

Considere a malha de pontos  $x_1, x_2, \dots, x_n$  nos quais a função  $g$  é conhecida. O interesse é avaliar o valor dessa função num ponto  $x$ , tal que  $x_i \leq x \leq x_{i+1}$ . Uma das formas mais simples e intuitiva é a interpolação linear, ou seja, para cada valor  $x \in [x_i, x_{i+1}]$ , ajusta-se uma reta que passa pelos pontos  $x_i$  e  $x_{i+1}$ . Com a equação da reta, encontra-se o valor de  $g(x)$ . Isto é,

$$\hat{g}(x) = g(x_i) + \frac{g(x_{i+1}) - g(x_i)}{x_{i+1} - x_i}(x - x_i), \quad \text{se } x_i \leq x \leq x_{i+1}. \quad (4.4)$$

O erro cometido ao utilizar essa aproximação é dado por  $e(x) = g(x) - \hat{g}(x)$ . Note que esse erro depende da concavidade de  $g(x)$ , ou seja, de  $g''(x)$  no intervalo  $[x_i, x_{i+1}]$ . Se  $g(x)$  têm derivada de segunda ordem contínua em  $[x_i, x_{i+1}]$ , então o erro é dado pela seguinte fórmula:

$$e(x) = (x - x_i)(x - x_{i+1})\frac{g''(\xi_x)}{2}, \quad \text{com } \xi_x \in [x_i, x_{i+1}].$$

No entanto, a fórmula acima tem uso limitado na prática, uma vez que, raramente,  $g''(\xi_x)$  é conhecida e o ponto  $\xi_x$  nunca é conhecido. Então, utiliza-se um limitante para o erro, como visto em Ruggiero e Lopes (1996):

$$\begin{aligned} |e(x)| &= |(x - x_i)(x - x_{i+1})|\frac{|g''(\xi_x)|}{2} \\ &\leq |(x - x_i)(x - x_{i+1})|\frac{M_2}{2} \\ &\leq \frac{(x_{i+1} - x_i)^2 M_2}{4}, \end{aligned} \quad (4.5)$$

em que  $M_2 = \max_{x \in [x_i, x_{i+1}]} |g''(x)|$ .

As estimativas obtidas para cada classe de erro serão apresentadas, consideran-

do-se tanto as bases de Daubechies como as ondaletas de Meyer. Os resultados estão resumidos por tabelas e gráficos. As tabelas apresentam os EQM's quando se compara a densidade estimada à verdadeira densidade de probabilidade e à densidade obtida a partir dos coeficientes teóricos. As quantidades entre parênteses, são os desvios-padrão dos EQM's. Além disso, aplica-se o teste de aderência de Kolmogorov-Smirnov para verificar a normalidade das estimativas, calcula-se o p-valor médio e a porcentagem de amostras rejeitadas para um nível de significância de 5% e 1%, respectivamente. Nos gráficos, têm-se as estimativas de densidade e as funções de distribuição acumulada.

## 4.2.1 Família Daubechies

### 4.2.1.1 Erros Regulares Ordinários

Nesse primeiro caso simulado, utiliza-se a distribuição Exponencial Dupla com esperança 0 e parâmetro de escala 2, idêntica à que foi utilizada nas simulações por função-núcleo. O erro de aproximação nesse caso é de  $0,8474 \times 10^{-3}$ .

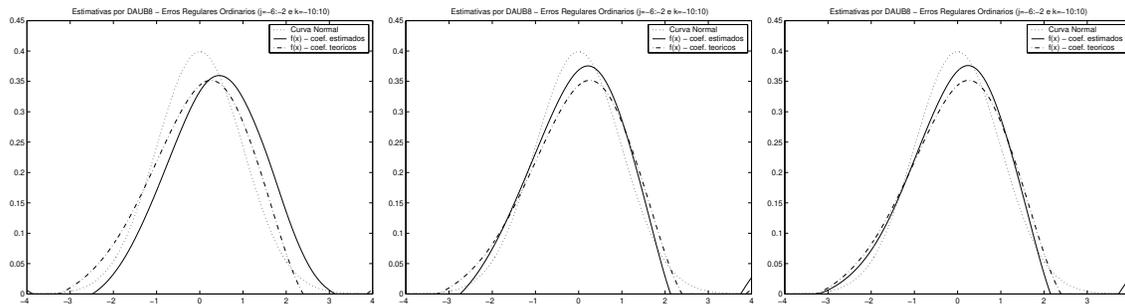


Figura 4.12: Estimativas de Densidade por *DAUB8*,  $n=128, 1024$  e  $4096$

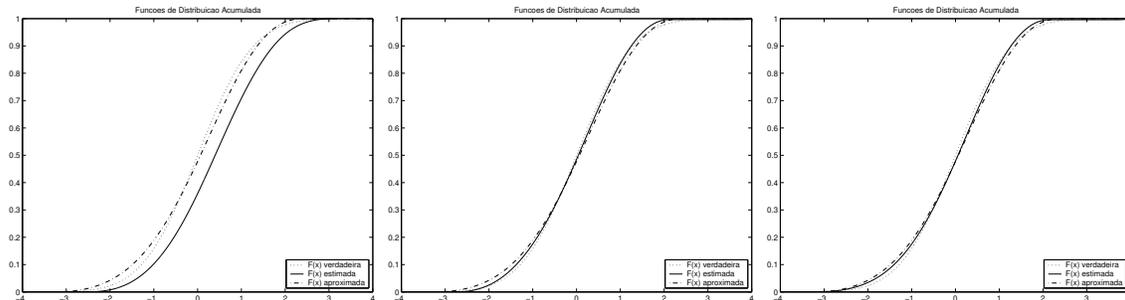


Figura 4.13: Funções de Distribuição Acumuladas por *DAUB8*,  $n=128, 1024$  e  $4096$

Tabela 4.5: Erros por *DAUB8* sem Limiar - Regulares Ordinários

n	EQM ( $\times 10^{-3}$ )		p-valor médio	% de Rejeição	
	$(f_X - \hat{f}_{j_1, j_0})$	$(\tilde{f}_{j_1, j_0} - \hat{f}_{j_1, j_0})$		$\alpha = 0, 05$	$\alpha = 0, 01$
128	7,2064 (6,9257)	6,0079 (6,1199)	0,0647	80,3%	71,5%
1024	1,9426 (1,3832)	1,0662 (1,1433)	0,0122	94,4%	88,8%
4096	1,1281 (0,4491)	0,2758 (0,3397)	$3,9 \times 10^{-5}$	100%	99,8%

Tabela 4.6: Erros por *DAUB8* com Limiar - Regulares Ordinários

n	EQM ( $\times 10^{-3}$ )		p-valor médio	% de Rejeição	
	$(f_X - \hat{f}_{j_1, j_0})$	$(\tilde{f}_{j_1, j_0} - \hat{f}_{j_1, j_0})$		$\alpha = 0, 05$	$\alpha = 0, 01$
128	7,2715 (7,0980)	6,1396 (6,2765)	0,0597	79,8%	68,9%
1024	1,8287 (1,3783)	1,2071 (1,2257)	0,0159	92,1%	86,0%
4096	0,9706 (0,4452)	0,4281 (0,4745)	0,0002	100,0%	99,5%

