

A FEA e a USP respeitam os direitos autorais deste trabalho. Nós acreditamos que a melhor proteção contra o uso ilegítimo deste texto é a publicação online. Além de preservar o conteúdo motiva-nos oferecer à sociedade o conhecimento produzido no âmbito da universidade pública e dar publicidade ao esforço do pesquisador. Entretanto, caso não seja do interesse do autor manter o documento online, pedimos compreensão em relação à iniciativa e o contato pelo e-mail [bibfea@usp.br](mailto:bibfea@usp.br) para que possamos tomar as providências cabíveis (remoção da tese ou dissertação da BDTD).

**UNIVERSIDADE DE SÃO PAULO**  
**FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE**  
**DEPARTAMENTO DE ADMINISTRAÇÃO**

**PREVISÃO DE SUCESSÕES CRONOLÓGICAS ECONÔMICO-FINANCEIRAS  
POR MEIO DE REDES NEURAS ARTIFICIAIS RECORRENTES DE TEMPO  
REAL E DE PROCESSOS ARMA-GARCH: UM ESTUDO COMPARATIVO  
QUANTO À EFICIÊNCIA DE PREVISÃO**

**Mauri Aparecido de Oliveira**

**Orientador: Prof. Dr. José de Oliveira Siqueira**

**São Paulo**

**2003**

T330.015195  
048p

**UNIVERSIDADE DE SÃO PAULO**  
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE  
DEPARTAMENTO DE ADMINISTRAÇÃO

**PREVISÃO DE SUCESSÕES CRONOLÓGICAS ECONÔMICO-FINANCEIRAS POR  
MEIO DE REDES NEURAS ARTIFICIAIS RECORRENTES DE TEMPO REAL E  
DE PROCESSOS ARMA-GARCH: UM ESTUDO COMPARATIVO QUANTO À  
EFICIÊNCIA DE PREVISÃO**

↙  
**Mauri Aparecido de Oliveira**

**Orientador: Prof. Dr. José de Oliveira Siqueira**

<b>USP - FEA - SBD</b>
DATA DA DEFESA <u>19 / 01 / 04</u>

Dissertação apresentada ao Departamento de Administração na Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, como requisito parcial para a obtenção do título de Mestre em Administração

DEDALUS - Acervo - FEA



20600025853

**São Paulo**  
**2003**

85883

CRRCR

# PREVISÃO DE SUCESSÕES CRONOLÓGICAS ECONÔMICO-FINANCEIRAS POR MEIO DE REDES NEURAS ARTIFICIAIS RECORRENTES DE TEMPO REAL E DE PROCESSOS ARMA-GARCH: UM ESTUDO COMPARATIVO QUANTO À EFICIÊNCIA DE PREVISÃO

## Resumo

*O principal objetivo desse trabalho é estudar o processamento de séries temporais para a realização de previsão utilizando redes neurais artificiais e os modelos ARIMA-GARCH.*

*Com relação as redes neurais foram estudados os algoritmos de processamento temporal utilizando redes neurais alimentadas adiante e as redes recorrentes. Sendo que nas redes recorrentes o algoritmo utilizado para análise da série temporal foi o algoritmo de aprendizagem recorrente em tempo real (RTRL). Para os modelos ARIMA foi utilizada a metodologia desenvolvida por Box e Jenkins. Foram utilizadas as séries temporais de retornos diários do IBOVESPA, Petrobrás, Nasdaq, IBM e saca de 60Kg de soja como exemplo de aplicação das metodologias.*

# FORECASTING OF ECONOMIC-FINANCIAL CHRONOLOGICAL SUCCESSION USING REAL TIME RECURRENT NEURAL NETWORKS AND ARMA-GARCH PROCESS: AN COMPARATIVE EFFICIENCY STUDY

## Abstract

*The main objective of this dissertation is the time series processing to perform forecasting using artificial neural networks and ARIMA models. Regarding to neural networks to perform time series processing my studies focused feedforward and recurrent networks. The main recurrent algorithm applied to time series analysis were real time recurrent learning (RTRL). The Box and Jenkins methodology was applied to ARIMA analysis. As an application example we analyzed the following daily returns time series: IBOVESPA, Petrobras, Nasdaq, IBM, 60Kg soybean bag.*

# **Dedicação**

Esta dissertação é dedicada à Cristina Mayumi Shiguemori Oliveira que teve paciência e sempre me apoiou durante os meus estudos. Sem o seu amor e encorajamento eu jamais teria conseguido chegar onde estou.

# Agradecimentos

Ao Prof. Dr. José de O. Siqueira, pelos ensinamentos, pelo apoio e pela dedicada orientação durante o curso de mestrado e durante a elaboração desse trabalho.

Aos meus pais, aos meus irmãos, e ao meu sobrinho Kepler Silveira pelo carinho e incentivo.

A todos os professores do curso de mestrado pelos seus ensinamentos.

Aos meus amigos do curso de mestrado pela amizade e pelos bons momentos.

À FEA pelo apoio e estrutura durante o curso de mestrado.

Ao programa de aperfeiçoamento de ensino (PAE) pela oportunidade do estágio.

À Prof. Solange Oliveira Rezende do ICMC-USP/São Carlos pelo incentivo desde o curso de graduação e pela minha iniciação científica que tanto contribuiu para minha formação acadêmica e profissional.

À Senhora Fumiko Shiguemori pelo auxílio incondicional durante todas as etapas de elaboração desse trabalho.

# Abreviações e Símbolos

## Abreviações

AIC	<i>Akaike Information Criteria</i> – Critério de Informação de Akaike
AR	Auto-regressivo
ARCH	<i>Autoregressive Conditional Heterocedasticity</i> – Heterocedasticidade Autoregressiva Condicionada
ARIMA	Auto-regressivo Integrado de Média Móvel
ARMA	Auto-regressivo Média Móvel
BIC/SBC	<i>Bayesian Information Criteria</i> – Critério de Informação Bayesiano <i>Schwartz Bayesian Criteria</i> – Critério de Informação de Schwartz
BP	<i>Back Propagation</i> - Retropropagação
BPTT	<i>Back Propagation Through Time</i> - Retropropagação Através do Tempo
EKF	<i>Extended Kalman Filter</i> - Filtro de Kalman Extendido
FAC	Função de Auto-correlação
FACP	Função de Auto-correlação Parcial

FIR	<i>Finite Impulse Response Filter</i> – Filtro de resposta a impulso de duração finita
FK	Filtro de Kalman
GARCH	<i>Generalized Autoregressive Conditional Heterocedasticity</i> – heterocedasticidade autoregressiva condicionada
GARCH-M	GARCH-in-Mean
IA	Inteligência Artificial
MA	<i>Moving Average</i> - Média Móvel
MAPE	<i>Mean Absolute Percentage Error</i> - erro absoluto médio percentual
MLP	<i>Multi-Layer Perceptron</i> - Perceptron de Camadas Múltiplas
MQO	Mínimos Quadrados Ordinários
MSE	<i>Mean Squared Error</i> - Erro Quadrático Médio
MSPE	<i>Mean Squared Prediction Error</i> - erro quadrático médio de predição
NID	Normal e Identicamente Distribuído
RB	Ruído Branco
RMLP	<i>Recurrent Multi-Layer Perceptron</i> - perceptron recorrente de camada múltipla
RMSE	<i>Root Mean Square Error</i> – raiz do erro quadrático médio
RMSPE	<i>Root Mean Squared Prediction Error</i> – raiz do erro quadrático médio de predição
RSPE	<i>Root Squared Prediction Error</i> – raiz do erro quadrado de predição
RTRL	<i>Real Time Recurrent Learning</i> - aprendizado recorrente em tempo real

SIC	<i>Schwarz Information Criteria</i> – critério de informação de Schwarz
SPE	<i>Square prediction error</i> – erro quadrado de predição
SSE	<i>Sum of Squared Errors</i> – soma quadrática dos erros
TDNN	<i>Time Delay Neural Network</i> – rede neural com atraso de tempo
TLFN	<i>Time Lagged Feedforward Network</i> – rede alimentada adiante atrasada no tempo

### Símbolos Importantes

Em negrito significa que são vetores

$\mathbf{r}(t)$	ruído de medida no tempo $t$
$\mathbf{R}(t)$	matriz de covariância do ruído de medida
$\mathbf{q}(t)$	processo-ruído
$\mathbf{Q}(t)$	matriz de covariância do processo ruído
$\mathbf{y}(t)$	vetor de medida
$\mathbf{x}(t)$	vetor de estado
$\hat{\mathbf{x}}(t t)$	estimativa de estado a posterior na teoria de Kalman
$\hat{\mathbf{x}}(t t-1)$	estimativa de estado prévio
$\mathbf{P}(t t)$	matriz de correlação estado-erro a posterior
$\mathbf{P}(t t-1)$	matriz de correlação estado-erro a prévio
$\mathbf{K}(t)$	ganho de Kalman
$\mathbf{H}(t)$	matriz de medida
$\mathbf{F}(t)$	matriz de transição de estado
$h_t(\bullet)$	função vetor de medida não linear

$f_t(\bullet)$	função vetor estado não linear
$\mathbf{H}_n(t)$	jacobiana – matriz de medida
$\mathbf{F}_t(t)$	jacobiana – matriz transição de estado
$\mathbf{I}$	matriz identidade
$t_d$	profundidade de truncamento
$P_p$	parâmetro que inicializa $\mathbf{P}(t t)$
$P_r$	parâmetro taxa de aprendizagem que inicializa $\mathbf{R}(t)$
$t_t$	número total de unidades na rede
$t_o$	número de unidades de saída na rede
$t_i$	número de unidades de entrada na rede
$t_w$	número total de pesos sinápticos na rede
$\mathbf{u}(t)$	vetor de entrada da rede neural
$\phi(\bullet)$	função de ativação
$\mathbf{w}(t)$	vetor de pesos sinápticos ordenados
$w_i$	pesos individuais
$s(t)$	t-ésima amostra de série temporal
$\mathbf{s}(t)$	vetor coordenado de atrasos
$\tau$	atraso

# Sumário

<b>Resumo</b>	<b>v</b>
<b>Abstract</b>	<b>vi</b>
<b>Dedicação</b>	<b>vii</b>
<b>Agradecimentos</b>	<b>viii</b>
<b>Abreviações e Símbolos</b>	<b>x</b>
<b>Lista de Tabelas</b>	<b>xvii</b>
<b>Lista de Figuras</b>	<b>xviii</b>

## **1 Introdução**

1.1 Motivação .....	1
1.2 Objetivos da Pesquisa e Considerações Iniciais .....	3
1.3 Delineamento da Dissertação .....	6

## **2 Modelos Econométricos**

2.1 Econometria .....	7
2.2 Séries Temporais .....	8
2.2.1 Processos Estocásticos .....	9
2.2.2 Modelos Estacionários .....	10
2.2.2.1 Modelos Auto-regressivos de Médias Móveis (ARMA) .....	10

2.2.2.2 Modelos Auto-regressivos (AR) .....	11
2.2.2.3 Modelos de Médias Móveis (MA) .....	12
2.2.3 Modelos Não-estacionários .....	13
2.2.4 Heterocedasticidade Condicionada .....	14
2.2.4.1 Modelos ARCH .....	15
2.2.4.2 Modelos GARCH .....	17
2.2.4.2.1 Modelo GARCH-M.....	19
2.3 Metodologia de Box e Jenkins .....	21
2.3.1 Identificação.....	21
2.3.1.1 Função de Auto-correlação.....	22
2.3.1.2 Função de Auto-correlação Parcial.....	26
2.3.2 Estimação .....	27
2.3.3 Diagnóstico.....	29
2.3.4 Previsão.....	31

### 3 Redes Neurais Artificiais

3.1 Características das Redes Neurais .....	37
3.2 Motivação para utilizar Redes Neurais na Previsão .....	40
3.2.1 Redes Neurais e Estatística.....	42
3.3 Modelo de um Neurônio Artificial .....	44
3.4 Algoritmo de Retropropagação ( <i>Backpropagation</i> ) .....	46
3.5 Redes Neurais Alimentadas Adiante ( <i>Feedforward</i> ) .....	50
3.5.1 Incorporação do Tempo na Rede Neural .....	51
3.5.2 Algoritmo de Retropropagação Temporal .....	56
3.6 Redes Neurais Recorrentes .....	62
3.6.1 Redes de Elman e o Modelo de Espaço de Estados .....	64
3.6.2 Algoritmo de Aprendizagem Recorrente em Tempo Real (RTRL) .....	66
3.6.3 Algoritmo de Retropropagação Através do Tempo (BPTT) .....	73
3.6.3.1 Algoritmo de Retropropagação Através do Tempo por Época .....	73
3.6.3.2 Algoritmo de Retropropagação Através do Tempo Truncada .....	76

3.6.4 Filtro de Kalman e as Redes Recorrentes de Perceptrons de	
Múltiplas Camadas .....	78
3.6.4.1 Filtro de Kalman .....	78
3.6.4.2 Filtro de Kalman Extendido .....	82
3.6.4.3 Aplicação do EKF a Redes Neurais .....	84
<b>4 Resultados Experimentais</b>	
4.1 Descrição da série de retornos do IBOVESPA diário .....	87
4.1.1 Testes de raiz unitária.....	89
4.1.2 Estimação dos modelos ARIMA-ARCH .....	90
4.1.3 Utilização de redes neurais artificiais para previsão da série do	
IBOVESPA.....	96
4.1.3.1 Considerações sobre o tratamento dos dados.....	96
4.1.3.2 Especificação da rede neural recorrente .....	97
4.1.4 Análise gráfica e estatísticas de previsão.....	98
4.2 Análise da série de retornos diários da Petrobrás.....	101
4.3 Análise da série de retornos diários da Nasdaq.....	108
4.4 Análise da série de preços da saca de 60kg da soja.....	113
4.5 Análise da série de retornos das ações da IBM.....	121
4.6 Comparação dos resultados obtidos pelas previsões realizadas pelos modelos	
ARIMA-GARCH e redes neurais.....	125
<b>5 Conclusões.....</b>	<b>127</b>
<b>6 Apêndice .....</b>	<b>133</b>
6.1 Previsão .....	133

6.1.1	Derivação de Previsões de Erro Médio Quadrático Mínimo.....	134
6.1.2	Três Formas Básicas para a Previsão.....	137
6.1.3	Função de Previsão.....	139
6.2	Filtro de Kalman: apresentação dos conceitos .....	140
6.3	Valores do IBOVESPA.....	150
6.4	Série da Petrobrás.....	153
6.5	Série da Nasdaq.....	158
6.6	Série de retornos da IBM.....	159
6.7	Série da Soja.....	162
6.8	Programa MATLAB para o algoritmo de aprendizagem recorrente em tempo real – ARTR.....	166
<b>7</b>	<b>Referências Bibliográficas .....</b>	<b>168</b>

# Lista de Tabelas

3.1 Terminologias usadas em modelos estatísticos e em redes neurais.....	43
3.2 Resumo do algoritmo de retropropagação temporal para uma rede alimentada adiante .....	62
3.3 Algoritmo RTRL (Williams e Zipser, 1989) 72.....	72
3.4 Equações do filtro de Kalman .....	80
3.5 Equações do filtro de Kalman estendido .....	83

# Lista de Figuras

1.1	Procedimento para modelagem de processos ARIMA.....	4
1.2	(a) Arquitetura de rede alimentada adiante (feedforward) (b) rede recorrente .....	5
3.1	Modelo de um neurônio artificial .....	44
3.2	Rede <i>Feedforward</i> – Alimentada Adiante .....	50
3.3	Exemplo de uma rede TDNN - <i>Time Delay Neural Network</i> .....	53
3.4	Filtro Neural Focado (a), Rede TLFN – Time Lagged Feedforward Network (b).....	53
3.5	Filtro FIR aplicado à uma rede neural .....	55
3.6	Representação do somatório da Equação (3.50).....	59
3.7	Derivadas $\partial v/\partial y$ .....	60
3.8	Rede de Jordan .....	63
3.9	Rede de Elman .....	64
3.10	Rede de Elman estendida .....	64
3.11	Representação gráfica do princípio da descida do gradiente.....	68
3.12	Desdobramento para construção do algoritmo BPTT .....	74
3.13	Gradientes locais no desdobramento para construção do algoritmo BPTT.	75
4.1	Série de pontos de fechamento do IBOVESPA diário (período de 03/01/2003 até 18/11/2003).....	88
4.2	Série dos retornos do IBOVESPA diário (período de 03/01/2003 até 18/11/2003).....	89

4.3	Estatísticas dos retornos do IBOVESPA diário.....	90
4.4	Correlograma da série de retornos DLIBOVESPA diário.....	91
4.5	Correlograma dos resíduos quadrados da série DLIBOVESPA, modelo AR([14], [16]).....	92
4.6	RMSE da RNA recorrente e do modelo AR([16])-GARCH(2,1).....	99
4.7	MAE da RNA recorrente e do modelo AR([16])-GARCH(2,1).....	99
4.8	Coeficiente de desigualdade de Theil para a RNA recorrente e do modelo AR([16])-GARCH(2,1).....	100
4.9	Valores para previsão da série NDLIBOVESPA utilizando a RNA 10- 2R-1.....	100
4.10	Valores para previsão do DLIBOVESPA utilizando o modelo AR([16])- GARCH(2,1).....	101
4.11	Série de pontos de fechamento das ações da Petrobrás.....	102
4.12	Série de retornos ações da Petrobrás.....	102
4.13	Estatísticas dos retornos diários da Petrobrás.....	103
4.14	Valores para previsão da série DLPETRO utilizando modelo AR([1])- GARCH(1,1).....	106
4.15	Valores para previsão da série NDLPETRO utilizando a RNA 10-3R-1....	107
4.16	Histograma dos resíduos da previsão da RNA 10-2R-1.....	107
4.17	Série de valores da Nasdaq .....	108
4.18	Série de retornos da Nasdaq .....	108
4.19	Estatísticas dos retornos diários da Nasdaq.....	109
4.20	Função de auto-correlação para a série de retornos diários da Nasdaq.....	109
4.21	Função de auto-correlação parcial para a série de retornos diários da Nasdaq.....	110
4.22	Ajuste da curva de previsão de uma rede neural 10-4R-1 para a série NDLNASDAQ.....	112
4.23	Histograma dos resíduos da previsão da RNA 10-4R-1 para a série NDLNASDAQ.....	113
4.24	Série de valores de preços diários da saca de 60Kg da soja. Fonte: Série histórica Indicador CEPEA/ESALQ (R\$ /sc 60 kg).....	113

4.25	Série de retornos diários da saca de 60Kg da soja	114
4.26	Estatísticas dos retornos diários da saca de 60Kg da soja.....	114
4.27	FAC e FACP da série de retornos diários (DLSOJA) da saca de 60Kg da soja.....	115
4.28	Correlograma dos resíduos quadrados da série DLSOJA.....	116
4.29	Valores para previsão da série DLSOJA utilizando modelo AR([1])-GARCH(1,1).....	120
4.30	Ajuste da curva de previsão de uma rede neural 10-2R-1 para a série NDLSOJA.....	120
4.31	Série de retornos diários DLIBM. Fonte: Analysis of Financial Time Series.....	121
4.32	Histograma da série de retornos diários DLIBM.....	121
4.33	Função de auto-correlação e auto-correlação parcial para a série de retornos diários da IBM.....	122
4.34	Ajuste da curva de previsão de uma rede neural 10-2R-1 para a série NDLIBM.....	124
4.35	Histograma dos resíduos da previsão da RNA 10-2R-1 para a série NDLIBM.....	124
4.36	Valores da série do IBOVESPA no nível e da curva de previsão do modelo ARIMA-GARCH.....	125
4.37	Valores da série do IBOVESPA no nível e da curva de previsão da RNA.	126
4.38	Valores da série da Petrobrás no nível e da curva de previsão do modelo ARIMA-GARCH.....	127
4.39	Valores da série da Petrobrás no nível e da curva de previsão da RNA.....	127
4.40	Valores da série do preço da saca de 60Kg de soja no nível e a curva de previsão do modelo ARIMA-GARCH.....	128
4.41	Valores da série do preço da saca de 60Kg de soja no nível e a curva de previsão da RNA.....	129
6.1	Representação de uma forma de aplicação do filtro de Kalman.....	142
6.2	Densidade condicional da posição baseada no valor da medida $z_1$ .....	146
6.3	Densidade condicional da posição baseada apenas na medida $z_2$ .....	147
6.4	Densidade condicional da posição baseada nos dados $z_1$ e $z_2$ .....	148

# Capítulo 1

## Introdução

### 1.1 Motivação

A previsão de séries temporais (sucessões cronológicas ou crônicas) tem sido uma das áreas-chave em Ciências Sociais Aplicadas. Dadas algumas observações passadas do comportamento de alguns sistemas, surge a questão de como fazer previsões sobre o comportamento futuro e quão precisas essas previsões podem ser.

As condições presentes determinam em algum grau o futuro de tal forma que pode haver muitas interações e complexas relações entre as variáveis envolvidas. Tais processos ocorrem em muitos dos ramos da ciência, sendo que alguns exemplos incluem meteorologia, oceanografia, epidemiologia, a economia e finanças.

Existindo informação suficiente disponível sobre o sistema em estudo, então uma abordagem matemática pode ser desejável, sendo que as equações construídas modelam os mecanismos responsáveis pela geração das séries temporais e como o seu comportamento evolui. No entanto, em muitos problemas reais de interesse não temos informações e/ou condições ideais suficientes para construir equações que governem o comportamento das

variáveis que queremos prever. Na ausência de informações suficientes para gerar as equações é mais atrativo usar uma abordagem baseada em modelos. Uma abordagem baseada em redes neurais artificiais é um exemplo de abordagem baseada em modelos, na qual uma rede é treinada para abstrair as equações geradoras de uma série temporal. Portanto, ao invés de explicitamente deduzir equações que descrevam o comportamento subjacente do sistema em estudo, um capturador automático de um modelo não linear usando um algoritmo de aprendizagem neural é construído para aproximar as equações ideais. As redes neurais são atrativas para modelar sistemas não lineares porque elas são inerentemente não lineares, devido as funções de ativação não lineares, e também são bem estabelecidas como funções de aproximação (Forsgren e Kling, 2003). Tem sido mostrado que uma rede neural com um número arbitrário de unidades (ou neurônios) e possuindo duas ou mais camadas pode aproximar qualquer função uniformemente contínua (Wan e Merwe, 2001).

Entre as muitas técnicas que desempenharam um papel importante no desenvolvimento teórico dos métodos para previsão destaca-se a metodologia de Box e Jenkins (Box, Jenkins e Reinsel, 1994). Esta metodologia foi desenvolvida no começo da década de 1970 para possibilitar a construção de modelos auto-regressivos integrados de média móvel (ARIMA) que uma vez estimados são utilizados para realizar previsão. Utilizaremos os modelos ARIMA como ponto inicial para a análise das séries temporais porque muitos modelos neurais para processamento temporal são construídos a partir de variações de modelos ARIMA.

Muitos algoritmos neurais para processamento temporal têm sido propostos, entre estes destacam-se o algoritmo de retropropagação através do tempo (BPTT, *back propagation through time*), aprendizado recorrente em tempo real (RTRL, *real-time recurrent learning*) e os algoritmos de redes recorrentes utilizando o filtro de Kalman ((Williams, 1992) (Williams e Zipser, 1995)).

## 1.2 Objetivos da Pesquisa e Considerações Iniciais

O objetivo deste trabalho é estudar os modelos estocásticos de séries temporais e os modelos de redes neurais utilizadas para processamento temporal e sua aplicação na previsão de séries univariadas econômico-financeiras. As redes neurais são poderosas técnicas de aprendizagem no sentido de serem capazes de extrair características relevantes de grandes conjuntos de dados e isto é muito útil em aplicações no mundo real envolvendo sistemas complexos que devem ser analisados baseando-se em conhecimentos prévios incompletos.

Embora muitas técnicas de regressão não linear estejam sendo desenvolvidas, as propriedades de não linearidade das redes neurais fazem com que sejam preferidas para previsão de séries temporais. Colocando as redes neurais no contexto dos métodos estatísticos tradicionais, elas podem ser consideradas como uma técnica de estimação não paramétrica, não linear, multivariada e múltipla, ou seja é uma função dos dados sem um modelo de descrição específico. Ser multivariada indica que a entrada da rede neural contém duas ou mais variáveis distintas, cujas interdependências e causalidades são exploradas para prever o comportamento futuro de uma ou mais variáveis intertemporais. Não paramétrica e sem um modelo específico é uma consequência de algumas suposições referentes as relações entre as variáveis de entrada e as extrapolações dentro de um horizonte de tempo futuro.

Um importante aspecto das redes neurais é a sua capacidade de construir relações não lineares entre os dados de entrada e as saídas desejadas. A linearidade (por exemplo,  $y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots$  sendo que  $y_t$  é uma série e os  $\alpha_i$  são parâmetros) apenas descreve a relação mais simples que pode ocorrer em uma série temporal. As redes neurais apresentam muitas características desejáveis como uma ferramenta para análise de dados e com um esquema de implementação relativamente eficiente em termos de velocidade de computação e memória necessárias para processamento. Estes modelos podem aprender de experiências anteriores, podem generalizar e “ver” através de ruídos e distorções e podem abstrair características essenciais na presença de dados irrelevantes. De acordo com Lippman (1987), estes modelos fornecem um alto grau de robustez e tolerância à falha. Em

adição à estas características, os MRNA (modelos de redes neurais artificiais) podem encontrar as transformações corretas para as variáveis, detectar relações lineares fracas e trabalhar com valores que estejam discrepantes das demais observações (*outliers*).

O procedimento para o modelamento ARIMA normalmente é realizado em quatro etapas distintas, como mostrado na Figura-1.1 (Jhee e Lee, 1993), conhecidas como identificação, estimação, diagnóstico e previsão.

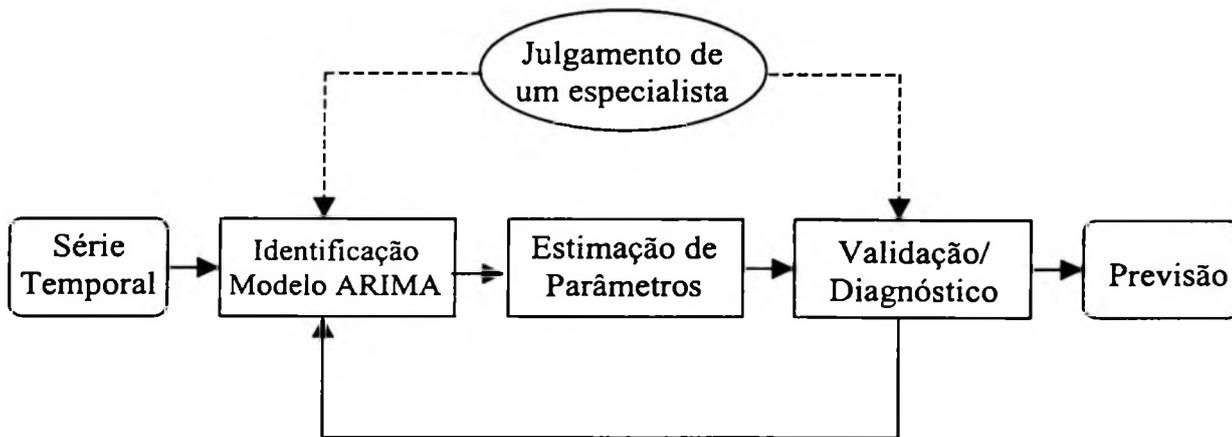


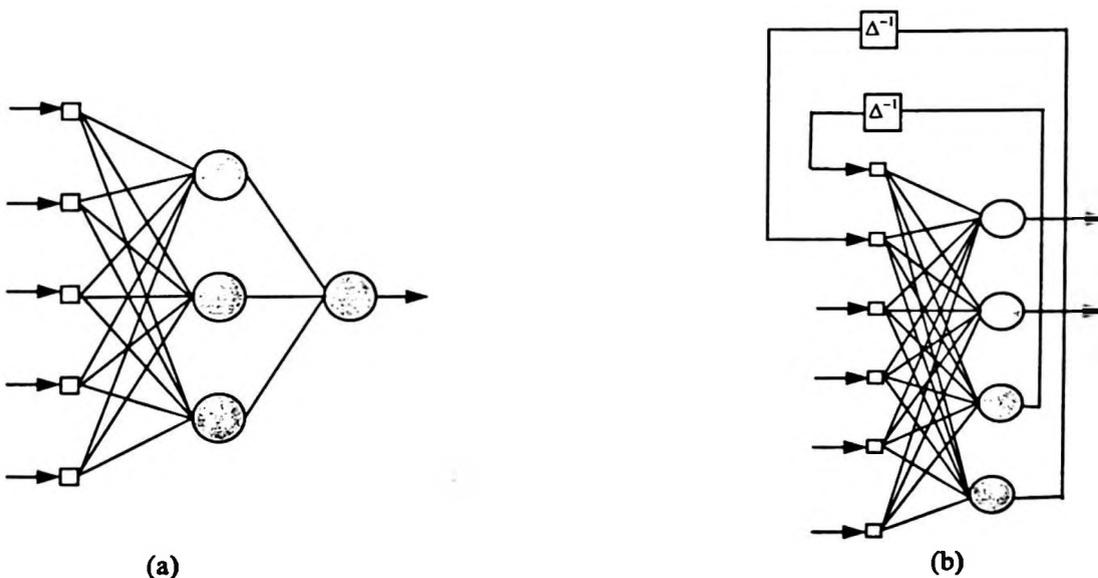
Figura-1.1 Procedimento para modelagem de processos ARIMA (Jhee e Lee, 1993)

O modelamento ARIMA é uma das ferramentas gerenciais para previsão mais amplamente usadas e é utilizada como um método de referência na criação de modelos lineares de dados temporais, principalmente por causa de sua elaboração teórica e precisão nas previsões de curto prazo. Os modelos ARIMA são construídos através de passos iterativos, como mostrados na Figura-1.1. Entre as etapas de modelamento, a identificação do modelo que determina a ordem  $p$  do processo AR e a ordem  $q$  do processo MA é crucial para representar a série temporal adequadamente, porque se a série temporal possuir ruído o modelo construído pode conter mais parâmetros que o necessário e dessa forma contrariar o princípio da parcimônia de Box e Jenkins. Portanto, é importante o julgamento do especialista que realiza o modelamento na interpretação de algumas informações estatísticas tal como as funções de autocorrelação (FAC) e autocorrelação parcial (FACP).

Alguns sistemas para realizar a tarefa de previsão são construídos de maneira a suprimir o julgamento do especialista, isso é feito a partir do uso de uma rede neural que atua como um classificador de padrões na etapa de identificação (Jhee e Lee, 1993).

No caso das redes neurais estamos interessados em estudar quais as arquiteturas de rede que devem ser utilizadas para o processamento temporal e principalmente como derivar os algoritmos de aprendizagem necessários para realizar o processo de treinamento. A forma mais simples de fazer com que uma rede neural realize processamento temporal é através da adição de memória à estrutura da rede, estas memórias podem ser de longo prazo ou de curto prazo. A memória é inserida na rede através de atrasos de tempo que podem ser implementados no nível sináptico dentro da rede ou na camada de entrada da rede. Uma rede com memória de curto prazo muito utilizada é *rede neural de atrasos de tempo* (TDNN, *time delay neural network*) e as redes com memórias de longo prazo mais conhecidas são as *redes recorrentes* de Elman(1990) e de Jordan(1986).

Os algoritmos de treinamento de redes neurais recorrentes mais utilizados são o algoritmo de aprendizagem em tempo real (RTRL) e o de retropropagação através do tempo (BPTT). No entanto, muitos autores têm notado que um algoritmo utilizando o filtro de Kalman estendido (EKF, *extended Kalman filter*) pode ser usado para treinar redes correntes com a vantagem de possuir uma velocidade de convergência superior e baixa tendência de ficar fixado num mínimo local quando comparado com os outros métodos baseados no método do gradiente descendente. A Figura-1.2 mostra as arquiteturas de redes neurais artificiais mais utilizadas para processamento temporal.



**Figura-1.2** (a) Arquitetura de rede alimentada adiante (*feedforward*) (b) rede recorrente

Na rede neural alimentada adiante a rede é alimentada com as observações das variáveis de interesse de entrada da série temporal a cada passo de tempo e o objetivo é prever um passo à frente. As redes recorrentes, são redes neurais com um ou mais laços de realimentação, a rede opera dependendo de suas próprias saídas a cada passo de tempo. Para prever vários passos de tempo à frente é de fato uma tarefa difícil e desafiadora para a rede neural realizar, pois os erros de saída não poderão ser amplificados e deverão tender a um valor pré-estabelecido a cada passo do processamento (Elman, 1990 e Mozer, 1988). Utilizando o algoritmo de retropropagação o erro quadrático médio será minimizado se (1) o modelo não ficar preso em um mínimo local e (2) se houver um número adequado de neurônios na camada escondida. Para assegurar que o erro seja minimizado é de costume construir a rede com uma camada escondida, ou intermediária, e dobrar de forma sucessiva os neurônios pertencentes a ela até que o erro seja reduzido (Ersoy e Hong, 1990).

### **1.3 Delineamento da Dissertação**

No Capítulo 2 vamos discorrer sobre modelos econométricos, como as séries temporais estão inseridas neste contexto e descrever a metodologia de análise de Box e Jenkins. Vamos apresentar quais os requisitos para construir um modelo que represente a série temporal e como realizar previsões. No Capítulo 3, vamos explorar as redes neurais e suas capacidades de processamento temporal. Serão analisadas as redes alimentadas adiante e as redes recorrentes. Os processos utilizados para treinamento das redes neurais que realizam processamento temporal serão discutidos detalhadamente. Apresentaremos os algoritmos de retropropagação, RTRL e o de Filtro de Kalman Extendido. O Capítulo 4 contém os resultados experimentais obtidos a partir da análise de séries econômicas e de séries financeiras. O Capítulo 5 é destinado a sumarizar o trabalho apresentado e a tirar algumas conclusões.

## Capítulo 2

# Modelos Econométricos

### 2.1 Econometria

A Econometria refere-se à análise de dados que descrevem os fenômenos econômicos. Esses dados econômicos vêm quase que exclusivamente de origens não experimentais, isto é, os dados são observacionais. Os cientistas sociais geralmente devem aceitar as condições sob as quais seus objetos de estudos estão submetidos e a forma como as respostas ocorrem. Estes pesquisadores não podem especificar ou escolher o nível de um estímulo para depois então registrar um novo resultado podendo apenas observar os experimentos naturais que acontecem (Ruud, 2000). A Econometria pode ser considerada como uma área da matemática que tem intersecção com teoria econômica e estatística. Um processo econométrico básico envolve os seguintes passos (McNelis, 2002):

- 1- Desenvolvimento de hipóteses testáveis da teoria econômica;
- 2- Especificação da hipóteses em forma matemática;
- 3- Especificação de um modelo estatístico ou econométrico;
- 4- Coleta de dados;

- 5- Estimaco;
- 6- Previso ou predico;
- 7- Anlise poltica ou escolha comportamental.

A anlise economtrica comea com um modelo econmico, isto , existe alguma relao a prvia que  assumida, baseada em um modelo terico, o qual  ento testado usando dados do mundo real. A Estatstica  ento utilizada, por exemplo, para estimar parmetros e testar estatisticamente os modelos. Com modelos de regresso o objetivo  *prever* os valores das variveis dependentes *dentro* do domnio das variveis independentes. Com modelos de sries temporais o objetivo  *prever* os valores das variveis dependentes *fora* do domnio das variveis independentes. Sendo que a varivel independente reduz-se ao tempo.

## 2.2 Sries Temporais

A primeira razo para o interesse na anlise de sries temporais  adquirir conhecimento de padres que se desenvolvem a medida que os eventos acontecem em um perodo de tempo. Isto  genericamente feito atravs de observaes e anlises dos dados passados sobre ocorrncia dos eventos. Tipicamente, tais processos envolvem a consolidao das experincias histricas em sistemas matemticos que descrevem o comportamento dos eventos atravs da passagem do tempo. Esta tarefa deve ser feita da forma mais concisa possvel e pode ser chamada de construo do modelo. Num segundo estgio, estes modelos matemticos so utilizados para projetar o que  provvel de ocorrer durante algum curto perodo de tempo. Tais processos so usualmente chamados de "previso", ou seja, uma extrapolao para alm do domnio temporal conhecido. A combinao destes dois processos forma essencialmente o que se convencionou chamar de anlise de sries temporais. Os resultados dessa anlise podem ser utilizados de modos muito diferentes, sendo que sua aplicao  de grande importncia nas reas de Economia e Finanas.

Uma srie temporal  um conjunto de observaes geradas seqencialmente no tempo (Box e Jenkins, 1976). Se o conjunto  contnuo, a srie temporal  chamada de contnuo. Se o conjunto  discreto, a srie temporal  chamada de discreta. As observaes de uma srie

temporal feitas nos tempos  $t_1, t_2, \dots, t_i, \dots, t_n$  podem ser denotadas por  $y(t_1), y(t_2), \dots, y(t_i), \dots, y(t_n)$ . Tendo  $n$  sucessivos valores de uma série disponível para análise também podemos escrever  $y_1, y_2, \dots, y_i, \dots, y_n$  para denotar as observações feitas em intervalos de tempo eqüidistantes  $t_0 + h, t_0 + 2h, \dots, t_0 + ih, \dots, t_0 + nh$ . Na prática, muitos dados de séries temporais econômicas e financeiras são coletadas em períodos de tempo discretos. Dessa forma, utilizaremos séries temporais discretas ao invés de contínuas, sendo que  $y_t$  representará uma variável aleatória observável da série temporal num dado instante  $t$ .

Normalmente, as séries econômicas, financeiras ou de algum outro negócio podem possuir cinco características chave: (i) tendência, (ii) sazonalidade, (iii) algum ponto de influência discrepante, (iv) uma variância que se altera devido as observações passadas - heterocedasticidade condicional e (v) não-linearidade. Tipicamente, uma série temporal econômica apresenta pelo menos duas ou três dessas características (Franses, 1998).

### 2.2.1 Processos Estocásticos

Se os valores futuros de uma série temporal são determinados exatamente por alguma função matemática, então esta série é denominada de determinística. Se os valores futuros podem ser descritos apenas em termos de uma distribuição de probabilidade a série temporal é chamada de não-determinística ou simplesmente série temporal estocástica. Um fenômeno estatístico que envolva leis probabilísticas no tempo é chamado de processo estocástico. Neste trabalho as séries temporais serão analisadas como uma realização particular produzidas por um mecanismo de probabilidade subjacente. Ou seja, quando utilizarmos a metodologia de Box e Jenkins para analisar as séries temporais estaremos nos referindo a uma realização particular de um processo estocástico, ver Apêndice 7.1.

Um bloco construtivo básico dos modelos estocásticos discretos de séries temporais é o processo de ruído branco. A seqüência  $\{\varepsilon_t\}$  é um processo ruído branco se cada valor da seqüência tiver esperança nula, variância constante e forem serialmente não correlacionados.

Se a notação  $E[x]$  denota o valor esperado teórico de  $x$ , a seqüência  $\{\varepsilon_t\}$  é um processo de ruído branco se para cada instante de tempo  $t$  tivermos (Enders, 2003):

$$E[\varepsilon_t] = E[\varepsilon_{t-1}] = \dots = 0$$

$$E[\varepsilon_t^2] = E[\varepsilon_{t-1}^2] = \dots = \sigma^2 < \infty$$

$$E[\varepsilon_t \cdot \varepsilon_{t-s}] = E[\varepsilon_{t-j} \cdot \varepsilon_{t-j-s}] = 0, t \neq s$$

## 2.2.2 Modelos Estacionários

Formalmente, a definição de estacionariedade é que para uma dada série temporal  $y_t$  as condições abaixo deverão ser satisfeitas (Enders, 2003):

$$E[y_t] = \mu \quad \text{para todo } t = 1, 2, \dots, n \quad (2.1)$$

$$E[(y_t - \mu)^2] = \gamma_0 \quad \text{para todo } t = 1, 2, \dots, n \quad (2.2)$$

$$E[(y_t - \mu)(y_{t-s} - \mu)] = \gamma_s \quad \text{para todo } t = 1, 2, \dots, n, t \neq s \quad (2.3)$$

$$\text{e para todo } s = \dots, -2, -1, 0, 1, 2, \dots$$

Sendo que  $\mu$ ,  $\gamma_0$ ,  $\gamma_s$  são todos números finitos (Enders, 2003). Ou seja, as equações acima significam que a média, auto-variância e auto-covariância devem ser constantes para que a série seja estacionária. Para uma dada série temporal é usualmente difícil verificar se estas três condições acontecem ao mesmo tempo. Intuitivamente, para verificar (2.1) com um determinado teste estatístico, nós precisaremos de um estimador da variância condicionada de  $y_t$ , o qual necessariamente deverá obedecer (2.2) que por sua vez depende da validade de (2.1). Na literatura, um processo de covariância estacionária é também referido como possuindo uma estacionariedade fraca. Para que um processo possua estacionariedade estrita (ou forte) é necessário que todos os momentos sejam iguais para uma dada distribuição conjunta das séries  $\{y_1, y_2, \dots, y_n\}$  e  $\{y_{1+h}, y_{2+h}, \dots, y_{n+h}\}$ , sendo  $h$  um número inteiro.

### 2.2.2.1 Modelos Auto-regressivos de Médias Móveis (ARMA)

Os modelos auto-regressivos (AR) e de médias móveis (MA) são casos especiais do modelo geral denominado auto-regressivo de médias móveis (ARMA).

ARMA é um modelo linear que tem a seguinte forma:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (2.4)$$

que também pode ser escrita como:

$$y_t = \phi_0 + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (2.5)$$

sendo que  $\{\varepsilon_t\} \sim \text{RB}$ .

A abordagem de Box e Jenkins necessita que o modelo seja invertível. Formalmente,  $\{y_t\}$  é invertível se puder ser representado por uma ordem finita ou o processo auto-regressivo for convergente. A invertibilidade é importante por causa do uso das funções de auto-correlação (FAC) e auto-correlação parcial (FACP) que implicitamente assumem que a seqüência  $\{y_t\}$  pode ser aproximada por um modelo auto-regressivo. Considerando um modelo MA(1) da forma  $y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$ , se  $|\theta_1| < 1$  teremos que  $y_t \geq (1 - \theta_1 B) \cdot \varepsilon_t$  ou  $y_t + \theta_1 y_{t-1} + \theta_1^2 y_{t-2} + \theta_1^3 y_{t-3} + \dots = \varepsilon_t$ . Se  $|\theta_1| < 1$  este modelo pode ser estimado usando a metodologia de Box e Jenkins. No entanto, se  $|\theta_1| \geq 1$ , então a seqüência  $\{y_t\}$  não pode ser representada por um processo de ordem finita AR e dessa forma não é invertível. Genericamente, para um modelo ARMA ter uma representação AR convergente, as raízes do polinômio  $1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$  devem estar fora do círculo unitário. Sendo que  $B$  é um operador de atraso unitário definido pela seguinte propriedade:

$$B^k y_t = y_{t-k} \text{ para } k = \dots, -2, -1, 0, 1, 2, \dots$$

Conseqüentemente,  $B^{-2} y_t$  significa  $y_{t+2}$  e  $B^0 y_t$  significa  $y_t$ . Quando  $0 < \alpha < 1$  podemos escrever:  $(1 - \alpha B)^{-1} = 1 + \alpha B + \alpha^2 B^2 + \alpha^3 B^3 + \dots$ , este operador é muito útil para sumarizar modelos complicados de séries temporais.

### 2.2.2.2 Modelos Auto-regressivos (AR)

Supondo que as observações em um série temporal  $y_t$  depende de  $p$  de suas observações atrasadas, ou seja, que  $y_t$  pode ser descrita pelo um modelo linear:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad t = p+1, p+2, \dots, n \text{ e } n > p \quad (2.6)$$

sendo que  $\phi_1, \phi_2$ , até  $\phi_p$  são parâmetros. Consideramos  $\varepsilon_t$  como sendo um processo ruído branco. Na prática, as observações sobre  $\varepsilon_t$  não são diretamente observadas e têm de ser estimadas dos dados, baseado nos pressupostos do modelo  $y_t$ . No caso da equação (2.6),  $y_t$

pode ser descrita por um modelo de regressão que inclua apenas variáveis  $y_t$  atrasadas, e conseqüentemente este modelo (2.6) é usualmente chamado de modelo auto-regressivo de ordem  $p$  [AR( $p$ )], ou um modelo de um processo auto-regressivo de ordem  $p$ . Uma restrição para que um processo AR(1) seja estacionário é que  $|\phi_1| < 1$ .

Com o operador de atraso  $B$ , a expressão (2.6) pode ser abreviada como:

$$\phi_p(B)y_t = \varepsilon_t \quad (2.7)$$

sendo que

$$\phi_p(B) = 1 + \phi_1 B + \dots + \phi_p B^p \quad (2.8)$$

o qual é chamado polinômio-AR em  $B$  de ordem  $p$ . Os pesos sobre os atrasos são os parâmetros  $\phi_1$  até  $\phi_p$  e estes expressam em qual extensão  $y_t$  depende do seu passado. Desde que a observação em qualquer período de tempo  $t$  depende de  $p$  observações passadas, então (2.6) assume que de algum modo  $y_t$  depende de todas as observações passadas. Para prever  $y_{n+h}$  deve-se antes de tudo assegurar que esta dependência do passado seja constante. De fato, se para qualquer tempo  $t$  esta dependência do passado diferir, haverá dificuldade na tentativa de prever  $y_{n+h}$  uma vez que para qualquer horizonte de tempo futuro  $h$  a função de previsão poderá ser viesada. Além disso, para que possamos fazer afirmações sensatas sobre  $y_{n+h}$  deverá ser considerado que o passado imediato é mais importante do que o menos recente. Em outras palavras, se para medir o impacto da observação no tempo  $t=10$ ,  $y_{10}$ , será necessário considerar observações não tão distantes, ou seja será preferível utilizar informações sobre  $y_{11}$  a informações dadas por  $y_{70}$ , por exemplo (Franses, 1998).

### 2.2.2.3 Modelos de Médias Móveis (MA)

Em alguns casos é conveniente considerar apenas uma variante simples de um modelo ARMA( $p,q$ ), ou seja o modelo MA( $q$ ) dado por:

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (2.9)$$

ou

$$y_t = \theta_q(B)\varepsilon_t \quad (2.10)$$

sendo que

$$\theta_q(B) = 1 + \theta_1 B + \dots + \theta_q B^q \quad (2.11)$$

Uma característica importante de um modelo MA(q), e conseqüentemente de um modelo ARMA(p,q) é que as variáveis na equação (2.9), ou seja,  $\varepsilon_{t-1}$  até  $\varepsilon_{t-q}$ , são não observadas e têm de ser estimadas usando a amostra de dados disponível. Para que isso não cause problemas é usual manter q o tanto menor quanto possível. Na prática, este valor de q é freqüentemente tomado como sendo 0, 1 ou 2. A primeira vista pode parecer que  $y_t$  não depende de seu próprio passado quando um modelo MA(q) descreve esta variável. No entanto, de forma similar ao que foi mostrado para o modelo auto-regressivo, a equação (2.10) pode ser escrita como:

$$[\theta_q(B)]^{-1} y_t = \varepsilon_t \quad (2.12)$$

o que mostra que  $y_t$  depende de todos os valores prévios de  $y_t$ . Por exemplo, para o modelo MA(1) temos que  $y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$  pode ser derivado da seguinte forma:

$$\begin{aligned} y_1 &= \varepsilon_1 \\ y_2 &= \varepsilon_2 + \theta_1 \varepsilon_1 = \varepsilon_2 + \theta_1 y_1 \\ y_3 &= \varepsilon_3 + \theta_1 \varepsilon_2 = \varepsilon_2 + \theta_1 (y_2 - \theta_1 y_1) \end{aligned} \quad (2.13)$$

...

e assim por diante. Similarmente ao conceito de raiz unitária no polinômio AR(p), pode haver uma ou mais raízes unitárias em um polinômio MA(q). Em geral, os modelos MA(q) são invertíveis quando as soluções para:

$$1 + \theta_1 z + \dots + \theta_q z^q = 0 \quad (2.14)$$

estão todas fora do círculo unitário.

### 2.2.3 Modelos Não-estacionários

Quando uma variável apresenta efeitos permanentes dos choques as séries são usualmente transformadas para uma série temporal com efeitos transitórios tomando-se as primeira diferença da série temporal  $y_t$ . A motivação é que quando  $y_t$  é uma série temporal denominada de passeio aleatório, ou seja apresenta a forma:

$$y_t = y_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i} \quad t = 1, 2, \dots, n \quad (2.15)$$

ela não apresenta estacionariedade. Por exemplo, no caso em que:

$$y_t = y_{t-1} + \varepsilon_t \quad (2.16)$$

todos os choques tem um efeito permanente. No entanto, para a série transformada  $z_t = y_t - y_{t-1}$ , a qual pode ser descrita por um simples modelo ruído branco:

$$z_t = y_t - y_{t-1} = \varepsilon_t \quad (2.17)$$

os choques passados et têm apenas efeitos transitórios.

Definindo um operador diferença  $\nabla$  onde  $\nabla = (1-B)$  temos que:

$$\nabla_j^d = (1 - B^j)^d \quad (2.18)$$

Quando a série temporal precisa ser diferenciada  $d$  vezes, ela é chamada de integrada de ordem  $d$ , ou de forma abreviada I(d). Quando  $y_t$  é uma série temporal I(d) e depois de diferenciada  $d$  vezes ela pode ser modelada usando um modelo AR(p), este modelo pode ser escrito como:

$$\nabla_1^d y_t = \phi_1 \nabla_1^d y_{t-1} + \dots + \phi_p \nabla_1^d y_{t-p} + \varepsilon_t \quad t = p+d, p+d+1, \dots, n \quad (2.19)$$

O modelo representado pela equação (2.18) é normalmente chamado de ARI(p,d).

## 2.2.4 Heterocedasticidade Condicionada

Em muitas séries econômico-financeiras tem-se que informações adicionais discrepantes e *outliers* provindos de inovações não acontecem com frequência. No entanto, as séries temporais financeiras são exceção. Uma vez que os dados refletem resultados de negócios entre compradores e vendedores, por exemplo, o mercado acionário, várias fontes de eventos econômicos exógenos podem ter um impacto sobre o padrão dos preços de capital. Dado que algumas notícias podem levar a várias interpretações diferentes e também que certos eventos econômicos específicos, tal como uma crise de petróleo, podem durar por algum tempo, freqüentemente observa-se que grandes quantidades positivas e grandes quantidades negativas de observações em séries temporais financeiras tendem a aparecer em clusters (agrupamentos). Uma abordagem para tratar desta situação é explorar o fato de que os outliers aparecem em clusters e tentar construir um modelo de série temporal para os próprios outliers. Uma vez que a seqüência de outliers pode ser considerada como refletindo a volatilidade do período, estes modelos de séries temporais podem ser utilizados para prever a volatilidade. Ou seja, devido a presença de conjuntos de valores discrepantes,

a variância de séries temporais financeiras varia em função do tempo e conseqüentemente os intervalos de previsão para cada nível também deverão variar. A intuição é que em períodos de maior volatilidade existe maior incerteza sobre a próxima observação do que em períodos de menor volatilidade e conseqüentemente em períodos que apresentam maior volatilidade os intervalos de previsão serão maiores.

### 2.2.4.1 Modelos ARCH

O modelo ARCH - Autoregressive Conditional Heteroskedasticity, ou seja de Heterocedasticidade Autoregressiva Condicionada foi proposto inicialmente por Engle (1982). Desde então, têm aparecido inúmeros estudos com refinamentos e modificações do modelo ARCH básico e também com aplicações empíricas em taxa de câmbio, mercado de ações e muitos outros tipos de ativos (Bollerslev, Engle e Nelson/1994).

Consideremos uma série temporal  $\varepsilon_t$ , a qual pode ser descrita pelo conjunto de equações:

$$\varepsilon_t = \eta_t \sqrt{h_t} \quad (2.20)$$

sendo que:

$$h_t = \text{Var}[\varepsilon_t | I_{t-1}]$$

$$\eta_t \sim \text{NID}(0;1) \quad (2.21)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2, \quad \alpha_0 > 0, \quad 0 < \alpha_1 < 1 \quad (2.22)$$

sendo que a equação (2.20) descreve o comportamento de  $\varepsilon_t$  condicionado a  $I_{t-1}$  e sendo que NID significa normal e identicamente distribuído. Na prática, a série  $\varepsilon_t$  pode ser igual a  $[\phi_p(B)/\theta_q(B)]x_t$ , onde  $x_t$  é a série realmente observada. Usualmente,  $x_t$  é o retorno sobre um ativo financeiro, ou seja,  $x_t = \log(z_t) - \log(z_{t-1})$ , sendo  $z_t$  a série no nível.

Dessa forma, tendo que  $\varepsilon_t | I_{t-1} = \eta_t \sqrt{h_t}$ ,  $\eta_t \sim \text{NID}(0,1)$  e a série  $y_t$  que pode ser escrita

como  $y_t = (1-B)\ln x_t = \nabla_1^1 \ln x_t$ , resulta então que:  $y_t = \frac{\theta_q(B)}{\phi_p(B)} \varepsilon_t$ ,  $\{\varepsilon_t\} \sim \text{RB}$ , isolando  $\varepsilon_t$

teremos  $\varepsilon_t = \frac{\phi_q(B)}{\theta_p(B)} y_t$ , portanto:

$$\varepsilon_t = \frac{\phi_q(B)}{\theta_p(B)} \nabla_1^1 \ln x_t$$

Devido a  $h_t$  depender do período atrasado  $\varepsilon_{t-1}^2$ , a série  $\varepsilon_t$  é então chamada de modelo ARCH de ordem 1. A expressão (2.22) indica que a variância condicional de  $\varepsilon_t$  é variante no tempo e é importante notar que não há um termo de erro adicional. A equação (2.21) assume que todas as observações  $\eta_t$  têm as mesmas propriedades distribucionais.

Para analisar as propriedades dos dados do tipo ARCH é conveniente escrever (2.20) e (2.22) como:

$$\varepsilon_t = \eta_t (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^{1/2} \quad (2.23)$$

$$\eta_t \sim NID(0;1) \quad (2.24)$$

Uma vez que  $E[\eta_t] = 0$ , sendo que  $E_t[\cdot]$  denota o operador esperança, pode-se verificar que a esperança condicionada de  $\varepsilon_t$  é igual a:

$$E_t[\varepsilon_t] = E[\sqrt{h_t} E[\eta_t]] = 0 \quad (2.25)$$

Para a variância condicional de  $\varepsilon_t$  tem-se da equação (2.23) que:

$$E[\varepsilon_t^2] = h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \quad (2.26)$$

Usando a expressão (2.23) tem-se que a variância condicional de  $\varepsilon_t$  é igual a:

$$E_t[\varepsilon_t^2] = E[h_t] = \alpha_0 + \alpha_1 E[\varepsilon_{t-1}^2] \quad (2.27)$$

Esta expressão mostra que  $\varepsilon_t^2$  imita um processo AR(1) com parâmetro  $\alpha_1$ , ou seja, a série temporal quadrática apresenta autocorrelação. Pode-se verificar, dadas as definições anteriores, que esta série não é estritamente um ruído branco. Uma vez que,  $\varepsilon_t$  pode ser uma série residual de um modelo ARMA para  $x_t$ , isto indica a diferença-chave entre séries temporais ARCH e aquelas que não são ARCH. Com  $0 < \alpha_1 < 1$ , a equação (2.27) pode ser resolvida como:

$$E[\varepsilon_t^2] = \frac{\alpha_0}{(1-\alpha_1)} \quad (2.28)$$

A expressão em (2.22) mostra que valores absolutos grandes (ou pequenos) de  $\varepsilon_t$  são esperados ser seguidos por valores absolutos grandes (ou pequenos), enquanto houver a igualdade:

$$E[\varepsilon_t \varepsilon_{t-h}] = 0 \quad (2.29)$$

ou seja, a série  $\varepsilon_t$  é não correlacionada. Conseqüentemente, um modelo ARCH pode descrever uma série temporal com seqüências de dados pontuais que parecem com outliers, onde o fato de que estes outliers aparecem em clusters é causado pela equação de variância e não pelas autocorrelações relevantes no nível da série temporal.

#### 2.2.4.2 Modelos GARCH

O modelo ARCH generalizado, conhecido como GARCH- *Generalized* ARCH foi primeiramente proposto por Bollerslev (1986).

Dado que um modelo AR(p)-ARCH(q) pode ser representado por:

$$\varepsilon_t = (h_t)^{1/2} \eta_t \quad (2.30)$$

$$\eta_t \sim NID(0;1) \quad (2.31)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 \quad (2.32)$$

com  $\alpha_0 > 0, \alpha_i \geq 0$  ( $i=1,2,\dots,q$ ) e  $\phi_p(1) > 0$  e  $\sum_{i=1}^q \alpha_i < 1$ . Para muitas séries temporais financeiras, o valor de  $q$  em (2.32) pode assumir inconvenientemente valores grandes de tal forma que seja necessário estimar muitos parâmetros. Além disso, pode ser inconveniente impor as restrições  $\alpha_i \geq 0$  para todos os  $i$  em (2.32). Portanto, pode ser útil aproximar o polinômio de  $q$ -ésima ordem em (2.32) por um polinômio de ordem  $(p,q)$  similar a um modelo padrão ARMA para uma série  $y_t$ .

O modelo GARCH é então expresso por:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 - \beta_1 h_{t-1} - \dots - \beta_p h_{t-p} \quad (2.33)$$

ou seja, é a equação GARCH(p,q), onde  $\alpha_0, \alpha_1, \dots, \alpha_q$  e  $\beta_1, \dots, \beta_p$  excede zero e  $\sum_{j=1}^q \alpha_j + \sum_{i=1}^p \beta_i < 1$ . Na prática, o valor de  $q$  em (2.33) é muito menor que em (2.32) e é típico encontrar que o modelo GARCH(1,1) produz uma descrição adequada de muitas

séries temporais financeiras (Bollerslev, Chou e Kroner/1992). Expressões explícitas para a curtose e as autocorrelações de uma série  $\varepsilon_t$  gerada pelas equações de (2.30) a (2.32) podem ser geradas mostrando que um modelo GARCH(p,q) pode descrever séries temporais com seqüências de grandes observações tanto positivas como negativas. A função de autocorrelação de  $\varepsilon_t^2$  pode ser encontrada aplicando as mesmas técnicas descritas anteriormente (Bollerslev/1988). Por exemplo, a equação GARCH(1,1):

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (2.34)$$

pode ser expressa como:

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 + \nu_t - \beta_1 \nu_{t-1} \quad (2.35)$$

onde  $\nu_t = \varepsilon_t^2 - h_t$ . Uma vez que a série  $\nu_t$  é não correlacionada com o seu próprio passado, a equação (2.35) indica que o modelo GARCH(1,1) implica em que a FAC de  $\varepsilon_t^2$  parece com a FAC de um modelo ARMA(1,1). Sendo que quando  $\alpha_1 = 0$ , o parâmetro  $\beta_1$  em (2.35) é não determinado. Conseqüentemente, para qualquer processo GARCH(p,q) o valor de q deverá ser no mínimo igual a 1.

Uma das motivações para considerar modelos do tipo GARCH é que tais modelos permitem a previsão de volatilidade condicional. Por exemplo, com um modelo GARCH(1,1), pode ser gerada a previsão de um passo à frente para  $h_t$  (no tempo n) como:

$$h_{n+1} = \alpha_0 + \alpha_1 \varepsilon_n^2 + \beta_1 h_n \quad (2.36)$$

Para comparar a performance de previsão para amostras de vários modelos GARCH rivais deve-se usar:

$$v_{n+1} = (x_{n+1} - \bar{x})^2 \quad (2.37)$$

como uma medida de verdade, mas não observada, volatilidade, onde  $\bar{x}$  é alguma média de um intervalo longo selecionado a priori de retornos passados (Day e Lewis, 1992) e (Pagan e Swert, 1990).

Nesse caso o erro do modelo ARIMA(p, 1, q) pode ser escrito como

$$\varepsilon_t^{ARIMA(p,1,q)} = \frac{\phi_q(B)}{\theta_p(B)} \nabla_1 \ln x_t, \text{ o erro do modelo GARCH é dado por } \varepsilon_t = \eta_t \sqrt{\text{Var}_{t-1}[\varepsilon_t]}, \text{ portanto}$$

podemos escrever que  $\text{Var}_{t-1} \varepsilon_t = \frac{(\varepsilon_t^{ARIMA(p,1,q)})^2}{\eta_t^2}$ . Outra importante expressão que ser

construída para os modelos GARCH em que se visualizam os principais parâmetros do

modelo é  $\frac{\phi_p^m(B)}{\theta_q^m(B)} \nabla_1^d \ln x_t = \eta_t \sqrt{\frac{\alpha_q^v(B)}{\beta_p^v(B)}} \varepsilon_t$ , isolando  $x_t$  temos o valor da série no nível:

$$x_t = \left( \frac{\theta_q^m(B)}{\phi_p^m(B)} \right) \eta_t \sqrt{\frac{\alpha_q^v(B)}{\beta_p^v(B)}} \varepsilon_t$$

#### 2.2.4.2.1 Modelo GARCH-M

Existem muitos tipos diferentes de formas funcionais não-lineares para serem utilizadas como alternativa aos modelos lineares. Muitos modelos não-lineares tentam capturar os processos não-lineares “verdadeiros” ou “subjacentes” através de “suposições paramétricas” com formas funcionais não-lineares específicas. Um exemplo desta abordagem é o modelo GARCH-in-Mean ou GARCH-M. Nesta abordagem a variância do termo de distúrbio afeta diretamente a média da variável dependente e evolui através do tempo como uma função do seu próprio valor passado e do passado do erro quadrático de previsão (Engle, Lilien e Robins, 1987). Por esta razão, a variância variando no tempo é chamada de *variância condicionada*. As seguintes equações descrevem um típico modelo paramétrico GARCH-M:

$$\sigma_t^2 = \delta_0 + \delta_1 \sigma_{t-1}^2 + \delta_2 \varepsilon_{t-1}^2 \quad (2.38)$$

$$\varepsilon_t^2 \sim N(0; \sigma_t^2)$$

$$y_t = \alpha + \beta \sigma_t + \varepsilon_t \quad (2.39)$$

sendo que  $y$  é a taxa de retorno sobre um ativo,  $\alpha$  é a taxa esperada de avaliação e  $\varepsilon_t$  é um termo de distúrbio normalmente distribuído com média zero e *variância condicionada*  $\sigma_t^2$ . O parâmetro  $\beta$  representa o efeito do *prêmio pelo risco* sobre o retorno do ativo, enquanto que os parâmetros  $\delta_0$ ,  $\delta_1$  e  $\delta_2$  definem a evolução da variância condicional.

O modelo GARCH-M é um sistema recursivo estocástico, dadas as condições iniciais  $\sigma_0^2$  e  $\varepsilon_0^2$ , bem como as estimativas para  $\alpha$ ,  $\beta$ ,  $\delta_0$ ,  $\delta_1$  e  $\delta_2$ . Uma vez que a variância condicional é dada, o choque aleatório é “arrancado” de uma distribuição normal, e o retorno do ativo é

completamente determinado como uma função de sua própria média, o choque aleatório e o efeito do prêmio pelo risco determinados por  $\beta\sigma_t$ .

Uma vez que a distribuição do choque é “normal “ pode-se utilizar a estimação de máxima verosimilhança para aproximar as estimativas para  $\alpha$ ,  $\beta$ ,  $\delta_0$ ,  $\delta_1$  e  $\delta_2$ . A função de verosimilhança  $L$  é uma função de probabilidade conjunta para  $\hat{y}_t = y_t$ , para  $t=1,2,\dots,T$ .

Para o modelo GARCH-M tem-se a seguinte forma:

$$L_t = \prod_{i=1}^T \frac{1}{2\pi\sigma_i} \exp\left[-\frac{(y_i - \hat{y}_i)^2}{2\hat{\sigma}_i^2}\right] \quad (2.40)$$

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}\hat{\sigma}_t$$

$$\hat{\varepsilon}_t = y_t - \hat{y}_t$$

$$\hat{\sigma}_t^2 = \hat{\delta}_0 + \hat{\delta}_1\hat{\sigma}_{t-1}^2 + \hat{\delta}_2\hat{\varepsilon}_{t-1}^2$$

sendo que os símbolos  $\hat{\alpha}$ ,  $\hat{\beta}$ ,  $\hat{\delta}_0$ ,  $\hat{\delta}_1$  e  $\hat{\delta}_2$  são estimadores dos parâmetros. Este método para obter as estimativas dos parâmetros maximiza a soma do logaritmo da função de verosimilhança sobre toda a amostra de tempo T, t = 1 até t = T. com respeito a escolha da estimativa dos coeficientes, sujeita a restrição de que a variância é maior que zero, dada a condição inicial  $\hat{\sigma}_0^2$  e  $\hat{\varepsilon}_{t-1}^2$ .

A tendência para o uso da abordagem GARCH-M é que ela remete à origem da não-linearidade do processo. A variância condicional é uma transformação não-linear dos valores passados, do mesmo modo que a medida de variância é uma transformação não-linear da predição dos erros passados. A justificativa do uso da *variância condicional* como uma variável afetando a variável dependente é que é que a *variância condicional* representa um “fator de risco” bem entendido o qual eleva a “taxa de retorno exigida” quando tem-se uma previsão dinâmica do preço do ativo.

Uma das maiores desvantagens do método GARCH-M é a minimização das funções do logaritmo da verosimilhança são freqüentemente difíceis de se alcançar. Se estivermos interessados na significância dos coeficientes estimados  $\hat{\alpha}$ ,  $\hat{\beta}$ ,  $\hat{\delta}_0$ ,  $\hat{\delta}_1$  e  $\hat{\delta}_2$  será difícil obter estimativas para os intervalos de confiança. Todas estas dificuldades são comuns nas abordagens por máxima verosimilhança para estimação de parâmetros.

A abordagem paramétrica GARCH-M para a especificação de processos não-lineares é, portanto, restritiva: tem-se um conjunto específico de parâmetros que se quer estimar, os quais tem uma interpretação e um significado muito bem definidos, sabe-se como estimar estes parâmetros e até mesmo se há alguma dificuldade.

A vantagem dos modelos GARCH-M é que eles capturam fenômenos bem observados em séries temporais financeiras, onde períodos de alta volatilidade são seguidos por alta volatilidade e períodos de baixa volatilidade são seguidos por períodos similares.

As limitações do modelo GARCH-M também são suas desvantagens: ficá-se limitado a conjuntos de parâmetros bem definidos, distribuições bem definidas, formas funcionais não-lineares específicas e a um método de estimação o qual nem sempre converge para parâmetros estimados que fazem algum sentido. A conclusão é que utilizando-se modelos não-lineares específicos, pode-se, portanto, faltar flexibilidade para se ajustar determinados processos não-lineares alternativos.

Com uma rede neural, utilizada como método de “aproximação”, pode-se aproximar processos não-lineares desconhecidos. Não há limite no número de parâmetros e nem, infelizmente, eles têm uma interpretação direta tal como nos modelos GARCH-M.

## **2.3 Metodologia de Box e Jenkins**

### **2.3.1 Identificação**

A metodologia de Box e Jenkins consiste basicamente em quatro estágios e é utilizada para propósitos de estimação e previsão de séries temporais univariadas (Enders, 1995, pg. 63). No estágio de identificação, o pesquisador examina visualmente os gráficos da série temporal, as funções de auto-correlação e correlação parcial. A partir do gráfico construído com cada observação da seqüência  $y_t$  é possível verificar a presença de valores discrepantes (*outliers*), valores faltantes e quebra estrutural nos dados. A aplicação da primeira diferença também pode ser realizada após a análise gráfica. As variáveis não-estacionárias podem ter uma tendência pronunciada ou desviar de forma não constante da média e da variância. Valores faltantes e *outliers* podem ser corrigidos neste ponto da análise. O modelo ARMA tem uma característica importante que o torna distinto de muitos outros modelos

econométricos. Este modelo é adequado para descrever determinadas séries temporais que são reconhecidas por características específicas de seus dados. Estas características são chamadas de auto-correlação (FAC) e auto-correlação parcial (FACP). Na fase de identificação a comparação entre amostras de FAC e FACP de vários processos ARMA podem sugerir muitos modelos plausíveis.

### 2.3.1.1 Função de Auto-correlação

A função de auto-correlação FAC é definida pela seguinte expressão:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (2.41)$$

sendo que  $\gamma_k$  é a k-ésima auto-covariância de  $y_t$ , ou seja:

$$\gamma_k = E[(y_t - \mu)(y_{t-k} - \mu)] \quad k = \dots, -2, -1, 0, 1, 2, \dots \quad (2.42)$$

Dada a equação (2.42) é fácil verificar que para as auto-correlações isto implica que  $\rho_0 = 1, \rho_{-k} = \rho_k$  e que  $-1 < \rho_k < 1$ .

