



Instituto de  
MATEMÁTICA  
E ESTATÍSTICA

UFRGS



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

DEPARTAMENTO DE ESTATÍSTICA

## ESTIMAÇÃO ROBUSTA DA FUNÇÃO DENSIDADE ESPECTRAL

BATISTA NUNES DE OLIVEIRA

Porto Alegre  
2016

**BATISTA NUNES DE OLIVEIRA**

**ESTIMAÇÃO ROBUSTA DA FUNÇÃO DENSIDADE ESPECTRAL**

Trabalho de conclusão de curso submetido  
como requisito parcial para a obtenção do  
grau de Bacharel em Estatística

Orientador Metodológico  
Dr. Marcio Valk

Porto Alegre  
2016

Instituto de Matemática e Estatística  
Departamento

**ESTIMAÇÃO ROBUSTA DA FUNÇÃO DENSIDADE ESPECTRAL**  
Batista Nunes de Oliveira

Banca examinadora:

Dr. Guilherme Pumi  
UFRGS

Dr. Marcio Valk  
UFRGS

## Agradecimentos

A Deus, de onde vem todas as forças para alcançarmos nossos objetivos.

A minha mãe, Dona Geny, que sempre é um exemplo de bondade, acolhimento e incentivo na minha caminhada.

A minha Esposa, Margarete, que muito viu-se privada da minha companhia, mas que sempre demonstrou paciência, afeição e carinho para comigo. Lembrando “A” na disciplina “cezinho” em casa.

Ao meu filho Renan, que constantemente perguntava como estava o curso e tento ser um exemplo.

Ao orientador Marcio Valk, que sempre esteve parceiro em todas as etapas deste TCC, seja nas disciplinas de Processo Estocástico e Regressão Linear, ora no Estágio e nesta atividade terminal de curso, que com paciência e sabedoria contribui imensamente com o meu crescimento.

Aos colegas de curso, que com cumplicidade fizeram a Estatística mais bela, em especial: Dani, Renan, Max, Paula, Daia, Mariana, Raí e Angélica os mais cúmplices.

Aos professores, que sempre deixaram ensinamentos e buscam o nosso melhor.

*"Claro que se arrepender de atitudes que tivemos na vida, nos faz ser pessoas melhores, mas se é pra se arrepender verdadeiramente que seja daquilo que deixamos de fazer"*

## Resumo

Em séries temporais os eventos extremos podem afetar consideravelmente a especificação de modelos, estimação de parâmetros e previsão, ou seja, basicamente afeta toda a parte de inferência. Neste trabalho exploramos aspectos da estimação da função da densidade espectral e da função autocorrelação/autocovariância, visto que existe uma relação direta entre as duas, num contexto de séries temporais contaminadas por outliers aditivos e também no contexto de processos não gaussianos, em que a probabilidade de ocorrência de eventos extremos é significativa. Além disso, estudamos computacionalmente o desempenho de uma técnica bayesiana para a estimação da densidade espectral e o comportamento de um método para estimar a correlação entre eventos extremos, especialmente útil para detectar clusters de volatilidade.

**Palavras-chave:** Séries Temporais, Autocorrelação, Outliers, Extremograma.

# Abstract

Extreme events in time series might affect significantly the whole inference process, send as model specification, parameter estimation and prediction. Since there is a direct relationship between spectral density and autocovariance function, we explore in this work aspects of spectral density estimation and autocorrelation/autocovariance estimation in a context of time series contaminated by additive outliers. Besides that, we consider spectral density estimation in non-Gaussian processes, where the probability of occurrence of extreme events is significantly higher. In addition, we study the performance of a Bayesian technique for spectral density estimation and we investigate the behavior of a method to estimate the correlation among extreme events, especially useful for detecting clusters of volatility.

**Keywords:** Time series, Autocorrelation, Outliers, Extremogram.

## Lista de Figuras

2.1	Exemplos de periodicidades . . . . .	4
3.1	Gráfico ARMA(1,1) com $\phi = 0.8$ e $\theta = -0.3$ , gerada no R. . . . .	10
3.2	Fac e espectro teórica de uma ARMA(1,1) . . . . .	13
5.1	Efeito de outliers aditivos em um processo ARMA(1,1), com tamanho amostral $n=200$ e para magnitude igual a 0, 5 e 10. . . . .	28
6.1	Estimação de um ARMA(2,1) . . . . .	32
6.2	Estimação de um GARCH ao quadrado . . . . .	34
6.3	Estimação da função densidade espectral de um AR(1) . . . . .	36
6.4	Estimação da função densidade espectral de um GARCH(1,1) ao quadrado . . . . .	36
7.1	Estimação da função densidade espectral de um GARCH ao quadrado em detalhe . . . . .	42
7.2	Estimação da densidade espectral de um GARCH(1,1) ao quadrado $n=20000$ . . . . .	43

## Lista de Tabelas

4.1	<i>Debug</i> da função Extremogram1.R . . . . .	23
6.1	Vício e EQM para três estimadores da função de autocorrelação considerando o modelo AR(1) com coeficiente $\phi = -0.6$ . . . . .	31
6.2	Vício e EQM para três estimadores da função de autocorrelação considerando o modelo ARMA(2,1) com coeficientes $\phi_1 = -0.4$ , $\phi_2 = 0.4$ e $\theta = 0.8$ . . . . .	32
6.3	Vício e EQM para três estimadores da função de autocorrelação considerando o modelo GARCH(1,1) ao quadrado com coeficientes $\phi_1 = 0.3$ e $\theta = 0.6$ . . . . .	34

## Sumário

1	Introdução . . . . .	1
2	Periodicidade em séries temporais . . . . .	3
2.1	Frequência e período . . . . .	3
2.2	Medida de periodicidade . . . . .	5
3	Modelos para séries temporais . . . . .	8
3.1	Modelos ARMA . . . . .	8
3.2	Modelos ARCH e GARCH . . . . .	14
4	Estimadores para a função densidade espectral . . . . .	18
4.1	Periodograma robusto . . . . .	19
4.2	Estimador bayesiano da densidade espectral . . . . .	21
4.3	Extremograma . . . . .	21
4.4	Exemplo de aplicação da função Extremogram1.R . . . . .	23
4.5	Extremoperiodograma, o periodograma para eventos extremos . . . . .	24
5	Séries temporais com outliers . . . . .	26
5.1	Outliers . . . . .	26
5.2	Outliers aditivos e modelos estacionários . . . . .	28
6	Estudo de simulação . . . . .	30
6.1	Estimação da função de autocorrelação de modelos ARMA com outliers aditivos . . . . .	30

6.2	Estimação da função de autocorrelação em modelos GARCH ao quadrado . . . . .	33
6.3	Estimação da função de densidade espectral . . . . .	35
7	Resultados e conclusões . . . . .	38
	Referências . . . . .	40
	Apêndice . . . . .	42

# 1 Introdução

Dados coletados em diversas áreas do conhecimento como economia, biologia, medicina e engenharia, estão muitas vezes relacionados com o tempo (ou espaço) e nos levam naturalmente a considerar modelos de séries temporais. Amostras colhidas de todos estes campos são normalmente observadas sequencialmente e indexadas no tempo. Por exemplo, observa-se dados diários, mensais ou anuais e examina-se as histórias de tais séries ao longo de vários períodos de tempo, onde observações adjacentes não são independentes. Assim, as técnicas usuais da estatística clássica, desenvolvidos principalmente para observações independentes e identicamente distribuídas, não são aplicáveis [Shumway and Stoffer, 2010].

Muitas vezes uma série temporal apresenta comportamento cíclico, isto é, dado um determinado, tempo elas se repetem. O menor tempo em que ela começa a se repetir é chamado de período da série [Shumway and Stoffer, 2010]. Para explorar essa característica desenvolveu-se a análise espectral, também conhecida como análise no domínio da frequência, que consiste em decompor uma série temporal estacionária em uma soma de senoides com coeficientes aleatórios não correlacionados [Brockwell and Davis, 2013]. Uma das principais ferramentas nesse contexto é a função de densidade espectral.

Existe uma ampla literatura a respeito de como estimar a função densidade espectral para diversos tipos de processos estacionários [Morettin and Tolo, 2006]. No entanto existe uma carência de estudos do comportamento desses estimadores quando a série temporal apresenta algum tipo de contaminação por outliers, como os descritos em [Fox, 1972], ou até mesmo quando os processos não são Gaussianos e podem apresentar eventos extremos, clusters de volatilidades e não linearidade, como é o caso dos processos ARCH e GARCH ao quadrado.

O objetivo desse trabalho é investigar algumas abordagens supostamente robustas a

outliers e até mesmo a eventos extremos, como o Bspec que é uma abordagem bayesiana para a estimação da densidade espectral, proposto por [Röver et al., Röver], o Extremo-periodograma, que é um periodograma construído a partir do Extremograma, introduzido em [Davis et al., 2009] e [Davis et al., 2012] é uma abordagem robusta baseada na estimação quantílica da função de autocovariância proposta por [Ma and Genton, 2000].

No Capítulo 2 exploramos os conceitos de frequência e período e introduzimos uma medida de periodicidade. No Capítulo 3 introduzimos alguns modelos clássicos e exploramos algumas de suas propriedades e no Capítulo 4 apresentamos alguns estimadores para a função de densidade espectral. No Capítulo 5 definimos o conceito de outlier aditivo e no Capítulo 6, resultados de um estudo de simulação para estudar o comportamento dos estimadores, é apresentado.

## 2 Periodicidade em séries temporais

Um dos objetivos da análise espectral é encontrar comportamento periódico em séries temporais estacionárias. Com esse objetivo, nesse capítulo exploramos os conceitos de frequência e período e introduzimos uma medida de periodicidade.

### 2.1 Frequência e período

A noção de periodicidade pode ficar mais precisa se introduzirmos os conceitos de frequência e período. Para especificar a taxa na qual uma série oscila, definimos o primeiro ciclo como um período completo da função seno ou cosseno, sobre um intervalo de comprimento  $2\pi$ . Como exemplo de série periódica podemos considerar o processo

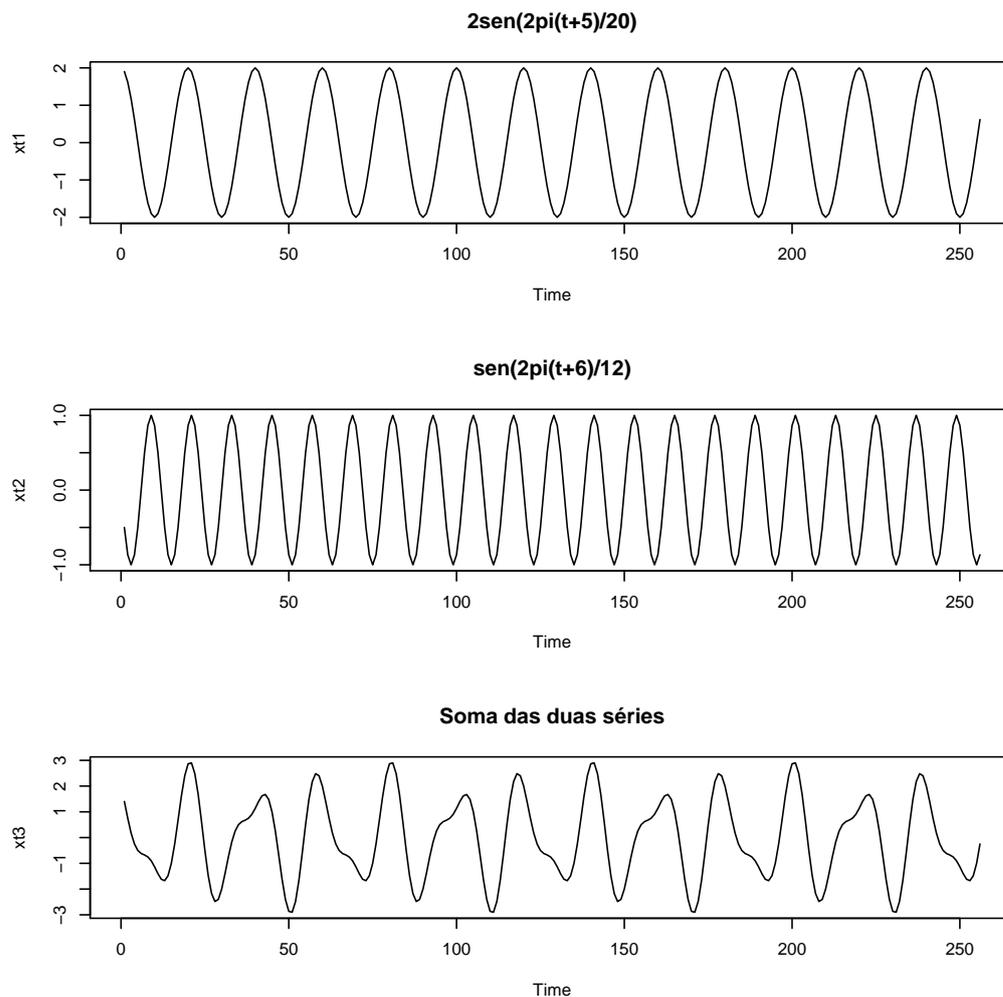
$$x_t = A \text{sen}(2\pi\omega t + \phi) \quad (2.1)$$

para  $t = 0, \pm 1, \pm 2, \dots$  onde  $\omega$  é o índice da frequência,  $A$  e  $\phi$  são arbitrários, normalmente chamados de amplitude e fase, respectivamente. A amplitude indica a altura do processo e a fase onde inicia a curva senoidal (ponto zero). O processo acima oscila com frequência  $\omega$ . Para  $\omega = 1$  a série faz um ciclo por unidade de tempo, se  $\omega = 0,5$  a série tem um ciclo para duas unidades de tempo, se  $\omega = 0,25$  a série faz um ciclo para quatro unidades. Assim, observa-se a relação  $\omega = 1/T_0$ , onde  $T_0$  é o período, ou seja, a frequência é o inverso do período [Shumway and Stoffer, 2010].

