



PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA

# MÉTODO AUXILIAR, BASEADO EM AUTOVALORES, PARA SELEÇÃO DE ATRASOS DA SAÍDA EM MODELOS LINEARES E NÃO-LINEARES

**Bruno Eduardo Lopes**

Dissertação submetida à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia do Centro Universitário do Leste de Minas Gerais, como parte dos requisitos necessários à obtenção do grau de Mestre em Engenharia Industrial.

**Área de Concentração:** Processos Industriais

**ORIENTADOR:** Dr. Marcelo Vieira Corrêa

**CO-ORIENTADOR:** Dr. Roselito de Albuquerque Teixeira

Coronel Fabriciano

2008

L864m Lopes, Bruno Eduardo.

Método auxiliar, baseado em autovalores, para seleção de atrasos da saída em modelos lineares e não-lineares / Bruno Eduardo Lopes. — 2008.

57 f.

Dissertação (mestrado) — Centro Universitário do Leste de Minas Gerais - Unileste-MG, 2008

“Orientação: Marcelo Vieira Corrêa.”

1. Autovalores – Método – Análise. I. Título.

CDU – 519.614

*Aos meus pais Vicente Lopes e Maria Eunice Lopes,  
às minhas irmãs Kátia Lopes e Karina Lopes e  
à minha esposa Patrícia Lopes*

# Agradecimentos

Em primeiro lugar, à DEUS por me amparar nos momentos difíceis, me dar força interior para superar as dificuldades, mostrar os caminho nas horas incertas e me suprir em todas as minhas necessidades.

Aos meus pais, Vicente Lopes Valente e Maria Eunice Lopes, pela educação que recebi, pelos ensinamentos de vida e por terem me dado todo o amor, carinho, atenção e condições necessárias para que eu pudesse chegar aonde cheguei.

Às minhas irmãs, Kátia Cristina Lopes e Karina Fernanda Lopes, por todos os momentos que passamos juntos, pela amizade e carinho.

À Patrícia Lopes, minha esposa, minha sincera gratidão pelo carinho, incentivo e compreensão mesmo nos momentos mais difíceis da minha vida.

Aos meus orientadores, Prof. Dr. Marcelo Viera Corrêa e Prof. Dr. Roselito Albuquerque, pela amizade, pela dedicação e pelo constante incentivo sempre indicando a direção a ser tomada durante toda a orientação.

Aos meus familiares que sempre me deram amor e força, valorizando meus potenciais.

A todos os meus amigos e amigas que sempre estiveram presentes me aconselhando e incentivando com carinho e dedicação.

*"A mente que se abre a uma nova idéia,  
jamais voltará ao seu tamanho original."*

***Albert Einstein***

# Resumo

Uma dificuldade no processo de identificação de sistemas é obter a melhor ordem de um modelo matemático. Como a melhor ordem para os modelos não é antecipadamente conhecida, torna-se necessário desenvolver critérios a fim de escolher dentre as ordens possíveis, a ideal. Com objetivo de introduzir uma solução para este problema é demonstrado um novo método chamado Método de Análise dos Autovalores (MAA). O MAA é utilizado, nesta dissertação, para verificar a existência do sobredimensionamento da ordem de um modelo linear e para determinar o número de atrasos de saída de uma rede neural por meio do princípio da análise dos autovalores. Para elaboração do método, uma investigação entre a ordem do modelo e a posição dos autovalores foi realizada. Como resultado, quando a ordem é sobredimensionada, pelo menos um autovalor aparece no lado esquerdo do plano  $Z$ , se a massa de dados utilizada é livre de ruído. Em modelos lineares, os resultados desta investigação sugerem a integração do MAA com os métodos baseados no critério de informação, melhorando o desempenho na determinação da ordem, principalmente quando os dados utilizados estão contaminados com ruído. Para modelos neurais, o desempenho do MAA foi testado utilizando dados obtidos de um processo térmico e dados simulados a partir de um caso teórico. O MAA indicou corretamente o número de atrasos de saída da rede neural em ambos os casos.

# Abstract

A difficulty in the identification systems process is to get the best order of a mathematical model. As the best order to the model is not beforehand known it is necessary to develop criteria to choose from the possible orders, the ideal. To introduce a solution to this problem a new method called Eigenvalue Analysis Method (MAA) is demonstrated. In this dissertation, the MAA is used to verify the existence of the linear model oversizing and to determine the output delays' number of a neural network by the eigenvalues analysis principle. To method's elaboration it was done an investigation between model order and the eigenvalues position. As a result, at least an eigenvalue appears at the left side of the  $Z$  plan when the order is oversizing, if the data used is noise free. In linear models, the results of this investigation suggest integrate the MAA with information criterion methods, improving performance in the order determination, mainly when the data used are contaminated with noise. To neural models, the performance of the MAA was tested using data obtained from a thermal process and simulated data from a theoretical case. The MAA correctly stated the output delays' number of a neural network in both cases.

# Sumário

<b>Dedicatória</b>	<b>ii</b>
<b>Agradecimentos</b>	<b>iii</b>
<b>Epígrafe</b>	<b>iv</b>
<b>Resumo</b>	<b>v</b>
<b>Abstract</b>	<b>vi</b>
<b>Lista de Figuras</b>	<b>x</b>
<b>Lista de Símbolos</b>	<b>xv</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Relevância . . . . .	2
1.2 Delimitação . . . . .	4
1.3 Apresentação do trabalho . . . . .	4



<b>2</b>	<b>Fundamentação Teórica</b>	<b>6</b>
2.1	Tempo de Amostragem . . . . .	6
2.2	Sinais de excitação . . . . .	8
2.3	Tamanho do conjunto de dados . . . . .	9
2.4	Representação Matemática de Modelos Lineares e Não-Lineares . . . . .	9
2.4.1	Modelo ARX . . . . .	10
2.4.2	Modelo ARMAX . . . . .	11
2.4.3	Representação Baseada em OE . . . . .	11
2.4.4	Representações Baseadas em Redes Neurais MLPs . . . . .	12
2.5	Estimação de parâmetros . . . . .	15
2.5.1	Propriedades das Estimativas e dos Estimadores . . . . .	16
2.5.2	Polarização . . . . .	16
2.5.3	Estimador de Mínimos Quadrados . . . . .	17
2.5.4	Estimador Estendido de Mínimos Quadrados . . . . .	18
2.6	Métodos de seleção de estrutura . . . . .	18
2.6.1	Métodos Baseados na Razão entre Determinantes . . . . .	19
2.6.2	Métodos Baseados em Medidas de Informação . . . . .	19
2.7	Comentários Finais . . . . .	21
<b>3</b>	<b>Método de Análise dos Autovalores: Considerações Preliminares</b>	<b>22</b>

3.1	Investigação da relação entre a ordem e os autovalores do modelo . . . . .	22
3.1.1	Análise matemática do sobredimensionamento . . . . .	23
3.1.2	Análise a partir de um sistema simulado . . . . .	25
3.1.3	Robustez ao ruído . . . . .	27
3.2	Procedimento para seleção de ordem de modelos lineares . . . . .	29
3.3	Aplicação do Método de Análise dos Autovalores em Modelos Lineares . . . . .	30
3.4	Comentários Finais . . . . .	33
<b>4</b>	<b>Método de Análise dos Autovalores aplicados a redes neurais recorrentes</b>	<b>34</b>
4.1	Autovalores de redes neurais recorrentes de entrada-saída . . . . .	34
4.2	Procedimento para seleção de atrasos de saída de redes neurais . . . . .	37
4.3	Considerações Finais . . . . .	38
<b>5</b>	<b>Aplicação do método de análise dos autovalores para modelos neurais recorrentes de entrada-saída</b>	<b>39</b>
5.1	Aplicação do MAA em dados obtidos a partir de um caso teórico simulado . . . . .	40
5.2	Aplicação do MAA em dados obtidos a partir de um processo térmico . . . . .	46
5.3	Considerações Finais . . . . .	51
<b>6</b>	<b>Conclusão</b>	<b>52</b>
6.1	Sugestões para trabalhos futuros . . . . .	52

**Referências Bibliográficas**

# Lista de Figuras

2.1	Transmissão e recuperação de sinais . . . . .	7
2.2	Rede neural artificial de $n$ entradas, 1 saída e $m$ neurônios na camada escondida . . . . .	13
2.3	Modelo auto-regressivo não-linear com entradas exógenas (NARX) . . . . .	14
2.4	Modelo neural auto-regressivo não-linear com média móvel e entradas exógenas (NARMAX) . . . . .	15
3.1	Autovalores e Resíduos estimados pelo Método MQ, sendo $p_3$ o autovalor sobredimensionado. . . . .	26
3.2	Autovalores e resíduos estimados da Equação (3.7) pelo Método MQ. . . . .	26
3.3	Autovalores e resíduos para o modelo (3.7) com SNR igual à 100, sendo $p_3$ o autovalor sobredimensionado. . . . .	27
3.4	Autovalores e resíduos para o modelo (3.7) com SNR igual à 50, sendo $p_3$ o autovalor sobredimensionado. . . . .	28
3.5	Autovalores e Resíduos para o modelo (3.7) com SNR igual à 50 e ordem igual a 4, sendo $p_3$ e $p_4$ o autovalor sobredimensionado. . . . .	28
4.1	Modelo neural recorrente entrada-saída . . . . .	36
5.1	Dados de identificação simulados a partir da equação (5.1) utilizados para treinamento da rede neural. (a) Dados $dad\_ctv1$ e (b) dados $dad\_ctv2$ . . . . .	41
5.2	Dados de identificação simulados a partir da equação (5.1) utilizados para validação da rede neural. (a) Dados $dad\_ctv1$ e (b) dados $dad\_ctv2$ . . . . .	41

5.3	Autovalores de modelos neurais NARMAX utilizando massa de dados <i>dad_ctv1</i> . (a) Modelo neural com $n_y = 1$ e (b) Modelo neural com $n_y = 2$ . . . . .	42
5.4	Autovalores de modelos neurais NARMAX utilizando massa de dados <i>dad_ctv2</i> . (a) Modelo neural com $n_y = 1$ e (b) Modelo neural com $n_y = 2$ . . . . .	43
5.5	Gráfico da contante de tempo em função do ponto de operação dos modelos neurais obtidos apartir da simulação da equação (5.1). (- -)Sistema simulado (-)Modelos neurais. . . . .	44
5.6	Histograma do sinal de entrada $u(t)$ dos dados de identificação da massa de dados <i>dad_ctv2</i> . . . . .	45
5.7	Predição livre utilizando modelo neural NARMAX. (a) dados <i>dad_ctv1</i> , $nn = 3$ , $n_y = 1$ e $n_u = 1$ e (b) dados <i>dad_ctv2</i> , $nn = 4$ , $n_y = 1$ e $n_u = 1$ . . . . .	46
5.8	Resposta aos ruído quantizado de um Processo Térmico. (a) Massa de dados <i>frq1</i> e (b) massa de dados <i>frq2</i> . . . . .	47
5.9	Resposta ao degrau de um Processo Térmico. (a) Massa de dados <i>fd1</i> e (b) massa de dados <i>fd2</i> . . . . .	47
5.10	Autovalores de modelos neurais recorrentes utilizando da massa de dados <i>frq1</i> . (a) Modelo neural com $n_y = 2$ e (b) Modelo neural com $n_y = 3$ . . . . .	48
5.11	Predição livre utilizando de um modelo neural recorrente com $nn = 1$ , $n_y = 2$ e $n_u = 1$ . (a) massa de dados <i>frq1</i> (b) massa de dados <i>frq2</i> (c) massa de dados <i>fd1</i> (d) massa de dados <i>fd2</i> . . . . .	50

# Lista de Tabelas

3.1	Estimação da ordem do modelo (3.7) utilizando de uma entrada PRBS de 7 bits	31
3.2	Estimação da ordem do modelo (3.8) utilizando de uma entrada PRBS de 7 bits	32
3.3	Estimação da ordem do modelo (3.9) utilizando de uma entrada aleatória . . . .	33
5.1	Valor RMSE variando o $n_y$ , $nn$ e os dados de validação do caso teórico simulado	43
5.2	Valor RMSE variando o $n_y$ , $nn$ e os dados de validação do processo térmico . .	49

# Lista de Símbolos

$A$	matriz de estimadores lineares do tipo $\hat{\theta} = Ay$ ;
$A(q^{-1}), B(q^{-1})$	polinômios em $q^{-1}$ ;
$e$	vetor de erro, pode ou não ser branco;
$e(k)$	erro no instante $k$ , pode ou não ser branco;
$E[\cdot]$	esperança matemática;
$e(t, \theta)$	erro de estimação
$j$	$\sqrt{-1}$
$z^{-1}$	operador atraso, $z^{-1}u(t) = u(t - 1)$ ;
$K$	ganho estático
$n_{\theta}$	número de parâmetros
$\nu(k)$	variável aleatória (sempre) branca
$N$	número de amostras;
$p$	número de parâmetros do modelo
$Re(x)$	parte real do número complexo $x$
$r_{y^*}(\tau)$	função de autocovariância linear;
$r_{y^*2'}(\tau)$	função de autocovariância não-linear;
$u(k)$	entrada de sistemas não-autônomos no instante $k$
$V(\theta)$	variância estimada do ruído branco
$y(k)$	sinal de saída no instante $k$
$\xi$	vetor de resíduos
$\xi(k)$	resíduo no instante $k$
$\lambda$	autovalor
$\theta$	vetor de parâmetros a estimar
$\hat{\theta}$	vetor de parâmetros estimado
$\hat{\theta}_{MQ}$	vetor de parâmetros estimado usando o estimador MQ
$\theta_i$	$i$ -ésimo parâmetro do vetor $\theta$
$\Delta$	taxa de decimação;

$\Theta$	vetor de parâmetros de um modelo NARMAX polinomial;
$\gamma_i$	elementos do polinômio da não-linearidade;
$\Xi(\cdot)$	esperança matemática;
$\tau$	constante de tempo;
$\tau_m$	Valor mínimo entre $\tau_y$ e $\tau_{y^2}$ ;
$\tau_y$	Tempo em que ocorre o primeiro mínimo da autocorrelação linear de um sinal;
$\tau_{y^2}$	Tempo em que ocorre o primeiro mínimo da autocorrelação não-linear de um sinal;
$cov[\cdot]$	covariância ou matriz de covariância
$f(\cdot)$	função genérica
$Re[\cdot]$	parte real
*	complexo conjugado
$\hat{\phantom{x}}$	valor estimado
$\overline{x(k)}$	média temporal



# Siglas e Abreviações

AIC	Critério de informação de Akaike ( <u>A</u> kaike <u>I</u> nformation <u>C</u> riterion)
AR	Modelos Auto-Regressivos ( <u>A</u> uto <u>R</u> egressive)
ARX	Modelos Auto-Regressivos com entradas exógenas ( <u>A</u> uto <u>R</u> egressive with <u>e</u> Xogenous inputs)
ARMAX	Modelo auto-regressivo de média móvel e entradas exógenas ( <u>A</u> uto <u>R</u> egressive <u>M</u> oving <u>A</u> verage model with <u>e</u> Xogenous inputs)
BIC	Critério B de informação ( <u>B</u> ayes <u>I</u> nformation) <u>C</u> riterion
CAT	Critério da função de transferência autoregressiva ( <u>C</u> riterion <u>A</u> utoregressive <u>T</u> ransfer)
DR	Razão entre Determinantes ( <u>D</u> eterminant <u>R</u> ation)
ERR	Taxa de redução do erro ( <u>E</u> rror <u>R</u> eduction <u>R</u> atio)
FPE	Erro de Previsão Final ( <u>F</u> inal <u>P</u> rediction <u>E</u> rror)
FNN	Falsos Vizinhos Próximos ( <u>F</u> alse <u>N</u> earest <u>N</u> eighbors)
GLRT	Teste da Razão de Verossimilhanças Generalizado ( <u>G</u> eneralized <u>L</u> ikelihood <u>R</u> atio <u>T</u> est)
MAA	<u>M</u> étodo de <u>A</u> nálise dos <u>A</u> utovalores
MDL	Critério de Comprimento Mínimo ( <u>M</u> inimum <u>D</u> escription <u>L</u> ength)
MISO	Múltiplas Entradas e uma saída ( <u>M</u> ultiple <u>I</u> nputs <u>S</u> ingle <u>O</u> utput)
MLP	Perceptron Multi-Camadas ( <u>M</u> ulti <u>L</u> ayer Perceptron)
MQ	<u>M</u> ínimos <u>Q</u> uadrados
MQE	<u>M</u> ínimos <u>Q</u> uadrados <u>E</u> stendidos
NARMAX	Modelo não-linear auto-regressivo de média móvel e entradas exógenas ( <u>N</u> onlinear <u>A</u> uto <u>R</u> egressive <u>M</u> oving <u>A</u> verage model with <u>e</u> Xogenous inputs)
NARX	Modelo não-linear auto-regressivo com entradas exógenas ( <u>N</u> onlinear <u>A</u> uto <u>R</u> egressive model with <u>e</u> Xogenous inputs)
$nn$	Número de neurônios
$n_u$	Número de atraso de entrada de uma rede neural recorrente de entrada-saída

$n_y$	Número de atraso de saída de uma rede neural recorrente de entrada-saída
$n_e$	Número de atraso dos termos em $e(k)$ de uma rede neural recorrente de entrada-saída
OE	Modelo de Erro na Saída ( <u>O</u> utput <u>E</u> rror)
PRBS	Sinal binário pseudo-aleatório ( <u>P</u> seudo <u>R</u> andom <u>B</u> inary <u>S</u> ignal)
RIC	Critério de Informação Residual ( <u>R</u> esidual <u>I</u> nformation <u>C</u> riterion)
RMSE	Erro Quadrático médio ( <u>R</u> oot <u>M</u> ean <u>S</u> quared <u>E</u> rror)
SNR	Relação Sinal/Ruído ( <u>S</u> ignal to <u>N</u> oise <u>R</u> atio)
SPE	Erro de Predição de Sinal ( <u>S</u> ignal <u>P</u> rediction <u>E</u> rror)
$T_s$	Tempo de amostragem

# Capítulo 1

## Introdução

O processo de identificação de um sistema nada mais é do que a utilização de técnicas de modelagem cujo objetivo maior é aproximar um determinado modelo matemático do comportamento dinâmico de um sistema real. O processo de identificação consiste de cinco etapas fundamentais (Reimbold et al., 2007):

- i) Coleta de dados
- ii) Representação matemática a ser usada
- iii) Determinação da estrutura
- iv) Estimação dos parâmetros
- v) Validação do modelo estimado

Para a coleta de dados destaca-se a escolha ideal do tempo de amostragem, pois em diversas aplicações científicas e tecnológicas faz-se necessário registrar variáveis contínuas de forma discreta no tempo.