A função de auto-correlação é útil para caracterizar modelos de séries temporais ARMA. Um simples exemplo é a série ruído branco  $\varepsilon_t$  para a qual  $E[\varepsilon_t] = 0$  e  $\rho_k = 0$  para todo  $k \neq 0$ . Para o modelo AR(1):

$$y_t - \mu = \phi_1 (y_{t-1} - \mu) + \varepsilon_t, \quad t = 2, 3, \dots, n \quad (2.43)$$

nós podemos derivar que:

$$E(y_t) = \mu + \phi_1 E(y_{t-1} - \mu) + E(\varepsilon_t) = (1 - \phi_1)\mu + \phi_1 E(y_{t-1}) \quad (2.44)$$

Quando  $E(y_{t-i}) = E(y_t)$ , que é o caso no qual as séries temporais correspondem aos modelos ARMA sendo que não há componente  $(1 - B)$  na parte AR, implicando em  $|\phi_1| < 1$ , a equação (2.44) pode ser escrita como:

$$E(y_t) = (1 - L)^{-1} (1 - \phi_1)\mu = \mu \quad (2.45)$$

Para calcular a ACF, começamos com:

$$\gamma_0 = E[(y_t - E(y_t))(y_t - E(y_t))] \quad (2.46)$$

Para o modelo AR(1), o lado direito da equação (2.46) é:

$$E[(y_t - \mu)(y_t - \mu)] = E[\phi_1 (y_{t-1} - \mu)\phi_1 (y_{t-1} - \mu)] + E[\varepsilon_t^2] + 2E[\phi_1 (y_{t-1} - \mu)\varepsilon_t] \quad (2.47)$$

A covariância de  $\mu$  com a série temporal é logicamente igual a zero. Considerando que o modelo AR(1) pode ser escrito como:

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + y_0 \quad (2.48)$$

onde os parâmetros são escalonados por  $\theta_0$ , na versão para um período de atraso temos que:

$$y_{t-1} = \varepsilon_{t-1} + \theta_1 \varepsilon_{t-2} + \theta_2 \varepsilon_{t-3} + \dots + y_0 \quad (2.49)$$

fica claro que  $E[y_{t-1}\varepsilon_t] = 0$ . Movendo a equação (2.49) cada vez mais atrás no tempo segue que  $E[y_{t-j}\varepsilon_t] = 0$  para qualquer  $j$  discreto maior que zero. A partir das equações

(2.48) e (2.49) fica evidente que  $E[y_{t-j}\varepsilon_{t-j}] = E[\varepsilon_t^2] = \sigma^2$ . Com estes resultados na equação (2.47) tem-se:

$$\gamma_0 = (1 - \phi_1^2)^{-1} \sigma^2, \quad \text{sujeito a condição } |\phi_1| < 1 \quad (2.50)$$

A auto-covariância de primeira ordem para uma série temporal com modelo AR(1) é:

$$\gamma_1 = E[(y_t - \mu)(y_{t-1} - \mu)] = E[\phi_1 (y_{t-1} - \mu)(y_{t-1} - \mu)] + E[\varepsilon_t (y_{t-1} - \mu)] = \phi_1 \gamma_0 \quad (2.51)$$

Conseqüentemente, o coeficiente de correlação  $\rho_1$  para o modelo AR(1) torna-se simplesmente:

$$\rho_1 = \gamma_1 / \gamma_0 = \phi_1 \quad (2.52)$$

Para calcular  $\rho_k$  é conveniente considerar a seguinte expressão para o modelo AR(1):

$$E[(y_t - \mu)(y_{t-k} - \mu)] = E[\phi_1 (y_{t-1} - \mu)(y_{t-1} - \mu)] \quad (2.53)$$

considerando o fato que  $E[y_{t-k}\varepsilon_t] = 0$ . Dividindo ambos os lados por  $\gamma_0$  resulta que:

$$\rho_k = \phi_1 \rho_{k-1}, \quad \text{para } k = 1, 2, 3, \dots \quad (2.54)$$

Por exemplo, quando  $\phi_1 = 0,8$  as primeira quatro auto-correlações são dadas por 0,8, 0,64, 0,512 e 0,4096. Na prática nós podemos determinar tais correlações para dados reais, e verificar se o padrão se encaixa com esta seqüência. Se isso acontecer, podemos então considerar u modelo AR(1) para realizar previsão.

Em princípio, a determinação dos coeficientes de auto-correlação para modelos auto-regressivos de ordens maiores seguem um procedimento similar ao que foi mostrado no caso do modelo AR(1). Por exemplo, considerando o modelo AR(2):

$$y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} = \varepsilon_t \quad (2.55)$$

multiplicando ambos os lados por  $y_{t-1}$ , tomando as esperanças e dividindo por  $\gamma_0$  resulta em:

$$\rho_1 - \phi_1 \rho_0 - \phi_2 \rho_1 = 0 \quad (2.56)$$

e sendo  $\rho_0 = 1$ , obtém-se:

$$\rho_1 = \frac{\phi_1}{1 - \phi_2} \quad (2.57)$$

para determinar um expressão para  $\rho_2$ , operações análogas são aplicadas sobre a equação (2.55), produzindo:

$$\rho_2 - \phi_1 \rho_1 - \phi_2 \rho_0 = 0 \quad (2.58)$$

Substituindo (2.57) em (2.58) encontra-se:

$$\rho_2 = \frac{\phi_1^2}{1 - \phi_2} + \phi_2 \quad (2.59)$$

Portanto, analogamente as equações (2.56) e (2.58) tem-se:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \quad \text{para } k = 2, 3, 4, \dots \quad (2.60)$$

Para encontrar expressões para a FAC para modelos AR(p) com  $p > 2$ , nós utilizamos a mesma técnica mostrada acima. Em geral isto implica que a FAC de um processo AR mostra um padrão de decaimento exponencial.

A FAC é mais utilizada para a determinação de modelos MA(q) (Shummway e Stoffer, 2000).

Considere, por exemplo, o modelo MA(2):

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \quad (2.61)$$

onde para a versão com k atrasos tem-se que:

$$y_{t-k} = \varepsilon_{t-k} + \theta_1 \varepsilon_{t-k-1} - \theta_2 \varepsilon_{t-k-2} \quad (2.62)$$

A variância  $\gamma_0$  é igual a:

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2) \sigma^2 \quad (2.63)$$

respeitando a condição que todas as covariâncias entre  $\varepsilon_t$  e os seus atrasos são iguais a zero. Com as equações (2.61) e (2.62) podemos verificar que:

$$\gamma_1 = E(y_t y_{t-1}) = (\theta_1 + \theta_1 \theta_2) \sigma^2 \quad (2.64)$$

$$\gamma_2 = E(y_t y_{t-2}) = \theta_2 \sigma^2 \quad (2.65)$$

$$\gamma_k = 0, \text{ para } k = 3, 4, \dots \quad (2.66)$$

conseqüentemente,  $\rho_k = 0$  para  $k = 3, 4, \dots$

Isto implica que quando na prática a FAC estimada é disponível, sendo que todos os valores são zero depois do q-ésimo lag, nós poderemos decidir analisar um modelo MA(q) para  $y_t$ . Isto segue do fato de que para um modelo MA(q) tem-se:

$$\gamma_k = \begin{cases} \left[ \sum_{i=0}^{q-k} \theta_i \theta_{i+k} \right] \sigma^2 & \text{para } k = 0, 1, \dots, q \\ 0 & \text{para } k > q \end{cases} \quad (2.67)$$

Para modelo ARMA(p,1) o padrão da FAC é uma mistura do padrão da FAC para modelos AR e MA puros. Por exemplo, considerando o modelo ARMA(1,1):

$$y_t = \phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} \quad (2.68)$$

pode-se derivar que:

$$\gamma_0 = \phi_1 \gamma_1 + \sigma^2 + \theta_1 E(y_t \varepsilon_{t-1}) = \phi_1 \gamma_1 + [1 + \theta_1 (\phi_1 + \theta_1)] \sigma^2$$

$$\gamma_1 = \phi_1 \gamma_0 + \theta_1 \sigma^2$$

$$\gamma_1 = \phi_1 \gamma_1$$

$$\gamma_k = \phi_1 \gamma_{k-1}$$

sendo  $k = 3, 4, 5, \dots$ , de tal forma que após alguma álgebra a função de correlação resulta em:

$$\rho_k = \frac{\phi_1^{k-1} (1 + \phi_1 \theta_1) (\phi_1 + \theta_1)}{(1 + \phi_1 \theta_1 + \theta_1^2)}, \text{ para } k = 1, 2, 3, \dots \quad (2.69)$$

Da expressão (2.69) pode-se verificar que  $\rho_k$  pode assumir uma ampla variedade de valores para escolhas distintas de  $\phi_1$  e  $\theta_1$ . Isto implica que a identificação de um modelo de série temporal ARMA a partir somente da análise dos padrões de comportamento da FAC é uma tarefa difícil.

### 2.3.1.2 Função de Auto-correlação Parcial

A FAC é utilizada para identificar se um modelo MA (de uma ordem qualquer) pode descrever  $y_t$ , mas é menos útil para identificar modelos AR.

A razão para isso é que, por exemplo para um modelo AR(1) tem-se:

$$y_t = \phi_1 y_{t-1} + \varepsilon_t \quad (2.70)$$

o qual pode ser escrito como:

$$y_t = \phi_1^2 y_{t-2} + \varepsilon_t + \phi_1 \varepsilon_{t-1} \quad (2.71)$$

a inclusão de  $y_{t-1}$  no modelo de regressão para  $y_t$ , também faz com que  $y_t$  dependa de  $y_{t-2}$  (sendo  $\phi_1^2 < \phi_1$ ) como pode ser observado na FAC da equação (2.54). É importante notar que para identificar um modelo AR, adicionando  $y_{t-2}$  à regressão em (2.70), poderá não auxiliar para explicar  $y_t$ , ou seja o parâmetro correspondente deverá ser igual a zero. Isso nos leva a construção da função de auto-correlação parcial (FACP). O valor da FACP no lag 1, digamos  $\psi_1$ , é dado por:

$$y_t = \psi_1 y_{t-1} + u_t \quad (2.72)$$

onde  $u_t$  é apenas uma série temporal de erro ruído branco quando o modelo para  $y_t$  é de fato um AR(1). Da equação (2.72) segue que sendo  $\psi_1$  igual a  $\gamma_1/\gamma_0$ , por construção nós temos que  $\psi_1 = \rho_1$  para todos os modelos de série temporal. O segundo valor da FACP resulta do modelo de regressão:

$$y_t = \eta_1 y_{t-1} + \psi_2 y_{t-2} + u_t \quad (2.73)$$

No caso de um AR(1), o  $\psi_2$  é igual zero. No caso de um AR(2) ou de ordem superior,  $\psi_2$  é diferente de zero. Para um modelo AR(2), acontece que  $\psi_3 = 0$  na regressão:

$$y_t = \eta_1 y_{t-1} + \eta_2 y_{t-2} + \psi_3 y_{t-3} + u_t \quad (2.74)$$

conseqüentemente, quando  $\psi_{p+1}$  é igual a zero, quando  $\psi_p$  não é, pode-se então considerar o modelo como sendo um AR de ordem p.

Uma outra forma de determinar a FACP é através das equações de Yule-Walker . Sendo  $\phi_{11}$  tanto a auto-correlação como a auto-correlação parcial entre  $y_t$  e  $y_{t-1}$ , teremos que a partir das auto-correlações podemos formar as auto-correlações parciais da seguinte forma:

$$\phi_{11} = \rho_1 \quad (2.75)$$

$$\phi_{22} = (\rho_2 - \rho_1^2) / (1 - \rho_1^2) \quad (2.76)$$

para lags adicionais temos uma recursão dada por:

$$\phi_{kk} = \left( \rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_{k-j} \right) \cdot \left( 1 - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_j \right)^{-1}, \text{ sendo } k=3,4,5,\dots \quad (2.77)$$

### 2.3.2 Estimação

A estratégia de especificação de uma série temporal ARMA começa com a inspeção dos valores da FAC e da FACP estimadas, as quais são dadas por:

$$\hat{\rho}_k = C_k / C_0 \quad (2.78)$$

sendo que:

$$C_k = \sum_{t=k+1}^n (y_t - \bar{y})(y_{t-k} - \bar{y}) \quad (2.79)$$

onde  $\bar{y}$  denota a média amostral relevante de  $y_t$ ,  $t = 1, 2, 3, \dots, n$ . Dessa forma,  $\hat{\rho}_k$  para  $k = 0, 1, 2, \dots$  forma a FAC estimada.

A FACP pode ser obtida aplicando-se mínimos quadrados ordinários em:

$$y_t - \bar{y} = \hat{\psi}_1 (y_{t-1} - \bar{y}) + \dots + \hat{\psi}_k (y_{t-k} - \bar{y}) + \nu_t \quad (2.80)$$

para quaisquer valores de k, onde  $\nu_t$  não é necessariamente uma série temporal ruído branco.

Para checar quais valores são significantes de tal forma que estruturas de modelos ARMA simples e razoáveis possam ser levados como hipótese, devemos estimar os parâmetros dos vários modelos e investigar se os resíduos estimados podem ser considerados

aproximadamente como ruído branco. Esta estratégia pode ser descrita como sendo de identificação, estimação e modificação.

Vamos analisar a estimação de modelos AR.

Os parâmetros no modelo AR( $p$ ), onde a escolha de  $p$  está baseada nos valores estimados da FACP que são significantes, dados por:

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad t = p+1, p+2, \dots, n \quad (2.81)$$

podem ser estimados pelo método dos mínimos quadrados ordinários (MQO), onde as observações de  $y_1$  até  $y_p$  são assumidas como sendo os valores iniciais. Pode ser mostrado que os estimadores de MQO dos parâmetros são consistentes e assintoticamente normal e a estatística  $t$  padrão pode ser utilizada para investigar a significância.

Quanto aos modelos ARMA existe uma grande variedade de rotinas de estimação para determiná-los. A principal razão para isto é a importante característica de que as variáveis  $\varepsilon_t$  atrasadas na parte MA do modelo são variáveis não observadas e conseqüentemente têm que ser estimadas. Por exemplo, para o modelo ARMA(1,1):

$$y_t = \phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} \quad (2.82)$$

os parâmetros  $\phi_1$  e  $\theta_1$  não são conhecidos e a variável  $\varepsilon_{t-1}$  também não é conhecida. O modelo (2.82) pode ser escrito como:

$$(1 + \theta_1 B)^{-1} y_t = \phi_1 (1 + \theta_1 B)^{-1} y_{t-1} + \varepsilon_t \quad (2.83)$$

Denotando  $z_t = (1 + \theta_1 B)^{-1} y_t$  tal que:

$$z_t = y_t - \theta_1 y_{t-1} + \theta_1^2 y_{t-2} - \theta_1^3 y_{t-3} + \dots \quad (2.84)$$

tem-se as observações (assumindo que  $y_0 = 0$ ):

$$z_1 = y_1$$

$$z_2 = y_2 - \theta_1 y_1$$

$$z_3 = y_3 - \theta_1 y_2 + \theta_1^2 y_1$$

...

e assim por diante. Atribuindo um valor para  $\theta_1$ , pode-se construir a série temporal  $z_t$  e estimar  $\phi_1$  via MQO aplicados em (2.83). Esta regressão fornece a série  $\hat{\varepsilon}_t$  estimada, a qual

quando atribuímos  $\varepsilon_1 = 0$  pode ser usada em (2.82) para atribuir um valor para  $\theta_1$  em um segundo passo. Uma rotina de otimização comumente aplicada que produz os estimadores finais para os parâmetros ARMA em um número não muito elevado de iterações é a rotina de Gauss-Newton.

### 2.3.3 Diagnóstico

Um requisito óbvio para um modelo de série temporal ARMA é que a série temporal de resíduo estimado seja aproximadamente um ruído branco. Se não for, pode-se não ter alcançado a estrutura dinâmica em  $y_t$  que incorpore um modelo ARMA. A FAC estimada dos resíduos é dada por:

$$r_k(\hat{\varepsilon}) = \frac{\left[ \sum_{i=k+1}^n \hat{\varepsilon}_i \hat{\varepsilon}_{i-k} \right]}{\left[ \sum_{i=1}^n \hat{\varepsilon}_i^2 \right]} \quad (2.85)$$

para  $k = 1, 2, 3, \dots$ . Dado um modelo adequado, pode-se mostrar que as populações equivalentes de  $r_k(\hat{\varepsilon})$  são assintoticamente não-correlacionados e têm variâncias que podem ser aproximadas por  $(n-k)/(n^2 + 2n) \cong n^{-1}$ . Conseqüentemente, sob a suposição adicional de normalidade, uma checagem ao nível de confiança de 95% pode testar se as auto-correlações residuais estimadas estão no intervalo  $\pm 2n^{-\frac{1}{2}}$ . Ljung e Box (1978) propõem um teste conjunto para a significância das  $m$  primeiras auto-correlações residuais, a qual é dada por:

$$LB(m) = n(n+2) \sum_{k=1}^m (n-k)^{-1} r_k^2(\hat{\varepsilon}) \quad (2.86)$$

a qual segue assintoticamente uma distribuição  $\chi^2(m-p-q)$  sob as hipóteses de não haver auto-correlação residual, sendo que  $m/n$  é um valor pequeno e  $m$  é moderadamente grande ( $>30$ ).

Realizando os estágios de identificação, estimação e diagnóstico resulta em um conjunto de modelos, estes modelos não podem ser rejeitados usando as medidas de diagnóstico mencionadas acima então será necessário adotar alguns critérios de seleção.

Um estudo sobre critérios de seleção de modelos pode ser encontrado em De Gooijer *et al.* (1985).

É sensato assumir que nenhum modelo pode ser preferível *a priori* e que conseqüentemente os modelos devem ser tratados simetricamente (Granger, King e White, 1995).

Em geral, isto implica que o modelo final selecionado é aquele que minimiza o valor de uma certa função de critério.

O coeficiente padrão de determinação  $R^2$  é pouco útil para avaliar modelos de séries temporais. Por exemplo, para o modelo AR(1)  $y_t = \phi_1 y_{t-1} + \varepsilon_t$ , teoricamente segue que:

$$\begin{aligned} R^2 &= 1 - \sigma^2 / \gamma_0 \\ &= 1 - \sigma^2 / \left[ \sigma^2 / (1 - \phi_1^2) \right] \\ &= \phi_1^2 \end{aligned} \tag{2.87}$$

Ou seja, o  $R^2$  depende somente do parâmetro AR. Esta expressão mostra que o  $R^2$  no caso em  $\phi_1 = 0,9$  é muito maior que no caso onde  $\phi_1 = 0,2$ , quando ambos parecem ser adequados (Nelson, 1976). Um critério  $R^2$  modificado pode ser estudado em Harvey (1989).

Dois critérios freqüentemente utilizados para selecionar modelos de séries temporais são o critério de informação Akaike (1974) e Schwarz (1978). Ambos os critérios avaliam o ajuste versus o número de parâmetros. Quando  $n$  denota o número de observações efetivas (as quais são observações necessárias para estimar os parâmetros) e  $k$  denota o número de parâmetros ARMA a serem estimados, o critério de informação de Akaike é dado por:

$$AIC(k) = n \log \hat{\sigma}_{ML}^2 + 2k$$

onde  $\hat{\sigma}_{ML}^2 = RSS/n$  e RSS é a soma dos resíduos quadráticos (*residual sum of squares*). O valor de  $k$  que minimiza  $AIC(k)$  é selecionado. A mesma regra de decisão aplica-se para o critério de Schwarz, o qual é dado por:

$$SIC(k) = n \log \hat{\sigma}_{ML}^2 + k \log n$$

Comparando as expressões para AIC e SIC, quando  $n \geq 8$ , o critério SIC penaliza a inclusão de regressores (e portanto, de parâmetros adicionais) mais que o AIC. Isto significa que a ordem do modelo selecionado com o critério SIC é usualmente menor que a ordem do modelo selecionado com o AIC, ou seja o critério SIC sempre seleciona um modelo mais parcimonioso que o AIC (Enders, 1995, pg. 88).

### 2.3.4 Previsão

Uma vez que um ou mais modelos de séries temporais foram selecionados devemos gerar previsões 1-passo ou h-passos à frente de previsão para  $y_t$  (onde  $y_t$  denota um variável diferenciada adequadamente). As previsões referem-se as observações  $y_{n+h}$  e conseqüentemente será denotada por  $\hat{y}_{n+h}$ , com  $h = 1, 2, 3, \dots, m$ . Em geral, a previsão um-passo à frente é definida por  $\hat{y}_{n+1}$ , a qual é baseada no conjunto de informação  $y_n$ . No caso onde  $m$  observações foram registradas para avaliar uma previsão um-passo à frente, avança-se o conjunto de informação com cada previsão. Por exemplo,  $\hat{y}_{n+2}$  é baseado em  $y_{n+1}$  e  $\hat{y}_{n+3}$  é baseado em  $y_{n+2}$ . A motivação intuitiva por trás dessa lógica é não é sempre possível ser capaz de re-estimar os parâmetros do modelo de séries temporais. Neste caso, são mantidos os estimadores para os parâmetros fixados para a amostra  $t = 1, 2, \dots, n$  enquanto faz-se previsões um-passo à frente para  $y_{n+1}$  até  $y_{n+m}$ . Se possível, podemos também decidir estimar os parâmetros para cada um dos conjuntos de informação  $y_{n+h-1}$ . Além disso, podemos então *deletar* os primeiros  $h$  dados pontuais, de tal forma que a amostra de estimação tem o mesmo tamanho para as  $m$  previsões. Finalmente, uma previsão h-passos (ou dinâmica) denota uma previsão de  $y_{n+h}$  (para  $h = 1, 2, \dots, m$ ) a qual é baseada em  $Y_n$ . Obviamente, quando  $h$  é maior que a memória do modelo (ou, por exemplo, o comprimento do polinômio AR), a previsão  $\hat{y}_{n+h}$  provavelmente inclui previsões para dados pontuais prévios como  $\hat{y}_{n+h-1}$  e também os erros de previsão tendem a ser maiores (Enders, 1995).

O princípio de previsão de modelos ARMA é muito simples. Considere o modelo MA(2):

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \quad t = 1, 2, \dots, n \quad (2.88)$$

para o qual é possível verificar que:

$$y_{n+1} = \varepsilon_{n+1} + \theta_1 \varepsilon_n + \theta_2 \varepsilon_{n-1} \quad (2.89)$$

No tempo  $n$ , o  $\varepsilon_{n+1}$  é desconhecido, e desde que para sua expectância no tempo  $n$  acontece que  $E_n[\varepsilon_{n+1}] = 0$ , uma previsão não-viesada de  $y_{n+1}$  é igual a:

$$\hat{y}_{n+1} = \theta_1 \varepsilon_n + \theta_2 \varepsilon_{n-1} \quad (2.90)$$

Na prática,  $\theta_1$ ,  $\theta_2$ ,  $\varepsilon_n$  e  $\varepsilon_{n-1}$  devem ser estimados, mas por conveniência assumindo que são dados, podemos comparar as expressões  $y_{n+1}$  e  $\hat{y}_{n+1}$ , com isso o erro de previsão (ou erro de predição) é dado por:

$$y_{n+1} = \hat{y}_{n+1} - \varepsilon_{n+1} \quad (2.91)$$

e conseqüentemente o erro de predição quadrático (*squared prediction error* – SPE) é  $\sigma^2$ .

Para dois passos à frente, tem-se:

$$y_{n+2} - \hat{y}_{n+2} = (\varepsilon_{n+2} + \theta_1 \varepsilon_{n+1} + \theta_2 \varepsilon_n) + (\theta_2 \varepsilon_n) \quad (2.92)$$

uma vez que no tempo  $n$ ,  $\varepsilon_{n+2}$  e  $\varepsilon_{n+1}$  não são conhecidos, para os quais o SPE é igual a  $(1 + \theta_1^2) \sigma^2$ . Para três passos à frente, temos:

$$y_{n+3} - \hat{y}_{n+3} = (\varepsilon_{n+3} + \theta_1 \varepsilon_{n+2} + \theta_2 \varepsilon_{n+1}) - (0) \quad (2.93)$$

e conseqüentemente o SPE é igual a  $(1 + \theta_1^2 + \theta_2^2) \sigma^2$ . Similarmente, para três passos à frente, não há memória no modelo MA(2) que possa auxiliar a previsão  $y_t$ , quatro passos à frente, uma vez que:

$$y_{n+4} - \hat{y}_{n+4} = (\varepsilon_{n+4} + \theta_1 \varepsilon_{n+3} + \theta_2 \varepsilon_{n+2}) - (0) \quad (2.94)$$

o qual novamente produz um SPE de  $(1 + \theta_1^2 + \theta_2^2) \sigma^2$ .

Em geral, para um modelo MA(q) a previsão h-passos à frente é igual a:

$$\hat{y}_{n+h} = \sum_{i=0}^q \theta_{i+h} \varepsilon_{n-i} \quad (2.95)$$

com  $\theta_0 = 1$  e  $\theta_{i+h} = 0$  para  $i + h > q$ , com um erro h-passos:

$$e_{n+h} = y_{n+h} - \hat{y}_{n+h} = \sum_{i=0}^{h-1} \theta_i \varepsilon_{n+h-i} \quad (2.96)$$

Dada a suposição de ruído branco sobre  $\varepsilon_t$ , segue que:

$$E(e_{n+h}) = 0 \text{ e} \quad (2.97)$$

$$SPE(h) = E[e_{n+h}^2] = \sigma^2 \sum_{i=0}^{h-1} \theta_i^2 \quad (2.98)$$

Da equação (2.74) pode ser visto que para uma série temporal com média zero que pode ser descrita como um modelo MA(q) isto implica que  $\hat{y}_{n+h} = 0$  quando  $h > q$ . Assumindo normalidade, um intervalo de previsão de 95% para  $y_{n+h}$  é limitado por  $\hat{y}_{n+h} - 1,95RSPE(h)$  e  $\hat{y}_{n+h} + 1,95RSPE(h)$ , onde RSPE é a raiz quadrada do SPE.

No modelo AR(p) acontece que  $y_t$  depende de todas as observações prévias, a previsão h-passos à frente tem propriedades similares ao que foi mostrado anteriormente. Considerando, por exemplo, o modelo AR(2):

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t \quad (2.99)$$

e a previsão um-passo à frente no tempo n, é dada por:

$$\hat{y}_{n+1} = \phi_1 y_n + \phi_2 y_{n-1} \quad (2.100)$$

e uma vez que

$$y_{n+1} = \phi_1 y_n + \phi_2 y_{n-2} + \varepsilon_{n+1} \quad (2.101)$$

o SPE para um-passo à frente, denotado por SPE(1) é  $\sigma^2$ . Para dois passos à frente, obtemos:

$$\hat{y}_{n+2} = \phi_1 \hat{y}_{n+1} + \phi_2 y_n = \phi_1 (\phi_1 y_n + \phi_2 y_{n-1}) + \phi_2 y_n \quad (2.102)$$

uma vez que:

$$y_{n+2} = \phi_1 y_{n+1} + \phi_2 y_n + \varepsilon_{n+2} = \phi_1 (\phi_1 y_n + \phi_2 y_{n-1} + \varepsilon_{n+1}) + \phi_2 y_n + \varepsilon_{n+2} \quad (2.103)$$

implica que:

$$SPE(2) = (1 + \phi_1^2) \sigma^2$$

Para três passos á frente tem-se que:

$$\hat{y}_{n+3} = \phi_1 \hat{y}_{n+2} + \phi_2 y_{n+1} = \phi_1 (\phi_1 (\phi_1 y_n + \phi_2 y_{n-1}) + \phi_2 y_n) + \phi_2 (\phi_1 y_n + \phi_2 y_{n-1}) \quad (2.104)$$

e sendo:

$$y_{n+3} = \phi_1 y_{n+2} + \phi_2 y_{n+1} + \varepsilon_{n+3} = \phi_1 (\phi_1 (\phi_1 y_n + \phi_2 y_{n-1} + \varepsilon_{n+1}) + \phi_2 y_n + \varepsilon_{n+2}) + \varepsilon_{n+3} \quad (2.105)$$

a variância do erro de previsão é:

$$SPE(3) = (1 + \phi_1^2 + \phi_2^2 + 2\phi_1^2\phi_2 + \phi_1^4) \sigma^2 \quad (2.106)$$

o qual mostra que  $SPE(3) > SPE(2)$ .

Em geral acontece que  $SPE(h) > SPE(h-1)$  para modelos AR(p) (Franses, 1998). A expressão (2.106) mostra que a expressão para os erros de previsão h-passos à frente pode ser notacionalmente pesada. Por isso é mais usual escrever um modelo AR(p) no formato MA e dessa forma usar fórmulas semelhantes a (2.98). Por exemplo, o modelo AR(2) representado por (2.99) pode ser escrito como:

$$y_t = \varepsilon_t + \eta_1 \varepsilon_{t-1} + \eta_2 \varepsilon_{t-2} + \eta_3 \varepsilon_{t-3} + \dots \quad (2.107)$$

para o qual tem-se:

$$SPE(h) = \sigma^2 \sum_{i=0}^{h-1} \eta_i^2, \text{ com } \eta_0 = 1 \quad (2.108)$$

Para o modelo AR(3), pode-se verificar que  $\eta_1 = \phi_1$  e  $\eta_2 = \phi_1^2 + \phi_2$ , as quais substituídas em (2.108) produzem o resultado apresentado em (2.106).

A previsão h-passos para frente para modelos ARMA é derivada de forma similar ao que foi realizado para os modelos AR e MA. Por exemplo, para o modelo ARMA(1,1) tem-se que (Hamilton, 1994):

$$\hat{y}_{n+1} = \phi_1 y_n + \theta_1 \varepsilon_n \quad (2.109)$$

sendo que  $SPE(1) = \sigma^2$  e:

$$\hat{y}_{n+2} = \phi_1 \hat{y}_{n+1} = \phi_1 (\phi_1 y_n + \theta_1 \varepsilon_n) \quad (2.110)$$

que comparado com:

$$y_{n+2} = \phi_1 y_{n+1} + \varepsilon_{n+2} + \theta_1 \varepsilon_{n+1} = \phi_1 (\phi_1 y_n + \varepsilon_{n+1} + \theta_1 \varepsilon_n) + \varepsilon_{n+2} + \theta_1 \varepsilon_{n+1} \quad (2.111)$$

produz  $SPE(2) = (1 + \theta_1^2 + \phi_1^2 + 2\phi_1\theta_1) \sigma^2$ . Esta última expressão também pode ser obtida escrevendo-se o modelo ARMA(1,1) como:

$$y_t = (1 - \phi_1 B)^{-1} (1 + \theta_1 B) \varepsilon_t = \varepsilon_t + \eta_1 \varepsilon_{t-1} + \eta_2 \varepsilon_{t-2} + \eta_3 \varepsilon_{t-3} + \dots \quad (2.112)$$

onde  $\eta_1 = \phi_1 + \theta_1$ .

Um outro ponto importante para se destacar na etapa de previsão concerne ao fato da realização de previsão de  $w_t$  quando o modelo é descrito pela transformação  $y_t = \log(w_t)$ .

Considerando novamente o modelo MA(2) e a sua previsão um-passo para a frente:

$$\hat{y}_{n+1} = \theta_1 \varepsilon_n + \theta_2 \varepsilon_{n-1} \quad (2.113)$$

Para prever  $w_{n+1}$  usando:

$$\hat{w}_{n+1} = \exp(\hat{y}_{n+1}) \quad (2.114)$$

é fácil verificar que  $\hat{w}_{n+1}$  é viesado para  $w_{n+1}$  uma vez que:

$$\begin{aligned} E[w_{n+1}] &= E[\exp(\varepsilon_{n+1} + \theta_1 \varepsilon_n + \theta_2 \varepsilon_{n-1})] \\ &= \exp\left(\frac{\sigma^2}{2}\right) E[\exp(\theta_1 \varepsilon_n + \theta_2 \varepsilon_{n-1})] \\ &= \exp\left(\frac{\sigma^2}{2}\right) \hat{w}_{n+1} \end{aligned} \quad (2.115)$$

Dessa forma, uma previsão não-viesada de  $w_{n+1}$  no caso de modelos utilizando logaritmos é dada por  $\exp(\sigma^2/2) \hat{w}_{n+1}$ , onde  $\hat{w}_{n+1}$  é chamada de *previsão ingênua*. Para dois passos à frente tem-se que:

$$\begin{aligned} E[w_{n+2}] &= E[\exp(\varepsilon_{n+2} + \theta_1 \varepsilon_{n+1} + \theta_2 \varepsilon_n)] \\ &= \exp\left[\frac{(1 + \theta_1^2) \sigma^2}{2}\right] E[\exp(\theta_2 \varepsilon_n)] \\ &= \exp\left[\frac{(1 + \theta_1^2) \sigma^2}{2}\right] \hat{w}_{n+2} \end{aligned} \quad (2.116)$$

Quando modelos ARMA são escritos na forma de MA, expressões apropriadas podem ser derivadas para o fator de correção das *previsões ingênuas* para  $w_{n+h}$  (Granger e Newbold/1976).

A seguir vamos discutir sobre como realizar comparação entre as previsões.

Um procedimento prático comum é manter  $m$  observações para avaliar previsões de modelos que foram ajustados para as  $n$  primeiras observações. Uma possibilidade é checar se 95% das previsões de fato encontram-se no intervalo de 95%. Se isso acontece, os modelos criados passam a ter confiança. Se não acontecer, é provável que a variância dos dados será sub-estimada.

Estudos empíricos mostram que os modelos que tendem a ser melhores para dados amostrais não necessariamente realizam previsões melhores para dados que estejam fora dessa amostra. Não há uma regra estrita para isso, mas experiências empíricas sugerem que pode ser melhor selecionar poucos modelos baseados no AIC e SIC, e avaliar estes sobre  $m$  dados desejados fora da amostra (Diebold e Mariano/1995). Esta avaliação pode ser baseada na raiz do erro de predição quadrático médio (*root mean squared prediction error* – RMSPE) também chamado de RMSE:

$$RMSPE = \sqrt{\left(\frac{1}{m}\right) \left[ \sum_{h=1}^m (\hat{y}_{n+h} - y_{n+h})^2 \right]} \quad (2.117)$$

Ou pode ser baseada no erro absoluto médio percentual (*mean absolute percentage error* – MAPE) dado por:

$$MAPE = \left(\frac{1}{m}\right) \sum_{h=1}^m \frac{|\hat{y}_{n+h} - y_{n+h}|}{y_{n+h}} \quad (2.118)$$

## Capítulo 3

# Redes Neurais

### 3.1 Características das Redes Neurais

Historicamente, o estudo das redes neurais tem progredido por várias décadas e é originado das pesquisas sobre Inteligência Artificial (IA), a qual tem o interesse de entender o cérebro e/ou criar uma máquina que mimetize suas funções. As primeiras investigações sistemáticas que tentaram definir um modelo computacional do processo de pensamento do cérebro iniciou com os trabalhos de McCulloch e Pitts, Hebb e Rosenblat entre os anos de 1940 e 1950. No entanto, o progresso subsequente foi gradual e sofreu da falta de algoritmos de treinamento eficientes, sendo que grande parte do desenvolvimento destes algoritmos foi desenvolvido na área de engenharia de controle. A renovação das atividades em redes neurais aconteceu grandemente devido ao desenvolvimento no treinamento de redes neurais com profundidade, ou camadas escondidas (também chamadas de camadas intermediárias), conhecidas como perceptrons de múltiplas camadas; MLP – *multilayer perceptron* (Rosenblatt, 1958). Este foi o resultado de pesquisas de psicologistas como Rummelhart que publicou em 1986 um novo método de treinamento, chamado de *backpropagation*

(retropropagação). A rápida disseminação desta técnica levou à uma nova onda de atividades de pesquisa. O MLP com *backpropagation* é o exemplo mais bem sucedido de uma rede neural, com a maioria das aplicações industriais (exemplo: controle e robótica) implementando esta arquitetura, a despeito de haver um grande número alternativo de arquiteturas (Azoff, 1994).

O nome “rede neural” deriva da estrutura neural do cérebro e foi inventada por biólogos e psicólogos que tentaram simular e modelar os neurônios do cérebro. Este assunto é hoje interdisciplinar, indo dos biólogos e psicólogos, os quais têm um grande entendimento sobre o processo computacional no cérebro, até físicos que têm formulado com sucesso modelos mecânico-estatísticos de redes, engenheiros interessados em projetos de filtros massivamente paralelos, matemáticos, cientistas de computação e estatísticos para os quais a fundamentação rigorosa, a lógica do raciocínio e o reconhecimento de padrão são áreas ativas de pesquisa.

Uma rede neural pode ser considerada como uma técnica de processamento de dados que mapeia, ou relaciona, algum tipo de entrada de informação à uma saída de dados. Por exemplo, a entrada pode estar na forma de uma imagem bidimensional com alguma parte faltante ou segmento distorcido e o processamento da saída pode produzir uma imagem corrigida. Nas aplicações em séries temporais, a entrada pode ser uma série temporal unidimensional e a saída a melhor estimativa para o próximo item na série (Azoff, 1994). Em geral, a tarefa executada pela rede neural pode ser dividida em quatro tipos distintos de aplicação (Duda e Hart, 1993):

1. Classificação: decidir em qual categoria uma determinada entrada deverá ser inserida;
2. Associação: atua como uma memória de conteúdo endereçável que recorda uma saída, dada alguma parte desta na forma de uma entrada;
3. Codificação: comprime e codifica uma entrada através da produção de uma saída com dimensão reduzida. Também pode ser realizada a operação inversa, ou seja, uma decodificação.
4. Simulação: a criação de uma nova saída para uma entrada que atua como um estímulo, no caso da rede ter sido exposta a uma amostra de estímulo possível.

Qualquer tarefa particular pode envolver um certo número das funções descritas acima. Por exemplo, a previsão de séries temporais por um *perceptron* pode envolver um elemento de compressão dos dados de entrada (codificação), um elemento de redução de ruído (associação), a capacidade de determinar padrões recorrentes que levem a comportamento preditível (classificação) e a previsão de um único evento (simulação).

Algumas tarefas de processamento desempenhadas pela rede neural podem também ser realizadas por técnicas tradicionais, tal como a análise de regressão (Azoff, 1994). No entanto, nenhuma compartilha das capacidades únicas da rede neural, que podem ser sumarizadas como:

- **Generalização:** a rede neural é inicialmente estabelecida a partir de uma fase de treinamento, onde entradas de exemplo são apresentadas a rede é treinada para extrair informações relevantes desses padrões. Subseqüentemente, a rede tem a capacidade de generalizar, ou seja, uma entrada ainda não vista e não utilizada também poderá ser processada;
- **Flexibilidade:** o limite de tarefas para as quais uma rede neural pode ser aplicada, excede qualquer uma das técnicas tradicionais;
- **Modelamento não-linear:** o processo de mapeamento da rede neural envolve funções não-lineares que podem conseqüentemente cobrir um limite maior da complexidade do problema. Embora outras técnicas não lineares existam, a rede neural é superior na sua generalidade e fácil implementação prática.

Existem numerosas arquiteturas de redes neurais, porém estes podem ser divididos em três grandes classes, tomando como base as técnicas usadas para treinar iterativamente os parâmetros livres na rede, conhecidos como pesos. As classes de aprendizagem são as seguintes (Haykin, 2001):

**Treinamento supervisionado:** as saídas desejadas da rede são conhecidas durante o treinamento, de tal forma que o erro, a diferença entre o valor desejado e saída real, podem ser realimentadas para a rede para melhorar o desempenho. Um exemplo dessa arquitetura é o perceptron de múltiplas camadas.

**Aprendizagem por reforço:** valores desejados para cada um dos exemplos de entrada da rede são não conhecidos (ou não usados), apenas um critério para verificar se o desempenho da rede está correto ou não é utilizado na retropropagação, nenhuma

informação referente a saída correta real está disponível para treinamento. Em alguns casos o julgamento para avaliação está disponível apenas depois que uma sequência de entradas tenha sido apresentada, a qual cria uma separação entre entradas individuais e o desempenho desejado da rede.

**Aprendizagem não supervisionada:** existe uma falta total de realimentação para dizer se o objetivo da aprendizagem foi alcançado ou não. Estes tipos de rede são altamente organizáveis, usando correlações entre os padrões de entrada para categorizá-los de acordo com as classes auto-descobertas. Um exemplo deste tipo de rede são os mapas de Kohonen auto-organizáveis.

### 3.2 Motivação para utilizar Redes Neurais na Previsão

A razão para o uso de redes neurais nesse trabalho é a previsão de um dado valor desejado ou variável de saída  $y$ , a partir de um conjunto de variáveis de entrada observadas  $x$ . Em séries temporais, o conjunto de variáveis de entrada  $x$  pode incluir variáveis atrasadas no tempo das variáveis atuais de  $x$ , bem como valores atrasados de  $y$ . No estudo de previsão, usualmente inicia-se com um modelo de regressão dado pela seguinte equação:

$$y_i = \sum_{k=1}^K \beta_k x_{k,i} + \varepsilon_i \quad (3.1)$$

$$\varepsilon_i \sim N(0, \sigma^2) \text{ NID}$$

sendo que a variável  $\varepsilon_i$  é considerado um distúrbio aleatório, assumido como sendo normalmente distribuído com média zero e variância  $\sigma^2$  e onde  $\{\beta_k\}$  representa os parâmetros a serem estimados. O conjunto de estimadores dos parâmetros é denotado por  $\{\hat{\beta}_k\}$ , enquanto que o conjunto de previsões de  $y$  geradas pelo modelo com o conjunto de coeficientes  $\{\hat{\beta}_k\}$  é denotado por  $\{\hat{y}_i\}$ . O objetivo é selecionar  $\{\hat{\beta}_k\}$  de tal forma que seja minimizada a soma das diferenças quadráticas entre as observações atuais  $y$  e as observações previstas pelo modelo linear,  $\hat{y}$ .

Em séries temporais, as variáveis de entrada e saída  $[x, y]$ , tem subscrito  $t$ , denotando a observação particular de um dado, com a observação mais antecipada iniciando em  $t=1$ . Em

econometria existe uma grande variedade de métodos para estimar o conjunto de parâmetros  $\{\beta_k\}$ , sujeitos à uma variedade de suposições alternativas sobre a distribuição dos distúrbios  $\varepsilon_t$ , sobre a constância de sua variância,  $\sigma^2$ , bem como sobre a independência da distribuição das variáveis de entrada  $x_k$  com relação as perturbações  $\varepsilon_t$ .

O objetivo do processo de estimação é encontrar o conjunto de parâmetros para o modelo de regressão, dado por  $\{\hat{\beta}_k\}$ , para minimizar E, definido como a soma das diferenças quadráticas, ou resíduos, entre o valor desejado ou variável de saída  $y$  e a variável gerada pelo modelo  $\hat{y}$ , sobre todas as observações. O problema de regressão é formulado da seguinte forma:

$$\underset{\hat{\beta}}{\text{Min}} E = \sum_{i=1}^T \varepsilon_i^2 = \sum_{i=1}^T (y_i - \hat{y}_i)^2 \quad (3.2)$$

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^T \varepsilon_i^2$$

$$y_i = \sum \beta_k x_{k,i} + \varepsilon_i$$

$$\hat{y}_i = \sum \hat{\beta}_k x_{k,i}$$

Um modelo linear comumente utilizado para previsão é o modelo autoregressivo:

$$y_i = \sum_{i=1}^m \beta_i y_{i-1} + \sum_{j=1}^k \gamma_j x_{j,i} + \varepsilon_i \quad (3.3)$$

no qual há  $k$  variáveis independentes  $k$  com  $\gamma_j$  coeficientes para  $x_j$  e  $m$  lags (ou valores atrasados, ou ainda, valores passados) para a variável independente  $y$ , portanto, tendo  $k + m$  parâmetros a serem estimados. Dessa forma, quanto maior for a estrutura de atrasos, maior a quantidade de parâmetros que devem ser estimados e menor o número de graus de liberdade das estimativas de regressão.

O número de variáveis de saída, obviamente, pode ser maior do que um. No entanto, no modelo de regressão padrão, pode ser estimada a previsão de cada variável de saída  $y_j$ ,  $j = 1, 2, \dots, j^*$ , com uma série de  $j^*$  modelos de regressão independentes. Para  $j^*$  variáveis de saída deve-se estimar  $j^* \cdot k$  parâmetros (Patel, 2000).

O modelo linear tem a propriedade útil de ter uma solução de forma fechada para resolver o problema de regressão, ou seja, de minimizar a diferença quadrática entre  $y$  e  $\hat{y}$ . Deste modo, o método de regressão linear (ou auto-regressão linear) é rápido. Para a previsão de um período curto o modelo linear é um ponto de partida razoável, ou “benchmark”, visto que em muitos mercados, observam-se apenas pequenas mudanças simétricas na variável a ser predita, ao longo de uma tendência de longo prazo. No entanto, isto pode não ser preciso o bastante para mercados financeiros voláteis em que pode haver um processo não-linear dos dados. Movimentos lentos no aumento dos preços dos ativos, seguidos por colapsos repentinos conhecidos como “bolhas” são comuns de acontecer. Dessa forma, o modelo linear pode falhar em capturar ou prever pontos de mudança acentuada nos dados.

Por esta razão é que são utilizadas técnicas não-lineares de previsão.

A rede neural é uma alternativa aos modelos lineares e a algumas abordagens não-paramétricas para a aproximação de sistemas não-lineares. A razão para o uso de uma rede neural é simples e direta. O objetivo é encontrar uma abordagem ou método o qual realize bem previsões para dados gerados para processos que freqüentemente são desconhecidos e altamente não-lineares, com uma quantidade pequena de parâmetros e o qual seja mais fácil de estimar que os modelos não-lineares paramétricos.

### 3.2.1 Redes Neurais e Estatística

Existem algumas similaridades que podem ser destacadas entre os modelos estatísticos e as redes neurais. Por exemplo, no caso da resolução de um problema utilizando a técnica de regressão linear que apresente o seguinte formato geral.

$Y_i = \beta_0 + \sum_{j=1}^{p-1} \beta_j X_{ij} + \varepsilon_i$ , verificamos uma similaridade de forma com a saída de um neurônio

que pode ser expressada por:  $v_j(n) = \sum_{i=0}^m w_{ji}(n)x_i(n)$ . Assim, podemos construir as

seguintes relações (Pereira e Rodrigues, 1998):

Redes Neurais	Modelos Estatísticos
Entrada da rede	Variáveis independentes, regressores
Saída da rede	Valores previstos
Valores de treinamento, alvos	Variáveis dependentes
Erros	Resíduos
Treinamento, aprendizagem, adaptação, organização própria	Estimação
Função de erro, função de custo	Critério de estimação
Padrões	Observações
Pesos	Estimativas dos parâmetros
Neurônios intermediários	Interações
Conexões funcionais	Transformações
Aprendizado supervisionado	Regressão e análise discriminante
Aprendizado não-supervisionado	Redução dos dados, análise de componentes principais
Aprendizado competitivo	Análise de grupos

**Tabela-3.1** Terminologias usadas em modelos estatísticos e em redes neurais

As redes neurais atuam como um tipo de modelo de regressão não paramétrica, possibilitando o modelamento de formas funcionais complexas. As redes neurais são úteis quando não sabemos o relacionamento funcional entre variáveis dependentes e independentes, a rede usa os dados para determinar o relacionamento funcional entre as variáveis dependentes e independentes. Sendo a rede neural dependente dos dados, ela apresenta um desempenho melhor em função do tamanho da amostra. A regressão tem desempenho melhor quando a teoria ou a experiência indicam um padrão subjacente. A regressão também pode ser a melhor alternativa para amostras com tamanho extremamente pequenos (Warner e Misra, 1996). O modelo de regressão por postos é a técnica mais adequada para ser utilizada quando estamos interessados em construir uma função que represente dados que apresentam monotonicidade linear (Iman e Conover, 1979). Uma das vantagens de utilizar regressão utilizando método dos mínimos quadrados (MMQ) é a de

utilizar pouca quantidade de dados observados, quando comparado a outros algoritmos, para chegar a valores assintóticos (Galindo-Flores, pág. 71, 1998).

### 3.3 Modelo de um Neurônio Artificial

Um neurônio é uma unidade de processamento que é fundamental para a operação de uma rede neural (Haykin, 2001, pg. 36). A Figura-3.1 apresenta o modelo de um neurônio, que forma a base para o projeto de redes neurais (artificiais).

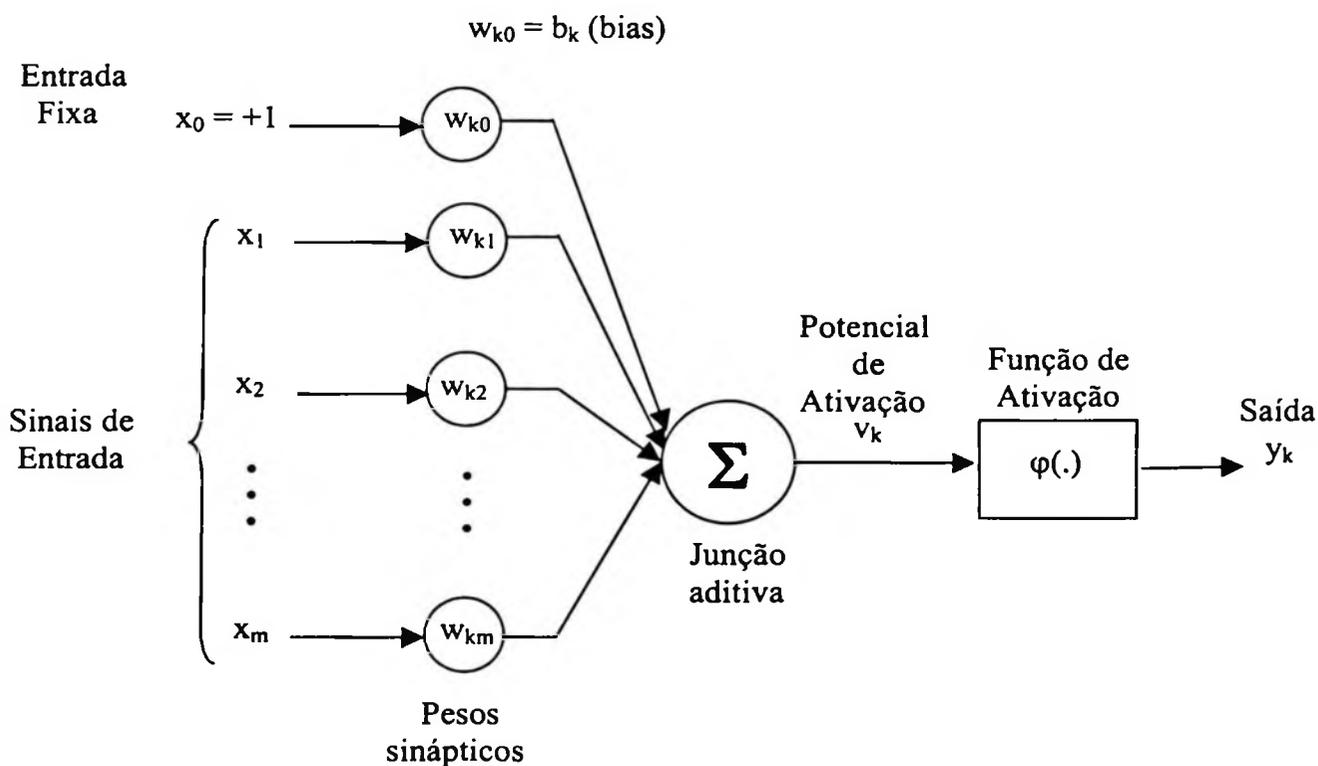


Figura-3.1 Modelo de um neurônio artificial (Haykin, 2001)

Neste modelo neural da Figura-3.1 temos:

- Um conjunto de *sinapses* ou *elos de conexão*, sendo que cada sinapse é caracterizada por um peso ou força própria. Especificamente, um sinal  $x_j$  na entrada da sinapse  $j$  conectada ao neurônio  $k$  é multiplicado pelo peso sináptico  $w_{kj}$ . É importante notar como são escritos

os índices do peso sináptico  $w_{kj}$ . O primeiro índice se refere ao neurônio em questão e o segundo refere ao terminal de entrada da sinapse à qual o peso se refere.

- Um *somador ou integrador* para somar os sinais de entrada, ponderados pelas respectivas sinapses do neurônio; as operações descritas aqui constituem um combinador linear.

- Uma *função de ativação* para restringir a amplitude da saída de um neurônio. A função de ativação é também referida como *função restritiva* já que restringe (limita) o intervalo permissível de amplitude do sinal de saída a um valor finito.

Tipicamente, o intervalo normalizado da amplitude da saída de um neurônio é escrito como intervalo unitário fechado  $[0,1]$  ou alternativamente  $[-1,1]$ .

O modelo neural da Figura-3.1 inclui também um *bias* aplicado externamente, representado por  $b_k$ . O bias  $b_k$  tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se ele é positivo ou negativo, respectivamente.

Em termos matemáticos, pode-se descrever um neurônio  $k$  a partir das seguintes equações:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (3.7)$$

$$y_k = \varphi(u_k + b_k) \quad (3.8)$$

sendo que  $x_1, x_2, \dots, x_m$  são os sinais de entrada;  $w_{k1}, w_{k2}, \dots, w_{km}$  são os pesos sinápticos do neurônio  $k$ ;  $u_k$  é a *saída do combinador linear* devido aos sinais de entrada;  $b_k$  é o bias;  $\varphi(\cdot)$  é a função de ativação e  $y_k$  é o sinal de saída do neurônio. O uso do bias  $b_k$  tem o efeito de aplicar uma *transformação afim* à saída  $u_k$  do combinador linear no modelo da Figura-3.1, como mostrado por:

$$v_k = u_k + b_k \quad (3.9)$$

Em particular, dependendo se o bias  $b_k$  é positivo ou negativo, a relação entre o *campo local induzido* ou *potencial de ativação*  $v_k$  do neurônio  $k$  e a saída do combinador linear  $u_k$  pode ser alterada.

### 3.4 Algoritmo de Retropropagação (*Backpropagation*)

A Figura-3.1 apresenta um perceptron de múltiplas camadas formado por cinco elementos principais: camada de entrada, pesos, camada oculta, bias e camada de saída. O primeiro elemento é a **camada de entrada**, mostrada como quadrados, sendo que estes nós é que recebem os valores dos dados de entrada. As linhas de interconexão indicam que o valor de saída de um neurônio é passado ao longo desta linha até o próximo neurônio. Quando todos os dados da camada de entrada tiverem passado através da última camada, conhecida como camada de saída, um ciclo ou época terá sido realizado. Cada linha de interconexão tem um valor, chamado de **peso**, os pesos são parâmetros os que operam sobre os dados associados a cada linha (ou seja, é realizada uma multiplicação do valor do dado pelo peso). Estes pesos são incrementalmente ajustados durante a fase de treinamento de tal forma a alcançar o resultado de saída desejado para um determinado dado de entrada. Inicializações típicas para estes pesos encontram-se no intervalo de  $[-1;1]$  ou menos.

A segunda camada, e todas as camadas subseqüentes, contêm nós de processamento, conhecidos como neurônios artificiais e são mostrados como círculos na Figura-3.1. qualquer camada entre as camadas de entrada e saída são chamadas de **camadas ocultas** (ou intermediárias). O processamento realizado em um neurônio artificial pode ser dividido em três passos:

- 1- os dados que passam ao longo das linhas de entrada para os neurônios são multiplicados pelos de pesos;
- 2- todos esses dados que foram multiplicados pelos pesos são somados dentro do neurônio;
- 3- o valor total dessa soma é passado através de uma função de transferência, cuja saída representa o valor de saída do neurônio.

A função de transferência mais simples é a função linear (adequada para um espaço linearmente separável), a qual iguala a saída à entrada. A utilização de funções de transferência não lineares (adequadas para um espaço não linearmente separável) é uma das características principais da rede neural.

Na Figura-3.1 também é mostrado um nó especial, este nó é conhecido como bias e tem o seu valor fixado como sendo 1 e alimenta todos os neurônios dentro da rede. A utilização

do bias elimina a necessidade de tratar o limiar como uma característica especial do neurônio e leva a uma implementação de algoritmo mais eficiente.

No final da rede encontra-se o neurônio de saída ou a camada de saída na qual a saída é o estado final da rede. A seleção do número de neurônios na camada de saída e a escolha de suas funções de transferência deverá ser feita atendendo os requisitos do problema de que se dispõem.

A seguir são apresentadas todas as etapas do algoritmo de *backpropagation* (Azoff, 1994).

1. Definir a função de pesos da rede,  $E(w)$ , conhecida como a energia ou função de custo, a qual deseja-se minimizar. Considerando a metade da soma quadrática dos erros (SSE – *sum of square errors*) tem-se que:

$$E_{total} = \sum_{\mu} \sum_{o} E(w) \quad (3.10)$$

$$E(w) = \frac{1}{2} [T - O]^2 \quad (3.11)$$

$$\frac{\partial E}{\partial O} = -(T - O) \quad (3.12)$$

sendo que a soma em (3.10) é realizada sobre os padrões  $\mu$  e os neurônios de saída  $o$ ,  $T$  é o valor alvo e  $O$  é a saída observada da rede, uma função dos pesos da rede  $w$ . A expressão (3.10) significa que deve-se minimizar o erro de todos os processadores da camada de saída para todos os padrões que são apresentados à rede.

2. Definir a função de transferência,  $\varphi(v)$ , para os neurônios. Escolhendo a função sigmóide tem-se que:

$$v = \sum_{i=0}^n w_i x_i \quad (3.13)$$

$$O = \varphi(v) \quad (3.14)$$

$$\varphi(v) = \frac{1}{1 + e^{-v}} \quad (3.15)$$

$$\varphi'(v) = O(1 - O) \quad (3.16)$$

tendo-se que  $v$  é a entrada da função de transferência, sendo que a soma é realizada sobre todos os neurônios de  $i = 1, \dots, n$ , ou seja, o produto de cada peso  $w$  pelo valor de entrada  $x$ .

E  $\varphi'(v)$  é a primeira derivada de  $\varphi(v)$ .

3. Definir as medidas de desempenho da rede. A raiz do erro quadrático médio (RMSE) é uma medida utilizada:

$$RMSE = \hat{\varepsilon} = \sqrt{\frac{\sum_{\mu} \sum_{o} (T - O)^2}{N_o N_e}} \quad (3.17)$$

$N_e$  é o número total de padrões em uma época e  $N_o$  é o número de neurônios na camada de saída. Um conjunto de apresentações de todos os vetores de treinamento constitui uma época.

Para cada medida é selecionado um nível de tolerância para determinar o ponto no qual o treinamento da rede deve parar.

4. Inicializar os pesos randomicamente, na prática utilizam-se pesos no intervalo de  $[-0,5; 0,5]$ . Se a constante de momento for utilizada na atualização dos pesos então tomar  $\Delta w$  inicialmente como zero.

5. Inicializar a lista de ordem das épocas, ou seja, a ordem de apresentação de cada padrão de treinamento, usando um procedimento de aleatoriedade. Tomar o primeiro padrão indicado na lista e apresentar aos nós de entrada da rede.

6. Propagar os valores dos nós de entrada adiante através da rede. Estimar para cada neurônio a sua entrada  $a$  na função de transferência. Estimar a saída da função de transferência  $\varphi(v)$ , este é o valor de saída do neurônio o qual pode ser propagado adiante através das sinapses entre neurônios. Este processo é continuado até os neurônios da camada de saída terem sido alcançados e a rede ter um novo estado de saída observável.

7. Comparar os valores observados nos neurônios da camada de saída,  $O$ , com os valores alvo,  $T$ , e atualizar as medidas de desempenho com o resultado.

8. Calcular o valor  $\delta_i$  para cada neurônio  $i$  na camada de saída, definido como:

$$\delta_i = -\varphi'(v_i) \frac{\partial E}{\partial O} \quad (3.18)$$

$$\delta_i = (1 - O_i) O_i (T_i - O_i) \quad (3.19)$$

9. Retropropagar os deltas através da rede até a camada precedente dos neurônios  $j$ , conectado a saída dos neurônios  $i$ , e calcular os novos valores de delta:

$$\delta_j = -\varphi'(v_j) \sum_i w_{ij} \delta_i \quad (3.20)$$

Dessa forma para um neurônio particular  $j$  na camada escondida anterior a camada de saída, o peso conectando-o ao neurônio  $i$  da camada de saída é multiplicado pelo último valor de delta e somado com todos os neurônios da camada de saída  $i$ . Esta soma é então multiplicada pela derivada da função de transferência do neurônio  $j$ , ou seja,  $\varphi'(v_j)$ . Este processo é continuado para trás na rede até que todos os deltas tenham sido calculados para cada neurônio. Se substituíssemos os neurônios  $i$  por  $k$  teríamos  $\delta_j = -\varphi'(v_j) \sum_k w_{kj} \delta_k$ .

10. O peso conectando quaisquer dois neurônios é atualizado pelo produto da saída dos neurônios da camada precedente,  $V_p$ , e o valor de delta dos neurônios da camada precedente,  $\delta_q$ :

$$\bullet p \rightarrow \bullet q \quad (3.21)$$

$$\Delta w_{qp} = \eta \delta_q V_p \quad (3.22)$$

$$w_{qp}^{novo} = w_{qp}^{antigo} + \Delta w_{qp} \quad (3.23)$$

sendo que a seta na expressão (3.21) indica uma estrutura de caminho adiante, e  $\eta$  é a taxa de aprendizagem. Todos os pesos são atualizados de acordo com esta regra.

11. A expressão de atualização de pesos acima para  $\Delta w$  pode ser modificada com a adição de um termo de momento:

$$\Delta w_{qp}^{novo} = \eta \delta_q V_p + \alpha \Delta w_{qp}^{antigo} \quad (3.24)$$

sendo que o coeficiente de momento  $\alpha$  recebe um valor no intervalo de  $[0;1]$ . A consequência disto é dar à atualização dos pesos uma memória de sua última atualização, fornecendo uma suavização nas forças que afetam as mudanças dos pesos.

12. Se a lista de épocas foi exaurida, devemos ir até o próximo passo, caso contrário apresentar o próximo conjunto de treinamento à camada de entrada e voltar ao passo 6.

13. Tendo completado uma época, e acumulado os erros de todos os padrões de entrada, os erros obtidos são avaliados e comparados com os níveis de tolerância. Se a rede tiver convergido então voltar até o passo 5, caso contrário encerrar o treinamento.

### 3.5 Redes Neurais Alimentadas Adiante (*Feedforward*)

A aplicação de redes neurais na previsão de séries temporais envolve três estágios principais. O primeiro é o estágio de pré-processamento no qual a arquitetura da rede é definida. O estágio de pré-processamento também inclui a construção de um vetor de entrada e sua transformação para uma forma adequada para treinamento da rede. O segundo é o procedimento de treinamento no qual a rede aprende as relações subjacentes que governam os dados de entrada e saída. Neste estágio os parâmetros da rede são adaptados de acordo com o algoritmo de aprendizagem de tal forma a minimizar o erro de aproximação na camada de saída. Finalmente, há o estágio de teste da rede onde um conjunto separado dos dados que não foi apresentado à rede é usado para estimar o desempenho levando a seleção do melhor da rede treinada. Muitas medidas de desempenho, tal como a raiz do erro quadrático médio (RMSE – *root mean square error*) e a medida de correlação podem ser computadas para medir o desempenho.

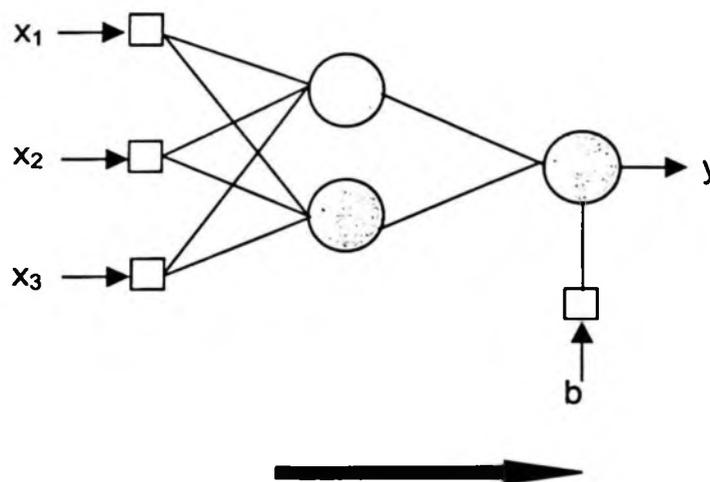


Figura-3.2 Rede *Feedforward* – Alimentada Adiante

Vamos descrever agora o algoritmo de aprendizagem de retropropagação temporal para uma rede neural alimentada adiante com atraso no tempo. A Figura-3.2 ilustra uma rede neural com uma camada escondida contendo dois neurônios, três variáveis de entrada  $\{x_i\}, i = 1, 2, 3$  e uma saída  $y$ .

O que dá à rede neural seu poder de previsão é o processamento paralelo. Adicionalmente ao processamento seqüencial dos sistemas lineares típicos, no qual apenas entradas observadas são utilizadas para prever uma saída observada, pela ponderação dos neurônios de entrada, os neurônios na camada oculta processam as entradas de forma paralela, de tal forma a melhorar as predições. Esta rede de camada simples alimentada adiante, ou multiperceptron, com uma camada escondida é a rede neural mais básica e comumente utilizada em aplicações econômicas e financeiras. Mais genericamente, a rede representa o modo como o cérebro humano processa os dados sensoriais de entrada, recebidos como neurônios de entrada, e internamente reconhecidos como um neurônio de saída. Com o desenvolvimento do cérebro, mais e mais neurônios são interconectados e mais sinapses, e os sinais de diferentes neurônios, trabalham de forma paralela, em mais e mais camadas escondidas, são combinados pelas sinapses para produzir *insights* e reações.

### 3.5.1 Incorporação do Tempo na Rede Neural

A incorporação do tempo na operação de uma rede neural pode ser realizada de duas formas (Haykin, 2001):

- 1ª. Forma - Representação implícita: o tempo é representado pelo efeito que tem sobre o processamento de sinais de uma maneira implícita. Por exemplo, o sinal de entrada é amostrado uniformemente, e a seqüência de pesos sinápticos de cada neurônio conectados à camada de entrada da rede sofre uma convolução com uma seqüência diferente de amostras de entrada. Fazendo assim, a estrutura temporal do sinal de entrada é inserida na estrutura espacial da rede.
- 2ª. Forma - Representação explícita: o tempo recebe sua própria representação particular. O sistema de ecolocalização de um morcego, por exemplo, opera emitindo um curto sinal de frequência modulada (FM), de modo que o mesmo nível de intensidade é mantido para cada canal de frequência restrito a um período muito curto dentro da varredura de FM. São realizadas múltiplas comparações entre várias frequências diferentes codificadas por um arranjo de receptores auditivos com a finalidade de extrair informação precisa sobre a distância (alcance) até o alvo.

Quando um eco é recebido do alvo com um atraso desconhecido, um neurônio (no sistema auditivo) com uma linha de atraso casada com este sinal responde, fornecendo dessa forma uma estimativa do alcance do alvo.

Neste trabalho utilizaremos a representação implícita do tempo, de tal forma que uma rede neural estática (por exemplo, um perceptron de múltiplas camadas) é suprida de propriedades dinâmicas. E dessa maneira a rede torná-se sensível à estrutura temporal dos sinais portadores de informação.

Para que uma rede neural seja dinâmica, ela deve ter memória.

A memória pode ser dividida em memória de *curto prazo* e de *longo prazo*, dependendo do tempo de retenção. Nesse estudo necessitaremos de alguma forma de memória de curto prazo para tornar a rede dinâmica. Uma forma simples de inserir memória de curto prazo na estrutura de uma rede neural é através de atrasos de tempo, que podem ser implementados a nível sináptico dentro da rede ou na camada de entrada da rede.

Portanto, o papel principal da memória é transformar uma rede estática em uma rede dinâmica. Em particular, incorporando memória na estrutura de uma rede estática como um perceptron de múltiplas camadas ordinário, a saída da rede se torna uma função do tempo. Sendo assim, a abordagem para construir um sistema dinâmico não-linear é direta porque fornece uma clara separação de responsabilidades: a rede estática é responsável pela não-linearidade, e a memória é responsável pelo tempo.

Esta abordagem é chamada de rede neural com atraso de tempo (TDNN – *time delay neural network*) e tem sido utilizada em diversas áreas, tais como reconhecimento de fala. Formalmente, os atrasos de tempo são a janelas de tempo e podem então ser visualizadas como modelos autoregressivos. A Figura-3.3 apresenta uma TDNN.

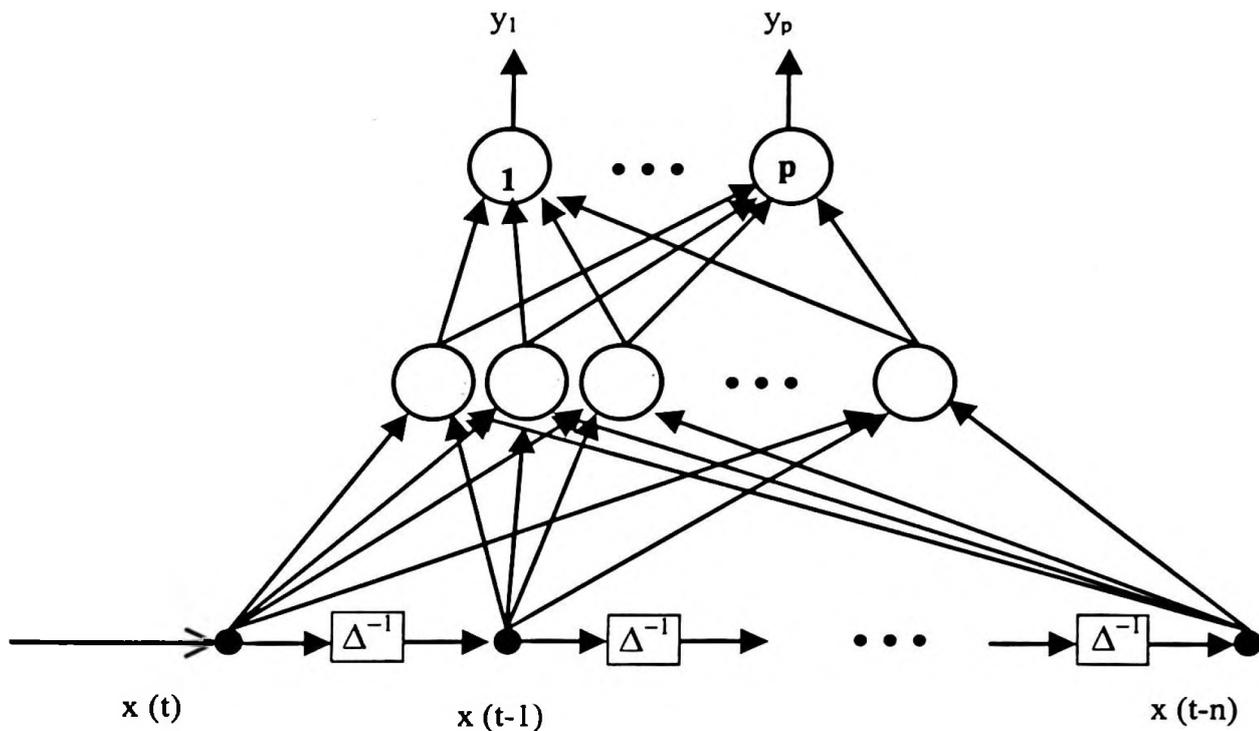


Figura-3.3 Exemplo de uma rede TDNN - *Time Delay Neural Network*

Na Figura-3.4a temos uma unidade de processamento temporal composta de uma memória de linha de atraso derivada com suas derivações conectadas às sinapses de um neurônio. A memória de linha de atraso derivada captura a informação temporal contida no sinal de entrada e o neurônio insere esta informação em seus próprios pesos sinápticos. Esta unidade de processamento é chamada de filtro neural focado, focado no sentido de que a estrutura inteira da memória é localizada no terminal de entrada da unidade.