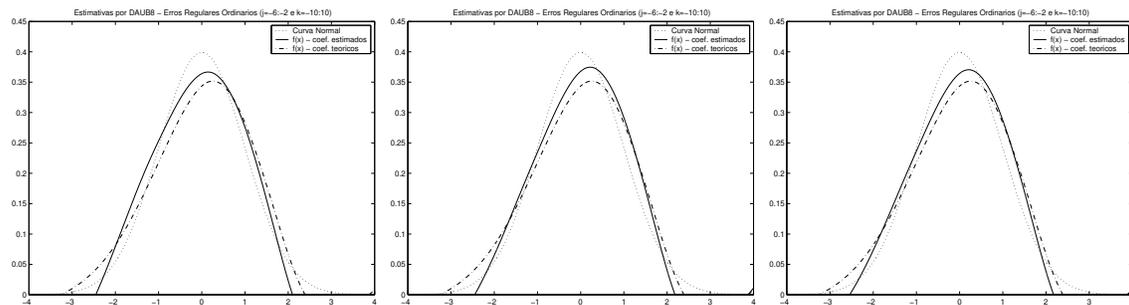


Figura 4.14: Estimativas de Densidade por *DAUB8* com Limiar

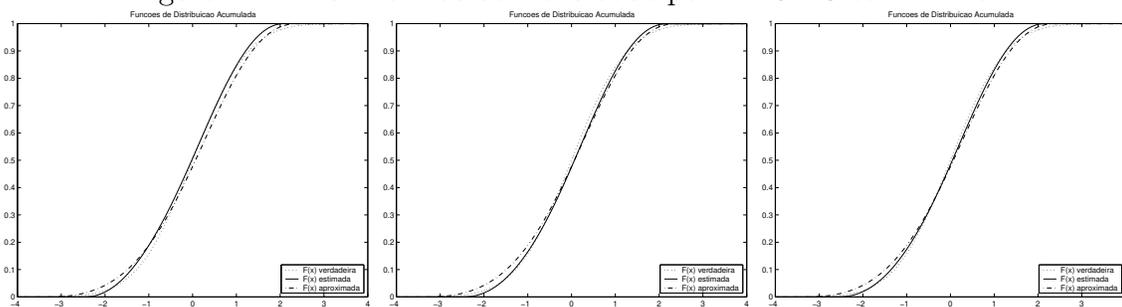


Figura 4.15: Funções de Distribuição Acumuladas por *DAUB8* com Limiar

### 4.2.1.2 Erros Super-Regulares

Para a classe dos erros super-regulares, utilizam-se erros que seguem uma distribuição normal padrão. Nesse caso, tem-se um grande aumento no erro de aproximação, que é de  $8,2728 \times 10^{-3}$ .

Tabela 4.7: Erros por *DAUB8* sem Limiar - Super-Regulares

n	EQM ( $\times 10^{-3}$ )		p-valor médio	% de Rejeição	
	$(f_X - \hat{f}_{j_1, j_0})$	$(\tilde{f}_{j_1, j_0} - \hat{f}_{j_1, j_0})$		$\alpha = 0,05$	$\alpha = 0,01$
128	8,3623 (0,2571)	0,0512 (0,0624)	$0,1372 \times 10^{-3}$	100%	99,9%
1024	8,2908 (0,0867)	0,0069 (0,0094)	0,0000	100%	100%
4096	8,2748 (0,0408)	0,0017 (0,0025)	0,0000	100%	100%

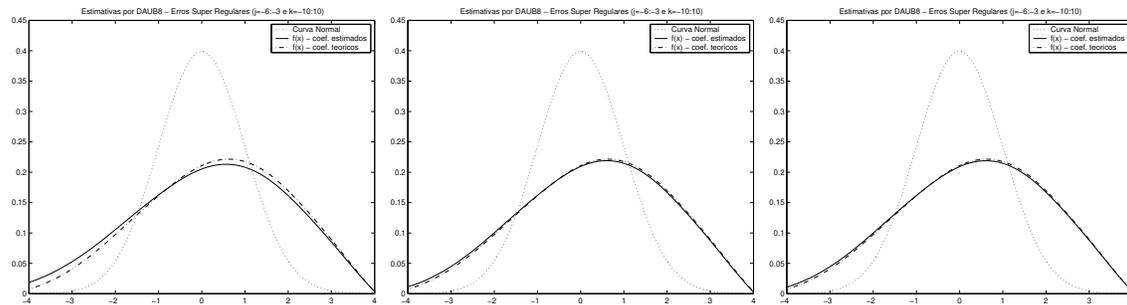


Figura 4.16: Estimativas de Densidade por *DAUB8*,  $n=128, 1024$  e  $4096$

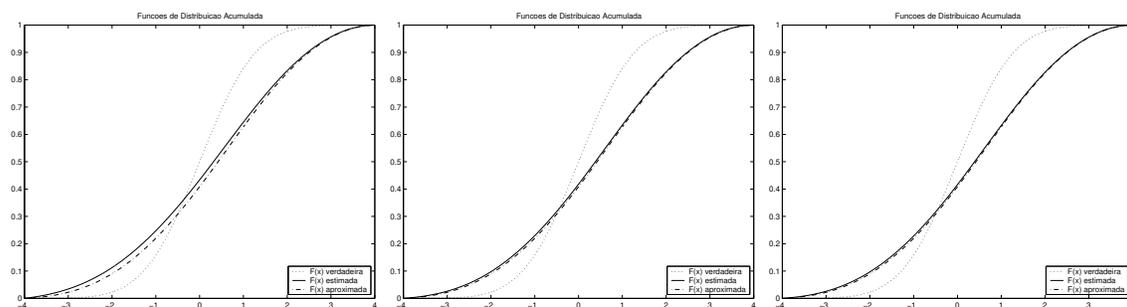


Figura 4.17: Funções de Distribuição Acumuladas por *DAUB8*,  $n=128, 1024$  e  $4096$

Tabela 4.8: Erros por *DAUB8* com Limiar - Super-Regulares

n	EQM ( $\times 10^{-3}$ )		p-valor médico	% de Rejeição	
	$(f_X - \hat{f}_{j_1, j_0})$	$(\tilde{f}_{j_1, j_0} - \hat{f}_{j_1, j_0})$		$\alpha = 0,05$	$\alpha = 0,01$
128	6,6647 (0,7243)	2,1409 (0,3610)	$0,7887 \times 10^{-4}$	100,0%	99,9%
1024	7,0087 (0,0983)	0,4302 (0,0759)	0,0000	100,0%	100,0%
4096	6,7777 (0,1447)	2,2155 (0,0656)	0,0000	100,0%	100,0%