A representação matemática a ser utilizada é muito importante no processo de identificação de sistemas. Nesta etapa é escolhido o tipo de modelo a ser utilizado, seja ele linear ou não-linear.

Os métodos de determinação da estrutura consistem na obtenção do número de parâmetros a serem estimados com o intuito de descrever a dinâmica do sistema.

A estimação de parâmetros é um procedimento numérico que determina os valores dos parâmetros do modelo, desconhecidos e não-observáveis. Após encontrar os valores dos parâmetros, aplicam-se métodos de validação de modelos, com o objetivo de qualificar o desempenho do modelo estimado. Estes métodos baseiam-se, por exemplo, na comparação entre a resposta real do sistema e aquela obtida por meio do modelo.

A escolha da ordem de um modelo aparece como um compromisso entre a capacidade de representação das dinâmicas do sistema e a simplicidade do modelo que representa um menor esforço computacional para os algoritmos de estimação. Com os métodos de seleção de ordem é possível julgar se o incremento na ordem do modelo traz uma contribuição significativa na capacidade de representar o sistema real (Haber e Unbehauen, 1990). É importante destacar que para modelos lineares é utilizado a nomenclatura “ordem” para representar o número de atrasos de saída do modelo.

A escolha de uma ordem muito baixa resulta em uma estimativa espectral excessivamente suavizada, subdimensionada <sup>1</sup>, enquanto a escolha de ordens excessivamente elevadas aumenta a resolução e introduz detalhes espúrios no espectro, sobredimensionada <sup>2</sup> (Kay, 1987).

## 1.1 Relevância

Para modelos lineares existem diversos estudos de métodos para estimação de ordem de um modelo, sendo que os mais conhecidos são os métodos baseados em critérios de informação. Estes métodos surgiram na década de 70 e até hoje são objetos de estudo.

Estes estudos buscam aprimorar os resultados obtidos pelos critérios de seleção de ordem submetidos a diferentes valores de relação sinal/ruído (SNR) e diferentes quantidade de amostras fornecidas para identificação do modelo. (Shi e Tsai, 2001; Shi e Tsai, 2004; Xiau e colegas, 2001)

No trabalho proposto por Shi e Tsai (2004) foi utilizado o Critério de Informação Residual (RIC). Este critério pode ao mesmo tempo determinar variáveis de regressão e ordens auto-regressivas.

Stoica e colegas (2004) demonstraram que a implementação do Teste Generalizado da Relação de Verossimilhança, do inglês *Generalized Likelihood Ratio Test* (GLRT) é equivalente aos métodos baseados em critérios de informação.

---

<sup>1</sup>Subdimensionada: Quando se estima uma ordem menor que a ordem do sistema original

<sup>2</sup>Sobredimensionada: Quando se estima uma ordem maior que a ordem do sistema original

Feil e colegas (2004) desenvolveram um novo método que utiliza um algoritmo de Falsos Vizinhos, do inglês *False Nearest Neighbors* (FNN) para estimação da ordem de modelos lineares e não-lineares. O método proposto utiliza uma matriz de covariância dos autovalores para estimar a estrutura do modelo.

Xiao e colegas (2005) introduziram um critério de seleção de ordem chamado Erro de Predição de Sinal (SPE) para a identificação de um modelo de regressão linear. Nesse artigo foi demonstrado que o método SPE depende muito do valor da relação sinal/ruído (SNR) e depende menos do número de amostras. Então, foi proposto uma melhoria do método para aumentar a robustez à relação sinal/ruído (SNR), chamado  $SPE^D$ .

Al-Smadi e Al-Zaben (2005) demonstraram uma nova abordagem baseada na detecção de extremidade para estimar a ordem de modelos ARMA de processo é apresentado. O método proposto realiza detecção de extremidade para seleccionar a ordem ideal de modelos ARMA, extraindo dados de contornos de uma matriz covariância derivada da sequência de dados observados.

Thanagasundram e Schindwein (2006) demonstraram uma regra prática para determinar a ordem de modelos AR através da análise de espectro dos dados de máquinas giratórias.

Palaniappan (2006) propôs o uso de algoritmo genético junto com Fuzzy ARTMAP para seleccionar a ordem apropriada de modelos AR. Resultados satisfatórios foram obtidos em sinais de eletroencefalograma.

Quinlan e colegas (2007) demonstram um método de seleção de ordem que traz uma boa performance mesmo com o baixo número de amostras. É demonstrado que, para ruído gaussiano branco e com baixo número de amostras é possível, através do perfil dos autovalores do ruído ordenado, seleccionar a ordem do modelo.

Quando o modelo é não-linear são poucas as ferramentas para seleção de ordem. Korenberg (1985) propôs a aplicação de uma taxa de redução do erro, ERR (Error Reduction Ratio). Esta estratégia permite a detecção de quais parcelas do modelo são mais relevantes para serem incluídas e quais podem ser consideradas desprezíveis.

Fogel (1991) aplicou o método AIC modificado para seleccionar a melhor rede, contudo o método é complexo e pode apresentar resultados imprecisos. Lin e colegas (1997) apresenta um algoritmo baseado em poda (pruning) que pode ser aplicado na determinação do número de atrasos da saída de uma rede neural. Esse método apresentou bons resultados para uma estrutura de modelo recorrente de entrada-saída. Embora não foi encontrado trabalhos recentes para seleção de ordem de modelos neurais a seleção de estrutura vem sendo discutida. (Rovithakis e

colegas, 2004; Ghiassi e colegas, 2005; Sorjamaa e colegas, 2005; Dyrholm e colegas, 2006)

Em Santos (2003) pode ser verificado a utilização do método da razão entre determinantes para sistemas de Hammerstein (não-Linear).

Kim e colegas (2005) apresentaram uma nova versão do AIC para seleção de ordem de modelos não-lineares. Lombardini e Gini (2005) investigaram alguns prolemas implantando métodos de critérios de informação baseados em autovalores para modelos não-lineares. János Madár e colegas (2005) propuseram um novo método para seleção de estrutura de modelos não-lineares, na qual utiliza programação genética para gerar modelos não-lineares de entrada-saída. O método proposto resulta em modelos robustos e interpretáveis.

## 1.2 Delimitação

Nesta dissertação é proposto um método que faz parte da seleção de estrutura. Trata-se de um novo método para seleção de ordem chamado Método de Análise dos Autovalores (MAA) que utiliza da análise dos autovalores para seleção de ordem de modelos lineares e não-lineares e para reduzir a tendência de sobredimensionamento de ordem apresentada por outros métodos de seleção de ordem. Ao contrário dos métodos de seleção de ordem destacados neste capítulo que possuem sua aplicação limitada a uma representação matemática específica, o MAA pode ser utilizado para diversas representações matemáticas, linear ou não-linear.

Uma aplicação importante do MAA é a possibilidade de utilização em conjunto com outros métodos de seleção de ordem com objetivo de evitar o sobredimensionamento da ordem dos modelos estimados. Em modelos lineares foi sugerido a integração do MAA com os métodos baseados no critério de informação, melhorando o desempenho na determinação adequada da ordem, principalmente quando os dados utilizados estão contaminados com ruído.

## 1.3 Apresentação do trabalho

Esta dissertação está organizada em seis capítulos da seguinte forma:

**Capítulo 2: Fundamentação Teórica.** Este capítulo apresenta uma revisão bibliográfica das teorias utilizadas neste trabalho com intuito de auxiliar o leitor no entendimento dessas teorias.

**Capítulo 3: Método de Análise dos Autovalores: Considerações Preliminares.** Neste capítulo é apresentado um estudo cujo objetivo é demonstrar a relação entre os autovalores com a ordem do modelo. É demonstrado o procedimento adotado pelo MAA para seleção de ordem

de modelos lineares e apresentado uma comparação dos resultados obtidos pelo MAA com os resultados obtidos pelos métodos baseados em medidas de informação.

**Capítulo 4: Método de Análise dos Autovalores em redes neurais.** Neste capítulo, é apresentado um procedimento para determinação dos autovalores de modelos neurais recorrentes de entrada-saída e demonstrado o procedimento adotado pelo MAA para seleção de atrasos de saída dos modelos neurais recorrentes de entrada-saída.

**Capítulo 5: Aplicação do método de análise dos autovalores para modelos neurais recorrentes de entrada-saída.** Este capítulo apresenta a aplicação do MAA em dois sistemas, sendo um sistema simulado e outro real. O objetivo é demonstrar a eficiência do MAA na seleção de atrasos de saída em modelos neurais recorrentes.

**Capítulo 6: Conclusão.** Neste capítulo é demonstrado considerações finais a respeito do trabalho e sugestões para trabalhos futuros.

# Capítulo 2

## Fundamentação Teórica

Para o desenvolvimento desta dissertação foi necessário fazer uma revisão de várias teorias relacionadas a identificação de sistemas, sendo que um breve resumo destas teorias são apresentadas neste capítulo.

### 2.1 Tempo de Amostragem

A maioria dos sistemas reais são processos contínuos no tempo. Em diversas aplicações é necessário registrar variáveis contínuas de forma discreta no tempo. O período entre duas amostras é chamado de período ou tempo de amostragem,  $T_s$ .

Para que um sinal amostrado retenha alguma das características fundamentais do sinal original, é necessário que o tempo de amostragem seja suficientemente curto. Porém, o efeito do uso de taxas de amostragem muito alta resultam:

- Redução da capacidade computacional.
- Problemas de instabilidade numérica.
- Estimação de parâmetros poderá se tornar mal condicionada.

A transmissão digital de sinais também passa pelo mesmo processo de amostragem e codificação, porém é importante que o sinal original possa ser reconstruído, a partir do digital transmitido. Torna-se então imperativo saber a frequência de amostragem do sinal para que, uma vez



discretizado<sup>1</sup>, se possa reconstruí-lo a partir de suas amostras coletadas, conforme demonstrado na Figura 2.1.

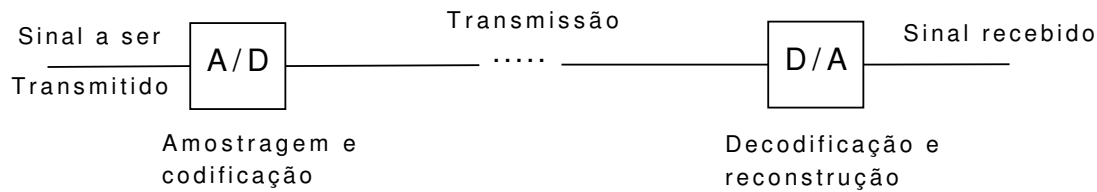


Figura 2.1: Transmissão e recuperação de sinais

O teorema de Shannon diz que um sinal que não contenha componentes de frequência acima de  $1/2T_s$  pode ser determinado unicamente a partir de amostras de tal sinal separados por  $T_s$ . Na prática, a frequência de amostragem é normalmente escolhida entre 5 a 10 vezes maior do que a maior frequência de interesse contida nos dados, e não apenas 2 vezes maior (chamada de frequência de Nyquist), como exigido pelo teorema de Shannon. Porém, o critério de se escolher a frequência de amostragem 5 a 10 vezes maior que a frequência de interesse nem sempre ajuda muito, uma vez que o conhecimento de tal frequência pode não existir *a priori*. Outra dificuldade é que, em muitos casos, será desejado verificar se um sinal previamente amostrado foi obtido com um tempo de amostragem satisfatório.

Um critério simples que tem dado uma boa indicação da frequência de amostragem e utilizada em diversas situações reais é apresentado abaixo.

- Em primeiro lugar, assume-se que um sinal  $y^*(k)$  tenha sido registrado utilizando-se um tempo de amostragem muito pequeno, ou seja, muito menor do que o necessário. Em tais casos é comum dizer que o sinal encontra-se amostrado. A questão passa a ser a definição de uma taxa pela qual o sinal observado  $y^*(k)$  será decimado de forma a gerar o sinal de trabalho  $y(k)$ , devidamente amostrado. Ou seja, deseja-se determinar  $\Delta \in \mathbb{N}$  de forma que  $y(k) = y^*(\Delta k)$ . Para fazer isto, será necessário verificar o grau de correlação (redundância) entre observações adjacentes do sinal  $y^*(k)$ . Deve ser observado que quanto mais superamostrado estiver o sinal  $y^*(k)$  maior será a redundância entre duas observações consecutivas.
- A fim de quantificar os efeitos causados pela sobreamostragem do sinal  $y^*(k)$ , serão calculadas as seguintes funções de autocovariância, uma linear e outra não-linear:

<sup>1</sup>Discretizado: Amostragem de um sinal no tempo

$$\begin{aligned} r_{y^*}(\tau) &= E \left[ \left( y^*(k) - \overline{y^*(k)} \right) \left( y^*(k - \tau) - \overline{y^*(k)} \right) \right], \\ r_{y^{*2'}}(\tau) &= E \left[ \left( y^{*2}(k) - \overline{y^{*2}(k)} \right) \left( y^{*2}(k - \tau) - \overline{y^{*2}(k)} \right) \right] \end{aligned} \quad (2.1)$$

sendo que  $E[\cdot]$  indica a esperança matemática, a barra indica a média com relação ao tempo e  $r_{y^*}(\tau)$   $r_{y^{*2'}}(\tau)$  representam os primeiros mínimos das funções de autocorrelação linear e não-linear, respectivamente. O menor desses mínimos passará a ser o valor de trabalho, ou seja,  $\tau_m^* = \min(\tau_{y^*}, \tau_{y^{*2'}})$ .

- Deseja-se escolher  $\Delta$  de forma que as funções de autocovariância do sinal decimado  $y(k) = y^*(\Delta k)$  satisfaçam

$$10 \leq \tau_m \leq 20, \quad (2.2)$$

sendo que os limites inferior e superior de (2.2) podem ser relaxados para 5 e 25, respectivamente. Mais detalhes deste critério pode ser verificado em Aguirre (2000).

## 2.2 Sinais de excitação

Em qualquer processo de identificação para se obter bons modelos matemáticos é necessário que o sinal de excitação possua características de: (i) persistência na excitação, (ii) excursão do sistema em toda a faixa de operação que se deseja modelar, (iii) tempo de amostragem compatível com as constantes de tempo envolvidas.

O sinal de excitação ideal deve possuir um amplo espectro de potência na faixa de frequências desejadas de forma a conseguir excitar um número elevado de frequências dominantes. Em outras palavras, as características dinâmicas e estáticas que não forem excitadas não aparecerão nos dados e o que não estiver nos dados não será identificado.

O sinal de excitação escolhido deverá provocar variações no processo que permitam a coleta de dados para identificar um modelo que seja capaz de reproduzir as características dinâmicas e estáticas necessárias para a aplicação desejada.

Sinais aleatórios <sup>2</sup>, ou “brancos” são comumente escolhidos porque seu espectro de potência está uniformemente distribuído numa ampla faixa de frequências. Porém, sinais aleatórios não são necessariamente fáceis de serem gerados e utilizados.

---

<sup>2</sup>Sinais aleatórios: Sinais sendo que o valor presente não possui nenhuma relação com valores passados

Uma classe de sinais largamente utilizada, e que possui características semelhantes, são os chamados *sinais pseudo-aleatórios*. Dentre esses, os sinais binários são bastante utilizados, pois são fáceis de serem gerados. Estes sinais possuem apenas dois valores possíveis,  $+V$  e  $-V$  e qualquer mudança só pode ocorrer em intervalos bem definidos  $t = 0, T_b, 2T_b, \dots$ . Este tipo de sinal é também conhecido como *sinais binários pseudo-aleatórios* (PRBS<sup>3</sup>).

## 2.3 Tamanho do conjunto de dados

O comprimento dos dados é a quantidade de amostras que deverá ser coletada durante o experimento de identificação. Quando multiplicado pelo período de amostragem determina a duração total do experimento.

A duração do experimento deverá ser longa o suficiente para que o sinal de entrada consiga excitar todas as frequências de interesse. Porém, uma dificuldade para determinar o comprimento ótimo dos dados é que o modelo real do processo é desconhecido *a priori*. Desta forma, a determinação do comprimento dos dados fica a critério do projetista. Por outro lado a duração do experimento deverá ser minimizada com o objetivo de reduzir variações indesejadas no processo e o tempo de engenharia necessário para acompanhar o experimento.

## 2.4 Representação Matemática de Modelos Lineares e Não-Lineares

Representar um sistema dinâmico ou simplesmente um fenômeno físico, por uma equação matemática é algo que muito desperta o interesse de pesquisadores. Com o desenvolvimento de processos industriais e a necessidade de controlá-lo, surgiu também a necessidade do desenvolvimento de modelos que melhor reproduzam suas características estáticas e dinâmicas (Corrêa, 2001).

Dentre as representações surgidas nas últimas décadas podem-se citar como modelos lineares, os modelos ARX, ARMAX e de erro na saída e como modelos não-lineares, os modelos NARMAX. A definição destas representações são demonstradas nas sub-seções a seguir.

---

<sup>3</sup>PRBS: Sigla em inglês *P*seudo *R*andom *B*inary *S*ignal.

### 2.4.1 Modelo ARX

Considere o seguinte modelo geral:

$$A(q)y(k) = \frac{B(q)}{F(q)}u(k) + \frac{C(q)}{D(q)}v(k), \quad (2.3)$$

sendo  $q^{-1}$  o operador de atraso, de forma que  $y(k)q^{-1} = y(k-1)$ ,  $v(k)$  ruído branco e  $A(q)$ ,  $B(q)$ ,  $C(q)$ ,  $D(q)$  e  $F(q)$  os polinômios definidos a seguir:

$$\begin{aligned} A(q) &= 1 + a_1q^{-1} + \dots + a_{n_y}q^{-n_y}, \\ B(q) &= b_1q^{-1} + \dots + b_{n_u}q^{-n_u}, \\ C(q) &= 1 + c_1q^{-1} + \dots + c_{n_v}q^{-n_v}, \\ D(q) &= 1 + d_1q^{-1} + \dots + d_{n_d}q^{-n_d}, \\ F(q) &= 1 + f_1q^{-1} + \dots + f_{n_f}q^{-n_f}. \end{aligned} \quad (2.4)$$

O modelo auto-regressivo com entradas externas pode ser obtido a partir do modelo geral (2.3) tomando-se  $C(q) = D(q) = F(q) = 1$  sendo  $A(q)$  e  $B(q)$  polinômios arbitrários:

$$A(q)y(k) = B(q)u(k) + v(k) \quad (2.5)$$

Uma vez que o ruído  $v(k)$  aparece diretamente na equação, o modelo ARX é normalmente classificado como pertencendo à classe de modelos de *erro na equação*. O modelo (2.5) pode ser reescrito da seguinte forma:

$$y(k) = \frac{B(q)}{A(q)}u(k) + \frac{1}{A(q)}v(k) \quad (2.6)$$

O ruído que aparece adicionado à saída,  $e(k) = v(k)/A(q)$ , não é branco. Nesta representação o ruído é modelado como um processo branco filtrado por um filtro auto-regressivo, com pólos idênticos aos do processo, que são as raízes do polinômio  $A(q)$ .