A primeira série mostrada na Figura 2.1 foi gerada tomando  $A=2$ ,  $\omega = 0,05$ ,  $\phi = \frac{10\pi}{20} = \frac{\pi}{2}$ , ou  $\phi = 90$  graus. A segunda  $A=1$ ,  $\omega = \frac{1}{12}$ ,  $\phi = \frac{12\pi}{12} = \pi$ , ou  $\phi = 180$  graus. A terceira é uma adição das duas anteriores, ou seja,

$$x_t = 2\text{sen}\left(\frac{2\pi(t+5)}{20}\right) + \text{sen}\left(\frac{2\pi(t+6)}{12}\right),$$

mostrando que a soma das séries é uma mistura das frequências das séries anteriores contendo frequências que oscilam com  $\omega=1/20$  e  $1/12$  ciclos por ponto. Para verificar isto basta observar que a série tem 10 ciclos em 200 pontos. É evidente que a soma de senos faz um trabalho razoável de emular o comportamento das duas séries observadas.



**Figura 2.1:** Exemplos de periodicidades. O primeiro gráfico apresenta uma série temporal cíclica com uma frequência de  $1/20$ , enquanto que o segundo gráfico apresenta uma série com frequência maior, de  $1/12$ . O terceiro gráfico representa a soma das duas primeiras séries. Fonte: <http://www.stat.ucdavis.edu/~shumway/sta137/ch3s.pdf>

A partir da relação entre período e frequência de uma série temporal, defini-se uma área da estatística chamada de análise espectral, que desempenha um papel fundamental na análise de séries temporais e é uma ferramenta importante para estimação de modelos. O

”power spectro”, ou função de densidade espectral, é uma medida de poder ou variância de uma série temporal, dada uma determinada frequência  $\omega$ . A função de densidade espectral, denotada por  $f(\omega)$  é essencial na especificação de modelos de séries temporais estacionárias e na construção de estimadores dos parâmetros destes modelos. A função  $f(\omega)$  pode ser definida como a Transformada de Fourier da função de autocovariância. No entanto, esse fato será apropriadamente abordado no Capítulo 4. Cabe salientar que a Transformada de Fourier (TF) é baseada na chamada série de Fourier que diz: “dada uma função periódica  $g(x)$ , ela pode ser escrita como uma combinação de senos e cossenos (raras exceções)”, ou seja,

$$g(x) = M + \sum_{n=1}^{\infty} \left( a_n \cos\left(\frac{nx\pi}{L}\right) + b_n \sen\left(\frac{nx\pi}{L}\right) \right) \quad (2.2)$$

onde  $M$  é uma constante;  $a_n, b_n$  são séries numéricas e  $L = T_0/2$ . A Transformada de Fourier utiliza esta mesma ideia, fazendo o período  $T_0 \rightarrow \infty$ , transformando o somatório em integral, obtendo

$$\hat{g}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega x} g(x) d(x). \quad (2.3)$$

Assim,  $\hat{g}(\omega)$  é chamado de Transformada de Fourier da função  $g(x)$  e claro que que a TF depende do seno e cosseno, uma vez que

$$e^{-i\omega x} = \cos(\omega x) - i \sen(\omega x).$$

## 2.2 Medida de periodicidade

A fim de estimar numericamente as quantidades conceituais utilizando dados da amostra se faz necessário o desenvolvimento de estimadores espectrais coerentes com base de dados experimentais observados. Discutimos essas quantidades abaixo, depois de definir a transformada de Fourier discreta e o periodograma como quantidades preliminares a partir do qual o espectro pode ser coerentemente calculado.

A Transformada de Fourier Discreta (DFT) de uma série temporal amostral,  $x_t$ ,  $t = 0, \dots, n - 1$ , é definida como

$$\tilde{x}(\omega_k) = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} x_t \exp(-i\omega_k t), \quad (2.4)$$

onde  $\omega_k = 2\pi k/n, k = 0, \dots, n-1$ , define o conjunto de frequências sobre as quais a DFT é calculado, denominadas *frequências de Fourier*. Note que rotulamos  $x_0, x_1, \dots, x_{n-1}$  por conveniência de notação, isto significa que as frequências calculadas em DFT são da forma  $\omega = 0, 2\pi/n, 4\pi/n, \dots, 2\pi(n-1)/n$ . O cálculo do produto em (2.4) usualmente assume que o comprimento da série, de símbolo  $n$ , é uma potência de 2. Se a série não for uma potência de 2, podemos acrescentar zeros estendendo a série temporal até que atinja uma potência de dois. A representação  $\exp(-i\omega t) = \cos(\omega t) - i \sin(\omega t)$  mostra que a transformada de Fourier (2.4) é o resultado de combinações de séries das funções seno e cosseno para diferentes frequências  $\omega$ , [Shumway and Stoffer, 2010].

É razoável que (2.4) deve ter valores que reforcem a periodicidade da série  $x_t$ . Em particular, seria de se esperar (2.4) grande quando  $\omega_k$  é um componente da série de frequências e deve ser pequeno caso contrário. Uma vez que (2.4) têm uma parte real e imaginária para toda frequência e que pode ser positiva ou negativa, é tradicional calcular primeiramente a magnitude quadrática de (2.4), que é chamado de periodograma.

O periodograma é definido como

$$I_n(\omega_k) = |\tilde{x}(\omega_k)|^2 \quad (2.5)$$

e é apenas a soma dos quadrados dos senos e cossenos da série transformada. Se considerarmos a densidade espectral  $f(\omega_k)$  como a quantidade a estimar, e temos a relação

$$E|\tilde{x}(\omega_k)|^2 \approx f(\omega_k), \quad (2.6)$$

para  $n$  grande, isto implica que o periodograma é um estimador assintoticamente não viesado da densidade espectral na frequência  $\omega_k$ .

Podes-se mostrar que a autocovariância entre duas frequências  $\omega_k$  e  $\omega_\ell$  é zero para todo  $k \neq \ell$  com ambas frequências múltiplas de  $2\pi/n$  [Shumway and Stoffer, 2010]. Também mostra-se que as funções seno e cosseno de (2.4) não são correlacionadas e são aproximadamente normais com variâncias igual a  $\frac{1}{2}f(\omega_k)$ , sendo  $n$  grande. Isto implica que os quadrados padronizados das componentes reais e imaginários têm distribuição qui-quadrado com 1 grau de liberdade. O periodograma é a soma de duas variáveis qui-quadradas, assim deve ter uma distribuição qui-quadrado com dois graus de liberdade.

Combinando os resultados acima com a definição e as propriedades de distribuição

do periodograma, nos permite definir o intervalo de confiança espectral  $(100(1 - \alpha)\%)$  utilizando

$$\frac{2I(\omega_k)}{\chi_2^2(\frac{\alpha}{2})} \leq f(\omega_k) \leq \frac{2I(\omega_k)}{\chi_2^2(1 - \frac{\alpha}{2})}. \quad (2.7)$$

Deve-se notar que existe uma tendência de introduzir periodicidades de longo alcance que inflam o espectro em baixas frequências (valores pequenos de  $\omega_k$ ), já que o periodograma assume que a tendência é parte de uma frequência muito longa. Este comportamento possivelmente irá obscurecer componentes importantes em frequências mais altas. Por esta razão é importante trabalhar em (2.4) com a transformação da série ajustada da média  $(x_t - \bar{x})$  ou retificada  $(x_t - \hat{a} - \hat{b}t)$ .

Uma outra modificação que, por vezes, melhora a aproximação em (2.5) é o afinilamento, um processo em que cada ponto  $x_t$  é substituído por  $a_t x_t$ , onde  $a_t$  é uma função, frequentemente o cosseno, que é máximo no ponto médio da série e decaí nos extremos. O afinilamento é usado em diferentes aplicações de engenharia, onde existem grande variações no valor da densidade espectral  $f(\omega)$ .

Mais algumas propriedades do periodograma são apresentadas no Capítulo 4.

### 3 Modelos para séries temporais

No Capítulo anterior, ficou caracterizado uma medida de periodicidade para séries temporais num contexto geral. No entanto, para especificar a função densidade espectral é necessário caracterizar o processo. Neste capítulo, introduzimos alguns modelos clássicos e exploramos algumas de suas propriedades.

#### 3.1 Modelos ARMA

Os modelos autorregressivos AR, são usados com frequência na área de Economia, quando o valor de uma variável no instante  $t$ , seja função de valores anteriores desta variável. Em outras áreas, o interesse no modelo autorregressivo pode ser o de estimar o espectro do processo [Morettin and Toloi, 2006]. Um processo pode ser representado pelo modelo de Médias Móveis (MA) puro, mas não é muito intuitivo, visto que dependem basicamente de ruídos brancos. Assim, muitas vezes o modelo solução para um processo sem um número muito grande de parâmetros é a combinação destes dois modelos gerando um processo chamado de ARMA(p,q) da forma

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \dots + \phi_p \tilde{Z}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}. \quad (3.1)$$

onde  $\varepsilon_t$  é um  $RB(0, \sigma_\varepsilon^2)$ . Se  $\phi(B) = 1 - \phi_1(B) - \phi_2(B^2) - \dots - \phi_p(B^p)$  e  $\theta(B) = 1 - \theta_1(B) - \theta_2(B^2) - \dots - \theta_q(B^q)$  são os operadores autorregressivos e de médias móveis, respectivamente, podemos escrever compactadamente

$$\phi(B)\tilde{Z}_t = \theta(B)\varepsilon_t. \quad (3.2)$$

Um modelo muito utilizado é o qual temos  $p = q = 1$ , ou seja, ARMA(1,1)

$$\tilde{Z}_t = \phi \tilde{Z}_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}. \quad (3.3)$$

Substituindo-se sequencialmente  $\tilde{Z}_{t-1}, \tilde{Z}_{t-2}, \dots$  em (3.3), obtemos  $\tilde{Z}_t$  na forma linear (médias móveis de ordem infinita). Vejamos primeiro substituindo  $\tilde{Z}_{t-1}$

$$\tilde{Z}_t = \phi(\phi \tilde{Z}_{t-2} + \varepsilon_{t-1} - \theta \varepsilon_{t-2}) + \varepsilon_t - \theta \varepsilon_{t-1}.$$

Substituindo  $\tilde{Z}_{t-2}$

$$\tilde{Z}_t = \phi(\phi(\phi \tilde{Z}_{t-3} + \varepsilon_{t-2} - \theta \varepsilon_{t-3}) + \varepsilon_{t-1} - \theta \varepsilon_{t-2}) + \varepsilon_t - \theta \varepsilon_{t-1}.$$

Continuando o processo obteremos

$$\tilde{Z}_t = \psi(B)\varepsilon_t, \quad \psi(B) = \sum_{j=1}^{\infty} \psi_j B^j$$

onde  $\psi_j = \phi^{j-1}(\phi - \theta)$  para  $j \geq 1$ . Assim o processo será estacionário se

$$\sum_{j=1}^{\infty} \psi_j = (\phi - \theta) \sum_{j=1}^{\infty} \phi^{j-1} < \infty,$$

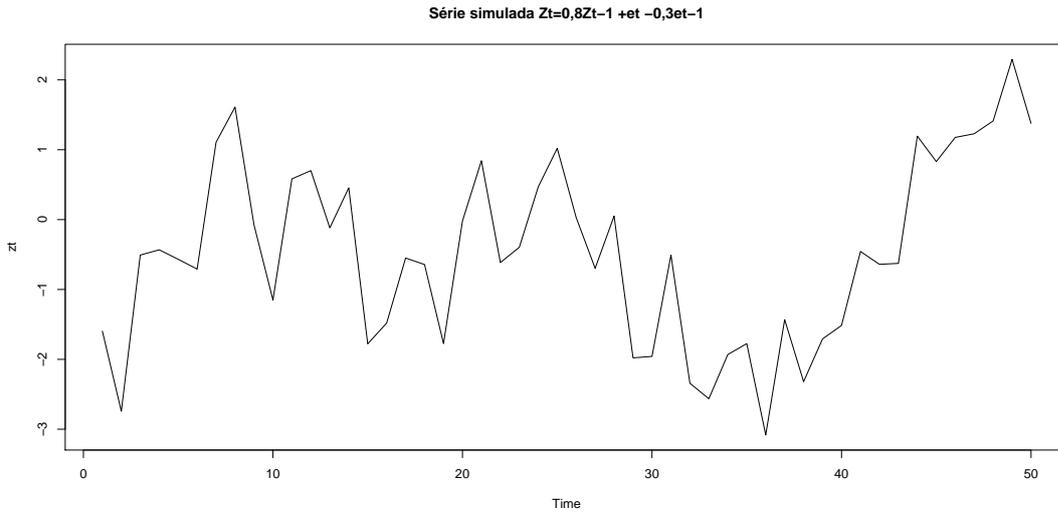
isto é, se  $|\phi| < 1$ . Para mais detalhes veja [Morettin and Tolo, 2006].

Analogamente, o modelo ARMA(1,1), pode ser escrito da seguinte maneira

$$\pi(B)\tilde{Z}_t = \varepsilon_t,$$

onde  $\pi_j = \theta^{j-1}(\phi - \theta)$ ,  $j \geq 1$ , e o processo é invertível se  $\sum \pi_j < \infty$ , isto é, se  $|\theta| < 1$ . Observa-se que a condição de estacionariedade e invertibilidade de um processo ARMA(1,1) é idêntica a estacionariedade de um processo AR(1) e de invertibilidade de um MA(1). Estes resultados podem ser generalizado para um processo ARMA(p,q) qualquer. Tomando como referência (3.2) podemos escrever

$$\tilde{Z}_t = \psi(B)\varepsilon_t = \theta(B)\phi^{-1}(B)\varepsilon_t$$



**Figura 3.1:** Gráfico ARMA(1,1) com  $\phi = 0.8$  e  $\theta = -0.3$ , gerada no R.

ou

$$\pi(B)\tilde{Z}_t = \phi(B)\theta^{-1}(B)\tilde{Z}_t = \varepsilon_t.$$

**Exemplo 1.** A Figura 3.1 mostra uma série de 50 observações, gerados por uma ARMA(1,1)

$$Z_t = 0,8Z_{t-1} + \varepsilon_t - 0,3\varepsilon_{t-1}.$$

Note que  $|\phi| = |0,8| < 1$  e  $|\theta| = |-0,3| < 1$ , portanto  $Z_t$  é estacionária e invertível.

