

Cristiano Ribeiro Ferreira Jácome

**Uso de Conhecimento Prévio na
Identificação de Modelos Polinomiais
NARMAX**

Dissertação submetida à banca examinadora designada pelo
Colegiado do Programa de Pós-Graduação em Engenharia
Elétrica da Universidade Federal de Minas Gerais, como parte dos
requisitos necessários à obtenção do grau de Mestre em
Engenharia Elétrica.

Programa de Pós-Graduação em Engenharia Elétrica
Universidade Federal de Minas Gerais

Índice Analítico

Resumo.....	i
Abstract.....	iii
Agradecimentos.....	v
Nomenclatura.....	ix
Capítulo 1	
1. Introdução.....	1
1.1 Introdução à Modelagem Matemática de Sistemas Dinâmicos.....	1
1.2 Definições.....	3
1.3 Modelagem Matemática de Sistemas Dinâmicos.....	5
1.4 Aspectos de Identificação de Sistemas.....	7
1.5 Apresentação do Trabalho.....	9
Capítulo 2	
2. Identificação de Sistemas Não-lineares.....	11
2.1 Introdução.....	11
2.2 Identificação de Sistemas.....	13
2.2.1 Obtenção de dados à Partir da Experimentação do Sistema	14
2.2.2 Detecção de Não-linearidades.....	16
2.2.3 Escolha e Representação de Estruturas.....	17
2.2.3.1 Representação de Estruturas.....	17
2.2.3.2 Detecção de Estrutura.....	20
2.2.4 Estimação de Parâmetros.....	23
2.2.5 Implementações Ortogonais do Algoritmo de Mínimos	
Quadrados.....	27
2.2.6 Detecção de Estrutura Utilizando o ERR.....	30
2.2.7 Validação de Modelos Identificados.....	32
2.2.7.1 Validação Estatística.....	33
2.2.7.2 Validação Dinâmica.....	34
2.3 Comentários Finais.....	35
Capítulo 3	
3. Seleção de Estrutura de Modelos Não-lineares: Agrupamentos de Termos e	
Coeficientes de Agrupamentos.....	37
3.1 Introdução.....	37
3.2 Principais Problemas com Modelos Muito Complexos.....	38
3.3 Agrupamentos de Termos e Coeficientes de Agrupamentos.....	39
3.4 Verificação da Importância de um Agrupamento.....	43
3.5 Comentários Finais.....	47

Capítulo 4

4. Análise de Modelos Discretizados de Sistemas com Constante de Tempo e Ganho Variáveis.....	50
4.1 Introdução.....	50
4.2 Sistemas Lineares com Parâmetros Variáveis.....	51
4.3 Ganho e Constante de Tempo em Modelos Discretizados.....	54
4.4 Representação de Sistemas com Parâmetros Variáveis.....	59
4.5 Modelos Discretizados com Parâmetros Variáveis.....	61
4.5.1 Modelo discretizado de Primeira Ordem.....	62
4.6 Generalização de Modelos Discretizados na Forma de Agrupamentos de Termos e Coeficientes de Agrupamentos.....	65
4.7 Comentários Finais.....	71

Capítulo 5

5. Análise de Modelos Polinomiais NARMAX com Constante de Tempo e Ganho Variáveis.....	72
5.1 Introdução.....	72
5.2 Constante de Tempo e Ganho Variáveis em Modelos Polinomiais NARX.....	72
5.2.1 Obtenção do Ganho e Constante de Tempo Variáveis para um Sistema de Primeira Ordem.....	73
5.2.1.1 Obtenção da Constante de Tempo Variável.....	74
5.2.1.2 Obtenção do Ganho Variável.....	76
5.3 Relações entre Ordem e Grau de Não-linearidade do Modelo Polinomial NARX.....	80
5.3.1 Ordem.....	80
5.3.2 Grau de Não-linearidade.....	80
5.4 Comentários Finais.....	82

Capítulo 6

6. Uso de Técnicas de Identificação na Modelagem de Sistemas com Características Estáticas Não-lineares.....	86
6.1 Introdução.....	86
6.2 Característica Estática.....	87
6.3 Modelo de Hammerstein.....	88
6.3.1 Simulação do Modelo e Obtenção dos Dados.....	88
6.3.2 Análise de Agrupamentos de Termos e Coeficientes de Agrupamentos.....	91
6.4 Modelo de Wiener.....	101
6.5 Comentários Finais.....	112

Capítulo 7

7. Identificação de uma Válvula de Controle.....	114
7.1 Introdução.....	114

7.2 Análise Preliminar do Processo.....	115
7.3 Identificação da Válvula de Controle.....	117
7.3.1 Informações <i>a priori</i>	117
7.3.2 Identificação Baseada em Conhecimento <i>a priori</i> da Válvula.....	120
7.3.3 Identificação do tipo "Caixa-preta".....	129
7.4 Comentários Finais.....	139
 Conclusão.....	 141
 Apêndice A	
A.1 Modelo Discretizado de Segunda Ordem.....	144
A.2 Modelo Discretizado de Terceira Ordem.....	145
 Apêndice B	
B.1 Sistemas de Segunda Ordem.....	146
B.2 Sistemas de Terceira Ordem.....	146
 Apêndice C	
C.1 Generalização para um Sistema de Segunda Ordem.....	147
C.2 Generalização para um Sistema de Terceira Ordem.....	150
 Referências Bibliográficas.....	 156

Resumo

O objetivo deste trabalho é investigar a aplicação de técnicas de identificação de sistemas dinâmicos não-lineares utilizando modelos polinomiais NARMAX, auxiliadas pelo uso de informações *a priori* do sistema. Os modelos polinomiais NARMAX são estruturas paramétricas do tipo entrada-saída capazes de representar uma ampla classe de sistemas não-lineares.

O procedimento de identificação de modelos não-lineares paramétricos pode ser dividido em cinco etapas principais. Estas cinco etapas são apresentadas e analisadas no âmbito de identificação de modelos não-lineares “caixa-preta”.

A seleção de estrutura é a etapa crucial da identificação de modelos não-lineares. Por esse motivo, ela é o foco principal deste trabalho, sendo analisada em maior profundidade. Os conceitos de agrupamentos de termos e coeficientes de agrupamentos são utilizados para derivar um procedimento preliminar, que visa auxiliar na seleção de estruturas de modelos não-lineares polinomiais. Este procedimento, juntamente com conhecimentos *a priori* do sistema, podem ser bastante eficazes em situações reais.

Em seguida, são analisados modelos discretizados de sistemas dinâmicos lineares até terceira ordem, com ganho e constantes de tempo fixos (não variáveis). Posteriormente são consideradas variações de ganho e de uma das constantes de tempo em função de um sinal do processo. Tais resultados são comparados com certos modelos polinomiais NARMAX, investigando-se assim, o acréscimo de novos termos no modelo. Estes procedimentos auxiliam na constituição de informações *a priori*, e são posteriormente aplicados em dois sistemas simulados: um modelo de Wiener e um modelo de Hammerstein, respectivamente.

Finalmente, o procedimento desenvolvido é aplicado a uma válvula de controle real. Partindo-se do modelo dinâmico obtido recupera-se a característica estática da válvula com boa exatidão.

Através desta aplicação, percebe-se que a obtenção de modelos polinomiais NARMAX concisos e dinamicamente válidos é possível mediante a conciliação de certas técnicas de identificação e de conhecimento *a priori* do sistema. É mostrado como tal conhecimento pode ser utilizado para melhor adequar a estrutura do modelo em situações onde as técnicas “caixa-preta” não mais podem ajudar.

Abstract

The main objective of this work is to investigate the use of techniques to identify nonlinear dynamical systems with prior knowledge about the system, using NARMAX polynomial models. Such models are parametric input-output structures able to represent dynamic behavior of a wide class of nonlinear systems.

The identification of nonlinear parametric models can be divided in five stages. These five stages are introduced and analysed in the scope of nonlinear black-box model identification.

Model structure selection is the crucial stage of nonlinear model identification and is the main focus of this work. Thus this issue has been analysed more deeply. The concepts of term clusters and clusters coefficients are used to derive a preliminar procedure to help in the structure selection in nonlinear polynomial models. This procedure associated with prior knowledge about the system being modelled can work efficiently in situations where other techniques fail.

In this way, discretized models of linear dynamic systems up to third order with constant gain and invariant time constants are analysed. After that, variations with process signals are included in the gain and one time constant. These results are compared with certain NARMAX polynomial models. This procedure is applied to two simulated systems: a Wiener model and a Hammerstein model.

Finally, the developed procedure is applied in the identification of a real control valve. The valve's static nonlinearity is recovered from the identified dynamic model.

Besides this practical application, it will be seen that is possible to get reduced dynamically valid NARMAX models, under certain assumptions and prior knowledge about the system. The importance the new results become evident in situation when usual identification techniques can not improve the model structure any further.

Agradecimentos

Eu gostaria de agradecer aos meus pais Ricardo e Marta, aos meus irmãos Thiago e Jayme, à minha irmã Débora e à minha esposa Ivandira pela paciência e pelo apoio irrestrito durante a realização deste trabalho.

Eu gostaria de agradecer ao professor Luis Antonio Aguirre pela orientação acadêmica atuante, pela extrema disposição e pela amizade durante a convivência diária.

Um agradecimento especial é dedicado a Herlon Ayres Camargo pelo inestimável auxílio com simulações e aquisição de dados, bem como pela amizade desenvolvida.

Os amigos do Laboratório de Controle de Processos Industriais (*LCPI*) e Centro de Pesquisa e Desenvolvimento em Engenharia Elétrica (*CPDEE*) garantiram a finalização desse trabalho.

O apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e da Pró-reitoria de Pesquisa dessa Universidade foram de vital importância.

Nomenclatura

Conjuntos Numéricos

Z	Conjunto dos números inteiros
Z⁺	Conjunto dos números inteiros positivos
N	Conjunto dos números naturais
R	Conjunto dos números reais

Variáveis

A	Matriz auxiliar, no procedimento de Gram-Schmidt, para estimação ortogonal de parâmetros.
α	Número de constantes de tempo variáveis em um sistema.
c_{ij}	Coefficientes de um termo $x(t)$ de um modelo polinomial NARMAX
d	Tempo de retardo de um sistema ou tempo morto.
d_e	Dimensão de imersão (embedding) de um sistema dinâmico.
$e(t)$	Ruído ou perturbações em um sistema dinâmico.
$F^l(\cdot)$	Função não-linear qualquer com grau de não-linearidade l .
$f(\cdot)$	Função escalar qualquer.
g	Parâmetro do modelo ortogonalizado, no procedimento de Gram-Schmidt.
$\hat{g}$	Estimativa do parâmetro g .
h	Tempo de atraso na definição das coordenadas de atraso.
$J_N(\cdot)$	Índice de desempenho ou função de custo.
K	ganho.
K_{re}	relação estática saída/entrada.
k	Intervalo de predição.
l	Grau de não-linearidade de um mapeamento não-linear.
m	Ordem dos termos $x(t)$ de um modelo polinomial NARMAX.
N	Comprimento do registro de dados disponíveis.
n	Ordem de um sistema dinâmico.
n_e	Atraso máximo nos tremos em $e(t)$, em modelos discretos.
n_p	Número de termos de processo em um modelo NARMAX polinomial.

n_y	Atraso máximo nos termos em $y(t)$, em modelos discretos.
n_ξ	Número de termos de ruído em um modelo NARMAX polinomial.
n_θ	Número total de termos em um modelo NARMAX polinomial.
Ψ^T	Matriz de regressores de um modelo NARMAX polinomial.
$E\{[\Psi\Psi^T]^{-1}\}$	Matriz de covariância dos parâmetros para $\text{cov}\{e\} = I$.
$\Psi(t)$	Regressor de um modelo NARMAX polinomial representado na configuração preditor de “um-passo-a-frente”.
Q	Matriz ortogonal (decomposição QR).
$\overline{Q}$	Matriz ortogonal aumentada (decomposição QR).
R	Matriz triangular superior.
R_τ	Maior expoente da variável $x_\tau(t)$.
R_k	Maior expoente da variável $x_k(t)$.
T_s	Período de amostragem.
t	Tempo ou número de amostras no caso discreto.
$u(t)$	Sinal medido de entrada de um sistema.
$v(t)$	Variável instrumental.
W	Matriz de regressores ortogonais.
$w(t)$	Regressor ortogonal.
$x(t)$	Termo de um modelo NARMAX polinomial.
$x_k(t)$	Variável ou função de variáveis das quais o ganho K depende.
$x_\tau(t)$	Variável ou função de variáveis das quais a constante de tempo τ depende.
Y	Vetor de saídas.
$y(t)$	Sinal medido de saída de um sistema.
$\hat{y}(t)$	Sinal do preditor de “um-passo-a-frente”.
$\mathbf{y}(t)$	Vetor de coordenadas de atraso.
y_1, y_2	Vetores auxiliares na estimação ortogonal de parâmetros pelo método da transformação de Householder.
Δt	Passo de integração numérica.
$\delta(\cdot)$	Função delta de Dirac.
Θ	Vetor de parâmetros de um modelo polinomial NARMAX.
θ	Parâmetro nominal de um sistema dinâmico representado por um modelo polinomial NARMAX.
$\hat{\theta}$	Parâmetro de um modelo NARMAX, estimado pelo método dos mínimos quadrados.
Ξ	Vetor de resíduos de modelagem em um modelo NARMAX polinomial.
$\xi(t)$	Resíduo de modelagem no instante t .

Σ_i	Coeficiente do agrupamento de termos Ω_i em um modelo NARMAX polinomial.
τ	Constante de tempo
τ_m	Tempo mínimo de amostragem estabelecido pelos critérios de correlação.
$\Phi_{y_1 y_1}(\cdot)$	Função auto-correlação do sinal $y_1(t)$.
$\Phi_{y_1 y_2}(\cdot)$	Função correlação cruzada dos sinais $y_1(t)$ e $y_2(t)$.
Ω	Agrupamento de termos em um modelo polinomial NARMAX.

Operadores

$E\{\cdot\}$	Esperança matemática.
$Var\{y(t)\}$	Variância de $y(t)$.
$\ \cdot\ $	Norma genérica.
$\ \cdot\ _2$	Norma Euclidiana.
$ \cdot $	Valor absoluto de um número.
$[\cdot]^T$	Transposição de vetor ou matriz.

Siglas

ARMAX	Modelo (linear) auto-regressivo, de média móvel e entrada externa (“ <u>A</u> uto- <u>r</u> egressive, <u>m</u> oving- <u>a</u> verage with <u>e</u> xogenous input model”).
ERR	Taxa de redução do erro (“ <u>E</u> rror <u>r</u> eduction <u>r</u> atio”).
FPE	Erro de predição Final (“ <u>F</u> inal prediction error”).
NAR	Modelo não-linear auto-regressivo (“ <u>N</u> on-linear <u>a</u> uto- <u>r</u> egressive model”).
NARMA	Modelo não-linear auto-regressivo, com média móvel (“ <u>N</u> on-linear <u>a</u> uto- <u>r</u> egressive, <u>m</u> oving- <u>a</u> verage model”).
NARMAX	Modelo não-linear auto-regressivo, com média móvel e entrada externa (“ <u>N</u> on-linear <u>a</u> uto- <u>r</u> egressive, <u>m</u> oving- <u>a</u> verage with <u>e</u> xogenous input model”).
NARX	Modelo não-linear auto-regressivo, com entrada externa (“ <u>N</u> on-linear <u>a</u> uto- <u>r</u> egressive with <u>e</u> xogenous input model”).
PRBS	Sinal binário pseudo-aleatório (“ <u>P</u> seudo <u>b</u> inary <u>r</u> andom <u>s</u> ignal”).
RBF	Função de base radial (“ <u>R</u> adial <u>b</u> asis <u>f</u> unction”).
SNR	Relação sinal/ruído (“ <u>S</u> ignal/ <u>n</u> oise <u>r</u> atio”).
TM	Marca registrada (“ <u>T</u> rade <u>m</u> ark”).

1 Introdução

1.1 Introdução à Modelagem Matemática de Sistemas Dinâmicos

No âmbito da engenharia e das ciências naturais existem noções gerais de *observação* e *medição*. Baseado na observação, o cientista constrói uma visão física do problema sob estudo, e daquela observação formula-se uma teoria por tentativa e erro. Esta teoria é um conceito proposto do comportamento natural do sistema observado, e guiado por este conceito, projetam-se novos experimentos. Os resultados destes experimentos podem confirmar a teoria em parte ou descartá-la totalmente (Eykhoff, 1974). Deste ponto de vista, e do ponto de vista da engenharia, a teoria e os experimentos são fundamentais. A formulação de uma teoria, que é um aspecto proposto do sistema sob estudo, pode ser chamada de “construção de modelos”. Por teoria entende-se um modelo matemático ou verbal da realidade (Eykhoff, 1974).

O conjunto de regras matemáticas que auxiliam as pessoas na organização de suas idéias é um dos temas mais fascinantes da cultura humana (Ljung, 1994). A modelagem matemática de sistemas é uma tarefa de vital importância para o desenvolvimento da ciência e da tecnologia. A evolução do conhecimento científico ao longo dos séculos baseou-se na construção e na análise de modelos para os sistemas em estudo. Tais modelos procuravam produzir padrões de comportamentos observados na natureza. A modelagem permitiu estudar sistemas minúsculos, como o átomo, e sistemas imensos, como o nosso sistema solar. Na engenharia, os modelos matemáticos permitem analisar e prever o comportamento de um sistema, sob diversas condições de operação, e ajustar o desempenho do mesmo caso não se mostre satisfatório (Rodrigues, 1996).

Existem tipos distintos de modelos para representar o comportamento de sistemas. Existem modelos do cotidiano que são naturais na realização de tarefas, indicando a sequência adequada de ações para atingir um dado objetivo, estes modelos são

chamados de *modelos mentais*. Os modelos mentais podem ser extremamente complexos, e são construídos a partir da experiência diária das pessoas. A habilidade de dirigir um veículo, por exemplo, é gerada pelo processamento de um modelo mental. Um outro tipo de representação de sistemas é o *modelo gráfico*. Nesse caso, o comportamento do sistema é descrito por tabelas numéricas ou curvas de desempenho. Um exemplo de modelo gráfico é a curva característica tensão-corrente de um dispositivo eletrônico. Por fim, uma das estruturas mais utilizadas para representar o comportamento de sistemas é o *modelo matemático* (Rodrigues, 1996). Um modelo matemático é constituído por um conjunto de equações diferenciais (tempo contínuo) ou equações de diferenças (tempo discreto) que descrevem a variação temporal e/ou espacial das variáveis de interesse no sistema. A complexidade de um modelo matemático dependerá da aplicação visada, seja na pesquisa, no projeto, ou no controle (Eykhoff, 1974).

O levantamento e a formulação matemática de todos os fenômenos que afetam o comportamento de um dado sistema é uma tarefa extremamente complexa. Com isso, pode ser difícil para o modelo reproduzir exatamente o comportamento do sistema original. Então, o modelo deve ser capaz de reproduzir o comportamento original da melhor maneira possível, pois em caso contrário todos os esforços posteriores para análise do modelo e controle do sistema serão pouco eficientes.

Este capítulo visa fornecer uma visão geral de modelagem matemática de sistemas dinâmicos, descrevendo os principais métodos utilizados para este fim, e busca inserir neste contexto o processo de identificação de sistemas e sua importância, que é a linha principal deste trabalho. O capítulo está organizado como se segue, na seção 1.2 são definidos alguns termos que serão utilizados ao longo do trabalho, na seção 1.3 mostra um breve histórico da modelagem matemática de sistemas dinâmicos, inserindo a identificação de sistemas neste processo. Na seção 1.4 são citados alguns aspectos básicos de identificação, e na seção 1.5 é feita uma breve apresentação do trabalho.

1.2 Definições

Esta seção apresenta a definição de alguns conceitos importantes em modelagem matemática de sistemas, e que são utilizados ao longo do capítulo e deste trabalho.

Os sinais de entrada e saída de um sistema monovariável genérico podem ser representados por $u(t)$ e $y(t)$, respectivamente. As definições a seguir são válidas também para sistemas multivariáveis, fazendo-se as devidas adaptações.

Definição 1.1

Uma *planta* é uma parte de equipamento, eventualmente um conjunto de itens de uma máquina que funcionam juntos, cuja finalidade é desempenhar uma dada operação (Ogata, 1993).

Definição 1.2

Um *processo* é uma operação ou um desenvolvimento natural, que evolui progressivamente, caracterizado por uma série de mudanças graduais que se sucedem, umas em relação às outras, de um modelo relativamente fixo conduzindo a um resultado particular. Neste trabalho, qualquer operação a ser controlada será chamada de processo.

Definição 1.3

Um *sistema* é uma combinação de componentes que atuam conjuntamente e realizam um objetivo (Ogata, 1993). Um sistema não está limitado a algo físico. O conceito de sistema pode ser aplicado a fenômenos abstratos, dinâmicos, tais como os encontrados em economia. A palavra sistema deve, portanto, ser interpretada para designar sistemas físicos, biológicos, mecânicos, econômicos, etc.

Definição 1.4

Um sistema é dito *não-autônomo* quando as equações diferenciais (ou equações de diferenças) que governam o seu comportamento dependem explicitamente do tempo t . Logo, as equações diferenciais que regem o comportamento de um sistema *autônomo* não dependem explicitamente do tempo t .

Definição 1.5

Um sistema é *invariante no tempo* quando seus parâmetros não variam com o tempo. Caso contrário, o sistema é dito *variante no tempo*.

Definição 1.6

Um sistema é dito *dinâmico* quando o valor atual de sua saída $y(t)$ depende do valor atual da entrada $u(t)$ aplicada, e também da evolução temporal da entrada e da saída. Um sistema é dito *não-dinâmico* quando o valor atual de sua saída $y(t)$ depende apenas do valor atual da entrada $u(t)$ aplicada.

Definição 1.7

Um sistema encontra-se em *regime estático* quando a sua saída atual, $y(t)$, é igual à saída anterior $y(t-1)$, e assim por diante, podendo ser escrita na forma estática Y , ou seja, $y(t) = y(t-1) = \dots y(t-n) = Y$. Da mesma maneira para a entrada $u(t) = u(t-1) = \dots u(t-n) = U$.

Definição 1.8

Um sistema é *determinístico* quando a resposta é prognosticável e repetível, caso contrário é estocástico (Ogata, 1993). Isto quer dizer que, em um sistema determinístico, o conhecimento da entrada aplicada $u(t)$ e das condições iniciais adequadas determina

unicamente a saída. Já um sistema estocástico, o conhecimento destes fatores não é suficiente para a determinação da saída.

Os sistemas considerados neste trabalho são determinísticos, apresentando comportamento dinâmico invariante no tempo.

1.3 Modelagem Matemática de Sistemas Dinâmicos

Existem duas vertentes básicas do desenvolvimento de modelos matemáticos para sistemas dinâmicos:

- modelagem pela física do processo (Doebelin, 1980; Pena, 1983);
- identificação de sistemas (Aström and Eykhoff, 1971; Eykhoff, 1974; Box and Jenkins, 1976; Norton, 1986; Ljung, 1987; Söderström and Stoica, 1989).

A modelagem pela física do processo é feita mediante a obtenção das equações baseadas na física do processo, que exige um grande conhecimento das leis empíricas e teóricas que regem o comportamento dinâmico deste. Quando o sistema a ser modelado é grande e complexo, e existe a necessidade de atuação e controle em tempo real, torna-se inviável a utilização de modelos derivados pela física do processo. Uma alternativa é o emprego de técnicas de identificação de sistemas, onde um modelo de entrada-saída do sistema é obtido a partir de dados (Ljung, 1987; Rodrigues, 1996). Os modelos gerados pela identificação de sistemas podem ser utilizados para inferir propriedades dinâmicas e estáticas do sistema original.

Um sistema dinâmico pode ser analisado no domínio do tempo e/ou domínio da frequência. Por esse motivo, a identificação deve ser capaz de derivar modelos (lineares ou não-lineares) que descrevam o comportamento do sistema original no domínio do tempo (equações diferenciais) ou no domínio da frequência (resposta em frequência, conforme o enfoque desejado).

A identificação de sistemas existe desde a primeira metade deste século, onde até meados da década de 70, destacavam-se pesquisadores como Kalman (Kalman, 1960), Aström (Aström and Eykhoff, 1971) e Eykhoff (Eykhoff, 1974). Naquele tempo, os sistemas eram tratados com um enfoque linear (Doebelin, 1980). A partir da segunda metade da década de 70, a comunidade de Controle teve um interesse ativo no desenvolvimento de técnicas de estimação de modelos de sistemas dinâmicos à partir dos dados de entrada e saída, que será visto no capítulo 2, e a esta tarefa deu-se o nome de “Identificação de Sistemas”. Os pesquisadores procuraram então, adaptá-la para a utilização na análise, predição e controle de sistemas reais (Norton, 1986; Söderström, 1989; Ljung and Söderström, 1983, Söderström et al., 1978).

Historicamente, os algoritmos de identificação e controle foram desenvolvidos assumindo-se linearidade dos sistemas dinâmicos (Box and Jenkins, 1976). Isso ocorreu devido a restrições teóricas e computacionais. Porém, a maioria dos sistemas físicos reais são significativamente não-lineares, e a utilização de modelos lineares para descrevê-los é limitada, pois estes modelos são incapazes de reproduzir uma gama de comportamentos dinâmicos apresentados pelos sistemas reais, tais como ciclos limites, saturações e bifurcações. Já que muitos dos sistemas em estudo atualmente apresentam tais fenômenos, o interesse pela identificação de modelos não-lineares cresceu, principalmente desde o início da década de 80. Atualmente, o crescente interesse pelo assunto é refletido pelo elevado número de trabalhos publicados (Diaz and Desrochers, 1988; Rodrigues et al., 1996; Aguirre and Billings, 1995a; Jang and Kim, 1994).

Os métodos desenvolvidos para a identificação de sistemas podem ser divididos em três grupos (Ljung, 1987):

- métodos paramétricos;
- métodos não-paramétricos;
- métodos no domínio da frequência;

Os métodos ditos *paramétricos* utilizam estruturas matemáticas parametrizadas para descrever o comportamento dinâmico original no domínio do tempo. Os parâmetros destas estruturas matemáticas são ajustados por algoritmos de estimação a partir dos dados medidos (Ljung, 1987). Os métodos ditos *não-paramétricos* também geram modelos no domínio do tempo. Nesse caso, o comportamento dinâmico do sistema é determinado através de funções de correlação calculadas sobre os dados disponíveis. Finalmente, os métodos no *domínio da frequência* geram modelos que reproduzem a resposta em frequência do sistema em estudo. Estes métodos utilizam a transformada de Fourier para calcular as respostas em frequência (Hougen, 1972).

1.4 Aspectos de Identificação de sistemas

Os dados utilizados na identificação de modelos dinâmicos são denominados *dados de identificação*. Os dados utilizados na validação dos modelos, que devem ser diferentes dos dados de identificação, são denominados *dados de validação*. Os dados de identificação devem ser gerados preferencialmente com o sistema bem excitado com relação aos espectros de amplitude e frequência. Já os dados de validação podem ser coletados durante a operação normal do sistema.

Os modelos matemáticos gerados pela identificação de sistemas podem ser divididos em duas classes:

- modelos entrada-saída;
- modelos em espaços de estados.

Os modelos de entrada-saída descrevem a resposta $y(t)$ do sistema em função da entrada $u(t)$ aplicada. Porém, estes modelos não permitem analisar o comportamento de variáveis internas do sistemas para uma dada operação, ao contrário dos modelos em espaço de estados (Rodrigues, 1996).

A escolha da estrutura a ser utilizada para representar os regimes dinâmicos descritos pelos dados é um problema crucial na identificação de sistemas. Este problema pode ser suavizado pela utilização de propriedades do sistema conhecidas *a priori* na seleção da estrutura do modelo que deverá representá-lo, conforme será visto ao longo do trabalho. Os modelos matemáticos gerados pela identificação de sistemas podem ser classificados em três grupos de acordo com o nível de conhecimento *a priori* utilizado na seleção de sua estrutura (Ljung, 1987):

- modelos caixa-branca (“white-box”);
- modelos caixa-cinza (“grey-box”);
- modelos caixa-preta (“black-box”).

Os modelos “caixa-branca” têm suas estruturas completamente ajustadas a partir de informações conhecidas *a priori*. Nesse caso, a forma da função matemática que descreve o comportamento dinâmico do sistema original é previamente conhecida. Os modelos “caixa-cinza” são identificados utilizando algum conhecimento *a priori* para simplificar os algoritmos de seleção de estrutura. Por fim, a identificação de modelos “caixa-preta” não utiliza nenhuma informação conhecida *a priori*. A estrutura destes modelos pode ser ajustada dentro de famílias conhecidas por apresentar boa flexibilidade na modelagem de sistemas.

Com o conhecimento da estrutura de um processo e um entendimento pelo menos parcial deste, é possível ter-se uma quantidade de informação *a priori* do mesmo (Eykhoff, 1974). Este conhecimento certamente auxiliará na seleção da estrutura ou na escolha dos parâmetros do modelo, seja físico ou matemático. A disponibilidade de tal conhecimento pode gerar uma diferença de complexidade estrutural e ordem de magnitude entre modelos. Este fato normalmente também é verdade para a distinção de aspectos lineares e não-lineares de processos reais, que conseqüentemente determinará o tipo de tratamento dado ao problema de identificação. Apesar da aproximação de modelos “caixa-preta” ser atrativo do ponto de vista teórico, este tipo de modelo em

alguns casos leva a soluções pobres, devido à não utilização de informações que podem estar disponíveis (Eykhoff, 1974).

Neste trabalho serão discutidas formas de incorporar alguma informação *a priori* do sistema no procedimento de identificação, auxiliando a seleção da estrutura do modelo.

1.5 Apresentação do Trabalho

O objetivo principal desse trabalho é desenvolver um procedimento que auxilie na detecção de estruturas de modelos polinomiais NARMAX (non-linear auto-regressive moving average with exogenous inputs), gerando estruturas dinamicamente válidas, formadas somente a partir de termos efetivos, possibilitando recuperar a partir do modelo informações estáticas, tais como saturações, etc.

Os modelos polinomiais NARMAX são estruturas paramétricas do tipo entrada-saída capazes de representar o comportamento dinâmico de uma ampla classe de sistemas dinâmicos não-lineares reais (Leontaritis and Billings, 1985a, 1985b).

O procedimento de identificação de modelos paramétricos não-lineares pode ser dividido em cinco etapas: (i) aquisição de dados do sistema que se deseja modelar; (ii) aplicação de testes para a detecção de não-linearidades aos dados medidos; (iii) seleção da estrutura do modelo que deverá representar a dinâmica original, que é o enfoque principal deste trabalho; (iv) estimação dos parâmetros da estrutura selecionada; (v) validação do modelo identificado. O capítulo 2 descreve brevemente as cinco etapas do procedimento de identificação de modelos NARMAX polinomiais, com base na literatura correlata.

O capítulo 3 descreve uma ferramenta importante no auxílio à seleção da estrutura dos modelos. Modelos não-lineares sobreparametrizados tendem a apresentar regimes dinâmicos espúrios; isto é, regimes dinâmicos não apresentados pelo sistema original. O

capítulo revisa a utilização do conceito de agrupamentos de termos e coeficientes de agrupamentos no auxílio à seleção do conjunto de termos importantes para reproduzir o comportamento do sistema. Em seguida, os termos de modelos NARMAX polinomiais podem ser escolhidos dentro desse conjunto utilizando técnicas usuais de seleção de estruturas, ou até mesmo um conhecimento *a priori* do sistema.

No capítulo 4, através da manipulação e análise de modelos discretizados, pela semelhança apresentada com modelos polinomiais NARMAX, tentar-se-á determinar a estrutura de um modelo que apresente o ganho e/ou uma constante de tempo variáveis. Através de certos procedimentos, isolar-se-ão as informações relativas a variação destes parâmetros. No capítulo 5, este resultado é comparado com estruturas de modelos polinomiais NARMAX, atentando-se para os agrupamentos presentes nos modelos. A partir destes modelos polinomiais NARMAX tentar-se-á recuperar informações como relação estática saída/entrada e constante de tempo variáveis em função de outros sinais. O resultado final é generalizado na forma de agrupamentos de termo e coeficientes de agrupamento, fornecendo uma visão mais ampla da estrutura dos modelos.

O capítulo 6, utiliza as técnicas de agrupamentos de termos e coeficientes de agrupamentos na definição da estrutura de sistemas constituídos de uma característica estática em série com um modelo linear dinâmico de primeira ordem.

Finalmente os resultados dos capítulos 4, 5 e 6 são utilizados no auxílio à seleção de estrutura do modelo de uma válvula de controle de igual porcentagem, presente em uma planta de tanques interativos, no Laboratório de Controle de Processos Industriais (LCPI) da Universidade Federal de Minas Gerais. Através do modelo dinâmico identificado tentar-se-á recuperar a característica estática da válvula.