### 2.4.2 Modelo ARMAX

O modelo auto-regressivo com média móvel e entradas externas possui a representação mostrada abaixo:

$$A(q)y(k) = B(q)u(k) + C(q)v(k) \quad (2.7)$$

ou alternativamente

$$\begin{aligned} y(k) &= \frac{B(q)}{A(q)}u(k) + \frac{C(q)}{A(q)}v(k) \\ y(k) &= H(q)u(k) + e(k), \end{aligned} \quad (2.8)$$

sendo  $e(k)$  não branco e, como no ARX, este modelo é determinado a partir da equação geral (2.3), fazendo os polinômios  $D(q) = F(q) = 1$  e  $A(q)$ ,  $B(q)$  e  $C(q)$  polinômios arbitrários.

Nesta estrutura a dinâmica do distúrbio é representada de forma independente da dinâmica do processo, através da inclusão do polinômio  $C(q)$ . À semelhança do modelo ARX, o modelo ARMAX pertence à classe de modelos de *erro na equação*. Neste caso o erro é modelado com um processo de média móvel (MA), e o ruído adicionado à saída,  $e(k)$ , é modelado com ruído branco filtrado pelo filtro ARMA,  $C(q)/A(q)$ . Por outro lado, se um modelo ARMAX  $A(q) = C(q) = F(q)$  tal modelo pode ser representado com um modelo de erro na saída.

### 2.4.3 Representação Baseada em OE

O modelo *erro na saída* são modelos que podem ser escritos na forma da equação (2.3), mas com o polinômio  $A(q) = 1$ <sup>4</sup>. Sua expressão é mostrada abaixo:

$$y(k) = \frac{B(q)}{F(q)}u(k) + v(k) \quad (2.9)$$

é determinado a partir da equação geral (2.3), fazendo os polinômios  $A(q) = C(q) = D(q) = 1$  e  $B(q)$  e  $C(q)$  polinômios arbitrários.

---

<sup>4</sup>Alternativamente, pode-se dizer que os modelos do tipo erro na equação são aqueles que as funções de transferência do processo e do ruído têm polinômio  $A(q)$  como fator comum.

Nesta estrutura o modelo descreve somente a dinâmica do sistema. Nenhum modelo do distúrbio é considerado, ou seja, o ruído adicionado a saída é branco.

#### 2.4.4 Representações Baseadas em Redes Neurais MLPs

A utilização de redes neurais na identificação de sistemas dinâmicos foi proposta por Narendra e Parthasarathy [1990].

As redes neurais artificiais inicialmente foram concebidas para mapeamento de funções estáticas e reconhecimento de padrões. Com o desenvolvimento da teoria de redes neurais, o surgimento de novas redes e o aprimoramento de algoritmos de treinamento, as redes neurais vêm cada vez mais sendo usadas na identificação de sistemas dinâmicos [Gençay e Liu, 1997; Henrique et al., 1998].

De uma maneira geral, as redes neurais são caracterizadas por sua arquitetura, pelas características computacionais de seus neurônios e pelo algoritmo de treinamento (Henrique et al., 1998). Outra caracterização de redes neurais diz respeito à direção do fluxo de informação internamente à rede. Redes nas quais os sinais internos fluem na direção da entrada para saída são denominadas redes *feedforward*. Por outro lado, quando ocorre realimentação (*feedback*) interna na rede, estas são denominadas redes recorrentes [Jordan, 1986; Elman, 1990]. Tanto redes *feedforward* quanto redes recorrentes vêm sendo usadas na identificação de sistemas [Gençay e Liu, 1997; Henrique et al., 1998].

##### 2.4.4.1 Redes Feedforward

A Figura 2.2 apresenta a arquitetura básica de uma rede *feedforward* com  $n$  entradas, 1 saída, 1 camada escondida com  $m$  neurônios. Como citado anteriormente, o fluxo de informação internamente na rede ocorre sempre da entrada para a saída. A entrada da rede é composta pelo vetor  $x$  que é constituído, no caso de identificação de sistemas dinâmicos, por valores atrasados da saída e da entrada.

Para obter uma rede neural que reproduza o comportamento de um sistema qualquer, as seguintes etapas devem ser cumpridas: (i) escolha do número de camadas escondidas, (ii) número de neurônios em cada camada, (iv) escolha da função de ativação e (v) treinamento da rede. Esses fatores determinam a chamada arquitetura da rede.

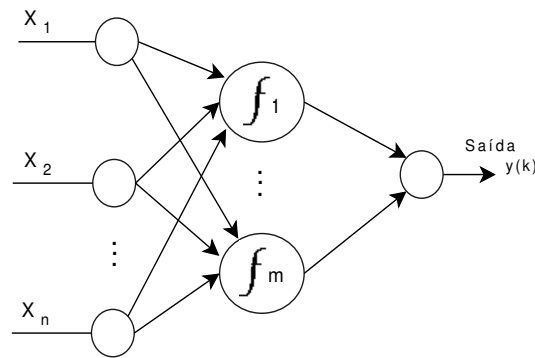


Figura 2.2: Rede neural artificial de  $n$  entradas, 1 saída e  $m$  neurônios na camada escondida

#### 2.4.4.2 Redes Neurais Recorrentes

Outra classe de redes usadas na representação de sistemas dinâmicos são as redes recorrentes. As redes neurais recorrentes são redes que possuem uma ou mais conexões de realimentação as quais proporcionam comportamento dinâmico à rede. A realimentação armazena, indiretamente, os valores prévios apresentados à rede, constituindo uma memória. Existem várias arquiteturas de rede recorrentes, cada uma realçando uma forma específica de realimentação global:

- Modelo Recorrente de Entrada-Saída - NARX [Narendra e Parthasarathly, 1992].
- Modelo de Espaço de Estados (Elman, 1990).
- Perceptron de Múltiplas Camadas Recorrentes (Puskorius et al., 1996).
- Redes de Segunda Ordem (Giles et al., 1990).
- Modelo auto-regressivo não-linear com média móvel e entradas exógenas - NARMAX [Narendra e Parthasarathly, 1992].

Todas estas redes incorporam um Perceptron de múltiplas camadas estático e exploram a capacidade de mapeamento não-linear deste Perceptron de múltiplas camadas. Para esta dissertação será utilizado apenas do modelo recorrente de entrada-saída, sendo que mais detalhes desta rede são demonstrados a seguir.

#### **Modelo recorrente de entrada-saída - NARX**

O modelo recorrente de entrada-saída possui:

- uma única entrada que é aplicada a uma memória de linha de atraso com  $n_y$  unidades
- uma única saída realimentada para a entrada através de uma outra memória de linha de atraso com  $n_u$  unidades

O conteúdo das duas memórias de linha de atraso são utilizados para alimentar a camada de entrada de uma rede Perceptron de múltiplas camadas.

A figura 2.3 mostra a arquitetura do modelo recorrente de entrada-saída, também conhecido como modelo não-linear auto-regressivo com entradas exógenas (NARX).

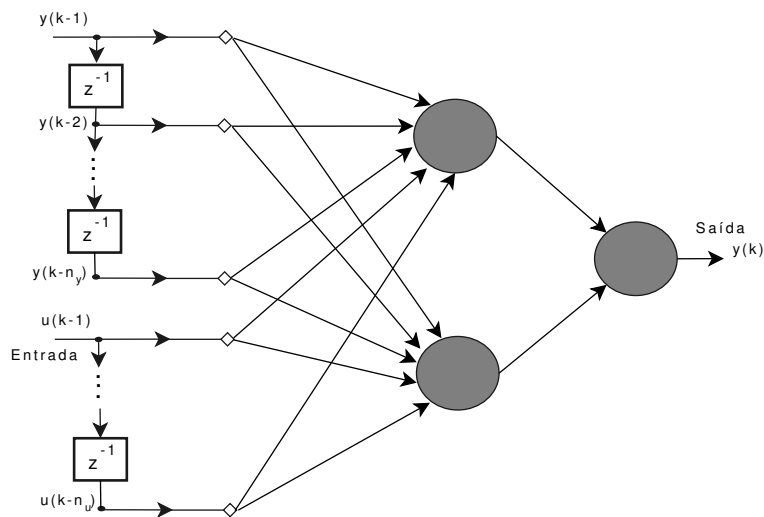


Figura 2.3: Modelo auto-regressivo não-linear com entradas exógenas (NARX)

### Modelo auto-regressivo não-linear com média móvel e entradas exógenas (NARMAX)

A estrutura apresentada pelo modelo neural NARMAX se aproxima do modelo NARX, tendo como maior diferença a inclusão do erro  $e(k)$  realimentado para a entrada por meio de uma memória de linha de atraso com  $n_e$  unidades.

A figura 2.4 mostra a arquitetura do modelo neural auto-regressivo não-linear com média móvel e entradas exógenas.



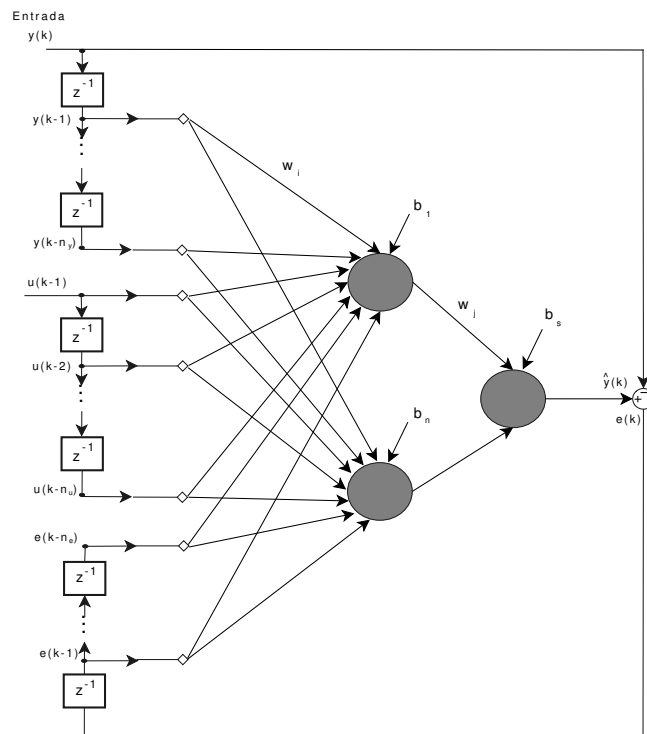


Figura 2.4: Modelo neural auto-regressivo não-linear com média móvel e entradas exógenas (NARMAX)

## 2.5 Estimação de parâmetros

Frequentemente uma amostra aleatória de uma variável aleatória  $X$  pode ser empregada com o objetivo de estimar um ou vários parâmetros desconhecidos associados à distribuição de probabilidade de  $X$ .

Considere uma função densidade <sup>5</sup> de probabilidade de uma variável aleatória  $X$ ,  $p(x, \theta)$ , que contenha um parâmetro  $\theta$  que seja desconhecido. Uma amostra aleatória  $(X_1, X_2, \dots, X_n)$  é então selecionada. Baseando-se em dados experimentais, faz-se uma estimativa do valor de  $\theta$ , por meio do cálculo de alguma estatística (chamada neste caso de estimador), por exemplo, a média  $\bar{X}$  da amostra. Diz-se então que  $\bar{X} = \hat{\mu}_x$ .

Embora não se espere que a estimativa de um parâmetro seja igual a ele, deve-se buscar sempre a melhor aproximação possível entre estes dois valores. Assim, cabem duas questões:

<sup>5</sup>Função densidade: É uma função utilizada para representar a distribuição de probabilidade caso a variável aleatória seja contínua

1. Quais características se deseja que uma boa estimativa apresente?
2. Como se consegue decidir que uma estimativa é melhor que a outra?

Existem algumas propriedades, vistas a seguir, que orientam a escolha das estimativas e estimadores.

### 2.5.1 Propriedades das Estimativas e dos Estimadores

1. Uma estimativa  $\hat{\theta}$  de um parâmetro  $\theta$  é dita ser não tendenciosa se:

$$E(\hat{\theta}) = \theta$$

2. Um estimador é dito ser consistente se a estimativa tende a se aproximar cada vez mais dos valores da população à medida que o tamanho da amostra aumenta. Ou seja:

$$E((\hat{\theta} - \theta)^2) = 0$$

3. O estimador mais eficiente, para o mesmo tamanho de amostra, é aquele que tem menor variância. Os critérios de imparcialidade (ser não tendencioso) e mínima variância não podem ser considerados separadamente, porque uma estimativa levemente tendenciosa com uma pequena variância pode ser preferível a uma estimativa não tendenciosa mas com grande variância.

### 2.5.2 Polarização

A estimação de parâmetros de um modelo linear nos parâmetros do tipo

$$\mathbf{Y} = \mathbf{P}\Theta + \mathbf{e}, \quad (2.10)$$

pode ser obtida por

$$\hat{\Theta} = \mathbf{A}\mathbf{Y} \quad (2.11)$$

Sendo  $\mathbf{P}$  a matriz dos regressores,  $\mathbf{e}$  um ruído e  $\mathbf{A}$  uma matriz cujos elementos dependem de alguma maneira dos regressores. Os parâmetros estimados não estarão polarizados se

$$\begin{aligned}
E[\mathbf{AY}] - \Theta &= 0, \\
&= E[\mathbf{A}(\mathbf{P}\Theta + \mathbf{e})] - \Theta, \\
&= E[\mathbf{AP} - \mathbf{I}]\Theta + E[\mathbf{Ae}], \\
&= (E[\mathbf{AP}] - \mathbf{I})\Theta + E[\mathbf{Ae}]
\end{aligned} \tag{2.12}$$

sendo  $\Theta$  determinístico. Da equação ( 2.12) verifica-se que a polarização será nula se

- $E[\mathbf{AP}] = \mathbf{I}$ ,
- os elementos de  $\mathbf{A}$  não forem correlacionados com o ruído,
- o ruído possuir média nula.

A condição  $E[\mathbf{AP}] = \mathbf{I}$  é satisfeita por meio da escolha adequada da matriz  $\mathbf{A}$ . Pode ser demonstrado que o estimador de mínimos quadrados atende à primeira condição. As duas últimas são satisfeitas pelo estimador estendido de mínimos quadrados (EMQ).

### 2.5.3 Estimador de Mínimos Quadrados

O método de mínimos quadrados é um dos mais conhecidos e mais utilizados nas mais diversas áreas de ciência e tecnologia (Aguirre, 2007).

O método dos mínimos quadrados é uma técnica de otimização matemática que procura encontrar o melhor ajustamento para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre a curva ajustada e os dados (tais diferenças são chamadas resíduos).

Um requisito implícito para o método dos mínimos quadrados trabalhar é que os erros em cada medida sejam distribuídos aleatoriamente com função densidade gaussiana, e que os resíduos sejam independentes. O Método dos Mínimos Quadrados (MMQ) é um processo de busca dos coeficientes que minimizam a soma dos quadrados dos resíduos. Para tanto, uma das condições é que a forma escolhida para a equação seja adequada. O MMQ minimiza a soma dos quadrados dos resíduos.

Seja um problema escrito na forma da equação abaixo:

$$Y = \mathbf{P}\hat{\Theta} + \xi, \tag{2.13}$$

O estimador de mínimos quadrados é um estimador do tipo  $\hat{\Theta} = \mathbf{A}\mathbf{Y}$ , sendo

$$\mathbf{A} = [\mathbf{P}^T \mathbf{P}]^{-1} \mathbf{P}^T, \quad (2.14)$$

$\xi$  é o resíduo de modelagem e  $\mathbf{P}$  a matriz de regressores. Para regressores incluídos na saída ocorrerá polarização do estimador de mínimos quadrados.

#### 2.5.4 Estimador Estendido de Mínimos Quadrados

O estimador estendido de mínimos quadrados (EMQ) contorna o problema de polarização acrescentando à matriz de regressores a parte modelada do ruído. Então, a nova matriz de regressores será composta por

$$\mathbf{P}^* = [\mathbf{P} \ \mathbf{P}_e], \quad (2.15)$$

sendo  $\mathbf{P}$  a matriz de regressores que contém os termos de processo, e  $\mathbf{P}_e$  a matriz que contém os termos de ruído. Dessa forma, garante-se que o estimador será não polarizado e a estimação não-polarizada de  $\hat{\Theta}$  será dada, através de um processo iterativo, por

$$\hat{\Theta}_i^* = [P_i^{*T} P_i^*]^{-1} P_i^{*T} y. \quad (2.16)$$

Na prática a função que modela a parte determinística do ruído não é conhecida. Porém, tem-se observado que modelos lineares de ruído com  $n_e$  entre 10 e 20 são suficientes para redução da polarização dos estimadores.

Sendo o ruído uma grandeza que não pode ser medida, a construção da matriz  $\mathbf{P}_e$  é feita com valores estimados do ruído, determinados por

$$\xi(k)_i = y - P_{i-1}^* \hat{\Theta}_{i-1}^*, \quad (2.17)$$

sendo que  $i$  indica o número de iterações do algoritmo EMQ. No primeiro passo,  $i = 1$  utiliza-se o estimador de mínimos quadrados. Em geral com quatro iterações,  $i = 4$ , é suficiente para se eliminar a polarização do estimador. Detalhes a respeito de estimadores não polarizados podem ser vistos em Aguirre (2007).

## 2.6 Métodos de seleção de estrutura

Um dos aspectos mais importantes na determinação da estrutura de um modelo é a escolha da ordem do modelo. A necessidade de se escolher um valor adequado para a ordem de um mo-

delo pode ser apreciada verificando-se, que se a ordem usada for menor que a ordem efetiva do sistema real, o modelo não possuirá a complexidade estrutural necessária para reproduzir a dinâmica do sistema. Por outro lado, se a ordem do modelo for muito maior do que a necessária, a estimação de parâmetros será provalvemente mal condicionada. Apresenta-se nesta seção alguns métodos que, embora não desenvolvidos com o objetivo específico para seleção de ordem, retornam a estrutura ideal e com ela a ordem.

### 2.6.1 Métodos Baseados na Razão entre Determinantes

Estes métodos baseiam-se na singularidade de uma matriz de informação cujos elementos contêm dados sobre a correlação entre os sinais de entrada e saída em diferentes instantes de tempo. A ordem da matriz de informação é função da ordem do modelo a ser obtido. Quando a ordem do modelo é selecionada maior que a do sistema real, a matriz de informação torna-se redundante, isto é, têm-se colunas linearmente dependentes causando a singularidade <sup>6</sup> (Castanho, 1991).

