

Séries Temporais, Aula 3 Minicurso

10 de outubro de 2013

I congreso Internacional de Estadística. Trujillo-Perú. 2013 I Congreso

1 Introdução

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 têm dependência e estamos interessados em analisar e modelar esta dependência. Enquanto em modelos de regressão por exemplo a ordem das observações é irrelevante para a análise, em séries temporais a ordem dos dados é crucial. Vale notar também que o tempo pode ser substituído por outra variável como espaço, profundidade, etc (Ehlers, 2007)

Como a maior parte dos procedimentos estatísticos foi desenvolvida para analisar observações independentes o estudo de séries temporais requer o uso de técnicas específicas. Dados de séries temporais surgem em vários campos do conhecimento como Economia (preços diários de ações, taxa mensal de desemprego, produção industrial), Medicina (eletrocardiograma), Epidemiologia (número mensal de novos casos de meningite), Meteorologia (precipitação pluviométrica, temperatura diária, velocidade do vento), etc.

Os objetivos principais da análise de séries temporais é investigar o mecanismo gerador da série, descrever o comportamento, e identificar a presença de periodicidade ou sazonalidade e finalmente fazer previsões de valores futuros da série.

2 Pacotes e funções para série temporais no R

R contém várias pacotes e funções que são usadas para análise de séries temporais.

1. **forecast**: pacote para modelos de séries temporais
2. **dse**: pacote para modelos de séries temporais multivariado (**dse1**, para estimação de sistemas dinâmicos; **dse2**, para extensão de sistemas dinâmicos como o métodos de raiz específica).
3. **acf()** e **pacf()**: função para as funções de autocorrelação e função de autocorrelação parcial respectivamente.
4. **auto.arima()** do pacote **forecast**, **arima.sim()** do pacote padrão **stats**, e **arimax()** do pacote **TSA**: estas funções são utilizadas para identificar modelos ARMA, ARMAX, ARIMA, SARIMA.
5. **fracdiff**: pacote para modelos de memória longa ou Arima fracionados (ARFIMA).
6. **HoltWinters()**: função para predição e ajustamento de algoritmos de HoltWinters com e sem sazonalidade (Aditivo e multiplicativo), métodos de suavização.
7. **uroot**: pacote para testes de raiz untaria como Dickey-Fuller, Phillips-Perron, Schmidt e Phillips, Hylleberg-Engle-Granger-Yoo, etc.
8. **tseries**: pacote para séries temporais e analise financeira, com modelos garch, e modelos arima, testes de Jarque -Bera, e testes de estacionaridade, etc.
9. **urca**: pacote para testar modelos de raiz unitária (Dicker e Fuller, Schmidt e Phillips, Phillips e Perron, e outros).

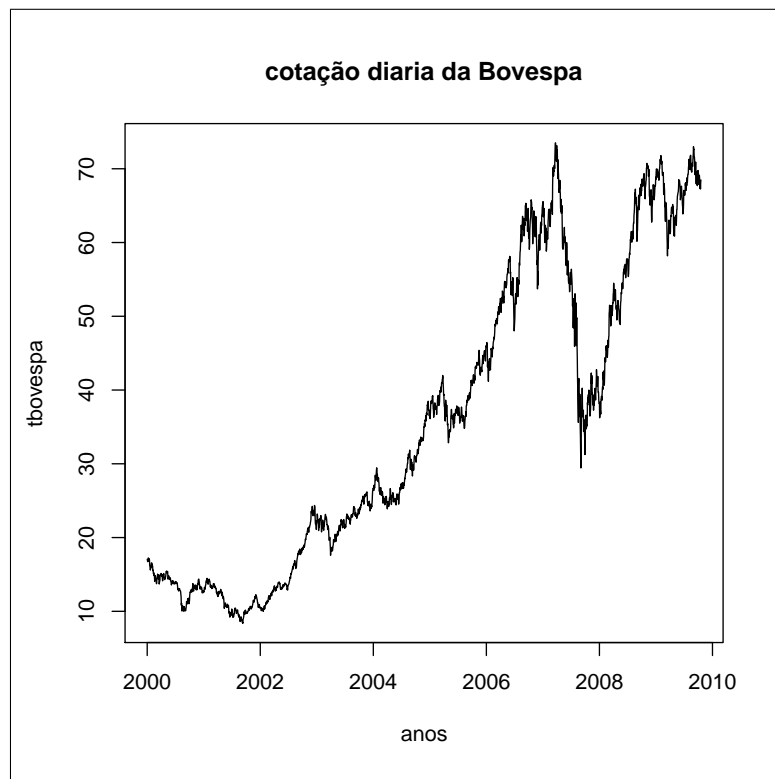
10. **seriesChaos**: função para modelos de séries temporais não linear, com sistema lorenz, ajuste recursivo, e espaço temporal.
11. **dynlm**: Pacote para Modelos Lineares Dinâmicos e Regressão com Séries Temporais.

3 Fontes

Dados para Séries temporais podem ser encontradas em páginas confiáveis na internet. Por exemplo Séries Econômicas do Brasil são encontradas em <http://www.ipeadata.gov.br>, outras séries temporais podem ser encontrados em fontes como o IBGE, no sitio: <http://www.ibge.gov.br>. Como exemplo, considere os dados baixados do ipeadata, correspondente a cotação dos dias úteis dos índices da BOVESPA, durante os anos de 2000 – 2010. Assumindo que há em média 245 dias úteis no Brasil, os dados podem se chamados no R da seguinte maneira.

```
> bov=read.table("bovespa.txt",header=T) # dados em formato txt
> tbovespa=ts(bov,start=c(2000,1),frequency=245)
> plot.ts(tbovespa, main="cotação diaria da Bovespa", xlab="anos")
```

Figura 27. Série da Bolsa de valores de São Paulo.



Um segundo exemplo pode ser os índices econômicos mensais do Brasil no período de 1980 até abril de 2011, denominado de **decon.txt** e disponíveis no **ipeadata**, com as seguintes variáveis:

1. **anomes**; corresponde ao ano e o mês da informação coletada,
2. **tcc**: taxa média de cambio do dólar compra, em reais,
3. **tcv**: taxa média de cambio do dólar venda, em reais,
4. **bc**: balança comercial em US (Dólares),
5. **exp**: exportações FOB em US (Dólares),

6. smr: salario mínimo real (IPEA) em reais,
7. igpd: índice geral de preços distribuição interna (igpdi).

Estas séries podem ser baixados um formato xls, simultaneamente do site do IPEA. A chamada dos dados e os gráficos no R e dada a seguir,

```
> ie=read.table("decon.txt",header=T)
> dim(ie)
[1] 376 7
> attach(ie)
> names(ie)
[1] "anomes" "tcc" "tcv" "bc" "exp" "smr" "igpd"
> ie.ts=ts(ie,start=c(1980,1),frequency=12)
> plot.ts(ie.ts, main="Indicadores Econômicos do Brasil",xlab="tempo")
```

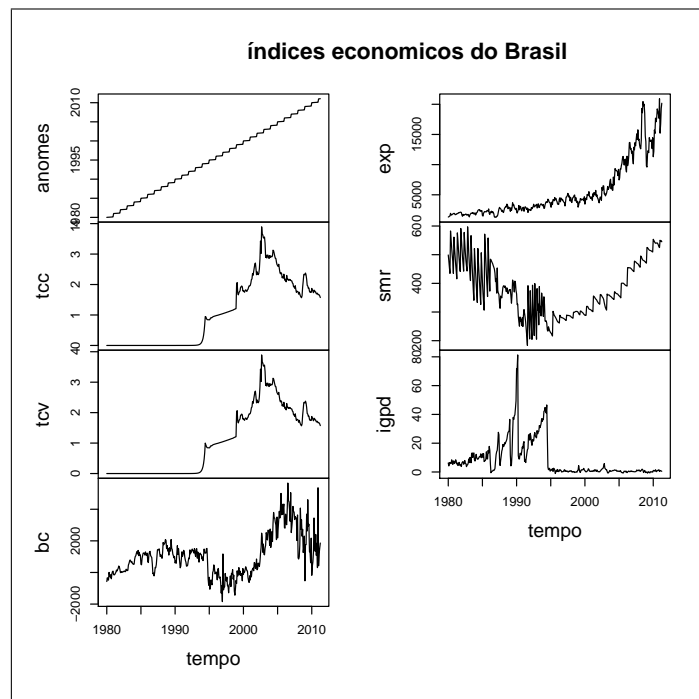


Figura 28. Série individuais do banco de dados ie.

Para obter uma serie de ie, podemos fazer:

```
> plot.ts(bc, main="Balança Comercial do Brasil em US$")
```

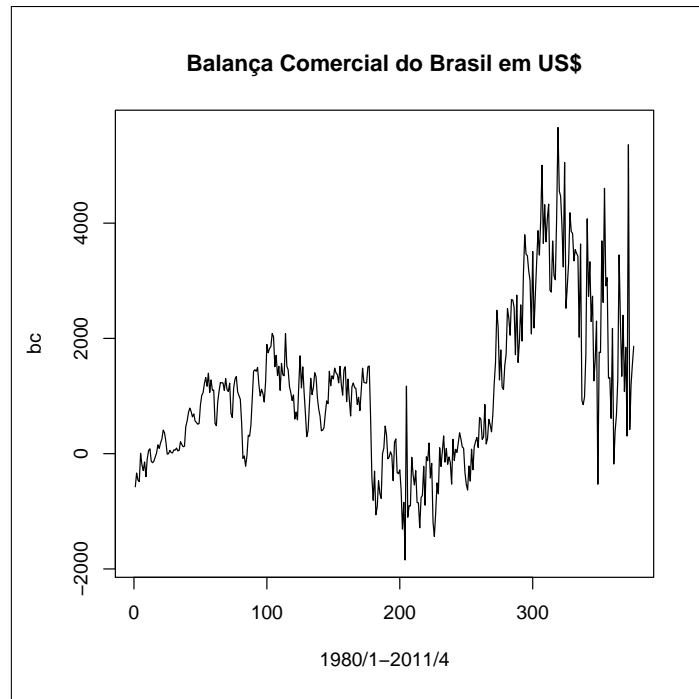


Figura 29. Uma série extraída do banco de dados ie.

4 Regressão linear com séries temporais

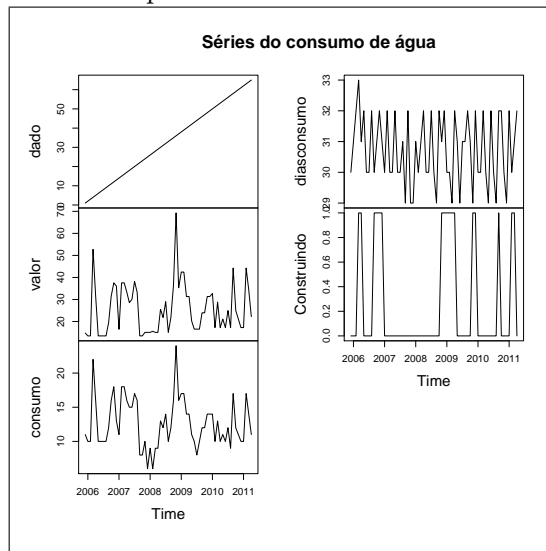
Em econometria frequentemente usamos o método de mínimos quadrados para estimar parâmetros em modelos de regressão e ajustar um conjunto de observações, usando a função **lm()**, porém podemos construir convenientemente dados do tipo séries temporais usando a função **ts()**. O objetivo de melhorar o ajuste das observações, levando em consideração o efeito temporal. Algumas propriedades são necessárias para implementar no formato de série temporal, como a data de início da série (*start*), ou o final da série, e que tipo ou período que a série apresenta, (observações diárias, mensais, anuais).

Para usar modelos dinâmicos e relacionar nossa série temporal com os *lags* (retardos da série ou observações anteriores), podemos usar uma função **dynlm()**, do pacote **dynlm**. Para exemplificar, consideremos o valor de consumo no instante t o qual pode estar relacionado com o valor

de consumo no instante imediatamente anterior $t - 1$ (mês anterior), da seguinte maneira: $valor_t = \beta_1 + \beta_2 diasconsumo_t + \beta_3 valor_{t-1}$, ou um modelo com as primeiras diferenças da seguinte forma: $valor_t - valor_{t-1} = \beta_1 + \beta_2(consumo_t - consumo_{t-1})$. Para exemplificar, primeiro usamos a função `ts()`:

```
> ca.ts=ts(ca,start=c(2005,12),frequency=12)
> plot(ca.ts,main="Séries do consumo de água")
```

Figura 30. Plot para as variáveis do consumo de água



Para construir os modelos descritos, temos

```
> mc1=dynlm(formula = valor ~ diasconsumo + L(valor))
> mc2=dynlm(d(valor)~d(consumo))
> summary(mc1)
```

Time series regression with "numeric" data:

Start = 1, End = 65

Call:

```
dynlm(formula = valor ~ diasconsumo + L(valor))
```

Residuals:

Min	1Q	Median	3Q	Max
-1.022e-14	-2.383e-16	1.106e-16	5.037e-16	3.328e-15

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	8.774e-15	5.205e-15	1.686e+00	0.09690 .
diasconsumo	-4.706e-16	1.718e-16	-2.740e+00	0.00801 **
L(valor)	1.000e+00	1.682e-17	5.944e+16	< 2e-16 ***

Residual standard error: 1.498e-15 on 62 degrees of freedom
 Multiple R-squared: 1, Adjusted R-squared: 1
 F-statistic: 1.846e+33 on 2 and 62 DF, p-value: < 2.2e-16

Para visualizar o ganho de ajuste do modelo mc1 se comparado com o modelo m1, temos:

```
> plot.ts(valor)
> lines(fitted(m1),col=4)
> lines(fitted(mc1),col=8)
```

Um segundo exemplo para séries da classe `ts()`, com as características definidas nos indicadores econômicos (`ie.ts`), e usar regressões com as primeiras diferenças de uma variável, por exemplo considere a relação da balança comercial com as exportações, pode ser construído a seguinte regressão temporal:

Modelo 1 (me1).

$$bc_t = \beta_0 + \beta_1 exp_t + \beta_2 exp_{t-1} + \epsilon_t \quad (1)$$

Modelo 2 (me2).

$$bc_t = \beta_1 exp_t + \beta_2 bc_{t-1} + \beta_3 bc_{t-2} + \epsilon_t \quad (2)$$

Os dois modelos gerados no R é:

```
> library(dynlm)
> me1=dynlm(bc~exp+L(exp),data=ie.ts)
Time series regression with "ts" data:
Start = 1980(2), End = 2011(4)
```

Call:

```
dynlm(formula = bc ~ exp + L(exp), data = ie.ts)
```


Residuals:

Min	1Q	Median	3Q	Max
-2963.0	-751.9	115.2	688.9	2805.8

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	151.18904	84.18325	1.796	0.0733 .
exp	0.41509	0.06194	6.701	7.66e-11 ***
L(exp)	-0.24954	0.06272	-3.979	8.34e-05 ***

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Residual standard error: 1044 on 372 degrees of freedom
 Multiple R-squared: 0.3881, Adjusted R-squared: 0.3848
 F-statistic: 118 on 2 and 372 DF, p-value: < 2.2e-16

#####

```
> me2=dynlm(bc~exp+L(bc,1)+L(bc,2)-1,data=ie.ts)
> summary(me2)
```

Time series regression with "ts" data:
 Start = 1980(3), End = 2011(4)

Call:

```
dynlm(formula = bc ~ exp +L(bc, 1) + L(bc, 2) - 1, data = ie.ts)
```

Residuals:

Min	1Q	Median	3Q	Max
-2488.2	-244.6	2.9	262.6	3774.3

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
exp	0.031757	0.007318	4.339	1.84e-05 ***
L(bc, 1)	0.414529	0.046141	8.984	< 2e-16 ***
L(bc, 2)	0.429927	0.046094	9.327	< 2e-16 ***

Residual standard error: 625.3 on 371 degrees of freedom
 Multiple R-squared: 0.8676, Adjusted R-squared: 0.8665
 F-statistic: 810.4 on 3 and 371 DF, p-value: < 2.2e-16

```
> deviance(me1)
[1] 405120372
> deviance(me2)
[1] 145081812
```

Os gráficos por separado das séries balança comercial e exportações, assim como o ajuste dois modelos me1 e me2 com a balança comercial é dado a seguir:

```
> plot(ie.ts[,c("bc","exp")],lty=c(3,1),plot.type="single",ylab="")
> legend("topleft",legend=c("balança","exportações"), lty=c(3,1),
bty="n")
> plot.ts(ie.ts[, "bc"],ylab="balança",xlab="tempo")
> f1=fitted(me1)
> f2=fitted(me2)
> lines(f1,lty=2)
> lines(f2,lty=3)
> legend(1980,5000, legend=c("me1","me2"), lty=2:3)
```

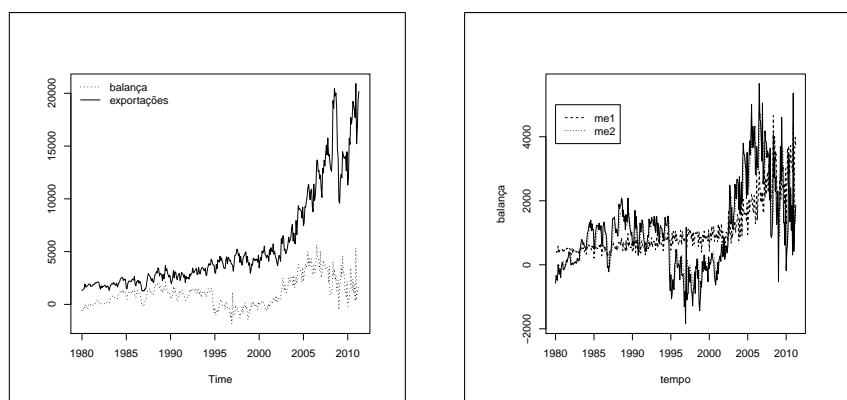


Figura 31. plot de duas séries e ajustes de dos modelos me1 e me2.

5 Deflacionando uma série

Em algumas séries é pertinente o uso da série real, desprovida de qualquer efeito inflacionário. Considere o IGPD de agosto de 1994 até abril 2011 (igp) coletada do IPEADATA (ver dados ie), para realizar este deflacionamento usamos:

```
> attach(ie)
> igp=igpd[-c(1:175)]
> igp.ts=ts(igp, start=c(1994,8),frequency=12)
> igpdi=igp.ts/100
> igpdi=igpdi/igpdi[length(igpdi)]
> plot.ts(igpdi, main="IGPDI do Brasil")
> lines(igp.ts,lty=3)
> legend("topleft", c("IGPDI real", "IGPDI deflacionado"),
lty=1:3, bty="n")
```

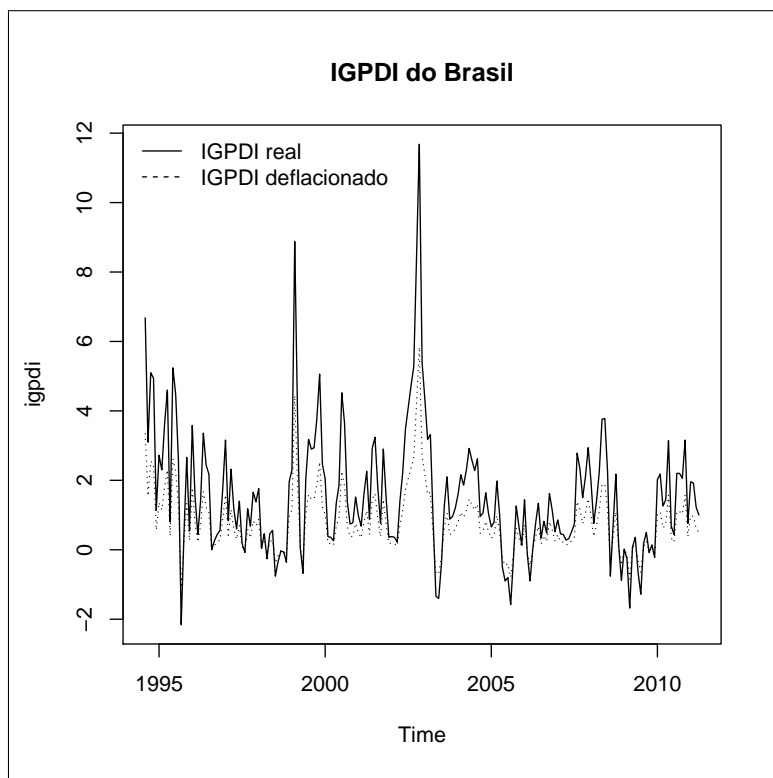


Figura 32. IGPDI - real e deflacionado.

6 Decomposição de uma série

Uma questão fundamental numa ou mais séries temporais é a representação gráfica dos dados sequencialmente ao longo do tempo, pois uma simples visualização pode revelar padrões de comportamento importantes,

como: tendências, ciclos, sazonalidade e aleatoriedade. Dentre as tendências temos as de crescimento ou decrescimento, quando séries apresentam apenas este comportamento, modelos de regressão tem bom desempenho, porem quando a série apresentam padrões cíclicas, outras alternativas de ajuste são pertinentes, os quais serão abordados neste capítulo.

A definição da tendência e sazonalidade é: Seja $X_t, t = 1, 2, \dots, T$, uma série temporal $X_t = T_t + S_t + \epsilon_t$, onde T_t , representa a tendência, S_t , a sazonalidade e ϵ_t , um ruído branco (erro aleatório), com variância σ_ϵ^2 . Inicialmente o gráfico temporal deve ser sempre o primeiro passo e antecede qualquer análise.

6.1 Retornos

Para avaliar os riscos de uma carteira de ativos financeiros, o risco é medido em termos de variações de preços. Considere P_t o preço de um ativo no instante t . A variação de preços entre os instantes $t - 1$ e t é dada por: $\Delta P_t = P_t - P_{t-1}$, e o *retorno líquido simples* deste ativo, no mesmo instante é definido por:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}} \quad (3)$$

R_t , também chamado de *taxa de retorno*.

O *retorno composto* ou simplesmente *retorno* é dado por: $\log(1+R_t)$. na pratica usa-se os retornos que tem propriedades desejáveis como estacionaridade e ergodicidade próprios da modelagem ARMA, ARIMA ou GARCH. exemplo considere a cotação da Bovespa, os retornos e algumas características gráficas destas duas séries, são apresentadas a seguir:

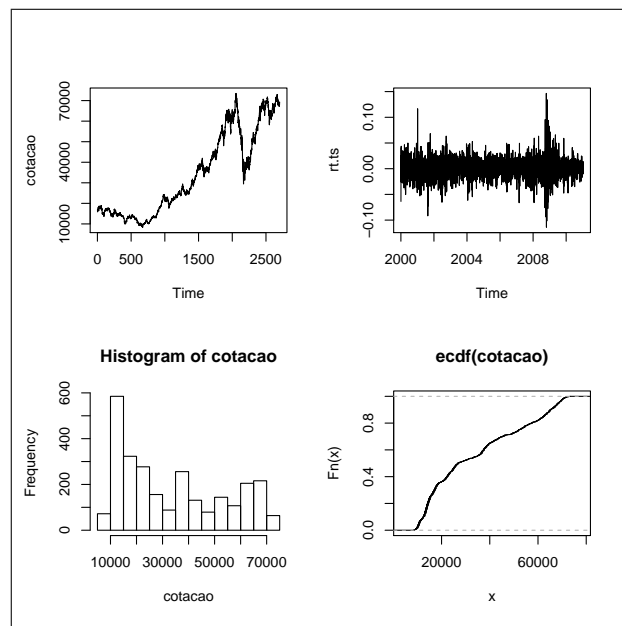


Figura 33. Características da cotação da Bovespa.

```

bov=read.table(file.choose(),header=T) # bovespa15072013.txt

attach(bov)

rt=diff(cotacao)/cotacao[-length(cotacao)]

cotacao.ts=ts(cotacao,end=c(2013,123),frequency=245) #15/07/2013

rt.ts=ts(rt,end=c(2013,123),frequency=245)

par(mfrow=c(2,2))

plot.ts(cotacao.ts)

plot.ts(rt.ts)

hist(cotacao.ts)

plot(ecdf(cotacao.ts))

```

6.2 Decomposição clássica

Muitas das propriedades observadas em uma série temporal X_t podem ser captadas assumindo-se a seguinte forma de decomposição:

$$X_t = T_t + C_t + \epsilon_t \quad (4)$$

onde T_t é uma componente de tendência, C_t é uma componente cíclica ou sazonal e ϵ_t é uma componente aleatória ou ruído (a parte não explicada, que espera-se ser puramente aleatória). A componente cíclica se repete a cada intervalo fixo s , i.e

$$\dots = C_{t-2s} = C_{t-s} = C_t = C_{t+s} = C_{t+2s} = \dots \quad (5)$$

Assim, variações periódicas podem ser captadas por esta componente. Para visualizar a decomposição da cotação e os retornos da Bovespa, usamos a função **decompose()**, ea função **stl()**, da seguinte maneira:

```
dcotacao=decompose(cotacao.ts)

stlcotacao=stl(cotacao.ts,s.window=245)

drt.ts=decompose(rt.ts)

stlrt.ts=stl(rt.ts, s.window=245)

plot(stlcotacao, main="Decomposição da cotação da bovespa")

plot(stlrt.ts, main="Decomposição dos retornos da bovespa")
```

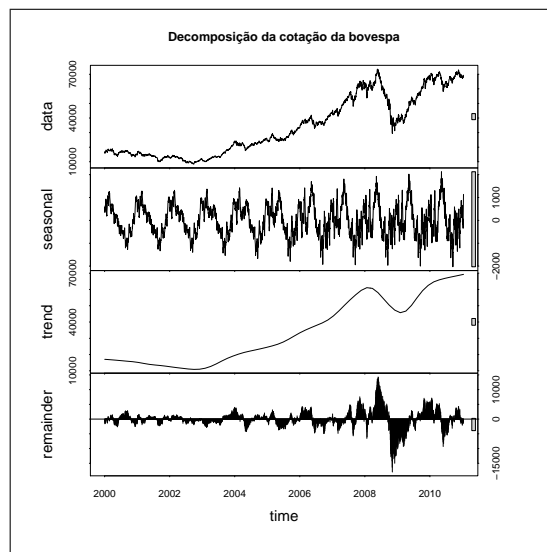


Figura 34. Cotação dos índices da Bovespa.

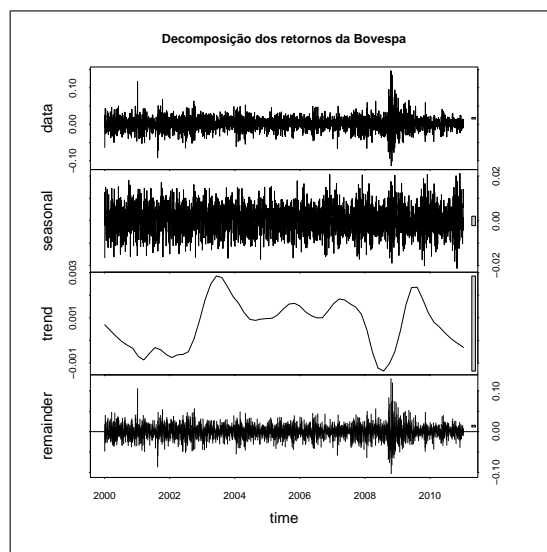


Figura 35. Retorno dos índices da Bovespa.

6.3 Autocorrelação

Uma importante ferramenta para se identificar as propriedades de uma série temporal é o coeficiente de autocorrelação.

Definição 1 Seja $X_t, t \in T / \text{Var}(x_t) < \infty \quad \forall t$. A função de Autocovariância $\gamma_X(.,.)$ de X_t é definida como:

$$\gamma_X(r, s) = E[(X_r - E(X_r))(X_s - E(X_s))], \quad r, s \in T. \quad (6)$$

no exemplo dos retornos da Bovespa, seja $s = r - 1$, então

$$\gamma_X(r, r - 1) = E[(X_r - E(X_r))(X_{r-1} - E(X_{r-1}))]. \quad (7)$$

```
> cov(rt[-length(rt)], rt[-1])
[1] 1.071587e-05
```

se $r = s$, então $\gamma_X(r, s) = \text{Var}(X_r)$.

```
> var(rt)
[1] 0.0005282069
```

Notação alternativa:

$$\gamma(h) = \text{Cov}(X_t, X_{t+h}) = \text{Cov}(X_t, X_{t-h}) \quad (8)$$

Teorema 2 Seja $\gamma(h), h = 0, 1, 2, \dots$, a função de autocovariância de um processo estacionário, então:

1. $\gamma(0) \geq 0$
2. $|\gamma(h)| \leq \gamma(0)$
3. $\gamma(.)$ é par.