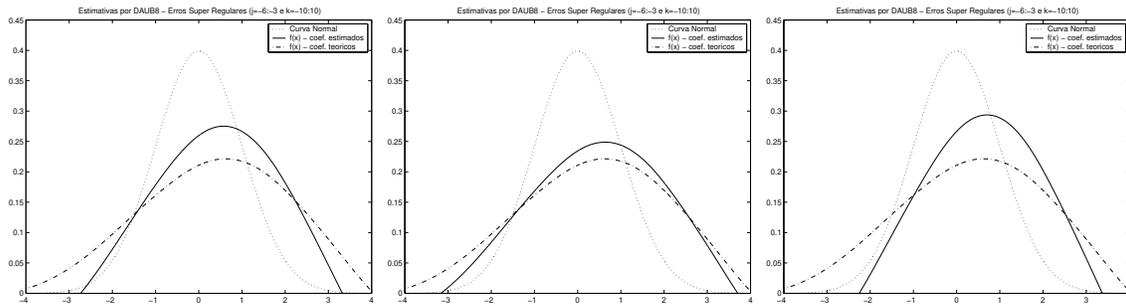


Figura 4.18: Estimativas de Densidade por *DAUB8* com Limiar

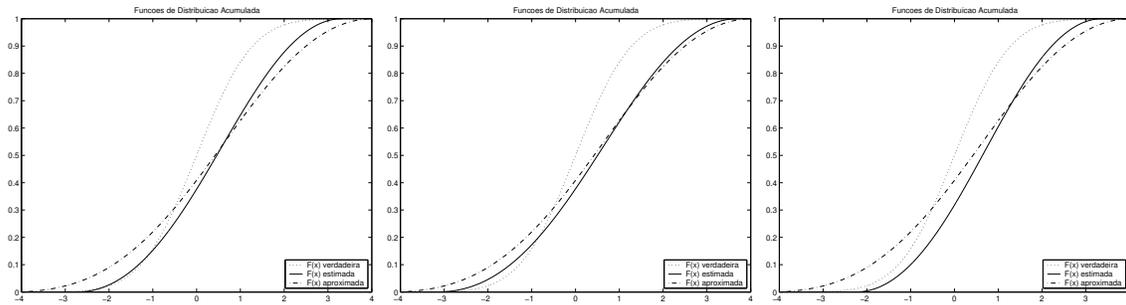


Figura 4.19: Funções de Distribuição Acumuladas por *DAUB8* com Limiar

## 4.2.2 Ondaletas de Meyer

### 4.2.2.1 Erros Regulares Ordinários

Assim como no caso das Daubechies, inicialmente utiliza-se uma distribuição para os erros pertencente à classe regular ordinária, ou seja,  $\varepsilon$  segue uma distribuição Exponencial Dupla. Aqui, o erro de aproximação com valor de  $2,345 \times 10^{-4}$  é bem menor se comparado com os da base de Daubechies.

Tabela 4.9: Erros por Meyer - Regulares Ordinários

n	EQM ( $\times 10^{-3}$ )		p-valor médio	% de Rejeição	
	$(f_X - \hat{f}_{j_1, j_0})$	$(\tilde{f}_{j_1, j_0} - \hat{f}_{j_1, j_0})$		$\alpha = 0,05$	$\alpha = 0,01$
128	5,5542 (6,3680)	5,2310 (5,7093)	0,1002	66,0%	52,2%
1024	1,0343 (1,1183)	0,8273 (0,9095)	0,0669	76,0%	64,7%
4096	0,4297 (0,3259)	0,2044 (0,2153)	0,0230	87,2%	77,0%

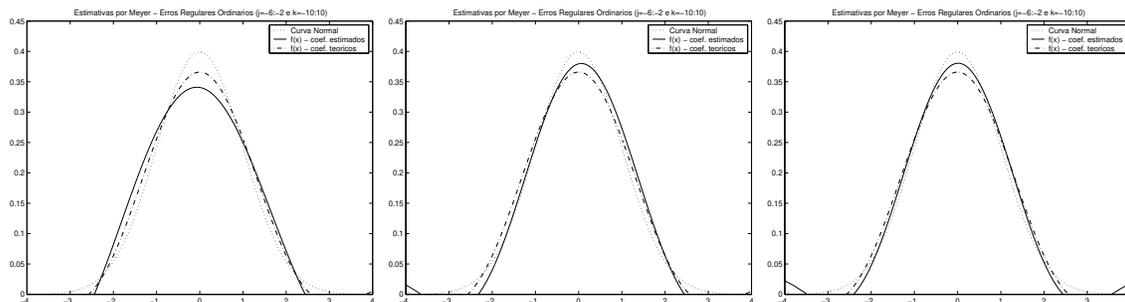


Figura 4.20: Estimativas de Densidade por Meyer,  $n=128, 1024$  e  $4096$

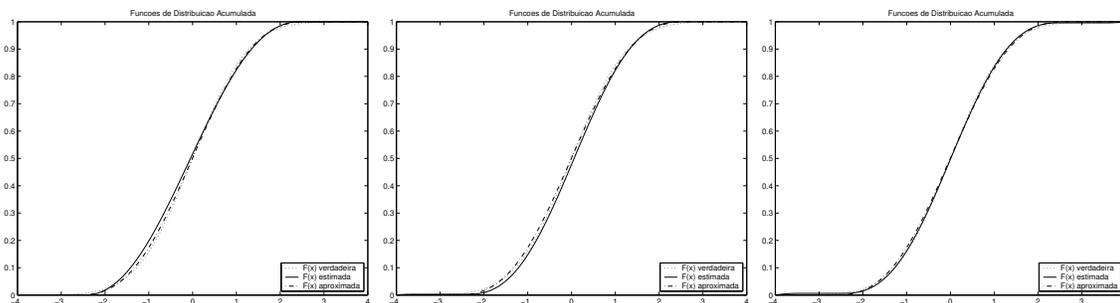


Figura 4.21: Funções de Distribuição Acumuladas por Meyer,  $n=128, 1024$  e  $4096$

### 4.2.2.2 Erros Super-Regulares

Para a classe dos erros super-regulares, utilizam-se erros que seguem uma distribuição normal padrão. O erro de aproximação é também de  $2,345 \times 10^{-4}$ .

Tabela 4.10: Erros por Meyer - Super-Regulares

n	EQM ( $\times 10^{-3}$ )		p-valor médio	% de Rejeição	
	$(f_X - \hat{f}_{j_1, j_0})$	$(\tilde{f}_{j_1, j_0} - \hat{f}_{j_1, j_0})$		$\alpha = 0,05$	$\alpha = 0,01$
128	0,9097 (0,9405)	0,7050 (0,7502)	0,2866	30,1%	15,3%
1024	0,3606 (0,2101)	0,1118 (0,1204)	0,1440	54,0%	33,8%
4096	0,2495 (0,0754)	0,0244 (0,0257)	0,0388	79,1%	58,3%

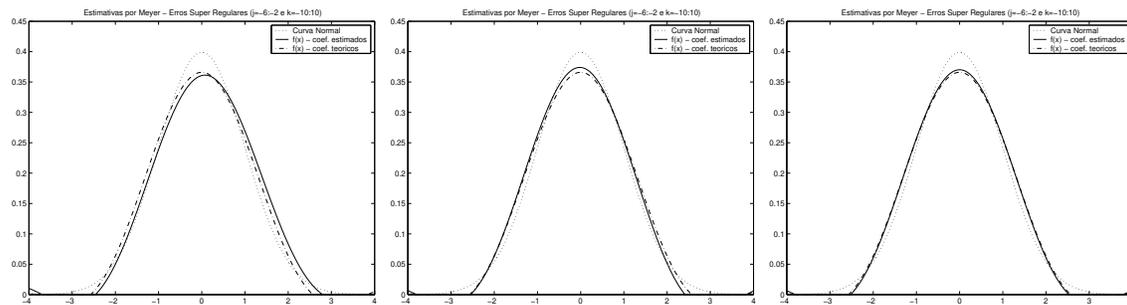


Figura 4.22: Estimativas de Densidade por Meyer,  $n=128, 1024$  e  $4096$

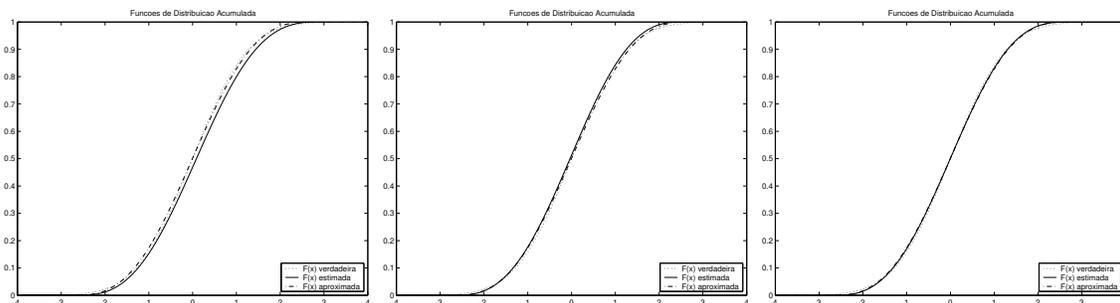


Figura 4.23: Funções de Distribuição Acumuladas por Meyer,  $n=128, 1024$  e  $4096$

Fan e Koo (2002), considerando a taxa de convergência, sugerem que ondaletas de Daubechies sejam utilizadas para erros regulares ordinários, enquanto ondaletas de Meyer são recomendadas para erros super-regulares. Através das simulações anteriores, é possível comprovar essa relação. No entanto, as ondaletas de Meyer,

embora obtenham desempenho superior para erros super-regulares, têm seus resultados para erros regulares ordinários ainda mais satisfatórios do que as bases de Daubechies. Essa diferença pode ter ocorrido devido ao fato de que as Transformadas de Fourier das ondaletas de Meyer possuem forma explícita, o que favorece essas bases na hora dos cálculos, visto que as Transformadas de Fourier para as bases de Daubechies são obtidas por aproximações.