Tendo uma visão geral, do modelo ARMA(p,q):

a) Estacionariedade e Invertibilidade

Pelo que foi visto acima, o processo é estacionário quando as raízes de  $\phi(B) = 0$  estão fora do círculo unitário e analogamente é invertível quando as raízes de  $\theta(B) = 0$  estão fora do círculo unitário.

b) Função de autocorrelação

A partir de (3.1) e multiplicando por  $Z_{t-j}$  e tomando as esperanças chegamos em

$$\gamma_j = E(\tilde{Z}_t \tilde{Z}_{t-j}) = E\{(\phi_1 \tilde{Z}_{t-1} + \dots + \phi_p \tilde{Z}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}) \tilde{Z}_{t-j}\}$$

segue que,

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p} + \gamma_{z\alpha}(j) - \theta_1 \gamma_{z\alpha}(j-1) - \dots - \theta_q \gamma_{z\alpha}(j-q), \quad (3.4)$$

onde  $\gamma_{z\alpha}(j)$  é a covariância cruzada entre  $Z_t$  e  $\varepsilon_t$ , definida por

$$\gamma_{z\alpha}(j) = E(\varepsilon_t \tilde{Z}_{t-j}). \quad (3.5)$$

Como  $\tilde{Z}_{t-j}$  depende apenas dos efeitos de  $\varepsilon_t$  ocorridos até o instante  $t-j$ , obtemos

$$\gamma_{z\alpha}(j) \begin{cases} = 0, & j > 0 \\ \neq 0, & j \leq 0. \end{cases} \quad (3.6)$$

Assim (3.4) fica

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p}. \quad (3.7)$$

A fac(função de autocorrelação) segue de (3.7):

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_p \rho_{j-p}, \quad j > q, \quad (3.8)$$

conclui-se que os parâmetros das médias móveis influenciam diretamente as autocorrelações de "lags"  $1, 2, \dots, q$ , porém para  $j > q$  as mesmas funcionam como nos modelos autor-regressivos.

### c) Função densidade espectral

Da proposição 5.1 de [Morettin and Toloi, 2006] e substituindo  $\psi(B)$  por  $\theta(B)\phi^{-1}(B)$ , temos que

$$f(\omega) = \frac{\sigma_\varepsilon^2 |1 - \theta_1 e^{-i\omega} - \theta_2 e^{-2i\omega} - \dots - \theta_q e^{-qi\omega}|^2}{2\pi |1 - \phi_1 e^{-i\omega} - \phi_2 e^{-2i\omega} \dots - \phi_p e^{-pi\omega}|^2}, \quad -\pi \leq \omega \leq \pi. \quad (3.9)$$

**Exemplo 2.** Lembrando o modelo ARMA(1, 1) de (3.3), com  $\phi = \phi_1$  e  $\theta = \theta_1$ .

De (3.4) temos

$$\gamma_1 = \phi_1\gamma_0 - \theta_1\gamma_{z\alpha}(0) = \phi_1\gamma_0 - \theta_1\sigma_\varepsilon^2. \quad (3.10)$$

Usando a definição da variância cruzada em (3.5) segue que

$$\gamma_{z\alpha}(0) = E(\varepsilon_t\tilde{Z}_t) = E[\varepsilon_t(\phi_1\tilde{Z}_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1})] = E(\varepsilon_t^2) = \sigma_\varepsilon^2.$$

Também,

$$\gamma_0 = \phi_1\gamma_1 + \gamma_{z\alpha}(0) - \theta_1\gamma_{z\alpha}(-1) = \phi_1\gamma_1 + \sigma_\varepsilon^2 - \theta_1(\phi_1 - \theta_1)\sigma_\varepsilon^2, \quad (3.11)$$

visto que

$$\begin{aligned} \gamma_{z\alpha}(-1) &= E(\varepsilon_t\tilde{Z}_{t+1}) = E[\varepsilon_t(\phi_1\tilde{Z}_t + \varepsilon_{t+1} - \theta_1\varepsilon_t)] \\ &= \phi_1E(\varepsilon_t\tilde{Z}_t) + E(\varepsilon_t\varepsilon_{t+1}) - \theta_1E(\varepsilon_t^2) \\ &= \phi_1E(\varepsilon_t^2) - \theta_1E(\varepsilon_t^2) = (\phi_1 - \theta_1)\sigma_\varepsilon^2. \end{aligned}$$

Resolvendo (3.10) e (3.11), chegamos que

$$\gamma_0 = \frac{1 + \theta_1^2 - 2\phi_1\theta_1}{1 - \phi_1^2}\sigma_\varepsilon^2, \quad \gamma_1 = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 - \phi_1^2}\sigma_\varepsilon^2, \quad (3.12)$$

assim,

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 - 2\phi_1\theta_1}. \quad (3.13)$$

Da fac de (3.8), temos para  $j > 1$

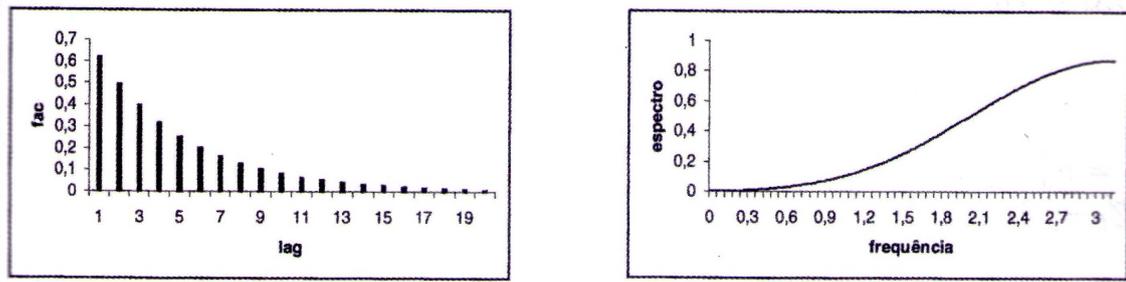
$$\rho_j = \phi_1\rho_{j-1}. \quad (3.14)$$

De (3.9) a função densidade espectral é dada por

$$f(\omega) = \frac{\sigma_\varepsilon^2 (1 + \theta_1^2 - 2\theta_1\cos(\omega))}{2\pi (1 + \phi_1^2 - 2\phi_1\cos(\omega))}, \quad -\pi \leq \omega \leq \pi. \quad (3.15)$$

Lembrando o exemplo 1, na qual  $\phi_1 = 0,8$  e  $\theta_1 = -0,3$ , temos que

$$\rho_1 = \frac{(1 - 0,24)(0,5)}{0,61} = 0,623,$$



**Figura 3.2:** Fac e espectro teórica de uma ARMA(1,1) Fonte: [Morettin and Toloi, 2006].

$$\rho_2 = \phi_1 \rho_1 = (0,8)\rho_1 = 0,498,$$

$$\rho_3 = (0,8)\rho_2 = 0,399,$$

e assim sucessivamente, e

$$f(\omega) = \frac{\sigma_\varepsilon^2 (1,09 - 0,6 \cos(\omega))}{2\pi (1,64 - 1,6 \cos(\omega))}, \quad -\pi \leq \omega \leq \pi.$$

A Figura 3.2 mostra a fac teórica de um modelo ARMA(1,1), com  $\phi_1 = 0,8$  e  $\theta_1 = -0,3$ , bem como a função de densidade espectral.

Se considerarmos a soma de dois processos ARMA independentes, temos um importante resultado;

**Teorema 3.1.1.** *Se  $X_t \sim ARMA(p_1, q_1)$  e  $Y_t \sim ARMA(p_2, q_2)$ ,  $X_t$  e  $Y_t$  independentes,  $Z_t = X_t + Y_t$ , então  $Z_t \sim ARMA(p, q)$ , onde  $p \leq p_1 + p_2$ , e  $q \leq \max\{p_1 + q_2, p_2 + q_1\}$ .*

Em resumo, a soma de dois modelos ARMA também será um ARMA, para verificar a demonstração, veja [Morettin and Toloi, 2006].

**Exemplo 3.** Sejam

$$(1 - \phi_1 B)X_t = \varepsilon_t$$

$$(1 - \phi_1 B)(1 - \phi_2 B)Y_t = a_t,$$

$$Z_t = X_t + Y_t.$$

Onde  $\varepsilon_t$  e  $a_t$  são ruídos brancos independentes e assim,  $(1 - \phi_1 B)(1 - \phi_2 B)Z_t = (1 - \phi_1 B)\varepsilon_t + a_t$ . Portanto

$$Z_t \sim ARMA(2, 1).$$

**Exemplo 4.** Suponha

$$(1 - \phi B)X_t = \varepsilon_t, \quad (1 + \phi B)Y_t = a_t,$$

com  $\sigma_\varepsilon^2 = \sigma_a^2 = \sigma^2$ . Se a série  $Z_t = X_t + Y_t$ , temos que

$$(1 - \phi B)(1 + \phi B)Z_t = (1 + \phi B)\varepsilon_t + (1 - \phi B)a_t.$$

Se definirmos  $Q_t = \varepsilon_t + \phi\varepsilon_{t-1} + a_t - a_{t-1}$ , então

$$\text{Var}(Q_t) = 2(1 + \phi^2)\sigma^2$$

e

$$E(Q_t Q_{t-k}) = 0, \quad k > 0,$$

o que nos leva que  $Q_t$  é ruído branco e portanto  $Z_t \sim AR(2)$ , isto é, não sendo uma ARMA(2,1), visto que  $q=0$ .

### 3.2 Modelos ARCH e GARCH

É conhecido na literatura que previsão de séries temporais, tais como preços de ativos financeiros, taxas de câmbio e inflação, por exemplo, oscilam substancialmente de um período para outro.

A magnitude dos erros de previsão varia, sendo relativamente grande em determinados períodos, depois relativamente pequenos e essa característica tende a aparecer agrupada. Para calcular os erros de previsão [De Losso, 2011] usa

$$\begin{aligned} \rho_T(1) &= y_{T+1} - \hat{y}_T(1) = \varepsilon_{T+1} \\ \rho_T(2) &= y_{T+2} - \hat{y}_T(2) = \varepsilon_{T+2} + \psi_1 \varepsilon_{T+1} \\ &\vdots \\ \rho_T(h) &= y_{T+h} - \hat{y}_T(h) = \varepsilon_{T+h} + \psi_1 \varepsilon_{T+h-1} + \dots + \psi_{h-1} \varepsilon_{T+1} \end{aligned}$$

Podemos observar que os erros de previsão dependem das perturbações do modelo  $\varepsilon_t$  e que pode haver algum tipo de autocorrelação entre esses erros. Assim, de certa forma fica caracterizado a existência de autocorrelação na variância de  $\varepsilon_t$ , além de que a variância

dos erros de previsão não é constante, o que é chamado de heterocedasticidade.

Um dos objetivos em finanças é a avaliação de riscos de uma carteira de ativos financeiros. O risco é frequentemente medido em termos de variações dos preços dos ativos. Se denotarmos por  $P_t$  o preço de um ativo no instante  $t$ , supondo que não tenha dividendos pago no período. A variação de preços entres os instantes  $t - 1$  e  $t$ , é dada por  $\Delta P_t = P_t - P_{t-1}$  e o retorno líquido simples é definido por

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

Na prática, é preferível trabalhar com *log-retorno*, ou simplesmente *retorno*, que é definido por

$$r_t = \log \left( \frac{P_t}{P_{t-1}} \right), \quad (3.16)$$

o qual possui propriedades estatísticas mais interessantes, como estacionariedade e ergodicidade ([Morettin and Tolo, 2006]).

Quando considerarmos a variável condicionada ao seu instante anterior (um retorno por exemplo), a variância dessa variável condicionada é chamada de volatilidade.

Modelos ARCH (Autoregressive conditional heteroskedasticity) foram introduzidos por [Engle, 1982] e sua generalização GARCH (Generalized ARCH) foi proposto por [Bollerslev, 1986]. Essa classe de modelos é muito utilizada para modelar a volatilidade de uma série de retornos.

Os retornos apresentam algumas características próprias que são chamadas de fatos estilizados, são eles:

- (i) Os retornos são, em geral, não autocorrelacionados;
- (ii) Os retornos apresentam clusters de volatilidade ao longo do tempo;
- (iii) Quando considera-se o quadrado dos retornos, estes possuem autocorrelação;
- (iv) Distribuição com caudas pesadas, não normal e assimetria;
- (v) Podem apresentar características de não linearidade.

Na análise de séries temporais não lineares as inovações (choques aleatórios)  $\varepsilon_t$  são em

geral, supostas i.i.d. e o modelo tem a seguinte forma:

$$r_t = \mu(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots) + \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots)\varepsilon_t.$$

Assim, se  $\mu(\cdot)$ , que representa a função da média condicional, for não linear, dizemos que o modelo é não linear na média. Se  $\sigma^2(\cdot)$ , que representa a variância condicional, for não linear, então diz-se que o modelo é não linear na variância. Um modelo típico para a volatilidade pode ser definido como

$$r_t = \mu_t + \sigma_t \varepsilon_t,$$

em que  $E_{t-1}(\varepsilon_t) = E(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = 0$  e  $Var_{t-1}(\varepsilon_t) = Var(\varepsilon_t | \varepsilon_{t-1}, \dots) = 1$ .

Assume-se, em geral, que  $\varepsilon_t$  é uma sequência i.i.d. com certa distribuição,  $\mu$  e  $\sigma^2$  são respectivamente, a média e a variância condicional de  $r_t$ .