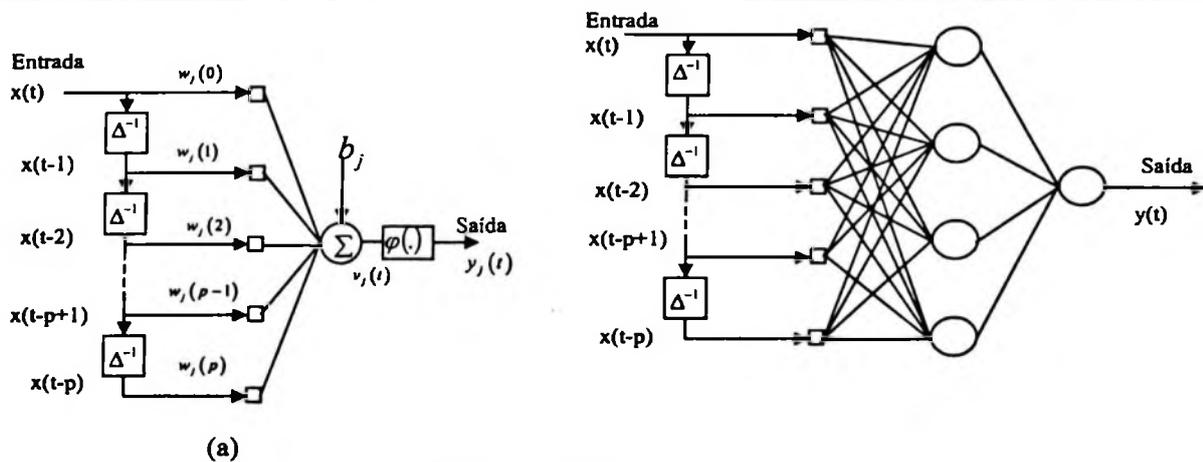


Figura-3.4 Filtro Neural Focado (a), Rede TLFN – *Time Lagged Feedforward Network* (b)

A rede representada na Figura-3.4b é chamada de rede alimentada adiante focada defasada no tempo (TLFN – *focused time lagged feedforward network*), consiste em uma memória de linha de atraso com derivações de ordem  $p$  e um perceptron de múltiplas camadas.

A saída do filtro na Figura-3.4a, em resposta à entrada  $x(t)$  e aos seus valores passados  $x(t-1), \dots, x(t-p)$ , é dada por:

$$y_j(t) = \varphi \left( \sum_{l=0}^p w_j(l) x(t-l) + b_j \right) \quad (3.25)$$

No caso da Figura-3.4b onde tem-se uma camada oculta a saída  $y(t)$  será dada por:

$$y(t) = \sum_{j=1}^{m_1} w_j \cdot y_j(t) \quad (3.26)$$

substituindo  $y_j(t)$  da expressão anterior implica que a saída na rede da Figura-3.4b será:

$$y(t) = \sum_{j=1}^{m_1} w_j \cdot \varphi \left( \sum_{l=0}^{m_1} w_j(l) x(t-l) + b_j \right) + b_0 \quad (3.27)$$

A Figura-3.5 mostra o diagrama de um filtro de resposta a impulso de duração finita (FIR – *finite impulse response filter*). O filtro FIR é um dos blocos construtivos básicos na área de processamento digital de sinais.

Temos então que a combinação de elementos de atraso unitário e os pesos sinápticos associados pode ser vista como um filtro FIR que gera uma saída  $s_j(t)$  que é a convolução entre os sinais de entrada e os pesos sinápticos. A convolução pode ser entendida como um método matemático que combina dois sinais para produzir um terceiro sinal.

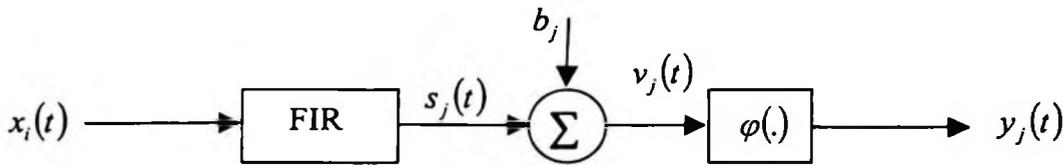
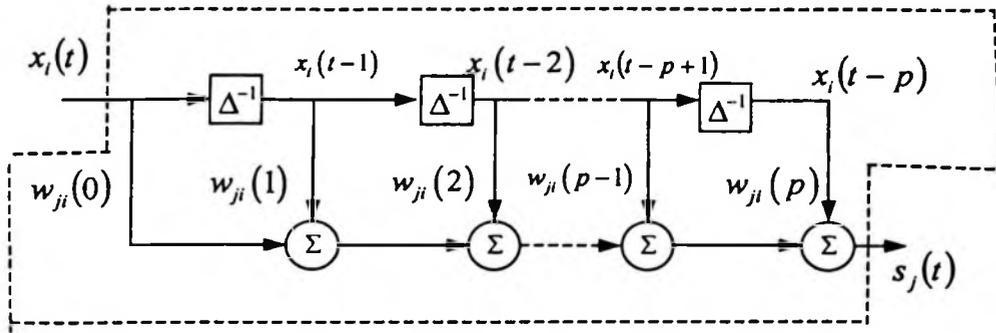


Figura-3.5 Filtro FIR aplicado à uma rede neural

A saída do filtro FIR da Figura-3.5 é dada por:

$$s_j(t) = \sum_{k=0}^p w_{ji}(k)x_i(t-k) \quad (3.28)$$

Podemos então, em termos matemáticos, descrever o processamento espaço-temporal realizado no filtro neural da Figura-3.5 expressando sua saída  $y_j(n)$ , como:

$$y_j(t) = \varphi \left( \sum_{i=1}^m \sum_{l=0}^p w_{ji}(l)x_i(t-l) + b_j \right) \quad (3.29)$$

sendo que  $w_{ji}(l)$  é o peso da  $l$ -ésima sinapse secundária pertencente à  $i$ -ésima sinapse primária,  $x_i(t)$  é a entrada aplicada à  $i$ -ésima sinapse primária no tempo  $t$  e  $b_j$  é o bias aplicado ao neurônio. Na Figura-3.5, o neurônio possui  $m$  sinapses primárias que são responsáveis pela dimensão espacial do processamento da série, sendo que cada sinapse primária possui  $(p+1)$  sinapses secundárias que são responsáveis pela dimensão temporal do processamento da série e o número de total de pesos sinápticos é igual a  $m(p+1)$ .

Uma limitação da TLFN é de ser adequada apenas para o uso em ambientes estacionários (ou seja, invariantes no tempo). Para superar esta limitação utiliza-se uma rede alimentada adiante atrasada no tempo (TLFN) distribuída. Esta rede é distribuída porque a influência implícita do tempo é distribuída através da rede.

O modelamento de uma rede neural alimentada adiante atrasada no tempo distribuída pode ser realizada conforme mostrado a seguir.

Seja  $w_{ji}(l)$  o peso conectado a  $l$ -ésima derivação do filtro FIR que modela a sinapse que conecta a saída do neurônio  $i$  ao neurônio  $j$ , então o sinal  $s_{ji}(t)$  que aparece na  $i$ -ésima sinapse do neurônio  $j$  é dado por:

$$s_{ji}(t) = \sum_{l=0}^p w_{ji}(l) x_i(t-l) \quad (3.30)$$

Escrevendo vetorialmente tem-se:

$$s_{ji}(t) = \mathbf{w}_{ji}^T \mathbf{x}_i(t) \quad (3.31)$$

sendo que:

$$\mathbf{x}_i(t) = [x_i(t), x_i(t-1), \dots, x_i(t-p)]^T \quad (3.32)$$

$$\mathbf{w}_{ji} = [w_{ji}(0), w_{ji}(1), \dots, w_{ji}(p)]^T \quad (3.33)$$

O sobrescrito  $T$  significa operador de transposição.

A saída do neurônio  $j$  será dada então pelas seguintes expressões:

$$v_j(t) = \sum_{i=1}^m s_{ji}(t) + b_j = \sum_{i=1}^m \mathbf{w}_{ji}^T \mathbf{x}_i(t) + b_j \quad (3.34)$$

$$y_j(t) = \varphi(v_j(t)) \quad (3.35)$$

sendo que  $v_j(t)$  representa o campo local induzido do neurônio  $j$  (Haykin, 2001, pg. 54),  $b_j$  é o bias aplicado externamente e  $\varphi(\square)$  representa a função de ativação não-linear do neurônio. É importante notar que se o vetor de pesos  $\mathbf{w}_{ji}$  e o vetor de estado  $\mathbf{x}_i(t)$  forem substituídos pelos escalares  $w_{ji}$  e  $x_i$ , respectivamente, o modelo dinâmico do neurônio se reduz ao modelo estático do perceptron de múltiplas camadas.

### 3.5.2 Algoritmo de Retropropagação Temporal

Para realizar o treinamento da TLFN distribuída é necessário um algoritmo de aprendizagem supervisionada no qual a resposta real de cada neurônio na camada de saída é

comparada com uma resposta desejada a cada instante de tempo (Lang e Hinton, 1988) (Waibel, et al. 1989).

Seja  $y_j(t)$  a resposta real do neurônio  $j$  da camada de saída no tempo  $t$  e  $d_j(t)$  a resposta desejada do neurônio  $j$  da camada de saída no tempo  $t$ . O valor instantâneo para a soma dos erros quadráticos será dado por:

$$E(t) = \frac{1}{2} \sum_j e_j^2(t) \quad (3.36)$$

o erro é definido como:

$$e_j(t) = d_j(t) - y_j(t) \quad (3.37)$$

O objetivo é minimizar o erro quadrático para todos os tempos:

$$E_{total} = \sum_t E(t) \quad (3.38)$$

Diferenciando em relação ao vetor de pesos:

$$\frac{\partial E_{total}}{\partial \mathbf{w}_{ji}} = \sum_t \frac{\partial E(t)}{\partial \mathbf{w}_{ji}} \quad (3.39)$$

Também podemos considerar um modo alternativo de expressar a derivada parcial da função de custo  $E_{total}$  em relação ao vetor de pesos  $\mathbf{w}_{ji}(t)$ :

$$\frac{\partial E_{total}}{\partial \mathbf{w}_{ji}} = \sum_t \frac{\partial E(t)}{\partial v_j(t)} \cdot \frac{\partial v_j(t)}{\partial \mathbf{w}_{ji}} \quad (3.40)$$

A derivada parcial  $\frac{\partial E(t)}{\partial v_j(t)}$  pode ser interpretada como a variação da função de custo  $E_{total}$

produzida por uma variação no campo local induzido  $v_j$  do neurônio  $j$  no tempo  $t$ .

Entretanto, é importante notar que:

$$\frac{\partial E_{total}}{\partial v_j(t)} \cdot \frac{\partial v_j(t)}{\partial \mathbf{w}_{ji}} \neq \frac{\partial E(t)}{\partial \mathbf{w}_{ji}} \quad (3.41)$$

A igualdade é válida somente quando é realizada uma soma sobre todo  $t$ , tal como nas Equações (3.39) e (3.40).

A atualização do vetor de pesos sinápticos será realizada da seguinte maneira:

$$\mathbf{w}_{ji}(t+1) = \mathbf{w}_{ji}(t) - \eta \frac{\partial E_{total}}{\partial v_j(t)} \frac{\partial v_j(t)}{\partial \mathbf{w}_{ji}(t)} \quad (3.42)$$

porém,

$$\frac{\partial v_j(t)}{\partial \mathbf{w}_{ji}(t)} = \mathbf{x}_i(t) \quad (3.43)$$

sendo que  $\mathbf{x}_i(t)$  é o vetor de entrada aplicado à sinapse  $i$  do neurônio  $j$ .

Definindo o gradiente local para o neurônio  $j$  como:

$$\delta_j(t) = -\frac{\partial E_{total}}{\partial v_j(t)} \quad (3.44)$$

a atualização do vetor de pesos sinápticos será obtida da seguinte maneira:

$$\mathbf{w}_{ji}(t+1) = \mathbf{w}_{ji}(t) + \eta \delta_j(t) \mathbf{x}_i(t) \quad (3.45)$$

Tal como na derivação do algoritmo de retropropagação padrão, a forma explícita do gradiente local  $\delta_j(t)$  depende se o neurônio  $j$  está na camada de saída ou na camada oculta da rede. Teremos que considerar os dois casos: o neurônio  $j$  está na camada de saída e depois a situação em que o neurônio  $j$  está em uma camada oculta.

**Caso 1** → O neurônio  $j$  está na camada de saída.

Neste caso temos simplesmente que:

$$\delta_j(t) = -\frac{\partial E_{total}}{\partial v_j(t)} = -\frac{\partial E(t)}{\partial v_j(t)} = -\frac{\partial E(t)}{\partial y_j(t)} \frac{\partial y_j(t)}{\partial v_j(t)} \quad (3.46)$$

Porém,  $E(t) = \frac{1}{2} \sum_j e_j^2(t)$  e  $e_j(t) = d_j(t) - y_j(t)$ , dessa forma as derivadas parciais na

expressão (3.46) podem ser substituídas por:

$$\frac{\partial E(t)}{\partial y_j(t)} = e_j(t) \frac{\partial e_j(t)}{\partial y_j(t)} = e_j(t) \frac{\partial [d_j(t) - y_j(t)]}{\partial y_j(t)} = -e_j(t) \quad (3.47)$$

$$\frac{\partial y_j(t)}{\partial v_j(t)} = \varphi'(v_j(t)) \quad (3.48)$$

No caso em que o neurônio  $j$  está na camada de saída o gradiente local é obtido através da seguinte expressão:

$$\delta_j(t) = e_j(t) \varphi'(v_j(t)) \quad (3.49)$$

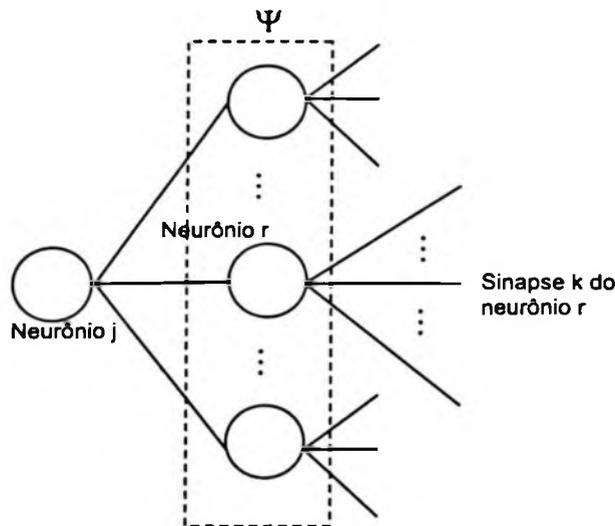
**Caso 2** → O neurônio  $j$  está na camada oculta.

Seja  $\Psi$  o conjunto de todos os neurônios cujas entradas são alimentadas pelo neurônio  $j$  e considerando  $v_r(t)$  o potencial do neurônio  $r \in \Psi$ . Podemos então escrever que:

$$\delta_j(t) = -\frac{\partial E_{total}}{\partial v_j(t)} = -\sum_{r \in \Psi} \sum_k \frac{\partial E_{total}}{\partial v_r(k)} \frac{\partial v_r(k)}{\partial v_j(t)} \quad (3.50)$$

Para fazermos a equivalência entre  $\frac{\partial E_{total}}{\partial v_j(t)}$  e  $\sum_{r \in \Psi} \sum_k \frac{\partial E_{total}}{\partial v_r(k)} \frac{\partial v_r(k)}{\partial v_j(t)}$  precisamos somar todos

os campos locais induzidos das  $k$  sinapses do neurônio  $r$  e depois somar todos os neurônios que pertencem ao conjunto de neurônios  $\Psi$ . A Figura-3.6 mostra esquematicamente como deve ser realizado o duplo somatório da Equação (3.50).



**Figura-3.6** Representação do somatório da Equação (3.50)

Fazendo  $\delta_r(k) = -\frac{\partial E_{total}}{\partial v_r(k)}$  implica em:

$$\delta_j(t) = \sum_{r \in \Psi} \sum_k \delta_r(k) \frac{\partial v_r(k)}{\partial v_j(t)} = \sum_{r \in \Psi} \sum_k \delta_r(k) \frac{\partial v_r(k)}{\partial y_j(t)} \frac{\partial y_j(t)}{\partial v_j(t)} \quad (3.51)$$

lembrando que  $y_j(t)$  é a saída do neurônio  $j$  e  $\frac{\partial y_j(t)}{\partial v_j(t)} = \varphi'(\partial v_j(t))$  tem-se:

$$\delta_j(t) = \varphi'(\partial v_j(t)) \sum_{r \in \Psi} \sum_k \delta_r(k) \frac{\partial v_r(k)}{\partial y_j(t)} \quad (3.52)$$

$\partial v_r(k)$  é o potencial do neurônio r que é alimentado pela saída do neurônio j dado por:

$$\partial v_r(k) = \sum_{j=0}^{m_0} \sum_{l=0}^p w_{rj}(l) y_j(k-l) \quad (3.53)$$

Com  $w_{r0}(l) = b_r$  e  $y_0(k-l) = 1$  representando o bias do neurônio r.

O cálculo de  $\frac{\partial v_r(k)}{\partial y_j(t)}$  é mostrado na Figura-3.7.

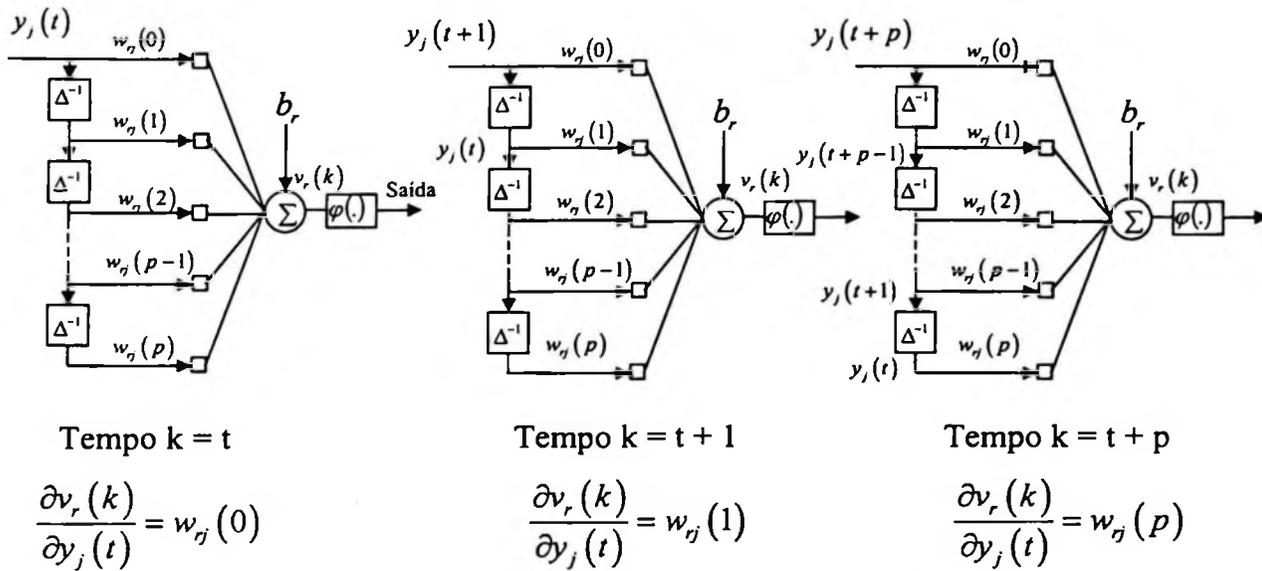


Figura- 3.7 Derivadas  $\partial v/\partial y$

Dessa forma a diferenciação da Equação (3.53) com relação a  $y_j$  pode ser sintetizada como:

$$\frac{\partial v_r(k)}{\partial y_j(t)} = \begin{cases} w_{rj}(k-t), & \text{para } t \leq k \leq t+p \\ 0, & \text{caso contrário} \end{cases} \quad (3.54)$$

Com isso o gradiente local será dado por:

$$\delta_j(t) = \varphi'(\partial v_j(t)) \sum_{r \in \Psi} \sum_k \delta_r(k) \frac{\partial v_r(k)}{\partial y_j(t)}$$

$$\delta_j(t) = \varphi'(\partial v_j(t)) \sum_{r \in \Psi} \sum_{k=t}^{t+p} \delta_r(k) w_{rj}(k-t)$$

$$\delta_j(t) = \varphi'(\partial v_j(t)) \sum_{r \in \Psi} \sum_{k=0}^p \delta_r(k+t) w_{rj}(k) \quad (3.55)$$

Definindo dois novos vetores (p+1)-por-1:

$$\Delta_r(t) = [\delta_r(t), \delta_r(t+1), \dots, \delta_r(t+p)]^T \quad (3.56)$$

$$\mathbf{w}_{rj}(t) = [w_{rj}(0), w_{rj}(1), \dots, w_{rj}(p)]^T \quad (3.57)$$

Dessa forma a Equação (3.55) pode ser reescrita em notação matricial:

$$\delta_j(t) = \varphi'(v_j(t)) \sum_{r \in \Psi} \Delta_r^T(t) \mathbf{w}_{rj} \quad (3.58)$$

sendo que  $\Delta_r^T \mathbf{w}_{rj}$  é o produto interno dos vetores  $\Delta_r^T$  e  $\mathbf{w}_{rj}$ , ambos tendo dimensão (p+1).

A equação (3.58) completa avaliação do gradiente local  $\delta_j(t)$  para o neurônio  $j$  na camada oculta.

A atualização dos pesos para a retropropagação temporal pode ser resumida a partir das seguintes relações (Wan, 1990, 1994):

$$\mathbf{w}_{ji}(t+1) = \mathbf{w}_{ji}(t) + \eta \delta_j(t) \mathbf{x}_i(t) \quad (3.59)$$

$\delta_j(t) = \begin{cases} e_j(t) \varphi'(v_j(t)), \\ \varphi'(v_j(t)) \sum_{r \in \Psi} \Delta_r^T(t) \mathbf{w}_{rj} \end{cases}$	o neurônio $j$ está na camada de saída	(3.60)
	o neurônio $j$ está em uma camada oculta	

O cálculo do gradiente local em (3.60) para um neurônio na camada oculta está em função dos valores futuros de  $\delta_j(t)$ , para tornar o vetor gradiente local  $\Delta_r$  dependente apenas dos valores presentes e passados é feita a seguinte reindexação:

$$\Delta_r(t-p) = [\delta_r(t-p), \delta_r(t+1-p), \dots, \delta_r(t)]^T \quad (3.61)$$

Perceba na Equação (3.61) que  $\Delta_r(t-p)$  representa um vetor de  $p+1$  elementos por 1 (Veja a Figura-3.7).

A Tabela-3.2 resume o algoritmo de retropropagação temporal para uma rede neural alimentada adiante diante - *feedforward*.

### Algoritmo de Retropropagação Temporal

1. Propagar a série de entrada para frente através da rede, camada por camada. Determinar o sinal de erro para o neurônio  $j$  da camada de saída subtraindo sua saída .

2. Para o neurônio  $j$  da camada de saída, calcular:

$$\delta_j(t) = e_j(t) \varphi'(v_j(t))$$

$$w_{ji}(t+1) = w_{ji}(t) + \eta \delta_j(t) x_i(t)$$

sendo que  $x_i(t)$  é o estado da sinapse  $i$  de um neurônio oculto conectado ao neurônio de saída  $j$ .

3. Para o neurônio  $j$  em uma camada oculta, calcular:

$$\delta_j(t-lp) = \varphi'(\partial v_j(t-lp)) \sum_{r \in \Psi} \Delta_r^T(t-lp) w_{rj}$$

$$w_{ji}(t+1) = w_{ji}(t) + \eta \delta_j(t-lp) x_i(t-lp)$$

sendo que  $p$  é a ordem de cada filtro sináptico FIR e o índice  $l$  identifica a camada oculta em questão. Para redes com múltiplas camadas ocultas,  $l=1$  corresponde a uma camada atrás da camada de saída,  $l=2$  corresponde a duas camadas atrás da camada de saída, e assim por diante.

**Tabela-3.2** Resumo do algoritmo de retropropagação temporal para uma rede alimentada adiante

## 3.6 Redes Neurais Recorrentes

Existe uma grande quantidade de algoritmos disponíveis para o treinamento dos pesos de redes neurais recorrentes baseados nas entradas e saídas correntes de dados. Talvez, os algoritmos mais amplamente utilizados sejam o algoritmo de aprendizagem recorrente em tempo real (RTRL – *real time recurrent learning*) e o algoritmo de retropropagação através do tempo (BPTT – *backpropagation through time*) (William e Zipser, 1995).

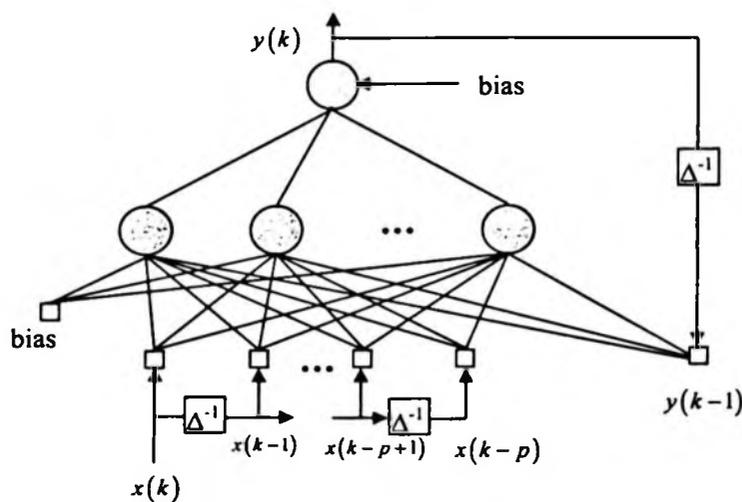
Recentemente, muitos autores tem notado que o filtro de Kalman estendido (EKF – *extended Kalman Filter*) também pode ser usado para o propósito de treinamento da rede para desenvolver mapeamentos de entrada-saída desejados.

Nesta seção vamos apresentar os conceitos principais sobre redes neurais recorrentes e os algoritmos mais aplicados.

Uma rede neural pode ser classificada de acordo com a arquitetura da rede. A forma mais simples de rede é que foi vista na seção anterior, a rede alimentada adiante – *feedforward*. Como foi mostrado, estas redes constituem-se de múltiplas camadas, onde a entrada de cada camada é a saída da camada predecessora e a interconexão dos neurônios é acíclica.

As redes recorrentes tendem a ser mais difíceis de treinar do que as redes alimentadas adiante, principalmente em função dos ciclos. As redes recorrentes mais utilizadas são a rede de Jordan (Jordan, 1986), a rede de Elman e a rede de Elman estendida conforme mostrado nas Figuras-3.8/3.9 e 3.10, respectivamente.

A diferença entre a rede de Elman e a rede de Jordan é que na rede de Jordan a realimentação ocorre apenas dos nós de saída para os nós de entrada, enquanto que na rede de Elman a realimentação sai dos nós das camadas escondidas. Existe também a rede de Elman estendida na qual a realimentação sai dos nós da camada escondida e também dos nós (ou neurônios) de saída.



**Figura- 3.8** Rede de Jordan

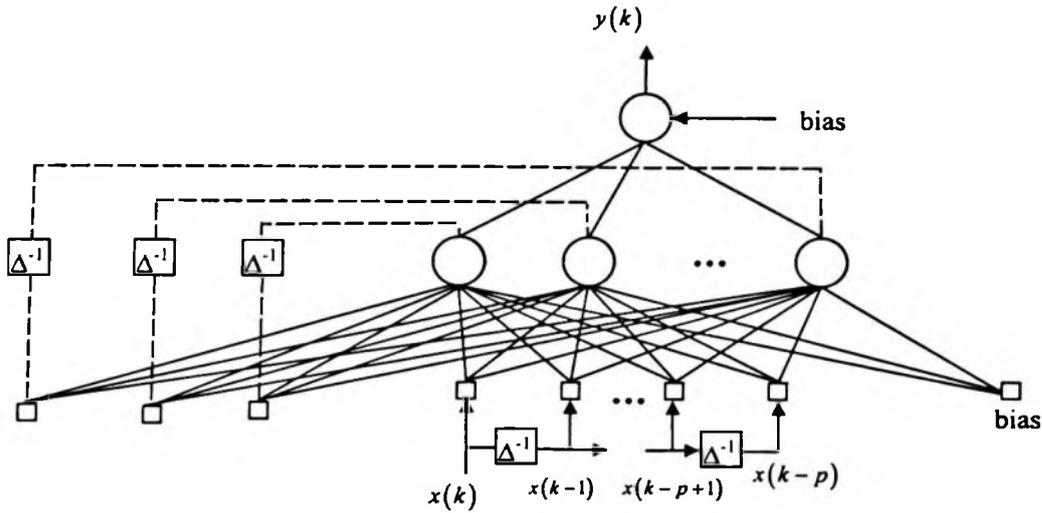


Figura- 3.9 Rede de Elman

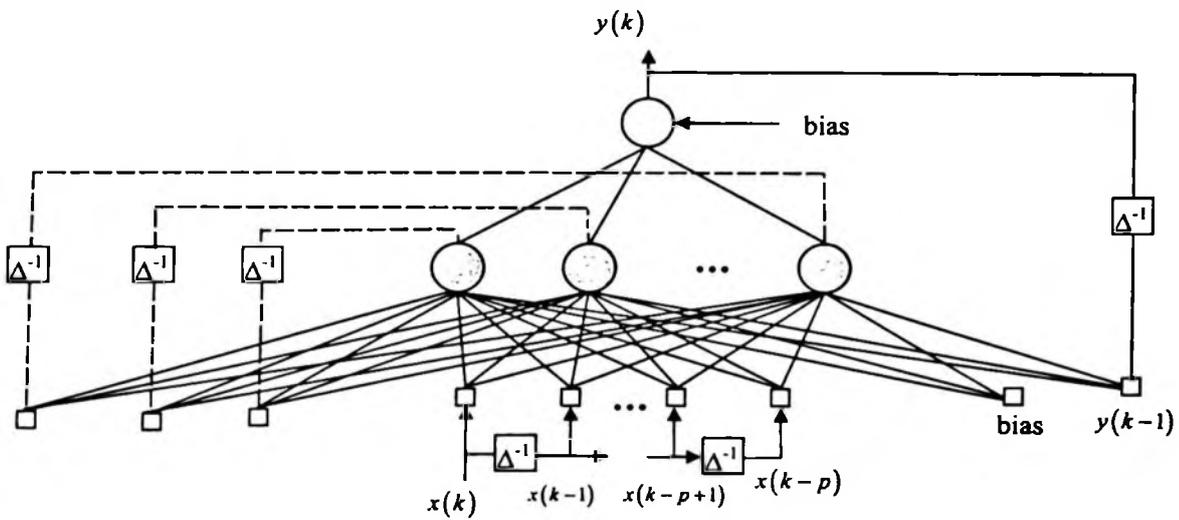


Figura- 3.10 Rede de Elman estendida

### 3.6.1 Redes de Elman e o Modelo de Espaço de Estados

Um outro método comum para processamento de séries temporais são os chamados *modelos de espaço de estados*. A suposição básica destes modelos é que uma série temporal pode ser descrita como uma transformação linear de um estado de tempo dependente, ou seja, dado um vetor de estado  $s$  :

$$\mathbf{x}(t) = \mathbf{C}\mathbf{s}(t) + \boldsymbol{\epsilon}(t) \quad (3.62)$$

sendo que  $\mathbf{C}$  é uma matriz de transformação (Dorffner, 1996). O vetor de estado de tempo dependente também é usualmente descrito por um modelo linear:

$$\mathbf{s}(t) = \mathbf{A}\mathbf{s}(t-1) + \mathbf{B}\boldsymbol{\eta}(t) \quad (3.63)$$

sendo que  $\mathbf{A}$  e  $\mathbf{B}$  são matrizes, e  $\boldsymbol{\eta}(t)$  é um ruído tal como  $\boldsymbol{\epsilon}(t)$  na expressão (3.53).

O modelo para a mudança de estado, neste caso, é basicamente um processo ARMA(1,1). A suposição básica subjacente neste modelo é a chamada *suposição de Markov*, significando que o próximo elemento da seqüência pode ser predito pelo estado de um sistema produzindo uma série temporal em si mesma, não importando como o estado foi conseguido. Em outras palavras, toda a história de uma série necessária para produzir um elemento da seqüência pode ser expresso por um vetor de estado. Desde que o vetor  $\mathbf{s}$  tenha um valor contínuo, todos os vetores de estado possíveis formam um *vetor de espaço* Euclidiano em  $R^n$ .

Se assumirmos que os estados são dependentes do vetor de seqüências passadas e negligenciarmos o termo de média móvel  $\mathbf{B}\boldsymbol{\eta}(t)$  teremos que:

$$\mathbf{s}(t) = \mathbf{A}\mathbf{s}(t-1) + \mathbf{D}\mathbf{x}(t-1) \quad (3.64)$$

dessa forma, obtemos um equação descrevendo um tipo de rede neural, conhecida como rede de Elman, mostrada na Figura- 3.9. A rede de Elman é um perceptron de múltiplas camadas (MLP) com uma camada de entrada adicional, chamada de *camada de estado*, recebendo como realimentação uma cópia das ativações da camada oculta no passo de tempo anterior. Se esta rede for utilizada para previsão, e igualando o vetor de ativação da camada oculta a  $\mathbf{s}$ , a única diferença para a Equação-(3.64) é o fato de que em um MLP a função de ativação sigmóide é aplicada a entrada de unidade escondida:

$$\mathbf{s}(t) = \varphi(\mathbf{A}\mathbf{s}(t-1) + \mathbf{D}\mathbf{x}(t-1)) \quad (3.65)$$

lembrando que  $\varphi(\mathbf{a})$  refere-se a aplicação da função sigmóide (ou logística)  $1/(1 + \exp(-a_i))$  a cada elemento  $a_i$  de  $\mathbf{a}$ . Em outras palavras, a transformação é não-linear mas a aplicação de um regressor logístico aos vetores de entrada.

A rede de Elman pode ser treinada com qualquer algoritmo de aprendizagem para MLPs, tal como o *backpropagation*. De forma semelhante à rede de Jordan, a rede de Elman

pertence a classe das redes recorrentes simples (SRN – *simple recurrent networks*). Embora ela contenha conexões de realimentação, ela não é vista como um sistema dinâmico nas quais as ativações são espalhadas indefinidamente. Em vez disso, as ativações para cada camada são computadas apenas uma vez a cada passo de tempo (cada apresentação de um vetor de seqüências).

Conforme mostrado acima, a forte relação com o processamento clássico de séries temporais pode ser explorado para introduzir-se novos algoritmos de aprendizagem. Por exemplo, o algoritmo do Filtro de Kalman, desenvolvido para o modelo de espaço de estado original é aplicado para redes neurais recorrentes gerais.

### 3.6.2 Algoritmo de Aprendizagem Recorrente em Tempo Real (RTRL)

Seja uma rede neural com  $n$  unidades e  $m$  linhas de entrada externas. Fazendo com que  $y(t)$  denote um conjunto de  $n$ -elementos das unidades da rede no tempo  $t$ , e fazendo com que  $x(t)$  denote um conjunto de  $m$ -elementos de sinais externos de entrada para a rede no tempo  $t$ . Os conjuntos  $y(t)$  e  $x(t)$  podem ser concatenados para formar o conjunto de  $(m+n)$ -elementos  $z(t)$ , com  $U$  denotando o conjunto de índices  $k$  de tal forma que  $z_k$  é a saída de uma unidade na rede e  $I$  o conjunto de índices  $k$  para os quais  $z_k$  é uma entrada externa. Os índices de  $y$  e  $x$  são escolhidos para corresponder àqueles de  $z$ , de tal forma que:

$$z_k(t) = \begin{cases} x_k(t) & \text{se } k \in I \\ y_k(t) & \text{se } k \in U \end{cases} \quad (3.66)$$

Fazendo  $W$  denotar a matriz de pesos para a rede, com um único peso entre todos os pares de unidades e também de cada linha de entrada à cada unidade. Pela adoção da convenção descrita, pode-se incorporar todos os pesos dentro de uma matriz  $n$ -por- $(m+n)$ . Para permitir que cada unidade tenha um peso para o bias, simplesmente pode-se incluir entre as  $m$  linhas de entrada uma outra entrada para a qual seu valor seja sempre um.

Para prosseguir, adota-se o uso de uma formulação de tempo discreta e assume-se que a rede consista inteiramente de unidades semi-lineares; isto permite estender a abordagem

para um tempo contínuo e outras formas de unidade computacional diferenciável (Williams e Zipser/89). Fazendo com que:

$$v_k(t+1) = \sum_{l \in U \cup I} w_{kl} z_l(t) \quad (3.67)$$

denote a rede de entrada para a  $k$ -ésima unidade no tempo  $t$ , para  $k \in U$ , com a sua saída no próximo passo de tempo sendo:

$$y_k(t+1) = \varphi_k(v_k(t+1)) \quad (3.68)$$

sendo que  $\varphi_k$  é a unidade da função de ativação.

Portanto, o sistema de equações (3.67) e (3.68), constituem o sistema dinâmico total da rede, onde os valores  $z_k$  são definidos pela equação (3.66). É importante notar que a entrada externa no tempo  $t$  não influencia a saída de qualquer unidade até o tempo  $t + 1$ , portanto a rede é um sistema dinâmico discreto. Usualmente, tanto as unidades (ou neurônios) escondidos e de saída terão funções de ativação não-lineares.

Algumas das unidades em  $U$  são unidades de saída, para as quais um valor desejado é definido. No entanto, pode não haver um valor desejado definido para cada simples entrada da rede.

Agora vamos derivar um algoritmo de treinamento para esta rede que é chamado de *aprendizagem temporal supervisionada*, significando que certos valores das unidades de saída devem igualar-se a valores desejados definidos em tempos específicos. Fazendo  $T(t)$  denotar o conjunto de índices  $k \in U$  para os quais existe um valor desejado especificado  $d_k(t)$  para o qual a saída da  $k$ -ésima unidade deverá igualar-se no tempo  $t$ . Então definindo um conjunto  $e$  de  $n$ -elementos variando no tempo, como:

$$e_k(t) = \begin{cases} d_k(t) - y_k(t) & \text{se } k \in T(t) \\ 0 & \text{caso contrário} \end{cases} \quad (3.69)$$

Esta formulação permite a possibilidade de que os valores desejados sejam especificados para diferentes unidades em tempos diferentes. O conjunto de unidades consideradas como “visíveis” podem, portanto, ser variantes no tempo.

Agora, fazendo com que:

$$E(t) = \frac{1}{2} \sum_{k \in U} [e_k(t)]^2 \quad (3.70)$$

denote o erro global da rede no tempo  $t$ . Por enquanto, deve-se assumir está iniciando no tempo  $t_0$  até algum tempo final  $t_1$ . O objetivo será a minimização do erro total, dado por:

$$E_{total}(t) = \sum_{t=t_0+1}^{t_1} E(t) \quad (3.71)$$

Isto pode ser feito utilizando-se o método do gradiente descendente, ajustando o peso ao longo de  $\nabla W E_{total}(t_0, t+1)$ , conforme mostrado na Figura-3.11.

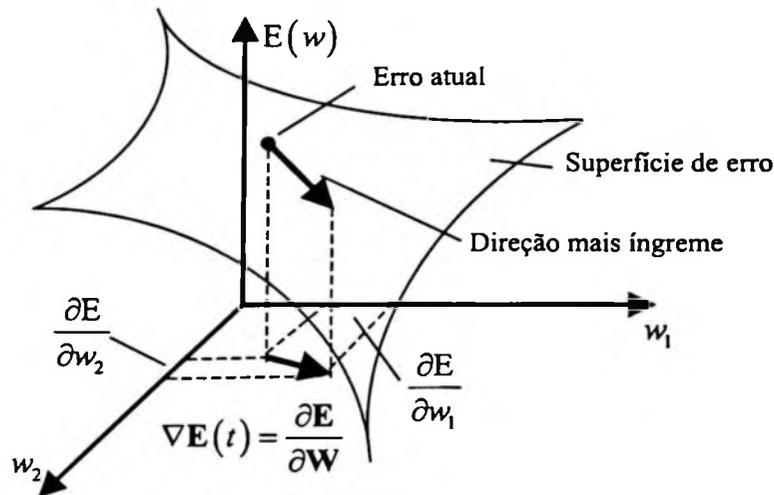


Figura-3.11 Representação gráfica do princípio da descida do gradiente (Szilas, 1997)

Uma vez que o erro total é apenas a soma dos erros nos passos individuais de tempo, uma forma para computar este gradiente é através da acumulação dos valores de  $\nabla W E(t)$  para cada passo de tempo ao longo da trajetória. A mudança de peso global para qualquer peso particular  $w_{ij}$  na rede pode, portanto, ser escrito como:

$$\Delta w_{ij} = \sum_{t=t_0+1}^{t_1} \Delta w_{ij}(t) \quad (3.72)$$

sendo que:

$$\Delta w_{ij}(t) = -\alpha \frac{\partial E(t)}{\partial w_{ij}} \quad (3.73)$$

e  $\alpha$  é alguma taxa de aprendizagem positiva.

Agora:

$$-\alpha \frac{\partial E(t)}{\partial w_{ij}} = \sum_{k \in U} e_k(t) \frac{\partial y_k(t)}{\partial w_{ij}} \quad (3.74)$$

O ponto principal para entender o algoritmo RTRL é entender o que o fator  $\frac{\partial y_k(t)}{\partial w_{ij}}$  expressa. Ele é essencialmente a medida de sensibilidade do valor de saída da unidade de entrada  $k$  no tempo  $t$  para uma pequena mudança no valor de  $w_{ij}$ , levando em conta que o efeito de uma tal mudança no peso sobre a trajetória de toda a rede de  $t_0$  até  $t$ . É importante notar que  $w_{ij}$  não precisa estar conectado a unidade  $k$ . Portanto, este algoritmo é *não local*, sendo que é necessário considerar o efeito de uma mudança em um local da rede sobre os valores computados em um outro local completamente diferente.

$\frac{\partial y_k(t)}{\partial w_{ij}}$  é computado pela diferenciação da rede dinâmica, equações (3.67) e (3.68), da

forma como segue:

$$\text{sabendo que } v_k(t+1) = \sum_{l \in U \cup I} w_{kl} z_l(t) \quad \text{e} \quad y_k(t+1) = \varphi_k(v_k(t+1))$$

então,

$$y_k(t+1) = \varphi_k \left( \sum_{l \in U \cup I} w_{kl} z_l(t) \right)$$

$$\begin{aligned} \frac{\partial y_k(t+1)}{\partial w_{ij}} &= \frac{\partial}{\partial w_{ij}} \left[ \varphi_k \left( \sum_{l \in U \cup I} w_{kl} z_l(t) \right) \right] \\ &= \varphi'_k \left( \sum_{l \in U \cup I} w_{kl} z_l(t) \right) \left[ \frac{\partial}{\partial w_{ij}} \left( \sum_{l \in U \cup I} w_{kl} z_l(t) \right) \right] \\ &= \varphi'_k(v_k(t+1)) \left[ \frac{\partial}{\partial w_{ij}} \left( \sum_{l \in U \cup I} w_{kl} z_l(t) \right) \right] \\ &= \varphi'_k(v_k(t+1)) \left[ \sum_{l \in U \cup I} \left( \frac{\partial}{\partial w_{ij}} w_{kl} z_l(t) \right) \right] \\ &= \varphi'_k(v_k(t+1)) \left[ \sum_{l \in U \cup I} \left( \frac{\partial}{\partial w_{ij}} w_{kl} z_l(t) \right) \right] \end{aligned}$$

$$= \varphi'_k(v_k(t+1)) \left[ \sum_{k \in U \cup I} \left( w_{kl} \frac{\partial z_l(t)}{\partial w_{ij}} + \frac{\partial w_{kl}}{\partial w_{ij}} z_l(t) \right) \right]$$

Neste ponto temos duas derivadas que precisam ser explicitadas, que são:

Derivada I:  $\frac{\partial z_l(t)}{\partial w_{ij}}$

e

Derivada II:  $\frac{\partial w_{kl}}{\partial w_{ij}}$

Na primeira derivada,  $\frac{\partial z_l(t)}{\partial w_{ij}}$ , é importante perceber que os dados de entrada não

dependem dos pesos da rede, portanto:

$$\frac{\partial z_l(t)}{\partial w_{ij}} = 0 \text{ para } l \in I$$

isso quer dizer que esta derivada somente existirá quando  $z_l(t)$  for uma saída, ou seja quando  $z_l(t)$  for igual a  $y_l(t)$ .

Para a segunda derivada  $\frac{\partial w_{kl}}{\partial w_{ij}}$  indica a variação do peso sináptico ligando o neurônio  $l$  ao  $k$

com relação a variação do peso sináptico ligando o neurônio  $j$  ao neurônio  $i$ . Nesse caso temos que quando  $i = k$  implica necessariamente em  $j = l$  para que a derivada exista e seja

igual a um. Ou seja,  $\frac{\partial w_{kl}}{\partial w_{ij}} = \delta_{ik} \delta_{jl}$ , sendo que  $\delta_{ab}$  é o delta de Kronecker caracterizado por:

$$\delta_{ab} = \begin{cases} 1, & \text{se } a = b \\ 0, & \text{caso contrário} \end{cases}$$

Necessariamente se tivermos  $i = k$  implica necessariamente em  $j = l$  e dessa forma

podemos suprimir  $\delta_{jl}$ , ficando apenas com  $\frac{\partial w_{kl}}{\partial w_{ij}} = \delta_{ik}$ , sendo que  $\delta_{ik}$  é o delta de

Kronecker.

Portanto, teremos que:

$$\frac{\partial y_k(t+1)}{\partial w_{ij}} = \varphi'_k(v_k(t+1)) \left[ \sum_{k \in U \cup I} w_{kl} \frac{\partial z_l(t)}{\partial w_{ij}} + \sum_{k \in U \cup I} \frac{\partial w_{kl}}{\partial w_{ij}} z_l(t) \right] \quad (3.75)$$

o segundo somatório pode ser suprimido e então o resultado final, de acordo com as explicações anteriores, será:

$$\frac{\partial y_k(t+1)}{\partial w_{ij}} = \varphi'_k(v_k(t+1)) \left[ \sum_{i \in U} w_{kl} \frac{\partial y_l(t)}{\partial w_{ij}} + \delta_{ik} z_j(t) \right] \quad (3.76)$$

A equação (3.76) é recursiva. Ou seja, sabendo-se o valor do lado esquerdo para o tempo 0, pode-se computar o valor para o tempo 1 e usar este valor para computar o valor no tempo 2, etc. Pelo fato de ter-se assumido que o estado inicial (em  $t = 0$ ) da rede não tem dependência funcional dos pesos, implica que o outro fato conhecido é a condição de contorno:

$$\frac{\partial y_k(t_0)}{\partial w_{ij}} = 0$$

Estas equações aplicam-se para todo  $k \in U, i \in U$  e  $j \in U \cup I$ .

Dessa forma cria-se um sistema dinâmico com variáveis  $p_{ij}^k$  para todo  $k \in U, i \in U$  e  $j \in U \cup I$ , dado por:

$$p_{ij}^k(t+1) = \varphi'_k(v_k(t+1)) \left[ \sum_{i \in U} w_{kl} p_{ij}^l(t) + \delta_{ik} z_j(t) \right] \quad (3.77)$$

com condições iniciais

$$p_{ij}^k(t_0) = 0 \quad (3.78)$$

seguinte-se que

$$p_{ij}^k(t) = \frac{\partial y_k(t)}{\partial w_{ij}}$$

para todo passo de tempo  $t$  e todo  $i, j$  e  $k$  apropriados.

Precisamente, o algoritmo constitui-se de computar, para cada passo de tempo  $t$  de  $t_0$  até  $t_1$ , as quantidades  $p_{ij}^k(t)$  usando as equações (3.77) e (3.78) e então usar as discrepâncias  $e_k(t)$  entre as saídas desejadas e as reais, para computar as mudanças de pesos:

$$\Delta w_{ij}(t) = \alpha \sum_{k \in U} e_k(t) p_{ij}^k(t)$$

A correção global para ser aplicada a cada peso  $w_{ij}$  na rede é, portanto, simplesmente a soma destes valores individuais  $\Delta w_{ij}(t)$  para cada passo de tempo  $t$  ao longo da trajetória.

No caso quando cada unidade na rede usar a função logística, então utiliza-se:

$$\varphi'_k(v_k(\tau)) = y_k(\tau)[1 - y_k(\tau)]$$

na equação (3.77).

A Tabela-3.3 apresenta o resumo do algoritmo de aprendizagem recorrente em tempo real (RTRL).

Algoritmo de Aprendizagem Recorrente em Tempo Real
<p>1. Definir a dimensão dos parâmetros</p> <ul style="list-style-type: none"> <li>- espaço de entrada;</li> <li>- espaço de estados;</li> <li>- espaço de saída.</li> </ul>
<p>2. Inicialização</p> <ul style="list-style-type: none"> <li>- atribuir aos pesos sinápticos valores pequenos selecionados de uma distribuição uniforme;</li> <li>- fazer o valor inicial do vetor de estado <math>x_k(t_0) = 0, k \in I</math>;</li> <li>- fazer <math>p_{ij}^k(t_0) = \frac{\partial y_k(t_0)}{\partial w_{ij}} = 0</math>.</li> </ul>
<p>3. Computar, para o passo de tempo <math>t = 0, 1, 2, \dots</math></p> $p_{ij}^k(t+1) = \varphi'_k(v_k(t+1)) \left[ \sum_{l \in U} w_{il} p_{ij}^l(t) + \delta_{ik} z_j(t) \right]$ $e_k(t) = d_k(t) - y_k(t)$ $\Delta w_{ij}(t) = \alpha \sum_{k \in U} e_k(t) p_{ij}^k(t)$

**Tabela-3.3** Algoritmo RTRL (Williams e Zipser, 1989)

### 3.6.3 Algoritmo de Retropropagação Através do Tempo (BPTT)

Nesta seção vamos derivar os algoritmos BPTT por época e o BPTT truncada para uma rede neural completamente recorrente geral.

O algoritmo de retropropagação através do tempo (*BPTT, back-propagation-through-time*) para o treinamento de uma rede recorrente é uma extensão do algoritmo de retropropagação padrão (Haykin, 2001).

#### 3.6.3.1 Algoritmo de Retropropagação Através do Tempo por Época

A abordagem de retropropagação através do tempo pode ser derivada pelo desdobramento da operação temporal de uma rede *feedforward* multicamada que cresce de uma camada a cada passo de tempo, veja a Figura- 3.12. Se o treinamento corrente é segmentado em épocas, então pode-se derivar uma versão específica que se chama *retropropagação através do tempo por época (epochwise backpropagation through time – epochwise BPTT)*. Este algoritmo é organizado da maneira como segue. Com  $t_0$  denotando o tempo inicial da época e  $t_1$  denotando o tempo final, o objetivo é computar o gradiente de  $E_{total}(t_0, t_1)$ . Isto é feito primeiro deixando a rede rodar através do intervalo  $[t_0; t_1]$  e salvando toda a história das entradas para a rede, o estado da rede e os vetores de valores desejados sobre este intervalo. Então, um único passo para trás, sobre esta história é realizado para computar o conjunto de valores  $\delta_k(\tau) = -\partial E_{total}(t_0, t_1) / \partial v_k(\tau)$ , para todo  $k \in U$  e  $\tau \in (t_0; t_1]$ , utilizando-se a equação (3.79), que é desenvolvida da seguinte maneira:

$$\frac{\partial E_{total}(t_0, t_1)}{\partial w_{kj}} = \sum_{\tau=t_0}^{t_1} \frac{\partial E_{total}(t_0, t_1)}{\partial w_{kj}(\tau)} = \sum_{\tau=t_0}^{t_1} \frac{\partial E_{total}(t_0, t_1)}{\partial v_k(\tau)} \frac{\partial v_k(\tau)}{\partial w_{kj}(\tau)} = \sum_{\tau=t_0}^{t_1} \frac{\partial E_{total}(t_0, t_1)}{\partial v_k(\tau)} x_j(\tau-1) \quad (3.79)$$

sendo que:

$$\frac{\partial E_{total}(t_0, t_1)}{\partial v_k(\tau)} = \frac{\partial E_{total}(t_0, t_1)}{\partial y_k(\tau)} \frac{\partial y_k(\tau)}{\partial v_k(\tau)} = \frac{\partial E_{total}(t_0, t_1)}{\partial y_k(\tau)} \varphi'_k(v_k(\tau)) \quad (3.80)$$

se  $\tau = t_1$  então,

$$\frac{\partial E_{total}(t_0, t_1)}{\partial y_k(\tau)} = -e_k(\tau) \quad (3.81)$$

se  $\tau < t_1$ , então,

$$\begin{aligned} \frac{\partial E_{total}(t_0, t_1)}{\partial y_k(\tau)} &= \frac{\partial E(\tau, t_1)}{\partial y_k(\tau)} = \frac{\partial e^2(\tau)}{\partial y_k(\tau)} + \frac{\partial E(\tau+1, t_1)}{\partial y_k(\tau)} = \\ &= -e_k(\tau) + \sum_{i \in U} \frac{\partial E(\tau+1, t_1)}{\partial v_i(\tau+1)} w_{ik} \\ &= -e_k(\tau) + \sum_{i \in U} \frac{\partial E(t_0, t_1)}{\partial v_i(\tau+1)} w_{ik} \end{aligned} \quad (3.82)$$

para cada  $k \in U$  e  $\tau \in [t_0, t_1]$  vamos definir o gradiente local:

$$\delta_k(\tau) = -\frac{\partial E_{total}(t_0, t_1)}{\partial v_k(\tau)}$$

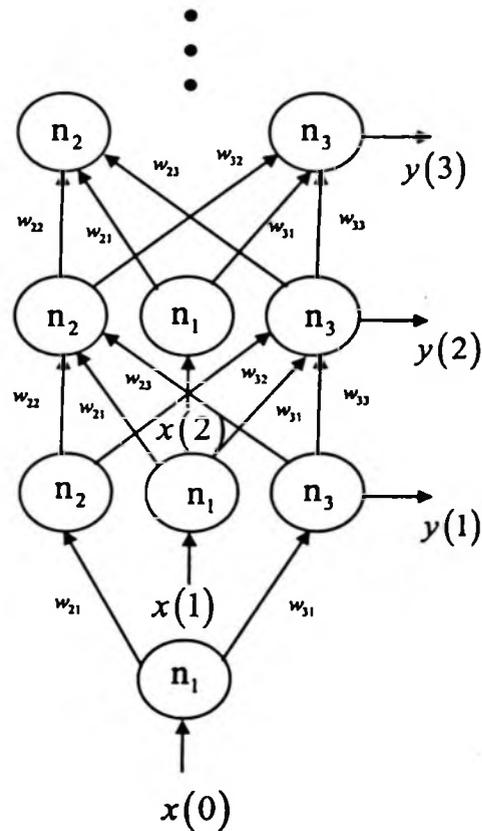
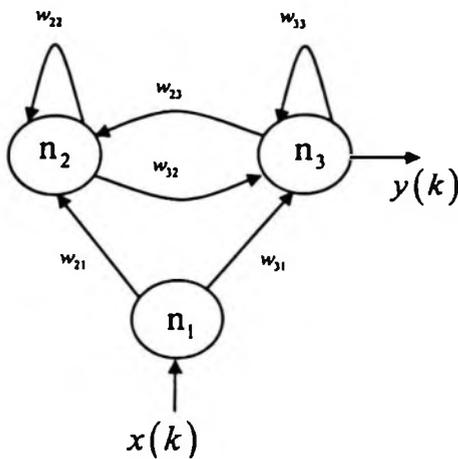


Figura - 3.12 Desdobramento para construção do algoritmo BPTT

Pelas equações (3.80), (3.81) e (3.82) temos finalmente que:

$$\delta_k(\tau) = \begin{cases} \varphi'_k(v_k(\tau))e_k(\tau) & \text{se } \tau = t_1 \\ \varphi'_k(v_k(\tau)) \left[ e_k(\tau) + \sum_{l \in U} w_{lk} \delta_l(\tau+1) \right] & \text{se } t_0 < \tau < t_1 \end{cases} \quad (3.83)$$

Isto pode ser visualizado como representando a computação de retropropagação aplicada à uma rede *feedforward* na qual os valores desejados são especificados para unidades em muitas camadas, não apenas para última, ver Figura- 3.12. O processo começa no último passo de tempo e prossegue para passos de tempo anteriores através do uso repetido destas equações. Na descrição desse algoritmo é útil destacar a “injeção dos erros” no tempo  $\tau$  significando o passo computacional da adição de  $e_k(\tau)$  para cada  $k$  à soma apropriada uando computando a expressão em colchetes na equação para  $\delta_k(\tau)$ . Também é considerada a computação de  $\delta_k(t_1)$  para envolver uma injeção de erro correspondente; neste caso, isto pode ser visualizado como a adição de  $e_k(t_1)$  a zero.

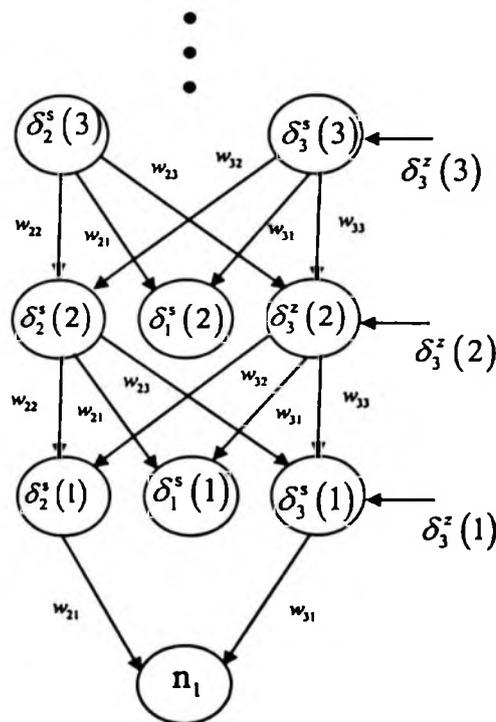


Figura – 3.13 Gradientes locais no desdobramento para construção do algoritmo BPTT

Na Figura-3.13 o índice  $s$  caracteriza o erro de atraso que vai retropropagar-se pela rede, análogo ao erro na camada escondida de um MLP. O índice  $z$  descreve o erro, o qual atua num dado instante de tempo na rede e que é subtraído do valor desejado.

Uma vez que a computação de retropropagação tenha sido realizada para o tempo  $t_0 + 1$ , as mudanças dos pesos podem ser feitas ao longo do gradiente negativo do erro global pelo uso da equação:

$$\Delta w_{ij} = -\alpha \frac{\partial E_{total}(t_0, t_1)}{\partial w_{ij}} = \alpha \sum_{\tau=t_0+1}^{t_1} \delta_i(\tau) x_j(\tau-1) \quad (3.84)$$

sendo  $\alpha$  é o parâmetro da taxa de aprendizagem e  $x_j(\tau-1)$  é a entrada aplicada à  $j$ -ésima sinapse do neurônio  $j$  no tempo  $\tau-1$ .

### 3.6.3.2 Algoritmo de Retropropagação Através do Tempo Truncada

A versão *on-line* (em tempo real) do algoritmo de aprendizagem BPTT é chamada de Retropropagação Através do Tempo Truncada (*Truncated BPTT*) o principal trabalho relatando o desenvolvimento e as características desse algoritmo é de (Williams e Peng, 1990) para redes neurais completamente recorrentes.

Neste caso a função de custo a ser minimizada é o erro quadrático instantâneo no passo de tempo atual, ou corrente, denotado por  $e^2(t)$ .

A regra de atualização, agora é dada por:

$$\Delta w_{ij} = -\alpha \sum_{\tau=t-h+1}^t \frac{\partial e^2(t)}{\partial w_{ij}(\tau)} \quad (3.85)$$

sendo que  $h$  é a história passada considerada para os cálculos. Para um cálculo exato  $h$  deverá ser igual ao número de passos de tempo ocorridos desde o começo, mas uma vez que  $h$  cresce com o passar do tempo, este método não pode ser implementado de forma *on-line*. Por essa razão,  $h$  deve ser fixado e escolhido o maior para uma melhor aproximação ao custo de uma maior complexidade (Campolucci, 1998). Qualquer informação anterior a  $h$  passos de tempo no passado é considerada irrelevante, portanto pode ser ignorada. Portanto, para que os ajustes sejam feitos continuamente, enquanto a rede está realizando o processamento, devemos salvar apenas a história relevante dos dados de entrada e do estado

da rede para um número fixo de passos de tempo, chamado de *profundidade de truncamento*:  $h$ .

Isso implica então que:

$$\frac{\partial e^2(t)}{\partial w_{ij}(\tau)} = \frac{\partial e^2(t)}{\partial v_i(\tau)} \frac{\partial v_i(\tau)}{\partial w_{ij}(\tau)} = \frac{\partial e^2(t)}{\partial v_i(\tau)} x_j(\tau-1) \quad (3.86)$$

sendo que:

$$\frac{\partial e^2(t)}{\partial v_i(\tau)} = \frac{\partial e^2(t)}{\partial y_i(\tau)} \frac{\partial y_i(\tau)}{\partial v_i(\tau)} = \frac{\partial e^2(t)}{\partial y_i(\tau)} \varphi'_i(v_i(\tau)) \quad (3.87)$$

se  $\tau = t$  então:

$$\frac{\partial e^2(t)}{\partial y_i(\tau)} = -e_i(\tau) \quad (3.88)$$

se  $\tau < t$  então:

$$\frac{\partial e^2(t)}{\partial y_i(\tau)} = \sum_{l \in U} \frac{\partial e^2(t)}{\partial s_l(\tau+1)} \frac{\partial s_l(\tau+1)}{\partial y_i(\tau)} = \sum_{l \in U} \frac{\partial e^2(t)}{\partial s_l(\tau+1)} w_{li}(\tau+1) \quad (3.89)$$

para cada  $i \in U$  e  $\tau \in [t_0, t_1]$  vamos definir o gradiente local:

$$\delta_i(\tau) = -\frac{\partial e^2(t)}{\partial s_i(\tau)}$$

Pelas equações (3.86), (3.87) e (3.88) temos finalmente que:

$$\delta_i(\tau) = \begin{cases} \varphi'_i(v_i(\tau)) e_i(\tau) & \text{se } \tau = t \\ \varphi'_i(v_i(\tau)) \sum_{l \in U} w_{li}(\tau+1) \delta_l(\tau+1) & \text{se } \tau < t \end{cases} \quad (3.90)$$

Utilizando a expressão (3.85) a regra de atualização dos pesos torna-se:

$$\Delta w_{ij} = -\alpha \sum_{\tau=t-h+1}^t \frac{\partial e^2(t)}{\partial w_{ij}(\tau)} = \alpha \sum_{\tau=t-h+1}^t \delta_i(\tau) x_j(\tau-1) \quad (3.91)$$

Por motivo de simplicidade, freqüentemente os pesos não são memorizados na prática mas apenas o valor atual do peso é utilizado com a mesma performance (Williams e Peng, 1990), dessa forma (3.91) torna-se:

$$\delta_i(\tau) = \begin{cases} \varphi'_i(v_i(\tau))e_i(\tau) & \text{se } \tau = t \\ \varphi'_i(v_i(\tau)) \sum_{l \in U} w_{li} \delta_l(\tau+1) & \text{se } \tau < t \end{cases} \quad (3.92)$$

Os algoritmos BPTT por época e o de BPTT truncada diferem na regra de atualização e no cálculo das variáveis  $\delta_i(\tau)$ , compare (3.84) com (3.91) e (3.83) com (3.90). Uma vez que no caso em tempo real os parâmetros são atualizados para cada passo de tempo, o sinal de erro injetado, ou seja, o termo  $e_i(\tau)$  em (3.83), deve ser computado apenas no passo de tempo corrente, por essa razão é que ele não está presente na segunda linha de (3.90).

### 3.6.4 Filtro de Kalman e as Redes Recorrentes de Perceptrons de Múltiplas Camadas

Nesta dissertação, vamos analisar a utilização do Filtro de Kalman (Kalman, 1960) para realizar o treinamento de redes neurais recorrentes de perceptrons de múltiplas camadas (RMLP – *recurrent multilayer perceptron*). Para descrever como proceder com esta abordagem, o Filtro de Kalman (FK) e o Filtro de Kalman Extendido são descritos em detalhe.

#### 3.6.4.1 Filtro de Kalman

Assumindo um sistema linear e que o ruído é branco e Gaussiano, o filtro de Kalman é ótimo no sentido em que o erro quadrático médio (MSE) é minimizado. O filtro de Kalman consiste em um conjunto de equações matemáticas que computam recursivamente o estado estimado de um sistema dinâmico linear no tempo  $t$ ,  $\mathbf{x}(t)$ , partindo das estimativas prévias do estado no tempo  $t-1$ ,  $\mathbf{x}(t-1)$  e dos novos dados de entrada  $\mathbf{y}(t)$ . Uma vez que ele é recursivo, não há necessidade de armazenar todos os dados passados observados para estimar o estado do sistema. Esta é uma propriedade importante e permite a possibilidade de implementação em tempo real.

O estado de um sistema é simplesmente um vetor coluna contendo variáveis que são de interesse do analista. Frequentemente, estas variáveis (ou, os estados) não podem ser

observadas diretamente. Pelo contrário, elas devem ser inferidas de um conjunto de medidas (muito provavelmente contaminadas com ruído) tomadas do sistema. Usando o filtro de Kalman estendido para treinar uma rede neural o vetor de estado deverá consistir dos pesos sinápticos da rede. Estas são as variáveis pelas quais temos interesse. Este é o objetivo do filtro de Kalman, usar medidas tomadas do sistema para deduzir as melhores estimativas para as variáveis de estado (Patel, 2000).

Para alicerçar o método do filtro de Kalman, considere o modelo de espaço de estado linear de um sistema dinâmico de tempo discreto descrito pelas seguintes duas equações:

$$\mathbf{y}(t) = \mathbf{H}(t)\mathbf{x}(t) + \mathbf{r}(t) \quad (3.93)$$

$$\mathbf{x}(t+1) = \mathbf{F}(t+1,t)\mathbf{x}(t) + \mathbf{q}(t) \quad (3.94)$$

sendo que,

- $\mathbf{y}(t)$  :  $M \times 1$  vetor de medida no tempo  $t$
- $\mathbf{H}(t)$  :  $M \times N$  matriz de medida no tempo  $t$
- $\mathbf{r}(t)$  :  $M \times 1$  ruído de medida no tempo  $t$
- $\mathbf{R}(t)$  :  $M \times M$  matriz de covariância do ruído de medida
- $\mathbf{x}(t)$  :  $N \times 1$  estado do sistema no tempo  $t$
- $\mathbf{x}(t+1)$  :  $N \times 1$  estado do sistema no tempo  $t + 1$
- $\mathbf{F}(t+1,t)$  :  $N \times N$  matriz de transição do estado do tempo  $t$  para o  $t + 1$
- $\mathbf{q}(t)$  :  $N \times 1$  ruído do processo no tempo  $t$
- $\mathbf{Q}(t)$  :  $N \times N$  matriz de covariância do ruído do processo

As Equações (3.93) e (3.94) são conhecidas como equações de medida e de processo, respectivamente. Nas Equações (3.93) e (3.94), tanto  $\mathbf{r}(t)$  como  $\mathbf{q}(t)$  são assumidos como sendo seqüências de ruído branco de média zero com matrizes de covariância diagonal (ou correlação)  $\mathbf{R}(t)$  e  $\mathbf{Q}(t)$  respectivamente, descritas como:

$$\mathbf{R}(t) = E[\mathbf{r}(t)\mathbf{r}^T(t)] \quad (3.95)$$

$$\mathbf{Q}(t) = E[\mathbf{q}(t)\mathbf{q}^T(t)] \quad (3.96)$$

Na formulação do filtro de Kalman é assumido que o sistema sob consideração pode ser descrito pelo modelo de espaço de estado acima e que os parâmetros do modelo  $\mathbf{F}(t+1,t)$ ,  $\mathbf{H}(t)$ ,  $\mathbf{Q}(t)$  e  $\mathbf{R}(t)$  são conhecidos. Na prática,  $\mathbf{Q}(t)$  e  $\mathbf{R}(t)$  podem não ser conhecidos mas devem também serem estimados considerando a informação *a priori* disponível ou serem aproximadas pela experiência específica que se tem sobre a aplicação (Patel, 2000).

