

# Modelos de séries temporais com coeficientes variando no tempo

Leandro Teixeira Lopes de Souza

# **Modelos de séries temporais com coeficientes variando no tempo**

Leandro Teixeira Lopes de Souza

Orientador: Profa. Dra. Maria Sílvia de Assis Moura.

Texto apresentado ao Departamento de Estatística da Universidade Federal de São Carlos - DEs/UFSCar para Dissertação de Mestrado, como parte dos requisitos para obtenção do título de Mestre em Estatística.

UFSCar - São Carlos

Agosto/2009

**Ficha catalográfica elaborada pelo DePT da  
Biblioteca Comunitária da UFSCar**

S729ms

Souza, Leandro Teixeira Lopes de.

Modelos de séries temporais com coeficientes variando no tempo / Leandro Teixeira Lopes de Souza. -- São Carlos : UFSCar, 2009.

72 f.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2009.

1. Análise de séries temporais. 2. Modelos tvAR. 3. Modelos tvARCH. I. Título.

CDD: 519.55 (20<sup>a</sup>)

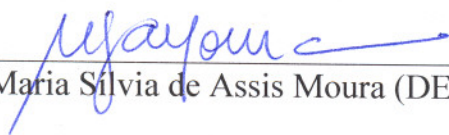
**Leandro Teixeira Lopes de Souza**

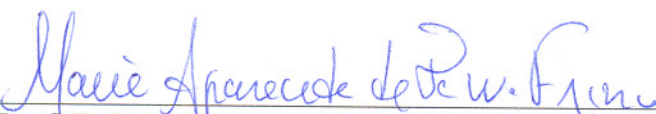
**Modelos de séries temporais com coeficientes variando no tempo**

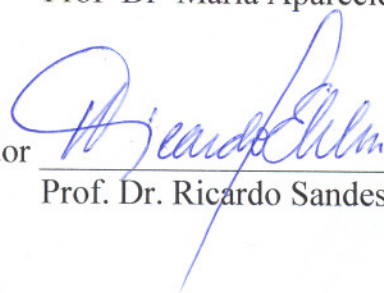
Dissertação apresentada à Universidade Federal de São Carlos, como parte dos requisitos para obtenção do título de Mestre em Estatística.

Aprovada em 26 de fevereiro de 2009.

**BANCA EXAMINADORA**

Presidente   
Prof.<sup>a</sup> Dr.<sup>a</sup> Maria Silvia de Assis Moura (DEs-UFSCar/ Orientadora)

1º Examinador   
Prof.<sup>a</sup> Dr.<sup>a</sup> Maria Aparecida de Paiva Franco (DEs-UFSCar)

2º Examinador   
Prof. Dr. Ricardo Sandes Ehlers (ICMC-USP)

# Agradecimentos

Primeiramente agradeço a Deus, que me dá saúde e força para superar os obstáculos todos os dias.

Aos meus pais e demais familiares, pela formação de caráter em boas maneiras de conduzir minha vida, e por me incentivaram na árdua tarefa de estudar.

A todos os meus amigos que sempre estiveram carinhosamente presentes contribuindo com críticas, sugestões e paciente tolerância.

Aos professores Maria Aparecida de Paiva Franco, Marinho Gomes de Andrade Filho e Ricardo Sandes Ehlers, membros da banca do exame de qualificação e defesa, pelas sugestões feitas.

Aos professores do Departamento de Estatística da Universidade Federal de São Carlos que me abriram as portas e me ofereceram o ambiente acolhedor e sadio para que eu pudesse realizar o meu mestrado.

Um agradecimento especial à minha amiga e orientadora Maria Sílvia de Assis Moura, cujo carinho, incentivo e dedicação transcenderam em muito as exigências do seu papel e da sua responsabilidade.

Finalmente, agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio concedido para este trabalho.

# Sumário

<b>Lista de Figuras</b> . . . . .	<b>v</b>
<b>Lista de Tabelas</b> . . . . .	<b>vii</b>
<b>Resumo</b> . . . . .	<b>ix</b>
<b>Abstract</b> . . . . .	<b>x</b>
<b>1 Introdução</b> . . . . .	<b>1</b>
1.1 Objetivo . . . . .	2
1.2 Organização dos Capítulos . . . . .	2
<b>2 Modelagem financeira</b> . . . . .	<b>4</b>
2.1 Notação . . . . .	4
2.2 Processos estocásticos . . . . .	4
2.3 Conceito de Volatilidade . . . . .	5
2.4 Dados em alta frequência . . . . .	5
2.5 Retornos . . . . .	6
2.6 Fatos estilizados sobre retornos . . . . .	7
<b>3 Modelos Auto Regressivos com coeficientes variando no tempo</b> . . . . .	<b>9</b>
3.1 Estimação . . . . .	9

3.1.1	Estimação de Yule-Walker . . . . .	9
3.1.2	Estimação via <i>wavelet</i> não-linear . . . . .	10
3.1.3	Um estudo de simulação . . . . .	11
3.2	Criação de Bandas de Confiança . . . . .	15
3.2.1	Bootstrap . . . . .	15
3.2.2	Banda de confiança . . . . .	16
3.2.3	Procedimento . . . . .	16
3.3	Métodos de previsão para modelos $tvAR(p)$ . . . . .	18
3.3.1	Método de extrapolação . . . . .	18
3.3.2	Método localmente estacionário . . . . .	19
3.3.3	Estudo de simulação . . . . .	20
3.3.4	Resultados e Discussão . . . . .	21
<b>4</b>	<b>Modelo ARCH com coeficientes variando no tempo . . . . .</b>	<b>22</b>
4.1	Algoritmos <i>Least Mean Square</i> . . . . .	23
4.1.1	Filtros adaptativos . . . . .	23
4.1.2	Algoritmo LMS . . . . .	23
4.1.3	Algoritmo NLMS . . . . .	24
4.2	Um algoritmo recursivo para estimação dos parâmetros do modelo $tvARCH(p)$ . . . . .	25
4.3	Método de previsão para o modelo $tvARCH(p)$ . . . . .	25
4.4	Estudo de Simulação . . . . .	27
4.4.1	Introdução . . . . .	27
4.4.2	Tamanho da Série . . . . .	27
4.4.3	Estudo por simulação do efeito do tamanho $N$ da série . . . . .	27

4.4.4	Resultados e Discussão . . . . .	30
4.4.5	Estudo por simulação do efeito de escolha dos parâmetros $\lambda_1$ e $\lambda_2$ . . . . .	30
4.4.6	Resultados e Discussão . . . . .	34
4.5	Uma proposta para escolha do parâmetro $\lambda$ em 4.13. . . . .	34
4.6	Um estudo de simulação $\lambda(t)$ . . . . .	35
4.6.1	Objetivo . . . . .	35
4.6.2	Especificação da simulação . . . . .	35
4.7	Resultados e Discussão . . . . .	40
<b>5</b>	<b>Aplicação . . . . .</b>	<b>41</b>
5.1	Definição do Problema . . . . .	41
5.1.1	Exemplo 1: Dados diários . . . . .	41
5.1.2	Exemplo 2: Dados de alta frequência . . . . .	44
5.2	Resultados e Discussão . . . . .	46
<b>6</b>	<b>Conclusões e Propostas Futuras . . . . .</b>	<b>48</b>
	<b>Referências Bibliográficas . . . . .</b>	<b>50</b>
<b>A</b>	<b>Wavelets . . . . .</b>	<b>52</b>
A.1	Visão geral . . . . .	52
A.2	Fatos Históricos . . . . .	53
A.3	Introdução . . . . .	54
A.4	Análise de Multirresolução . . . . .	56
A.5	<i>Threshold</i> . . . . .	58
<b>B</b>	<b>Biblioteca R . . . . .</b>	<b>65</b>



B.1	Ajuste de modelos $tvAR(p)$ . . . . .	65
B.2	Ajuste de modelos $tvARCH(p)$ . . . . .	68
B.3	Banda de Confiança para modelos $tvAR(p)$ . . . . .	70

# Lista de Figuras

3.1	<i>Série localmente estacionária simulada com coeficientes dados por 3.7.</i>	12
3.2	<i>Estimativas das funções verdadeiras <math>a_i(\mu)</math>, <math>i = 1, 2</math>, a partir de simulação do primeiro modelo, por mínimos quadrados e por transformação wavelet.</i>	13
3.3	<i>Série localmente estacionária simulada com coeficientes dados por 3.1.3.</i>	14
3.4	<i>Estimativas das funções <math>a_i</math>, <math>i = 1, 2</math>, para o segundo exemplo.</i>	15
3.5	<i>Bandas de Confiança para as funções <math>a_i(t)</math>, <math>i = 1, 2</math></i>	18
3.6	<i>Série localmente estacionária simulada com coeficientes dados por 3.21.</i>	20
4.1	<i>Série simulada de tamanho <math>N = 1000</math> de um modelo <math>tvARCH(1)</math>.</i>	28
4.2	<i>Estimativas dos coeficientes <math>\alpha_j(t)</math>, <math>j = 0, 1</math>, do modelo <math>tvARCH(1)</math> obtidas recursivamente por 4.14 a partir de uma série de tamanho <math>N = 1000</math>.</i>	28
4.3	<i>Série simulada de tamanho <math>N = 5000</math> de um modelo <math>tvARCH(1)</math>.</i>	29
4.4	<i>Estimativas dos coeficientes <math>\alpha_j(t)</math>, <math>j = 0, 1</math>, do modelo <math>tvARCH(1)</math> usando <math>\lambda_1 = 0.01</math> e <math>\lambda_2 = 0.001</math> a partir de uma série de tamanho <math>N = 5000</math>.</i>	29
4.5	<i>Série simulada de acordo com o modelo <math>tvARCH(2)</math></i>	30

4.6	<i>Estimativas dos coeficientes <math>\alpha_i</math>, <math>i = 0, 1, 2</math>, do modelo <math>tvARCH(2)</math> com <math>\lambda_1 = 0.1</math> e <math>w = 0.1</math> . . . . .</i>	31
4.7	<i>Estimativas dos coeficientes <math>\alpha_i</math>, <math>i = 0, 1, 2</math>, do modelo <math>tvARCH(2)</math> com <math>\lambda_1 = 0.1</math> e <math>w = 0.9</math> . . . . .</i>	31
4.8	<i>Estimativas dos coeficientes <math>\alpha_i</math>, <math>i = 0, 1, 2</math>, do modelo <math>tvARCH(2)</math> com <math>\lambda_1 = 0.01</math> e <math>w = 0.1</math> . . . . .</i>	32
4.9	<i>Estimativas dos coeficientes <math>\alpha_i</math>, <math>i = 0, 1, 2</math>, do modelo <math>tvARCH(2)</math> com <math>\lambda_1 = 0.01</math> e <math>w = 0.9</math> . . . . .</i>	32
4.10	<i>Estimativas dos coeficientes <math>\alpha_i</math>, <math>i = 0, 1, 2</math>, do modelo <math>tvARCH(2)</math> com <math>\lambda_1 = 0.001</math> e <math>w = 0.1</math> . . . . .</i>	33
4.11	<i>Estimativas dos coeficientes <math>\alpha_i</math>, <math>i = 0, 1, 2</math>, do modelo <math>tvARCH(2)</math> com <math>\lambda_1 = 0.001</math> e <math>w = 0.9</math> . . . . .</i>	33
4.12	<i>Série simulada de acordo com o modelo <math>tvARCH(1)</math> . . . . .</i>	36
4.13	<i>Gráfico da função <math>\hat{m}c(t)</math> e do parâmetro <math>w = 0.1</math>. . . . .</i>	36
4.14	<i>Estimativas do modelo <math>tvARCH(1)</math> utilizando 4.14 com <math>\lambda = 0.01</math> e <math>w = 0.1</math>. . . . .</i>	37
4.15	<i>Estimativas do modelo <math>tvARCH(1)</math> utilizando 4.21 com <math>\lambda(t) = \delta_w(\hat{m}c(t))</math>. . . . .</i>	37
4.16	<i>Série simulada de acordo com o modelo <math>tvARCH(1)</math> . . . . .</i>	38
4.17	<i>Gráfico da função <math>\hat{m}c(t)</math> e do parâmetro <math>w = 0.1</math>. . . . .</i>	38
4.18	<i>Estimativas do modelo <math>tvARCH(2)</math> utilizando 4.14 com <math>\lambda = 0.01</math> e <math>w = 0.1</math>. . . . .</i>	39
4.19	<i>Estimativas do modelo <math>tvARCH(2)</math> utilizando 4.21 com <math>\lambda(t) = \delta_w(\hat{m}c(t))</math>. . . . .</i>	39
5.1	<i>Série de fechamento do pregão da companhia Vale do Rio Doce. . . . .</i>	42
5.2	<i>(a) Série de retornos logarítmicos. (b) Série de retornos simples. . . . .</i>	42
5.3	<i>Função de Auto Correlação da série de retornos logarítmicos. . . . .</i>	43

5.4	(a) Estimativas da função $a_1(t)$ e (b) Estimativas da função $a_2(t)$ .	43
5.5	Série cambial entre GBP vs USD. . . . .	44
5.6	Série de retorno para GBP vs USD. . . . .	45
5.7	Função de Auto Correlação (ACF) para série de retorno. . . . .	45
5.8	Estimativa da função $a_1$ do modelo $tvAR(1)$ . . . . .	46
5.9	Função de Auto Correlação (ACF) para amostra reduzida. . . . .	46
A.1	Exemplo de uma função wavelet de Daubelets. . . . .	55
A.2	Gráfico da wavelet de Haar. . . . .	56
A.3	A função escala, $\phi$ , associada a wavelet de Haar. . . . .	56
A.4	Transformação da função $f(x) = x$ pelas funções Hard e Soft Threshold com $\lambda = 3$ . . . . .	59
A.5	Série de ressonância magnética nuclear. . . . .	61
A.6	Coefficientes de wavelet para os dados de ressonância magnética nuclear. . . . .	62
A.7	Estimativa do sinal $rmn$ com a aplicação do hard threshold e cálculo do parâmetro $\lambda$ pelo método universal. . . . .	63
A.8	Análise de multirresolução do sinal $rmn$ . . . . .	64

# Lista de Tabelas

3.1	EQMP para os modelos e métodos. . . . .	21
5.1	EQM para os modelos. . . . .	43
5.2	EP para os modelos e métodos. . . . .	44

# Resumo

No presente trabalho são apresentadas extensões dos modelos Auto Regressivo e Auto Regressivo Condicionalmente Heteroscedasticos com coeficientes variando ao longo do tempo. Estes têm sido utilizados como modelos para séries temporais reais não estacionárias, em especial as séries financeiras.

O objetivo desse trabalho é apresentar os modelos e as técnicas envolvidas para estimar esses coeficientes que variam no tempo, além disso, é feita uma introdução a modelagem financeira e algumas sugestões para facilitar a aplicação e utilização dos modelos com coeficientes variando no tempo. Os estudos de simulação e a aplicação em dados reais mostraram que os modelos têm um grande potencial a ser explorados na análise de séries não estacionárias.

As sugestões de banda de confiança e previsão para os modelos Auto Regressivos com coeficientes variando no tempo viabilizam a utilização dos modelos em dados financeiros e outras séries que apresentam uma característica de não estacionariedade.

As modificações no algoritmo de estimação dos modelos ARCH variando no tempo foram para aumentar a taxa de convergência. A criação de um método para previsão dos modelos ARCH necessitam de um estudo mais profundo, porém o algoritmo mostrou resultados promissores no estudo de simulação, dando alguns indícios de que pode ser aplicada na prática.

Por fim, as contribuições na criação de funções para um software livre que facilitam a utilização e a análise dos modelos estudados bem como a utilização dos métodos propostos.

**Palavras-chave:** Modelos tvAR e tvARCH, Processos não estacionários, Séries Temporais.

# Abstract

In this work they are presented extensions of Auto Regressive and Auto Regressive Conditional Heteroscedasticity models with coefficients varying in time. These coefficients have been used as models for non stationary real time series, specially for financial series.

The objective of this work is to present the models and the techniques involved in estimating time-varying coefficients, moreover, it is made an introduction to financial modeling and some suggestions in order to facilitate implementation and use of models with time-varying coefficients. The simulation studies and the application on real data showed that the models have great potential to be exploited in the analysis of non-stationary series.

The suggestions in confidence band and forecasting for the Auto regressive models with time-varying coefficients enable the use of models in financial data and other series that show a non-stationary characteristic.

The modified algorithm for estimation of ARCH models varying in time was to increase the rate of convergence. The creation of a method for forecasting for ARCH models require a deeper study, although the algorithm has shown promising results in simulation study, giving some evidences that it can be applied in real situation.

Finally, the contributions in the creation of functions for a free software that facilitate the use and the analysis of the models studied and the use of the proposed methods.

**Key-Words:** Model tvAR and tvARCH, Non-stationary processes, time series.

# Capítulo 1

## Introdução

A observação da realidade é conseqüência da manifestação de eventos complexos e incertos no decorrer do tempo. Embora estes eventos não sejam exatamente os mesmos, eles também não são completamente diferentes. Existe uma linha de continuidade, similaridade e preeditibilidade nestes eventos que permite estudar a origem, complexidade e generalizar eventos futuros, freqüentemente de modo correto, a partir de experiências passadas.

A possibilidade de prever o futuro dos acontecimentos sempre despertou interesse nas mais diversas áreas do conhecimento humano. No mercado acionário, especificamente, a possibilidade de se realizar algum prognóstico permite a tomada de decisão antecipada de um investidor em relação às expectativas do mercado.

Quando um determinado fenômeno é expresso através de equações precisas, as quais podem a princípio ser resolvidas, é possível prever o comportamento futuro dos eventos pela aplicação de um modelo matemático. Neste sentido, vários procedimentos de previsão de séries temporais foram desenvolvidos com o objetivo de solucionar o problema da previsão. Entretanto, apesar da diversidade de métodos de previsão de séries temporais, nem sempre a solução encontrada pelo emprego destes métodos é satisfatória. Um ponto possível desta deficiência talvez seja a forma com que o modelo descreva o fenômeno, e a mais provável causa seja a ausência de não estacionariedade nos modelos.

Os modelos para séries estacionárias sempre foram o principal interesse no estudo teórico, portanto existe uma classe grande de séries a serem exploradas, as não estacionárias. Entre os modelos estacionários destacam-se os Auto Regressivos (AR) por modelarem uma grande faixa de situações como, por exemplo, dados que mostram um comportamento periódico e também os modelos Auto Regressivos Condicionalmente Heteroscedásticos (ARCH) por modelarem bem a volatilidade de dados financeiros em períodos de tempo relativamente curtos.

Uma suposição bastante comum na análise de séries temporais é a que diz respeito à



estacionariedade fraca da série. Admitindo-se esta suposição, a média da série é constante no tempo. Todavia na prática os dados não apresentam esse comportamento, mostrando algum tipo de não-estacionariedade.

Este trabalho investiga o potencial dos modelos citados acima com coeficientes variando no tempo considerando uma estacionariedade local.

## 1.1 Objetivo

O principal objetivo desse trabalho é apresentar o estudo teórico feito sobre alguns modelos Auto Regressivos não estacionários para séries temporais, que têm ainda a característica de possuírem coeficientes variando no tempo. Em segundo lugar, o objetivo é expor as contribuições pessoais no desenvolvimento do tema, que consistiu em

1. Formulação de um procedimento para construção de uma banda de confiança para os coeficientes do modelo Auto Regressivos com coeficientes variando no tempo,
2. Métodos de previsão,
3. Criação de funções computacionais em software livre para estimação pontual e construção da banda de confiança de tais coeficientes e para previsão,
4. Realização de estudos por simulação de propriedades dos estimadores, e na utilização em aplicação prática em um conjunto de dados reais.

Uma justificativa para o presente trabalho pode ser encontrado em estudos recentes sobre o comportamento não estacionário de dados financeiros feitos por Mikosch et al. (2004).

## 1.2 Organização dos Capítulos

Os capítulos desse trabalho estão apresentados da seguinte maneira: No capítulo 2 apresentamos uma introdução aos dados financeiros e suas características e uma introdução aos estudos de *wavelets* voltado para séries temporais. O capítulo 3 faz uma introdução aos modelos Auto Regressivo com coeficientes variando no tempo e traz algumas sugestões. O capítulo 4 apresenta os modelos ARCH com coeficientes variando no tempo e alguns estudos de simulação.

No capítulo 5, a abordagem dos modelos apresentados anteriormente é aplicada na manipulação de séries de cotações de ações e é feito, uma investigação da habilidade desses modelos em prever valores futuros de ações e no estudo da estrutura dos dados financeiros. Por fim, o último capítulo tem algumas conclusões e propostas futuras.

No apêndice A são apresentadas as principais funções desenvolvidas para a manipulação e análise de séries temporais utilizando os modelos com coeficientes variando no tempo.

# Capítulo 2

## Modelagem financeira

### 2.1 Notação

Nesse tópico será feita a formalização e a notação para dados de séries temporais. Uma série temporal é uma coleção de observações feitas sequencialmente ao longo do tempo. A característica mais importante deste tipo de dados é que as observações vizinhas são dependentes e estamos interessados em analisar e modelar esta dependência. (Ehlers, 2006).

Os objetivos da análise de séries temporais podem ser os mais variados possíveis, como por exemplo, fazer previsões de valores futuros, descrever o comportamento da série ou ainda procurar periodicidade nos dados.

Uma série temporal e a estrutura de seu mecanismo gerador estão relacionadas com o intervalo de ocorrência das observações no tempo. Podemos classificar as séries em dois tipos básicos: contínua se é possível observar a série em qualquer tempo e discreta se as observações são feitas em intervalos de tempo.

Na maioria das séries, as observações são tomadas em uma seqüência de tempos equidistantes. Uma série discreta pode ser representada por  $X_T = \{x_1, x_2, \dots, x_T\}$ , onde observação  $x_t$  está associada ao instante de tempo  $t$ .

### 2.2 Processos estocásticos

De modo geral os modelos utilizados para descrever séries temporais são processos estocásticos, isto é, processos controlados por leis probabilísticas. Um processo estocástico pode ser definido como uma coleção de variáveis aleatórias definidas em um mesmo espaço de probabilidade  $(\Omega, A, P)$ , e indexadas por um parâmetro  $t \in T \subset \mathfrak{R}$  sendo  $T$  um conjunto

qualquer, enumerável ou não. A variável aleatória da coleção que corresponde ao índice  $t$ , será denotada por  $X_t$ , se  $T$  for um conjunto enumerável e por  $X(t)$  se  $T$  for não enumerável. Usualmente  $t$  representa um instante de tempo.

Em análise de séries temporais será possível observar o processo em diversos valores de  $t$ , mas freqüentemente não será possível fazer mais do que uma observação em cada valor de  $t$ . Portanto, tem-se apenas uma realização do processo estocástico e uma única observação da variável aleatória  $X(t)$ .

Para o presente trabalho, de forma geral, uma série diz-se estacionária se oscilar ao redor de uma média constante, com uma variância também constante (Morettin, 2006). A estrutura de correlação depende apenas da defasagem na variável aleatória  $X(t)$ .

## 2.3 Conceito de Volatilidade

O termo volatilidade é utilizado para designar a sensibilidade de uma determinada ação<sup>1</sup> ou carteira de ações às variações dos mercados financeiros nacionais e internacionais. O mesmo termo volatilidade pode ser também utilizado para designar uma medida estatística da variação da cotação da ação durante um determinado período de tempo. A medida estatística mais comum para a volatilidade é a variância.