Um modelo ARCH(p) é definido por

$$r_t = \sigma_t \varepsilon_t, \tag{3.17}$$

em que  $\varepsilon_t \sim IID(0, 1)$  com variância condicional

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \alpha_2 r_{t-2}^2 + \dots + \alpha_p r_{t-p}^2, \tag{3.18}$$

onde  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $i = 1, \dots, m - 1$  e  $\alpha_m > 0$ .

Definindo

$$v_t = r_t^2 - \sigma_t^2$$

e substituindo em (3.18), obtemos

$$r_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \alpha_2 r_{t-2}^2 + \dots + \alpha_p r_{t-p}^2 + v_t,$$

ou seja,  $r_t^2$  segue um AR(p) com

$$v_t = r_t^2 - \sigma_t^2 = \sigma_t^2(\varepsilon_t^2 - 1)$$

sendo não Gaussiano mesmo quando  $\varepsilon_t \sim N(0, 1)$ .

Um modelo GARCH(p,q) para a volatilidade em (3.17) é definido como

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i r_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad (3.19)$$

em que  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $i = 1, \dots, p-1$  e  $\alpha_p > 0$ ,  $\beta_i \geq 0$ , para  $i = 1, \dots, q-1$ , e  $\beta_q > 0$ .

Sendo  $q^* = \max(p, q)$ , a condição

$$\sum_{i=1}^{q^*} (\alpha_i + \beta_i) < 1$$

é suficiente mas não necessária para que  $\sigma_t^2 > 0$ .

Novamente, definindo  $v_t = r_t^2 - \sigma_t^2$  e substituindo em (3.19), segue que

$$r_t^2 = \alpha_0 + \sum_{i=1}^{q^*} (\alpha_i + \beta_i) r_{t-i}^2 - \sum_{j=1}^q \beta_j v_{t-j} + v_t$$

de onde segue que  $r_t^2$  é um processo ARMA( $q^*$ ,q). Observe que  $v_t$  não é, em geral, um processo iid.

Para um processo GARCH(1,1) com  $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$ ,  $\varepsilon_t \sim \text{iid}(0,1)$  e quarto momento finito, a função de autocorrelação é dada por

$$\rho_1 = \left( \alpha + \frac{\alpha^2 \beta}{1 - 2\alpha\beta - \beta^2} \right)$$

e

$$\rho_k = \left( \alpha + \frac{\alpha^2 \beta}{1 - 2\alpha\beta - \beta^2} \right) (\alpha + \beta)^{k-1}, \quad \text{para } k \geq 2, \quad (3.20)$$

ou seja,

$$\rho_k = \rho_1 (\alpha + \beta)^{k-1}$$

( ver [Baillie and Chung, 2001]).

## 4 Estimadores para a função densidade espectral

No Capítulo 2 abordamos a teoria de periodicidade em séries temporais, definimos a função de densidade espectral e apresentamos o periodograma como um estimador da densidade espectral, obtido diretamente da transformada discreta de Fourier. No entanto, é conhecido na literatura, que existe uma relação direta entre a função de autocovariância e a função de densidade espectral. Nesse capítulo exploramos essa relação para definir alguns estimadores da densidade espectral.

Para um processo estacionário <sup>1</sup>  $\{x_t\}_{t \in \mathbb{Z}}$  com média zero, a função de autocovariância é definida como

$$\gamma(h) = \int_{(-\pi, \pi]} e^{ihx} dF(x), \quad h \in \mathbb{Z}, \quad (4.1)$$

em que  $F(\omega) = \int_{-\pi}^{\omega} f(x) dx$ , com  $\omega \in (-\pi, \pi]$  é chamada *distribuição espectral*, em que  $f(\cdot)$  é a *função densidade espectral* ou *espectro*. Se  $\{x_t\}_{t \in \mathbb{Z}}$  tem espectro contínuo, então podemos escrever o espectro da forma

$$f(\omega) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \gamma(h) \cos(h\omega), \quad \omega \in (-\pi, \pi]. \quad (4.2)$$

Seja  $x_t$ ,  $t = 1, \dots, n$ , uma realização do processo  $\{x_t\}_{t \in \mathbb{Z}}$ . Um estimador natural para a densidade espectral é dado por  $\hat{f}(\omega_j) = I(\omega_j)$ , em que  $I(\cdot)$  é o *periodograma* que pode ser escrito como

$$I(\omega_j) = \frac{1}{2\pi} \hat{\gamma}(0) + \frac{1}{\pi} \sum_{h=1}^{n-1} \hat{\gamma}(h) \cos(h\omega_j) \quad (4.3)$$

---

<sup>1</sup>Entende-se estacionariedade a estacionariedade de covariância ou de segunda ordem

em que  $\omega_j$  são as frequências de Fourier e

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_t - \bar{x})(x_{t+|h|} - \bar{x}). \quad (4.4)$$

A função periodograma é um estimador assintoticamente não viesado para  $2\pi f(\omega_j)$ , e sua variância é dada por  $\text{Var}[I(\omega_j)] = (2\pi)^2 f^2(\omega_j) + O(n^{-\frac{1}{2}})$ . Note que quando  $n \rightarrow \infty$  a variância do estimador não tende a zero o que torna o periodograma um estimador não consistente para o espectro (ver [Brockwell and Davis, 1991]). É também conhecido que o periodograma é fortemente afetado por outliers aditivos, (ver [Fajardo et al.]). Uma opção consistente é o chamado *estimador suavizado de covariâncias*, denotado por  $I^*(\cdot)$ , é dado por

$$I^*(w_j) = \frac{1}{2\pi} \hat{\gamma}(0) + \frac{1}{\pi} \sum_{h=1}^m \kappa\left(\frac{h}{m}\right) \hat{\gamma}(h) \cos(hw_j), \quad \text{para } w_j \in (0, \pi], \quad (4.5)$$

onde  $\hat{\gamma}(\cdot)$  é a função de autocovariância amostral dada pela expressão (4.4),  $m = n^\beta$ , para  $0 \leq \beta \leq 1$  e  $\kappa(u)$  é uma função par, contínua por intervalos de  $u$  satisfazendo as condições

A.i)  $0 \leq \kappa(u) \leq \kappa(0) = 1$ ,

A.ii)  $\kappa(-u) = \kappa(u)$ , para todo  $u$ ,

A.iii)  $\kappa(u) = 0$ , para  $|u| > 1$ .

A função  $\kappa(\cdot)$  é chamada de *função de ponderação*, *função peso* ou *núcleo*.

Note que, pela propriedade A.iii) da função de ponderação, o produto  $\kappa\left(\frac{h}{m}\right) \hat{\gamma}_x(h) = 0$ , para  $|h| > m$ .

## 4.1 Periodograma robusto

Uma opção robusta para o periodograma foi proposto por [Fajardo et al.], o qual utiliza a função de autocovariância robusta no cálculo do estimador da densidade espectral. Essa função de autocovariância robusta é apresentada por [Ma and Genton, 2000]

$$\hat{\gamma}_R(h) = \frac{1}{4} [Q_{n-h}^2(u+v) + Q_{n-h}^2(u-v)] \quad (4.6)$$

em que  $u$  e  $v$  são vetores contendo as primeiras  $n - h$  e as últimas  $n - h$  observações, respectivamente.  $Q_n(\cdot)$  é denominado *robust scale estimator function* baseado na  $\eta$ -ésima estatística de ordem entre as  $\binom{n}{2}$  distâncias  $\{|x_i - x_j|, i < j\}$ , ou seja,  $Q_n(x)$  é a  $\eta$ -ésima estatística de ordem do vetor  $|x_i - x_j|, i < j$ , e pode ser escrita como

$$Q_n(x) = c \times \{|x_i - x_j|, i < j\}_{(\eta)} \quad (4.7)$$

em que  $x_t, t = 1, \dots, n$ , é uma série temporal,  $c$  é uma constante utilizada para garantir consistência do estimador da função de autocovariâncias (para a distribuição normal,  $c=2.2191$ ) e  $\eta = \lfloor (\binom{n}{2} + 2)/4 \rfloor + 1$ .

Um estimador robusto para a densidade espectral é obtido substituindo-se a função de autocovariância robusta na equação (4.3). Desta forma fica definido um estimador robusto da densidade espectral

$$I_R(\omega_j) = \frac{1}{2\pi} \hat{\gamma}_R(0) + \frac{1}{\pi} \sum_{h=1}^{n-1} \hat{\gamma}_R(h) \cos(h\omega_j). \quad (4.8)$$

Observe que para  $h = n - 1$  os vetores  $u$  and  $v$  têm tamanho 1. Assim, a estimação de  $\gamma_R(h)$  para  $h$  perto de  $n - 1$  é bem pobre. Esse problema pode ser contornado utilizando-se a ideia da versão suavizada do periodograma (ver [Fajardo et al.]) em que introduz-se uma janela espectral para tornar consistente a estimativa do periodograma. Assim, omite-se as estimativas correspondentes as caudas da autocovariância robusta da seguinte forma:

$$I_R^*(\omega_j) = \frac{1}{2\pi} \hat{\gamma}_R(0) + \frac{1}{\pi} \sum_{h=1}^m \kappa(h) \hat{\gamma}_R(h) \cos(h\omega_j). \quad (4.9)$$

em que  $\kappa(\cdot)$  é definida como:

$$\kappa(h) = \begin{cases} 1, & \text{se } h \leq m \\ 0, & \text{se } h > m. \end{cases}$$

Observe que  $\kappa(\cdot)$  é um caso particular da função de ponderação usado na teoria clássica de análise espectral para obter consistência do periodograma, e  $m$  é utilizado como uma função da amostra do tipo  $m = n^\beta$ , para  $\beta \in (0, 1)$ .

## 4.2 Estimador bayesiano da densidade espectral

Uma abordagem bayesiana para a estimação da densidade espectral é proposto por [Röver et al., Röver]. Neste trabalho a suposição de ruído branco Gaussiano estacionário, comumente utilizada na aproximação de Whittle, é flexibilizada para uma distribuição  $t$ , para acomodar eventos extremos. Na abordagem de Whittle, a suposição explícita para ruído  $\varepsilon_t$  é que sua transformada discreta de Fourier  $\tilde{\varepsilon}(\omega)$  é independente com distribuição Gaussiana com média zero e variância proporcional a densidade espectral

$$\text{Var}(\text{Re}(\tilde{\varepsilon}(\omega_j))) = \text{Var}(\text{Im}(\tilde{\varepsilon}(\omega_j))) = \frac{n}{4\Delta_t} \rho(\omega_j)$$

onde  $\Delta_t = \frac{1}{2\pi}$ ,  $\omega_j$  são as frequências de Fourier  $f(\omega_j)$  é a densidade espectral,  $j = 1, 2, \dots, \frac{n}{2}$ .

Segundo [Röver], para uma série observada  $x_t$  assume-se que ela pode ser decomposta em uma parte parametrizada com vetor de parâmetros e outra um ruído. A decomposição seria

$$x_t = s_t(\theta) + \varepsilon_t$$

se, e somente se, a transformada discreta de Fourier é

$$\tilde{x}_\omega = \tilde{s}_\omega(\theta) + \tilde{\varepsilon}_\omega,$$

ou seja, a aditividade vale no domínio do tempo e da frequência.

A função de verossimilhança correspondente é dada por

$$p(x|\theta) \propto \exp\left(-\frac{1}{2} \sum_j \frac{|\tilde{x}(w_j) - \tilde{s}_{w_j}(\theta)|^2}{\frac{n\pi}{2} f(\omega_j)}\right)$$

A maximização da verossimilhança é feita no pacote “bspec”. Um estudo de simulação é feito no Capítulo 6 para verificar a performance desse método comparado com outros.

## 4.3 Extremograma

O Extremograma é uma ferramenta quantitativa flexível para medir vários tipos de dependência extrema em séries temporais estacionárias. Em muitos aspectos o Extremograma pode ser visto como um análogo da função de autocorrelação (FAC) para eventos

extremos. Na análise clássica de séries temporais a função de autocorrelação e suas versões amostrais, são utilizadas para medir dependência linear na família de processos lineares. A FAC pode ser utilizada para medir dependência em séries temporais não lineares ou funções não lineares dos dados, tais como valores absolutos e valores quadráticos. No entanto a FAC tem limitações para obter uma medida de dependência entre eventos extremos. Por outro lado, o Extremograma considera somente observações ou grupos de observações que são grandes. Para definir o Extremograma é necessário introduzir o conceito de *conjunto limitado longe de zero*.

**Definição 4.3.1.** Um conjunto  $C$  é dito ser limitado longe de zero (*bounded away from zero*) se  $C \subset \{y : |y| > r\}$  para algum  $r > 0$  e  $C$  limitado.

**Definição 4.3.2.** Para um processo  $d$ -dimensional estritamente estacionário,  $(x_t)$ , o Extremograma é definido, para dois conjuntos  $A$  e  $B$  limitados longe de zero, por

$$\varrho_{A,B}(h) = \lim_{s \rightarrow \infty} P(s^{-1}x_h \in B \mid s^{-1}x_0 \in A), \quad (4.10)$$

desde que esse limite exista.

Como  $A$  e  $B$  são limitados longe de zero, os eventos  $\{s^{-1}x_h \in B\}$  e  $\{s^{-1}x_0 \in A\}$  são extremos no sentido de que a probabilidade converge a zero quando  $s \rightarrow \infty$ . Particularmente, para séries temporais univariadas, se  $A = B = (1, \infty)$ , então o Extremograma é o coeficiente de dependência caudal (superior) entre  $x_0$  e  $x_h$  frequentemente utilizado em teoria de valor extremo [McNeil, 1999].