---

Computar o ganho de Kalman

$$\mathbf{K}(t) = \frac{\mathbf{P}(t|t-1)\mathbf{H}^T(t)}{\mathbf{H}(t)\mathbf{P}(t|t-1)\mathbf{H}^T(t) + \mathbf{R}(t)} \quad (3.97)$$

Atualizar (corrigir) Medidas

$$\hat{\mathbf{x}}(t|t) = \hat{\mathbf{x}}(t|t-1) + \mathbf{K}(t)[\mathbf{y}(t) - \mathbf{H}(t)\hat{\mathbf{x}}(t|t-1)] \quad (3.98)$$

$$\mathbf{P}(t|t) = \mathbf{P}(t|t-1) - \mathbf{K}(t)\mathbf{H}(t)\mathbf{P}(t|t-1) \quad (3.99)$$

Atualizar tempo (predição)

$$\hat{\mathbf{x}}(t+1|t) = \mathbf{F}(t+1,t)\hat{\mathbf{x}}(t|t) \quad (3.100)$$

$$\mathbf{P}(t+1|t) = \mathbf{F}(t+1,t)\mathbf{P}(t|t)\mathbf{F}^T(t+1,t) + \mathbf{Q}(t) \quad (3.101)$$

---

Tabela-3.4 Equações do filtro de Kalman

As equações do filtro de Kalman na Tabela-3.4 podem ser classificadas em dois grupos: equações de atualização de tempo e equações de atualização de medida. Estas equações mantêm ligadas as estimativas do estado  $\hat{\mathbf{x}}$  e a estimativa da matriz de correlação estado-erro  $\mathbf{P}$ . As equações de atualização de medida corrigem as estimativas baseadas nas novas medidas  $\mathbf{y}(t)$  que é feita disponível no passo  $t$ . As equações de atualização de tempo predizem valores das estimativas no futuro (um-passo à frente) antes que a próxima medida seja tornada disponível. Este processo é repetido recursivamente até a estimativa de estado convergir.

Com a utilização desta notação temos então que,  $\mathbf{x}(t)$  é o estado verdadeiro do sistema a ser determinado,  $\hat{\mathbf{x}}(t|t-1)$  refere-se a estimativa de estado *a priori* no passo  $t$  dado o

conhecimento das medidas até o passo  $t - 1$ , e  $\hat{\mathbf{x}}(t|t)$  refere-se à estimativa de estado a posteriori no passo  $t$  dadas as medidas até o passo  $t$ . Similarmente,  $\mathbf{P}(t|t-1)$  é a matriz de correlação estado-erro a priori e  $\mathbf{P}(t|t)$  é a matriz de correlação estado-erro a posteriori, descrita como segue:

$$\mathbf{P}(t|t-1) = E \left[ (\mathbf{x}(t) - \hat{\mathbf{x}}(t|t-1)) (\mathbf{x}(t) - \hat{\mathbf{x}}(t|t-1))^T \right] \quad (3.102)$$

$$\mathbf{P}(t|t) = E \left[ (\mathbf{x}(t) - \hat{\mathbf{x}}(t|t)) (\mathbf{x}(t) - \hat{\mathbf{x}}(t|t))^T \right] \quad (3.103)$$

Para iniciar o procedimento de interação, descrito acima, a estimativa de estado a priori e a estimativa de covariância estado-erro a priori devem primeiro ser inicializadas com valores apropriados,  $\hat{\mathbf{x}}(1|0)$  e  $\mathbf{P}(1|0)$ , respectivamente.

No centro do filtro de Kalman está a noção que quando uma nova observação ou medida é tornada disponível, tanto o vetor de estado a priori  $\hat{\mathbf{x}}(t|t-1)$  como a matriz de covariância estado-erro a priori  $\mathbf{P}(t|t-1)$  são atualizadas. Isto é realizado via as equações de atualização de medida. As equações de atualização de medida obtêm as estimativas a posteriori das estimativas a priori baseadas nas novas medidas  $\mathbf{y}(t)$  no passo  $t$ . O termo de diferença  $\mathbf{y}(t) - \mathbf{H}(t)\hat{\mathbf{x}}(t|t-1)$  na Equação (3.98) é conhecido como o *vetor de inovações*, uma vez ele representa novas informações na medida atual  $\mathbf{y}(t)$  quando comparada a medida predita  $\mathbf{H}(t)\hat{\mathbf{x}}(t|t-1)$  (Haykin, 1996). Próximo, as equações de atualização de tempo predizem as estimativas a priori no passo  $t + 1$  baseadas nas estimativas a posteriori nos passos prévios  $t$  e então o procedimento inteiro é repetido. Juntas, as Equações (3.99) e (3.101) são chamadas *equações de diferença finita de Riccati* e elas computam recursivamente a matriz de correlação estado-erro  $\mathbf{P}$  (Haykin, 1996).

Este procedimento recursivo direto e de atualização leva à minimização do traço da matriz de covariância estado-erro  $\mathbf{P}$ . Uma vez que as diagonais da matriz  $\mathbf{P}$  são as médias quadradas dos erros no estado, como é dado pelas Equações (3.102) e (3.103), isto significa que o filtro de Kalman encontra a estimativa de erro quadrático mínimo do estado.

A entrada para o filtro de Kalman é o vetor de medida no tempo  $t$ ,  $y(t)$ . Portanto, para operar o filtro de Kalman a cada passo de tempo, a medida correta deve ser tornada disponível. Por essa razão, pode-se dizer que o filtro de Kalman é de fato um algoritmo supervisionado.

### 3.6.4.2 Filtro de Kalman Estendido

Há duas suposições básicas na derivação do filtro de Kalman. A primeira é que o sistema é descrito por um modelo de espaço de estado linear e a segunda é que os ruídos são brancos e Gaussianos com média zero, e eles também são assumidos serem não correlacionados um com o outro e com o estado inicial. Quando estas suposições são satisfeitas, o filtro de Kalman é ótimo no sentido do erro quadrático médio. Quando o sistema sob consideração é não linear, a primeira condição é violada e o filtro de Kalman estendido (EKF – *extended Kalman filter*) é aplicado como um filtro sub-ótimo. No EKF os termos não lineares são aproximados pelos termos lineares de primeira ordem pela expansão em uma série de Taylor. Para começar, considere um sistema não linear descrito como:

$$y(t) = h_t(t) x(t) + r(t) \quad (3.104)$$

$$x(t+1) = f_t(t+1, t) x(t) + q(t) \quad (3.105)$$

As Equações (3.104) e (3.105) são as contra-partidas não lineares das Equações (3.93) e (3.94). Elas são as equações de medida e processo para o caso não linear, onde,  $h_n$  e  $f_n$  são funções de vetores não lineares do estado. Procedendo tal como foi afirmado acima, nós linearizamos ambas as equações usando a expansão por série de Taylor sobre  $\hat{x}(t|t-1)$  e  $\hat{x}(t|t)$  como segue, respectivamente:

$$h_t(x(t)) = h_t(\hat{x}(t|t-1)) + H_t(t)(x(t) - \hat{x}(t|t-1)) + \dots \quad (3.106)$$

$$f_t(x(t)) = f_t(\hat{x}(t|t)) + F_t(t+1, t)(x(t) - \hat{x}(t|t)) + \dots \quad (3.107)$$

onde as matrizes Jacobianas  $F_t(t+1, t)$  e  $H_t(t)$  são definidas como:

$$F_t(t+1, t) = \frac{\partial f_t(\hat{x}(t|t))}{\partial x} \quad \text{e} \quad H_t(t) = \frac{\partial h_t(\hat{x}(t|t-1))}{\partial x}$$

Ignorando os termos de maior ordem nas expansões de Taylor acima, as equações de medida e processo não lineares podem ser aproximadas como segue:

$$\mathbf{y}(t) = \mathbf{H}_i(t) \mathbf{x}(t) + \mathbf{u}(t) + \mathbf{r}(t) \quad (3.108)$$

$$\mathbf{x}(t+1) = \mathbf{F}_n(t+1, t) \mathbf{x}(t) + \mathbf{v}(t) + \mathbf{q}(t) \quad (3.109)$$

sendo que:

$$\mathbf{u}(t) = h_i(\hat{\mathbf{x}}(t|t-1)) - \mathbf{H}_n(t) \hat{\mathbf{x}}(t|t-1)$$

$$\mathbf{v}(t) = f_i(\hat{\mathbf{x}}(t|t-1)) - \mathbf{F}_i(t+1, t) \hat{\mathbf{x}}(t|t-1)$$

Computar o ganho de Kalman

$$\mathbf{K}(t) = \frac{\mathbf{P}(t|t-1) \mathbf{H}_i^T(t)}{\mathbf{H}_i(t) \mathbf{P}(t|t-1) \mathbf{H}_i^T(t) + \mathbf{R}(t)} \quad (3.110)$$

Atualizar (corrigir) Medidas

$$\hat{\mathbf{x}}(t|t) = \hat{\mathbf{x}}(t|t-1) + \mathbf{K}(t) [\mathbf{y}(t) - h_i \hat{\mathbf{x}}(t|t-1)] \quad (3.111)$$

$$\mathbf{P}(t|t) = \mathbf{P}(t|t-1) - \mathbf{K}(t) \mathbf{H}_i(t) \mathbf{P}(t|t-1) \quad (3.112)$$

Atualizar tempo (predição)

$$\hat{\mathbf{x}}(t+1|t) = f_i \hat{\mathbf{x}}(t|t) \quad (3.113)$$

$$\mathbf{P}(t+1|t) = \mathbf{F}_i(t+1, t) \mathbf{P}(t|t) \mathbf{F}_i^T(t+1, t) + \mathbf{Q}(t) \quad (3.114)$$

Tabela-3.5 Equações do filtro de Kalman estendido

Comparando as equações do filtro de Kalman na Tabela-3.4 com aquelas do filtro de Kalman estendido na Tabela-3.5, apenas umas poucas diferenças são notadas. Primeiro, os termos lineares  $\mathbf{H}(t) \hat{\mathbf{x}}(t|t-1)$  e  $\mathbf{F}_i(t+1, t) \hat{\mathbf{x}}(t|t)$  no filtro de Kalman são substituídas por  $h_i(\hat{\mathbf{x}}(t|t-1))$  e  $f_i(\hat{\mathbf{x}}(t|t-1))$  no EKF, respectivamente. A matriz de transição de estado  $\mathbf{F}(t+1, t)$  e a matriz de medida  $\mathbf{H}(t)$  no filtro de Kalman também são substituídas pelas matrizes Jacobianas  $\mathbf{F}_i(t+1, t)$  e  $\mathbf{H}_i(t)$  no EKF, respectivamente. As matrizes de derivadas devem ser recalculadas para cada iteração do filtro de Kalman.

### 3.6.4.3 Aplicação do EKF a Redes Neurais

O algoritmo do filtro de Kalman estendido foi primeiro aplicado para o treinamento de perceptrons multicamadas alimentados adiante por (Singhal e Wu, 1989). Eles mostram que o algoritmo convergia em menos iterações que o algoritmo de retropropagação padrão usando uns poucos exemplos artificiais. Eles também mostraram que em alguns casos que quando o algoritmo de retropropagação falhava, o EKF convergia para uma boa solução. Seguindo esta demonstração, muitos autores usaram este método para treinar redes neurais. Puskorius e Feldkamp (1991) aplicaram o algoritmo para o treinamento de redes neurais recorrentes para problemas reais na área de engenharia. Williams (1992) também estudou o seu uso em redes recorrentes. Williams (1992) examina a relação entre EKF e os algoritmos de gradiente mais amplamente utilizados o RTRL e o BPTT. É observado que o algoritmo RTRL é uma forma degenerada do método EKF dando um considerável aumento de velocidade em termos dos passos requeridos para convergência quando comparado com o BPTT e o RTRL.

O treinamento dos pesos em redes neurais pode ser abordado como um problema de estimação de estado (Patel, 2000). As redes neurais operam inerentemente de uma forma não linear e de tal forma que esta estimação de estado é realizada via o filtro de Kalman estendido como contrário ao algoritmo do filtro de Kalman, o qual é aplicada para sistemas lineares. Quando aplica-se o algoritmo EKF para o treinamento de redes neurais, os pesos da rede são organizados como um vetor  $w(t)$  e são tratados como o vetor de estado  $x(t)$  do sistema, descrito acima. Dessa forma, o vetor de estado poderá ser referido como  $w(t)$  em lugar de  $x(t)$ . O algoritmo EKF atualiza o estado de um passo de tempo para um outro e também atualiza a matriz  $P$  de correlação estado-erro como descrito anteriormente.

Primeiro, considere uma rede neural arbitrária (*feedforward* ou recorrente) dado um vetor de objetivos desejado  $d(t)$  de tamanho  $n_0 \times 1$  para nós de saída em cada passo. O objetivo é encontrar os pesos de tal forma que a diferença entre a saída da rede e o valor desejado (isto é, o erro) seja minimizado no sentido do erro quadrático médio. Seja uma rede com um total de  $t_w$  pesos (incluindo os pesos de bias) e  $t_0$  unidades de saída que têm valores desejados associados a elas. Pode-se lançar o problema de treinamento da rede neural

dentro de uma estrutura de EKF, primeiro anotando as equações não lineares que governam a operação da rede. O modelo da rede neural do sistema de espaço de estado é então dado por:

$$\mathbf{d}(t) = h_i(\mathbf{w}(t), \mathbf{u}(t)) + \mathbf{r}(t) \quad (3.115)$$

$$\mathbf{w}(t+1) = \mathbf{w}(t) \quad (3.116)$$

Comparando as Equações (3.115) e (3.116) com as Equações de medida e processo (3.104) e (3.105), notamos que  $\mathbf{F}_i(t+1, t) = I$ , onde  $I$  é a matriz identidade e  $\mathbf{Q} = 0$ . Também notamos que o vetor de valores desejados  $\mathbf{d}(t)$  é tratado como o vetor de medida na terminologia Kalman e que a função vetor não linear  $h_i$  descreve a relação entre as entradas, pesos e saídas da rede neural. Note que  $\mathbf{u}(t)$  é o vetor de entrada para a rede no passo  $t$ . Com a formulação de espaço de estado de uma rede neural arbitrária, dada pelas equações (3.115) e (3.116), as equações do EKF da Tabela-3.5 podem ser escritas como:

$$\mathbf{K}(t) = \frac{\mathbf{P}(t|t-1) \mathbf{H}_i^T(t)}{\mathbf{H}_i(t) \mathbf{P}(t|t-1) \mathbf{H}_i^T(t) + \mathbf{R}(t)} \quad (3.117)$$

$$\hat{\mathbf{w}}(t|t) = \hat{\mathbf{w}}(t|t-1) + \mathbf{K}(t) [\mathbf{d}(t) - h_i(\hat{\mathbf{w}}(t|t-1), \mathbf{u}(t))] \quad (3.118)$$

$$\mathbf{P}(t|t) = \mathbf{P}(t|t-1) - \mathbf{K}(t) \mathbf{H}_i(t) \mathbf{P}(t|t-1) \quad (3.119)$$

$$\hat{\mathbf{w}}(t+1|t) = \hat{\mathbf{w}}(t|t) \quad (3.120)$$

$$\mathbf{P}(t+1|t) = \mathbf{P}(t|t) \quad (3.121)$$

pode ser mais simplificado em uma forma compacta, primeiro simplificando a notação como:

$$\hat{\mathbf{w}}(t) = \hat{\mathbf{w}}(t|t) \quad \text{e} \quad \mathbf{P}(t) = \mathbf{P}(t|t)$$

e então substituindo as Equações (3.120) e (3.121) nas Equações (3.117), (3.118) e (3.119), resultando em:

$$\mathbf{K}(t) = \frac{\mathbf{P}(t) \mathbf{H}_i^T(t)}{\mathbf{H}_i(t) \mathbf{P}(t) \mathbf{H}_i^T(t) + \mathbf{R}(t)} \quad (3.122)$$

$$\hat{\mathbf{w}}(t+1) = \hat{\mathbf{w}}(t) + \mathbf{K}(t) [\mathbf{d}(t) - h_i(\hat{\mathbf{w}}(t), \mathbf{u}(t))] \quad (3.123)$$

$$\mathbf{P}(t+1) = \mathbf{P}(t) - \mathbf{K}(t) \mathbf{H}_i(t) \mathbf{P}(t) \quad (3.124)$$

Para linearizar a Equação (3.115), a matriz Jacobiana  $H_t(t)$  deve ser calculada como:

$$H_t(t) = \frac{\partial h_n(\mathbf{w}(t), \mathbf{u}(t))}{\partial \mathbf{w}}$$

$$= \begin{bmatrix} \frac{\partial h_1}{\partial w_1} & \frac{\partial h_1}{\partial w_2} & \dots & \frac{\partial h_1}{\partial w_{i_w}} \\ \frac{\partial h_2}{\partial w_1} & \frac{\partial h_2}{\partial w_2} & \dots & \frac{\partial h_2}{\partial w_{i_w}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_{i_0}}{\partial w_1} & \frac{\partial h_{i_0}}{\partial w_2} & \dots & \frac{\partial h_{i_0}}{\partial w_{i_w}} \end{bmatrix} \quad (3.125)$$

sendo que,  $h_t(\bullet) = [h_1, h_2, \dots, h_{i_0}]$  são as  $i_0$  saídas da rede, e  $\mathbf{w} = [w_1, w_2, \dots, w_{i_w}]$  são os  $i_w$  pesos ordenados da rede. As derivadas são avaliadas na estimativa corrente do estado  $\hat{\mathbf{w}} = (t)$ . A relação  $h_t(\bullet)$  depende da arquitetura da rede que é empregada e desempenha um papel crucial no método pelo qual a matriz Jacobiana  $H_t$  é calculada. A diferença entre a aplicação do EKF para o treinamento de redes feedforward e recorrentes reside no cálculo da Jacobiana (Haykin, 2001; Werbos, 1990; Puskorius e Feldkamp, 1994 e Patel, 2000).

## Capítulo 4

### Resultados Experimentais

Nesse capítulo vamos analisar alguns resultados empíricos obtidos a partir da utilização de metodologias econométricas tradicionais e de redes neurais para modelagem da série temporal dos retornos do IBOVESPA para realizar previsão.

Para a análise da série temporal utilizaremos modelos ARIMA e modelos ARCH, no caso das redes neurais usaremos redes neurais artificiais recorrentes de William e Zipser, estas redes neurais são chamadas de redes de aprendizagem recorrente em tempo real (ARTR).

#### 4.1 Descrição da série de retornos do IBOVESPA diário

A série do IBOVESPA apresentada na Figura-4.1 é claramente não estacionária, notamos a presença de fortes oscilações em determinados períodos, como por exemplo em setembro de 2001 (atentados terroristas aos Estados Unidos) e outubro de 2002 (fato).

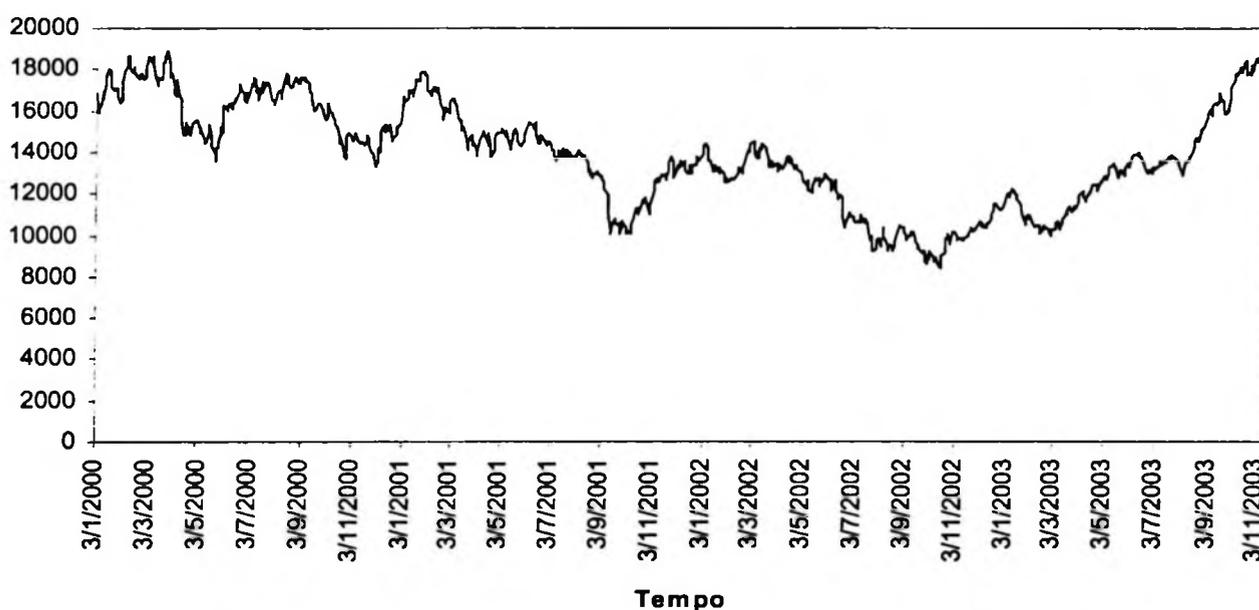
Aplicando a primeira diferença ao logaritmo da série obtemos a série de retornos do IBOVESPA para dados diários, que nesse caso chamamos de DLIBOVESPA.

O comportamento da série DLIBOVESPA é apresentado na Figura-4.2. Podemos identificar nessas figuras que flutuações de mesma magnitude do retorno tendem a se agrupar, ou seja, grandes variações tendem a ser seguidas por outras grandes variações, o mesmo acontecendo para o caso de pequenas variações sendo seguidas por pequenas

variações. De acordo com Enders (1995, pág. 145) uma das características dos modelos ARCH é a de capturar o efeito onde: choques grandes (em valores absolutos) estão associados com aumento persistente da variância da seqüência de erros.

Para avaliar a eficiência dos métodos de previsão, foram estimados alguns modelos para prever o retorno do IBOVESPA. A série estuda refere-se ao valor do fechamento do IBOVESPA no período de 3 de janeiro de 2000 até 18 de novembro de 2003, no total são 965 valores.

Para realizar as análises de desempenho de previsão de cada um dos métodos estudados a série foi dividida em duas partes: uma para aprendizagem correspondendo a 955 valores e uma parcela para previsão no total de 10 valores.



**Figura-4.1** Série de pontos de fechamento do IBOVESPA diário (período de 03/01/2003 até 18/11/2003)

A seguir é mostrada a série resultante da aplicação do log e da primeira diferença à série de fechamento do IBOVESPA diário mostrado na Figura-4.1.

A Figura-4.2 representa a série DLIBOVESPA que é a base para o modelamento ARIMA-ARCH.

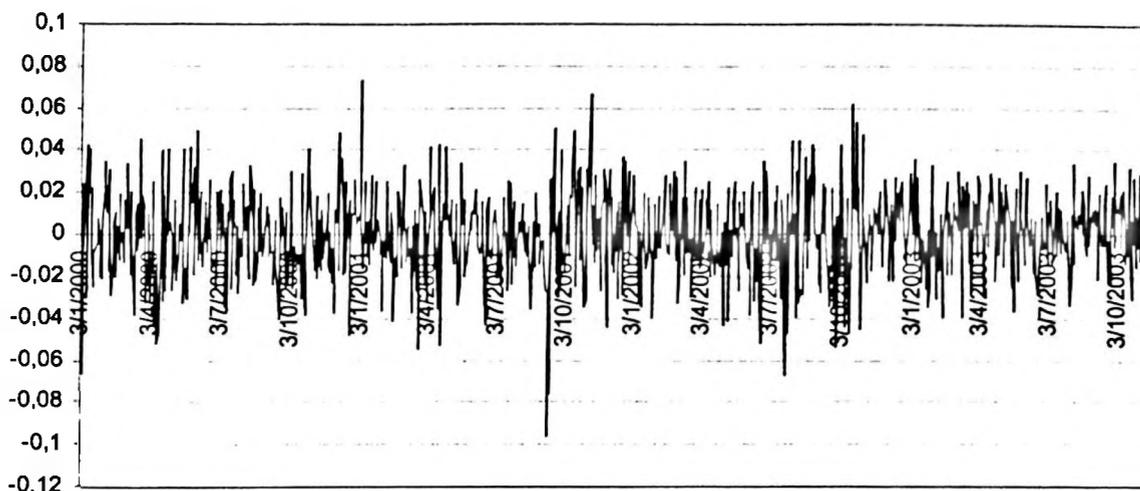


Figura-4.2 Série dos retornos do IBOVESPA diário (período de 03/01/2003 até 18/11/2003)

### 4.1.1 Testes de raiz unitária

A seguir são apresentados os testes para verificar se a série DLIBOVESPA apresenta uma ou múltiplas raízes unitárias.

Para o teste de uma raiz unitária foi utilizado o teste de Dickey-Fuller<sup>1</sup> (ADF), que apresentou os seguintes resultados, utilizando o lag de 5:

Estatística do teste ADF*	valores críticos 1%	-2,5679
	valores críticos 5%	-1,9397
	valores críticos 10%	-1,6158
-13,13676		

\*Sem intercepto e sem constante

Estatística do teste ADF**	valores críticos 1%	-3,4399
	valores críticos 5%	-2,8650
	valores críticos 10%	-2,5686
-13,13189		

\*\*Com intercepto e sem constante

Estatística do teste ADF***	valores críticos 1%	-3,9726
	valores críticos 5%	-3,4168
	valores críticos 10%	-3,1304
-13,22663		

\*\*\*Com intercepto e com constante

Portanto, rejeitamos a hipótese de raiz unitária pelo teste de Dickey-Fuller Aumentado.

Para verificar a presença de múltiplas raízes unitárias utilizamos o teste de Dikey-Pantula (1987), que apresentou os seguintes resultados com 5% de nível de significância:

$$^1 \Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=2}^p \delta_i y_{t-i+1} + \varepsilon_t$$

Hipóteses testadas	Estatística	valores críticos a 5%
$H_0 : 3RU$	-66,78856	-2,76
$H_a : 2RU$		
$H_0 : 2RU$	-21,92759	-3,37
$H_a : 1RU$	-37,83719	-3,77
$H_0 : 1RU$	-30,87212	-3,42
$H_a : 0RU$	-14,06487	-3,80
	-17,74525	-4,16

Valores críticos obtidos a partir de (Fuller, 1976).

Portanto, temos que todas as hipóteses de raízes unitárias são rejeitadas para a série DLIBOVESPA.

#### 4.1.2 Estimação dos modelos ARIMA-ARCH

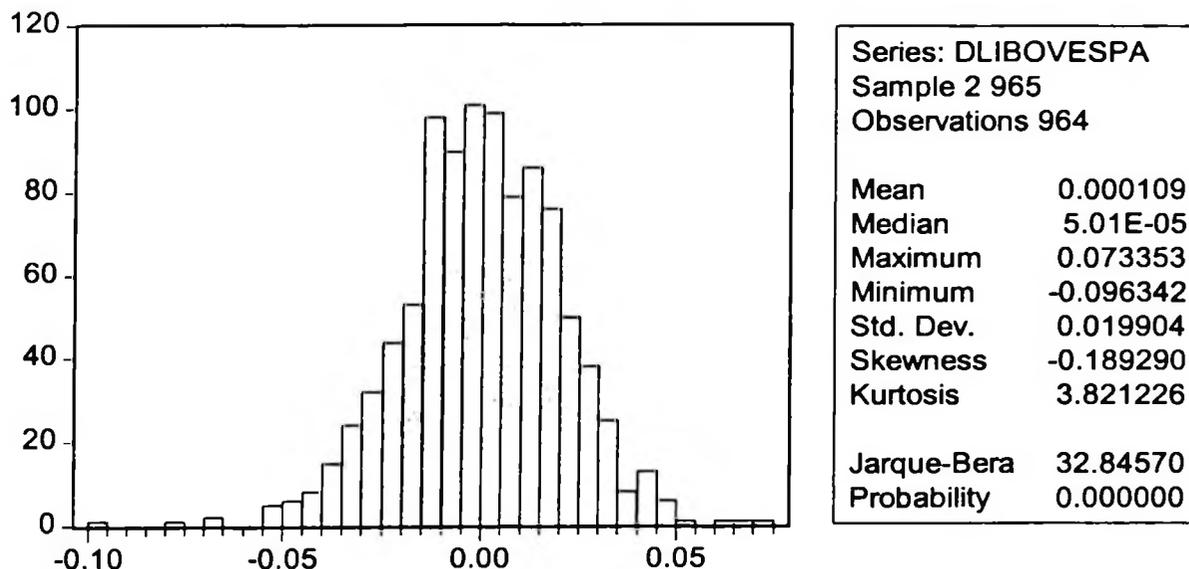


Figura-4.3 Estatísticas dos retornos do IBOVESPA diário

Na Figura-4.3 temos as estatísticas da série dos retornos do fechamento do IBOVESPA diário, verificamos que rejeita-se a hipótese de normalidade pelo teste de Jarque-Bera, observamos excesso de curtose e assimetria negativa. Estas constatações indicam a necessidade de especificação de um modelo ARCH (Engle et al., 1987).

Para estimação dos modelos ARIMA-ARCH vamos realizar a identificação a partir do correlograma da série DLIBOVESPA.

A Figura-4.4 apresenta o correlograma (para 25 lags) da série DLIBOVESPA diário.

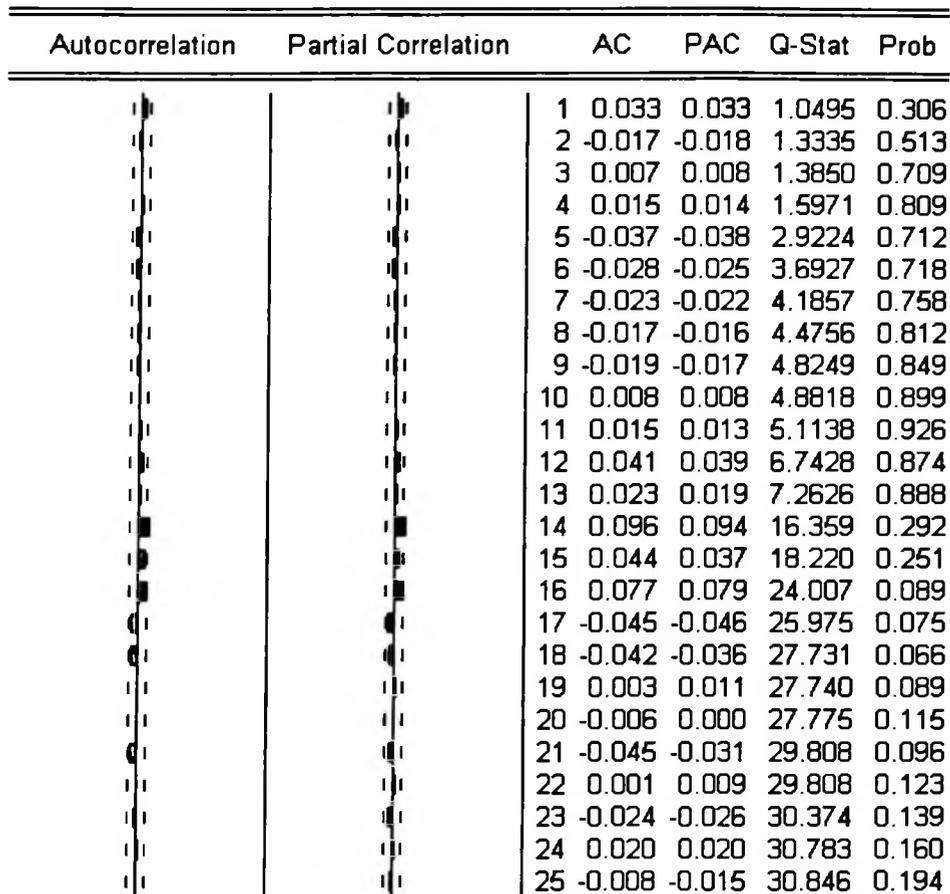


Figura-4.4 Correlograma da série de retornos DLIBOVESPA diário

Os modelos da família ARIMA que serão estimados para representar a série DLIBOVESPA são: AR([14]), AR([16]) e AR([14], [16]).

Estes modelos também podem ser escritos como: ARIMA([14], 1, 0), ARIMA([16], 1, 0) e ARIMA([14], [16], 1, 0), respectivamente.

Estimação do modelo AR(16).

Variable	Coefficient	Std. Error	t-Statistic	Prob.	Akaike info criterion	-5,008156
AR([16])	0,080447	0,032257	2,493972	0,0128	Schwarz criterion	-5,002992

Estimação do modelo AR(14).

Variable	Coefficient	Std. Error	t-Statistic	Prob.	Akaike info criterion	-5,013936
AR([14])	0,101066	0,032154	3,143148	0,0017	Schwarz criterion	-5,008780

Estimação do modelo AR(14, 16).

Variable	Coefficient	Std. Error	t-Statistic	Prob.	Akaike info criterion	-5,016765
AR([14])	0,102725	0,032310	3,179333	0,0015	Schwarz criterion	-5,006438
AR([16])	0,082285	0,032106	2,562897	0,0105		

A Figura-4.5 apresenta o correlograma (para 23 lags) dos resíduos quadrados da série de retornos DLIBOVESPA diário.

Verificamos a presença de auto-correlação dos resíduos quadrados da série DLIBOVESPA para as três formas autoregressivas: AR([14]), AR([16]) e AR([14], [16]). A Figura-4.5 apresenta apenas o correlograma do modelo AR([14], [16]).

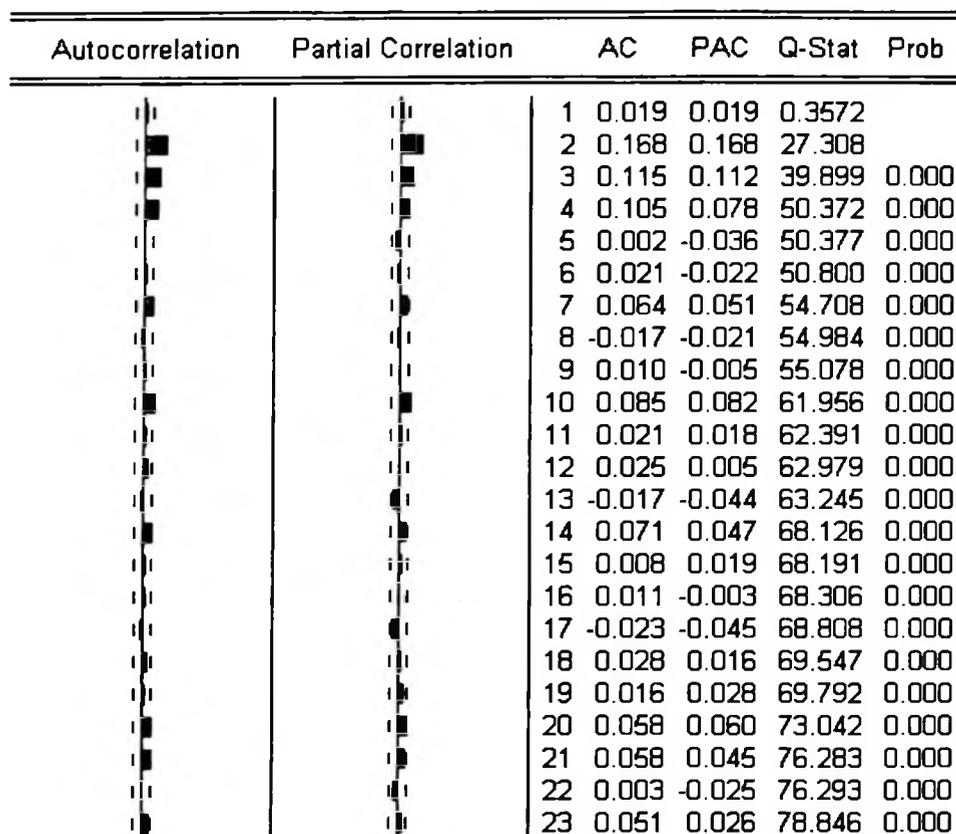


Figura-4.5 Correlograma dos resíduos quadrados da série DLIBOVESPA, modelo AR([14], [16])

O teste ARCH-LM foi proposto por Engle (1982) para testar a presença de heterocedasticidade condicional auto-regressiva nos resíduos da série.

Para a série DLIBOVESPA obtemos o seguinte resultado:

Teste ARCH-LM com 4 lags.

ARCH Test:  
Obs\*R-squared            41,46725    Probability            0,000000

Verificamos que apresenta heterocedasticidade condicional.

### Modelos ARCH

Análise dos parâmetros estimados do modelo AR([14], [16])-ARCH(3)

$$RIBOVESPA_t = \phi_4 RIBOVESPA_{t-14} + \phi_6 RIBOVESPA_{t-16} + \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \alpha_3 \varepsilon_{t-3}^2$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\phi_4$	0,112137	0,032354	3,465936	0,0005
$\phi_6$	0,054656	0,031970	1,709625	0,0873
$\alpha_0$	0,000283	2,40E-05	11,80853	0,0000
$\alpha_1$	-0,002337	0,027743	-0,084246	0,9329
$\alpha_2$	0,110740	0,040929	2,705623	0,0068
$\alpha_3$	0,162278	0,044514	3,645547	0,0003

Análise dos parâmetros estimados do modelo AR([14], [16])-ARCH(2)

$$RIBOVESPA_t = \phi_4 RIBOVESPA_{t-14} + \phi_6 RIBOVESPA_{t-16} + \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\phi_4$	0,103368	0,034339	3,010188	0,0026
$\phi_6$	0,070572	0,032494	2,171846	0,0299
$\alpha_0$	0,000345	2,59E-05	13,33414	0,0000
$\alpha_1$	0,005519	0,026988	0,204487	0,8380
$\alpha_2$	0,095884	0,042986	2,230580	0,0257

Análise dos parâmetros estimados modelo AR([14], [16])-GARCH(1,1)

$$RIBOVESPA_t = \phi_4 RIBOVESPA_{t-14} + \phi_6 RIBOVESPA_{t-16} + \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1}$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\phi_4$	0,105130	0,034367	3,059042	0,0022
$\phi_6$	0,069089	0,031985	2,160030	0,0308
$\alpha_0$	2,19E-05	1,93E-05	1,136060	0,2559
$\alpha_1$	0,052290	0,025021	2,089843	0,0366
$\beta$	0,890852	0,068292	13,04467	0,0000

Análise dos parâmetros estimados modelo AR([16])-GARCH(2,1)

$$RIBOVESPA_t = \phi_6 RIBOVESPA_{t-16} + \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \beta h_{t-1}$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\phi_6$	0,050352	0,032379	1,555069	0,1199
$\alpha_0$	6,80E-05	4,07E-05	1,673217	0,0943
$\alpha_1$	-0,008724	0,025343	-0,344241	0,7307
$\alpha_2$	0,118611	0,038920	3,047544	0,0023
$\beta$	0,716345	0,133359	5,371547	0,0000

Estatísticas dos modelos: critérios de informação e log de verosimilhança

Modelo	Log de Verosimilhança	AIC	BIC
AR([14], [16])-ARCH(3)	<b>2370,197</b>	<b>-5,040932</b>	-5,009949
AR([14], [16])-ARCH(2)	2361,162	-5,023800	-4,997981
AR([14], [16])-GARCH(1,1)	2367,837	-5,038031	<b>-5,012211</b>
AR([16])-GARCH(2,1)	2365,060	-5,032112	-5,006292

### Testes de Diagnóstico

Abaixo são apresentados os teste de Ljung-Box (Q (K) e  $Q^2(K)$ ) e o teste ARCH-LM .

Estatística	AR([14], [16])-ARCH(3)	P-value	AR([14], [16])-ARCH(2)	P-value	AR([14], [16])-GARCH(1,1)	P-value	AR([16])-GARCH(2,1)	P-value
Q(3)	1,9390	0,164	1,8014	0,180	1,7535	0,185	2,3928	0,302
Q(6)	6,1376	0,189	5,6888	0,224	5,1514	0,272	6,0950	0,297
Q(24)	17,172	0,754	17,559	0,732	14,982	0,863	27,882	0,220
$Q^2(3)$	0,1516	0,697	<b>15,050</b>	<b>0,000</b>	<b>7,8374</b>	<b>0,005</b>	0,4553	0,796
$Q^2(6)$	4,5084	0,342	<b>20,026</b>	<b>0,000</b>	<b>9,6827</b>	<b>0,046</b>	2,6621	0,752
$Q^2(24)$	32,922	0,063	<b>53,352</b>	<b>0,000</b>	25,942	0,254	25,381	0,331
LM	0,151936	0,984948	0,048921	0,975836	1,374118	0,241106	0,098423	0,951980

Estatísticas de previsão<sup>2</sup> (método dinâmico para 10 passos à frente):

Modelos	
AR([14], [16])-ARCH(3)	
Estatísticas de Previsão	Valor
RMSE	0,013156
MAE	0,010752
TIC	0,964406
AR([14], [16])-ARCH(2)	
Estatísticas de Previsão	Valor
RMSE	0,013175
MAE	0,010715
TIC	0,968152
AR([14], [16])-GARCH(1,1)	
Estatísticas de Previsão	Valor
RMSE	0,013183
MAE	0,010730
TIC	0,967701
AR([16])-GARCH(2,1)	
Estatísticas de Previsão	Valor
RMSE	0,011976
MAE	0,009980
TIC	0,964499

Para os critérios RMSE (Root Mean Squared Error – raiz do erro quadrático médio) e MAE (Mean Absolute Error – erro absoluto médio) temos que quanto menor o erro, melhor a capacidade de previsão do modelo.

<sup>2</sup>Erro Absoluto Médio (Mean Absolute Error – MAE):  $MAE = \sum_{t=T+1}^{T+h} |\hat{y}_t - y_t| / h$

Raiz do Erro Quadrático Médio (Root Mean Square Error – RMSE):  $RMSE = \sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h}$

Coefficiente de Desigualdade de Theil (Theil Inequality Coefficient - TIC):  $TIC = \frac{\sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h}}{\sqrt{\sum_{t=T+1}^{T+h} \hat{y}_t^2 / h + \sum_{t=T+1}^{T+h} y_t^2 / h}}$

As duas primeiras estatísticas de erro de previsão dependem da escala da variável dependente, devem ser utilizadas como medidas relativas para comparar previsões para a mesma série quando utilizam-se diferentes modelos; quanto menor o erro, melhor a capacidade de previsão do modelo de acordo com aquele critério. O coeficiente de desigualdade de Theil sempre estará entre zero e um, sendo que zero indica um ajuste perfeito.

O critério MAPE (Mean Absolute Percentage Error – erro percentual absoluto médio) é invariante com relação a escala, quanto menor este critério melhor o modelo.

O coeficiente de desigualdade de Theil (TIC) sempre está entre os valores zero e um, sendo que zero indica o ajuste perfeito.

O viés de proporção (bias proportion) nos diz o quanto a média da previsão está distante da média real da série.

A proporção de variância (variance proportion) nos diz o quanto a variação da previsão está da variação da série real.

A proporção de covariância (covariance proportion) mede os erros de previsão assistemáticos remanescentes.

Se a previsão é “boa”, o viés e as proporções de variância deverão ser pequenas de tal forma que a maior parte do viés deverá estar concentrado nas proporções de covariância.

Para os modelos AR([14], [16])-ARCH(3) e AR([16])-GARCH(2,1) ajustados para a série DLIBOVESPA temos os seguintes resultados para uma previsão de 10 passos à frente:

AR([14], [16])-ARCH(3)		AR([16])-GARCH(2,1)	
Estatísticas de Previsão	Valor	Estatísticas de Previsão	Valor
MAPE	103,4526	MAPE	92,61260
Proporção do viés	0,013848	Proporção do viés	0,014015
Proporção de variância	0,517383	Proporção de variância	0,796099
Proporção de covariância	0,468769	Proporção de covariância	0,189887

### 4.1.3 Utilização de redes neurais artificiais para previsão da série do IBOVESPA

A seguir são apresentados os resultados obtidos utilizando-se redes neurais artificiais para o tratamento da série temporal do IBOVESPA.

#### 4.1.3.1 Considerações sobre o tratamento dos dados

Os dados da série de fechamento do IBOVESPA são divididos em três partes, a primeira parte contém 150 valores que são utilizados para treinar os pesos sinápticos da rede, o segundo subconjunto, de 40 valores da série é utilizado para realizar a validação e o terceiro subconjunto de 10 valores é utilizado para a tarefa de previsão.

Para que os dados sejam inseridos na rede foi utilizado um processo de normalização. É aconselhável utilizar a normalização dos dados quando usamos o algoritmo ARTR (Aprendizagem Recorrente em Tempo Real), de acordo com (Azoff, 1994, pág. 25) existem quatro abordagens para normalização. Neste caso vamos utilizar o processo que considera um vetor de entrada de dimensão  $n$ , ou seja, os elementos no vetor variam de  $i = 1, \dots, n$ . Os elementos  $t_i^j$ , para a sinapse  $i$  e a amostra total de tamanho  $N$ , do vetor indo de  $j = 1, \dots, N$  são normalizados ao longo da extensão  $[0,1]$ . A Figura-4.2 mostra a série de dados que são normalizados e utilizados nesta análise.

A normalização tem a seguinte forma:

$$NDLIBOVESPA_t = \frac{\Delta(\log(IBOVESPA_t)) - \min(\Delta(\log(IBOVESPA_t)))}{\max(\Delta(\log(IBOVESPA_t))) - \min(\Delta(\log(IBOVESPA_t)))}$$

$$NDLIBOVESPA_t = \frac{DLIBOVESPA_t - \min(DLIBOVESPA_t)}{\max(DLIBOVESPA_t) - \min(DLIBOVESPA_t)}$$

#### 4.1.3.2 Especificação da rede neural recorrente

As redes que foram utilizadas apresentam as seguintes características: 2, 10 e 20 neurônios recorrentes na camada escondida, vetor de 150 valores no subconjunto de treinamento, matriz inicial de pesos igual a zero, função de ativação do tipo sigmóide, taxa de aprendizagem de 0,005 e número de ciclos de treinamento de 1000, 5000, 10000, 15000 e 20000. O RMSE foi armazenado e se não houvesse variação no seu valor (grau de variação de 0,00001) o treinamento era encerrado.

Foram obtidos os seguintes resultados para algumas redes, para previsão de 10 passos à frente:

Topologia da RNA	RMSE	MAE	TIC
30-2R-1	0,18759	0,1523	0,1603
30-3R-1	0,18681	0,15148	0,1597
30-10R-1	0,18453	0,14911	0,1578
30-20R-1	0,18369	0,14834	0,1570
100-2R-1	0,18635	0,15155	0,1567
10-2R-1	0,14037	0,12178	0,12533

#### 4.1.4 Análise gráfica e estatísticas de previsão

A partir dos resultados obtidos do RMSE, MAE e coeficiente de desigualdade de Theil podemos verificar o comportamento das estatísticas de previsão dos modelos ARIMA-ARCH e das redes neurais recorrentes.

Estatísticas de previsão e ajuste para rede neural artificial – ARTR 10-2R-1			
Passos a Frente	rmse	mae	TIC
1	0,008511	0,00851	0,00919
2	0,15507	0,13625	0,13767
3	0,13548	0,10678	0,11959
4	0,12779	0,10636	0,11714
5	0,12657	0,10342	0,12061
6	0,15131	0,13464	0,13294
7	0,14103	0,12013	0,12569
8	0,13886	0,12137	0,12039
9	0,14787	0,1314	0,1339
10	0,14037	0,12178	0,12533

Estatísticas de previsão dinâmica e ajuste para o modelo AR([16])-GARCH(2, 1)				Estatísticas de previsão estática e ajuste para o modelo AR([16])-GARCH(2, 1)			
Passos a Frente	rmse	mae	TIC	Passos a Frente	rmse	mae	TIC
1	0,009212	0,006618	0,937288	1	0,009212	0,006618	0,937288
2	0,012449	0,010139	0,971665	2	0,012449	0,010139	0,971665
3	0,010898	0,008400	0,971373	3	0,010898	0,008400	0,971373
4	0,009991	0,007701	0,96339	4	0,009991	0,007701	0,963390
5	0,010075	0,008165	0,961569	5	0,010075	0,008165	0,961569
6	0,012122	0,009924	0,957842	6	0,012122	0,009924	0,957842
7	0,011351	0,00887	0,954657	7	0,011351	0,008870	0,954657
8	0,011645	0,009414	0,959543	8	0,011645	0,009414	0,959543
9	0,012312	0,010192	0,963697	9	0,012312	0,010192	0,963697
10	0,011976	0,009980	0,964499	10	0,011976	0,009980	0,964499

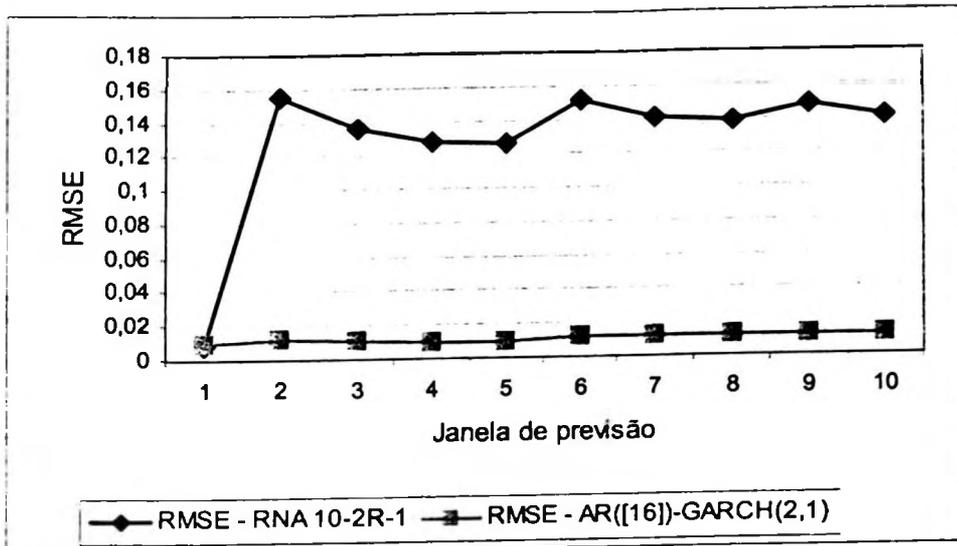


Figura-4.6 RMSE da RNA recorrente e do modelo AR([16])-GARCH(2,1)

Para construir a Figura-4.6 foram realizadas previsões para intervalos consecutivos de 1 a 10. Verificamos que o RMSE do modelo AR([16])-GARCH(2,1) apresenta valores menores para todas as janelas de previsões que foram realizadas.

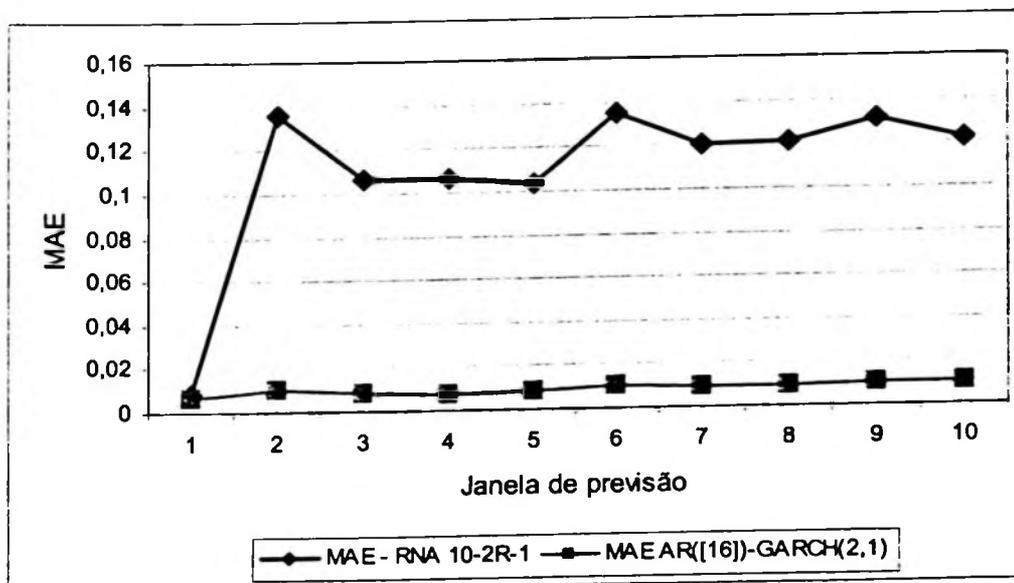


Figura-4.7 MAE da RNA recorrente e do modelo AR([16])-GARCH(2,1)

Plotando os valores do erro absoluto médio para cada janela de previsão, também podemos verificar que o modelo AR([16])-GARCH(2,1) apresenta valores menores do aqueles que forma obtidos com a RNA 10-2R-1.

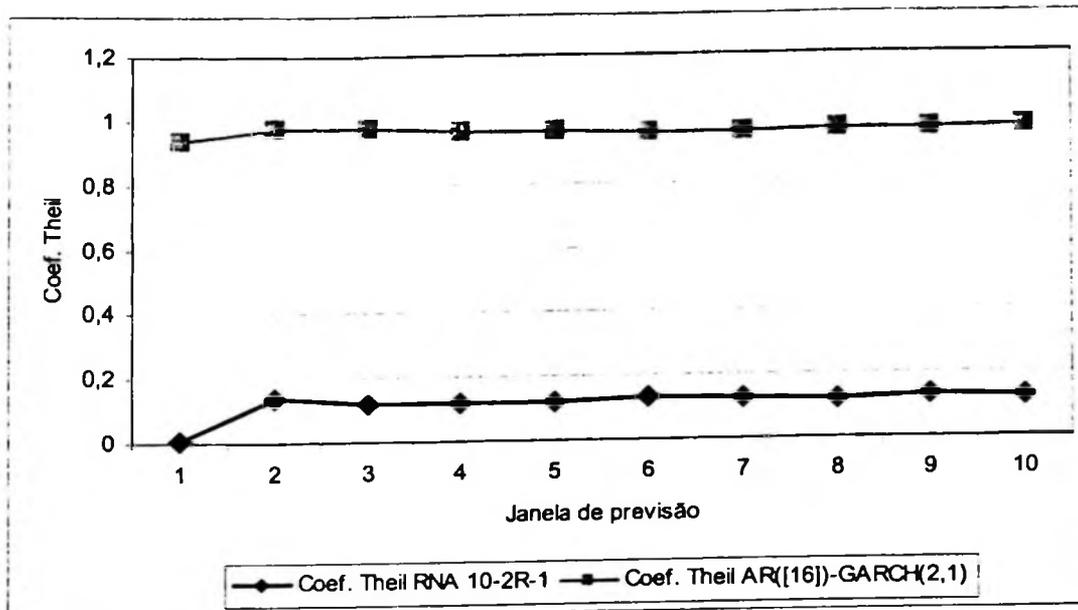


Figura-4.8 Coeficiente de desigualdade de Theil para a RNA recorrente e do modelo AR([16])-GARCH(2,1)

O coeficiente de desigualdade de Theil nesse caso foi utilizado para medir o ajuste entre a curva dos valores previstos e os valores desejados. Nesse caso os valores obtidos para o ajuste da rede neural recorrente foram menores do que aqueles obtidos no modelo AR([16])-GARCH(2,1).

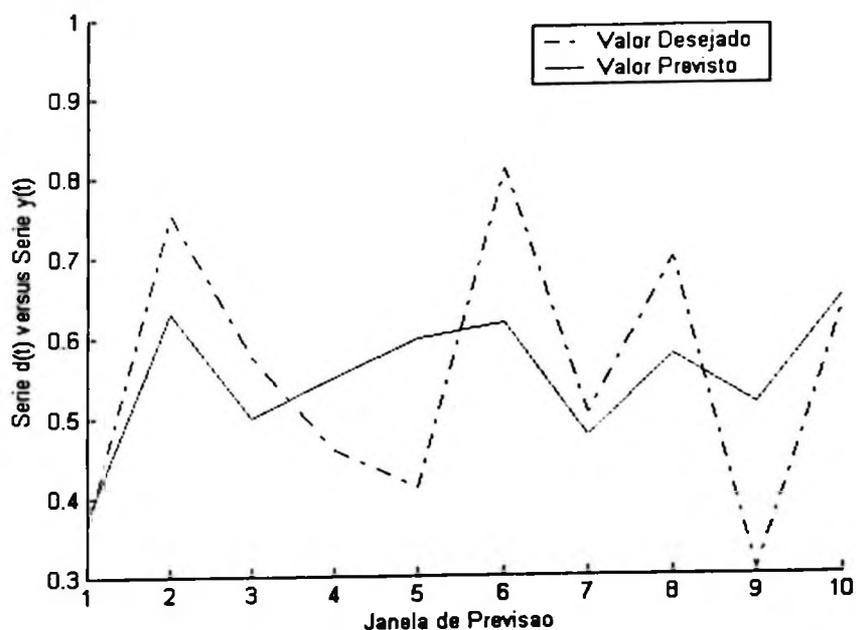


Figura-4.9 Valores para previsão da série NDLIBOVESPA utilizando a RNA 10-2R-1

O valor do MAPE da rede neural para previsão de 10 passos à frente é de 21,56.

A Figura-4.9 apresenta o ajuste obtido entre os valores previstos e os valores desejados para a série DLIBOVESPA.

Na Figura-10 temos o ajuste obtido entre valores previstos e valores desejados quando utilizamos o modelo AR([16])-GARCH(2,1).

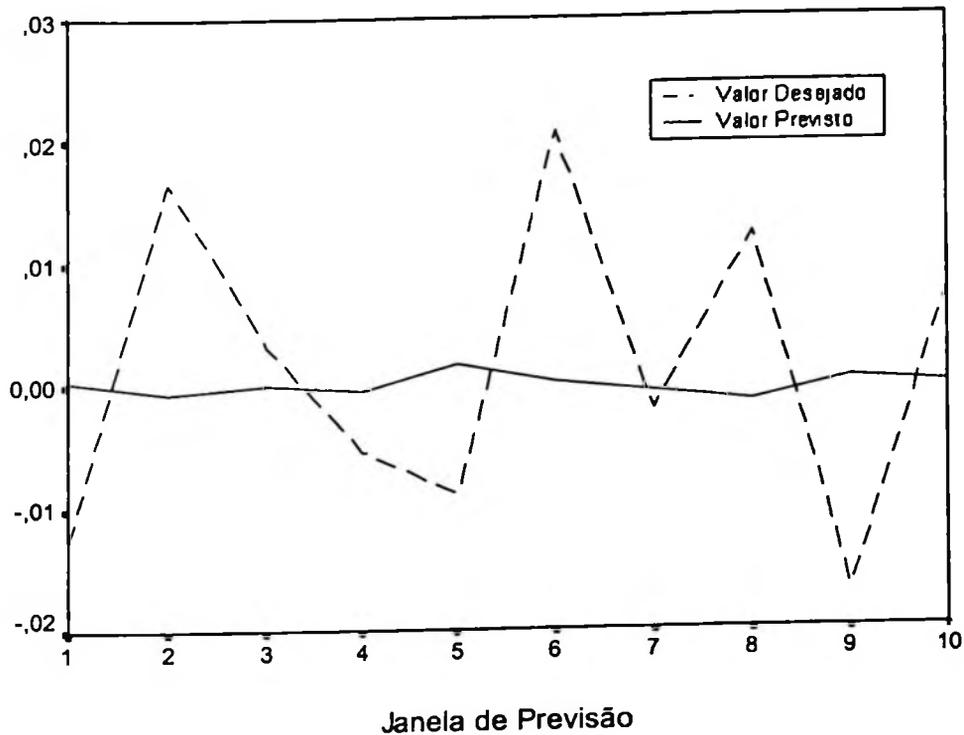


Figura-4.10 Valores para previsão do DLIBOVESPA utilizando o modelo AR([16])-GARCH(2,1)

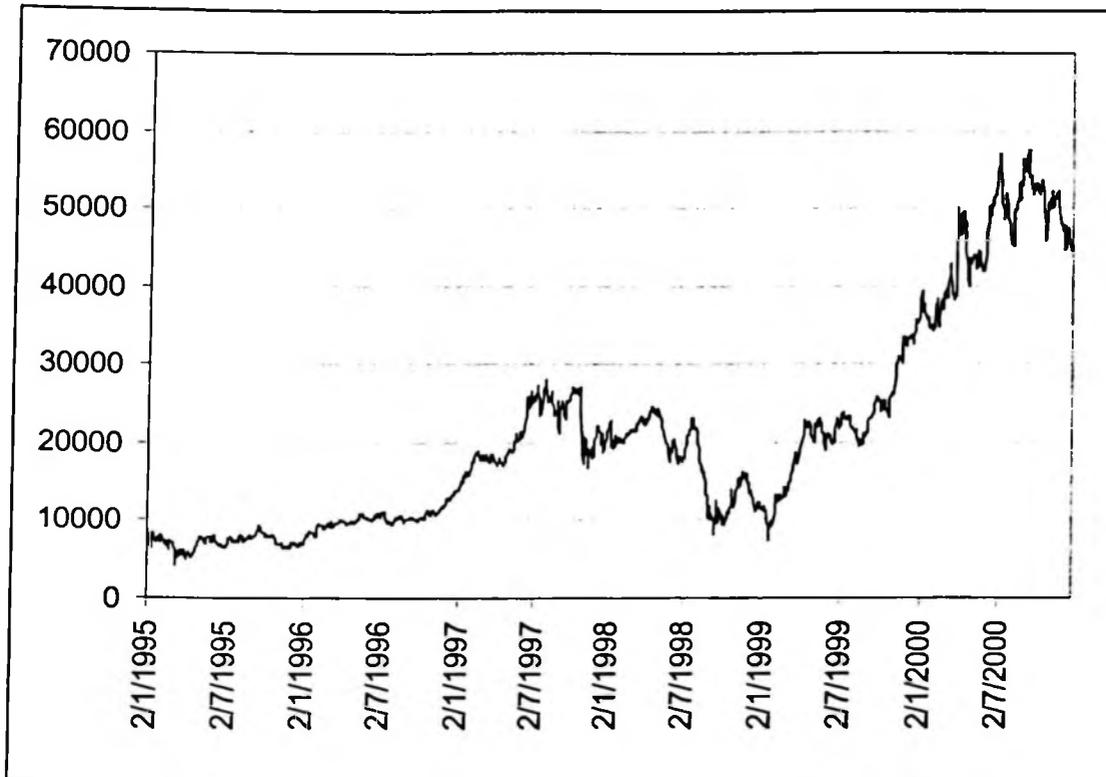
Na Figura-4.9 podemos verificar o ajuste obtido pela rede neural e na Figura-4.10 temos o ajuste do modelo ARIMA-GARCH.

## 4.2 Análise da série de retornos diários da Petrobrás

A seguir apresentaremos uma análise da série de retornos diários da Petrobrás.

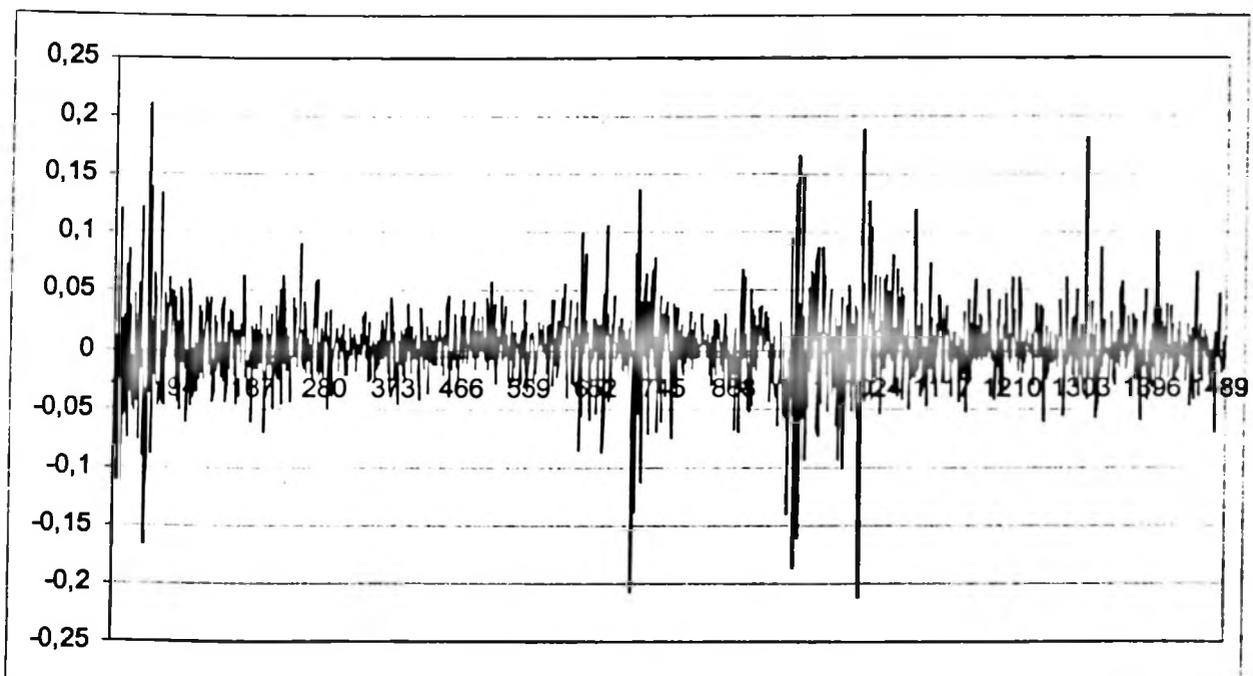
A série de valores de pontos de fechamento da Petrobrás mostrada na Figura-4.11 refere-se ao período de 2 de janeiro de 1995 até 27 de dezembro de 2000, no total são 1500 valores.

85283



**Figura-4.11** Série de pontos de fechamento das ações da Petrobrás

Aplicando a primeira diferença ao logaritmo da série da Figura-4.11 obtemos a série de retornos para as ações da Petrobrás apresentadas na Figura-4.12.



**Figura-4.12** Série de retornos ações da Petrobrás

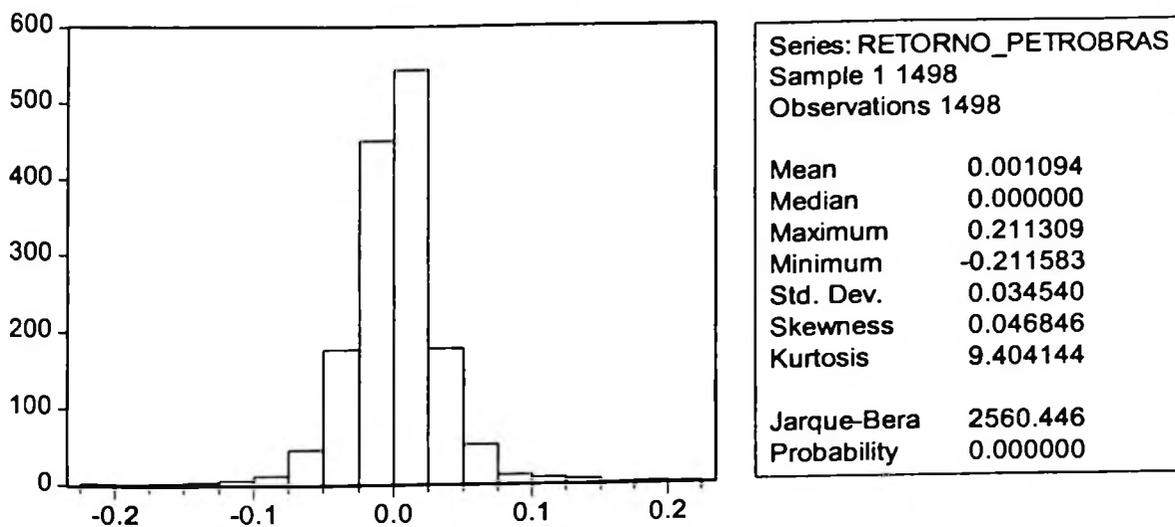


Figura-4.13 Estatísticas dos retornos diários da Petrobrás

Para a série de retornos da Petrobrás rejeitamos a hipótese de normalidade pelo teste de Jarque-Bera e observamos excesso de curtose o que indica a necessidade de especificação de um modelo ARCH para a série.

Realizando os testes de raiz unitária temos que a série de retornos da Petrobrás é estacionária, conforme pode ser visto nas duas tabelas abaixo dos testes ADF e Dickey-Pantula.

Estadística do teste ADF	-19,08329	1% Critical Value*	-3,4376
		5% Critical Value	-2,8639
		10% Critical Value	-2,5681

Verificamos que o teste ADF rejeita a hipótese de uma raiz unitária.

Hipóteses testadas	Estadística	valores críticos a 5%
$H_0 : 2RU$	-28,97595	-3,37
$H_a : 1RU$	-43,58847	-3,77
$H_0 : 1RU$	-41,01864	-3,42
$H_a : 0RU$	-15,58016	-3,80
$H_a : 0RU$	-23,23530	-4,16

Rejeitamos as hipóteses de presença de duas e uma raiz unitária pelo teste de Dickey-Pantula.

Serão analisadas as seguintes estimativas para a série de retornos da Petrobrás: AR([1])-GARCH(1,1), AR([3], [9])-GARCH(1,1) e AR([9])-GARCH(1,1).

Temos então os seguintes resultados:

Estimativa do parâmetros do modelo AR([1])-GARCH(1,1)

$$RPETROBRAS_t = \phi_1 RPETROBRAS_{t-1} + \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1}$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\phi_1$	0,132377	0,030020	4,409637	0,0000
$\alpha_0$	0,0000355	0,0000113	3,151247	0,0016
$\alpha_1$	0,143232	0,045241	3,166004	0,0015
$\beta$	0,829494	0,044863	18,48959	0,0000

Estimativa do parâmetros do modelo AR([3], [9])-GARCH(1,1)

$$RPETROBRAS_t = \phi_3 RPETROBRAS_{t-3} + \phi_9 RPETROBRAS_{t-9} + \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1}$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\phi_3$	-0,009704	0,033803	-0,287086	0,7740
$\phi_9$	0,030401	0,025315	1,200906	0,2298
$\alpha_0$	0,0000350	0,0000113	3,102241	0,0019
$\alpha_1$	0,141412	0,045490	3,108616	0,0019
$\beta$	0,831714	0,045542	18,26251	0,0000

Estimativa do parâmetros do modelo AR([9])-GARCH(1,1)

$$RPETROBRAS_t = \phi_9 RPETROBRAS_{t-9} + \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1}$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\phi_9$	0,030836	0,025359	1,215988	0,2240
$\alpha_0$	0,0000349	0,0000113	3,103468	0,0019
$\alpha_1$	0,141418	0,045334	3,119483	0,0018
$\beta$	0,831791	0,045446	18,30293	0,0000

Dos três modelos estimados o único modelo que apresentou todos os coeficientes significantes é o modelo AR([1])-GARCH(1,1).