A volatilidade é uma das mais importantes ferramentas para quem atua no mercado de opções, pois neste mercado estamos interessados na direção do mesmo e também na velocidade com que ele vai se movimentar. Para maiores informações e o uso de volatilidade em cálculo de risco consultar Morettin (2006).

## 2.4 Dados em alta freqüência

Dados financeiros em alta freqüência são observações sobre variáveis financeiras ações, taxas de juros, taxas de câmbio, opções, tomadas diariamente ou em escala intra-diária, freqüentemente irregularmente espaçadas no tempo. Esses dados possuem características únicas, que não aparecem em dados com freqüências mais baixas (semanais, mensais etc). Entre outras destacamos algumas características:

1. Dados não-sincronizados: negociações de ações não aparecem de forma sincronizada; ações diferentes têm freqüências de negócios diferentes e mesmo para uma mesma ação,

---

<sup>1</sup>Ação ou Ações são títulos financeiros que conferem propriedade parcial de uma companhia. Por colocarem seu capital em risco ao entregá-lo à administração da companhia para que esta conduza os negócios, os acionistas têm direito a uma parte da receita, depois de cumpridas as outras obrigações [Matthew Bishop].

- a intensidade de negociação varia de hora para hora do dia. Para mais detalhe ver Wood et al. (1985) e Harris (1986) que estão entre os pioneiros a investigar as propriedades intra-diárias.
2. O número de observações de uma série de dados em alta frequência é usualmente enorme, por exemplo, o número diário de cotações para taxa de câmbio USD/EUR no mercado “spot”<sup>2</sup> é da ordem de 20.000.
  3. Os dados são geralmente registrados com erros e têm que ser corrigidos (“limpos”) antes de serem analisados. Os dados são irregularmente espaçados, com número aleatório de observações por dia.

Existe uma grande evidência de que a incorporação de informações contidas em dados em alta frequência melhora as estimativas da volatilidade dos retornos diários, ao obter uma precisão do processo real seguido pela volatilidade dos ativos financeiros. Destacamos que entre os motivos pelo qual aplicações empíricas de dados em alta frequência são raras ainda na literatura estão o custo de obter, selecionar, armazenar e recuperar esses dados. Entretanto, com o crescente desenvolvimento tecnológico computacional, e principalmente a disponibilidade de dados em alta frequência isso vem tornando-se um problema secundário (Moreira et al., 2004).

Estudo recente realizado por Moreira et al. (2004) é exemplo do uso de dados em alta frequência para modelagem de risco. Eles investigaram seu uso na estimação do valor em risco (VaR) para o índice IBOVESPA.

## 2.5 Retornos

Nessa seção apresentamos algumas considerações sobre a modelagem de retornos de ativos financeiros. A principal motivação para ajustar retornos e não preços dos ativos na maioria dos estudos financeiros é que retornos possuem propriedades estatísticas mais interessantes, destacando-se que retornos são em geral, estacionários e ergódicos (Morettin et al., 2006). Além disso os retornos consistem em uma medida livre de escala da oportunidade de investimento. Denotemos por  $P_t$  o preço de um ativo em  $t$ . Supondo por simplicidade que não são pagos dividendos no intervalo  $(t, t - 1)$ , define-se o retorno simples (ou retorno líquido simples) de manter o ativo no período como:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}. \quad (2.1)$$

---

<sup>2</sup>O termo “spot” é usado nas bolsas de mercadorias para se referir a negócios realizados com pagamento à vista e pronta entrega da mercadoria, em oposição aos mercados a futuro e a termo. A entrega, aqui, não significa entrega física, mas sim a entrega de determinado montante de dinheiro correspondente à quantidade de mercadoria negociada.

O retorno composto continuamente de um período (retorno contínuo ou simplesmente log-retorno de um período) é definido por 2.2:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(1 + R_t) \quad (2.2)$$

Observando que para valores pequenos de  $R_t$ ,  $\log(1 + R_t) \approx R_t$ , via de regra retornos simples e compostos são bastante semelhantes para o mesmo período. Finalmente, o log-retorno de  $k$  períodos é dado pela seguinte expressão:

$$r_t[k] = \log\left(\frac{P_t}{P_{t-k}}\right) = \log(1 + R_t[k]) = \sum_{j=0}^{k-1} \log(1 + R_{t-j}) = \sum_{j=0}^{k-1} r_{t-j} \quad (2.3)$$

Por 2.3, o retorno composto continuamente acumulado em  $k$  períodos é obtido pela soma dos log-retornos de um período. Para uma discussão mais elaborada ver Campbell et al. (1997).

## 2.6 Fatos estilizados sobre retornos

Existe uma vasta literatura sobre fatos estilizados sobre retornos e estes estão documentados em vários estudos empíricos e teóricos. Uma pesquisa bastante ampla está em Count (2001), que mostra entre outros fenômenos:

1. Ausência de autocorrelação: em geral retornos de ativos não apresentam dependência linear, exceto em dados de altas frequências. Cont (2001) sugere intervalos de até cerca de 20 minutos onde os efeitos de microestruturas tendem a causar uma modesta autocorrelação negativa em defasagens pequenas, normalmente um período.
2. Caudas pesadas: a distribuição não condicional dos retornos parece possuir uma distribuição de caudas pesadas (leptocúrtica), ou seja, com maior frequência de ocorrência de valores extremos do que no caso da distribuição normal de mesma média e variância, embora a forma precisa das caudas seja difícil de determinar.
3. Assimetria: a distribuição não condicional é negativamente assimétrica, sugerindo que valores extremos negativos são mais frequentes que os positivos.
4. Agrupamento de volatilidade: frequentemente grandes variações de preço são seguidas por outras grandes variações, isto é, existe uma dependência temporal não linear nos incrementos dos preços, e podemos dizer que a volatilidade dos retornos é serialmente correlacionada. Este fato está ligado ao fato dos retornos serem leptocúrticos (distribuição não condicional com curtose superior a da normal).
5. Efeito de alavanca: a maioria das medidas de volatilidade de um ativo é negativamente correlacionada com os retornos passados do mesmo. Desta forma, em geral a volatilidade é maior em momentos de baixa de seus papéis.

6. Memória longa e persistência da volatilidade: além do agrupamento de volatilidade notada acima, um resultado comum está relacionado à alta persistência da volatilidade condicional, especialmente para amostras grandes, como em períodos de tempo longos com observações diárias ou dados em alta frequência. Estas séries possuem um comportamento que exibe dependência entre observações bastante espaçadas no tempo denominado de memória longa. Definições formais de memória longa podem ser encontradas em Palma (2007).

# Capítulo 3

## Modelos Auto Regressivos com coeficientes variando no tempo

Uma série temporal  $\{X_t\}$ ,  $t = 1, \dots, T$ , segue um modelo Auto Regressivo de coeficientes variando no tempo de ordem  $p$  ( $tvAR(p)$ ) se satisfaz a seguinte equação:

$$X_{t,T} + \sum_{i=1}^p a_i \left( \frac{t}{T} \right) X_{t-i,T} = \sigma \left( \frac{t}{T} \right) \epsilon_t \quad t = 1, \dots, T \quad (3.1)$$

onde os  $\epsilon_t$ ,  $t = 1, \dots, T$ , são independentes e identicamente distribuídos com  $E(\epsilon_t) = 0$  e  $E(\epsilon_t^2) = 1$ . Supõe-se ainda que as primeiras observações  $X_0, \dots, X_{1-p}$  sigam um processo estacionário  $AR(p)$  com parâmetros  $a_1(0), \dots, a_p(0)$ .

O modelo Auto Regressivo de coeficientes variando no tempo é um processo localmente estacionário e modela uma classe maior de séries. Para definição formal sobre processos localmente estacionários ver Dahlhaus (1997).

### 3.1 Estimação

Uma das suposições básicas feitas para estimar um modelo Auto Regressivo é que o processo gerador dos dados seja um processo estacionário. O problema para o caso dos modelos  $tvAR(p)$  consiste em estimar as funções  $a_i(\cdot)$ , apresentadas em 3.1, dado que o modelo  $tvAR$  é localmente estacionário. No primeiro momento será apresentada a estimativa de Yule-Walker.

#### 3.1.1 Estimação de Yule-Walker

Nesse tópico apresentaremos uma maneira de criar uma estrutura tratável para usar a estimação clássica de Yule-Walker. As funções  $a_i(\cdot)$  do modelo 3.1 podem ser estima-



das fixando um ponto  $t_0/T$  no tempo e em uma vizinhança de  $t_0$ . Com isso, o modelo é aproximadamente um modelo estacionário na vizinhança de  $t_0$ , por exemplo, fixando  $t_0/T$  estimaríamos  $a_1(t_0/T), \dots, a_p(t_0/T)$  com a estimativa clássica de Yule-Walker sobre o segmento  $X_{t_0-N,T}, \dots, X_{t_0+N,T}$ , para  $N/T$  pequeno. As funções  $a_i(\cdot)$  são estimadas em pequenos intervalos tais que em cada intervalo o processo é aproximadamente estacionário.

Este método tem a desvantagem de produzir automaticamente uma estimativa suave de  $a_i(\cdot)$ . Mudanças repentinas na função  $a_i(\cdot)$  não são detectadas e com isso este método fica restrito a uma classe de funções suaves. Além disso, o desempenho deste método depende da escolha a priori e apropriada do parâmetro  $N$ . Existe uma alternativa que evita essa escolha  $N$  e que se adapta às características locais de suavidade, usando funções não paramétricas.

### 3.1.2 Estimação via *wavelet* não-linear

Esse método, consiste em fazer uma aproximação *wavelet* das funções  $a_i(\cdot)$ . Este conceito, baseado em expansões ortogonais da série, tem sido explorado no problema de estimação da regressão não paramétrica para detalhes ver apêndice A.

Normalmente não se conhece a função  $a_i(\cdot)$ , que é expandida em série ortogonal com respeito a uma base ortonormal especialmente escolhida de  $L_2[0, 1]$ , uma base *wavelet*, por exemplo, Haar. As funções da base são geradas por dilatação e por translação de  $\phi$  que são conhecidas como função escala *wavelet*, são localizadas na posição espacial, (aqui temporal) e na frequência. Estas funções da base, ao contrário da maioria das transformações, têm a capacidade de compressão otimizada e as funções têm suavidade completamente homogênea sobre todo o domínio. A melhor compressão de um sinal é dado pelo menor número de coeficientes que o representa.

O procedimento todo requer estimadores consistentes para a variação dos coeficientes empíricos. Um estimador consistente é dado, por exemplo, pelo quadrado do resíduo de um ajuste Auto Regressivo local.

### Resultado principal

Em Cohen et al. (1993) mostraram que as funções  $\phi(t)$  e  $\psi(t)$  respectivamente conhecidas como *wavelet* pai e *wavelet* mãe formam um base ortonormal para  $L_2[0, 1]$ , conhecidas como bases CDV. Podemos expandir a função  $a_i(t/T)$  em uma série ortogonal para mais detalhes ver apêndice A,

$$a_i(t/T) = \sum_k \alpha_{0,k}^{(i)} \phi_{0,k}(t/T) + \sum_k \sum_j \beta_{jk}^{(i)} \psi_{jk}(t/T) \quad (3.2)$$

onde

- (i)  $\phi(t/T)$  e  $\psi(t/T)$  são  $C^r[0, 1]$  e têm suporte compacto  
(ii)  $\int \phi(t/T)dt = 1, \int \psi(t/T)t^k dk = 0$  para  $0 \leq k \leq r$ .

Substituindo 3.2 em 3.1 e temos o seguinte resultado

$$X_{t,T} + \sum_{i=1}^p \left[ \sum_k \alpha_{0,k}^{(i)} \phi_{0,k}(t/T) + \sum_k \sum_j \beta_{jk}^{(i)} \psi_{jk}(t/T) \right] X_{t-i,T} = \sigma \left( \frac{t}{T} \right) \epsilon_t \quad (3.3)$$

A soma dos quadrados de  $\sigma(t/T)\epsilon_t$  em 3.3 para  $t = p+1, \dots, T$  é:

$$\sum_{t=p+1}^T \left( X_{t,T} + \sum_{i=1}^p \left[ \sum_k \alpha_{0,k}^{(i)} \phi_{0,k}(t/T) + \sum_k \sum_j \beta_{jk}^{(i)} \psi_{jk}(t/T) \right] X_{t-i,T} \right)^2. \quad (3.4)$$

Minimizando a soma 3.4 obtém-se os coeficientes empíricos  $\tilde{\beta}_{jk}^{(i)}$ , pelo método de mínimos quadrados. Para maiores detalhes ver Dahlhaus et al. (1999).

Os estimadores lineares podem ser ótimos com respeito a taxa de convergência contanto que a suavidade subjacente da função  $a_i(t)$  não seja muito heterogênea. Esta situação muda consideravelmente se a suavidade variar fortemente no domínio. Nesse caso temos um efeito novo que não pode ser explicado por um número pequeno de coeficientes visto que a maioria deles é menor que o nível de ruído.

Portanto aplica-se uma função *threshold* (*hard* ou *soft*) definidas em A.26 e A.27 nos coeficientes para zerar aqueles em que os níveis de ruído são maiores que os coeficientes e para reduzir o número de coeficientes.

$$\hat{\beta}_{jk}^{(i)} = \delta_\lambda^S(\tilde{\beta}_{jk}^{(i)}) \quad (3.5)$$

Finalmente, constroe-se uma estimativa da função  $\hat{a}_i(t)$  aplicando a transformação inversa de *wavelet* nos coeficientes obtidos em 3.5. Observamos que esse método é não linear.

### 3.1.3 Um estudo de simulação

O interesse deste capítulo é apresentar algumas séries localmente estacionárias e analisar o método de estimação *wavelet* para as funções  $a_i(t)$  para diferentes modelos  $tvAR(p)$ .

Foi aplicado o procedimento de estimação *wavelet* descrito na seção acima a duas séries  $tvAR(2)$  simuladas, ambas com comprimento  $T = 1024$ , a partir do modelo

$$X_{t,T} + a_1 \left( \frac{t}{T} \right) X_{t-1,T} + a_2 \left( \frac{t}{T} \right) X_{t-2,T} = \epsilon_t, \quad t = 1, \dots, T, \quad (3.6)$$

onde os  $\epsilon_t$  para  $t = 1, \dots, T$  são observações i.i.d de uma distribuição normal padrão.

Em ambos os exemplos, os parâmetros Auto Regressivos  $a_i(\mu) = a_i(t/T)$ ,  $i = 1, 2$  e  $0 < \mu \leq 1$ , são as funções que mudam ao longo do tempo, isto é, os exemplos simulados são uma realização de um processo não estacionário que segue o modelo descrito em 3.6.

Exemplo 1. Esse exemplo foi retirado de Dahlhaus et al. (1999). O primeiro coeficiente é uma função escada com um salto em  $\mu = 0.6$  e o segundo coeficiente é uma função constante. As funções  $a_1(\mu)$  e  $a_2(\mu)$  são:

$$a_1(\mu) = \begin{cases} -1.69, & \text{se } 0 \leq \mu \leq 0.6 \\ 0.38, & \text{se } 0.6 < \mu \leq 1 \end{cases} \quad a_2(\mu) = 0.81 \quad 0 \leq \mu \leq 1. \quad (3.7)$$

A série simulada a partir do exemplo 1 encontra-se na Figura 3.1.

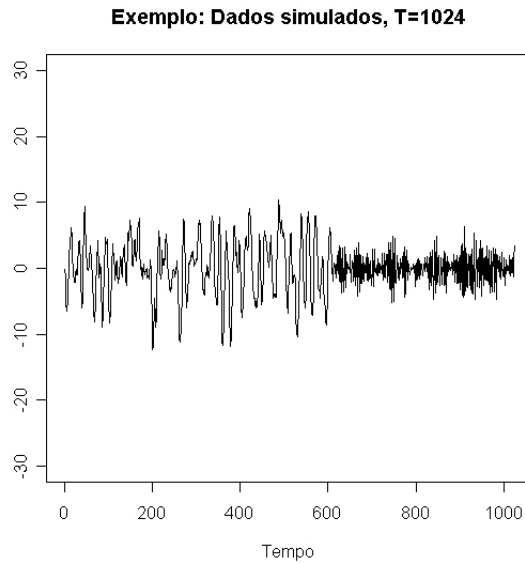


FIGURA 3.1: *Série localmente estacionária simulada com coeficientes dados por 3.7.*

Aplicando o procedimento de estimação usando *wavelets* de Haar, com  $J = 5$ . Usando a transformação rápida *wavelet*, geramos os coeficientes da *wavelet*  $\tilde{\beta}_{jk}^{(i)}$  em escalas  $j = 0, \dots, 4$  e aplicamos a transformação inversa até a escala 10, da amostra original.

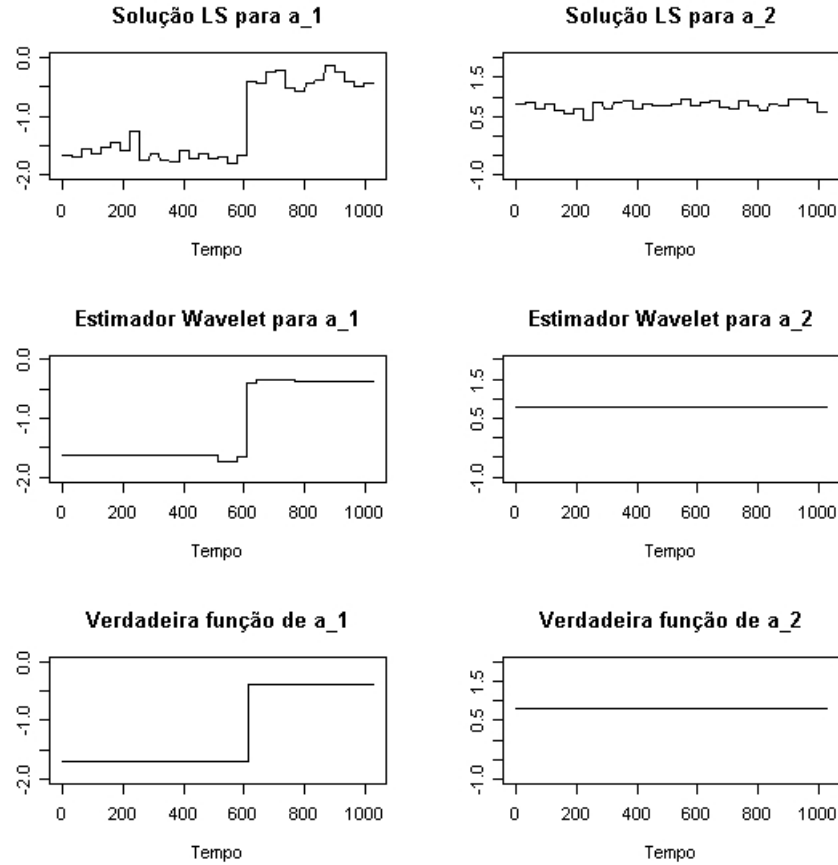


FIGURA 3.2: Estimativas das funções verdadeiras  $a_i(\mu)$ ,  $i = 1, 2$ , a partir de simulação do primeiro modelo, por mínimos quadrados e por transformação wavelet.

Exemplo 2. Esse exemplo é uma pequena modificação do exemplo 1 e pode ser encontrado em Dahlhaus et al. (1999). O segundo coeficiente Auto Regressivo é outra vez uma constante em relação ao tempo, entretanto, o primeiro mostra uma variação no tempo com uma oscilação diferente entre saltos impostos no ponto  $u = 0.25$  e no ponto  $u = 0.75$ . Isto foi possível escolhendo  $a_1(\mu)$  e  $a_2(\mu)$  da seguinte forma,

$$a_1(\mu) = \begin{cases} -1.8\cos(1.5 - \cos(4\pi\mu + \pi)), & \text{se } 0 \leq \mu \leq 0.25 \text{ ou } 0.75 < \mu \leq 1 \\ -1.8\cos(3 - \cos(4\pi\mu + \pi/2)), & \text{se } 0.25 < \mu \leq 0.75 \end{cases} \quad (3.8)$$

$$a_2(\mu) = 0.81 \quad 0 \leq \mu \leq 1. \quad (3.9)$$

Observamos que este processo tem um comportamento não estacionário. A Figura 3.3 mostra uma simulação da série do processo.

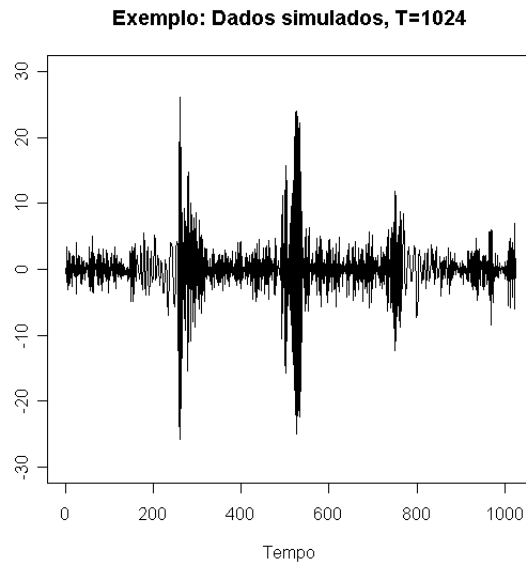


FIGURA 3.3: *Série localmente estacionária simulada com coeficientes dados por 3.1.3.*

Para a estimação dos coeficientes por *wavelet*, foi escolhido base de *wavelet* Daubechies que é periódica com momentos de ordem  $N = 4$ , e com  $J = 6$ .

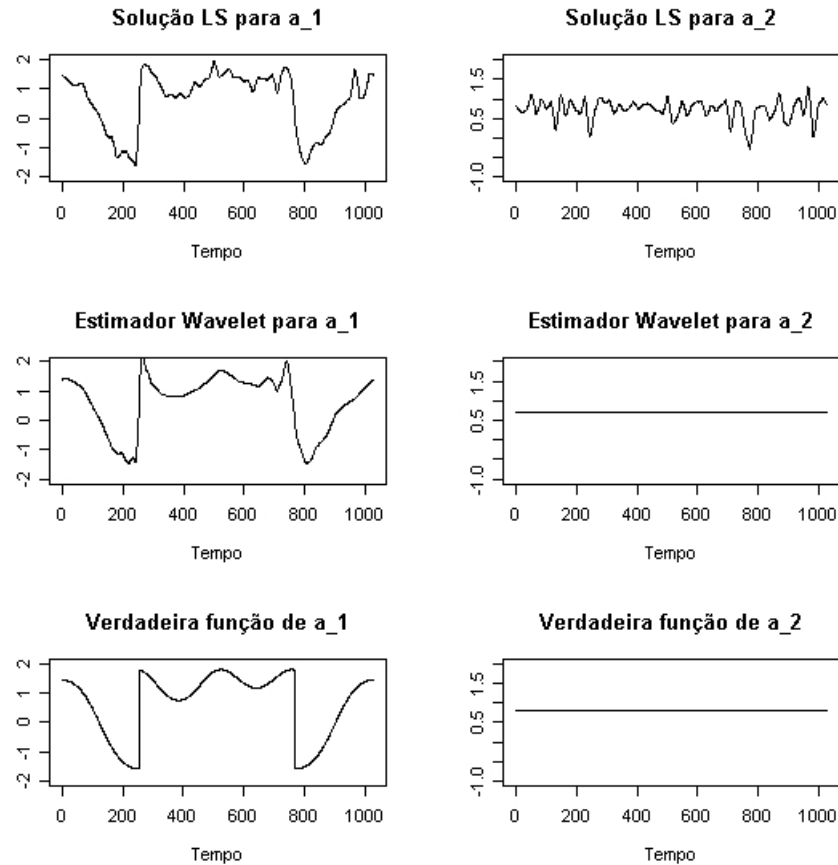


FIGURA 3.4: Estimativas das funções  $a_i$ ,  $i = 1, 2$ , para o segundo exemplo.