Definição 3 Seja $X_t, t \in T$ um processo estacionário. A função de Autocorrelação do processo é $\rho(h) = \text{corr}(X_t, X_{t+h}) = \gamma(h)/\gamma(0)$, e onde $-1 \leq \rho(h) \leq 1$

6.4 O Correlograma

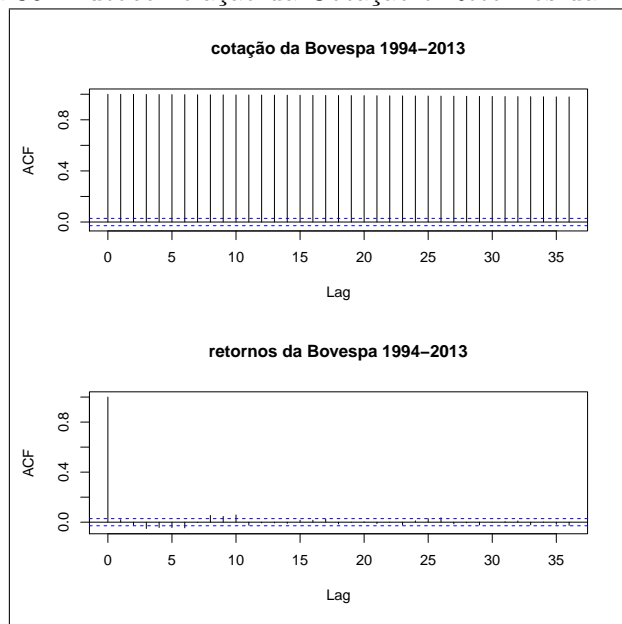
Um gráfico com os k primeiros coeficientes de autocorrelação como função de k é chamado de correlograma e pode ser uma ferramenta poderosa para identificar características da série temporal. Porém isto requer uma interpretação adequada do correlograma, i.e. devemos associar certos padrões

do correlograma como determinadas características de uma série temporal. Esta nem sempre é uma tarefa simples e a seguir são dadas algumas indicações. A primeira questão que podemos tentar responder através do correlograma é se uma série temporal é aleatória ou não. Para uma série completamente aleatória os valores defasados são não correlacionados e espera-se que $\rho(h) = 0$.

Considere as autocorrelações da cotação da Bovespa e os Retornos da Bovespa

```
> par(mfrow=c(2,1))
> acf(bov, main="cotação da Bovespa 1994-2013")
> acf(rt, main="retornos da Bovespa 1994-2013")
```

Figura 36. Autocorrelação da Cotação e Retornos da Bovespa.



6.5 Processos Estacionários

Suposições de estacionaridade (significa flutuação em torno de uma reta) ou transformação de uma série em estacionária é comum em séries temporais.

Definição 4 Um Processo estocástico é uma família de variáveis aleatórias $\{X_t(\omega)\}_{t \in T}$, definidas no mesmo espaço de probabilidade $(\Omega, \mathfrak{F}, P)$, onde $\omega \in \Omega$ e T é um conjunto arbitrário. Aqui, Ω é o espaço amostral, \mathfrak{F} é uma σ -álgebra de Ω e P é uma medida de probabilidade em \mathfrak{F} .

Definição 5 Uma série temporal $X_t, t \in \mathbb{N}$ é dita ser fortemente estacionária se as funções de distribuição $\forall n \in \mathbb{N}$ e $\forall t_1, \dots, t_n, t_{1+n}, \dots, t_{n+h} \in T$ são idênticas, onde T é o conjunto de índices em geral. O conjunto T é comumente tomado como o conjunto dos números inteiros $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$. Uma série temporal é uma realização de um certo processo estocástico. Os dois primeiros momentos de $\{X_t(\omega)\}_{t \in \mathbb{Z}}$ são definidos como

$$E[X_t] = \mu_t \text{ e } E(X_t - \mu_t)^2 = \sigma_t^2,$$

e a correlação é dada por

$$\frac{\text{Cov}(X_t, X_{t+h})}{\sqrt{\sigma_t^2 \sigma_{t+h}^2}} \quad \text{para } h \in \mathbb{Z}, \quad (9)$$

Definição 6 Estacionaridade. Um processo estocástico é dito ser fracamente estacionário se e somente se:

1. $E[X_t] = \mu$, para todo $t \in \mathbb{Z}$,
2. $E(X_t - \mu)^2 = \sigma^2, 0 < \sigma^2 < \infty$, para todo $t \in \mathbb{Z}$,
3. $R(h) = \text{Cov}(X_t, X_{t+h})$ depende apenas de h , para todo $t \in \mathbb{Z}$.

As autocorrelações $\rho(h)$ são obtidas normalizando as autocovariâncias através da sua divisão pelo produto dos respectivos desvios padrão, i.e., $\rho(h) = \frac{R(h)}{R(0)}$. O exemplo mais simples de um processo estacionário é o processo de ruído branco (RB), definido como uma sequência de variáveis aleatórias não-correlacionadas com média constante e variância constante (estritamente positiva e finita) ao longo do tempo.

Definição 7 Processo linear geral pode ser representado por : $\{X_t\}$

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j \epsilon_{t-j}, \quad t \in \mathbb{Z},$$

onde $\{\epsilon_t\} \sim RB(0, \sigma^2)$ e $\{\psi_j\}$ é uma sequência de constantes com $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$.

Definição 8 Função geratriz de autocovariâncias. Seja: $\{X_t\}$, um processo estacionário com função de autocovariâncias $R(h)$ que satisfaz $\sum_{h=-\infty}^{\infty} |R(h)| < \infty$. A função geratriz de autocovariâncias de $\{X_t\}$ é definida como

$$g(z) = \sum_{h=-\infty}^{\infty} R(h) z^h$$

onde z é um escalar complexo.

6.6 Passeio Aleatório

Seja: ϵ_t um processo discreto puramente aleatório com média μ e variância σ^2 . Um processo X_t é chamado de passeio aleatório se $X_t = X_{t-1} + \epsilon_t$.

7 Modelos de séries temporais

O estudo de séries temporais pode ser motivado pelo interesse em investigar o mecanismo gerador de um conjunto de dados observados ao longo do tempo para descrever sua dinâmica com o objetivo de gerar previsões acerca de seu comportamento futuro. Para tanto são construídos modelos probabilísticos que pertencem a um domínio temporal previamente estabelecido [28].

7.1 Alisamento exponencial

É uma classe de algoritmos de previsão, os mais conhecidos são:

7.1.1 Algoritmo simples

A componente não observável deste algoritmo é o nível, o interesse é estimar este nível e projetá-lo para gerar previsões futuras. As estimativas do nível é:

$$N_t = \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + \dots, \quad 0 < \alpha < 1 \quad (10)$$

α é constante de alisamento. A forma de recorrência é:

$$N_t = \alpha y_t + (1 - \alpha)N_{t-1} \quad (11)$$

A forma de correção de erros é:

$$N_t = N_{t-1} + \alpha \epsilon_t \quad (12)$$

Onde ϵ_t : erro de previsão um passo à frente

7.1.2 Algoritmo de Holt

Este algoritmo apresenta duas componentes não observáveis: o nível e a tendência, sendo que esta última não precisa ser globalmente fixa, podendo

evoluir ao longo do tempo. A forma de recorrência é:

$$N_t = \alpha y_t + (1 - \alpha)[N_{t-1} + T_{t-1}] \quad (13)$$

$$T_t = \beta(N_t - N_{t-1}) + (1 - \beta)T_{t-1} \quad (14)$$

Escolhe-se $0 < \alpha, \beta < 1$, que minimizem $\sum \epsilon_t^2$. A forma de correção de erros é:

$$N_t = N_{t-1} + T_{t-1} + \alpha \epsilon_t \quad (15)$$

$$T_t = T_{t-1} + \alpha \beta \epsilon_t \quad (16)$$

7.1.3 Algoritmo de Holt-Winters(HW)

A suposição é que a série possui três componentes não observáveis: nível, tendência e sazonalidade. Há dois enfoques para este algoritmo, o aditivo e o multiplicativo.

7.1.4 HW aditivo

$$N_t = \alpha(y_t - F_{t-s}) + (1 - \alpha)[N_{t-1} + T_{t-1}] \quad (17)$$

$$T_t = \beta(N_t - N_{t-1}) + (1 - \beta)T_{t-1} \quad (18)$$

$$F_t = \gamma(y_t - N_t) + (1 - \gamma)F_{t-s} \quad (19)$$

Escolhe-se $0 < \alpha, \beta, \gamma < 1$. A forma de correção de erros é dado por:

$$N_t = N_{t-1} + T_{t-1} + \alpha \epsilon_t \quad (20)$$

$$T_t = T_{t-1} + \alpha \beta \epsilon_t \quad (21)$$

$$F_t = F_{t-s} + \gamma(1 - \gamma)\epsilon_t \quad (22)$$

7.1.5 HW multiplicativo

$$N_t = \alpha \frac{y_t}{F_{t-s}} + (1 - \alpha)[N_{t-1} + T_{t-1}] \quad (23)$$

$$T_t = \beta(N_t - N_{t-1}) + (1 - \beta)T_{t-1} \quad (24)$$

$$F_t = \gamma \frac{y_t}{N_t} + (1 - \gamma)F_{t-s} \quad (25)$$

Escolhe-se $0 < \alpha, \beta, \gamma < 1$. A forma de correção de erros é dado por:

$$N_t = N_{t-1} + T_{t-1} + \frac{\alpha}{F_{t-s}} \epsilon_t \quad (26)$$

$$T_t = T_{t-1} + \frac{\alpha\beta}{F_{t-s}} \epsilon_t \quad (27)$$

$$F_t = F_{t-s} + \frac{\gamma(1-\gamma)}{N_t} \epsilon_t \quad (28)$$

Uma aplicação de estes algoritmos é dado a seguir:

```
> hwc=HoltWinters(cotacao.ts)
> hwr=HoltWinters(rt.ts)
> hwcm=HoltWinters(cotacao.ts, seasonal="multiplicative")
```

Para os retornos dos índices da Bovespa não é pertinente o uso de algoritmos de Holt-Winters multiplicativo, por apresentar valores zero. Para visualizar gráficos e o resumo dos parâmetros dos modelos acima propostos, usamos:

```
> plot.ts(cotacao.ts)
> lines(fitted(hwc)[,1],col=2)
> f=cbind(hwc,hwr,hwcm)
> f
```

	hwc	hwr	hwcm
fitted	Numeric,9832	Numeric,9828	Numeric,9832
x	Integer,2703	Numeric,2702	Integer,2703
alpha	0.9307564	0.001454172	0.8457155
beta	0.001083805	0.001273331	0
gamma	1	0.2754765	1
coefficients	Numeric,247	Numeric,247	Numeric,247
seasonal	"additive"	"additive"	"multiplicative"
SSE	1572141540	1.249631	2218260761
call	Expression	Expression	Expression

Para realizar previsões futuras da cotação 6 dias à frente, usamos a função: `predict(hwc,6)`.

8 ARMA e ARIMA

Para abordar os processo ARMA ou auto-regressivos e de médias móveis, precisamos identificar algumas propriedades que são fundamentais neste tipo

de processos.

$$\Delta x_t = x_t - x_{t-1} \quad (29)$$

Para uma n -ésima diferença temos

$$\Delta^n x_t = \Delta^{n-1} x_t - \Delta^{n-1} x_{t-1} = \sum_{r=1}^n (-1)^r \binom{n}{r} x_{t-r} \quad (30)$$

Como exemplo considere:

$$\Delta^2 x_t = \Delta x_t - \Delta x_{t-1} = x_t - x_{t-1} - (x_{t-1} - x_{t-2}) = x_t - 2x_{t-1} + x_{t-2}$$

Operador de defassagem (B)

$$Bx_t = x_{t-1}, \quad B^2 x_t = x_{t-2}, \quad B^n x_t = x_{t-n} \quad (31)$$

Na metodologia de Box e Jenkin, é preciso identificar os estágios do ciclo iterativo, os quais são:

1. Uma classe de modelo de séries temporais é especificada (proposta)
2. Uma etapa de identificação a partir das funções de autocorrelação, autocorrelação parcial, função de autocovariância.
3. A fase de estimar os parâmetros do modelo proposto
4. O diagnostico do modelo através de critérios de seleção de modelos, com por exemplo AIC, BIC, estabilidade estrutural do modelo são requeridas

8.1 Processos autorregressivos

Um processo autorregressivo de primeira ordem $AR(1)$ é dado por:

$$x_t = c + \phi x_{t-1} + \nu_t, \quad \nu \approx RB(0, \sigma^2), \quad c, \phi \in \Re \quad (32)$$

Este processo é estacionário se $|\phi| < 1$

O processo autorregressivo de ordem p , $AR(p)$ é dado por:

$$x_t = c + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \nu_t \quad (33)$$

Onde $\nu \approx RB(0, \sigma^2)$, $c, \phi \in \Re$. Uma forma alternativa de escrever estes processos é dado por:

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) x_t = c + \nu_t \quad (34)$$

Centralizando a variável em torno de c , podemos escrever como $\phi(B)x_t = \nu_t$. No caso de um $AR(p)$, é estacionário se todas as raízes de $\phi(z) = 0$ estiverem fora do círculo unitário.

Todo processo autorregressivo é invertível. Exemplo, para identificar um modelo como sendo autorregressivo, Box-Jenkins consideraram que se a função de autocorrelação têm um decaimento exponencial (observar a função de autocorrelação, FAC), somente ϕ é significativo (observar a função de autocorrelação parcial, FACP) á indícios da presença de um processo $AR(1)$. Ver detalhes para identificação de modelos $AR(p)$, em Moretin e Toloi [28].

8.2 Processo de médias móveis

O processo de médias moveis de primeira ordem $MA(1)$ é dado por:

$$x_t = \mu + \theta_1 \epsilon_{t-1} + \epsilon_t, \quad \epsilon \approx RB(0, \sigma^2), \quad \mu, \theta \in \mathbb{R} \quad (35)$$

Este processo é invertível se $|\theta| < 1$.

A autocovariância dada por:

$$\rho_1 = \frac{\theta}{1 + \theta^2} \quad (36)$$

O processo de médias moveis de ordem q , $MA(q)$ é dado por:

$$x_t = \mu + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} \quad (37)$$

Onde $\epsilon \approx RB(0, \sigma^2)$. Uma forma alternativa de escrever estes processos é dado por:

$$(1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \epsilon_t = x_t + \mu \quad (38)$$

Centralizando a variável em torno de μ , podemos escrever como $\theta(B)\epsilon_t = x_t$. No caso de um $MA(q)$, é invertível se todas as raízes de $\theta(z) = 0$ estiverem fora do círculo unitário. Todo processo de médias móveis é estacionário. Exemplo, para identificar se um modelo como sendo de médias moveis $MA(1)$, somente ρ_1 é significativo ou diferente de zero (ver função de autocorrelação) e a função de autocorrelação parcial deve ter um decaimento exponencial dominante. Ver detalhes para identificação de modelos $MA(q)$, em Moretin e Toloi (2006).

8.3 Processo ARMA

Seja $\{X_t\}$ um processo que satisfaz a equação em diferenças dada por

$$\Phi(B)X_t = \Theta(B)\epsilon_t, \quad (39)$$

onde $\{\epsilon_t\}$ é ruído branco, ie., $\{\epsilon_t\} \sim RB(0, \sigma_\epsilon^2)$, B é operador de defasagem definido como $B^m X_t = X_{t-m}$, $m = 1, \dots, p$, $\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p$ e $\Theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q$.

O processo $\{X_t\}$ definido em (39) é chamado processo auto-regressivo e de médias móveis, ARMA(p, q).

Definição 9 *Invertibilidade.* Um processo $\{X_t\}$ com representação ARMA(p, q) é invertível se existem constantes $\{\pi_j\}$ tais que $\sum_0^\infty |\pi_j| < \infty$ e $\epsilon_t = \sum_0^\infty \pi_j X_{t-j}$, para todo $t \in \mathbb{Z}$.

Seguindo as definições 3 e 5 o processo (39) é estacionário e invertível se as raízes de $\Phi(z) = 0$ e $\Theta(z) = 0$ são não comuns e encontram-se fora do círculo unitário.

Definição 10 *Causalidade.* Um processo $\{X_t\}$ com representação ARMA(p, q) é causal, ou função causal de $\{\epsilon_t\}$, se existem constantes $\{\psi_j\}$ tais que $\sum_{j=0}^\infty |\psi_j| < \infty$, e $X_t = \sum_{j=0}^\infty \psi_j \epsilon_{t-j}$ para todo $t \in \mathbb{Z}$.

Note que as propriedades de invertibilidade e causalidade não são apenas do processo $\{X_t\}$, mas também da relação entre os processos $\{X_t\}$ e $\{\epsilon_t\}$ da definição da equação ARMA apresentada em (39). Invertibilidade e causalidade garantem que há uma solução única estacionária, com probabilidade um, para a equação ARMA.

Função de Autocovariâncias e densidade espectral de um processo ARMA(p, q).

O cálculo da função de autocovariância as para um processo $\{X_t\}$ com representação ARMA(p, q) causal é realizado através das equações

$$R(k) - \phi R(k-1) - \dots - \phi R(k-p) = \sigma_\epsilon^2 \sum_{j=0}^\infty \theta_{k+j} \psi_j, \quad 0 \neq k < m,$$

$$R(k) - \phi R(k-1) - \dots - \phi R(k-p) = 0, \quad k \geq m,$$

onde $m = \max(p, q+1)$, $\psi_j - \sum_{k=1}^p \phi_k \psi_{j-k} = \theta_j$, $j = 0, 1, \dots, \psi_j = 0$ para $j < 0$, $\theta_0 = 1$ e $\theta_j = 0$ para $j \in \{0, 1, \dots, q\}$.

O espectro de $\{X_t\}$ é dado por

$$f_{ARMA}(\lambda) = \frac{\sigma_\epsilon^2 |\Theta(\exp^{-i\lambda})|^2}{2\phi |\Phi(\exp^{-i\lambda})|^2}, \quad \lambda \in [-\pi, \pi] \quad (40)$$

8.4 Processo ARIMA(p, d, q)

Seja d um inteiro não negativo. $\{x_t\}$ é um processo auto-regressivo integrado e de médias móveis ARIMA(p, d, q) se $Y_t = (1 - B)^d X_t$ é um processo ARMA(p, q) causal. Esta definição sugere que $\{X_t\}$ satisfaz a equação em diferenças da forma

$$\Phi(B)(1 - B)^d X_t = \Theta(B)\epsilon_t, \quad \{\epsilon_t\} \sim RB(0, \sigma_\epsilon^d).$$

8.4.1 Exemplo 1

Para exemplificar a modelagem arima usaremos a função **auto.arima()** do pacote **forecast**, para a cotação e os retornos dos índices da bovespa, disponíveis em **bov**.

```
> library(forecast)
> bov=read.table(file.choose(),header=T) #bovespa15072013.txt
> attach(bov)

> #Arima para a cotação do Bovespa

> cotarima=auto.arima(cotacao)
> summary(cotarima)
Series: cotacao
ARIMA(2,1,2)

Coefficients:
          ar1      ar2      ma1      ma2
      -0.2896  0.7015  0.2468 -0.7336
s.e.    0.0805  0.0800  0.0758  0.0747

sigma^2 estimated as 0.4172:  log likelihood=-4624.42
AIC=9258.83   AICc=9258.85   BIC=9291.12

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
```

```

Training set 0.01047288 0.6458221 0.4019317 0.03303618 1.582666 0.9999299
              ACF1
Training set 0.009322901
> par(mfrow=c(2,1))
> plot.ts(cotacao)
> lines(fitted(cotarima),col=4)
> plot.ts(cotacao,xlim=c(4700,4730), ylim=c(45,48))
> abline(v=4711,lty=2,col=8)
> previsto=predict(arima(cotacao,order=c(2,1,2)),4)
> previsto
$pred
Time Series:
Start = 4712
End = 4715
Frequency = 1
[1] 46.70404 46.68037 46.66341 46.65171

$se
Time Series:
Start = 4712
End = 4715
Frequency = 1
[1] 0.6458907 0.8940655 1.0797783 1.2303234

> lines(previsto$pre,col=4)
> bovespa161907=c(46.869,47.407,47.656,47.400)
> lines(4712:4715,bovespa161907,col=2)
> mapexterno=mean(abs(bovespa161907-previsto$pred)/bovespa161907)*100
> mapexterno
[1] 1.386548

```

A saída gráfica é dada por:

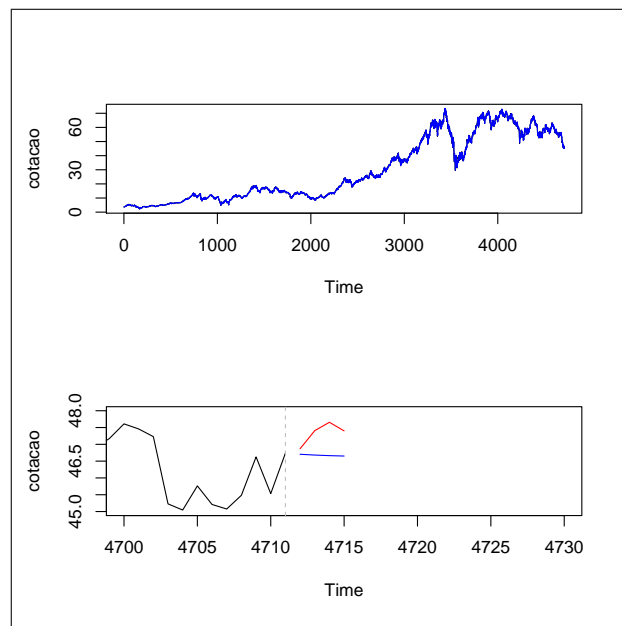


Figura a. Cotação da Bovespa e previsões

Para os retornos usamos:

```
> #Arima para os retornos da Bovespa
```

```
> rt=diff(cotacao)/cotacao[-length(cotacao)]
```

```
> rtarima=auto.arima(rt)
```

```
> summary(rtarima)
```

Series: rt

ARIMA(3,0,3) with non-zero mean

Coefficients:

	ar1	ar2	ar3	ma1	ma2	ma3	intercept
	0.4257	0.6206	-0.7684	-0.4073	-0.6461	0.7077	8e-04
s.e.	0.0501	0.0642	0.0682	0.0556	0.0703	0.0711	3e-04

sigma² estimated as 0.0005216: log likelihood=11117.27

AIC=-22218.53 AICc=-22218.5 BIC=-22166.87

Training set error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
----	------	-----	-----	------	------	------

```

Training set 2.430886e-05 0.02283862 0.0157432 NaN Inf 0.7074504 -0.005039146
> par(mfrow=c(2,1))
> plot.ts(rt)
> lines(fitted(rtarima),col=4)
> plot.ts(rt,xlim=c(4700,4730), ylim=c(-0.1,0.1))
> abline(v=4710,lty=2,col=8)
> previstor=predict(arima(rt,order=c(3,0,3)),4)
> previstor
$pred
Time Series:
Start = 4711
End = 4714
Frequency = 1
[1] -6.696700e-04 -9.413669e-04 -1.080333e-03 3.628575e-05

$se
Time Series:
Start = 4711
End = 4714
Frequency = 1
[1] 0.02283862 0.02284251 0.02284604 0.02288278

> lines(previstor$pre,col=4)
> retorno161907=c(0.026464323,0.002802858,0.011478803,0.005252389)
> lines(4711:4714,retorno161907,col=2)
> error=mean(abs(retorno161907-previstor$pred))*100
> error
[1] 1.216336

```

A saída gráfica é dada por:

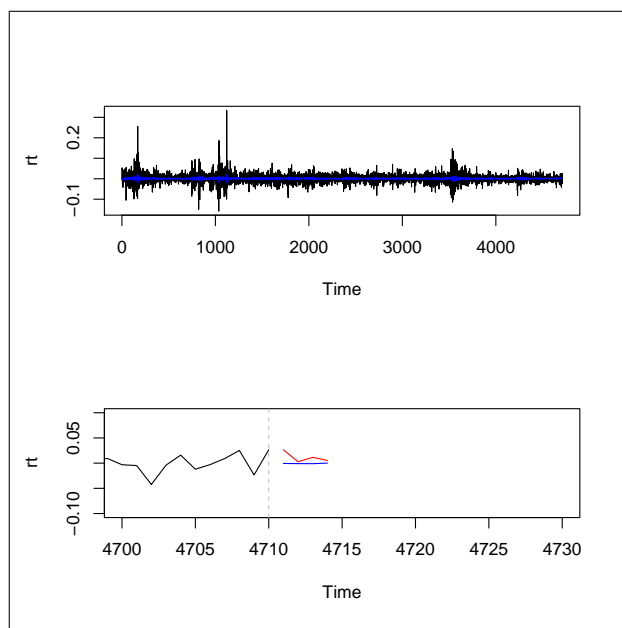


Figura b. Retornos da Bovespa e previsões.