Notamos que a soma dos coeficientes ARCH e GARCH estão muito próximos de um, indicando que os choques de volatilidade são persistentes. Este resultado é frequentemente observado em séries financeiras de alta frequência (Engle e Patton, 2001).

Estatísticas dos modelos: critérios de informação e log de verosimilhança

Modelo	Log de Verosimilhança	AIC	BIC
AR([1])-GARCH(1,1)	3154,824	-4,237826	-4,223557
AR([3], [9])-GARCH(1,1)	3137,989	-4,236632	-4,218717
AR([9])-GARCH(1,1)	3137,927	-4,237900	-4,223568

### Testes de Diagnóstico

Abaixo são apresentados os teste de Ljung-Box (Q (K) e  $Q^2(K)$ ) e o teste ARCH-LM .

Estatística	AR([1])- GARCH(1,1)	P-value	AR([3], [9])- GARCH(1,1)	P-value	AR([9])- GARCH(1,1)	P-value
Q(3)	1,8654	0,393	28,227	0,000	28,302	0,000
Q(6)	4,4229	0,490	30,776	0,000	30,908	0,000
Q(24)	24,248	0,390	51,766	0,000	51,908	0,001
$Q^2(3)$	2,8467	0,241	2,3364	0,126	2,3683	0,306
$Q^2(6)$	5,5183	0,356	5,1460	0,273	5,1558	0,397
$Q^2(24)$	12,587	0,960	11,452	0,968	11,560	0,977
LM	6,225720	0,795956	6,206194	0,797652	6,233802	0,795253

Para os modelos AR([9])-GARCH(1,1) e AR([1])-GARCH(1,1) ajustados para a série DLPETRO temos os seguintes resultados para uma previsão de 10 passos à frente:

AR([9])-GARCH(1,1)		AR([1])-GARCH(1,1)	
Estatísticas de Previsão	Valor	Estatísticas de Previsão	Valor
MAPE	98,35472	MAPE	98,02898
Proporção do viés	0,001560	Proporção do viés	0,001052
Proporção de variância	0,910995	Proporção de variância	0,974487
Proporção de covariância	0,087445	Proporção de covariância	0,024461

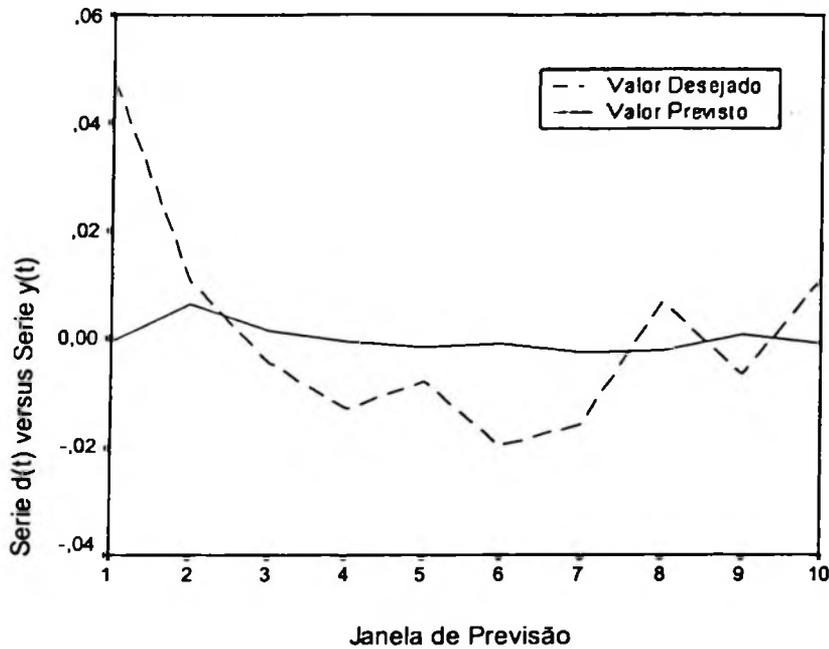


Figura-4.14 Valores para previsão da série DLPETRO utilizando modelo AR([1])-GARCH(1,1)

Estatísticas de previsão para o modelo AR([1])-GARCH(1,1)

Estatísticas de previsão dinâmica e ajuste para o modelo AR([1])-GARCH(1, 1)				Estatísticas de previsão estática e ajuste para o modelo AR([1])-GARCH(1, 1)			
Passos a Frente	rmse	mae	TIC	Passos a Frente	rmse	mae	TIC
1	0,034017	0,025322	0,985842	1	0,034261	0,025495	0,990578
2	0,035030	0,029512	0,999967	2	0,028081	0,018409	0,874595
3	0,028715	0,021149	0,999916	3	0,024493	0,015265	0,875343
4	0,025702	0,019107	0,999632	4	0,022598	0,014691	0,876262
5	0,023247	0,016834	0,999543	5	0,020775	0,013246	0,870330
6	0,022715	0,017335	0,999069	6	0,020506	0,014042	0,871919
7	0,021857	0,017110	0,998844	7	0,019736	0,013928	0,859999
8	0,020852	0,015808	0,998808	8	0,018836	0,013356	0,861570
9	0,019533	0,014794	0,998773	9	0,018029	0,012778	0,863537
10	0,018861	0,014467	0,998684	10	0,017567	0,012708	0,867996

Utilizando uma rede neural 10-3R-1 obtemos o seguinte resultado:

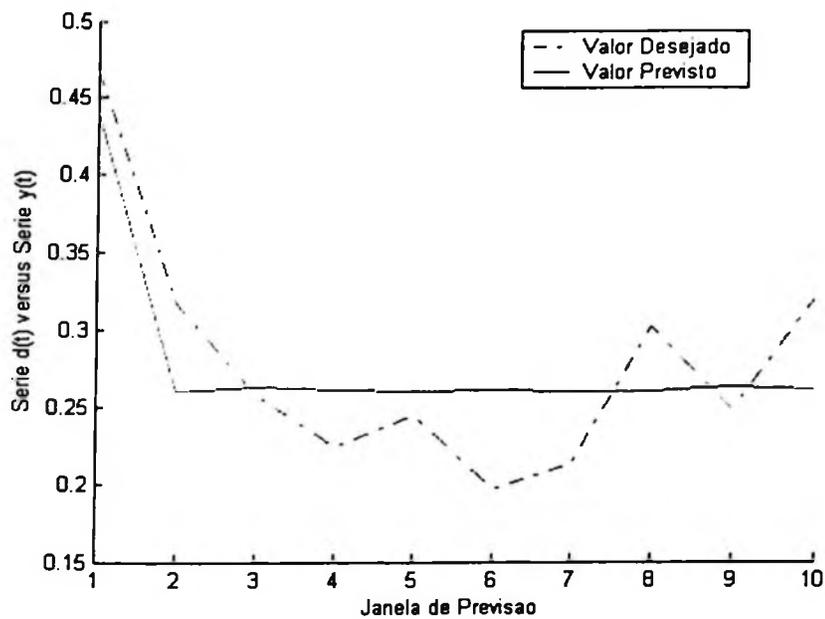


Figura-4.15 Valores para previsão da série NDLPETRO utilizando a RNA 10-3R-1

Estatísticas de previsão:

Topologia da RNA	RMSE	MAE	TIC
10-3R-1	0,041825	0,036845	0,072983

O valor do MAPE da rede neural para previsão de 10 passos à frente é de 13,40.

O histograma dos resíduos gerados da previsão da rede 10-3R-1 é mostrado na figura abaixo:

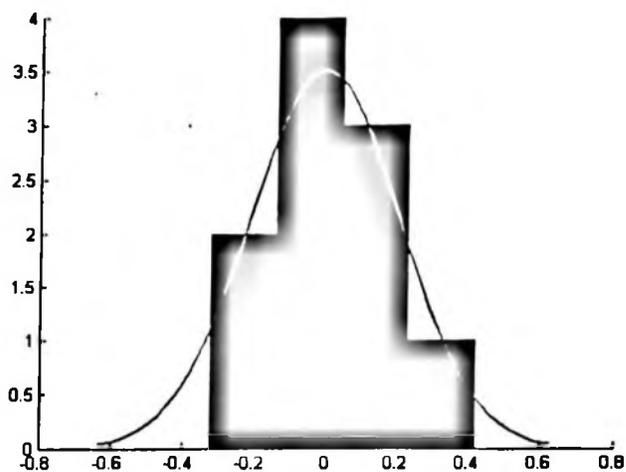


Figura-4.16 Histograma dos resíduos da previsão da RNA 10-2R-1

A linha vermelha na Figura-4.16 indica uma curva normal.

### 4.3 Análise da série de retornos diários da Nasdaq

A série de valores de fechamento do índice Nasdaq é mostrada na Figura-4.17. A série de retornos diários para 887 valores da Nasdaq é mostrada na Figura-4.18.

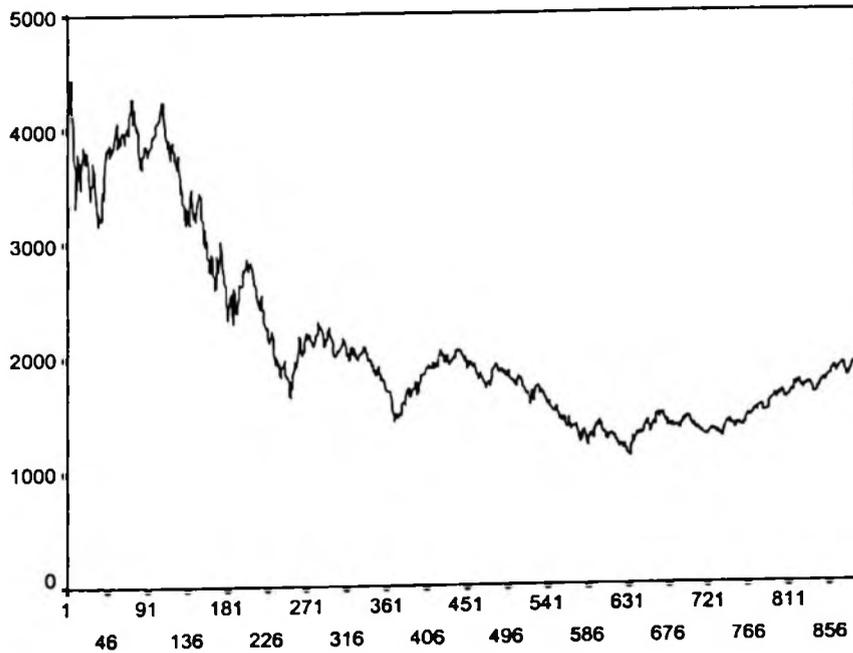


Figura-4.17 Série de valores da Nasdaq

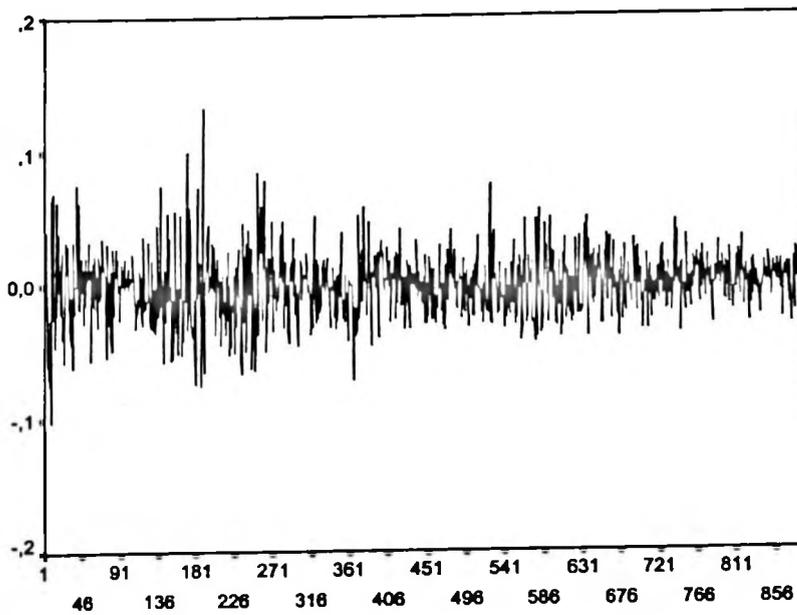
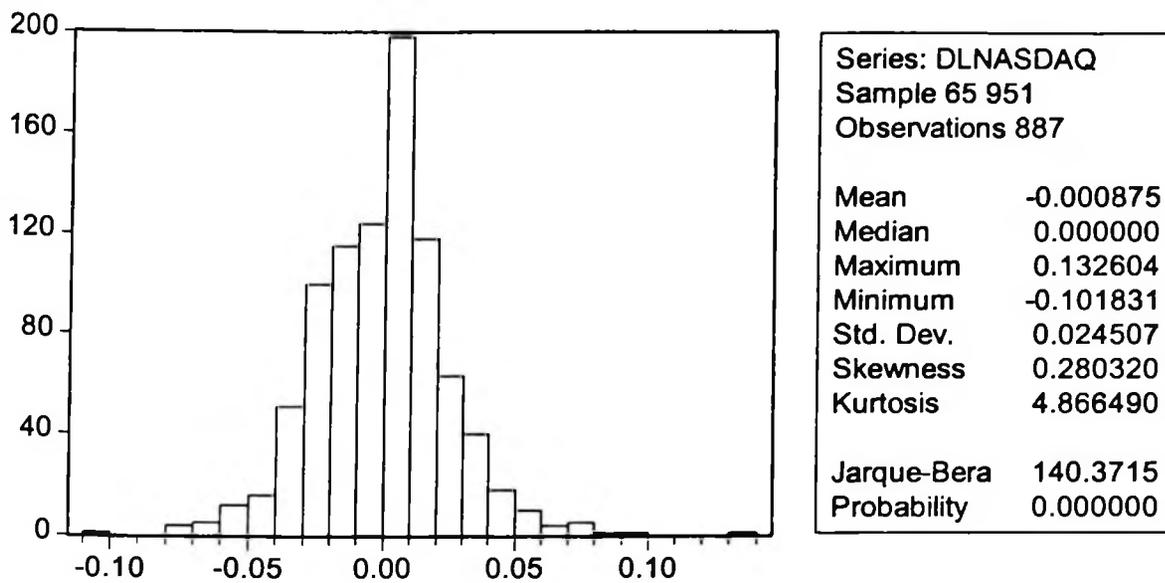
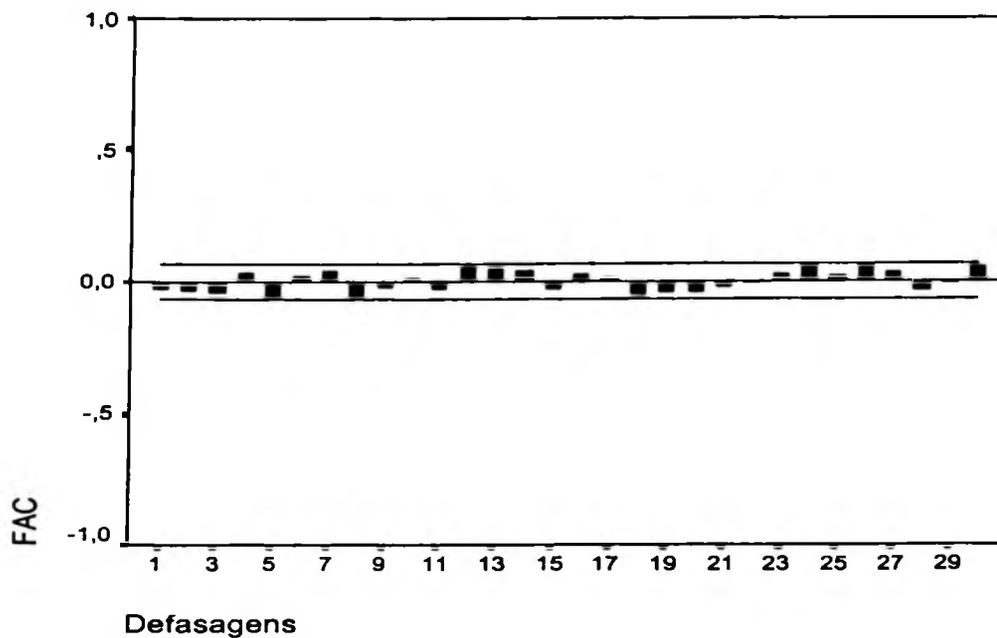


Figura-4.18 Série de retornos da Nasdaq



**Figura-4.19** Estatísticas dos retornos diários da Nasdaq

Rejeitamos a hipótese de normalidade pelo teste de Jarque-Bera e observamos excesso de curtose o que indica a necessidade de especificação de um modelo ARCH para a série.



**Figura-4.20** Função de auto-correlação para a série de retornos diários da Nasdaq

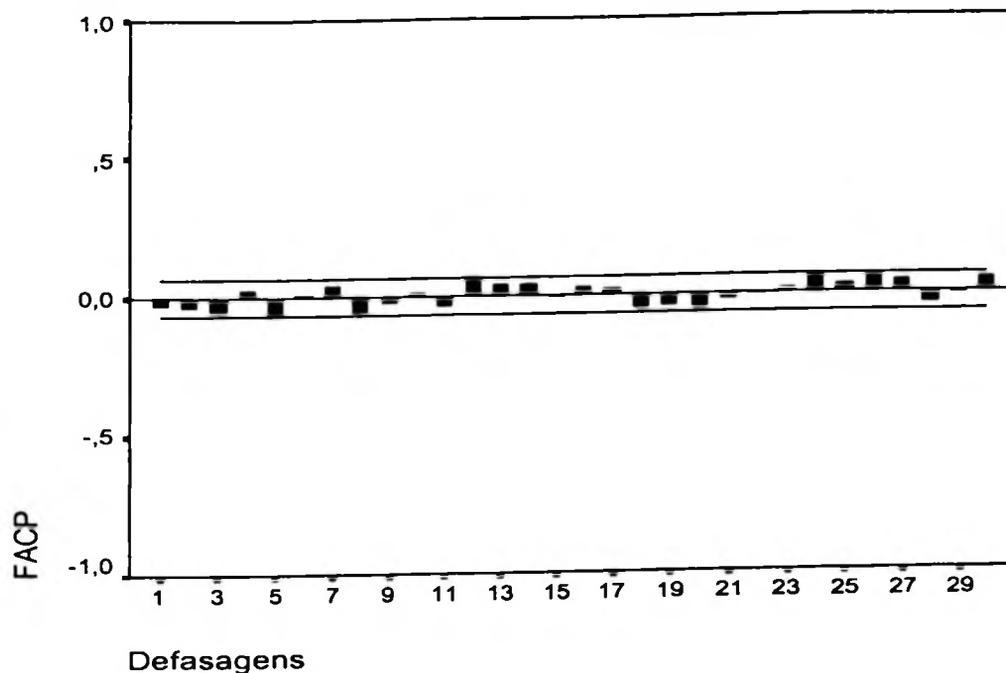


Figura-4.21 Função de auto-correlação parcial para a série de retornos diários da Nasdaq

As funções FAC e FACP para a série de retornos diários da Nasdaq mostram que esta série tem comportamento que não pode ser capturado por uma modelagem ARIMA.

Para a série de retornos diários da Nasdaq serão estimados os modelos: GARCH(1,1) e ARCH(1).

Estimativa dos parâmetros do modelo GARCH(1,1)

$$RNASDAQ_t = \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1}$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\alpha_0$	0,00000722	0,00000362	1,996764	0,0459
$\alpha_1$	0,071304	0,018290	3,898610	0,0001
$\beta$	0,915770	0,018565	49,32847	0,0000

Neste caso também observamos que a soma dos coeficientes ARCH e GARCH estão muito próximos de um, indicando que os choques de volatilidade são persistentes. Vale ressaltar que este resultado é frequentemente observado em séries financeiras de alta frequência.

Estimativa dos parâmetros do modelo ARCH(1):

$$RNASDAQ_t = \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\alpha_0$	0,000525	0,0000423	12,38797	0,0000
$\alpha_1$	0,130409	0,067316	1,937259	0,0527

Estatísticas dos modelos: critérios de informação e log de verosimilhança

Modelo	Log de Verosimilhança	AIC	BIC
GARCH(1,1)	2078,788	-4,733838	-4,717499
ARCH(1)	2014,694	-4,589952	-4,579059

### Testes de Diagnóstico

Abaixo são apresentados os teste de Ljung-Box (Q (K) e  $Q^2(K)$ ) e o teste ARCH-LM .

Estatística	GARCH(1,1)	P-value	ARCH(1)	P-value
Q(3)	1,0847	0,781	1,6314	0,652
Q(6)	2,4919	0,869	5,3199	0,503
Q(24)	17,991	0,803	29,186	0,213
$Q^2(3)$	5,2566	0,154	48,964	0,000
$Q^2(6)$	7,6004	0,269	99,504	0,000
$Q^2(24)$	22,732	0,536	285,31	0,000
LM	12,33484	0,263272	31,65727	0,000000

Nesse caso verificamos que o modelo GARCH(1,1) é melhor por ter capturado a dependência temporal na volatilidade da série de retornos da Nasdaq.

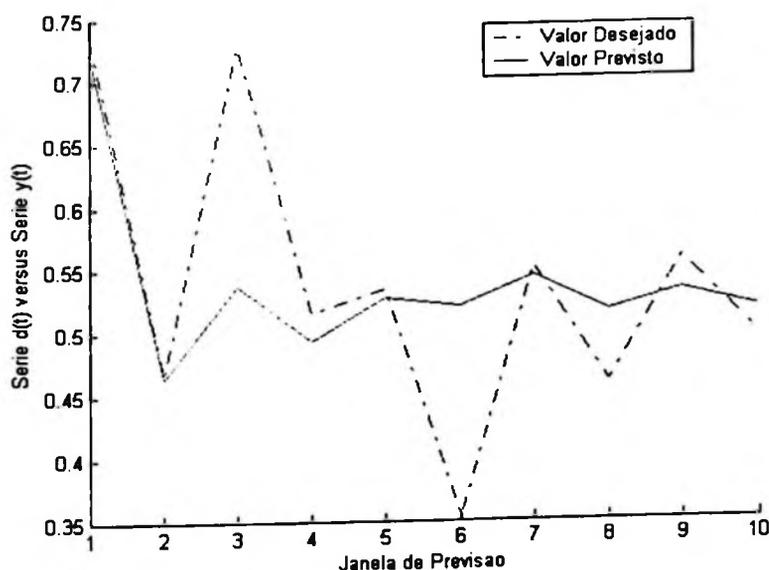
Resultados para uma previsão de 10 passos à frente:

GARCH(1,1)	
Estatísticas de Previsão	Valor
MAPE	100,0000
Proporção do viés	0,176634
Proporção de variância	0,823366
Proporção de covariância	0,000000

### Estatísticas de previsão para o modelo GARCH(1,1)

Estatísticas de previsão e ajuste para o modelo GARCH(1, 1)			
Passos a Frente	rmse	mae	TIC
1	0,023036	0,022956	1,00000
2	0,018851	0,016031	1,00000
3	0,020325	0,018077	1,00000
4	0,018400	0,015733	1,00000
5	0,017103	0,014427	1,00000
6	0,016077	0,013418	1,00000
7	0,015406	0,012923	1,00000
8	0,014535	0,011661	1,00000
9	0,014138	0,011482	1,00000
10	0,013552	0,010861	1,00000

Utilizando uma rede neural 10-4R-1 obtemos os seguintes resultados mostrados abaixo.



**Figura-4.22** Ajuste da curva de previsão de uma rede neural 10-4R-1 para a série NDNLASDAQ

O valor do MAPE da rede neural para previsão de 10 passos à frente é de 10,33.

Estatísticas de previsão:

Topologia da RNA	RMSE	MAE	TIC
30-2R-1	0,16053	0,12901	0,1684
10-3R-1	0,082286	0,051076	0,075398
10-4R-1	0,08266	0,050974	0,075726

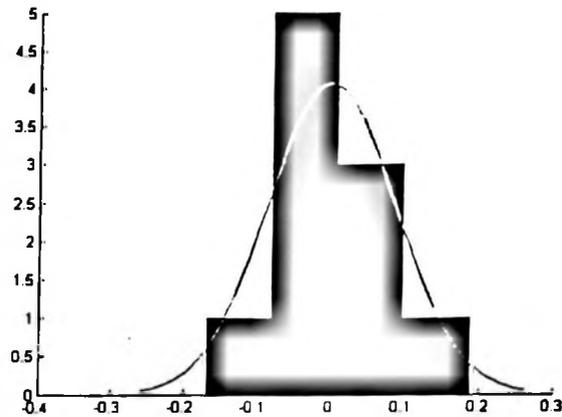


Figura-4.23 Histograma dos resíduos da previsão da RNA 10-4R-1 para a série NDLNASDAQ

#### 4.4 Análise da série de preços da saca de 60kg da soja

A série de valores diários da saca de 60Kg da soja é mostrada na Figura-4.24.

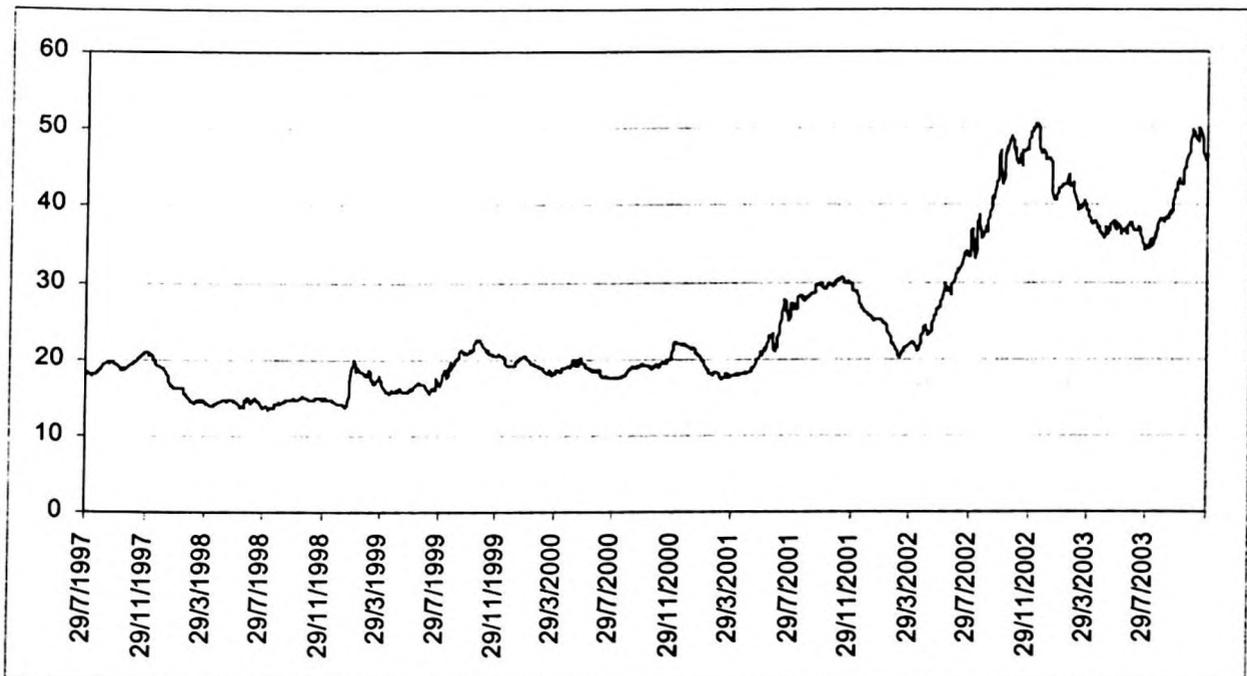


Figura-4.24 Série de valores de preços diários da saca de 60Kg da soja. Fonte: Série histórica Indicador CEPEA/ESALQ (R\$ /sc 60 kg)

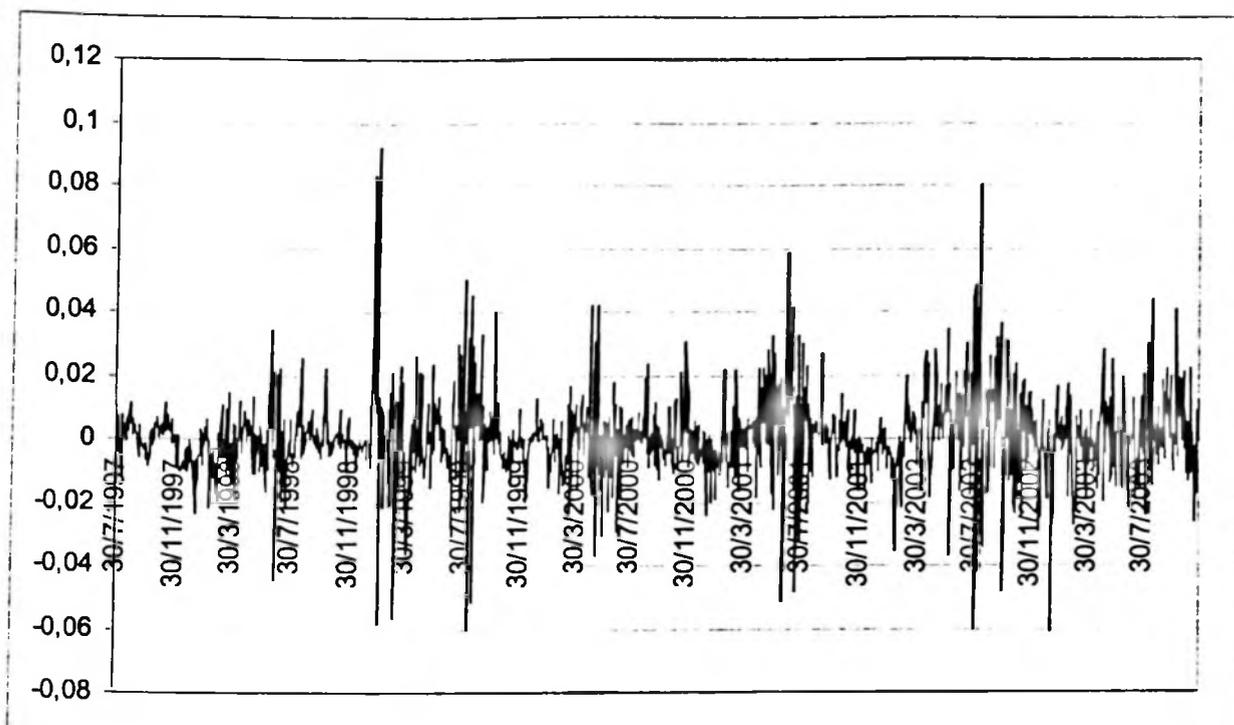


Figura-4.25 Série de retornos diários da saca de 60Kg da soja

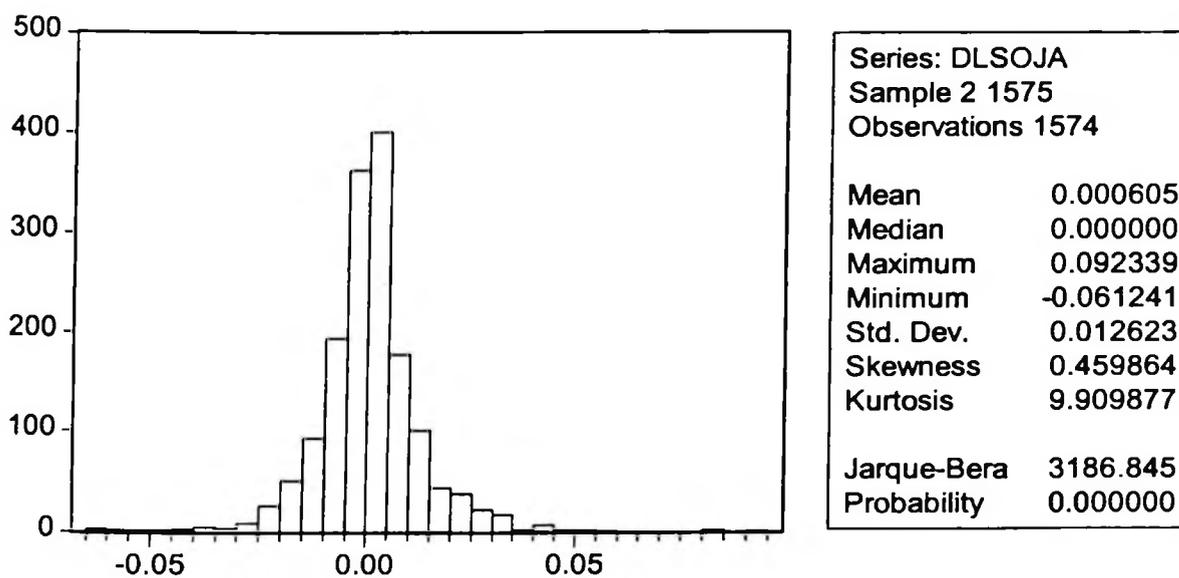


Figura-4.26 Estatísticas dos retornos diários da saca de 60Kg da soja

Rejeitamos a hipótese de normalidade pelo teste de Jarque-Bera e observamos excesso de curtose o que indica a necessidade de especificação de um modelo ARCH para a série.

### Testes de raiz unitária para a série DLISOJA

Abaixo temos o resultado do teste ADF.

ADF Test Statistic	-15,87229	1% Critical Value*	-3,4374
		5% Critical Value	-2,8638
		10% Critical Value	-2,5680

Rejeitamos a hipótese de presença de uma raiz unitária na série DLISOJA.

O resultado do teste de Dikey-Pantula para uma e duas raízes unitárias é mostrado a seguir.

Hipóteses testadas	Estatística	valores críticos a 5%
$H_0 : 2RU$	-30.36328	-3,37
$H_a : 1RU$	-43.07428	-3,77
$H_0 : 1RU$	-38.76627	-3,42
$H_a : 0RU$	-21.56586	-3,80
	-19.13812	-4,16

Rejeitamos as hipóteses de presença de duas e uma raiz unitária pelo teste de Dickey-Pantula.

A Figura-4.25 representa a série de retornos diários da saca de 60Kg da soja, podemos verificar que a função de autocorrelação (Autocorrelation) apresenta comportamento declinante e a função de autocorrelação parcial (Partial Correlation) é truncada na primeira defasagem, isso é indicação de um modelo auto-regressivo de primeira ordem AR([1]).

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.284	0.284	127.50	0.000
		2	0.126	0.049	152.61	0.000
		3	0.066	0.020	159.57	0.000
		4	0.047	0.019	163.03	0.000
		5	0.004	-0.020	163.06	0.000
		6	0.001	-0.001	163.06	0.000
		7	-0.016	-0.017	163.45	0.000
		8	0.021	0.032	164.14	0.000
		9	0.023	0.014	165.01	0.000
		10	0.048	0.038	168.63	0.000
		11	0.057	0.035	173.84	0.000
		12	0.036	0.003	175.92	0.000
		13	0.014	-0.007	176.24	0.000
		14	-0.027	-0.041	177.41	0.000

Figura-4.27 FAC e FACP da série de retornos diários (DLISOJA) da saca de 60Kg da soja

Especificando um modelo AR([1]) para representar a série de retornos diários do preço da saca de 60Kg da soja, verificamos que neste modelo existe correlação dos resíduos quadrados, conforme pode ser verificado pelo correlograma da Figura-4.28.

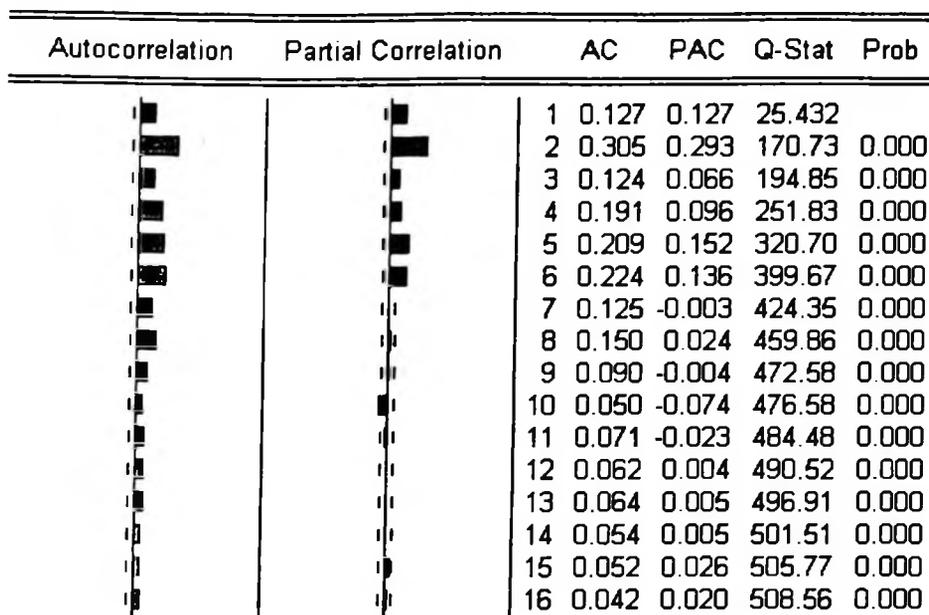


Figura-4.28 Correlograma dos resíduos quadrados da série DLSOJA

Resultado do teste ARCH-LM para a série DLSOJA AR([1]), para uma defasagem:

ARCH Test:

Obs*R-squared	25,36882	Probability	0,000000
---------------	----------	-------------	----------

Portanto, rejeitamos a hipótese nula “não existe efeitos ARCH” no modelo

Para esse caso vamos estimar os modelos AR([1])-ARCH(1), AR([1])-ARCH(5) e AR([1])-GARCH(1,1).

Estimativa dos parâmetros do modelo AR([1])-ARCH(1)

$$RSOJA_t = \phi_1 RSOJA_{t-1} + \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\phi_1$	0,393339	0,076654	5,131363	0,0000
$\alpha_0$	0,0000967	0,00000915	10,56474	0,0000
$\alpha_1$	0,411514	0,089432	4,601411	0,0000

Estimativa dos parâmetros do modelo AR([1])-ARCH(5)

$$RSOJA_t = \phi_1 RSOJA_{t-1} + \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \alpha_3 \varepsilon_{t-3}^2 + \alpha_4 \varepsilon_{t-4}^2 + \alpha_5 \varepsilon_{t-5}^2$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\phi_1$	0,322080	0,030163	10,67814	0,0000
$\alpha_0$	3,18E-05	5,37E-06	5,915502	0,0000
$\alpha_1$	0,195430	0,045639	4,282137	0,0000
$\alpha_2$	0,284174	0,062529	4,544689	0,0000
$\alpha_3$	0,067347	0,040369	1,668284	0,0953
$\alpha_4$	0,135541	0,043637	3,106127	0,0019
$\alpha_5$	0,246190	0,059919	4,108738	0,0000

Estimativa dos parâmetros do modelo AR([1])-GARCH(1,1)

$$RSOJA_t = \phi_1 RSOJA_{t-1} + \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1}$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\phi_1$	0,344720	0,027821	12,39087	0,0000
$\alpha_0$	0,00000333	0,00000106	3,153767	0,0016
$\alpha_1$	0,160294	0,028825	5,561005	0,0000
$\beta$	0,831960	0,028134	29,57108	0,0000

Estadísticas dos modelos: critérios de informação e log de verosimilhança

Modelo	Log de Verosimilhança	AIC	BIC
AR([1])-GARCH(1,1)	4955,847	-6,336337	-6,322634
AR([1])-ARCH(1)	4768,803	-6,098276	-6,087999
AR([1])-ARCH(5)	4943,905	-6,317216	-6,293237

## Testes de Diagnóstico

Abaixo são apresentados os teste de Ljung-Box (Q (K) e  $Q^2(K)$ ) e o teste ARCH-LM .

Estadística	AR([1])- GARCH(1,1)	P-value	AR([1])- ARCH(1)	P-value	AR([1])- ARCH(5)	P-value
Q(3)	14,659	0,001	9,5132	0,009	8,6335	0,013
Q(6)	25,956	0,000	15,583	0,008	20,081	0,001
Q(24)	47,975	0,002	37,563	0,028	42,178	0,009
$Q^2(3)$	0,2489	0,883	62,670	0,000	1,3597	0,507
$Q^2(6)$	1,4315	0,921	122,79	0,000	7,3704	0,195
$Q^2(24)$	10,194	0,990	207,13	0,000	20,927	0,586
LM	4,533526	0,920087	41,08599	0,000000	13,42752	0,200741

Estatísticas de previsão dinâmica para os modelos AR([1])-GARCH(1,1) e AR([1])-ARCH(5) são mostrados na tabela a seguir.

Estatísticas de previsão dinâmica e ajuste para o modelo AR([1])-GARCH(1,1)				Estatísticas de previsão dinâmica e ajuste para o modelo AR([1])-ARCH(5)			
Passos a Frente	rmse	mae	TIC	Passos a Frente	rmse	mae	TIC
1	0,003967	0,003777	0,982368	1	0,003800	0,003613	0,980396
2	0,009714	0,007806	0,841912	2	0,009705	0,007718	0,851135
3	0,015316	0,012254	0,898241	3	0,015324	0,012195	0,904949
4	0,014337	0,011695	0,901838	4	0,014346	0,011650	0,908379
5	0,013122	0,010132	0,902033	5	0,013130	0,010096	0,908568
6	0,013810	0,011167	0,912462	6	0,013817	0,011136	0,918299
7	0,013279	0,010858	0,914576	7	0,013285	0,010831	0,920271
8	0,012884	0,010666	0,916745	8	0,012890	0,010642	0,922291
9	0,012224	0,009643	0,916749	9	0,012229	0,009621	0,922295
10	0,012189	0,009842	0,920033	10	0,012194	0,009822	0,925354

Estatísticas de previsão estática para os modelos AR([1])-GARCH(1,1) e AR([1])-ARCH(5) são mostrados na tabela abaixo.

Estatísticas de previsão estática e ajuste para o modelo AR([1])-GARCH(1,1)				Estatísticas de previsão estática e ajuste para o modelo AR([1])-ARCH(5)			
Passos a Frente	rmse	mae	TIC	Passos a Frente	rmse	mae	TIC
1	0,003580	0,002924	0,909858	1	0,003446	0,002841	0,908627
2	0,010097	0,007529	0,881060	2	0,010051	0,007465	0,886009
3	0,013349	0,010690	0,720385	3	0,013446	0,010723	0,733218
4	0,011943	0,008692	0,622087	4	0,012039	0,008822	0,636570
5	0,010909	0,007396	0,614477	5	0,010994	0,007472	0,628926
6	0,011886	0,008708	0,656195	6	0,011962	0,008780	0,669767
7	0,011161	0,007963	0,626384	7	0,011243	0,008070	0,640601
8	0,011270	0,008423	0,652288	8	0,011321	0,008499	0,665243
9	0,010750	0,007936	0,651295	9	0,010793	0,007986	0,664040
10	0,010868	0,008304	0,667854	10	0,010906	0,008349	0,680224

Resultados para uma previsão de 10 passos à frente:

AR([1])-GARCH(1,1)		AR([1])-ARCH(5)	
Estatísticas de Previsão	Valor	Estatísticas de Previsão	Valor
MAPE	162,4107	MAPE	157,1177
Proporção do viés	0,060723	Proporção do viés	0,067702
Proporção de variância	0,501960	Proporção de variância	0,526861
Proporção de covariância	0,437317	Proporção de covariância	0,405437

Para ambos os modelos a proporção do viés e a proporção de variância tem valores próximos, da tabela acima verificamos que o viés do modelo AR([1])-GARCH(1,1) está mais concentrado na proporção da covariância.

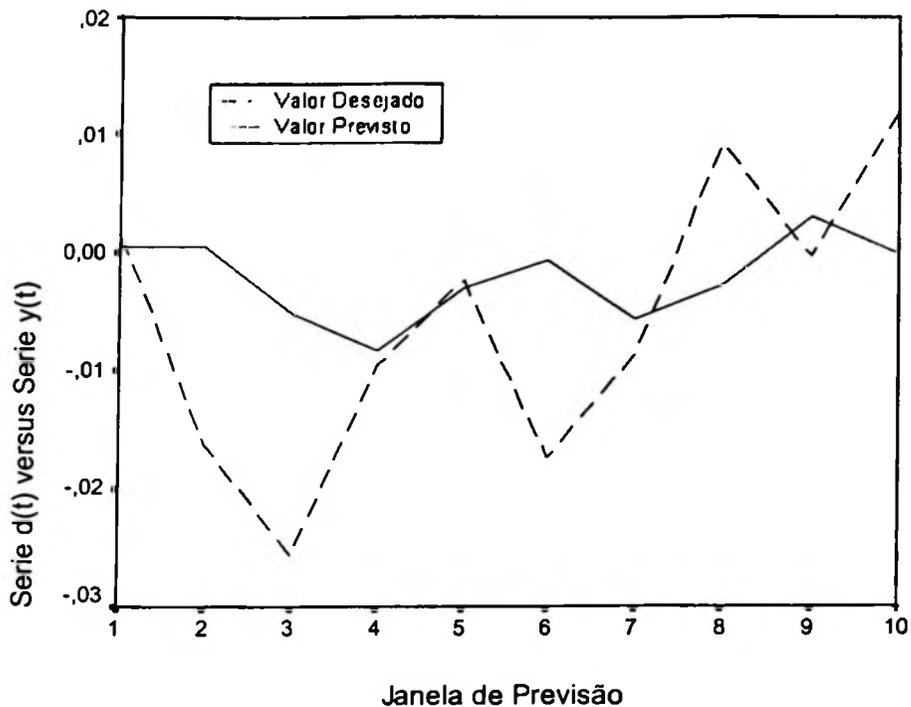


Figura-4.29 Valores para previsão da série DLSOJA utilizando modelo AR([1])-GARCH(1,1)

Utilizando uma rede neural 10-2R-1 obtemos os seguintes resultados mostrados abaixo.

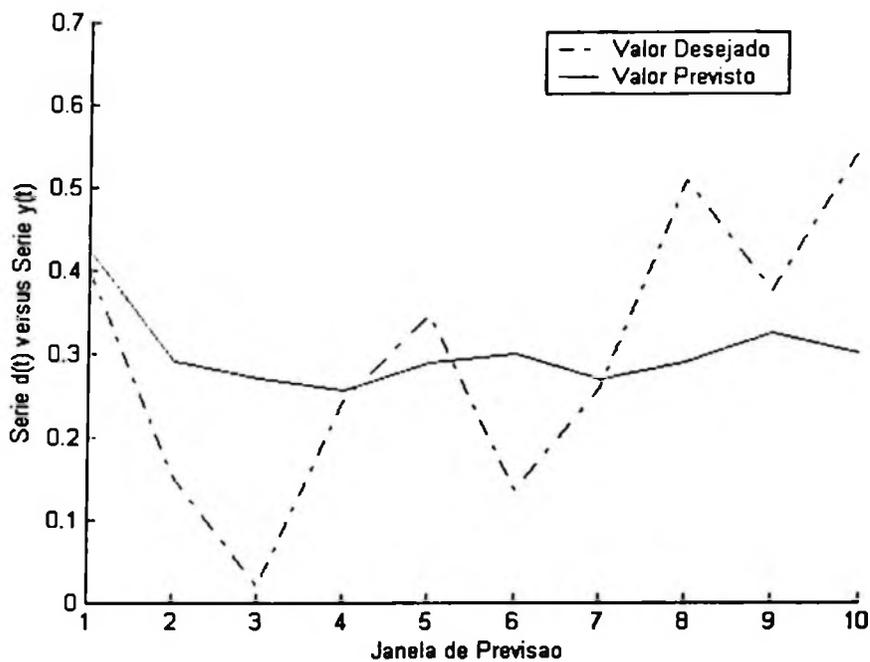


Figura-4.30 Ajuste da curva de previsão de uma rede neural 10-2R-1 para a série NDLSOJA

Estatísticas de previsão :

Topologia da RNA	RMSE	MAE	TIC
30-2R-1	0,15035	0,11749	0,23349

O valor do MAPE da rede neural para previsão de 10 passos à frente é de 165,00.

#### 4.5 Análise da série de retornos das ações da IBM

A série DLIBM de 1500 valores de retornos diários das ações da IBM é mostrada abaixo.

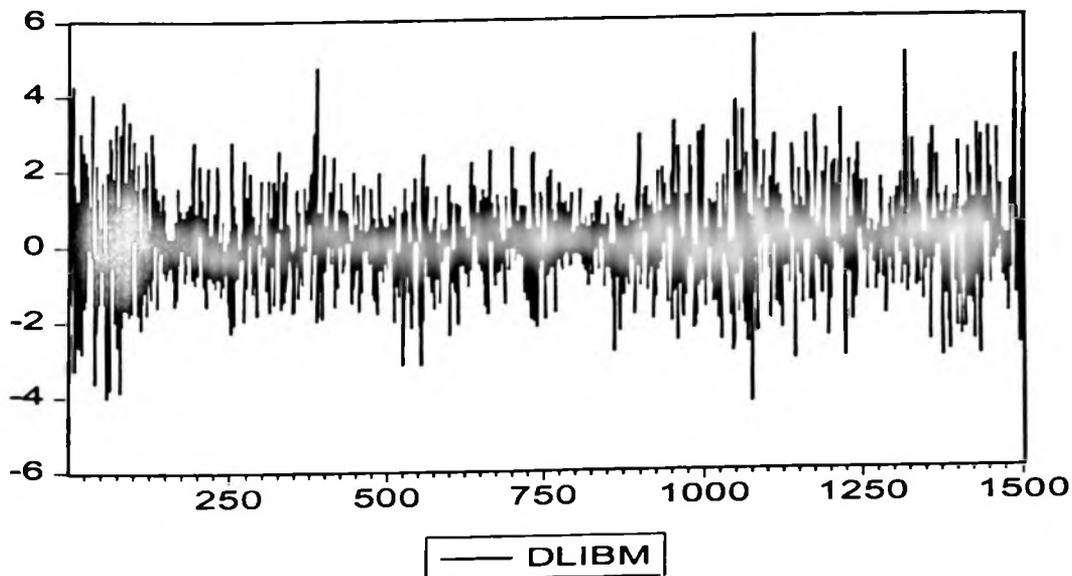


Figura-4.31 Série de retornos diários DLIBM. Fonte: Analysis of Financial Time Series (Tsay, 2003)

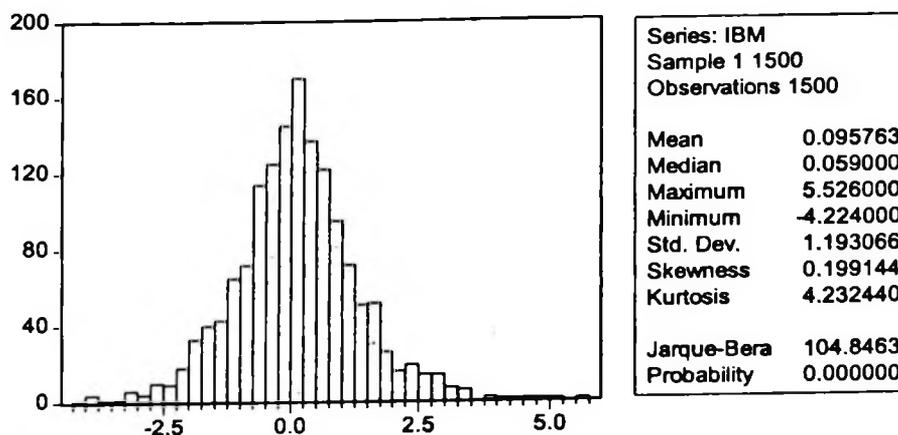


Figura-4.32 Histograma da série de retornos diários DLIBM

Na Figura-4.32 verificamos excesso de curtose e rejeitamos a hipótese de normalidade pelo teste de Jarque-Bera.

Teste de raiz unitária ADF para a série DLIBM:

Estatística do teste ADF	-17,42389	1% Critical Value	-2,5671
		5% Critical Value	-1,9396
		10% Critical Value	-1,6157

Sem constante e sem intercepto.

Estatística do teste ADF	-17,74311	1% Critical Value*	-3,4376
		5% Critical Value	-2,8639
		10% Critical Value	-2,5681

Com intercepto.

Estatística do teste ADF	-17,76384	1% Critical Value*	-3,9694
		5% Critical Value	-3,4153
		10% Critical Value	-3,1295

Com constante e com intercepto.

Rejeitamos a hipótese de presença de uma raiz unitária na série DLIBM pelo teste ADF.

Teste de Dikey-Pantula para uma e duas raízes unitárias é mostrado a seguir.

Hipóteses testadas	Estatística	valores críticos a 5%
$H_0 : 2RU$	-26,17949	-3,37
$H_a : 1RU$	-48,19363	-3,77
$H_0 : 1RU$	-36,19443	-3,42
$H_a : 0RU$	-17,28853	-3,80
	-21,21214	-4,16

Para a série DLIBM rejeitamos as hipóteses de presença de duas e uma raiz unitária pelo teste de Dickey-Pantula.

As funções FAC e FACP da série DLIBM são mostradas na Figura-4.33a e Figura-4.33b, respectivamente. A estatística Q para os valores dos retornos da IBM indicam que cada uma das auto-correlações é menor que dois desvios padrões, dessa forma não vamos estimar um modelo ARIMA para esta série.

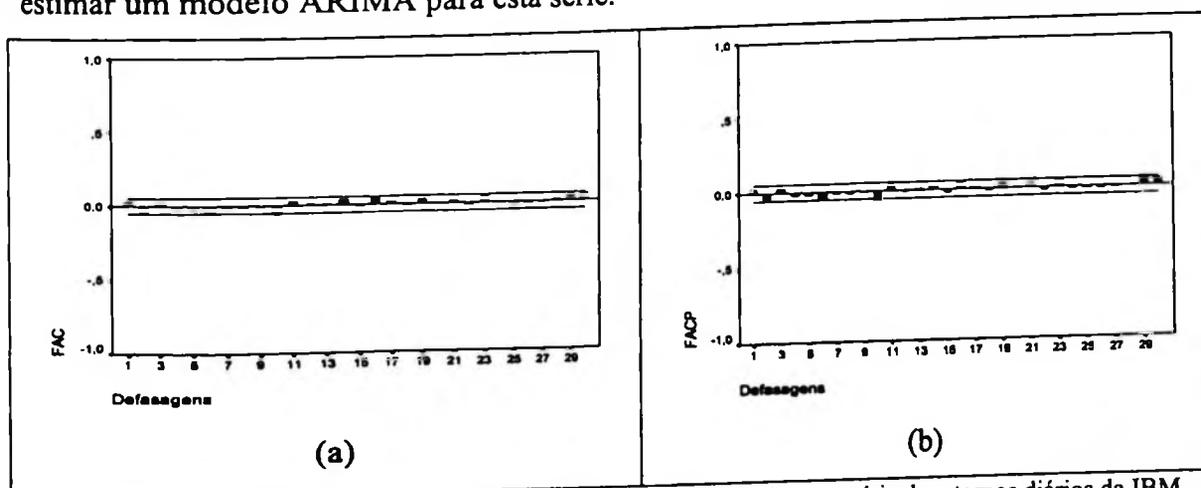


Figura-4.33 Função de auto-correlação e auto-correlação parcial para a série de retornos diários da IBM

Estimando um modelo GARCH(1,1) para esta série obtemos os seguintes resultados:

$$RIBM_t = \varepsilon_t$$

$$\varepsilon_t^2 | I_{t-1} \sim N(0; h_t)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1}$$

Parâmetros	Coefficiente	Erro-padrão	Estatística z	Significância
$\alpha_0$	0,036268	0,016843	2,153276	0,0313
$\alpha_1$	0,071879	0,016850	4,265875	0,0000
$\beta$	0,903874	0,023981	37,69100	0,0000

O teste ARCH-LM para o modelos GARCH(1,1), para duas defasagens produz o seguinte resultado:

ARCH Test:

Obs\*R-squared      4,021653      Probability      0,133878

Portanto, não rejeitamos a hipótese “não existe efeitos ARCH” no modelo.

### Testes de Diagnóstico

Abaixo são apresentados os teste de Ljung-Box (Q (K) e  $Q^2(K)$ ) e o teste ARCH-LM para o modelo GARCH(1,1).

Estatística	GARCH(1,1)	P-value
Q(3)	8,7926	0,032
Q(6)	13,600	0,034
Q(24)	26,632	0,322
$Q^2(3)$	4,5271	0,210
$Q^2(6)$	6,9506	0,325
$Q^2(24)$	11,192	0,988
LM	7,183048	0,708059

### Estatísticas de previsão para o modelo GARCH(1,1)

Estatísticas de previsão dinâmica e ajuste para o modelo GARCH(1,1)				Estatísticas de previsão estática e ajuste para o modelo GARCH(1,1)			
Passos a Frente	rmse	mae	TIC	Passos a Frente	rmse	mae	TIC
1	1,352766	1,351500	1,00000	1	1,352766	1,351500	1,00000
2	1,201487	1,174000	1,00000	2	1,201487	1,174000	1,00000
3	1,075230	1,016000	1,00000	3	1,075230	1,016000	1,00000
4	0,965948	0,853200	1,00000	4	0,965948	0,853200	1,00000
5	0,885672	0,744833	1,00000	5	0,885672	0,744833	1,00000
6	1,320935	1,029857	1,00000	6	1,320935	1,029857	1,00000
7	1,305059	1,049625	1,00000	7	1,305059	1,049625	1,00000
8	1,511562	1,225667	1,00000	8	1,511562	1,225667	1,00000
9	1,434175	1,110300	1,00000	9	1,434175	1,110300	1,00000
10	1,696810	1,312273	1,00000	10	1,696810	1,312273	1,00000

Utilizando uma rede 10-2R-1 obtemos o seguinte resultado:

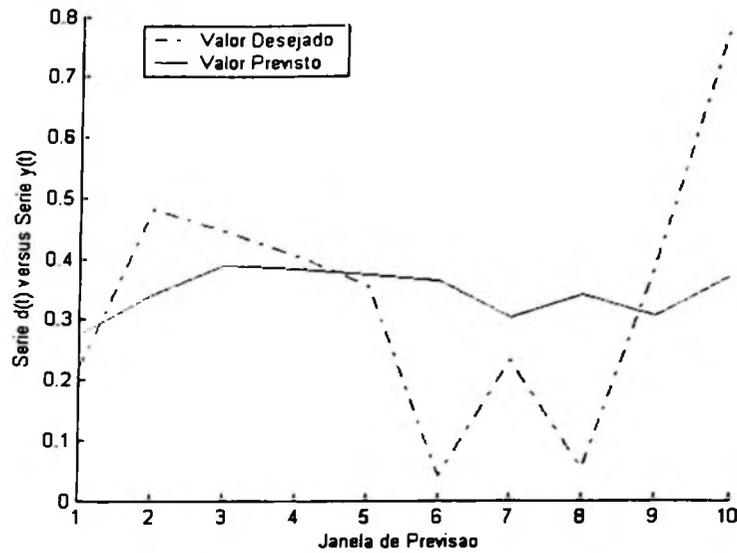


Figura-4.34 Ajuste da curva de previsão de uma rede neural 10-2R-1 para a série NDLIBM

Estatísticas de previsão :

Topologia da RNA	RMSE	MAE	TIC
10-2R-1	0,20018	0,14745	0,2682

O valor do MAPE da rede neural para previsão de 10 passos à frente é de 151,904.

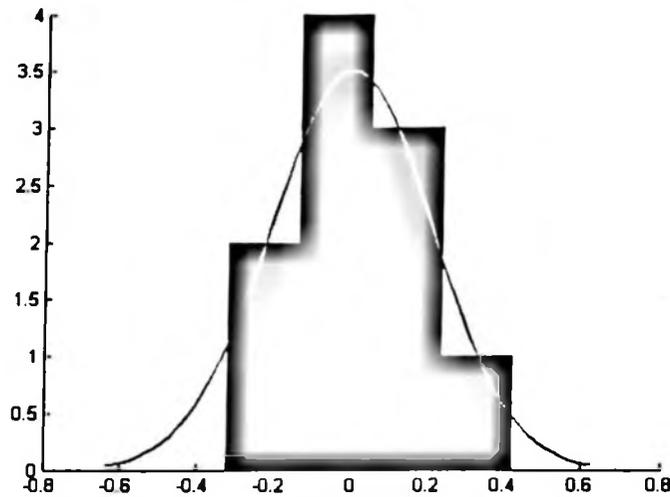


Figura-4.35 Histograma dos resíduos da previsão da RNA 10-2R-1 para a série NDLIBM

## 4.6 Comparação dos resultados obtidos pelas previsões realizadas pelos modelos ARIMA-GARCH e redes neurais

Neste tópico vamos fazer uma comparação entre as previsões obtidas pela metodologia de processos ARIMA-GARCH e de RNAs recorrentes. Para isso serão utilizadas as séries do IBOVESPA, Petrobrás e de preço da saca de 60Kg de soja.

### Série IBOVESPA

A Figura-4.36 apresenta os resultados das previsões obtidas para a série no nível, ou seja, é feita uma comparação entre a série de valores brutos e os valores previstos pela metodologia ARIMA-GARCH.

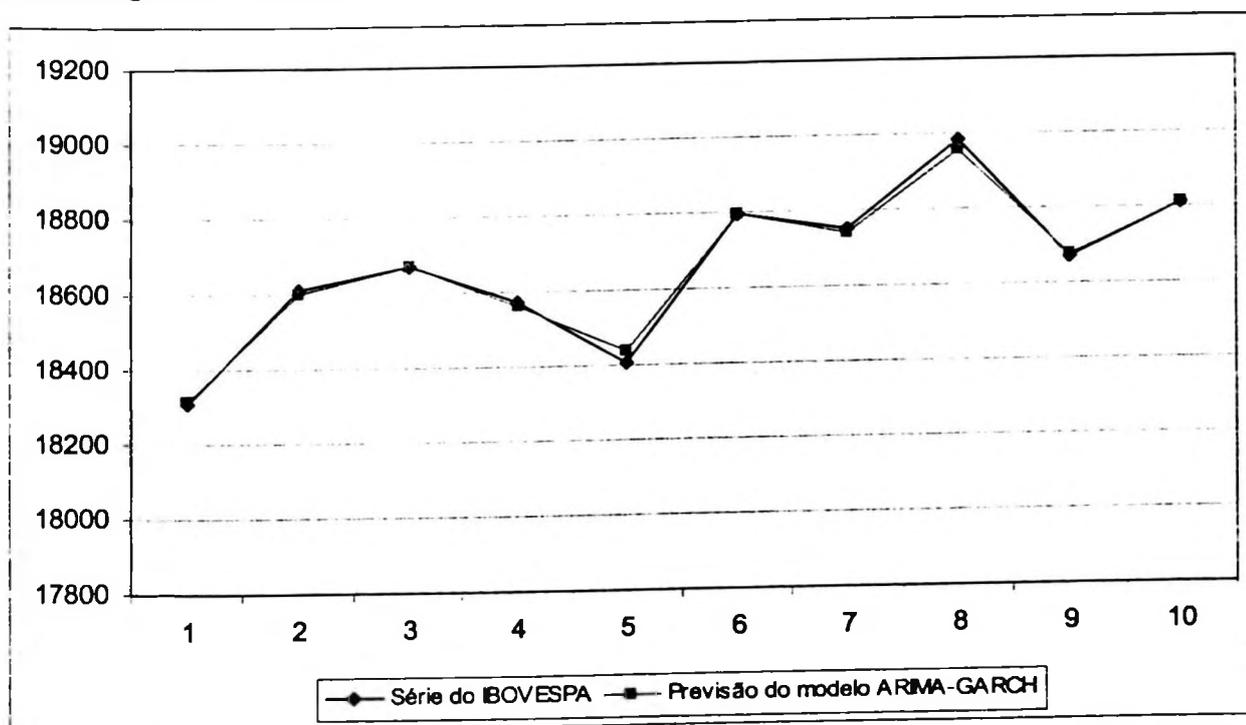


Figura-4.36 Valores da série do IBOVESPA no nível e da curva de previsão do modelo ARIMA-GARCH

Os valores do coeficiente de desigualdade de Theil (TIC) e do erro percentual absoluto médio (MAPE) para a série do IBOVESPA são:

$$\text{MAPE-ARIMA} = 0,000591761$$

$$\text{TIC-ARIMA} = 1,64869\text{E-}11$$

Os valores acima resultantes da previsão do modelo AR([16])-GARCH(2,1).

A Figura-4.37 apresenta a curva dos valores da previsão obtida pela modelagem econométrica e pela RNA.

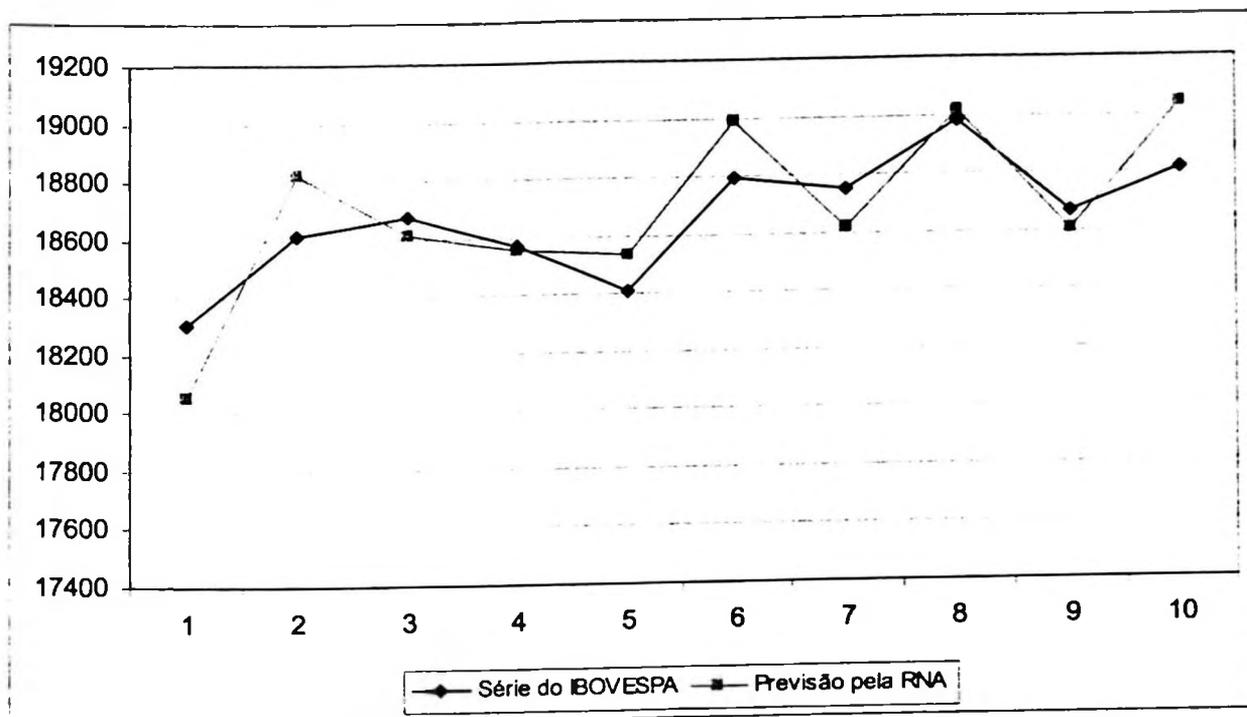


Figura-4.37 Valores da série do IBOVESPA no nível e da curva de previsão da RNA

No caso das redes neurais o valor de MAPE e do TIC são:

$$\text{MAPE} = 0,00706$$

$$\text{TIC} = 1,86325\text{E-}09$$

Comparando os valores obtidos para a série no nível do IBOVESPA podemos verificar que as previsões dos modelos ARIMA-GARCH se ajustaram melhor e apresentam MAPE menor que os valores obtidos pela RNA.

### Série Petrobrás

No caso da série da Petrobrás a Figura-4.38 mostra a curva da série no nível e os valores previstos pelo modelo AR([1])-GARCH(1,1).

Para a série da Petrobrás as estatísticas obtidas no modelo ARIMA-GARCH são:

$$\text{MAPE} = 0,001799008$$

$$\text{TIC} = 6,45694\text{E-}11$$

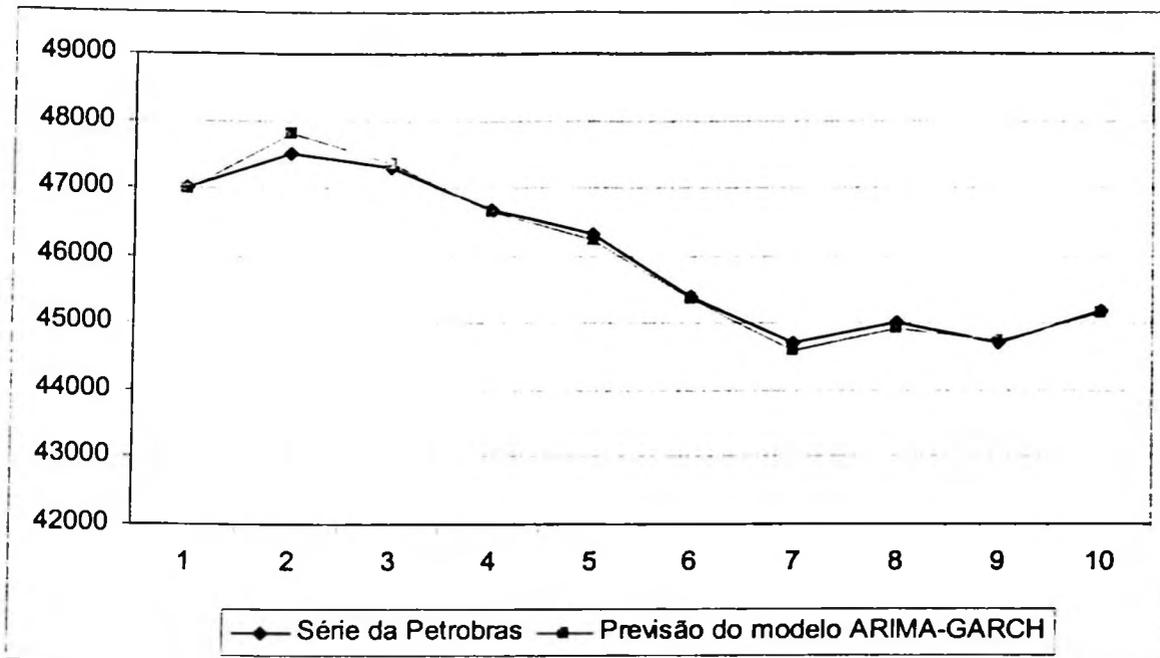


Figura-4.38 Valores da série da Petrobrás no nível e da curva de previsão do modelo ARIMA-GARCH

A comparação entre os valores previstos pela RNA e os valores da série no nível da Petrobrás é mostrada abaixo.