Após os sete capítulos descritos, são apresentadas as conclusões desse trabalho e algumas propostas de continuação. Por fim, os apêndices A, B e C apresentam resultados extendidos de sistemas de primeira ordem, para sistemas de segunda e terceira ordem.

2 Identificação de Sistemas Não-Lineares

2.1 Introdução

Modelos matemáticos de sistemas dinâmicos (sistemas que apresentam uma relação entre a sua história passada e seu comportamento atual) são de vital importância em engenharia, física, medicina, economia, biologia e na maioria das áreas do conhecimento humano. Estes modelos ajudam as pessoas a organizarem seus pensamentos, e, principalmente em engenharia, a necessidade de desenvolvimento de modelos matemáticos de sistemas tem se tornado cada vez mais evidente (Ljung, 1987; Ljung 1994). Através de modelos matemáticos pode-se prever o comportamento de um sistema quando excitado por entradas diferentes, atentando-se para comportamentos não desejáveis e permitindo um ajuste no mesmo. Portanto, o modelo deve representar da melhor maneira possível o comportamento ou aspectos do sistema que representa. A partir de modelos pode-se, por exemplo projetar-se um controle para o mesmo (Pröll and Karim, 1994). Existem dois grandes grupos de métodos de modelagem, a modelagem pelas equações físicas que regem o processo, e a identificação de sistemas à partir de dados de entrada e saída.

A modelagem física necessita de um conhecimento detalhado das leis físicas que regem o processo a ser modelado, estas leis são agrupadas e organizadas de forma a simular e analisar o sistema. Além disto, um sistema grande e complexo, ao ser modelado, pode apresentar uma interação muito grande entre as variáveis envolvidas, tornando-se assim difícil de ser descrito matematicamente. Uma solução para este problema, seria a utilização de modelos identificados a partir de dados de entrada e saída do sistema.

Porém, nada impede que técnicas dos dois grupos sejam utilizadas em conjunto, o que pode elucidar regimes de comportamento que ora são exibidos por um, ora por outro. Portanto, a informação disponível sobre o processo a ser modelado, o tempo útil para se

desenvolver o modelo, as ferramentas disponíveis para o mesmo e o grau de aproximação desejado do modelo (uma vez que ele não reflete o comportamento exato do sistema) determinará qual dos métodos deverá ser utilizado.

Muito embora sistemas reais sejam não-lineares, as facilidades encontradas na análise e no projeto de sistemas lineares (além das restrições computacionais) levaram, no passado, à representação destes sistemas através de modelos lineares. Toda a teoria do “Controle Clássico” e a maior parte do chamado “Controle Moderno”, cujo melhor nome seria “Controle por variáveis de Estado”, assumem que os sistemas são lineares (Kalman, 1960; Åström and Eykhoff, 1971; Eykhoff, 1974). Originalmente, a preocupação primária era desenvolver novas ferramentas matemáticas para se entender os fenômenos não-lineares que não poderiam ser analisados utilizando-se as técnicas já existentes para sistemas lineares, uma vez que a utilização de modelos lineares não consegue reproduzir toda a gama de comportamentos dinâmicos apresentados pelos sistemas reais, tais como ciclos limite, bifurcações, comportamento quase-periódicos e caos (Aguirre and Aguirre, 1995a; Aguirre, 1994a; Genesio and Tesi, 1993).

Muitos processos em estudo atualmente apresentam tais regimes dinâmicos, o interesse pela modelagem não-linear de sistemas tem crescido significativamente (Billings and Leontaritis, 1981; Billings and Voon, 1984; Billings and Fadzil, 1985; Diaz and Desrochers, 1988; Billings and Chen, 1989a; Chen *et al.*, 1989; Zhao and Lu, 1991; Pröll and Karim, 1994; Parameswaran and Raol, 1994; Aguirre and Billings, 1995a; Çinar, A. 1995; Aguirre and Mendes, 1996), bem como a aplicação das técnicas de identificação à processos reais (Kortmann et al., 1988; Valverdú et al., 1992; Jang and Kim, 1994; Rodrigues et al., 1996).

Em problemas de modelagem de sistemas é muito importante que o modelo final seja válido. Tem-se notado que modelos não-lineares normalmente apresentam melhores resultados na predição, quando submetidos a excitações diferentes, que modelos lineares (Rodrigues, 1996).

O objetivo deste capítulo é apresentar uma revisão dos principais tópicos relacionados à identificação de sistemas não lineares, baseando-se para isso em algoritmos descritos na literatura. A capítulo está organizado da seguinte maneira: na seção 2.2 são mostradas as etapas do processo de identificação, seja linear ou não-linear; na seção 2.2.1 comenta-se sobre a obtenção dos dados para a identificação, e sobre o tempo de amostragem dos sistemas. Na seção 2.2.2 é apresentada uma maneira de se discernir a necessidade ou não de um modelo não-linear a partir da massa de dados, utilizando-se testes para detecção de não-linearidades. Na seção 2.2.3 discute-se as formas de representação de modelos e escolha das estruturas dos mesmos. Na seção 2.2.4 apresenta-se um algoritmo de estimação de parâmetros do modelo, neste caso o algoritmo dos mínimos quadrados. Na seção 2.2.5 mostra-se como é feita a implementação ortogonal dos mínimos quadrados, a fim de contornar problemas numéricos. Na seção 2.2.6 apresenta-se uma forma de detecção de estruturas baseada na técnica do ERR, e na seção 2.2.7 comenta-se a respeito da validação de modelos, sendo mostradas algumas técnicas de validação estatística. Finalmente na seção 2.3 é feito um comentário geral sobre o capítulo.

2.2 Identificação de Sistemas

A identificação de sistemas pode ser dividida em cinco partes principais (Godfrey, 1986; Ljung, 1987):

- obtenção dos dados de experimentação do sistema que se deseja modelar;
- aplicação de testes aos dados obtidos para verificação de não-linearidades;
- escolha da melhor estrutura na representação do modelo;
- estimação dos parâmetros do modelo;
- validação do modelo obtido.

O procedimento descrito acima é utilizado na identificação tanto de sistemas lineares quanto de sistemas não-lineares. As principais diferenças se devem à maneira com que cada passo é implementado.

2.2.1 Obtenção dos Dados à partir da Experimentação do Sistema

Nesta etapa são obtidos dados de entrada e saída do sistema, que serão utilizados para a estimação de parâmetros e ajuste do modelo. A obtenção destes dados, é provavelmente o passo mais importante no estudo de identificação de sistemas, porque todos os resultados subsequentes a esta etapa dependem da qualidade e da quantidade de informação contidas nos dados coletados (Billings and Fadzil, 1985). Enquanto dados provenientes da operação normal do sistema podem ser utilizados como uma base para identificação, é preferível, sempre que possível, que sejam injetadas no sistema entradas externas, a fim de melhor excitar a sua dinâmica (Eykhoff, 1974; Billings and Voon, 1983).

Estas entradas são geralmente projetadas baseando-se em experimentações preliminares do processo, para que estas sejam “persistentemente excitantes” (Norton, 1986). Para alguns sistemas não-lineares, a entrada deve ser selecionada de forma à ter um largo espectro de frequência e o perfil de amplitude de interesse no sistema (Aguirre and Billings, 1995b). Entretanto, é muito difícil encontrar uma solução analítica para as condições experimentais, mas utilizando-se resultados da Teoria de Informação pode ser mostrado que (Leontaritis and Billings, 1987):

- i) para uma restrição de amplitude ou potência na entrada, esta deve ser uma sequência independente;
- ii) para uma restrição de potência a entrada deve ser Gaussiana;
- iii) para uma restrição de amplitude, a entrada deve ser uniformemente distribuída;
- vi) para sistemas multivariáveis as entradas devem ser mutuamente independentes

Dentre as entradas utilizadas na identificação de sistema como um todo pode-se citar sinais binários pseudo-aleatórios (PRBS) e ruído ‘branco’, pois apresentam um espectro de frequências relativamente largo.

O uso de dados provenientes da operação normal da planta é tentador quando a entrada parece suficientemente excitante, mas problemas podem aparecer quando existe “realimentação” no sistema, o que não é óbvio pela simples inspeção visual dos dados. Para sistemas com realimentação, é altamente aconselhável o uso de uma perturbação aplicada externamente. Sem este sinal, a dinâmica estimada pode ser pouco parecida com a dinâmica do ramo direto da planta (Söderström and Stoica, 1989, capítulo 10).

Outro fator importante na experimentação do sistema é o tempo de amostragem utilizado. Esta importância é devida ao fato de que tempos de amostragem diferentes proporcionam modelos diferentes. A obtenção de dados superamostrados, onde as sucessivas medições tendem a estar altamente correlacionadas, pode levar a um mal-condicionamento numérico e falta de recursos computacionais na gravação e processamento de tais dados (Billings and Aguirre, 1995f). Por outro lado, a obtenção de dados subamostrados pode levar a uma perda de informação dinâmica entre uma amostra e outra, e a complexidade do modelo estimado pode reduzir-se. A obtenção de modelos a partir de dados amostrados, dependendo do espaçamento das amostras, pode incluir parâmetros estimados com baixa precisão. É aconselhável então, realizar um teste de sensibilidade no modelo, variando cada parâmetro e examinando seu efeito na saída, verificando-se assim a sua importância (Godfrey, 1986).

Uma maneira de se obter o tempo de amostragem do sistema seria através do teorema da amostragem de Shannon (Phillips, 1995). Este teorema indica que o tempo de amostragem adequado para um sistema seria com uma frequência no mínimo duas vezes maior que a maior frequência de interesse. Na prática, usa-se uma frequência de cinco a dez vezes maior que a frequência de interesse.

Uma outra maneira de escolher o tempo de amostragem seria tomar o menor entre os primeiros mínimos (τ_m) das funções de autocorrelação linear, $\Phi_{yy}(h)$, e não-linear, $\Phi_{y^2y^2}(h)$, dos dados de saída e utilizar um tempo de amostragem que esteja entre o seguinte intervalo (Aguirre, 1995):

$$\frac{\tau_m}{20} \leq T_s \leq \frac{\tau_m}{10}$$

onde

$$\begin{aligned}\Phi_{yy}(t) &= E\{[y(k) - \overline{y(k)}][y(k-h) - \overline{y(k)}]\}, & h = 0, 1, \dots \\ \Phi_{y^2y^2}(t) &= E\{[y^2(k) - \overline{y^2(k)}][y^2(k-h) - \overline{y^2(k)}]\}, & h = 0, 1, \dots\end{aligned}$$

A barra indica o valor médio e E a esperança matemática.

2.2.2 Detecção de não-linearidades

Uma tarefa também primária na identificação é decidir através dos dados de entrada e saída, se o processo sob consideração é linear ou não linear. Com o intuito de determinar tal característica do processo, pode-se aplicar testes de detecção de não-linearidade nos dados (Billings and Voon, 1983; Billings and Fadzil, 1985). Estes testes verificam se os dados do sistema satisfazem (dentro de certos limites) certas propriedades peculiares a sistemas lineares. Adota-se um modelo não-linear quando tais características não forem verificadas.

O teste de não-linearidade deve ser não paramétrico, ou seja, deve ser independente da estrutura não-linear e da ordem da dinâmica da parte linear do processo; ele também não deve requerer muito tempo computacional. Portanto, os métodos mais utilizados na detecção de não-linearidades são (Haber, 1985):

- teste no domínio do tempo;
- método da correlação cruzada não linear ;
- método da autocorrelação de ordem elevada.

Existem outros métodos, tais como o método da saída média e o método de regime permanente, porém estes procedimentos só detectam características não-lineares do processo em regime permanente (Haber and Unbehauen, 1990).

De acordo com Haber e Unbehauen (1990), os sistemas que apresentam termos constantes além dos termos lineares e não-lineares, podem encobrir os verdadeiros resultados dos testes de não-linearidade. A fim de se eliminar o efeito do termo constante, os métodos de correlação devem ser realizados com os dados normalizados, e não diretamente com os dados medidos.

2.2.3 Escolha e Representação de Estruturas

2.2.3.1 Representação de Estruturas

Um modelo de entrada/saída é um meio matemático de descrever a relação entre a entrada e a saída de um sistema, relacionando-os de maneira direta e proporcionando uma aproximação adequada a um pequeno custo computacional.

Uma forma de representação de sistemas seria através da linearização de modelos em torno de um ponto de operação, que é chamado de representação local (Billings, 1980). Este método consiste em tomar-se um conjunto de modelos lineares locais simples, representando cada ponto de operação desejável do sistema, e interpolá-los a fim de obter o modelo completo. Mas este método requer que o sistema seja validado em torno de cada um destes pontos de operação, validando-se consequentemente os modelos locais.

Quando da escolha da representação, alguns aspectos devem ser considerados para auxiliarem na seleção de estruturas do modelo, devido a variedade de representações existentes (Çinar, 1995). Por exemplo, algumas representações como “linear por partes” (Billings and Voon, 1987) e funcionais requerem arquivos de dados muito maiores que

outras formas de representações. Outras representações tais como redes neurais (Elsner, 1992) e funções de base radial (Casdagli, 1989), demonstram certas vantagens na manipulação de grandes vetores de dados. Por outro lado, representações analíticas simples como funções racionais e polinomiais (Billings and Chen, 1989a; Haber and Unbenhauen, 1990) são mais fáceis de serem analisadas em relação às outras. Isto não quer dizer, entretanto, que uma determinada representação só tem um bom desempenho em alguns poucos problemas práticos. Na verdade, sabe-se que representações globais através de polinômios não lineares são adequadas para a modelagem e quantificação de vários fenômenos complexos (Aguirre, 1994), desde que a estrutura do modelo seja apropriada.

Os modelos a serem utilizados neste trabalho são os modelos polinomiais do tipo “non-linear auto-regressive moving average with exogenous inputs” (NARMAX) (Leontaritis and Billings, 1985a; Leontaritis and Billings, 1985b), que constituem uma representação natural para uma grande classe de sistemas não lineares (Billings and Chen, 1989). Esta escolha de representação pode ser justificada por um grande número de razões, dentre as quais:

- tal representação é global, e sendo assim mais apropriada para a análise da dinâmica global do sistema;
- a complexidade do modelo é facilmente quantificada pelo grau de não linearidade, número de termos e máximo atraso utilizado ;
- níveis moderados de ruído podem ser manipulados com certa facilidade;
- facilidade relativa na análise;
- sendo linear nos parâmetros, algoritmos simples e robustos podem ser usados na estimação de parâmetros.

Um modelo NARMAX com período de amostragem normalizado, pode ser representado como se segue (Leontaritis and Billings, 1985a; Leontaritis and Billings, 1985b):

$$y(k) = F^l(y(k-1), y(k-2), \dots, y(k-n_y), u(k-d), u(k-d-1), \dots, u(k-d-n_u+1), e(k-1), e(k-2), \dots, e(k-n_e)) + e(k) \quad (2.1)$$

onde n_y , n_u , e n_e são respectivamente os máximos atrasos considerados nos termos de saída, entrada e ruído; e d é o atraso puro de tempo medido em intervalos de amostragem T_s . Mais ainda, $u(k)$ e $y(k)$ são respectivamente as séries temporais de entrada e saída obtidas pela amostragem dos dados contínuos de $u(t)$ e $y(t)$ nos instantes T_s . A função F^l representa uma função não-linear qualquer, com grau de não-linearidade l .

Como dito anteriormente, o escopo deste trabalho são as equações polinomiais de modelos NARMAX. Um modelo polinomial não-linear de grau l pode ser representado da seguinte maneira (Leontaritis and Billings, 1985a; Leontaritis and Billings, 1985b):

$$y(t) = \theta_0 + \sum_{i_1=1}^n \theta_{i_1} x_{i_1} + \sum_{i_1=1}^n \sum_{i_2=1}^n \theta_{i_1 i_2} x_{i_1}(t) x_{i_2}(t) + \sum_{i_1=1}^n \dots \sum_{i_n=i_1-1}^n \theta_{i_1 \dots i_n} x_{i_1}(t) \dots x_{i_n}(t) + e(t) \quad (2.2)$$

onde $x_1(t) = y(t-1)$, $x_2 = y(t-2)$, $\dots$, $x_{n_y+1}(t) = u(t-d)$, $\dots$,
 $x_{n_y+n_u+1}(t) = e(t-1)$, $\dots$ $x_n(t) = e(t-n_e)$, $n = n_u + n_y + n_e$

Os θ 's são os parâmetros a serem estimados de forma que a estrutura escolhida para o modelo se ajuste aos dados de estimação e represente a dinâmica do sistema.

Neste trabalho será utilizada apenas a parte determinística do modelo polinomial NARMAX, isto é, a parte NARX, uma vez a parte da média móvel será omitida. Tal parte tem a função de evitar a polarização dos parâmetros estimados, sendo descartados em qualquer outra utilização do modelo em que sua presença não se faz necessária.

2.2.3.2 Detecção da Estrutura

Uma dificuldade na escolha da estrutura de modelos polinomiais NARMAX se deve à relação entre o número de termos no modelo, com a ordem e o grau de não-linearidade do mesmo. Isto porque o número de termos possíveis em modelos polinomiais cresce rapidamente com o aumento do grau de não-linearidade l e da ordem do modelo, n_y . Na equação (2.2) pode ser visto também que o número de termos depende de n_u e n_e . Logo, o número de termos possíveis em modelos polinomiais pode ser expresso pela seguinte fórmula (Korenberg *et al.*, 1988) :

$$n_\theta = M + 1 \quad \text{onde } n_\theta = \text{número de termos do modelo e,}$$

$$M = \sum_{i=1}^l n_i ,$$

$$n_i = \frac{n_{i-1}(n_y + n_u + n_e + i - 1)}{i}, \quad n_0 = 1$$

O valor de n_θ pode se tornar impraticavelmente grande para modelos polinomiais, embora o problema aqui não seja tão grave quanto no caso das séries de Volterra ou modelos de Wiener, onde o número de termos pode facilmente chegar a 10^{10} para sistemas relativamente simples (Billings, 1980). A união de todos os termos possíveis de serem incluídos em um modelo polinomial é denominado *conjunto de termos candidatos*. O procedimento de se encontrar os termos a serem incluídos no modelo, dentre os termos candidatos, é chamado de *detecção de estrutura*.

A detecção de estrutura é a tarefa mais difícil e importante no processo de identificação. Esta importância encontra-se no fato de que as características dinâmicas e estáticas do sistema estão intimamente ligadas a sua estrutura. Um modelo cuja estrutura é deficiente, certamente não irá reproduzir bem o comportamento do sistema original.

No processo de detecção da estrutura, os conhecimentos *a priori* do sistema podem auxiliar na escolha, como será visto nos próximos capítulos. Neste contexto a seleção da estrutura do modelo não consiste em escolher a estrutura “verdadeira”, pois esta não existe, e sim encontrar uma que melhor descreva aspectos dinâmicos do sistema, e se possível possa incluir também os aspectos estáticos, tais como a saturação.

Na tentativa de ajustar melhor os dados e representar dinamicamente o sistema, alguns modelos podem incluir termos que não são necessários. Estes termos, chamados espúrios, podem incluir características não condizentes com a realidade do sistema. À presença de termos não necessários denomina-se sobreparametrização do modelo (Aguirre and Billings, 1995b). Frequentemente, o uso de modelos não-lineares com maior complexidade e tamanho que o necessário, são mais utilizados, e as principais causas disto são:

- i) dificuldade de escolher uma representação adequada;
- ii) alguns algoritmos de seleção de estruturas gastam muito tempo computacional;
- iii) modelos maiores são mais flexíveis e tendem à acomodar melhor os dados, provavelmente reduzindo a variância dos resíduos;
- iv) disponibilidade de computadores mais rápidos e mais baratos, que facilitou a manipulação de modelos maiores.

Entretanto, modelos flexíveis que ajustam melhor os dados não necessariamente capturam a dinâmica do sistema apropriadamente, logo, o item (iii) deve ser visto com cautela.

Devido ao explosivo crescimento do conjunto de termos candidatos ao modelo com a ordem e retardo máximo de tempo assumidos para este modelo, uma procura exaustiva pela melhor estrutura nem sempre é fácil. Com o intuito de condicionar uma seleção de estruturas adequada, diversos métodos foram desenvolvidos, tais como algoritmos genéticos (Fonseca et al., 1993), onde cada modelo possível foi codificado como um

indivíduo de uma cadeia de genes, e cada gene indica um termo constituinte do modelo. Através do uso de técnicas “evolucionárias” é selecionado o indivíduo mais apto; ou seja, o melhor modelo para o sistema em questão.

Outro método seria a validação cruzada (Sorenson, 1970), que sob determinadas condições irá selecionar uma estrutura ótima com relação ao critério de validação cruzada estabelecido. Na interpretação da validação cruzada dos modelos atribui-se a eles uma aparência intuitiva, e permite assim adaptá-los para aplicações específicas pela escolha apropriada do parâmetro deste critério, sendo uma decisão do usuário.

Outros métodos que poderiam ser utilizados no auxílio a seleção de estruturas seriam o critério de informação de Akaike (Akaike, 1974), Agrupamento de termos (Aguirre, 1994; Aguirre and Billings, 1995a) (será citado no capítulo 3) e Error Reduction Ratio (ERR) (Chen *et al.*, 1989; Korenberg *et al.*, 1988) que se apresenta como uma técnica muito útil e eficaz na detecção de estruturas para modelos NARMAX. O ERR de cada termo é um número que indica a melhoria obtida na explicação da variância dos dados, com a inclusão daquele termo no modelo. O algoritmo de detecção de estrutura baseado no ERR será descrito na seção 2.2.6, após a apresentação do algoritmo de estimação de parâmetros.

Alguns métodos mais recentes, que tentam auxiliar no problema de seleção de estruturas são estudados em Thouverez and Jezequel (1996), e Wang and Cluett (1996).

Um objetivo deste trabalho é desenvolver técnicas e procedimentos, que aliadas à um conhecimento *a priori* do sistema, auxilie na escolha de uma estrutura concisa e dinamicamente válida para o modelo. Verificar-se-á quais tipos de informações estão incluídas e podem ser eventualmente representadas a partir dos modelos polinomiais NARMAX.

2.2.4 Estimação de Parâmetros

Determinada a representação e a estrutura do modelo, deve-se estimar seus parâmetros, a partir de dados obtidos em experimentação, buscando aproximar o comportamento dinâmico exibido pelo sistema. Normalmente, em modelos polinomiais NARMAX, que são lineares nos parâmetros, o método de estimação de parâmetros utilizado é o dos mínimos quadrados, que pode ser solucionado por várias técnicas numéricas bem desenvolvidas como a transformação de Householder e a decomposição de Choleski (uma revisão destas técnicas pode ser vista em Ljung (1987) e Chen et al. (1989)).

A estrutura apresentada na equação (2.2) pode ser representada como (Chen *et al.*, 1989):

$$y(t) = \sum_{i=1}^{n_\theta} \varphi_i(t) \theta_i + e(t) \quad (2.3)$$

onde n_θ é o número de termos no modelo, $\varphi_i(t)$ corresponde aos regressores, que são os diferentes termos do polinômio representando os monômios de $x_1(t)$ a $x_n(t)$ até o grau l . O termo $e(t)$ representa o ruído e erros de modelagem em geral, e θ_i representa os parâmetros do modelo.

A equação (2.3) pode ser escrita ainda na forma do erro de predição (EP) (Leontaritis and Billings, 1985b):

$$y(t) = \sum_{i=1}^{n_\theta} \varphi_i(t) \hat{\theta}_i + \xi(k) \quad (2.4)$$

onde os termos em chapéu representam os valores estimados, e $\xi(k)$ é definido da seguinte forma:

$$\xi(t) = y(t) - \hat{y}(t, \hat{\theta}) \quad (2.5)$$

$$\therefore \hat{y}(t, \hat{\theta}) = \sum_{i=1}^{n_{\theta}} \varphi_i(t) \hat{\theta}_i \quad (2.6)$$

O vetor de resíduos $\{\xi(t), t = 1, 2, \dots, N\}$ engloba tanto os erros de modelagem quanto o ruído do sistema e incertezas de ordem qualquer. A equação $\hat{y}(t, \hat{\theta}) = \sum_{i=1}^{n_{\theta}} \varphi_i(t) \hat{\theta}_i$ é denominada preditor “um-passo-a-frente” e $\hat{y}(t)$ é a predição “um-passo-a-frente” de $y(t)$.

Os parâmetros θ_i do modelo devem ser escolhidos de modo a minimizar um índice de desempenho pré-estabelecido, como por exemplo (Ljung, 1987):

$$J_N(\theta) = \frac{1}{N} \sum_{k=1}^N f(\xi(t, \theta)) \quad (2.7)$$

onde $f(\xi)$ é uma função escalar qualquer.

O conjunto de parâmetros estimado será diferente para cada f considerada. Usando $f(\xi) = \xi^T$ (tem-se o chamado método dos mínimos quadrados. Nesse caso, a equação (2.7) se transforma em:

$$J_N(\theta) = \frac{1}{N} \sum_{k=1}^N \xi^T \xi(t, \theta) \quad (2.8)$$

que deverá ser minimizado para determinar-se o conjunto de parâmetros correspondente.

Colocando-se a equação (2.4) em notação matricial, tem-se :

$$Y = \Psi^T \Theta + \Xi \quad (2.9)$$

onde

$$\begin{aligned}
 Y &= [y(1) \ y(2) \ \dots \ y(N)]^T \\
 &= [\xi(1) \ \xi(2) \ \dots \ \xi(N)]^T \\
 \Psi^T &= \begin{bmatrix} \varphi_1(1) & \varphi_2(1) & \dots & \varphi_{n_\theta}(1) \\ \varphi_1(2) & \varphi_2(2) & \dots & \varphi_{n_\theta}(2) \\ \vdots & \vdots & & \vdots \\ \varphi_1(N) & \varphi_2(N) & \dots & \varphi_{n_\theta}(N) \end{bmatrix} \\
 \Theta &= [\theta_1 \ \theta_2 \ \dots \ \theta_{n_\theta}]^T
 \end{aligned}$$

e T indica a transposição da matriz (ou vetor).

A solução dos mínimos quadrados é dada por (Ljung, 1987):

$$\hat{\theta}_{LS} = [\Psi \Psi^T]^{-1} \Psi Y \quad (2.10)$$

A matriz $\Psi \Psi^T$ é denominada matriz de covariância dos regressores. Ela é simétrica e positiva definida, exceto em casos de mal-condicionamento grave.

A solução $\hat{\theta}_{LS}$ existe e é única desde que $\Psi \Psi^T$ seja não-singular. Para que isso seja verificado, a matriz Ψ^T deve ter posto pleno de colunas. Nessa situação, as colunas de Ψ^T são linearmente independentes e formam a base do espaço vetorial formado por $\Psi^T \Theta$ (espaço imagem de Ψ^T). Quando Y não pertencer ao espaço vetorial descrito pelas colunas de Ψ^T (como irá acontecer em quase totalidade dos casos), os dados do sistema não poderão ser completamente explicados pelos regressores; $\Psi^T \hat{\theta}_{LS}$, nesse

caso, representará a projeção ortogonal de Y no espaço imagem de Ψ^T e os resíduos da estimação $\xi(t)$ serão mínimos (Ljung, 1987).

A estimativa é dita não-polarizada se o vetor de resíduos $\{\xi(t), t = 1, 2, \dots, N\}$ for ruído branco e não apresentar correlação com os regressores; isto é (Korenberg et al., 1988):

$$E\{\hat{\theta}_{LS}\} = \theta \quad (2.11)$$

onde E representa esperança matemática e θ é o vetor nominal de parâmetros. Quando isso não acontece, os resíduos apresentam alguma dinâmica que não foi devidamente absorvida pelo modelo. Novos termos podem ser incluídos no mesmo para que as estimativas se tornem não-polarizadas e toda a dinâmica dos dados seja refletida no modelo. Na inclusão de mais termos deve-se estar atento para a sobreparametrização do modelo, que podem gerar o aparecimento de regimes dinâmicos que não estão presentes no sistema real (Aguirre and Billings, 1995b). Normalmente, o que se faz é incluir termos de ruído apenas para evitar a polarização dos parâmetros. Obtido o modelo do ruído, despreza-se então a parte estocástica e considera-se apenas a parte determinística para simulação e análise (Aguirre, 1994).

O modelo não-linear deve ser estimado utilizando-se dados puros. Estimação baseada em dados com média nula $y'(\cdot) = y(\cdot) - \bar{y}$, $u'(\cdot) = u(\cdot) - \bar{u}$ como geralmente se faz para o caso linear, irá resultar em um modelo que é “sensível à entrada” (Billings and Voon, 1983). Isto significa que os parâmetros do modelo tornam-se uma função da variância e estatística de ordem elevada do sinal de entrada. Um modelo estimado para uma certa entrada, não será válido para predição baseada em qualquer outra entrada que apresente uma estatística diferente da primeira. Este problema pode ser evitado utilizando-se os dados com seus valores ‘dc’ originais. De acordo com Billings and Fadzil (1985), deve-se evitar subtrair a média dos dados, porque pode induzir à um grande número de termos não-lineares no modelo, que aparecem na tentativa de ‘explicar’ fenômenos relativos à componente ‘dc’ retirada dos dados.

Existem métodos de estimação de parâmetros que são não-polarizados, mesmo quando o ruído do sistema não for branco. Um desses métodos é o das Variáveis Instrumentais (Billings and Voon, 1984) no qual é definida uma nova variável aleatória $v(k)$ que apresenta correlação com os regressores do modelo e é não-correlacionada com o ruído e . Uma revisão neste método pode ser visto em Ljung (1987).

2.2.5 Implementações Ortogonais do Algoritmo de Mínimos Quadrados

Quando Ψ^T é mal condicionada, a formação da matriz de covariância dos regressores, $\Psi\Psi^T$, torna-se sujeita a problemas numéricos (propagação de erros, precisão insuficientes nos cálculos) que podem afetar a estabilidade do algoritmo dos mínimos quadrados. Neste caso, a solução da equação normal (2.10) torna-se inviável. Mal-condicionamento numérico acontece quando as colunas de regressores, em $\Psi\Psi^T$, são altamente correlacionadas entre si.

Uma alternativa para aliviar tal problema é a ortogonalização da matriz Ψ . Nessa situação, as colunas de Ψ serão não-correlacionadas e formarão uma base ortogonal para o espaço vetorial solução de $\Psi^T\Theta$. Os algoritmos utilizados na solução do problema dos mínimos quadrados ortogonais são o procedimento de Gram-Schmidt (clássico e modificado) e o método da transformação de Householder (Ljung, 1987; Chen et al., 1989).

Descrever-se-á, em seguida, o método da transformação de Householder (Chen et al. 1989), a ser utilizado na estimação ortogonal de parâmetros neste trabalho.

Se a matriz Ψ^T tiver posto pleno de colunas, ela pode ser decomposta em:

$$\Psi^T = QR \quad (2.12)$$

onde Q é uma matriz ortogonal ($Q^T Q = I$) N (n_θ e R é uma matriz triangular superior de dimensão $n_\theta \times n_\theta$).

Criando a matriz aumentada $\bar{Q}$ obtém-se :

$$\bar{Q} = [Q | q_{n_\theta+1} \quad \dots \quad q_N] \quad (2.13)$$

$$\Psi^T = \bar{Q} \cdot \begin{bmatrix} R \\ 0 \end{bmatrix} \quad (2.14)$$

onde $\bar{Q}$ tem dimensão N (N). As novas colunas acrescentadas a Q podem ser quaisquer, desde que $\bar{Q}$ tenha posto pleno de colunas. Aplicando $\bar{Q}^T$ a Y (vetor de saídas do sistema):

$$\bar{Q}^T \cdot Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad (2.15)$$

onde y_1 e y_2 têm dimensão $n_\theta \times 1$ e $(N - n_\theta) \times 1$, respectivamente.

Assim:

$$\|\Xi\|_2 = \|Y - \Psi^T \Theta\|_2 = \left\| \bar{Q}^T (Y - \Psi^T \Theta) \right\|_2 = \|y_1 - R\Theta\|_2 + \|y_2\|_2 \quad (2.16)$$

onde $\|\cdot\|_2$ é a norma euclidiana.

A estimativa dos parâmetros pode ser obtida solucionando-se o seguinte sistema linear triangular (usando substituição reversa, por exemplo):

$$R\Theta = y_1 \quad (2.17)$$

A fatoração QR da matriz Ψ^T pode ser implementada utilizando-se transformações de Householder, devidamente projetadas para tanto (Ljung, 1987; Chen *et al.*, 1989).

No procedimento de Gram-Schmidt, a matriz Ψ^T é decomposta no produto de duas matrizes, uma ortogonal e a outra triangular superior, como na fatoração de Householder. A forma de implementar tal decomposição é o que diferencia os dois métodos.