A seleção da ordem do modelo é feita po meio do cálculo do determinante da matriz de informação e comparando-o com o de ordem imediatamente superior  $(n + 1)$ . Quando a razão entre os determinantes,  $DR$ , apresenta um aumento significativo admite-se, então, como  $n$  a ordem mais adequada para o modelo.

$$DR(n) = \left| \frac{\det(\mathbf{Q}(\phi, n))}{\det(\mathbf{Q}(\phi, n + 1))} \right| \quad (2.18)$$

sendo,  $\mathbf{Q}$  a matriz de informação. Embora, geralmente, não seja muito significativa no cálculo do determinante da matriz  $\mathbf{Q}$ , é desejável que a seleção da ordem do polinômio da não-linearidade  $m$ , seja determinada antes do cálculo da  $DR$ . Isto pode ser feito através da determinação da característica estática do processo e do ajuste de um polinômio para sua representação. A ordem escolhida é a menor capaz de representar a não-linearidade estática na faixa de operação de interesse.

### 2.6.2 Métodos Baseados em Medidas de Informação

Os métodos baseados em medidas de informação são os mais conhecidos e utilizados na atualidade e se baseiam na utilização de uma fórmula matemática que se divide basicamente em duas parcelas. Uma parcela quantifica a diminuição da variância dos resíduos resultante da in-

---

<sup>6</sup>Singularidade é geralmente um ponto no qual um dado objeto matemático não é definido.

clusão de um termo, ao passo que a outra parcela penaliza a inclusão de cada termo. Sendo que os diferentes métodos baseados em medidas de informação se diferem basicamente em como ponderar estas duas parcelas. Os mais conhecidos são:

1. Critério de informação Akaike (AIC) (Akaike, 1969)
2. Critério do Erro de Previsão Final (FPE) (Akaike, 1969)
3. Critério da função de transferência autoregressiva (CAT) (Parzen, 1974)
4. Critério de Comprimento Mínimo (MDL) (Rissanen, 1978)
5. Critério B de informação (BIC) (Akaike, 1977, 1978; Schwartz, 1978)

### **Critério do Erro de Previsão Final (FPE)**

Método desenvolvido por Akaike (1969) consiste, entre modelos pré-determinados, na escolha daquele que apresente o menor valor de FPE, conforme a função

$$FPE(n) = \frac{N + (p + 1)}{N - (p + 1)} \hat{\rho}_p \quad (2.19)$$

sendo  $p$  o número de parâmetros do modelo (considera-se que  $p = 2n$ ),  $N$  o número de amostras e  $\hat{\rho}_p$  é a variância estimada do ruído branco de entrada do processo para o modelo de ordem  $p$ .

### **Critério de informação Akaike (AIC)**

Outro critério de seleção sugerido por Akaike (1974), que utiliza a abordagem de máxima verossimilhança, sendo conhecido como *Akaike Information Criterion* (AIC). Baseia-se na minimização do erro entre o modelo aproximado e a verdadeira função de transferência do sistema. O modelo escolhido como o mais adequado é o que apresenta o menor valor de AIC conforme a função.

$$AIC(n) = \ln(\hat{\rho}_p) + \frac{2p}{N} \quad (2.20)$$

O termo  $2p$  representa uma penalidade por utilizar coeficientes AR extras os quais não resultam na redução da variância do erro de predição.

Para valores grandes de  $N$ , os critérios AIC e FPE fornecem resultados praticamente idênticos. Segundo Marple (1987), os dois critérios funcionam bem para processos puramente autoregressivos, mas não apresentam resultados confiáveis quando utilizados para estimar a ordem

ótima de sinais reais. Além disso, Kashyap (1980) chamou atenção para o fato de que o AIC é estatisticamente inconsistente, uma vez que a probabilidade de erro na escolha da ordem ótima não tende a zero à medida que  $N$  tende a infinito. Por isso, esse critério tende a sugerir ordens mais elevadas à medida que se aumenta a duração do sinal, assim como o FPE.

### **Critério de Comprimento Mínimo (MDL)**

O critério de comprimento mínimo de descrição (MDL, *Minimum Description Length*), proposto por Rissanen (1978), procura solucionar a inconsistência estatística do AIC de superestimar a ordem do modelo quando os trechos analisados têm um número grande de amostras, a função a ser minimizada é fornecida por:

$$MDL(p) = N \ln(\hat{\rho}_p) + p \ln(N) \quad (2.21)$$

sendo que  $MDL$  representa *Minimum Description Length*. Este método é considerado estatisticamente consistente porque o 2º termo da equação,  $p \ln(N)$ , aumenta mais rápido com relação à  $N$  do que com  $p$ .

## **2.7 Comentários Finais**

Neste capítulo é abordado os diversos conceitos que serão utilizados nos capítulos seguintes. Importantes considerações sobre representações matemáticas de modelos lineares e não-lineares foram realizadas, sendo que uma discussão mais detalhada é feita sobre as representações baseadas em redes neurais MLPs.

No final deste capítulo é demonstrado um breve resumo dos métodos baseados em critério de informação. No capítulo seguinte é demonstrado os resultados da aplicação dos métodos baseados em medidas de informação em 3 modelos distintos.

## Capítulo 3

# Método de Análise dos Autovalores: Considerações Preliminares

Como demonstrado no capítulo anterior, os métodos mais conhecidos e utilizados para seleção de ordem de modelos lineares são os métodos baseados em critério de informação, porém, estes métodos possuem uma pequena tendência de sobredimensionamento. É demonstrado neste capítulo um procedimento que utiliza o Método de Análise dos Autovalores (MAA) para diminuir a tendência de sobredimensionamento de ordem dos métodos baseados em medidas de informação.

Apresenta-se aqui uma investigação da relação entre a ordem e autovalores de um modelo linear discreto. Os resultados desta investigação possibilitaram a criação do MAA.

### 3.1 Investigação da relação entre a ordem e os autovalores do modelo

Corrêa (2001) sugere uma análise dos autovalores do modelo como indicação de sobredimensionamento da ordem do mesmo. Em um sistema SISO contínuo com taxa de amostragem adequada, seus autovalores estarão sempre no semicírculo direito e, a existência de autovalores do lado esquerdo pode ser indicativo de sobredimensionamento da ordem do modelo.

Nesta seção é apresentado um estudo cujo objetivo é demonstrar a relação entre os autovalores com a ordem do modelo. Para realizar este estudo, utiliza-se sistemas representados por função de transferência pulsada (discreta).



Um sistema linear pode ser representado na forma de função de transferência ou de espaço de estado. Os pólos são obtidos a partir de representações na forma de função de transferência e autovalores são obtidos a partir das representações na forma de espaço de estado. Porém, é importante ressaltar que os pólos de uma função de transferência correspondem a autovalores na representação de espaço de estado. Nesta dissertação, toda referência feita para autovalor de uma função transferência estará fazendo referência ao autovalor desta função de transferência representada na forma de espaço de estado.

### 3.1.1 Análise matemática do sobredimensionamento

Considere o modelo de segunda ordem descrito pela Equação (3.1) como sendo o modelo exato de um dado sistema,

$$F_1(z) = \frac{F(z)}{(z - p_1)(z - p_2)} \quad (3.1)$$

sendo  $p_1$  e  $p_2$  os pólos do modelo  $F_1(z)$ . Porém, tomando uma estrutura sobredimensionada conforme a Equação (3.2),

$$F_2(z) = \frac{\widehat{F}(z)}{(z - \widehat{p}_1)(z - \widehat{p}_2)(z - \widehat{p}_3)} \quad (3.2)$$

sendo  $\widehat{p}_1$ ,  $\widehat{p}_2$  e  $\widehat{p}_3$  pólos estimados do modelo  $F_2(z)$ , pretende-se verificar o comportamento de  $\widehat{p}_3$ , considerando este o pólo sobredimensionado.

Assim, será considerada a aproximação de  $F_1(z)$  a partir de  $F_2(z)$ , supondo que ambos os modelos são estáveis.

$$\frac{F(z)}{(z - p_1)(z - p_2)} \approx \frac{\widehat{F}(z)}{(z - \widehat{p}_1)(z - \widehat{p}_2)(z - \widehat{p}_3)} \quad (3.3)$$

Expandindo os modelos (3.1) e (3.2) em frações parciais, tem-se:

$$\begin{aligned} F_1(z) &= \frac{a_1}{(z - p_1)} + \frac{a_2}{(z - p_2)} \\ F_2(z) &= \frac{\widehat{a}_1}{(z - \widehat{p}_1)} + \frac{\widehat{a}_2}{(z - \widehat{p}_2)} + \frac{\widehat{a}_3}{(z - \widehat{p}_3)} \end{aligned} \quad (3.4)$$

No tempo discreto,

$$\begin{aligned} f_1(k) &= a_1(p_1)^k + a_2(p_2)^k \\ f_2(k) &= \hat{a}_1(\hat{p}_1)^k + \hat{a}_2(\hat{p}_2)^k + \hat{a}_3(\hat{p}_3)^k \end{aligned} \quad (3.5)$$

sendo  $a_1$  e  $a_2$  os resíduos do modelo  $F_1(z)$  e  $\hat{a}_1, \hat{a}_2$  e  $\hat{a}_3$  os resíduos estimados do modelo  $F_2(z)$ .

Fazendo uma hipótese que o estimador de parâmetros estime o valor de  $p_1 = \hat{p}_1$  e  $p_2 = \hat{p}_2$  e analisando a equação no tempo discreto, pode-se perceber que para fazer com que o modelo (3.2) tenda ao modelo (3.1) é necessário que  $\hat{a}_3(\hat{p}_3)^k \rightarrow 0$ . Assim,

- $a_1 \rightarrow \hat{a}_1, a_2 \rightarrow \hat{a}_2$  e  $\hat{a}_3 \rightarrow 0$ ;
- e ainda  $\sum_{k=1}^{\infty} (\hat{p}_3)^k \rightarrow 0$

Utilizando-se da Equação 3.4 e calculando  $\hat{a}_3$  pelo método dos resíduos, tem-se:

$$\hat{a}_3 = \frac{\hat{F}(\hat{p}_3)}{(\hat{p}_3 - \hat{p}_1)(\hat{p}_3 - \hat{p}_2)} \quad (3.6)$$

Pode-se verificar através da Equação (3.6) que, um valor de  $\hat{p}_3$  negativo, contribui para minimizar  $\hat{a}_3$ .

Para que  $\sum_{k=1}^{\infty} (p_3)^k \rightarrow 0$ , em função da variação de  $k$  é necessário, também, que  $\hat{p}_3$  seja negativo.

Logo pode-se concluir que  $\hat{p}_3$  negativo ou seu autovalor correspondente no lado esquerdo do plano  $z$  leva à minimização da contribuição de  $\hat{a}_3(\hat{p}_3)^k$  em  $f_2(k)$  (Equação (3.5)).

Ressalta-se que, nesta seção, os modelos descritos pelas equações (3.1) e (3.2) foram apresentados no domínio  $z$ . Porém, tanto para simulação quanto para estimação de parâmetros foram utilizadas equações de diferenças.

Sem perda de generalidade toda análise feita utilizando da representação na forma de função de transferência e seus resultados são os mesmos obtidos para representação em espaço de estado.

### 3.1.2 Análise a partir de um sistema simulado

Com intuito de obter uma relação entre a ordem do modelo e a posição dos autovalores é feito uso de um sistema “real” representado pela Equação (3.7), que possui autovalores iguais a 0,82 e 0,36. A partir da simulação da Equação (3.7) obtém-se um conjunto de dados representado por  $[\mathbf{u}, \mathbf{y}]$ .

$$\begin{aligned} y(k) = & 1,18y(k-1) - 0,2952y(k-2) \\ & - 0,25u(k-1) + e(k). \end{aligned} \quad (3.7)$$

Utilizando o conjunto de dados  $[\mathbf{u}, \mathbf{y}]$ , foi identificado um modelo com ordem de uma unidade superior ao sistema simulado. Para a estimação dos parâmetros foi utilizado do método dos Mínimos Quadrados (MQ) e 100 realizações com 200 amostras em cada uma. O modelo estimado terá um autovalor a mais que o sistema simulado.

Utilizando um ruído igual à zero, sinal de entrada aleatório e ordenando os autovalores obtidos pelo estimador de mínimos quadrados, pode-se perceber por meio da Figura 3.1 que os valores estimados de 0,82 e 0,36 para o primeiro e segundo autovalor, respectivamente, são iguais aos autovalores do sistema original.

Em relação ao autovalor sobredimensionado, observa-se em todas as realizações um valor com sinal negativo e o seu resíduo tendendo a zero.

Para um melhor entendimento da relação entre a ordem do modelo e a posição dos autovalores é demonstrado na Figura 3.2 uma representação gráfica da localização dos autovalores e resíduos no plano  $z$  de dois modelos obtidos pelo método MQ utilizando do conjunto de dados  $[\mathbf{u}, \mathbf{y}]$ . A Figura 3.2(a) representa um modelo com ordem igual a ordem do sistema original e a Figura 3.2(b) representa um modelo com uma unidade de ordem superior. Pode-se verificar que todos os autovalores da Figura 3.2(a) encontram-se no semicírculo direito, porém, na Figura 3.2(b) o autovalor sobredimensionado encontra-se no semicírculo esquerdo.

Os resultados com o sistema simulado sugerem que a hipótese levantada é razoável. Naturalmente, embora não identificado, pode existir situações que a hipótese não seja verdadeira. Todos os resultados apresentados nesta seção corroboram a análise matemática apresentada na seção 3.1.1.

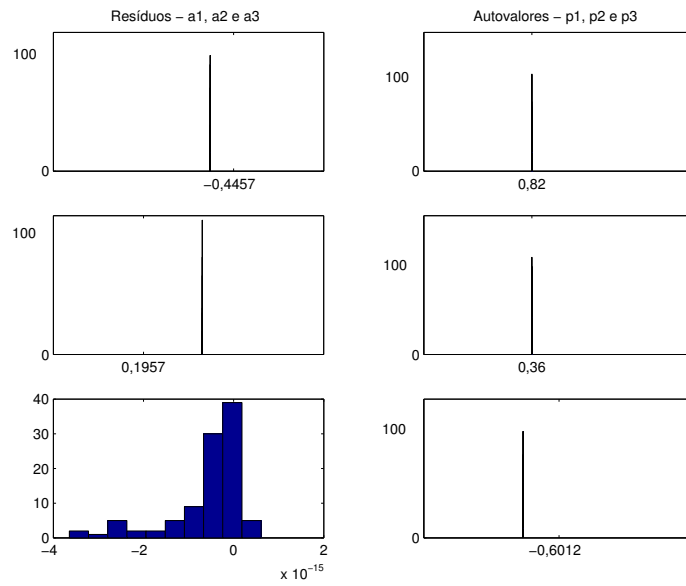


Figura 3.1: Autovalores e resíduos estimados pelo Método MQ, sendo  $p_3$  o autovalor sobredimensionado. O eixo das ordenadas representa a quantidade de amostras e o eixo das abcissas representa a faixa de histograma do autovalor ou resíduo do modelo estimado.

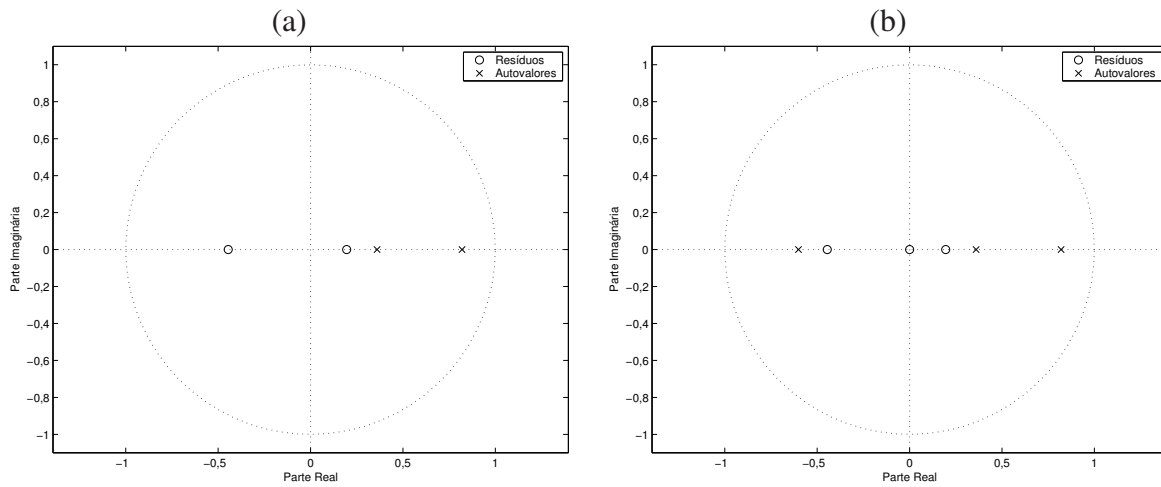


Figura 3.2: Autovalores e resíduos estimados da Equação (3.7) pelo Método MQ. (a) Modelo com ordem igual a ordem do sistema original (b) Modelo com uma unidade de ordem superior a ordem do sistema original.

### 3.1.3 Robustez ao ruído

Nesta seção avalia-se o que ocorre com os autovalores do modelo estimado quando os dados de identificação estão contaminados com ruído. Para esta análise o modelo (3.7) foi simulado adicionando um ruído, com distribuição de probabilidade teoricamente normal, por meio da variável  $e(k)$ , obtendo-se diferentes valores de SNR. Novamente foram feitas 100 realizações com 200 amostras para cada uma delas. Os parâmetros de um modelo sobredimensionado de uma ordem foram estimados por meio do MQ.

Os resultados obtidos são apresentados nas Figuras 3.3 e 3.4, respectivamente para SNR=100 e SNR=50. Observa-se na Figura 3.3 que após a adição do ruído, o autovalor sobredimensionado, em média, tende à zero. Este fato também é verificado na Figura 3.4 quando se tem um valor de SNR mais baixo, porém nesse caso tem-se um aumento da variância. Portanto, não se observou na totalidade das realizações a estimação de valor negativo para o autovalor sobredimensionado.

A Figura 3.5 apresenta o resultado para um modelo sobredimensionado na ordem em duas unidades. Note-se neste caso que o segundo autovalor sobredimensionado tende a um valor negativo.