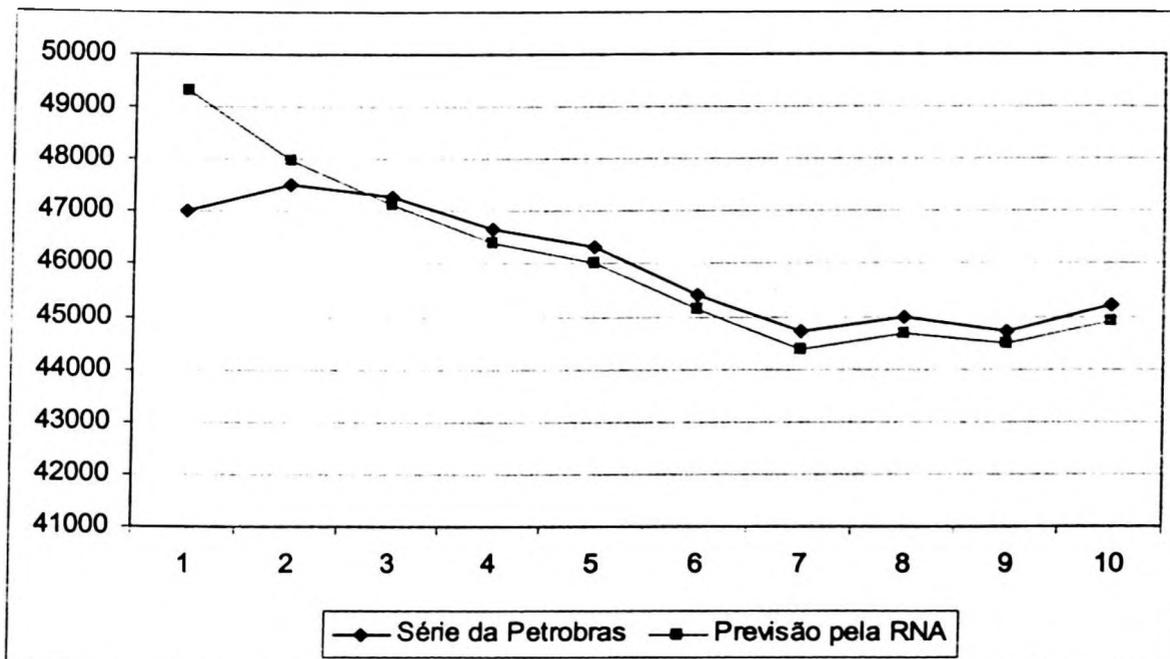


Figura-4.39 Valores da série da Petrobrás no nível e da curva de previsão da RNA

Os valores do MAPE e do TIC para as RNAs são:

MAPE = 0,01063

TIC = 2,8438E-09

Os valores das estatísticas de previsão da série da Petrobrás do modelo ARIMA-GARCH são menores que aqueles da RNA.

### Série da saca de 60Kg de soja

Na figura abaixo é mostrado o resultado obtido pelo modelamento ARIMA-GARCH da série do preço da saca de 60Kg soja.

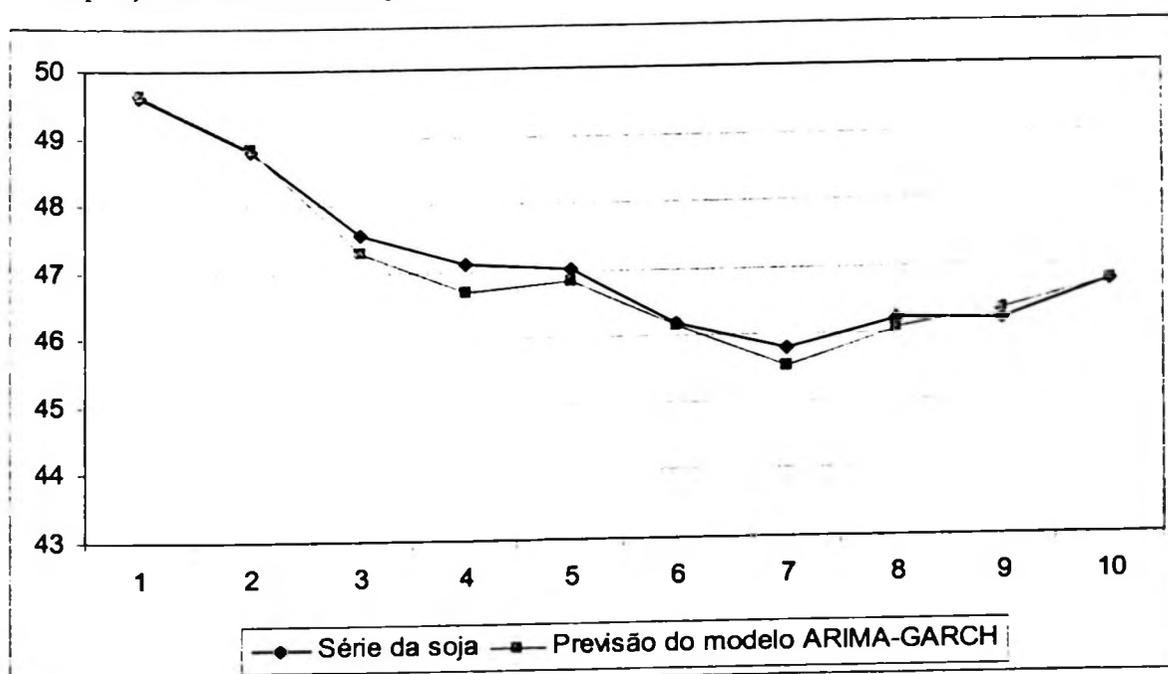


Figura-4.40 Valores da série do preço da saca de 60Kg de soja no nível e a curva de previsão do modelo ARIMA-GARCH

As estatísticas MAPE e TIC obtidas pelo modelo AR([1])-GARCH(1,1) para a série da saca de 60Kg de soja são:

MAPE = 0,003196428

TIC = 1,8902E-07

A RNA construída para prever os valores da série da soja forneceu os seguintes valores:

MAPE = 0,006485

TIC = 5,64399E-07

Na Figura-4.41 tem-se o ajuste obtido pela rede 10-2R-1.

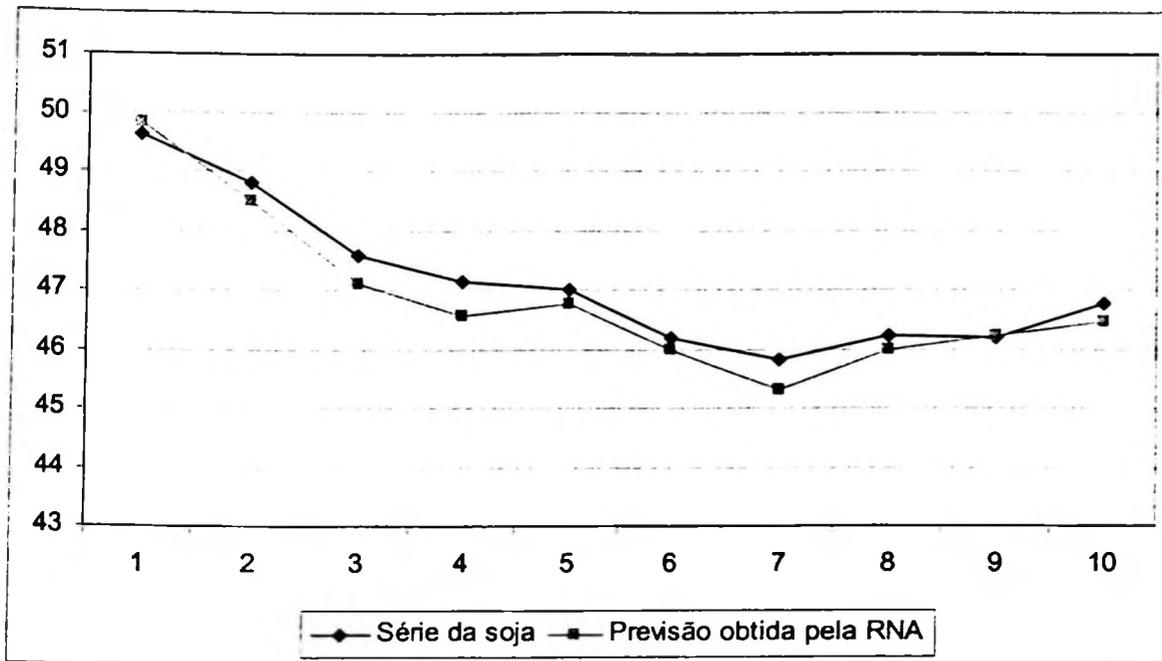


Figura-4.41 Valores da série do preço da saca de 60Kg de soja no nível e a curva de previsão da RNA

No caso da série do preço da soja o modelo  $AR([1])$ -GARCH(1,1) forneceu uma previsão mais precisa que a rede neural 10-2R-1.

## Conclusões 5

Neste trabalho o objetivo principal foi estudar os métodos tradicionais da econometria e os algoritmos de redes neurais para realizar a tarefa de previsão de séries temporais econômico-financeiras. Dessa forma, foram analisadas as séries univariadas de retornos do IBOVESPA, Nasdaq, IBM, Petrobrás e da saca de 60Kg de soja. Para estas séries, conforme mostrado no capítulo 4, foram identificados efeitos de heterocedasticidade condicionada auto-regressiva.

As previsões foram realizadas através de modelos ARIMA-GARCH e redes neurais artificiais (RNAs) que utilizam o algoritmo de aprendizagem recorrente em tempo real (ARTR). As metodologias estudadas foram aplicadas separadamente para cada série, assim pode-se avaliar que a seqüência empregada na construção dos modelos ARIMA-GARCH tem maior rigor estatístico em cada etapa do processo para realizar a previsão, no caso das redes neurais o trabalho principal constituiu-se basicamente na implementação de um algoritmo de otimização. Uma vez que o algoritmo da rede neural foi transformado em um programa de computador os dados de entrada são inseridos e tem-se um valor de resposta, ou seja, uma saída. Em muitos casos, essa rede neural implementada pode ser utilizada para a tarefa de previsão de séries temporais com comportamentos distintos, por exemplo, tanto para séries de alta, quanto de baixa frequência, por essa falta de critério no uso da RNA para resolver um problema específico tem-se rotulado em muitas aplicações a RNA como uma caixa preta.

Uma diferença importante que deve ser destacada entre as RNAs recorrentes e os modelos ARIMA-GARCH é a capacidade das RNA de capturar padrões não-lineares subjacentes ao comportamento da série, enquanto que isso não acontece na modelagem ARIMA-GARCH.

No caso da série de retornos do IBOVESPA pode-se observar que o modelo ARIMA-GARCH gerou uma curva de valores de previsão que acompanhou de forma mais próxima os valores desejados. O modelamento ARIMA-GARCH, para a série de retornos do IBOVESPA, apresentou estatística MAPE menor que a RNA recorrente e o TIC indica um ajuste mais perfeito que o obtido pela rede neural.

Na série de retornos da Petrobrás o comportamento da curva de previsão tanto da rede neural quanto do modelo ARIMA-GARCH tenderam a média condicionada da série. Foram realizadas previsões de dez passos à frente para a série de retornos da Petrobrás e foi construído um histograma dos resíduos obtidos pela RNA, apesar do número de resíduos obtidos ser pequeno, pode-se verificar uma tendência à forma normal. O melhor resultado das previsões da série da Petrobrás no nível foi encontrado utilizando-se o modelamento ARIMA-GARCH, sendo que o TIC desse modelo indica um ajuste mais perfeito que o obtido pela RNA.

No modelamento da série de retornos da Nasdaq pode-se constatar que não foi necessário um modelo ARIMA para a série, ou seja, a série foi modelada diretamente como um processo GARCH, uma explicação para esse fato pode ser que a hipótese de mercado eficiente aplica-se para a série da Nasdaq.

Para as séries de retornos verificou-se que as curvas de valores de previsão dos modelos ARIMA-GARCH tendiam à um valor médio da série a ser prevista, como a série de retornos da Nasdaq apresentou um comportamento de ruído com média zero, os valores previstos foram iguais a zero e dessa forma o coeficiente de desigualdade de Theil (TIC), que mede o ajuste da curva de previsão, indicou que o ajuste foi o menos perfeito dentro da escala. A RNA em contrapartida apresentou um ajuste muito melhor considerando o TIC.

Na série de retornos da saca de 60Kg de soja o modelo ARIMA-GARCH obteve um valor de erro percentual absoluto médio menor que o obtido pela RNA, e o ajuste da curva de previsão foi melhor quando comparado com os modelos ARIMA-GARCH das demais séries analisadas. Nesta série de retornos da soja verificou-se claramente um comportamento auto-regressivo de primeira ordem pelas funções FAC e FACP. Quanto a série do preço da soja no nível o modelo ARIMA-GARCH também apresentou melhores resultados que a rede recorrente.

A série de retornos da IBM apresentou um comportamento parecido com a série de retornos da Nasdaq, ou seja, a partir das funções FAC e FACP não foi possível identificar um modelo ARIMA e o valor do TIC indicou o ajuste menos perfeito dentro dessa escala.

Para a análise de séries temporais, ou sucessões cronológicas, utilizando modelos ARIMA existe uma grande quantidade de programas de computadores disponíveis que fornecem vários testes estatísticos e procedimento de estimação que seguem padrões muito parecidos. No caso das redes neurais artificiais ainda não existem tantas opções e a determinação da arquitetura adequada de uma RNA não é uma tarefa fácil. Por exemplo, não existem regras claras para determinar o número de neurônios na camada intermediária, o número adequado de camadas e a topologia da rede. Para a implementação das redes neurais recorrentes utilizadas neste trabalho foi usado um programa de computador de baixa flexibilidade para manutenção. A complexidade de implementação é uma das desvantagens atuais para a utilização da RNA, sendo que o aumento do número de neurônios e de conexões também aumenta o tempo de processamento para se obter um valor de saída.

A aplicação do filtro de Kalman (FK) em redes neurais é uma das técnicas que têm sido estudadas para reduzir o número de iterações do algoritmo da RNA para convergir a um certo valor, uma extensão desse trabalho é a implementação de redes que utilizem o FK para realizar a tarefa de previsão.

Uma conclusão importante desse trabalho é que a unificação da notação entre modelos econométricos e algoritmos de redes neurais é fundamental para que as metodologias possam ser utilizadas de maneira mais eficiente e para que sejam construídas metodologias híbridas que explorem as principais características de cada técnica.

O modelamento de séries econômico-financeiras fornece informações úteis para a tomada de decisão, principalmente para planejamentos estratégicos. No entanto, com os resultados obtidos nesse trabalho não podemos concluir de forma precisa que uma metodologia seja superior a outra. A capacidade das RNAs de lidarem com séries não-lineares é uma vantagem, as curvas de valores previstos que foram obtidas são uma indicação dessa valiosa característica, porém a falta de procedimentos estatísticos para avaliar os parâmetros estimados da rede é uma limitação que merece atenção e torna-se um desafio para trabalhos futuros.

## Apêndice 6

### 6.1 Previsão

Neste tópico mostraremos como os modelos ARIMA podem ser utilizados para prever valores futuros de uma série temporal observada.

Considere o seguinte processo:

$$\varphi(B)z_t = \theta(B)a_t \quad (6.1)$$

sendo que  $\varphi(B)z_t = \phi(B)\nabla^d$  e  $\{a_t\} \sim \text{RB}$ . Vamos considerar a previsão de um valor  $z_{t+l}$ ,  $l \geq 1$ , quando estamos no momento de tempo contemporâneo  $t$ . Esta previsão é dita ser realizada na *origem*  $t$  para um passo de tempo  $l$ .

**Três forma explícitas para o modelo:** uma observação  $z_{t+l}$  gerada pelo processo (6.1) pode ser expressada como segue:

1. Diretamente em termos da equação de diferença por:

$$z_{t+l} = \varphi_1 z_{t+l-1} + \dots + \varphi_{p+d} z_{t+l-p-d} - \theta_1 a_{t+l-1} - \dots - \theta_q a_{t+l-q} + a_{t+l} \quad (6.2)$$

2. Como uma soma ponderada infinita dos choques atuais e anteriores  $a_j$ ,

$$z_{t+l} = \sum_{j=0}^{\infty} \psi_j a_{t+l-j} \quad (6.3)$$

sendo que  $\psi_0 = 1$  e os pesos  $\psi$  devem ser obtidos igualando-se os coeficientes em:

$$\varphi(B)(1 + \psi_1 B + \psi_2 B^2 + \dots) = \theta(B) \quad (6.4)$$

Equivalentemente, para  $l$  positivo, com referência a origem  $k < t$ , o modelo pode ser escrito na forma truncada:

$$\begin{aligned} z_{t+l} &= a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1} + \psi_l a_t + \dots + \psi_{t+l-k-1} a_{k+1} + C_k(t+l-k) \\ &= a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1} + C_l(l) \end{aligned} \quad (6.5)$$

sendo que  $C_k(t+l-k)$  é a função complementar relativa a origem finita  $k$  do processo.

$C_l(l)$  está associada com a soma infinita truncada (Box, Jenkins e Reinsel, 1994)

$$C_l(l) = \sum_{j=1}^{\infty} \psi_j a_{t+l-j} \quad (6.6)$$

3. Como uma soma ponderada infinita de observações prévias, mais um choque aleatório:

$$z_{t+l} = \sum_{j=0}^{\infty} \pi_j z_{t+l-j} + a_{t+l} \quad (6.7)$$

Também, se  $d \geq 1$ ,

$$\bar{z}_{t+l}(\pi) = \sum_{j=0}^{\infty} \pi_j z_{t+l-j} \quad (6.8)$$

será uma média ponderada, desde que  $\sum_{j=1}^{\infty} \pi_j = 1$ . Os pesos  $\pi$  podem ser obtidos de:

$$\varphi(B) = (1 - \pi_1 B - \pi_2 B^2 - \dots) \theta(B) \quad (6.9)$$

### 6.1.1 Derivação de Previsões de Erro Médio Quadrático Mínimo

Supondo, que estando na origem  $t$ , faremos um previsão  $\hat{z}_t(l)$  de  $z_{t+l}$  a qual é uma função linear de observações atuais e passadas  $z_t, z_{t-1}, z_{t-2}, \dots$ . Então, seja também uma função linear de choques atuais e passados  $a_t, a_{t-1}, a_{t-2}, \dots$ .

Suponha, então, que a melhor previsão é:

$$\hat{z}_t(l) = \psi_l^* a_t + \psi_{l+1}^* a_{t-1} + \psi_{l+2}^* a_{t-2} + \dots$$

sendo que os pesos  $\psi_l^*, \psi_{l+1}^*, \dots$  devem ser determinados. Então, usando (6.3) o erro quadrático médio da previsão é:

$$E[z_{t+l} - \hat{z}_t(l)]^2 = (1 + \psi_1^2 + \dots + \psi_{l-1}^2) \sigma_a^2 + \sum_{j=0}^{\infty} (\psi_{l+j} - \psi_{l+j}^*)^2 \sigma_a^2 \quad (6.10)$$

a qual é minimizada fazendo com que  $\psi_{l+j}^* = \psi_{l+j}$ .

Temos então que:

$$z_{t+l} = (a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1}) + (\psi_l a_t + \psi_{l+1} a_{t-1} + \dots) \quad (6.11)$$

$$= e_t(l) + \hat{z}_t(l) \quad (6.12)$$

sendo que  $e_t(l)$  é o erro da previsão  $\hat{z}_t(l)$  no passo de tempo  $l$ .

Alguns fatos importantes emergem. Denotaremos  $E[z_{t+l} | z_t, z_{t-1}, \dots]$ , a esperança condicionada de  $z_{t+l}$  dado o conhecimento de todos os  $z$ 's além do tempo  $t$ , como sendo  $E_t[z_{t+l}]$ . Assumiremos que os  $a_t$  são uma seqüência de variáveis aleatórias independentes.

1. Então,  $E[a_{t+j} | z_t, z_{t-1}, \dots] = 0$ ,  $j > 0$ , e de (6.3):

$$\hat{z}_t(l) = \psi_l a_t + \psi_{l+1} a_{t-1} + \dots = E_t[z_{t+l}] \quad (6.13)$$

Portanto, o erro quadrático médio mínimo de previsão na origem  $t$ , para um passo de tempo  $l$ , é a esperança condicionada de  $z_{t+l}$  no tempo  $t$ . Quando  $\hat{z}_t(l)$  é uma função de  $l$  para um  $t$  fixado, será chamada de função de previsão para a origem  $t$ . Nós notamos que o requisito mínimo para os choques aleatórios  $a_t$  no modelo (6.1) para a esperança condicionada  $E_t[z_{t+l}]$ , a qual sempre se iguala ao erro quadrático médio mínimo da previsão linear é que  $E_t[a_{t+j}] = 0, j > 0$ . Esta situação pode não acontecer para certos tipos de processos intrinsecamente não lineares. Tais processos podem, de fato, possuir uma representação linear como em (6.1), mas os choques  $a_t$  não serão independentes, serão apenas não-correlacionados, e a melhor previsão  $E_t[z_{t+l}]$  poderá não coincidir com a melhor previsão linear  $\hat{z}_t(l)$  tal como foi obtido em (6.11).

O desenvolvimento anterior, o qual produziu (6.13), baseado apenas informalmente na forma do choque aleatório infinito (6.3) para os casos não estacionários quando  $d \geq 1$ , pode também ser derivado formalmente a partir da forma truncada do modelo (6.5). Para isto, nós assumimos que a função complementar relativa a origem inicial fixa  $k < t$ ,  $C_k(t-k)$ , a

qual já é determinada no tempo  $k$ , é independente de todos os choques aleatórios subsequentes  $a_{k+i}, i > 0$ . Então, tomando a esperança condicionada do lado direito da primeira equação em (6.5) nós obtemos:

$$E_t[z_{t+l}] = \hat{z}_t(l) = \psi_l a_t + \psi_{l+1} a_{t-1} + \dots + \psi_{l+k-1} a_{k+1} + C_k(t+l-k)$$

como o erro de previsão quadrático médio mínimo na origem  $t$ . Esta expressão é portanto informalmente representada pelo segundo termo do lado direito de (6.11).

2. O erro de previsão para  $l$  passos à frente é:

$$e_t(l) = a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1} \quad (6.14)$$

Desde que:

$$E_t[e_t(l)] = 0 \quad (6.15)$$

a previsão é não viesada. Também, a variância do erro de previsão é:

$$V[l] = Var[e_t(l)] = (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{l-1}^2) \sigma_a^2 \quad (6.16)$$

3. Pode ser mostrado que, não somente,  $\hat{z}_t(l)$  é o erro de previsão quadrático médio mínimo de  $z_{t+l}$ , mas que qualquer função linear  $\sum_{i=1}^L w_i \hat{z}_t(i)$  das previsões é um erro de previsão quadrático mínimo da correspondente função linear  $\sum_{i=1}^L w_i z_{t+i}$  das observações futuras. Por exemplo, suponha que usando (6.13), nós obtivemos, para dados mensais, erros de previsão quadráticos mínimos  $\hat{z}_t(1)$ ,  $\hat{z}_t(2)$  e  $\hat{z}_t(3)$  das vendas de um produto 1, 2 e 3 meses à frente. Então é verdade que  $\hat{z}_t(1) + \hat{z}_t(2) + \hat{z}_t(3)$  é o erro de previsão quadrático médio mínimo das vendas  $z_{t+1} + z_{t+2} + z_{t+3}$ , durante o próximo trimestre.

4. *Os choques como erros de previsão de um passo à frente.* Usando (6.14), o erro de previsão de um passo à frente é:

$$e_t(1) = z_{t+1} - \hat{z}_t(1) = a_{t+1} \quad (6.17)$$

Conseqüentemente, os choques  $a_t$ , os quais geraram o processo, e os quais nós introduzimos meramente como um conjunto de variáveis aleatórias independentes ou choques, passaram a ser os erros de previsão um passo à frente.

Segue que para um erro de previsão quadrático médio mínimo, os erros de previsão um passo à frente devem ser não correlacionados. Isto é eminentemente sensato, porque se os

erros de um passo à frente forem correlacionados, os erros de previsão  $a_{t+l}$  poderiam, em alguma extensão, serem preditos de erros de previsão disponíveis  $a_t, a_{t-1}, a_{t-2}, \dots$ . Se a predição obtida foi  $\hat{a}_{t+l}$ , então  $\hat{z}_t(l) + \hat{a}_{t+l}$  poderá ser a melhor previsão de  $z_{t+l}$  do que  $\hat{z}_t(l)$ .

5. *Correlação entre os erros de previsão.* Embora os erros de previsão ótimos no passo de tempo 1 sejam não correlacionados, os erros de previsão para passos de tempo mais longos em geral serão correlacionados.

### 6.1.2 Três Formas Básicas para a Previsão

O erro de previsão quadrático médio mínimo  $\hat{z}_t(l)$  para o passo de tempo  $l$  é a esperança condicionada  $E_t[z_{t+l}]$ , de  $z_{t+l}$  na origem  $t$ . Usando este fato, nós podemos derivar expressões para a previsão em qualquer um dos três diferentes modos, correspondendo aos três diferentes modos de expressar o modelo sumarizados anteriormente. Por simplicidade na notação, adotaremos temporariamente a convenção de que os colchetes implicam que a esperança condicionada, no tempo  $t$ , está sendo tomada. Portanto,

$$[a_{t+l}] = E_t[a_{t+l}] \quad [z_{t+l}] = E_t[z_{t+l}]$$

Para  $l > 0$ , os três diferentes modos de expressar as previsões são:

**Previsões de equações diferentes.** Tomando esperanças condicionadas no tempo  $t$  em (6.2), nós obtemos:

$$[z_{t+l}] = \hat{z}_t(l) = \varphi_1 [z_{t+l-1}] + \dots + \varphi_{p+d} [z_{t+l-p-d}] - \theta_1 [a_{t+l-1}] - \dots - \theta_q [a_{t+l-q}] + [a_{t+l}] \quad (6.18)$$

**Previsões na forma integrada.** Usando (6.3) teremos:

$$[z_{t+l}] = \hat{z}_t(l) = [a_{t+l}] + \psi_1 [a_{t+l-1}] + \dots + \psi_{l-1} [a_{t+1}] + \psi_l [a_t] + \psi_{l+1} [a_{t-1}] + \dots \quad (6.19)$$

produzindo a forma (6.13) que nós já havíamos encontrado. Alternativamente, utilizando a forma truncada do modelo (6.5), nós temos:

$$[z_{t+l}] = \hat{z}_t(l) = [a_{t+l}] + \psi_1 [a_{t+l-1}] + \dots + \psi_{l+k-1} [a_{k+1}] + C_k (t+l-k) \quad (6.20)$$

$$= [a_{t+l}] + \psi_1 [a_{t+l-1}] + \dots + \psi_{l-1} [a_{t+1}] + C_l(l)$$

sendo que  $C_t(l)$  é a função complementar na origem  $t$ .

**Previsões como médias ponderadas de observações prévias e previsões realizadas em passos de tempo prévios da mesma origem.** Finalmente, tomando esperanças condicionadas em (6.7) produz:

$$[z_{t+l}] = \hat{z}_t(l) = \sum_{j=1}^{\infty} \pi_j [z_{t+l-j}] + [a_{t+l}] \quad (6.21)$$

Deve ser notado que o erro de previsão quadrático médio mínimo é definido em termos da esperança condicionada:

$$[z_{t+l}] = E_t [z_{t+l}] = E [z_{t+l} | z_t, z_{t-1}, \dots]$$

a qual teoreticamente requer conhecimento dos  $z$ 's anteriores do passado infinito. No entanto, a condição de invertibilidade, a qual é imposta é aos ao modelo ARIMA geral, assegura que os pesos  $\pi$  em (6.21) formam um série convergente. Conseqüentemente, para a computação de uma previsão para um dado grau de acuracidade, para algum  $k$ , a dependência de  $z_{t-j}$  para  $j > k$  pode ser ignorada. Na prática, os pesos  $\pi$  usualmente decaem rapidamente, tal que, qualquer que seja a forma do modelo empregado na computação, apenas um tamanho moderado da série  $z_t, z_{t-1}, \dots, z_{t-k}$  é necessário para calcular as previsões para uma acuracidade suficiente (Box, Jenkins e Reinsel, 1994). Os métodos discutidos são facilmente modificados para calcular as previsões exatas de amostras finitas,  $E [z_{t+l} | z_t, z_{t-1}, \dots, z_1]$ , baseadas num tamanho finito de dados  $z_t, z_{t-1}, \dots, z_1$ .

Para calcular as esperanças condicionadas nas expressões (6.18) a (6.21), nós notamo que se  $j$  é um inteiro positivo:

$$\begin{aligned} [z_{t-j}] &= E_t [z_{t-j}] = z_{t-j} & j = 0, 1, 2, \dots \\ [z_{t+j}] &= E_t [z_{t+j}] = \hat{z}_t(j) & j = 1, 2, \dots \\ [a_{t-j}] &= E_t [a_{t-j}] = a_{t-j} = z_{t-j} - \hat{z}_{t-j-1}(1) & j = 0, 1, 2, \dots \\ [a_{t+j}] &= E_t [a_{t+j}] = 0 & j = 1, 2, \dots \end{aligned} \quad (6.22)$$

Portanto, para obter a previsão  $\hat{z}_t(l)$ , escrevemos o modelo para  $z_{t+l}$  em qualquer uma das três formas acima e tratamos os termos a direita de acordo com as seguintes regras:

1. Os  $z_{t-j}$  ( $j = 0, 1, 2, \dots$ ), os quais já ocorreram na origem  $t$ , são mantidos inalterados.
2. Os  $z_{t+j}$  ( $j = 1, 2, \dots$ ), os quais não aconteceram ainda, são substituídos pelas suas previsões  $\hat{z}_t(j)$  na origem  $t$ .
3. Os  $a_{t-j}$  ( $j = 0, 1, 2, \dots$ ), os quais aconteceram, estão disponíveis a partir de  $z_{t-j} - \hat{z}_{t-j-1}(1)$ .
4. Os  $a_{t+j}$  ( $j = 1, 2, \dots$ ), os quais não aconteceram ainda, são substituídos por zeros.

Para rotina de cálculo a forma simples é trabalhar diretamente com a equação de diferença na forma (6.18). Conseqüentemente, as previsões são explicitamente calculadas, para  $l = 1, 2, \dots$ , recursivamente como:

$$\hat{z}_t(l) = \sum_{j=1}^{p+d} \varphi_j \hat{z}_t(l-j) - \sum_{j=1}^q \theta_j a_{t+l-j}$$

sendo que  $\hat{z}_t(-j) = [z_{t-j}]$ , denota o valor observado  $z_{t-j}$  para  $j \geq 0$ , e os termos de média móvel não estão presentes para passos de tempo  $l > q$ .

### 6.1.3 Função de Previsão

As previsões são mais simplesmente calculadas pelo uso direto da equação de diferença (Ver Apêndice 7.1). No entanto, do ponto de vista do estudo da natureza das previsões, é vantajoso considerar em detalhes as formas alternativas discutidas na seção anterior, em particular, considerar a forma explícita da função de previsão.

No tempo  $t + l$  os modelos ARIMA podem ser escritos como:

$$z_{t+l} - \varphi_1 z_{t+l-1} - \dots - \varphi_{p+d} z_{t+l-p-d} = a_{t+l} - \theta_1 a_{t+l-1} - \dots - \theta_q a_{t+l-q} \quad (6.23)$$

Tomando esperanças condicionadas no tempo  $t$  em (6.23), nós temos, para  $l > q$ ,

$$\hat{z}_t(l) - \varphi_1 \hat{z}_t(l-1) - \dots - \varphi_{p+d} \hat{z}_t(l-p-d) = 0 \quad l > q \quad (6.24)$$

onde é entendido que  $\hat{z}_t(-j) = z_{t-j}$  para  $j \geq 0$ . A equação de diferença (6.24) tem a solução:

$$\hat{z}_t(l) = b_0^{(l)} f_0(l) + b_1^{(l)} f_1(l) + \dots + b_{p+d-1}^{(l)} f_{p+d-1}(l) \quad (6.25)$$

para  $l > q - p - d$ . Note que a previsão  $\hat{z}_t(l)$  é a função complementar (Ver Apêndice 6.1).

Em (6.25),  $f_0(l), f_1(l), \dots, f_{p+d-1}(l)$  são funções do passo de tempo  $l$ . Em geral, elas podem ser polinomiais, exponenciais, senos, cossenos, e produtos destas funções. Para uma dada origem  $t$ , os coeficientes  $b_j^{(t)}$  são constantes aplicadas para todos os passos de tempo  $l$ , mas eles mudam de uma origem para a próxima, *adaptando* a si próprias apropriadamente a uma parte particular da série que está sendo considerada.

Enquanto o operador autoregressivo decide a natureza da função de previsão, o operador média móvel influencia na determinação de como as funções se ajustam aos dados e conseqüentemente como os coeficientes  $b_0^{(t)}, b_1^{(t)}, \dots, b_{p+d-1}^{(t)}$  são calculados e atualizados.

Por exemplo, considerando o processo IMA(0,2,3):

$$z_{t+l} - 2z_{t+l-1} + z_{t+l-2} = a_{t+l} - \theta_1 a_{t+l-1} - \theta_2 a_{t+l-2} - \theta_3 a_{t+l-3}$$

Usando o argumento da esperança condicionada, a função de previsão é definida por:

$$\hat{z}_t(1) = 2z_t - z_{t-1} - \theta_1 a_t - \theta_2 a_{t-1} - \theta_3 a_{t-2}$$

$$\hat{z}_t(2) = 2\hat{z}_t(1) - z_t - \theta_2 a_t - \theta_3 a_{t-1}$$

$$\hat{z}_t(3) = 2\hat{z}_t(2) - \hat{z}_t(1) - \theta_3 a_t$$

$$\hat{z}_t(l) = 2\hat{z}_t(l-1) - \hat{z}_t(l-2) \quad l > 3$$

Portanto, uma vez que  $\varphi(B) = (1-B)^2$  neste modelo, a função de previsão é a linha reta única:

$$\hat{z}_t(l) = b_0^{(t)} + b_1^{(t)}l \quad l > 1$$

a qual passa por  $\hat{z}_t(2)$  e  $\hat{z}_t(3)$ .

## 6.2 Filtro de Kalman: apresentação dos conceitos

O filtro de Kalman é simplesmente um “algoritmo de processamento de dados recursivo ótimo” (Maybeck, 1979 e Kalman, 1960).

Existem várias formas de definir um critério de otimização avaliar performance. Um aspecto dessa “otimalidade” é que o filtro de Kalman incorpora todas as informações que

possam ser fornecidas a ele. Ele processa todas as medidas disponíveis, a despeito de suas precisões, para estimar o valor corrente das variáveis de interesse, com o uso de:

- (1) conhecimento do sistema e das formas de medição;
- (2) descrição estatística dos ruídos do sistemas, erros de medição e incerteza nos modelos;
- (3) qualquer informação disponível sobre as condições iniciais das variáveis de interesse.

Para um sistema qualquer, ao invés de ignorar qualquer informação ou saída, o filtro de Kalman pode ser construído para combinar todos esses dados e conhecimentos dos vários sistemas dinâmicos para gerar a melhor estimativa global.

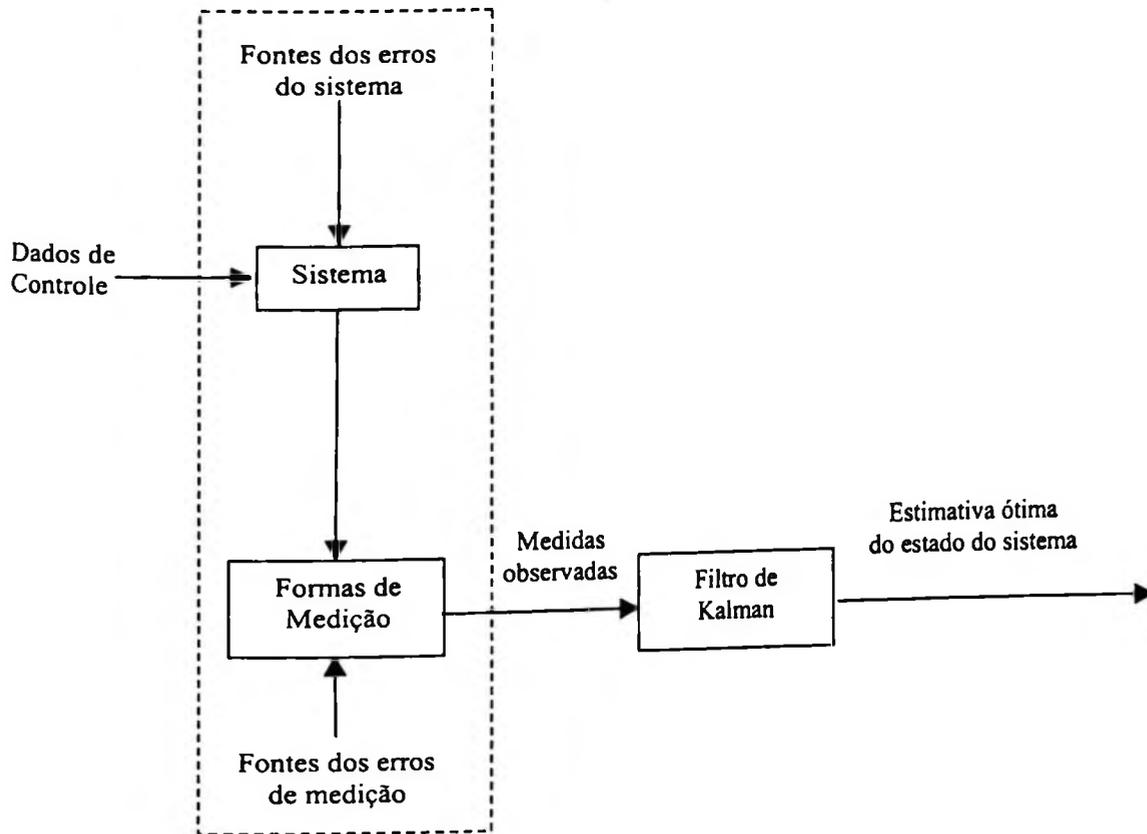
A palavra “recursiva”, na descrição feita previamente, significa que, ao contrário de certos conceitos de processamento de dados, o filtro de Kalman não necessita de todos os dados prévios mantidos armazenados e re-processados a cada vez que uma nova medida é feita. Isto é de vital importância para a praticidade de implementação do filtro.

O “filtro” é um algoritmo de processamento de dados.

A despeito da típica conotação do filtro como uma “caixa preta” contendo redes elétricas, o fato é que na maioria das aplicações práticas, o filtro é apenas um programa de computador. Dessa forma, ele incorpora amostras de medidas de tempo-discreto ao invés de entradas de tempo contínuo.

A Figura-3.13 descreve uma situação típica na qual o filtro de Kalman pode ser usado vantajosamente. O sistema de alguma forma é guiado por algum controle conhecido, e as formas de medição fornecem o valor de certas quantidades pertinentes. O conhecimento dessas entradas e saídas do sistema é tudo o que está explicitamente disponível do sistema para propósitos de estimação.

A “necessidade” para um filtro agora torna-se aparente.



**Figura-6.1** Representação de uma forma de aplicação do filtro de Kalman

Freqüentemente as variáveis de interesse, alguns números finitos de quantidades para descrever o “estado” do sistema, não podem ser medidas diretamente e alguns meios de inferir estes valores dos dados disponíveis devem ser gerados. Esta inferência é complicada pelo fato de que o sistema é tipicamente guiado por entradas outras que não são os nossos próprios controles conhecidos e que as relações entre as várias variáveis de “estado” e saídas medidas são conhecidas apenas com algum grau de incerteza.

Além disso, qualquer medida será corrompida de algum grau pelo ruído, vieses, e imprecisões das formas de medição, e assim um meio de extrair informação valiosa de um sinal ruidoso deve ser fornecida. Deve haver também um número de diferentes formas de medição, cada uma com a sua dinâmica particular e erros característicos, que forneçam alguma informação sobre a variável particular, e poderá ser desejável combinar suas saídas de uma maneira ótima e sistemática. Um filtro de Kalman combina todos os dados de medidas disponíveis, mais o conhecimento a priori sobre o sistema e as formas de medição, para produzir uma estimativa de variáveis desejadas de tal maneira que o erro seja

minimizado estatisticamente. Em outras palavras, se nós estivéssemos para rodar um número de filtros candidatos muitas vezes para a mesma aplicação, então o resultado médio do filtro de Kalman deverá ser melhor que os resultados médios de qualquer outro (Maybeck, 1979).

Conceitualmente, o que qualquer tipo de filtro tenta fazer é obter uma estimativa ótima de quantidades desejadas dos dados fornecidos por um ambiente ruidoso, “ótimo” significando que ele minimiza os erros de alguma forma.

Um filtro de Kalman realiza a propagação de densidade de probabilidade condicional para problemas nos quais o sistema pode ser descrito através de um (i) *modelo linear* e no qual os ruídos do sistema e de medição são (ii) *brancos* e (iii) *Gaussianos*. Sob estas condições, a média, moda, mediana e virtualmente qualquer escolha razoável para uma estimativa “ótima” tem todos coincidindo, então existe de fato uma estimativa “melhor” do conjunto de valores dos quais dispomos. Sob estas três restrições (i, ii e iii), o filtro de Kalman pode ser mostrado como sendo o melhor filtro de qualquer forma concebível. Algumas das restrições podem ser relaxadas, produzindo um filtro qualificado melhor. Por exemplo, se a suposição *Gaussiana* for removida, o filtro de Kalman pode ser mostrado como sendo o melhor filtro (variância de erro mínima) fora da classe de filtros lineares não viesados. Porém, estas três suposições podem ser justificadas para muitas aplicações potenciais.

Suposições Básicas: neste ponto é importante olhar para três suposições básicas na formulação do filtro de Kalman. Numa primeira inspeção, elas podem aparentar ser de sobremaneira restritivas e irrealistas. Para minimizar quaisquer dúvidas desta espécie, esta seção discutirá brevemente as implicações destas suposições.

Um modelo de sistema linear é justificável por um número de razões. Frequentemente um tal modelo é adequado para o propósito a mão (que se tem no momento), e quando não-linearidades existem, uma abordagem adotada é linearizar sobre algum ponto nominal ou trajetória, alcançando um modelo de perturbação ou modelo de erro. Sistemas lineares são desejáveis por serem mais facilmente manipulados com ferramentas de análise e a teoria de sistemas lineares (ou equações diferenciais) é muito mais completa e prática que a não-linear. O fato é que existem meios de estender o conceito de filtro de Kalman para algumas aplicações não-lineares ou desenvolvimento de filtros não-lineares diretamente, mas estes são considerados apenas se os modelos lineares provarem ser inadequados.

Ser “branco” para um ruído implica que o valor do ruído não está correlacionado no tempo. Afirmado de maneira mais simples, se você sabe qual o valor do ruído agora, este conhecimento não trás a você nenhum benefício na predição de qual valor será em qualquer outro tempo. Isto transforma a matemática envolvida no filtro, tornando-a vastamente simplificada (de fato, tornou tratável). Portanto, o modelo ruído branco é usado.

Considerando a “brancura” pertencente as relações de tempo ou freqüência de um ruído, ser *Gaussiano* tem a haver com sua amplitude. Deste modo, em qualquer ponto individual no tempo, a densidade de probabilidade de uma amplitude de ruído *Gaussiano* toma forma de uma curva normal em forma de sino.

Esta suposição pode ser justificada pelo fato de que um sistema ou ruído de medida é tipicamente causado por um número de pequenas fontes. Isto pode ser mostrado matematicamente, que quando um número de variáveis aleatórias independentes são adicionadas conjuntamente, o efeito somado pode ser descrito muito estreitamente por uma densidade de probabilidade *Gaussiana*, apesar da forma das densidades individuais.

Existe também uma justificativa prática para usar densidades *Gaussianas*.

Similarmente a “brancura”, ela torna a matemática tratável. Mas, mais do que isto, tipicamente um analista desejará saber, a primeira e segunda ordem estatística (média e variância ou desvio padrão) de um processo de ruído. Na ausência de qualquer ordem estatística maior, não há uma melhor forma para assumir que a densidade é *Gaussiana*. A primeira e a segunda ordem estatística determinam completamente uma densidade *Gaussiana*, diferente de outras densidades que requerem um número muito grande de ordens estatísticas para poder especificar sua forma completamente. Dessa maneira, o filtro de Kalman, o qual propaga a primeira e a segunda ordem estatísticas, inclui todas as informações contidas na densidade de probabilidade condicional, ao invés de apenas algumas delas, como poderia ser o caso com uma forma de densidade diferente.

**Exemplo paradigmático:** para ver como o filtro de Kalman trabalha, vamos desenvolver um exemplo.

Suponha que você esteja perdido no mar durante a noite e não tenha idéia exata de sua localização. Então você toma uma estrela brilhando para estabelecer a sua posição (por motivo de simplicidade, considere uma localização unidimensional). Em algum tempo  $t_1$

você determina sua localização sendo  $z_1$ . Porém, por causa das imprecisões inerentes do instrumento de medida, erros humanos, e assim por diante, o resultado de sua medição é algo incerto. Digamos que você decida que a precisão é de tal forma que o desvio padrão (valor de um sigma) envolvido é  $\sigma_{z_1}$  (ou equivalentemente, a variância, ou segunda ordem estatística, seja  $\sigma_{z_1}^2$ ). Então, você pode estabelecer a probabilidade condicional de  $x(t_1)$ , sua posição no tempo  $t_1$ , condicionada ao valor observado da medida sendo  $z_1$ , como mostrado na Figura-6.2. Esta é uma plotagem de  $f_{x(t_1)|z_1}(x|z_1)$  como uma função da localização  $x$ : isto diz a você a probabilidade de estar em qualquer localização, baseada sobre a medida que você tomou. Note que  $\sigma_{z_1}$  é uma medida direta de incerteza: quanto maior  $\sigma_{z_1}$ , mais largo o pico de probabilidade, estendendo o peso de probabilidade sobre um limite maior dos valores de  $x$ . Para uma densidade Gaussiana, 68,3% do peso de probabilidade está contido na banda de  $\sigma$  unidades para cada lado da média, a porção sombreada na Figura-6.2.

Baseado na densidade de probabilidade condicional, o melhor estimador da sua posição é:

$$\hat{x}(t_1) = z_1$$

e a variância do erro no estimador é:

$$\sigma_x^2(t_1) = \sigma_{z_1}^2$$

note que  $\hat{x}$  é tanto a moda (pico) como a mediana (valor com  $\frac{1}{2}$  do peso de probabilidade de cada área), bem como a média (centro de massa).

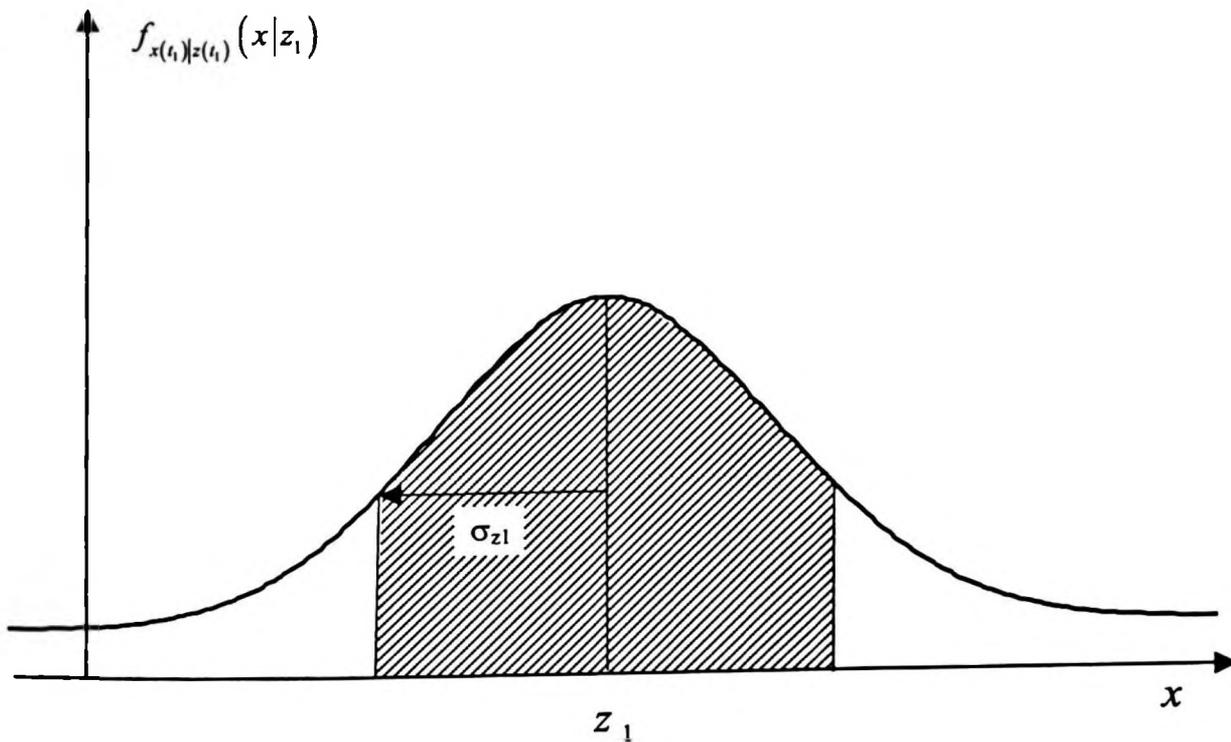


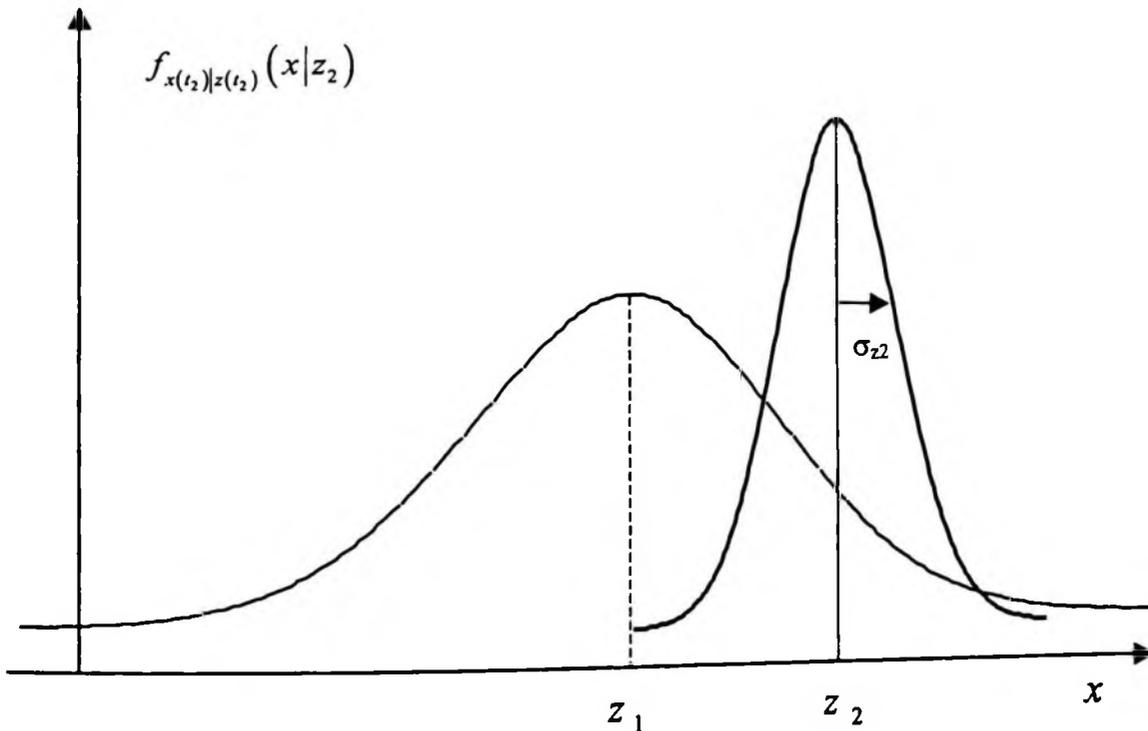
Figura-6.2 Densidade condicional da posição baseada no valor da medida  $z_1$

Agora digamos que um amigo que seja um navegador treinado toma uma fixação independente diretamente após você o fazer, no tempo  $t_2 \cong t_1$  (de tal modo que a posição verdadeira não mudou no total), e obtendo a medida  $z_2$  com uma variância  $\sigma_{z_2}^2$ . Devido a ele ter uma maior habilidade, assuma que a variância na medida dele seja um tanto menor que a sua. A Figura-6.3 apresenta a densidade condicional da sua posição no tempo  $t_2$ , baseada apenas no valor medido  $z_2$ . Note o pico mais estreito devido a menor variância, indicando que você está mais propriamente certo da sua posição baseado na medida dele. Neste ponto, você tem duas medidas disponíveis para estimar sua posição. A questão é, como você combina estes dados?

A densidade condicional da sua posição no tempo  $t_2 \cong t_1$ ,  $x(t_2)$ , dada por  $z_1$  e  $z_2$ , é uma densidade Gaussiana com média  $\mu$  e variância  $\sigma^2$  como indicado na Figura-6.3, com:

$$\mu = \left[ \frac{\sigma_{z_2}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} \right] z_1 + \left[ \frac{\sigma_{z_1}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} \right] z_2 \quad (6.26)$$

$$\frac{1}{\sigma^2} = \frac{1}{\sigma_{z_1}^2} + \frac{1}{\sigma_{z_2}^2} \quad (6.27)$$



**Figura-6.3** Densidade condicional da posição baseada apenas na medida  $z_2$

Note que, de (6.27),  $\sigma$  é menor que  $\sigma_{z_1}$  ou  $\sigma_{z_2}$ , o qual está dizendo que a incerteza no seu estimador da posição decresceu combinando os dois pedaços de informação.

Dada esta densidade, o melhor estimador é:

$$\hat{x}(t_2) = \mu$$

Com uma variância de erro associada de  $\sigma^2$ . Ela é a moda e a média (ou, desde que ela é a média da densidade condicional ela também é chamada de média condicional). Além disso, ela é também o estimador de máxima verossimilhança, o estimador de mínimos quadrados ponderado, e o estimador linear para o qual a variância é menor que qualquer um outro estimador não-viesado. Em outras palavras, é o “melhor” que você pode fazer de acordo com qualquer um outro critério razoável.

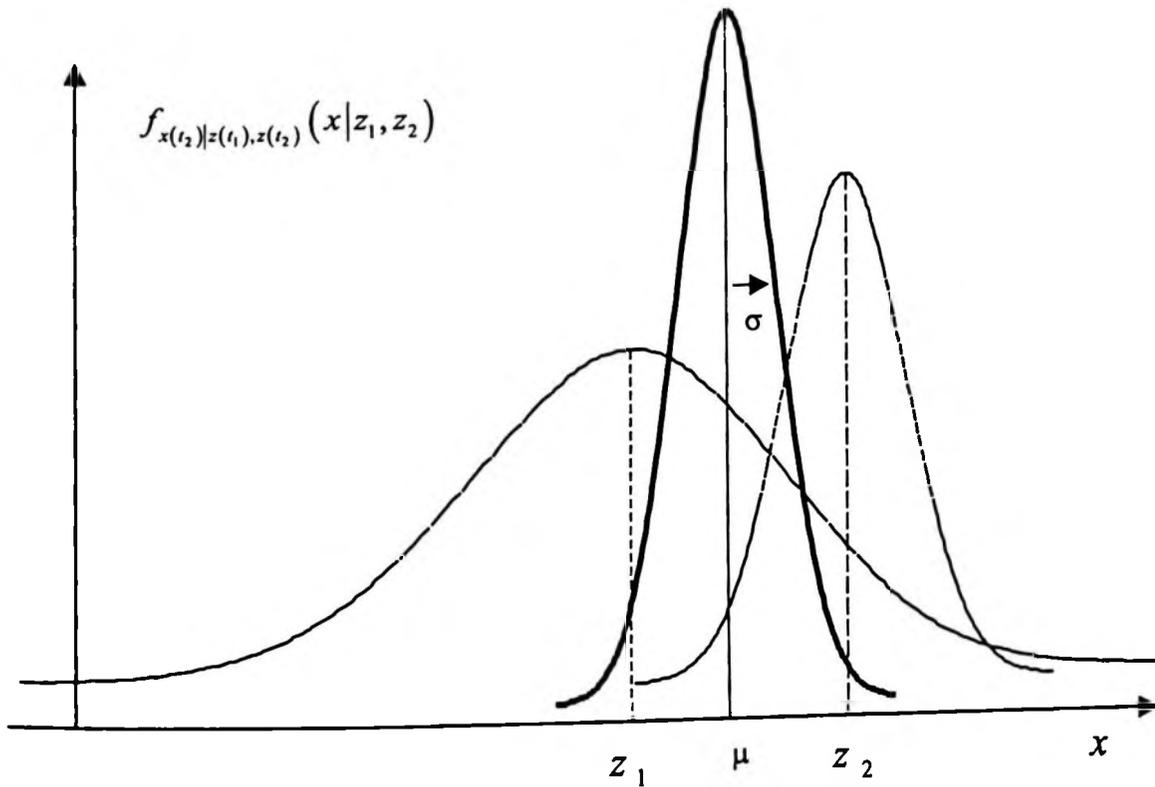


Figura-6.4 Densidade condicional da posição baseada nos dados  $z_1$  e  $z_2$

Depois de algum estudo, a forma de  $\mu$  na equação (6.26) faz bom sentido. Se  $\sigma_{z_1}$  fosse igual a  $\sigma_{z_2}$ , o qual faz você pensar que as medidas são de igual precisão, a equação diz que a estimativa ótima de posição é simplesmente a média das duas medidas, como você poderia esperar. Por outro lado, se  $\sigma_{z_1}$  fosse maior que  $\sigma_{z_2}$ , o qual diz que a incerteza envolvida na medida  $z_1$  é maior que na  $z_2$ , então a equação impõem “ponderadamente”  $z_2$  mais pesada que  $z_1$ . Finalmente, a variância da estimativa é menor que  $\sigma_{z_1}$  até mesmo se  $\sigma_{z_2}$  é muito grande: até mesmo dados com qualidade pobre fornecem alguma informação, e poderiam então incrementar a precisão da saída do filtro.

A equação para  $\hat{x}(t_2)$  pode ser re-escrita como:

$$\hat{x}(t_2) = \left[ \frac{\sigma_{z_2}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} \right] z_1 + \left[ \frac{\sigma_{z_1}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} \right] z_2$$

$$= z_1 + \left[ \frac{\sigma_{z_1}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} \right] \cdot (z_2 - z_1) \quad (6.28)$$

ou, na forma final que é realmente usada nas implementações do filtro de Kalman [notando que  $\hat{x}(t_1) = z_1$  ],

$$\hat{x}(t_2) = \hat{x}(t_1) + K(t_2)[z_2 - \hat{x}(t_1)] \quad (6.29)$$

sendo que:

$$K(t_2) = \frac{\sigma_{z_1}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} \quad (6.30)$$

Estas equações dizem que a estimativa ótima no tempo  $t_2$ ,  $\hat{x}(t_2)$ , é igual a melhor predição do seu valor antes de  $z_2$  ser tomado,  $\hat{x}(t_1)$ , mais o termo de correção do valor de ponderação ótimo vezes a diferença entre  $z_2$  e a melhor predição do seu valor antes dele ser realmente tomado,  $\hat{x}(t_1)$ . Dessa forma, pode-se entender conceitualmente a estrutura “preditor-corretor” do filtro. Baseado nas informações prévias, uma predição do valor que as variáveis desejadas e medidas terão na próxima vez que a medida é feita. Então, quando a próxima medida é tomada, a diferença e o seu valor predito é usado para “corrigir” a predição das variáveis desejadas.

Usando  $K(t_2)$  na equação (6.30), a equação de variância dada pela equação (6.27) pode ser re-escrita como:

$$\sigma_x^2(t_2) = \sigma_x^2(t_1) - K(t_2)\sigma_x^2(t_1) = \sigma_x^2(t_1) \cdot [1 - K(t_2)] \quad (6.31)$$

Note que os valores de  $\hat{x}(t_2)$  e  $\sigma_x^2(t_2)$  incluem todas as informações em:

$$f_{x(t_2)|z(t_1), z(t_2)}(x|z_1, z_2).$$

Afirmando de forma diferente, pela propagação destas duas variáveis, a densidade condicional da variável no tempo  $t_2$ , dado  $z_1$  e  $z_2$ , está completamente especificada.