8.5 Simulações de modelos Arima

O R contém algumas funções para simulação de modelos autorregressivos e de médias móveis. Exemplo: simulação de séries temporais ($T = 200$) de modelo estacionário com processo autorregressivo gaussiano AR(1), com parâmetros 0.5 e 0.95.

```
> y = arima.sim(n=200, list(ar=0.5), innov=rnorm(200))
> y = arima.sim(n=200, list(ar=0.95), innov=rnorm(200))
> plot.ts(y) #gráfico correspondente
> acf(y)
```

Identifique matematicamente as simulações de modelos AR, MA e ARIMA abaixo descritos.

```
> AR2=arima.sim(n = 100, list( ar = c(0.8897, -0.4858)))
> AR1=arima.sim(n = 100, list( ar = 0.8897))
> MA1=arima.sim(n = 100, list( ma = 1))
> MA2=arima.sim(n = 100, list( ma = c(-2, 1.5)))
> ARMA11=arima.sim(n = 100, list(ar=0.8897, ma=-0.2279))
> ARMA22=arima.sim(n = 100, list(ar=c(0.8897, -0.4858),
```

```
ma=c(-0.2279, 0.2488)))  
> ARIMA110=arima.sim(list(order=c(1,1,0),ar = 0.7),n = 100)  
> ARIMA111=arima.sim(list(order = c(1,1,1), ar = 0.7,  
ma = 2), n = 100)  
> par(mfrow=c(2,2), pty="s")  
> plot(AR1, main="AR(0.8897)")  
> plot(AR2, main="AR(0.8897, -0.4858)")  
> plot(MA1, main="MA(1.000)")  
> plot(MA2, main="MA(-2, 1.5)")
```

9 Modelos Sarima

Os modelos SARIMA são a extensão dos modelos ARIMA (*sazonal multiplicativo*) e contêm uma parte não sazonal da série, com parâmetros (p, d, q) , e uma sazonal, com parâmetros (P, D, Q) . O modelo mais geral é dado pela equação:

$$\Phi(B)\Phi(B)(1-B)^D(1-B)^dX_t = \theta(B)\Theta(B)\epsilon_t, \quad \{\epsilon_t\} \sim RB(0, \sigma_\epsilon^d).$$

Para exemplificar considere uma série x_t , observada mês a mês por um modelo $ARIMA(1, 0, 0)$:

$$x_t = \phi x_{t-1} + \epsilon_t,$$

para modelar por estações do ano, através de um processo $AR(1) = ARIMA(1, 0, 0)$ sazonal:

$$x_t = \Phi_4 x_{t-4} + \epsilon_t,$$

para modelar ano a ano através do processo $MA(1) = ARIMA(0, 0, 1)$ sazonal:

$$x_t = \epsilon_t - \Theta \epsilon_{t-12}.$$

9.1 Exemplo 2

Considere a cotação da Bovespa do exemplo 1, para identificar o modelo sarima, usamos

```
bov=read.table(file.choose(),header=T) #bovespa15072013.txt  
attach(bov)  
dim(bov)
```

```

library(forecast)
cotacao.ts=ts(cotacao,end=c(2013,123),frequency=245)
ndiffs(cotacao.ts)
acotacao=auto.arima(cotacao.ts)
summary(acotacao)
Series: cotacao.ts
ARIMA(2,1,2)

Coefficients:
          ar1      ar2      ma1      ma2
      -0.2896  0.7015  0.2468 -0.7336
s.e.    0.0805  0.0800  0.0758  0.0747

sigma^2 estimated as 0.4172:  log likelihood=-4624.42
AIC=9258.83   AICc=9258.85   BIC=9291.12

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
Training set 0.01047288 0.6458221 0.4019317 0.03303618 1.582666 0.056673
              ACF1
Training set 0.009322901
mapecotacao=mean(abs((cotacao.ts-fitted(acotacao))/cotacao.ts))*100
mapecotacao
[1] 1.582666

rt.ts=diff(cotacao.ts)/cotacao.ts[-length(cotacao.ts)]
aretorno=auto.arima(rt.ts)
summary(aretorno)

Series: rt.ts
ARIMA(3,0,3) with non-zero mean

Coefficients:
          ar1      ar2      ar3      ma1      ma2      ma3  intercept
      0.4257  0.6206 -0.7684 -0.4073 -0.6461  0.7077         8e-04
s.e.    0.0501  0.0642  0.0682  0.0556  0.0703  0.0711         3e-04

```

```
sigma^2 estimated as 0.0005216: log likelihood=11117.27
AIC=-22218.53 AICc=-22218.5 BIC=-22166.87
```

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	2.430886e-05	0.02283862	0.0157432	NaN	Inf	0.6847435	-0.005039146

```
erroretorno=mean(abs(rt.ts-fitted(aretorno)))
erroretorno
0.0157432
```

Por ser uma série com 245 efeitos sazonais, a saída do modelo retorno um ARIMA, pois alocar todos estes efeitos tem um custo computacional enquanto a memória disponível. assim é conveniente realizar modelos com efeitos sazonais mensais (12), como será abordado a seguir:

9.2 Exemplo 3

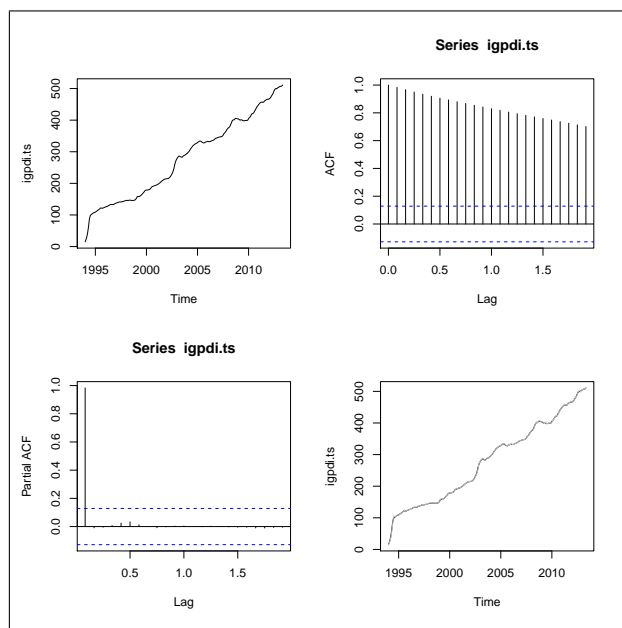


Figura 37. Características da série do IGPDI

Considere a série acima do índice geral de preços distribuição interna do Brasil por mês desde janeiro de 1944 até maio de 2013, disponíveis no site do IPEA.


```

library(forecast)
igpdiicms=read.table(file.choose(),header=T)
# dadosigpdicmssebase1994fim201305.txt
attach(igpdiicms)
names(igpdiicms)
igpdi.ts=ts(igpdi,start=c(1994,1), frequency=12)
par(mfrow=c(2,2))
plot.ts(igpdi.ts)
acf(igpdi.ts)
pacf(igpdi.ts)
plot.ts(igpdi.ts)
lines(fitted(auto.arima(igpdi.ts)),col=8)

```

Para realizar previsões com o respectivo gráfico para os meses de junho a outubro de 2013, fazemos

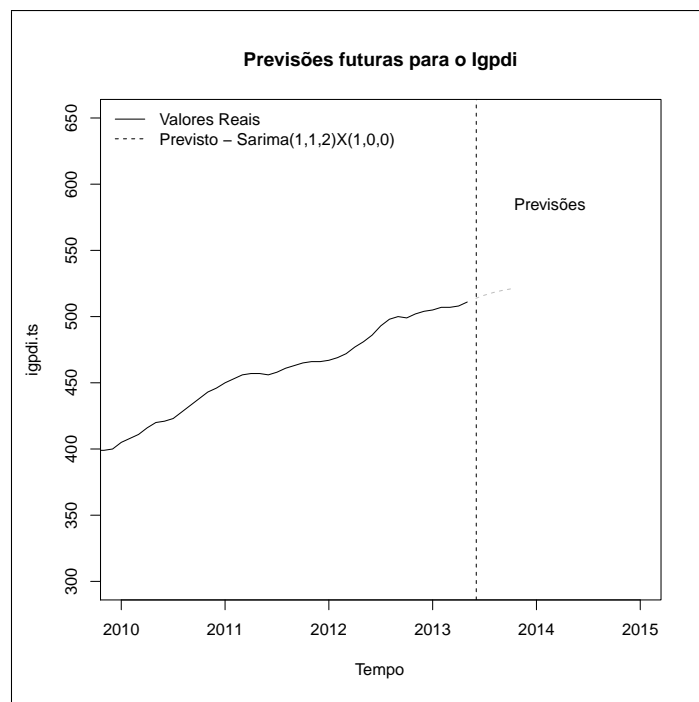


figura 38. Série do IGPDI e previsões.

```

> ma=auto.arima(igpdi.ts)
> pa=predict(arima(igpdi.ts,order=c(1,1,2),seasonal=list(order=c(1,0,0))),5)

```

```

> pa
$pred
      Jun      Jul      Aug      Sep      Oct
2013 513.8452 516.1959 518.1510 519.7032 520.8895

$se
      Jun      Jul      Aug      Sep      Oct
2013 1.617845 3.884224 6.259942 8.729258 11.242843

par(mfrow=c(1,1))
plot.ts(igpdi.ts,xlab="Tempo", main=
"Previsões futuras para o Igpdi",xlim=c(2010,2015),ylim=c(300,650))
abline(v=2013.42, lty=2)
legend("topleft", c("Valores Reais",
"Previsto - Sarima(1,1,2)X(1,0,0)"),lty = 1:2, bty="n")
legend(2013.5,600, legend="Previsões", bty="n")
lines(pa$pred, type="l", col=8, lty=2)

plot(rwf(igpdi.ts),xlim=c(2012,2013.7)) #Gráfico com IC para as Previsões

```

9.3 Deflacionando o ICMS - SE

O ICMS (imposto sobre operações relativas à circulação de mercadorias e sobre prestações de serviços de transporte interestadual, intermunicipal e de comunicação) é de competência dos Estados, em particular trataremos o estado de Sergipe. Esta informação está disponível em IPEADATA, em miles de reais - Ministério da Fazenda, Conselho Nacional de Política Fazendária, Comissão Técnica Permanente do ICMS, Boletim de Arrecadação (Min. Fazenda/Cotepe). Este indicador econômico tem como período base julho de 1994. Para mudança do período Base (exemplo: maio de 2013), realizamos os seguintes comandos:

```

igpdir=igpdi/igpdi[length(igpdi)]
icmseser=icmsse/igpdir
plot.ts(icmsse)
lines(icmseser,col=2)

```

9.4 Tarefa

1. Construa um modelo do tipo linear do IGPDI em função dos meses e encontre o MAPE (verifique que o MAPE é: 9.51), quais suas conclusões?.
2. Identifique o fator integrador desta série (quando ela é estacionária?). Use a função `ndiffs()`.
3. Construa modelos Holt Winters, ARIMA e SARIMA para o ICMS-SE tanto na série real como a série deflacionada (desprovida do efeito inflacionário) e interprete os resultados.
4. Construa o MAPE, o Erro Médio (MAE no R) e o Erro Médio Percentual (MAE*100) para os modelos em 3, e interprete.
5. Realize previsões para junho - agosto de 2013 para os modelos propostos em 3 e interprete.
6. Gere uma saída gráfica (em conjunto) para os modelos propostos em 3.

9.5 Arima vs Sarima

Consideremos o salário mínimo real mensal do Brasil desde janeiro de 1980, até abril de 2011, do banco de dados `ie`

```
> attach(ie)
> names(ie)
[1] "anomes" "tcc" "tcv" "bc" "exp" "smr" "igpd"
> smr1=auto.arima(smr)
> smr.ts=ts(smr,start=c(1980,1),frequency=12)
> smr2=auto.arima(smr.ts)

> summary(smr1) # Modelo ARIMA
Series: smr
ARIMA(5,1,5)

Coefficients:
      ar1      ar2      ar3      ar4      ar5     ma1     ma2     ma3     ma4     ma5
    -0.475 -0.551 -0.942 -0.487 -0.509  0.044 -0.09  0.603  0.228 -0.165
s.e.  0.086  0.081  0.029  0.083  0.072  0.096  0.08  0.070  0.098  0.074
```

```
sigma^2 estimated as 950.2:  log likelihood = -1819.82
AIC = 3661.64    AICc = 3662.36    BIC = 3704.83
```

```
In-sample error measures:
```

ME	RMSE	MAE	MPE	MAPE
0.3716551	30.7837859	20.7415191	-0.5311600	5.7073384

```
> summary(smr2) # Modelo SARIMA
```

```
Series: smr.ts
```

```
ARIMA(4,1,2)(2,0,1)[12]
```

```
Coefficients:
```

ar1	ar2	ar3	ar4	ma1	ma2	sar1	sar2	sma1
-0.0434	0.2968	-0.0186	0.3641	-0.3192	-0.6237	1.1152	-0.2260	-0.5691
s.e. 0.0896	0.0641	0.0516	0.0518	0.0881	0.0805	0.1775	0.1273	0.1638

```
sigma^2 estimated as 843:  log likelihood = -1799.3
AIC = 3618.6    AICc = 3619.2    BIC = 3657.87
```

As previsões correspondentes para os meses de maio a outubro de 2011 são:

```
> predict(smr1,6)
```

```
Time Series:
```

```
Start = 377
```

```
End = 382
```

```
Frequency = 1
```

```
pred: 541.4022 536.4573 550.9956 550.1748 550.2760 541.2245
se   : 30.82479 35.46528 35.69889 36.44613 40.53671 41.60850
```

```
> predict(smr2,6)
```

```
2011
```

	May	Jun	Jul	Aug	Sep	Oct
pred:	545.9977	545.2510	546.5741	544.9905	544.5428	540.7435
se :	29.03376	34.43006	35.70931	36.24331	39.00929	40.13379

10 Processos ARFIMA

O modelo $ARIMA(p, d, q)$ é algumas vezes denominado como processo geral com diferenciação fracionária (General Fractional Differenced process)

(ARFIMA) quando o parâmetro d (grau de diferenciação) assume valores não inteiros. Este modelo é utilizado para dados de séries hidrológicas. A mais importante característica do modelo ARFIMA(p, d, q) é a propriedade de longa dependência "long-memory" quando $d \in (0.0, 0.5)$, e pequena dependência "short-memory" quando $d \in (-0.5, 0.0)$.

Longa dependência (ou persistência) é caracterizada pela presença, na série, de uma significativa dependência entre as observações mesmo para distantes "lags". Esta característica tem sido observada em séries de diferentes áreas de estudos tais como, meteorologia, astronomia, hidrologia e economia (Cribari, 2002). No modelo ARIMA(p, d, q), $d \in (-0.5, 0.5)$, as características de longa e curta dependência podem ser notadas pelo comportamento da função espectral e da função de autocorrelação.

No domínio do tempo, as autocorrelações decaem lentamente de uma forma hiperbólica, isto é, $\rho_k \sim k^{-d}$, o oposto das autocorrelações produzidas pelo modelo ARMA(p, q) (Box and Jenkins (1976)), as quais têm decaimento exponencial $\rho_k \sim a^{-d}$, $0 < a < 1$. No domínio da frequência a função espectral tende a infinito quando a frequência se aproxima de zero, i.e., $f(\lambda) \rightarrow \infty$ quando $\lambda \rightarrow 0$.

Se $d \in (-0.5, 0.0)$, o processo tem a propriedade de curta dependência. No domínio da frequência isto é indicado pelo comportamento da função espectral que se aproxima de zero quando a frequência também se aproxima de zero. No domínio do tempo a função de autocorrelação poderá exibir dependências negativas entre observações distantes, portanto, o tipo de dependência é essencialmente determinado pelo valor fracionário de d . Se $d = 0$, o processo é chamado de memória curta (Geweke, 1983). Formalmente o processo ARFIMA(p, d, q) é definido como a seguir:

Seja $d \in \mathbb{R}$. $\{X_t\}$ segue um processo ARFIMA(p, d, q) se satisfaz a equação em diferenças da forma

$$\Phi(B)(1 - B)^d X_t = \Theta(B)\epsilon_t, \quad (41)$$

com $\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p$ e $\Theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q$, $\{\epsilon_t\}$ sendo um processo ruído branco com média 0 e variância σ_ϵ^2 . O filtro de diferenciação fraccionaria $(1 - B)^d$ é definido pela expansão binomial

$$(1 - B)^d = \sum_{j=0}^{\infty} \pi_j B^j,$$

onde $\pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}$, $j = 0, 1, \dots$, e $\Gamma(\cdot)$ é a função gamma definida em $\mathbb{R} - \mathbb{Z}^-$:

$$\Gamma(x) = \begin{cases} \int_0^\infty t^{x-1} \exp^{-t} dt, & x > 0; \\ \infty, & x = 0; \\ x^{-1}\Gamma(1+x), & x < 0. \end{cases}$$

Quando $d \in (-0.5, 0.5)$ e as raízes dos polinômios $\Phi(z) = 0$ e $\Theta(z) = 0$ são não-comuns e estão fora do círculo unitário, o processo definido em (41) é estacionário e invertível e com função de densidade espectral dada por

$$f_{ARFIMA}(\lambda) = f_{ARMA}(\lambda) \left\{ 2 \sin \left(\frac{\lambda}{2} \right) \right\}^{-2d}, \quad \lambda \in [-\pi, \pi], \quad (42)$$

onde $f_{ARMA}(\lambda)$ está definida em (40).

Hosking, 1981, mostrou que valores $d \geq 0.5$ $\{X_t\}$ é não estacionário e invertível, e ainda que séries com representação $ARFIMA(p, d, q)$ com $d \in (0, 0.5)$ apresentam estacionariedade e memória longa. Assim, no que se segue nós consideraremos o processo $ARFIMA(p, d, q)$ $d \in (0, 0.5)$.

10.1 Estimação do parâmetro d

Existem vários estimadores do parâmetro de diferenciação fracionária d , que podem ser classificados em semi-paramétricos e paramétricos. Os primeiros envolvem a estimação simultânea dos parâmetros do modelo, em geral utilizando o método de máxima verossimilhança. Nos procedimentos semi-paramétricos, a estimação dos parâmetros do modelo é realizada em dois passos: primeiro estima-se o parâmetro de memória longa d , por exemplo, através de um modelo de regressão do logaritmo da função periodograma e, posteriormente, estimam-se os parâmetros auto-regressivos e de médias móveis. O estimador mais popular dentro dessa classe é o estimador proposto por Geweke e Porter-Hudak (GPH); variantes foram desenvolvidas por Reisen, 2007 e Ehlers, 2007.

10.1.1 Usando a função periodograma (GPH)

Seja $f(\lambda_j)$ a função definida em (42), para $\lambda_j = \frac{2\pi j}{n}$, $j = 0, 1, \dots, \lfloor \frac{n}{2} \rfloor$, onde n é o tamanho amostral. O logaritmo de $f(\lambda_j)$ pode ser escrito como:

$$\ln f(\lambda_j) = \ln f_u(0) - d \ln \left\{ 2 \sin \left(\frac{\lambda_j}{2} \right) \right\}^2 + \ln \frac{f_u(\lambda_j)}{f_u(0)}, \quad (43)$$

onde $f_u(\lambda)$ é a densidade espectral de $U_t = (1 - B)^d \{X_t\}$. Aqui, $\lfloor \cdot \rfloor$ denota a parte inteira.

Geweke e Porter-Hudak sugerem um estimador semi-paramétrico de d , adicionando $\ln I(\lambda)$ em ambos os lados da equação (43) e considerando as frequências próximas de zero, obtendo a aproximação:

$$\ln I(\lambda_j) \approx \ln f_u(0) - d \ln \left\{ 2 \sin \left(\frac{\lambda_j}{2} \right) \right\}^2 + \ln \frac{I(\lambda_j)}{f(\lambda_j)}, \quad (44)$$

que sugere a equação de regressão dada por

$$\ln I(\lambda_j) \approx \beta_0 + \beta_1 \ln \left\{ 2 \sin \left(\frac{\lambda_j}{2} \right) \right\}^2 + \epsilon_j, \quad j = 1, 2, \dots, g(n),$$

onde $\beta_0 = \ln f_u(0)$, $\beta_1 = -d$ e $g(n)$ é a amplitude de banda (bandwidth), que corresponde ao número de frequências utilizadas na regressão. Os erros $\{\epsilon_j\}$ são assintoticamente independentes com distribuição Gumbel de média 0 e variância $\frac{\pi}{6}$. O estimador GPH é dado por

$$d_{GPH} = - \frac{\sum_{i=1}^{g(n)} (x_i - \bar{x}) \ln I(\lambda_j)}{\sum_{i=1}^{g(n)} (x_i - \bar{x})^2} \quad (45)$$

onde $x_i = \ln \{2 \sin(\frac{\lambda_j}{2})\}^2$. Geweke e Porter-Hudak sugerem considerar $g(n) = n^\alpha$, $0 < \alpha < 1$. Algumas propriedades assintóticas do estimador dado em (45) foram derivadas por Hurvich, Deo e Brodsky e Velasco.

10.1.2 Usando a função periodograma suavizado (GPH_S)

Tomando à expressão (42) e usando resultados assintóticos (Reisen, 2007).

Para $\ln \{I_s(\lambda)/f(\lambda)\}$ pode-se escrever a equação de regressão da forma

$$\ln f_s(\lambda_j) = \ln f_u(0) - d \ln \left\{ 2 \sin \left(\frac{\lambda_j}{2} \right) \right\}^2 + \ln \frac{f_s(\lambda_j)}{f(\lambda_j)} + \ln \frac{f_u(\lambda_j)}{f_u(0)}, \quad (46)$$

Restringindo o domínio de j , $1 < j < g(n)$, e escolhendo a função $g(n)$ como anteriormente, pode-se reescrever (46):

$$\ln I_s(\lambda_j) \approx \ln f_u(0) - d \ln \left\{ 2 \sin \left(\frac{\lambda_j}{2} \right) \right\}^2 + \ln \frac{f_s(\lambda_j)}{f(\lambda_j)}, \quad (47)$$

A equação (47) é uma forma aproximada da equação de regressão linear simples, isto é,

$$y_j = a + bx_j + \epsilon_j \quad j = 1, 2, \dots, g(n), \quad (48)$$

onde $y_j = \ln f_s(\lambda_j)$, $b = d$, $x_j = \ln(2 \sin(\lambda_j/2))^2$, $\epsilon_j = \ln(f_s(\lambda_j)/f(\lambda_j))$ e $a = \ln f_u(0)$.

Como notado anteriormente, quando $-0.5 < d < 0.0$ no modelo ARIMA(p, d, q), Os erros $\{\epsilon_j\}$ são assintoticamente independentes com distribuição Gumbel de média 0 e variância $\frac{\pi}{6}$ (Reisen, 2007).

O estimador de d obtido pelo método de regressão utilizando a função periodograma suavizado com a janela Parzen é dado por:

$$d_{SP} = - \frac{\sum_{i=1}^{g(n)} (x_i - \bar{x}) \ln I_s(\lambda_j)}{\sum_{i=1}^{g(n)} (x_i - \bar{x})^2} \quad (49)$$

onde $x_i = \ln\{2 \sin(\frac{\lambda_j}{2})\}^2$.

10.2 Estimação de máxima verossimilhança.

A função de verossimilhança de $\mathbf{X} = (X_1, \dots, X_n)$ proveniente de um processo ARFIMA(p, d, q) pode ser expressa na forma

$$L(\eta, \sigma_a^2) = (2\pi\sigma_a^2)^{-1/2} \exp \left[-\frac{1}{\sigma_a^2} \sum_{j=1}^n (X_j - \hat{X}_j)^2 / r_{j-1} \right] \quad (50)$$

em que $\eta = (d, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$, \hat{X}_j , $j = 1, \dots, n$, são as previsões um passo à frente e $r_{j-1} = (\sigma_a^2)^{-1} E(X_j - \hat{X}_j)^2$. Os estimadores de máxima verossimilhança dos parâmetros são dados por

$$\hat{\sigma}_{MV}^2 = n^{-1} S(\hat{\eta}_{MV}), \quad (51)$$

onde

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

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

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

entretanto, o cálculo de $\ell(\eta)$ é bastante lento. Um procedimento alternativo é considerar uma aproximação para $\ell(\eta)$ dada por

$$\ell(\eta) \simeq \ell_*(\eta) = \ln \frac{1}{n} \sum_j \frac{I_n(w_j)}{2\pi f(w_j; \eta)}, \quad (52)$$

em que

$$I_n(w_j) = \frac{1}{n} \left| \sum_{t=1}^n X_t e^{-itw_j} \right|^2$$

é o periodograma dos dados,

$$f(w_j; \eta) = \frac{\sigma_a^2 |1 - \theta_1 e^{-iw_j} - \dots - \theta_q e^{qiw_j}|^2}{2\pi |1 - \phi_1 e^{iw_j} - \dots - \phi_p e^{piw_j}|^2} \cdot |1 - e^{iw_j}|^{-2}$$

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

Hannan e Fox e Taququ mostram que:

1. O estimador $\hat{\eta}_{MV}$ que minimiza (52) é consistente;
2. se $d > 0$,

$$\hat{\eta} \mathcal{N}(\eta, n^{-1} A^{-1}(\eta)),$$

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

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

3. a variância é estimada por

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

Para o exemplo aumentamos o tamanho da série da cotação da Bovespa desde 4 de julho de 1994 até 15 de julho de 2013, totalizando 4711 observações de dias úteis. Para a estimação dos modelos ARFIMA usando o método de Geweke e Porter-Hudak (GPH), usamos os dados da cotação em primeira diferença. Uma segunda série com os retornos dos índices da Bovespa também é analisada.

```
> library(fracdiff)
> fdGPH(diff(cotacao)) #GEWEKE PORTER HUDAK
$d
[1] 0.06387949

$sd.as
[1] 0.08627253

$sd.reg
[1] 0.08854559

> fdSperio(diff(cotacao)) #REISEN E VALDERIO
$d
[1] 0.0246452

$sd.as
[1] 0.03235789

$sd.reg
[1] 0.03616914
```

Concluimos que usando o método de Geweke e Porter-Hudak (GPH), identificamos que as primeiras diferenças da cotação dos índices da Bovespa segue um processo de memória longa. Para os retornos da Bovespa identificamos um processo de memória curta. Usando o método de Reisen e Valderio:

```
> fdSperio(rt)
$d
[1] -0.07206497

$sd.as
[1] 0.03235789
```

```
$sd.reg  
[1] 0.0330011
```

```
> fdGPH(rt)  
$d  
[1] -0.07779845
```

```
$sd.as  
[1] 0.08627253
```

```
$sd.reg  
[1] 0.07612321
```

Com este método de estimação confirmamos que as diferenças da cotação dos índices da Bovespa apresentam memória longa, e os retronos memória curta. Para estimar os parâmetros do ARFIMA(p,d,q) automaticamente, podemos usar o algoritmo de Hyndman-Khandakar(2008) para selecionar p , e q , além de uma estimação de máxima verossimilhança baseado em Haslett e Raftery(1989) para a inclusão do parâmetro d . A função usada é `arfima()` do pacote `forecast`.

```
> library(forecast)  
This is forecast 4.06  
  
> arfima2=arfima(cotacao, drange = c(0, 0.5))  
> summary(arfima2)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
d	0.4999135	0.0001061	4709.87	<2e-16 ***
ma.ma1	-0.6735015	0.0140289	-48.01	<2e-16 ***
ma.ma2	-0.6011644	0.0164523	-36.54	<2e-16 ***
ma.ma3	-0.4682584	0.0140321	-33.37	<2e-16 ***
ma.ma4	-0.3670649	0.0163978	-22.39	<2e-16 ***
ma.ma5	-0.2362165	0.0174452	-13.54	<2e-16 ***

```

sigma[eps] = 0.7797594
[d.tol = 0.0001221, M = 100, h = 5.814e-05]
Log likelihood: -5517 ==> AIC = 11048.2 [7 deg.freedom]

> predict(arfima2,5)
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
4712      47.18915 46.19422 48.18408 45.66754 48.71076
4713      47.69328 46.15938 49.22718 45.34738 50.03918
4714      47.96619 45.95155 49.98083 44.88506 51.04732
4715      48.17970 45.76732 50.59207 44.49028 51.86911
4716      48.57182 45.82972 51.31392 44.37814 52.76550
> plot.ts(cotacao)
> lines(fitted(arfima2),col=2)
mape2=mean(abs((cotacao-fitted(arfima2))/cotacao))*100
> mape2
[1] 2.532338

> arfima3=arfima(rt, drange = c(-0.5, 0))
> summary(arfima3)

Coefficients:
      Estimate
d          -0.394
ar.ar1      0.522
ar.ar2      0.518
ar.ar3     -0.073
ma.ma1      0.114
ma.ma2      0.480
ma.ma3      0.109
ma.ma4      0.059
ma.ma5      0.019
sigma[eps] = 0.02284145
[d.tol = 0.0001221, M = 100, h = 0.0001171]
Log likelihood: 1.112e+04 ==> AIC = -22213 [10 deg.freedom]

> predict(arfima3,5)
> predict(arfima3,5)

```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2013.502	0.0003268554	-0.02900929	0.02966300	-0.04453891	0.04519262
2013.506	0.0001542161	-0.02918473	0.02949316	-0.04471583	0.04502426
2013.510	-0.0003340129	-0.02968493	0.02901690	-0.04522236	0.04455433
2013.514	0.0001457671	-0.02924235	0.02953388	-0.04479948	0.04509101
2013.518	0.0003005578	-0.02910565	0.02970677	-0.04467236	0.04527347

Para construir modelos alternativos ARFIMA(p,d,q), podemos usar a função `fracdiff()`, do pacote `fracdiff`.