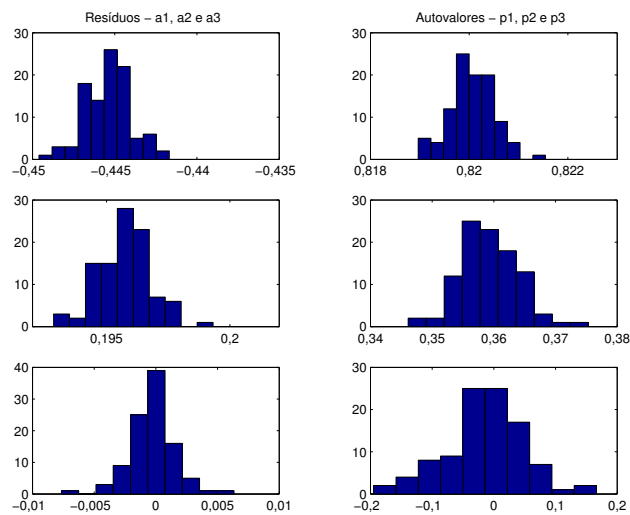


Figura 3.3: Autovalores e resíduos para o modelo (3.7) com SNR igual à 100, sendo  $p_3$  o autovalor sobredimensionado. O eixo das ordenadas representa a quantidade de realizações e o eixo das abscissas representa a faixa de histograma do autovalor ou resíduo do modelo estimado.

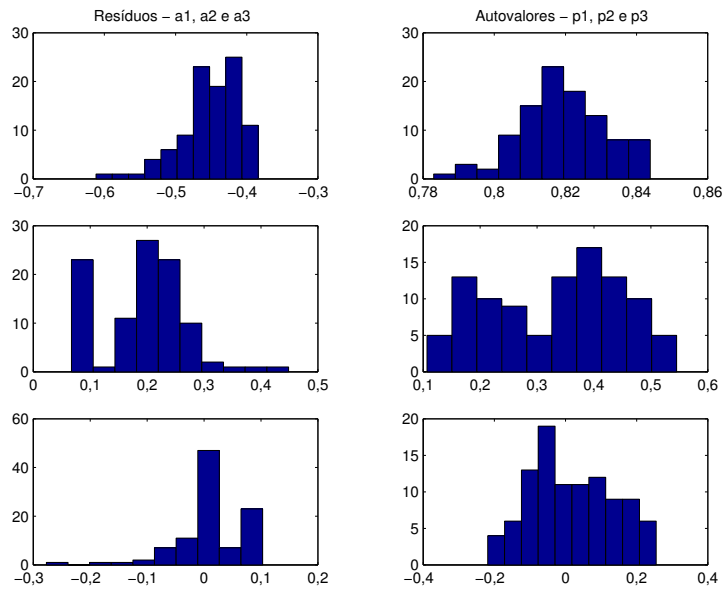


Figura 3.4: Autovalores e resíduos para o modelo (3.7) com SNR igual à 50, sendo  $p_3$  o autovalor sobredimensionado. O eixo das ordenadas representa a quantidade de realizações e o eixo das abcissas representa a faixa de histograma do autovalor ou resíduo do modelo estimado.

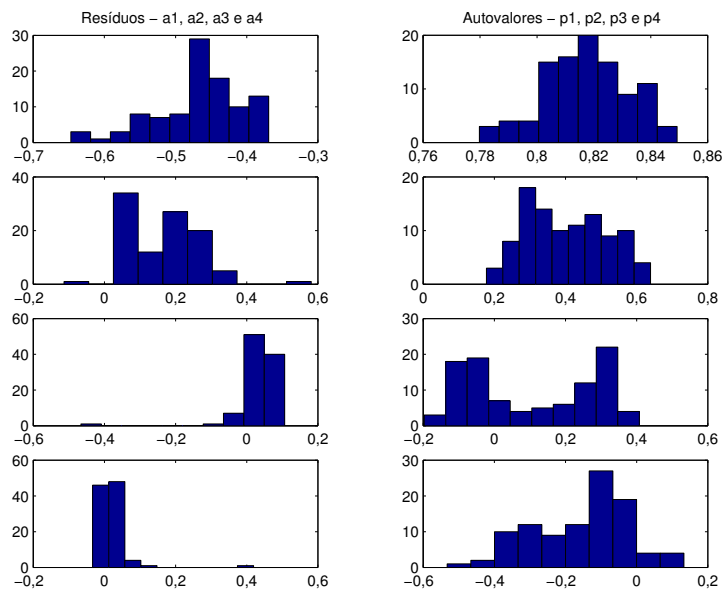


Figura 3.5: Autovalores e Resíduos para o modelo (3.7) com SNR igual à 50 e ordem igual a 4, sendo  $p_3$  e  $p_4$  o autovalor sobredimensionado. O eixo das ordenadas representa a quantidade de realizações e o eixo das abcissas representa a faixa de histograma do autovalor ou resíduo do modelo estimado.

Com os resultados obtidos nesta seção é possível perceber que quando o modelo (3.7) foi simulado adicionando ruído aos dados de identificação, o autovalor sobredimensionado tentará modelar este ruído. Desta forma, este autovalor na média, tende à zero.

Com os resultados apresentados, conclui-se que o comportamento dos autovalores do modelo estimado, utilizando dados de identificação contaminados com ruído, difere dos resultados apresentados na seção 3.1.2. Neste caso, o autovalor sobredimensionado não tenderá a um valor negativo e sim à zero. Faz-se necessário o aumento da ordem do modelo simulado em duas unidades para que se perceba a presença de um autovalor com tendência de valor negativo.

## 3.2 Procedimento para seleção de ordem de modelos lineares

Por meio da análise matemática do sobredimensionamento de um modelo e análise das figuras obtidos na Seção 3.1 é proposto um método que utiliza o princípio de análise dos autovalores para estimação da ordem de modelos lineares.

A seguir o MAA para seleção de ordem de modelos lineares é descrito:

- a) Parte-se do pressuposto que o sistema seja estável identificado em malha aberta e que o valor de frequência de amostragem seja adequado, ou seja, 5 a 10 vezes maior que a frequência de interesse contida nos dados. Após discretizado com um valor de frequência de amostragem adequada os autovalores do sistema original estarão no semicírculo direito do plano discreto  $z$ ;
- b) Utilizar do modelo ARX para identificação do processo e incluir os termos candidatos na seguinte ordem:  $y(k-1)$ ,  $u(k-1)$ ,  $y(k-2)$ ,  $u(k-2)$ ,  $y(k-3)$ ,  $u(k-3)$  e assim por diante;
- c) A cada inclusão de termo, desenvolver a análise dos autovalores no plano  $z$ . Quando aparecer um autovalor no semicírculo esquerdo do plano  $z$  será um indicativo de ordem sobredimensionada;
- d) A ordem a se escolher para o modelo será a ordem na qual todos os autovalores se encontram no semicírculo direito;

O MAA pode também ser utilizado em conjunto com outros métodos de seleção de ordem para modelos lineares, diminuindo uma tendência de sobredimensionamento que este método possa vir apresentar. Para isso:

- a) Utilize do método de seleção de ordem para obter a ordem do modelo;
- b) Após obtido o resultado utilize do MAA para verificar a posição dos autovalores no plano  $z$ ;
- c) Caso apareça algum autovalor no semicírculo esquerdo diminua a ordem;
- d) A ordem a se escolher para o modelo será a ordem na qual todos os autovalores se encontram no semicírculo direito;

Na próxima seção serão apresentados os resultados do MAA utilizado em conjunto com métodos baseados em critério de informação. Nesta dissertação utiliza-se, sem perda de generalidade, o FPE e MDL.

### 3.3 Aplicação do Método de Análise dos Autovalores em Modelos Lineares

Para avaliar o procedimento proposto para seleção de ordem foram utilizados dados obtidos a partir da simulação dos modelos (3.7), (3.8) e do modelo de erro na saída (3.9). Os resultados do MAA são comparados com FPE e MDL, assim como a utilização conjunta dos métodos FPE e MDL com o MAA, respectivamente aqui chamados de FPE-A e MDL-A.

$$\begin{aligned}
 y(k) &= 1,2y(k-1) - 0,8y(k-2) + 0,4y(k-3) \\
 &\quad + u(k-1) - 0,4u(k-2) + 0,1u(k-3) \\
 &\quad - u(k-4) + e(k)
 \end{aligned} \tag{3.8}$$

$$\begin{aligned}
 w(k) &= 1,5w(k-1) - 0,75w(k-2) - u(k-1) \\
 y(k) &= w(k) + e(k)
 \end{aligned} \tag{3.9}$$

Uma característica importante a ser analisada é o comportamento do método com a variação do sinal de entrada, pois conforme demonstrado em Feil e colegas [2004] o sinal de entrada tem um grande impacto sobre a qualidade dos resultados. Portanto, para os modelos (3.7) e (3.8) foi utilizada uma entrada PRBS de 7 bits e o Método de MQ para estimação dos parâmetros,



enquanto que para o modelo (3.9) foi utilizada uma entrada Aleatória e o Método de Mínimos Quadrados Estendido (MQE).

Para poder comparar o desempenho dos diversos métodos na seleção de ordem dos modelos (3.7), (3.8) e (3.9) foram realizadas simulações variando a quantidade de amostras ( $N$ ) e o valor SNR. Os resultados das simulações são apresentados nas Tabelas (3.1), (3.2) e (3.3).

A Tabela (3.1) apresenta o resultado para a situação em que os dados foram obtidos a partir do modelo (3.7) de ordem 2, com erro na equação. A Tabela (3.2) mostra o resultado para o modelo (3.8) de ordem 3, também com erro na equação. Já a Tabela (3.3) apresenta os resultados para o modelo (3.9) de ordem 2, com erro na saída.

Tabela 3.1: Estimação da ordem do modelo (3.7) utilizando de uma entrada PRBS de 7 bits

SNR	N	Critério	Ordem				
			1	2	3	4	5
10	200	FPE	0	<b>87</b>	13	0	0
		MDL	0	<b>97</b>	3	0	0
		MAA	0	<b>51</b>	43	6	0
		FPE-A	0	<b>94</b>	6	0	0
		MDL-A	0	<b>99</b>	1	0	0
	50	FPE	0	<b>88</b>	9	2	1
		MDL	0	<b>96</b>	3	1	0
		MAA	0	<b>52</b>	38	10	0
		FPE-A	2	<b>94</b>	4	0	0
		MDL-A	2	<b>98</b>	0	0	0
1	200	FPE	0	<b>89</b>	9	2	0
		MDL	0	<b>97</b>	3	0	0
		MAA	0	<b>46</b>	45	9	0
		FPE-A	0	<b>98</b>	2	0	0
		MDL-A	0	<b>100</b>	0	0	0
	50	FPE	0	<b>89</b>	11	0	0
		MDL	0	<b>95</b>	5	0	0
		MAA	1	<b>46</b>	47	0	0
		FPE-A	2	<b>96</b>	2	0	0
		MDL-A	2	<b>97</b>	1	0	0

Por meio da Tabela (3.1), verifica-se que o MAA apresentou resultados inferiores ao FPE e MDL. Neste caso verifica-se que o MAA teve uma tendência de sobredimensionamento de ordem em uma unidade. Os melhores resultados foram obtidos com os métodos FPE-A e MDL-A.

Na Tabela (3.2) novamente verifica-se que os métodos FPE-A e MDL-A apresentaram desempenho superior ao FPE e MDL. Porém, neste caso o MAA se mostrou mais robusto aos testes com pequeno número de amostras. Conforme observado neste capítulo, o MAA apresentou tendência em sobreestimar a ordem em uma unidade para os casos de dados com ruído.

Tabela 3.2: Estimação da ordem do modelo (3.8) utilizando de uma entrada PRBS de 7 bits

SNR	N	Critério	Ordem				
			1	2	3	4	5
10	200	FPE	0	6	<b>14</b>	5	75
		MDL	0	14	<b>25</b>	13	48
		MAA	0	0	<b>48</b>	38	14
		FPE-A	0	6	<b>48</b>	31	15
		MDL-A	0	14	<b>54</b>	25	7
	50	FPE	0	57	<b>30</b>	9	4
		MDL	0	73	<b>16</b>	9	2
		MAA	0	1	<b>39</b>	49	11
		FPE-A	0	61	<b>32</b>	7	0
		MDL-A	0	76	<b>18</b>	6	0
1	200	FPE	0	15	<b>44</b>	19	22
		MDL	0	44	<b>47</b>	7	2
		MAA	0	0	<b>43</b>	49	8
		FPE-A	0	15	<b>61</b>	23	1
		MDL-A	0	44	<b>51</b>	5	0
	50	FPE	0	71	<b>22</b>	6	1
		MDL	0	86	<b>13</b>	1	0
		MAA	0	3	<b>57</b>	38	2
		FPE-A	0	72	<b>24</b>	4	0
		MDL-A	0	86	<b>14</b>	0	0

No caso apresentado na Tabela (3.3), observa-se que, em geral, o melhor desempenho obtido foi o MAA. Novamente o uso conjunto do MAA com FPE e MDL possibilitou uma sensível melhoria na estimação correta da ordem dos modelos. Embora não apresentados, testes com outros modelos e com diferentes métodos baseados em critério de informação, apresentaram resultados semelhantes.

É importante ressaltar que o MAA não apresentou o mesmo problema dos métodos AIC e FPE de ser estatisticamente inconsistente para  $N$  tendendo ao infinito. Pode-se constatar para os modelos simulados que as variações de  $N$  influenciaram pouco nos resultados apresentados pelo MAA.

Tabela 3.3: Estimação da ordem do modelo (3.9) utilizando de uma entrada aleatória

SNR	N	Critério	Ordem				
			1	2	3	4	5
10	200	FPE	5	<b>60</b>	23	8	4
		MDL	6	<b>62</b>	24	6	2
		MAA	1	<b>89</b>	10	0	0
		FPE-A	6	<b>94</b>	0	0	0
		MDL-A	7	<b>93</b>	0	0	0
	50	FPE	48	<b>37</b>	12	3	0
		MDL	55	<b>36</b>	8	1	0
		MAA	1	<b>83</b>	16	0	0
		FPE-A	49	<b>51</b>	0	0	0
		MDL-A	55	<b>45</b>	0	0	0
1	200	FPE	33	<b>32</b>	23	10	2
		MDL	37	<b>35</b>	20	8	0
		MAA	4	<b>88</b>	8	0	0
		FPE-A	33	<b>67</b>	0	0	0
		MDL-A	37	<b>63</b>	0	0	0
	50	FPE	41	<b>35</b>	19	5	0
		MDL	46	<b>35</b>	17	2	0
		MAA	8	<b>80</b>	12	0	0
		FPE-A	44	<b>56</b>	0	0	0
		MDL-A	47	<b>53</b>	0	0	0

### 3.4 Comentários Finais

Foi demonstrado neste capítulo um novo método que utiliza da análise dos autovalores para seleção de ordem de modelos lineares. Este método foi nomeado MAA e para a sua criação foi realizada uma análise matemática do sobredimensionamento de um modelo.

Os resultados deste capítulo demonstram que o MAA possui uma tendência de sobredimensionar a ordem de modelos lineares em uma unidade. Quando o MAA foi utilizado em conjunto com os métodos baseados em critérios de informação, ele diminuiu a tendência de sobredimensionamento de ordem destes métodos. A melhora de desempenho da utilização conjunta do MAA com os métodos baseados em critérios de informação pode ser observada em todos os modelos simulados.

# Capítulo 4

## Método de Análise dos Autovalores aplicados a redes neurais recorrentes

Seleção de ordem de modelos de entrada-saída de sistemas dinâmicos é um passo fundamental no processo de identificação de sistemas. Diversos métodos foram criados para modelos lineares, contudo, quando o modelo é não-linear poucas pesquisas tem sido realizadas (Lin et al., 1997). Demonstra-se, neste capítulo, a utilização do MAA em modelos neurais recorrentes de entrada-saída para estimar o número de atrasos da saída. Como o MAA utiliza de análise dos autovalores, é demonstrado neste capítulo um procedimento para auxiliar na determinação dos autovalores de redes neurais recorrentes de entrada-saída.

### 4.1 Autovalores de redes neurais recorrentes de entrada-saída

Quando se faz referência a autovalores de redes neurais está se referindo a autovalores do modelo neural linearizado através do jacobiano, em torno do ponto de operação. É demonstrado a seguir um procedimento para obtenção destes autovalores.

Um modelo NARMAX pode ser representado pela equação (4.1):

$$\begin{aligned} y(k) = & F[y(k-1), y(k-2), \dots, y(k-n_y), \\ & u(k-1), \dots, u(k-n_u), \\ & e(k-1), \dots, e(k-n_e)] + e(k), \end{aligned} \quad (4.1)$$

sendo que  $u(k)$  e  $y(k)$  representam, respectivamente, a entrada e a saída nos instantes discre-

tos  $k$  ( $k = 1, 2, \dots$ ),  $e(k)$  representa um ruído independente não-observável,  $n_u$ ,  $n_y$  e  $n_e$  são, respectivamente, os máximos atrasos dos termos da entrada, saída e ruído. Neste trabalho,  $F$  representa uma função não-linear mapeada pela rede neural.

O procedimento adotado para obter os autovalores de modelos neurais recorrentes é: (i) representar o modelo na forma de equação de estados, linearizando-o em torno dos pontos fixos através da matriz jacobiana, (ii) a partir dessa matriz montar a equação característica, determinando seus autovalores em função do ponto de operação e, (iii) mapear os autovalores do modelo discreto no plano  $z$ .

Para deduzir a função de autovalores, será considerada apenas a parte determinística do modelo (4.1), ou seja o modelo NARX, representado no espaço de estados linearizando-o em torno de cada estado estacionário. Chamando  $[y(k - n_y) \ y(k - n_y + 1) \ \dots \ y(k - 2) \ y(k - 1)]^T$  de vetor de estado, a Equação (4.1) pode ser linearizada na forma:

$$\begin{bmatrix} y(k - n_y + 1) \\ y(k - n_y + 2) \\ \vdots \\ y(k - 1) \\ y(k) \end{bmatrix} = D_f \begin{bmatrix} y(k - n_y) \\ y(k - n_y + 1) \\ \vdots \\ y(k - 2) \\ y(k - 1) \end{bmatrix} + BU(k), \quad (4.2)$$

sendo,

$$D_f = \begin{bmatrix} 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ \Delta_d & \Delta_{d-1} & \dots & \Delta_2 & \Delta_1 \end{bmatrix}, \quad (4.3)$$

$$\Delta_d = \left. \frac{\partial y(k)}{\partial y(k-d)} \right|_{\bar{y}, \bar{u}}, \quad (4.4)$$

e  $d$  é o atraso da saída  $y$ . Note-se que a expressão de  $\Delta_d$  é determinada em função da derivada parcial, avaliada para  $y = \bar{y}$  e  $u = \bar{u}$ , ou seja, no ponto de operação. Sendo que  $\bar{y}$  e  $\bar{u}$  representam o estado estacionário do sistema, ponto em que  $y(k-1) = y(k-2) = \dots = y(k-n_y) = \bar{y}$

e  $u(k-1) = u(k-2) = \dots = u(k-n_y) = \bar{u}$ .

Com objetivo de se obter uma expressão geral para  $\Delta_d$  é utilizado o modelo neural recorrente ilustrado na Figura 4.1.