Para estimação de (4.10) o limite em  $s$  é substituído por um quantil alto  $a_m$  do processo. Definindo  $a_m = (1 - 1/m)$ -quantil da distribuição estacionária de  $|x_t|$ , a versão amostral do Extremograma, baseado nas observações  $x_1, \dots, x_n$ , é dada por

$$\hat{\varrho}_{A,B}(h) = \frac{\sum_{t=1}^{n-h} I\{a_m^{-1}x_{t+h} \in B, a_m^{-1}x_t \in A\}}{\sum_{t=1}^n I\{a_m^{-1}x_t \in A\}} \quad (4.11)$$

Para que sejam válidas propriedades assintóticas, é necessário que  $m = m_n \rightarrow \infty$ , com  $m_n/n \rightarrow 0$ , quando  $n \rightarrow \infty$ . Na prática não se conhece  $a_m$ , mas podemos substituí-lo pelo quantil empírico.

#### 4.4 Exemplo de aplicação da função `Extremogram1.R`

No pacote *Extremogram* do R, existem algumas possibilidades para calcular o Extremograma. Uma delas é a função *Extremogram1.R*. Para entendermos o procedimento para calcular a estimativa do correlograma para eventos extremos, rodamos um exemplo simples que consiste em definir alguns parâmetros necessários para a utilização da função, que são  $level = quantile(x, prob = 0.7)$ , que define o quantil 0.7 como um ponto de corte para um evento extremo, ou seja, todos os valores superiores ao quantil 0.7 de  $x$ , são considerados pontos extremos, a saber, qualquer valor de  $x$  maior que  $level = 0.2880798$  é considerado um valor extremo nesse exemplo. A função “Extremogram1.R” tem o seguinte código

```
{ rhat[1] = 1
  for ( i in 1:(maxlag-1)){
    rhat[i+1]=length((1:(n-i))[x[1:(n-i)] > level & x[(i+1):n]> level])      (1)
    rhat[i+1]=rhat[i+1]/length((1:n)[x[1:(n-i)]>level])                        (2)
  }
}
```

Apresentamos o passo a passo do exemplo na seguinte tabela:

i	x	0.231	0.225	1.032	0.419	-0.425	-1.455	-0.971	-0.502	-0.929	0.724	(1)	(2)
1	x[(i+1):n] x[1:(n-i)]	0.225	1.032	<b>0.419</b>	-0.420	-1.455	-0.971	-0.502	-0.929	0.724		1	0.500
2	x[(i+1):n] x[1:(n-i)]	1.032	0.419	-0.420	-1.455	-0.971	-0.502	-0.929	0.724			0	0
⋮													
7	x[(i+1):n] x[1:(n-i)]	-0.502	-0.929	<b>0.724</b>								1	1

**Tabela 4.1:** *Debug* da função `Extremogram1.R`

Na tabela, para  $i = 1$ , por exemplo, contamos as colunas em que ambas as linhas ultrapassam o valor do  $level = 0.2880798$ , que no caso ocorre só na terceira coluna, e dividimos pelo número de valores maior que 0.2880798 na segunda linha, que no caso são 2 valores (1.032 e 0.410). Assim,  $\hat{\rho}(1)$  representado pela coluna (2) é  $\frac{1}{2} = 0.5$ . Desta forma, continuamos para  $i = 2, 3, \dots, 7$ .

Durante a verificação da simulação, em que geramos os dados a partir de uma distribuição normal padrão, verificamos a existência de um pequeno *bug* na rotina do código. A lógica apresentava uma falha a partir de um determinado ponto, o que não se repetia se mudássemos a amostra. Foi reportado ao desenvolvedor do pacote, juntamente com uma

sugestão de correção da rotina para que a possibilidade de falha fosse eliminada. A sugestão de modificar “ $(1 : (n))[x[1 : (n - i)] > level]$ ” por “ $(1 : (n - i))[x[1 : (n - i)] > level]$ ” foi aceita.

Algumas propriedades amostrais do extremograma são exploradas em [Davis et al.].

## 4.5 Extremoperiograma, o periodograma para eventos extremos

O Extremoperiograma, que é um periodograma construído a partir do Extremograma, é introduzido em [Davis et al., 2009] e [Davis et al., 2012]. Para definir o Extremoperiograma e apresentar alguns resultados sobre este estimador, é necessário definir a seguinte condição de mistura/dependência de uma sequência  $(x_t)$ .

Condição (M): A sequência  $(x_t)$  é fortemente misturado em função da taxa  $(\xi_t)$ . Se existem  $m = m_n \rightarrow \infty$  e  $r_n \rightarrow \infty$  tal que  $m_n/n \rightarrow 0$  e  $r_n/m_n \rightarrow 0$  e

$$\lim_{n \rightarrow \infty} m_n \sum_{h=r_n}^{\infty} \xi_h = 0, \quad (4.12)$$

e para todo  $\epsilon > 0$ ,

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} m_n \sum_{h=k}^{r_n} P(|x_h| > \epsilon a_m, |x_0| > \epsilon a_m) = 0. \quad (4.13)$$

Para  $A = [1, \infty)$ , escrevendo  $I_t = I_{\{x_t/a_m \in A\}}$ ,  $\tilde{I}_t = I_t - p_0$ ,  $p_0 = EI_t = P(a_m^{-1}x \in A)$   $t = 1, \dots, n$  para alguma sequência  $m = m_n \rightarrow \infty$ , tal que  $m_n/n \rightarrow 0$  como na condição (M) acima, pode-se introduzir os estimadores

$$I_{nA}(\omega) = \frac{m_n}{n} \left| \sum_{t=1}^n \tilde{I}_t e^{-it\omega} \right|^2, \quad \omega \in [0, \pi] \quad e \quad \hat{P}_m(A) = \frac{m_n}{n} \sum_{t=1}^n I_t. \quad (4.14)$$

Segue do Teorema 3.1 em [Davis et al., 2009] que

$$\hat{P}_m(A) = \frac{m_n}{n} \sum_{t=1}^n I_t \xrightarrow{L^2} \mu_0(A) = \lim_{n \rightarrow \infty} m_n P(a_m^{-1}x \in A), \quad (4.15)$$

desde que  $A$  é um  $\mu_0$ -mensurável, com respeito a medida limitada  $\mu_0$ . As condições

$m_n \rightarrow \infty$  e  $m_n/n \rightarrow 0$  não podem ser ignoradas, pois é necessário que  $E\hat{P}_m(A) = m_n P(a_m^{-1}x \in A) \rightarrow \mu_0(A)$  e então obtemos  $Var\hat{P}_m(A) = O(m_n/n)$ . Em [Davis et al., 2009], Teorema 5.1, também prova que o estimador lag-window ou periodograma truncado

$$\hat{f}_{nA}(\omega) = \tilde{\gamma}_n(0) + 2 \sum_{h=1}^{r_n} \cos(\omega h) \tilde{\gamma}_n(h) \quad (4.16)$$

com  $\tilde{\gamma}_n(0) = (m/n) \sum_{t=1}^n I_t$ , e  $\tilde{\gamma}_n(h) = (m/n) \sum_{t=1}^{n-h} \tilde{I}_t \tilde{I}_{t+h}$ ,  $h > 0$ , para  $\omega \in (0, \pi)$  fixo, satisfazendo as relações

$$E\hat{f}_{nA}(\omega) \rightarrow \mu_0(A) f_A(\omega) \quad e \quad E(\hat{f}_{nA}(\omega) - \mu_0 f_A(\omega))^2 \rightarrow 0. \quad (4.17)$$

Sob a condição (M), se  $A$  é um  $\mu_0$ -mensurável e o conjunto  $Ax \bar{R}_0^{h-1} x A$  são contínuos com respeito a  $\mu_h$ ,  $h \geq 1$ ,  $m_n r_n^2 = O(n)$  e  $\bar{R}_0^{h-1}$  representa o espaço  $(h-1)$ -dimensional dos reais estendidos, com respeito a  $\mu_h$ , que está relacionado com a convergência de  $\hat{\rho}(h)$ . Se combinarmos (2.6) e (2.8) teremos para  $\omega \in (0, \pi)$  fixo,

$$\frac{\hat{f}_{nA}(\omega)}{\hat{P}_m(A)} \xrightarrow{P} f_A(\omega) \quad (4.18)$$

Um estimador natural auto-normalizado da densidade espectral  $f_A(\omega)$  é o seguinte análogo do periodograma

$$\tilde{I}_{nA}(\omega) = \frac{I_{nA}(\omega)}{\hat{P}_m(A)} = \frac{|\sum_{t=1}^n \tilde{I}_t e^{-it\omega}|^2}{\sum_{t=1}^n I_t}, \quad \omega \in [0, \pi]. \quad (4.19)$$

Em contraste com  $\hat{f}_{nA}(\omega)$ , não é preciso conhecer as quantidades  $m_n$  e  $r_n$  que aparecem na definição de  $\hat{f}_{nA}(\omega)$  e são difíceis de determinar para fins de estimativas práticas. Denotamos  $\tilde{I}_{nA}(\omega)$  o periodograma padronizado. No entanto, sabemos da teoria para periodograma clássico de um processo estacionário  $(x_t)$  não é um estimador consistente da densidade espectral  $f(\omega)$  do processo  $(x_t)$  mesmo no caso em que  $(x_t)$  é i.i.d. e tem variância finita, por exemplo, a proposição 10.3.2 em [Brockwell and Davis, 2013]. Para conseguirmos um estimador consistente de  $f(\omega)$  necessitamos o truncamento do periodograma, similarmente a  $\hat{f}_{nA}(\omega)$ , ou aplicar técnicas de suavização na vizinhança ordenada do periodograma. Uma similar observação aplica-se ao periodograma para eventos extremos. Neste trabalho não foi utilizado o suavizado.

## 5 Séries temporais com outliers

Segundo [Valk, 2011] as observações que desviam significativamente da média são usualmente referidas na literatura como outliers ou valores atípicos. Uma prática comum era eliminar essas observações dos conjuntos de dados, pois assumia-se que eram impurezas devido a erros de anotações. Esse assunto ganhou maior notoriedade a partir de estudos como de [Anscombe, 1960] e atualmente existe uma variedade de aplicações, como detecção de fraudes, detecção de invasão de sistemas, diagnósticos de falhas, [Ferdousi and Maeda, 2006] e [Maruyama and Matsuoka, 2010].

A partir do trabalho de [Fox, 1972], a relevância dos outliers aumentou consideravelmente na área de séries temporais. Nesse trabalho foram introduzidos os conceitos de outlier do tipo I e do tipo II, conhecidos como outliers aditivos e outliers de inovação. Posteriormente, os outliers também começaram a ser considerados no desenvolvimento da teoria denominada análise de intervenção. Essa teoria é constituída de métodos que lidam com séries temporais que foram afetadas por acontecimentos inesperados e incontroláveis, como, por exemplo, greves, mudanças políticas, guerras e catástrofes ambientais, como os derramamentos de petróleo e secas. Essa área de estudo foi desenvolvida primeiramente por [Box and Tiao] e, posteriormente, [Tiao, 1985], entre outros.

Na análise de séries temporais, outliers podem comprometer os procedimentos usuais de modelagem, o que pode levar à identificação incorreta do modelo e a estimação viesada dos parâmetros do modelo. [Fajardo et al.] estudam o impacto de outliers aditivos nas estimativas de parâmetros de modelos estacionários, especialmente o modelo ARMA.

### 5.1 Outliers

[Box and Tiao] propõem um modelo de intervenção mais geral que inclui três tipos

de intervenções importantes em séries temporais, que são: mudanças de nível, outliers de inovação e outliers aditivos. No entanto, para este trabalho consideramos somente o modelo desenvolvido por [Fox, 1972], o qual abrange os efeitos dos outliers aditivos.

O outlier aditivo é caracterizado por ser ele a única observação afetada. Um modelo para este tipo de outlier, descrito por [Fox, 1972], é

$$u_t = \sum_{i=1}^p \phi_i u_{t-i} + \varepsilon_t, \text{ para } t = 1, \dots, n, \quad (5.1)$$

em que  $\phi_i$ , para  $i = 1, \dots, p$ , são os parâmetros autorregressivos e  $\varepsilon_t$  são i.i.d.  $N(0, \sigma_\varepsilon^2)$ . As observações  $y_t$  são tais que

$$y_t = \begin{cases} u_t, & \text{se } t \neq T, \\ u_T + \Delta, & \text{se } t = T. \end{cases} \quad (5.2)$$

Assume-se que a série não tem tendência e os  $\phi_i$ 's são tomados de tal forma que o processo  $\{u_t\}_{t \in \mathbb{Z}}$  seja estacionário. A ordem  $p$  da regressão é assumida conhecida. É possível testar se  $y_T$ , para um determinado valor de  $T$ , é um outlier ou testar se existe algum outlier na amostra  $y$ . Estes procedimentos podem ser encontrados em [Fox, 1972].

Outro tipo de outlier comum é o outlier de inovação. [Fox, 1972] definiu o seguinte modelo para estes outliers

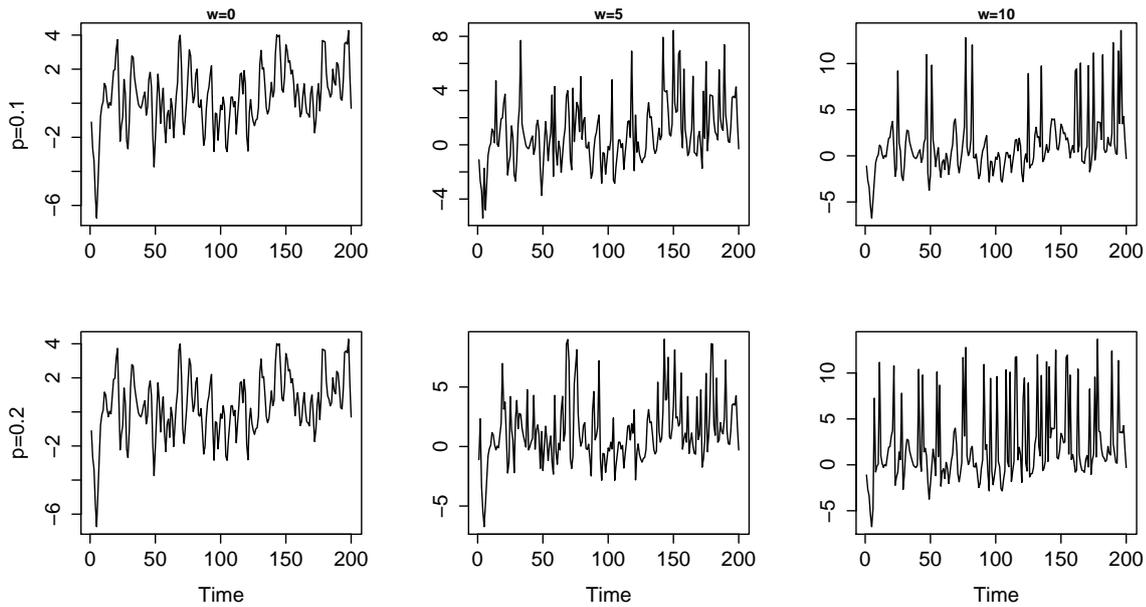
$$y_t = \sum_{i=1}^p y_{t-i} + \Delta_t + \varepsilon_t \quad (5.3)$$

em que

$$\Delta_t = \begin{cases} 0, & \text{se } t \neq T, \\ \Delta, & \text{se } t = T. \end{cases} \quad (5.4)$$

e os  $\phi_i$ 's e  $\varepsilon_t$  são definidos como em (5.1). Então, o outlier  $\Delta$  afeta, não somente a observação  $y_T$ , mas também as observações subsequentes  $y_{T+1}, \dots, y_n$ .