```
> fcot3=fracdiff(cotacao,nar=0,nma=9)
> summary(fcot3)
```

Call:

```
fracdiff(x = cotacao, nar = 0, nma = 9)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
d	0.4998193	0.0002495	2003.683	< 2e-16 ***
ma1	-0.6728187	0.0144660	-46.510	< 2e-16 ***
ma2	-0.6250666	0.0197965	-31.575	< 2e-16 ***
ma3	-0.5541535	0.0191476	-28.941	< 2e-16 ***
ma4	-0.5171683	0.0145801	-35.471	< 2e-16 ***
ma5	-0.4736796	0.0170956	-27.708	< 2e-16 ***
ma6	-0.4010545	0.0199475	-20.105	< 2e-16 ***
ma7	-0.2989917	0.0172495	-17.333	< 2e-16 ***
ma8	-0.2369755	0.0189120	-12.530	< 2e-16 ***
ma9	-0.1089231	0.0199350	-5.464	4.66e-08 ***

```
sigma[eps] = 0.7303714
```

```
[d.tol = 0.0001221, M = 100, h = 5.489e-05]
```

```
Log likelihood: -5208 ==> AIC = 10438.53 [11 deg.freedom]
```

```
> predict(fcot3,5)
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
4712	47.24627	46.28938	48.20315	45.78284	48.70970
4713	47.91661	46.44193	49.39130	45.66128	50.17195
4714	48.23249	46.28072	50.18426	45.24751	51.21747
4715	48.44416	46.05991	50.82840	44.79777	52.09054
4716	48.69737	45.90856	51.48618	44.43226	52.96248

```
> mape3=mean(abs((cotacao-fitted(fcot3))/cotacao))*100
> mape3
[1] 2.224464
```

Para a cotação dos índices da Bovespa o modelo alternativo (fcot3), apresentou melhor MAPE que o modelo arfima2. Uma sugestão para comparar modelos dos retornos da Bovespa é usar o erro relativo percentual.

Para simular um modelo arfima, podemos usar a função `arfima.sim()`, do pacote `forecast`, exemplo:

```
x <- fracdiff.sim( 1000, ma = -.98, d = .48)$series
```

10.3 Tarefa

1. Verifique se o IGPDI, ICMS-SE são processos de memória longa.
2. Pesquise funções que apresentem um processo de memória longa ou memória curta.

11 Modelos não lineares estruturais

Estes modelos são apropriadas para séries financeiras onde a variância condicional evolui no tempo, e modelos lineares do tipo ARIMA, não são adequados para descrever o comportamento de estas séries. Modelos com alta volatilidade mais conhecidos são o ARCH (Heterocedasticidade condicional autorregressiva) e o GARCH (generalização dos ARCH) são alguns dentre uma variedade de modelos.

iniciaremos apresentando um teste de ausência de autocorrelação residual ou independência dos resíduos.

11.1 Teste de Ljung e Box

Após estimar o modelo, precisamos identificar o comportamento dos resíduos estimados do modelo. Um teste para as autocorrelações dos resíduos estimados, que a pesar de não detectar quebras específicas no comportamento do ruído branco, pode indicar se esses valores são muito altos é dado pela estatística:

$$Q(k) = n(n-2) \sum_{j=1}^k \frac{\hat{r}_j^2}{(n-j)} \quad (53)$$

A qual terá uma distribuição χ^2 com $k - p - q$ graus de liberdade. A distribuição assintótica é obtida sobre a hipóteses de que $k = k(n) \rightarrow \infty$ quando $n \rightarrow \infty$. Uma variação é o teste de Box-Pierce (Moretin e Toloi, 2006).

11.1.1 Manifestações da volatilidade numa série financeira

1. A volatilidade aparece em grupos, de maior ou menor variabilidade.
2. A volatilidade evolui continuamente no tempo podendo ser considerada estacionária.
3. Ela reage positivamente ou negativamente, podendo ser considerada estacionária.

11.2 Retornos

Para avaliar os riscos de uma carteira de ativos financeiros, o risco é medido em termos de variações de preços. Considere P_t o preço de um ativo no instante t . A variação de preços entre os instantes $t - 1$ e t é dada por: $\Delta P_t = P_t - P_{t-1}$, e o *retorno líquido simples* deste ativo, no mesmo instante é definido por:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}} \quad (54)$$

R_t , também chamado de *taxa de retorno*.

O *retorno composto* ou simplesmente *retorno* é dado por: $\log(1+R_t)$. na pratica usa-se os retornos que tem propriedades desejáveis como estacionaridade e ergodicidade próprios da modelagem ARMA, ARIMA ou GARCH. exemplo considere a cotação da Bovespa, os retornos e algumas características gráficas destas duas séries, são apresentadas a seguir:

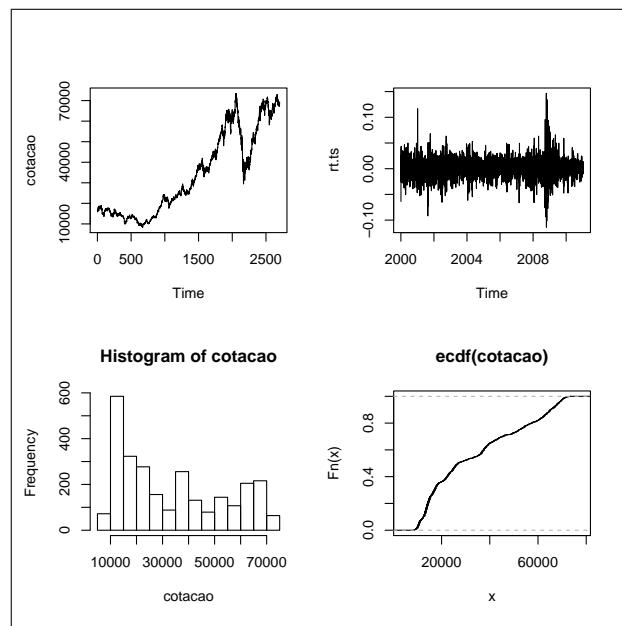


Figura 33. Características da cotação da Bovespa.

```

bov=read.table(file.choose(),header=T) # bovespa15072013.txt

attach(bov)

rt=diff(cotacao)/cotacao[-length(cotacao)]

cotacao.ts=ts(cotacao,end=c(2013,123),frequency=245) #15/07/2013

rt.ts=ts(rt,end=c(2013,123),frequency=245)

par(mfrow=c(2,2))

plot.ts(cotacao.ts)

plot.ts(rt.ts)

hist(cotacao.ts)

plot(ecdf(cotacao.ts))

```


11.3 ARCH

Utilizado para modelar séries financeiras que apresentam a variância condicional evoluindo no tempo. Um destes modelos não lineares é o ARCH (autorregresive conditional heterocedasticity), estes modelos são não lineares no que se refere a variância. O objetivo é modelar a variância condicional conhecida como volatilidade, comumente usa-se nos retornos de uma carteira de ativos. Considere uma série de retornos X_t é não correlacionada serialmente, não entanto sua volatilidade depende dos retornos passados considerando uma função quadrática. A definição é dado por:

$$X_t = \sqrt{h_t}\epsilon_t, h_t = \alpha_0 + \sum_{i=1}^r \alpha_i X_{t-i}^2 \quad (55)$$

Onde ϵ_t é uma sequência de variáveis aleatórias independentes e identicamente distribuídas (iid) com média zero e variância um. $\alpha_0 > 0, \alpha_i \geq 0, i > 0$. Na prática ϵ_t , tem distribuição normal padrão ou distribuição t. Para implementar no R, temos que alertar que estes modelos pertencem a uma classe denominada de 4, que fornece as saídas com @ ao invés de \$ próprios da classe 3 desenvolvidas até aqui.

Um primeiro passo para a construção destes modelos é ajustar modelos ARMA, para remover a correlação serial na série original, e posteriormente assumir que nosso resíduo é um ARCH(r), ou $\epsilon \approx ARCH(r)$. Para implementar um exemplo com os retornos da Bovespa, consideraremos a função `garchFit()`, do pacote `fGarch`. Consideremos o preço de um ativo financeiro, como o índice de cotação da Bovespa. A ideia é modelar a série dos retornos por um ARMA(1,0), e os resíduos estimados são gerados por um ARMA(2,0). O retorno desta série financeira usando o R com a função `garchFit()` do pacote `fGarch` é dada por: A ideia é sugerir que a série dos retornos segue um ARMA(0,1), e os resíduos um ARCH(2)

```
library(fGarch)
library(forecast)
library(tseries)
arch=garchFit(~arma(0,1)+garch(2,0),rt)
summary(arch)
resumo=arch@fit
resumo
plot(arch)
ajustearchz=arch@fit$series$z #estimação
ajustearchx=arch@fit$series$x #valor real
```

```
ajustearch=arch@fit$series$h #nível
predict(ajustearchz,5) #previsões de 5 dias á frente
jarque.bera.test(residuals(arch))
```

O comando anterior fornece o teste de não normalidade dos resíduos no modelo ARCH, o qual é dado a seguir:

Jarque Bera Test

```
data: residuals(arch)
X-squared = 50433.67, df = 2, p-value < 2.2e-16
```

Uma saída gráfica para a densidade dos resíduos é proposta por:

```
plot(density(residuals(arch)))
curve(dnorm(x,mean=mean(residuals(arch)),sd=sd(residuals(arch))),add=T,col=2)
```

Uma saída gráfica para comparar a série real dos retornos com a suas estimativas é dado por:

```
plot.ts(rt)
lines(arch@fit$series$z,col=2)
```

Interpretação: nesta saída resumo, estão incluídas todas as estatísticas correspondentes ao modelo, incluindo convergência, iterações, a log-verosimilhança os coeficientes do modelo, a matriz hessiana das estimativas dos parâmetros, critérios de informação, a série estimada dos retornos, além da série estimada dos resíduos.

Tarefa: crie cenários alternativos para estimar os retornos da Bovespa utilizando modelos ARCH.

11.4 Simulando modelos ARCH

Para simular modelos do tipo ARCH no R, propomos os seguintes parâmetros:

$\alpha_0 = 0.1, \alpha_1 = 0.4, \alpha_2 = 0.2$:

```
set.seed(20)
alpha0=0.1
alpha1=0.4
alpha2=0.2
w=rnorm(10000)
a=rep(0,10000)
```

```

h=rep(0,10000)
for(i in 3:10000){
h[i]=alpha0+alpha1*(a[i-1]^2)+alpha2*(a[i-2]^2)
a[i]=w[i]*sqrt(h[i])
}

```

A saída gráfica desta simulação pode ser visualizada com os seguintes comandos:

```

acf(a)
pacf(a^2)
plot(a,type="l")

```

11.5 GARCH

Este modelo é mais parcimonioso que os modelos ARCH, e é definido por

$$X_t = \sqrt{h_t} \epsilon_t, h_t = \alpha_0 + \sum_{i=1}^r \alpha_i X_{t-i}^2 + \sum_{j=1}^s h_{t-j} \quad (56)$$

Onde ϵ_t é normal ou t-student, além disso uma sequência de variáveis aleatórias independentes e identicamente distribuídas (iid) com média zero e variância um. $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0, \sum_i = 1^q (\alpha_i + \beta_j) < 1, q = \max(r, s)$. A função usada é **garch()**, do pacote **tseries** e o modelo padrão do R é um GARCH(1,1), com a seguinte estrutura:

$$y_t = \sigma_t \nu_t, \quad \nu \approx N(0, 1) i.i.d., \sigma_t = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2, \quad \omega, \alpha > 0, \beta \geq 0 \quad (57)$$

Para aplicar o modelo mais simples, para a série de retornos da Bovespa usamos os seguintes comandos:

```

> library(fGarch)
> garchr=garchFit(~arma(1,0)+garch(2,3),rt)
> plot(residuals(rar))
> plot(residuals(rarima))
> garchr=garchFit(~arma(1,0)+garch(1,1),rt)

> ajustegarch=garchr@fit$series$h
> predict(ajustegarch,5)

```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
4286	0.0003738443	0.0002657738	0.0004819149	2.085646e-04	0.0005391241
4287	0.0003771122	0.0002227740	0.0005314504	1.410723e-04	0.0006131521

```

4288    0.0003803800 0.0001886256 0.0005721345 8.711687e-05 0.0006736432
4289    0.0003836479 0.0001587702 0.0006085255 3.972721e-05 0.0007275685
4290    0.0003869157 0.0001314454 0.0006423860 -3.792403e-06 0.0007776238

```

Um modelo mais elaborado, é considerar que a série dos retornos segue um ARMA(10,0), e os resíduos um GARCH(1,1), sugerido por Moretin e Toloi, (2006) é:

```

garchmoretin=garchFit(formula = ~arma(10, 0) + garch(1, 1), data = rt)
> summary(garchmoretin)

```

Title:

GARCH Modelling

Call:

```
garchFit(formula = ~arma(10, 0) + garch(1, 1), data = rt)
```

Mean and Variance Equation:

```
data ~ arma(10, 0) + garch(1, 1)
```

```
<environment: 0x03740cc0>
```

```
[data = rt]
```

Conditional Distribution:

norm

Coefficient(s):

mu	ar1	ar2	ar3	ar4	ar5	ar6
1.55e-03	8.77e-03	-2.43e-02	-3.16e-02	-3.27e-02	-2.98e-02	-4.76e-03
ar7	ar8	ar9	ar10	omega	alpha1	beta1
-1.36e-02	1.54e-02	1.16e-02	4.16e-02	1.28e-05	1.22e-01	8.54e-01

Std. Errors:

based on Hessian

Error Analysis:

	Estimate	Std. Error	t value	Pr(> t)
mu	1.545e-03	2.791e-04	5.537	3.08e-08 ***
ar1	8.767e-03	1.629e-02	0.538	0.59042
ar2	-2.432e-02	1.603e-02	-1.517	0.12926
ar3	-3.155e-02	1.585e-02	-1.991	0.04650 *

```

ar4      -3.272e-02   1.599e-02   -2.046   0.04075 *
ar5      -2.985e-02   1.587e-02   -1.881   0.06000 .
ar6      -4.755e-03   1.583e-02   -0.300   0.76388
ar7      -1.357e-02   1.557e-02   -0.872   0.38335
ar8       1.542e-02   1.555e-02    0.991   0.32147
ar9       1.162e-02   1.546e-02    0.751   0.45252
ar10     4.164e-02   1.526e-02    2.729   0.00636 **
omega    1.269e-05   2.455e-06    5.168  2.36e-07 ***
alpha1   1.224e-01   1.160e-02   10.559  < 2e-16 ***
beta1    8.544e-01   1.404e-02   60.833  < 2e-16 ***

```

Log Likelihood:

```
10752.44    normalized:  2.509322
```

Description:

```
Fri Nov 04 16:34:36 2011 by user: user
```

Standardised Residuals Tests:

			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	426.7988	0
Shapiro-Wilk Test	R	W	0.9884059	0
Ljung-Box Test	R	Q(10)	6.86544	0.7380843
Ljung-Box Test	R	Q(15)	10.90783	0.7591006
Ljung-Box Test	R	Q(20)	14.24748	0.8177295
Ljung-Box Test	R^2	Q(10)	15.45189	0.1164347
Ljung-Box Test	R^2	Q(15)	16.84624	0.3281391
Ljung-Box Test	R^2	Q(20)	23.55899	0.262187
LM Arch Test	R	TR^2	15.2171	0.2297781

Information Criterion Statistics:

AIC	BIC	SIC	HQIC
-5.012109	-4.991320	-5.012130	-5.004766

```
gm=garchmoretin@fit$series$h #nível
```

```
predict(gm,5)
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
4286	0.0003759518	0.0002674082	0.0004844955	2.099486e-04	0.0005419550
4287	0.0003789948	0.0002228952	0.0005350944	1.402611e-04	0.0006177286
4288	0.0003820378	0.0001876322	0.0005764435	8.472005e-05	0.0006793556
4289	0.0003850809	0.0001568135	0.0006133482	3.597610e-05	0.0007341856
4290	0.0003881239	0.0001286022	0.0006476455	-8.780257e-06	0.0007850280

```
gmz=garchmoretin@fit$series$z #estimativa
```

```
predict(gmz,5)
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
4286	-0.0006119849	-0.03072869	0.02950472	-0.04667151	0.04544754
4287	-0.0006119849	-0.03072869	0.02950472	-0.04667151	0.04544754
4288	-0.0006119849	-0.03072869	0.02950472	-0.04667151	0.04544754
4289	-0.0006119849	-0.03072869	0.02950472	-0.04667151	0.04544754
4290	-0.0006119849	-0.03072869	0.02950472	-0.04667151	0.04544754

```
plot(ecdf(residuals(garchmoretin)))
```

```
x=seq(-0.3,0.3,0.01)
```

```
lines(x,pnorm(x,mean=mean(residuals(garchmoretin)), sd=sd(residuals(garchmoretin)))
```

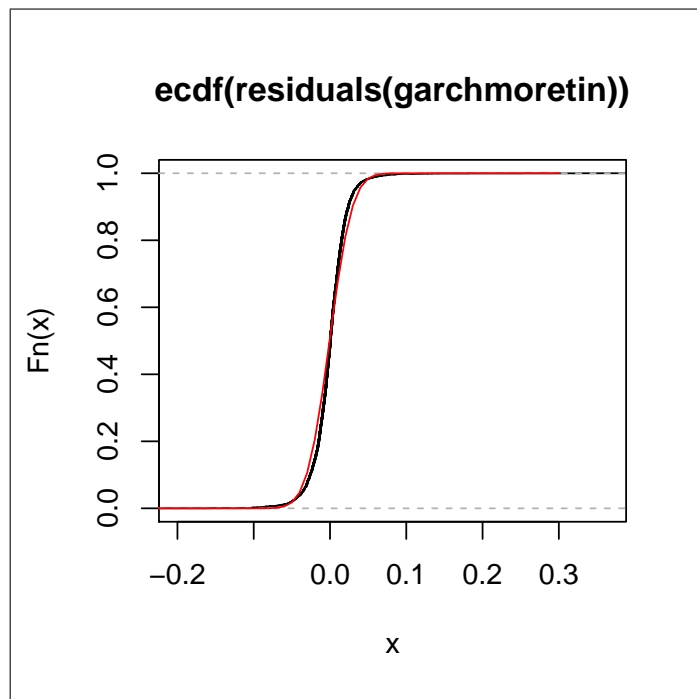


Figura 39. Distribuição dos resíduos padronizados dos retornos.

Interpretação: o ajustamento do modelo acima verificamos que os parâmetros do modelo são significantes. Os valores da estatística de Ljung-Box para os resíduos padronizados são dados por $Q(10)=6,86$ (0,74) e $Q(20)=14,25$ (0,82) enquanto que para os quadrados dos resíduos $Q(10)=15,45$ (0,12) e

$Q(20)=23,55$ (0,23). Analisando os níveis descritos podemos concluir que o modelo é adequado para modelar os retornos diários da IBovespa.

11.6 Simulando modelos GARCH

Para simular modelos do tipo GARCH no R, propomos os seguintes parâmetros: $\alpha_0 = 0.1, \alpha_1 = 0.4, \beta_1 = 0.2$:

```
set.seed(20)
alpha0=0.1
alpha1=0.4
beta1=0.2
w=rnorm(10000)
a=rep(0,10000)
h=rep(0,10000)
for(i in 2:10000){
  h[i]=alpha0+alpha1*(a[i-1]^2)+beta1*h[i-1]
  a[i]=w[i]*sqrt(h[i])
}
```

A saída gráfica desta simulação pode ser visualizada com os seguintes comandos:

```
acf(a)
pacf(a^2)
plot(a,type="l")
library(tseries)
a.garch=garch(a,grad="numerical", trace=FALSE)
confint(a.garch)
```

12 Raízes unitárias

A maioria das séries econômicas possuem raiz unitária. Quando a hipótese de raiz unitária for verdadeira para uma série, os choques aleatórios que ela sofre geram na mesma um efeito permanente. As flutuações neste caso não são transitórias, derrubando, por exemplo, as teorias de que os ciclos econômicos seriam flutuações temporárias em torno de uma tendência 1. Portanto, quando uma variável apresenta raiz unitária, os pressupostos estatísticos de que a média e a variância devem ser constantes ao longo do tempo são violados comprometendo, dessa forma, os resultados obtidos com

a utilização de modelos econométricos. A regressão, neste caso, é considerada espúria (sem significado econômico), Granger e Newbold (1974).

Em modelos de séries temporais em econometria (aplicação de métodos estatísticos na economia), a unidade de raiz é uma característica dos processos que evoluem ao longo do tempo e que podem causar problemas na inferência estatística, se não for tratada adequadamente. Um processo estocástico linear tem uma raiz unitária se 1 é raiz da equação característica do processo. Tal processo é não-estacionário. Se as outras raízes de equação característica dentro do círculo unitário - ou seja, têm um módulo (valor absoluto) menos de um -, então a primeira diferença do processo estiver parado. Embora uma série possua raiz unitária, se aplicarmos uma modelagem a esta série adequadamente, o seus resíduos tem que garantir a ausência de raiz unitária ($I(0)$).

Para modelos com tendência determinística é possível encontrar uma raiz unitária. Consideremos em particular um ruído branco, dizemos que uma série é $I(0)$, se segue um processo autorregressivo de primeira ordem $AR(1)$. Para gerar os dados considere o ruído branco ϵ_t dado por:

$$\epsilon_t = \alpha_0 \epsilon_{t-1} + \nu_t, \quad \nu \approx N(0, \sigma^2) \quad (58)$$

O objetivo é testar as seguintes hipóteses:

$H_0 : \alpha_0 = 1 \rightarrow \epsilon_t \approx I(1)$

$H_a : |\alpha| < 1 \rightarrow \epsilon_t \approx I(0)$

Há varias estatísticas para testar estas hipóteses, em particular as mais comuns são Phillips-Perron, Schmidt-Phillips e Dickey e Fuller, para implementar este teste considere o consumo de energia elétrica comercial mensal no Brasil em qde. - GWh - Eletrobras, no período de janeiro de 1976 a junho de 2013, extraídas do IPEADATA:

```
cee=read.table(file.choose(),header=T) #consumoenergia19761.txt
attach(cee)
plot.ts(ceh)
cee.ts=ts(ceh,start=c(1976,1),frequency=12)
mcee=auto.arima(cee.ts)
plot.ts(cee.ts, main="Consumo de Energia Elétrica\n em GW hora")
```

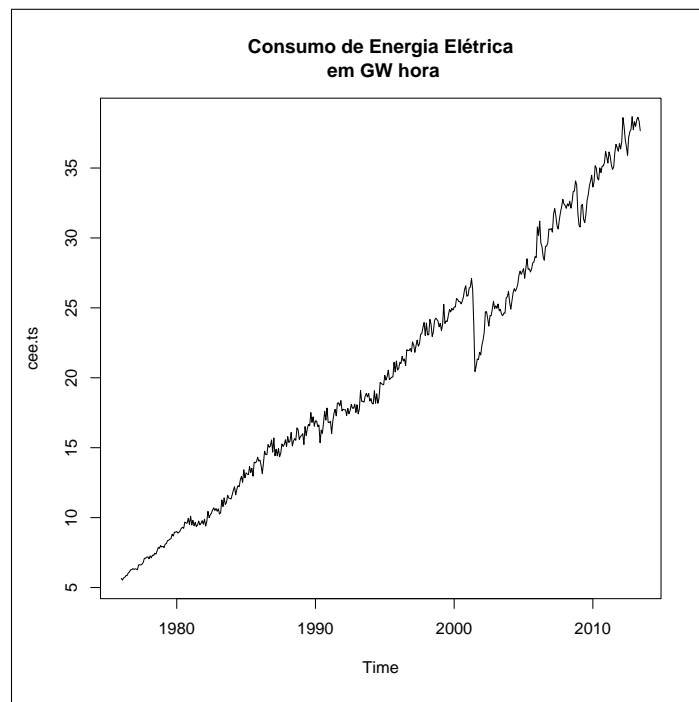



Figura 40. Consumo de Energia em GWh no Brasil.

Mostraremos inicialmente que esta série de consumo é Integrada de ordem 1, $I(1)$.

```
library(urca)
t= ur.pp(mcee) # teste de Phillips-Perron
summary(t)
```

```
t= ur.pp(ceh) # teste de Phillips-Perron
summary(t)
```

```
#####
# Phillips-Perron Unit Root Test #
#####
```

Test regression with intercept

Call:

```
lm(formula = y ~ y.l1)
```

```
Residuals:
```

Min	1Q	Median	3Q	Max
-3.3444	-0.3011	0.0073	0.3237	2.1227

```
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.088571	0.066269	1.337	0.182
y.l1	0.999150	0.002982	335.111	<2e-16 ***

```
Residual standard error: 0.5639 on 447 degrees of freedom  
Multiple R-squared: 0.996, Adjusted R-squared: 0.996  
F-statistic: 1.123e+05 on 1 and 447 DF, p-value: < 2.2e-16
```

```
Value of test-statistic, type: Z-alpha is: 0.0111
```

	aux. Z statistics
Z-tau-mu	1.4272

Como nosso segundo interesse é verificar o comportamento dos resíduos do modelo, em particular do mcee, podemos usar vários testes do pacote URCA, da seguinte forma:

```
library(urca)  
t1= ur.pp(residuals(mcee)) # teste de Phillips-Perron  
summary(t1)
```

```
#####  
# Phillips-Perron Unit Root Test #  
#####
```

```
Test regression with intercept
```

```
Call:  
lm(formula = y ~ y.l1)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-3.5462	-0.2173	-0.0022	0.2755	2.3182

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.01799	0.02256	0.797	0.426
y.l1	-0.03789	0.04726	-0.802	0.423

Residual standard error: 0.4778 on 447 degrees of freedom

Multiple R-squared: 0.001436, Adjusted R-squared: -0.0007984

F-statistic: 0.6426 on 1 and 447 DF, p-value: 0.4232

Value of test-statistic, type: Z-alpha is: -487.598

	aux. Z statistics
Z-tau-mu	0.7967

```
t2= ur.sp(residuals(mcee)) # teste de Schmidt-Phillips
summary(t2)
```

```
#####
# Schmidt-Phillips Unit Root Test #
#####
```

Call:

```
lm(formula = sp.data)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-3.5772	-0.2304	0.0265	0.2666	2.2849

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.0377262	0.0453056	-0.833	0.405
y.lagged	-0.0422358	0.0473110	-0.893	0.372

```
trend.exp1    0.0002469  0.0001741   1.418    0.157
```

Residual standard error: 0.4773 on 446 degrees of freedom

Multiple R-squared: 0.005915, Adjusted R-squared: 0.001457

F-statistic: 1.327 on 2 and 446 DF, p-value: 0.2664

Value of test-statistic is: -22.5192

Critical value for a significance level of 0.01

is: -3.59

```
t3= ur.df(residuals(mcee)) # teste de Dickey e Fuller
summary(t3)
```

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression none

Call:

```
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-3.4599	-0.2368	0.0148	0.2865	2.3084

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
z.lag.1	-0.93256	0.06783	-13.748	<2e-16 ***
z.diff.lag	-0.10030	0.04711	-2.129	0.0338 *

Residual standard error: 0.4763 on 446 degrees of freedom

Multiple R-squared: 0.5231, Adjusted R-squared: 0.521

F-statistic: 244.6 on 2 and 446 DF, p-value: < 2.2e-16

Value of test-statistic is: -13.7477

Critical values for test statistics:

	1pct	5pct	10pct
tau1	-2.58	-1.95	-1.62

Os três testes, confirmam que não há evidências para aceitar H_0 , em consequência, não é integrada de ordem 1, ou super consistente, $I(1)$, e sim integrada de ordem 0, $I(0)$. A partir das saídas dos testes, não há evidências para rejeitar H_0 . Por tanto há evidências da presença de um $AR(1)$.

13 Bolsa de Lima

13.1 IGBVL

O Índice Geral da Bolsa de Valores de Lima (IGBVL) é um indicador que mede o desempenho do mercado de ações e serve para comparação dos ganhos alcançados pelos diferentes setores (industrial, banca, agricultura, mineração, serviços públicos, etc.) participantes da Bolsa de Valores de Lima, em um determinado período de tempo. Este índice se determina partir de uma carteira composta por mais significativa de negociação de ações, selecionados com base em sua negociação de frequência, a quantidade de comércio e número de operações. Atualmente as 33 empresas que compõem o índice são: Alicorp, Alturas Minerals, Agroindustrial Pomalca, Atacocha, Austral Group, Banco Continental, Candente, Agroindustrial Casa Grande, Cementos Lima, Cementos Pacasmayo, Cerro Verde, Coporación Aceros Arequipa, Corporación Lindley, Credicorp, Edegel, El Brocal, Ferreyros, Gold Fields - La Cima, Graña y Montero, Intergroup Financial Services, Luz del Sur, Maple Energy, Milpo, Minera IRL, Minsur, Relapasa, Rio Alto Mining, Scotiabank Perú, Sider Perú, Simsa, Southern Copper Corporation, Telefónica, Volcan.

13.2 ISBVL

Índice seletivo da Bolsa de Valores de Lima (ISBVL): foi calculado desde julho de 1993, um indicador que mede as variações nos preços das 15 ações mais representante da Bolsa de Valores de Lima. Esse índice é usado para mostrar a tendência do mercado de ações em termos das mudanças que ocorrem nos preços das 15 ações mais representativas. Como o Índice Geral, a base é 100 e é datado em 30 de dezembro de 1991.

13.3 Índice de Nacional de Capitalização - INCA

O Índice Nacional de Capital (INCA) é uma estatística que pretende refletir o comportamento dos preços das principais ações listadas na Bolsa de Valores de Lima (BVL). O INCA é um índice de capitalização composta de uma carteira 20 ações diversificadas empresas peruanas. As ações INCA são parte de ações mais líquidas negociadas na LSE e os pesos das ações na carteira diversificada obtido em base la capitalização do mercado livre dessas ações. Para visualizar estas três carteiras, podemos apresentar o seguinte plot:

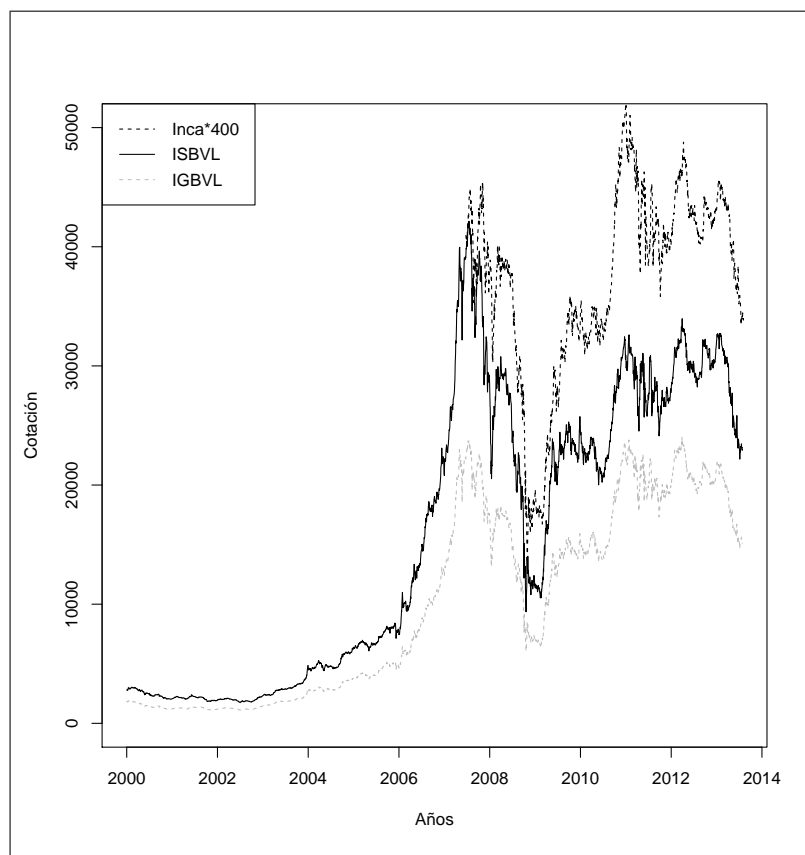


Figura 41. Indicadores na Bolsa de Lima.

Como podemos observar, visualmente as três séries são não estacionárias, podemos provar isto com os testes de raízes unitárias, onde:

```
> summary(ur.pp(igbvldiario.ts))
```

```
#####
# Phillips-Perron Unit Root Test #
#####
```

Test regression with intercept

```
Call:
lm(formula = y ~ y.l1)
```

```
Residuals:
      Min       1Q   Median       3Q      Max
-2641.03  -31.62   -5.34   43.04  1585.38
```

```
Coefficients:
              Estimate Std. Error  t value Pr(>|t|)
(Intercept)  8.14335     5.60833   1.452    0.147
y.l1         0.99958     0.00044 2272.000 <2e-16 ***
```

```
Residual standard error: 202.9 on 3390 degrees of freedom
Multiple R-squared:  0.9993,    Adjusted R-squared:  0.9993
F-statistic: 5.162e+06 on 1 and 3390 DF,  p-value: < 2.2e-16
```

Value of test-statistic, type: Z-alpha is: -1.8873

```
          aux. Z statistics
Z-tau-mu          1.4247
```

```
> summary(ur.pp(isbvldiario.ts))
```

```
#####
# Phillips-Perron Unit Root Test #
#####
```

Test regression with intercept

```
Call:
```

```
lm(formula = y ~ y.l1)
```

Residuals:

Min	1Q	Median	3Q	Max
-3639.9	-58.6	-8.7	74.4	2824.7

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.401e+01	9.638e+00	1.454	0.146
y.l1	9.995e-01	4.907e-04	2036.680	<2e-16 ***

Residual standard error: 342.9 on 3390 degrees of freedom

Multiple R-squared: 0.9992, Adjusted R-squared: 0.9992

F-statistic: 4.148e+06 on 1 and 3390 DF, p-value: < 2.2e-16

Value of test-statistic, type: Z-alpha is: -2.2872

	aux. Z statistics
Z-tau-mu	1.4506

```
> summary(ur.pp(incaibvldiario.ts))
```

```
#####
# Phillips-Perron Unit Root Test #
#####
```

Test regression with intercept

Call:

```
lm(formula = y ~ y.l1)
```

Residuals:

Min	1Q	Median	3Q	Max
-13.4795	-0.7167	0.0448	0.8191	7.9852

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.261788	0.185889	1.408	0.159

13.3 Índice de Nacional de Capitalização - INCA 13 BOLSA DE LIMA

y.11 0.997062 0.001954 510.355 <2e-16 ***

Residual standard error: 1.536 on 1528 degrees of freedom
Multiple R-squared: 0.9942, Adjusted R-squared: 0.9942
F-statistic: 2.605e+05 on 1 and 1528 DF, p-value: < 2.2e-16

Value of test-statistic, type: Z-alpha is: -5.3277

 aux. Z statistics
Z-tau-mu 1.5429

Uma comparação gráfica entre o Índice da Bolsa de São Paulo (Bovespa) e a Bolsa de Lima (IGBVL) é proposta, da seguinte maneira:

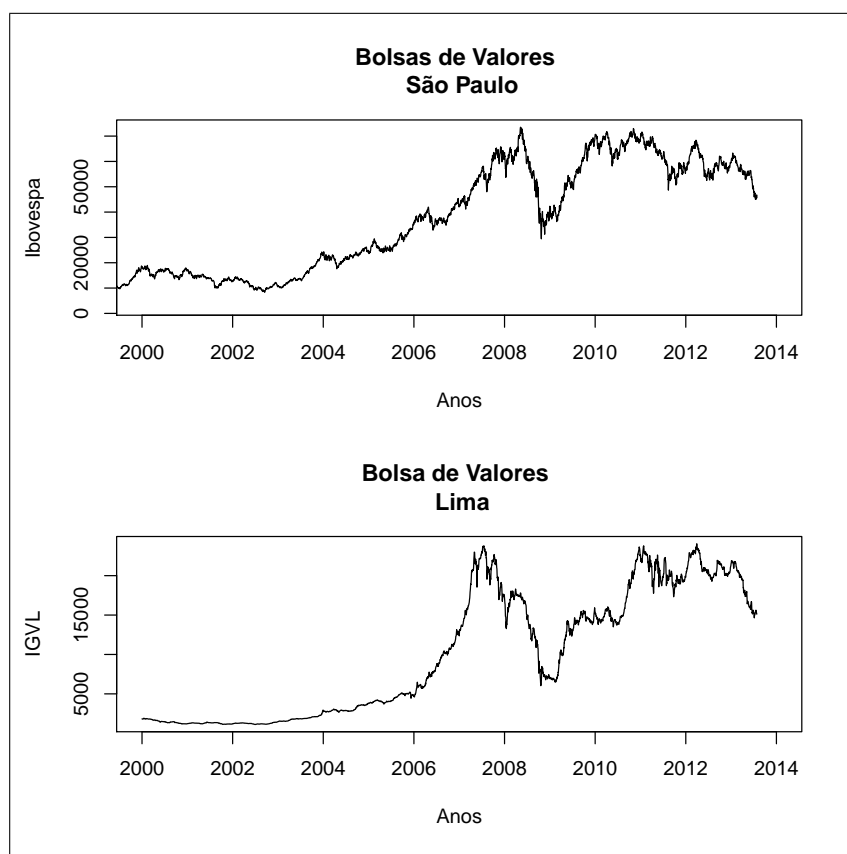


Figura 42. Indicadores da Bolsa de São Paulo e Bolsa de Lima.

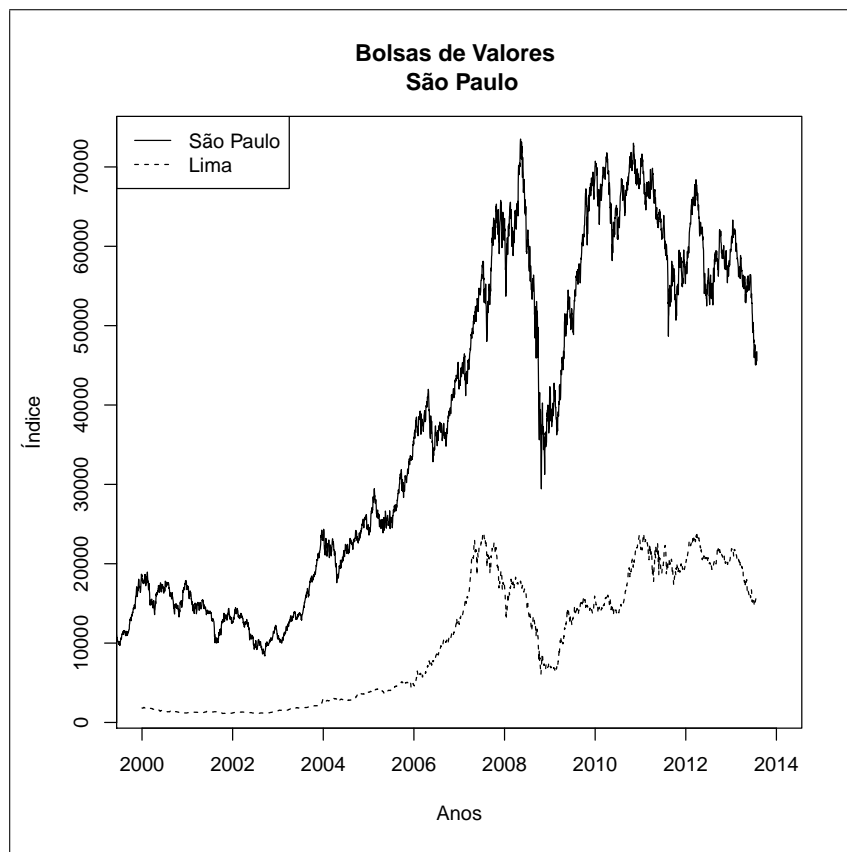


Figura 43. Sobreposição das Bolsas de São Paulo e Lima.

13.4 A crises mundial de 1994

A crise econômica do México de 1994, mais conhecida como Efeito Tequila, teve repercussões mundiais. Foi provocada pela falta de reservas internacionais, causando desvalorização do peso, durante os primeiros dias da presidência de Ernesto Zedillo. Tratou-se de uma crise de balança de pagamentos associada a especulação financeira e fuga de capitais, resultantes de uma crise política interna do México. Até esse momento, a economia mexicana era uma referência para o sistema financeiro internacional, por ser considerado um país moderno e alinhado as reformas do consenso de Washington, tendo recebido grande quantidade de investimento externo. Com a crise política resulta numa crise de confiança e na desestabilização da

economia e uma fuga dos investimentos para o exterior. bEm 20 de dezembro de 1994, apenas três semanas após o início do governo Ernesto Zedillo, o Ministério da Fazenda ampliou a banda cambial em 15,3%. Os investidores entraram em pânico e iniciaram uma corrida ao peso, no dia seguinte, a peg (paridade) foi abandonada e o peso flutuou livremente em relação ao dólar. O valor do peso imediatamente à metade de seu valor nominal, mergulhando o México numa depressão surpreendentemente profunda. Uma depressão que fez com que em todo o mundo caíssem as cotações dos títulos dos países emergentes. Mais de 200 mil mexicanos perderam seus empregos e milhares de empresas fecharam as portas. A taxa de desemprego era o dobro do ano anterior. Em 1995, a chute do PIB mexicano será de 7%. Revista Veja, Março de 1995.

13.5 A crises de 1997

A Crise financeira asiática foi um período que atingiu grande parte da Ásia, tendo começado no verão de 1997 gerando temor de uma crise em escala mundial e contágio financeiro. Essa crise é comumente conhecida como Crise monetária do sudeste asiático. A crise começou na Tailândia com o colapso financeiro do Thai baht causado pela decisão do governo tailandês de tornar o câmbio flutuante, desatrelando o baht do dólar, após exaustivos esforços para evitar a massiva fuga de capitais em parte devido ao estado. Na época, a Tailândia adquiriu uma enorme dívida externa que acabou por deixar o país falido logo após esse colapso monetário. A drástica redução das importações resultante da desvalorização tornou a reabilitação das reservas cambiais impossíveis a longo e médio prazo sem uma audaciosa intervenção internacional. Após o agravamento da situação, a crise se espalhou para o Sudeste Asiático e o Japão, afundando cotações monetárias, desvalorizando mercados de ações, e precipitando a dívida privada. O que parecia ser uma crise regional com o tempo se converteu no que se denominou a primeira grande crise dos mercados globalizados, de cujos efeitos existe uma grande incerteza sobre a verdadeira magnitude de seu impacto na economia mundial.

13.6 A crises mundial de 1998

Conhecida como Moratória Russa, onde a Rússia passou por uma profunda crise econômica nos anos 1990, com altas taxas de endividamento, desemprego e inflação e baixos índices de crescimento econômico (PIB). Em grande medida este processo foi resultado de uma transição acelerada e mal

sucedida de uma economia planificada para uma economia de mercado, em meio ao colapso político da União Soviética. A crise da economia planificada soviética tem início nos anos 1970, mas foi maquiada com a alta no preço das commodities agrícolas e minerais, especialmente do petróleo, após a crise petrolífera de 1973 e a crise petrolífera de 1979-1980. Estes produtos eram exportados em grande quantidade pela então URSS, que também havia aumentado as exportações militares aos países do Terceiro Mundo. Com uma economia aquecida e excesso de moedas fortes nas contas do país, os efeitos negativos da economia planificada não eram percebidos como sérios. Entretanto, a queda no preço das commodities agrícolas, minerais e energéticas (petróleo, gás natural) a partir de 1984-1985, deixou claro os limites daquele modelo. A Perestroika era um plano ousado para realizar uma transição controlada para uma economia de mercado, que fracassou devido ao colapso econômico do país ainda nos anos 1980. A crise econômica foi agravada pelo colapso político da URSS e a desintegração territorial da União Soviética em 1991. A partir de 1992, a Rússia procura implementar uma política de choque econômico em direção ao capitalismo de mercado que foi desastroso, pois nem conseguiu reestruturar os setores produtivos tradicionais nem implementar outros novos. A falência de milhares de empresas levou ao aparecimento de milhões de desempregados, acompanhado de altas taxas de violência urbana. O período pós-1992 foi de grande turbulência econômica, quando o país mergulhou em profunda crise econômica, apresentando taxas negativas de crescimento do PIB, altas taxas de inflação e elevado desemprego. O consumo total de energia na ex-URSS caiu em quase 50% quando comparada ao período soviético (até 1991). O desemprego atingia 15% da população economicamente ativa, e 35% dos russos passaram a viver abaixo da linha da pobreza. Por Revista Veja, 1999.

13.7 A crises mundial de 2008

Em 2008 os Estados Unidos se envolveu com duas guerras, Iraque e Afeganistão, gastando mais do que deveria, ao mesmo tempo em que o país investia dinheiro na guerra, a economia interna já não ia muito bem, uma das razões é que os Estados Unidos estavam importando mais do que exportando. Em vez de conter os gastos, os americanos receberam ajuda de países como China e Inglaterra. Com o dinheiro injetado pelo exterior, os bancos passaram a oferecer mais crédito, inclusive a clientes considerados de risco. Aproveitando-se da grande oferta a baixas taxas de juros, os consumidores compraram muito, principalmente imóveis, que começaram a valorizar. A expansão do crédito financiou a bolha imobiliária, já que a grande procura

elevou o preço dos imóveis. Porém, depois disso, chegou uma hora em que a taxa de juros começou a subir, diminuindo a procura pelos imóveis e derrubando os preços. Com isso, começou a inadimplência, afinal, as pessoas já não viam sentido em continuar pagando hipotecas exorbitantes quando as propriedades estavam valendo cada vez menos. A Casa Branca decidiu não mais interferir, deixando o banco Lehman Brothers quebrar em setembro de 2008. O fechamento do quarto maior banco de crédito dos Estados Unidos causou pânico e travou o crédito. Chegou a crise, que prejudica também muitos países em particular os latinoamericanos como Brasil e Perú, que tiveram um retrocesso de 2 anos na queda dos seus indicadores, ver figura 42 e 43. Sem crédito internacional, também diminui o crédito no Brasil e no Perú, caem as exportações e o preço das nossas mercadorias aumenta o risco e a taxa de juros.

13.8 A crises mundial de 2010

Com prolongamento da grande crise que se iniciou em 2008, a crise da zona do Euro, desatada pelos acontecimentos na Grécia em 2010, recolocou na agenda mundial o embate sobre a regulação versus livre mercado. O choque de fundo que se manifesta na zona do Euro é essencialmente sobre a forma de organizar a economia e a sociedade. A origem da crise da zona do Euro, que teve seu epicentro na Grécia, tem a ver com o ingresso do país na Comunidade Econômica Europeia (CEE). Como todo país que ingressou no euro, a Grécia teve que, além de cumprir uma série de metas fiscais, monetárias e financeiras, renunciar à possibilidade de emitir sua própria moeda. Esse privilégio ficou nas mãos do Banco Central Europeu (BCE), entidade supranacional que funciona como um banco central independente. Logo, os países ficam seriamente restringidos em suas políticas econômicas pela dificuldade de obter créditos. Encontram-se nessa situação os chamados países denominados PIIGE (anagrama para Portugal, Irlanda, Itália, Grécia e Espanha). O mesmo não ocorre, por exemplo, com países que também estão na zona do Euro, mas que optaram por continuar com moeda própria. É o caso da Inglaterra. A Grécia violou a regra da zona do Euro de que o déficit orçamentário não deveria ultrapassar 3% do PIB do país. Outros países já violaram. Na Grã-Bretanha que, como se disse, não está na zona do euro, esse déficit chega a 13% do PIB. Na Espanha ele chega a 11,2% por cento, na Irlanda a 14,3% por cento e na Itália a 5,3%. A guerra cambial é outro desdobramento da crise econômica mundial de 2008 que estouro em 2010, que dá força às teses, como dissemos anteriormente, aos que afirmavam

que as consequências dessa crise se fariam sentir por muito tempo. A essência da guerra cambial está relacionada à excessiva liquidez de dólares no mundo e às resistências da China em valorizar a sua moeda. Trata-se, portanto, de uma guerra de gigantes como destacou o economista Guilherme Delgado em entrevista à IHU On-Line. Por Rudá Ricci, doutorado em Ciências Sociais pela Universidade Estadual de Campinas, no (2002).

Será mesmo que vivemos uma nova crise? Ou é apenas uma crise continuada de 2008? A verdade é que não importa se é ou não a mesma, mas sim qual a magnitude que os atuais problemas irão trazer.

13.9 Continuação da crise 2008

Para entendermos melhor, voltemos no tempo. A crise financeira de 2008 foi gerada por uma bolha imobiliária americana. Mas agora o problema é outro!. Os próprios países estão extremamente endividados. Portugal, Irlanda, Itália, Grécia e Espanha têm dívidas grandes, por exemplo a Itália deve 120% de seu PIB. Há ainda um agravante a Europa está estagnada. É um continente de idade avançada da população. Desemprego está alto o que dificulta o crescimento econômico. Quem resgata hoje os estados? Quem sofrerá com um calote da dívida grega? Bancos franceses e alemães têm boa parte da dívida. Eles sofreriam prejuízos enormes em seus balanços. Poderíamos então ter um efeito dominó. Bancos novamente quebrando e estados sem recursos para resgatá-los. O pior dos mundos. Para agravar mais a situação, a maior economia mundial, os Estados Unidos, não conseguem engrenar. Os estímulos não estão funcionando e o desemprego não é reduzido. Novos estímulos estão sendo anunciados, mas o ceticismo é grande diante do que realmente vai acontecer com os novos pacotes americanos.

No âmbito Peruano, a dívida externa tem-se reduzido gradualmente ano a ano, por exemplo a dívida do país se ha reduzido comparada com o crescimento de nossa economia; em 1985 represento el 80% do produto interno bruto e em 2001 era de 50, 4%, em dezembro de 2011 foi de 21, 7%. Mas isto não foi impedimento para ter um impacto significativo na bolsa de Lima no ano de 2011 e 2012. Em 2013 a crise econômica peruana continua, e 73% das empresas deixaram de exportar devido a paisagem interna e externa difícil que atravessa Peru. As dificuldades encontradas na economia são muitas, derrapagens de custos logísticos, trâmites burocráticos, a discrição dos funcionários e da instabilidade da taxa de câmbio, as quais reduzem a competitividade do comércio exterior e a alerta do ex-ministro da Economia e Finanzas, Pedro Pablo Kuczynski, sobre o perigo de uma desvalorização excessiva do Nuevo Sol, a qual tiveram uma desvalorização

de 10% e 11% nos últimos anos, o que enfraqueceu ainda mais a moeda.

13.10 O impacto das crises na Bolsa

A crise mundial de 1998, conhecida como a moratória da Rússia, teve impacto positivo no Brasil, chegando a ter uma volatilidade de aproximadamente 0.112, ou um lucro financeiro na ordem de 33,41% dos ativos num dia de pregão. Uma outra crise que favoreceu o Brasil num dia de pregão é a crise do México, chegando a ter uma volatilidade de aproximadamente de 0.06, com um lucro financeiro de 25,63% dos ativos. A crise dos Imobiliária do Estados Unidos, trouxe dias de lucros e perdas, deixando o mercado instável e nervoso, nesta época num dia de pregão o Brasil teve uma volatilidade de 0,021 ou lucros na ordem de 14,65% e nos dias posteriores uma volatilidade 0.013 ou de perdas na ordem de 11,4%. Para visualizar este impacto nas bolsas do Brasil, utilizamos os quadrados dos retornos financeiros da cotação da bolsa (Variância condicional), e os retornos, no caso do Brasil, temos:

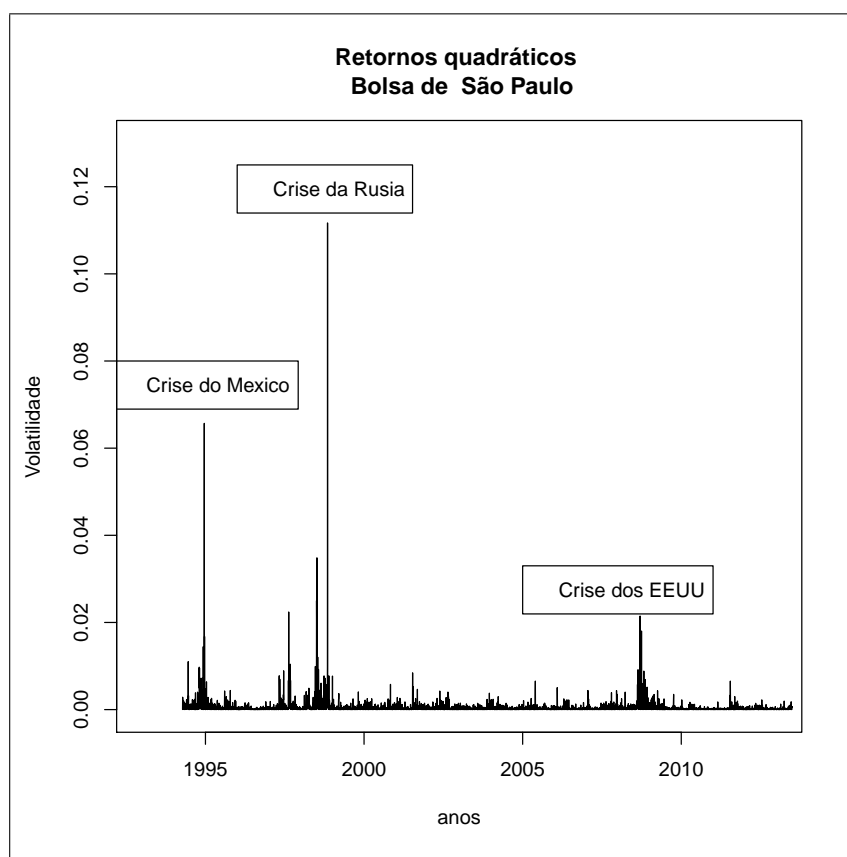


Figura 44.1. Identificação das crises no Brasil usando a variância condicional.

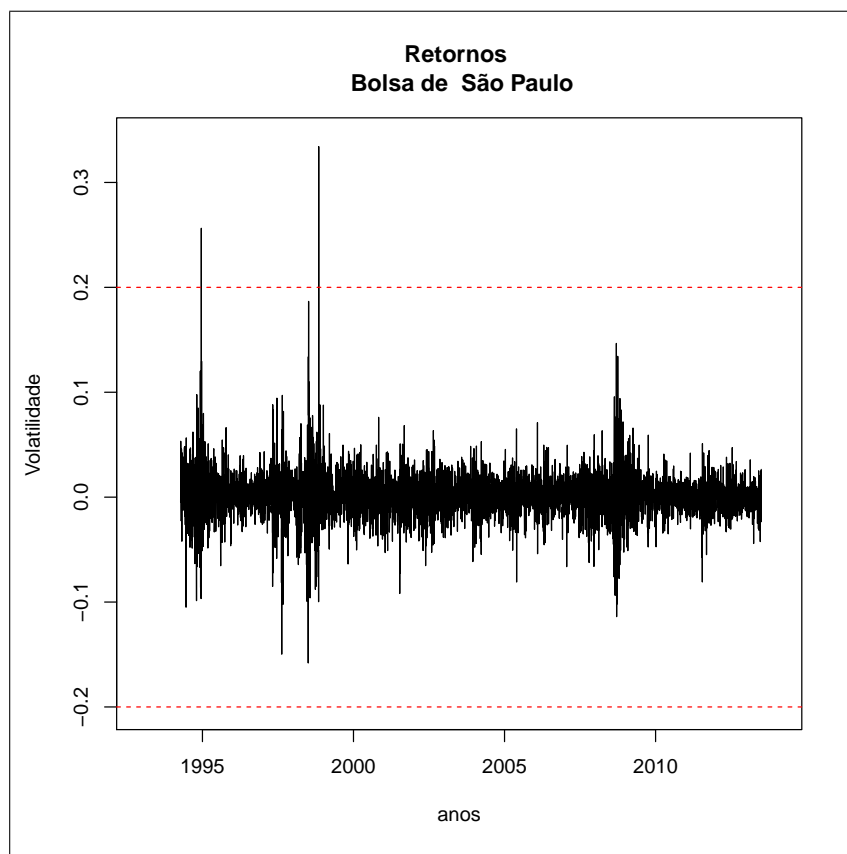


Figura 44.2. Identificação das crises no Brasil usando os retornos.

Os riscos do mercado de ações no Perú (desvio padrão do retorno financeiro), frente as crises tiveram os seguintes impactos: em 2011, com volatilidade 0,0154 e com uma perda de 0,12% num dia de pregão. A crise de 2008 também trouxe dias de lucros e perdas, deixando o mercado instável e nervoso, nesta época a bolsa de Lima obteve num dia de pregão um lucro máximo de 13,67% nos ativos, após que no dia anterior teve uma perda de 10,81% nestes mesmos ativos. Isto pode ser observado no gráfico dos retornos quadráticos:

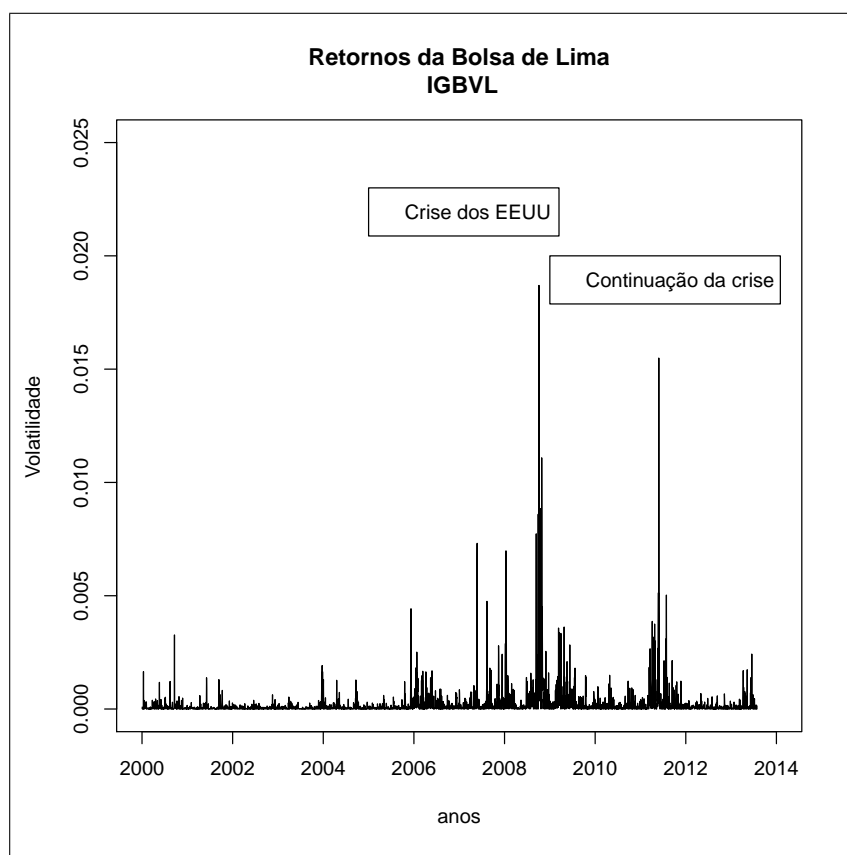


Figura 45. Identificação das crises no Perú usando a variância condicional.

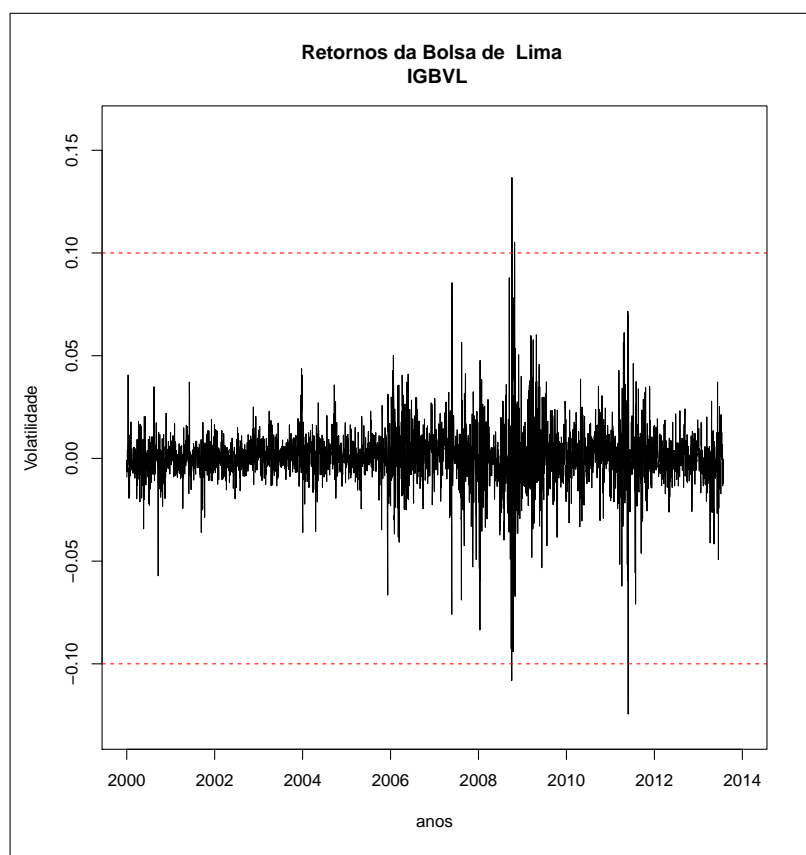


Figura 46. Identificação das crises no Perú usando os retornos.

Para visualizar que as carteiras de ativos no mercado de ações (bursatil) peruano apresenta riscos diferenciados que no Brasil, comparamos os dois retornos:

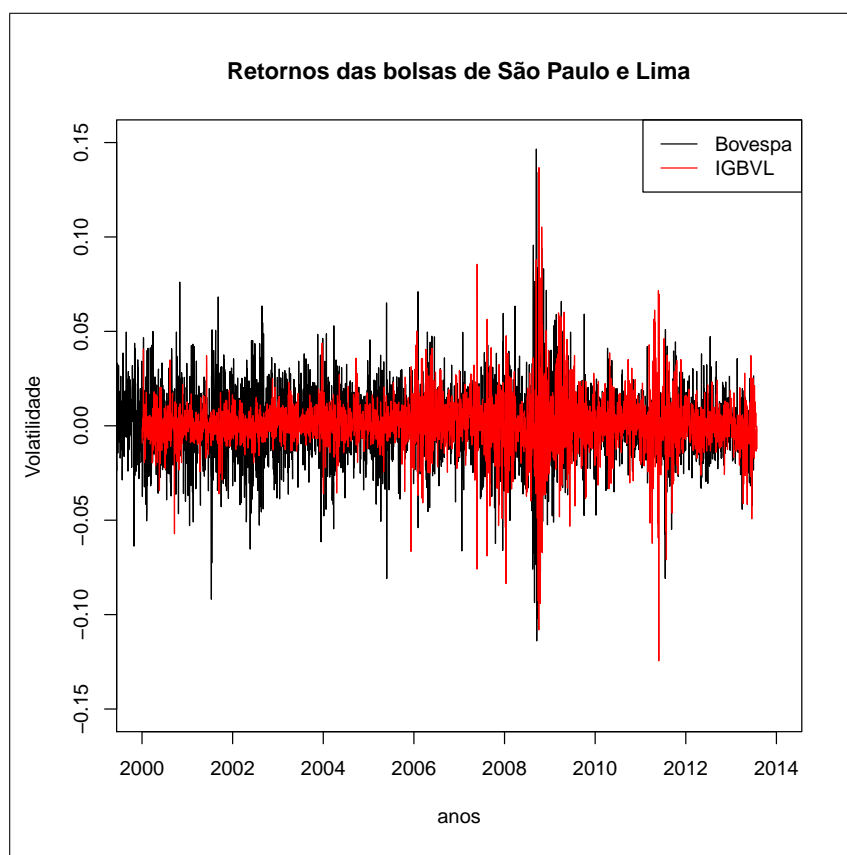


Figura 47. Retornos da IBOVESPA e LIMA.

Para visualizar com mais cuidado criamos dois cenários: O primeiro de 2000 a 2006 e o segundo de 2006 a 2013, ver figuras 48 e 49.

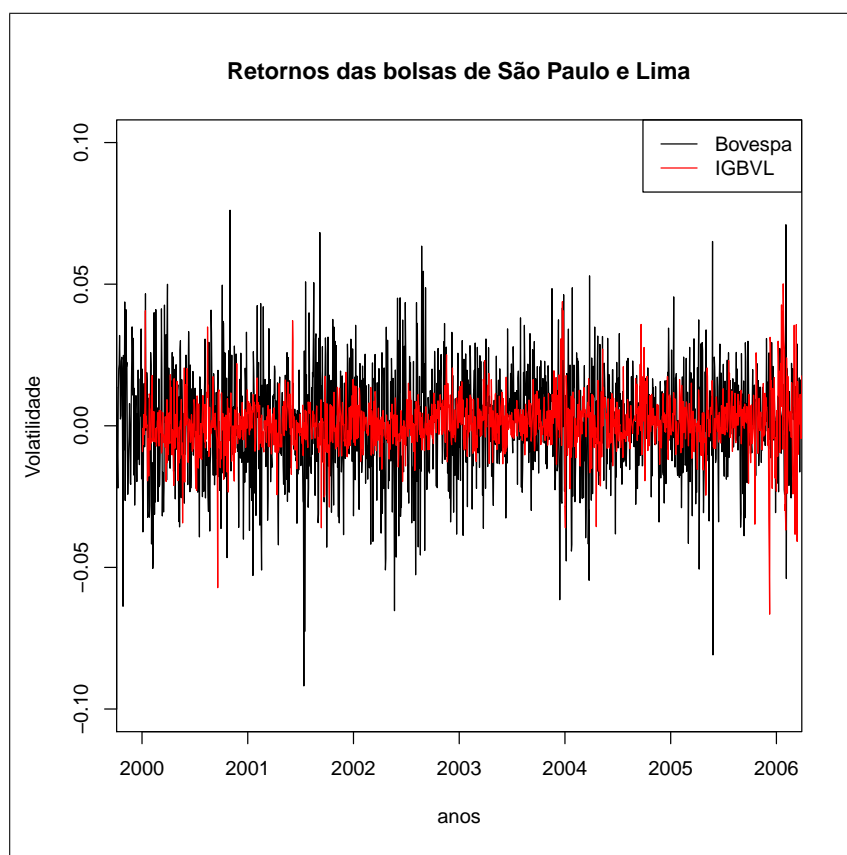


Figura 48. Cenário 1. 2000-2006.

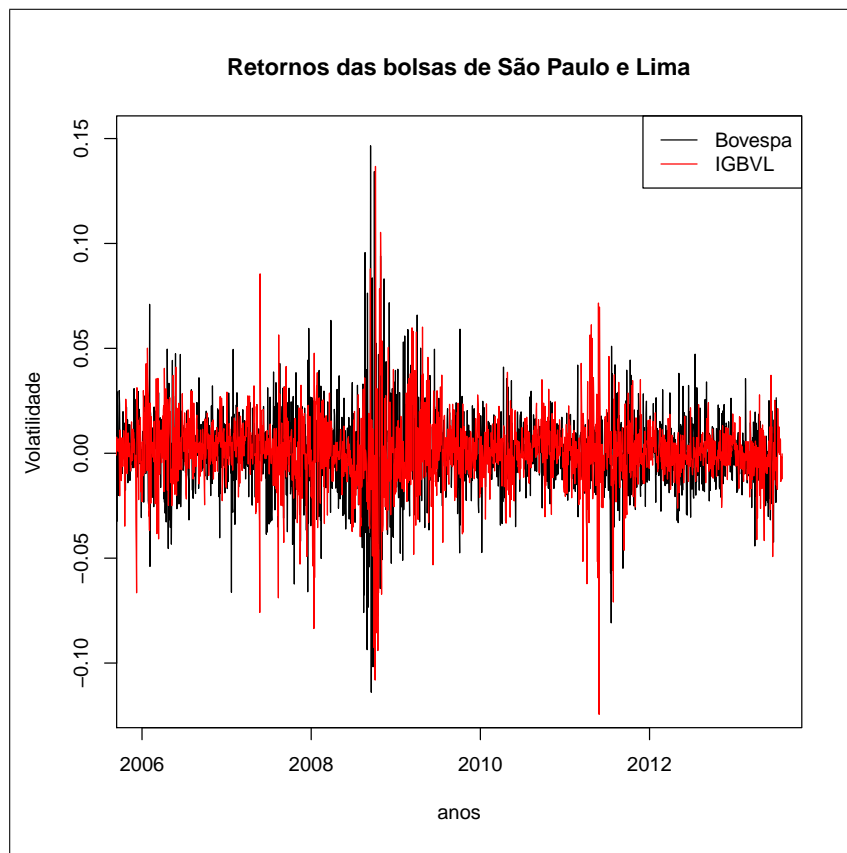


Figura 49. Cenário 2. 2006-2013.

14 Analisando as bolsas de Lima

Nas seções anteriores verificamos que as bolsas de lima **IGBVL**, **ISBVL**, e **INCA**, são não estacionárias, em consequência estas séries apresentam além de sazonalidade também tendência, além de apresentar efeito temporal. Para incluir o efeito temporal das três séries, usamos:

```
library(forecast)
igbvldiario=scan()
incaibvldiario=scan()
isbvldiario=scan()
igbvldiario.ts=ts(igbvldiario,start=c(2000,1),frequency=249)
isbvldiario.ts=ts(isbvldiario,start=c(2000,1),frequency=249)
incaibvldiario.ts=ts(incaibvldiario,start=c(2007,120),frequency=249)
```

14.1 Usando Holt Winters

Após de incluir o efeito temporal, uma proposta de aplicar modelos de alisamento exponencial pode ser nosso pontapé inicial para abordar estas séries. No R usamos a função `HoltWinters()`, para os modelos de Holt-Winters, podendo ser aditivos e multiplicativos da seguinte forma:

```
hwig=HoltWinters(igbvldiario.ts)
hwis=HoltWinters(isbvldiario.ts)
hwin=HoltWinters(incaibvldiario.ts)
```

Para visualizar as estatísticas dos parâmetros e a saída gráfica usamos os comandos:

```
hwig
hwis
hwin
```

```
plot(igbvldiario.ts, main="Cotação do IGBVL e Previstos")
lines(fitted(hwig)[,1],col=2)
plot(isbvldiario.ts, main="Cotação do ISBVL e Previstos")
lines(fitted(hwis)[,1],col=2)
plot(incaibvldiario.ts, main="Cotação do INCA e Previstos")
lines(fitted(hwin)[,1],col=2)
```

Por exemplo para o IGBVL, o valor das cotações e suas estimativas são:

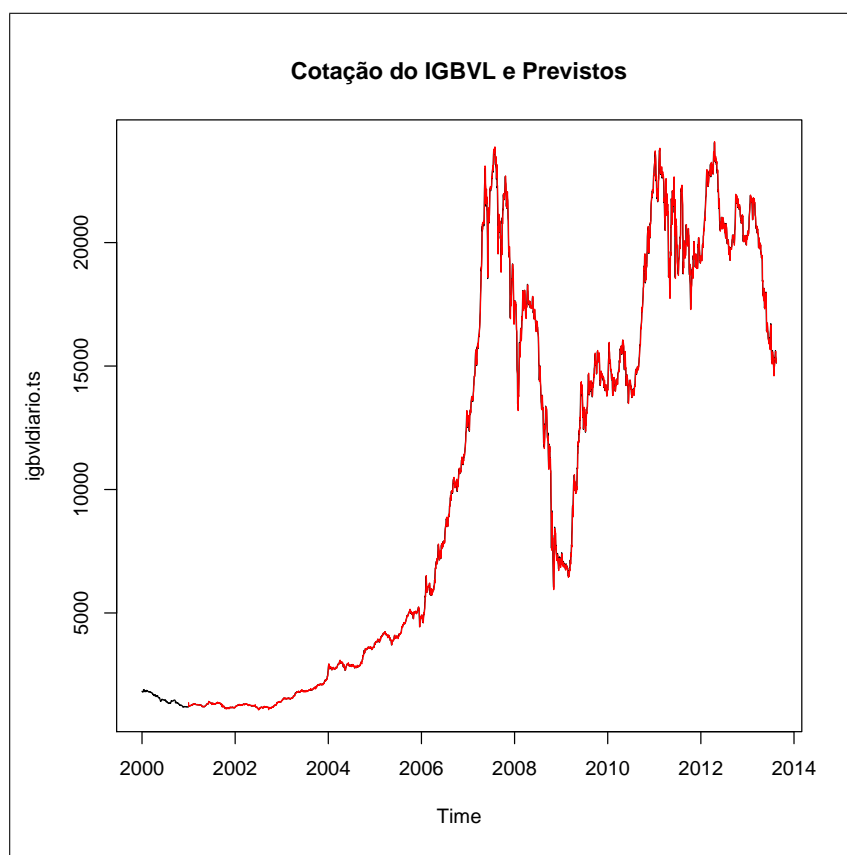


Figura 50. Modelo de HoltWinters para o IGBVL.

As estimativas de projeção para 5 dias úteis são:

```
cbind(predict(hwig,5),predict(hwis,5),predict(hwin,5))
Time Series:
Start = c(2013, 157)
End = c(2013, 161)
Frequency = 249
```

	predict(hwig, 5)	predict(hwis, 5)	predict(hwin, 5)
2013.627	15105.78	23042.65	85.21063
2013.631	15093.40	23054.76	86.64467
2013.635	15044.83	23037.48	85.77179
2013.639	15011.66	23032.96	85.99027
2013.643	14988.13	23047.16	86.17019

Contudo uma medida universal de erro de medida que serve para ajustar uma série é o MAPE, que no caso é:

```
mapehw1=mean(abs((igbvldiario.ts[-c(1:249)]-fitted(hwig))/
igbvldiario.ts[-c(1:249)]))*100
mapehw2=mean(abs((isbvldiario.ts[-c(1:249)]-fitted(hwis))/
isbvldiario.ts[-c(1:249)]))*100
mapehw3=mean(abs((incaibvldiario.ts[-c(1:249)]-fitted(hwin))/
incaibvldiario.ts[-c(1:249)]))*100

mapehw1
mapehw2
mapehw3

[1] 50.73032
[1] 50.78463
[1] 53.36111
```

A partir desta medida podemos afirmar que o erro de previsão é maior que 5% em consequência, abordaremos outros modelos para estimar estas séries.

14.2 Regressão dinâmica

A proposta que a cotação da bolsa de hoje depende da cotação de dias prévios, é usada como hipóteses para estimar modelos dinâmicos:

```
library(dynlm)
library(forecast)
library(tseries)
migbvl=dynlm(igbvldiario.ts~L(igbvldiario.ts))

summary(migbvl)

Time series regression with "ts" data:
Start = 2000(2), End = 2013(156)

Call:
dynlm(formula = igbvldiario.ts ~ L(igbvldiario.ts))

Residuals:
```


Min	1Q	Median	3Q	Max
-2641.03	-31.62	-5.34	43.04	1585.38

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	8.14335	5.60833	1.452	0.147
L(igbvldiario.ts)	0.99958	0.00044	2272.000	<2e-16 ***

Residual standard error: 202.9 on 3390 degrees of freedom

Multiple R-squared: 0.9993, Adjusted R-squared: 0.9993

F-statistic: 5.162e+06 on 1 and 3390 DF, p-value: < 2.2e-16

```

migbvls=dynlm(igbvldiario.ts~L(igbvldiario.ts)-1)
summary(migbvls)
mi2gbvl=dynlm(igbvldiario.ts~L(igbvldiario.ts,2))
summary(mi2gbvl)
mi3gbvl=dynlm(igbvldiario.ts~L(igbvldiario.ts)+L(igbvldiario.ts,2))
summary(mi3gbvl)

```

Os respectivos MAPES para estes modelos são:

```

mapemig=mean(abs((igbvldiario.ts-fitted(migbvls))/igbvldiario.ts))*100
mapemig
[1] 1.001344

```

```

mapemigs=mean(abs((igbvldiario.ts-fitted(migbvls))/igbvldiario.ts))*100
mapemigs
[1] 0.9581274

```

```

mapemig2=mean(abs((igbvldiario.ts-fitted(mi2gbvl))/igbvldiario.ts))*100
mapemig2
[1] 1.613056

```

```

mapemig3=mean(abs((igbvldiario.ts-fitted(mi3gbvl))/igbvldiario.ts))*100
mapemig3
[1] 0.9872054

```

As previsões para alguns destes modelos são:

```
predict(fitted(migbvls),5)
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2013.627	15703.66	15428.62	15974.03	15281.50	16127.60
2013.631	15653.58	15239.54	16058.36	15018.99	16284.23
2013.635	15982.03	15451.35	16490.13	15176.96	16769.45
2013.639	16073.75	15433.85	16670.94	15149.75	16980.29
2013.643	16148.04	15418.54	16833.60	15090.59	17222.62

```
predict(fitted(mi2gbvl),5)
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2013.627	15426.47	15165.24	15703.79	15016.91	15838.98
2013.631	15458.19	15055.39	15862.70	14848.62	16079.93
2013.635	15386.53	14871.85	15888.92	14620.33	16165.20
2013.639	15696.28	15078.31	16297.59	14794.96	16639.66
2013.643	15772.08	15070.89	16473.45	14742.32	16832.86

```
predict(fitted(mi3gbvl),5)
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2013.627	15714.89	15402.54	16020.30	15245.10	16162.62
2013.631	15634.05	15192.48	16067.54	14952.58	16307.12
2013.635	15996.08	15446.96	16555.45	15160.61	16847.44
2013.639	16048.30	15399.52	16688.59	15045.84	17067.27
2013.643	16112.99	15370.79	16853.41	14987.48	17234.76

```
plot.ts(igbvldiario.ts)
lines(fitted(migbvl),col=2)
```

Para O ISBVL, temos:

```
misbvl=dynlm(isbvldiario.ts~L(isbvldiario.ts))
summary(misbvl)
```

Time series regression with "ts" data:
Start = 2000(2), End = 2013(156)

Call:
dynlm(formula = isbvldiario.ts ~ L(isbvldiario.ts))

Residuals:

Min	1Q	Median	3Q	Max
-3639.9	-58.6	-8.7	74.4	2824.7

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.401e+01	9.638e+00	1.454	0.146
L(isbvldiario.ts)	9.995e-01	4.907e-04	2036.680	<2e-16 ***

Residual standard error: 342.9 on 3390 degrees of freedom
 Multiple R-squared: 0.9992, Adjusted R-squared: 0.9992
 F-statistic: 4.148e+06 on 1 and 3390 DF, p-value: < 2.2e-16

```
> misbvls=dynlm(isbvldiario.ts~L(isbvldiario.ts)-1)
> mapemisbvl=mean(abs((isbvldiario.ts-fitted(misbvl))/isbvldiario.ts))*100
> mapemisbvl
[1] 1.111277
mapemisbvls=mean(abs((isbvldiario.ts-fitted(misbvls))/isbvldiario.ts))*100
mapemisbvls
[1] 1.070792
predict(fitted(misbvl),5)
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
2013.627      24032.69 23578.85 24488.52 23347.73 24715.71
2013.631      23963.75 23277.75 24604.66 22924.88 24995.82
2013.635      24491.02 23609.29 25335.43 23189.48 25778.21
2013.639      24629.88 23581.67 25646.26 23081.89 26198.33
2013.643      24756.26 23527.34 25906.47 22961.26 26552.20
```

Para o INCA, os modelos dinamicos propostos são:

```
mincabvl=dynlm(incaibvldiario.ts~L(incaibvldiario.ts))
```

```
mincabvls=dynlm(incaibvldiario.ts~L(incaibvldiario.ts)-1)
```

```
mapemincabvl=mean(abs((incaibvldiario.ts-fitted(mincabvl))/incaibvldiario.ts))*100
mapemincabvl
[1] 1.248479
```

```
mapemincbvls=mean(abs((incaibvldiario.ts-fitted(mincabvls))/incaibvldiario.ts))*100
```

```
mapemincabvls  
[1] 1.247656
```

```
predict(fitted(mincabvl),5)  
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95  
2013.627      86.84267 85.21644 88.46891 84.35557 89.32978  
2013.631      86.45729 84.05508 88.85950 82.78343 90.13115  
2013.635      88.40474 85.39183 91.41765 83.79690 93.01259  
2013.639      88.81045 85.27023 92.35067 83.39615 94.22475  
2013.643      89.12009 85.10638 93.13380 82.98165 95.25853
```

14.3 Abordagem Arima

Para abordar estes modelos, Box-Jenkins, propuseram várias etapas de trabalho, como identificação, estimação, validação, e previsões. Na Identificação usamos as funções de autocorrelação e autocorrelação parcial. Para implementar estas funções no R, para as três bolsas de lima, usamos:

```
par(mfrow=c(3,2))  
acf(igbvldiario,main="Autocorrelación \n bolsa de LIMA (IGBVL)")  
pacf(igbvldiario, main="Autocorrelación Parcial \n bolsa de LIMA (IGBVL)")  
acf(isbvldiario, main="Autocorrelación \n bolsa de LIMA (ISGBVL)")  
pacf(isbvldiario, main="Autocorrelación Parcial \n bolsa de LIMA (ISBVL)")  
acf(incaibvldiario,main="Autocorrelación \n bolsa de LIMA (INCA)")  
pacf(incaibvldiario, main="Autocorrelación Parcial\n bolsa de LIMA (INCA)")
```

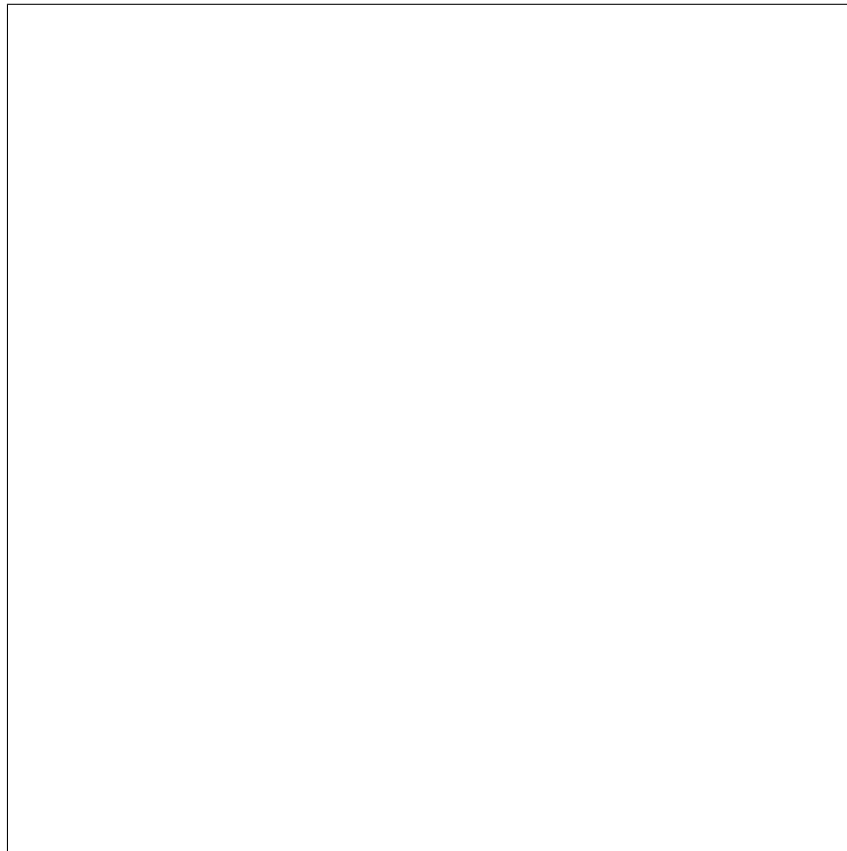


Figura 51. Bolsas de Lima.

Os gráficos apresentam longa dependência da série. Para aplicar modelos do tipo ARIMA, com critério de seleção de modelos o menor AIC, usamos a função `auto.arima()`, da biblioteca `forecast`.

```
mig=auto.arima(igbvldiario)
mis=auto.arima(isbvldiario)
min=auto.arima(incaibvldiario)
```

```
> summary(mig)
Series: igbvldiario
ARIMA(1,1,2)
```

```
Coefficients:
          ar1          ma1          ma2
```

```

      0.8677 -0.7459 -0.0770
s.e.  0.1254  0.1273  0.0316

```

```

sigma^2 estimated as 40449:  log likelihood=-22803.89
AIC=45615.78  AICc=45615.79  BIC=45640.3

```

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	2.917412	201.0906	106.8228	0.03976854	0.9446441	0.988452

ACF1

Training set -0.001169171

> summary(mis)

Series: isbvldiario

ARIMA(1,1,0)

Coefficients:

```

      ar1
      0.1325
s.e.  0.0170

```

```

sigma^2 estimated as 115552:  log likelihood=-24584.13
AIC=49172.25  AICc=49172.26  BIC=49184.51

```

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	5.170027	339.8793	184.8479	0.04268242	1.060715	0.9910698

ACF1

Training set -0.001845832

> summary(min)

Series: incaibvldiario

ARIMA(0,1,1)

Coefficients:

```

      ma1
      0.0684
s.e.  0.0261

```

```

sigma^2 estimated as 2.349:  log likelihood=-2824.32
AIC=5652.65  AICc=5652.65  BIC=5663.31

```

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	-0.01064868	1.532189	1.065254	-0.02813419	1.243135	0.9964838
	ACF1					
Training set	-0.001585991					

Após estimação, podemos validar esta modelagem pois apresentou boa performance no erro de medida MAPE, para as 3 séries da bolsa. Finalmente realizaremos previsões para 5 dias úteis da bolsa, da seguinte forma:

```
parimaig=predict(mig,5)
parimais=predict(mis,5)
parimainca=predict(min,5)

> parimaig
$pred
Time Series:
Start = 3394
End = 3398
Frequency = 1
[1] 15114.65 15109.87 15105.72 15102.12 15099.00

$se
Time Series:
Start = 3394
End = 3398
Frequency = 1
[1] 201.1202 302.2504 380.6575 448.0947 508.6768

> parimais
$pred
Time Series:
Start = 3394
End = 3398
Frequency = 1
[1] 23015.36 23016.88 23017.08 23017.11 23017.11

$se
Time Series:
Start = 3394
```

```
End = 3398
Frequency = 1
[1] 339.9294 513.5655 645.4273 754.9979 850.6170
```

```
> parimainca
$pred
Time Series:
Start = 1532
End = 1536
Frequency = 1
[1] 84.87093 84.87093 84.87093 84.87093 84.87093
```

```
$se
Time Series:
Start = 1532
End = 1536
Frequency = 1
[1] 1.532688 2.242884 2.777036 3.223868 3.615898
```

Podemos observar que as previsões para as três séries usando modelos ARIMA, no geral tendem a ser constantes, mostrando ausência de variabilidade na cotação das bolsas.