Considere:

$$\Psi^T = WA \quad (2.18)$$

onde A é uma matriz triangular superior com diagonal unitária, de dimensão $n_c(n_c \times n_c)$ e W é uma matriz ortogonal ($W^T W = D$, sendo D uma matriz diagonal), de dimensão N ($n_c \times N$) a ser construída a partir dos dados de operação.

Da equação (2.9):

$$Y = \Psi^T (A^{-1} A) \Theta + \Xi \quad (2.19)$$

Define-se:

$$W = \Psi^T A^{-1} \quad (2.20)$$

$$g = A \Theta \quad (2.21)$$

$$Y = Wg + \Xi \quad (2.22)$$

onde as colunas de W constituem os novos regressores (ditos ortogonais).

A solução dos mínimos quadrados ortogonais (2.22) é (Ljung, 1987; Chen *et al.*, 1989; Stoica *et al.*, 1986) :

$$\hat{g} = D^{-1}W^T Y \quad (2.23)$$

Determinado $\hat{g}$ através de (2.23), os parâmetros do modelo original podem ser obtidos segundo (2.21).

O procedimento de estimação descrito acima pode ser implementado simultaneamente com um esquema de detecção de estrutura baseado no ERR, conforme será descrito na próxima seção.

2.2.6 Detecção de Estrutura Utilizando o ERR

Usando (2.22), define-se a variância do erro de modelagem $\xi(t)$ como:

$$Var\{\xi(t)\} = \frac{1}{N} \Xi^T \Xi = \frac{1}{N} \left[Y^T Y - \sum_{i=1}^{n_g} g_i^2 w_i^T w_i \right] \quad (2.24)$$

onde g_i indica os elementos do vetor de parâmetros g e w_i indica os regressores ortogonais (colunas da matriz W). Supondo que nenhum termo fosse incluído no modelo, a variância de $\xi(t)$ seria igual à variância de saída. A cada novo termo acrescentado ao modelo, a variância dos resíduos $\xi(t)$ é decrescida de um fator $\frac{1}{N} (g_i^2 w_i^T w_i)$, onde w_i indica o termo incluído e g_i o seu respectivo parâmetro. A redução no valor da variância pode ser normalizada com relação à variância total do sinal de saída ($y(t)$). Isto resulta no já citado “Error Reduction Ratio” (ERR) de cada termo, o qual é definido formalmente como (Chen *et al.*, 1989) :

$$[ERR]_i = \frac{g_i^2 w_i^T w_i}{Y^T Y}, \quad 1 \leq i \leq n_\theta \quad (2.25)$$

Portanto, o ERR indica a porção da variância da saída explicada pela inclusão de um novo termo no modelo. Ele pode ser utilizado na detecção de estrutura de modelos. Escolhe-se o número de termos desejados para o modelo (utilizando-se por exemplo, critérios de informação) e normalmente considerar-se-á aqueles que possuem os maiores ERR's.

O valor do ERR depende da posição em que o termo considerado é incluído no modelo. Em Billings et al. (1989) é apresentado um esquema de detecção de estrutura, baseado no ERR, que visa contornar este defeito. No algoritmo, denominado “regressão-direta”, todos os regressores $\phi_i(t)$ são considerados na determinação de cada $w_i(t)$, sendo escolhido aquele que apresentar o maior ERR naquela posição.

Existem critérios de informação que podem ser úteis na determinação do número de termos a ser considerado no modelo, e serão utilizados nos algoritmos de estimação deste trabalho. Esses critérios baseiam-se na estimação da ordem do modelo a partir dos dados de operação. Entre tais métodos destacam-se o critério de informação de Akaike (AIC) (Akaike, 1974), o critério de informação de Bayes (BIC) (Kashyap, 1977), LILC (Khinchin's law of iterated logarithm criterion) (Hannan and Quinn, 1979) e técnica da predição final do erro (FPE) (Akaike, 1970).

Nos algoritmos descritos nesta seção, a estimação de parâmetros é feita de maneira “off-line”. Os dados do sistema são obtidos, através dos procedimentos de aquisição de dados, e em seguida, processados para a obtenção do vetor de parâmetros do modelo. A estimação pode também ser implementada recursivamente. Nesse caso, o vetor de parâmetros é inicializado e recalculado (segundo uma expressão recursiva) a cada intervalo de amostragem, quando um novo registro de entrada e saída do sistema é obtido.

Os algoritmos recursivos de estimação são normalmente utilizados quando o sistema que está sendo modelado é variante no tempo (Pröll and Karim, 1994). Nessa situação, os parâmetros precisam ser continuamente estimados e mudanças refletidas nos dados do sistema devem ser capazes de forçar a atualização dos parâmetros. Para que isso aconteça, os algoritmos recursivos utilizam um fator de esquecimento, geralmente exponencial, que privilegia os dados mais recentes em detrimento dos mais antigos. A estimação recursiva também é utilizada no projeto de controladores adaptativos.

Um objetivo deste trabalho é obter resultados para auxiliar o ERR, que é um critério puramente estatístico na escolha de termos constituintes do modelo, através da limitação do conjunto de termos candidatos.

2.2.7 Validação de Modelos Identificados

Quando os parâmetros são estimados, não há garantias de que foram selecionados apenas os termos significantes no modelo NARMAX que representa o sistema real (Billings and Voon, 1986a). A ausência de algum termo importante no modelo pode provocar a polarização dos parâmetros, como já dito anteriormente. Desta maneira torna-se necessário submeter o modelo obtido a alguns testes que deverão avaliar a sua capacidade em representar o sistema original; estes testes são os testes de validação.

Existem dois tipos básicos de validação, a estatística e a dinâmica, que serão apresentados a seguir.

2.2.7.1 Validação Estatística

Algumas condições observadas a fim de que a estimação seja não polarizada diz respeito à correlação do ruído. Entretanto, como este não pode ser medido, é necessário utilizar os *resíduos* que servem como uma estimativa do ruído.

A motivação para se testar os resíduos de um modelo baseia-se no fato de que se um modelo ajustar bem os dados (não polarizado), ele deve ser capaz de explicar toda a dinâmica presente nos dados. Se isto ocorrer, não haverá dinâmica restante nos resíduos, que serão considerados brancos. Portanto, validar estatisticamente equivale a verificar a aleatoriedade dos resíduos de tal modelo.

Sabe-se que os testes lineares tradicionais que calculam a autocorrelação dos resíduos e a correlação cruzada da entrada e os resíduos, geralmente falham ao indicar a deficiência do modelo quando o processo é não-linear (Billings and Leontaritis, 1981; Billings and Voon, 1983). Esta falha pode levar à hipótese de que não há mais informação nos resíduos, quando na verdade deveriam existir mais termos não-lineares que ao serem omitidos podem gerar estimativas polarizadas dos parâmetros do modelo.

Portanto, como mencionado anteriormente, quando o sistema é não-linear, os resíduos devem ser imprevisíveis a partir de todas as combinações de entradas e saídas passadas (Billings and Voon, 1983), e esta condição se satisfaz caso:

$$\begin{aligned}\Phi_{\xi\xi}(h) &= E\{\xi(k-h)\xi(k)\} = \delta(0) \\ \Phi_{u\xi}(h) &= E\{u(k-h)\xi(k)\} = 0, \forall h \\ \Phi_{\xi\xi u}(h) &= E\{\xi(k-1-h)\xi(k)u(k-1-h)\} = 0, \forall h \\ \Phi_{u^2\xi}(h) &= E\{(u^2(k-h) - \overline{u^2(k)})\xi(k)\} = 0, \forall h \\ \Phi_{u^2\xi^2}(h) &= E\{(u^2(k-h) - \overline{u^2(k)})\xi^2(k)\} = 0, \forall h\end{aligned}$$

onde $\xi(k)$ representa os resíduos, (a função de correlação, (é o impulso unitário e a barra indica o valor médio. As condições acima devem ser sempre satisfeitas no caso de modelos não-polarizados.

Como o vetor de dados (e resíduos) é finito, as igualdades acima não poderão ser verificadas exatamente. Entretanto, define-se “limites de confiança” de 95% em $\pm 1,96/\sqrt{N}$, onde N é o número de dados. Assim se houver picos significativos fora dos limites de confiança das funções de correlação, há 95% de chance daqueles resíduos não serem brancos.

Para o caso em que o modelo não-linear estimado é autônomo, os seguintes testes deverão ser usados (Billings and Tao, 1991):

$$\begin{aligned}\Phi_{\xi\xi}(h) &= E\{(\xi(k) - \overline{\xi(k)})(\xi(k-h) - \overline{\xi(k)})\} = \delta(0) \\ \Phi_{\xi\xi^2}(h) &= E\{(\xi(k) - \overline{\xi(k)})(\xi(k-h) - \overline{\xi(k)})^2\} = 0 \\ \Phi_{\xi^2\xi^2}(h) &= E\{(\xi^2(k) - \overline{\xi^2(k)})(\xi^2(k-h) - \overline{\xi^2(k)})\} = \delta(0)\end{aligned}$$

No caso de sistemas lineares não-autônomos, apenas os testes Φ_{ξ} e Φ_u são considerados; no caso de sistemas lineares autônomos apenas o teste Φ_{ξ} é normalmente utilizado (Söderström and Stoica, 1989).

2.2.7.2 Validação Dinâmica

A validação dinâmica de um modelo visa verificar se este modelo realmente reproduz certas características dinâmicas do sistema identificado. Algumas propriedades dinâmicas que podem ser utilizadas para validar modelos são (Aguirre and Billings, 1995b; Aguirre and Billings, 1994a):

- expoentes de Lyapunov;
- seções (ou mapas) de Poincaré;
- diagramas de bifurcação.

Os invariantes dinâmicos acima citados descrevem aspectos qualitativos e quantitativos da dinâmica do sistema, sendo que os mais indicados para a validação dinâmica são os diagramas de bifurcação, pois são mais sensíveis às mudanças na estrutura do modelo.

Uma prática também comum consiste em tomar uma massa de dados diferentes dos dados de estimação e aplicar os dados de entrada destes no modelo obtido. O resultado fornecido pelo modelo é comparado com os dados de saída reais, então é feita uma avaliação visual da capacidade do modelo de acompanhar novas entradas.

Vale a pena salientar que em algumas aplicações não é possível utilizar os invariantes citados acima, porque tais valores não são facilmente obtidos a partir do sistema original.

2.3 Comentários Finais

Nesse capítulo, foram revistos os principais passos da identificação de sistemas, que será aplicada ao problema de obtenção de modelos para sistemas não-lineares.

O processo de obtenção de modelos não-lineares é, em essência, semelhante à identificação de sistemas lineares (problema este conhecido na literatura). Os dois procedimentos são diferenciados pela maneira com que cada etapa do processo é implementada.

O projeto de excitações (sinais de entrada) para coleta de dados de sistemas não-lineares (dados a serem utilizados para identificação) deve levar em consideração que esses sistemas podem apresentar comportamentos dinâmicos completamente distintos, dependendo da região de operação.

Os algoritmos de detecção de não-linearidades têm como função informar o grau de não-linearidade apresentado nos dados de experimentação do sistema. De acordo com a

indicação destes algoritmos, será feita a opção entre a utilização de um modelo linear ou um modelo não-linear. Assim, os dados de experimentação devem refletir o comportamento do sistema em todas as suas faixas de operação. É desejável, sempre que possível, utilizar modelos lineares, os quais são mais simples e fáceis de serem analisados e controlados.

A grande dificuldade apresentada pelo modelamento não-linear é a escolha da estrutura a ser utilizada para representar a dinâmica do sistema. Essa estrutura deve ser suficientemente complexa para permitir a representação de uma classe ampla de sistemas não-lineares. Por outro lado é desejável que os modelos obtidos sejam passíveis de manipulação computacional posterior.

Os algoritmos de identificação devem ser capazes de selecionar os termos necessários (dentro de um conjunto de termos candidatos) para reproduzir o comportamento dinâmico refletido nos dados de experimentação. A sobreparametrização de modelos (não-lineares) faz que estes apresentem características não contidas nos dados. A estimação de parâmetros deve ajustar a estrutura escolhida para representar o sistema em consideração. Por fim, os modelos obtidos devem passar por testes que irão avaliar a capacidade daqueles em ajustar os dados de identificação e reproduzir a dinâmica apresentada pelo sistema.

Deve ser ressaltado boa parte das técnicas apresentadas nesse capítulo foram desenvolvidas nos últimos 30 anos, e que esse campo de pesquisa não se encontra encerrado visto a grande quantidade de publicações recentes nesta área (Nagy and Ljung, 1992; Lindskog and Ljung, 1994; Ljung and Guo, 1995; Schoukens et al., 1994). Um método recente que auxilia na detecção de estrutura será citado no próximo capítulo.

3 Seleção de Estrutura de Modelos Não-Lineares: Agrupamentos de Termos e Coeficientes de Agrupamentos

3.1 Introdução

Uma das maiores dificuldades na identificação de sistemas é a escolha da estrutura do modelo, isto é, a escolha da base do modelo antes da estimação de parâmetros. A seleção de uma estrutura adequada é de vital importância no desenvolvimento de modelos não-lineares, especialmente os modelos polinomiais (Aguirre, 1994b).

Uma das dificuldades que aparecem na escolha da estrutura nas representações polinomiais, é que o número de termos possíveis de integrarem o modelo cresce rapidamente com o aumento da ordem e o do grau de não-linearidade do modelo. Então, mesmo para pequenos valores destas grandezas, o número de termos em um modelo polinomial pode se tornar impraticável. O mais importante é que, mesmo se o tempo de computação do modelo e os requisitos de memória estiverem disponíveis, a inclusão de termos no modelo além dos necessários normalmente leva-o a apresentar regimes dinâmicos espúrios (Aguirre and Billings, 1995b).

Neste capítulo serão discutidos os principais efeitos da inclusão de termos espúrios no modelo, e em seguida será apresentada a técnica de agrupamentos de termos e coeficientes de agrupamentos, que pode auxiliar na seleção de termos em modelos polinomiais. O capítulo está organizado como se segue: na seção 3.2 comenta-se a respeito dos problemas apresentados por modelos polinomiais NARMAX muito complexos. Na seção 3.3 é apresentada a técnica de agrupamentos de termos e coeficientes de agrupamentos. Na seção 3.4 é verificada a importância de um agrupamento, e são citados exemplos. E por fim, na seção 3.5 são feitos os comentários finais a respeito do capítulo.

3.2 Principais Problemas de Modelos muito Complexos

Já foi mencionado anteriormente que os modelos polinomiais, neste caso os modelos polinomiais NARMAX apresentam um aumento muito rápido do número possível de termos com o aumento da ordem e do grau de não-linearidade. Então para que tal modelo seja dinamicamente válido e conciso, é preciso que a estrutura final seja restringida (Aguirre and Billings, 1995b).

Normalmente, parece haver uma tendência ao uso de modelos que são mais complexos do que o necessário. As possíveis razões para esta tendência já foram comentadas no capítulo anterior. Entretanto, o fato de modelos maiores ajustarem melhor os dados não quer dizer que eles necessariamente capturam a dinâmica do sistema apropriadamente, afirmação esta que poderá ser comprovada nos capítulos posteriores.

O número mínimo de termos que devem ser incluídos no modelo pode ser definido como o número mínimo de termos requeridos para que o modelo estimado seja aprovado por certos testes de validação. Em outras palavras, o número mínimo de termos em um modelo está diretamente associado com o tamanho do mais simples modelo válido de acordo com algum critério específico.

Na prática, este número pode ser encontrado através da introdução gradual de termos no modelo e posterior verificação se este modelo reproduz bem a dinâmica desejada. Quando o modelo é não-linear, normalmente é difícil de se determinar o seu melhor tamanho, levando ao acréscimo de mais termos do que o necessário. Neste ponto, os modelos não-lineares em geral são sensíveis à sobreparametrização.

Além do mal-condicionamento numérico e possível instabilidade, modelos sobreparametrizados podem apresentar um grave problema que é a presença de regimes espúrios. Ou seja, o acréscimo de termos espúrios ao modelo podem levá-lo a regimes dinâmicos não condizentes com o sistema real. A idéia do conceito de agrupamento de termos e coeficientes de agrupamentos é justamente ajudar os critérios já existentes na

identificação de tais termos espúrios, possibilitando um modelo conciso e dinamicamente válido.

3.3 Agrupamentos de Termos e Coeficientes de Agrupamentos

Existem diversas técnicas de escolha de estruturas para modelos polinomiais (Billings et. al., 1988; Haber and Unbehauen, 1990; Wang and Cluett, 1996). A maioria destas técnicas proporcionam uma maneira detalhada de selecionar os termos mais importantes, normalmente a partir de um conjunto muito grande de termos candidatos ou, em alguns casos, pela eliminação de termos não necessários com base em um modelo já estimado. Entretanto, o desempenho de tais métodos normalmente depende de fatores tais como o tempo de amostragem (Billings and Aguirre, 1995) e a presença de ruído (Aguirre, 1994b).

O objetivo desta seção é revisar os conceitos de agrupamento de termos que pode ser utilizado como uma ferramenta em conjunto com as técnicas já existentes na seleção de estruturas, de forma a reduzir os efeitos da sobreparametrização. Os conceitos de agrupamentos de termos e coeficientes de agrupamentos são definidas para modelos polinomiais NARMAX que é a representação escolhida para os modelos neste trabalho. É mostrado também como tais conceitos podem ajudar na escolha de estruturas.

Definição 3.1 (Aguirre and Billings, 1995a)

Seja o modelo polinomial NARMAX genérico representado pela equação (2.1). A parte determinística do modelo NARMAX, isto é, a parte NARX, pode ser expandida como a soma de termos com graus de não-linearidade na faixa de $1 \leq m \leq l$, onde l é o grau de não-linearidade e m é a ordem do termo. Cada termo de ordem m contém um fator de ordem p em $y(k - n_i)$ e um fator de ordem $(m - p)$ em $u(k - n_i)$, e é multiplicado por um coeficiente $c_{p,m-p}(n_1, \dots, n_m)$ como segue:

$$y(k) = \sum_{m=0}^l \sum_{p=0}^m \sum_{n_1, n_m}^{n_y, n_u} c_{p, m-p}(n_1, \dots, n_m) \prod_{i=1}^p y(k - n_i) \prod_{i=p+1}^m u(k - n_i) \quad (3.1)$$

onde

$$\sum_{n_1, n_m}^{n_y, n_u} \equiv \sum_{n_1=1}^{n_y} \dots \sum_{n_m=1}^{n_u} \quad (3.2)$$

Em (3.2) limite superior é n_y se a soma dos fatores se refere a $y(k - n_i)$ ou n_u para fatores em $u(k - n_i)$. Expandindo-se a equação (3.1) até a segunda ordem, isto é, até $l=2$ tem-se:

$$\begin{aligned} y(k) = & c_{0,0} + \sum_{n_1=1}^{n_y} c_{1,0}(n_1)y(k - n_1) + \sum_{n_1=1}^{n_u} c_{0,1}(n_1)u(k - n_1) + \\ & + \sum_{n_1=1}^{n_y} \sum_{n_2=1}^{n_y} c_{2,0}(n_1, n_2)y(k - n_1)y(k - n_2) + \sum_{n_1=1}^{n_y} \sum_{n_2=1}^{n_u} c_{1,1}(n_1, n_2)y(k - n_1)u(k - n_2) + \\ & + \sum_{n_1=1}^{n_y} \sum_{n_2=1}^{n_u} c_{0,2}(n_1, n_2)u(k - n_1)u(k - n_2) \end{aligned} \quad (3.3)$$

Note-se que os coeficientes dos termos dependem do tempo de amostragem e deveriam ser então representados como $c_{p, m-p}(T_s, n_1, \dots, n_m)$. Entretanto, para facilitar o problema, o argumento T_s é descartado.

Exemplo 3.1

O modelo:

$$\begin{aligned} y(t) = & 2,1579y(t-1) - 1,3203y(t-2) + 0,16239y(t-3) + 0,22480 \times 10^{-3}y(t-3)^3 - \\ & 0,48196 \times 10^{-2}y(t-1)^3 + 0,19463 \times 10^{-2}u(t-2) + 0,3416 \times 10^{-3}u(t-1) + \\ & 0,3523 \times 10^{-2}y(t-1)^2y(t-2) - 0,12162 \times 10^{-2}y(t-1)y(t-2)y(t-3) \end{aligned} \quad (3.4)$$

pode ser descrito na forma da equação (3.1) com

$$\begin{aligned} c_{1,0}(1) &= 2,1579 & c_{1,0}(2) &= -1,3203 \\ c_{1,0}(3) &= 0,16239 & c_{3,0}(3,3,3) &= 0,22480 \times 10^{-3} \end{aligned}$$

$$\begin{aligned}
 c_{3,0}(1,1,1) &= -0,48196 \times 10^{-2} & c_{0,1}(2) &= 0,1963 \times 10^{-2} \\
 c_{0,1}(1) &= 0,34160 \times 10^{-3} & c_{3,0}(1,1,2) &= 0,35230 \times 10^{-2} \\
 c_{3,0}(1,2,3) &= -0,12162 \times 10^{-2} & &
 \end{aligned} \tag{3.5}$$

Se o tempo de amostragem T_s for suficientemente pequeno, pode-se dizer que

$$\begin{aligned}
 y(k-1) &\approx y(k-2) \approx \dots \approx y(k-n_y) \\
 u(k-1) &\approx u(k-2) \approx \dots \approx u(k-n_u)
 \end{aligned} \tag{3.6}$$

ou seja, as equações (3.6) são análogas às equações representativas do regime estático de um sistema, definido no capítulo 1. Assim, pode-se considerar que o cálculo dos agrupamentos de termos e coeficientes de agrupamentos são feitos mediante a aproximação do sistema em regime estático. Então a equação (3.1) pode ser reescrita como

$$y(k) \approx \sum_{n_1, n_m}^{n_y, n_u} c_{p, m-p}(n_1, \dots, n_m) \sum_{m=0}^l \sum_{p=0}^m y(k-1)^p u(k-1)^{m-p} \tag{3.7}$$

Definição 3.2 (Aguirre and Billings, 1995a)

As constantes $\sum_{n_1, n_m}^{n_y, n_u} c_{p, m-p}(n_1, \dots, n_m)$ são os coeficientes dos agrupamentos de termos $\Omega_{y^p u^{m-p}}$ que contém termos da forma $y(k-i)^p u(k-j)^{m-p}$ para $m = 0, \dots, l$ e $p = 0, \dots, m$. Tais coeficientes, ditos coeficientes de agrupamentos, são representados por $\Sigma_{y^p u^{m-p}}$.

As aproximações na equação (3.6) foram feitas no sentido de apontar o raciocínio em que se baseia a análise de agrupamentos de termos, bem como introduzir uma definição mais formal de agrupamentos de termos e coeficientes de agrupamentos. Entretanto, tais conceitos são válidos na maioria das aplicações nas quais o tempo de amostragem é

selecionado de forma a permitir uma estimação de parâmetros confiável (Billings and Aguirre, 1995).

Portanto, um conjunto de termos candidatos para um modelo NARX nada mais é do que a união de todos os agrupamentos possíveis até o grau l , isto é:

$$\begin{aligned} \{\text{todos os termos possíveis}\} = & \bigcup_{\substack{p=0\dots m \\ m=0\dots l}} \Omega_{y^p u^{m-p}} = \text{constante} \cup \Omega_y \cup \Omega_u \cup \Omega_{y^2} \cup \Omega_{yu} \cup \Omega_{u^2} \cup \dots \\ & \dots \cup \text{todas as combinações possíveis até o grau } l \end{aligned} \quad (3.8)$$

Exemplo 3.2

Os coeficientes de agrupamentos do modelo da equação (3.4) são:

$$\begin{aligned} c_{1,0}(1) + c_{1,0}(2) + c_{1,0}(3) &= \Sigma_y = 0,999999 \\ c_{3,0}(3,3,3) + c_{3,0}(1,1,1) + c_{3,0}(1,1,2) &= \Sigma_{y^3} = -2,2880 \times 10^{-3} \\ c_{0,1}(1) + c_{0,1}(2) &= \Sigma_u = 2,2879 \times 10^{-3} \end{aligned} \quad (3.9)$$

que correspondem aos agrupamentos de termos $\Omega_y, \Omega_{y^3}, \Omega_u$ respectivamente.

Logo, um agrupamento de termos, $\Omega_{y^p u^{m-p}}$, é um conjunto de termos da forma $y(k-i)^p u(k-j)^{m-p}$ para $m = 0, \dots, l$ e $p = 0, \dots, m$, e o seu respectivo coeficiente de agrupamento, $\Sigma_{y^p u^{m-p}}$, é o somatório dos coeficientes de todos os termos no modelo que estão contidos em tais agrupamentos. É interessante notar que, no limite quando $T_s \rightarrow 0$, todos os termos de um agrupamento tornam-se indistinguíveis, e todos os coeficientes de agrupamentos anulam-se, exceto Σ_y , que no limite é unitário, isto é:

$$\lim_{T_s \rightarrow 0} \Sigma_y = 1, \quad \lim_{T_s \rightarrow 0} \Sigma_{y^p u^{m-p}} = 0 \text{ para todos os outros agrupamentos} \quad (3.10)$$

Na prática, uma das formas de utilizar agrupamentos de termos consiste em estimar-se uma família de modelos variando-se o número de termos de um valor mínimo até um valor máximo de acordo com o interesse, calculando-se os valores dos coeficientes de

agrupamento de cada modelo. Ao final traça-se um gráfico com o histórico da variação dos coeficientes de agrupamentos em função do número de termos no modelo. A identificação de agrupamentos espúrios é explicada a seguir.

3.4 Verificação da importância de um agrupamento

Nesta seção será discutido que o efeito dinâmico dos resíduos de termos provenientes de agrupamentos espúrios é desprezível. A fim de mostrar tal consideração, uma aproximação intuitiva será utilizada (Aguirre and Billings, 1995a).

Assuma que uma massa de dados foi obtida pela amostragem da saída e possivelmente da entrada de um sistema dinâmico contínuo. Se o modelo

$$y(k) = \Psi^T(k-1)\hat{\Theta} + \xi(k) \quad (3.12)$$

é ajustado a tais dados, e os parâmetros não são polarizados, sabe-se que os resíduos, $\xi(k)$, devem ser imprevisíveis a partir de todas as combinações de termos lineares e não-lineares, ou seja, ele pode ser considerado “branco”. Em outras palavras, um modelo é dito não polarizado se ele explica toda a dinâmica relevante presente nos dados. Se alguma dinâmica não é explicada pelo modelo, ela irá aparecer nos resíduos e normalmente será detectada por testes apropriados de validação não-linear (ver seção 2.2.7.1).

Se considerarmos que a estrutura de um modelo polinomial NARMAX pode ser composta de termos efetivos e termos espúrios, a equação (3.11) pode ser reescrita da seguinte maneira (Aguirre and Billings, 1995a):

$$y(k) = \Psi_E^T \hat{\Theta}_E + [\Psi_s^T \hat{\Theta}_s + \xi(k)] \quad (3.13)$$

onde $\Psi^T(k-1) = [\Psi_E^T \quad \Psi_S^T]$ e $\hat{\Theta} = [\hat{\Theta}_E^T \quad \hat{\Theta}_S^T]^T$. Os subscritos E e S denotam, respectivamente, termos efetivos e espúrios no modelo.

Sabe-se que o efeito de agrupamentos espúrios é desprezível na medida em que os resíduos são “brancos”. Então, se um agrupamento de termos é espúrio, isto será revelado através de um coeficiente de agrupamento desprezível, devido ao fato de que o efeito de $\Psi_S^T \hat{\Theta}_S$ nos resíduos ser cancelado dentro de cada agrupamento no modelo. O aparecimento no modelo de termos pertencentes a um mesmo agrupamento se cancelando, indica que tais termos poderiam ser espúrios e conseqüentemente retirados do modelo. Este cancelamento seria o análogo não-linear ao cancelamento de pólo-zero nos modelos lineares (Aguirre and Billings, 1995a).

O coeficiente de um agrupamento não é drasticamente alterado pela presença de ruído nos dados de identificação (desde que a relação sinal/ruído seja suficientemente elevada). Por outro lado, o coeficiente dos agrupamentos espúrios irá apresentar grandes variações quando o nível de ruído nos dados é alterado. Este fato poderá ser utilizado como indicação de que tais agrupamentos não devem estar presentes no modelo final.

Portanto, esta seção pode ser resumida em alguns tópicos importantes no auxílio a identificação de agrupamentos de termos espúrios. Ou seja, agrupamentos de termos espúrios normalmente apresentam as seguintes características:

- coeficientes de agrupamento de valor muito baixo comparado aos valores dos demais agrupamentos;
- coeficientes de agrupamentos oscilatórios através de uma faixa variável de valores, podendo eventualmente oscilar em torno do zero;
- coeficientes de agrupamentos que não apresentam uma tendência definida.

Exemplo 3.3 (Aguirre and Billings, 1995a)

A equação do oscilador de Duffing-Ueda foi utilizada para gerar um conjunto de 900 dados amostrados em $T_s = \frac{\pi}{30}$. Fixando o grau de não-linearidade $l = 3$ e $n_y = n_u = 3$, uma família de modelos com um número crescente de termos foi estimada. Os termos foram selecionados automaticamente baseados no critério do ERR. A figura 3.1 mostra os coeficientes de agrupamento Σ_y , Σ_u , Σ_{y^3} , Σ_{u^3} . Foi implementado um programa em MATLABTM (MathWorks, 1990a) para calcular e traçar a variação dos coeficientes de agrupamento (Jácome e Aguirre, 1996b), sendo utilizado em conjunto com as rotinas de identificação (Rodrigues e Aguirre, 1996).

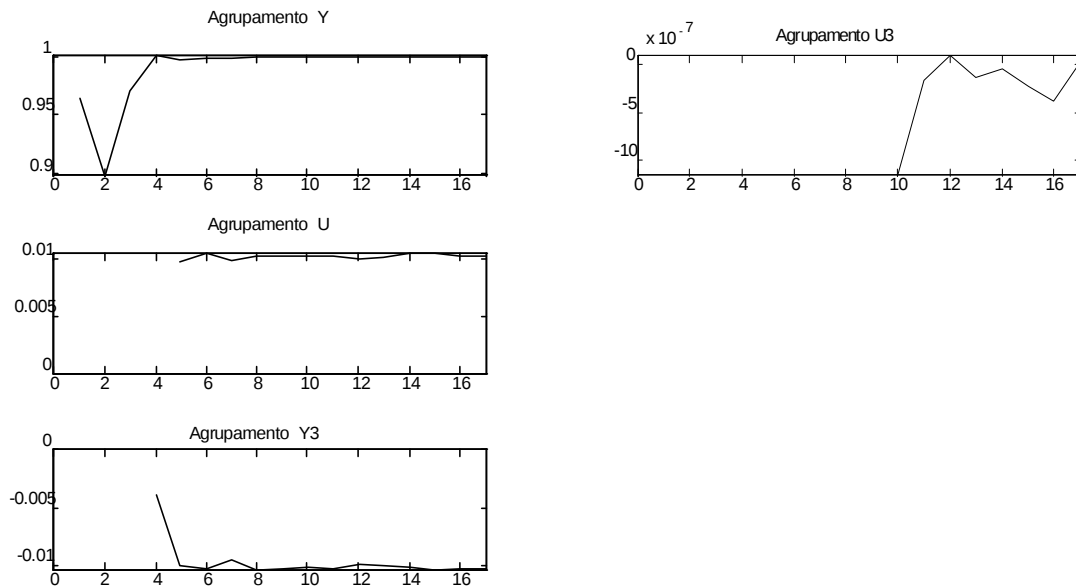


Figura 3.1- Variação dos coeficientes de agrupamento em função do número de termos no modelo.

Três pontos relativos ao coeficiente de agrupamento Σ_{u^3} devem ser observados; (i) o primeiro termo deste agrupamento foi o décimo a integrar o modelo, enquanto que os demais já faziam parte do modelo com cinco termos; (ii) o valor de Σ_{u^3} tende a oscilar, e em torno do zero; e (iii) a magnitude de tal coeficiente é bem menor que os demais coeficientes dos outros agrupamentos, na verdade Σ_{u^3} atinge valores tão baixos quanto

$-6,4637 \times 10^{-8}$. Baseando-se nestas informações parece apropriado excluir todos os agrupamentos de termos do conjunto de termos candidatos, exceto os agrupamentos Ω_y, Ω_u e Ω_{y^3} .

Tomou-se então a mesma massa de dados e estimou-se um modelo com 14 termos, com os mesmos valores de l , n_y e n_u , porém o conjunto de termos candidatos era composto apenas pelos agrupamentos Ω_y, Ω_u e Ω_{y^3} . Com base em certos critérios de validação dinâmica, percebeu-se que o modelo estimado sem a presença dos agrupamentos espúrios refletia melhor a dinâmica do sistema real.