Como o desempenho das ondaletas de Meyer é satisfatório, apenas essas bases serão utilizadas no restante do texto.

### 4.2.3 Séries Simuladas - GARCH(1,1) e AR(1)-SV

Os estimadores definidos na Seção 3.4 por (3.35) e (3.37) são aplicados na série real de cotação diária da taxa de câmbio Iene/Dólar, apresentada na Seção 1.2. A série contém 3502 observações, correspondentes ao período de 02 de janeiro de 1990 a 05 de dezembro de 2003. Para verificar as propriedades dos estimadores, geramos séries de tamanhos variados a partir de dois modelos bem conhecidos na literatura, o AR(1)-SV e o GARCH(1,1). Os parâmetros utilizados na geração das séries são os parâmetros estimados para cada modelo aplicado nos dados reais, que serão vistos no próximo capítulo. Dessa forma, temos a geração das séries a partir de dois critérios:

- (i) pelo modelo GARCH(1,1), com as estimativas dos parâmetros obtidas por Palaro (2004); e
- (ii) pelas estimativas dos parâmetros de um modelo AR(1)-SV.

Não discutiremos detalhes sobre os modelos GARCH. Excelentes referências são Engle (1982), Bollerslev (1986), Bollerslev, Engle e Nelson (1994). Os modelos de Volatilidade Estocástica são apresentados na Subseção 1.3.1 e também podem ser vistos em maiores detalhes em Taylor (1994), Kim, Shephard e Chib (1998).

Foram geradas séries de retornos  $\{R_t\}$  de tamanhos 128, 256, 1024 e 8192, desconsiderando-se as primeiras 7500 observações, com 1000 replicações de cada caso, a partir dos modelos AR(1)-SV e GARCH(1,1) a seguir:

- Modelo AR(1)-SV

$$\begin{cases} R_t = \sigma_t \xi_t \\ \sigma_t = \exp(X_t/2) \\ X_t = \alpha_0 + \alpha_1 X_{t-1} + \eta_t, \end{cases} \quad (4.6)$$

em que  $\{\eta_t\}$  e  $\{\xi_t\}$  são independentes com distribuição  $N(0, \sigma_\eta^2)$  e  $N(0, 1)$ , respectivamente. A log-volatilidade,  $X_t$ , segue uma distribuição  $N(\alpha_0/(1 - \alpha_1), \sigma_\eta^2/(1 - \alpha_1^2))$ .

As estimativas dos parâmetros foram obtidas para a série de retornos verdadeira através do aplicativo SVPack, desenvolvido para a linguagem Ox, que faz os cálculos apresentados em Kim, Shephard e Chib (1998).

Os valores encontrados foram  $\hat{\alpha}_0 = -0,7803$ ,  $\hat{\alpha}_1 = 0,92336$  e  $\hat{\sigma}_\eta = 0,26169$ .

- Modelo GARCH(1,1)

$$\begin{cases} R_t = \mu + \epsilon_t \\ \epsilon_t = \sigma_t \xi_t \\ \sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \sigma_{t-1}^2, \end{cases} \quad (4.7)$$

em que  $\xi_t \sim N(0, 1)$ . A variância não condicional de  $\epsilon_t$  é dada por  $Var(\epsilon_t) = \alpha_0/(1 - \alpha_1 - \alpha_2)$  e será utilizada como valor inicial para  $\sigma_0^2$ .

As estimativas dos parâmetros desse modelo foram obtidas por Palaro (2004) por meio de uma rotina própria criada no aplicativo Matlab.

Os valores encontrados foram  $\hat{\mu} = -5,40 \times 10^{-5}$ ,  $\hat{\alpha}_0 = 7,65 \times 10^{-7}$ ,  $\hat{\alpha}_1 = 0,0383$  e  $\hat{\alpha}_2 = 0,9459$ .

As Figuras 4.24 e 4.25 apresentam uma série de retornos de tamanho 1024, obtida de cada modelo e suas respectivas volatilidades.

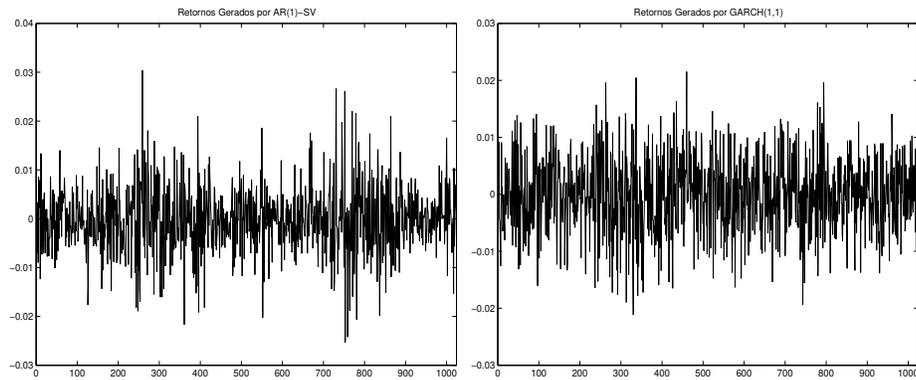


Figura 4.24: Retornos gerados por AR(1)-SV e GARCH(1,1)

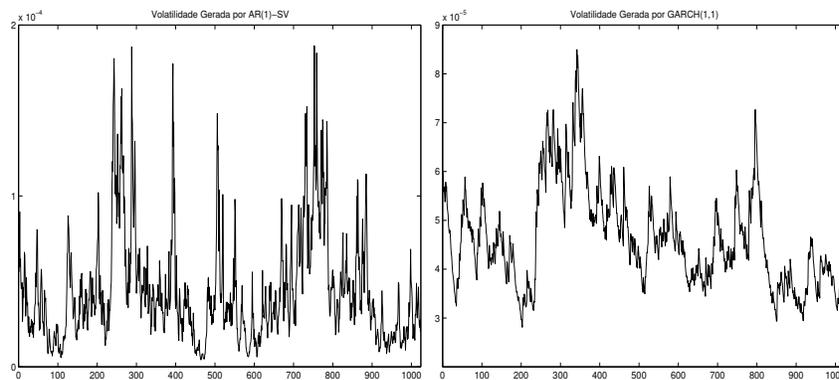


Figura 4.25: Volatilidade gerada por AR(1)-SV e GARCH(1,1)

Mesmo quando a série de retornos é gerada a partir de um modelo GARCH, as estimativas são obtidas considerando-se que seu comportamento da série possa ser explicado por um modelo de Volatilidade Estocástica, para o qual foi desenvolvido o estimador por desconvolução. Considerando-se esse fato, as estimativas obtidas para os dois modelos serão comparadas com a densidade  $N(-10, 181; 0, 465)$ , que é a verdadeira densidade da log-volatilidade.

As estimativas de densidade da log-volatilidade e da volatilidade para a série gerada por um AR(1)-SV são apresentadas, respectivamente, pelas Figuras 4.26 e 4.27 para diferentes tamanhos amostrais. As Figuras 4.28 e 4.29 apresentam as características análogas para a série gerada por um GARCH(1,1).