## 3.2 Criação de Bandas de Confiança

A construção de uma banda de confiança para coeficientes de um modelo  $tvAR(p)$  através de técnicas de reamostragem será descrita a seguir. As situações consideradas na simulação anterior foram novamente submetidas a estudo, segundo a técnica de reamostragem bootstrap paramétrica. A primeira parte desta seção apresenta detalhes sobre a técnica empregada. A próxima parte fornece as especificações e procedimentos para a construção de uma banda de confiança e, por fim, os resultados são comentados e comparados.

### 3.2.1 Bootstrap

O método bootstrap de estimação tem a vantagem de evitar desenvolvimentos analíticos e tem sido uma das técnicas mais utilizadas e expandidas, com aplicações nas mais diversas áreas. O método foi criado por Bradley Efron na década de 80, existindo duas variantes do mesmo o paramétrico e não-paramétrico. No presente trabalho vamos discutir o bootstrap

paramétrico. Para maiores informações ver Efron et al. (1993) e Davison et al. (1997).

Sejam  $t = (t_1, \dots, t_n)$  realizações das variáveis aleatórias i.i.d.  $T_1, \dots, T_n$ , com distribuição de probabilidade  $F$ , que depende de um parâmetro desconhecido  $\theta$ , dado por  $\theta = s(F)$ , sendo  $s(\cdot)$  a função que define  $\theta$ . O parâmetro  $\theta$  é estimado por  $\hat{\theta} = s(\theta)$ . A idéia é aproximar a distribuição de  $\theta$  pela distribuição de  $\hat{\theta}^* = s(t^*, F)$ , onde  $t^*$  é uma amostra aleatória de tamanho  $n$  retirada de  $F$ . A amostra  $t^* = (t_1^*, t_2^*, \dots, t_n^*)$  é chamada de amostra bootstrap. Essa amostra é gerada através da distribuição suposta  $F$  por meio das estimativas obtidas. Isto é, as estimativas fazem o papel dos verdadeiros parâmetros para a geração da amostra. Repetindo este procedimento um número suficientemente grande de vezes, calcula-se uma distribuição empírica de  $\hat{\theta}^*$  e, a partir desta, obtém-se média, erro-padrão, intervalo de confiança etc.

### 3.2.2 Banda de confiança

Em análise estatística um intervalo de confiança representa a incerteza da estimativa do parâmetro. Uma banda de confiança é utilizada para representar a incerteza da estimativa de uma curva ou função.

Suponha que o objetivo é estimar uma função  $f(x)$ , onde  $x$  tome valores em um conjunto discreto. Se  $x$  é medido com uma dada precisão, é possível construir um intervalo de confiança 95% para  $f(x)$  para cada valor de  $x$ . Um intervalo de confiança para  $f(x)$ , com probabilidade 0,95. Tomados em conjunto, esses intervalos de confiança de 95% constituem uma banda de confiança pontual para  $f(x)$ .

Simbolicamente,  $\hat{f}(x) \pm w(x)$  é uma banda de confiança pontual, com probabilidade de cobertura  $(1 - \alpha)$ , se satisfaz a seguinte condição para cada valor de  $x$ :

$$P\left(\hat{f}(x) - w(x) \leq f(x) \leq \hat{f}(x) + w(x)\right) = 1 - \alpha \quad (3.10)$$

sendo  $\hat{f}(x)$  é a estimativa do ponto  $f(x)$ .

Cada intervalo de confiança para  $x$  tem probabilidade de cobertura 0.95, então a probabilidade de cobertura simultânea é inferior a 0.95.

### 3.2.3 Procedimento

Para criação de uma banda de confiança desenvolvemos um procedimento baseado em uma propriedade dos coeficientes de *wavelet* empíricos  $\tilde{\beta}_{jk}^{(i)}$ , obtidos do ajuste de mínimos quadrados como visto em 3.4. Dahlhaus et al. (1999) mostraram que as estimativas dos coeficientes de *wavelet* são não viciadas e seguem uma distribuição aproximadamente normal, ou seja,

$$\tilde{\beta}_{jk}^{(i)} \sim N(\beta_{jk}^{(i)}, \sigma^2) \quad (3.11)$$

Com essa informação aplicamos o método bootstrap paramétrico para cada coeficiente de *wavelet*, obtendo um intervalo de confiança para cada  $\tilde{\beta}_{jk}^{(i)}$ , da definição em 3.10 temos uma banda de confiança pontual para os coeficientes de *wavelet*, aplicando a transformada inversa de *wavelet*, nos limites inferiores e superiores desses intervalos construímos uma banda de confiança para a função estimada pelos coeficientes  $\tilde{\beta}_{jk}^{(i)}$ .

A proposta de Dahlhaus et al. (1999) é um estimador de *wavelet* não linear obtido através da aplicação de um *threshold* (*Hard* ou *Soft*) nos coeficientes de *wavelet* empíricos. Para criar a banda de confiança para os parâmetros é usado o mesmo *threshold* nos limites inferiores e superiores.

Vamos discutir a escolha do parâmetro  $\lambda$  para a aplicação do *threshold* que é feito através de uma pequena modificação do universal proposta por Dahlhaus et al. (1999).

Suponha que a função  $a_i(t)$  tenha uma expansão em uma série ortogonal como definido em 3.2. Seja  $\xi_T = \{(j, k) | l \leq j, 2^j \leq T^{1/2}, k \in I_j\}$  o conjunto de todos os pares ordenados  $(j, k)$  para os coeficientes  $\tilde{\beta}_{jk}^{(i)}$  e  $\sigma_{ijk}$  a variância do coeficiente de *wavelet* empírico  $\tilde{\beta}_{jk}^{(i)}$ .

Dahlhaus et al. (1999) definiram  $\lambda$  “universal” como sendo

$$\lambda_T^{(i)} = \sigma_T^{(i)} \sqrt{2 \log(\#\xi_T)}, \quad \sigma_T^{(i)} = \max_{(j,k) \in \xi_T} \{\sigma_{ijk}\} \quad (3.12)$$

onde  $\#\xi_T$  é o número de coeficientes de *wavelets*. Observa-se que essa modificação em  $\lambda_T^{(i)}$  faz com que cada coeficiente  $a_i(t)$  tenha seu valor de  $\lambda$  e  $\sigma_T^{(i)}$  é dado pelo maior variância entre todos as estimativas de  $\tilde{\beta}_{jk}^{(i)}$ .

## Algoritmo

- 1 Aplica-se o estimador de mínimos quadrados visto em 3.4, reserva as estimativas de  $\tilde{\beta}_{jk}^{(i)}$  que são os coeficientes empíricos.
- 2 Assumindo que essas estimativas têm uma distribuição como vista 3.11 aplica a metodologia bootstrap e guarda os coeficientes referentes aos quantis desejados da distribuição que foi aproximada.
- 3 Utiliza-se um *threshold* e calcula o parâmetro  $\lambda$  como visto em 3.12, nos coeficientes obtidos em 2.
- 4 Finalmente aplica a transformada inversa de *wavelets* nos coeficientes obtidos em 3, criando uma banda de confiança.

Aplicando esse algoritmo para o primeiro exemplo mostrado na Figura 3.5 temos o seguinte resultado:

Analisando o comportamento das bandas de confiança poderíamos verificar a qualidade do ajuste e selecionarmos o melhor modelo, através de uma inferência em cada uma



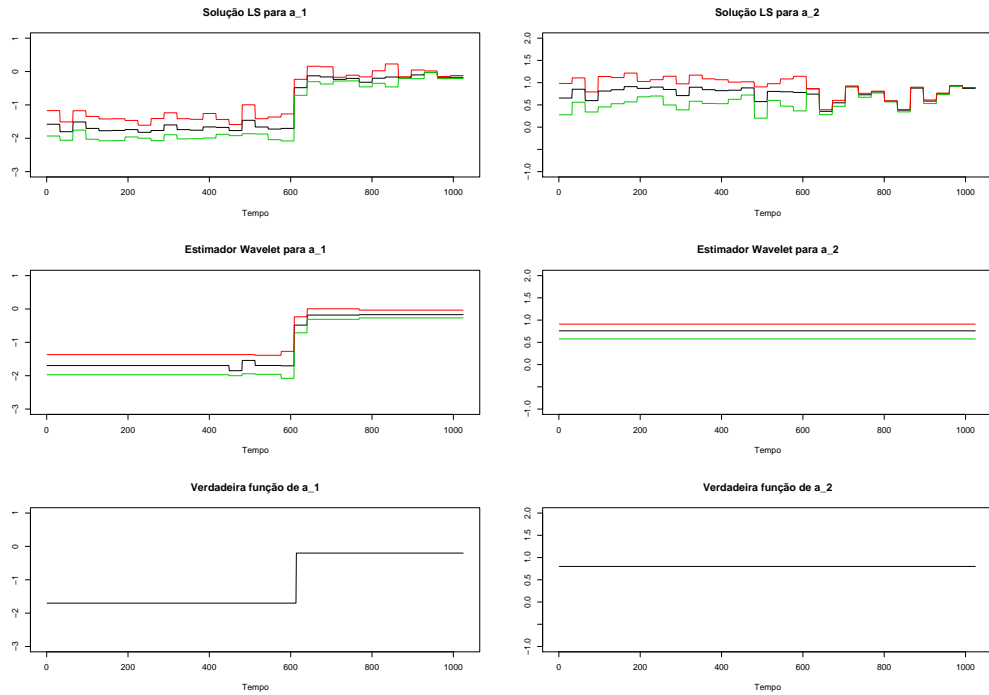


FIGURA 3.5: *Bandas de Confiança para as funções  $a_i(t)$ ,  $i = 1, 2$*

das funções do modelo. O procedimento bootstrap paramétrico se mostra adequado e consistente além de ser rápido. Para conclusões de grande expressão é necessário um estudo mais aprofundado de simulação.

### 3.3 Métodos de previsão para modelos $tvAR(p)$

O objetivo dessa seção é sugerir dois métodos para fazer previsões em modelos Auto Regressivos com coeficientes variando no tempo. Wheelwright et al. (1989) definem um método de previsão como sendo o conjunto de procedimentos usados no desenvolvimento de uma determinada previsão.

Os métodos de previsão de séries temporais, classificados como métodos quantitativos, baseiam suas previsões na extrapolação de características de observações passadas e no inter-relacionamento entre essas observações, fornecendo previsões acuradas se o futuro apresentar comportamento similar ao passado (Wheelwright et al., 1989).

#### 3.3.1 Método de extrapolação

O primeiro método de previsão é fundamentado na extrapolação das características de observações passadas e das funções  $a_i$ , para o modelo  $tvAR(p)$  descrito em 3.1. Para extração

das características dos coeficientes  $a_i(t)$ , consideramos que  $\{a_{i,t}, t = 1, \dots, T \text{ e } i = 1, \dots, p\}$  seja um conjunto de séries temporais, e usamos uma decomposição clássica para cada  $a_i(t)$ ,

$$a_{i,t} = T_{i,t} + S_{i,t} + \eta_{i,t} \quad (3.13)$$

onde  $T_{i,t}$  e  $S_{i,t}$  representam a tendência e sazonalidade, respectivamente, enquanto  $\eta_{i,t}$  é uma componente aleatória.

Por simplicidade, considere um modelo em que componente sazonal  $S_{i,t}$  não está presente, dados por:

$$a_{i,t} = T_{i,t} + \eta_{i,t} \quad (3.14)$$

Há vários métodos para estimar  $T_{i,t}$ . No presente trabalho optamos por um tendência polinomial de grau 1, assim,

$$a_{i,t} = \beta_{i,0} + \beta_{i,1}t + \epsilon_{i,t} \quad (3.15)$$

Aplicando o método de mínimos quadrados em 3.15, obtemos os estimadores  $\hat{\beta}_{i,0}$  e  $\hat{\beta}_{i,1}$  usuais.

Logo as previsões para o modelo  $tvAR(p)$ , dado pela expressão 3.1 são obtidas recursivamente.

$$\hat{X}_T(1) = a_1(1)X_T + a_2(1)X_{T-1} + \dots + a_p(1)X_{T-p} \quad (3.16)$$

é a previsão de  $X_{T+1}$ , com origem fixada no instante  $T$ . As previsões  $k$  passos à frente, com origem em  $T$ , são dadas por

$$\hat{X}_T(k) = \sum_{i=1}^p \hat{a}_i(k) \hat{X}_T(k-i) \quad \text{e} \quad \hat{a}_i(k) = \hat{\beta}_{i,0} + \hat{\beta}_{i,1}(T+k) \quad (3.17)$$

em que  $\hat{X}_T(k-i) = X_{T+k-i}$ , se  $k-i \leq 0$ .

### 3.3.2 Método localmente estacionário

Freqüentemente, as pessoas têm tentado utilizar os modelos Auto Regressivo para dados que mostram certo tipo de comportamento não-estacionário pela aplicação desses modelos em pequenos segmentos.

O segundo método de previsão é baseado nessa idéia e na característica de estacionariedade local do modelo  $tvAR$ , ou seja, o último segmento do modelo é estacionário para um  $N$  pequeno. Portanto, a previsão um passo à frente é dada por

$$\hat{X}_T(1) = \bar{a}_1 X_T + \bar{a}_2 X_{T-1} + \dots + \bar{a}_p X_{T-p} \quad (3.18)$$

e

$$\bar{a}_i = \frac{1}{N} \sum_{t=1}^N \hat{a}_i(T+1-t) \quad i = 1, \dots, p \quad (3.19)$$

sendo portanto  $\bar{a}_i$  a média das estimativas de  $a_i(t)$  nos  $N$  últimos valores de  $t$  em que a série foi observada. As previsões  $k$  passos à frente, com origem em  $T$ , são dadas por

$$\hat{X}_T(k) = \sum_{i=1}^p \bar{a}_i \hat{X}_T(k-i), \quad (3.20)$$

em que  $\hat{X}_T(k-i) = X_{T+k-i}$ , se  $k-i \leq 0$ .

### 3.3.3 Estudo de simulação

Para ilustrar os métodos de previsão apresentados nessa seção foi simulada uma série  $tvAR(2)$  seguindo o modelo 3.1 e ajustados os modelos  $tvAR(2)$ ,  $AR(2)$  e  $AR(6)$ . O modelo  $AR(6)$  foi o melhor ajuste. Para aferir sobre os métodos de previsão e comparar com as previsões obtidas dos modelos  $AR$  foi utilizado o Erro Quadrado Médio da Previsão (EQMP).

A Figura 3.6 é uma série simulada de um  $tvAR(2)$  com tamanho  $T = 2053$  e coeficientes  $a_i(\mu)$  dados por,

$$a_1(\mu) = \begin{cases} 0.1, & \text{se } \mu \leq 0.7 \\ 1.7, & \text{se } \mu > 0.7. \end{cases} \quad a_2(\mu) = 0.77 \quad 0 \leq \mu \leq 1, \quad (3.21)$$

onde  $\mu = t/T$  para  $t = 1, \dots, T$ .

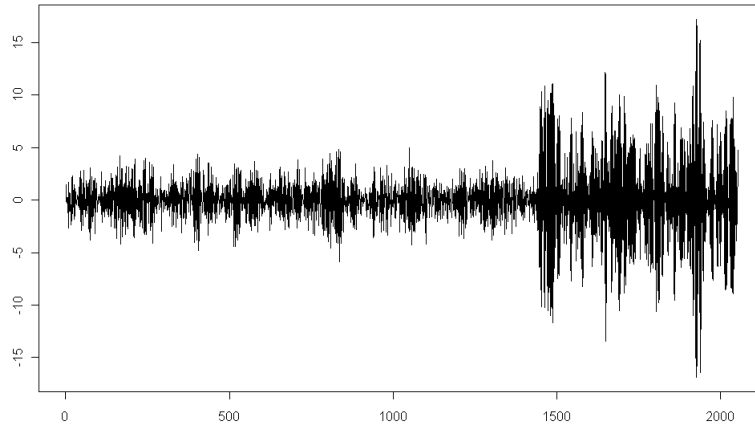


FIGURA 3.6: *Série localmente estacionária simulada com coeficientes dados por 3.21.*

Foram retiradas as cinco últimas observações para fazer previsões da série. Às demais observações da série foram ajustados três modelos: um  $tvAR(2)$ , um  $AR(2)$  e um  $AR(6)$ .

Foi feito a previsão sob os três modelos para uma janela  $h = 5$ . A Tabela 3.1 mostra o Erro Quadrado Médio de Previsão (EQMP para os modelos  $AR(6)$  e  $AR(2)$ ). O EQMP para o ajuste do  $tvAR(2)$  foi calculado segundo os métodos de previsão de extrapolação e localmente estacionário e são indicados por  $tvAR_1(2)$  e  $tvAR_2(2)$  respectivamente. Para o método de previsão, o localmente estacionário consideramos  $N = 2$ .

TABELA 3.1: EQMP para os modelos e métodos.

	$AR(6)$	$AR(2)$	$tvAR_1(2)$	$tvAR_2(2)$
<i>EQMP</i>	6.96	7.99	6.57	14.45

Observando a Tabela 3.1 o EQMP foi menor no modelo  $tvAR(2)$  utilizando o método de previsão de extrapolação. O método de previsão localmente estacionário foi o que apresentou o maior EQMP para série simulada.

### 3.3.4 Resultados e Discussão

As estimativas dos modelos  $tvAR$  são baseadas em expansões *wavelets*, que possuem como já mencionado anteriormente a capacidade de capturar bem o comportamento de funções. Contudo, a estimativa é não paramétrica ficando difícil qualquer tipo de previsão. Por isso propusemos alguns procedimentos para criar métodos de previsão baseados nesse estimador.

O método de previsão por extrapolação conseguiu o melhor desempenho para a série simulada em termo do EQMP. Esse método teve um desempenho superior que o melhor ajuste Auto Regressivo, que tem um número significativo de coeficientes a menos. Por outro lado, o método de previsão localmente estacionário, apresentou uma acuidade inferior aos demais. Outros estudos de simulação devem ser feitos para comprovar se o método pode ser descartado.

O objetivo do presente estudo foi apresentar alguns métodos e sugestões para fazer previsão em modelos  $tvAR$ . Os métodos estão abertos a discussões, podemos salientar que há muito estudo a ser feito, na proposta de extrapolação não foi considerado a sazonalidade, o ajuste da tendência foi polinomial e de grau 1. São necessários estudos mais aprofundados tanto no sentido teórico como empírico para consolidar algum método ou descartá-lo.

Para uma discussão no contexto de procedimentos para melhorar a previsão em processos estacionários ver Kabaila et al. (2008).

# Capítulo 4

## Modelo ARCH com coeficientes variando no tempo

O modelo ARCH (Auto Regressive Conditional Heteroscedasticity) foi proposto por Engle (1983). Apesar de ter sido desenvolvido para modelar e prever inflação, as propriedades do modelo foram consideradas úteis para análise de séries financeiras em geral. Muitas séries temporais exibem períodos de grande volatilidade seguidos de períodos de relativa tranquilidade. Nestes casos, a suposição de variância constante (homocedasticidade) pode não ser apropriada. A estratégia usada por Engle foi modelar  $\sigma_t^2$  como dependente dos quadrados dos retornos passados,  $X_{t-1}, X_{t-2}, \dots$ , e definiu  $X_t$  como um produto de  $\sigma_t^2$  por

$$X_t = \sqrt{\sigma_t^2} \epsilon_t \quad (4.1)$$

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^p \alpha_j X_{t-j}^2, t \in Z \quad (4.2)$$

onde  $\{\epsilon_t\}$  é uma seqüência de variáveis aleatórias i.i.d. com média zero e variância unitária,  $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, p$ . Este modelo será denotado por  $ARCH(p)$ . Em que uma extensão desses modelos proposta por Dahlhaus et al. (2007) supõe-se que os parâmetros  $\alpha_j$  variem no tempo, ou seja,  $\alpha_j(t)$ . A fim de obter uma estrutura tratável para desenvolver uma teoria assintótica, assim como em regressão não paramétrica (Donoho et al., 1998) e para processos localmente estacionários (Dahlhaus, 1997). Os parâmetros  $\alpha_j(t)$  para  $t = 1, \dots, N$ , devem ser re-parametrizados. Assim

$$\begin{aligned} X_{t,N} &= \sigma_{t,N} \epsilon_t \\ \sigma_{t,N}^2 &= \alpha_0 \left( \frac{t}{N} \right) + \sum_{j=1}^p \alpha_j \left( \frac{t}{N} \right) X_{t-j,N}^2, t = 1, \dots, N \end{aligned} \quad (4.3)$$

Chamamos a seqüência  $\{X_{t,N} : t = 1, \dots, N\}$  que satisfaz a equação 4.3, um processo ARCH com coeficientes variando no tempo de ordem  $p$  (tvARCH(p)).

Dahlhaus et al. (2006) mostraram que o processo tvARCH pode ser localmente aproximado por um processo ARCH estacionário. Além disso, eles estenderam várias propriedades do processo ARCH para um processo tvARCH (Dahlhaus et al., 2006).

Antes de apresentarmos o método de estimação dos parâmetros do modelo tvARCH(p) proposto por Dahlhaus et al. (2007), vamos fazer uma breve revisão bibliográfica sobre algoritmo *Least Mean Square* (LMS).

## 4.1 Algoritmos *Least Mean Square*

Esse tipo de algoritmo está relacionado com processamento de sinal e tem uma grande aplicação em telefonia celular e comunicações. O desempenho dos algoritmos adaptativos depende do comprimento do filtro e do parâmetro de convergência  $\mu$ .

### 4.1.1 Filtros adaptativos

Um sistema é dito ser adaptativo ao tentar ajustar seus parâmetros com o auxílio de algumas metas bem definidas que dependem do estado do sistema e dos seus arredores (Honig et al., 1987). Portanto, o sistema ajusta-se de modo a responder a alguns fenômenos que estão ocorrendo ao seu redor.

Suponha um filtro adaptável simples, representado na equação 4.4, que tenha uma única entrada,  $x(n)$  e uma saída  $y(n)$ . A saída  $y(n)$  é gerada como uma combinação linear da seqüência de entrada  $x(n)$ , de acordo com a equação

$$y(n) = \sum_{i=0}^{N-1} w_i(n)x(n-i) \quad (4.4)$$

onde  $N$  é o comprimento do filtro,  $\{w_i(n)\}$  são os pesos do filtro (coeficientes) que variam no tempo e que são controlados pela adaptação do algoritmo e  $\{x(n-i)\}$  são as observações de entrada. (Haykin, 2001).