Exemplo 3.4

Utilizando dados de uma série temporal obtida experimentalmente, foram identificados modelos NARMA (autônomos) polinomiais com $l = 3$ e $n_y = 3$, com número de termos no modelo variando de 14 a 20. Os resultados dos coeficientes de agrupamento são mostrados na figura 3.2.

Através dos gráficos da figura 3.2 verifica-se que o coeficiente do agrupamento Ω_{y^2} apresenta uma magnitude muito reduzida com relação aos demais. Note-se também que o coeficiente Σ_0 apresenta uma troca de sinal, e permanece em torno do zero, para $15 \leq n_\theta \leq 19$. Se for observada a magnitude dos demais agrupamentos, verifica-se que estas variam relativamente pouco em torno de um certo valor diferente de zero. Esses fatos sugerem que os agrupamentos Ω_0 e Ω_{y^2} podem ser retirados do conjunto de termos candidatos, bem como os demais agrupamentos que não apareceram durante a estimação.

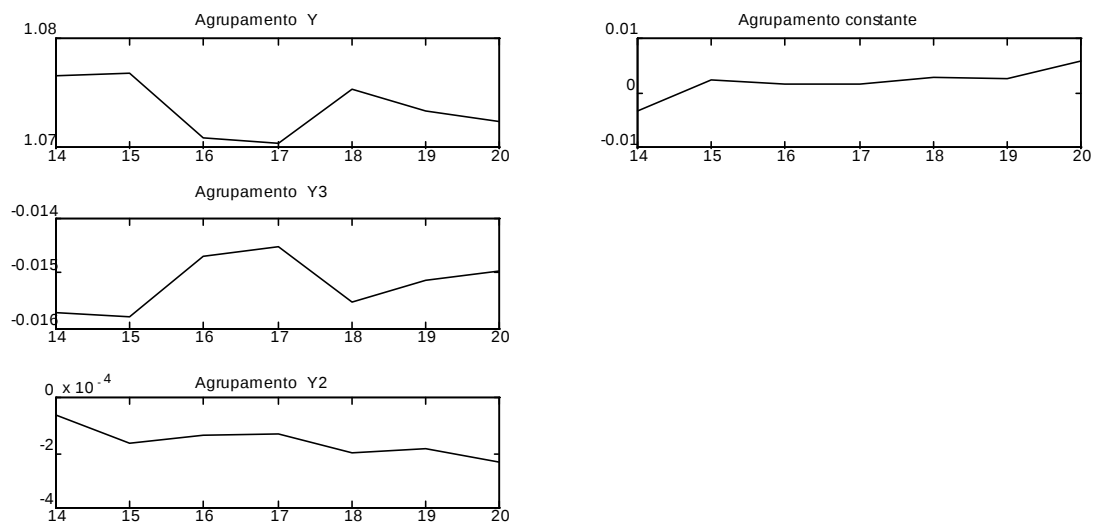


Figura 3.2 - Variação dos coeficientes de agrupamento em função do número de termos no modelo.

3.5 Comentários Finais

Neste capítulo foram citadas algumas utilidades da técnica de agrupamento de termos e coeficientes de agrupamento no *auxílio* à seleção de estruturas de modelos polinomiais. Foram vistas as principais características de agrupamentos de termos que podem fornecer uma idéia de serem espúrios ou não. Porém será visto nos capítulos posteriores que *nem sempre o quase cancelamento de termos de um mesmo agrupamento indica que tal agrupamento seja necessariamente espúrio*.

Através de resultados experimentais como o exemplo 3.3 verifica-se que os modelos polinomiais não-lineares identificados a partir do conjunto de agrupamentos efetivos têm melhores chances de reproduzir as características dinâmicas do sistema original, bem como certas características estáticas, como será visto posteriormente também. Além disso, estes modelos tendem a ser mais robustos aos efeitos prejudiciais da sobreparametrização, principalmente pelo fato de modelos polinomiais serem bem sensíveis a este efeito (Aguirre and Billings, 1995b).

Modelos levemente sobreparametrizados, em que o excesso de termos são parte dos agrupamentos efetivos, ainda podem reproduzir a dinâmica original sem introduzirem regimes dinâmicos espúrios. Este fato se deve ao período de amostragem ser curto o bastante, de forma que termos pertencentes a um mesmo agrupamento são muito semelhantes. A inclusão de alguns termos não terá o efeito de compartilhar informação entre os termos durante a estimação de parâmetros. Entretanto, se estes termos excedentes são retirados dos agrupamentos espúrios, o tipo de não-linearidade que estes termos estão aptos a modelar e que não estão presentes nos dados, aparecem na estrutura do modelo incluindo regimes dinâmicos espúrios.

Uma das vantagens da utilização desta técnica é que o tempo de computação pode ser sensivelmente diminuído, com a utilização de somente termos efetivos na identificação, levando a modelos concisos e dinamicamente válidos. No exemplo 3.3, o conjunto de termos candidatos considerando-se todos os agrupamentos possíveis era de 64 termos. Após a estimação dos modelos, e identificação dos agrupamentos espúrios através da técnica apresentada neste capítulo, retirou-se do conjunto de termos candidatos os agrupamentos espúrios, passando a ser composto apenas de 16 termos considerados efetivos. Isto corresponde a uma redução de 75% no tamanho original do conjunto de termos candidatos. Pode ser notado que maiores reduções podem acontecer em sistemas de ordem mais elevada.

Os resultados apresentados neste capítulo sugerem que os coeficientes de agrupamentos de termos espúrios são normalmente menores do que de termos efetivos. Uma questão intuitiva que pode aparecer é: Quanto é “muito menor?” (Aguirre and Billings, 1995a). Pode-se considerar que o coeficiente de agrupamentos espúrios são 10^4 vezes menor que o maior coeficiente presente. Porém tal resultado não pode ser aplicado cegamente, porque na presença de valores da entrada e/ou saída um pouco maiores, pode levar o coeficiente de um agrupamento a ter um valor pequeno, para compensar a amplitude de certos agrupamentos como por exemplo Ω_u^3 . Uma forma de minimizar este problema é através da normalização dos dados.

O que se faz é, se o coeficiente de um agrupamento considerado espúrio não for convincentemente menor que os demais, geralmente traça-se o gráfico da variação dos coeficientes em função do número de termos no modelo, tal como as figuras 3.1 e 3.2. Agrupamentos espúrios, como já foi dito, normalmente revelam-se insignificantes ou oscilatórios. Mais ainda, se for acrescentado ruído aos dados, ou se uma janela diferente de dados for utilizada, os coeficientes de agrupamentos efetivos geralmente não mudam drasticamente. Em geral este não será o caso de agrupamentos espúrios.

Em Aguirre e Mendes (1996), é mostrado que o conceito de agrupamento de termos e coeficientes de agrupamentos têm fortes vínculos com os aspectos dinâmicos do modelo. Sendo caracterizado que o número de pontos fixos de um modelo polinomial depende dos agrupamentos de termos presentes na estrutura do modelo, e que a localização de tais pontos fixos é completamente determinada pelos respectivos coeficientes de agrupamentos. No capítulo 6, será visto que em certos casos, uma análise semelhante a dos pontos fixos é aplicada, na tentativa de representar certas propriedades do sistema empírico presente no modelo.

Portanto, as principais propriedades da técnica de agrupamentos de termos e coeficientes de agrupamentos podem ser resumidas a seguir:

- (i) se um modelo é composto de termos selecionados a partir dos agrupamentos considerados efetivos, tal modelo estará mais apto a reproduzir fielmente as características do sistema real;
- (ii) os modelos estimados são mais robustos à sobreparametrização devido ao fato de que termos em excesso não pertencem a agrupamentos considerados espúrios;
- (iii) este procedimento é robusto à presença do ruído.

4 Análise de Modelos Discretizados de Sistemas com Constante de Tempo e Ganho Variáveis

4.1 Introdução

Dentre as etapas do processo de identificação de modelos matemáticos que representem sistemas reais, uma das mais importantes e difíceis é a seleção da estrutura do modelo, ou seja, escolhida uma representação deve-se determinar quais termos irão compor o modelo.

Como já foi dito no capítulo 2, modelos não-lineares, que são o foco deste trabalho, tendem a representar melhor certos sistemas do que modelos lineares (Rodrigues, 1996). Porém, nos modelos não-lineares o número de termos possíveis cresce rapidamente com o aumento da ordem e do grau de não-linearidade, dificultando a escolha de quais termos devem compor o modelo. Além disso, modelos com muitos termos apresentam restrições como grande tempo de computação e possível mal-condicionamento numérico, além de uma forte tendência a apresentarem regimes dinâmicos espúrios decorrentes da sobreparametrização (Aguirre and Billings, 1995b).

Seria interessante então que, conhecimentos *a priori* do sistema (vide capítulo 1) pudesse ser usado para reduzir o conjunto de termos candidatos, e talvez o número de termos do modelo. Tal procedimento aumentaria as chances de obter modelos polinomiais NARMAX com representações concisas e dinamicamente válidas.

Através da manipulação de modelos discretizados pode-se perceber matematicamente o que acontece com as constantes de tempo e com o ganho do modelo, durante a inclusão de novos termos no modelo.

As equações discretizadas de sistemas contínuos lineares podem ser colocadas em uma forma semelhante aos modelos polinomiais ARX. O enfoque deste capítulo é incluir nestas equações discretizadas certas equações que representam a variação do ganho e da constante de tempo, notando-se que o modelo discretizado tomará a forma NARX. A partir do modelo discretizado NARX, analisar-se-á o aparecimento de novos termos nos modelos, bem como tentar-se-á agrupar as características de variação do ganho e constante de tempo de modo que estas possam ser representadas a partir do modelo. Os resultados deste capítulo servirão de base comparativa para modelos polinomiais NARX genéricos do próximo capítulo.

Neste capítulo será feita uma análise de modelos discretizados de equações dinâmicas genéricas no domínio de Laplace, para sistemas até terceira ordem, e estará organizado como se segue. Na seção 4.2 será apresentado o tipo de relação genérica de variação do ganho e de uma constante de tempo do sistema com algum sinal. Na seção 4.3 são feitas as discretizações das equações dinâmicas de sistemas até terceira ordem, com ganho e constante de tempo constantes. Na seção 4.4 é feita uma comparação quanto à representação entre modelos discretizados e modelos polinomiais NARMAX, quando os parâmetros em questão são variáveis. Na seção 4.5 são incluídos aos modelos discretizados da seção 4.3 as variações de ganho e constante de tempo da seção 4.2, onde será verificado o aparecimento de novos termos nos modelos discretizados. Finalmente, na seção 4.6 faz-se uma generalização dos modelos discretizados em termos de agrupamentos de termos e coeficientes de agrupamentos. Na seção 4.7 são feitos os comentários finais do capítulo.

4.2 Sistemas Lineares com Parâmetros variáveis

Um sistema ‘linear’ com parâmetros variáveis, significa que qualquer parâmetro de uma equação diferencial linear depende de um sinal ou variável de estado do processo, por exemplo:

- ganho;
- constante de tempo;
- coeficiente de amortecimento;
- zero;
- pólo.

Estes parâmetros podem depender de:

- um sinal de entrada;
- um sinal de saída;
- um sinal externo e/ou
- qualquer sinal calculado baseado em sinais medidos.

Este sinal por sua vez pode ser:

- variante no tempo;
- variável com uma função de um outro sinal;

e assim por diante. A vantagem desta descrição é que pode-se entender mais facilmente o fato de um parâmetro de um sistema linear não ser constante, mas o sistema ainda pode ser representado por uma equação diferencial não-linear, com parâmetros constantes. Por outro lado, algoritmos de controle lineares podem ser facilmente aplicados a estes modelos (Haber and Keviczky, 1985).

Neste trabalho, vai-se tratar apenas com a variação do ganho e de uma constante de tempo do sistema. Isto porque em alguns sistemas o ganho ou a constante de tempo dominante possuem uma relação de dependência com uma variável ou uma função de variáveis do processo, como dito anteriormente. Esta relação de dependência, em um sistema dinâmico, são responsáveis pelo aparecimento de termos específicos no modelo, que se encontram relacionados com o sinal de entrada ou de saída considerados no processo (Haber and Keviczky, 1985).

Um sistema dinâmico ‘linear’ com parâmetros variáveis pode ser basicamente descrito por duas equações:

- uma equação diferencial não-linear do sistema dinâmico com os parâmetros constantes, e
- uma equação estática que descreve como estes parâmetros dependem de um sinal medido ou calculado.

Seja um sistema dinâmico com constante de tempo τ e K_{re} variável de acordo com a equação abaixo:

$$K_{re}(x_k) = K_0 + K_1 x_k \quad (4.1)$$

onde $K_{re}(x_k)$ é a relação estática variável, sendo K_0 e K_1 constantes e x_k é um sinal, ou uma função de sinais dos quais $K_{re}(x_k)$ é dependente.

Suponha também que ao invés de K_{re} variável, o sistema apresente uma das suas constantes de tempo variáveis de acordo com a equação abaixo:

$$\tau(x_\tau) = \tau_0 + \tau_1 x_\tau \quad (4.2)$$

onde $\tau(x_\tau)$ é a constante de tempo variável, sendo τ_0 e τ_1 constantes, e x_τ é o sinal ou uma função de sinais dos quais $\tau(x_\tau)$ depende.

Cabe ressaltar que x_k e x_τ não precisam ser necessariamente iguais.

4.3 Ganho e Constante de Tempo em Modelos discretizados

Nesta seção será feita a discretização, pelo método de Euler, de modelos dinâmicos genéricos até terceira ordem no domínio de Laplace, considerando-se primeiramente seus parâmetros (ganho e constantes de tempo) constantes. O objetivo desta seção é verificar a relação de tais parâmetros com os sinais de entrada e saída através do modelo discretizado.

Seja a seguinte função de transferência de ganho "K" e ordem "n" no domínio de Laplace, com pólos complexos ou reais:

$$\frac{Y(s)}{U(s)} = \frac{K}{\prod_{i=1}^n (\tau_i s + 1)}$$

A função de transferência anterior pode ser escrita da seguinte forma:

$$Y(s) \prod_{i=1}^n (\tau_i s + 1) = KU(s)$$

Da equação anterior, percebe-se que, as constantes de tempo do sistema podem ser relacionadas diretamente com a saída, enquanto que o ganho com a entrada. Pode-se discretizar uma equação no domínio de Laplace, por exemplo, utilizando-se a transformação de Euler abaixo (Haber and Keviczky, 1985):

$$s = \frac{z - 1}{T_s} = \frac{1 - z^{-1}}{z^{-1}T_s}$$

onde T_s , neste caso, é o tempo de discretização.

Então, seja um sistema de primeira ordem genérico no domínio de Laplace:

$$\frac{Y(s)}{U(s)} = \frac{K}{(\tau s + 1)},$$

Aplicando-se a transformação de Euler:

$$y(z) \left(\tau \left(\frac{1 - z^{-1}}{z^{-1} T_s} \right) + 1 \right) = K u(z)$$

Desenvolvendo-se a equação anterior, chega-se à seguinte expressão discretizada:

$$\tau(y(k) - y(k-1)) + T_s y(k-1) = K T_s u(k-1) \quad (4.3)$$

Pode-se discretizar uma equação em tempo contínuo, por diversos métodos, tais como, transformação bilinear e transformação de Euler, como na equação (4.3) (Phillips and Nagle, 1995). Para facilitar o entendimento e abreviar o desenvolvimento, adotar-se-á a transformação de Euler. Resultados semelhantes podem ser obtidos através da transformação bilinear. Logo, as equações discretizadas para modelos de primeira a terceira ordem são:

Para um sistema de primeira ordem, de acordo com a equação (4.3):

$$\frac{\tau}{T_s} [y(k) - y(k-1)] + \tilde{y}(k) = K \tilde{u}(k) \quad (4.4)$$

$$\text{onde } \begin{cases} \tilde{y}(k) = y(k-n) \\ \tilde{u}(k) = u(k-w), \end{cases}$$

sendo “n” a ordem do modelo, e “w” o máximo atraso presente no sinal de entrada. Pelo fato do tipo de discretização utilizada ser o método de Euler, tem-se que “n” e “w” coincidem, podendo então ser suprimidos.

Para um sistema de segunda ordem, partindo-se da seguinte expressão no domínio de Laplace:

$$\frac{Y(s)}{U(s)} = \frac{K}{(\tau_1 s + 1)(\tau_2 s + 1)}$$

e aplicando-se o mesmo desenvolvimento feito para sistemas de primeira ordem, chega-se a:

$$\begin{aligned} \frac{\tau_1 \tau_2}{T_s^2} [y(k) - 2y(k-1) + y(k-2)] + \frac{(\tau_1 + \tau_2)}{T_s} [y(k) - y(k-1)] + \\ + y(k-2) = Ku(k-2) \end{aligned} \quad (4.5)$$

Da mesma maneira, para um sistema de terceira ordem, partindo-se de:

$$\frac{Y(s)}{U(s)} = \frac{K}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1)}$$

chega-se à seguinte expressão:

$$\begin{aligned} \frac{\tau_1 \tau_2 \tau_3}{T_s^3} [y(k) - 3y(k-1) + 3y(k-2) - y(k-3)] + \frac{(\tau_1 \tau_2 + \tau_3 \tau_2 + \tau_1 \tau_3)}{T_s^2} [y(k-1) - 2y(k-2) \cdot \\ + \frac{(\tau_1 + \tau_2 + \tau_3)}{T_s} [y(k-2) - y(k-3)] + y(k-3) = Ku(k-3) \end{aligned} \quad (4.6)$$

Chamando:

$$\begin{cases} \Delta y(k) = [y(k) - y(k-1)] \\ \nabla y(k) = [y(k) - 2y(k-1) + y(k-2)] \\ \Delta y(k) = [y(k) - 3y(k-1) + 3y(k-2) - y(k-3)] \end{cases}$$

podem-se reescrever as equações (4.4), (4.5) e (4.6) respectivamente como:

$$\frac{\tau}{T_s} \Delta y(k) + y(k-1) = Ku(k-1) \quad (4.7)$$

$$\frac{\tau_1 \tau_2}{T_s^2} \nabla y(k) + \frac{(\tau_1 + \tau_2)}{T_s} \Delta y(k) + y(k-2) = Ku(k-2) \quad (4.8)$$

$$\begin{aligned} \frac{\tau_1 \tau_2 \tau_3}{T_s^3} \Delta y(k) + \frac{(\tau_1 \tau_2 + \tau_2 \tau_3 + \tau_1 \tau_3)}{T_s^2} \nabla y(k-1) + \frac{(\tau_1 + \tau_2 + \tau_3)}{T_s} \Delta y(k-2) + \\ + y(k-3) = Ku(k-3) \end{aligned} \quad (4.9)$$

Considerando $\Delta y(k) = y(k) - y(k-1) \approx \Delta y(k-1) = y(k-1) - y(k-2) \approx \Delta y(k-2) = y(k-2) - y(k-3)$, as equações (4.7), (4.8) e (4.9) podem ser manipuladas de forma a agrupar seus parâmetros τ e K .

Para a equação (4.7), isolando-se o termo $\Delta y(k)$ do lado esquerdo da equação tem-se:

$$\Delta y(k) = \frac{T_s}{\tau} [Ku(k-1) - y(k-1)] \quad (4.10)$$

Para a equação (4.8), tendo-se em mente que:

$$\nabla y(k) = \frac{1}{T_s^2} [y(k) - 2y(k-1) + y(k-2)] = \frac{1}{T_s^2} [\Delta y(k) - \Delta y(k-1)] \approx \frac{1}{T_s^2} [\Delta y(k) - \Delta y(k)] = 0$$

e da mesma maneira que a equação (4.10) chega-se em:

$$\Delta y(k) = \frac{T_s}{(\tau_1 + \tau_2) + T_s} [Ku(k-2) - y(k-1)] \quad (4.11a)$$

ou

$$\Delta y(k) = \frac{T_s}{(\tau_1 + \tau_2)} [Ku(k-2) - y(k-2)] \quad (4.11b)$$

dependendo da ordem do termo em y que se deseja evidenciar.

Da mesma maneira para a equação (4.9):

$$\Delta y(k) = \frac{T_s}{(\tau_1 + \tau_2 + \tau_3) + 2T_s} [Ku(k-3) - y(k-1)] \quad (4.12a)$$

ou

$$\Delta y(k) = \frac{T_s}{(\tau_1 + \tau_2 + \tau_3)} [Ku(k-3) - y(k-3)] \quad (4.12.b)$$

dependendo também da ordem do termo em y que se deseja explicitar.

As das equações (4.10) a (4.12), mostram como se pode isolar as constantes de tempo e o ganho partindo-se da equação discretizada onde tais parâmetros são constantes. No capítulo 5, proceder-se-á de maneira análoga, porém utilizando-se modelos NARX genéricos, tentando-se agrupar em um ponto, da mesma maneira que os modelos discretizados, tais parâmetros que porventura estejam contidos nestes modelos.

4.4 Representação de Sistemas com Parâmetros Variáveis

Nesta seção será discutido o problema da presença de parâmetros variáveis na representação NARMAX escolhida para o desenvolvimento deste trabalho, com base nos modelos discretizados anteriormente.

Um dos problemas que se deve ter em mente, é que o modelo polinomial obtido, é do tipo regressivo, ou seja, só há termos com atrasos maiores ou iguais a 1.

Para que as expressões discretizadas (eqs 4.4, 4.5 e 4.6) de equações contínuas tomem a forma de um modelo polinomial NARX, deve-se colocá-los na forma preditiva, ou seja, com o termo em $y(k)$ separado do lado esquerdo da igualdade, e todos os demais termos do lado direito, resultando nos seguintes modelos discretizados :

Para sistemas de primeira ordem:

$$y(k) = \frac{\tau y(k-1) - T_s y(k-1) + Ku(k-1)}{\tau} \quad (4.13)$$

Para sistemas de segunda ordem:

$$y(k) = \frac{T_s^2}{\tau_1 \tau_2 + T_s(\tau_1 + \tau_2)} \left(\begin{aligned} &\frac{\tau_1 \tau_2}{T_s^2} [2y(k-1) - y(k-2)] + \\ &+ \frac{(\tau_1 + \tau_2)}{T_s} y(k-1) - y(k-2) + \\ &+ Ku(k-2) \end{aligned} \right) \quad (4.14)$$

Para sistemas de terceira ordem :

$$y(k) = \frac{T_s^3}{\tau_1 \tau_2 \tau_3} \left(\begin{aligned} & \frac{\tau_1 \tau_2 \tau_3}{T_s^3} [-3y(k-1) + 3y(k-2) - y(k-3)] + \\ & + \frac{(\tau_1 \tau_2 + \tau_3 \tau_2 + \tau_1 \tau_3)}{T_s^2} [y(k-1) - 2y(k-2) + y(k-3)] + \\ & + \frac{(\tau_1 + \tau_2 + \tau_3)}{T_s} [y(k-2) - y(k-3)] + y(k-3) + Ku(k-3) \end{aligned} \right) \quad (4.15)$$

Observando as equações anteriores nota-se a presença de τ no denominador da equação (4.13), τ_1 e τ_2 no denominador da equação (4.14), e τ_1 , τ_2 e τ_3 no denominador da equação (4.15). A existência de parâmetros τ_1 , τ_2 e τ_3 no denominador das equações provêm da colocação de $y(k)$ em evidência do lado esquerdo destas, a fim de que tomem a forma preditiva de um modelo polinomial NARX. Então, desde que τ_1 , τ_2 e τ_3 sejam constantes, percebe-se que as equações (4.13), (4.14) e (4.15) mantêm a forma polinomial NARX, apresentando-se na forma preditiva.

Este resultado pode ser enunciado na forma de um lema.

Lema 4.1

Se uma das constantes de tempo de um sistema for variável, obtém-se uma representação NARX na forma preditiva, sendo do tipo **racional**.

prova:

Seja um sistema de primeira ordem como o da equação (4.13). Suponha que a constante de tempo varie de acordo com a equação (4.2).

Substituindo a equação (4.2) na equação (4.13), e mantendo-a na forma preditiva tem-se:

$$y(k) = \frac{\tau_0 y(k-1) + \tau_1 x_\tau y(k-1) - T_s y(k-1) + Ku(k-1)}{\tau_0 + \tau_1 x_\tau} \quad (4.16)$$

onde x_τ é uma função de sinais do processo como já dito anteriormente. Da equação (4.16) percebe-se que a existência de pelo menos uma constante de tempo variável não permite a obtenção de um modelo semelhante ao modelo polinomial NARX na forma preditiva, porém possibilita a obtenção de um modelo racional. Isto porque a constante de tempo que aparecia multiplicando o termo $y(k)$ passou para o lado esquerdo da equação (4.16) como um divisor, a fim de que a forma preditiva fosse mantida.

A seguir verificar-se-á quais termos devem estar presentes em um modelo discretizado quando o ganho e a constante de tempo não forem constantes. Tais resultados serão então comparados a resultados derivados para modelos polinomiais NARX genéricos no próximo capítulo.

4.5 Modelos Discretizados com Parâmetros Variáveis

Sabe-se que modelos polinomiais NARX quando sobreparametrizados normalmente apresentam termos espúrios. Por termos espúrios entende-se: termos que não deveriam fazer parte da estrutura do modelo. Como geralmente é difícil saber exatamente quais termos devem compor o modelo, foram elaboradas técnicas estatísticas que possibilitam encontrar um número bom de termos para este (vide seções 2.2.3.2 e 2.2.6) e quais termos devem constituir-lo (Billings et al., 1989). Uma revisão da aplicação de tais técnicas pode ser encontrada em Pottmann et al. (1993). Como dito no capítulo 3, a técnica de agrupamentos de termos e coeficientes de agrupamentos visa auxiliar na escolha da estrutura do modelo polinomial. Foi dito também que termos de um mesmo agrupamento que aparecem se cancelando normalmente são considerados espúrios. Com esta idéia em mente será feita uma análise das equações discretizadas da seção 4.3 sendo acrescentadas a estas as variações representadas pelas equações (4.1) e (4.2), observado-se

então o aparecimento de novos termos no modelo. Obtidas tais equações, serão feitas observações relativas aos agrupamentos de termos e coeficientes de agrupamentos presentes nestas.

4.5.1 Modelo Discretizado de primeira Ordem

Substituindo-se as equações (4.1) e (4.2) na equação (4.13) tem-se:

$$y(k) = \frac{(\tau_0 + \tau_1 x_\tau(k))y(k-1) - T_s y(k-1) + (K_0 + K_1 x_k(k))u(k-1)}{(\tau_0 + \tau_1 x_\tau(k))}$$

que pode ser reescrita da seguinte forma:

$$(\tau_0 + \tau_1 x_\tau(k))y(k) = (\tau_0 + \tau_1 x_\tau(k))y(k-1) - T_s y(k-1) + (K_0 + K_1 x_k(k))u(k-1)$$

Desenvolvendo-se a equação anterior:

$$\begin{aligned} y(k) = & -\left(\frac{T_s}{\tau_0} - 1\right)y(k-1) - \frac{\tau_1}{\tau_0} x_\tau(k)[y(k) - y(k-1)] + K_0 \frac{T_s}{\tau_0} u(k-1) \\ & + K_1 \frac{T_s}{\tau_0} x_k(k)u(k-1) \end{aligned} \quad (4.17)$$

Baseando-se na equação (4.17), pode-se afirmar que em um sistema de primeira ordem, além dos agrupamentos de termos lineares, isto é, Ω_y e Ω_u , para uma dependência da constante de tempo, o agrupamento $\Omega_{x_\tau y}$ deve estar presente. Semelhantemente, para que o modelo seja capaz de representar a dependência do ganho, o agrupamento $\Omega_{x_k u}$ deve estar presente. Note-se ainda que os coeficientes dos termos do agrupamento $\Omega_{x_\tau y}$ possuem a mesma amplitude, porém com magnitudes contrárias. Este fato, de acordo com uma análise ‘tradicional’ de agrupamento de termos sugeriria que tal agrupamento

fosse espúrio (Aguirre and Billings, 1995a). Entretanto, a equação 4.17 mostra que $\Sigma_{x_\tau y} \rightarrow 0$ nos casos em que a constante de tempo varia, independentemente do ganho. Isto pode ser notado quando considera-se o sistema em regime estático, que corresponde a aplicação do teorema do valor final em seu modelo. Para o caso discreto, o teorema do valor final pode ser calculado da seguinte maneira:

$$y(\infty) = \lim_{z \rightarrow 1} (z-1)Y(z) \Big|_{\frac{Uz}{z-1}} \quad \text{se } y(\infty) \text{ existe.} \quad (4.18)$$

O modelo da equação (4.17) pode ser reescrito na forma:

$$yz = -\left(\frac{T_s}{\tau_0} - 1\right)yz^{-1} - \frac{\tau_1}{\tau_0} x_\tau z [yz - yz^{-1}] + K_0 \frac{T_s}{\tau_0} uz^{-1} + K_1 \frac{T_s}{\tau_0} x_k z uz^{-1}$$

Colocando-se y e u em evidência nos dois lados da equação:

$$y \left(z + \left(\frac{T_s}{\tau_0} - 1 \right) z^{-1} + \frac{\tau_1}{\tau_0} x_\tau z [z - z^{-1}] \right) = + \left(K_0 \frac{T_s}{\tau_0} z^{-1} + K_1 \frac{T_s}{\tau_0} x_k z z^{-1} \right) u$$

Explicitando-se o y do lado esquerdo da equação, tem-se:

$$y = \frac{\left(K_0 \frac{T_s}{\tau_0} z^{-1} + K_1 \frac{T_s}{\tau_0} x_k z z^{-1} \right)}{\left(z + \left(\frac{T_s}{\tau_0} - 1 \right) z^{-1} + \frac{\tau_1}{\tau_0} x_\tau z [z - z^{-1}] \right)} u \quad (4.19)$$

Logo, aplicando-se o teorema do valor final (eq. 4.18) à equação 4.19 tem-se:

$$Y = \lim_{z \rightarrow 1} (z-1) \left[\frac{\left(K_0 \frac{T_s}{\tau_0} z^{-1} + K_1 \frac{T_s}{\tau_0} X_k z z^{-1} \right)}{\left(z + \left(\frac{T_s}{\tau_0} - 1 \right) z^{-1} + \frac{\tau_1}{\tau_0} X_\tau z [z - z^{-1}] \right)} \right] U \frac{z}{(z-1)} \quad (4.20a)$$

Onde $U \frac{z}{(z-1)}$ é a transformada z da entrada degrau, de amplitude U, a partir da qual é definido o teorema do valor final.

O resultado da aplicação do teorema do valor final na equação (4.20a) é:

$$Y = (K_0 + K_1 X_k)U$$

Que pode ser reescrita da seguinte forma:

$$\frac{Y}{U} = K_0 + K_1 X_k \quad (4.20b)$$

A equação (4.20b) corresponde à relação estática saída/entrada do sistema. Logo, através da aplicação do teorema do valor final, pode-se calcular tal relação.

Pela equação (4.20a) percebe-se que $\Sigma_{x,y} = 0$, a partir do cancelamento dos termos em negrito, no denominador desta equação. Portanto, em regime estático, tem-se que $\Sigma_{x,y} = 0$, porém este agrupamento não é espúrio, uma vez que representa a parcela variável da constante de tempo. Uma consequência de tal cancelamento pode ser o não influenciamento por parte deste agrupamento, na recuperação da equação para a relação estática, como será visto mais adiante.

A mesma observação é feita para modelos de segunda e terceira ordem, e pode ser visto no apêndice A.

Exemplo 4.1

considere um sistema cuja constante de tempo varia com a saída da seguinte maneira (Haber and Unbehauen, 1990):

$$\tau(t) = 10 - 2,5y(t)$$

Neste caso tem-se $\tau_0 = 10$, $\tau_1 = -2,5$ e $x_\tau(t) = y(t)$.

O modelo obtido para representar tal sistema foi (Haber and Unbehauen, 1990):

$$y(k) = 0,82y(k-1) + 0,356u(k) + 0,00401u(k-1) + 0,1152y^2(k) - 0,1153y^2(k-1)$$

Pelo modelo anterior pode perceber-se que os coeficientes dos termos do agrupamento Ω_{y^2} são aproximadamente iguais e com sinais contrários. Se tais termos fossem considerados espúrios e retirados do modelo, o modelo se tornaria linear e a equação variável da constante de tempo não poderia ser recuperada (Haber and Unbehauen, 1990). O fato do agrupamento Ω_{y^2} não ser espúrio é confirmado pelo resultado anterior.

Na próxima seção, as equações discretizadas de sistemas de primeira à terceira ordem, com parâmetros constantes serão generalizadas em termos de agrupamentos de termos e coeficientes de agrupamento.