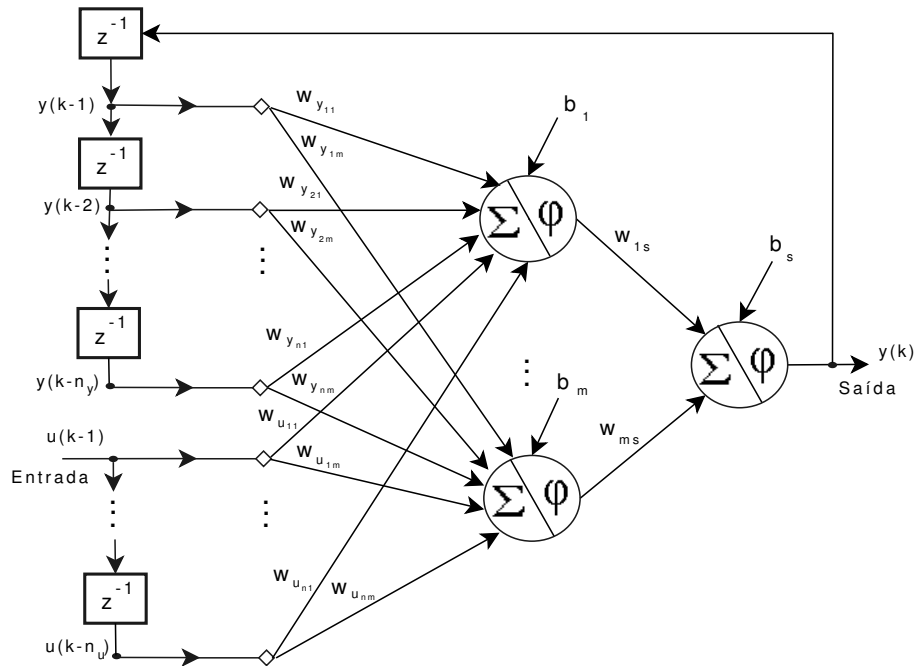


Figura 4.1: Modelo neural recorrente entrada-saída

A saída linear é representado por  $\mu_1$  à  $\mu_m$  para os neurônios da camada escondida e  $\mu_s$  para o neurônio da saída, devendo este possuir uma função de ativação linear. Sendo que  $\mu_m$  representa a saída do neurônio  $m$ . As equações são representadas abaixo:

$$\begin{aligned}
 \mu_1 &= y(k-1)w_{y_{11}} + y(k-2)w_{y_{21}} + \dots + \\
 &\quad y(k-n_y)w_{y_{n1}} + u(k-1)w_{u_{11}} + \dots + \\
 &\quad u(k-n_u)w_{u_{n1}} + b_1 \\
 \mu_m &= y(k-1)w_{y_{1m}} + y(k-2)w_{y_{2m}} + \dots + \\
 &\quad y(k-n_y)w_{y_{nm}} + u(k-1)w_{u_{1m}} + \dots + \\
 &\quad u(k-n_u)w_{u_{nm}} + b_m \\
 \mu_s &= \varphi(\mu_1)w_{1s} + \dots + \varphi(\mu_m)w_{ms}
 \end{aligned} \tag{4.5}$$

$$\begin{aligned}
\Delta_1 &= \frac{\partial y(k)}{\partial y(k-1)} = \varphi'(\mu_1)w_{y_{11}}w_{1s} + \dots + \varphi'(\mu_m)w_{y_{1m}}w_{ms} \\
\Delta_2 &= \frac{\partial y(k)}{\partial y(k-2)} = \varphi'(\mu_1)w_{y_{21}}w_{1s} + \dots + \varphi'(\mu_m)w_{y_{2m}}w_{ms} \\
\Delta_{n_y} &= \frac{\partial y(k)}{\partial y(k-n_y)} = \varphi'(\mu_1)w_{y_{n_1}}w_{1s} + \dots + \varphi'(\mu_m)w_{y_{nm}}w_{ms}
\end{aligned} \tag{4.6}$$

Para  $d = 1, \dots, n_y$  a Equação (4.4) pode ser escrita,

$$\Delta_d = \begin{bmatrix} \varphi'(\mu_1) & \dots & \varphi'(\mu_m) \end{bmatrix} \begin{bmatrix} w_{y_{d1}}w_{1s} \\ \vdots \\ w_{y_{dm}}w_{ms} \end{bmatrix} \tag{4.7}$$

sendo as raízes da Equação (4.8) os autovalores do modelo neural recorrente no ponto de operação  $(\bar{u}, \bar{y})$ .

$$\lambda^d - \Delta_1\lambda^{d-1} - \dots - \Delta_{d-1}\lambda - \Delta_d = 0, \tag{4.8}$$

## 4.2 Procedimento para seleção de atrasos de saída de redes neurais

No Capítulo 3 foi introduzido o conceito de um novo método para seleção de ordem de modelos lineares, chamado MAA. o MAA pode ser, também, utilizado em modelos não-lineares. Busca-se nesse trabalho uma extensão da aplicação do MAA em redes neurais recorrentes. Para isso, o procedimento demonstrado na Seção 3.2 será aplicado aos modelos neurais recorrentes com a seguinte observação:

- O procedimento deverá ser utilizado a um modelo neural recorrente de entrada-saída com apenas 1 (um) neurônio oculto para a identificação do processo.

Foi observado que a variação do número de neurônios influenciou nos resultados apresentados pelo MAA, ajustando uma outra localização para os autovalores, por este motivo, é sugerido fixar o número de neurônios oculto da rede neural em 1 (um).

Após utilizado do procedimento para identificação do número de atrasos da saída do modelo neural recorrente é necessário realizar o treinamento da rede neural ajustando o número de neurônios escondidos para uma boa capacidade de generalização e validação adequada do modelo.

### **4.3 Considerações Finais**

Neste capítulo, um procedimento para determinação de autovalores de modelos neurais recorrentes de entrada-saída foi discutido, com o objetivo de possibilitar a utilização do MAA na seleção do número de atrasos de saída. No próximo capítulo será aplicado o procedimento aqui proposto.



## Capítulo 5

# Aplicação do método de análise dos autovalores para modelos neurais recorrentes de entrada-saída

São poucos os procedimentos que tratam da seleção de atraso em redes neurais, no geral o atraso é definido de forma empírica. O método apresentado nesta dissertação tem como principal objetivo a seleção de atraso de modelos não-lineares, incluindo as redes neurais. Para analisar a eficiência do MAA na seleção de atraso em redes neurais serão utilizados, neste capítulo, dois sistemas, sendo um sistema simulado e outro real. O objetivo é demonstrar a eficiência do MAA na seleção de atrasos de saída em modelos neurais recorrentes. Por meio dos sistemas simulados são analisados a robustez do método à presença de ruídos nos dados e a capacidade de recuperar a posição dos autovalores usando a representação neural. Por meio do sistema real busca-se demonstrar a utilização prática do método em dados coletados a partir do funcionamento de uma planta didática.

Em aplicações reais geralmente as saídas dos sistemas são contaminadas por algum tipo de ruído. Para reduzir o efeito do ruído no treinamento da rede neural é utilizado a representação NARMAX, demonstrado na Figura 2.4.

Em todos os dois casos é utilizado procedimento de identificação caixa-preta, ou seja, a rede neural é treinada a partir de dados de entrada e saída. Para seleção de ordem dos modelos foi utilizada de uma rede neural MLP NARMAX com a seguinte estrutura:

- A rede neural possui duas camadas de neurônios;

- O número de neurônios da primeira camada (escondida) foi fixada em 1 (conforme sugerido pelo procedimento) e a função de ativação utilizada é a tangente hiperbólica;
- Para a segunda camada (saída) o número de neurônios foi fixado em 1 com a função de ativação linear;
- Para treinamento da rede neural foi utilizado da função que utiliza o algoritmo gradiente conjugado escalonado;

Para esta dissertação a escolha de  $n_e$  não é crítico e pode ser escolhido um valor suficiente para garantir que os resíduos sejam brancos. Para todas as simulações realizadas observou-se que valores de  $n_e = 10$  é o suficiente.

## 5.1 Aplicação do MAA em dados obtidos a partir de um caso teórico simulado

Nesta seção é apresentado o desempenho do MAA na seleção de atrasos de modelos neurais utilizando de um sistema teórico simulado. Investiga-se também o efeito de ruído aditivo. O sistema teórico simulado é um sistema de primeira ordem representado por Haber e colegas [1990]:

$$[1 + \alpha u(t)] \frac{dy}{dt} + y(t) - u(t) = 0, \quad (5.1)$$

ou seja, um sistema cujo ganho é constante e a constante de tempo varia com o ponto de operação na seguinte forma  $\tau(u) = 1 + \alpha u$ . A Equação (5.1) foi simulada com  $\alpha = 0,5$ , gerando massas de dados para identificação e para validação.

A partir da massa de dados gerada para identificação, originou-se mais duas massas de dados, sendo que nestas duas últimas foi adicionado ruído, simulando erro de medição. Então, para o processo de estimação de parâmetros será usado: (i) Massa de dados *dad\_ctv1* - sinal com ruído adicionado  $y_2(k) = y(k) + \epsilon$ , sendo que  $\epsilon$  possui distribuição gaussiana de média zero e variância  $\sigma^2 = 0,01$ ; (ii) Massa de dados *dad\_ctv2* - sinal com ruído adicionado  $y_3(k) = y(k) + \epsilon$ , sendo que  $\epsilon$  possui distribuição gaussiana de média zero e variância  $\sigma^2 = 0,05$ . Desta forma, verificou-se o efeito do ruído na estimação de parâmetros. Trechos dos sinais usados na identificação estão apresentados nas figuras 5.1 e 5.2, sendo que os dados apresentados na Figura 5.1 foram utilizadas para treinamento da rede neural e os dados apresentados na Figura 5.2 utilizados para validação.

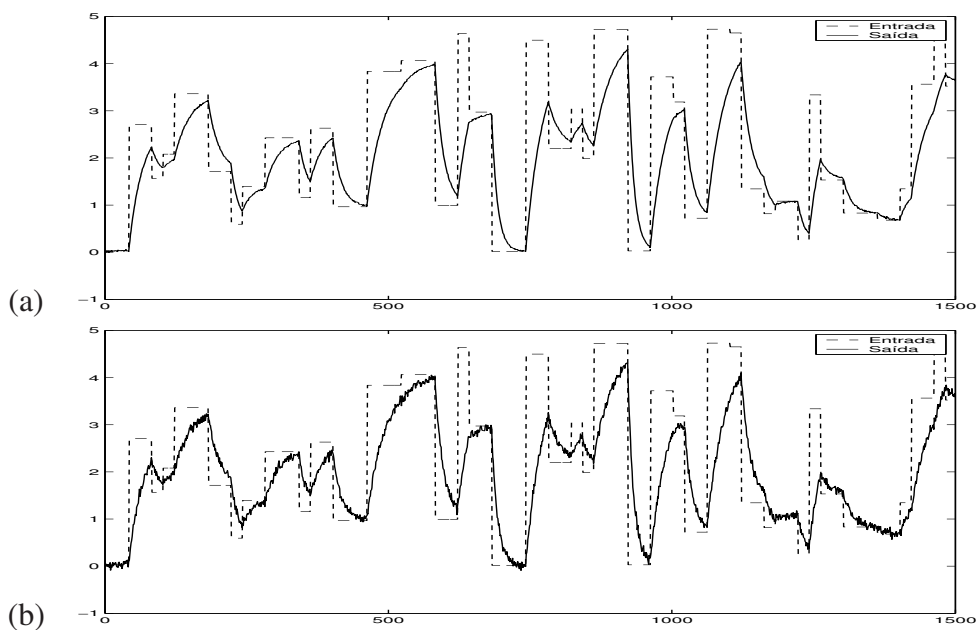


Figura 5.1: Dados de identificação simulados a partir da equação (5.1) utilizados para treinamento da rede neural. (a) Dados *dad\_ctv1* e (b) dados *dad\_ctv2*.

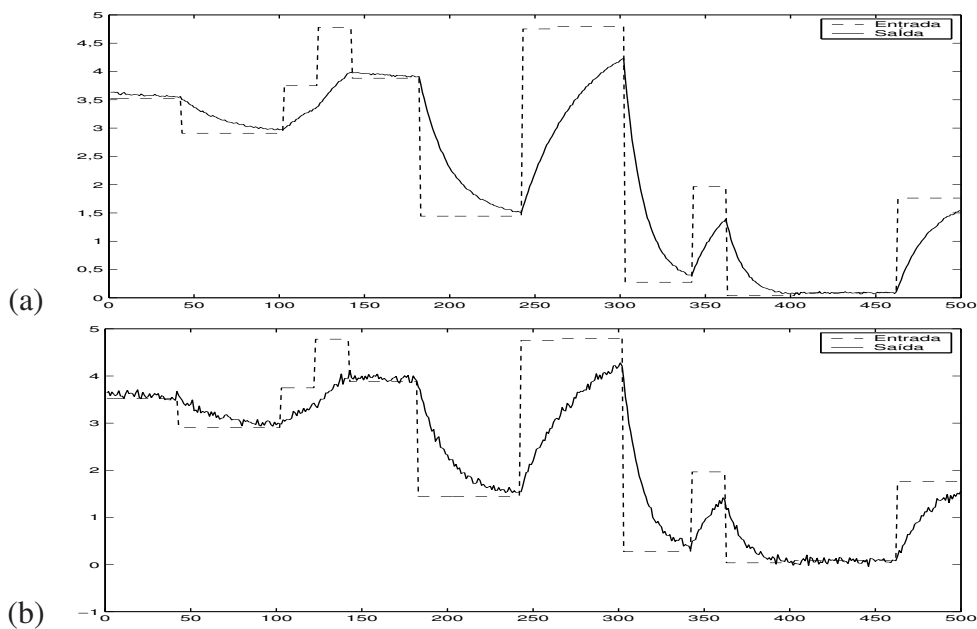


Figura 5.2: Dados de identificação simulados a partir da equação (5.1) utilizados para validação da rede neural. (a) Dados *dad\_ctv1* e (b) dados *dad\_ctv2*.

Como pode ser observado, a entrada possui amplitude aleatória, excursionando o sistema em toda a sua faixa de operação com duração de cada patamar também aleatório. O tempo de amostragem usado foi 0,1 segundos e a massa de dados total possui 2000 amostras, sendo que 1500 amostras foram utilizadas para treinamento da rede neural e o restante para validação.

Com os dados de treinamento  $dad\_ctv1$  e  $dad\_ctv2$  foi utilizado do MAA para estimar o número de atraso de saída ( $n_y$ ) dos modelos neurais obtidos. O valor estimado pelo MAA foi  $n_y = 1$  em ambos os casos. Esta conclusão pode ser obtido analisando as Figuras 5.3 e 5.4.

As Figuras 5.3(a) e 5.4(a) demonstram os autovalores de modelos neurais, utilizando  $n_y = 1$  e massa de dados  $dad\_ctv1$  e  $dad\_ctv2$ , respectivamente. Todos os autovalores destes modelo encontram-se no semicírculo direito do plano  $z$ . Após incluir um termo de atraso de saída e desenvolver a análise dos autovalores no plano  $z$ , os novos resultados são demonstrados nas Figuras 5.3(b) e 5.4(b), na qual apresentam 1 autovalor no semicírculo esquerdo do plano  $z$ .

Com os resultados apresentados nas Figuras 5.3 e 5.4 é possível perceber que o MAA selecionou corretamente o número de atrasos de saída da rede neural, pois o valor de  $n_y$  da Equação (5.1) é igual a 1, o mesmo obtido pelo MAA.

Após obtido  $n_y$  é necessário realizar o treinamento da rede neural ajustando o número de neurônios escondidos para obter a estrutura de rede neural com uma boa capacidade de generalização e validação adequada do modelo.

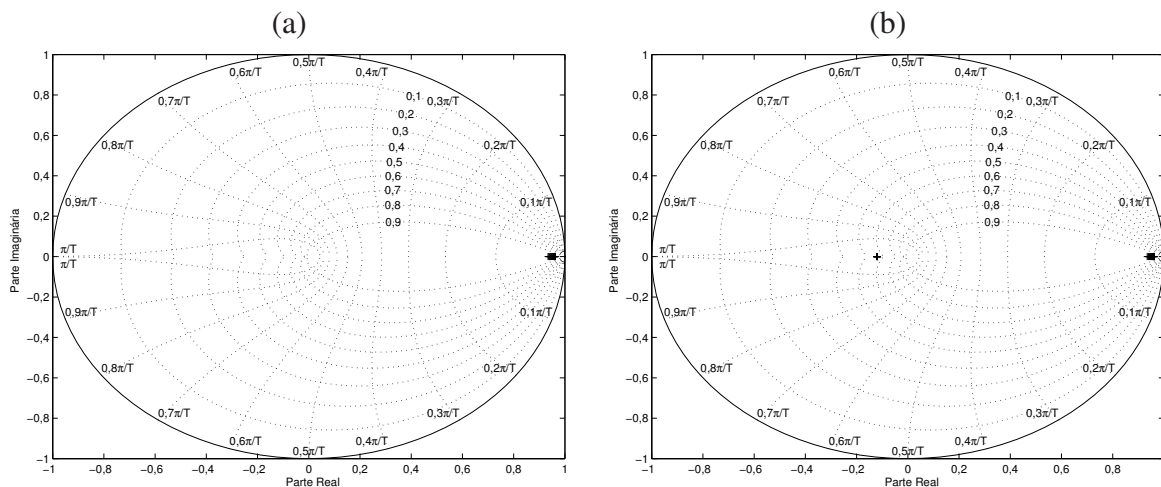


Figura 5.3: Autovalores de modelos neurais NARMAX utilizando massa de dados  $dad\_ctv1$ . (a) Modelo neural com  $n_y = 1$ , todos os autovalores encontram-se no semicírculo direito do plano  $z$  (b) Modelo neural com  $n_y = 2$ , possui 1 autovalor no semicírculo esquerdo do plano  $z$

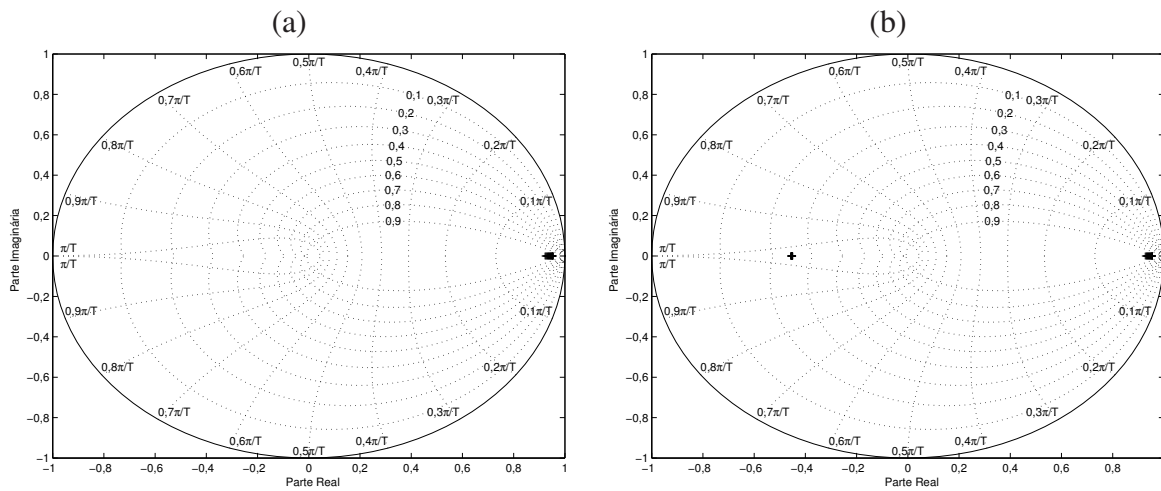


Figura 5.4: Autovalores de modelos neurais NARMAX utilizando massa de dados *dad\_ctv2*. (a) Modelo neural com  $n_y = 1$ , todos os autovalores encontram-se no semicírculo direito do plano  $z$  (b) Modelo neural com  $n_y = 2$ , possui 1 autovalor no semicírculo esquerdo do plano  $z$

Com intuito de obter a estrutura da rede neural que possibilita aproximar o comportamento dinâmico observado nos dados gerados pela Equação (5.1) foram utilizadas diversas redes neurais e comparado com o valor RMSE, apresentados na Tabela (5.1). Por meio do RMSE, pode-se constatar que os melhores modelos obtidos foram os que utilizam  $n_y = 1$ .