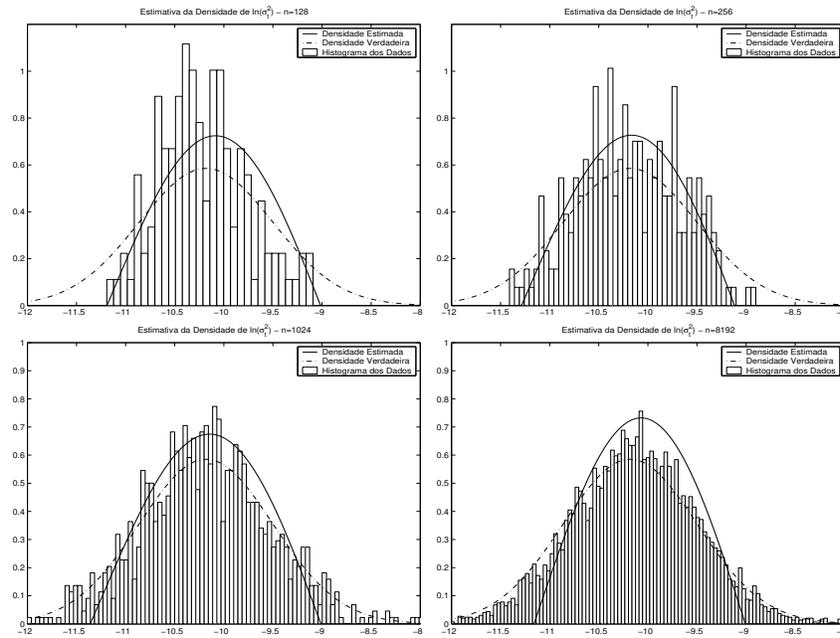


Figura 4.26: Estimativa da Densidade da Log-Volatilidade - AR(1)-SV

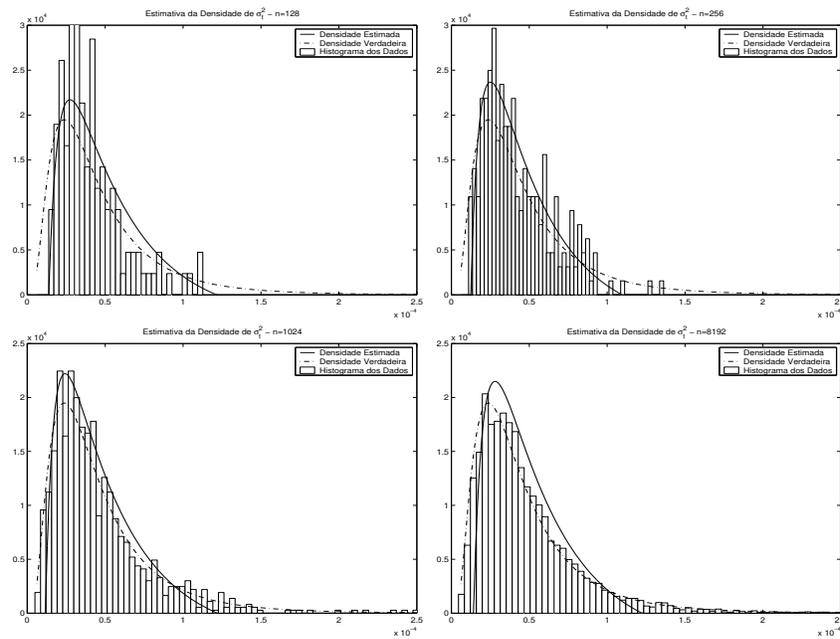


Figura 4.27: Estimativa da Densidade da Volatilidade - AR(1)-SV

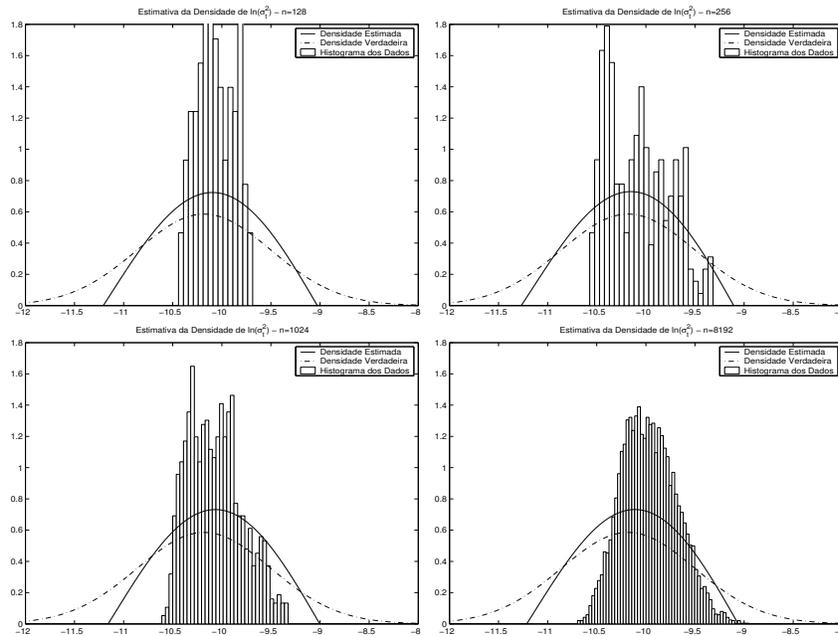


Figura 4.28: Estimativa da Densidade da Log-Volatilidade - GARCH(1,1)

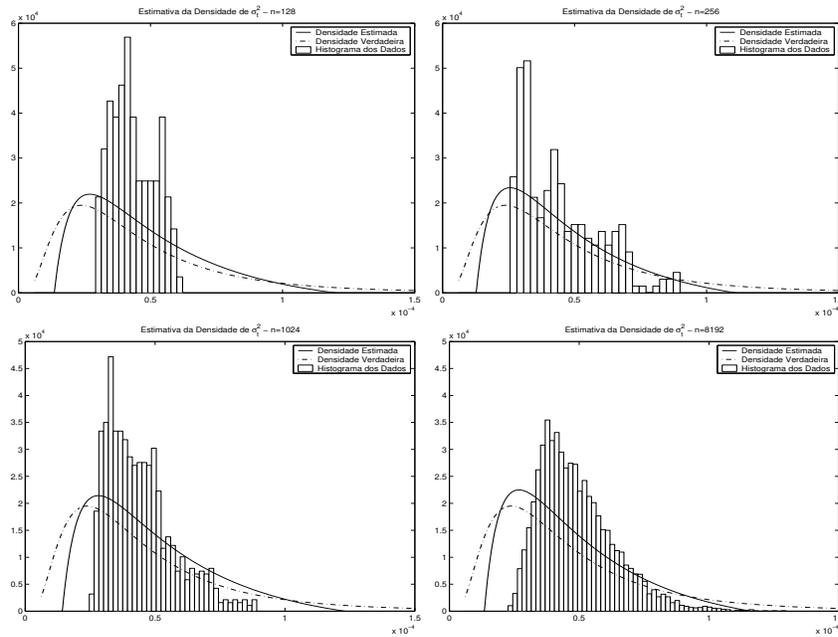


Figura 4.29: Estimativa da Densidade da Volatilidade - GARCH(1,1)

A Tabela 4.11 apresenta uma medida de erro para comparar a verdadeira densidade com as estimativas obtidas para as séries geradas pelos modelos AR(1)-SV e GARCH(1,1).