4.6 Generalização dos Modelos Discretizados na Forma de Agrupamentos de Termos e Coeficientes de Agrupamentos

Pode-se escrever as equações (4.17), (A.2) e (A.3) em regime estático, em termos de coeficientes de agrupamentos e agrupamentos de termos das equações discretizadas para sistemas de primeira a terceira ordem da seguinte maneira:

$$y = \frac{1}{\Sigma_y} \left[(\Sigma_{x_\tau y}) x_\tau y + (\Sigma_{x_k u}) x_k u + (\Sigma_u) u \right] \quad (4.21)$$

onde para um sistema de primeira ordem:

$$\Sigma_y = 1 + \frac{T_s}{\tau_0} - 1;$$

$$\Sigma_{x_\tau y} = -\frac{\tau_1}{\tau_0} + \frac{\tau_1}{\tau_0};$$

$$\Sigma_u = K_0 \frac{T_s}{\tau_0};$$

$$\Sigma_{x_k u} = K_1 \frac{T_s}{\tau_0}.$$

Para sistemas de segunda e terceira ordem, tal generalização é mostrada no apêndice B.

A relação estática de sistemas como aqueles das equações (4.17), (A.2) e (A.3) pode ser calculada através de agrupamentos de termos e coeficientes de agrupamentos, ou aplicar pelo teorema do valor final. Nestas condições, e baseando-se nos valores dos coeficientes Σ_y , $\Sigma_{x_\tau y}$, Σ_u e $\Sigma_{x_k u}$, e substituindo-os na equação (4.21) tem-se:

$$Y = \frac{1}{\Sigma_y} [(\Sigma_{x_\tau y})X_\tau Y + (\Sigma_{x_k u})X_k U + (\Sigma_u)U] = K_0 U + K_1 X_k U = (K_0 + K_1 X_k)U$$

cujas relação estática pode ser calculada da seguinte maneira:

$$K_{re} = \frac{Y}{U} = K_0 + K_1 X_k \quad (4.22)$$

que nada mais é do que a relação estabelecida pela equação (4.1).

Como visto na seção 4.5.1, percebe-se que em regime estático, o coeficiente do agrupamento $\Omega_{x_{\tau}y}$, isto é, $\Sigma_{x_{\tau}y}$, tende a zero. Uma das consequências deste fato é que tais termos não influenciarão na relação estática, como pode ser visto na equação (4.22).

Baseado nesta afirmativa pode-se concluir que os agrupamentos originados por variações da constante de tempo tendem a desaparecer em regime estático, o que não necessariamente indica que sejam espúrios.

Com base na introdução de termos nos modelos de primeira à terceira ordem da seção 4.5, provenientes de dependências do ganho e da constante de tempo, pode-se elaborar um lema que define a forma de combinação de tais termos nos modelos discretizados.

Lema 4.2

Seja um sistema de ordem n , cujo ganho e constantes de tempo variam de acordo com as equações (4.1) e (4.2) respectivamente. Supondo ainda que tal sistema tenha α constantes de tempo variáveis, $\{\tau_1, \dots, \tau_\alpha\}$, e $(n - \alpha)$ constantes de tempo fixas, além dos agrupamentos lineares, aparecerão os seguintes termos no modelo:

- termos relativos à variação das constantes de tempo:

$$\Omega_{(C_{x_{\tau_i}, m})y} \quad \therefore m, i = 1 \dots \alpha.$$

onde $C_{x_{\tau_i}, m}$ é a combinação das variáveis x_{τ_i} m a m , e α é o número de constantes de tempo variáveis.

- termos relativos à variação do ganho:

$$\Omega_{x_k u}.$$

Apesar de tediosa, a prova deste lema é simples. A seguir mostra-se o caso para um sistema de segunda ordem. A extensão para outros sistemas é direta.

Prova: Seja um modelo de segunda ordem como o da equação (A.2), com o ganho e uma constante de tempo variáveis (neste caso $\alpha = 1$). Os agrupamentos de termos presentes neste caso são: Ω_y , Ω_u , $\Omega_{x_k u}$ e $\Omega_{x_{\tau_1} y}$. Suponha que a constante de tempo τ_2 seja também variável de acordo com a seguinte equação $\tau_2 = \tau_{c2} + \tau_{v2} x_{\tau_2}$. Substituindo esta equação na equação (A.2) tem-se:

$$\begin{aligned} \left(\frac{\tau_{c1}\tau_{c2} + \tau_{c1}\tau_{v2}x_{\tau_2}}{T_s^2} + \frac{\tau_{c1} + \tau_{c2} + \tau_{v2}x_{\tau_2}}{T_s} \right) y(k) = & \left(\frac{\tau_{c1}\tau_{c2} + \tau_{c1}\tau_{v2}x_{\tau_2}}{T_s^2} + \frac{\tau_{c1} + \tau_{c2} + \tau_{v2}x_{\tau_2}}{T_s} \right) y(k-1) - \\ & - \frac{(\tau_{v1}\tau_{c2} + \tau_{v1}\tau_{v2}x_{\tau_2})}{T_s^2} x_{\tau_1} [y(k) - y(k-1)] - \frac{\tau_{v1}}{T_s} x_{\tau_1} [y(k) - y(k-1)] + \\ & + \frac{(\tau_{v1}\tau_{c2} + \tau_{v1}\tau_{v2}x_{\tau_2})}{T_s^2} x_{\tau_1} [y(k-1) - y(k-2)] + \frac{(\tau_{c1}\tau_{c2} + \tau_{c1}\tau_{v2}x_{\tau_2})}{T_s^2} [y(k-1) - y(k-2)] - \\ & - y(k-2) + K_0 u(k-2) + K_1 x_K u(k-2) \end{aligned} \quad (4.23)$$

Da equação (4.23), além dos agrupamentos existentes na equação (A.2), são introduzidos também os agrupamentos $\Omega_{x_{\tau_2} y}$ e $\Omega_{x_{\tau_1} x_{\tau_2} y}$. Assim, para a equação (4.23), se $x_{\tau_1} \neq x_{\tau_2}$, os seguintes agrupamentos de termos aparecerão no modelo além dos agrupamentos lineares: $\Omega_{x_k u}$, $\Omega_{x_{\tau_1} y}$, $\Omega_{x_{\tau_2} y}$ e $\Omega_{x_{\tau_1} x_{\tau_2} y}$, o que está de acordo com o lema.

Cabe ressaltar que se existirem mais de uma constante de tempo variáveis e iguais, elas aparecerão em um mesmo agrupamento de termos. Isto é, se $x_{\tau_1} \dots x_{\tau_n} = x_\tau$, os agrupamentos de termos presentes no modelo serão:

$$\Omega_{x_\tau^q y}, \text{ onde } q = 1 \dots \alpha.$$

Exemplo 4.2

Para a equação (4.23), se $x_{\tau_1} = x_{\tau_2} = x_\tau$, os seguintes agrupamentos aparecerão no modelo, além dos agrupamentos lineares:

$$\Omega_{x_k u}, \Omega_{x_\tau y} \text{ e } \Omega_{x_\tau^2 y}.$$

Existem termos no modelo polinomial NARX que podem pertencer tanto ao agrupamento $\Omega_{x_\tau y}$ quanto ao agrupamento $\Omega_{x_k u}$. Tal fato é mostrado a seguir.

Exemplo 4.3:

O termo $u(k-1)y(k-1)$ pode fazer parte tanto do agrupamento de termos $\Omega_{x_\tau y}$ (se $x_\tau = u$), quanto do agrupamento de termos $\Omega_{x_k u}$ (se $x_\tau = y$). O agrupamento $\Omega_{x_\tau y}$ será induzido por uma constante de tempo variável com y , ao passo que $\Omega_{x_k u}$ será induzido por um ganho variável com u .

Exemplo 4.4:

Um termo do tipo $u^3(k-1)$, por sua vez, só poderá pertencer ao agrupamento $\Omega_{x_k u}$ (com $x_k = u^2$) e não, por exemplo, ao agrupamento $\Omega_{x_\tau y}$, pelo fato de não possuir o termo y .

Com base nos exemplos anteriores, pode-se formular as seguintes definições:

Definição 4.1

Termos de dependência transitivos são termos do modelo pertencentes a agrupamentos que podem estar relacionados tanto à uma variação do ganho quanto à uma variação da constante de tempo.

Definição 4.2

Termos de dependência não-transitivos são termos do modelo pertencentes unicamente a um tipo de agrupamento de termos.

Portanto, baseado nas observações feitas para os modelos discretizados, sabe-se que se o termo transitivo fizer parte do agrupamento de termos $\Omega_{x_t y}$, quando for considerado o modelo em regime estático, este agrupamento de termos tenderá a zero¹, de forma a não se relacionar com o ganho, e consequentemente, não influir na equação da relação estática, como visto anteriormente, ao contrário do agrupamento de termos $\Omega_{x_k u}$.

A presença de termos transitivos em modelos polinomiais NARMAX será vista posteriormente nos capítulos 6 e 7.

4.7 Comentários Finais

¹ Cabe ressaltar ainda que podem haver casos em que o ganho do sistema apresente uma relação interativa com a constante de tempo, quando, por exemplo, ambos dependerem de um mesmo sinal. Nesta situação, ao considerar o sistema em regime estático, o agrupamento $\Omega_{x_t y}$ provavelmente não irá se cancelar, mostrando assim que o ganho e a constante de tempo se relacionam diretamente, através de suas dependências.

Neste capítulo foram analisados os modelos discretizados de sistemas de primeira a terceira ordem, com parâmetros, ganho e constante de tempo, constantes e variáveis. A análise de modelos discretizados servirá de subsídio para as considerações que serão feitas no próximo capítulo.

Notou-se que à medida que foram incluídas nos modelos expressões representativas das variações do ganho e da constante de tempo, novos termos são incluídos no modelo conforme enunciado no lema 2.

Na seção 4.6, foi feita uma análise de agrupamento de termos para os modelos discretizados e verificou-se que os termos do agrupamento $\Omega_{x_t y}$ se anulam. Tal efeito pode não ser devido ao fato deste agrupamento ser realmente espúrio. Foi dito que os coeficientes dos termos deste agrupamento foram distribuídos de forma que ao analisar os modelos em regime estático, o que corresponde a aplicação do teorema do valor final, tais coeficientes se anulam, indicando que a constante de tempo não se relaciona com o ganho, uma consequência deste cancelamento é a não interferência na relação estática. Porém, os coeficientes do agrupamento $\Omega_{x_t y}$ contêm as informações referentes a dependência da constante de tempo, não podendo ser assim descartados do conjunto de termos candidatos ao modelo. Convém dizer que em modelos reais tal cancelamento não é exato, entretanto pode vir a ser evidente através da proximidade dos valores e pela existência de sinais opostos.

Uma observação importante foi feita no fim da seção 4.6 com relação à variação de K e τ com um mesmo sinal. Sendo ressaltado que em certos casos pode haver uma interação entre tais parâmetros, nesse caso não ocorre o cancelamento do agrupamento $\Omega_{x_t y}$ em regime estático.

5 Análise de Modelos Polinomiais NARMAX com Constante de Tempo e Ganho Variáveis

5.1 Introdução

No capítulo anterior, foram feitas observações a respeito da introdução de termos em modelos discretizados provenientes de variações do ganho e da constante de tempo. Tais modelos, pela semelhança com a forma polinomial NARX servirão de base para o desenvolvimento que será feito neste capítulo, onde parte-se do modelo na forma preditiva. Deseja-se investigar como descrever as variações da constante de tempo e do ganho a partir do modelo.

Os modelos serão comparados com as formas discretizadas da seção 4.3, na tentativa de recuperar as informações a respeito da característica estática e da constante de tempo, agrupando-as de uma maneira que possam ser explicitadas. Será observado como os termos decorrentes da variação do ganho e da constante de tempo estão relacionados nos modelos NARX.

O capítulo está organizado como se segue. Na seção 5.2 será proposta uma maneira de recuperar equações para a relação estática e constante de tempo variáveis a partir de modelos polinomiais NARX genéricos até terceira ordem. Na seção 5.3 serão estabelecidas relações entre a ordem, o grau de não-linearidade e o aparecimento de termos decorrentes de variações do ganho ou constante de tempo. A seção 5.4 apresenta um comentário geral a respeito do capítulo.

5.2 Constante de Tempo e Ganho Variáveis em modelos Polinomiais NARX

Nesta seção será desenvolvida uma formulação para a representação da relação estática e constante de tempo variáveis a partir modelos NARX, baseando-se em agrupamentos de termos e coeficientes de agrupamento.

Serão discutidos os principais problemas na representação em um modelo da variação do ganho e da constante de tempo. Nesta seção, o problema em questão é: tendo-se um modelo NARX em mãos, como é possível retirar informações a respeito da característica estática e das constantes de tempo?

5.2.1 Obtenção da Rrelação Estática e Constante de Tempo Variáveis para um Sistema de Primeira Ordem

Suponha um modelo polinomial NARX de primeira ordem na forma:

$$y(k) = Ay(k-1) + Bu(k-1) + Ex_{\tau}(k-1)y(k-1) + Fx_k(k)u(k-1) \quad (5.1)$$

Note-se que o modelo (5.1) é bastante genérico uma vez que os termos $x_{\tau}(k-1)y(k-1)$ e $x_k(k)u(k-1)$ podem ser quadráticos, cúbicos, etc.

A variável $x_{\tau}(k-1)$, representa termos somente com atraso de 1. A variável $x_k(k)$, sem atraso, significa que engloba termos com qualquer valor de atraso. A limitação no caso da variável x_{τ} é necessária devido ao grande número de opções que seriam possíveis se esta restrição não fosse feita.

Uma maneira de se conseguir explicitar a relação estática e a constante de tempo deste sistema, é colocar-se a equação (5.1) na forma da equação (4.10). Porém, a equação (4.10) foi derivada a partir de um modelo discretizado, que apresenta certas diferenças do modelo da equação (5.1). Diferenças entre modelos discretizados obtidas por discretização e identificação de um mesmo sistema contínuo são comuns (Aguirre and Billings, 1994b).

Ao contrário do modelo discretizado do capítulo anterior (eq. 4.17), o modelo da equação (5.1) não apresenta dois termos do agrupamento $\Omega_{x_{\tau}y}$. Com a ausência de outro

termo deste agrupamento, ao analisar-se o modelo da equação (5.1) em regime estático como citado na seção 4.5, tem-se que $\Sigma_{x_\tau y} \neq 0$. A menos que $\Sigma_{x_\tau y} \approx 0$, não será possível explicitar-se matematicamente uma equação para a relação estática, e nem para a constante de tempo, como será visto mais adiante.

Para que $\Sigma_{x_\tau y}$ seja muito pequeno, deve haver um outro termo deste agrupamento para cancelar com o primeiro. Em um modelo polinomial NARX, isto só seria possível se o algoritmo aumentasse a ordem do modelo para acrescentar um outro termo, como forma de suprir tal carência (Aguirre and Billings, 1995b). Outra alternativa ocorre quando a constante de tempo não é variável, resultando em $\Sigma_{x_\tau y} = 0$.

A seguir, serão derivadas equações representativas da relação estática e da variação constante de tempo para o modelo polinomial NARX da equação (5.1). Para isto, proceder-se-á com os modelos polinomiais NARX a fim de se encontrar uma equação de forma semelhante às equações da seção 4.3 do capítulo anterior. Então, por meio da comparação das equações obter-se-ão as representações do ganho e constante de tempo variáveis em modelos polinomiais NARX.

5.2.1.1 Obtenção da Constante de Tempo Variável

Pode-se manipular equação (5.1) de forma que fique com um aspecto semelhante à equação (4.10). Feito isto, pode-se encontrar uma equação de representação para a relação estática e constante de tempo variáveis, análogas às equações (4.1) e (4.2) respectivamente. Então:

$$y(k) = Ay(k-1) + Bu(k-1) + E_{x_\tau}(k-1)y(k-1) + F_{x_k}(k)u(k-1)$$

$$y(k) - Ay(k-1) = Bu(k-1) + E_{x_\tau}(k-1)y(k-1) + F_{x_k}(k)u(k-1)$$

Pode-se somar $+y(k-1)$ e $-y(k-1)$, do lado esquerdo do sinal $=$, a fim de obter-se o $\Delta y(k) = y(k) - y(k-1)$. O mesmo procedimento é feito com relação ao termo $E_{x_\tau}(k-1)y(k-1)$.

Assim:

$$y(k) - \mathbf{y(k-1)} + \mathbf{y(k-1)} - Ay(k-1) = Bu(k-1) + E_{x_\tau}(k-1)y(k) - E_{x_\tau}(k-1)y(k) + E_{x_\tau}(k)y(k-1) + F_{x_k}(k)u(k-1)$$

A seguir, estão indicados em negrito, os termos que se unirão na forma de $\Delta y(k)$:

$$\mathbf{y(k)} - \mathbf{y(k-1)} + y(k-1) - Ay(k-1) = Bu(k-1) + E_{x_\tau}(k)y(k) - \mathbf{E_{x_\tau}(k-1)y(k)} + \mathbf{+E_{x_\tau}(k-1)y(k-1)} + F_{x_k}(k)u(k-1)$$

Daí:

$$\Delta \mathbf{y(k)} + (1 - A)y(k-1) = Bu(k-1) - \mathbf{E_{x_\tau}(k-1)\Delta y(k)} + E_{x_\tau}(k-1)y(k) + F_{x_k}(k)u(k-1)$$

Explicitando-se então o $\Delta y(k)$ do lado esquerdo da equação tem-se:

$$[1 + E_{x_\tau}(k-1)]\Delta y(k) = - (1 - A)y(k-1) + Bu(k-1) + E_{x_\tau}(k-1)y(k) + F_{x_k}(k)u(k-1)$$

Que pode ser colocada em uma forma semelhante á equação (4.10):

$$\Delta y(k) = \frac{(1-A)}{[1 + E_{x_\tau}(k-1)]} \left[\frac{Bu(k-1) + E_{x_\tau}(k-1)y(k) + F_{x_k}(k)u(k-1)}{(1-A)} - y(k-1) \right] \quad (5.2)$$

Tomando-se o coeficiente em evidência na equação (5.2), isto é, $(1-A)/[1-E_{x_\tau}(k-1)]$, e comparando-o ao mesmo da equação (4.10), tem-se por analogia (Haber and Keviczky, 1985):

$$\frac{T_s}{\tau} = \frac{(1 - A)}{(1 + Ex_\tau(k - 1))}$$

que em regime estático, onde $u(k) = u(k-1) = U$, $y(k) = y(k-1) = Y$ e $x_\tau(k) = X_\tau$ (ver definição 1.7), pode ser escrita na forma:

$$\frac{T_s}{\tau} = \frac{(1 - A)}{(1 + EX_\tau)}$$

Logo:

$$\frac{\tau}{T_s} = \frac{1}{(1 - A)} + \frac{E}{(1 - A)} X_\tau$$

que pode ser reescrita em função dos coeficientes de agrupamento de termos como se segue:

$$\frac{\tau}{T_s} = \frac{1}{(1 - \Sigma_y)} + \frac{E}{(1 - \Sigma_y)} X_\tau \quad (5.3)$$

onde, neste caso $\Sigma_y = A$. A equação (5.3) está na forma da equação (4.2),

com $\tau_0 = \frac{T_s}{1 - \Sigma_y}$ e $\tau_1 = \frac{E}{1 - \Sigma_y} X_\tau$.

5.2.1.2 Obtenção da Relação Estática Variável

Como visto na seção 4.6, para calcular a equação para K_{re} , considera-se o sistema em regime estático, aplicando-se, por exemplo o teorema do valor final. Logo, considerando-se a equação (5.2) em regime estático tem-se que :

$$\Delta y(k) = (y - y) = 0 = \frac{(1 - A)}{(1 + EX_\tau)} \left[\frac{BU + FX_k U + EX_\tau Y}{(1 - A)} - Y \right]$$

Donde:

$$(1 - A)Y - EX_\tau Y = BU + FX_k U$$

Portanto:

$$K_{re} = \frac{Y}{U} = \frac{B}{(1 - \Sigma_y) - EX_\tau} + \frac{F}{(1 - \Sigma_y) - EX_\tau} X_k \quad (5.4)$$

que é a equação da relação estática saída/entrada, representada de maneira análoga a equação (4.1).

Note que para um sistema com um integrador (pólo em $z = 1$), $\Sigma_y = 1$ (Aguirre and Billings, 1995b), a constante de tempo da equação (5.3) tende para o infinito, conforme esperado.

Caso o modelo não apresente termos do tipo $x_\tau y$ ou $x_k u$, ele será linear e τ , K_{re} e consequentemente o ganho podem ser calculados diretamente.

As generalizações para sistemas de segunda e terceira ordem são feitas no apêndice C.

Logo, baseando-se nos resultados desta seção e do apêndice C, as generalizações para modelos polinomiais NARX de primeira, segunda e terceira ordem podem ser resumidas em uma fórmula geral:

- Para a constante de tempo:

$$\frac{\sum_{i=1}^n \tau_i}{T_s} = \frac{(n - (n-1)_+ A - (n-2)_+ B)}{(1 - \Sigma_y)} + \frac{(E + (2)_* F + (3)_* K)}{(1 - \Sigma_y)} X_\tau \quad (5.5)$$

onde n é a ordem do sistema, e:

$$(x)_+ = \begin{cases} x & \forall \quad x \geq 0 \\ 0 & \forall \quad x < 0 \end{cases}$$

$$(y)_* = \begin{cases} y & \forall \quad y \geq n \\ 0 & \forall \quad y < n \end{cases}$$

- Para a relação estática:

$$K_{re} = \frac{Y}{U} = \frac{(\Sigma_u)}{(1 - \Sigma_y) - (\Sigma_{x_\tau y})X_\tau} + \frac{(\Sigma_{x_K u})X_K}{(1 - \Sigma_y) - (\Sigma_{x_\tau y})X_\tau} \quad (5.6)$$

Note-se que na equação (5.5) os parâmetros A , B , E , F e K , correspondentes aos coeficientes dos termos do modelo NARMAX, mudam de termo dependendo da ordem do modelo.

Cabe ressaltar que, pode-se estender tal generalização para modelos com ordens superiores, mediante a inclusão de novos termos na equação (5.5).

Exemplo 5.1

Seja um modelo polinomial NARX de primeira ordem, obtido à partir de uma massa de dados amostrada de 0,25 segundos, representado pela seguinte equação:

$$y(k) = 0,752y(k-1) + 0,0482u(k-1) + 0,377u^2(k-1) \quad (5.7)$$

Suponha que o sistema apresente constante de tempo unitária e ganho variável de acordo com a equação abaixo:

$$K = 0,2 + 1,5 U \quad (5.8)$$

Observando-se a equação (5.7) percebe-se que o único termo não-linear presente é o termo relativo ao agrupamento Ω_{u^2} . Como dito no capítulo 4, termos relativos à variação da constante de tempo pertencem ao agrupamento $\Omega_{x_\tau y}$, e como na equação (5.7) não existe nenhum termo nesta forma, pode-se dizer que em tal modelo a constante de tempo não é variável. Esta observação confirma o sistema a partir do qual os dados foram coletados, que como dito no enunciado do exemplo, possui $\tau = 1$, logo $\Sigma_{x_\tau y} = 0$.

Portanto comparando-se a equação (5.7) com a equação (5.1), e baseando-se nas equações (5.5) e (5.6) tem-se:

$$\frac{\tau}{0,25} = \frac{1}{(1 - 0,752)} + \frac{0}{(1 - 0,752)} X_\tau \quad (5.9)$$

$$K_{re} = \frac{0,0482}{(1 - 0,752) - (0)X_\tau} + \frac{0,377U}{(1 - 0,752) - (0)X_\tau} \quad (5.10)$$

Então, da equação (5.9) tem-se:

$$\tau = 1,0081$$

E, da equação (5.10) tem-se:

$$K = K_{re} = 0,1944 + 1,52 U$$

Portanto, com base nos resultados encontrados, percebe-se que as informações a respeito do ganho e da constante de tempo puderam ser recuperadas com boa precisão.

5.3 Relações Entre Ordem e Grau de Não-Linearidade do Modelo Polinomial NARX

Nesta seção serão discutidas a influência na ordem e no grau de não-linearidade de termos variáveis de ganho ou de constante de tempo. Existe uma relação muito estreita entre a ordem do modelo, o grau de não-linearidade e a presença de termos decorrentes da variação do ganho ou da constante de tempo. Este fato será visto a seguir.

5.3.1 Ordem

O lema 4.2 mostra quais agrupamentos de termos são induzidos por α constantes de tempo variáveis. Claramente se vê que quanto maior for o número de constantes de tempo variáveis, maior será a inclusão de termos no modelo.

Obviamente, num sistema real, o aumento da ordem dinâmica do modelo acarretará um aumento no número de constantes de tempo variáveis, α .

5.3.2 Grau de Não-linearidade

Lema 5.1

Seja x_{τ_i} a parte variável da i -ésima constante de tempo, e x_k a parte variável do ganho de um modelo com α constantes de tempo variáveis. Seja R_{τ_i} o maior expoente das variáveis em x_{τ_i} . Similarmente, seja R_k o maior expoente das variáveis em x_k .

Então, o grau de não linearidade do modelo é dado por $\max(L_\tau, L_k)$ onde:

- $L_\tau = \left(\sum_{i=1}^{\alpha} R_{\tau_i} \right) + 1.$

e

- $L_k = R_k + 1.$

Prova:

O maior grau de não-linearidade presente nos agrupamentos de um modelo ocorrerá para o maior entre $\Omega_{x_{\tau_1} x_{\tau_2} \dots x_{\tau_\alpha} y}$ ou $\Omega_{x_k u}$:

Exemplo 5.1

Suponha um modelo de segunda ordem com as duas constantes de tempo variáveis de forma que $x_{\tau_1} = y$, e $x_{\tau_2} = u + y^2$. Suponha também que o ganho varie de forma que $x_k = u^4$. Portanto, o grau de não-linearidade do sistema será dado pelo maior entre:

$$L_\tau = (1 + 2) + 1 = 4, \text{ resultante de } \Omega_{x_{\tau_1} x_{\tau_2} y},$$

onde 1 é o maior expoente de x_{τ_1} , 2 é o maior expoente de x_{τ_2} e 1 é o expoente de y .

e

$$L_k = 4 + 1 = 5, \text{ resultante de } \Omega_{x_k u}.$$

onde 4 é o expoente de x_k e 1 é o expoente do termo u .

Sendo então, o grau de não-linearidade do modelo igual a 5.

Exemplo 5.2

Seja um modelo polinomial NARX de segunda ordem e ganho unitário, onde $x_{\tau_1} = y(k-1) + u^2(k-1)$ exprime a dependência da primeira constante de tempo, e $x_{\tau_2} = y^3(k-1)$ a dependência da segunda constante de tempo. O modelo é:

$$y(k) = u(k-1) + 0,5u^2(k-1)y(k-1) - 0,23y^4(k-1)y(k-2) - 0,14y^3(k-1)u^2(k-1)y(k-2) \\ + 1,32u^2(k-1)y^4(k-1) - 0,76y^5(k-1) + 0,5y^2(k-1) - 1,2y^3(k-1)y(k-2)$$

Do modelo anterior tem-se que $R_{\tau_1} = 2$ e $R_{\tau_2} = 3$, donde $L_{\tau} = 6$. Como o ganho é unitário, tem-se que $R_k = 0$ e consequentemente $L_k = 1$. Portanto, o modelo possui grau de não-linearidade 6, conforme visto.

Portanto, pode-se dizer que existe uma relação estreita entre o grau de não-linearidade, a ordem, e forma de variação dos parâmetros ganho e constante de tempo do modelo.

Note que caso o sistema não apresente nem ganho e nem constante de tempo variáveis, R_{τ} e R_k são nulos, resultando em um sistema com grau de não-linearidade 1, ou seja, sistema linear.

5.4 Comentários Finais

O foco deste capítulo foi analisar certos critérios que possibilitassem inferir quais termos deveriam fazer parte de um modelo polinomial NARX para a situação em que o ganho e uma das constantes de tempo varie segundo as equações (4.1) e (4.2)

respectivamente. Tentou-se também representar as características estáticas de tais parâmetros.

Para modelos com ordem superior a um, torna-se difícil explicitar separadamente o valor de cada constante de tempo, mesmo quando estas não são variáveis, caso não se tenha um conhecimento *a priori* de nenhuma delas. Isto se deve ao fato do número de incógnitas ser maior que o número de restrições disponíveis. A menos que o sistema apresente constantes de tempo iguais, não se consegue em geral calcular o valor destas individualmente. Esta dificuldade pode ser observada nas equações (C.5) e (C.11), onde sendo considerada apenas uma constante de tempo variável, as outras são consideradas constantes, e por este motivo aparecerão relacionadas com a parte constante da equação de variação da constante de tempo, ou seja, estarão diretamente associadas a τ_0 .

Observando-se as equações (5.6), (C.6) e (C.12), que representam as variações do ganho, percebe-se que, para quando os termos do agrupamento $\Omega_{x,y}$ não exercem nenhuma influência sobre o ganho, $\Sigma_{x,y} \approx 0$. Nestas condições, para um sistema como o da equação (5.1), a menos que τ seja constante ou o coeficiente do termo variável seja muito pequeno, ou seja, $\Sigma_{x,y} \approx 0$, a relação estática, e consequentemente o ganho, variarão também em função da constante de tempo.

Para o sistema da equação (C.1), a condição necessária para que $\Sigma_{x,y} \approx 0$, é que $E \approx -F$. Nesta situação, ocorre que a informação relativa a parte variável da constante de tempo, aparecerá dividida de maneira aproximadamente igual entre os termos, permitindo a recuperação da característica da constante de tempo, e impedindo que esta influencie na relação estática.

Já para um sistema de terceira ordem como o da equação (C.7), a condição para que os termos relativos a variação da constante de tempo não interfiram na relação estática é: $K + E + F \approx 0$.

Note que as considerações anteriores, ou seja, que $\Sigma_{x,y} \approx 0$, mostram que o agrupamento de termos $\Omega_{x,y}$, em uma análise de agrupamentos de termos e coeficientes de agrupamentos, pode parecer espúrio devido ao cancelamento. Entretanto, tal cancelamento é necessário para não interferir no ganho. Neste caso, o cancelamento não indica que este agrupamento seja espúrio.

Logo, com base nas suposições anteriores percebe-se que a variedade de interpretação do tipo de variação das constantes de tempo de um modelo polinomial NARX é muito grande. No entanto, pode-se manipulá-lo da mesma forma que o modelo polinomial discretizado tentando-se agregar em um ponto as constantes de tempo. Esta análise, mesmo quando não se consiga representar e calcular as constantes de tempo variáveis, permite saber quais tipos de termos devem ou não compor o modelo, para permitir que se represente pelo menos a característica estática do ganho de maneira direta e explícita.

A análise da característica estática do ganho, aliada aos conhecimentos das dependências transitivas e não-transitivas, permite saber quais termos devem ser incluídos ou não no modelo a fim de se recuperar a característica estática do ganho.

Então, pode-se resumir em uma tabela quais termos podem estar presentes nos modelos lineares (tabela I) e não-lineares (tabela II), para os dois tipos de dependência: dependência do ganho e dependência de uma constante de tempo.

I. Tabela de Agrupamento de termos para modelos lineares.

Tipo de dependência	Agrupamento de termos possíveis	Agrupamento de termos impossíveis
Constante de Tempo	Ω_y, Ω_u	todos os demais (não há dependência)
Ganho	Ω_y, Ω_u	todos os demais (não há dependência)

II. Tabela de Agrupamento de termos para modelos não-lineares.

Tipo de dependência	Agrupamento de termos possíveis	Agrupamento de termos impossíveis
Constante de Tempo	$\Omega_y, \Omega_u, \Omega_{x_\tau y}$	$\Omega_{x_k y}$
Ganho	$\Omega_y, \Omega_u, \Omega_{x_k u}$	$\Omega_{x_\tau u}$

Pela tabela I, percebe-se que sistemas lineares não apresentam ganho e constantes de tempo variáveis, ao contrário dos sistemas não-lineares da tabela II.

6

Uso de Técnicas de Identificação na Modelagem de Sistemas com Características Estáticas Não-lineares

6.1. Introdução

Existem certas técnicas de representação de modelos matemáticos de sistemas reais que representam sistemas não-lineares como dois blocos em cascata, um bloco referente à característica estática não-linear, e outro bloco referente ao modelo linear dinâmico (Billings and Fakhouri, 1982). Se esta característica estática antecede o sistema linear, o modelo é chamado “modelo de Hammerstein” (Billings and Fakhouri, 1982; Eskinat et al., 1991), caso esta característica seja posterior ao sistema dinâmico linear, o modelo é dito “modelo de Wiener” (Wiener, 1958; Billings and Fakhouri, 1978).