Para o cálculo do RMSE, foram utilizados dados de validação e comparado com os valores de predição livre obtidos pela rede neural. Foram variados o número de neurônios ( $nn$ ) e o número de atraso de saída ( $n_y$ ). O número de atraso de entrada ( $n_u$ ) foi fixado em 1.

Tabela 5.1: Valor RMSE variando o  $n_y$ ,  $nn$  e os dados de validação do caso teórico simulado

Dados	$n_y$	$nn$			
		2	3	4	5
<i>dad_ctv1</i>	1	0,4827	<b>0,4198</b>	0,4300	0,4650
	2	0,4836	0,4624	0,4848	0,5120
	3	0,5710	0,5310	0,5760	0,5990
<i>dad_ctv2</i>	1	0,4819	0,4810	<b>0,4006</b>	0,4720
	2	0,4088	0,4768	0,4364	0,4928
	3	0,5780	0,5608	0,5540	0,5860

Para a massa de dados *dad\_ctv1* a rede neural que mais se aproximou dos dados de validação possui  $nn = 3$  e  $n_y = 1$ . Para a massa de dados *dad\_ctv2* foi obtido  $nn = 4$  e  $n_y = 1$ .

Para verificar a eficiência da rede neural na estimação de parâmetros, foi observado o comportamento no gráfico da constante de tempo em função do ponto de operação para as redes neurais que obtiveram menor valor RMSE. O resultado é demonstrado nas figuras 5.5(a) e 5.5(b), na qual apresentam graficamente uma comparação dos valores obtidos pelo sistema simulado e pelas redes neurais.

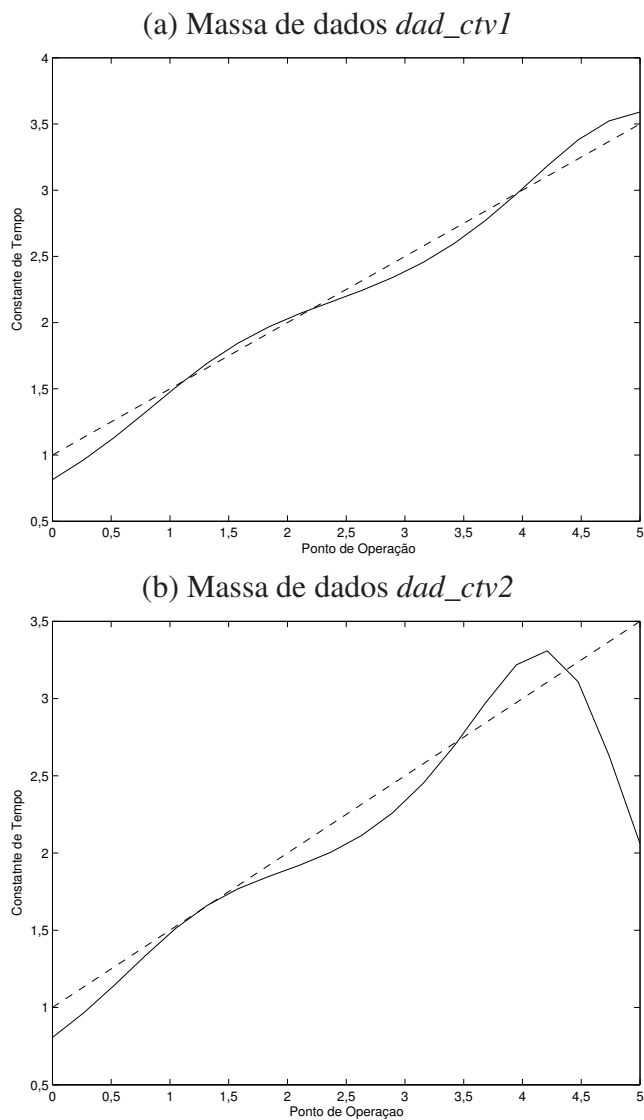


Figura 5.5: Gráfico da contante de tempo em função do ponto de operação dos modelos neurais obtidos apartir da simulação da equação (5.1). (- -)Sistema simulado (-)Modelos neurais.

Analisando as figuras 5.5(a) e 5.5(b), pode-se concluir que a rede neural conseguiu recuperar de

forma satisfatória a constante de tempo dos dados  $dad\_ctv1$  e  $dad\_ctv2$ , pois não foi observado diferença significativa nos valores de constante de tempo obtidos através do sistema simulado com os valores de constante de tempo estimados pela rede neural. É importante ressaltar que o aumento do ruído ocasionou em afastamento na curva da constante de tempo. Esse afastamento pode ser explicado analisando a Figura 5.6, onde demonstra poucos dados contidos entre os valores 4,5 e 5.

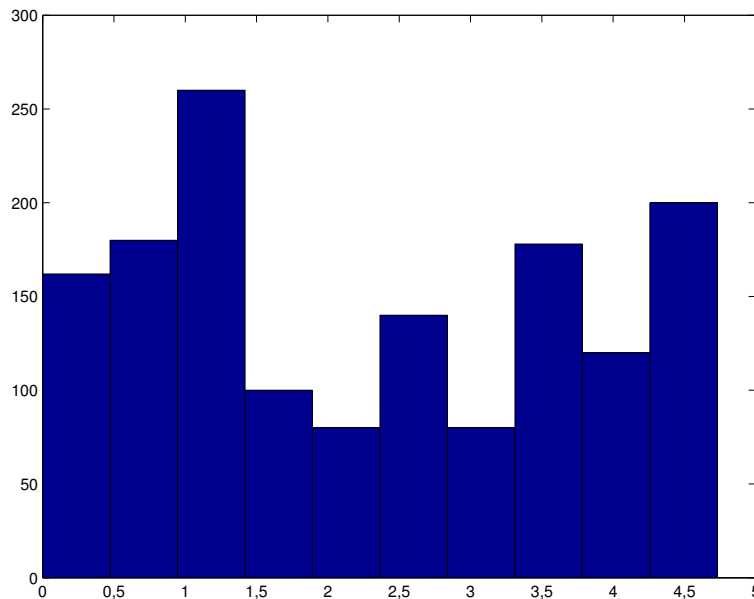


Figura 5.6: Histograma do sinal de entrada  $u(t)$  dos dados de identificação da massa de dados  $dad\_ctv2$

A seguir são demonstrados os resultados da predição livre das redes neurais que apresentaram menor valor RMSE. Os resultados apresentados confirmam a eficiência do MAA na escolha correta de  $n_y$ . Através das figuras 5.7(a) e 5.7(b) pode-se verificar que as duas redes neurais escolhidas ajustaram adequadamente o comportamento observado pelos dados de validação. Para este exemplo conclui-se que o MAA selecionou corretamente o valor real de  $n_y$ .

Embora não apresentados nesta dissertação, testes com outros dados obtidos a partir de casos teóricos simulados obtiveram resultados semelhantes.

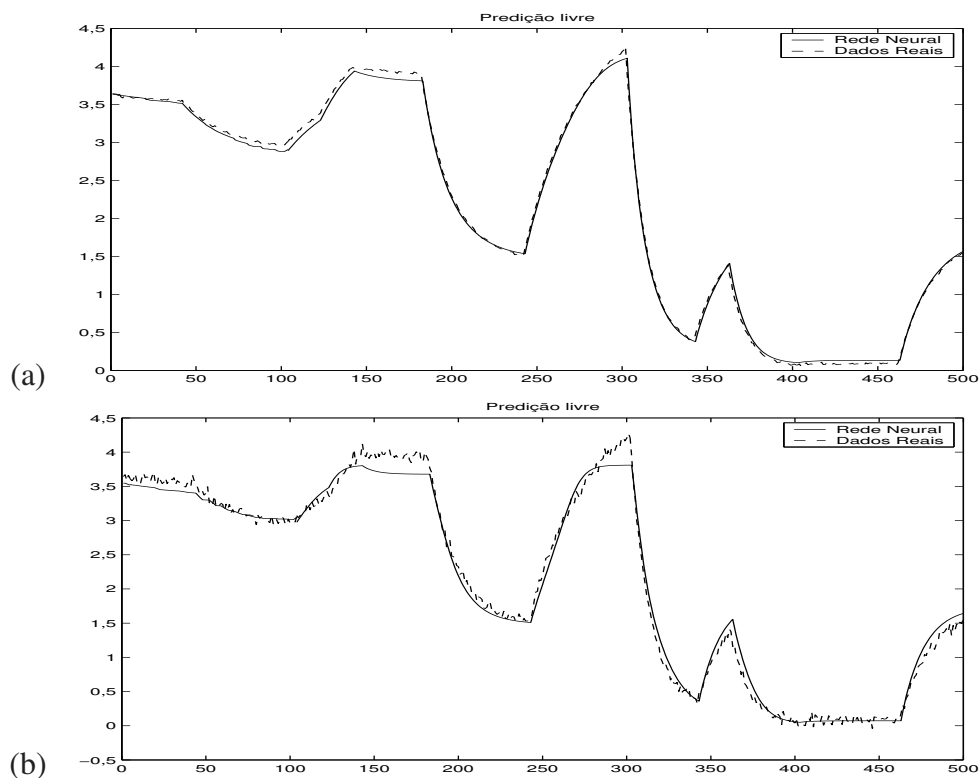


Figura 5.7: Predição livre utilizando modelo neural NARMAX. (a) dados *dad\_ctv1*,  $nn = 3$ ,  $n_y = 1$  e  $n_u = 1$  e (b) dados *dad\_ctv2*,  $nn = 4$ ,  $n_y = 1$  e  $n_u = 1$ .

## 5.2 Aplicação do MAA em dados obtidos a partir de um processo térmico

Com o objetivo de verificar o desempenho do MAA a partir de dados reais é utilizado nesta seção dados de um processo térmico que se refere a um forno elétrico do Laboratório de Controle de Processos Industriais (LCPI) do centro de Pesquisa e Desenvolvimento em Engenharia Elétrica (CPDEE) (Abreu, 1993). Essa é uma caixa metálica com dimensões 15x10x31 cm, construído a partir de uma chapa de alumínio com espessura igual a 2 mm. O elemento de aquecimento interno é uma lâmpada elétrica de 200 W. O forno não é isolado termicamente, de modo que variações na temperatura ambiente afetam o seu comportamento dinâmico.

Os dados usados na identificação dos modelos aqui apresentados são: Figura 5.8 - respostas ao ruído quantizado e Figura 5.9 - respostas ao degrau.



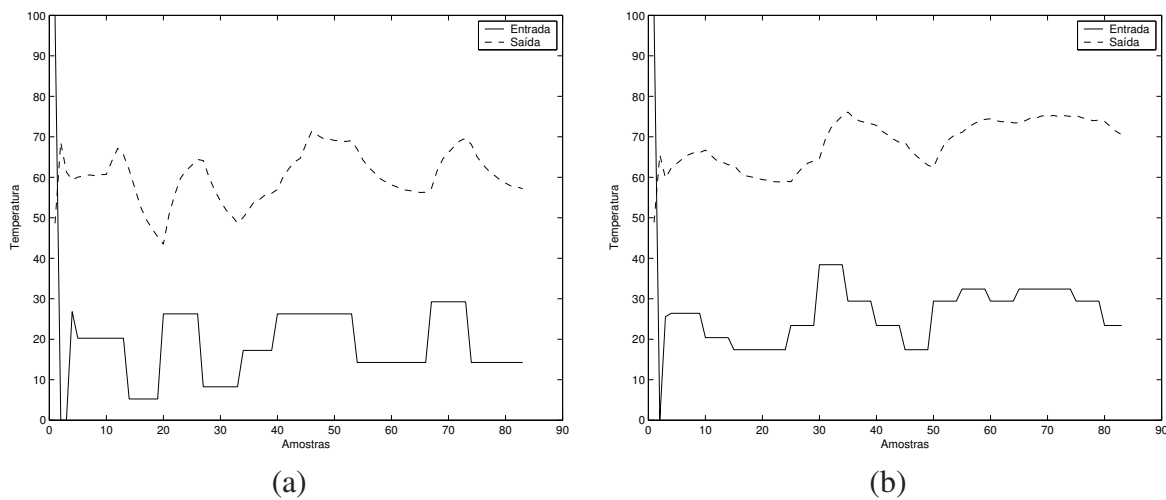


Figura 5.8: Resposta aos ruído quantizado de um Processo Térmico. Eixo- $x$  são as amostras. Linha contínua é a entrada e a tracejada é a temperatura de saída, ambos normalizados de 0 a 100%. (a) Massa de dados  $frq1$  e (b) massa de dados  $frq2$

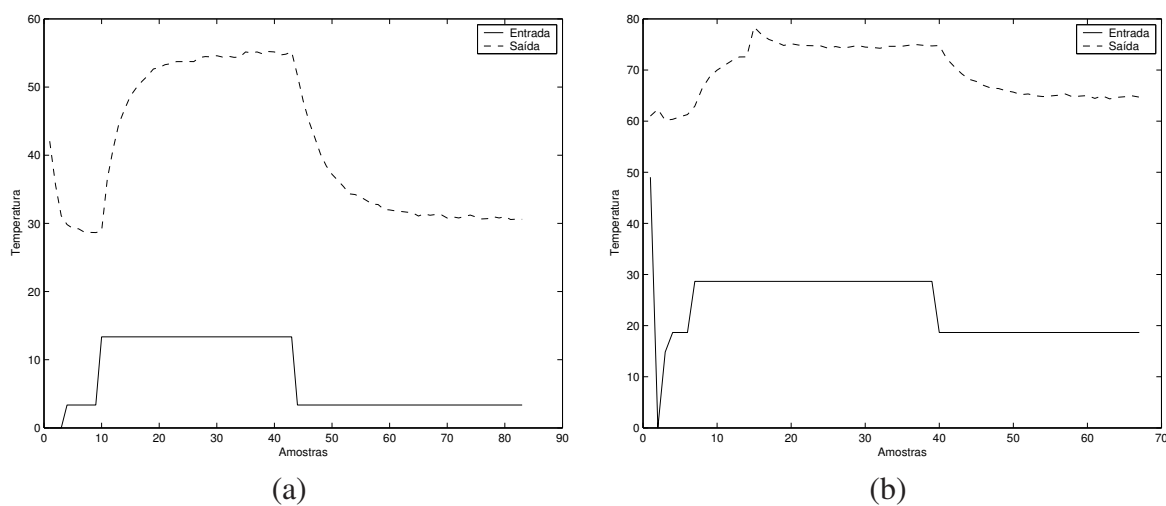


Figura 5.9: Resposta ao degrau de um Processo Térmico. Eixo- $x$  são as amostras. Linha contínua é a entrada e a tracejada é a temperatura de saída, ambos normalizados de 0 a 100%. (a) Massa de dados  $fd1$  e (b) massa de dados  $fd2$

O sinal de saída é a temperatura da superfície externa do forno, normalizado em relação a uma temperatura máxima possível e a entrada é o sinal que define a tensão aplicada ao comando de um controlador que controla a potência média fornecida à lâmpada. O intervalo de amostragem utilizado foi de  $T_s = 210s$  e o eixo  $x$  na figura, indica o número de amostras. A figura 5.9

revela que o forno possui constante de tempo de aquecimento diferente da constante de tempo de resfriamento.

Outra característica observada nas massas de dados  $fd1$  e  $fd2$  relatada em Fernandes (2006), é o lento e contínuo aumento da temperatura da saída do sistema enquanto o sinal de controle é mantido constante, intervalo  $25 < k < 40$ . Este fato mostra que o sistema possui duas constantes de tempo de aquecimento, uma rápida e outra lenta.

Outra observação importante diz respeito à região excitada por cada massa de dados. Na massa de dados  $frq1$  a temperatura do sistema oscila entre 50 e 70% do valor normalizado, em  $fd1$  está entre 30 a 50%, em  $fd2$  entre 70 a 80% e em  $frq2$  está entre 60 a 75%. Esta observação é importante, pois a massa de dados  $frq1$  foi utilizada para identificação dos modelos e as demais massas de dados foram utilizadas na validação dinâmica dos mesmos.

Utilizando do MAA para estimar o número de atrasos da saída de um modelo neural, através da massa de dados  $frq1$ , obtém-se  $n_y = 2$ . Esta conclusão pode ser obtida através da análise da Figura 5.10. A Figura 5.10(a) mostra os autovalores de um modelo neural utilizando  $n_y = 2$ , todos os autovalores deste modelo encontram-se no semicírculo direito do plano  $z$ . Após incluir um termo de atraso de saída e desenvolver a análise dos autovalores no plano  $z$ , o novo resultado é mostrado na Figura 5.10(b). Este modelo apresenta autovalor no semicírculo esquerdo do plano  $z$  e conforme o MAA este é um indicativo de número de atrasos de saída sobredimensionado.