Na Figura 5.1, apresentamos os efeitos de outliers aditivos e de outliers de inovação em um modelo ARMA(1,1), com parâmetro autorregressivo  $\phi_1 = 0.5$  e com parâmetro de média móvel  $\theta = 0.8$ . Adicionamos outliers de magnitude 0, 5 e 10, com  $p_j$  definidos em (5.6), dados por  $p_j = 0.10$  e  $p_j = 0.20$  e com inovações normalmente distribuídas.



**Figura 5.1:** Efeito de outliers aditivos em um processo ARMA(1,1), com tamanho amostral  $n=200$  e para magnitude igual a 0, 5 e 10.

## 5.2 Outliers aditivos e modelos estacionários

Um tópico bastante abordado na literatura é a detecção de outliers. Observações com valores extremos podem afetar a modelagem e por este motivo é importante saber o grau de impacto de outliers sobre as estimativas dos parâmetros. [Fajardo et al.] estudam o impacto de outliers sobre as estimativas dos parâmetros de modelos estacionários. Embora o caso abordado seja o modelo ARFIMA, uma decomposição interessante do periodograma para séries temporais com outliers é apresentada. Essa decomposição é interessante pela alta dependência da magnitude do outlier em um de seus termos. No trabalho de [Fajardo et al.] são considerados processos  $\{u_t\}_{t \in \mathbb{Z}}$  fracamente estacionários e processos  $\{y_t\}_{t \in \mathbb{Z}}$  contaminados por outliers aditivos, os quais podem ser descritos por

$$y_t = u_t + \sum_{j=1}^m \delta_j x_{j,t} \quad (5.5)$$

em que  $m$  é o número máximo de outliers, o parâmetro  $\delta_j$  representa a magnitude do  $j$ -ésimo outlier e  $x_{j,t}$  ( $\equiv x_j$ ) é uma variável aleatória com distribuição de probabilidade

$$Pr(x_j = -1) = Pr(x_j = 1) = p_j/2 \text{ e } Pr(x_j = 0) = 1 - p_j, \quad (5.6)$$

em que  $\mathbb{E}X_j = 0$  e  $\mathbb{E}X_j^2 = \text{Var}X_j = p_j$ . Esses outliers afetam a função de densidade espectral, as autocorrelações e, conseqüentemente, o periodograma, que pode ser escrito como

$$I_y(\omega) = I_u(\omega) + \Delta(\delta), \quad \omega \in (-\pi, \pi], \quad (5.7)$$

em que  $I_y(\omega)$  é o periodograma da série  $y$ ,  $I_u$  é o periodograma de  $u_t$ ,

$$\Delta(\delta) = \frac{\delta^2}{2\pi n} \pm \frac{\delta}{\pi n} \left\{ (u_T - \bar{u}) + \sum_{h=1}^{n-1} (u_{T-h} + u_{T+h} - 2\bar{u}) \cos(h\omega) \right\} + o_p(n^{-1}),$$

e  $T$  é o tempo em que ocorre o outlier de magnitude  $\delta$ .

## 6 Estudo de simulação

Embora o objetivo do trabalho seja estimar robustamente a função de densidade espectral, vale lembrar que existe uma relação direta entre a densidade espectral e a função de autocovariância. Por esse motivo fizemos um estudo de alguns estimadores da função de autocorrelação em diferentes modelos com contaminação. A contaminação é assumida ser aditiva, como definida na Seção 5.2, com probabilidade de contaminação variando em  $\{0.00, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30\}$ . Consideramos somente um tipo de outlier com magnitude fixada em  $\delta = 10$ . Foram considerados três estimadores, Fac Padrão, Fac Robusta definida em (4.6) e o Extremograma definido em (4.11). As séries foram simuladas com tamanho  $n = 2000$  e foram feitas 100 replicações Gaussianas.

### 6.1 Estimação da função de autocorrelação de modelos ARMA com outliers aditivos

Na Tabela 6.2 e Figura 6.1 apresentamos os resultados da simulação para a estimativa da função de autocorrelação em um processo ARMA(2,1) com coeficientes  $\phi_1 = -0.4$ ,  $\phi_2 = 0.4$  e  $\theta = 0.8$ . Comparamos com a autocorrelação teórica desse modelo, a qual é especificada na equação (3.5). Podemos observar que sem contaminação tanto a Fac Robusta e a Fac padrão estimam bem a Fac teórica, porém na medida de 10% de contaminação e lag's pequenos somente a Fac Robusta estima razoavelmente a Fac teórica, acima disto nenhuma estima razoavelmente. Além disso observa-se que a estimativa do Extremograma em todas as situações deixa a desejar. Note que na Figura 6.1 quando temos 20% de contaminação a Fac Robusta superestima a Fac teórica, porém com 30% de contaminação ela subestima, isto é devido ao fato que Fac Robusta é baseada em quantis, e

**Tabela 6.1:** Vício e EQM para três estimadores da função de autocorrelação considerando o modelo AR(1) com coeficiente  $\phi = -0.6$ .

Pob.Cont.		h=1		h=2		h=3		h=4		h=5	
		Vício	EQM								
0.00	FAC	0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	0.00	0.00	0.00
	ROB	0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	0.00	0.00
	EXT	-0.36	0.13	0.28	0.08	-0.13	0.02	0.10	0.01	-0.04	0.00
0.05	FAC	0.45	0.20	-0.27	0.07	0.16	0.03	-0.10	0.01	0.06	0.00
	ROB	0.01	0.00	0.01	0.00	0.02	0.00	0.01	0.00	0.02	0.00
	EXT	-0.36	0.13	0.23	0.05	-0.12	0.02	0.09	0.01	-0.04	0.00
0.10	FAC	0.51	0.26	-0.31	0.10	0.18	0.03	-0.11	0.01	0.06	0.00
	ROB	0.07	0.01	0.04	0.00	0.08	0.01	0.06	0.01	0.08	0.01
	EXT	-0.36	0.13	0.24	0.06	-0.13	0.02	0.09	0.01	-0.04	0.00
0.15	FAC	0.53	0.29	-0.32	0.10	0.19	0.04	-0.12	0.01	0.07	0.01
	ROB	0.23	0.06	0.05	0.00	0.21	0.05	0.11	0.01	0.17	0.03
	EXT	-0.36	0.13	0.25	0.06	-0.13	0.02	0.09	0.01	-0.04	0.00
0.20	FAC	0.55	0.30	-0.33	0.11	0.20	0.04	-0.12	0.01	0.07	0.01
	ROB	0.56	0.32	-0.05	0.00	0.28	0.08	0.05	0.00	0.18	0.03
	EXT	-0.36	0.13	0.25	0.06	-0.13	0.02	0.09	0.01	-0.04	0.00
0.25	FAC	0.55	0.31	-0.33	0.11	0.20	0.04	-0.12	0.02	0.07	0.01
	ROB	0.63	0.40	-0.22	0.05	0.28	0.08	-0.04	0.00	0.15	0.02
	EXT	-0.36	0.13	0.25	0.07	-0.13	0.02	0.09	0.01	-0.04	0.00
0.30	FAC	0.56	0.31	-0.34	0.11	0.20	0.04	-0.12	0.01	0.07	0.01
	ROB	0.59	0.36	-0.38	0.15	0.27	0.07	-0.10	0.01	0.13	0.02
	EXT	-0.36	0.13	0.25	0.07	-0.13	0.02	0.10	0.01	-0.04	0.00

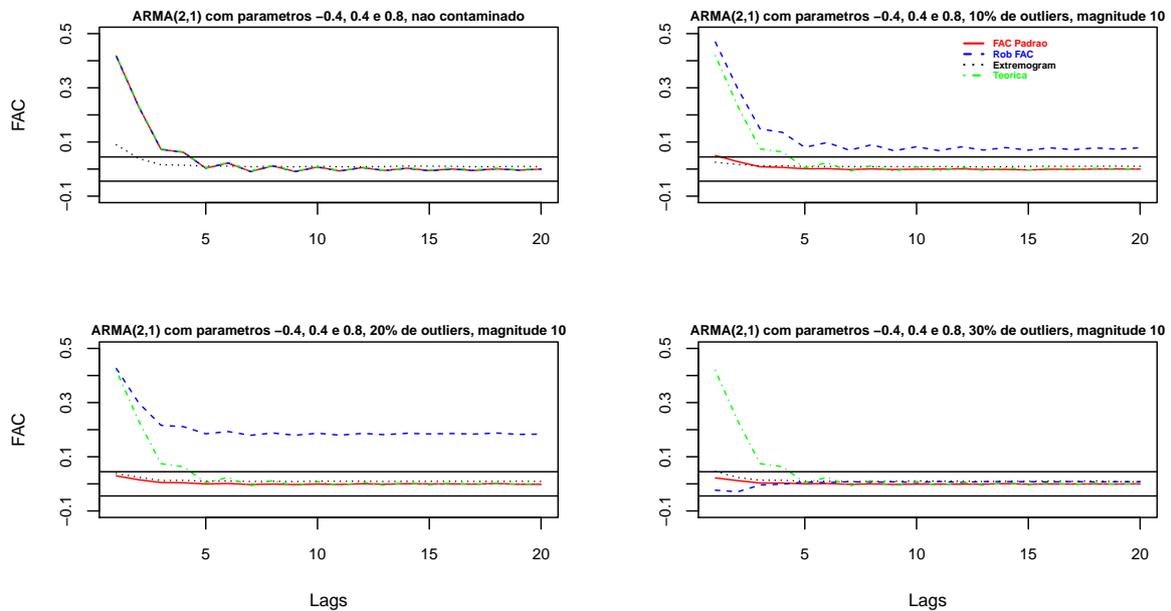
provavelmente que o *breakdown point*<sup>1</sup> deve estar entre 20% e 30%. Como é conhecido que o processo ARMA(2,1) possui curta dependência e isto se traduz em decaimento rápido da autocorrelação podemos observar que para h=4 ou 5 a diferença entre os estimadores diminui, mas isto se deve a quase inexistência de autocorrelação para h grande.

---

<sup>1</sup>É o percentual de contaminação para o qual a partir deste a estimativa fica viesada e apresenta alta variabilidade.

**Tabela 6.2:** Vício e EQM para três estimadores da função de autocorrelação considerando o modelo ARMA(2,1) com coeficientes  $\phi_1 = -0.4$ ,  $\phi_2 = 0.4$  e  $\theta = 0.8$ .

Pop.Cont.		h=1		h=2		h=3		h=4		h=5	
		Vício	EQM								
0.00	FAC	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00
	ROB	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00
	EXT	-0.14	0.02	-0.04	0.00	-0.05	0.00	0.01	0.00	-0.01	0.00
0.05	FAC	-0.33	0.11	-0.18	0.03	-0.06	0.00	-0.05	0.00	-0.00	0.00
	ROB	0.01	0.00	0.01	0.00	0.02	0.00	0.02	0.00	0.02	0.00
	EXT	-0.21	0.05	-0.06	0.00	-0.05	0.00	0.01	0.00	-0.01	0.00
0.10	FAC	-0.37	0.14	-0.21	0.04	-0.07	0.00	-0.06	0.00	-0.00	0.00
	ROB	0.05	0.00	0.06	0.00	0.07	0.01	0.07	0.01	0.08	0.01
	EXT	-0.21	0.04	-0.06	0.00	-0.05	0.00	0.01	0.00	-0.01	0.00
0.15	FAC	-0.38	0.15	-0.21	0.05	-0.07	0.01	-0.06	0.00	-0.00	0.00
	ROB	0.08	0.01	0.13	0.02	0.16	0.03	0.17	0.03	0.18	0.03
	EXT	-0.20	0.04	-0.06	0.00	-0.05	0.00	0.01	0.00	-0.01	0.00
0.20	FAC	-0.39	0.15	-0.22	0.05	-0.07	0.01	-0.06	0.00	-0.00	0.00
	ROB	0.01	0.00	0.07	0.01	0.14	0.02	0.15	0.02	0.18	0.03
	EXT	-0.19	0.04	-0.05	0.00	-0.05	0.00	0.01	0.00	-0.01	0.00
0.25	FAC	-0.39	0.16	-0.22	0.05	-0.07	0.01	-0.06	0.00	-0.00	0.00
	ROB	-0.18	0.04	-0.09	0.01	0.02	0.00	0.03	0.00	0.08	0.01
	EXT	-0.19	0.04	-0.05	0.00	-0.05	0.00	0.01	0.00	-0.01	0.00
0.30	FAC	-0.40	0.16	-0.22	0.05	-0.07	0.01	-0.06	0.00	-0.00	0.00
	ROB	-0.44	0.20	-0.26	0.07	-0.08	0.01	-0.06	0.01	0.00	0.00
	EXT	-0.19	0.04	-0.05	0.00	-0.05	0.00	0.01	0.00	-0.01	0.00



**Figura 6.1:** Fac, Fac robusta, Extremograma e Fac teórica para um processo ARMA(2,1) com coeficientes  $\phi_1 = -0.4$ ,  $\phi_2 = 0.4$  e  $\theta = 0.8$ .