### 4.1.2 Algoritmo LMS

O algoritmo LMS (*Least Mean Square*) baseia-se no método dos gradientes descendentes. O algoritmo calcula iterativamente os coeficientes do filtro que tende para a solução ótima dos pesos. A idéia por trás dos filtros LMS é encontrar um vetor de coeficiente  $w_n$  que minimiza uma função custo no sentido do gradiente dela (Sridharan et al., 1986). O algoritmo LMS pode ser resumido como

Parâmetros:  $p$  = tamanho do filtro;  $\mu$  = taxa de aprendizado

Inicialização:  $w_0 = 0$

Computacionalmente:

Para  $n = 0, 1, 2, \dots$

$$\bar{x}(n) = [x(n), x(n-1), \dots, x(n-p+1)]^T \quad (4.5)$$

$$e(n) = y(n) - \hat{y}(n) \quad (4.6)$$

$$\hat{y}(n) = \bar{w}_n^T \bar{x}(n) \quad (4.7)$$

$$\bar{w}_n = \bar{w}_{n-1} + \mu e(n) \bar{x}(n) \quad (4.8)$$

A função custo  $J$  é dada por

$$J = E [(e(n))^2]$$

que tem um ponto de mínimo quando  $\nabla J = 0$  sendo  $\nabla J = [\partial J / \partial w_0, \dots, \partial J / \partial w_{N-1}]^T$ .

Na equação 4.5,  $\bar{x}(n)$  é o vetor de entrada, enquanto  $x(n), x(n-1), \dots, x(n-p+1)$  são as entradas individuais. Na equação 4.6,  $e(n)$  e  $\hat{y}(n)$  denotam o erro da saída e a estimativa do sinal desejado respectivamente.

Por fim, na equação 4.7 a fórmula para a atualização do vetor de coeficiente após a  $n$ -ésima iteração.

### 4.1.3 Algoritmo NLMS

A principal desvantagem do algoritmo LMS, é a grande sensibilidade ao tamanho dos valores de entrada  $x(n)$ . Isso torna difícil a escolha da taxa de aprendizado  $\mu$  que garanta a estabilidade do algoritmo (Lim et al., 1987). Uma variante do algoritmo LMS é o *Normalized Least Mean Squares* (NLMS) que resolve este problema através da normalização do vetor de entrada (Lim et al., 1987). O algoritmo NLMS pode ser resumido como:

Parâmetros:  $p =$  tamanho do filtro;  $\mu =$  taxa de aprendizado

Inicialização:  $w_0 = 0$

Computacionalmente:

Para  $n = 0, 1, 2, \dots$

$$\bar{x}(n) = [x(n), x(n-1), \dots, x(n-p+1)]^T \quad (4.9)$$

$$e(n) = y(n) - \hat{y}(n) \quad (4.10)$$

$$\hat{y}(n) = \bar{w}_n^T \bar{x}(n) \quad (4.11)$$

$$\bar{w}_n = \bar{w}_{n-1} + \frac{\mu e(n) \bar{x}(n)}{\bar{x}^T(n) \bar{x}(n)} \quad (4.12)$$

A diferença entre os algoritmos LMS e NLMS está na equação 4.8 e 4.12 que atualizam os pesos  $\bar{w}_n$  (Haykin et al., 2001).

## 4.2 Um algoritmo recursivo para estimação dos parâmetros do modelo tvARCH(p)

Com base nos algoritmos LMS e NLMS, Dahlhaus et al. (2007) propuseram o seguinte estimador recursivo dos coeficientes  $\alpha(t) = \alpha_t$  do modelo  $tvAR(p)$ ,

$$\hat{\alpha}_{t,N} = \hat{\alpha}_{t-1,N} + \lambda \left\{ X_{t,N}^2 - \hat{\alpha}_{t-1,N}^T \chi_{t-1,N} \right\} \frac{\chi_{t-1,N}}{|\chi_{t-1,N}|_1^2} \quad (4.13)$$

onde

$$\chi_{t-1,N}^T = (1, X_{t-1,N}^2, \dots, X_{t-p,N}^2), |\chi_{t-1,N}|_1 = 1 + \sum_{j=1}^p X_{t-j,N}^2$$

sendo a condição inicial dada por  $\hat{\alpha}_{p,N} = (0, \dots, 0)_{p+1}$ . O algoritmo é linear nos estimadores, apesar da não-linearidade do processo tvARCH.

Para diminuir o vício do estimador dado acima, Dahlhaus et al. (2007) propõe um estimador obtido da combinação linear de dois estimadores do tipo 4.13, calculados segundo dois valores de  $\lambda$ , indicados por  $\lambda_1$  e  $\lambda_2$ , com  $\lambda_2 = w\lambda_1$  e  $0 < w < 1$ .

$$\check{\alpha}_{t_0,N}(w) = \frac{1}{1-w} \hat{\alpha}_{t_0,N}(\lambda_1) - \frac{w}{1-w} \hat{\alpha}_{t_0,N}(\lambda_2) \quad (4.14)$$

As propriedades deste estimador são estudadas em Dahlhaus et al. (2007) em um contexto não-estacionário e a normalidade assintótica e uma expressão para o vício devido à não estacionariedade estão estabelecidas.

Para usar 4.14, é necessário escolher os valores de  $\lambda_1$  e  $w$ . A escolha desses valores depende do grau não-estacionariedade, de modo que  $\lambda_1$  deve ser grande se as características do processo mudam rapidamente.

## 4.3 Método de previsão para o modelo tvARCH(p)

A análise de uma série temporal real compreende vários itens. A princípio, há a necessidade de descobrir a estrutura da série, e uma vez escolhido por algum critério um modelo para a série, o objetivo pode ser o de fazer previsões de valores futuros da série.

Nesta seção o foco é propor uma maneira de prever valores futuros de uma série gerada segundo um modelo  $tvARCH(p)$ .



Para séries  $ARCH(p)$  dadas por 4.1 e 4.2 a previsão do valor futuro um passo à frente e  $k$  passos à frente são obtidas a partir das expressões 4.15 e 4.16 respectivamente.

$$\hat{\sigma}_t^2(1) = \alpha_0 + \alpha_1 X_t^2 + \alpha_2 X_{t-1}^2 + \dots + \alpha_p X_{t-p+1}^2 \quad (4.15)$$

onde  $t$  é a origem fixada. As previsões  $k$  passos à frente, com origem em  $t$ , são dadas por

$$\hat{\sigma}_t^2(k) = \alpha_0 + \sum_{i=1}^p \alpha_i \hat{\sigma}_t^2(k-i), \quad (4.16)$$

em que  $\hat{\sigma}_t^2(k-i) = X_{t+k-i}^2$ , se  $k-i \leq 0$ . [Morettin et al. (2004)]

Para fazer previsões  $k$  passos à frente de uma série  $tvARCH(p)$  observada até um ponto  $t$ , inspirado na construção de previsões para  $ARCH(p)$  propomos que obtenha-se previsões de  $\sigma_{t+k}^2 = \sigma^2(t+k)$  e  $\hat{\alpha}_{i,t+k} = \hat{\alpha}_i(t+k)$  de forma recursiva, começando pela previsão de  $\sigma_{t+1}^2 = \sigma_t^2(1)$ , dado por

$$\hat{\sigma}_t^2(1) = \alpha_0(t) + \alpha_1(t) X_t^2 + \alpha_2(t) X_{t-1}^2 + \dots + \alpha_p(t) X_{t-p}^2 \quad (4.17)$$

A seguir, faz-se a previsão de  $X_{t+1}$  usando  $\hat{X}_{t+1} = \hat{\sigma}_{t+1}^2$  para usar a previsão de  $\alpha_i(t)$ .

Estima-se a seguir cada função  $\alpha_i(t)$  em  $(t+1)$ , calculando  $\hat{\alpha}_i(t+1)$  através da expressão 4.13 para dois pares de  $\lambda$ , e usando a seguir 4.14.

Para a estimação de  $\alpha_i(t+2)$  com  $i = 1, \dots, p$ , procede-se de forma semelhante. Inicialmente faz-se a previsão de  $\sigma^2(t+2) = \sigma^2(2)$  usando 4.18.

$$\hat{\sigma}_t^2(2) = \hat{\alpha}_0(t+1) + \sum_{i=1}^p \hat{\alpha}_i(t+1) \hat{\sigma}_t^2(k-i), \quad (4.18)$$

onde  $\hat{\sigma}_t^2(k-i) = X_{t+k-i}^2$ , se  $k-i \leq 0$ . E em seguida faz-se a previsão  $\hat{X}_{t+2} = \sigma_{t+2}^2$ . Esta previsão é usada na expressão 4.13.

De forma geral a previsão  $k$  passos à frente é dado por duas equações recursivas. Assim,

$$\hat{\sigma}_t^2(k) = \hat{\alpha}_0(t+k-1) + \sum_{i=1}^p \hat{\alpha}_i(t+k-1) \hat{\sigma}_t^2(k-i), \quad (4.19)$$

onde  $\hat{\alpha}_0(t+k-1)$  são calculados por

$$\hat{\alpha}_{t+k-1,N}(w) = \frac{1}{1-w} \hat{\alpha}_{t+k-1,N}(\lambda_1) - \frac{w}{1-w} \hat{\alpha}_{t+k-1,N}(\lambda_2) \quad (4.20)$$

Os parâmetros  $\lambda_1$  e  $w$  são o principal interesse do nosso estudo de simulação, assim como o comportamento do estimador proposto por Dahlhaus et al. (2007) mostrado em 4.14.

## 4.4 Estudo de Simulação

### 4.4.1 Introdução

Para avaliar o comportamento assintótico do estimador recursivo (proposto em 4.14) dos parâmetros do modelo  $tvARCH(p)$  descrito em 4.3, é necessário fazer estudos de simulação do modelo em diversas composições de número de parâmetros ( $p$ ), tamanho  $N$  da série e valores de  $\lambda_1$  e  $w$ .

### 4.4.2 Tamanho da Série

Quando estamos falando de uma teoria assintótica algumas perguntas são frequentemente levantadas a respeito do tamanho da amostra (no nosso caso o tamanho da série) e sua influência. Entretanto, é muito comum que, na prática estejam disponíveis apenas amostras relativamente pequenas para a análise. Dessa forma, é natural que apareça um interesse em saber como o estimador se comporta em tal situação, analisar o comportamento dos estimadores para amostras pequenas e grandes e se possível determinar ou indicar um tamanho mínimo para o uso do estimador.

### 4.4.3 Estudo por simulação do efeito do tamanho $N$ da série

Em um primeiro estudo, o objetivo foi comparar as estimativas dos coeficientes de um modelo  $tvARCH(p)$  a partir de duas realizações com tamanhos  $N = 1000$  e  $N = 5000$  respectivamente.

O modelo gerado é dado por 4.3 e os coeficientes são:

$$\alpha_0(t) = 0.9 \quad \forall t \quad \alpha_1(t) = \begin{cases} 0.1, & \text{se } t \leq N/2 \\ 0.8, & \text{se } t > N/2. \end{cases}$$

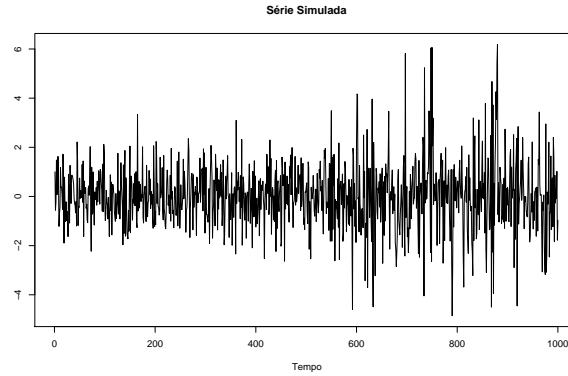


FIGURA 4.1: *Série simulada de tamanho  $N = 1000$  de um modelo  $tvARCH(1)$ .*

A partir das duas realizações foi feita a estimação dos coeficientes usando a proposta dado em 4.14, e tendo sido escolhido  $\lambda_1 = 0.01$  e  $\lambda_2 = 0.001$  ( $w = 0.1$ ) estão apresentadas na Figura 4.2.

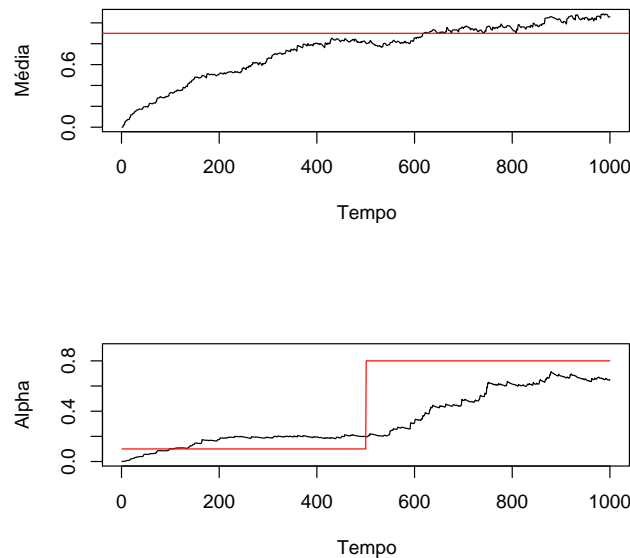


FIGURA 4.2: *Estimativas dos coeficientes  $\alpha_j(t)$ ,  $j = 0, 1$ , do modelo  $tvARCH(1)$  obtidas recursivamente por 4.14 a partir de uma série de tamanho  $N = 1000$ .*

A Figura 4.2 apresenta as estimativas de  $\alpha_j(t)$ ,  $j = 0, 1$  em preto e os verdadeiros valores de  $\alpha_j(t)$ ,  $j = 0, 1$  em vermelho. Observamos que as estimativas gastam muitas observações para se aproximar dos verdadeiros valores de  $\alpha_j(t)$ ,  $j = 0, 1$ .

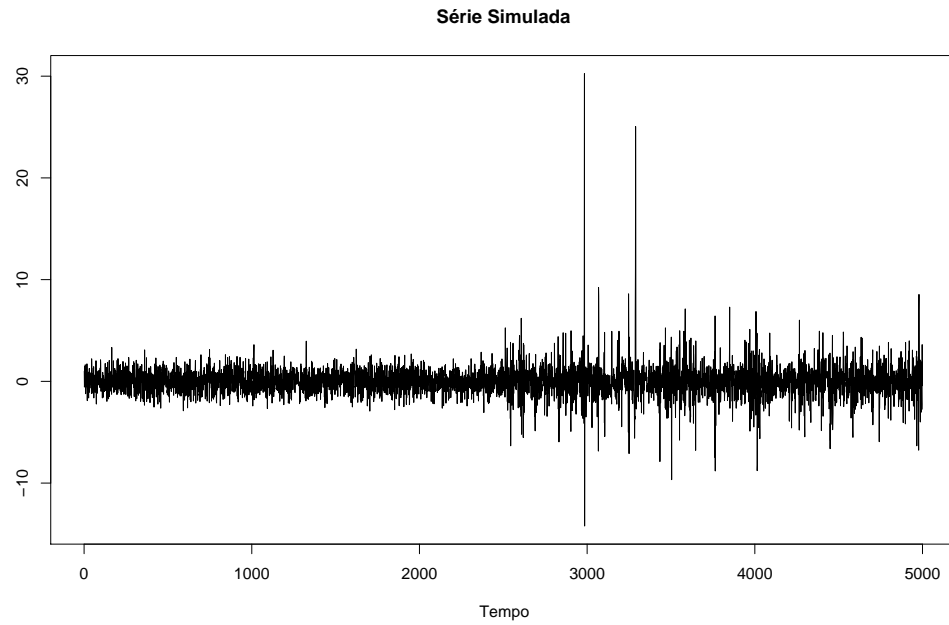


FIGURA 4.3: *Série simulada de tamanho  $N = 5000$  de um modelo  $tvARCH(1)$ .*

A Figura 4.4 mostra a estimativa dos coeficientes  $\alpha_j$ ,  $j = 0, 1$ , obtidas segundo 4.14 com  $\lambda_1 = 0.01$  e  $\lambda_2 = 0.001$ , ( $w = 0.1$ ).

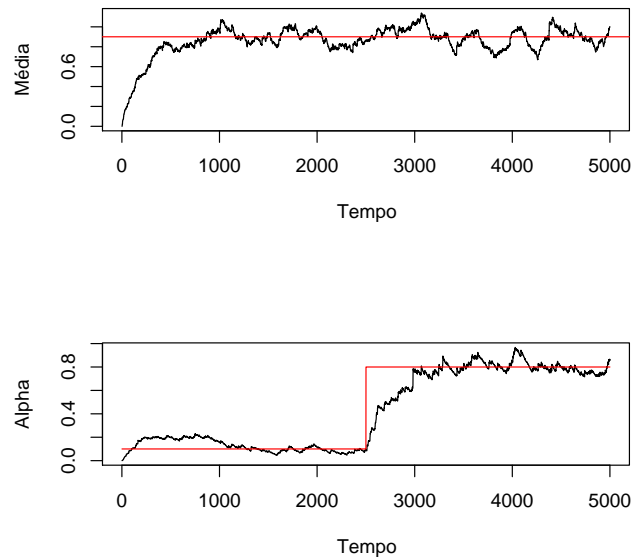


FIGURA 4.4: *Estimativas dos coeficientes  $\alpha_j(t)$ ,  $j = 0, 1$ , do modelo  $tvARCH(1)$  usando  $\lambda_1 = 0.01$  e  $\lambda_2 = 0.001$  a partir de uma série de tamanho  $N = 5000$ .*

Observamos na Figura 4.4 que o estimador sempre precisa de, por volta de 500

iterações para “esquentar”.

#### 4.4.4 Resultados e Discussão

O tamanho da série influencia as propriedades do estimador. As Figuras 4.2 e 4.4 mostram uma melhora significativa na estimação dos coeficientes  $\alpha_i(t)$  quando se aumenta o tamanho da série. É claro que as propriedades deste estimador sofrem influência do tamanho da série, mas existem ainda outros parâmetros a se considerar, como os valores de  $\lambda_1$  e  $\lambda_2$ . Esse assunto será abordado na próxima simulação.

#### 4.4.5 Estudo por simulação do efeito de escolha dos parâmetros $\lambda_1$ e $\lambda_2$ .

Uma segunda abordagem do estudo para simulação foi a procura dos efeitos de escolha de  $\lambda_1$  e  $\lambda_2$  sobre as propriedades dos estimadores de  $\alpha_i(t)$ .

O estudo consistiu na análise de uma série de tamanho  $N = 5000$  gerada sob o modelo 4.3, com parâmetros definidos por:

$$\alpha_0(t) = 0.5 \quad \forall t \quad \alpha_1(t) = \begin{cases} 0.2, & \text{se } t \leq N/2 \\ 0.8, & \text{se } t > N/2. \end{cases} \quad \alpha_2(t) = 0.3 \quad \forall t$$

Foi simulada uma única realização do processo, e foram calculadas as estimativas dos parâmetros, seguindo o algoritmo dado em 4.13 e 4.14. A série simulada está apresentada na Figura 4.5.

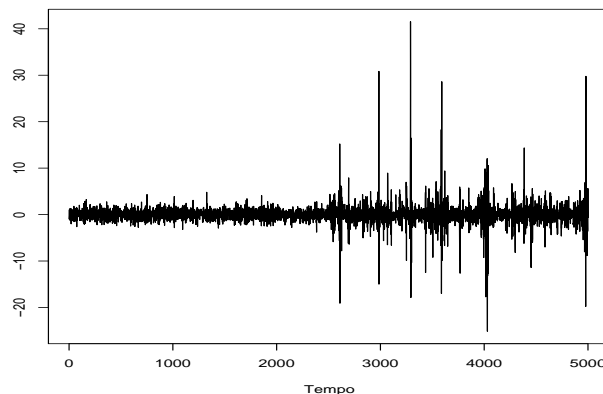


FIGURA 4.5: *Série simulada de acordo com o modelo tvARCH(2)*

Para usar o algoritmo 4.14 foram considerados 3 valores de  $\lambda_1$ , (0.1, 0.01, 0.001) e para cada valor de  $\lambda_1$  foram tomados 2 valores de  $w$ , (0.1, 0.9). O algoritmo foi usado portanto

6 vezes, estão apresentados nas Figuras 4.6-4.11.