14.4 Abordagem ARFIMA

Uma questão que fico por resolver a persistência da autocorrelação que apresentam as séries de cotação das bolsas, típicos de processos de memória longa, os quais podem ser identificados por uma função semi-paramétrica `fdGPH()`, ou `fdSperio()`, do pacote `fracdiff`, da seguinte forma:

```
> fdGPH(diff(igbvldiario.ts)) # Geweke Porter Hudak
$d
[1] 0.3058209

$sd.as
[1] 0.09452347

$sd.reg
[1] 0.07539007
```

```
> fdSperio(diff(igbvldiario.ts)) # Valderio e Reisen
```



```
$d
[1] 0.230356

$sd.as
[1] 0.03603882

$sd.reg
[1] 0.04208924

> fdGPH(diff(isbvldiario.ts))
$d
[1] 0.2705833

$sd.as
[1] 0.09452347

$sd.reg
[1] 0.07415209

> fdSperio(diff(isbvldiario.ts))
$d
[1] 0.2095247

$sd.as
[1] 0.03603882

$sd.reg
[1] 0.04657447

> fdGPH(diff(incaibvldiario.ts))
$d
[1] 0.05776458

$sd.as
[1] 0.1194045

$sd.reg
[1] 0.1257685

> fdSperio(diff(incaibvldiario.ts))
```

```
$d
[1] 0.01165445
```

```
$sd.as
[1] 0.04735413
```

```
$sd.reg
[1] 0.04776419
```

Uma abordagem para estimar modelos ARFIMA, foi proposta seguinte forma:

```
arfimaig=arfima(igbvldiario, drange = c(0, 0.5))
arfimais=arfima(isbvldiario, drange = c(0, 0.5))
arfimain=arfima(incaibvldiario, drange = c(0, 0.5))
summary(arfimaig)
summary(arfimais)
summary(arfimain)
marfig=predict(arfimaig,5)
marfis=predict(arfimais,5)
marfin=predict(arfimain,5)
```

```
> marfig
      Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
2013.627      15103.07 14845.39 15360.75 14708.99 15497.16
2013.631      15088.23 14701.01 15475.45 14496.03 15680.43
2013.635      15077.28 14590.94 15563.63 14333.48 15821.08
2013.639      15063.89 14490.49 15637.30 14186.94 15940.85
2013.643      15056.75 14404.71 15708.80 14059.53 16053.97
> marfis
      Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
2013.627      23006.92 22571.52 23442.33 22341.03 23672.81
2013.631      22997.65 22340.89 23654.41 21993.22 24002.08
2013.635      22988.31 22162.89 23813.73 21725.94 24250.69
2013.639      22979.05 22011.28 23946.83 21498.97 24459.14
2013.643      22969.90 21876.40 24063.39 21297.54 24642.25
> marfin
      Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
2013.627       84.88803 82.92613 86.84993 81.88756 87.88850
2013.631       84.91187 82.04500 87.77875 80.52736 89.29638
```

2013.635	84.93594	81.37942	88.49247	79.49671	90.37518
2013.639	84.96040	80.82457	89.09624	78.63518	91.28562
2013.643	84.98525	80.34073	89.62976	77.88207	92.08842

Os erros de previsão usado é o MAPE, com os seguintes valores:

```
mapeig=mean(abs((igbvldiario-fitted(arfimaig))/igbvldiario))*100
mapeig
[1] 0.9680164
mapeis=mean(abs((isbvldiario-fitted(arfimaais))/isbvldiario))*100
mapeis
[1] 1.084695
mapein=mean(abs((incaibvldiario-fitted(arfimain))/incaibvldiario))*100
mapein
[1] 1.245779
```

15 Comparação de modelos

As cotações das bolsas de Lima foram extraídas até o dia 31 de julho de 2013, foram colhidas cotações da bolsa para 5 dias úteis posteriores á serie analisada, com o interesse de avaliar o MAPE externo dos modelos, para tanto foi construída uma tabela de comparação reais e previstos da bolsa. Os dias úteis de agosto foram 01,02,05,06,e 07, conforme abaixo:

15.1 Para IGBVL

Tabela 1. MAPE externo, para a cotação da bolsa de Lima IGBVL.

Data	IGBVL	H-W	R Dinâmica	ARIMA	ARFIMA
01/08/2013	15308,76	15105,78	15703,00	15114,00	15103,00
02/08/2013	15298,04	15093,40	15654,00	15110,00	15088,00
05/08/2013	15267,87	15044,83	15982,00	15105,00	15077,00
06/08/2013	15045,40	15011,66	16073,00	15102,00	15063,00
07/08/2013	15082,58	14988,13	16148,00	15099,00	15056,00
MAPE		0,99	4,69	0,81	0,85

15.2 Para ISBVL

Tabela 2. MAPE externo, para a cotação da bolsa de Lima ISBVL.

Data	ISBVL	H-W	R Dinâmica	ARIMA	ARFIMA
01/08/2013	23282,66	23042,65	24032,00	23015,00	23007,00
02/08/2013	23202,29	23054,76	23964,00	23016,00	22998,00
05/08/2013	23119,94	23037,48	24491,00	23017,00	22988,00
06/08/2013	22763,89	23032,96	24629,00	23017,00	22979,00
07/08/2013	22820,16	23047,16	24756,00	23017,00	22969,00
MAPE		0,84	5,82	0,87	0,85

15.3 Para INCA

Tabela 3. MAPE externo, para a cotação da bolsa de Lima INCA.

Data	INCA	H-W	R Dinâmica	ARIMA	ARFIMA
01/08/2013	85,87	85,21	86,84	84,87	84,89
02/08/2013	85,52	86,64	86,46	84,87	84,91
05/08/2013	85,36	85,77	88,40	84,87	84,94
06/08/2013	84,62	85,99	88,81	84,87	84,96
07/08/2013	84,66	86,17	89,12	84,87	84,99
MAPE		1,19	3,20	0,61	0,63

15.4 Analisando os Retornos

Os retornos representam os lucros (quando positivo) e as perdas (quando negativo) de uma carteira de ativos, para construir estes indicadores no R é bastante simples. A construção e saída gráfica dos retornos da Bolsa de Lima, é dada por:

```

rigbvl=diff(igbvldiario.ts)/igbvldiario.ts[-length(igbvldiario.ts)]
risbvl=diff(isbvldiario.ts)/isbvldiario.ts[-length(isbvldiario.ts)]
rinca=diff(incaibvldiario.ts)/incaibvldiario.ts[-length(incaibvldiario.ts)]

plot.ts(risbvl, xlab="Años", ylab="Retorno",
main="Retornos en la Bolsa de Lima \n 1/1/2000-31/07/2013")
lines(rigbvl,col=2,lty=2)
lines(rinca,col=8)
legend("bottomleft",legend=c("INCA","ISBVL","IGBVL"),
col=c(8,1,2),lty=c(1,1,2))

```

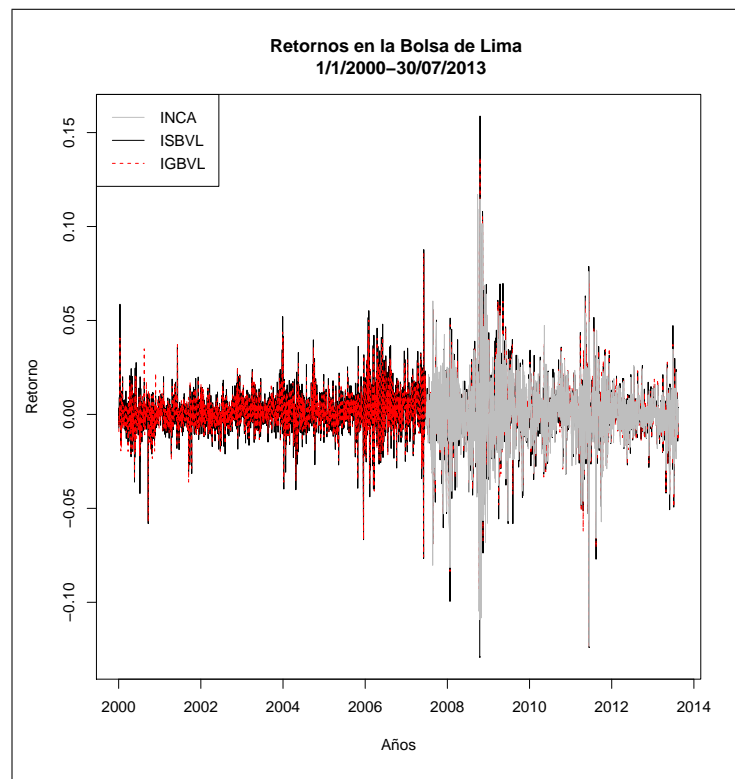


Figura 52. Volatilidade da Bolsa de Lima, 3 indicadores.

15.4.1 Interpretação dos retornos

```
mean(rigbvl)
0.0007318932
sd(rigbvl)
[1] 0.0147403
> mean(rigbvl[3237:3392])
[1] -0.001773055
> sd(rigbvl[3237:3392])
[1] 0.01193148
```

Durante las 13 anos, en media a cada 100 000 soles invertidos hay un retorno positivo (lucro) de 73,18 soles por dia, y el riesgo del investimento es de 1 470 soles por cada 100 000 soles. Si consideramos 2013, hay en media una perdida de 177 soles por cada 100 000 invertidos.

```
mean(risbvl)
```

```
[1] 0.0007562155
sd(risbvl)
[1] 0.01641675
> mean(risbvl[3237:3392])
[1] -0.001654203
> sd(risbvl[3237:3392])
[1] 0.01307777
```

Durante las 13 anos, en media a cada 100 000 soles invertidos hay un retorno positivo (lucro) de 75,6 soles por dia, y el riesgo de investimento es de 1 640 soles por cada 100 000 soles. Si consideramos 2013, hay en media una perdida de 165 soles por cada 100 000 invertidos.

```
mean(rinca)
5.654552e-05
sd(rinca)
0.01888527
> mean(rinca[1375:1530])
[1] -0.001421642
> sd(rinca[1375:1530])
[1] 0.01077818
```

Durante los 6 anos, en media a cada 100 000 soles invertidos en la bolsa, hay un retorno positivo (lucro) de 5,65 soles por dia, y el riesgo de investimento es de 1 880 soles por cada 100 000 soles. Si consideramos 2013, hay en media una perdida de 142 soles por cada 100 000 invertidos.

15.5 Modelos Arima para os retornos

No geral (nem sempre) um retorno financeiro é estacionário, assim uma estimação por modelos ARIMA, foi aplicada com os seguintes resultados:

15.5.1 IGBVL

```
maig=auto.arima(rigbvl) #Modelo

summary(maig) #Estatísticas do modelo

Series: rigbvl
ARIMA(5,0,4) with zero mean

Coefficients:
```

15.5 Modelos Arima para os retornos *5* COMPARAÇÃO DE MODELOS

```

          ar1      ar2      ar3      ar4      ar5      ma1      ma2      ma3
-0.0302 -0.0086  0.2446  0.8132 -0.1091  0.2003  0.0108 -0.2266
s.e.    0.0127      NaN  0.0300  0.0723   0.0240  0.0152      NaN      NaN
          ma4
-0.8111
s.e.    0.0683

```

```

sigma^2 estimated as 0.0002081:  log likelihood=9564.43
AIC=-19108.85   AICc=-19108.79   BIC=-19047.56

```

Training set error measures:

```

              ME      RMSE      MAE  MPE MAPE      MASE
Training set 0.0003765849 0.01442674 0.00942125 -Inf  Inf 0.658092
              ACF1
Training set -0.0009916489

```

```
pig=predict(maig,5) #Previsões 5 dias á frente
```

```

pig              # Retornos Estimados
$pred
Time Series:
Start = c(2013, 157)
End = c(2013, 161)
Frequency = 249
[1] -1.212592e-04 -1.400969e-04 -1.478581e-03 -7.241694e-04 -6.719469e-05

```

```

$se              # Risco do mercado
Time Series:
Start = c(2013, 157)
End = c(2013, 161)
Frequency = 249
[1] 0.01442674 0.01463399 0.01463405 0.01463601 0.01464928

```

15.5.2 ISBVL

```
mais=auto.arima(risbvl)
```

```
summary(mais)
```

15.5 Modelos Arima para os retornos *5* COMPARAÇÃO DE MODELOS

```
Series: risbvl  
ARIMA(2,0,2) with zero mean
```

```
Coefficients:
```

```
          ar1      ar2      ma1      ma2  
          0.7697  0.2161 -0.6114 -0.3589  
s.e.    0.1056  0.1031  0.1009  0.0963
```

```
sigma^2 estimated as 0.0002618:  log likelihood=9175.35  
AIC=-18340.69  AICc=-18340.68  BIC=-18310.05
```

```
Training set error measures:
```

```
              ME      RMSE      MAE      MPE      MAPE      MASE  
Training set 0.0003473046 0.01618042 0.01057556 98.2501 153.0638 0.6617089  
              ACF1  
Training set -4.670323e-05
```

```
pis=predict(mais,5)
```

```
pis
```

```
$pred
```

```
Time Series:
```

```
Start = c(2013, 157)
```

```
End = c(2013, 161)
```

```
Frequency = 249
```

```
[1] 0.0003003678 -0.0012119085 -0.0008679191 -0.0009299816 -0.0009034082
```

```
$se
```

```
Time Series:
```

```
Start = c(2013, 157)
```

```
End = c(2013, 161)
```

```
Frequency = 249
```

```
[1] 0.01618042 0.01638202 0.01638551 0.01638814 0.01638885
```

15.5.3 INCA

```
maic=auto.arima(rinca)
```

```
summary(maic)
```



```

Series: rinca
ARIMA(3,0,3) with zero mean

Coefficients:
          ar1      ar2      ar3      ma1      ma2      ma3
      0.4682 -0.8231  0.5827 -0.3794  0.7588 -0.4888
s.e.  0.1755  0.0550  0.1668  0.1837  0.0673  0.1711

sigma^2 estimated as 0.0003441:  log likelihood=3929.55
AIC=-7845.09  AICc=-7845.02  BIC=-7807.76

Training set error measures:
              ME      RMSE      MAE  MPE  MAPE      MASE
Training set 4.931522e-05 0.01854885 0.01241549 NaN  Inf 0.6511565
              ACF1
Training set -0.0008223587

pic=predict(maic,5)

pic
$pred
Time Series:
Start = c(2013, 157)
End = c(2013, 161)
Frequency = 249
[1] -0.0006483197  0.0015514397  0.0004212950 -0.0014574619 -0.0001251222

$se
Time Series:
Start = c(2013, 157)
End = c(2013, 161)
Frequency = 249
[1] 0.01854885 0.01862191 0.01862666 0.01862762 0.01867976

```

15.6 Decomposição dos Retornos

Muitas propriedades de uma série temporal podem ser capturadas, assumindo que a decomposição desta série apresenta componentes de tendência (trend), componente sazonal (seasonal), e uma componente aleatória ou ruído (remainder). Para visualizar a cotação da bolsa de Lima graficamente,

usamos os comandos:

```
#### IGBVL #####
drg=decompose(rigbvl.ts)
sltrg=stl(rigbvl.ts,s.window=13)
plot(sltrg,main="IGBVL")

#### ISBVL #####
drs=decompose(risbvl.ts)
sltrs=stl(risbvl.ts,s.window=13)
plot(sltrs,main="ISBVL")

#### INCA #####
drc=decompose(rinca.ts)
sltrc=stl(rinca.ts,s.window=13)
plot(sltrc,main="INCA")
```

Com a seguinte saída gráfica:

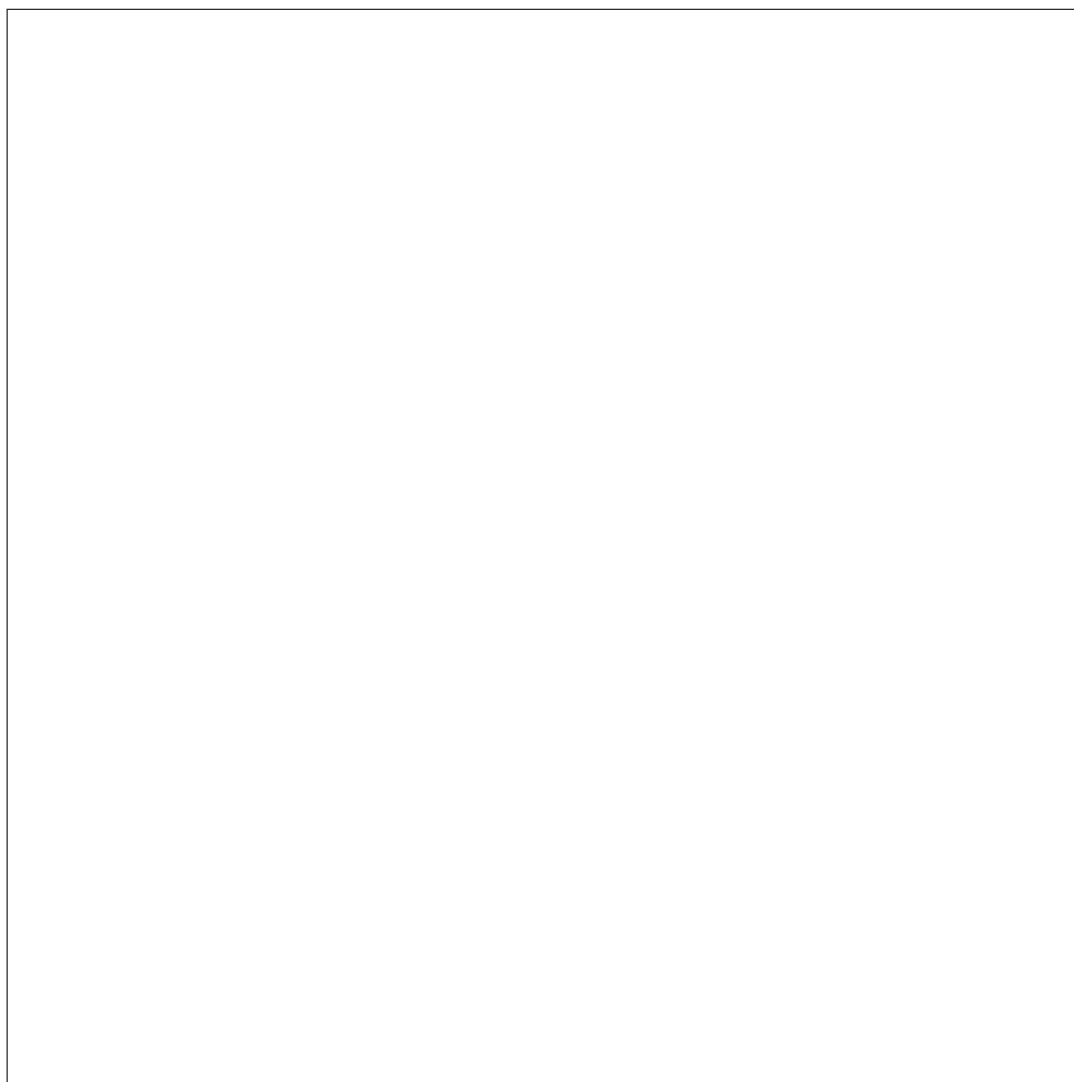


Figura 53. Decomposição da Bolsa de Lima.

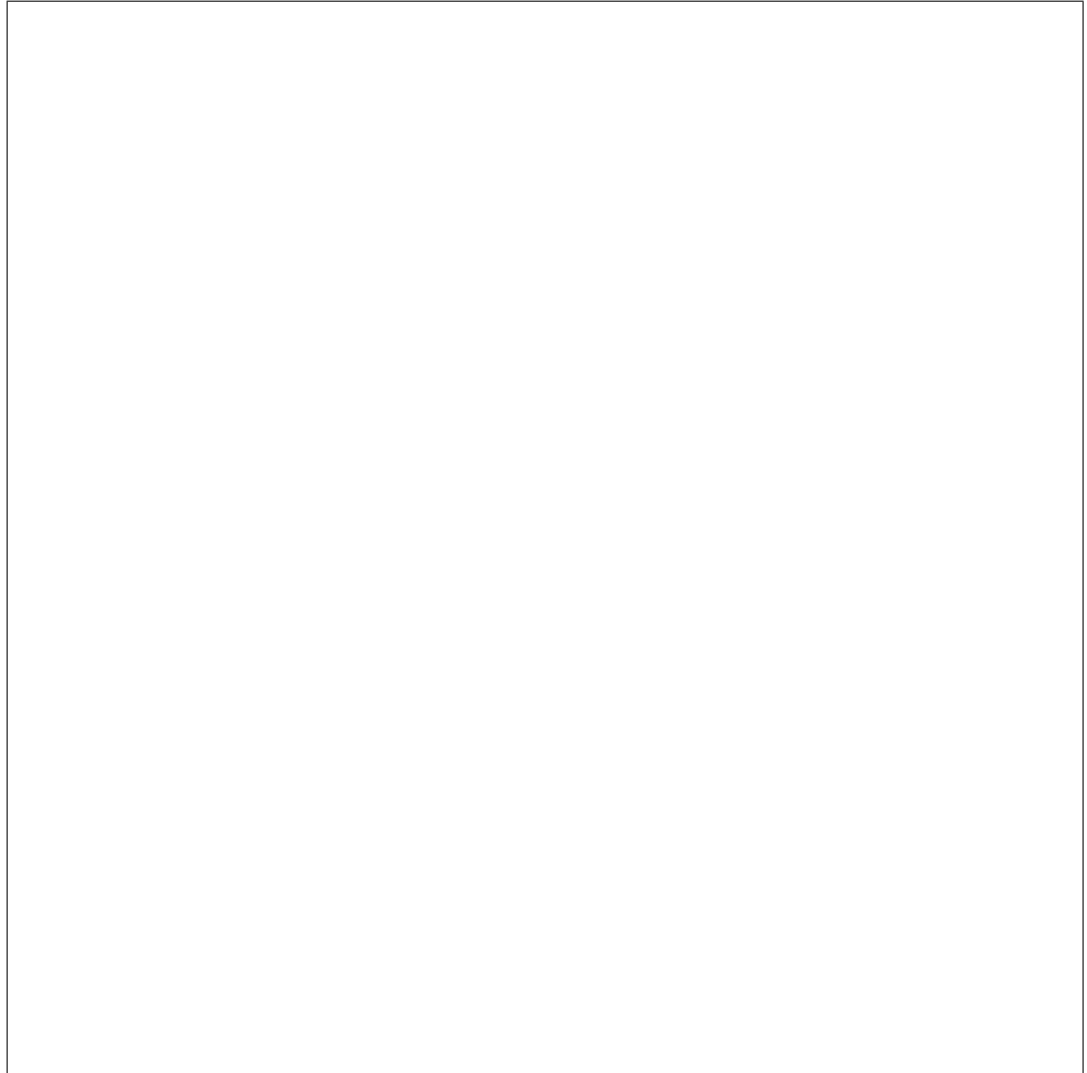


Figura 54. Decomposição da Bolsa de Lima.

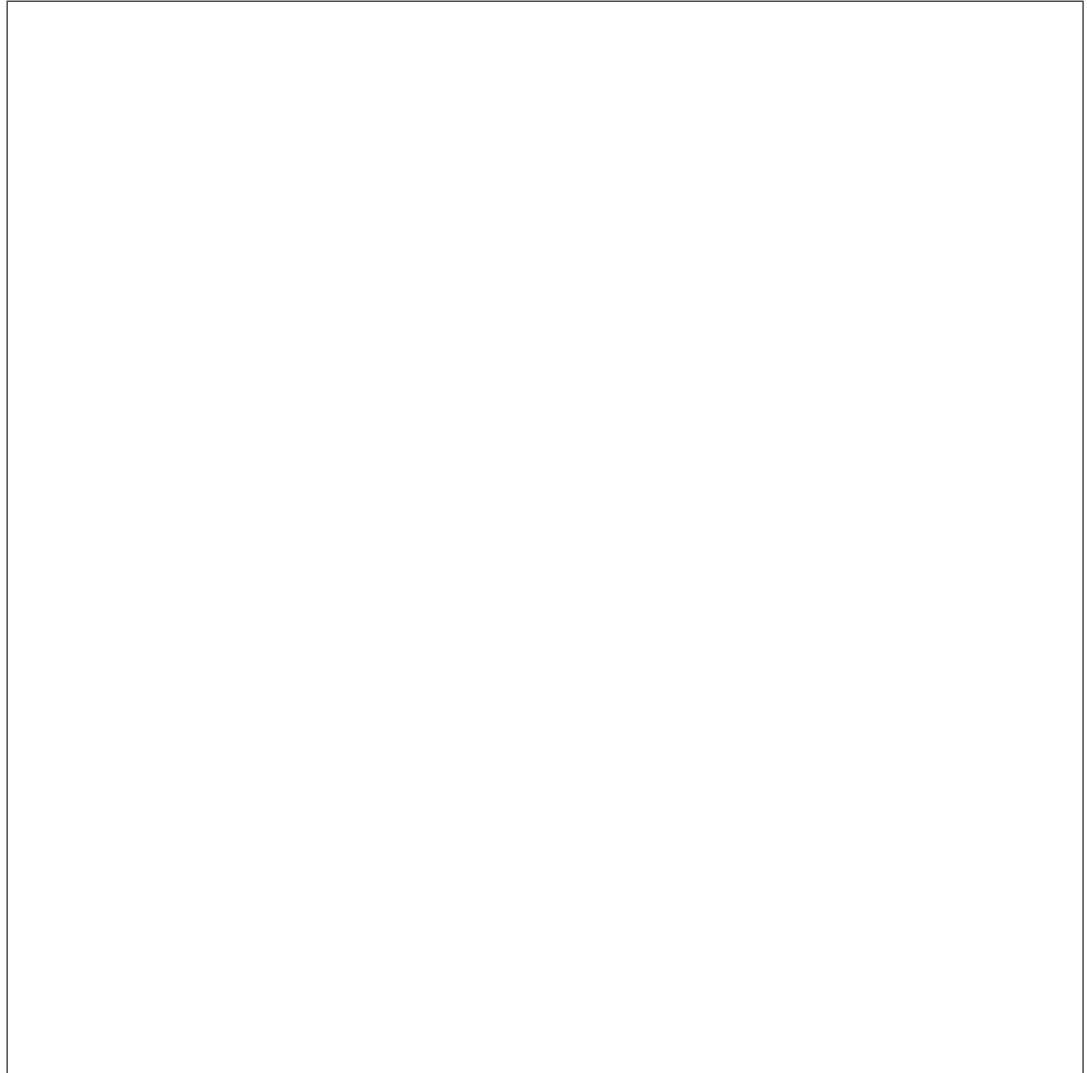


Figura 55. Decomposição da Bolsa de Lima.

15.7 Modelos ARCH e GARCH

Contudo, modelos do tipo ARIMA, não levam em consideração a volatilidade do mercado de ações (bursatil), ou as influencias de mercados externos, isto pode ser verificado na decomposição das séries de retornos. Para construir modelos para os retornos da carteira de ativos da bolsa de Lima é adequado

usar modelagem não linear na estrutura ou variância heterocedástica, conhecidos como modelos para séries com alta volatilidade, em particular usamos os modelos do tipo ARCH e GARCH, que são ideais para séries não lineares estruturais.