## 6.2 Estimação da função de autocorrelação em modelos GARCH ao quadrado

Como descrito na Seção 3.2 os processos GARCH não possuem autocorrelação, mas quando consideramos o quadrado de um processo GARCH, essa característica aparece. Segundo [Morettin and Toloi, 2006] dado um GARCH(p,q) e definindo

$$X_t = \sqrt{h_t} \varepsilon_t, \quad (6.1)$$

$$X_t^2 = \alpha_0 + \sum_{i=1}^{q^*} (\alpha_i + \beta_i) X_{t-i}^2 + v_t - \sum_{j=1}^q \beta_j v_{t-j} \quad (6.2)$$

onde  $\varepsilon_t$  i.i.d.  $(0,1)$ ,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$ ,  $\sum_{i=1}^{q^*} (\alpha_i + \beta_i) < 1$ ,  $q^* = \max(p, q)$  e  $v_t = X_t^2 - h_t$ .

Segue que no caso do GARCH(1,1), temos

$$E(X_t^2) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$$

com  $E(X_t) = 0$ . Assim,

$$Var(X_t^2) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1},$$

ou seja,

$$\gamma_0 = \frac{\alpha_0}{1 - \alpha - \beta}.$$

e, portanto,

$$\gamma_h = \rho_h \cdot \gamma_0$$

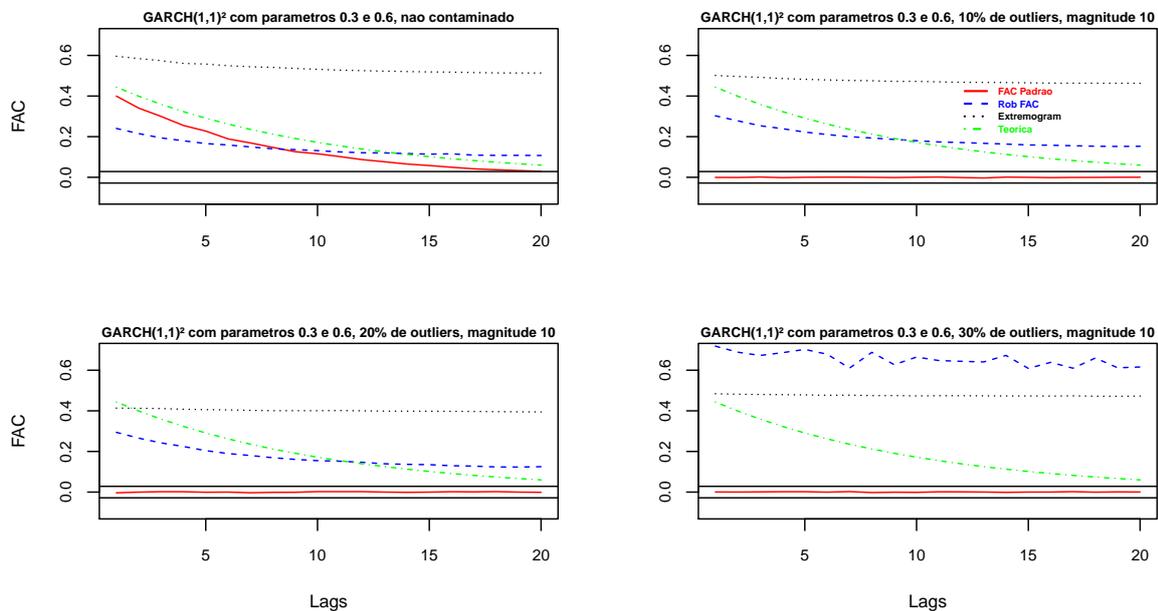
uma vez que  $\rho_h = \gamma_h / \gamma_0$ . Desta forma, podemos definir a função densidade espectral para o GARCH(1,1) ao quadrado.

A Figura 6.2 mostra as estimativas da Fac de um processo GARCH(1,1) ao quadrado com coeficientes  $\alpha_1 = 0.3$  e  $\beta_1 = 0.6$  obtidas a partir de três estimadores distintos, que são, o estimador padrão da Fac, a Fac robusta e o Extremograma.

Analisando a Figura 6.2 para dados não contaminados, temos que a melhor estimação é obtida pelo estimador padrão da Fac, porém ao contaminarmos com outliers ela não consegue estimar a autocorrelação do GARCH ao quadrado, enquanto que a Fac robusta

**Tabela 6.3:** Vicio e EQM para três estimadores da função de autocorrelação considerando o modelo GARCH(1,1) ao quadrado com coeficientes  $\phi_1 = 0.3$  e  $\theta = 0.6$ .

Pop.Cont.		h=1		h=2		h=3		h=4		h=5	
		Vicio	EQM								
0.00	FAC	-0.09	0.01	-0.10	0.02	-0.10	0.01	-0.10	0.02	-0.10	0.01
	ROB	-0.25	0.06	-0.23	0.05	-0.21	0.04	-0.18	0.03	-0.16	0.03
	EXT	0.15	0.07	0.19	0.09	0.22	0.10	0.24	0.12	0.27	0.13
0.05	FAC	-0.49	0.24	-0.44	0.20	-0.40	0.16	-0.35	0.13	-0.33	0.11
	ROB	-0.23	0.05	-0.20	0.04	-0.18	0.03	-0.15	0.02	-0.13	0.02
	EXT	0.09	0.09	0.12	0.09	0.16	0.11	0.19	0.12	0.22	0.13
0.10	FAC	-0.49	0.24	-0.44	0.20	-0.40	0.16	-0.36	0.13	-0.32	0.10
	ROB	-0.19	0.04	-0.17	0.03	-0.15	0.02	-0.12	0.01	-0.10	0.01
	EXT	0.06	0.08	0.10	0.08	0.13	0.09	0.16	0.10	0.19	0.11
0.15	FAC	-0.49	0.24	-0.44	0.20	-0.40	0.16	-0.36	0.13	-0.33	0.11
	ROB	-0.17	0.03	-0.14	0.02	-0.12	0.01	-0.10	0.01	-0.08	0.01
	EXT	0.03	0.07	0.07	0.07	0.10	0.08	0.14	0.09	0.17	0.10
0.20	FAC	-0.50	0.25	-0.44	0.20	-0.40	0.16	-0.36	0.13	-0.32	0.11
	ROB	-0.20	0.04	-0.18	0.03	-0.16	0.03	-0.13	0.02	-0.12	0.01
	EXT	-0.03	0.07	0.01	0.06	0.05	0.07	0.08	0.07	0.11	0.08
0.25	FAC	-0.49	0.24	-0.44	0.20	-0.40	0.16	-0.36	0.13	-0.32	0.10
	ROB	-0.48	0.24	-0.45	0.21	-0.45	0.21	-0.43	0.19	-0.41	0.18
	EXT	-0.01	0.06	0.03	0.06	0.07	0.06	0.10	0.07	0.13	0.08
0.30	FAC	-0.49	0.24	-0.44	0.20	-0.40	0.16	-0.36	0.13	-0.32	0.10
	ROB	0.23	0.50	0.25	0.57	0.27	0.59	0.33	0.60	0.38	0.61
	EXT	0.04	0.06	0.08	0.07	0.12	0.08	0.16	0.09	0.19	0.10



**Figura 6.2:** Fac, Fac robusta, Extremograma e Fac teórica para um processo GARCH(1,1) ao quadrado com coeficientes  $\alpha_1 = 0.3$  e  $\beta_1 = 0.6$ .

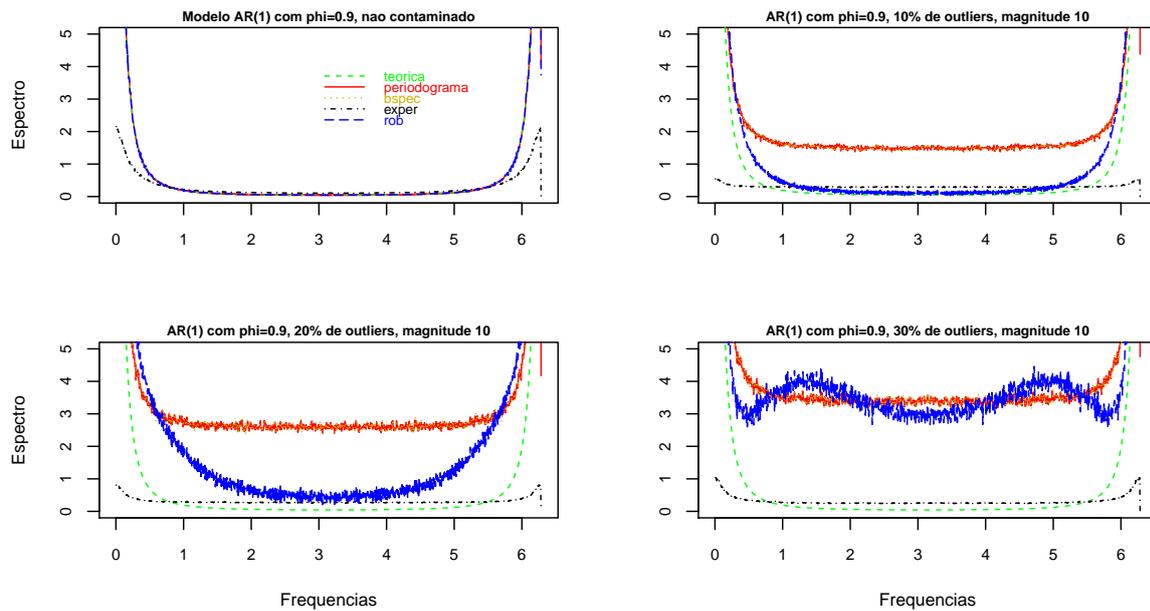
consegue estimar razoavelmente com contaminações de 10% e 20% e tendo o seu ponto de ruptura entre 20% e 30%. Por outro lado esperávamos muito do Extremograma, pois em

[Davis et al., 2009] e [Davis et al., 2012] comenta-se que ele seria apropriado para captar cluster de volatilidade, ou seja, cluster de eventos extremos, então acreditamos que teria alguma característica interessante nos processos GARCH ou no GARCH ao quadrado. Apesar de estarmos trabalhando com GARCH ao quadrado o comportamento é similar ao do GARCH, observamos que ele captura aproximadamente uma constante em todos os gráficos, por exemplo no gráfico com 30% de contaminação, a constante é de 0,5 que indica que acontecido um evento extremo a chance de sair um novo evento extremo no próximo passo é de 50% e isto se mantém para um segundo, terceiro, e sucessivos passos, que é estranho e não condiz com a Fac teórica. Na Tabela 6.3 observamos uma piora na estimação do Extremograma tanto no vício quanto no EQM com o aumento do lags “ $h$ ”, visto que o extremograma é praticamente constante, enquanto a FAC teórica decai. Em contra partida o FAC e a robusta melhoram com o aumento dos lags, que é um comportamento esperado devido a parte AR embutida no GARCH, exceto a robusta a 30%, visto que ocorre a sua ruptura.

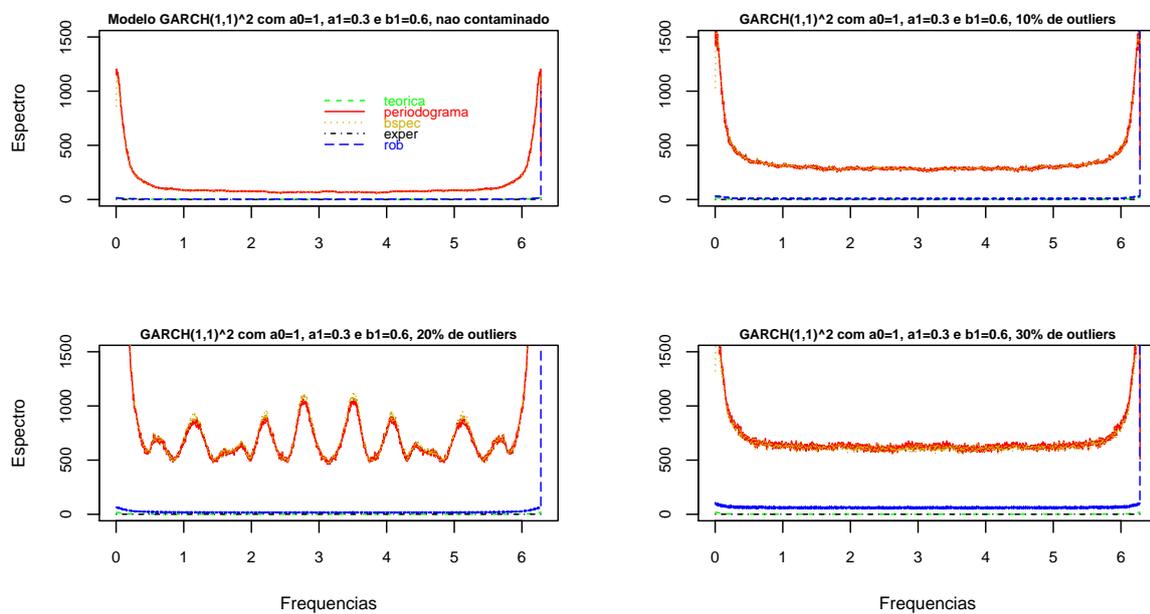
### 6.3 Estimação da função de densidade espectral

Observando a Figura 6.3, notamos que no gráfico do AR(1) com  $\phi = 0.9$  sem contaminação as estimações robusta (azul), o periodograma (vermelho) e o bspec (laranja) ficam muito próximas da teórica, já o Extremoperiodograma (preto) é o pior na estimação sem contaminação e com contaminações 10%, 20% e 30% ele fica praticamente o mesmo gráfico sem alterações. Isto mostra que os outliers não causam efeito relevante sobre ele. Já com contaminação de 10% e 20% o que se mantém mais próximo da teórica é a estimação robusta, enquanto que a Bspec e periodograma que andam juntas se perdem a medida que aumentamos a contaminação. Por fim observando a contaminação de 30% é visto que tanto o periodograma, Bspec e a robusta não estimam razoavelmente, sendo que esta última deve ter passado seu ponto de ruptura.