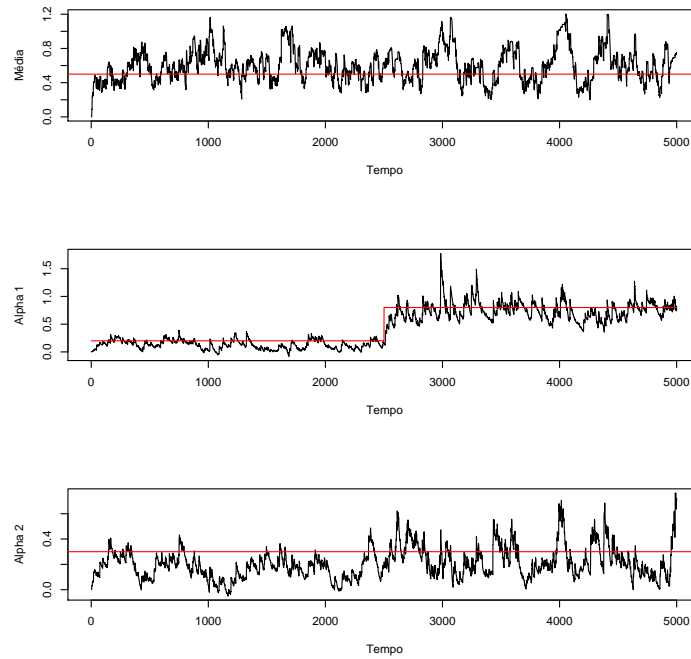


FIGURA 4.6: *Estimativas dos coeficientes  $\alpha_i$ ,  $i = 0, 1, 2$ , do modelo  $tvARCH(2)$  com  $\lambda_1 = 0.1$  e  $w = 0.1$*

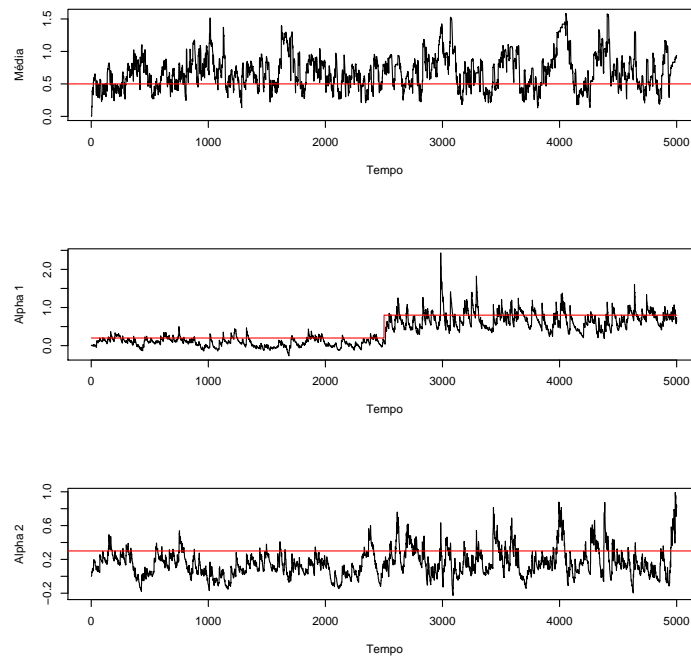


FIGURA 4.7: *Estimativas dos coeficientes  $\alpha_i$ ,  $i = 0, 1, 2$ , do modelo  $tvARCH(2)$  com  $\lambda_1 = 0.1$  e  $w = 0.9$*

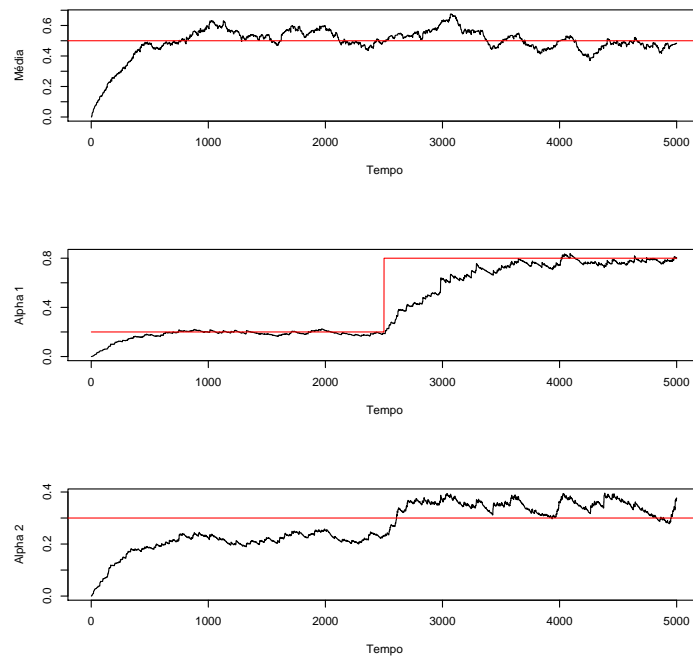


FIGURA 4.8: *Estimativas dos coeficientes  $\alpha_i$ ,  $i = 0, 1, 2$ , do modelo  $tvARCH(2)$  com  $\lambda_1 = 0.01$  e  $w = 0.1$*

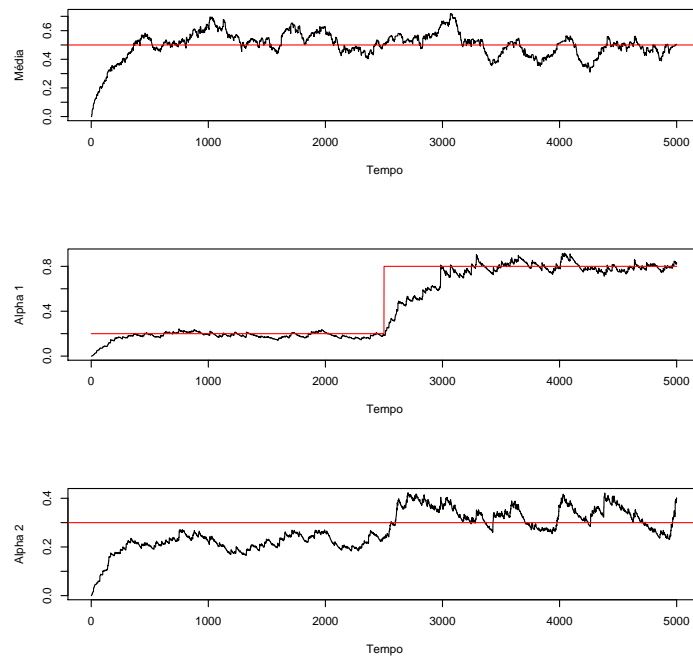


FIGURA 4.9: *Estimativas dos coeficientes  $\alpha_i$ ,  $i = 0, 1, 2$ , do modelo  $tvARCH(2)$  com  $\lambda_1 = 0.01$  e  $w = 0.9$*

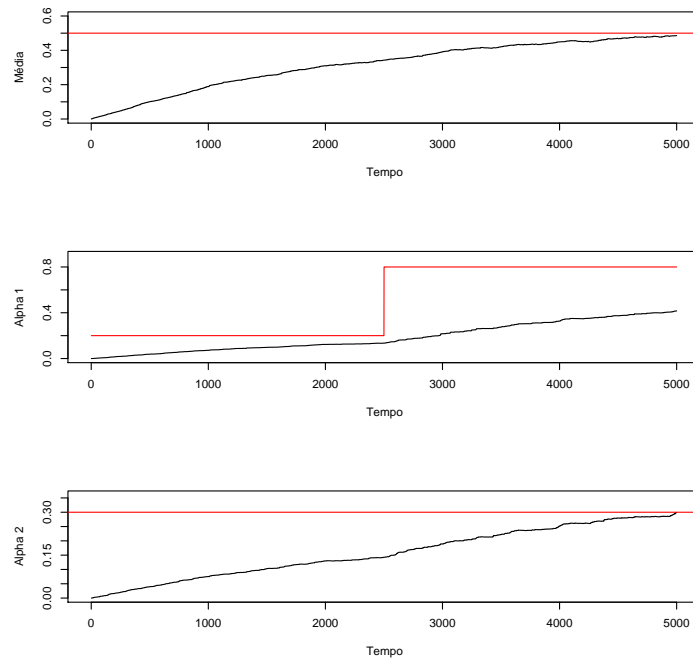


FIGURA 4.10: Estimativas dos coeficientes  $\alpha_i$ ,  $i = 0, 1, 2$ , do modelo  $tvARCH(2)$  com  $\lambda_1 = 0.001$  e  $w = 0.1$

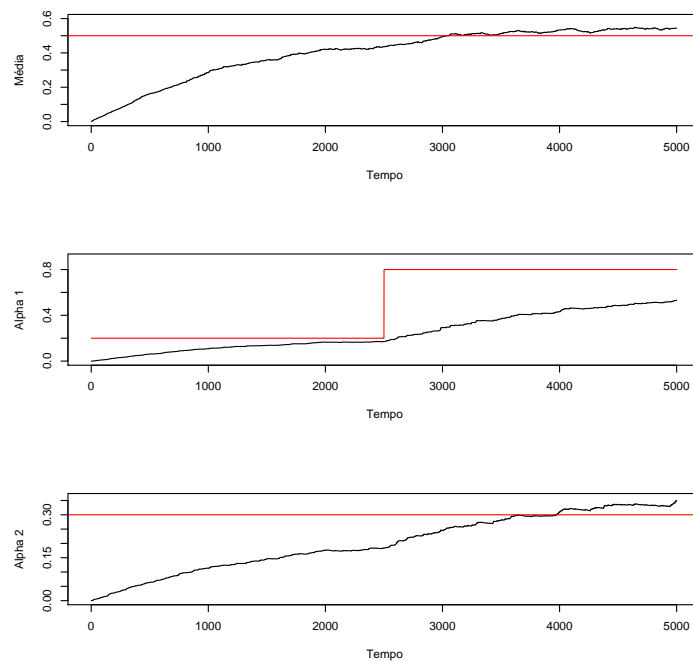


FIGURA 4.11: Estimativas dos coeficientes  $\alpha_i$ ,  $i = 0, 1, 2$ , do modelo  $tvARCH(2)$  com  $\lambda_1 = 0.001$  e  $w = 0.9$



### 4.4.6 Resultados e Discussão

Observa-se nos gráficos que fixando o valor de  $\lambda_1$  os valores escolhidos de  $w$  não apresentam uma grande influência na estimativa de  $\alpha_i$ . A situação muda completamente quando é fixado o valor de  $w$  e diversos valores para  $\lambda_1$  são considerados. A estimativa obtida sofre uma grande influência em dois aspectos. Em primeiro lugar nota-se que, a variância do estimador está relacionada com o valor de  $\lambda_1$  e em segundo lugar observa-se que a taxa de convergência também está relacionada com o valor de  $\lambda_1$ . A convergência da estimação é mais rápida quando são usados valores grandes para  $\lambda_1$ . Para valores pequenos de  $\lambda_1$ , as estimativas são curvas suavizadas mas levam um grande período de tempo até a convergência. Além disso, variações em  $\alpha_i$  não são detectadas.

São necessários mais estudos sobre o tema, tanto o ponto de vista teórico como de simulações. Já se pode deduzir, porém, que o valor ótimo de  $\lambda_1$  deve estar relacionado com o grau de heteroscedasticidade da série.

## 4.5 Uma proposta para escolha do parâmetro $\lambda$ em 4.13.

Como era previsto e observando o estudo de simulação (4.4) fica evidente que existe uma relação forte entre o valor de  $\lambda$  e a convergência do estimador. Dahlhaus et al. (2007) deixa claro que  $\lambda$  deve ser escolhido pelo grau de não-estacionariedade. Isto é,  $\lambda$  deve ser grande se as características do processo mudam mais rápido. No entanto, uma sugestão mais específica exige investigações teóricas e por simulações.

Nossa sugestão para o problema da escolha de  $\lambda$  em 4.13 e 4.14 é fazê-lo também variar no tempo  $\lambda = \lambda(t)$ . Assim o novo estimador de  $\alpha_i(t)$  será dado por

$$\hat{\alpha}_{t,N} = \hat{\alpha}_{t-1,N} + \lambda_{t,N} \left\{ X_{t,N}^2 - \hat{\alpha}_{t-1,N}^T \chi_{t-1,N} \right\} \frac{\chi_{t-1,N}}{|\chi_{t-1,N}|_1^2} \quad (4.21)$$

onde

$$\chi_{t-1,N}^T = (1, X_{t-1,N}^2, \dots, X_{t-p,N}^2), |\chi_{t-1,N}|_1 = 1 + \sum_{j=1}^p X_{t-j,N}^2$$

A condição inicial é dada por  $\hat{\alpha}_{p,N} = (0, \dots, 0)$ .

Os valores de  $\lambda(t)$  podem ser obtidos a partir de uma medida de característica do processo  $\{X_t\}$  em intervalos de tempo com comprimento menor que  $T$ .

Usamos a variância da série como medida de característica ( $mc$ ) e o tamanho do intervalo de tempo igual a duas vezes a quantidade de parâmetros ( $p$ ), ou seja,

$$mc(t) = \text{var}[X_{t-1}, X_{t-2}, \dots, X_{t-2p}] \quad (4.22)$$

Foi usado uma padronização de  $mc(t)$  dado por

$$\hat{mc}(t) = \frac{mc(t) - a}{b - a} \quad (4.23)$$

onde  $a = \min\{mc(t)\}$  e  $b = \max\{mc(t)\}$  para todo  $t$ .

Com base no estudo de simulação realizado na seção (4.1) o valor ótimo de  $\lambda$  deveria estar contido em um intervalo  $[0.01, 0.1]$ . Portanto a idéia inicial foi definir  $\lambda(t)$  usando um *threshold* em  $\hat{mc}(t)$ ;

$$\lambda(t) = \delta_w(\hat{mc}(t)) \quad (4.24)$$

e

$$\delta_w(x) = \begin{cases} 0.01, & \text{se } |x| \leq w \\ 0.1, & \text{se } |x| > w, \end{cases} \quad \text{com } w = 0.1. \quad (4.25)$$

O valor  $w = 0.1$  no *threshold* 4.25 foi escolhido de forma empírica. Sua justificativa esta associada ao fato de  $b$  ser a maior característica da série onde existe grande volatilidade e qualquer valor superior a 10% do valor de  $b$  é considerado uma característica forte, ou seja, o valor de  $\lambda$  associado deve ser grande. Para valores inferiores a 10% do valor de  $b$  são considerados como sendo um padrão da série e seu comportamento deve ser suave ao longo do tempo com isso valores de  $\lambda = 0.01$  devem ser mais adequados.

Realizamos um estudo empírico desse novo estimador proposto em 4.21, com  $\lambda$  variando no tempo. O cálculo de  $\lambda(t)$  foi feito como sugerido em 4.22-4.25.

## 4.6 Um estudo de simulação $\lambda(t)$

### 4.6.1 Objetivo

Nesse estudo estamos interessados em avaliar o comportamento do estimador proposto em 4.21 onde o parâmetro  $\lambda(t)$  varia no tempo e tem uma relação com as características da série. Além disso, podemos comparar o estimador 4.14 sugerido por Dahlhaus et al. (2007) com o estimador sugerido em 4.21.

### 4.6.2 Especificação da simulação

Foram simuladas realizações de 3 modelos  $tvARCH(p)$  definido em 4.3, com  $p = \{1, 2, 3\}$ . E coeficientes dos modelos dado por funções escadas. O tamanho de cada série foi  $N = 5000$ . Foram calculadas as estimativas de  $\alpha_i(t)$  dadas em 4.14 e 4.21. Os resultados são comparados para discussões dos estimadores.

Simulação 1: Geramos uma série  $\{X_t\}$  que segue o modelo tvARCH(1) definido em 4.14 com coeficientes dados por,

$$\alpha_0 = 0.5 \quad \alpha_1 = \begin{cases} 0.2, & \text{se } t \leq 2500 \\ 0.9, & \text{se } t > 2500. \end{cases}$$

O gráficos da série  $\{X_t\}$  da função  $\hat{m}c(t)$  dada em 4.23 e das estimativas 4.14 e 4.21 estão nas Figuras 4.12 - 4.15 respectivamente.

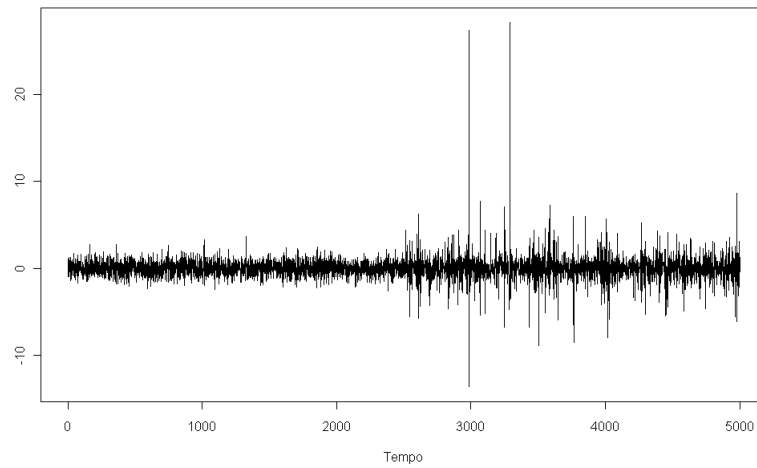


FIGURA 4.12: *Série simulada de acordo com o modelo tvARCH(1)*

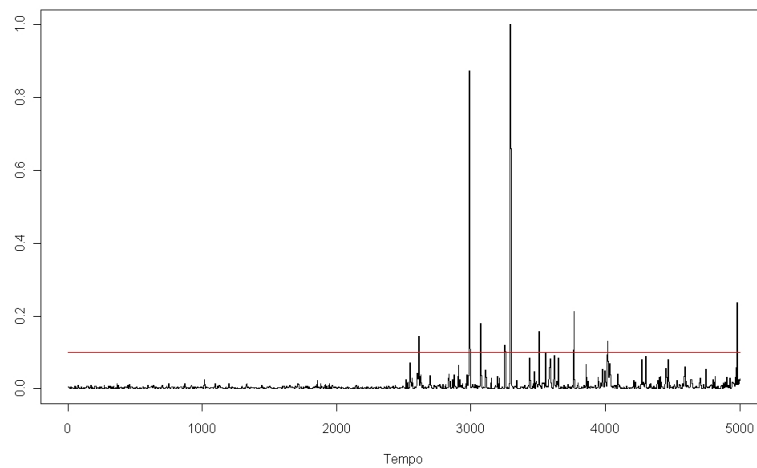


FIGURA 4.13: *Gráfico da função  $\hat{m}c(t)$  e do parâmetro  $w = 0.1$ .*

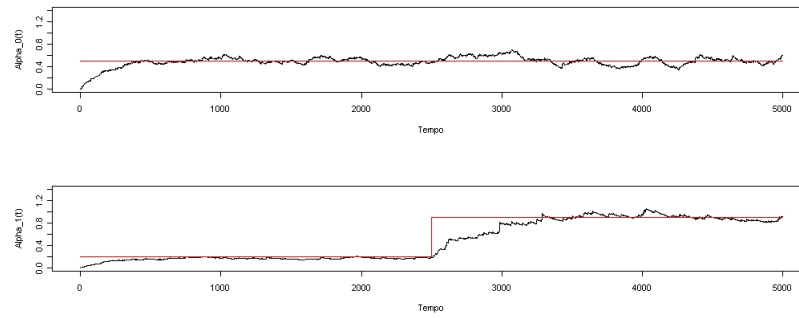


FIGURA 4.14: *Estimativas do modelo tvARCH(1) utilizando 4.14 com  $\lambda = 0.01$  e  $w = 0.1$ .*

Observamos na Figura 4.14 que o estimador de  $\alpha_1(t)$  precisou de 800 iterações para captar o salto no coeficiente  $\alpha_1(t)$ .

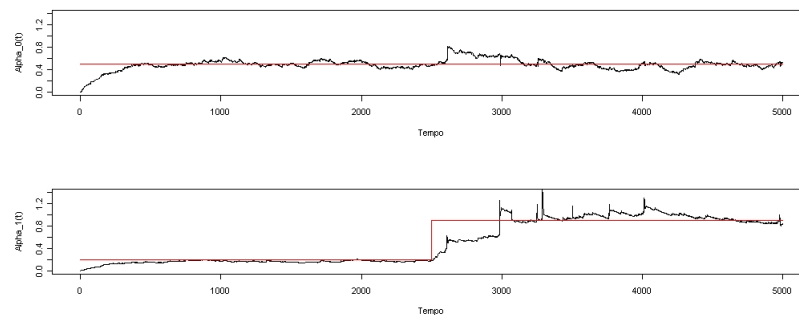


FIGURA 4.15: *Estimativas do modelo tvARCH(1) utilizando 4.21 com  $\lambda(t) = \delta_w(\hat{m}c(t))$ .*

Por outro lado, observamos na Figura 4.15 que utilizando 4.21 com a sugestão de relacionar  $\lambda$  com o grau de heteroscedasticidade da série, o estimador precisou de menos iterações para captar o salto no coeficiente  $\alpha_1(t)$ .

Simulação 2: Geramos uma série  $\{X_t\}$  que segue o modelo tvARCH(2) definido em

4.3 com coeficientes dados por,

$$\alpha_0 = 0.5 \quad \alpha_1 = \begin{cases} 0.2, & \text{se } t \leq 2500 \\ 0.8, & \text{se } t > 2500. \end{cases} \quad \alpha_2 = 0.3$$

O gráficos da série  $\{X_t\}$  da função  $\hat{m}c(t)$  dada em 4.23 e das estimativas 4.14 e 4.21 estão nas Figuras 4.16 - 4.19 respectivamente.

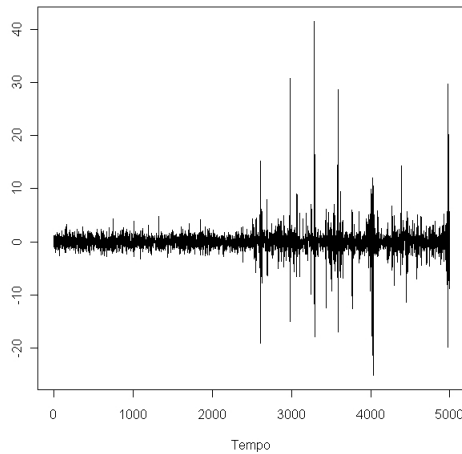


FIGURA 4.16: *Série simulada de acordo com o modelo tvARCH(1)*

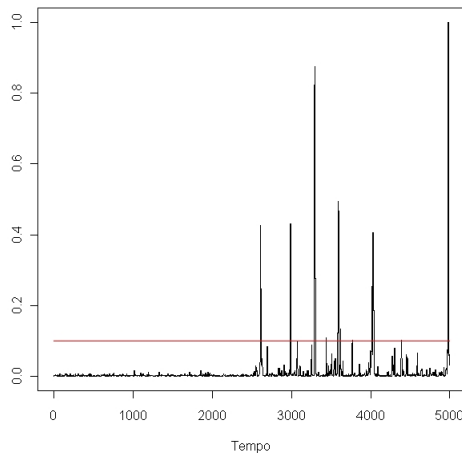


FIGURA 4.17: *Gráfico da função  $\hat{m}c(t)$  e do parâmetro  $w = 0.1$ .*

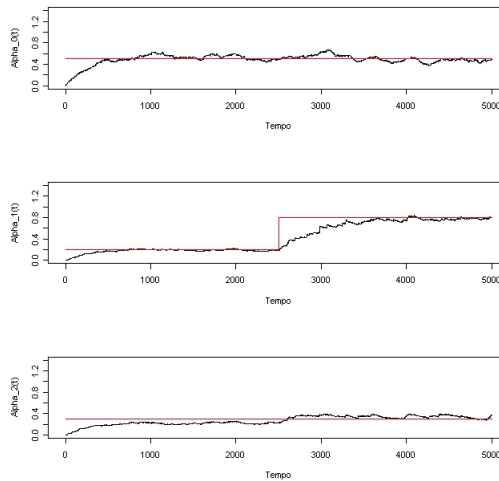


FIGURA 4.18: *Estimativas do modelo tvARCH(2) utilizando 4.14 com  $\lambda = 0.01$  e  $w = 0.1$ .*

Observamos novamente na Figura 4.18 que o estimador de  $\alpha_1(t)$  precisou de muitas iterações para captar o salto no coeficiente  $\alpha_1(t)$ .

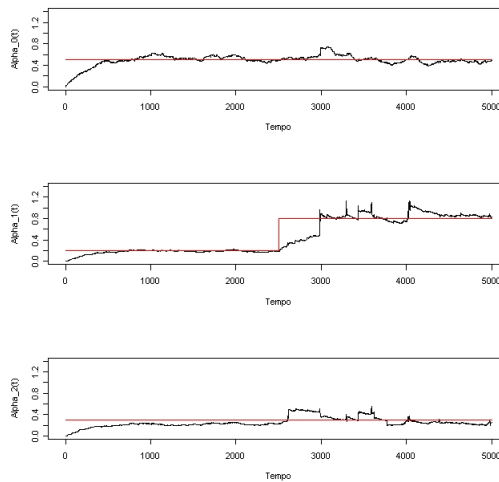


FIGURA 4.19: *Estimativas do modelo tvARCH(2) utilizando 4.21 com  $\lambda(t) = \delta_w(\hat{m}c(t))$ .*

Observamos na Figura 4.19 que utilizando 4.21 com a sugestão de relacionar  $\lambda$  com o grau de heteroscedasticidade da série, o estimador precisou de menos iterações para captar o salto no coeficiente  $\alpha_1(t)$ .

## 4.7 Resultados e Discussão

Analisando o estudo de simulação realizado em 4.6 o estimador proposto em 4.21 é mais sensível às mudanças nos coeficientes do modelo do que o estimador 4.13. Isto se deve ao fato de que o parâmetro  $\lambda$  no estimador 4.21 contém características da série em diferentes tempos. O mesmo não ocorre com o estimador 4.13 que é construído através de uma combinação linear de apenas dois valores de  $\lambda$ .

A proposta de  $\lambda$  variar no tempo mostra maior coerência em relacionar vários valores ao grau de heteroscedasticidade do que apenas um único valor ou uma combinação de valores para toda série.