Tabela 4.11: Desempenho dos Estimadores

n	$\left[ \mathbb{E}_f \  \hat{f}_{j_0, j_1} - f \ _2^2 \right]^{1/2}$	
	AR(1)-SV	GARCH(1,1)
128	0,5705	0,2963
256	0,4632	0,2603
1024	0,2888	0,2116
8192	0,2133	0,2012

A Figura 4.30 mostra graficamente os valores da Tabela 4.11 e também a curva  $C(\log(n))^{-\sigma'/s}$ , vista no Teorema 3.4.4, para que possamos comparar as estimativas obtidas com a propriedade teórica dada por esse teorema.

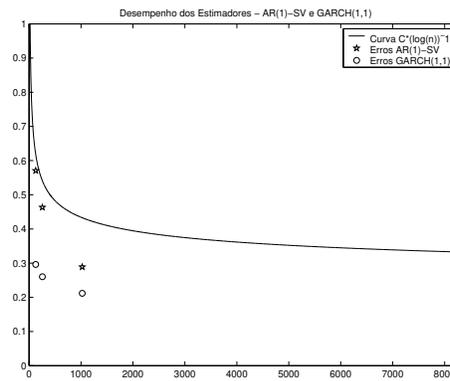


Figura 4.30: Desempenho do Estimadores

Uma das características das estimativas é a de que os pontos de maior densidade na curva normal verdadeira da log-volatilidade são sempre superestimados enquanto que as caudas são perdidas.

## 5 Aplicação em Dados Reais

Como já dito anteriormente, a técnica de desconvolução via ondaletas será aplicada na série de retornos das taxas diárias de câmbio Iene/Dólar Americano. Essas taxas são de fechamento diário de venda e correspondem ao câmbio comercial oficial do FED (Banco Central Norte-Americano). A série de retornos contém 3502 observações, correspondentes ao período de 02 de janeiro de 1990 a 05 de dezembro de 2003. A série original pode ser vista na Figura 1.1 e os retornos são mostrados novamente na Figura 5.31.

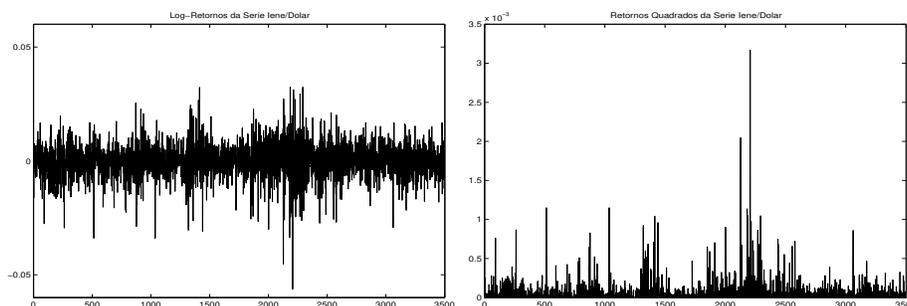


Figura 5.31: Gráfico dos Log-Retornos e Retornos Quadrados da Série Iene/Dólar

A metodologia apresentada foi desenvolvida para estimar a densidade da volatilidade em modelos de volatilidade estocástica. Sendo assim, a suposição de que a série de retornos pode ser modelada por uma das formas desse modelo é feita aqui, especificamente o caso em que a log-volatilidade segue um processo autoregressivo de ordem 1.

Palaro (2004) ajusta o modelo GARCH(1,1) para essa série. O modelo AR(1)-SV também foi ajustado com o uso do software SVPack, desenvolvido para a linguagem Ox. As estimativas dos parâmetros para ambos os modelos estão a seguir.

Para o modelo AR(1)-SV visto em (4.6), as estimativas são:

$$\begin{aligned}\hat{\alpha}_0 &= -0,7803; \\ \hat{\alpha}_1 &= 0,92336; \text{ e} \\ \hat{\sigma}_\eta &= 0,26169.\end{aligned}$$

Para o modelo GARCH(1,1) visto em (4.7), as estimativas são:

$$\begin{aligned}\hat{\mu} &= -5,40 \times 10^{-5}; \\ \hat{\alpha}_0 &= 7,65 \times 10^{-7}; \\ \hat{\alpha}_1 &= 0,0383; \text{ e} \\ \hat{\alpha}_2 &= 0,9459\end{aligned}$$

As estimativas para a densidade da log-volatilidade e da volatilidade da série real foram calculadas inicialmente para a série toda, como apresentadas na Figura 5.32. Depois, tomamos séries de tamanho 512, 1024 e 2048 a partir da série original com uma defasagem de 10 observações de uma série para outra e empregamos o estimador de densidade proposto. Sendo assim: para séries de tamanho 512, têm-se trezentas estimativas (Figura 5.33); para as de tamanho 1024, têm-se 248 estimativas (Figura 5.34); e para as de tamanho 2048, têm-se 146 estimativas (Figura 5.35). Com isso, é possível identificar possíveis mudanças associadas ao tamanho amostral e a mudanças de regime.