Na Figura 6.4 temos os gráficos da GARCH(1,1) ao quadrado com coeficientes  $\alpha_0 = 1$   $\alpha_1 = 0.3$  e  $\beta_1 = 0.6$ , sendo o primeiro sem contaminação e os outros 10%, 20% e 30% de contaminação de outliers, respectivamente. Observamos que o periodograma e o Bspec que andam aproximadamente juntos, não são bons estimadores para esta modelo e a medida que aumentamos a contaminação pior eles estimam, é de se estranhar no caso de contaminação de 20% em que oscilam muito, mas nas simulações com 15% eles iniciam



**Figura 6.3:** Função densidade espectral teórica, Periodograma, Periodograma Robusto, Periodograma Bayesiano e Extremoperiodograma de um processo AR(1) com coeficiente  $\phi = 0.9$ .



**Figura 6.4:** Função densidade espectral teórica, Periodograma, Periodograma Robusto, Periodograma Bayesiano e Extremoperiodograma de um processo GARCH(1,1) ao quadrado com coeficientes,  $\alpha_0 = 1$ ,  $\alpha_1 = 0.3$  e  $\beta_1 = 0.6$ .

a oscilação tendo seu auge com 20% e depois com 25% diminui a oscilação. A estimação robusta se mantém muito bem em todos os gráficos, somente na contaminação de 30% que ela desvia-se um pouco da estimação teórica. O Extremoperiodograma não apresenta resultados relevantes, como observa-se no Apêndice que este possui praticamente todas as estimativas próximo de zero, coerente com o que foi observado no Extremograma.

## 7 Resultados e conclusões

Neste trabalho apresentamos o conceito de periodicidade em séries temporais, caracterizando a noção de período, frequência e definimos uma medida para periodicidade, conhecida como função densidade espectral. Modelos clássicos para séries temporais foram definidos para que fosse explorada a forma explícita da função de densidade espectral. Apresentamos alguns estimadores da densidade espectral, que são o periodograma padrão, o Bspec baseado em uma abordagem bayesiana, uma versão robusta do periodograma baseado em quantis e o Extremoperiodograma, que é baseado no Extremograma que é um correlograma para eventos extremos. Além disso, abordamos o conceito de séries temporais contaminadas por outliers, especialmente outliers aditivos.

O objetivo inicial do trabalho foi estudar o comportamento de alguns estimadores da função de densidade espectral em séries temporais contaminadas por outliers e também em séries temporais não Gaussianas, em que existe uma maior probabilidade de ocorrência de eventos extremos, típico de processos ARCH e GARCH. Verificamos a performance desses estimadores na presença de outliers aditivos variando os percentuais de contaminação, aplicando por exemplo no AR(1), ARMA(2,1) e GARCH(1,1) ao quadrado.

Outro objetivo foi estudar o estimador Extremograma, que é apresentado como um correlograma para eventos extremos, isto é, ocorrido um evento extremo o quanto ele influenciaria a nova ocorrência de outro evento extremo. No entanto, nos casos considerados no estudo, o estimador não apresentou bom desempenho. Era esperado que, em processos GARCH e GARCH ao quadrado, este estimador apresentasse padrões e características típicas dos processos GARCH. Contudo, usando os pacotes propostos pelos autores, feito as comparações, onde estimador Extremograma supostamente deveria ser eficaz, este mostrou-se ineficiente. Um padrão singular de comportamento observado, é que ele é quase constante no decorrer do tempo, principalmente no modelo GARCH, ou

seja, dado que aconteceu um evento extremo a probabilidade de sair outro evento extremo é aproximadamente constante no tempo, que é de difícil interpretação. Além disso, os autores comentam que o Extremograma necessita de amostra grande. Por esse motivo, realizamos uma simulação com uma amostra de tamanho  $n = 20000$ , porém continuou apresentando a mesma ineficiência. Os resultados podem ser encontrados no Apêndice.

Por outro lado o estimador robusto se mostrou eficaz pois aguentou relativamente bem até uma contaminação de 20% em todas as situações estudadas, possivelmente por ser um estimador baseado em quantis. A partir do momento que temos muitos outliers, é claro que teremos um quantil alto, fazendo com este estimador apresente baixa eficiência, mostrando um ponto de ruptura, que em alguns casos verificamos que deve estar por volta de 25% a 30%.

Já o estimador Bspec, que também é baseado em um pacote que exploramos apresentou um comportamento muito similar ao periodograma padrão. Apesar de não analisarmos suas propriedades e de como é feita a sua estimação, o fato deste apresentar um comportamento muito similar ao periodograma, nos levou a não investirmos estudos mais profundos sobre o Bspec.

Como futuros trabalhos poderemos estudar mais detalhadamente o ponto de ruptura da estimação robusta, buscar uma explicação para a oscilação do periodograma no GARCH(1,1) ao quadrado quando temos uma contaminação de 20% e tentar responder por qual motivo o Extremograma apresenta uma decaimento lento, quase que constante. Essas foram situações que ficaram momentaneamente sem resposta. Ainda no estimador Bspec, não detectamos a possibilidade de escolher uma priori, ficando esta indagação para novos estudos.

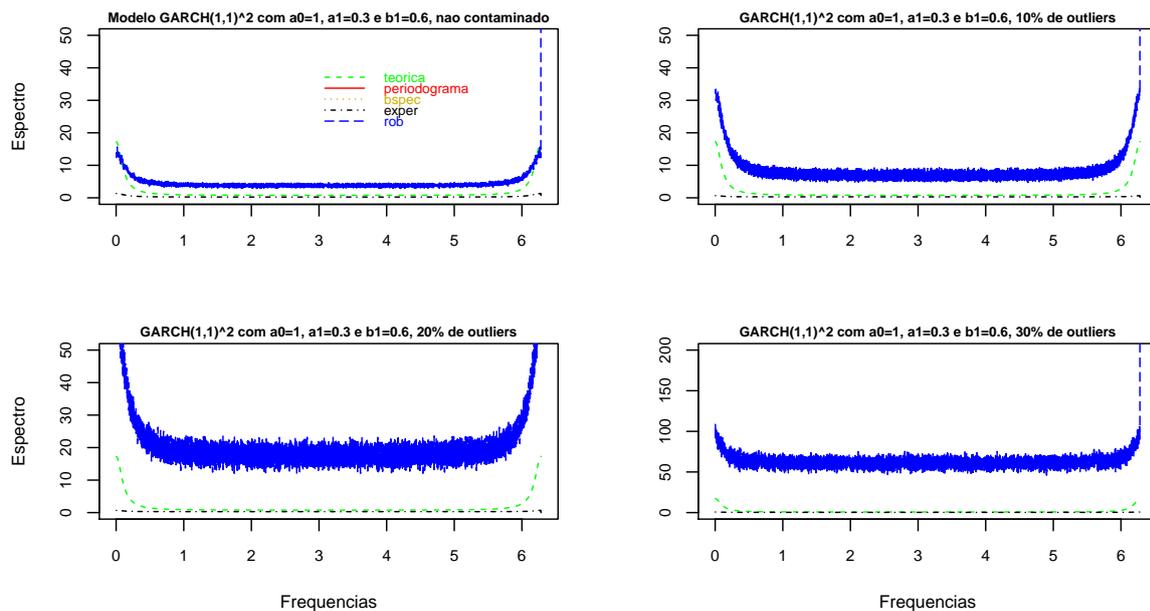
## Referências

- F. J. Anscombe. Rejection of outliers. *Technometrics*, 2:123–147, 1960.
- R. T. Baillie and H. Chung. Estimation of garch models from the autocorrelations of the squares of a process. *Journal of Time Series Analysis*, 22(6):631–650, 2001.
- T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.
- G. E. P. Box and G. C. Tiao. Intervention analysis with applications to economic and environmental problems.
- P. J. Brockwell and R. A. Davis. *Time series: theory and methods*. Springer-Verlag, 1991.
- P. J. Brockwell and R. A. Davis. *Time series: theory and methods*. Springer Science & Business Media, 2013.
- R. A. Davis, T. Mikosch, and I. Cribben. Estimating Extremal Dependence in Univariate and Multivariate Time Series via the Extremogram. *ArXiv e-prints*.
- R. A. Davis, T. Mikosch, et al. The extremogram: A correlogram for extreme events. *Bernoulli*, 15(4):977–1009, 2009.
- R. A. Davis, T. Mikosch, and I. Cribben. Towards estimating extremal serial dependence via the bootstrapped extremogram. *Journal of Econometrics*, 170(1):142–152, 2012.
- R. De Losso. *Econometria de Séries Temporais*. Cengage, 2011.
- R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

- F. Fajardo, V. Reisen, and F. Cribari-Neto. Robust estimation in long-memory processes under additive outliers.
- Z. Ferdousi and A. Maeda. Unsupervised outlier detection in time series data. In *Proceedings of the 22Nd International Conference on Data Engineering Workshops*, pages 121–. IEEE Computer Society, 2006.
- A. J. Fox. Outliers in time series. *Journal of the Royal Statistical Society. Series B. Methodological*, 34:350–363, 1972.
- Y. Ma and M. G. Genton. Highly robust estimation of the autocovariance function. *Journal of time series analysis*, 21(6):663–684, 2000.
- N. Maruyama and S. Matsuoka. Model-based fault localization: Finding behavioral outliers in large-scale computing systems. *New Generation Computing*, 28(3):237–255, 2010.
- A. J. McNeil. Extreme value theory for risk managers. *Departement Mathematik ETH Zentrum*, 1999.
- P. A. Morettin and C. Toloï. *Análise de séries temporais*. Blucher, 2006.
- C. Röver. Student-t based filter for robust signal detection. *Physical Review D*, 84(12).
- C. Röver, R. Meyer, and N. Christensen. Modelling coloured residual noise in gravitational-wave signal processing. *Classical and Quantum Gravity*, 28(1).
- R. H. Shumway and D. S. Stoffer. *Time series analysis and its applications: with R examples*. Springer Science & Business Media, 2010.
- G. Tiao. Autoregressive moving average models, intervention problems and outlier detection in time series. *Handbook of statistics*, 5:85–118, 1985.
- M. Valk. O uso de quase u-estatísticas para séries temporais uni e multivariadas. *Unicamp*, 2011.

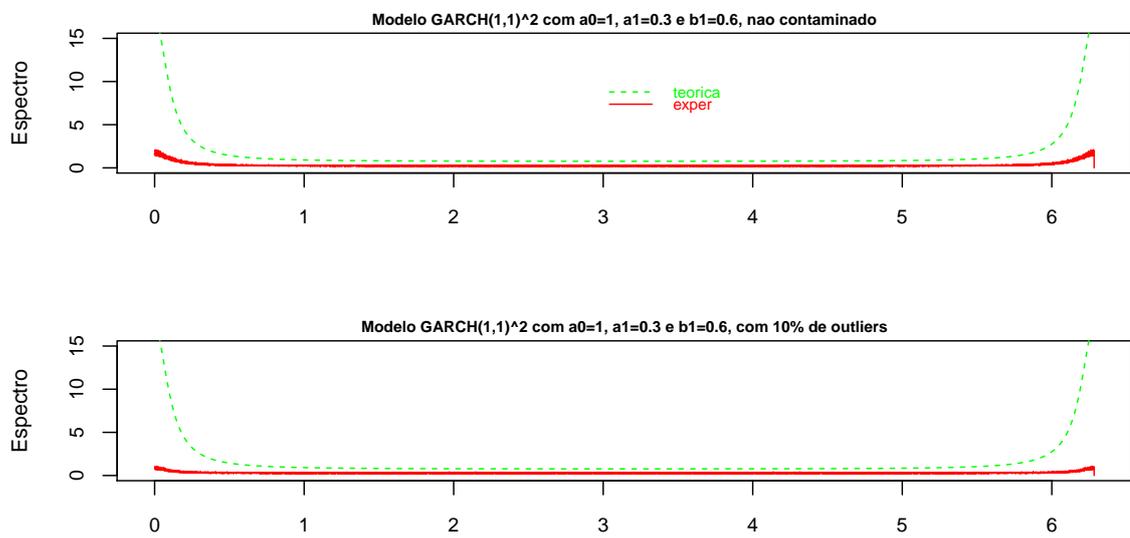
## Apêndice

A figura 7.1 é basicamente um "zoom" da figura 6.4 e mostra que o extremograma e o estimador robusto são os que melhor estimam a densidade espectral de um GARCH ao quadrado. No entanto é visível que essa estimativa não é satisfatória, especialmente quando temos contaminação.



**Figura 7.1:** Função densidade espectral teórica, Periodograma, Periodograma Robusto, Periodograma Bayesiano e Extremoperiodograma de um processo GARCH(1,1) ao quadrado com coeficientes,  $\alpha_0 = 1$ ,  $\alpha_1 = 0.3$  e  $\beta_1 = 0.6$ .

A figura 7.2 mostra que mesmo para tamanho amostral muito grande, que é uma suposição para o bom desempenho do Extremo Periodograma, ele não consegue capturar a informação em um processo GARCH(1,1) ao quadrado.



**Figura 7.2:** Função densidade espectral teórica e Extremoperiodograma de um processo GARCH(1,1) ao quadrado com coeficientes,  $\alpha_0 = 1$ ,  $\alpha_1 = 0.3$  e  $\beta_1 = 0.6$  com tamanho amostral  $n=20000$ .