## 6.3 Valores do IBOVESPA

3/1/2000	16930	16/3/2000	17642	29/5/2000	14772	8/8/2000	16884	19/10/2000	14845
4/1/2000	15851	17/3/2000	17511	30/5/2000	15248	9/8/2000	16533	20/10/2000	14529
5/1/2000	16245	20/3/2000	17525	31/5/2000	14958	10/8/2000	16834	23/10/2000	13989
6/1/2000	16106	21/3/2000	18342	1/6/2000	15450	11/8/2000	17395	24/10/2000	13799
7/1/2000	16309	22/3/2000	18314	2/6/2000	16222	14/8/2000	17810	25/10/2000	13665
10/1/2000	17022	23/3/2000	18331	5/6/2000	16022	15/8/2000	17744	26/10/2000	14223
11/1/2000	16572	24/3/2000	18674	6/6/2000	15946	16/8/2000	17330	27/10/2000	14691
12/1/2000	16616	27/3/2000	18951	7/6/2000	16271	17/8/2000	17702	30/10/2000	14891
13/1/2000	17298	28/3/2000	18337	8/6/2000	16398	18/8/2000	17313	31/10/2000	14867
14/1/2000	17657	29/3/2000	18053	9/6/2000	16342	21/8/2000	17083	1/11/2000	14791
17/1/2000	18053	30/3/2000	17646	12/6/2000	16096	22/8/2000	17224	3/11/2000	14533
18/1/2000	17903	31/3/2000	17820	13/6/2000	16354	23/8/2000	17450	6/11/2000	14800
19/1/2000	17470	3/4/2000	17245	14/6/2000	16366	24/8/2000	17310	7/11/2000	14968
20/1/2000	17176	4/4/2000	16757	15/6/2000	16561	25/8/2000	17642	8/11/2000	14656
21/1/2000	17034	5/4/2000	18714	16/6/2000	16424	28/8/2000	17460	9/11/2000	14524
24/1/2000	16971	6/4/2000	17146	19/6/2000	16850	29/8/2000	17354	10/11/2000	14825
26/1/2000	17105	7/4/2000	17513	20/6/2000	16843	30/8/2000	17414	13/11/2000	14371
27/1/2000	17081	10/4/2000	16783	21/6/2000	17254	31/8/2000	17346	14/11/2000	14540
28/1/2000	18734	11/4/2000	16539	23/6/2000	16977	1/9/2000	17577	16/11/2000	14489
31/1/2000	16388	12/4/2000	16323	26/6/2000	16744	4/9/2000	17613	17/11/2000	14323
1/2/2000	16522	13/4/2000	15500	27/6/2000	16512	5/9/2000	17424	20/11/2000	14499
2/2/2000	16868	14/4/2000	14794	28/6/2000	16842	6/9/2000	17589	21/11/2000	14783
3/2/2000	17457	17/4/2000	14815	29/6/2000	16407	8/9/2000	17433	22/11/2000	14576
4/2/2000	17932	18/4/2000	15408	30/6/2000	16727	11/9/2000	17287	23/11/2000	14314
7/2/2000	18125	19/4/2000	14926	3/7/2000	17088	12/9/2000	16882	24/11/2000	14335
8/2/2000	18685	20/4/2000	15203	4/7/2000	17279	13/9/2000	16996	27/11/2000	14008
9/2/2000	18361	24/4/2000	14794	5/7/2000	17134	14/9/2000	16768	28/11/2000	13913
10/2/2000	18603	25/4/2000	15400	6/7/2000	17348	15/9/2000	16563	29/11/2000	13787
11/2/2000	18083	26/4/2000	15446	7/7/2000	17598	18/9/2000	15913	30/11/2000	13287
14/2/2000	17819	27/4/2000	15440	10/7/2000	17471	19/9/2000	16188	1/12/2000	13437
15/2/2000	17920	28/4/2000	15537	11/7/2000	16880	20/9/2000	16077	4/12/2000	13510
16/2/2000	18112	2/5/2000	15524	12/7/2000	17053	21/9/2000	16146	5/12/2000	14181
17/2/2000	17979	3/5/2000	15108	13/7/2000	16444	22/9/2000	16353	6/12/2000	13945
18/2/2000	17625	4/5/2000	14969	14/7/2000	16880	25/9/2000	16318	7/12/2000	14459
21/2/2000	17699	5/5/2000	15217	17/7/2000	17387	26/9/2000	16254	8/12/2000	14982
22/2/2000	17489	8/5/2000	14890	18/7/2000	17368	27/9/2000	15848	11/12/2000	15187
23/2/2000	17833	9/5/2000	14581	19/7/2000	16927	28/9/2000	16013	12/12/2000	14906
24/2/2000	17775	10/5/2000	14433	20/7/2000	17203	29/9/2000	15928	13/12/2000	15290
25/2/2000	17786	11/5/2000	14498	21/7/2000	17318	2/10/2000	15559	14/12/2000	15259
28/2/2000	17542	12/5/2000	14458	24/7/2000	17385	3/10/2000	15633	15/12/2000	14987
29/2/2000	17660	15/5/2000	15055	25/7/2000	17049	4/10/2000	15876	18/12/2000	15082
1/3/2000	17953	16/5/2000	15360	26/7/2000	17121	5/10/2000	16362	19/12/2000	15336
2/3/2000	18015	17/5/2000	14870	27/7/2000	16948	6/10/2000	15945	20/12/2000	14622
3/3/2000	18631	18/5/2000	14626	28/7/2000	16485	9/10/2000	15855	21/12/2000	14505
8/3/2000	18282	19/5/2000	14326	31/7/2000	16454	10/10/2000	15747	22/12/2000	14652
9/3/2000	18650	22/5/2000	13891	1/8/2000	16289	11/10/2000	15528	26/12/2000	14794
10/3/2000	18280	23/5/2000	13587	2/8/2000	16314	13/10/2000	15385	27/12/2000	15186
13/3/2000	17594	24/5/2000	14166	3/8/2000	16720	16/10/2000	15241	28/12/2000	15259
14/3/2000	17221	25/5/2000	14185	4/8/2000	16781	17/10/2000	14870	2/1/2001	15425
15/3/2000	17381	26/5/2000	14542	7/8/2000	16995	18/10/2000	14420	3/1/2001	16599
4/1/2001	16675	20/3/2001	14903	31/5/2001	14648	13/8/2001	13771	24/10/2001	11487
5/1/2001	16409	21/3/2001	14852	1/6/2001	14781	14/8/2001	13829	25/10/2001	11723
8/1/2001	16562	22/3/2001	14067	4/6/2001	15103	15/8/2001	13658	26/10/2001	11780
9/1/2001	16975	23/3/2001	14435	5/6/2001	15394	16/8/2001	13509	29/10/2001	11377
10/1/2001	16918	26/3/2001	14712	6/8/2001	15328	17/8/2001	13044	30/10/2001	11023
11/1/2001	17023	27/3/2001	14835	7/6/2001	15464	20/8/2001	13115	31/10/2001	11364
12/1/2001	16850	28/3/2001	14631	8/6/2001	15368	21/8/2001	12891	1/11/2001	11387
15/1/2001	16962	29/3/2001	14282	11/6/2001	15224	22/8/2001	12952	5/11/2001	12164
16/1/2001	16720	30/3/2001	14438	12/6/2001	15173	23/8/2001	12750	8/11/2001	12415
17/1/2001	17191	2/4/2001	13981	13/8/2001	15403	24/8/2001	13001	7/11/2001	12615
18/1/2001	17521	3/4/2001	13736	15/6/2001	14985	27/8/2001	12985	8/11/2001	12553
19/1/2001	17530	4/4/2001	13854	18/8/2001	14355	28/8/2001	13018	9/11/2001	12730
22/1/2001	17391	5/4/2001	14442	19/8/2001	14400	29/8/2001	13078	12/11/2001	12587
23/1/2001	17832	6/4/2001	14484	20/8/2001	14571	30/8/2001	12892	13/11/2001	12918

24/1/2001	17771	9/4/2001	14718	21/8/2001	14822	31/8/2001	12840	14/11/2001	12825
28/1/2001	17889	10/4/2001	15047	22/8/2001	14682	3/9/2001	12800	18/11/2001	12882
29/1/2001	17883	11/4/2001	14818	25/8/2001	14539	4/9/2001	12766	19/11/2001	12987
30/1/2001	17722	12/4/2001	14969	28/8/2001	14459	5/9/2001	12590	20/11/2001	12838
31/1/2001	17872	16/4/2001	14444	27/8/2001	14308	6/9/2001	12255	21/11/2001	12793
1/2/2001	17038	17/4/2001	14338	28/8/2001	14390	10/9/2001	11922	22/11/2001	13019
2/2/2001	16914	18/4/2001	14955	29/8/2001	14559	11/9/2001	10827	23/11/2001	13423
5/2/2001	16731	19/4/2001	14431	2/7/2001	14555	12/9/2001	11113	26/11/2001	13759
6/2/2001	17008	20/4/2001	13696	3/7/2001	14352	13/9/2001	10306	27/11/2001	13602
7/2/2001	16812	23/4/2001	13892	4/7/2001	14058	14/9/2001	10034	28/11/2001	13019
8/2/2001	17243	24/4/2001	14068	5/7/2001	14047	17/9/2001	10544	29/11/2001	12738
9/2/2001	17138	25/4/2001	14162	6/7/2001	13906	18/9/2001	10554	30/11/2001	12931
12/2/2001	16917	26/4/2001	14757	10/7/2001	13569	19/9/2001	10744	3/12/2001	13336
13/2/2001	17095	27/4/2001	14927	11/7/2001	13811	20/9/2001	10543	4/12/2001	13146
14/2/2001	17120	30/4/2001	14917	12/7/2001	13912	21/9/2001	10418	5/12/2001	13382
15/2/2001	16937	2/5/2001	14897	13/7/2001	14078	24/9/2001	10532	8/12/2001	13539
16/2/2001	16259	3/5/2001	15141	16/7/2001	13811	25/9/2001	10229	7/12/2001	13298
19/2/2001	16080	4/5/2001	15092	17/7/2001	14188	26/9/2001	10005	10/12/2001	13448
20/2/2001	15910	7/5/2001	14875	18/7/2001	13790	27/9/2001	10404	11/12/2001	13348
21/2/2001	15593	8/5/2001	14747	19/7/2001	13781	28/9/2001	10635	12/12/2001	13562
22/2/2001	15910	9/5/2001	14778	20/7/2001	14092	1/10/2001	10501	13/12/2001	13148
23/2/2001	16157	10/5/2001	14993	23/7/2001	14067	2/10/2001	10350	14/12/2001	12959
28/2/2001	15891	11/5/2001	14493	24/7/2001	13737	3/10/2001	10233	17/12/2001	12912
1/3/2001	16416	14/5/2001	14132	25/7/2001	13955	4/10/2001	10061	18/12/2001	13397
2/3/2001	16581	15/5/2001	14226	26/7/2001	13810	5/10/2001	10197	19/12/2001	13290
5/3/2001	16537	16/5/2001	14714	27/7/2001	13910	8/10/2001	10094	20/12/2001	12918
6/3/2001	16324	17/5/2001	14793	30/7/2001	13703	9/10/2001	10284	21/12/2001	13368
7/3/2001	16395	18/5/2001	14884	31/7/2001	13754	10/10/2001	10462	26/12/2001	13358
8/3/2001	16226	21/5/2001	15127	1/8/2001	13743	11/10/2001	10784	27/12/2001	13757
9/3/2001	16123	22/5/2001	14828	2/8/2001	13822	15/10/2001	11329	28/12/2001	13577
12/3/2001	15527	23/5/2001	14691	3/8/2001	13838	16/10/2001	11257	2/1/2002	13872
13/3/2001	15584	24/5/2001	14523	6/8/2001	14046	17/10/2001	11271	3/1/2002	14265
14/3/2001	15244	25/5/2001	14310	7/8/2001	14035	18/10/2001	10994	4/1/2002	14331
15/3/2001	15060	28/5/2001	14302	8/8/2001	13919	19/10/2001	11331	7/1/2002	14378
18/3/2001	15237	29/5/2001	14453	9/8/2001	13834	22/10/2001	11699	8/1/2002	14187
19/3/2001	14835	30/5/2001	14489	10/8/2001	13915	23/10/2001	11613	9/1/2002	14020
10/1/2002	13570	26/3/2002	13589	7/8/2002	12282	19/8/2002	9418	28/10/2002	9573
11/1/2002	13587	27/3/2002	13424	10/6/2002	12599	20/8/2002	9263	29/10/2002	9600
14/1/2002	13121	28/3/2002	13254	11/8/2002	12210	21/8/2002	9437	30/10/2002	10088
15/1/2002	13010	1/4/2002	13467	12/8/2002	12132	22/8/2002	9702	31/10/2002	10187
18/1/2002	13083	2/4/2002	13245	13/8/2002	11962	23/8/2002	9678	1/11/2002	10140
17/1/2002	13336	3/4/2002	13084	14/8/2002	11698	26/8/2002	10097	4/11/2002	9912
18/1/2002	13372	4/4/2002	13360	17/8/2002	11937	27/8/2002	10371	5/11/2002	9861
21/1/2002	13154	5/4/2002	13276	18/8/2002	11821	28/8/2002	10379	6/11/2002	9702
22/1/2002	13003	8/4/2002	13159	19/8/2002	11493	29/8/2002	10455	7/11/2002	9799
23/1/2002	13232	9/4/2002	13195	20/8/2002	10908	30/8/2002	10382	8/11/2002	9860
24/1/2002	13162	10/4/2002	13416	21/8/2002	10397	2/9/2002	10378	11/11/2002	9885
28/1/2002	13001	11/4/2002	13411	24/8/2002	10759	3/9/2002	10135	12/11/2002	9720
29/1/2002	12501	12/4/2002	13750	25/8/2002	10705	4/9/2002	9998	13/11/2002	9763
30/1/2002	12532	15/4/2002	13471	26/8/2002	10690	5/9/2002	9723	14/11/2002	9884
31/1/2002	12721	16/4/2002	13822	27/8/2002	11013	6/9/2002	9716	18/11/2002	9970
1/2/2002	12658	17/4/2002	13732	28/8/2002	11139	9/9/2002	9954	19/11/2002	9971
4/2/2002	12512	18/4/2002	13573	1/7/2002	10892	10/9/2002	9960	20/11/2002	10087
5/2/2002	12747	19/4/2002	13478	2/7/2002	10845	11/9/2002	10182	21/11/2002	10289
6/2/2002	12750	22/4/2002	13224	3/7/2002	10635	12/9/2002	10172	22/11/2002	10404
7/2/2002	12683	23/4/2002	13188	4/7/2002	10655	13/9/2002	10180	25/11/2002	10245
8/2/2002	12597	24/4/2002	13380	5/7/2002	10523	16/9/2002	9831	26/11/2002	10131
13/2/2002	12980	25/4/2002	13272	8/7/2002	10687	17/9/2002	9650	27/11/2002	10226
14/2/2002	13245	26/4/2002	13075	10/7/2002	10555	18/9/2002	9505	28/11/2002	10239
15/2/2002	13229	29/4/2002	13090	11/7/2002	10806	19/9/2002	9372	29/11/2002	10508
18/2/2002	13121	30/4/2002	13085	12/7/2002	10987	20/9/2002	9585	2/12/2002	10873
19/2/2002	12991	2/5/2002	12538	15/7/2002	10833	23/9/2002	9284	3/12/2002	10862
20/2/2002	13302	3/5/2002	12609	16/7/2002	10577	24/9/2002	9148	4/12/2002	10840
21/2/2002	13530	6/5/2002	12430	17/7/2002	10754	25/9/2002	9227	5/12/2002	10413
22/2/2002	13582	7/5/2002	12342	18/7/2002	10812	26/9/2002	9199	6/12/2002	10569
25/2/2002	13977	8/5/2002	12618	19/7/2002	10583	27/9/2002	8715	9/12/2002	10338
26/2/2002	13983	9/5/2002	12101	22/7/2002	9892	30/9/2002	8622	10/12/2002	10331
27/2/2002	14212	10/5/2002	12130	23/7/2002	9745	1/10/2002	8997	11/12/2002	10813
28/2/2002	14033	13/5/2002	12002	24/7/2002	9937	2/10/2002	8820	12/12/2002	10588
1/3/2002	14414	14/5/2002	12204	25/7/2002	9685	3/10/2002	9139	13/12/2002	10584
4/3/2002	14471	15/5/2002	12349	26/7/2002	9217	4/10/2002	9259	16/12/2002	10771

5/3/2002	14011	16/5/2002	12660	29/7/2002	9240	7/10/2002	8863	17/12/2002	10832
6/3/2002	13838	17/5/2002	12699	30/7/2002	9341	8/10/2002	8848	18/12/2002	10984
7/3/2002	13725	20/5/2002	12667	31/7/2002	9782	9/10/2002	8714	19/12/2002	11207
8/3/2002	13961	21/5/2002	12701	1/8/2002	9759	10/10/2002	8866	20/12/2002	11489
11/3/2002	13695	22/5/2002	12367	2/8/2002	9852	11/10/2002	8854	23/12/2002	11470
12/3/2002	14181	23/5/2002	12555	5/8/2002	9469	14/10/2002	8450	28/12/2002	11318
13/3/2002	14244	24/5/2002	12573	6/8/2002	9755	15/10/2002	8506	27/12/2002	11234
14/3/2002	14117	27/5/2002	12698	7/8/2002	9869	16/10/2002	8370	30/12/2002	11288
15/3/2002	14365	28/5/2002	12728	8/8/2002	10315	17/10/2002	8901	2/1/2003	11602
18/3/2002	14244	29/5/2002	12985	9/8/2002	9985	18/10/2002	9022	3/1/2003	11600
19/3/2002	14117	31/5/2002	12861	12/8/2002	9723	21/10/2002	9128	6/1/2003	12019
20/3/2002	14089	3/6/2002	12659	13/8/2002	9444	22/10/2002	9331	7/1/2003	11878
21/3/2002	13717	4/6/2002	12600	14/8/2002	9343	23/10/2002	9840	8/1/2003	11785
22/3/2002	13312	5/6/2002	12589	15/8/2002	9183	24/10/2002	9799	9/1/2003	11911
25/3/2002	13277	6/6/2002	12112	16/8/2002	9526	25/10/2002	10014	10/1/2003	12242
13/1/2003	12110	26/3/2003	11206	9/6/2003	13845	20/8/2003	14468	29/10/2003	17944
14/1/2003	12174	27/3/2003	11233	10/6/2003	13893	21/8/2003	14669	30/10/2003	18093
15/1/2003	11971	28/3/2003	11396	11/6/2003	13876	22/8/2003	14613	31/10/2003	17982
16/1/2003	11951	31/3/2003	11273	12/6/2003	13982	25/8/2003	14473	3/11/2003	18517
17/1/2003	11675	1/4/2003	11592	13/6/2003	13733	26/8/2003	14877	4/11/2003	18541
20/1/2003	11648	2/4/2003	11872	16/6/2003	13832	27/8/2003	15142	5/11/2003	18307
21/1/2003	11434	3/4/2003	12006	17/6/2003	13776	28/8/2003	15064	6/11/2003	18612
22/1/2003	11142	4/4/2003	12065	18/6/2003	13510	29/8/2003	15174	7/11/2003	18672
23/1/2003	11162	7/4/2003	12135	20/6/2003	13130	1/9/2003	15352	10/11/2003	18572
24/1/2003	10783	8/4/2003	11778	23/6/2003	12991	2/9/2003	15454	11/11/2003	18408
27/1/2003	10529	9/4/2003	11758	24/6/2003	13108	3/9/2003	15633	12/11/2003	18793
28/1/2003	10516	10/4/2003	11591	25/6/2003	13025	4/9/2003	15704	13/11/2003	18754
29/1/2003	10863	11/4/2003	11718	26/6/2003	13111	5/9/2003	15899	14/11/2003	18985
30/1/2003	10750	14/4/2003	11873	27/6/2003	13024	8/9/2003	16050	17/11/2003	18668
31/1/2003	10941	15/4/2003	12106	30/6/2003	12972	9/9/2003	15718	18/11/2003	18812
3/2/2003	10910	16/4/2003	12043	1/7/2003	13291	10/9/2003	15983		
4/2/2003	10577	17/4/2003	12395	2/7/2003	13310	11/9/2003	16292		
5/2/2003	10591	22/4/2003	12452	3/7/2003	13134	12/9/2003	16421		
6/2/2003	10565	23/4/2003	12394	4/7/2003	13273	15/9/2003	16341		
7/2/2003	10380	24/4/2003	12120	7/7/2003	13402	16/9/2003	16269		
10/2/2003	10479	25/4/2003	12126	8/7/2003	13618	17/9/2003	16492		
11/2/2003	10509	28/4/2003	12462	10/7/2003	13501	18/9/2003	16889		
12/2/2003	10509	29/4/2003	12677	11/7/2003	13320	19/9/2003	16851		
13/2/2003	10107	30/4/2003	12558	14/7/2003	13588	22/9/2003	16485		
14/2/2003	10080	2/5/2003	12810	15/7/2003	13614	23/9/2003	16443		
17/2/2003	10188	5/5/2003	12832	16/7/2003	13487	24/9/2003	16057		
18/2/2003	10449	6/5/2003	12643	17/7/2003	13622	25/9/2003	15806		
19/2/2003	10251	7/5/2003	12956	18/7/2003	13794	26/9/2003	15810		
20/2/2003	10216	8/5/2003	12920	21/7/2003	13675	29/9/2003	16108		
21/2/2003	10330	9/5/2003	13214	22/7/2003	13835	30/9/2003	16010		
24/2/2003	10253	12/5/2003	13320	23/7/2003	13798	1/10/2003	16578		
25/2/2003	10191	13/5/2003	13420	24/7/2003	13761	2/10/2003	16893		
26/2/2003	9994	14/5/2003	13459	25/7/2003	13750	3/10/2003	17089		
27/2/2003	10150	15/5/2003	13129	28/7/2003	13643	6/10/2003	17273		
28/2/2003	10280	16/5/2003	13224	29/7/2003	13623	7/10/2003	17470		
5/3/2003	10305	19/5/2003	12745	30/7/2003	13474	8/10/2003	17804		
6/3/2003	10615	20/5/2003	12745	31/7/2003	13571	9/10/2003	17708		
7/3/2003	10723	21/5/2003	13033	1/8/2003	13129	10/10/2003	17676		
10/3/2003	10308	22/5/2003	13100	4/8/2003	12938	13/10/2003	18061		
11/3/2003	10339	23/5/2003	13142	5/8/2003	13059	14/10/2003	18178		
12/3/2003	10582	26/5/2003	12852	6/8/2003	12887	15/10/2003	17942		
13/3/2003	10783	27/5/2003	13246	7/8/2003	13328	16/10/2003	17955		
14/3/2003	10817	28/5/2003	13294	8/8/2003	13499	17/10/2003	17790		
17/3/2003	10874	29/5/2003	13405	11/8/2003	13560	20/10/2003	18369		
18/3/2003	11150	30/5/2003	13421	12/8/2003	13602	21/10/2003	18446		
19/3/2003	11006	2/6/2003	13228	13/8/2003	13682	22/10/2003	18234		
20/3/2003	11158	3/6/2003	13350	14/8/2003	13812	23/10/2003	17689		
21/3/2003	11377	4/6/2003	13718	15/8/2003	13889	24/10/2003	17813		
24/3/2003	11052	5/6/2003	13779	18/8/2003	14146	27/10/2003	17750		
25/3/2003	11232	6/6/2003	13923	19/8/2003	14157	28/10/2003	18227		

## 6.4 Série da Petrobrás

DATA	PETROBRAS						
2/jan/95	8780,2951	30/dez/97	21345,90102	4/jul/96	11212,89818	2/jul/99	23662,4113
3/jan/95	8278,563975	31/dez/97	21848,15752	5/jul/96	10961,04516	5/jul/99	23452,17926
4/jan/95	7944,076542	2/jan/98	22475,97813	9/jul/96	11085,71659	7/jul/99	23870,96818
5/jan/95	8027,6984	5/jan/98	22810,8158	10/jul/96	10793,70096	8/jul/99	23284,6637
6/jan/95	7400,534463	6/jan/98	22182,99518	11/jul/96	10541,84794	9/jul/99	23284,6637
9/jan/95	7191,479817	7/jan/98	21513,31985	12/jul/96	10417,17651	12/jul/99	23033,39035
10/jan/95	6438,883092	8/jan/98	20927,35394	15/jul/96	9956,979956	13/jul/99	22865,87478
11/jan/95	6606,126808	9/jan/98	19420,58446	16/jul/96	9705,963655	14/jul/99	23201,74349
12/jan/95	7442,345392	12/jan/98	19336,87504	17/jul/96	9789,635755	15/jul/99	23033,39035
13/jan/95	8278,563975	13/jan/98	20090,25979	18/jul/96	9705,963655	16/jul/99	23368,42148
16/jan/95	8069,509329	14/jan/98	20173,9692	19/jul/96	9873,307856	19/jul/99	23368,42148
17/jan/95	8278,563975	15/jan/98	19588,00329	22/jul/96	9763,697404	20/jul/99	23368,42148
18/jan/95	8027,6984	16/jan/98	20299,53332	23/jul/96	9622,291554	21/jul/99	23242,7848
19/jan/95	7400,534463	19/jan/98	21011,06336	24/jul/96	9413,111303	22/jul/99	23284,6637
20/jan/95	7609,589108	20/jan/98	20173,9692	25/jul/96	9538,619454	23/jul/99	22279,5703
23/jan/95	7191,479817	21/jan/98	19922,84095	26/jul/96	9664,127605	26/jul/99	21609,50803
24/jan/95	7358,723533	22/jan/98	20090,25979	29/jul/96	9413,111303	27/jul/99	21860,78138
25/jan/95	7358,723533	23/jan/98	19922,84095	30/jul/96	9580,455504	28/jul/99	21900,98512
26/jan/95	7902,265613	26/jan/98	20090,25979	31/jul/96	9496,783404	29/jul/99	21359,07226
27/jan/95	7609,589108	27/jan/98	20508,80686	1/ago/96	9623,128275	30/jul/99	21534,96361
30/jan/95	7066,047029	28/jan/98	20759,93511	2/ago/96	9998,816007	2/ago/99	21107,79891
31/jan/95	7693,210967	29/jan/98	20508,80686	5/ago/96	10082,48811	3/ago/99	21776,18602
1/fev/95	7944,076542	30/jan/98	20090,25979	6/ago/96	10249,83231	4/ago/99	21777,0236
2/fev/95	7818,643754	2/fev/98	19922,84095	7/ago/96	10333,50441	5/ago/99	21567,62914
3/fev/95	8111,320258	3/fev/98	20132,11449	8/ago/96	10125,16088	6/ago/99	20478,77796
6/fev/95	7902,265613	4/fev/98	19922,84095	9/ago/96	10207,15954	9/ago/99	19892,47348
7/fev/95	7525,96725	5/fev/98	20090,25979	12/ago/96	10417,17651	10/ago/99	19532,31501
8/fev/95	7484,156321	6/fev/98	20383,24274	13/ago/96	10333,50441	11/ago/99	20361,51707
9/fev/95	7107,857958	9/fev/98	20676,2257	14/ago/96	10333,50441	12/ago/99	19850,59459
10/fev/95	7024,2361	10/fev/98	20592,51628	15/ago/96	10459,01256	13/ago/99	20277,75928
13/fev/95	7358,723533	11/fev/98	21261,35451	16/ago/96	10584,52071	16/ago/99	19976,23126
14/fev/95	7149,668888	12/fev/98	21094,77277	19/ago/96	10499,17517	17/ago/99	19808,7157
15/fev/95	6856,992383	13/fev/98	21094,77277	20/ago/96	10500,84861	18/ago/99	19683,9166
16/fev/95	6564,315879	16/fev/98	20927,35394	21/ago/96	10500,84861	19/ago/99	19976,23126
17/fev/95	6940,614242	17/fev/98	21262,19161	22/ago/96	10291,66836	20/ago/99	21190,71912
20/fev/95	6773,370525	18/fev/98	20927,35394	23/ago/96	9956,979956	23/ago/99	21190,71912
21/fev/95	6273,311812	19/fev/98	21262,19161	26/ago/96	9538,619454	24/ago/99	21190,71912
22/fev/95	6522,50495	20/fev/98	21763,61101	27/ago/96	9873,307856	25/ago/99	20981,32466
23/fev/95	7358,723533	25/fev/98	21680,73869	28/ago/96	10040,65206	26/ago/99	20771,9302
24/fev/95	7234,126965	26/fev/98	21513,31985	29/ago/96	10166,16021	27/ago/99	21190,71912
1/mar/95	7234,126965	27/fev/98	21429,61044	30/ago/96	10166,16021	30/ago/99	21106,96134
2/mar/95	6606,126808	2/mar/98	21638,88398	2/set/96	9999,652728	31/ago/99	21525,75025
3/mar/95	6564,315879	3/mar/98	21571,91644	3/set/96	10040,65206	1/set/99	22195,81252
6/mar/95	6271,639375	4/mar/98	21429,61044	4/set/96	10124,32416	2/set/99	22781,27942
7/mar/95	5309,988004	5/mar/98	21513,31985	5/set/96	10039,81534	3/set/99	23283,82612
8/mar/95	4682,824067	6/mar/98	21555,17456	6/set/96	10082,48811	6/set/99	23075,26924
9/mar/95	4264,714775	9/mar/98	21680,73869	9/set/96	10290,83164	8/set/99	23159,02702
10/mar/95	5268,177075	10/mar/98	22350,41401	10/set/96	10374,50374	9/set/99	23451,34168
13/mar/95	5059,122429	11/mar/98	22434,12343	11/set/96	10000,48945	10/set/99	23452,17926
14/mar/95	5811,719154	12/mar/98	22768,96109	12/set/96	10081,65139	13/set/99	23787,21039
15/mar/95	5603,500727	13/mar/98	22936,37992	13/set/96	10165,32349	14/set/99	23745,3315
16/mar/95	6104,395658	16/mar/98	23020,08934	16/set/96	10291,66836	15/set/99	23619,69483
17/mar/95	5978,962871	17/mar/98	23187,50817	17/set/96	10207,99626	16/set/99	23569,44016
20/mar/95	5477,231721	18/mar/98	23061,94405	18/set/96	10083,32483	17/set/99	24792,30379
21/mar/95	5686,286367	19/mar/98	23061,94405	19/set/96	10166,16021	20/set/99	25336,72938
22/mar/95	5561,689798	20/mar/98	23355,88322	20/set/96	10082,48811	21/set/99	25294,01291
23/mar/95	5935,479505	23/mar/98	23272,17037	23/set/96	9957,816677	22/set/99	25797,39719
24/mar/95	6188,017517	24/mar/98	23026,17132	24/set/96	10082,48811	23/set/99	25546,12384
27/mar/95	6071,875428	25/mar/98	22523,78213	25/set/96	9956,979956	24/set/99	25378,60827
28/mar/95	5896,242668	26/mar/98	22774,13941	26/set/96	9831,471806	27/set/99	25378,60827
29/mar/95	5728,973372	27/mar/98	22942,43979	27/set/96	9789,635755	28/set/99	25244,59582
30/mar/95	5478,069429	30/mar/98	22272,58753	30/set/96	10039,81534	29/set/99	25085,45603
31/mar/95	5227,165485	31/mar/98	22691,24519	1/out/96	10123,48744	30/set/99	24038,48374
3/abr/95	5018,078866	1/abr/98	22607,51366	2/out/96	10290,83164	1/out/99	23862,5924
4/abr/95	5728,973372	2/abr/98	23193,63438	4/out/96	10374,50374	4/out/99	24457,27266
5/abr/95	5728,973372	3/abr/98	22691,24519	7/out/96	10375,34046	5/out/99	24122,24153
6/abr/95	5812,60802	6/abr/98	23277,36591	8/out/96	10249,83231	6/out/99	25144,92406
7/abr/95	5603,5214	7/abr/98	23444,82898	9/out/96	10207,99626	7/out/99	25294,05049
10/abr/95	5352,617457	8/abr/98	23947,21817	10/out/96	9999,652728	8/out/99	25336,72938
11/abr/95	5101,713514	13/abr/98	24197,57545	11/out/96	10123,48744	11/out/99	24415,39377
12/abr/95	5185,348162	14/abr/98	24699,96464	14/out/96	10166,16021	13/out/99	23870,96818
17/abr/95	5436,252105	15/abr/98	24700,80196	15/out/96	10333,50441	14/out/99	24038,48374
18/abr/95	5561,704077	16/abr/98	24365,87583	16/out/96	10710,02886	15/out/99	24038,48374
19/abr/95	5678,792583	17/abr/98	24282,1443	17/out/96	11210,38802	18/out/99	23284,6637

20/abr/95	5687,156048	20/abr/98	24030,9497	18/out/96	11039,69694	19/out/99	23745,3315
24/abr/95	5979,877315	22/abr/98	23612,29204	21/out/96	11379,40566	20/out/99	24289,75709
25/abr/95	6357,905923	23/abr/98	23695,18626	22/out/96	11002,88121	21/out/99	24541,03044
26/abr/95	6690,771821	24/abr/98	24198,41277	23/out/96	10877,37306	22/out/99	25713,6394
27/abr/95	7025,310412	27/abr/98	23402,1259	24/out/96	10877,37306	25/out/99	26048,67054
28/abr/95	7108,94506	28/abr/98	24114,68123	25/out/96	10709,19214	26/out/99	25964,91275
2/mai/95	7276,214356	29/abr/98	24030,9497	28/out/96	10626,35676	27/out/99	26291,56811
3/mai/95	7527,118299	30/abr/98	24282,1443	29/out/96	10874,8629	28/out/99	26308,31966
4/mai/95	7903,474214	4/mai/98	23444,82898	10/out/96	10793,70096	29/out/99	26006,79165
5/mai/95	7778,022242	5/mai/98	23444,82898	31/out/96	11128,38936	1/nov/99	26383,70167
8/mai/95	7527,118299	6/mai/98	22858,70825	1/nov/96	11253,06079	3/nov/99	28058,85733
9/mai/95	7861,65689	7/mai/98	22105,12446	4/nov/96	11170,22541	4/nov/99	28896,43516
10/mai/95	7736,204918	8/mai/98	22021,39293	5/nov/96	11002,88121	5/nov/99	29901,52856
11/mai/95	7401,666327	11/mai/98	21519,00374	6/nov/96	10919,20911	8/nov/99	30655,34861
12/mai/95	7568,935623	12/mai/98	21100,34608	7/nov/96	10793,70096	9/nov/99	31074,13752
15/mai/95	7192,579708	13/mai/98	20974,74878	8/nov/96	10500,84861	10/nov/99	31157,89531
16/mai/95	6983,493089	14/mai/98	20932,88302	11/nov/96	10793,70096	11/nov/99	30965,2524
17/mai/95	7359,849004	15/mai/98	21109,67491	12/nov/96	10793,70096	12/nov/99	30780,98528
18/mai/95	7652,570271	18/mai/98	19928,10463	13/nov/96	10876,53634	16/nov/99	30864,74306
19/mai/95	7568,935623	19/mai/98	20137,43346	14/nov/96	10877,37306	17/nov/99	30320,31747
22/mai/95	7861,65689	20/mai/98	19844,3731	18/nov/96	11337,56961	18/nov/99	32246,74649
23/mai/95	7861,65689	21/mai/98	19676,91004	19/nov/96	11212,06146	19/nov/99	33586,87101
24/mai/95	7945,291538	22/mai/98	19509,44697	20/nov/96	11295,73356	22/nov/99	33586,87101
25/mai/95	7819,839566	25/mai/98	18839,59471	21/nov/96	11630,42197	23/nov/99	33586,87101
26/mai/95	7819,839566	26/mai/98	17583,62173	22/nov/96	11880,60155	24/nov/99	33084,32432
29/mai/95	7694,387595	27/mai/98	18002,27939	25/nov/96	11881,43827	25/nov/99	32456,14094
30/mai/95	7234,397032	28/mai/98	19258,25237	26/nov/96	11755,93012	26/nov/99	32669,72329
31/mai/95	7108,94506	29/mai/98	18588,40012	27/nov/96	11672,25802	29/nov/99	32330,50427
1/jun/95	7443,483651	1/jun/98	18839,59471	28/nov/96	11714,09407	30/nov/99	32539,89873
2/jun/95	7903,474214	2/jun/98	20011,83616	29/nov/96	11881,43827	1/dec/99	32657,15962
5/jun/95	8238,012805	3/jun/98	19844,3731	2/dec/96	12006,94642	2/dec/99	33168,0821
6/jun/95	7861,65689	4/jun/98	20095,5677	3/dec/96	12341,63482	3/dec/99	33377,47656
7/jun/95	7568,935623	5/jun/98	20345,92498	4/dec/96	12341,63482	6/dec/99	33335,59767
8/jun/95	7569,771969	8/jun/98	20514,22536	5/dec/96	12843,66742	7/dec/99	33503,11323
9/jun/95	7276,214356	9/jun/98	20011,83616	6/dec/96	12926,5028	8/dec/99	33419,35545
12/jun/95	7108,94506	10/jun/98	19174,52084	9/dec/96	12843,66742	9/dec/99	33335,59767
13/jun/95	7150,762384	12/jun/98	18253,47399	10/dec/96	12759,99532	10/dec/99	33754,38658
14/jun/95	6899,858441	15/jun/98	17333,26445	11/dec/96	12550,81507	13/dec/99	33754,38658
16/jun/95	6933,3123	16/jun/98	17416,15867	12/dec/96	12759,99532	14/dec/99	33419,35545
19/jun/95	6694,117207	17/jun/98	18337,20552	13/dec/96	12676,32322	15/dec/99	32414,26205
20/jun/95	6565,31985	18/jun/98	17499,8902	16/dec/96	13011,01162	16/dec/99	32833,05097
21/jun/95	6816,223793	19/jun/98	18086,01093	17/dec/96	13011,01162	17/dec/99	33503,11323
22/jun/95	6648,954498	22/jun/98	17667,35327	18/dec/96	13221,0288	20/dec/99	34178,20096
23/jun/95	6607,137174	23/jun/98	18169,74246	19/dec/96	13220,19187	21/dec/99	35555,17892
26/jun/95	6481,685202	24/jun/98	18086,01093	20/dec/96	13261,1912	22/dec/99	35848,33116
27/jun/95	6690,771821	25/jun/98	17918,54786	23/dec/96	13303,86398	23/dec/99	35552,66618
28/jun/95	6607,137174	26/jun/98	17583,62173	24/dec/96	13303,86398	24/dec/99	35552,66618
29/jun/95	6690,771821	29/jun/98	17943,66732	26/dec/96	13680,38843	27/dec/99	35345,78446
30/jun/95	6523,502526	30/jun/98	18002,27939	27/dec/96	13764,06053	28/dec/99	35932,08894
3/jul/95	6523,502526	1/jul/98	18337,20552	30/dec/96	13847,73263	29/dec/99	37272,21347
4/jul/95	6773,570123	2/jul/98	18420,93705	31/dec/96	13847,73263	30/dec/99	38528,58022
5/jul/95	6983,493089	3/jul/98	19090,78931	2/jan/97	13722,22448	31/dec/99	38528,58022
6/jul/95	7292,941285	6/jul/98	19760,64157	3/jan/97	14015,07683	3/jan/00	39363,64531
7/jul/95	7527,118299	7/jul/98	20514,22536	6/jan/97	14643,4543	4/jan/00	37020,94012
10/jul/95	7627,479876	8/jul/98	20932,88302	7/jan/97	15060,97809	5/jan/00	36811,54566
11/jul/95	7527,118299	9/jul/98	20932,88302	8/jan/97	15144,65019	6/jan/00	36685,90899
12/jul/95	7861,65689	10/jul/98	21016,61455	9/jan/97	14891,96044	7/jan/00	36853,42455
13/jul/95	7694,387595	13/jul/98	21016,61455	10/jan/97	14935,46993	10/jan/00	37608,08218
14/jul/95	7860,820544	14/jul/98	21686,4668	13/jan/97	14809,96178	11/jan/00	36687,58414
17/jul/95	7736,204918	15/jul/98	21770,19834	14/jan/97	15688,51884	12/jan/00	36015,84672
18/jul/95	7401,666327	16/jul/98	22356,31906	15/jan/97	15730,35489	13/jan/00	35680,81559
19/jul/95	7150,762384	17/jul/98	22523,78213	16/jan/97	15604,84674	14/jan/00	36183,36229
20/jul/95	7359,849004	20/jul/98	23109,06553	17/jan/97	16232,38749	17/jan/00	35764,57337
21/jul/95	7231,887992	21/jul/98	23277,36591	20/jan/97	16316,05959	18/jan/00	35932,08894
24/jul/95	7401,666327	22/jul/98	23277,36591	21/jan/97	16190,55144	19/jan/00	35680,81559
25/jul/95	7359,849004	23/jul/98	22272,58753	22/jan/97	15730,35489	20/jan/00	35262,02668
26/jul/95	7234,397032	24/jul/98	21937,6614	23/jan/97	16023,20724	21/jan/00	34843,23776
27/jul/95	7276,214356	27/jul/98	21770,19834	24/jan/97	15730,35489	24/jan/00	34591,96441
28/jul/95	7277,050702	28/jul/98	21853,92987	27/jan/97	15730,35489	25/jan/00	34591,96441
31/jul/95	7276,214356	29/jul/98	21686,4668	28/jan/97	15897,69909	26/jan/00	35010,75333
1/ago/95	7108,94506	30/jul/98	21938,49872	29/jan/97	15814,02699	27/jan/00	34926,99554
2/ago/95	7192,579708	31/jul/98	21351,54068	30/jan/97	16065,04329	28/jan/00	34424,44885
3/ago/95	7443,483651	3/ago/98	20523,43582	31/jan/97	16818,0922	31/jan/00	34173,1755
4/ago/95	7778,022242	4/ago/98	19844,3731	3/fev/97	17487,469	1/fev/00	34843,23776
7/ago/95	8112,560833	5/ago/98	19258,25237	4/fev/97	17320,1248	2/fev/00	35513,30003
8/ago/95	7903,474214	6/ago/98	18588,40012	5/fev/97	17403,7969	3/fev/00	37020,94012
9/ago/95	7736,204918	7/ago/98	17415,32135	6/fev/97	17403,7969	4/fev/00	37439,72904
10/ago/95	7778,022242	10/ago/98	16913,76948	7/fev/97	18072,33698	7/fev/00	38026,03352

11/ago/95	7443,483651	11/ago/98	16662,57488	12/fev/97	18700,71446	8/fev/00	38528,58022
14/ago/95	7234,397032	12/ago/98	16201,21414	13/fev/97	18575,20631	9/fev/00	37691,00239
15/ago/95	7318,03168	13/ago/98	16578,84335	14/fev/97	18574,36958	10/fev/00	36853,42455
16/ago/95	7483,628282	14/ago/98	16662,57488	17/fev/97	18658,87841	11/fev/00	34843,23776
17/ago/95	7359,849004	17/ago/98	16202,05145	18/fev/97	18909,89471	14/fev/00	34759,47998
18/ago/95	7276,214356	18/ago/98	16118,31992	19/fev/97	18407,8621	15/fev/00	35848,33116
21/ago/95	7527,118299	19/ago/98	15826,09688	20/fev/97	17988,66488	16/fev/00	38109,7913
22/ago/95	7694,387595	20/ago/98	14987,94424	21/fev/97	17654,8132	17/fev/00	37691,00239
23/ago/95	7945,291538	21/ago/98	14150,62892	24/fev/97	18115,00975	18/fev/00	37230,33458
24/ago/95	7778,022242	24/ago/98	13397,04513	25/fev/97	18575,20631	21/fev/00	37186,78053
25/ago/95	7903,474214	25/ago/98	13062,119	26/fev/97	18073,1737	22/fev/00	37020,94012
28/ago/95	8027,253493	26/ago/98	12141,07215	27/fev/97	17738,4853	23/fev/00	37273,05105
29/ago/95	7669,2972	27/ago/98	10550,17304	28/fev/97	17989,5016	24/fev/00	38278,14444
30/ago/95	7527,118299	28/ago/98	10466,44151	3/mar/97	17738,4853	25/fev/00	39283,23784
31/ago/95	7652,570271	31/ago/98	10047,78385	4/mar/97	17989,5016	28/fev/00	38444,82243
1/set/95	7778,022242	1/set/98	11052,56223	5/mar/97	18156,8458	29/fev/00	38528,58022
4/set/95	7903,474214	2/set/98	11722,41449	6/mar/97	18240,5179	1/mar/00	39616,59382
5/set/95	7962,018467	3/set/98	11136,29376	7/mar/97	18324,19	2/mar/00	38444,82243
6/set/95	7945,291538	4/set/98	10718,47342	10/mar/97	18407,8621	3/mar/00	40455,00923
8/set/95	7861,65689	8/set/98	10132,3527	11/mar/97	18324,19	8/mar/00	41418,22373
11/set/95	8028,089839	9/set/98	9880,320784	12/mar/97	17738,4853	9/mar/00	42875,60916
12/set/95	8112,560833	10/set/98	8205,690142	13/mar/97	17822,1574	10/mar/00	41627,61819
13/set/95	8405,282101	11/set/98	9461,663123	14/mar/97	18115,00975	13/mar/00	40036,22031
14/set/95	8948,907311	14/set/98	10633,90457	17/mar/97	17612,97715	14/mar/00	38947,36913
15/set/95	9115,34026	15/set/98	12558,89249	18/mar/97	17529,30505	15/mar/00	39784,94696
18/set/95	9283,445902	16/set/98	11303,75683	19/mar/97	17738,4853	16/mar/00	39701,18918
19/set/95	9074,359283	17/set/98	9629,126187	20/mar/97	18491,5342	17/mar/00	39282,40026
20/set/95	8906,253641	18/set/98	10005,91808	21/mar/97	18491,5342	20/mar/00	38109,7913
21/set/95	8614,36872	21/set/98	9712,857719	24/mar/97	18163,56008	21/mar/00	38863,61135
22/set/95	8530,734072	22/set/98	9964,052316	25/mar/97	18582,07529	22/mar/00	38955,74491
25/set/95	8447,099425	23/set/98	11554,95142	26/mar/97	18247,26312	23/mar/00	38778,17841
26/set/95	8321,647453	24/set/98	10885,09917	31/mar/97	17619,49031	24/mar/00	46500
27/set/95	8279,830129	25/set/98	10298,97844	1/abr/97	17577,63879	27/mar/00	50000
28/set/95	8279,830129	28/set/98	10591,20149	2/abr/97	17200,9751	28/mar/00	47900
29/set/95	8447,099425	29/set/98	10550,17304	3/abr/97	17075,42054	29/mar/00	47701
2/out/95	7945,291538	30/set/98	10214,4096	4/abr/97	17284,67814	30/mar/00	46600
3/out/95	8070,74351	1/out/98	9294,200059	7/abr/97	17535,78727	31/mar/00	47200
4/out/95	7945,291538	2/out/98	9796,589252	8/abr/97	17786,89639	3/abr/00	46801
5/out/95	8112,560833	5/out/98	9336,065825	9/abr/97	18038,00552	4/abr/00	46550
6/out/95	8112,560833	6/out/98	9796,589252	10/abr/97	18079,85704	5/abr/00	48500
9/out/95	7987,108862	7/out/98	9629,126187	11/abr/97	17576,80176	6/abr/00	49200
10/out/95	7861,65689	8/out/98	9712,857719	14/abr/97	17535,78727	7/abr/00	49600
11/out/95	7987,108862	9/out/98	10381,87266	15/abr/97	17493,93575	10/abr/00	48400
13/out/95	7903,474214	13/out/98	11010,69647	16/abr/97	17410,23271	11/abr/00	47201
16/out/95	7819,839566	14/out/98	10926,96493	17/abr/97	17075,42054	12/abr/00	48200
17/out/95	8114,233526	15/out/98	11554,95142	18/abr/97	17200,9751	13/abr/00	47198
18/out/95	8028,926186	16/out/98	11638,68296	22/abr/97	17619,49031	14/abr/00	44590
19/out/95	7970,381932	19/out/98	11889,87755	23/abr/97	17954,30248	17/abr/00	42900
20/out/95	8070,74351	20/out/98	12810,92441	24/abr/97	18246,42609	18/abr/00	42603
23/out/95	7945,291538	21/out/98	13983,16585	25/abr/97	18330,96617	19/abr/00	40700
24/out/95	7736,204918	22/out/98	14065,22276	28/abr/97	18456,52073	20/abr/00	40800
25/out/95	7401,666327	23/out/98	13145,85053	29/abr/97	18707,62985	24/abr/00	40000
26/out/95	6899,858441	26/out/98	12225,641	30/abr/97	18707,62985	25/abr/00	43600
27/out/95	7025,310412	27/out/98	12057,34062	2/mai/97	19126,14506	26/abr/00	43000
30/out/95	6858,041117	28/out/98	12224,80368	5/mai/97	19084,29354	27/abr/00	43000
31/out/95	6942,512111	29/out/98	11680,54872	6/mai/97	19042,44202	28/abr/00	43010
1/nov/95	6858,041117	30/out/98	12559,72981	7/mai/97	19000,5905	2/mai/00	44100
3/nov/95	6941,675765	3/nov/98	11690,10549	8/mai/97	19000,5905	3/mai/00	43900
6/nov/95	6858,041117	4/nov/98	14233,52314	9/mai/97	18624,76384	4/mai/00	44000
7/nov/95	6816,223793	5/nov/98	15155,4073	12/mai/97	19000,5905	5/mai/00	43900
8/nov/95	7024,474066	6/nov/98	15404,92727	13/mai/97	19419,94274	8/mai/00	43950
9/nov/95	7025,310412	9/nov/98	15574,06496	14/mai/97	18916,88746	9/mai/00	43990
10/nov/95	6889,935475	10/nov/98	15239,13884	15/mai/97	19335,40267	10/mai/00	43200
13/nov/95	6439,867878	11/nov/98	14987,94424	16/mai/97	20171,59606	11/mai/00	42162
14/nov/95	6523,502526	12/nov/98	14234,36045	19/mai/97	20297,98765	12/mai/00	42450
16/nov/95	6766,043004	13/nov/98	14644,64496	20/mai/97	21176,86959	15/mai/00	43200
17/nov/95	6732,589145	16/nov/98	15406,6019	21/mai/97	20549,09678	16/mai/00	44600
20/nov/95	6565,31985	17/nov/98	15908,15378	22/mai/97	20507,24525	17/mai/00	44151
21/nov/95	6482,521548	18/nov/98	16242,24259	23/mai/97	20256,13613	18/mai/00	44300
22/nov/95	6439,867878	19/nov/98	15657,7965	26/mai/97	20423,54221	19/mai/00	43300
23/nov/95	6774,406469	20/nov/98	15424,18552	27/mai/97	20423,54221	22/mai/00	42050
24/nov/95	6648,954498	23/nov/98	15574,06496	28/mai/97	20925,76046	23/mai/00	42330
27/nov/95	6607,137174	24/nov/98	15489,49612	30/mai/97	21427,97872	24/mai/00	41800
28/nov/95	6314,415906	25/nov/98	15908,99109	2/jun/97	20925,76046	25/mai/00	42500
29/nov/95	6439,867878	26/nov/98	15406,6019	3/jun/97	20674,65134	26/mai/00	42010
30/nov/95	6858,041117	27/nov/98	15281,0046	4/jun/97	20590,9483	29/mai/00	42400
1/dez/95	7108,108714	30/nov/98	14318,09198	5/jun/97	20842,05742	30/mai/00	42530
4/dez/95	7067,127736	1/dez/98	14611,15235	6/jun/97	21761,95385	31/mai/00	41900

5/dez/95	6983,493089	2/dez/98	14904,21271	9/jun/97	22181,30609	1/jun/00	43901
6/dez/95	6899,858441	3/dez/98	13564,50819	10/jun/97	22265,00913	2/jun/00	46503
7/dez/95	7025,310412	4/dez/98	13313,3136	11/jun/97	23436,01469	5/jun/00	46499
8/dez/95	6899,858441	7/dez/98	13899,43432	12/jun/97	24776,10039	6/jun/00	48150
11/dez/95	6941,675765	8/dez/98	13397,04513	13/jun/97	25278,31864	7/jun/00	49998
12/dez/95	7024,474066	9/dez/98	12810,92441	16/jun/97	25780,53689	8/jun/00	50207
13/dez/95	7033,673877	10/dez/98	12810,92441	17/jun/97	25947,94298	9/jun/00	50200
14/dez/95	6732,589145	11/dez/98	12308,53521	18/jun/97	25864,23993	12/jun/00	49900
15/dez/95	6690,771821	14/dez/98	11136,29376	19/jun/97	24859,80343	13/jun/00	50100
18/dez/95	6648,954498	15/dez/98	11554,95142	20/jun/97	24441,28822	14/jun/00	49300
19/dez/95	6605,464481	16/dez/98	11052,56223	23/jun/97	24859,80343	15/jun/00	49100
20/dez/95	6899,858441	17/dez/98	10968,8307	24/jun/97	25947,94298	16/jun/00	50499
21/dez/95	6858,041117	18/dez/98	11135,45645	25/jun/97	26701,27035	19/jun/00	50400
22/dez/95	7098,908903	21/dez/98	11764,28026	26/jun/97	25947,94298	20/jun/00	50900
26/dez/95	7067,127736	22/dez/98	12141,07215	27/jun/97	25278,31864	21/jun/00	51500
27/dez/95	7042,037342	23/dez/98	12308,53521	30/jun/97	25028,04655	23/jun/00	51510
28/dez/95	6940,839418	24/dez/98	12308,53521	1/jul/97	25445,72472	26/jun/00	51600
29/dez/95	6940,839418	28/dez/98	11680,54872	2/jul/97	26115,34906	27/jun/00	53200
2/jan/96	7108,945006	29/dez/98	11220,0253	3/jul/97	25362,02168	28/jun/00	54500
3/jan/96	7778,022242	30/dez/98	11471,21989	4/jul/97	25864,23993	29/jun/00	53200
4/jan/96	7903,474214	31/dez/98	11471,21989	7/jul/97	26366,45818	30/jun/00	54500
5/jan/96	8028,926186	4/jan/99	11136,29376	8/jul/97	27454,59773	3/jul/00	55750
8/jan/96	8112,560833	5/jan/99	11553,27679	9/jul/97	27454,59773	4/jul/00	57200
9/jan/96	8028,926186	6/jan/99	11931,74332	10/jul/97	27203,48866	5/jul/00	55800
10/jan/96	7903,474214	7/jan/99	11621,93665	11/jul/97	26533,86427	6/jul/00	55810
11/jan/96	8110,051794	8/jan/99	11210,81483	14/jul/97	25613,13081	7/jul/00	53650
12/jan/96	8447,099425	11/jan/99	10508,30727	15/jul/97	23520,55476	10/jul/00	53850
15/jan/96	8572,551396	12/jan/99	9964,052316	16/jul/97	25947,94298	11/jul/00	50710
16/jan/96	8406,118447	13/jan/99	9001,139697	17/jul/97	24524,99126	12/jul/00	49820
17/jan/96	8363,464777	14/jan/99	7284,64329	18/jul/97	23687,96085	13/jul/00	48450
18/jan/96	8279,830129	15/jan/99	8790,973551	21/jul/97	23687,96085	14/jul/00	50000
19/jan/96	8363,464777	18/jan/99	9210,468527	22/jul/97	25696,83385	17/jul/00	49290
22/jan/96	8279,830129	19/jan/99	9377,931591	23/jul/97	26450,16123	18/jul/00	49000
23/jan/96	8321,647453	20/jan/99	9922,18655	24/jul/97	26366,45818	19/jul/00	49400
24/jan/96	8154,378157	21/jan/99	9712,857719	25/jul/97	26784,13636	20/jul/00	52010
25/jan/96	8154,378157	22/jan/99	9294,200059	28/jul/97	27454,59773	21/jul/00	51700
26/jan/96	7945,291538	25/jan/99	9294,200059	29/jul/97	27621,16678	24/jul/00	50500
29/jan/96	7945,291538	26/jan/99	9588,097737	30/jul/97	28040,51902	25/jul/00	48450
30/jan/96	8363,464777	27/jan/99	9587,260421	31/jul/97	27538,30077	26/jul/00	48860
31/jan/96	8530,734072	28/jan/99	10884,26185	1/ago/97	25947,94298	27/jul/00	47990
1/fev/96	9032,541959	29/jan/99	11387,48836	4/ago/97	25948,78001	28/jul/00	46650
2/fev/96	9576,167169	1/fev/99	12643,46134	5/ago/97	25110,91256	31/jul/00	46820
5/fev/96	9408,897874	2/fev/99	13271,44783	6/ago/97	25780,53689	1/ago/00	45360
6/fev/96	9491,696175	3/fev/99	13062,119	7/ago/97	25529,42777	2/ago/00	46310
7/fev/96	9291,809367	4/fev/99	12978,38747	8/ago/97	25278,31864	3/ago/00	46640
8/fev/96	9325,263226	5/fev/99	12434,13251	11/ago/97	25361,18465	4/ago/00	46000
9/fev/96	9283,445902	8/fev/99	13229,58207	12/ago/97	25445,72472	7/ago/00	47540
12/fev/96	9449,878851	9/fev/99	12936,5217	13/ago/97	26115,34906	8/ago/00	45800
13/fev/96	9408,897874	10/fev/99	12978,38747	14/ago/97	26616,73028	9/ago/00	45140
14/fev/96	9282,609556	11/fev/99	12852,79017	15/ago/97	25195,45263	10/ago/00	49900
15/fev/96	9074,359283	12/fev/99	12560,56712	18/ago/97	24608,69431	11/ago/00	49900
16/fev/96	8865,272663	17/fev/99	12727,19287	19/ago/97	24273,88214	14/ago/00	49700
21/fev/96	8865,272663	18/fev/99	12978,38747	20/ago/97	23436,85172	15/ago/00	50400
22/fev/96	9157,993931	19/fev/99	13480,77666	21/ago/97	23939,907	16/ago/00	51400
23/fev/96	9367,08055	22/fev/99	14318,09198	22/ago/97	23771,66389	17/ago/00	53220
26/fev/96	9534,349845	23/fev/99	14150,62892	25/ago/97	23855,36693	18/ago/00	51800
27/fev/96	9533,513499	24/fev/99	13690,10549	26/ago/97	24106,47605	21/ago/00	50950
28/fev/96	9617,148147	25/fev/99	12978,38747	27/ago/97	25278,31864	22/ago/00	51500
29/fev/96	9157,993931	26/fev/99	13489,98713	28/ago/97	24273,88214	23/ago/00	52690
1/mar/96	9325,263226	1/mar/99	14318,09198	29/ago/97	22265,00913	24/ago/00	52500
4/mar/96	9576,167169	2/mar/99	14150,62892	1/set/97	21176,86959	25/ago/00	53500
5/mar/96	9910,70576	3/mar/99	13917,01794	2/set/97	23521,39179	28/ago/00	53010
6/mar/96	9534,349845	4/mar/99	14820,48118	3/set/97	23772,50092	29/ago/00	54250
7/mar/96	9283,445902	5/mar/99	15071,67577	4/set/97	24106,47605	30/ago/00	56400
8/mar/96	9157,993931	8/mar/99	15825,25956	5/set/97	24859,80343	31/ago/00	55100
11/mar/96	9197,302215	9/mar/99	15783,39379	8/set/97	24943,50647	1/set/00	55250
12/mar/96	9199,811254	10/mar/99	16578,84335	9/set/97	25277,48161	4/set/00	55160
13/mar/96	9367,08055	11/mar/99	16997,50101	10/set/97	24190,1791	5/set/00	54810
14/mar/96	9534,349845	12/mar/99	16914,60679	11/set/97	23353,14868	6/set/00	56900
15/mar/96	9699,946448	15/mar/99	18337,20552	12/set/97	23856,3464	8/set/00	56500
18/mar/96	9704,12818	16/mar/99	18337,20552	15/set/97	23020,08934	11/set/00	57700
19/mar/96	9534,349845	17/mar/99	18169,74246	16/set/97	24107,47465	12/set/00	56650
20/mar/96	9617,148147	18/mar/99	18337,20552	17/set/97	24944,56881	13/set/00	55190
21/mar/96	9617,148147	19/mar/99	18672,13165	18/set/97	25112,82473	14/set/00	53990
22/mar/96	9705,126934	22/mar/99	18002,27939	19/set/97	25278,56937	15/set/00	54300
25/mar/96	9831,471806	23/mar/99	17751,0848	22/set/97	25447,66239	18/set/00	52900
26/mar/96	9831,471806	24/mar/99	17205,99252	23/set/97	25195,69705	19/set/00	52620
27/mar/96	10039,81534	25/mar/99	18426,71228	24/set/97	25196,53415	20/set/00	53510

28/mar/96	10040,65206	26/mar/99	18384,83339	25/set/97	25029,11532	21/set/00	52510
29/mar/96	9873,307856	29/mar/99	18996,2652	26/set/97	25531,37181	22/set/00	51790
1/abr/96	9956,979956	30/mar/99	19851,43217	29/set/97	25447,66239	25/set/00	52000
2/abr/96	9747,799705	31/mar/99	19934,35237	30/set/97	26117,33772	26/set/00	53210
3/abr/96	9746,962984	5/abr/99	21023,20355	1/out/97	26453,01248	27/set/00	53150
8/abr/96	9536,946012	6/abr/99	21942,86401	2/out/97	27162,86832	28/set/00	52900
9/abr/96	9454,110633	7/abr/99	22949,63256	3/out/97	27204,72303	29/set/00	52700
10/abr/96	9413,111303	8/abr/99	22874,25056	6/out/97	26786,17595	2/out/00	52360
11/abr/96	9413,111303	9/abr/99	21944,53917	7/out/97	26285,59365	3/out/00	52600
12/abr/96	9580,455504	12/abr/99	22195,81252	8/out/97	26284,75655	4/out/00	53000
15/abr/96	9538,619454	13/abr/99	21994,79384	9/out/97	26703,30363	5/out/00	53240
16/abr/96	9622,291554	14/abr/99	22112,05473	10/out/97	26620,43131	6/out/00	51800
17/abr/96	9704,290213	15/abr/99	22028,29695	13/out/97	26284,75655	9/out/00	51700
18/abr/96	9705,126934	16/abr/99	22572,72254	14/out/97	26786,17595	10/out/00	52850
19/abr/96	9831,471806	19/abr/99	22195,81252	15/out/97	26703,30363	11/out/00	53000
22/abr/96	9873,307856	20/abr/99	21148,00265	16/out/97	26116,50063	13/out/00	53640
23/abr/96	9789,635755	22/abr/99	21190,71912	17/out/97	26117,33772	16/out/00	51900
24/abr/96	9622,291554	23/abr/99	20939,44577	20/out/97	26787,01305	17/out/00	49500
25/abr/96	9664,127605	26/abr/99	20436,89907	21/out/97	27121,85071	18/out/00	50000
26/abr/96	9664,127605	27/abr/99	20101,03036	22/out/97	26954,43188	19/out/00	50900
29/abr/96	9664,127605	28/abr/99	20311,2624	23/out/97	25029,95241	20/out/00	50000
30/abr/96	9664,127605	29/abr/99	19851,43217	24/out/97	24191,18406	23/out/00	47950
2/mai/96	9789,635755	30/abr/99	22371,70386	27/out/97	19671,71271	24/out/00	45980
3/mai/96	9747,799705	3/mai/99	22447,08586	28/out/97	21346,73812	25/out/00	45800
6/mai/96	9873,307856	4/mai/99	21944,53917	29/out/97	21261,35451	26/out/00	48900
7/mai/96	10165,32349	5/mai/99	22276,21999	30/out/97	18500,61798	27/out/00	49370
8/mai/96	10166,16021	6/mai/99	22488,12718	31/out/97	17160,43023	30/out/00	50710
9/mai/96	10500,84861	7/mai/99	22949,63256	3/nov/97	19671,71271	31/out/00	50610
10/mai/96	10625,52004	10/mai/99	22530,84365	4/nov/97	20759,93511	1/nov/00	51180
13/mai/96	10793,70096	11/mai/99	22447,08586	5/nov/97	20507,96977	3/nov/00	50100
14/mai/96	10710,02886	12/mai/99	22279,5703	6/nov/97	19419,74737	6/nov/00	51190
15/mai/96	10752,70163	13/mai/99	22167,4028	7/nov/97	18416,07147	7/nov/00	52200
16/mai/96	10875,69962	14/mai/99	23033,39035	10/nov/97	19169,45621	8/nov/00	51310
17/mai/96	10584,52071	17/mai/99	22371,70386	11/nov/97	18750,90913	9/nov/00	50000
20/mai/96	10835,53701	18/mai/99	21609,50803	12/nov/97	16741,04606	10/nov/00	50310
21/mai/96	10541,84794	19/mai/99	21106,96134	13/nov/97	16993,0114	13/nov/00	50800
22/mai/96	10542,68466	20/mai/99	21149,67781	14/nov/97	17745,55905	14/nov/00	51000
23/mai/96	10291,66836	21/mai/99	20604,41464	17/nov/97	18918,32796	16/nov/00	51800
24/mai/96	10292,50508	24/mai/99	19932,67722	18/nov/97	18834,61855	17/nov/00	51900
27/mai/96	10375,34046	25/mai/99	19004,64098	19/nov/97	19002,03738	20/nov/00	51610
28/mai/96	10457,33912	26/mai/99	20436,89907	20/nov/97	18667,19972	21/nov/00	52200
29/mai/96	10208,83298	27/mai/99	20854,85041	21/nov/97	19420,58446	22/nov/00	52100
30/mai/96	10249,83231	28/mai/99	20520,65685	24/nov/97	18082,0709	23/nov/00	51000
31/mai/96	10082,48811	31/mai/99	21064,24487	25/nov/97	18332,36205	24/nov/00	50510
3/jun/96	10166,16021	1/jun/99	21106,96134	26/nov/97	19002,03738	27/nov/00	48760
4/jun/96	10166,16021	2/jun/99	20771,9302	27/nov/97	20340,55094	28/nov/00	48400
5/jun/96	9998,816007	4/jun/99	21023,20355	28/nov/97	20341,38803	29/nov/00	48380
7/jun/96	10040,65206	7/jun/99	20521,49443	1/dex/97	20592,51628	30/nov/00	47500
10/jun/96	9915,143906	8/jun/99	20353,14129	2/dex/97	22266,7046	1/dex/00	47200
11/jun/96	9956,979956	9/jun/99	19691,4548	3/dex/97	21931,86693	4/dex/00	47100
12/jun/96	10207,15954	10/jun/99	20101,86794	4/dex/97	21764,4481	5/dex/00	47790
13/jun/96	10124,32416	11/jun/99	19934,35237	5/dex/97	21931,02984	6/dex/00	44600
14/jun/96	10375,34046	14/jun/99	20018,94773	8/dex/97	22182,99518	7/dex/00	45190
17/jun/96	10500,84861	15/jun/99	20110,24372	9/dex/97	21764,4481	8/dex/00	45200
18/jun/96	10459,01256	16/jun/99	21065,08244	10/dex/97	21094,77277	11/dex/00	44950
19/jun/96	10375,34046	17/jun/99	21441,99247	11/dex/97	19671,71271	12/dex/00	44800
20/jun/96	10416,33979	18/jun/99	21777,0236	12/dex/97	19922,84095	13/dex/00	47000
21/jun/96	10751,86491	21/jun/99	22715,94835	15/dex/97	20843,64453	14/dex/00	47500
24/jun/96	10835,53701	22/jun/99	23284,6637	16/dex/97	21178,48219	15/dex/00	47290
25/jun/96	10626,35676	23/jun/99	22991,51145	17/dex/97	20843,64453	18/dex/00	46680
26/jun/96	10710,02886	24/jun/99	22363,32808	18/dex/97	19922,84095	19/dex/00	46320
27/jun/96	10709,19214	25/jun/99	21777,0236	19/dex/97	18750,90913	20/dex/00	45410
28/jun/96	10333,50441	28/jun/99	22003,16961	22/dex/97	19336,87504	21/dex/00	44700
1/jul/96	10291,66836	29/jun/99	22196,65009	23/dex/97	19922,84095	22/dex/00	45000
2/jul/96	10751,86491	30/jun/99	23033,39035	24/dex/97	19922,84095	26/dex/00	44700
3/jul/96	11002,88121	1/jul/99	23326,54259	26/dex/97	20507,13268	27/dex/00	45200

## 6.5 Série da Nasdaq

-0,017915	-0,00247	-0,004342	-0,019418	-0,005456	-0,011236	-0,03212	-0,034289	-0,017074
0,004809	0,006902	0,029298	0,002009	0	0,008999	-0,025546	0	0,024903
0,023469	0,002698	0,006668	0,018887	0,006543	-0,024371	0,016604	-0,008759	0,0006
0,040859	0,005375	-0,037054	0,013692	0,027869	0,013675	-0,015843	-0,003672	0,005381
-0,059782	0,02455	0,009393	-0,011725	0,004746	-0,024056	0	0,021115	-0,019266
-0,032026	0,006635	0,020289	0,007832	-0,001053	0,017811	-0,039555	-0,033703	-0,020885
-0,073122	0	0	0,006803	0,018263	-0,00114	0,02191	-0,012749	-0,003108
-0,024978	-0,021727	-0,02317	0,004831	-0,027787	0,035288	-0,032248	0,012749	-0,00187
-0,101831	-0,031881	0,0036	0,023811	-0,003195	-0,003308	0,034587	0,011852	0,019154
0,063578	-0,00876	-0,044827	0,018799	0	-0,004982	0,007892	-0,026867	-0,005523
0,069577	-0,020829	-0,006787	-0,005568	0,014823	-0,002779	0,011429	-0,000757	-0,001232
-0,023468	-0,011877	0,007914	-0,00373	0,018772	-0,021373	-0,003795	0,002268	0,01042
-0,016871	0,011364	-0,021245	0	-0,002579	-0,016624	-0,026977	-0,013688	0,023502
-0,045475	0,005123	-0,017795	-0,028905	-0,025106	-0,009875	0,008557	-0,003087	-0,009575
0,063695	-0,02039	-0,036165	-0,037223	0,023555	0,000584	-0,011687	0	0
-0,022069	-0,028566	0,00766	0,011412	-0,001035	-0,029605	-0,022188	-0,01548	0,034282
0,038903	0,036617	-0,025215	-0,032083	-0,013556	-0,004216	0,0032	0,011633	0,014422
0,022791	0,007987	0,025616	0,004574	0,029992	0,012594	-0,029175	-0,001543	0,000573
-0,01988	-0,017603	0,024585	0,051395	0,041901	-0,019862	0,004103	-0,012432	-0,017904
-0,020823	-0,006551	-0,051438	0,004326	0,003901	-0,019645	-0,024877	-0,001565	0,010435
0,003501	-0,0167	0	-0,027226	-0,016683	-0,021938	-0,007585	0,025514	0,012038
0,025741	-0,013998	-0,045127	0,018555	-0,014453	-0,002538	0,033281	0	-0,00114
-0,039546	-0,008986	-0,021366	-0,024983	0,005008	0,074652	0	0,027853	-0,002856
-0,023161	0,032825	-0,010634	0,01526	0,004485	-0,027497	-0,019001	-0,009698	-0,029021
-0,057405	-0,028186	0,007986	-0,008832	-0,032342	-0,030147	-0,022776	-0,002251	0,006457
0,033409	-0,028723	-0,050294	-0,019911	0,003077	0,031963	0,035209	0,013433	-0,01652
0,008252	-0,032174	0,014302	-0,015198	0,017259	0,039151	-0,022492	-0,020218	0,014763
0,022139	0,019201	-0,030228	0,012681	0,009018	0,004064	-0,01785	0,005281	0,007591
0,030032	-0,014582	0,011734	0,019467	-0,011033	0,002315	-0,022551	-0,01899	-0,010526
-0,01983	-0,032492	0,028067	0,002962	-0,032807	0,006336	-0,018593	0,015221	0,017483
-0,029512	-0,001489	0,009033	-0,005436	0,013972	-0,022656	0,008897	-0,011394	0,002308
-0,043014	-0,034868	-0,025041	0,004445	0	-0,02258	-0,013375	0,003812	-0,002308
-0,007699	-0,022781	-0,054964	0,020025	0,007679	0,005394	0,043046	-0,008407	-0,005794
-0,060978	0,04596	-0,065416	0,009146	0,00782	0,014833	0,039617	0,001534	0,008102
0,032943	-0,008173	0,046733	-0,010113	0,005551	-0,022031	0,00905	-0,020907	-0,011011
-0,020384	-0,023371	-0,021571	-0,01581	0	0	0,048751	-0,005492	-0,001166
0	-0,013154	-0,015845	-0,002954	-0,004034	-0,005433	-0,039782	0,006275	-0,023614
0	0,074986	-0,026098	-0,031049	0,032317	-0,01218	0,031952	0,047337	-0,012624
0,076267	0,018546	0,031236	-0,001527	0,007312	-0,009852	0,0125	-0,000746	-0,000605
-0,01691	-0,004028	-0,04938	-0,003572	-0,010742	-0,033347	0,018937	0,038072	-0,004854
0,052131	-0,014226	-0,019576	0,013205	0,009284	0,009551	-0,013062	0,006445	0,010889
0,062216	-0,057158	0,041415	-0,008614	-0,005365	0,010716	0,020687	-0,002859	0,01493
0,002358	0,012919	0,016201	-0,023688	0,000978	-0,025398	-0,016037	0,004286	0
-0,017419	0,001832	-0,005719	0,005716	-0,012288	-0,012945	0,024338	0,013452	0,007676
0,021857	-0,026899	0,027765	-0,033187	-0,01545	-0,002609	-0,011334	-0,037254	0,001176
-0,003392	0,054578	-0,081703	0,007471	0,00501	-0,022458	-0,011464	0,015212	0,021506
0,012726	-0,01104	-0,01796	-0,026941	-0,028899	0,014589	0,019788	-0,00288	0,012572
-0,028001	0,035081	0,01038	0,015714	0,021375	-0,014589	0,002258	-0,002185	0
0,021788	-0,010483	-0,031468	-0,009182	-0,028803	0,00533	0,023041	-0,010167	0,009607
-0,014122	0	-0,063679	0,039367	0	0,031396	0,026107	-0,021395	-0,007338
0,012822	-0,055369	-0,020532	-0,002612	-0,024654	-0,00646	0,002859	0,005208	-0,011398
0,003893	-0,00995	0,085332	-0,024889	0,0205	-0,030265	0,012766	0,035705	0,01479
0,032865	-0,054918	-0,037094	-0,011866	0,010887	-0,021608	-0,030045	0	0,008192
0,005748	-0,020681	0,015003	-0,028082	-0,002577	-0,016518	-0,013158	-0,008349	0,01005
0,012629	0,056034	0,058939	0,007228	0,003091	0,013099	-0,029875	0,004326	0,00554
-0,055394	-0,034363	0,025061	0	-0,026585	-0,024967	0,023231	-0,005049	0
0,017275	-0,00185	0,032127	-0,019016	0,01051	0,003505	0,008115	-0,018979	0,016982
-0,013641	-0,051172	-0,026351	-0,006799	0,010918	0,020776	0,036079	0,00661	0,006497
0,020773	-0,00174	0,008783	-0,041201	-0,011984	0,002738	-0,012121	-0,005138	0,008598
-0,016119	-0,041243	0,078001	0,004138	-0,029203	-0,041164	-0,013724	0,018951	-0,005903
0,022696	0	0,048355	0	-0,009202	-0,033312	0,031499	0,004323	0,016547
0,006534	0,052672	-0,008746	0	-0,014239	0,01607	0,033949	0,002872	-0,008508
0	-0,008299	-0,049276	0	-0,017247	0	0,000681	0,021979	-0,02851
-0,030263	-0,051659	-0,020609	0	0,028763	0,0481	0,008811	0,01738	0,011989
0,022487	-0,01029	0,021095	-0,070258	0,013539	-0,029435	-0,025975	0,010285	0,004864
0,015531	-0,041099	-0,01221	-0,015949	-0,008644	-0,017941	0,030016	-0,008158	0,003229
-0,010495	0,017929	0,019947	-0,017518	-0,020834	-0,025871	0	-0,015215	0,013874
-0,0063	-0,011025	0,019085	-0,038017	0	0,020589	-0,008087	0,018641	-0,002122
0,035754	0,09961	0,04843	-0,033175	-0,030374	0	0,004049	0,008137	0,014237

0.018127	-0.032709	-0.034352	0.052031	0.014177	0.006529	-0.024541	-0.00477	-0.002098
0.016863	-0.015856	0.021209	0.001999	-0.034368	-0.005801	-0.013199	0.026291	-0.016398
0.006807	0.057864	-0.008246	-0.025625	0.005231	0.015873	-0.013378	0.000685	0.014297
0	0.033044	0.011434	-0.002051	0.025752	-0.029051	0.007766	0.01321	-0.030989
-0.052587	-0.027915	-0.019284	0.025677	-0.001696	-0.028403	-0.039446	-0.011218	-0.01475
0.031309	-0.037885	-0.013086	-0.012756	-0.008525	-0.027673	0.017404	-0.011345	-0.013854
-0.021984	-0.033865	-0.010387	0.008075	-0.012059	-0.043	0.004304	0.019934	0.018248
-0.027738	-0.028244	-0.011936	0.05794	0.040753	0.048441	0.002145	0.013721	-0.021042
0.011982	-0.01061	0.001919	0.010089	0.030587	-0.039531	-0.027518	-0.000649	0.02487
-0.010477	-0.044002	0.037634	0.004997	0.003758	0.017586	0.027518	-0.003252	0.002181
-0.037297	-0.073924	0.012844	0.000623	0.01278	0.056234	-0.005731	0.01037	0.024214
-0.047711	0.002996	0.002276	-0.022671	-0.004242	0.006719	-0.021787	-0.007767	0.006359
0.027996	0.072917	0.047512	0.035047	0.025185	-0.011976	-0.005891	-0.030345	0.007893
-0.021737	-0.00918	0.003463	0.046268	-0.000518	-0.038814	0.006825	-0.00134	-0.007385
-0.007625	0.017682	-0.031163	-0.004119	-0.016728	-0.025318	0.013844	-0.000671	0.009459
0.027502	0.007455	0.017238	0.015214	-0.018623	-0.034233	-0.007262	0.012008	0.001568
0.007155	-0.034803	-0.013678	-0.045138	-0.004306	0.043803	-0.00292	0.001325	0.009873
0.01987	-0.075198	0	0.004244	0.007523	0.018529	-0.014728	0	0.004643
-0.003631	0.132604	-0.033886	0.01083	0.004806	0.027715	-0.005952	0.030651	
0.001298	-0.019291	-0.042719	0.021901	0.002129	-0.008387	0	0.003848	
-0.024693	-0.063941	0.011919	-0.002345	-0.02585	0.000765	0.033031	0.007648	
0.007683	-0.004996	0.018315	0.018288	0.01945	-0.029505	0.001443	0.013245	
0.015971	0.018607	0.003252	0.024524	-0.009677	0.049953	0.024218	-0.003138	
0.000519	0.033437	0.035539	-0.003386	-0.021295	0.008212	0.007711	0.008138	
0.002334	0.045313	-0.007188	-0.039786	0.006601	0.011826	-0.021886	0.019142	
0.020508	-0.005315	0.020527	-0.019603	0.001643	0.024675	0.026067	0.006705	
-0.002795	0	-0.021881	0.013703	0.009804	-0.012987	0.00693	-0.01181	
0.005835	-0.00305	-0.020065	0.032599	0.009709	0.022973	-0.001382	-0.014237	
0.001264	0.024143	-0.000461	0.02712	-0.032182	0.009887	0.01032	0.014852	
0.013302	0.031189	-0.022368	0.022597	-0.011148	-0.029959	-0.015173	0.010996	
0.010417	0.000722	-0.045309	0.001634	0.003358	0.007934	-0.010479	0.004848	