Sabe-se que algumas das não-linearidades de um processo podem aparecer devido a componentes que apresentam características estáticas não-lineares, como no caso de certos atuadores. Estas características podem trazer problemas para o controle do processo. Ou seja, um controlador ajustado para trabalhar em uma certa faixa de operação, onde o sistema se encontra em um determinado ponto de operação da característica estática, pode não apresentar um bom desempenho quando ao mudar de ponto de operação. Então, tendo-se informação a respeito da não-linearidade estática presente no sistema, é possível, pelo menos em princípio, levar tal informação em conta na implementação de uma lei de controle.

Neste capítulo desenvolver-se-á um procedimento, utilizando técnicas de agrupamentos de termos e coeficientes de agrupamentos na determinação de um conjunto de termos efetivos para um modelo de Wiener e um modelo de Hammerstein. Será visto também como, a partir do modelo dinâmico, recuperar a informação relativa às características estáticas dos sistemas.

O capítulo está organizado como se segue: na seção 6.2 será mostrada a característica estática considerada neste capítulo, na seção 6.3 será apresentado um procedimento de redução na estrutura de um modelo de Hammerstein utilizando a característica estática da seção 6.2. Na seção 6.4 será feita a mesma análise para um modelo de Wiener, e finalmente na seção 6.5 os resultados serão comparados e os comentários finais serão feitos.

6.2 Característica Estática

Dentre os diversos tipos de características estáticas existentes nos processos reais, escolheu-se a característica representada pela tangente hiperbólica de um sinal 'u', cuja amplitude varia entre $-2 \leq u \leq 2$, de acordo com a figura 6.1.

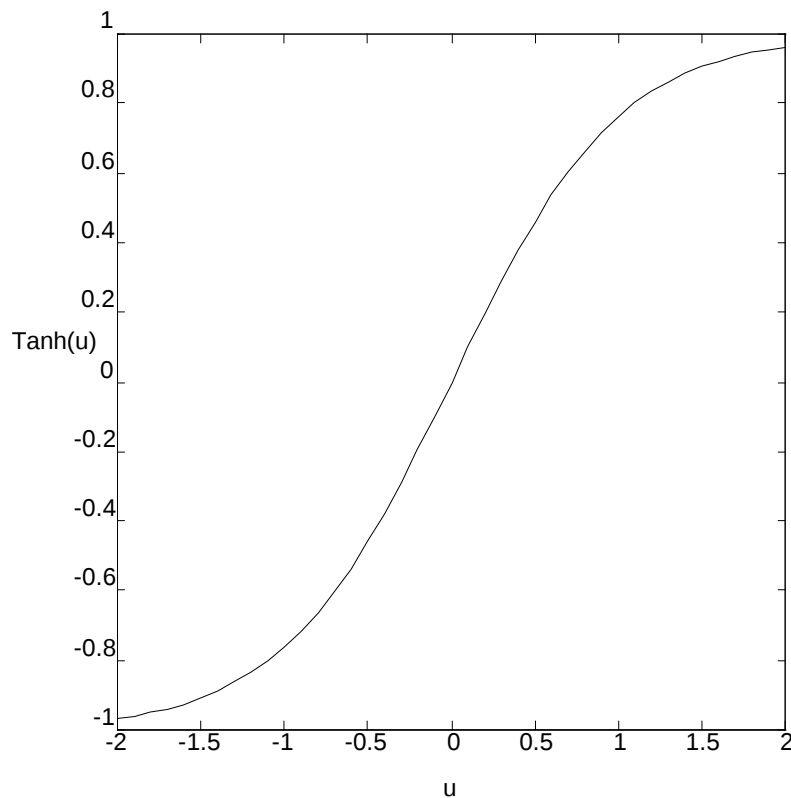


Figura 6.1 - Característica estática de $u \times \tanh(u)$.

Tal característica estática foi escolhida devido a semelhança apresentada com características estáticas presentes em diversos processos industriais e a sua comum utilização na literatura (Kortmann and Unbehauen, 1988; Genesio and Tesi, 1993; Gray et al., 1996).

6.3 Modelo de Hammerstein

Nesta seção, será considerado um modelo de Hammerstein, e através de dados coletados a partir deste identificar-se-á um modelo polinomial NARMAX preliminar. Obtido tal modelo, serão feitas considerações e estipulados procedimentos que visem adequar a estrutura deste.

6.3.1 Simulação do modelo e Obtenção de Dados

Seja o modelo de Hammerstein da figura 6.2:

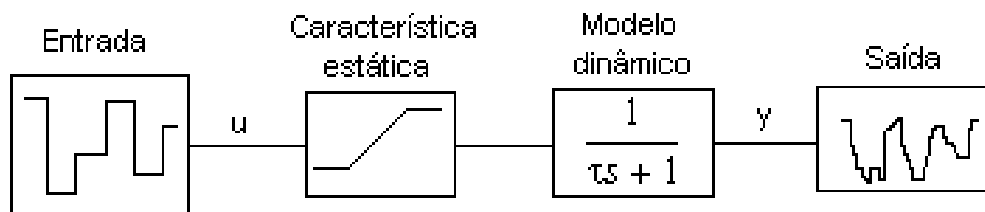


Figura 6.2 - Modelo com característica estática.

O modelo não-linear é expresso pela característica estática representada pela tangente hiperbólica dos dados de entrada, em série com um modelo linear dinâmico de primeira ordem, com ganho e constante de tempo unitários.

O sinal de entrada é mostrado na figura 6.3. Este sinal varia entre $-2 \leq u \leq 2$, com duração de pulsos de 40 amostras. O tempo de amostragem utilizado foi de 0,1, que é 10

vezes menor que a constante de tempo do modelo dinâmico. A aplicação de tal entrada no modelo resultou na saída da figura 6.4.

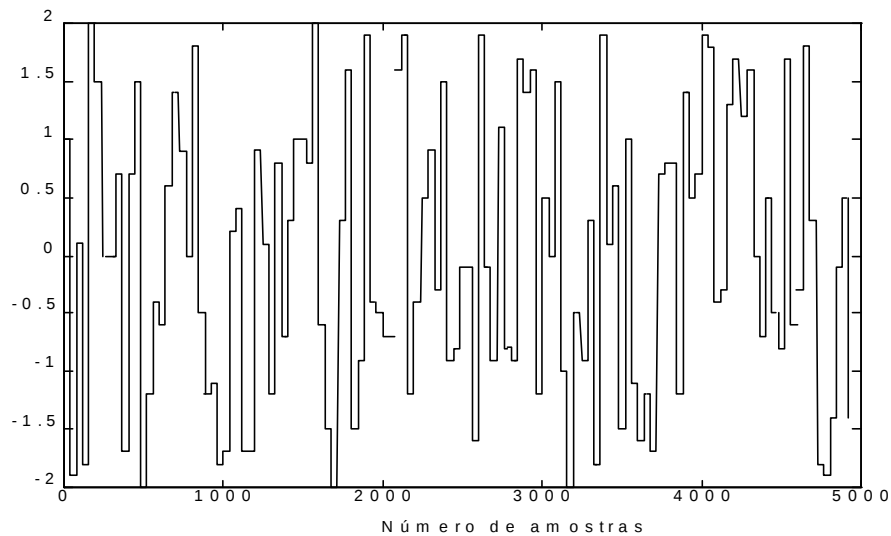


Figura 6.3 - Dados de entrada em função do número de amostras.

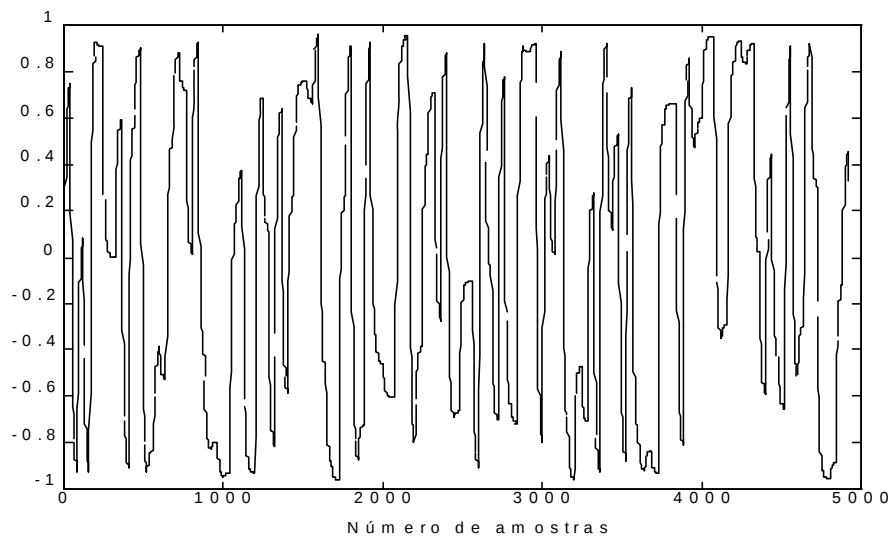


Figura 6.4 - Dados de saída em função do número de amostras.

Os dados foram dizimados de cinco em cinco amostras, acordo com os testes estatísticos de correlação, citados no capítulo 2, onde encontrou-se a seguinte faixa de variação do tempo de amostragem: $0,2 \leq T_s \leq 0,5$.

Ao sinal de saída, depois de dizimado, foi adicionado ruído resultando numa relação sinal/ruído (SNR) de 80 decibéis. A finalidade de acrescentar ruído é que sistemas reais normalmente apresentam este tipo de interferência, e, a presença destes nos dados pode induzir termos espúrios no modelo. Os dados dizimados são apresentados nas figuras 6.5 e 6.6.

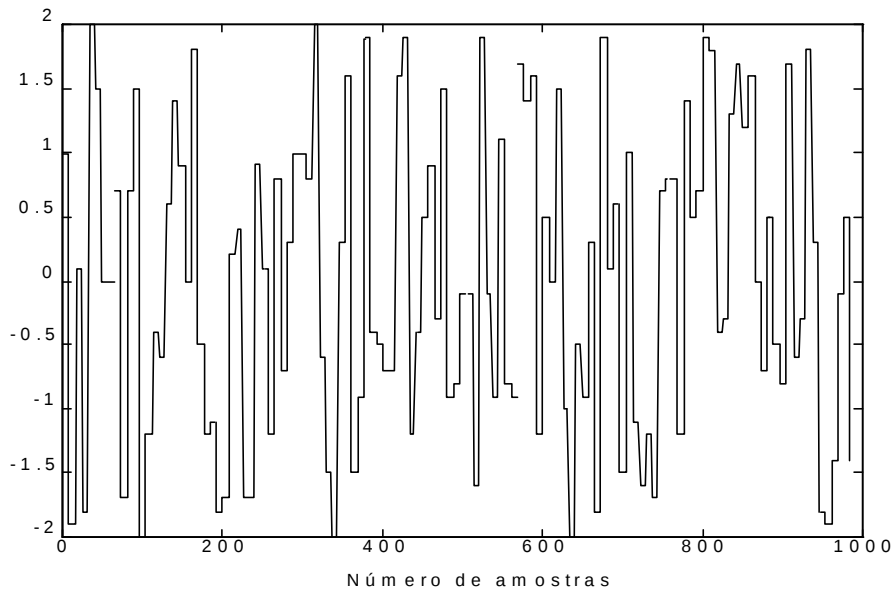


Figura 6.5 - Dados de entrada dizimados.

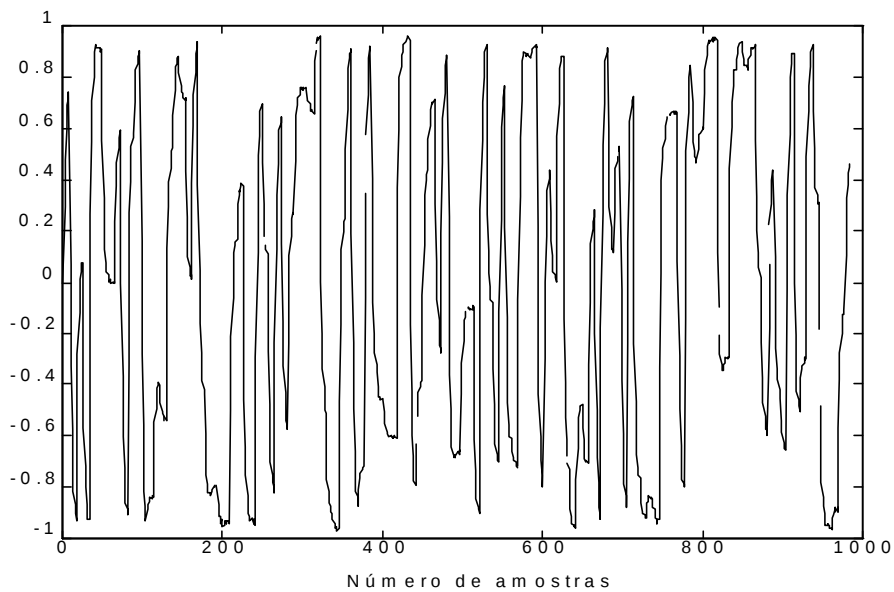


Figura 6.6 - Dados de saída dizimados e com ruído de SNR = 80 dB.

6.3.2 Análise de Agrupamentos de Termos e Coeficientes de Agrupamentos

Para se fazer a análise de agrupamentos de termos e coeficientes de agrupamento¹ do sistema proposto, foram estimados 38 modelos polinomiais NARMAX variando-se de 3 a 40 termos no modelo, sendo considerados grau de não-linearidade 3 e ordem 3, que normalmente são suficientes para a identificação de modelos para um grande número de sistemas reais. Paralelamente ao cálculo dos coeficientes de agrupamento de termos foram usadas algumas técnicas estatísticas que auxiliassem na escolha de um número adequado de termos para o modelo. Tais técnicas, que foram citadas no capítulo 2 são: variância dos resíduos, critério de Informação de Akaike (AIC), Final Prediction Error (FPE), Lei dos logaritmos Interativos (LILC), e o critério de Informação Bayesiano (BIC). Os resultados são apresentados nas figuras 6.7 a 6.9.

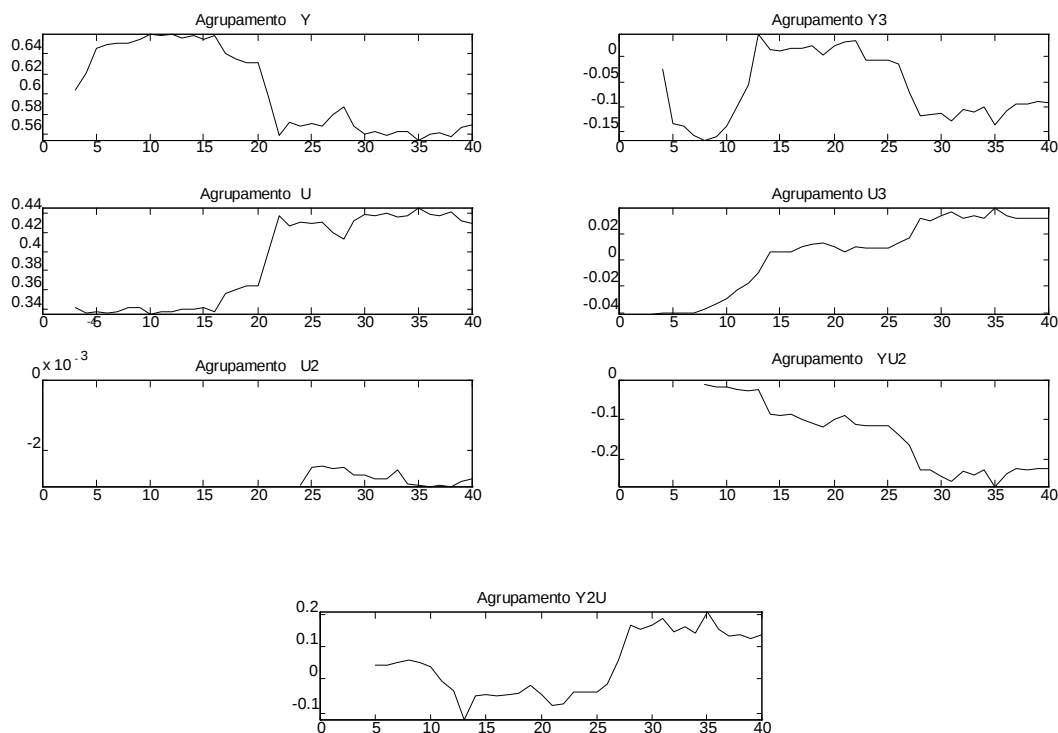


Figura 6.7 - Coeficientes de agrupamento em função do número de termos no modelo.

¹ As rotinas MATLAB[™] utilizadas neste trabalho, para cálculo e simulação de agrupamentos de termos e coeficientes de agrupamentos podem ser encontradas em Jácome e Aguirre (1996b).

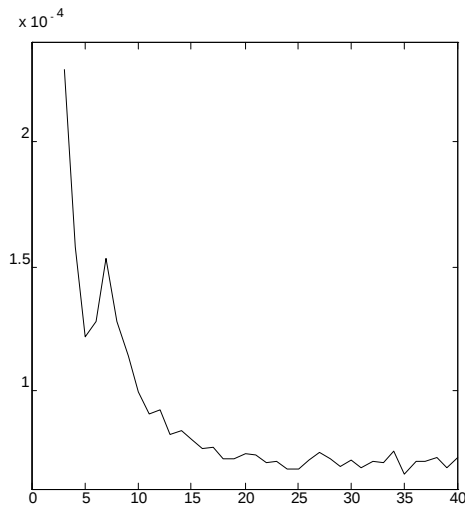


Figura 6.8 - Variância dos resíduos.

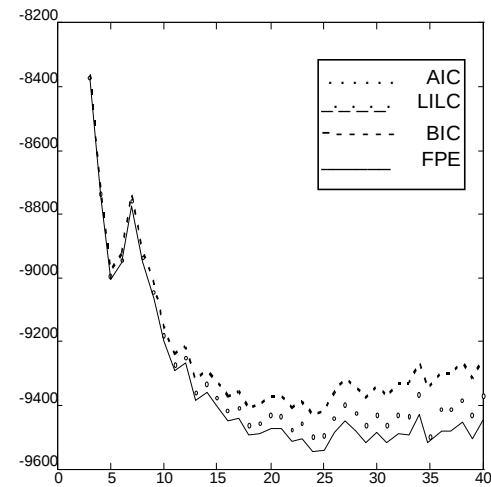


Figura 6.9 - Critérios de informação

Pela observação dos critérios de informação, e pela variância dos resíduos, percebe-se que estes sugerem aproximadamente 24 termos no modelo. Isto porque o primeiro mínimo ocorre com 24 termos, não havendo melhorias após este valor. Portanto, o modelo obtido com 24 termos foi:

$$\begin{aligned}
 y(k) = & 0,548 y(k-1) + 0,3365 u(k-1) - 0,03982 u(k-1)u(k-1)u(k-1) - \\
 & 0,005782 y(k-1)y(k-1)y(k-1) + 0,07324 u(k-2)y(k-3)y(k-1) - \\
 & 0,03973 u(k-2)y(k-3)y(k-3) + 0,07381 u(k-3)u(k-2)u(k-1) - \\
 & 0,1891 u(k-2)u(k-1)y(k-1) - 0,2894 u(k-2)y(k-1)y(k-1) + \\
 & 0,01887 y(k-3) + 0,05182 u(k-2)u(k-2)u(k-2) + \\
 & 0,1141u(k-3)y(k-3)y(k-1) - 0,03524 u(k-2)u(k-1)y(k-2)- \\
 & 0,1035 u(k-2)u(k-2)y(k-1) - 0,07721 u(k-3)u(k-2)u(k-2) + \\
 & 0,1438 u(k-1)y(k-1)y(k-1) + 0,06438 u(k-2) + \\
 & 0,2271 u(k-2)u(k-2)y(k-2) + 0,04713 u(k-1)y(k-3)y(k-2) + \\
 & 0,1896 u(k-3)y(k-1)y(k-1) + 0,031 u(k-3) - 0,2746 u(k-3)y(k-2)y(k-1) - \\
 & 0,0145 u(k-3)u(k-1)y(k-3) - 0,0003006 u(k-1) u(k-1) + \text{Termos de ruído.}
 \end{aligned}$$

Cabe ressaltar que a parte MA (moving average) do modelo, relativo aos termos de ruído, somente é utilizada no processo de identificação para evitar a polarização do modelo, sendo descartada para outros fins.

Pela figura 6.10 percebe-se que o modelo com 24 termos ajusta bem os dados de identificação.

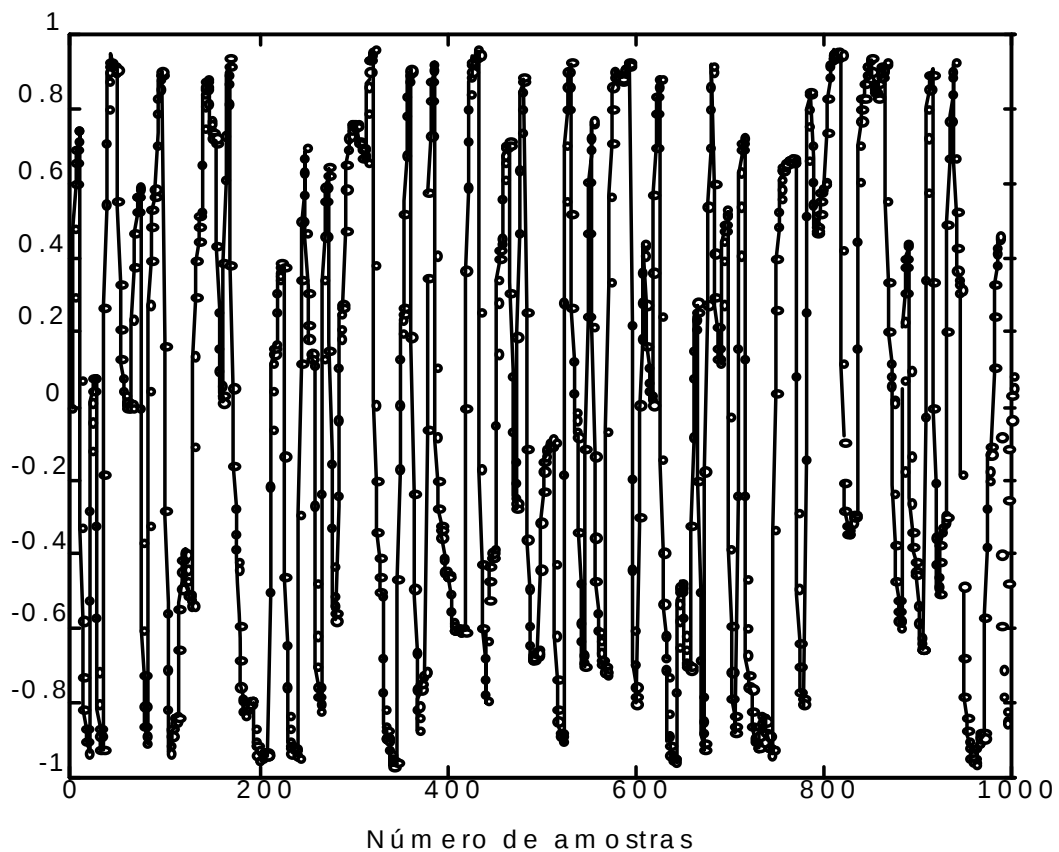


Figura 6.10 - (.....) Dados de identificação (____) Modelo com 24 termos.

A validação estatística do modelo é apresentada na figura 6.11, abaixo.

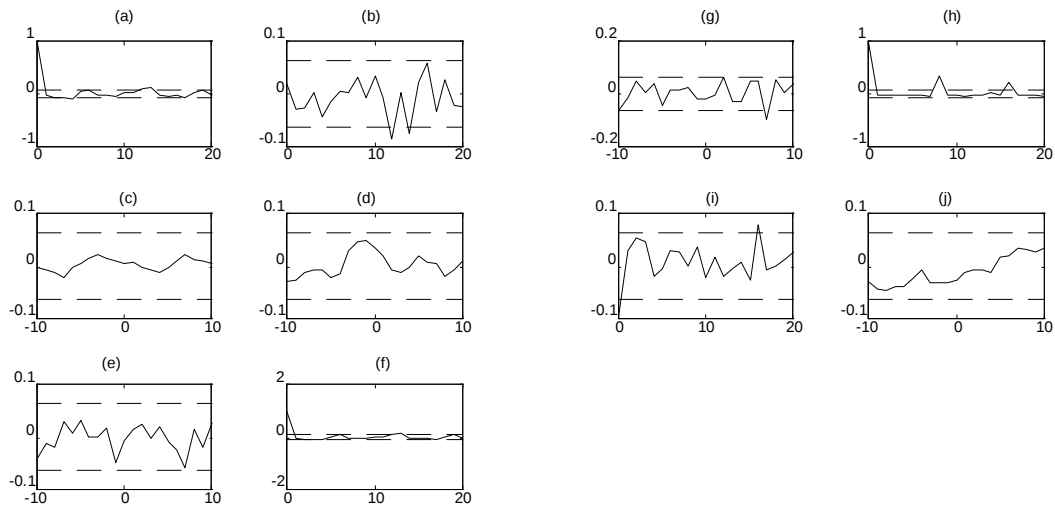


Figura 6.11 - Validação estatística do modelo com 24 termos. Correlações calculadas: (a) $\xi(t) \times \xi(t)$; (b) $u(t) \times \xi(t)$; (c) $\xi(t) \times \xi(t) \times u(t)$; (d) $u^2(t) \times \xi(t)$; (e) $u^2(t) \times \xi^2(t)$; (f) $\xi'(t) \times \xi'(t)$; (g) $\xi'(t) \times \xi^2(t)$; (h) $\xi^2(t) \times \xi^2(t)$; (i) $(y(t) \times e(t))' \times \xi^2(t)$; (j) $(y(t) \times e(t))' \times u^2(t)$.

A figura 6.12 mostra a característica estática obtida do modelo, *via simulação*.

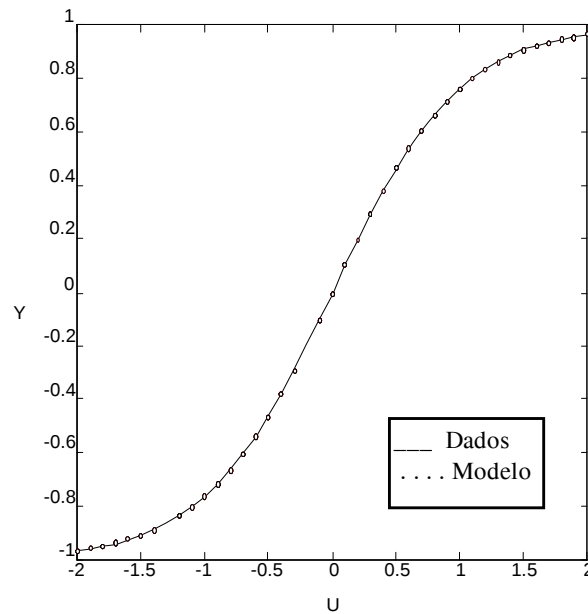


Figura 6.12 - Características estáticas.

A figura 6.13 abaixo, mostra o modelo com 24 termos sendo utilizado como preditor.

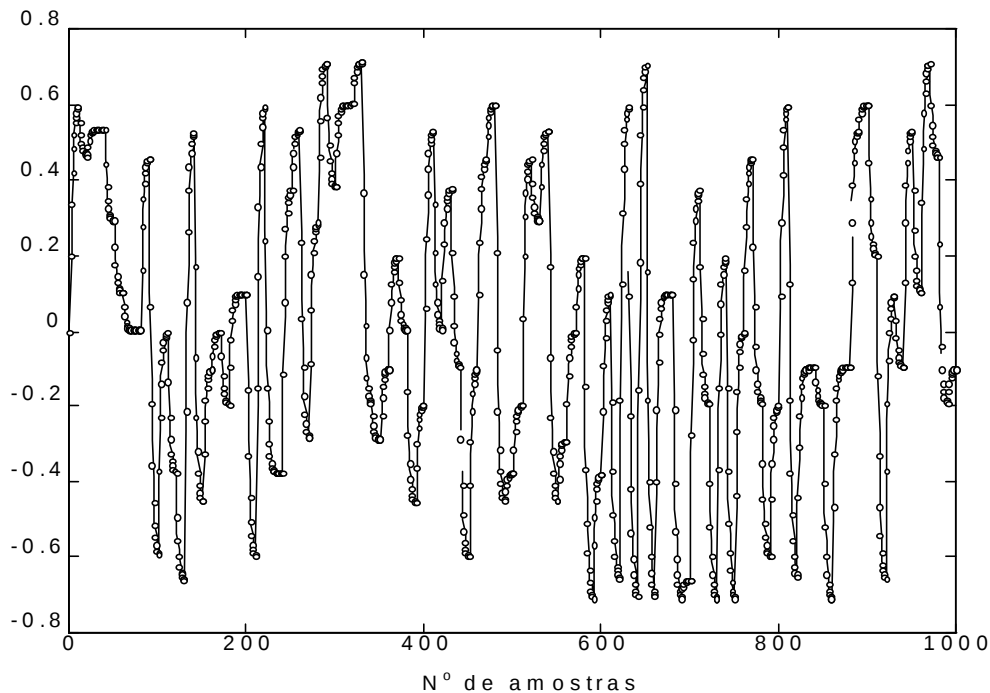


Figura 6.13 - (—) Dados e (. . .) predição do modelo.

Com base nos resultados anteriores pode-se dizer que o modelo com 24 termos é um bom modelo. Porém é um modelo grande, podendo estar sobreparametrizado. A questão a ser colocada é: Será que se pode limitar o conjunto de termos candidatos, de modo que o algoritmo de identificação escolha somente os termos realmente necessários para o modelo, sendo este de tamanho aceitável e dinamicamente válido?

O que vai ser feito a seguir, é uma análise baseada em agrupamentos de termos e coeficientes de agrupamentos, que auxiliará na redução de termos no modelo, restando apenas aqueles considerados efetivos.

Através da análise da variação dos coeficientes de agrupamentos da figura 6.7 percebe-se que os coeficientes dos agrupamentos Ω_{y^3} e Ω_{y^2u} oscilam em torno do zero, o que, de acordo com o capítulo 3, indica que tais agrupamentos podem ser descartados do modelo. O agrupamento Ω_{u^2} além de aparecer no modelo somente a partir do 24º termo, possui uma amplitude muito baixa com relação aos demais agrupamentos, sendo

desta forma descartado do modelo também. Os demais agrupamentos apresentam uma tendência, sendo portanto mantidos, por enquanto, na estrutura do modelo.

Com a redução do número de termos candidatos ao modelo, foi refeita a análise de coeficientes de agrupamento, porém variando-se o número de termos no modelo até um máximo de 30, isto porque o número de termos candidatos ao modelo tornou-se menor com a eliminação de certos agrupamentos. Os resultados são apresentados na figura 6.14.

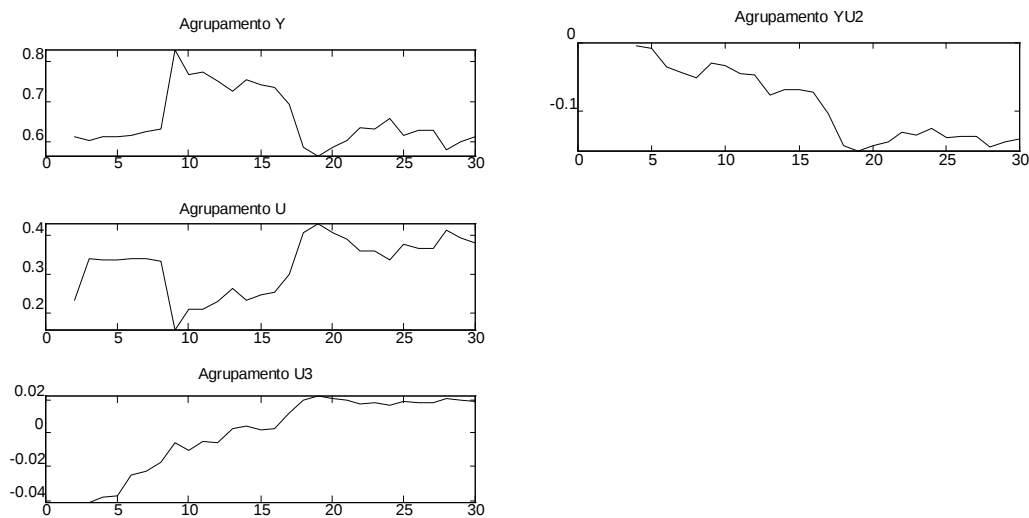


Figura 6.14 - Coeficientes de agrupamento em função do número de termos no modelo.