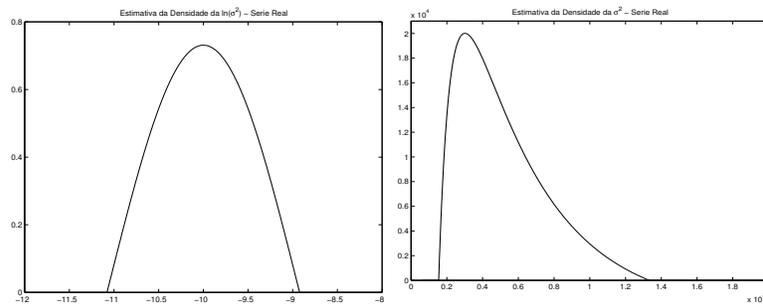


Figura 5.32: Estimativas de Densidade para a Série Completa

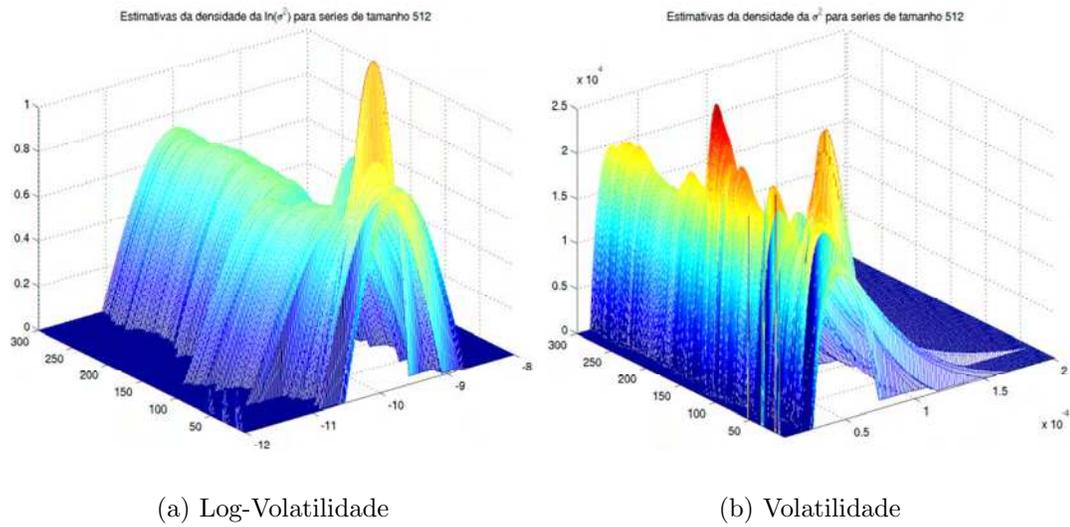


Figura 5.33: Estimativas de Densidade para as séries de tamanho 512

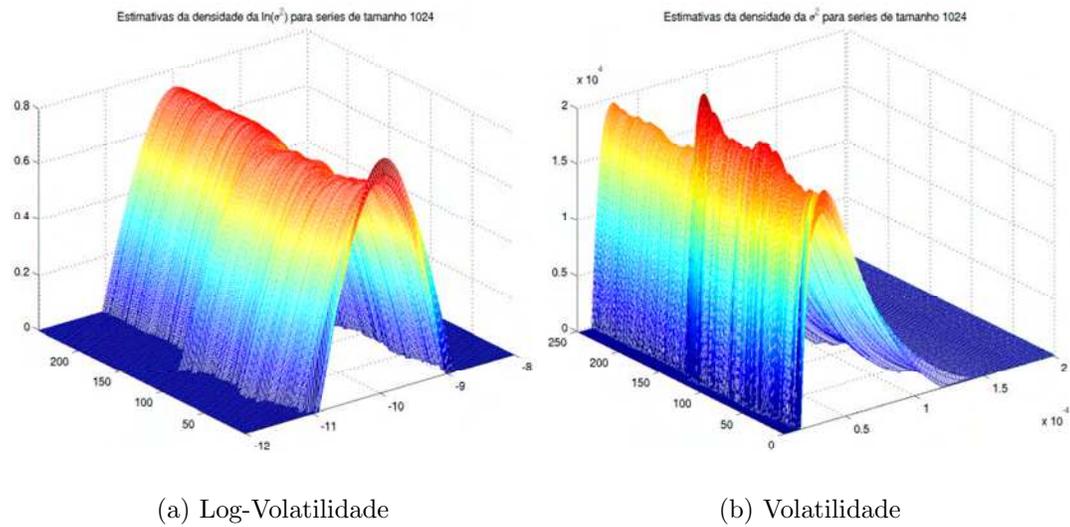


Figura 5.34: Estimativas de Densidade para as séries de tamanho 1024

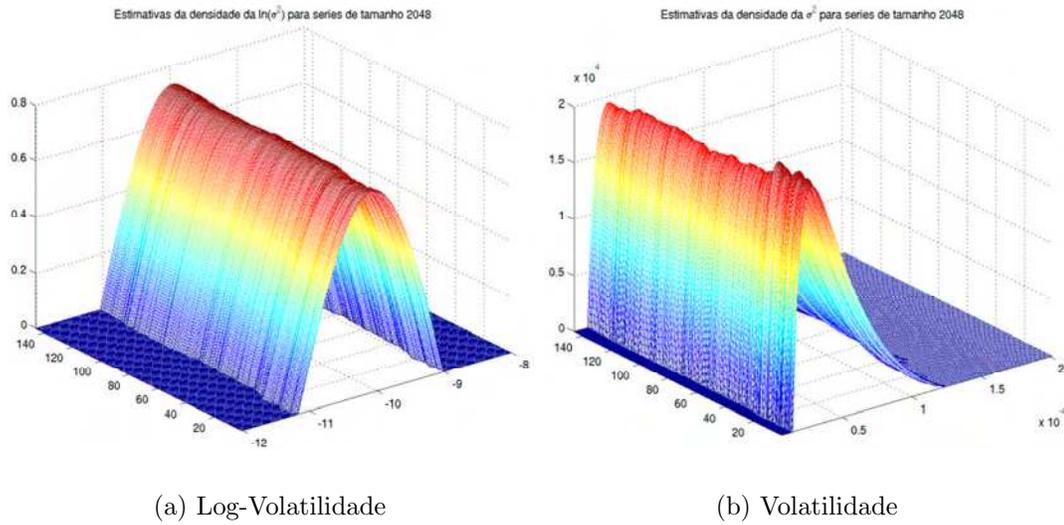


Figura 5.35: Estimativas de Densidade para as séries de tamanho 2048

Duas medidas de distância foram utilizadas: distância  $\mathcal{L}^2$  e distância de Hellinger. Sejam  $f$  e  $g$  duas funções de densidade; então

$$D_{\mathcal{L}^2}(f, g) = \sqrt{\int (f(x) - g(x))^2 dx}; \quad e \quad (5.1)$$

$$D_H(f, g) = \sqrt{\int [(f(x))^{1/2} - (g(x))^{1/2}]^2 dx}. \quad (5.2)$$

As Figuras 5.36, 5.37 e 5.38 mostram a matriz de distâncias entre as estimativas.

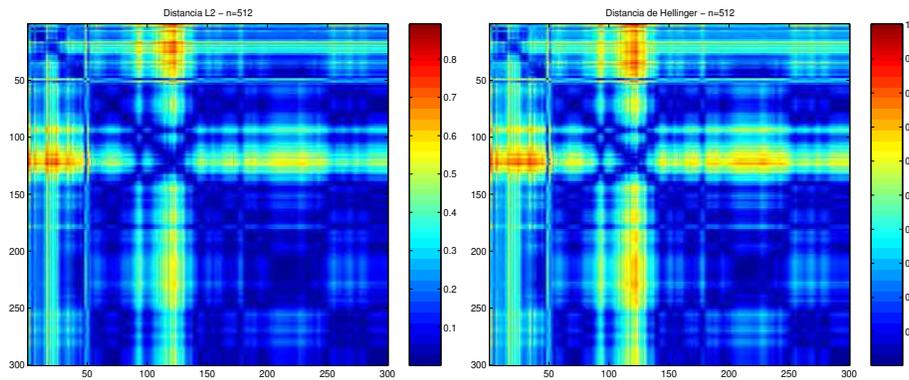


Figura 5.36:  $D_{\mathcal{L}^2}$  e  $D_H$  entre Estimativas de Densidades para  $n=512$

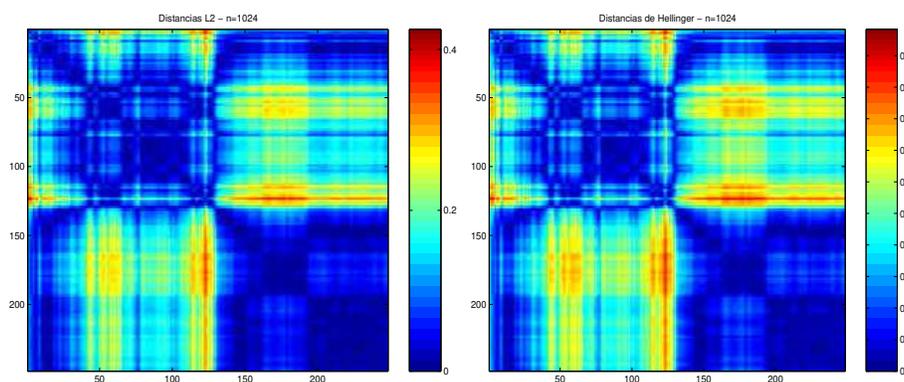


Figura 5.37:  $D_{\mathcal{L}^2}$  e  $D_H$  entre Estimativas de Densidades para  $n=1024$

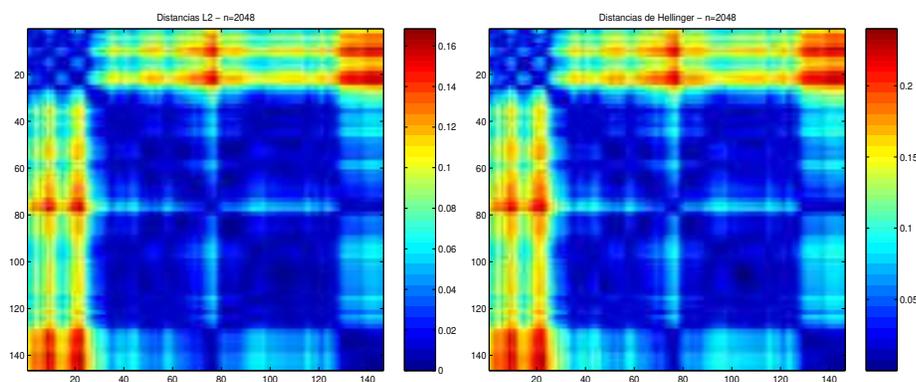


Figura 5.38:  $D_{\mathcal{L}^2}$  e  $D_H$  entre Estimativas de Densidades para  $n=2048$

Note, pelas estimativas de densidade e pelas distâncias, que na parte final da série, mais especificamente na segunda metade, a densidade de volatilidade se mantém estável, independentemente do tamanho da série. Esse fato já não ocorre no início da série, onde vê-se uma variação grande nas estimativas, principalmente no caso de séries com tamanho 512. Para séries de tamanho 2048, as estimativas de densidade são bem próximas da estimativa obtida utilizando a série completa. Isto mostra que um tamanho amostral de 2048 é suficiente para estimar a densidade da volatilidade.