A idéia básica nossa é  $\lambda = \lambda(t) = f(t)$ , onde  $f(t)$  relacione o grau de não estacionariedade, características da vizinhança de  $t$ , o tamanho da série  $T$  entre outros fatores que possam melhorar o estimador. O que deixa dúvidas e questões em aberto é quanto à forma de se calcular e propor  $f(t)$ . As idéias propostas aqui podem ser discutidas e melhoradas. Entretanto não há dúvidas no que diz respeito do ganho de informação do estimador 4.21 e na maior sensibilidade aos coeficientes do modelo quando utilizamos  $\lambda(t)$  no lugar de  $\lambda$ .

# Capítulo 5

## Aplicação

O objetivo deste capítulo consiste na aplicação da abordagem dos modelos  $tvAR$  e  $AR$  na manipulação de séries de cotações de ações e na investigação da habilidade desses modelos em prever os preços futuros das ações. Os métodos de previsão para modelo  $tvAR$  são os considerados e estabelecidos no capítulo 3. As previsões são comparadas com aquelas obtidas pelos métodos convencionais de previsão do modelo Auto Regressivo.

### 5.1 Definição do Problema

O propósito desses exemplos consiste na identificação de não estacionariedade presente em séries de retornos dos preços de ações após o ajuste de um modelo  $tvAR$ . Assumindo que a não estacionariedade possa ser detectada pelos coeficientes do modelo  $tvAR$ , e que elas persistem na série histórica das cotações das ações, é possível realizar a previsão dos valores futuros dessas ações, investigando então, a habilidade preditiva dos métodos propostos para as séries temporais econômicas irregulares, normalmente observados no mercado acionário.

O software R foi utilizado para efetuar a modelagem do  $tvAR$ . Para esse aplicativo foram desenvolvidos funções para ajustar modelos  $tvAR(p)$  e fazer previsões para um horizonte  $h$ . As funções desenvolvidas estão disponíveis e explicadas no apêndice A.

#### 5.1.1 Exemplo 1: Dados diários

Como primeiro exemplo, analisando a série é composta de cotações da companhia Vale do Rio Doce. Os valores utilizados são o de fechamento do pregão no período de 03/11/2003 a 28/12/2007, correspondendo a 1050 observações diárias.



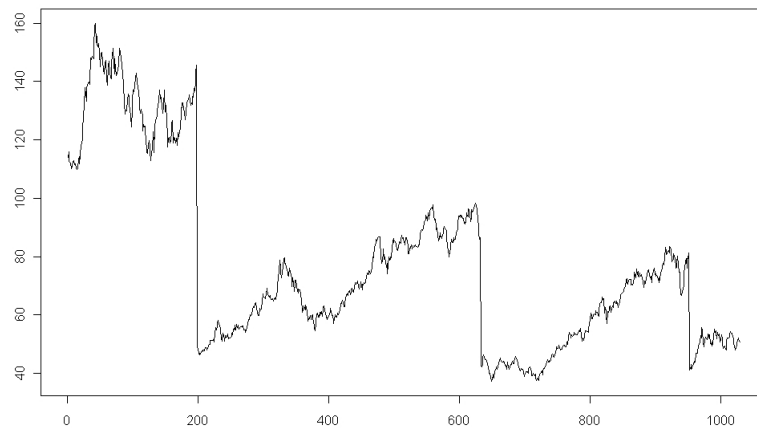


FIGURA 5.1: *Série de fechamento do pregão da companhia Vale do Rio Doce.*

De acordo com o procedimento usual, são utilizados os retornos logarítmicos em toda a nossa análise.

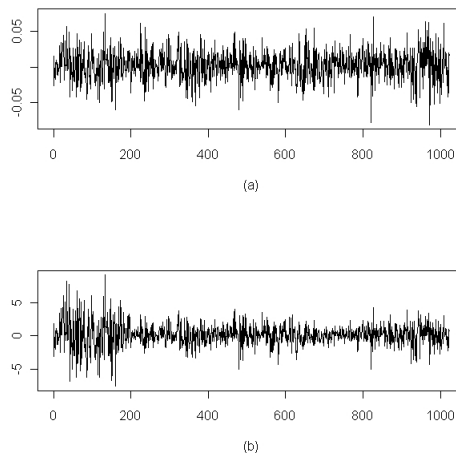


FIGURA 5.2: (a) *Série de retornos logarítmicos.* (b) *Série de retornos simples.*

Para a série de retornos foram retiradas três observações outliers e o conjunto de dados limpo foi dividido em dois, um para o ajuste dos modelos e outro para a validação das previsões.

Para escolha do número de parâmetros foi construído o gráfico da Função de Auto Correlação (ACF) da série de retornos que está na Figura 5.3.

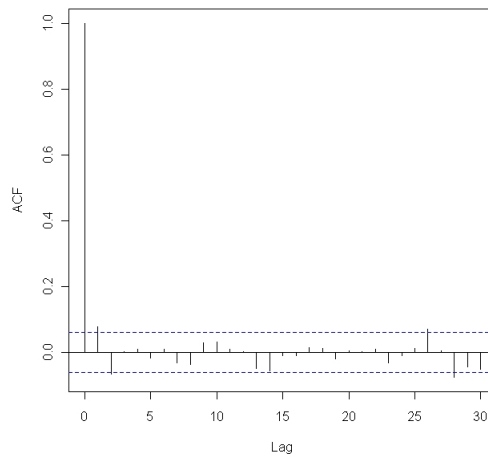


FIGURA 5.3: *Função de Auto Correlação da série de retornos logarítmicos.*

Observando a Figura 5.3, parece não existir correlação. Os modelos ajustados foram um  $tvAR(2)$  e um  $AR(2)$  e as estimativas estão na Figura 5.4 e na Tabela 5.1 respectivamente.

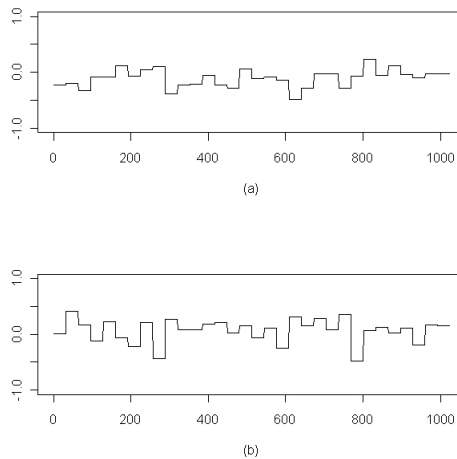


FIGURA 5.4: (a) *Estimativas da função  $a_1(t)$*  e (b) *Estimativas da função  $a_2(t)$ .*

TABELA 5.1: EQM para os modelos.

	$a_1$	$a_2$	$EQM$
$AR(2)$	-0.088	0.067	0.00045
$tvAR(2)$	(a)	(b)	0.00042

Observamos na Tabela 5.1 que o EQM foi menor no modelo  $tvARCH(2)$  e maior no modelo clássico  $AR(2)$ .

A Tabela 5.2 mostra o Erro da Previsão (EP) um passo à frente para os métodos sugeridos no capítulo 3 para o modelo  $tvAR$  e o erro da previsão para o modelo AR.

TABELA 5.2: EP para os modelos e métodos.

	$tvAR_1(2)$	$tvAR_2(2)$	$AR(2)$
$EP$	0.0101	0.0108	0.0097

Observamos na Tabela 5.2 que o EP foi maior nos métodos sugeridos no capítulo 3. O EP foi menor no modelo clássico  $AR(2)$  mesmo apresentando maior EQM. Outras discussões estão no final dessa seção.

### 5.1.2 Exemplo 2: Dados de alta frequência

O segundo exemplo está relacionado com a modelagem  $tvAR$  no contexto de dados de alta frequência. Nessa aplicação vamos utilizar o modelo  $tvAR$  para identificar mudança estrutural da série. A série avaliada é a taxa de câmbio entre a Libra Britânica e o Dólar Americano (GBP vs USD). O período observado foi de 29/04/2009 à 21/05/2009, com uma

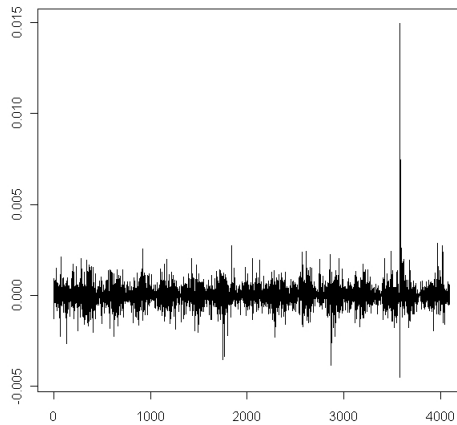
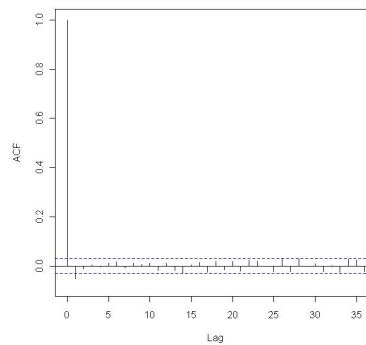


FIGURA 5.5: *Série cambial entre GBP vs USD.*

observação a cada 5 minutos, resultando em um total de 4096 observações. Novamente foram calculados os retornos logarítmicos. A Figura 5.6 mostra o retorno para a série GBP vs USD.

Para verificar a existência de correlação na série de retorno GBP vs USD, observamos a ACF apresentada na Figura 5.7.

Observando a Figura 5.7, não fica claro a presença de correlação na série e esse fato pode estar associado à quantidade de dados. Querendo identificar uma possível mudança na

FIGURA 5.6: *Série de retorno para GBP vs USD.*FIGURA 5.7: *Função de Auto Correlação (ACF) para série de retorno.*

estrutura da série ajustamos um modelo  $tvAR(1)$ .

Analisando a Figura 5.8, verifica-se que a função  $a_1$  parece ter um salto na metade das observações, ou seja, a função oscila em torno de uma média para a primeira parte que difere para a segunda parte. Isso fica mais evidente com um *threshold* nos coeficientes *wavelets*. Considerando essa mudança no comportamento da função  $a_1$ , dividimos a análise em dois conjuntos, amostra completa que é a serie total dos retornos e amostra reduzida como sendo a metade final da série de retornos. Calculamos novamente a ACF, agora para a amostra reduzida o resultado está na Figura 5.9.

Pela Figura 5.9 fica claro que os dados são correlacionados e que a estrutura da amostra reduzida não é a mesma que da amostra completa.

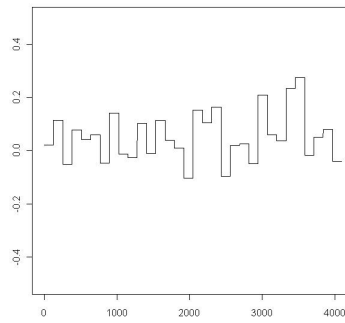


FIGURA 5.8: *Estimativa da função  $a_1$  do modelo tvAR(1).*

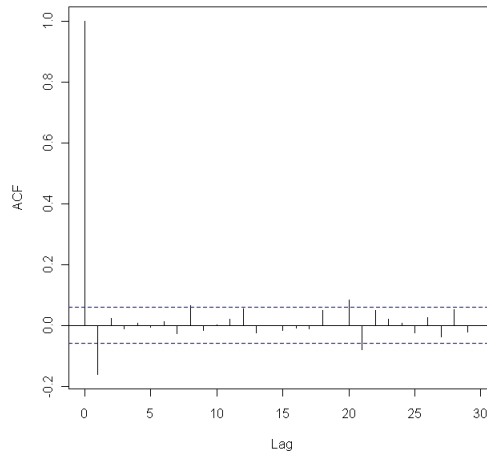


FIGURA 5.9: *Função de Auto Correlação (ACF) para amostra reduzida.*

## 5.2 Resultados e Discussão

O objetivo desse capítulo foi a apresentação de aplicações dos modelos tvAR em dados reais. No primeiro exemplo observa-se que o modelo tvAR se ajusta melhor aos dados, mostrando um EQM menor em relação ao modelo AR. Este resultado era esperado dados as características do modelo tvAR e o fato do seu estimador ser baseado em *wavelets*. Um resultado surpreendente foi o EP ser maior para os métodos sugeridos no capítulo 3 em relação à previsão do AR. Uma possível justificativa para esse resultado, visto que no capítulo 3 os resultados foram melhores para os métodos sugeridos é o fato da aplicação logarítmica na série estabilizar a variância o que é visto na Figura 5.2. Lembrando que uma das características dos modelos tvAR é a capacidade de modelar séries localmente estacionárias.

No segundo exemplo utilizamos o modelo tvAR para identificar mudanças estruturais em dados reais e o resultado foi muito bom se compararmos as ACF das amostras reduzida e completa. Esse problema poderia ser explorado de uma forma diferente como sendo a seleção

---

de um conjunto de dados para aplicação de algumas técnicas. Como já citado em dados de alta frequência existe muita informação disponível e não sabemos como selecionar o banco de dados.

Nessa mesma linha de pesquisas sugerindo outras técnicas para modelar dados de alta frequência e comparar com os métodos tradicionais citamos o trabalho de Kim et al. (2007), como referência e a apoio as nossas aplicações.

# Capítulo 6

## Conclusões e Propostas Futuras

O presente trabalho teve como objetivo principal apresentar os modelos localmente estacionários, sugestões para esses modelos e aplicações a dados financeiros.

Os modelos tvAR se mostraram bastante promissores para modelar séries com características não estacionárias, que estão presente em séries financeiras. Com relação à aplicação em dados reais seu ajuste foi superior e pode ser uma possível alternativa aos modelos AR.

As contribuições em métodos de previsão, comparação de ajuste em relação aos modelos AR, (Souza et al., 2009), criação de banda de confiança para os parâmetros do modelo foram passos importantes para viabilizar o uso dos modelos tvAR. Outro ponto importante do trabalho foi a criação de funções para ajuste, criação de banda de confiança e previsão para os modelos tvAR.

Por outro, para os modelos tvARCH estudamos o estimador recursivo com base em algoritmos LMS, onde existe o problema de convergência do parâmetro  $\lambda$ . Para esse problema sugerimos uma modificação no estimador para obter uma taxa de convergência mais rápida. Além disso, relacionamos essas convergência com o grau de heteroscedasticidade da série. Os resultados indicam um desempenho melhor do estimador.

Alguns resultados obtidos no presente trabalho ainda necessitam de estudos mais aprofundados no sentido teórico e empírico para conclusões.

Para os modelos tvARCH, os resultados obtidos da análise em dados reais não foram satisfatório como os obtidos nos estudos de simulação. Em nossas simulações, constatamos a existência de sensibilidade do estimador proposto por Dahlhaus et al. (2007) com relação aos dados de entrada. Portanto, existe ainda uma necessidade de um estudo maior sobre esse problema que só foi constatado na aplicação prática do modelo.

Ainda há muito a ser feito a respeito da classe de modelos não estacionários. Espera-se

que este trabalho possa ser o primeiro passo nas pesquisas futuras com o intuito de se ampliar os conhecimentos por este adquiridos.



# Referências Bibliográficas

- [1] Benedetto, J.J., Frazier, M.W., 1994. *Wavelets: Mathematics and Applications*. Boca Raton: CRC Press.
- [2] Bruce, A.G., Gao, H., 1995. *Waveshrink: Shrinkage functions and thresholds*. In Andrew F. Laine and Michael A. Unser, editores, *Wavelet Applications in Signal and Image Processing III*, San Diego, CA, July.
- [3] Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. *The Econometrics of Financial Markets*. Princeton: Princeton University Press.
- [4] Cohen, A., Daubechies, I., Vial, P., 1993. *Wavelets on the interval and fast wavelet transform*. Appl. Comput. Harmon. Anal., 1, 54-81.
- [5] Cont, R., 2001. *Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues*. Quantitative Finance, 1, 223-236.
- [6] Dahlhaus, R., 1997. *Fitting time series models to nonstationary processes*. Ann. Statist., 25, 1-37.
- [7] Dahlhaus, R., Neumann, M.H., von Sachs, R., 1999. *Nonlinear wavelet estimation of timevarying autoregressive processes*. Bernoulli, 5, 873-906.
- [8] Dahlhaus, R., Rao, S.S., 2006. *Statistical inference for time-varying ARCH processes*. Ann. Statist, 34, 1075-1114.
- [9] Dahlhaus, R., Rao, S.S., 2007. *A recursive online algorithm for the estimation of time-varying ARCH parameters*. Bernoulli, 13, 389-422.
- [10] Davison, A.C., Hinkley, D.V., 1997. *Bootstrap Methods and Their Application*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge.
- [11] Donoho, D.L., 1993. *Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data*. In Ingrid Daubechies, editor, *Proceedings os the Symposia in Applied Mathematics*. American Mathematical Society.
- [12] Donoho, D.L., Johnstone, I.M., 1994. *Ideal spatial adaptation via wavelet shrinkage*. Biometrika, 81, 425-455.
- [13] Donoho, D.L., Johnstone, I.M., 1995a. *Adapting to unknown smoothness via wavelet shrinkage*. Jornal of the American Statistical Association, 90, 1200-1224, December.
- [14] Donoho, D.L., Johnstone, I.M., 1998. *Minimax estimation via wavelet shrinkage*. Ann. Statist., 26, 879-921.
- [15] Donoho, D.L., Johnstone, I.M., Kerkyacharian, G., Picard, D., 1995b. *Wavelet Shrinkage: Asymptopia?* Journal Royal Statistical Society, B, 57, 301-369.
- [16] Donoho, D.L., Johnstone, I.M., Kerkyacharian, G., Picard, D., 1996. *Universal Near Minimality of Wavelet Shrinkage*. Festschrift for Lucien Le Cam. Springer Verlag.
- [17] Efron, B., Tibshirani, R.J., 1993. *An Introduction to the Bootstrap*. Volume 57 of Monographs on Statistics and Applied Probability. Chapmann Hall, New York, p. 456.

- [18] Ehlers, R.S., 2006. *Análise de Séries Temporais*. Departamento de Estatística, UFPR.
- [19] Engle, R.F., 1982. *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation*. *Econometrica* 50, 987-1007.
- [20] Harris, L., 1986. *A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns*. *Journal of Financial Economics*, Vol. 16, 99-117.
- [21] Haykin, S., 2001. *Adaptive Filter Theory*. Third Edition, Englewood Cliffs, NJ: Prentice-Hall.
- [22] Honig, M.L., Messerschmitt, David G. 1987. *Adaptive Filters: Structures, Algorithms and Applications*. *The Journal of the Acoustical Society of America*, 85, 356.
- [23] Kabaila, P., Syuhada, K., 2008. *Improved Prediction Limits for AR(p) and ARCH(p) Processes*. *Journal of Series Analysis*, Vol. 29, 2, 381-397.
- [24] Kim, K., Yoon, S., Kim, S., Chang, K., Kim, Y., 2007. *Dynamical structures of high-frequency financial data*. *Physica A*, 376, 525-531.
- [25] Lim, J.S., Oppenheim, A.V., 1987. *Advanced Topics in Signal Processing*. Prentice Hall, p. 518.
- [26] Meyer. Y., 1993. *Wavelets: Algorithms and Applications*. Society for Industrial and Applied Mathematics, Philadelphia, 13-31.
- [27] Mikosch, T., Starica, C., 2004. *Non-stationarities in financial time series, the longrange dependence, and the IGARCH effects*. *Rev. Econometrics Statist.*, 86, 378-390.
- [28] Moreira, J.M.S., Lemgruber, E.F., 2004. *O uso de dados de alta frequência na estimação da volatilidade e do valor em risco para o IBOVESPA*. *Revista Brasileira de Economia*, Jan./Mar., Vol. 58, 1, 100-120.
- [29] Morettin, P.A., 1999. *Ondas e Ondaletas: Da Análise de Fourier à Análise de Ondaletas*. São Paulo: Editora da Universidade de São Paulo.
- [30] Morettin, P.A., 2006. *Econometria Financeira: Um Curso em Séries Temporais Financeiras*. São Paulo: Associação Brasileira de Estatística.
- [31] Morettin, P.A., Toloí, C.M.C., 2006. *Análise de Séries Temporais*. 2 ed. São Paulo: Editora Blucher, vol. 1, p. 535.
- [32] Palma, W., 2007. *Long-Memory Time Series*. John Wiley & Sons, New Jersey.
- [33] Percival, D.B., Walden, A.T., 2000. *Wavelet Methods for Time Series Analysis*. Cambridge, England: Cambridge University Press.
- [34] Souza, L.T.L., Moura, M.S.A., 2009. *Comparação entre modelos Auto Regressivos e Auto Regressivos com coeficientes variando no tempo*. In: *Escola de Séries Temporais e Econometria*, 13, São Carlos.
- [36] Wheelwright, S.C., Makridakis, S., 1989. *Forecasting Methods for Management*. Fifth Edition. New York: John Wiley Sons Inc, p. 470.
- [36] Wood, R.A., McInish, T.H., Ord, J.K., 1985. *An Investigation of Transaction Data for NYSE Stocks*. *Journal of Finance*, 25, 723-739.

# Apêndice A

## Wavelets

### A.1 Visão geral

*Wavelets* são funções que satisfazem determinados requisitos e são utilizadas para decompor e representar outras funções.

As transformadas de *wavelets* podem ser vistas como mecanismos para decompor ou quebrar sinais ou séries temporais nas suas partes constituintes, permitindo analisar os dados em diferentes domínios de frequências com a resolução de cada componente amarrada à sua escala. Além disso, na análise de *wavelets*, podemos usar funções que estão contidas em regiões finitas, tornando-as convenientes na aproximação de dados com descontinuidades. A idéia de aproximar funções por uma combinação de outras funções não é nova, mas vem de 1800 quando Joseph Fourier descobriu que poderia combinar senos e cossenos para representar outras funções.

Os algoritmos de *wavelets* processam dados em diferentes escalas ou resoluções e independentemente da função de interesse, *wavelets* oferecem uma técnica elegante na representação dos níveis de detalhes presentes em diferentes escalas e resoluções.

Se olharmos para um sinal com uma grande “janela”, gostaríamos de ver grandes características desse sinal. Do mesmo modo, se olharmos para um sinal com uma pequena “janela”, gostaríamos de observar pequenas características. Poderíamos dizer grosso modo que o objetivo na análise *wavelet* é “ver a floresta e as árvores”.

Existem inúmeras aplicações para o uso de *wavelets* entre as quais destacamos compressão de dados, análise funcional em matemática, singularidades ou oscilações locais de funções, soluções de equações diferenciais, reconhecimento de padrões, biologia, medicina, astronomia, acústica, problemas de computação gráfica, reconhecimento de voz, visão humana.

## A.2 Fatos Históricos

As origens da teoria *wavelet* remontam aos anos de 1930, quando, ainda sem corpo, podia ser discretamente identificada em trabalhos de Análise Funcional e outros ramos matemáticos. Os esforços separados não parecem ser parte de uma teoria coerente.