Os gráficos da figura 6.14, sugerem que os agrupamentos são todos efetivos. Todos os coeficientes apresentam uma tendência em torno de um certo valor, sendo que mais nenhum agrupamento necessita ser retirado do conjunto de termos candidatos. Portanto, uma análise dos critérios estatísticos de informação das figuras 6.15 e 6.16 sugerem um modelo com dezoito termos.

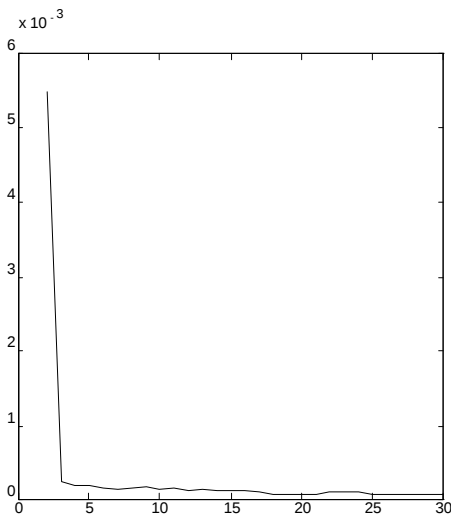


Figura 6.15 - Variância dos resíduos.

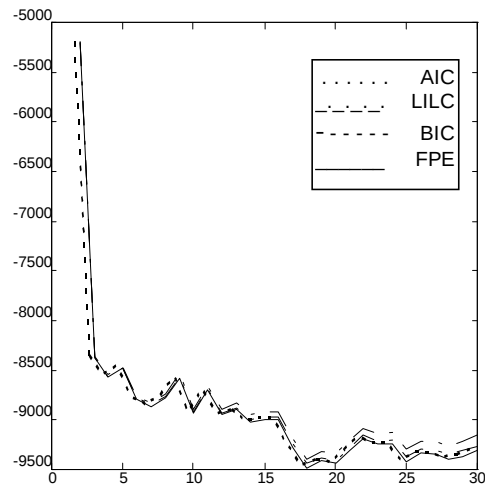


Figura 6.16 - Critérios estatísticos.

Logo, o modelo com 18 termos é:

$$\begin{aligned}
 y(k) = & 0,72640 y(k-1) + 0,33591 u(k-1) - 0,03926 u(k-1)u(k-1)u(k-1) + \\
 & + 0,02008 u(k-3)u(k-3)y(k-1) - 0,00504 u(k-3)u(k-3)u(k-2) - \\
 & - 0,09509 u(k-3)u(k-3)y(k-2) + 0,04599 u(k-2)u(k-2)u(k-2) - \\
 & - 0,23620 y(k-2) + 0,00280 u(k-2) + 0,05018 u(k-3)u(k-3)y(k-3) + \\
 & + 0,00083 u(k-3)u(k-2)y(k-2) + 0,098 y(k-3) + \\
 & + 0,00652 u(k-3)u(k-3)u(k-3) + 0,06938 u(k-3) + \\
 & + 0,01175 u(k-2)u(k-1)u(k-1) - 0,21081 u(k-1)u(k-1)y(k-1) + \\
 & + 0,19340 u(k-1)u(k-1)y(k-2) - 0,10882 u(k-2)u(k-2)y(k-1)
 \end{aligned} \quad (6.1)$$

cuja validação estatística é apresentada na figura 6.17.

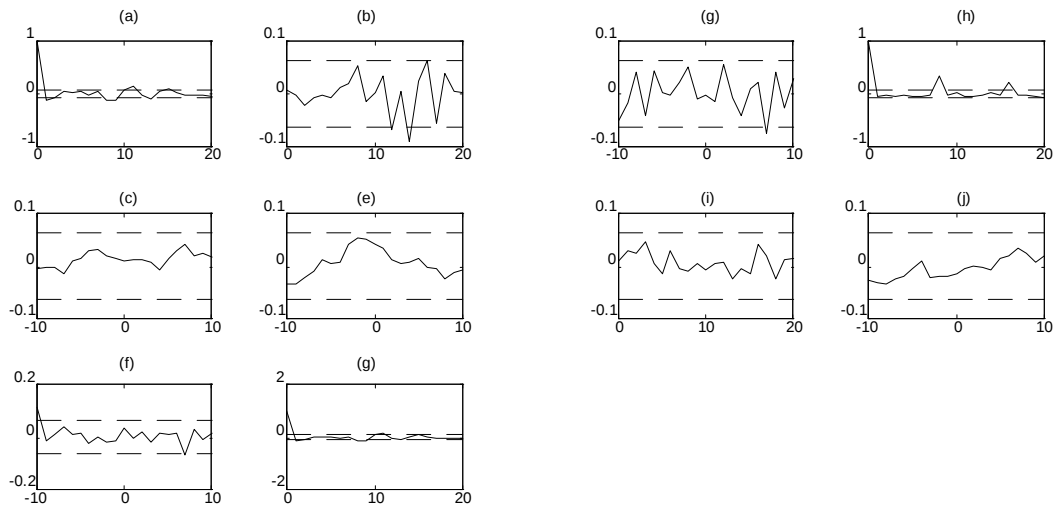


Figura 6.17 - Validação estatística do modelo com 18 termos. Correlações calculadas: (a) $\xi(t) \times \xi(t)$; (b) $u(t) \times \xi(t)$; (c) $\xi(t) \times \xi(t) \times u(t)$; (d) $u^2(t) \times \xi(t)$; (e) $u^2(t) \times \xi^2(t)$; (f) $\xi'(t) \times \xi'(t)$; (g) $\xi'(t) \times \xi^2(t)$; (h) $\xi^2(t) \times \xi^2(t)$; (i) $(y(t) \times e(t))' \times \xi^2(t)$; (j) $(y(t) \times e(t))' \times u^2(t)$.

Este modelo com 18 termos também foi utilizado também na predição, com uma *massa de dados diferente* da massa de identificação, de forma semelhante ao modelo com 24 termos. O resultado é apresentado na figura 6.18.

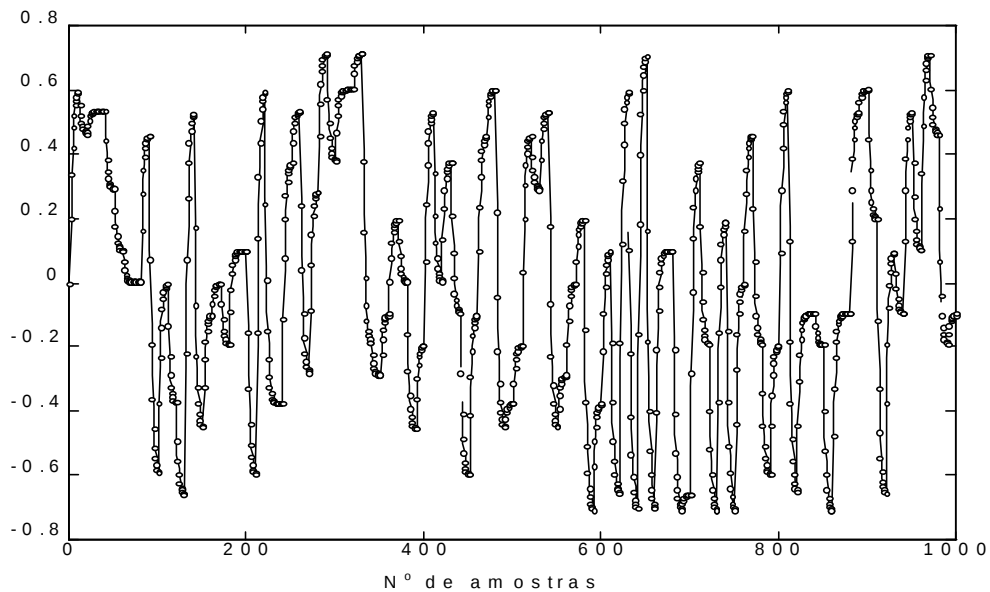


Figura 6.18 - (____) Dados e (.....) predição do modelo.

Com base nos critérios de correlação da figura 6.17, e na predição da figura 6.18, pode-se dizer que o modelo com 18 termos continua bom. Note que os dados de predição são diferentes dos dados de identificação. Portanto, resta saber se tal modelo é capaz de recuperar analiticamente a característica estática ao contrário do primeiro modelo identificado, e talvez uma equação para a variação da relação estática saída/entrada. Considerando o modelo da equação (6.1) em regime estático, tem-se:

$$0,4118Y = 0,40853U + 0,01996U^3 - 0,1502YU^2 \quad (6.2)$$

Que pode ser reescrita da seguinte forma:

$$Y = \left(\frac{0,40853U + 0,01996U^3}{0,4118 + 0,1502U^2} \right) \quad (6.3)$$

Atribuindo-se valores para U na equação (6.3), obedecendo sua faixa de variação de $-2 \leq u \leq 2$, chega-se na característica estática da figura 6.19.

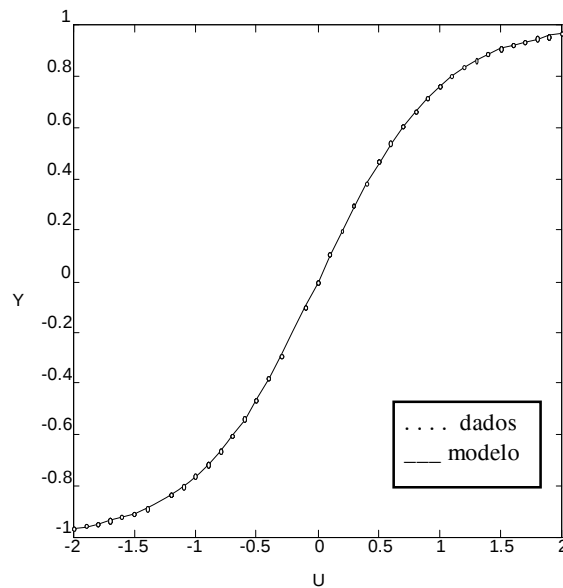


Figura 6.19 - Características estáticas.

Da figura 6.19 percebe-se que o modelo com 18 termos conseguiu reproduzir bem a característica estática, com uma vantagem em relação ao modelo de 24 termos. Esta vantagem baseia-se em que a característica estática do modelo com 18 termos foi encontrada analiticamente atribuindo-se valores para u na equação (6.3), enquanto que para o modelo com 24 termos tal característica foi encontrada via simulação.

Cabe ressaltar que, caso fossem escolhidos 18 termos, apenas truncando-se o modelo com 24 termos e reestimando-se os parâmetros, obter-se-ia um modelo que prediz bem os dados, porém a característica estática recuperada a partir de tal modelo é mostrada na figura 6.20. Note-se ainda que tal característica foi obtida via simulação do modelo, devido a impossibilidade de conseguir uma representação na forma $y = f(u)$.

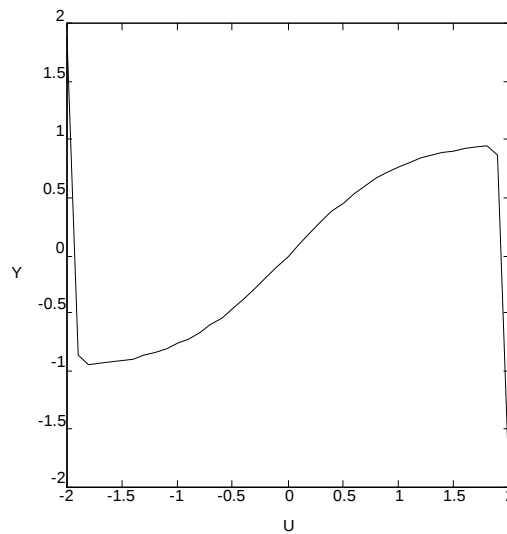


Figura 6.20 - Característica estática de um modelo truncado com 18 termos.

Tomando-se a equação (6.3) e fazendo $\frac{Y}{U}$ pode-se encontrar uma equação da relação estática saída/entrada:

$$K_{re} = \frac{Y}{U} = \left(\frac{0,40853 + 0,01996U^2}{0,4118 + 0,1502U^2} \right) \quad (6.4)$$

A forma gráfica da relação estática com a entrada, $K_{re} = g(u)$, é mostrada na figura 6.21.

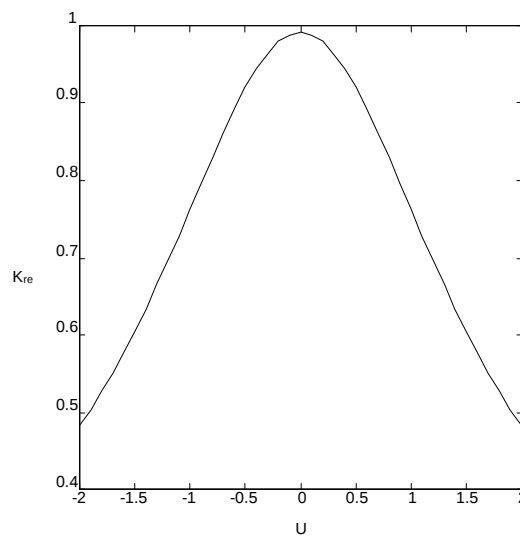


Figura 6.21 - Relação estática em função da entrada.

6.4 Modelo de Wiener

Nesta seção, verificar-se-ão os resultados para a mesma característica estática de seção 6.2, porém, desta vez a característica é colocada depois do modelo dinâmico, constituindo a representação de Wiener.

Portanto, seja o modelo na forma de Wiener da figura 6.22:

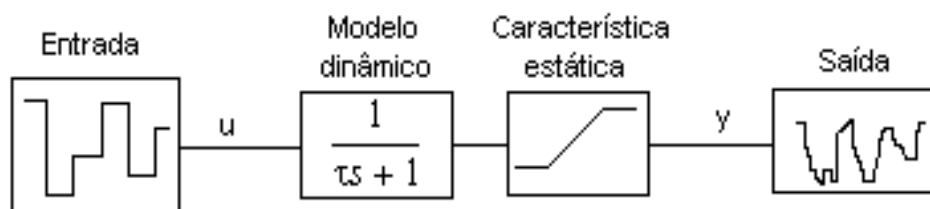


Figura 6.22 - Representação de Wiener

Os dados de identificação foram gerados com um intervalo de integração de 0,1. A entrada aplicada é mostrada na figura 6.23, variando de $-2 \leq u \leq 2$. Através de testes de correlação, escolheu-se o tempo de amostragem de 0,5. Isto implica que os dados obtidos devem ser dizimados de 5 em 5. Após a dizimação, foi somado ruído aos dados, com relação sinal/ruído de 80 decibéis, cujos motivos já foram explicados na seção 6.3. A figura 6.23 mostra os dados de identificação.

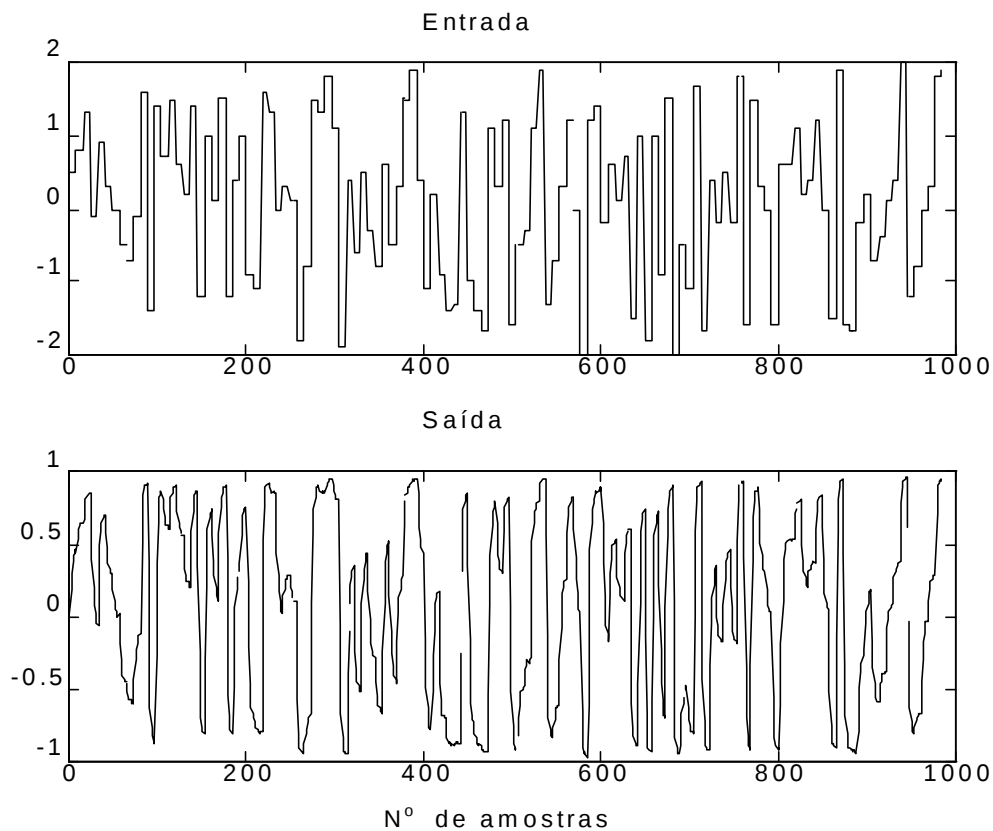


Figura 6.23 - Dados de identificação.

Nesta seção será desenvolvido um procedimento semelhante ao da seção anterior, porém neste caso, a característica estática estará localizada posteriormente ao modelo dinâmico, de acordo com a figura 6.22.

Logo, com os dados obtidos, estimou-se 38 modelos polinomiais NARMAX, variando-se o número de termos de 2 a 40, onde foram calculados os coeficientes de agrupamentos de acordo com a figura 6.24.

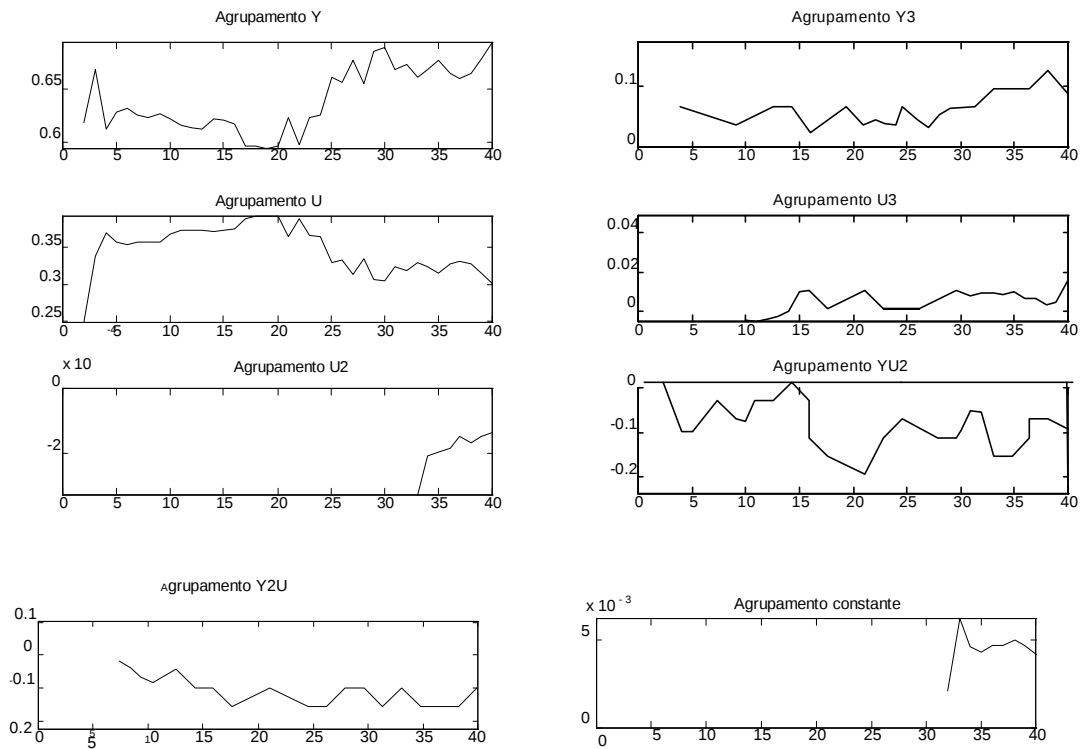


Figura 6.24 - Coeficientes de agrupamento em função do número de termos no modelo.

Com base nos resultados dos critérios estatísticos de informação, escolheu-se um modelo com 17 termos de processo.

Portanto, o modelo com 17 termos é:

$$\begin{aligned}
 y(k) = & 0,5958 y(k-1) + 0,3666 u(k-1) - 0,2852 u(k-2)u(k-1)y(k-1) + \\
 & 0,04127 u(k-3)u(k-3)y(k-2) - 0,1195 u(k-1)u(k-1)y(k-2) + \\
 & 0,1056 u(k-2)u(k-2)y(k-2) + 0,02799 u(k-3)u(k-1)y(k-2) - \\
 & 0,01303 u(k-1)y(k-3)y(k-2) + 0,1338 u(k-1)y(k-1)y(k-1) - \\
 & 0,004879 u(k-1)u(k-1)u(k-1) + 0,04169 y(k-3)y(k-3)y(k-2) + \\
 & 0,03943 u(k-3)u(k-1)u(k-1) + 0,01216 u(k-3)u(k-2)u(k-2) - \\
 & 0,06622 u(k-3)y(k-1)y(k-1) + 0,04266 u(k-2)u(k-2)u(k-1) - \\
 & 0,04529 u(k-3)u(k-3)u(k-2) + 0,02469 u(k-2)
 \end{aligned}$$

Foram aplicados ao modelo os testes de correlação estatística, verificando-se que todos se mantinham dentro da faixa de confiança. A figura 6.25 apresenta o modelo sendo utilizado na predição, com uma *massa de dados diferente* da massa de dados de identificação.

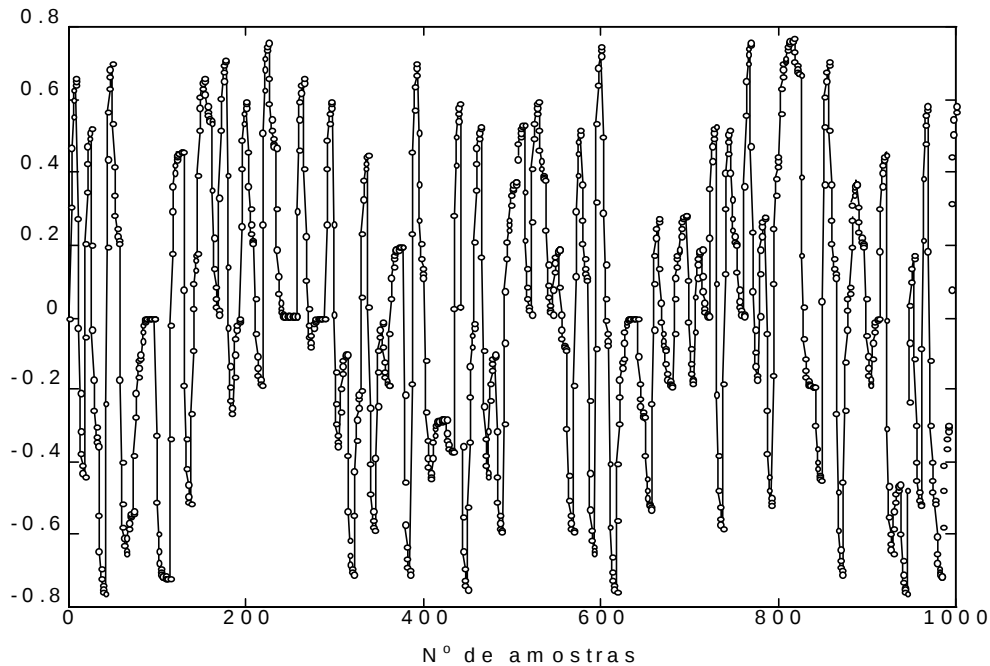


Figura 6.25 - (____) Dados e (. . .) predição do modelo.

A seguir, tentar-se-á simplificar a estrutura do modelo através da retirada de alguns agrupamentos de termos do modelo, esta retirada será feita com base na análise da variação dos coeficientes de agrupamento da figura 6.24.

Observando-se os agrupamentos Ω_0 e Ω_{u^2} na figura 6.24, percebe-se que estes somente começam a fazer parte do modelo a partir do 31º e 33º termos respectivamente, e com valores de coeficientes muito baixos. Estes fatores indicam que termos pertencentes a estes agrupamentos poderiam ser retirados do modelo. O agrupamento Ω_{u^3} permanece variando em torno do zero, e o agrupamento Ω_{yu^2} oscila muito, além de haver passagens pelo eixo zero também. Estas condições também indicam que tais agrupamentos podem ser espúrios, sendo então retirados do conjunto de termos candidatos ao modelo,

juntamente com os agrupamentos Ω_0 e Ω_{u^2} . Assim, reestimaram-se 28 modelos variando-se agora o número de termos de 2 a 30. Esta redução deve-se à diminuição do conjunto do conjunto de termos candidatos ao modelo. Os novos coeficientes de agrupamento podem ser vistos na figura 6.26.

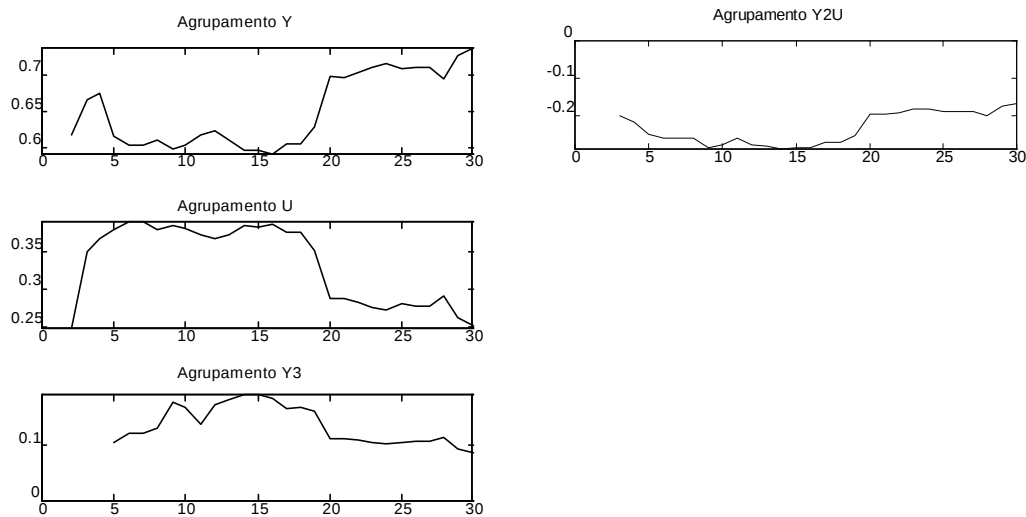


Figura 6.26 - Coeficientes de agrupamento em função do número de termos no modelo.

Foram calculados os respectivos critérios de informação, que sugerem um modelo com 14 termos de processo. O modelo obtido foi:

$$\begin{aligned}
 y(k) = & 0,5969y(k-1) + 0,3742u(k-1) + 98,83u(k-1)y(k-1)y(k-1) - \\
 & 197,9u(k-1)y(k-2)y(k-1) + 1,342y(k-1)y(k-1)y(k-1) + \\
 & 98,83u(k-1)y(k-2)y(k-2) + 0,002191u(k-1)y(k-3)y(k-3) + \\
 & 0,01185u(k-3) - 99,62u(k-2)y(k-1)y(k-1) + 198,4u(k-2)y(k-2)y(k-1) - \\
 & 98,86u(k-2)y(k-2)y(k-2) - 1,772y(k-2)y(k-1)y(k-1) + \\
 & 0,5906y(k-2)y(k-2)y(k-1) + 0,02912y(k-3)y(k-3)y(k-3)
 \end{aligned}$$

O modelo com 14 termos, como preditor é mostrado na figura 6.27.

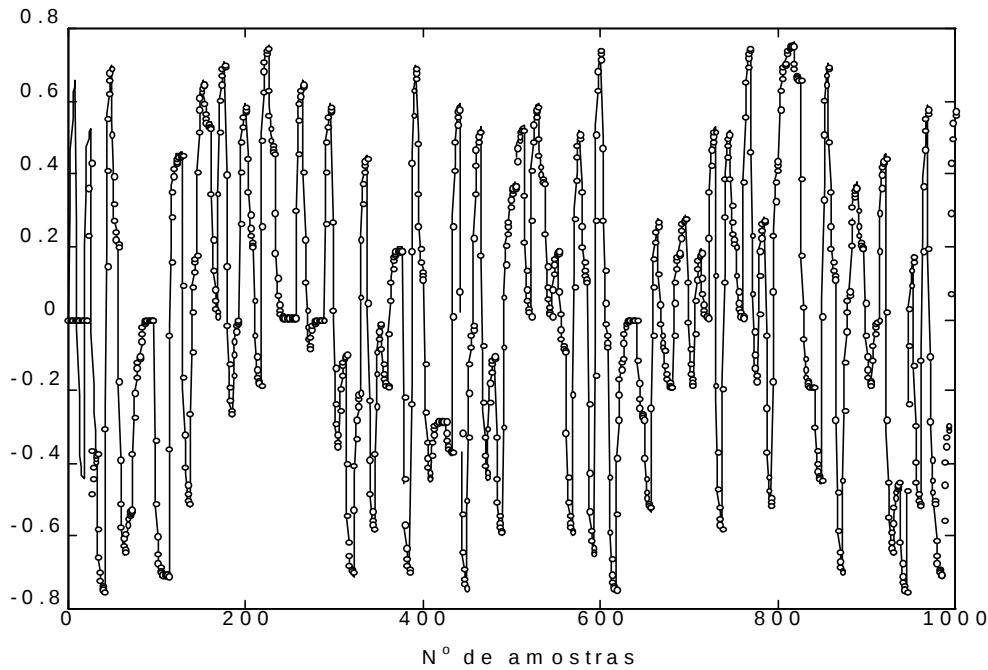


Figura 6.27 - (____) Dados e (. . . .) predição do modelo.

Nesta seção, a característica estática encontra-se posterior ao modelo dinâmico linear, conseqüentemente, a saída de tal característica estática varia com a saída do processo. Nestas condições o ganho apresentar-se-ia como uma dependência da saída do processo, o que implicaria em $x_k = f(y)$.

Então, como se sabe que a constante de tempo do sistema não varia, de acordo com a tabela II do capítulo 5, os agrupamentos passíveis de constituírem o modelo seriam: Ω_y , Ω_u e $\Omega_{x_k u}$. Portanto, os termos do agrupamento Ω_{y^3} serão descartados do conjunto de termos candidatos ao modelo.

Logo, reestimando-se um novo conjunto de modelos, variando-se o número de termos de 2 a 20, devido à diminuição do conjunto de termos candidatos, tem-se:

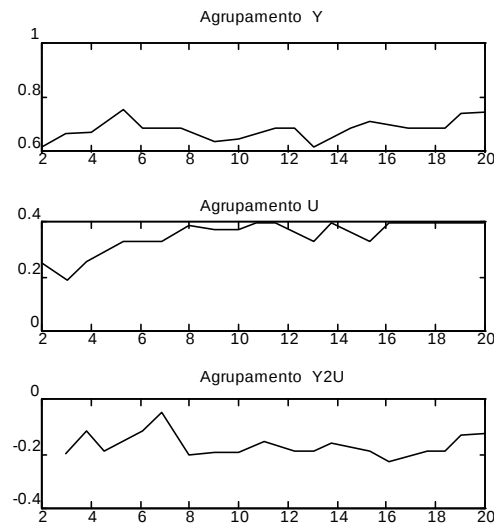


Figura 6.28 - Coeficientes de agrupamento em função do número de termos no modelo.

Então, com base nos critérios estatísticos de informação, escolheu-se um modelo com oito termos, representado na equação (6.5) abaixo. Pela variação dos coeficientes de agrupamento da figura 6.28, também pode se perceber que à partir de oito termos no modelo, os coeficientes permanecem em torno de um mesmo valor.

$$\begin{aligned}
 y(k) = & 0,62470 y(k-1) + 0,38791 u(k-1) - 0,41541 u(k-1)y(k-1)y(k-1) + \\
 & + 0,26182 u(k-1)y(k-2)y(k-1) + 0,03011 u(k-2)y(k-3)y(k-1) - \\
 & - 0,10170 u(k-1)y(k-2)y(k-2) + 0,02378 u(k-2)y(k-3)y(k-3) + \\
 & + 0,00131 u(k-3)y(k-2)y(k-1)
 \end{aligned} \tag{6.5}$$

Na figura 6.29, o modelo (6.5) é utilizado na predição, com uma *massa de dados diferente* da massa de dados de identificação.