O estimador proposto é consistente para a densidade da volatilidade, pois mesmo quando a série é gerada a partir de um modelo GARCH(1,1), as estimativas, que são obtidas apenas utilizando as informações dos retornos, se aproximam da densidade teórica. Para detectar valores extremos, o estimador proposto não apresenta um desempenho satisfatório. Como visto nas simulações, as estimativas de densidade sempre perdem as informações das caudas.

## Considerações Finais

Os estimadores propostos nesse trabalho podem ser de grande utilidade em Modelos de Volatilidade Estocástica. O desempenho dos estimadores é razoável, mas ainda não é totalmente satisfatório, porque não é possível assumir qualquer distribuição para os erros. As dificuldades de implementação estão na complexidade computacional dos cálculos envolvidos para obter tais estimativas. Como trabalhamos com transformadas de Fourier, é necessário o cálculo de integrais complexas, o que impede a generalização do estimador. Mesmo para os erros com os quais trabalhamos, fez-se necessário o uso de interpolação e discretização das integrais para obtermos as estimativas.

O estimador por desconvolução via função-núcleo consegue reproduzir razoavelmente bem as forma das curvas, principalmente na região central da amostra. Nota-se que nos valores extremos a qualidade das estimativas não é muito satisfatória.

Fan e Koo (2002) indicam o emprego das ondaletas de Meyer no caso de erros super-regulares e bases de Daubechies são empregadas para erros regulares ordinários. No entanto, na prática, as melhores estimativas, para ambos os tipos de erros, foram obtidas via ondaletas de Meyer.

Empregando-se as bases de Daubechies, nota-se uma diferença sensível entre os estimadores lineares e os não-lineares. Para as ondaletas de Meyer, as estimativas lineares e não-lineares são equivalentes e têm desempenho muito superior às ondaletas de Daubechies. Esse resultado empírico pode ter ocorrido devido às aproximações numéricas nos cálculos das estimativas. Com base nos resultados desse trabalho, recomendamos o uso das ondaletas de Meyer para ambas as classes de erro.

A Figura 4.30 mostra que as estimativas da densidade da log-volatilidade para dados gerados a partir do modelo GARCH(1,1) se comportam de acordo com a propriedade dada pelo Teorema 3.4.4, embora para tamanhos de amostra suficientemente grande, os erros das estimativas para séries geradas tanto pelo modelo

GARCH(1,1) quanto pelo AR(1)-SV são bem próximos.

O estimador por desconvolução via ondaletas é consistente para a densidade da volatilidade. Mesmo quando a série é gerada a partir de um modelo GARCH(1,1), as estimativas, que são obtidas apenas utilizando as informações dos retornos, se aproximam da densidade teórica.

Para detectar valores extremos, os estimadores propostos não apresentam um desempenho satisfatório. Como visto nas simulações, as estimativas de densidade via ondaletas sempre perdem as informações das caudas. No caso de função-núcleo, também não temos um bom desempenho nas caudas da função estimada.

Alguns dos problemas apresentados podem ser corrigidos se utilizarmos rotinas mais precisas, que diminuam o erro devido à interpolação. No entanto, para um estudo prospectivo de valores usuais, os estimadores já se mostram satisfatórios, sendo alternativas não-paramétricas viáveis aos modelos usualmente paramétricos da literatura.

# Bibliografia

- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, v. 31, n. 3, p. 307–327, 1986.
- BOLLERSLEV, T.; ENGLE, R. F.; NELSON, D. B. Arch models. In: *Handbook of econometrics, Vol. IV*. Amsterdam: North-Holland, 1994, (Handbooks in Econom., v. 2). p. 2959–3038.
- CHUI, C. K. *An introduction to wavelets*. Boston, MA: Academic Press Inc., 1992. 264 p. (Wavelet Analysis and its Applications, v. 1). ISBN 0-12-174584-8.
- COMTE, F. Kernel deconvolution of stochastic volatility models. *Technical Report*, 2000.
- ENGLE, R. F. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, v. 50, n. 4, p. 987–1007, 1982.
- FAN, J.; KOO, J.-Y. Wavelet deconvolution. *IEEE Trans. Inform. Theory*, v. 48, n. 3, p. 734–747, 2002.
- FAN, J.; TRUONG, Y. K. Nonparametric regression with errors in variables. *Ann. Statist.*, v. 21, n. 4, p. 1900–1925, 1993.
- FAN, J.; YAO, Q. Efficient estimation of conditional variance functions in stochastic regression. *Biometrika*, v. 85, n. 3, p. 645–660, 1998.
- HÄRDLE, W. *Applied nonparametric regression*. Cambridge: Cambridge University Press, 1990. 333 p. (Econometric Society Monographs, v. 19). ISBN 0-521-38248-3.
- HÄRDLE, W. *Smoothing techniques, with implementations in S*. New York: Springer, 1991. 261 p. (Springer series in statistics, v. 19).
- HÄRDLE, W. et al. *Wavelets, approximation, and statistical applications*. New York: Springer-Verlag, 1998. 265 p. (Lecture Notes in Statistics, v. 129). ISBN 0-387-98453-4.

KIM, S.; SHEPHARD, S.; CHIB, S. Stochastic volatility: likelihood inference and comparison with arch models. *Review of Economic Studies*, v. 65, p. 361–93, 1998.

PALARO, H. P. *Aplicação de acoplamento no cálculo do valor em risco*. Dissertação (Mestrado) — Universidade Estadual de Campinas, Departamento de Estatística, 2004.

PRESS, W. H. et al. *Numerical recipes in C*. Second edition. Cambridge: Cambridge University Press, 1992. 994 p. The art of scientific computing. ISBN 0-521-43108-5.

RUGGIERO, M. A. G.; LOPES, V. L. R. *Cálculo Numérico*. 2a. ed. São Paulo: Makron Books, 1996. 405 p. Aspectos Teóricos e Computacionais. ISBN 85-346-0204-2.

STEFANSKI, L.; CARROLL, R. J. Deconvoluting kernel density estimators. *Statistics*, v. 21, n. 2, p. 169–184, 1990.

TAYLOR, S. F. Modelling stochastic volatility: a review and comparative study. *Mathematical Finance*, n. 4, p. 183–204, 1994.

TAYLOR, S. J. Conjectured models for trends in financial prices, tests and forecasts. *J. Roy. Statist. Soc. Ser. A*, v. 143, n. 3, p. 338–362, 1980.

VIDAKOVIC, B. *Statistical modeling by wavelets*. New York: John Wiley & Sons Inc., 1999. 382 p. (Wiley Series in Probability and Statistics: Applied Probability and Statistics). A Wiley-Interscience Publication. ISBN 0-471-29365-2.

WEI, W. W. S. *Time series analysis*. Redwood City, CA: Addison-Wesley Publishing Company Advanced Book Program, 1990. 478 p. Univariate and multivariate methods. ISBN 0-201-15911-2.