Antes de 1930, o principal ramo da matemática conducente a *wavelets* começou com Joseph Fourier (1807), com suas teorias sobre análise de frequência, conhecida como Análise de Fourier. Ele afirmou que qualquer função  $f(x)$ ,  $2\pi$  - periódica pode ser dado pela série (de Fourier) definida em A.1.

$$f(x) = a_0 + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx)) \quad (\text{A.1})$$

Os coeficientes  $a_0$ ,  $a_k$ , e  $b_k$  são dados por A.2

$$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_0^{2\pi} f(x) dx, \\ a_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx, \\ b_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx \end{aligned} \quad (\text{A.2})$$

O primeiro relato sobre *wavelets* foi feito em 1909, por A. Haar, mas a *wavelet* de Haar ficou no anonimato por muitos anos. Uma propriedade da *wavelet* de Haar é que possui suporte compacto, o que significa que ela desaparece fora de um intervalo finito. A *wavelet* de Haar não é continuamente diferenciável algo que limita sua aplicação e por um período muito longo, ela continuou ser a única *wavelet* conhecida.

Na década de 30, vários pesquisadores trabalharam em representar funções em bases de função escala. Compreender os conceitos de bases de função escala é fundamental para a compreensão de *wavelets* e esse assunto será mais bem definido adiante.

Utilizando base de função escala associada a *wavelet* de Haar (mais detalhes adiante) Paul Levy (1930), um físico, investigou o movimento browniano, um tipo de sinal aleatório. Ele concluiu que funções base Haar são superiores às funções base de Fourier para o estudo de pequenos detalhes complicados no movimento Browniano.

Outras pesquisas realizadas em 1930 por Littlewood, Paley, e Stein envolviam o cálculo da energia<sup>3</sup> de uma função  $f(x)$ :

$$\text{energia} = \frac{1}{2} \int_0^{2\pi} |f(x)|^2 dx \quad (\text{A.3})$$

O cálculo produzia resultados diferentes se a energia estava concentrada em torno de um ponto ou em alguns pontos distribuídos ao longo de um intervalo maior. Este resultado

<sup>3</sup>O termo energia vem da física e refere-se à quantidade  $E_{\mathbf{X}} = \|\mathbf{X}\|^2$  como sendo a energia de uma série temporal  $\{X_t\}$ .

perturbava porque os cientistas indicaram que a energia não poderia ser conservada. Os pesquisadores descobriram uma função que poderia variar em escala e poderia conservar a energia quando calculada a energia da função. Esse trabalho foi realizado por David Marr com um algoritmo efetivo para processar imagens usando *wavelets* no início de 1980.

Em 1985, Stephane Mallat deu uma significativa contribuição para o uso de *wavelets* através do seu trabalho em processamento de sinal digital. Ele descobriu algumas relações entre quadratura e filtros, algoritmos piramidais, e base *wavelets* ortonormal (assunto que será discutido mais tarde). Inspirado em parte por estes resultados, Y. Meyer construiu a primeira *wavelet* não-trivial. Ao contrário da *wavelet* de Haar, a *wavelet* de Meyer é continuamente diferenciável, porém ela não tem suporte compacto. Alguns anos mais tarde, Ingrid Daubechies Mallat utilizando o trabalho na construção de um conjunto de funções base *wavelet* ortonormal, fez grandes avanços onde se tornou referência para aplicações *wavelets* de hoje. Para discussões mais elaboradas ver Meyer, (1993) e Percival et al. (2000).

### A.3 Introdução

$L^2(\mathfrak{R})$  é um espaço vetorial normado, com a norma quadrática  $\|f\|^2$  definida a partir do produto interno de dois elementos de  $L^2(\mathfrak{R})$ . Para mais detalhes ver Benedetto et al. (1994).

Sejam  $f$  e  $g$  dois elementos de  $L^2(\mathfrak{R})$ . O produto interno de  $f$  e  $g$  é dado por

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx \quad (\text{A.4})$$

A norma quadrática de  $f$  é dado por

$$\|f\|^2 = \langle f, f \rangle = \int_a^b f^2(x)dx \quad (\text{A.5})$$

Definição 1: Uma *wavelet* é uma função  $\psi(t) \in L^2(\mathfrak{R})$ , tal que a família de funções

$$\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k)$$

onde  $j$  e  $k$  são inteiros arbitrários, seja uma base ortonormal para  $L^2(\mathfrak{R})$ .

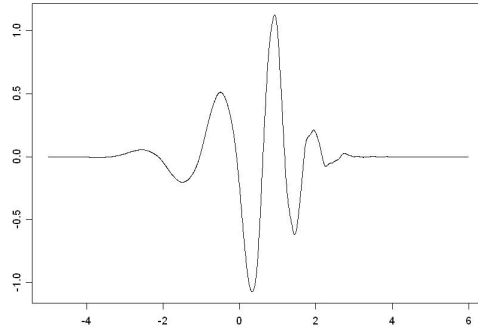


FIGURA A.1: *Exemplo de uma função wavelet de Daubechies.*

Da definição 1 segue que se  $\psi(t)$  é uma *wavelet*, então,  $\psi_{j,k}$  também o será para qualquer  $j, k \in Z$ . Além disso, a base ortonormal é construída a partir de translações e dilatações da função  $\psi(t)$ , que é conhecida como *wavelet* mãe. As funções  $\{\psi_{j,k}, j, k \in Z\}$  são ortonormais em  $L^2(\mathfrak{R})$ , isto é,

$$\langle \psi_{j,k}, \psi_{l,m} \rangle = \delta_{j,l} \delta_{k,m}, \quad j, k, l, m \in Z. \quad (\text{A.6})$$

Os detalhes para construção e verificação que  $\{\psi_{j,k}, j, k \in Z\}$  forma uma base ortonormal para  $L^2(\mathfrak{R})$  podem ser vistos em Benedetto et al. (1994). Toda função  $f(t) \in L^2(\mathfrak{R})$  pode ser dada por uma combinação linear de  $\psi_{j,k}$ .

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} d_{j,k} \psi_{j,k}(t). \quad (\text{A.7})$$

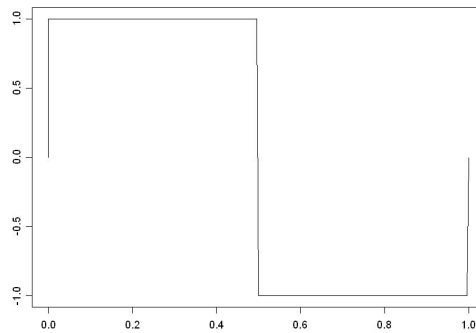
Dizemos que A.7 é uma série de *wavelets* de  $f(t)$  e os coeficientes de *wavelets* são dados por

$$d_{j,k} = \langle f, \psi_{j,k} \rangle = \int_{-\infty}^{\infty} f(t) \psi_{j,k}(t) dt. \quad (\text{A.8})$$

O exemplo mais antigo de *wavelet* foi dado por Haar e consiste na função definida por

$$\psi(t) = \begin{cases} 1, & \text{se } t \in [0, \frac{1}{2}) \\ -1, & \text{se } t \in [\frac{1}{2}, 1) \\ 0, & \text{caso contrário.} \end{cases}$$

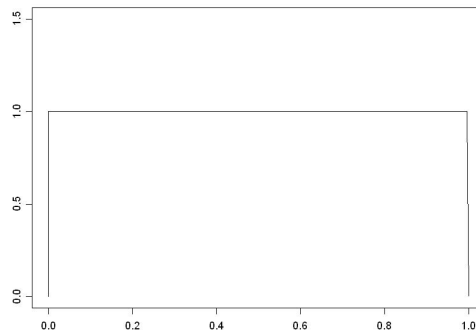
O gráfico da *wavelet* de Haar é dado na Figura A.2.

FIGURA A.2: *Gráfico da wavelet de Haar.*

A função

$$\phi(t) = \begin{cases} 1, & \text{se } 0 \leq t \leq 1 \\ 0, & \text{caso contrário,} \end{cases}$$

desempenha um papel muito importante no que se segue. Seu gráfico aparece na Figura A.3, ela é denominada função escala associada a wavelet de Haar.

FIGURA A.3: *A função escala,  $\phi$ , associada a wavelet de Haar.*

## A.4 Análise de Multirresolução

Formulada em 1986 por Mallat e Meyer, a Análise de Multirresolução fornece um referencial onde bases de *wavelets* são naturalmente compreendidas, bem como permite a construção de novas bases. Quando Mallat trabalhou com *wavelets* de Meyer pela primeira vez, ele estava trabalhando com análise de imagens, uma área de pesquisa em que o estudo de imagens em várias escalas era comum. Isto o estimulou a ver bases ortonormais de *wavelets* como

uma ferramenta para descrever matematicamente o “incremento na informação” necessário para se ir de uma aproximação grosseira para uma aproximação com maior resolução.

Definição 2: Uma Análise de Multirresolução (AMR) é uma seqüência,  $\{V_j\}_{j \in \mathbb{Z}}$ , de subespaços vetoriais fechados normados de  $L^2(\mathbb{R})$ , representando os sucessivos níveis de resoluções, satisfazendo as seguintes condições:

$$\dots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \dots; \quad (\text{A.9})$$

$$\bigcap_{j \in \mathbb{Z}} V_j = \{0\}; \quad (\text{A.10})$$

$$\bigcup_{j \in \mathbb{Z}} V_j \text{ é denso em } L^2(\mathbb{R}); \quad (\text{A.11})$$

$$f(x) \in V_j \Leftrightarrow f(2^j x) \in V_0; \quad (\text{A.12})$$

$$f(x) \in V_0 \Rightarrow f(x - n) \in V_0, \text{ para todo } n \in \mathbb{Z}; \quad (\text{A.13})$$

Existe  $\phi \in L^2(\mathbb{R})$ , denominada função escala, tal que

$$\{\phi(x - n)\}_{n \in \mathbb{Z}} \text{ constitui uma base ortonormal para } V_0. \quad (\text{A.14})$$

A idéia que está por trás das 3 primeiras relações é a seguinte. Temos uma função  $f \in L^2(\mathbb{R})$ , e queremos obter vários níveis de resolução. Cada aproximação da função ou resolução, será feita através da projeção ortogonal de  $f$  sobre  $V_j$ . O fato de  $V_j \subset V_{j-1}$  significa que ganhamos informações ao passar do nível  $j$  para o nível  $j - 1$ .

A propriedade A.12 expressa que todos os espaços estão relacionados por escala a um mesmo espaço  $V_0$ . As condições 2.14 e 2.16 implicam que  $\{\phi_{j,n}\}_{j,n \in \mathbb{Z}}$  é uma base ortonormal para  $V_j$  para todo  $j \in \mathbb{Z}$ .

A condição A.12 significa também que nenhuma escala é privilegiada. Os detalhes adicionais necessários para aumentar a resolução de  $2^j$  para  $2^{j-1}$  são dados pela projeção de  $f$  sobre o complemento ortogonal de  $V_j$  em relação a  $V_{j-1}$ , que denotado por  $W_j$ :  $V_j \oplus W_j = V_{j-1}$ .

Em virtude das condições 2.12 e 2.13, o espaço todo,  $L^2(\mathbb{R})$ , pode ser decomposto como uma soma direta de subespaços mutuamente ortogonais, os  $W'_j$ s. Além disso, os subespaços  $W_j$  herdam a propriedade de escala 2.14 de  $V_j$ :

$$f(x) \in W_j \Leftrightarrow f(2^j x) \in W_0; \quad (\text{A.15})$$

Por causa de A.15, se  $\{\psi(x - k)\}_k$  for uma base ortonormal para  $W_0$ , então,  $\{\psi_{j,k}(x)\}_k$  será uma base ortonormal para  $W_j$ .

Das definições 1 e 2, convém considerar o sistema ortonormal

$$\{\phi_{j,k}, \psi_{j,k}, j, k \in \mathbb{Z}\}, \quad (\text{A.16})$$



de modo que qualquer  $f(t) \in L^2(\mathfrak{R})$ , passa ser dada como em A.19

$$f(t) = \sum_{j,k} d_{j,k} \psi_{j,k}(t) \quad (\text{A.17})$$

$$= \sum_k c_{0,k} \phi_{0,k}(t) + \sum_{j \geq 0} \sum_k d_{j,k} \psi_{j,k}(t) \quad (\text{A.18})$$

$$= \sum_k c_{j_0,k} \phi_{j_0,k}(t) + \sum_{j \geq j_0} \sum_k d_{j,k} \psi_{j,k}(t), \quad (\text{A.19})$$

onde

$$c_{j,k} = \int_{-\infty}^{\infty} f(t) \phi_{j,k}(t) dt, \quad (\text{A.20})$$

$$d_{j,k} = \int_{-\infty}^{\infty} f(t) \psi_{j,k}(t) dt. \quad (\text{A.21})$$

para algum inteiro  $j_0$ . Para mais detalhes sobre AMR ver Morettin, (1999) e Percival et al. (2000).

## Transformação de *Wavelet*

Para  $f \in L^2(\mathfrak{R})$ , a transformada de *wavelet*, com respeito a  $\psi$ , é definida por

$$G_f(a, b) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) dt. \quad (\text{A.22})$$

que é uma função de dois parâmetros reais,  $a$  e  $b$ , onde  $a$  se refere a escala ou resolução e  $b$  é um parâmetro de translação ou localização da função *wavelet* mãe. Se definirmos  $\psi_{a,b}(t)$  como:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right), \quad (\text{A.23})$$

Podemos reescrever a transformada como o produto interno das funções  $f(t)$  e  $\psi_{a,b}(t)$ :

$$G_f(a, b) = \langle f(t), \psi_{a,b}(t) \rangle = \int_{-\infty}^{\infty} f(t) \psi_{a,b}(t) dt. \quad (\text{A.24})$$

Reciprocamente, conhecida a função  $G_f(a, b)$ , pode se obter em correspondência a função  $f(t)$  através da expressão A.25

$$IG_f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{|a|^2} G_f(a, b) \psi_{a,b}(t) da db, \quad (\text{A.25})$$

O parâmetro  $C_\psi$  é uma constante e depende da função *wavelet* escolhida. O parâmetro  $C_\psi$  necessita ser finito e positivo. Esta restrição sobre o valor de  $C$  é a condição de admissibilidade. Para maiores informações consultar Benedetto et al. (1994) e Percival et al. (2000).

## A.5 *Threshold*

A idéia básica da aplicação de um *threshold* nos coeficientes *wavelets* tem por objetivo reduzir o ruído presente em um sinal e diminuir o número de coeficientes, ou seja,

1. Aplica-se a transformada *wavelet*, obtendo-se os coeficientes *wavelets*, estes por sua vez captam desde a forma geral a detalhes do sinal. Em nossa análise os ruídos são os detalhes do sinal.
2. Escolhe-se um *threshold* para reduzir ou anular aqueles coeficientes abaixo de certo valor, obtendo novos coeficientes que descrevem apenas a forma geral do sinal, ou seja, livre de ruídos
3. Por fim, usa-se a transformada inversa de *wavelet*, com os coeficientes obtidos em 2, para recompor o sinal sem a presença do ruído.

Os dois métodos clássicos de *threshold* são conhecidos como *Hard* e *Soft*:

1. *Hard threshold*, é definido por

$$\delta_{\lambda}^H(x) = \begin{cases} 0, & \text{se } |x| \leq \lambda \\ x, & \text{se } |x| > \lambda \end{cases} \quad (\text{A.26})$$

2. *Soft threshold*, é definido por

$$\delta_{\lambda}^S(x) = \begin{cases} 0, & \text{se } |x| \leq \lambda \\ \text{sign}(x)(|x| - \lambda), & \text{se } |x| > \lambda \end{cases} \quad (\text{A.27})$$

As funções *Hard* e *Soft threshold* com  $\lambda = 3$  foram aplicadas a função identidade  $f(x) = x$  para  $x \in \mathfrak{R}$ . Os resultados estão na Figura A.4.

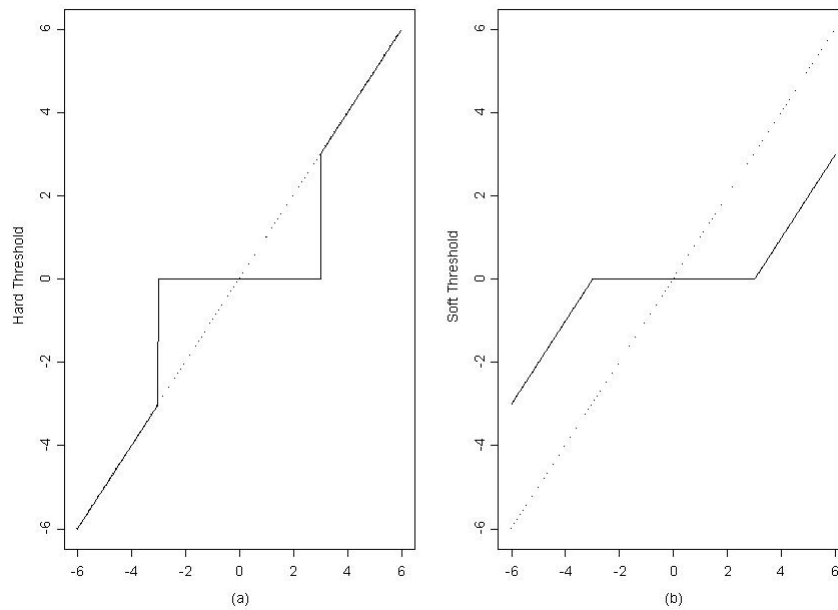


FIGURA A.4: Transformação da função  $f(x) = x$  pelas funções *Hard* e *Soft Threshold* com  $\lambda = 3$ .

O gráfico da transformada da função identidade *Hard threshold* tem pontos de descontinuidade em  $-\lambda$  e  $\lambda$ , e valores que estão entre  $-\lambda$  e  $\lambda$  são zerados. Por outro lado, o *Soft threshold* além de zerar os valores de  $f(x)$  para  $x$  entre  $-\lambda$  e  $\lambda$  modifica os demais valores da função. A motivação do uso de *Soft threshold* vem do princípio de que o ruído afeta todos os coeficientes *wavelets*.

O valor de  $\lambda$  das definições A.27 e A.26, pode ser utilizado em todos os níveis  $j$  (global) da equação A.19, ou então alguns valores de  $\lambda$  utilizados em diferentes níveis  $j$  (escala ou bloco). Alguns métodos para a escolha do parâmetro  $\lambda$  são:

- Universal

No método Universal, proposto por Donoho et al. (1994), o parâmetro  $\lambda$  é definido por

$$\lambda = \lambda_j = \sigma \sqrt{2 \log T}, \quad (\text{A.28})$$

Em A.28  $T$  representa o número de coeficientes *wavelets* e  $\sigma$  o nível de ruído que deve ser estimado a partir dos dados.

- Adaptativo (SureShrink)

Proposto também por Donoho et al. (1995a), de acordo com o procedimento SureShrink, no qual, a cada nível de resolução, se minimiza um estimador não-viesado do risco de Stein (Sure = Stein unbiased risk estimator, para detalhes ver Donoho et al., 1995a). Se no nível  $j$  tivermos  $n_j$  coeficientes, define-se os parâmetros da seguinte forma

$$\lambda_j = \operatorname{argmin}_{t \geq 0} \text{SURE}(\mathbf{d}_j, t) \quad (\text{A.29})$$

onde

$$\text{SURE}(\mathbf{d}_j, t) = n_j - 2 \sum_{k=1}^{n_j} I_{\{|d_{j,k}| \leq t\sigma_j\}} + \sum_{k=1}^{n_j} \min \{(d_{j,k}/\sigma_j)^2, t^2\} \quad (\text{A.30})$$

$\sigma_j$  é o nível de ruído para cada escala  $j$ . O segundo termo do lado direito é igual a duas vezes o número de coeficientes menores do que o limiar considerado.

Existem inúmeros métodos para calcular o parâmetro  $\lambda$  e outros tipos de *threshold*. Para maiores detalhes ver Bruce et al. (1995). Um método proposto por Dahlhaus, (1999) está descrito em 3.12.

Para entender a análise de *wavelet* apresentamos um exemplo envolvendo os conceitos vistos até o momento.

No presente trabalho estamos interessados em sinais  $f(t)$  finitos, ou seja,  $t = 1, \dots, T$ , logo a aproximação *wavelet* obtida a partir de A.19 é

$$f(t) \approx \sum_{k=1}^{n_j} s_{J,k} \phi_{J,k}(t) + \sum_{j=1}^J \sum_{k=1}^{n_j} d_{j,k} \psi_{j,k}(t) \quad (\text{A.31})$$

onde  $J$  é o número de componentes da análise de multirresolução, escala ou resolução,  $n_j$  é o número de coeficientes na escala  $j$ , os coeficientes  $s_{J,k}$  são responsáveis pela forma geral e os  $d_{j,k}$  são responsáveis pelos detalhes do sinal.

Pela construção da função  $\psi$ , vista na definição 1, quando  $T$  for divisível por  $2^J$ , existirão  $T/2$  coeficientes na escala  $d_{1,k}$ ,  $T/4$  coeficientes na escala  $d_{2,k}$  e assim sucessivamente. Na última escala existirão  $T/2^J$  coeficientes  $s_{J,k}$ .

Uma aplicação da análise de *wavelet* em regressão não paramétrica Seja

$$y_i = f(t_i) + \epsilon_i, \quad i = 1, \dots, T,$$

onde os  $\epsilon_i \sim N(0, \sigma^2)$  e  $t_i = (i - 1)/T$ , ou seja  $t_i \in [0, 1)$

Supõe-se que  $f$  pertença a determinada classe de funções, satisfazendo certas propriedades de regularidade. Se  $f_i = f(t_i)$ , o objetivo é estimar  $\mathbf{f} = (f_1, \dots, f_T)'$ , com menor erro quadrático médio. Para uma discussão mais ampla os principais trabalhos nessa área são Donoho et al. (1994, 1995a, 1995b, 1996). O problema pode também ser visto em como retirar o ruído ( $\epsilon_i$ ) presente no sinal  $y_i$ .

Exemplo 1 - Um sinal de ressonância magnética nuclear (rmn) com 1024 valores, que está representado na Figura A.5. Esse exemplo foi tirado de Donoho, (1993).

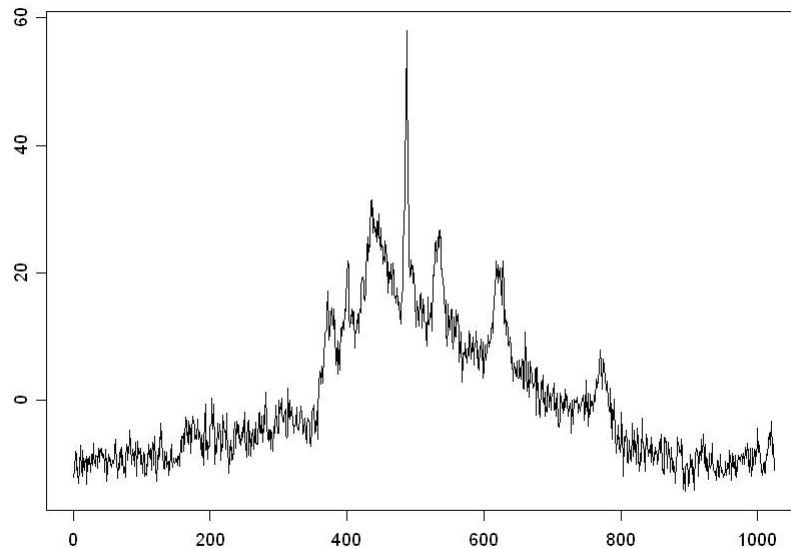


FIGURA A.5: *Série de ressonância magnética nuclear.*

No primeiro passo do ajuste do modelo, escolhe-se uma base *wavelet*  $\psi$  (Haar, Daubechies, Symmlets e Coiflets) e a quantidade de níveis de resolução  $J$  e com essas informações

aplica-se a transformada de *wavelet* nos dados. Nesse exemplo escolhemos a base de Daubechies e um nível  $J = 6$ . O resultado está na Figura A.6.