15.7.1 ARCH para IGBVL

```
archigbvl2=garchFit(~arma(5,4)+garch(3,0),rigbvl.ts)
summary(archigbvl2)
resumo2=archigbvl2@fit
resumo2
plot(archigbvl2)
ajustearchigbvl2z=archigbvl2@fit$series$z # estimación
ajustearchigbvl2x=archigbvl2@fit$series$x # valor real
ajustearchigbvl2h=archigbvl2@fit$series$h # nível
ajustearchigbvl2z.ts=ts(ajustearchigbvl2z,
end=c(2013,156),frequency=249)
ajustearchigbvl2h.ts=ts(ajustearchigbvl2h,
end=c(2013,156),frequency=249)

#Previsão

predict(ajustearchigbvl2z.ts,5) #previsoes de 5 dias a frente
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2013.627	-0.004791800	-0.02020415	0.010620554	-0.028362961	0.01877936
2013.631	0.016868704	0.00145635	0.032281057	-0.006702457	0.04043986
2013.635	0.000145070	-0.01526728	0.015557424	-0.023426091	0.02371623
2013.639	0.003519564	-0.01189279	0.018931918	-0.020051597	0.02709073
2013.643	-0.008098307	-0.02351066	0.007314047	-0.031669468	0.01547285

#Previsão do nível dos retornos

```
predict(ajustearchigbvl2h.ts,5) #previsoes de 5 dias a frente
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2013.627	2.046063e-04	-5.503944e-05	0.0004642520	-0.0001924876	0.0006017002
2013.631	7.080826e-05	-2.963679e-04	0.0004379844	-0.0004907392	0.0006323557
2013.635	-1.361412e-04	-5.858309e-04	0.0003135484	-0.0008238822	0.0005515997
2013.639	-2.199195e-04	-7.391720e-04	0.0002993330	-0.0010140477	0.0005742087

```
2013.643 -3.308021e-04 -9.113411e-04 0.0002497370 -0.0012186599 0.0005570558
```

Uma medida mais apropriada para comparar erros de previsão para os retornos é o MAE (Mean Absolute Error), ou erro absoluto medio, que para o modelo proposto é dado por:

```
MAEig=mean(abs((rigbvl-ajustearchigbvlz.ts)))
MAEig
[1] 0.002279143
```

Comparando o erro absoluto médio externo

```
retornoex=c(-0.000839414,-0.000820226,-0.003228352,-0.002209616,-0.001569909)
retigpredarch=c(-0.00479,0.01687,0.0001451,0.0035,-0.00809)
```

```
MAEexarch=mean(abs(retornoex-retigpredarch))
MAEexarch
[1] 0.007448794
```

15.7.2 GARCH para IGBVL

```
garchrigbvl=garchFit(~arma(0,1)+garch(2,1),rigbvl.ts)
ajustegarchrigbvlh=garchrigbvl@fit$series$h
ajustegarchrigbvlz=garchrigbvl@fit$series$z
ajustegarchigbvlz.ts=ts(ajustegarchrigbvlz,
end=c(2013,156),frequency=249)
ajustegarchigbvlh.ts=ts(ajustegarchrigbvlz,
end=c(2013,156),frequency=249)
```

```
MAEigarch=mean(abs(rigbvl-ajustegarchigbvlz.ts))
MAEigarch
[1] 0.002056131
```

```
predict(ajustegarchigbvlz.ts,5)
predict(ajustegarchigbvlh.ts,5)
```

```
> predict(ajustegarchigbvlz.ts,5)
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
2013.627 -0.006410675 -0.0218489248 0.009027574 -0.030021440 0.01720009
```

```

2013.631    0.014605647 -0.0008332457 0.030044540 -0.009006102 0.03821740
2013.635   -0.001682088 -0.0171216249 0.013757448 -0.025294822 0.02193065
2013.639    0.001535126 -0.0139050545 0.016975306 -0.022078592 0.02514884
2013.643   -0.008810521 -0.0242513449 0.006630303 -0.032425223 0.01480418
> predict(ajustegarchigbvlh.ts,5)
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
2013.627   -0.006410675 -0.0218489248 0.009027574 -0.030021440 0.01720009
2013.631    0.014605647 -0.0008332457 0.030044540 -0.009006102 0.03821740
2013.635   -0.001682088 -0.0171216249 0.013757448 -0.025294822 0.02193065
2013.639    0.001535126 -0.0139050545 0.016975306 -0.022078592 0.02514884
2013.643   -0.008810521 -0.0242513449 0.006630303 -0.032425223 0.01480418

```

Comparando o erro absoluto médio externo

```

retornoex=c(-0.000839414,-0.000820226,-0.003228352,-0.002209616,-0.001569909)
retigpredgarch=c(-0.00641,0.0146,-0.00168,0.001535,-0.00881)

```

```

MAEexgarch=mean(abs(retornoex-retigpredgarch))
MAEexgarch

```

```
[1] 0.006704774
```

Saída gráfica da estimação GARCH dos retornos

```
plot(garchrigbvl) #Todas as saídas gráficas do GARH.
```

Gráfico dos resíduos:

```
##### Construção gráfica dos resíduos#####
```

```

jarque.bera.test(residuals(garchrigbvl))
plot(density(residuals(garchrigbvl)))
curve(dnorm(x,mean=mean(residuals(garchrigbvl)),
sd=sd(residuals(garchrigbvl))),add=T,col=2)

```

Construção gráfica da série de retornos do IGBVL real e estimado por um GARCH.

```

plot.ts(rigbvl)
ajustegarchrigbvlz=ts(ajustegarchrigbvlz,start=c(2000,1),
frequency=249)
lines(ajustegarchrigbvlz,col=2)

```



```
legend("bottomleft", legend=c("IGBVL real","IGBVL estimado"),
lty=1:1,col=c(1,2))
```



Figura 56. Estimação GARCH do IGBVL.

15.7.3 ARCH para ISBVL

```
archisbvl2=garchFit(~arma(2,2)+garch(3,0),risbvl)
#summary(archisbvl2)
#resumo3=archisbvl2@fit
#resumo3
#plot(archisbvl)
ajustearchisbvl2z=archisbvl2@fit$series$z # estimación
ajustearchisbvl2x=archisbvl2@fit$series$x #valor real
ajustearchisbvl2h=archisbvl2@fit$series$h #nível

ajustearchisbvl2z.ts=ts(ajustearchisbvl2z,end=c(2013,156),frequency=249)
ajustearchisbvl2h.ts=ts(ajustearchisbvl2h,end=c(2013,156),frequency=249)
```

```
ajustearchisbvl2x.ts=ts(ajustearchisbvl2x,end=c(2013,156),frequency=249)
```

```
predict(ajustearchisbvl2z.ts,5)
predict(ajustearchisbvl2h.ts,5)
predict(ajustearchisbvl2x.ts,5) #previsoes de 5 dias a frente
```

```
#Previsão
```

```
> predict(ajustearchisbvl2z.ts,5)
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
2013.627 -0.0047905258 -0.021932361 0.012351309 -0.031006700 0.02142565
2013.631  0.0190716003  0.001929765 0.036213435 -0.007144574 0.04528777
2013.635  0.0006986929 -0.016443142 0.017840528 -0.025517481 0.02691487
2013.639  0.0042388932 -0.012902942 0.021380728 -0.021977281 0.03045507
2013.643 -0.0091567491 -0.026298584 0.007985086 -0.035372924 0.01705943
```

```
#Previsão do nível dos retornos
```

```
> predict(ajustearchisbvl2h.ts,5)
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
2013.627  2.296103e-04 -0.0000852690 0.0005444896 -0.0002519560 0.0007111766
2013.631  1.561297e-04 -0.0002891545 0.0006014140 -0.0005248738 0.0008371333
2013.635 -9.374813e-05 -0.0006390987 0.0004516024 -0.0009277898 0.0007402936
2013.639 -2.163224e-04 -0.0008460337 0.0004133889 -0.0011793827 0.0007467380
2013.643 -3.460672e-04 -0.0010501023 0.0003579680 -0.0014227960 0.0007306616
```

Uma medida mais apropriada para comparar erros de previsão para os retornos é o MAE (Mean Absolute Error), ou erro absoluto medio, que para o modelo proposto é dado por:

```
MAEis=mean(abs((risbvl-ajustearchisbvl2z.ts)))
MAEis
```

```
[1] 0.002225938
```

Comparando o erro absoluto médio externo

```
retornoex=c(0.001684268,0.000525271,-0.000750082,-0.000196241,0.000616128)
retispredarch=c(-0.00479, 0.019, 0.00069,0.0042,-0.0091)
```

```
MAEexarchis=mean(abs(retornoex-retispredarch))
MAEexarchis
[1] 0.00810029
```

15.7.4 GARCH para ISBVL

```
garchrisbvl2=garchFit(~arma(0,0)+garch(1,2),risbvl)
ajustegarch2risbvlh=garchrisbvl2@fit$series$h
ajustegarch2risbvlz=garchrisbvl2@fit$series$z
ajustegarchi2risbvlz.ts=ts(ajustegarch2risbvlz,end=c(2013,156),frequency=249)
ajustegarchi2risbvlh.ts=ts(ajustegarch2risbvlh,end=c(2013,156),frequency=249)
```

```
predict(ajustegarchi2risbvlz.ts,5)
predict(ajustegarchi2risbvlh.ts,5)
```

```
> predict(ajustegarch2risbvlh,5)
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
2013.622  8.287338e-05 -0.0001122361 0.0002779829 -0.0002155209 0.0003812676
2013.627 -6.161649e-06 -0.0002775339 0.0002652106 -0.0004211897 0.0004088664
2013.631  2.905501e-05 -0.0003014266 0.0003595366 -0.0004763730 0.0005344830
2013.635 -6.601885e-05 -0.0004465360 0.0003144983 -0.0006479696 0.0005159319
2013.639 -1.445904e-04 -0.0005692885 0.0002801077 -0.0007941101 0.0005049293
```

```
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
2013.622 -0.005777522 -0.023061607 0.011506562 -0.03221125 0.02065620
2013.627  0.015807360 -0.001477905 0.033092625 -0.01062817 0.04224289
2013.631  0.001854954 -0.015431490 0.019141399 -0.02458238 0.02829229
2013.635  0.001883691 -0.015403933 0.019171316 -0.02455545 0.02832283
2013.639 -0.010393277 -0.027682081 0.006895527 -0.03683422 0.01604767
```

Uma medida mais apropriada para comparar erros de previsão para os retornos é o MAE (Mean Absolute Error), ou erro absoluto medio, que para o modelo proposto é dado por:

```
MAEis=mean(abs(risbvl-ajustegarch2risbvlz))
```

```
MAEis
```

```
[1] 0.0009264017
```

Comparando o erro absoluto médio externo

```
retornoex=c(0.001684268,0.000525271,-0.000750082,-0.000196241,0.000616128)
```

```
retispredgarch=c(-0.005777, 0.0158, 0.00185, 0.00188, -0.01039)
```

```
MAEexgarch=mean(abs(retornoex-retispredgarch))
```

```
MAEexgarch
```

```
[1] 0.00768369
```

Saída gráfica da estimação GARCH dos retornos

```
plot(garchrisbvl2) #Todas as saídas gráficas do GARH.
```

Gráfico dos resíduos:

```
##### Construção gráfica dos resíduos#####
```

```
jarque.bera.test(residuals(garchrisbvl2))
```

```
plot(density(residuals(garchrisbvl2)))
```

```
curve(dnorm(x,mean=mean(residuals(garchrisbvl2)),
sd=sd(residuals(garchrisbvl2))),add=T,col=2)
```

Construção gráfica da série de retornos do IGBVL real e estimado por um GARCH.

```
plot.ts(risbvl)
```

```
lines(ajustegarchi2risbvlz.ts,col=2)
```

```
legend("bottomleft", legend=c("ISBVL real","ISBVL estimado"),
```

```
lty=1:1,col=c(1,2))
```

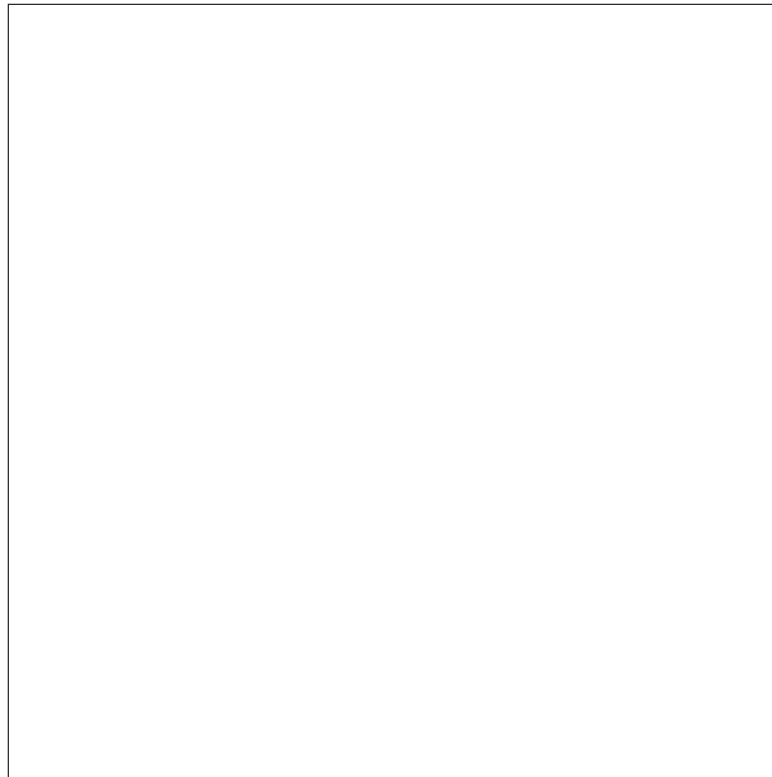


Figura 57. Estimação GARCH do ISBVL.

15.7.5 ARCH para INCA

```
archinca2=garchFit(~arma(3,3)+garch(2,0),rinca)
summary(archinca2)
resumo1=archinca2@fit
ajustearchinca2z=archinca2@fit$series$z # estimación
ajustearchinca2x=archinca2@fit$series$x #valor real
ajustearchinca2h=archinca2@fit$series$h #nível
ajustearchinca2z.ts=ts(ajustearchinca2z,end=c(2013,156),frequency=249)
ajustearchinca2h.ts=ts(ajustearchinca2h,end=c(2013,156),frequency=249)

#Previsão

> predict(ajustearchinca2z.ts,5) #previsoes de 5 dias a frente
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
```

```

2013.627 -0.0080052715 -0.027981666 0.011971123 -0.03855653 0.02254599
2013.631 0.0221930490 0.002216654 0.042169444 -0.00835821 0.05274431
2013.635 -0.0002673193 -0.020243714 0.019709075 -0.03081858 0.03028394
2013.639 0.0042366907 -0.015739704 0.024213085 -0.02631457 0.03478795
2013.643 -0.0111857978 -0.031162193 0.008790597 -0.04173706 0.01936546

```

#Previsão do nível dos retornos

```

> predict(ajustearchinca2h.ts,5) #previsoes do nível 5 dias a frente
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
2013.627 0.0008359591 0.0002171897 0.0014547286 -0.0001103671 0.001782285
2013.631 0.0002661811 -0.0004053549 0.0009377171 -0.0007608446 0.001293207
2013.635 0.0001969394 -0.0005235088 0.0009173875 -0.0009048910 0.001298770
2013.639 0.0003891983 -0.0003770461 0.0011554427 -0.0007826714 0.001561068
2013.643 -0.0001428075 -0.0009522612 0.0006666463 -0.0013807602 0.001095145

```

Uma medida mais adequada para comparar erros de previsão para os retornos é o MAE (Mean Absolute Error), ou erro absoluto medio, que para o modelo proposto é dado por:

```

MAEin=mean(abs((rinca-ajustearchinca2z.ts)))
MAEin

```

```
[1] 0.002508965
```

Comparando o erro absoluto médio externo

```

retornoex=c(0.004224856,0.016505078,-0.010143407,0.002558437,0.002088894)
reticpredarch=c(0.00083, 0.000266, 0.000196, 0.000389, -0.000143)

```

```

MAEexarchin=mean(abs(retornoex-reticpredarch))
MAEexarchin
[1] 0.006874934

```

15.7.6 GARCH para INCA

```
garchrinca=garchFit(~arma(3,3)+garch(2,1),rinca)
```

```

ajustegarchrincah=garchrinca@fit$series$h
ajustegarchrincax=garchrinca@fit$series$x
ajustegarchrincaz=garchrinca@fit$series$z
ajustegarchrincah.ts=ts(ajustegarchrincah,end=c(2013,156),frequency=249)
ajustegarchrincaz.ts=ts(ajustegarchrincaz,end=c(2013,156),frequency=249)
ajustegarchrincax.ts=ts(ajustegarchrincax,end=c(2013,156),frequency=249)

predict(ajustegarchrincah.ts,5)
predict(ajustegarchrincaz.ts,5)
predict(ajustegarchrincax.ts,5)

> predict(ajustegarchrincah.ts,5)
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
2013.627  2.331121e-04  6.238188e-05  0.0004038423 -2.799723e-05  0.0004942214
2013.631  1.827390e-04 -5.869793e-05  0.0004241758 -1.865069e-04  0.0005519848
2013.635  2.099594e-04 -8.573429e-05  0.0005056530 -2.422651e-04  0.0006621838
2013.639  1.654861e-04 -1.759487e-04  0.0005069209 -3.566933e-04  0.0006876656
2013.643  6.244548e-05 -3.192883e-04  0.0004441793 -5.213660e-04  0.0006462569

> predict(ajustegarchrincaz.ts,5)
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
2013.627 -0.0059157113 -0.0257559330  0.013924510 -0.036258712  0.02442729
2013.631  0.0208314860  0.0009912641  0.040671708 -0.009511515  0.05117449
2013.635  0.0006617769 -0.0191784450  0.020501999 -0.029681224  0.03100478
2013.639  0.0030253363 -0.0168148857  0.022865558 -0.027317665  0.03336834
2013.643 -0.0110966649 -0.0309368871  0.008743557 -0.041439666  0.01924634

```

Uma medida mais apropriada para comparar erros de previsão para os retornos é o MAE (Mean Absolute Error), ou erro absoluto medio, que para o modelo proposto é dado por:

```

MAEic=mean(abs(rinca-ajustegarchrincaz))
MAEic
[1] 0.0018666

```

Comparando o erro absoluto médio externo

```
retornoex=c(0.004224856,0.016505078,-0.010143407,0.002558437,0.002088894)
reticpredgarch=c(-0.00559,0.0208,0.00066,0.003025,-0.01109)

MAEexgarch=mean(abs(retornoex-reticpredgarch))
MAEexgarch
[1] 0.007711728
```

Saída gráfica da estimação GARCH dos retornos

```
plot(garchrinca) #Todas as saídas gráficas do GARH.
```

Gráfico dos resíduos:

```
##### Construção gráfica dos resíduos#####
```

```
jarque.bera.test(residuals(garchrinca))
plot(density(residuals(garchrinca)))
curve(dnorm(x,mean=mean(residuals(garchrinca)),
sd=sd(residuals(garchrinca))),add=T,col=2)
```

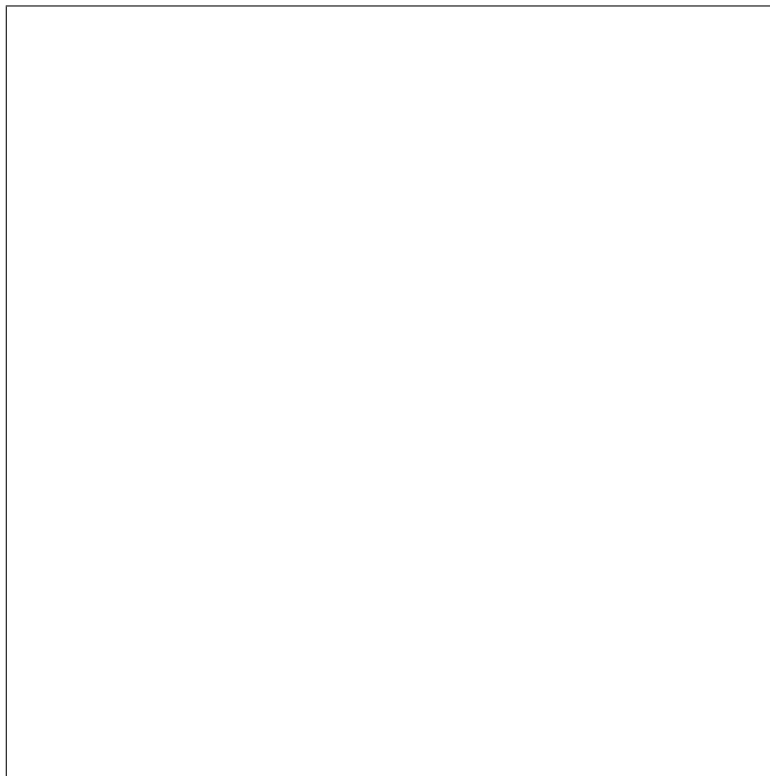


Figura 57. Resíduos do GARCH para INCA.

Para construção gráfica da série de retornos do IGBVL real e estimado por um GARCH.

```
plot.ts(rinca)
lines(ajustegarchrincaz.ts,col=2)
legend("toprigh", legend=c("ISBVL real","ISBVL estimado"),
lty=1:1,col=c(1,2))
```

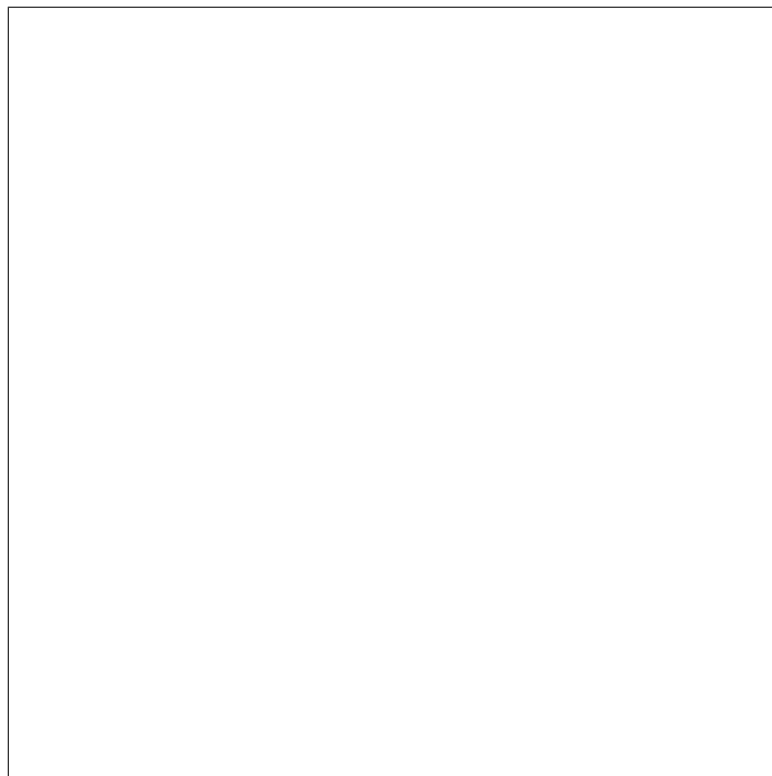


Figura 58. Estimação GARCH do INCA.

Referências

- [1] Ballie, RT., Chung, CF. e Tieslau, MA.(1996):*Analising inflation by the fractionally integrated ARFIMA-GARCH model*. Journal of applied econometrics, 11, 23 – 40.

-
- [2] Beran, J. (1994): *Statistics for long memory process*. New York: Chapman and Hall.
- [3] Borde, Arvind (1992). *T_EX by example. A beginner's guide*. Ed. Academic Press Professional. United States of America.
- [4] Carneiro, Orlando (1997). *Ecomometria Básica. Teoria e Aplicações*. Editora Atlas. Segunda Edição. São Paulo.
- [5] Conover, W. J. (1999) *Practical nonparametric statistics*. 3.ed. New York.
- [6] Cribari Neto, F.; Cabral de Araújo Gois Matheus (2002). *Uma Análise de Monte Carlo do Desempenho de Estimadores de Matrizes de Covariância sob Heterocedasticidade de Forma Desconhecida*. RBE, Rio de Janeiro 56(2), p.309-334.
- [7] Cribari Neto, F (2002). *C for Econometricians*. Computational Economics 14. p.135-149.
- [8] Reisen, V., Cribari-Neto, F., Jensen, Mark. *Long Memory Inflationary Dynamics: The Case of Brazil*. Studies in Nonlinear Dynamics & Econometrics. Vol. 7, Issue 3.(2003)
- [9] Ehlers,Ricardo.(2007): Análise de séries Temporais. Universidade Federal do Paraná.
- [10] Ferreira , M (2006). *Análise da sensibilidade dos testes de normalidade de Jarque-Bera e Lilliefors em modelos de regressão Linear*. Rev. Mat. Estat, v.24, n.4, p.89-98. São Paulo.
- [11] Frery, Alejandro; Cribari-Neto, Francisco.(2011).*Elementos de Estatística Computacional Usando Plataformas de Software Livre/Gratuito*. Publicações Matemáticas. Editora IMPA.
- [12] Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4, 221 – 238.
- [13] Goldreich, O. and Micali, S. (1986). *How to construct random functions*. Journal of the Association for Computing Machinery, 33(4), p.792–807.
- [14] Granger, Clive; Newbold, Paul. *Spurious regressions in econometrics*. Journal of Econometrics, Nottingham, v.2, p. 111-120, Jul. 1974.

-
- [15] Gujarati, Damodar N (2000). *Ecomometria Básica*. Makron Books. Terceira Edição. São Paulo.
- [16] Hill, Carter R.; Griffiths, William E.; Judge, George G (2003). *Ecomometria*. Editora Saraiva. Segunda Edição. São Paulo.
- [17] Hosking, J. R. M.(1981). Fractional differencing. *Biometrika*, **68**, 165 – 176.
- [18] Knuth, D.E. (1981). *The Art of Computer Programming, Third Edition. Volume 2: Seminumerical. Algorithms*, Addison-Wesley Publishing Company, Massachusetts.
- [19] Korgi, Rodrigo de Castro (2003). *El universo L^AT_EX*. Editora Facultad de Ciencias. Segunda Edição. Bogota.
- [20] Krawczyk, H. *How to Predict Congruential Generators*. In: *Journal of Algorithms*, V. 13, N. 4, 1992.
- [21] L'Ecuyer, P (1994). *Uniform random number generation*. *Annals of Operations Research*, pp. 77 – 120.
- [22] L'Ecuyer, P (2001). *Software for uniform random number generation: distinguishing the good and the bad*. In *Proceedings of the 2001 Winter Simulation Conference*.
- [23] Marsaglia, G. *The Marsaglia Random Number including the DIEHARD Battery of Tests of Randomness* (Disponível em: <http://stat.fsu.edu/pub/diehard/>).
- [24] Marsaglia, G. *A Current View of Random Number Generators, Keynote Address, Computer Science and Statistics*. In: 16th Symposium on the Interface, Atlanta. Published by Elsevier Press, 1984.
- [25] Marsaglia, G.(1993) *Monkey Tests for Random Number Generators*. In: *Computers and Mathematics with Applications*, 9, p.1-10, 1993.
- [26] McCullough (1998). *Benchmarking Statistical software*. *American Statistician*, Fortcoming.
- [27] Mingoti(2005). *Analise De Dados Atraves De Metodos De Estatistica Multivariada - Uma Abordagem Aplicada*. Editora UFMG. Belo Horizonte.

-
- [28] Morettin, A. P., Clélia, M. C.(2006) *Análise de séries temporais*. Editora Egard Blucher, 2^{da} edição.
- [29] Papoulis, A. e Pillai, S. (2002). *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, 4th edition.
- [30] R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- [31] Reisen, V., Lemonte, A.(2006). The fracdiff package Version 1.3 – 1.
- [32] Reisen, V. (2007): Minicurso: MODELO ARFIMA.
- [33] Soto, J. *Statistical Testing of Random Number Generators*. In: Proceedings of the 22nd National Information Systems Security Conference, Crystal City, Virginia, 1999.
(Disponível em: <http://csrc.nist.gov/rng/nissc-paper.pdf>).
- [34] VenablesWN, Ripley BD (2000). *S Programming*. Springer-Verlag, New York.
- [35] Venables WN, Ripley BD (2002). *Modern Applied Statistics with S*. 4th edition. Springer-Verlag, New York.
- [36] Wong, P. C (1990). *Random number generation without multiplication*. In 9th Annual Intl Phoenix Conf on Computers and Communications, p.217-221.
- [37] Zafon, G; Manacero Jr, Aleardo (2006). *Construção de Geradores independentes de números aleatórios para diferentes distribuições probabilísticas*. Anais do XXVI Congresso de SBC, p.16-21. Campo Grande. MS