## 6.6 Série de retornos da IBM

0.428	2.134	-0.851	0.000	-0.149	-0.276	-1.324	-1.455
-0.428	1.544	0.000	1.122	-0.249	-1.671	2.077	0.487
-1.439	0.871	0.186	-0.648	0.447	-0.281	-0.781	0.243
1.510	0.325	0.557	-0.950	-0.099	-1.347	2.105	-0.649
2.398	0.351	0.370	-0.478	1.429	0.356	0.547	-0.817
4.283	1.258	0.735	-0.120	-0.294	-0.965	-0.055	-2.492
1.228	-0.320	1.455	-0.985	-0.049	0.680	0.435	0.837
-0.728	-1.019	0.270	-1.220	0.539	1.135	1.458	-0.586
1.647	0.269	-0.451	0.123	0.536	-0.141	0.374	-0.370
-3.289	0.803	0.181	1.097	-0.536	1.089	-0.160	-0.933
-1.328	-0.696	0.000	1.085	0.487	0.869	1.116	0.933
0.814	-1.570	0.270	-0.120	0.388	-1.044	1.676	-2.305
1.217	0.653	0.225	-2.082	-0.194	-0.351	-0.384	-1.391
0.268	2.096	2.352	-0.215	-0.681	0.140	3.484	0.000
-2.649	0.106	0.960	0.459	0.681	-1.555	0.000	2.422
1.163	0.952	-1.576	1.034	-0.486	0.957	-0.221	0.171
0.543	-0.793	0.440	-0.546	-0.293	1.018	0.104	-1.548
1.711	-0.159	-0.794	-0.091	0.195	0.000	-0.052	1.398
3.014	0.318	-0.089	0.091	-1.079	-1.548	-0.104	1.674
-0.097	-0.425	-1.089	0.485	0.443	-0.855	-0.312	0.670
-2.817	-0.640	0.803	0.483	0.441	1.526	-1.099	-1.556
2.494	-0.914	0.973	0.000	0.390	-0.106	-3.044	-0.382
-0.195	0.914	0.000	-0.241	-0.488	-1.994	1.079	1.100
-1.307	-0.214	0.088	-0.727	0.780	-0.144	0.288	-0.676
-1.591	-0.322	1.180	0.606	0.774	0.380	0.053	-1.709
2.252	-0.323	1.681	-0.182	-0.822	-1.228	-0.752	-0.519
-1.385	-1.086	0.597	-0.547	0.243	-0.949	2.132	-2.634
-0.199	0.436	-0.682	0.183	-0.194	-2.564	0.211	1.370
0.795	2.151	-0.171	-0.426	0.533	0.113	0.210	3.068
1.896	1.531	-0.645	-1.074	0.530	1.641	-0.844	-0.426
0.129	-0.842	-0.432	-0.247	0.240	1.760	0.000	-0.945
-0.065	0.158	-0.478	1.138	1.098	-0.328	0.581	2.303
0.838	-0.847	0.953	2.178	0.000	0.036	-2.075	0.421
1.085	0.000	0.344	1.721	0.284	-0.482	1.708	-0.843
-0.254	-0.373	0.770	-0.649	-0.284	-0.683	-0.636	-3.006
4.054	-1.290	0.340	0.237	-0.666	0.663	0.000	2.775
0.366	0.593	0.085	0.177	1.233	2.682	-0.801	0.275
-0.366	0.696	0.887	0.939	0.188	-0.430	2.070	-1.578
-0.245	-1.020	0.168	1.565	0.282	-1.302	-1.482	0.899
-3.619	-0.054	0.481	1.257	0.281	-0.438	-0.749	-1.157

-0.637	-0.867	0.084	0.848	0.351	-1.919	2.545	0.944
-0.835	-0.183	-0.251	-0.058	0.490	0.000	0.574	0.341
2.168	-0.328	-1.477	-0.169	-0.188	-1.958	0.571	1.100
-2.233	-0.494	-0.640	0.843	0.372	0.078	0.619	0.871
-1.168	0.165	0.128	-0.280	-1.120	-0.380	0.308	0.417
1.297	-1.437	1.946	-0.508	0.141	-1.150	0.817	1.979
-1.493	-0.502	-1.095	1.177	0.234	-2.855	1.565	2.976
0.522	-1.807	-0.723	0.111	0.853	-2.612	-0.050	0.632
-0.587	-1.031	-0.342	-0.111	0.093	3.756	0.100	0.828
0.424	0.460	-0.300	-0.671	0.786	0.547	-0.301	-0.235
-0.033	-2.260	0.557	0.112	0.597	0.000	0.000	-1.183
1.231	2.775	-0.600	-0.507	-0.643	-0.158	-0.302	1.887
1.724	0.738	1.365	1.459	0.000	-0.863	0.101	-0.158
-0.762	0.733	0.676	-0.727	-0.369	-0.712	1.001	1.240
-1.478	-2.043	-1.698	-1.015	-0.928	-1.842	1.581	-0.118
-0.910	1.112	-0.171	1.239	0.000	1.842	-1.283	-0.503
-1.580	-0.883	-0.344	-1.580	0.000	1.731	-0.898	-0.311
-3.994	-0.689	0.429	0.114	-1.411	-0.469	0.100	-0.468
1.983	-0.578	-1.207	0.199	-0.570	1.401	-0.401	-0.548
-3.795	0.231	-1.090	-1.802	-1.729	3.493	0.301	-0.079
-1.559	0.634	1.609	-0.695	-2.850	-1.655	0.550	2.944
0.995	-0.518	0.774	0.580	0.992	-0.152	-0.100	-0.306
-3.784	-0.812	0.128	2.512	-0.892	-2.151	0.050	0.458
2.896	-0.058	-0.128	0.898	1.286	-2.119	0.398	0.458
-0.537	-0.292	-0.128	0.112	0.588	1.222	0.485	1.431
2.620	-0.293	0.214	0.334	0.827	-0.039	-0.595	1.337
0.696	0.409	-0.086	1.106	0.435	2.324	0.199	1.247
-0.139	0.117	0.427	0.275	-0.290	0.308	-0.248	1.447
1.791	0.581	0.128	-0.110	-1.955	-2.630	0.199	-0.967
-1.965	0.347	-0.469	0.165	-2.297	-1.182	-0.248	-0.291
-0.874	2.283	0.512	-0.550	0.805	0.711	-1.302	0.255
-0.211	-1.937	0.636	-1.053	-0.100	-2.109	0.402	-0.109
3.254	0.115	1.052	0.722	0.650	-0.443	0.100	-0.657
-2.623	-0.892	1.505	0.497	0.992	0.161	-0.604	-0.698
1.113	1.322	0.770	1.312	-0.197	-1.544	-0.912	-1.150
-1.779	1.136	0.408	-0.709	-1.344	-4.224	1.666	0.149
-1.991	-1.053	-1.281	0.546	0.200	5.526	1.145	-0.299
-1.483	0.428	-0.311	1.082	0.449	-1.384	0.297	-0.976
-3.864	1.020	-1.148	0.590	-0.850	2.589	0.246	-1.905
2.985	0.450	0.732	-0.698	0.401	1.270	1.321	-0.039
-0.442	1.945	-1.416	-0.216	-0.200	-0.316	-0.977	1.565
-2.090	0.604	0.474	0.323	-1.108	0.158	-0.295	-0.303
3.847	-0.859	-1.163	-0.323	-0.304	2.650	0.295	-0.763
2.083	-0.331	-1.771	0.108	0.658	-2.335	0.167	0.915
-1.648	-0.111	0.486	-0.487	0.554	0.315	0.294	0.906
1.719	-1.224	-0.649	-0.489	-0.050	1.093	-1.132	0.898
1.480	0.503	1.933	-0.658	-0.302	0.000	-1.899	3.009
1.390	0.832	-0.640	-1.547	0.302	0.850	-0.709	3.202
-0.069	-0.888	0.427	1.272	-0.453	1.452	-0.509	4.919
-1.675	0.334	0.744	0.000	0.604	1.057	0.102	-1.410
1.675	-1.794	0.738	-0.110	-0.503	0.562	0.914	-1.293
3.335	0.339	0.880	0.330	-0.708	-0.862	-0.507	0.819
-1.824	-0.226	0.572	0.383	-1.844	2.381	1.112	0.542
2.091	0.507	-0.938	0.436	-1.039	1.315	0.100	0.202
-0.535	0.618	0.418	-0.655	1.039	0.174	-0.806	-0.203
1.598	0.112	0.280	-0.110	1.539	0.724	0.404	-2.740
-1.732	0.558	-0.417	0.438	0.609	0.719	-0.911	-1.188
1.169	1.780	0.521	2.586	2.893	0.143	-1.177	-2.634
0.232	-0.437	-0.052	-0.108	-0.493	0.855	-0.052	0.072
2.810	0.328	-0.628	0.531	-0.694	-0.142	1.127	3.332
0.193	-0.219	-0.157	0.108	-1.101	-0.713	0.406	
1.532	-0.109	-0.052	-0.108	-0.201	0.855	1.360	
0.694	-0.219	0.835	1.472	-0.810	1.688	0.050	
0.877	-0.110	0.622	-0.104	-0.407	-0.419	0.399	
-0.625	0.027	0.823	-0.419	1.267	-0.562	0.497	
-1.772	-1.132	0.613	0.628	0.252	-1.991	-0.348	
2.210	0.940	-0.408	0.156	0.950	0.000	0.248	
-0.378	0.494	-0.258	-0.261	1.458	2.273	1.770	
-0.063	-0.439	-1.082	0.000	0.809	2.839	0.777	
-0.251	1.745	-0.208	0.000	0.923	1.154	0.000	
-2.182	-0.978	-1.993	0.209	1.488	0.673	-1.167	
-0.388	-1.707	1.054	1.036	-1.008	0.801	0.000	
1.800	-0.389	0.418	-0.310	-1.210	-1.575	0.488	
-1.471	0.777	0.000	0.516	0.988	-0.848	0.776	
0.388	-0.499	-1.049	0.103	-0.799	1.420	-0.679	
-1.182	-0.111	-0.370	1.028	-0.317	0.938	5.027	
-0.717	1.710	-0.159	-1.028	-0.489	1.519	0.389	
2.551	0.980	-0.319	-1.245	-0.590	0.392	1.917	
0.000	1.719	-1.339	0.209	-0.099	-1.712	-1.365	

-1.769	-1.502	-0.486	1.010	1.080	1.188	0.412
2.246	0.000	-0.217	-0.181	0.584	0.328	0.455
-0.064	0.647	0.892	-0.518	-1.981	-0.196	-0.364
-0.701	0.749	-0.675	-1.464	-1.095	-2.253	-1.008
0.383	1.481	0.108	0.316	0.748	-0.470	1.145
-0.639	0.419	-0.488	0.419	0.248	-0.337	0.000
-1.355	2.531	-3.149	-1.052	-1.397	0.808	0.726
3.008	0.407	1.508	-0.849	-0.101	-0.268	2.676
0.126	0.507	-1.395	-1.936	1.893	-0.674	0.614
-0.822	-0.152	0.951	2.363	-0.594	-0.543	-0.263
2.446	-1.633	0.000	-0.853	-0.698	-0.477	1.739
0.062	-0.619	-0.783	-0.107	-0.050	0.341	0.000
0.679	0.413	0.671	2.435	-0.250	0.544	1.029
0.307	1.737	0.886	-2.007	0.799	0.675	-0.513
1.582	0.757	-1.223	1.378	-0.499	-0.675	-1.208
0.181	-1.214	0.112	-0.846	-0.200	-1.090	1.550
-1.579	-1.435	0.056	0.740	1.984	0.954	1.105
1.277	1.993	-0.954	-0.053	0.588	2.545	0.967
-0.303	-1.120	-2.165	-1.220	0.000	2.094	-1.137
0.363	-0.668	1.146	-2.158	0.972	0.774	-0.085
0.181	-0.830	-0.887	-1.207	1.536	1.277	-0.170
-0.121	0.312	-1.534	0.386	-0.813	-0.509	0.887
0.422	0.621	-0.175	-0.607	-0.337	1.141	0.545
0.599	0.515	0.525	1.373	-1.553	-3.073	-0.168
1.424	0.205	1.786	-0.273	0.974	0.906	-0.504
-0.650	0.307	0.626	1.250	0.964	0.770	-1.571
-0.535	-0.512	0.113	-0.054	0.574	-0.320	-0.300
1.008	-0.412	0.227	-0.433	0.950	0.064	-0.241
0.000	-0.932	0.508	-0.435	0.377	1.905	0.172
-0.829	-1.732	-0.028	-0.218	1.681	0.502	-0.778
0.416	-0.053	-0.876	-0.438	1.858	-0.880	0.519
-0.655	-1.709	-1.142	-1.940	-2.043	0.252	1.286
0.179	0.269	-2.119	-0.956	0.647	1.127	0.510
0.238	1.281	-0.206	1.681	3.220	0.744	-0.510
-0.537	1.422	-3.165	1.762	-1.303	-1.244	1.354
0.179	1.505	1.086	0.653	-1.091	-1.640	2.490
-0.239	1.381	2.402	0.756	-0.366	0.887	0.979
0.060	-0.509	1.714	-0.811	0.549	0.629	-2.632
0.239	-1.337	-1.978	-0.599	1.269	0.025	2.956
-0.209	0.927	0.235	1.946	0.000	0.000	-0.894
-0.811	-0.617	-0.235	0.587	-2.555	2.848	0.245
-0.848	-0.103	1.052	-0.587	2.555	1.093	0.487
-0.183	-0.622	0.464	-0.214	1.075	0.000	-0.487
-1.535	-0.208	1.550	0.908	-1.075	1.409	0.000
0.678	-0.995	-0.801	0.318	0.449	1.800	-0.328
-0.369	0.000	0.516	0.581	0.447	1.278	0.651
-1.366	0.158	-0.229	0.000	-0.492	-1.571	2.247
-0.627	1.563	-0.402	-1.781	-1.583	-0.382	-0.159
1.561	0.000	-0.403	-0.673	-1.933	-0.265	0.594
-0.435	-0.103	-0.637	-0.433	-1.404	0.118	1.139
0.775	0.103	-1.228	1.079	-1.951	-0.354	0.701
0.953	-0.103	0.118	-0.054	0.814	0.590	0.155
-0.061	0.310	0.703	-0.376	1.279	-2.140	-0.700
-0.061	1.231	0.582	1.603	-1.374	3.310	0.389
-0.061	0.810	-0.699	0.845	-0.959	2.299	0.697
0.184	1.108	-1.649	0.733	0.193	0.284	-1.673
0.305	1.590	0.237	-0.576	-0.675	-0.284	-3.067
0.790	1.953	0.826	1.070	2.488	0.000	-0.081
-0.812	-1.363	1.284	1.011	2.552	-0.228	1.207
0.122	0.294	-0.699	-0.673	-0.380	-0.514	-0.441
0.971	-0.687	-0.880	0.415	1.990	-0.057	-1.416
-0.667	0.979	0.059	0.103	1.644	0.855	0.488
0.727	2.977	-0.354	-0.310	-0.538	0.170	1.369
-0.181	0.848	-0.385	0.310	0.135	-1.370	-0.080
0.683	0.581	-1.524	1.130	-0.135	1.596	0.400
1.135	-0.187	0.381	0.611	-2.249	1.794	-2.177
-0.059	4.745	0.598	-0.051	-2.180	0.056	-1.230
0.059	-1.980	0.357	-0.305	-1.857	-1.396	-2.888
0.000	0.544	-0.119	0.407	-1.014	0.169	0.719
-0.835	-0.181	-0.418	-0.560	-1.910	-0.056	0.756
1.131	-1.828	0.000	0.407	-0.894	2.445	0.167
1.790	0.918	0.419	0.608	1.486	1.200	0.915
2.782	0.456	-0.060	1.354	2.907	-0.653	0.824
1.599	-0.639	0.119	-0.550	0.000	-0.219	1.143
-0.979	-1.105	0.000	0.050	-1.153	0.055	1.130
-0.535	-1.117	0.000	0.250	0.289	-1.376	-1.130
0.872	-1.890	0.119	-0.050	-1.017	-2.469	0.000
-0.056	0.950	1.598	-0.250	1.053	0.227	0.889
-0.731	2.428	1.340	0.200	3.094	0.791	-0.081

-0.509 -0.185 -0.232 0.349 1.529 2.004 -2.280  
 0.000 -0.046 -2.348 0.447 0.138 -1.528 2.602

## 6.7 Série da Soja

Data	Valor	19/2/1998	15.08	21/9/1998	14,74	27/4/1999	16,09	22/11/1999	20,54	19/6/2000	18,35	16/1/2001	20,71
29/7/1997	18,04	20/2/1998	15,03	22/9/1998	14,73	28/4/1999	16,04	23/11/1999	20,47	20/6/2000	18,38	17/1/2001	20,5
30/7/1997	17,97	25/2/1998	14,77	23/9/1998	14,74	29/4/1999	15,91	24/11/1999	20,44	21/6/2000	18,71	18/1/2001	20,4
31/7/1997	18,17	26/2/1998	14,66	24/9/1998	14,8	30/4/1999	15,88	25/11/1999	20,41	23/6/2000	18,37	19/1/2001	20,36
1/8/1997	18,1	27/2/1998	14,37	25/9/1998	14,81	3/5/1999	15,77	26/11/1999	20,34	26/6/2000	18,38	22/1/2001	20,35
4/8/1997	18,23	2/3/1998	14,4	28/9/1998	14,79	4/5/1999	16,09	29/11/1999	20,37	27/6/2000	18,58	23/1/2001	19,95
5/8/1997	18,16	3/3/1998	14,37	29/9/1998	14,73	5/5/1999	16,12	30/11/1999	20,58	28/6/2000	18,48	24/1/2001	19,95
6/8/1997	18,15	4/3/1998	14,3	30/9/1998	14,69	6/5/1999	16,11	1/12/1999	20,56	29/6/2000	18,54	25/1/2001	19,85
7/8/1997	18,07	5/3/1998	14,22	1/10/1998	14,61	7/5/1999	15,97	2/12/1999	20,57	30/6/2000	18,27	26/1/2001	19,77
8/8/1997	18,05	6/3/1998	14,27	2/10/1998	14,58	10/5/1999	15,78	3/12/1999	20,52	3/7/2000	17,74	29/1/2001	19,59
11/8/1997	17,85	9/3/1998	14,41	5/10/1998	14,59	11/5/1999	15,7	6/12/1999	20,51	4/7/2000	17,71	30/1/2001	19,38
12/8/1997	17,99	10/3/1998	14,46	6/10/1998	14,56	12/5/1999	15,67	7/12/1999	20,51	5/7/2000	17,76	31/1/2001	19,02
13/8/1997	18	11/3/1998	14,62	7/10/1998	14,67	13/5/1999	15,6	8/12/1999	20,45	6/7/2000	17,63	1/2/2001	18,8
14/8/1997	18,07	12/3/1998	14,54	8/10/1998	14,64	14/5/1999	15,63	9/12/1999	20,4	7/7/2000	17,81	2/2/2001	18,74
15/8/1997	18,15	13/3/1998	14,48	9/10/1998	14,97	17/5/1999	15,81	10/12/1999	20,29	10/7/2000	17,47	5/2/2001	18,65
18/8/1997	18,17	16/3/1998	14,4	13/10/1998	14,95	18/5/1999	15,79	13/12/1999	20,05	11/7/2000	17,64	6/2/2001	18,49
19/8/1997	18,25	17/3/1998	14,54	14/10/1998	15	19/5/1999	15,84	14/12/1999	19,92	12/7/2000	17,8	7/2/2001	18,36
20/8/1997	18,24	18/3/1998	14,59	15/10/1998	14,99	20/5/1999	15,85	15/12/1999	19,53	13/7/2000	17,75	8/2/2001	18,23
22/8/1997	18,35	19/3/1998	14,6	16/10/1998	15,08	21/5/1999	15,81	17/12/1999	19,17	14/7/2000	17,83	9/2/2001	18,02
25/8/1997	18,48	20/3/1998	14,44	19/10/1998	15	24/5/1999	15,61	17/12/1999	19,17	17/7/2000	17,5	12/2/2001	18,04
26/8/1997	18,53	23/3/1998	14,43	20/10/1998	14,99	25/5/1999	15,98	20/12/1999	19,06	18/7/2000	17,54	13/2/2001	18,01
27/8/1997	18,64	24/3/1998	14,64	21/10/1998	14,92	26/5/1999	15,97	21/12/1999	19,06	19/7/2000	17,52	14/2/2001	18,05
28/8/1997	18,77	25/3/1998	14,81	22/10/1998	14,91	27/5/1999	15,9	22/12/1999	19,07	20/7/2000	17,6	15/2/2001	18,02
29/8/1997	18,99	26/3/1998	14,55	23/10/1998	14,84	28/5/1999	16,14	23/12/1999	19,08	21/7/2000	17,55	16/2/2001	17,9
1/9/1997	19,09	27/3/1998	14,34	26/10/1998	14,77	31/5/1999	16,09	27/12/1999	19,09	24/7/2000	17,59	19/2/2001	18,29
2/9/1997	19,21	30/3/1998	14,32	27/10/1998	14,74	1/6/1999	16,04	28/12/1999	19,13	25/7/2000	17,43	20/2/2001	18,45
3/9/1997	19,29	31/3/1998	14,18	28/10/1998	14,73	4/6/1999	16,3	30/12/1999	19,16	26/7/2000	17,53	21/2/2001	18,32
4/9/1997	19,32	1/4/1998	14,15	29/10/1998	14,73	4/6/1999	16,3	30/12/1999	19,16	27/7/2000	17,64	22/2/2001	18,21
5/9/1997	19,37	2/4/1998	14,22	30/10/1998	14,73	7/6/1999	16,52	3/1/2000	19,24	28/7/2000	17,59	23/2/2001	18,23
8/9/1997	19,46	3/4/1998	14,09	3/11/1998	14,76	8/6/1999	16,52	4/1/2000	19,24	31/7/2000	17,5	28/2/2001	17,96
9/9/1997	19,54	6/4/1998	13,82	4/11/1998	14,78	9/6/1999	16,66	5/1/2000	19,38	1/8/2000	17,56	1/3/2001	17,73
10/9/1997	19,6	7/4/1998	13,88	5/11/1998	14,79	10/6/1999	16,76	6/1/2000	19,63	2/8/2000	17,68	2/3/2001	17,59
11/9/1997	19,81	8/4/1998	13,87	6/11/1998	14,77	11/6/1999	16,84	7/1/2000	19,76	3/8/2000	17,6	5/3/2001	17,43
12/9/1997	19,64	13/4/1998	13,78	9/11/1998	14,78	14/6/1999	16,97	10/1/2000	19,75	4/8/2000	17,56	6/3/2001	17,47
15/9/1997	19,61	14/4/1998	13,95	10/11/1998	14,92	15/6/1999	16,96	11/1/2000	19,77	7/8/2000	17,48	7/3/2001	17,31
16/9/1997	19,5	15/4/1998	14,1	11/11/1998	15	16/6/1999	16,84	12/1/2000	19,86	8/8/2000	17,52	8/3/2001	17,36
17/9/1997	19,59	18/4/1998	14,16	12/11/1998	15,05	17/6/1999	16,77	13/1/2000	19,99	9/8/2000	17,55	9/3/2001	17,54
18/9/1997	19,63	17/4/1998	14,26	13/11/1998	15,01	18/6/1999	16,83	14/1/2000	20,08	10/8/2000	17,54	12/3/2001	17,64
19/9/1997	19,66	20/4/1998	14,2	16/11/1998	15,03	21/6/1999	16,79	17/1/2000	20,15	11/8/2000	17,6	13/3/2001	17,65
22/9/1997	19,62	22/4/1998	14,15	17/11/1998	15,06	22/6/1999	16,76	18/1/2000	20,25	14/8/2000	17,64	14/3/2001	17,45
23/9/1997	19,59	23/4/1998	14,28	18/11/1998	15,06	23/6/1999	16,81	19/1/2000	20,27	15/8/2000	17,73	15/3/2001	17,83
24/9/1997	19,56	24/4/1998	14,41	19/11/1998	15,05	24/6/1999	16,73	20/1/2000	20,28	16/8/2000	17,68	16/3/2001	17,96
25/9/1997	19,5	27/4/1998	14,47	20/11/1998	15,05	25/6/1999	16,83	21/1/2000	20,3	17/8/2000	17,89	19/3/2001	18,03
26/9/1997	19,46	28/4/1998	14,51	23/11/1998	14,9	28/6/1999	16,58	24/1/2000	20,31	18/8/2000	17,87	20/3/2001	17,95
29/9/1997	19,4	29/4/1998	14,52	24/11/1998	14,84	29/6/1999	16,54	26/1/2000	20,31	21/8/2000	17,91	21/3/2001	17,66
30/9/1997	19,36	30/4/1998	14,5	25/11/1998	14,76	30/6/1999	16,45	27/1/2000	20,33	22/8/2000	17,94	22/3/2001	17,66
1/10/1997	19,24	4/5/1998	14,57	26/11/1998	14,72	1/7/1999	16,27	28/1/2000	20,3	23/8/2000	17,92	23/3/2001	17,77
2/10/1997	19,16	5/5/1998	14,59	27/11/1998	14,78	2/7/1999	16,05	31/1/2000	19,99	24/8/2000	18,03	26/3/2001	17,81
3/10/1997	19,13	6/5/1998	14,68	30/11/1998	14,88	5/7/1999	15,98	1/2/2000	19,92	25/8/2000	18,08	27/3/2001	17,84
6/10/1997	18,97	7/5/1998	14,73	1/12/1998	14,86	6/7/1999	15,54	2/2/2000	19,87	28/8/2000	18,26	28/3/2001	17,79
7/10/1997	18,86	8/5/1998	14,69	2/12/1998	14,84	7/7/1999	15,64	3/2/2000	19,79	29/8/2000	18,61	29/3/2001	17,85
8/10/1997	18,72	11/5/1998	14,53	3/12/1998	14,84	8/7/1999	15,73	4/2/2000	19,59	30/8/2000	18,62	30/3/2001	17,9
9/10/1997	18,62	12/5/1998	14,58	4/12/1998	14,8	12/7/1999	16,02	7/2/2000	19,48	31/8/2000	18,79	2/4/2001	17,85
10/10/1997	18,65	13/5/1998	14,78	7/12/1998	14,79	13/7/1999	16,12	8/2/2000	19,31	1/9/2000	18,83	3/4/2001	17,9
13/10/1997	18,71	14/5/1998	14,79	8/12/1998	14,72	14/7/1999	16,11	9/2/2000	19,3	4/9/2000	18,75	4/4/2001	17,97
14/10/1997	18,75	15/5/1998	14,76	9/12/1998	14,89	15/7/1999	16,18	10/2/2000	19,42	5/9/2000	19,19	5/4/2001	17,91
15/10/1997	18,8	18/5/1998	14,79	10/12/1998	14,68	16/7/1999	15,98	11/2/2000	19,37	6/9/2000	19,02	6/4/2001	17,9
16/10/1997	18,76	19/5/1998	14,63	11/12/1998	14,64	19/7/1999	16,05	14/2/2000	19,38	8/9/2000	18,98	9/4/2001	18,01
17/10/1997	18,73	20/5/1998	14,61	14/12/1998	14,62	20/7/1999	16,28	15/2/2000	19,33	11/9/2000	18,93	10/4/2001	17,83
20/10/1997	18,74	21/5/1998	14,63	15/12/1998	14,48	21/7/1999	16,77	16/2/2000	19,19	12/9/2000	18,98	11/4/2001	17,85
21/10/1997	18,84	22/5/1998	14,63	16/12/1998	14,45	22/7/1999	17,01	17/2/2000	19,22	13/9/2000	19	12/4/2001	17,89
22/10/1997	18,91	25/5/1998	14,61	17/12/1998	14,38	23/7/1999	17,47	18/2/2000	19,17	14/9/2000	18,91	16/4/2001	17,96
23/10/1997	19,04	26/5/1998	14,46	18/12/1998	14,31	28/7/1999	17,07	21/2/2000	19,18	15/9/2000	18,87	17/4/2001	18,05
24/10/1997	19,16	27/5/1998	14,33	21/12/1998	14,29	27/7/1999	16,68	22/2/2000	19,2	18/9/2000	19,01	18/4/2001	18,03
27/10/1997	19,21	28/5/1998	14,39	22/12/1998	14,28	28/7/1999	16,44	23/2/2000	19,08	19/9/2000	18,85	19/4/2001	18,06
28/10/1997	19,27	29/5/1998	14,41	23/12/1998	14,25	29/7/1999	16,45	24/2/2000	19,08	20/9/2000	19,07	20/4/2001	18,16
29/10/1997	19,26	1/6/1998	14,29	28/12/1998	14,21	30/7/1999	16,45	25/2/2000	18,78	21/9/2000	19,1	23/4/2001	18,17
30/10/1997	19,32	2/6/1998	14,25	29/12/1998	14,16	2/8/1999	17,3	28/2/2000	18,8	22/9/2000	19,31	24/4/2001	18,12
31/10/1997	19,33	3/6/1998	14,23	30/12/1998	14,15	3/8/1999	17,83	29/2/2000	18,91	25/9/2000	19,13	25/4/2001	18,21
3/11/1997	19,4	4/6/1998	14,16	4/1/1999	14,14	4/8/1999	18,19	1/3/2000	18,85	26/9/2000	19,22	26/4/2001	18,24
4/11/1997	19,49	5/6/1998	14,03	5/1/1999	14,05	5/8/1999	18,01	2/3/2000	18,82	27/9/2000	19,27		

8/11/1997	19.62	9/6/1998	13.71	7/1/1999	13.9	9/8/1999	18.14	8/3/2000	18.38	29/9/2000	19.1	2/5/2001	18.43
7/11/1997	19.67	10/6/1998	13.75	8/1/1999	13.85	10/8/1999	18.73	9/3/2000	18.29	2/10/2000	19.17	3/5/2001	18.43
10/11/1997	19.9	12/6/1998	13.76	11/1/1999	13.83	11/8/1999	18.59	10/3/2000	18.08	3/10/2000	19.12	4/5/2001	18.43
11/11/1997	19.95	15/6/1998	13.63	12/1/1999	13.71	12/8/1999	17.5	13/3/2000	17.96	4/10/2000	19.14	7/5/2001	18.27
12/11/1997	20.01	16/6/1998	13.76	13/1/1999	13.86	13/8/1999	17.49	14/3/2000	18.18	5/10/2000	19.13	8/5/2001	18.56
13/11/1997	20.12	17/6/1998	14.07	14/1/1999	13.89	16/8/1999	17.93	15/3/2000	18.12	6/10/2000	19.15	9/5/2001	18.75
14/11/1997	20.21	18/6/1998	14.12	15/1/1999	13.91	17/8/1999	18.76	16/3/2000	18.34	9/10/2000	19.03	10/5/2001	18.84
17/11/1997	20.28	19/6/1998	14.81	18/1/1999	14.39	18/8/1999	18.83	17/3/2000	18.64	10/10/2000	19.02	11/5/2001	18.97
18/11/1997	20.44	22/6/1998	14.53	19/1/1999	15.13	19/8/1999	19.39	20/3/2000	18.22	11/10/2000	18.95	14/5/2001	19.4
19/11/1997	20.49	23/6/1998	14.84	20/1/1999	15.62	20/8/1999	19.65	21/3/2000	18.04	13/10/2000	18.87	15/5/2001	19.61
20/11/1997	20.55	24/6/1998	14.92	21/1/1999	16.97	23/8/1999	18.67	22/3/2000	18.06	16/10/2000	18.72	16/5/2001	20
21/11/1997	20.64	25/6/1998	14.86	22/1/1999	17.16	24/8/1999	19.12	23/3/2000	17.82	17/10/2000	18.83	17/5/2001	19.99
24/11/1997	20.71	26/6/1998	14.81	26/1/1999	18.82	25/8/1999	19.29	24/3/2000	17.98	18/10/2000	18.75	18/5/2001	19.97
25/11/1997	20.74	29/6/1998	14.17	27/1/1999	19.31	26/8/1999	19.14	27/3/2000	18.07	19/10/2000	18.91	21/5/2001	19.94
26/11/1997	20.87	30/6/1998	14.46	28/1/1999	19.47	27/8/1999	19.23	28/3/2000	18.13	20/10/2000	19.1	22/5/2001	19.9
27/11/1997	20.84	1/7/1998	14.46	29/1/1999	19.9	30/8/1999	19.32	29/3/2000	18.14	23/10/2000	19.02	23/5/2001	20.47
28/11/1997	20.86	2/7/1998	14.46	1/2/1999	19.31	31/8/1999	19.6	30/3/2000	18.34	24/10/2000	19.13	24/5/2001	20.79
1/12/1997	20.85	3/7/1998	14.47	2/2/1999	18.22	1/9/1999	19.83	31/3/2000	18.53	25/10/2000	19.33	25/5/2001	20.92
2/12/1997	20.87	6/7/1998	14.32	3/2/1999	18.46	2/9/1999	20.01	3/4/2000	18.68	26/10/2000	19.26	28/5/2001	20.56
3/12/1997	20.68	7/7/1998	14.64	4/2/1999	18.36	3/9/1999	20.3	4/4/2000	18.58	27/10/2000	19.13	29/5/2001	20.38
4/12/1997	20.63	8/7/1998	14.85	5/2/1999	18.54	6/9/1999	20.29	5/4/2000	18.59	30/10/2000	18.93	30/5/2001	20.33
5/12/1997	20.56	10/7/1998	14.56	8/2/1999	18.73	8/9/1999	20.97	6/4/2000	18.31	31/10/2000	19.01	31/5/2001	20.75
8/12/1997	20.59	13/7/1998	14.6	9/2/1999	18.87	9/9/1999	20.99	7/4/2000	18.38	1/11/2000	19.09	1/6/2001	21.43
9/12/1997	20.61	14/7/1998	14.5	10/2/1999	18.47	10/9/1999	20.99	10/4/2000	18.32	3/11/2000	19.31	4/6/2001	20.96
10/12/1997	20.6	15/7/1998	14.29	11/2/1999	18.44	13/9/1999	21.22	11/4/2000	18.28	8/11/2000	19.56	5/6/2001	21.52
11/12/1997	20.4	16/7/1998	14.38	12/2/1999	18.28	14/9/1999	21.26	12/4/2000	18.44	7/11/2000	19.59	6/6/2001	21.68
12/12/1997	20.2	17/7/1998	14.45	17/2/1999	18.3	15/9/1999	20.84	13/4/2000	18.7	8/11/2000	19.61	7/6/2001	21.84
15/12/1997	19.99	20/7/1998	14.14	18/2/1999	18.22	16/9/1999	20.78	14/4/2000	18.66	9/11/2000	19.42	8/6/2001	21.94
16/12/1997	19.85	21/7/1998	14.01	19/2/1999	18.07	17/9/1999	20.88	17/4/2000	18.77	10/11/2000	19.33	11/6/2001	22.32
17/12/1997	19.69	22/7/1998	13.8	22/2/1999	17.83	20/9/1999	20.77	18/4/2000	18.81	13/11/2000	19.31	12/6/2001	22.61
18/12/1997	19.52	23/7/1998	13.77	23/2/1999	18.09	21/9/1999	20.85	19/4/2000	18.85	14/11/2000	19.52	13/6/2001	22.96
19/12/1997	19.38	24/7/1998	13.71	24/2/1999	17.71	22/9/1999	20.65	20/4/2000	18.82	16/11/2000	19.49	15/6/2001	22.7
22/12/1997	19.2	27/7/1998	13.69	25/2/1999	17.7	23/9/1999	20.64	24/4/2000	19.01	17/11/2000	19.9	18/6/2001	22.82
23/12/1997	19.13	28/7/1998	13.69	26/2/1999	17.72	24/9/1999	20.64	25/4/2000	19.11	20/11/2000	20.03	19/6/2001	23.28
26/12/1997	19.08	29/7/1998	13.75	1/3/1999	17.92	27/9/1999	20.77	26/4/2000	19.11	21/11/2000	19.9	20/6/2001	23.43
29/12/1997	18.89	30/7/1998	13.84	2/3/1999	18.13	28/9/1999	20.95	27/4/2000	19.04	22/11/2000	20.03	21/6/2001	22.25
30/12/1997	18.76	31/7/1998	13.77	3/3/1999	18.51	29/9/1999	20.94	28/4/2000	18.81	23/11/2000	20.15	22/6/2001	21.35
2/1/1998	18.75	3/8/1998	13.68	4/3/1999	18.31	30/9/1999	20.95	2/5/2000	19.62	24/11/2000	20.78	25/6/2001	20.91
5/1/1998	18.67	4/8/1998	13.57	5/3/1999	17.3	1/10/1999	21.1	3/5/2000	19.97	27/11/2000	21.18	26/6/2001	21.15
6/1/1998	18.55	5/8/1998	13.51	8/3/1999	17.2	4/10/1999	21.03	4/5/2000	19.96	28/11/2000	21.22	27/6/2001	21.68
7/1/1998	18.45	6/8/1998	13.43	9/3/1999	17.06	5/10/1999	21.02	5/5/2000	19.72	29/11/2000	21.65	28/6/2001	21.38
8/1/1998	18.2	7/8/1998	13.4	10/3/1999	16.76	6/10/1999	21.34	8/5/2000	19.37	30/11/2000	22.05	29/6/2001	22.67
9/1/1998	18.07	10/8/1998	13.52	11/3/1999	16.96	7/10/1999	21.24	9/5/2000	19.22	1/12/2000	22.37	2/7/2001	23.22
12/1/1998	17.86	11/8/1998	13.58	12/3/1999	17.01	8/10/1999	22.1	10/5/2000	19.82	4/12/2000	22.23	3/7/2001	23.59
13/1/1998	17.44	12/8/1998	13.6	15/3/1999	16.76	11/10/1999	22.12	11/5/2000	19.82	5/12/2000	22.04	4/7/2001	23.66
14/1/1998	17.33	13/8/1998	13.64	16/3/1999	16.85	13/10/1999	22.37	12/5/2000	19.77	6/12/2000	21.96	5/7/2001	24.43
15/1/1998	17.02	14/8/1998	13.69	17/3/1999	17.24	14/10/1999	22.42	15/5/2000	19.08	7/12/2000	22.03	6/7/2001	25.25
16/1/1998	16.75	17/8/1998	13.75	18/3/1999	17.19	15/10/1999	22.39	16/5/2000	19.11	8/12/2000	22.01	10/7/2001	25.93
19/1/1998	16.55	18/8/1998	13.84	19/3/1999	17.28	18/10/1999	22.44	17/5/2000	19.03	11/12/2000	22.06	11/7/2001	27.03
20/1/1998	16.41	19/8/1998	14.2	22/3/1999	17.59	19/10/1999	22.5	18/5/2000	19.85	12/12/2000	22.1	12/7/2001	27.58
21/1/1998	16.33	20/8/1998	14.26	23/3/1999	17.72	20/10/1999	22.5	19/5/2000	19.93	13/12/2000	22.07	13/7/2001	27.67
22/1/1998	16.25	21/8/1998	14.18	24/3/1999	17.81	21/10/1999	22.4	22/5/2000	20.1	14/12/2000	21.97	16/7/2001	27.24
23/1/1998	16.3	24/8/1998	14.21	25/3/1999	17.39	22/10/1999	22.26	23/5/2000	19.95	15/12/2000	21.87	17/7/2001	27.6
26/1/1998	16.32	25/8/1998	14.26	26/3/1999	17.32	25/10/1999	22.07	24/5/2000	19.73	18/12/2000	21.99	18/7/2001	26.3
27/1/1998	16.33	26/8/1998	14.32	29/3/1999	16.95	26/10/1999	21.95	25/5/2000	19.42	19/12/2000	21.85	19/7/2001	26.43
28/1/1998	16.31	27/8/1998	14.26	30/3/1999	16.75	27/10/1999	21.87	26/5/2000	19.44	20/12/2000	21.79	20/7/2001	25.41
29/1/1998	16.18	28/8/1998	14.2	31/3/1999	16.15	28/10/1999	21.78	29/5/2000	19.35	21/12/2000	21.88	23/7/2001	24.94
30/1/1998	16.13	31/8/1998	14.19	5/4/1999	15.95	29/10/1999	21.59	30/5/2000	18.77	22/12/2000	21.72	24/7/2001	25.53
2/2/1998	16.12	1/9/1998	14.22	8/4/1999	15.94	1/11/1999	21.46	31/5/2000	18.76	26/12/2000	21.81	25/7/2001	26.31
3/2/1998	16.18	2/9/1998	14.27	7/4/1999	15.78	3/11/1999	21.3	1/6/2000	18.87	27/12/2000	21.72	26/7/2001	27.18
4/2/1998	16.19	3/9/1998	14.31	8/4/1999	15.64	4/11/1999	21.08	2/6/2000	18.98	28/12/2000	21.54	27/7/2001	27.28
5/2/1998	16.29	4/9/1998	14.39	9/4/1999	15.72	5/11/1999	21.07	5/6/2000	18.77	2/1/2001	21.53	30/7/2001	26.41
8/2/1998	16.34	8/9/1998	14.39	12/4/1999	15.78	8/11/1999	21.07	6/6/2000	18.76	3/1/2001	21.36	31/7/2001	26.5
9/2/1998	16.28	9/9/1998	14.4	13/4/1999	15.49	9/11/1999	21.1	7/6/2000	18.67	4/1/2001	21.41	1/8/2001	27.3
10/2/1998	16.25	10/9/1998	14.47	14/4/1999	15.6	10/11/1999	20.79	8/6/2000	18.77	5/1/2001	21.35	2/8/2001	26.82
11/2/1998	16.18	11/9/1998	14.6	15/4/1999	15.54	11/11/1999	20.78	9/6/2000	18.76	8/1/2001	21.39	3/8/2001	26.41
12/2/1998	15.83	14/9/1998	14.74	16/4/1999	15.57	12/11/1999	20.71	12/6/2000	18.34	9/1/2001	21.43	6/8/2001	26.39
13/2/1998	15.58	15/9/1998	14.78	19/4/1999	15.6	16/11/1999	20.56	13/6/2000	18.47	10/1/2001	21.26	7/8/2001	26.59
16/2/1998	15.44	16/9/1998	14.76	20/4/1999	16.01	17/11/1999	20.82	14/6/2000	18.47	11/1/2001	20.75	8/8/2001	27.02
17/2/1998	15.37	17/9/1998	14.78	22/4/1999	15.85	18/11/1999	20.58	15/6/2000	18.8	12/1/2001	20.74	9/8/2001	27.38
18/2/1998	15.19	18/9/1998	14.78	23/4/1999	15.79	19/11/1999	20.57	16/6/2000	18.57	15/1/2001	20.69	10/8/2001	28.01
				26/4/1999	15.76								
13/8/2001	28.29	14/3/2002	21.33	7/10/2002	43.58	8/5/2003	36.22						
14/8/20													

20/8/2001	27.85	21/3/2002	21.64	14/10/2002	46.97	15/5/2003	37.4
21/8/2001	28.02	22/3/2002	21.82	15/10/2002	47.82	16/5/2003	37.75
22/8/2001	27.97	25/3/2002	21.96	16/10/2002	48.04	19/5/2003	37.84
23/8/2001	27.63	26/3/2002	22.01	17/10/2002	48.46	20/5/2003	38.11
24/8/2001	27.73	27/3/2002	22.07	18/10/2002	47.91	21/5/2003	37.99
27/8/2001	27.85	28/3/2002	22.1	21/10/2002	48.88	22/5/2003	37.61
28/8/2001	27.91	1/4/2002	22.15	22/10/2002	48.92	23/5/2003	37.1
29/8/2001	27.99	2/4/2002	22.13	23/10/2002	48.82	26/5/2003	36.68
30/8/2001	28.05	3/4/2002	21.98	24/10/2002	48.37	27/5/2003	37.61
31/8/2001	28.21	4/4/2002	21.78	25/10/2002	47.39	28/5/2003	37.67
3/9/2001	28.35	5/4/2002	21.68	28/10/2002	46.3	29/5/2003	37.33
4/9/2001	28.4	8/4/2002	21.53	29/10/2002	47.4	30/5/2003	37.39
5/9/2001	28.48	9/4/2002	21.35	30/10/2002	47.23	2/6/2003	36.9
6/9/2001	28.49	10/4/2002	21	31/10/2002	46.37	3/6/2003	36.73
10/9/2001	28.55	11/4/2002	21.05	1/11/2002	45.91	4/6/2003	36.17
11/9/2001	28.5	12/4/2002	21.29	4/11/2002	45.27	5/6/2003	36.25
12/9/2001	28.61	15/4/2002	21.42	5/11/2002	45.92	6/6/2003	36.32
13/9/2001	29.38	16/4/2002	21.91	6/11/2002	46.14	9/6/2003	36.58
14/9/2001	29.57	17/4/2002	22.52	7/11/2002	46.56	10/6/2003	36.72
17/9/2001	29.51	18/4/2002	22.78	8/11/2002	46.25	11/6/2003	36.8
18/9/2001	29.59	19/4/2002	23.07	11/11/2002	45.14	12/6/2003	36.85
19/9/2001	29.55	22/4/2002	23.37	12/11/2002	45.91	13/6/2003	36.71
20/9/2001	29.64	23/4/2002	23.63	13/11/2002	46.77	16/6/2003	36.14
21/9/2001	29.77	24/4/2002	24.19	14/11/2002	47.23	17/6/2003	36.41
24/9/2001	29.78	25/4/2002	24.27	18/11/2002	46.77	18/6/2003	37.11
25/9/2001	29.87	26/4/2002	23.92	19/11/2002	47.25	20/6/2003	37.45
26/9/2001	29.79	29/4/2002	23.48	20/11/2002	47.44	23/6/2003	37.53
27/9/2001	29.68	30/4/2002	23.19	21/11/2002	47.54	24/6/2003	37.76
28/9/2001	29.31	2/5/2002	23.03	22/11/2002	47.47	25/6/2003	37.75
1/10/2001	29.09	3/5/2002	23.04	25/11/2002	47.18	26/6/2003	37.77
2/10/2001	28.98	6/5/2002	23.27	26/11/2002	47.85	27/6/2003	37.45
3/10/2001	29.1	7/5/2002	23.25	27/11/2002	48.54	30/6/2003	36.64
4/10/2001	29.16	8/5/2002	23.57	28/11/2002	48.61	1/7/2003	36.64
5/10/2001	29.38	9/5/2002	24.24	29/11/2002	49.1	2/7/2003	36.66
8/10/2001	29.49	10/5/2002	24.67	2/12/2002	49.5	3/7/2003	36.68
9/10/2001	29.66	13/5/2002	25.1	3/12/2002	49.49	4/7/2003	36.64
10/10/2001	29.78	14/5/2002	25.19	4/12/2002	49.57	7/7/2003	36.56
11/10/2001	29.96	15/5/2002	25.45	5/12/2002	49.96	8/7/2003	36.72
15/10/2001	29.61	16/5/2002	25.6	6/12/2002	49.75	10/7/2003	36.83
16/10/2001	29.53	17/5/2002	25.59	9/12/2002	49.96	11/7/2003	37.31
17/10/2001	29.42	20/5/2002	25.9	10/12/2002	50.25	14/7/2003	37.07
18/10/2001	29.59	21/5/2002	26.34	11/12/2002	50.48	15/7/2003	36.36
19/10/2001	29.58	22/5/2002	26.55	12/12/2002	50.45	16/7/2003	36.09
22/10/2001	29.54	23/5/2002	26.81	13/12/2002	50.51	17/7/2003	35.58
23/10/2001	29.7	24/5/2002	26.84	16/12/2002	50.15	18/7/2003	35.07
24/10/2001	30.11	27/5/2002	26.84	17/12/2002	49.57	21/7/2003	34.57
25/10/2001	30.17	28/5/2002	27.23	18/12/2002	48.14	22/7/2003	34.77
28/10/2001	30.32	29/5/2002	27.44	19/12/2002	47.68	23/7/2003	34.74
29/10/2001	30.36	31/5/2002	27.89	20/12/2002	46.51	24/7/2003	34.2
30/10/2001	30.3	3/6/2002	28.13	23/12/2002	46.51	25/7/2003	34.03
31/10/2001	30.22	4/6/2002	29.11	26/12/2002	47.08	28/7/2003	34.49
1/11/2001	30.33	5/6/2002	29.48	27/12/2002	47.07	29/7/2003	34.48
5/11/2001	30.6	6/6/2002	29.99	30/12/2002	46.75	30/7/2003	34.43
6/11/2001	30.59	7/6/2002	29.83	2/1/2003	45.88	31/7/2003	34.26
7/11/2001	30.59	10/6/2002	28.75	3/1/2003	46	1/8/2003	34.89
8/11/2001	30.3	11/6/2002	28.97	6/1/2003	45.99	4/8/2003	35.61
9/11/2001	29.91	12/6/2002	29.23	7/1/2003	45.81	5/8/2003	34.76
12/11/2001	29.85	13/6/2002	28.95	8/1/2003	46.08	6/8/2003	34.94
13/11/2001	29.87	14/6/2002	28.93	9/1/2003	46.08	7/8/2003	34.69
14/11/2001	29.89	17/6/2002	28.71	10/1/2003	45.95	8/8/2003	34.78
16/11/2001	29.8	18/6/2002	28.44	13/1/2003	45.62	11/8/2003	35.03
19/11/2001	30.07	19/6/2002	28.37	14/1/2003	42.91	12/8/2003	36.09
20/11/2001	29.93	20/6/2002	28.87	15/1/2003	41.85	13/8/2003	36.43
21/11/2001	29.88	21/6/2002	29.49	16/1/2003	40.69	14/8/2003	36.08
22/11/2001	29.84	24/6/2002	29.95	17/1/2003	40.75	15/8/2003	35.72
23/11/2001	29.78	25/6/2002	29.94	20/1/2003	40.58	18/8/2003	37.33
26/11/2001	29.54	26/6/2002	30.36	21/1/2003	40.58	19/8/2003	36.78
27/11/2001	29.05	27/6/2002	30.75	22/1/2003	40.92	20/8/2003	37.5
28/11/2001	28.92	28/6/2002	31	23/1/2003	40.96	21/8/2003	37.92
29/11/2001	28.77	1/7/2002	31.46	24/1/2003	41.03	22/8/2003	38.14
30/11/2001	28.78	2/7/2002	31.75	27/1/2003	41.71	25/8/2003	38.21
3/12/2001	28.72	3/7/2002	31.23	28/1/2003	42.2	28/8/2003	38.28
4/12/2001	28.7	4/7/2002	31.18	29/1/2003	42.28	27/8/2003	38.08
5/12/2001	28.58	5/7/2002	31.59	30/1/2003	42.24	28/8/2003	38.09
8/12/2001	28.53	8/7/2002	31.68	31/1/2003	42.31	29/8/2003	38.27
7/12/2001	28.43	10/7/2002	32.33	3/2/2003	42.8	1/9/2003	38.17
10/12/2001	28.15	11/7/2002	32.24	4/2/2003	42.71	2/9/2003	37.85
11/12/2001	27.73	12/7/2002	32.21	5/2/2003	42.57	3/9/2003	38.04
12/12/2001	27.37	15/7/2002	33.18	6/2/2003	42.75	4/9/2003	38.42

13/12/2001	26.94	16/7/2002	33.63	7/2/2003	42.59	5/9/2003	38.25
14/12/2001	26.79	17/7/2002	33.68	10/2/2003	42.76	8/9/2003	38.19
17/12/2001	26.72	18/7/2002	33.7	11/2/2003	42.65	9/9/2003	38.32
18/12/2001	26.59	19/7/2002	33.43	12/2/2003	42.73	10/9/2003	38.19
19/12/2001	26.37	22/7/2002	34.12	13/2/2003	43.13	11/9/2003	38.9
20/12/2001	26.29	23/7/2002	33.52	14/2/2003	43.36	12/9/2003	39.18
21/12/2001	26.21	24/7/2002	33.8	17/2/2003	44.11	15/9/2003	39.01
26/12/2001	25.84	25/7/2002	33.43	18/2/2003	43.76	16/9/2003	38.78
27/12/2001	25.84	26/7/2002	33.39	19/2/2003	42.97	17/9/2003	39.31
28/12/2001	25.69	29/7/2002	33.33	20/2/2003	42.45	18/9/2003	39.5
2/1/2002	25.61	30/7/2002	34.85	21/2/2003	42.55	19/9/2003	40.38
3/1/2002	25.58	31/7/2002	36.58	24/2/2003	42.35	22/9/2003	41.08
4/1/2002	25.49	1/8/2002	36.83	25/2/2003	42.92	23/9/2003	41.7
7/1/2002	25.36	2/8/2002	34.68	26/2/2003	42.7	24/9/2003	41.7
8/1/2002	25.1	5/8/2002	33.87	27/2/2003	42.36	25/9/2003	41.88
9/1/2002	25.11	8/8/2002	35.28	28/2/2003	41.87	26/9/2003	42.05
10/1/2002	24.96	7/8/2002	34.52	5/3/2003	40.75	29/9/2003	42.89
11/1/2002	25.12	8/8/2002	33.1	6/3/2003	39.98	30/9/2003	42.96
14/1/2002	25.24	9/8/2002	34.02	7/3/2003	39.18	1/10/2003	43.3
15/1/2002	25.2	12/8/2002	36.86	10/3/2003	39.45	2/10/2003	43.48
16/1/2002	25.26	13/8/2002	37.85	11/3/2003	39.45	3/10/2003	43.04
17/1/2002	25.26	14/8/2002	38.38	12/3/2003	39.56	6/10/2003	42.71
18/1/2002	25.14	15/8/2002	38.85	13/3/2003	39.61	7/10/2003	42.77
21/1/2002	25.1	16/8/2002	37.79	14/3/2003	39.76	8/10/2003	42.69
22/1/2002	25.18	19/8/2002	36.51	17/3/2003	39.66	9/10/2003	42.71
23/1/2002	25.21	20/8/2002	35.6	18/3/2003	40.02	10/10/2003	44.49
24/1/2002	25.21	21/8/2002	35.69	19/3/2003	40.17	13/10/2003	45.09
25/1/2002	25.12	22/8/2002	35.71	20/3/2003	40.52	14/10/2003	45.1
28/1/2002	24.88	23/8/2002	35.92	21/3/2003	40.33	15/10/2003	44.85
29/1/2002	24.79	26/8/2002	36.06	24/3/2003	40.11	16/10/2003	45.09
30/1/2002	24.67	27/8/2002	36.56	25/3/2003	39.76	17/10/2003	45.55
31/1/2002	24.6	28/8/2002	37.13	26/3/2003	39.19	20/10/2003	46.43
1/2/2002	24.52	29/8/2002	37.31	27/3/2003	39.3	21/10/2003	46.39
4/2/2002	24.39	30/8/2002	36.68	28/3/2003	39.22	22/10/2003	47.02
5/2/2002	24.26	2/9/2002	36.47	31/3/2003	38.47	23/10/2003	47.9
6/2/2002	23.95	3/9/2002	37.44	1/4/2003	38.16	24/10/2003	48.22
7/2/2002	23.72	4/9/2002	37.91	2/4/2003	38.3	27/10/2003	49.24
8/2/2002	23.57	5/9/2002	38.56	3/4/2003	38.1	28/10/2003	49.09
13/2/2002	22.75	6/9/2002	39.01	4/4/2003	37.84	29/10/2003	49.58
14/2/2002	22.25	9/9/2002	39.61	7/4/2003	37.52	30/10/2003	49.85
15/2/2002	22.02	10/9/2002	40	8/4/2003	37.56	31/10/2003	49.31
18/2/2002	21.89	11/9/2002	41.03	9/4/2003	37.7	3/11/2003	49.52
19/2/2002	21.71	12/9/2002	40.65	10/4/2003	38.04	4/11/2003	49.16
20/2/2002	21.62	13/9/2002	40.54	11/4/2003	37.95	5/11/2003	48.5
21/2/2002	21.84	16/9/2002	40.92	14/4/2003	37.93	6/11/2003	48.76
22/2/2002	21.59	17/9/2002	41.03	15/4/2003	37.52	7/11/2003	48.71
25/2/2002	21.3	18/9/2002	42.34	16/4/2003	37.17	10/11/2003	48.2
26/2/2002	20.85	19/9/2002	42.59	17/4/2003	36.94	11/11/2003	48.85
27/2/2002	20.55	20/9/2002	42.88	22/4/2003	37.12	12/11/2003	49.95
28/2/2002	20.38	23/9/2002	43.2	23/4/2003	36.54	13/11/2003	49.46
1/3/2002	20.35	24/9/2002	44.79	24/4/2003	38.31	14/11/2003	49.54
4/3/2002	20.23	25/9/2002	45.47	25/4/2003	36.06	17/11/2003	49.61
5/3/2002	20.14	26/9/2002	46.45	28/4/2003	35.99	18/11/2003	48.81
6/3/2002	20.18	27/9/2002	47.16	29/4/2003	35.57	19/11/2003	47.57
7/3/2002	20.58	30/9/2002	46.51	30/4/2003	35.64	20/11/2003	47.12
8/3/2002	20.69	1/10/2002	45.19	2/5/2003	36.1	21/11/2003	47.01
11/3/2002	21	2/10/2002	43.06	5/5/2003	37.12	24/11/2003	46.2
12/3/2002	21.05	3/10/2002	42.6	6/5/2003	36.94	25/11/2003	45.8
13/3/2002	21.11	4/10/2002	43.03	7/5/2003	36.94	26/11/2003	46.22
						27/11/2003	46.2
						28/11/2003	46.75

## 6.8 Programa MATLAB para o algoritmo de aprendizagem recorrente em tempo real - ARTR

Este programa é baseado nos trabalhos de Williams e Zipser (1989) e de Georgios C. Anagnostopoulos (2002).

```
load ibovespa965_200
% Numero de neuronios na camada intermediaria
H = 2;
% Estado inicial - Um neuronio de saida e H unidades intermediarias
InitialState = [0.0 zeros(1,H)];
% Uso da funcao de ativacao logistica
sigmoid_type = 0;
alpha = 1.0;
learning_rate = 0.005;
epochs = 1000;
epsilon = 0.0001;
Inputs = Inputs(191:200);
DesiredOutputs = DesiredOutputs(191:200);
[WeightMatrix]=rtrl_train(Inputs,DesiredOutputs,H,InitialState,sigmoid_type,alpha,learning_rate,epochs,epsilon);
[Outputs] = rtrl_perf(Inputs,InitialState,H,WeightMatrix,sigmoid_type,alpha);
[samples columns] = size(Inputs);
t = (1:samples)';
figure
hold on
plot(t,DesiredOutputs,'b-.');
plot(t,Outputs,'r');
legend('Valor Desejado','Valor Previsto',0)
xlabel('Janela de Previsao')
ylabel('Serie d(t) versus Serie y(t)')
```

```

figure
hold on
plot(t, DesiredOutputs - Outputs)
legend('Erro de previsao',0)
xlabel('Janela de Previsao')
ylabel('Erro de previsao, e(t)')
figure
hold on
plot(t, Outputs)
legend('Valores Previstos',0)
xlabel('Janela de Previsao')
ylabel('Valores Previstos, e(t)')
figure
hold on
histfit(DesiredOutputs - Outputs)
e = DesiredOutputs - Outputs
perf_mae = mae(e)
% perf2_mse = mse(e)
rmse = sqrt(mse(e))
% Calculo de valores necesarios para obter o Coeficiente de Desigualdade
% de Theil
eq = (DesiredOutputs - Outputs)*(DesiredOutputs - Outputs)'
dq = (DesiredOutputs)*(DesiredOutputs)'
oq = (Outputs)*(Outputs)'
deq = diag(eq)
ddq = diag(dq)
doq = diag(oq)
% Coeficiente de Desigualdade de Theil
theil = (sqrt(sum(diag(eq))/10))/((sqrt(sum(diag(oq))/10))+sqrt(sum(diag(dq))/10)))

```

## Referências Bibliográficas 7

ANAGNOSTOLOUPOS, G. C. Real Time Recurrent Learning, University of Central Florida, 2002.

AKAIKE, H. A new look at the statistical model identification, IEEE Transactions on Automatic Control, AC-19, 716-723, 1974.

AZOFF, E. M. Neural network time series forecasting of financial markets, John Wiley & Sons Ltd., 1994.

BOLLERSLEV, T. On the correlation structure for the generalized ARCH process, Journal of Time Series Analysis, 9, 121-131, 1988.

BOLLERSLEV, T., CHOU, R. E KRONER, K. ARCH modeling in finance: a review of the theory and empirical evidence, Journal of Econometrics, 52, 5-59, 1992.

BOLLERSLEV, T., ENGLE, R. F., E NELSON, D. B. ARCH models, Handbook of Econometrics, 63, 37-60, 1994.

BOX, G. E. P. E JENKINS, G. M. Time series analysis: forecasting and control, San Francisco: Holden Day, 1976.

CAMPOLUCCI, P. A circuit theory approach to recurrent neural network architectures and learning methods, Tesi di Dottorato, Università Degli Studi di Bologna, 1998.

DAY, T. E LEWIS, C. Stock market volatility and the information content o stock index options, Journal of Econometrics, 52, 267-287, 1992.

De GOOIJER, J. G., ABRAHAM, B., GOULD, A. E ROBINSON, L. Methods for determining the order of an autoregressive-moving average process: a survey, International Statistical Review, 85, 301-329, 1985.

- DIEBOLD, F. X. E MARIANO, R. S. Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13, 253-263, 1995.
- DORFFNER, G. *Neural Networks for time series processing*, University of Vienna, 1996.
- DUDA, R. e HART, P. *Pattern Classification and Scene Analysis*. Segunda edição, John Wiley, New York, 1993.
- ELMAN, J. L. Finding structure in time, *Cognitive Science*, vol. 14, pp. 179-211, 1990.
- ENDERS, W. *Applied econometric time series*, Wiley, 1995.
- ENGLE, R. F., LILIEN, D.M., e ROBINS, R. P. Estimating Time-Varying Risk Premia in the Term Structure: The ARCH-M Model, *Econometrica*, 55, 391-407.
- ENGLE, R. F. e PATTON, A. J. What good is a volatility model? *Quantitative Finance*, Volume 1, pag. 237-245, 2001.
- ERSOY, K. E HONG, D. Parallel, self-organizing, hierarchical neural networks, pp. 167-178, 1990.
- FORSGREN, A. E KLING, R. An Implementation of Recurrent Neural Networks for Prediction and Control of Nonlinear Dynamic Systems, março, 2003.
- FRANSES, P. H. *Time series models for business and economic forecasting*, Cambridge University Press, 1998.
- GALINDO-FLORES, J. *Statistical and Machine Learning Frameworks for Economics: Analysis of Error Curves and Applications to Derivatives Pricing and Credit Risk Assessment*, Harvard University, 1998.
- GRANGER, C. W. J., KING, M. L. E WHITE, H. Comments on testing economic theories and the use of model selection criteria, *Journal of Econometrics*, 67, 173-188, 1995.
- GRANGER, C. W. J. e NEWBOLD, P. Forecasting transformed time series, *Journal of the Royal Society B*, 38, 189-203, 1976.
- HAMILTON, J. D. *Time series analysis*, Princeton: Princeton University Press, 1994.
- HARVEY, A. C. *Forecasting, structural time series models and the Kalman filter*, Cambridge: Cambridge University Press, 1989.
- HAYKIN, S. *Adaptive Filter Theory*, Prentice Hall, 1996.
- HAYKIN, S. *Redes neurais: princípios e prática*, Bookman, Porto Alegre, 2001.

- IMAN, R. L. E CONOVER, W. J. The use of the rank transform in regression, *Technometrics*, vol. 21, no. 4, novembro 1979.
- JHEE, W. C. E LEE, J. K. Performance of neural networks in managerial forecasting, 1993.
- JORDAN, M. I. Attractor dynamics and parallelism in a connectionist sequential machine, 1986.
- KALMAN, R. E. A new approach to linear filtering and prediction problems, *Transactions of the ASME, Journal of Basic Engineering*, vol. 82, pp. 35-45, 1960.
- LANG, K. J. E HINTON, G. E., The development of the time-delay neural network architecture for speech recognition, Technical Report, Carnegie-Mellon University, Pittsburgh, PA, 1988.
- LIPPMAN, R. P. An introduction to computing with neural nets, *IEEE ASSP Magazine*, vol.4, pp. 4-22, 1987.
- LJUNG, G. E BOX, G. On a measure of lack of fit in time series models, *Biométrica* 65, 297-303, 1978.
- MAYBECK, P. S. Stochastic models, estimation, and control, vol. 1, New York: Academic Press, 1979.
- McNELIS, P. D. Financial Forecasting with Neural Networks, Georgetown University, 2002.
- MOZER, M. C. A focused back-propagation algorithm for temporal pattern recognition, University of Toronto, 1988.
- NELSON, C. R. The interpretation of  $R^2$  in autoregressive moving average time series models, *American Statistician*, 30, 175-180, 1976.
- OLIVEIRA, M. A. E FÁVERO, L. P. L. Uma Breve Descrição de algumas Técnicas para Análise de Séries Temporais: Séries de Fourier, Wavelets, Arima, Modelos Estruturais para Séries de Tempo e Redes Neurais, *Anais do VI Semead, FEA-USP*, 2002.
- OLIVEIRA, M. A. Aplicação do algoritmo de rede neural de aprendizagem recorrente em tempo real (RTRL) para previsão da série do IBOVESPA, 10ª. Escola de Séries Temporais e Econometria, 2003.
- OLIVEIRA, M. A. e SIQUEIRA, J. O. Regressão linear, regressão não-linear, classificação de padrão e redes neurais: uma visão unificada, Working Paper, FEA/USP, 2003.
- PAGAN, A. R. E SCHWERT, G. W. Alternative models for conditional stock market volatility, *Journal of Econometrics*, 45, 267-290, 1990.

- PATEL, G. S. Modeling Linear dynamics with extended Kalman filter trained recurrent multilayer perceptrons, McMaster University, 2000.
- PATI, Y. C. e KRISHNAPRASAD, P. S. Analysis and synthesis of feedforward neural networks using discrete affine wavelet transformations, Technical Research Report, University of Maryland, 1993
- PEREIRA, B. B. E RODRIGUES, C. V. S. Redes Neurais em Estatística, 13°. SINAPE, Caxambu, 1998.
- PUSKORIUS, G. V. E FELDKAMP, L. A. Decoupled extended Kalman filter training of feedforward layered networks, Proceeding of the International Joint Conference on Neural Networks, vol.1, pp. 771-777, Seattle, WA 1991.
- ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain, Psychological Review, vol. 65, pp. 386-408.
- RUUD, P. An introduction to classical econometric theory, Oxford University Press, Oxford, 2000.
- SCHWARZ, G. Estimating the dimension of a model, Annals of Statistics, 6, 461-464, 1978.
- SHUMWAY, R. H. E STOFFER, D. S. Time series analysis and its applications, Springer-Verlag New York, In. 2000.
- SINGHAL, S. E WU, L. Training feed-forward networks with extended Kalman filter, IEEE Transactions on Acoustics, Speech, and Signal Processing, pp. 1187-1190, 1989.
- SZILAS, N. Les réseaux récurrents supervisés: une revue critique, fevereiro, 1997.
- TSAY, R. S. Analysis of financial time series, Wiley Series in Probability and Statistics, março, 2002.
- WAIBEL, A., HANAZAWA, T., HINTON, G., SHIKANO, K. E LANG, K. J. Phoneme recognition using time-delay neural networks, IEEE Transactions on Acoustics, Speech, and Signal Processing, 1989.
- WAN, E. A. Time series prediction by using a connectionist network with internal delay lines, 1994.
- WAN, E. A. Temporal backpropagation for FIR neural networks, IEEE International Joint Conference on Neural Networks, vol. I, pp. 575-580, San Diego, CA, 1990.

WARNER, B. e MISRA, M. Understanding Neural Networks as Statistical Tools, The American Statistician, vol. 50, n<sup>o</sup> 4, novembro, 1996.

WERBOS, P. J. Backpropagation through time: what it does and how to do it, Proceedings of the IEEE, vol. 78, pp. 1550-1560, 1990.

WILLIAMS, R. J. Training recurrent networks using the extended Kalman filter, Proceeding of the International Joint Conference on Neural Networks, vol. 4, pp. 241-246, Baltimore 1992.

WILLIAMS, R. J. E PENG, J. An efficient gradient-based algorithm for on-line training of recurrent network trajectories, Neural Computation, vol. 2, pp. 490-501, 1990.

WILLIAMS, R. J. E ZIPSER, D. A learning algorithm for continually running fully recurrent neural networks, Neural Computation, vol. 1, pp. 270-280, 1989.

WILLIAMS, R. J. E ZIPSER, D. Gradient-based learning algorithms for recurrent networks and their computational complexity, 1995.

YIM, J. Previsão de Séries de Tempo: Modelos ARIMA, Modelos Estruturais e Redes Neurais Artificiais. Dissertação (Economia) - Universidade de São Paulo, 2001.

WAN, E.A. e MERWE, R. A new approach for filtering nonlinear systems. Kalman Filtering and Neural Networks. Wiley, 2001.