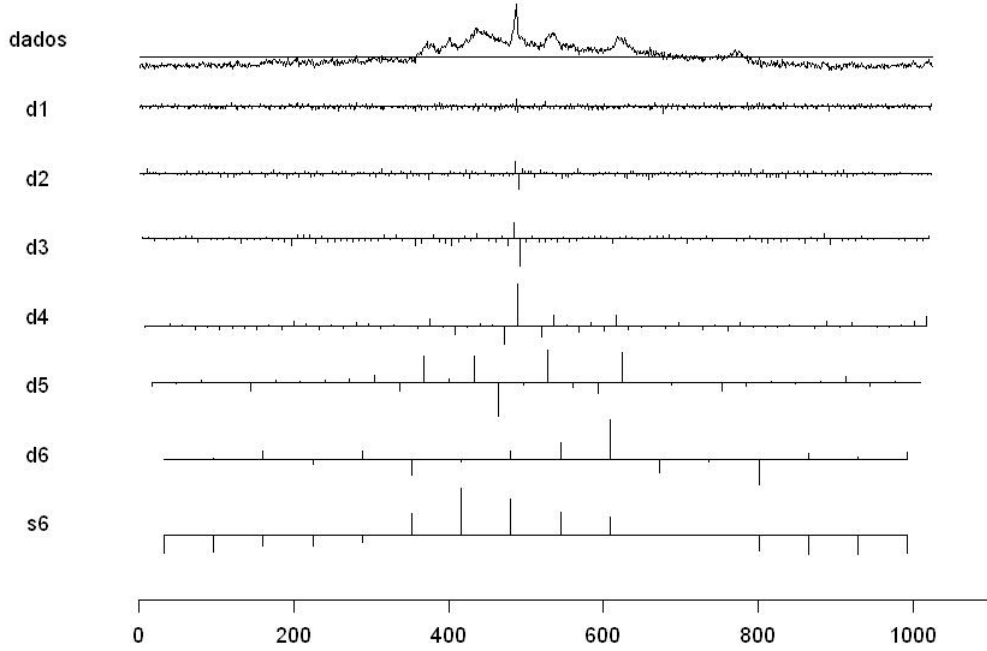


FIGURA A.6: *Coeficientes de wavelet para os dados de ressonância magnética nuclear.*

A seguir, é necessário escolher um *threshold* para remover o ruído que está presente nos coeficientes. Aplicamos o *Hard threshold* e o parâmetro  $\lambda$  calculado pelo método universal, definidos em A.26 e A.28 respectivamente. Como  $\sigma^2$  não é conhecido vamos estimá-lo através da proposta de Donoho et al. (1995a).

$$\hat{\sigma} = \frac{\text{med} \left\{ |d_{1,0}|, |d_{1,1}|, \dots, |d_{1, \frac{T}{2}-1}| \right\}}{0.6745} \quad (\text{A.32})$$

onde *med* inidica a mediana, a escala  $j = 1$  foi a escolhida por ser a mais fina e o fator 0.6745 está relacionado ao fato que  $0.6745 < \Phi(1) - \Phi(-1)$ .

Utilizando o nosso exemplo temos que  $\hat{\sigma} = 1.645$  logo  $\lambda = 6.126$ . Portanto aplicamos o *hard threshold* nas quatro primeiras escalas, com os novos coeficientes usamos a transformada inversa e temos uma estimativa da função  $f(t)$ . O resultado está na Figura A.7

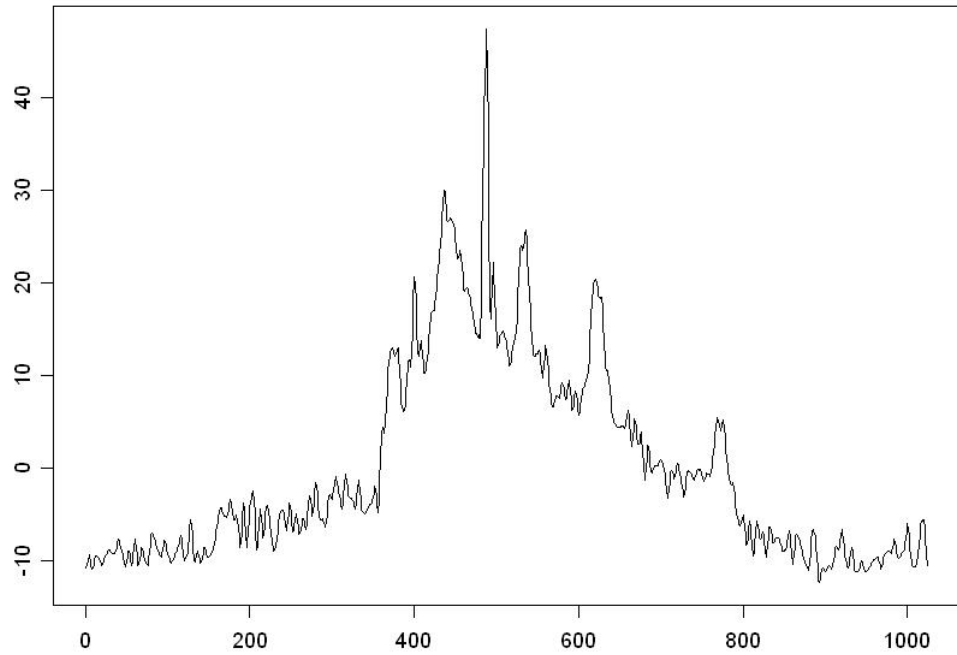


FIGURA A.7: *Estimativa do sinal  $rmn$  com a aplicação do hard threshold e cálculo do parâmetro  $\lambda$  pelo método universal.*

Continuando com esse exemplo podemos fazer uma análise de multirresolução utilizando não apenas os coeficientes *wavelets*, mas a representação deles através da base escolhida. Chamamos de sinal suave e detalhes do sinal respectivamente as seguintes funções:

$$S_J(t) = \sum_k s_{J,k} \phi_{J,k}(t) \quad (\text{A.33})$$

e

$$D_j(t) = \sum_k s_{j,k} \psi_{j,k}(t) \quad (\text{A.34})$$

O sinal  $f(t)$  pode ser aproximado por essas funções,

$$f(t) \approx S_J(t) + D_J(t) + D_{J-1}(t) + \dots + D_1(t) \quad (\text{A.35})$$

Aplicando essas novas funções a análise de multirresolução, para o exemplo 1 temos o seguinte resultado

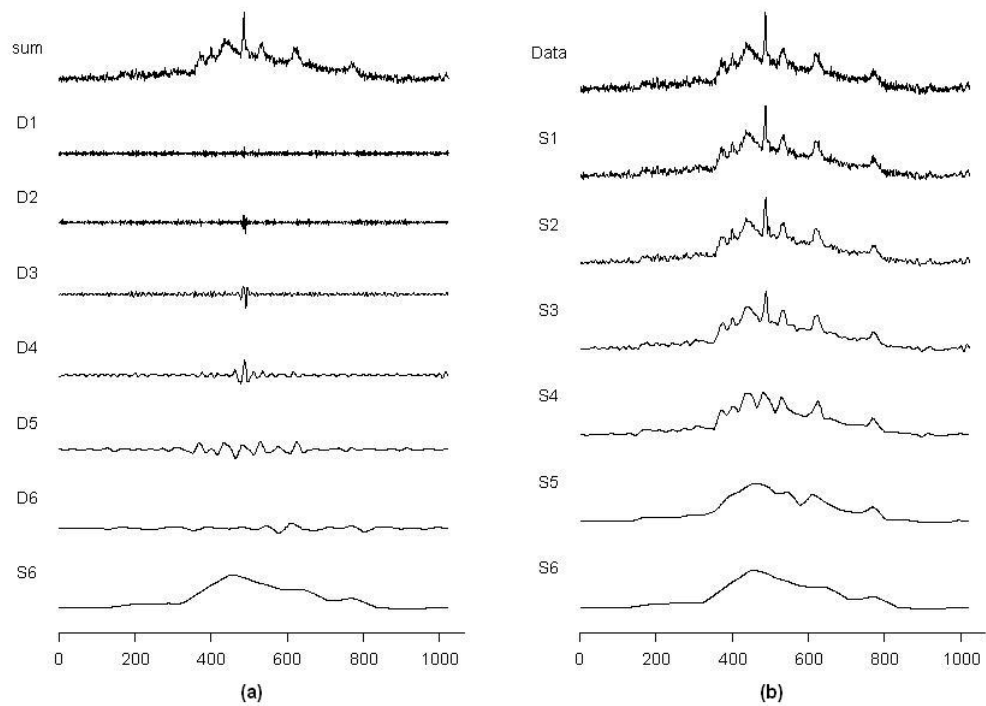


FIGURA A.8: *Análise de multirresolução do sinal rmn.*

Podemos observar que no gráfico (a) da Figura A.8 temos a função suavizada e mais os detalhes em diferentes níveis onde a nossa função  $f(t)$  será dada como a soma de todos os níveis como mostrado na Figura A.8. Por outro lado no gráfico (b) está a perda de informação da função  $f(t)$  quando passamos de uma classe mais fina para uma classe mais grosseira.

As Figuras desse capítulo e o exemplo foram feitos em S-Plus com o auxílio do pacote S-Wavelet. Para maiores aplicações e detalhes de como funcionam as *wavelets* ver Bruce et al. (1995), onde existe uma boa referência bibliográfica.

# Apêndice B

## Biblioteca R

Esse apêndice tem as principais funções desenvolvidas no software R 2.9.0 para modelagem, análise e previsão dos modelos tvAR e tvARCH.

### B.1 Ajuste de modelos $tvAR(p)$

Modelagem tvAR para Séries Temporais

Descrição:

Ajusta modelos tvAR(p) para séries temporais univariada.

Uso:

```
tvAR (x, J = 0, p = 2, wave = "haar")
```

Argumentos:

x: uma série temporal univariada. Deve ser um vetor de comprimento diádico (potência de 2).

J: especifica o nível de decomposição. Caso não especificado o valor assumido é  $\log_2(T^{1/2}) + 1$ , onde  $T$  é o tamanho da série  $x$ .

p: um número inteiro representando a ordem do modelo de ajuste.

wave: nome da base wavelet para uso na decomposição. Por definição está a wavelet de Haar, mas a base ortonormal de Daubechies pode ser usada.

Valor:



Uma lista com as seguintes componentes:

m.coef: uma matriz ( $p \times T$ ) com os valores das funções  $a_i$  do modelo para cada instante  $t$ , onde  $T$  representa o tamanho da série  $x$ .

ajuste: a série ajustada.

resíduo: a série dos resíduos.

Código R para criação da função

```
tvAR <- function(x, J=0, p=2, wave="haar") {
  if (J==0) J=as.integer(logb(sqrt(length(x)),base=2))+1
  pot=logb(length(x), base = 2) # potencia da série
  N=length(x) # tamanh da série
  n.coef= 2^(J-1) # 2^(j-1) no de coeficientes
  delta=p*n.coef # J no de escala
  # p no de parametros do AR(p)

  library(waveslim)
  library(wavethresh)
  matrix.coef=function(k,n.lvl,wav,escala){
    vetor=rep(0,N)
    vetor[k]=1
    vetor.Psi=numeric()
    wav.dwt=dwt(vetor,wf=wav,n.levels=n.lvl)
    ind=seq(n.lvl+1,((n.lvl+2)-escala),-1)
    for (i in ind){
      vetor.Psi=c(vetor.Psi,wav.dwt[[i]])
    }
    vetor.Psi=matrix(vetor.Psi)
    vetor.Psi
  }

  vetor.indice=seq(1,N,1)
  Psi=apply(as.array(vetor.indice),1,matrix.coef, n.lvl=pot, wav=wave, escala=J)
  xp=matrix(0,N,p)
  for (i in 1:p) xp[((i+1):N),i]=x[1:(N-i)]
  Xwave=matrix(0,N,p*n.coef)
  for (i in 1:delta){
    Xwave[,i]=Psi[((i-1)%n.coef+1),]*xp[,((i-1)%n.coef+1)]
  }
  Infor=lsfit(Xwave,x,intercept=FALSE)
  Beta=Infor$coef
  Beta=as.matrix(Beta)
  sigBeta<-ls.diag(Infor)
  sigBeta1<-sigBeta$cov.scaled
  Beta.sig<-rep(0,delta)
  for (i in 1:delta) Beta.sig[i]<-sigBeta1[i,i]
  a=matrix(0,n.coef,p)
  for (i in 1:p) a[,i]=Beta[((i-1)*n.coef+1):(i*n.coef)]
  f=matrix(0,p,N)
```

```

for (i in 1:p)
f[i,]=-t(Psi)%*%a[,i]
a.f=a
lambda=matrix(0,J-2,p)
for (i in 1:p){
  for (k in 1:(J-2)) lambda[k,i]=(mad(a[(2^(k)+1):(2^(k+1)) ,i])/0.6745)*sqrt(2*log(2^(J-1)))
    for (m in 1:(J-2)) {
      for (n in (2^(m)+1):(2^(m+1))) {
        if (abs(a.f[n,i]) < lambda[m,i] ) a.f[n,i]=0
      }
    }
}
f.f=matrix(0,p,N)
for (i in 1:p)
f.f[i,]=-t(Psi)%*%a.f[,i]
f.f.menos=matrix(0,p,N)
for (i in 1:p)
f.f.menos[i,]=-t(Psi)%*%(a.f[,i]-2*max(lambda[,i])/sqrt(2*log(2^(J-1))))
f.f.mais=matrix(0,p,N)
for (i in 1:p)
f.f.mais[i,]=-t(Psi)%*%(a.f[,i]+2*max(lambda[,i])/sqrt(2*log(2^(J-1))))
k=nrow(f)
inteiro=k%%3
rest=k%%3
ini=1
graf <- function(alpha,ini,fim){
x11()
par(mfrow = c(3, 1))
for (i in ini:fim) {
  plot(alpha[i,],type='l',ylab="Alpha",xlab="Tempo", ylim = c(min(alpha[i,]),max(alpha[i,])))
}
if (inteiro!=0){
for (i in 1:inteiro){
  graf(f,ini,ini+2)
  ini= ini+3} }
if (rest>0){
  graf(f,ini,ini+rest-1)
}
k=nrow(f)
inteiro=k%%3
rest=k%%3
ini=1
if (inteiro!=0){
for (i in 1:inteiro){
  graf(f.f,ini,ini+2)
  ini= ini+3} }
if (rest>0){
  graf(f.f,ini,ini+rest-1)
}
}

```

```

yh = numeric()
yh.f = numeric()
for (i in (p+1):N){
  yh[i]=xp[i,]*%*-f[,i]
  yh.f[i]=xp[i,]*%*-f.f[,i]
}
resi=x-yh
resi.f=x-yh.f
list(m.coef=f,Beta=Beta,Beta.sig=Beta.sig,J=J,Xwave=Xwave,Psi=Psi,
m.coef.f=f.f, aj=yh, aj.f=yh.f, resi=resi,resi.f=resi.f,a=a,a.f=a.f,
m.coef.f.mais=f.f.mais,m.coef.f.menos=f.f.menos)
}

```

## B.2 Ajuste de modelos $tvARCH(p)$

Modelagem tvARCH para Séries Temporais

Descrição:

Ajusta modelos tvARCH(p) para séries temporais univariada.

Uso:

tvARCH (x, lambda1, p)

Argumentos:

x: uma série temporal univariada.

lambda1: deve ser um vetor com o valor do parâmetro  $\lambda$  em cada instante de tempo  $t$ .

w: é o peso para o calculo do valor de  $\lambda_2$  deve ser um número entre  $(0 < w < 1)$ . Se  $w = 0$  o ajuste é feito segundo a seção 4.5. E os valores de  $\lambda(t)$  podem ser obtidos a partir de uma função de medida de característica que segue logo abaixo.

p: um número inteiro representando a ordem do modelo de ajuste.

Valor:

Uma lista com as seguintes componentes:

alpha: uma matriz (p x T) com os valores das funções  $\alpha_i$  do modelo para cada instante  $t$ , onde  $T$  representa o tamanho da série  $x$ .

sig: erro quadrático médio do ajuste do modelo.

Código R para criação da função

```

tvARCH <- function(x, lambda1, w=0, p) {
lambda2 = lambda1 * w
X=x
xxx <- function(t,X,p){

```

```

xx = rep(1,(p+1))
X = X^2
xx[2:(p+1)] = X[t:(t-p+1)]
xx
}
alpha = matrix(0, p+1, length(X))
for (t in (p+1):length(X)){
alpha.1 = alpha[,t-1] + lambda1[t]*(X[t]^2 - alpha[,t-1] %*%
xxx(t-1,X,p) )*( xxx(t-1,X,p)/((sum(xxx(t-1,X,p)))^2) )
alpha.2 = alpha[,t-1] + lambda2[t]*(X[t]^2 - alpha[,t-1] %*%
xxx(t-1,X,p) )*( xxx(t-1,X,p)/((sum(xxx(t-1,X,p)))^2) )
alpha[,t] = ((1/(1-w))*(alpha.1)) - ((w/(1-w))*(alpha.2))
}
k=nrow(alpha)
inteiro=k%%3
rest=k%%3
ini=1
graf <- function(alpha,ini,fim){
x11()
par(mfrow = c(3, 1))
for (i in ini:fim) {
plot(alpha[i,],type='l',ylab=paste(paste("Alpha_", (i-1), sep = ""),'(t)',sep='')
,xlab="Tempo", ylim = c(min(alpha[i,]),max(alpha[i,]))) }}
for (i in 1:inteiro){
graf(alpha,ini,ini+2)
ini= ini+3}
if (rest>0){
graf(alpha,ini,ini+rest-1)
}
n = length(X)
yh = rep(0,n)
yh[1] = X[1]^2
k=1
for (i in 2:n) {
if (k<p+1) k=k+1
for (j in 2:k){
yh[i] = alpha[j,i]*(X[i-(j-1)]^2) + yh[i]
}
yh[i]=yh[i]+alpha[1,i]
}
yh=sqrt(yh)
X2=sqrt(X^2)
sigma=(sum((X2-yh)^2))/n
list(alpha=alpha,sig=sigma)
}
#=====
# função para lambda(t) com mc
#=====
mc <- function(x, p){

```

```

lamb=numeric()
lamb[1]=0
for (i in 2:length(x)){
lamb[i]=lamb[i-1]
  if ( (i%(2*p)) == 0){
    lamb[i]=var(x[i:(i-(2*p-1))])}
}
a=min(lam)
b=max(lam)
k=(lam-a)/(b-a)
k[k<0.1]=0.01
k[k>=0.1]=0.1
k
}

```

### B.3 Banda de Confiança para modelos $tvAR(p)$

Criação de Banda de Confiança para os coeficientes do modelo  $tvAR(p)$

Descrição:

Cria uma banda de confiança para cada coeficientes do modelo  $tvAR(p)$ .

Uso:

```
tvAR.Band (B1=1000, pb = 0.05, obj)
```

Argumentos:

B1: tamanho da amostra bootstrap.

pb: grau de confiança para cada intervalo.

obj: um objeto tipo  $tvAR$ .

Valor:

Uma lista com as seguintes componentes:

a.mais: uma matriz ( $p \times T$ ) com os valores do limite superior da banda de confiança das funções  $a_i$  do modelo para cada instante  $t$ , onde  $T$  representa o tamanho da série  $x$ .

a.menos: uma matriz ( $p \times T$ ) com os valores do limite inferior da banda de confiança das funções  $a_i$  do modelo para cada instante  $t$ , onde  $T$  representa o tamanho da série  $x$ .

a.media: uma matriz ( $p \times T$ ) com os valores da média da amostra bootstrap das funções  $a_i$  do modelo para cada instante  $t$ , onde  $T$  representa o tamanho da série  $x$ .

Código R para criação da função

```

tvAR.Band <- function(B1=1000, pb=0.05, obj ) {
pot=logb(dim(obj$m.coef)[2], base = 2) # potencia da série

```

```

N=dim(obj$m.coef)[2]                # tamanh da série
J=obj$J
n.coef= 2^(J-1)                    # 2^(j-1) no de coeficientes
p=dim(obj$m.coef)[1]
delta=p*n.coef                     # J no de escala
                                   # p no de parametros do AR(p)

x.star <- matrix(0,N,N,p)
Xwave1<-matrix(0,N,delta)  ##boot
a.star <- matrix(0,n.coef,p)
a.boot <- matrix(0,B1,N)
a.boot.T <- vector("list",p)
for (i in 1:p) a.boot.T[[i]]=a.boot #vetor de matriz
Betastar<-rep(0,delta)
for (j in 1:B1)
{
for (i in 1:delta) Betastar[i]<-rnorm(1,obj$Beta[i],obj$Beta.sig[i])
xstar<-obj$Xwave%*%Betastar
xp.xstar=matrix(0,N,p)
for (i in 1:p) xp.xstar[((i+1):N),i]=xstar[1:(N-i)]
Xwave1=matrix(0,N,p*n.coef)
for (i in 1:delta){
  Xwave1[,i]=obj$Psi[((i-1)%n.coef+1),]*xp.xstar[,((i-1)%n.coef+1)]
}
Infor<-lsfit(Xwave1,xstar,intercept=FALSE)
Betast<-Infor$coef
a.star=matrix(0,n.coef,p)
for (i in 1:p) a.star[,i]=Betast[((i-1)*n.coef+1):(i*n.coef)]
f=matrix(0,p,N)
for (i in 1:p)
f[i,]=-t(obj$Psi)%*%a.star[,i]
for (n.co in 1:p)
a.boot.T[[n.co]][j,]=f[n.co,]
}
a.mais=matrix(0,p,N)
a.menos=matrix(0,p,N)
a.media=matrix(0,p,N)
for (i in 1:p){
a.mais[i,]=apply(a.boot.T[[i]], 2, quantile, probs=1-pb/2)
a.menos[i,]=apply(a.boot.T[[i]], 2, quantile, probs=pb/2)
a.media[i,]=apply(a.boot.T[[i]],2,mean)
}
k=nrow(a.mais)
inteiro=k%%3
rest=k%%3
ini=1
graf <- function(alpha,a.mais,a.menos,a.media, ini,fim){
x11()
par(mfrow = c(3, 1))
for (i in ini:fim) {

```

```
plot(alpha[i,], type='l',ylab="Alpha",xlab="Tempo", ylim = c(min(alpha[i,],max(alpha[i,])),1))
lines(a.mais[i,],col=2)
lines(a.menos[i,],col=3)
lines(a.media[i,],col=4)
}}
if (inteiro!=0){
for (i in 1:inteiro){
graf(obj$m.coef,a.mais,a.menos,a.media, ini,ini+2)
ini= ini+3} }
if (rest>0){
graf(obj$m.coef,a.mais,a.menos,a.media,ini,ini+rest-1)
}
list(a.mais=a.mais, a.menos=a.menos, a.media=a.media, a.boot.T=a.boot.T)
}
```