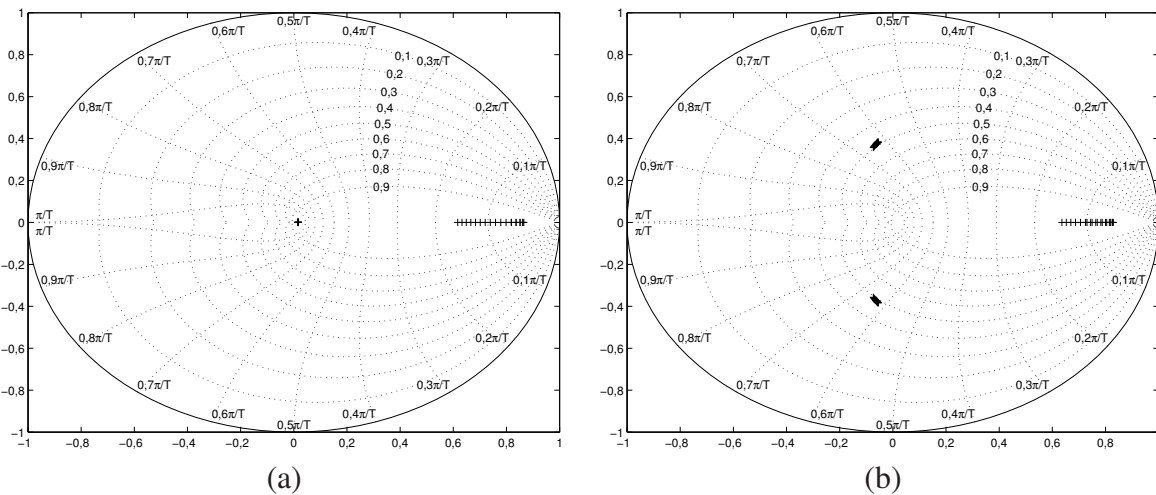


Figura 5.10: Autovalores de modelos neurais recorrentes utilizando da massa de dados  $frq1$ . (a) Modelo neural com  $n_y = 2$ . Todos os autovalores encontram-se no semicírculo direito do plano  $z$  (b) Modelo neural com  $n_y = 3$ . Possui autovalor no semicírculo esquerdo do plano  $z$ .

Ajustando o número de neurônios escondidos e utilizando dos dados de validação pode-se constatar que os melhores modelos obtidos foram os que utilizam 2 atrasos de saída. Esta afirmação pode ser comprovada pelos valores RMSE, apresentados na Tabela (5.2). Os menores valores de RMSE, destacados em negrito, são para redes neurais NARMAX que utilizam de 2 atrasos de saída.

Para o cálculo do RMSE, foram utilizados valores da saída na predição livre do modelo neural recorrente de entrada-saída, variando  $nn$ ,  $n_y$  e os dados de validação. O valor  $n_u$  foi fixado em 1.

Tabela 5.2: Valor RMSE variando o  $n_y$ ,  $nn$  e os dados de validação do processo térmico

Dados	$n_y$	$nn$			
		1	2	4	8
<i>frq2</i>	1	0,5154	0,5546	0,6305	0,6452
	2	<b>0,3836</b>	0,6635	0,6722	0,6904
	3	0,4526	0,5918	0,6208	0,6408
<i>fd1</i>	1	0,4281	0,8298	0,9104	0,9950
	2	<b>0,2178</b>	0,8914	0,9905	1,124
	3	0,4600	1,1851	1,212	1,345
<i>fd2</i>	1	0,9285	0,9768	0,9912	1,038
	2	0,8982	<b>0,8852</b>	0,9608	1,008
	3	0,9768	1,003	1,108	1,320

Consultando a Tabela (5.2) pode-se verificar que para os dados *frq2* e *fd1* as redes neurais que apresentaram menor valor RMSE foram as que continham  $nn = 1$  e  $n_y = 2$  e para os dados *fd2* foi a rede neural com  $nn = 2$  e  $n_y = 2$ .

No geral a rede neural que mais se aproximou dos dados de validação foi a rede que possui  $nn = 1$  e  $n_y = 2$ . Para analisar o desempenho desta rede neural foi realizada a predição livre e o resultado é mostrado na Figura 5.11. Para a figura 5.11(a) utilizou-se dos dados *frq1*, dados utilizados para treinamento, e para as Figuras 5.11(b), 5.11(c) e 5.11(d) utilizou-se os dados *frq2*, *fd1* e *fd2*, respectivamente.

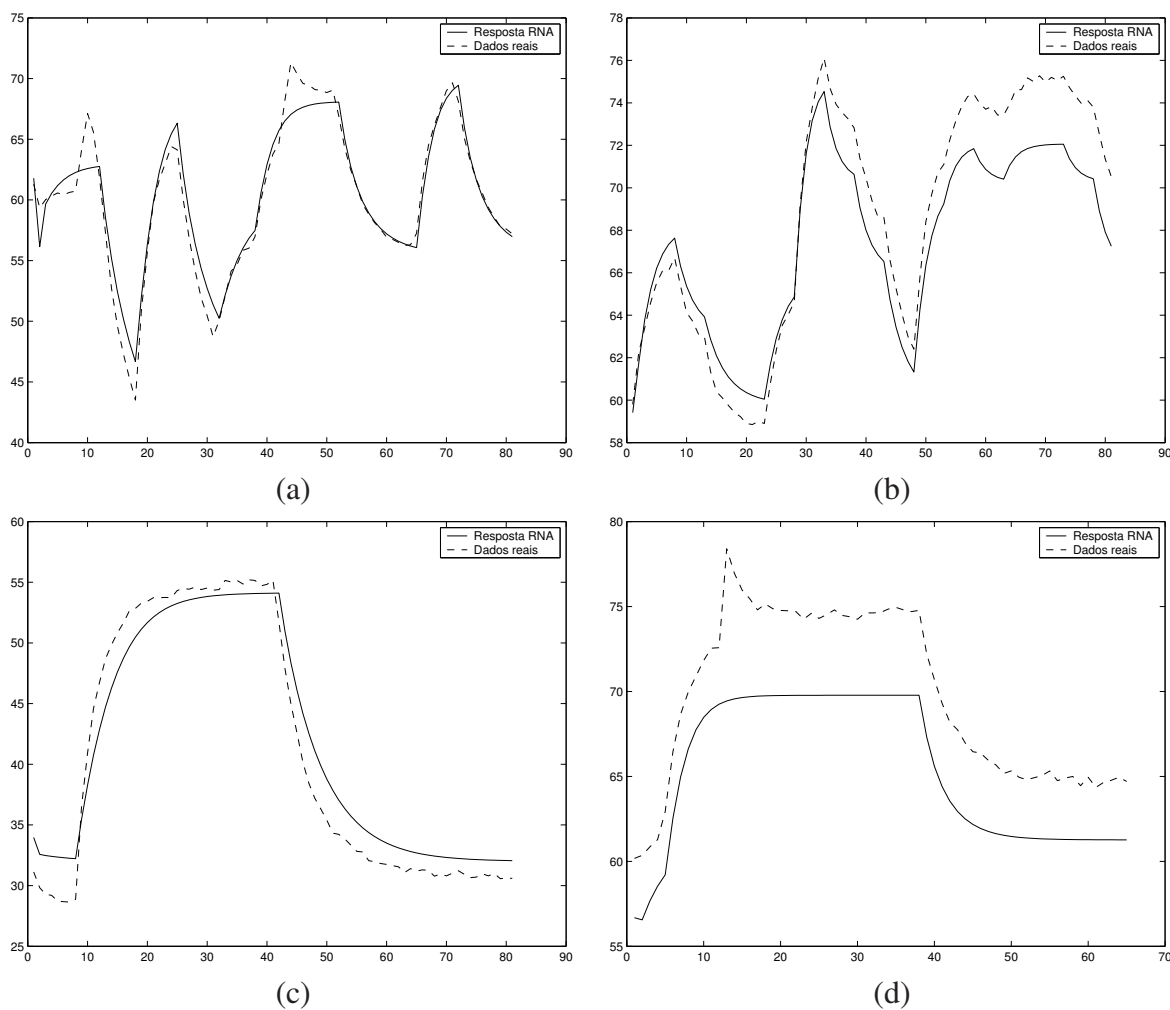


Figura 5.11: Predição livre utilizando de um modelo neural recorrente com  $nn = 1$ ,  $n_y = 2$  e  $n_u = 1$ .  
 (a) massa de dados *frq1* (b) massa de dados *frq2* (c) massa de dados *fd1* (d) massa de dados *fd2*.

Nota-se que as tendências das predições livres apresentadas nas Figuras 5.11(a), 5.11(b), 5.11(c) e 5.11(d) foram seguidas, porém é possível perceber um problema de ponto fixo nas Figuras 5.11(b) e 5.11(d). Uma explicação deste fato deve-se ao baixo número de padrões dos dados de treinamento.

Para este exemplo conclui-se que o MAA apresentou uma boa aproximação do valor real de  $n_y$  e embora não apresentados nesta dissertação, testes com outros dados obtidos a partir de sistemas reais obtiveram resultados semelhantes.

## 5.3 Considerações Finais

Neste capítulo, é demonstrado o desempenho do MAA utilizando de dados obtidos a partir de dados obtidos por um caso teórico simulado e de dados um processo térmico. O processo térmico em questão, se refere a um forno elétrico do Laboratório de Controle de Processos Industriais (LCPI) do centro de Pesquisa e Desenvolvimento em Engenharia Elétrica (CPDEE).

Os resultados deste capítulo demonstram que o MAA é uma ferramenta eficaz no processo de seleção de atrasos de saída de modelos neurais recorrentes entrada-saída. Resultados satisfatórios foram obtidos e demonstrados tanto para dados obtidos pelo caso teórico simulado quanto para dados obtidos de um processo térmico.

De uma maneira geral, pode-se afirmar que o MAA se mostrou bastante eficiente e sua utilização pode ser de grande ajuda na determinação da estrutura de um modelo neural.

# Capítulo 6

## Conclusão

Apresentou-se um novo critério de seleção de ordem chamado MAA. Este utiliza da análise dos autovalores para seleção de ordem de modelos lineares e não-lineares.

Para modelos lineares pode-se verificar que através da utilização conjunta do FPE e MDL com o MAA foi possível melhorar os resultados obtidos pelos métodos FPE e MDL. A utilização conjunta do FPE e MDL com MAA, aqui chamados FPE-M e MDL-M respectivamente, diminui a tendência do sobredimensionamento de ordem dos modelos apresentados pelos métodos FPE e MDL em algumas realizações. Em todas as simulações pode-se observar que os resultados apresentados pelo FPE-M e MDL-M são superiores aos apresentados pelo FPE e MDL.

Para modelos não-lineares, representado nesse trabalho por modelos neurais recorrentes, pode-se observar que para dados obtidos a partir de um caso teórico simulado o MAA apresentou bom resultado. Os modelos neurais que melhor se aproximaram aos dados foram os que utilizaram o número de atrasos de saída estimados pelo MAA. Resultado satisfatório foi, também, obtido utilizando dados de um processo térmico. Nesse caso, o MAA selecionou corretamente o número de atrasos de saída.

Pelos resultados apresentados pode-se afirmar que o MAA é uma ferramenta de simples aplicação e que pode desempenhar um importante papel na seleção de ordem de modelo linear e não-linear.

### 6.1 Sugestões para trabalhos futuros

Como propostas de trabalhos futuros destacam-se:

1. Aperfeiçoar o MAA na estimação de ordem em modelos lineares de forma a fazer com que não haja a necessidade de utilização conjunta de um segundo método para a obtenção de resultados melhores do que os apresentados pelos métodos baseados em critério de informação;
2. Verificar para a rede neural se, além do número de atrasos de saída, é possível obter alguma outra relação da análise dos autovalores com a estrutura da rede neural, como por exemplo o número de neurônios;
3. Utilizar o MAA para seleção de ordem de outros modelos não-lineares, além das redes neurais, e analisar a eficiência do método.

# Referências Bibliográficas

- Abreu, S. D. G. (1993). Avaliação de desempenho de controladores auto-ajustáveis. Dissertação de Mestrado, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Aguirre, L. A. (2000). *Introdução a Identificação de Sistemas. Técnicas Lineares e Não Lineares Aplicadas a Sistemas Reais*. Editora UFMG, Belo Horizonte - MG. Brasil, 2ª edição.
- Aguirre, L. A. (2007). *Introdução a Identificação de Sistemas. Técnicas Lineares e Não Lineares Aplicadas a Sistemas Reais*. Editora UFMG, Belo Horizonte - MG. Brasil, 3ª edição.
- Akaike, H. F. (1969). Autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, 21:243–247.
- Akaike, H. F. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, AC-19(6):716–723.
- Akaike, H. F. (1977). On entropy maximization principle. *Applications of Statistics: Proceedings*, pp. 27–41.
- Akaike, H. F. (1978). A bayesian analysis of the minimum aic procedure. *Annals of the Institute of Statistical Mathematics*, 30:9–14.
- Al-Smadi, A. e Al-Zaben, A. (2005). Arma model order determination using edge detection: a new perspective. *Circuits systems signal processing*, 24(6):723–732.
- Castanho, M. C. (1991). Métodos para determinação da estrutura de modelos de sistemas monovariáveis. Dissertação de Mestrado, UNICAMP.
- Corrêa, M. V. (2001). *Identificação caixa-cinza de sistemas não-lineares utilizando representações NARMAX racionais e polinomiais*. Tese de Doutorado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.



- Dyrholm, M., Makeig, S., e Hansen, L. K. (2006). Model structure selection in convolutive mixtures. *Independent Component Analysis and Blind Signal Separation*, 3889:74–81.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14:179–211.
- Feil, B., Abonyi, J., e Szeifert, F. (2004). Model order selection input-output models a clustering based approach. *Journal of Process Control*, pp. 593–602.
- Fernandes, F. S. (2006). Identificação por predição de erro e síntese de controladores robustos. Dissertação de Mestrado, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Fogel, D. B. (1991). An information criterion for optimal neural network selection. *IEEE Transactions on Neural Networks*, 2(5):490–497.
- Gençay, R. e Liu, T. (1997). Nonlinear modelling and prediction with feedforward and recurrent networks. *Physica D*, 108:119–134.
- Ghiassi, M., Saidane, H., e Zimbra, D. (2005). A dynamic artificial neural network model for forecasting time series events. *International Journal of Forecasting*, 21(2):341–362.
- Giles, C. L., Sun, G. Z., Chen, H. H., Lee, Y. C., e Chen, D. (1990). Higher order recurrent networks and grammatical inference. *Advances in Neural Information Processing Systems*, 2:380–387.
- Haber, R. e Unbehauen, H. (1990). Structure identification of nonlinear dynamic systems. a survey on input output approaches. *Automatica*, 26:651–677.
- Henrique, H. M., Lima, E. L., e Pinto, J. C. (1998). A bifurcation study on neural network models for nonlinear dynamic systems. *Latin American Applied Research*, 28:187–200.
- Jordan, M. I. (1986). Serial order: A parallel distributed processing approach. technical report. *Institute for Cognitive Science Report*, page 8604.
- Kashyap, R. (1980). Inconsistency of the aic rule for estimating the order of autoregressive models. *IEEE Transactions on Automatic Control*, AC-25(6):996–998.
- Kay, S. (1987). Modern spectral estimation: Theory and application. *Englewood Cliffs: Prentice-Hall*, pp. 234–237.
- Kim, H.-J. e Cavanaugh, J. E. (2005). Model selection criteria based on kullback information measures for nonlinear regression. *Journal of Statistical Planning and Inference*, pp. 332–349.

- Korenberg, M. (1985). Orthogonal identification of nonlinear difference equation models. *Midwest Symposium on Circuits and Systems*, pp. 303–328.
- Lin, T., Horne, B. G., e Kung, S. Y. (1997). A delay damage model selection algorithm for narx neural networks. *IEEE Transactions on Signal Processing*, 45(11):2719–2730.
- Lombardini, F. e Gini, F. (2005). Model order selection in multi-baseline interferometric radar systems. *EURASIP Journal on Applied Signal Processing*, pp. 3206–3219.
- Madar, J., Abonyi, J., e Szeifert, F. (2005). Genetic programming for the identification of nonlinear input-output models. *Industrial and Engineering Chemistry Research*, 44(9):3178–3186.
- Marple, S. (1987). Digital spectral analysis with applications. *Englewood Cliffs: Prentice-Hall*.
- Narendra, K. S. e Parthasarathly, K. (1990). Identification and control of dynamical systems using neural networks. *IEEE Transactions Neural Networks*, 27:1–4.
- Narendra, K. S. e Parthasarathly, K. (1992). Neural networks and dynamical systems. *International Journal of Approximate Reasoning*, 6:109–131.
- Palaniappan, R. (2006). Towards optimal model order selection for autoregressive spectral analysis of mental tasks using genetic algorithm. *IJCSNS International Journal of Computer Science and Network Security*, 6(1A):153–162.
- Parzen, E. (1974). Some recent advances in time series modeling. *IEEE Transactions on Automatic Control*, AC-19(6):723–730.
- Puskorius, G. V., Feldkamp, L. A., e Davis, L. I. (1996). Dynamic neural network methods applied to on-vehicle idle speed control. *Proceedings of the IEEE*, 84(1):1407–1420.
- Quinlan, A., Barbot, J.-P., Larzabal, P., e Haardt, M. (2007). Model order selection for short data: An exponential fitting test (eft). *EURASIP Journal on Advances in Signal Processing*, 2007:11.
- Reimbold, M. M. P., Leandro, G. V., Chong, W., e Ribas, R. P. (2007). Identificação do coeficiente de elasticidade de dispositivos mems utilizando modelo arx e estimador recursivo de mínimos quadrados. *Congresso Nacional de Matemática Aplicada e Computacional*, pp. 1–6.
- Rissanen, J. (1978). Modelling by shortest data description. *Automatica*, 14:465–471.

- Rovithakis, G., Chalkiadakis, I., e Zervakis, M. (2004). High-order neural network structure selection for function approximation applications using genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics Society*, 34:150–158.
- Santos, J. E. (2003). *Controle preditivo não-linear para sistemas de Hammerstein*. Tese de Doutorado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina.
- Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.
- Shi, P. e Tsai, C.-L. (2001). Regression model selection - a residual likelihood approach. *Journal of the Royal Statistical Society: Series B*, pp. 237–252.
- Shi, P. e Tsai, C.-L. (2004). A joint regression variable and autoregressive order selection criterion. *Journal of Time Series Analysis*, 25(6):923–941.
- Sorjamaa, A., Reyhani, N., e Lendasse, A. (2005). Input and structure selection for k-nn approximator. *Computational Intelligence and Bioinspired Systems*, 3512:985–992.
- Stoica, P., Selén, Y., e Li, J. (2004). On information criteria and the generalized likelihood ratio test of model order selection. *IEEE Signal Processing Letters*, 11(10):794–797.
- Thanagasundram, S. e Schlindwein, F. S. (2006). Autoregressive order selection for rotating machinery. *International Journal of Acoustics and Vibration*, 11(3):144–154.
- Xiao, X., Li, Y., e Mukkamala, R. (2005). A model order selection criterion with applications to cardio-respiratory-renal systems. *IEEE Transactions on Biomedical Engineering*, 52(3):445–453.