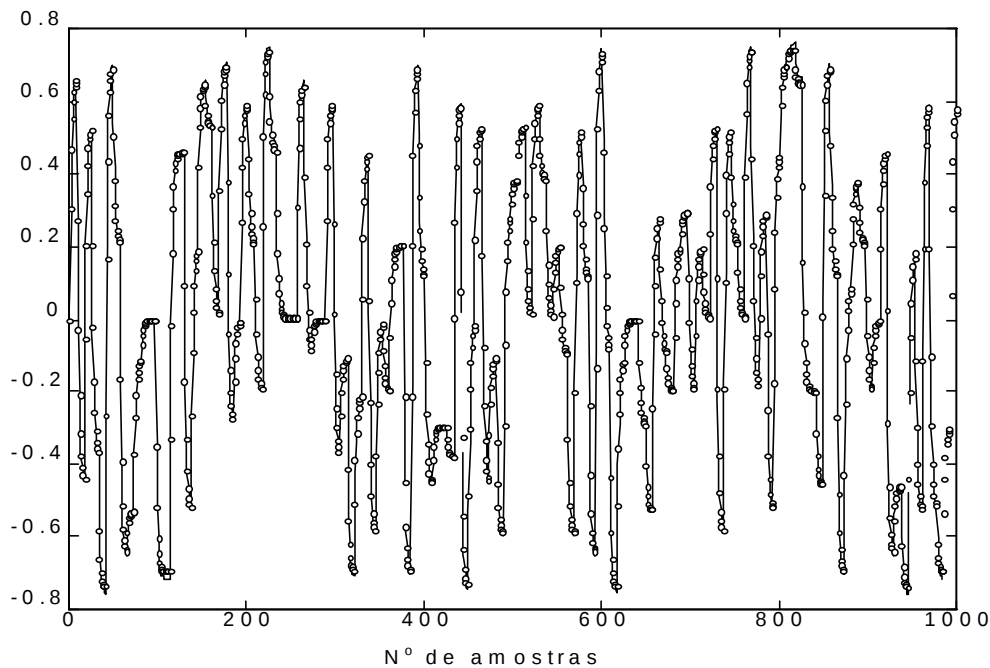


Figura 6.29 - (—) Dados e (- - -) predição do modelo.

Visto que o modelo prediz os dados relativamente bem, verificar-se-á a possibilidade e recuperação da característica estática. Então, da equação (6.5), considerando regime estático, tem-se:

$$Y = 0,6247Y + 0,3879U - 0,2002 Y^2 U$$

Logo:

$$Y = 1,03357U - 0,5334 Y^2 U \quad (6.6)$$

Porém, da equação (6.6), não é possível obter-se o valor na saída da característica estática (Y) analiticamente, devido à existência do termo $Y^2 U$. Entretanto, pode-se calcular os pontos fixos (para o caso não-autônomo) do modelo da equação (6.6).

Tais pontos fixos correspondem às raízes do polinômio (6.6) com U constante. Uma revisão sobre pontos fixos em modelos NARMAX pode ser encontrada em Aguirre e

Mendes (1996). Logo, para cada valor da entrada $-2 \leq U \leq 2$, calculam-se as raízes da equação (6.6). Como o modelo estático na equação (6.6) é um polinômio de segunda ordem, obter-se-ão duas raízes para cada ponto de operação U , que são traçadas na figura 6.30.

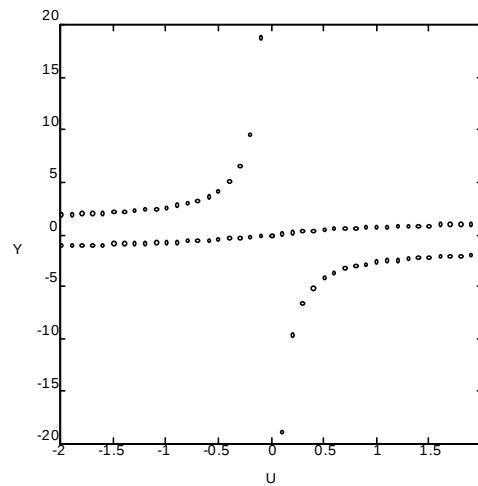


Figura 6.30 - Raízes da equação 6.6.

Tomando-se a equação da característica estática (fig. 6.1), e traçando-a sobre as curvas da figura 6.30, obtém-se:

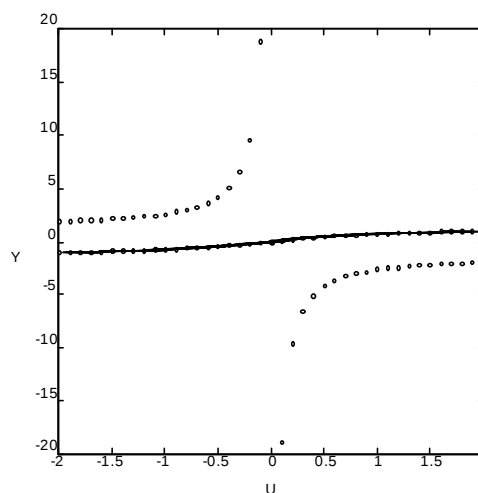


Figura 6.31 - Característica estática do sistema sobre as raízes da equação estática (6.6).

Traçando-se somente uma das raízes e a característica estática:

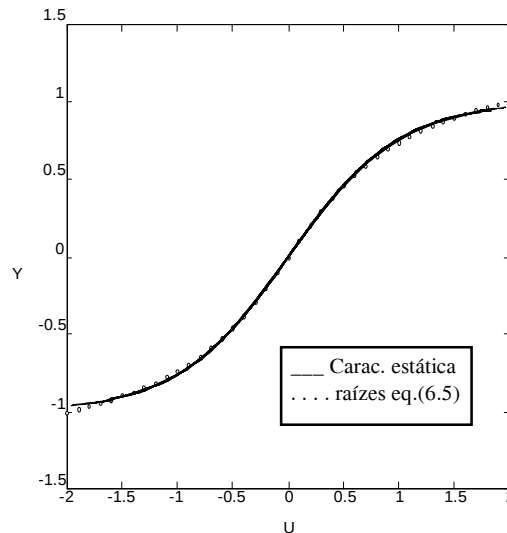


Figura 6.32 - Característica estática e raízes da equação (6.6).

Portanto, o modelo contém a informação a respeito da característica estática do sistema, no entanto ela só poderá ser encontrada mediante o cálculo das raízes da equação estática do modelo, como na figura 6.32. Cabe ressaltar que a obtenção da característica estática através das raízes, ainda é mais vantajosa que a obtenção desta via simulação. Note-se que a característica estática do sistema corresponde à raiz estável do modelo.

Contudo, tomando-se a equação (6.6), e fazendo-se $\frac{\partial y}{\partial u}$ pode-se encontrar a equação de variação do ganho, para este sistema:

$$K = 1,03357 - 0,5334Y^2 \quad (6.7a)$$

Ou ainda, da equação (6.6) tem-se que:

$$\Sigma_u = 0,3879, \Sigma_y = 0,6247 \text{ e } \Sigma_{x_k u} = \Sigma_{y^2 u} = -0,2002$$

Logo, aplicando-se estes coeficientes de agrupamento na equação (5.6) tem-se:

$$K_{re} = \frac{0,3879}{1-0,6247} - \frac{0,2002}{1-0,6247} Y^2 = 1,03357 - 0,53344 Y^2 \quad (6.7b)$$

Que está de acordo com a equação (6.7a).

Portanto, a equação (6.7a) representa o ganho, que varia da mesma forma que a relação estática, com a saída. Porém, desta vez a relação estática varia com a saída do processo, ou seja:

$$K_{re} = h(y) \quad (6.8)$$

Tomando-se a faixa de variação os valores da saída, seja do processo ou do modelo dinâmico (para o caso de modelos dinâmicos lineares), e aplicando-a na equação (6.7b), pode-se encontrar uma curva de $K_{re} \times Y$, onde sabendo-se a saída do sistema em um determinado momento, sabe-se também o valor da relação estática naquele instante. Esta curva é apresentada na figura 6.39.

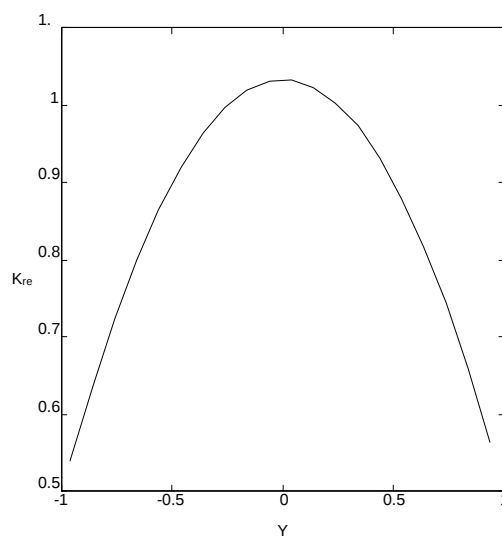


Figura 6.33 - Gráfico da relação estática pela saída do modelo de 8 termos.

Da mesma maneira que na seção 6.3, tomando-se o modelo inicial com 17 termos, truncando-o nos oito primeiros e reestimando-se os parâmetros, encontra-se um modelo cujas raízes são mostradas na figura 6.34.

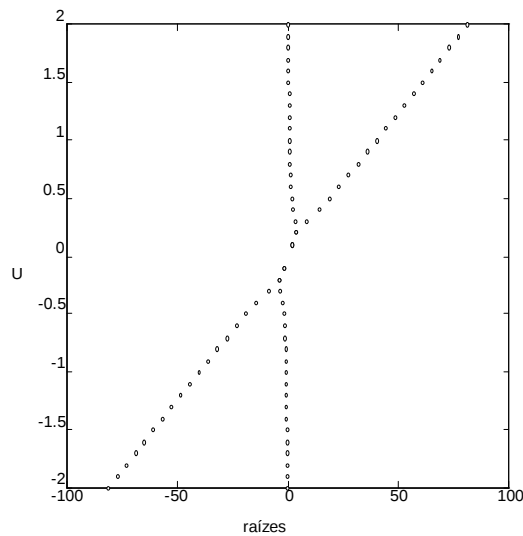


Figura 6.34 - Raízes do modelo de 17 termos truncado no oitavo termo.

Pela figura 6.34 percebe-se que tais raízes não representam a característica estática, sendo então ressaltada a importância da retirada de termos do modelo.

6.5 Comentários Finais

Neste capítulo foi visto como um modelo pode ser reduzido com base em certas observações, mesmo quando os critérios usuais de identificação não puderem fazer mais nenhuma redução no modelo. No modelo (6.1) percebeu-se a presença do agrupamento transitivo Ω_{yu^2} , bem como a sua importância no modelo.

Uma observação que pode ser feita, e que confirma a presença de termos cúbicos no modelo, é quanto ao formato da característica estática da figura 6.1. Para que uma curva do formato observado na figura 6.1 seja obtida, o modelo necessita, de termos cúbicos,

caso contrário tal informação não poderia ser recuperada. A presença ou ausência de termos quadráticos e lineares não alteram profundamente a característica estática (Jácome e Aguirre, 1996). Observou-se a presença dos termos cúbicos nos agrupamentos de termos Ω_{u^3} e Ω_{yu^2} para a equação (6.1), e no agrupamento Ω_{y^2u} do modelo da equação (6.5), e confirma assim esta idéia *a priori* da presença destes no modelo.

Outro conhecimento *a priori* que pode ser confirmado nos resultados deste capítulo, é que modelos de Hammerstein (figura 6.2), apresentam um ganho dependente do sinal de entrada, podendo obter-se um modelo, na forma $y = f(u)$ (Jácome e Aguirre, 1996), que implica em $K = g(u)$. Logo, com base nestes conhecimentos, e de acordo com a tabela II do capítulo 5, esperava-se encontrar apenas dentre os seguintes agrupamentos de termos no modelo: Ω_y , Ω_u , Ω_{u^2} , Ω_{u^3} e Ω_{yu^2} . Esta expectativa foi confirmada como pode ser visto na equação (6.1).

Analogamente, para modelos de Wiener, a saída da característica estática dependerá da saída do processo, obtendo-se $K = h(y)$. Nestas condições, com base na tabela II do capítulo 5, esperava-se encontrar apenas dentre os seguintes agrupamentos de termos no modelo: Ω_y , Ω_u , Ω_{yu} e Ω_{y^2u} . Esta expectativa foi confirmada como visto na equação (6.2).

No capítulo, percebeu-se também que a redução do tamanho do modelo não implicou na perda de qualidade do mesmo, uma vez que os modelos foram utilizados na predição e validados pelos critérios estatísticos. Nestes critérios os valores fora da faixa de confiança, devido a sua pequena amplitude podem ser desprezados.

7 Identificação de uma Válvula de Controle

7.1 Introdução

Nos capítulos anteriores foram desenvolvidos procedimentos de auxílio à detecção de estruturas em modelos polinomiais NARMAX. Estes procedimentos foram baseados inicialmente em modelos discretizados, mediante a inclusão de variações no ganho e na constante de tempo. Posteriormente, tais procedimentos foram comparados a modelos polinomiais NARMAX, e foram feitas generalizações nas estruturas de tais modelos, para variações do ganho e da constante de tempo. Os resultados foram utilizados no auxílio à identificação de dois sistemas simulados no capítulo 6.

Neste capítulo, tentar-se-á identificar um modelo polinomial NARMAX, para uma válvula de controle real, de característica de igual porcentagem, operando normalmente fechada com atuador tipo pistão, instalada numa planta de tanques interativos (Jota et al., 1995), no Laboratório de Controle de Processos Industriais da Universidade Federal de Minas Gerais. O interesse na modelagem de tal válvula provém do fato de que a presença de uma característica estática não-linear em tais válvulas é responsável pela variação do ganho da malha de controle dependendo do ponto de operação (Wigren, 1990). Serão estabelecidos quais tipos de termos que devem compor o modelo de tal válvula, e posteriormente tentar-se-á recuperar a característica estática da mesma. Esta característica é semelhante à característica estática dos sistemas do capítulo anterior.

O capítulo está organizado como segue: na seção 7.2 é feita uma análise preliminar do processo, para ilustrar o funcionamento da válvula de controle. Na seção 7.3 são sintetizados os conhecimentos *a priori* do processo, e posteriormente aplicados na

identificação de um modelo para a válvula. Obteve-se também, um outro modelo, utilizando-se os procedimentos usuais de identificação. Através do resultados desta seção, os dois métodos podem ser comparados. Finalmente, na seção 7.4, são feitos os comentários finais a respeito do capítulo.

7.2 Análise Preliminar do Processo

A válvula em consideração é acionada por um atuador pneumático. O sinal de controle enviado pelo controlador é um sinal de corrente que é transformado num sinal de pressão, pelo sistema posicionador da válvula. O elemento atuador é do tipo pistão. Basicamente, o sinal de pressão aplicado no elemento atuador, movimenta o pistão, que determina a vazão de saída da válvula.

O tempo gasto entre uma variação do sinal de corrente, e o movimento resultante da válvula caracteriza o tempo morto do conjunto. As constantes de tempo da válvula variam com o seu tamanho e com a velocidade de atuação (McMillan, 1994; Shinskey, 1988). No contexto de controle de processos, o valor da constante de tempo dominante da válvula, dentro de uma malha de controle pode representar a própria constante de tempo da malha ou simplesmente um tempo morto, dependendo das características dinâmicas da malha (McMillan, 1994).

O ganho de uma válvula pode ser definido como uma variação na vazão através desta, pela variação da posição do atuador, definidos pelo sinal aplicado na entrada (Shinskey, 1988).

Antes de se coletar os dados de qualquer processo a ser identificado, sempre que possível deve-se acumular o máximo de conhecimento *a priori* do sistema. No caso da

válvula, os sinais de entrada e saída são respectivamente, o sinal de controle enviado ao atuador e a vazão na saída. Cabe ressaltar que as características dinâmicas do medidor de vazão estarão incluídas nos dados, mas elas são, idealmente, pouco relevantes em comparação com a dinâmica da válvula.

Através de amostras do sinal de controle para o atuador, e da vazão medida na saída da válvula, encontrou-se a característica estática da figura 7.1a.

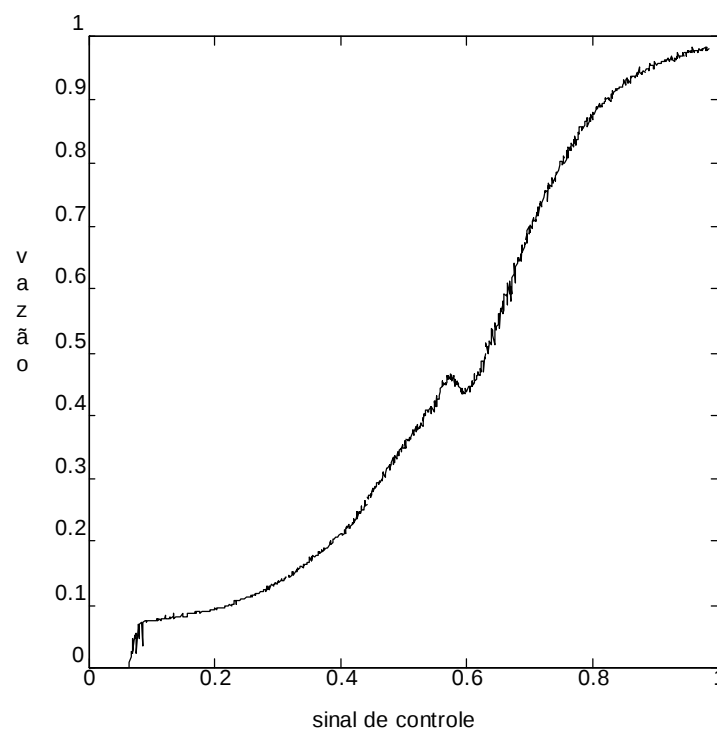


Figura 7.1a - Característica estática do conjunto válvula/medidor.

Pela figura 7.1a, percebe-se que a vazão de saída da válvula, dependente do sinal de controle na entrada do atuador. Isto mostra que o ganho estático do sistema varia, dependendo do sinal de controle e consequentemente, da vazão da saída.

Uma motivação para identificar um modelo dinâmico para a válvula, a partir do qual poderia se recuperar sua característica estática, baseia-se no fato de que uma válvula

ajustada para operar em um ponto da curva da figura 7.1a provavelmente não estaria bem ajustada para outro ponto de operação, trazendo problemas para o controle.

Cabe ressaltar, que a característica estática da figura 7.1a, pode ser estimada também através dos dados de identificação diretamente. O resultado deste procedimento é mostrado na figura 7.1b.

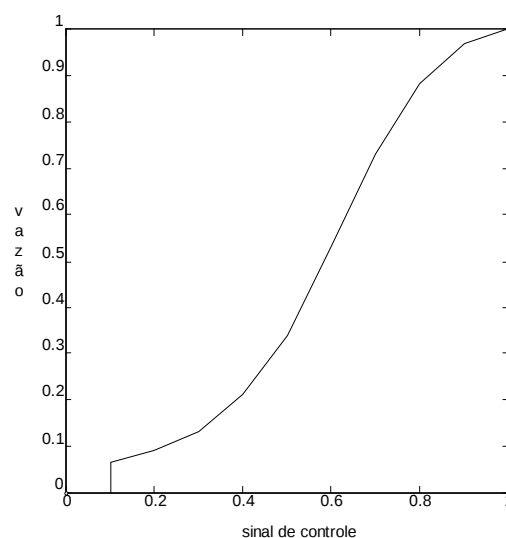


Figura 7.1b - Característica estática da válvula obtida a partir dos dados de identificação.

7.3 Identificação da Válvula de Controle

7.3.1 Informações *a priori*

As constantes de tempo da válvula variam com o seu tamanho e com a velocidade de atuação. Como o tamanho da válvula não varia, supõe-se que uma constante de tempo seja fixa. Entretanto, as demais constantes de tempo variam com a velocidade de atuação, que por sua vez dependem da amplitude do sinal de entrada (McMillan, 1994). Isto pode ser verificado através dos próprios dados de identificação, mostrados

posteriormente na figura 7.4, ou seja, a válvula apresenta constantes de tempo diferentes para cada amplitude do sinal de entrada.

Tomando-se o valor aproximado das constantes de tempo, para cada ponto de operação da válvula, encontra-se a curva da figura 7.2.

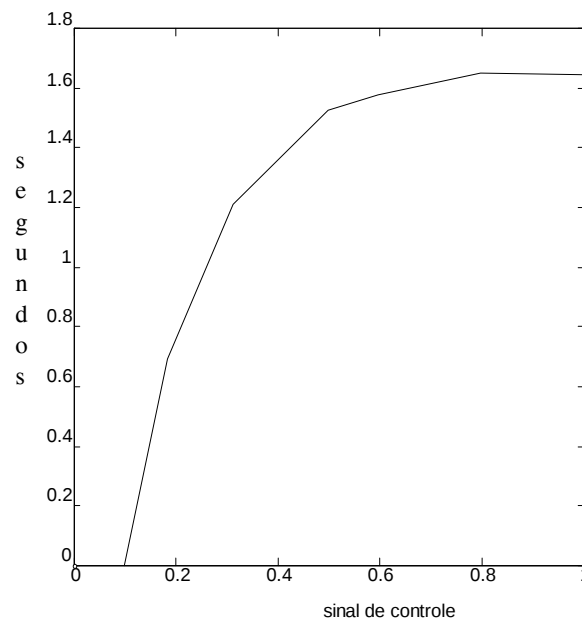


Figura 7.2 - Curva da constante de tempo variável em função sinal de controle.

O ganho da válvula depende da posição do atuador, que determina uma vazão. Por sua vez, a posição do atuador está diretamente relacionada ao sinal de entrada aplicado na válvula, o que provoca uma variação do ganho com a entrada. O ganho da válvula é dado pela inclinação da reta tangente à característica estática, e depende do ponto de operação, de acordo com a figura 7.3. Como esta inclinação varia ao longo da curva, o ganho também varia, conforme esperado.

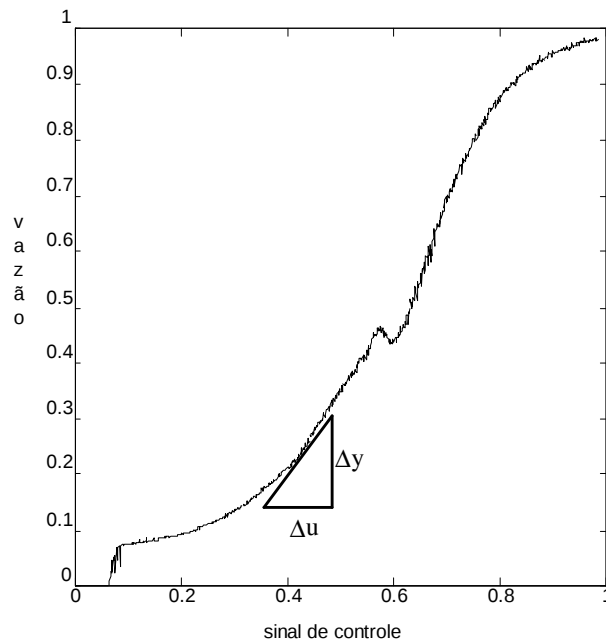


Figura 7.3 - Variação do ganho da válvula.

Pelas considerações anteriores, percebe-se que o ganho e as constantes de tempo têm uma relação de dependência final com o sinal de controle aplicado na válvula, indicando então que estas variáveis do processo podem estar relacionadas entre si, este fato será discutido posteriormente.

7.3.2 Identificação Baseada em Conhecimento *a priori* da Válvula

Os dados de identificação foram amostrados com um intervalo de 0,2 segundos, e são mostrados na figura 7.4. Os dados de entrada representam o sinal de controle normalizado que é enviado para o atuador, e os dados de saída correspondem à medição de vazão de saída da válvula, como já dito anteriormente. Procurou-se manter o sistema sob constante excitação, dentro do possível, de modo que sua dinâmica estivesse bem representada nos dados.

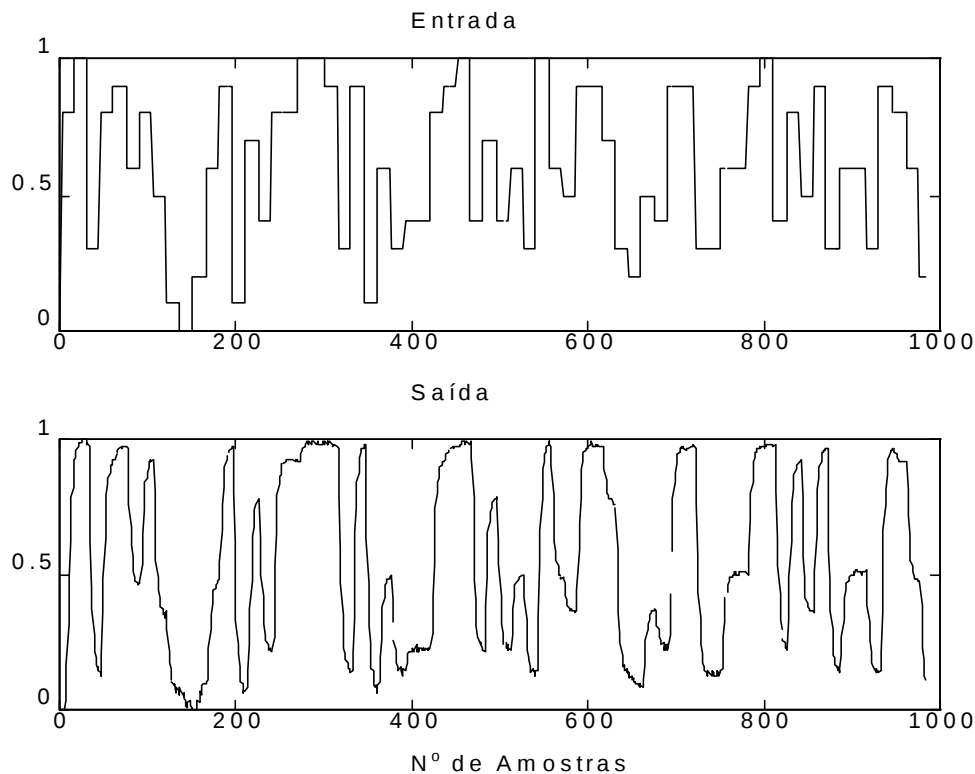


Figura 7.4 - Dados de identificação.

A aplicação de testes de não-linearidade (Haber, 1985) nos dados indicou a necessidade do uso de um modelo não-linear. Logo, o modelo polinomial NARMAX é um candidato a ser utilizado neste problema. A ordem máxima para os modelos foi $n = 3$. Esta escolha é empírica, como normalmente se faz para sistemas não-lineares, uma vez que para estes sistemas um método de escolha da ordem não foi ainda formalmente estabelecido. O procedimento de identificação desenvolvido a seguir será baseado em informações *a priori* do sistema. Será feita inicialmente uma análise para o ganho, desconsiderando a variação da constante de tempo. Em seguida analisar-se-á a variação da constante de tempo.

Com relação ao ganho do sistema, tem-se a seguinte informação *a priori*:

1. A característica estática da válvula é conforme mostrado no gráfico 7.1a.

2. O ganho da válvula é variável em função de u (sinal de controle);

Como dito no capítulo 6, sistemas que apresentam característica estática semelhante à característica da figura 7.1a, apresentam o ganho variável, conforme pode ser visto na figura 7.3. Para que tal informação possa ser reproduzida no modelo deste sistema, necessita-se de termos cúbicos neste. A necessidade de tais termos deve-se ao ponto de inflexão existente na curva da característica estática, sugerindo assim a necessidade de um polinômio de ordem cúbica. Logo, de acordo com a tabela II do capítulo 5, o agrupamento de termos relativo à variação do ganho, $\Omega_{x_k u}$, deve ser de grau cúbico.

Se o ganho da válvula varia com o sinal de controle, pode-se obter uma relação estática na forma $K_{re} = v(u)$, o que implica em $X_k = g(u)$. A partir da equação (5.6) vê-se que o grau de não-linearidade de X_k é uma unidade menor que o de $\Omega_{x_k u}$. Portanto, como $v(u)$ é uma função cúbica, tem-se que $X_k = U^2$. Note-se que isto não impede que X_k dependa também de termos de grau inferior a dois, ou seja, $X_k = f(c_0, U, U^2)$, onde c_0 é uma constante, seria também aceitável.

Portanto, pode-se dizer que o agrupamento Ω_{u^3} deve fazer parte do modelo da válvula, para uma variação do ganho. Os agrupamentos Ω_u e Ω_{u^2} podem também integrar o modelo, mas não são indispensáveis.

Cabe ressaltar que a característica estática da figura 7.2, não parte da origem. Logo, também é necessária a presença do termo constante, Ω_0 , no modelo da válvula.

Para a constante de tempo variável, a seguinte informação *a priori* está disponível:

1. A constante de tempo varia com o sinal de controle conforme mostrado na figura 7.2

Pela figura 7.2 percebe-se que a curva de variação da constante de tempo é semelhante a um braço de parábola, o que indica a necessidade de termos quadráticos para representar a variação da constante de tempo. Logo, de acordo com a equação (5.5), tem-se que X_τ deve ser um termo quadrático. Isto implica que o agrupamento de termos $\Omega_{x_\tau y}$, deve ser de grau cúbico.

A constante de tempo varia com o sinal de controle, ou seja, $\tau = h(u)$. Para que a equação (5.5) represente uma curva como a da figura 7.2, tem-se que $X_\tau = U^2$. Note-se que isto não impede que X_τ possua termos de grau inferior a dois, ou seja, $X_\tau = q(c_0, U, U^2)$, seria também aceitável.

Portanto, com base nos possíveis valores de X_τ , pode-se dizer que o agrupamento Ω_{yu^2} , deve fazer parte do modelo da válvula, na representação da variação da constante de tempo. Os demais agrupamentos, originados a partir do agrupamento $\Omega_{x_\tau y}$, como Ω_y e Ω_{yu} , podem estar presentes no modelo final da válvula, não sendo indispensáveis.

Portanto, com base no desenvolvimento anterior, pode-se limitar o conjunto de termos candidatos ao modelo da válvula aos seguintes agrupamentos: Ω_0 , Ω_y , Ω_u , Ω_{u^3} , Ω_{u^2} , Ω_{yu^2} e Ω_{yu} . Então, escolhidos os agrupamentos considerados efetivos, utilizaram-se as técnicas de agrupamentos de termos e coeficientes de agrupamentos, para auxiliar na escolha de quais destes agrupamentos constituirão o modelo definitivo da válvula. Para isso estimou-se um conjunto de modelos variando-se o número de termos de 2 a 30, e mantendo-se 22 termos de ruído a fim de evitar a polarização das estimativas, como dito no capítulo anterior. Os critérios estatísticos de informação também foram utilizados no

auxílio à escolha de um número adequado de termos para o modelo. Os resultados são apresentados nas figuras 7.5 a 7.7.

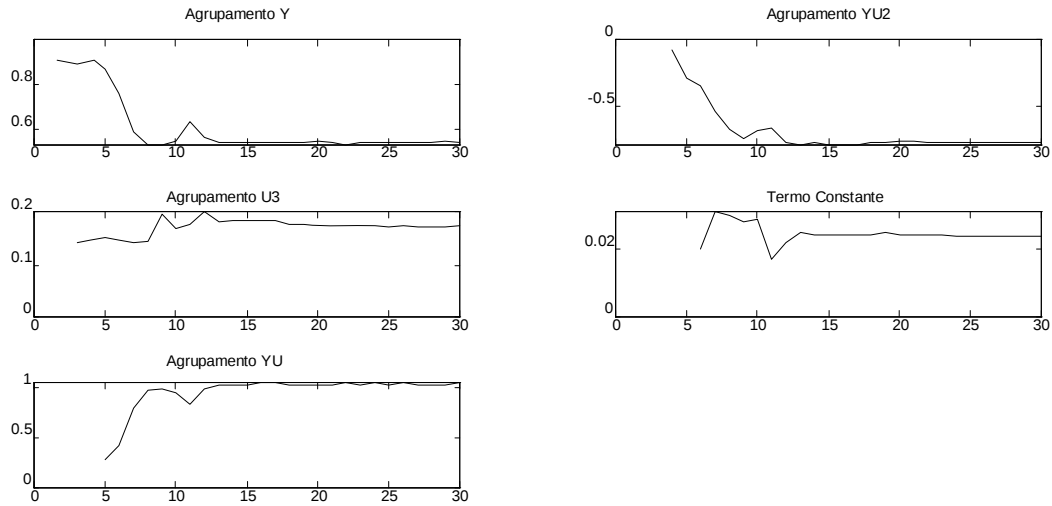


Figura 7.5 - Coeficientes de agrupamentos em função do número de termos no modelo.

A análise de agrupamentos de termos feita inicialmente, indicou que Ω_u e Ω_{u^2} poderiam ser descartados do modelo.

As figuras 7.6 e 7.7 mostram os critérios de auxílio na escolha do número de termos adequado para o modelo.

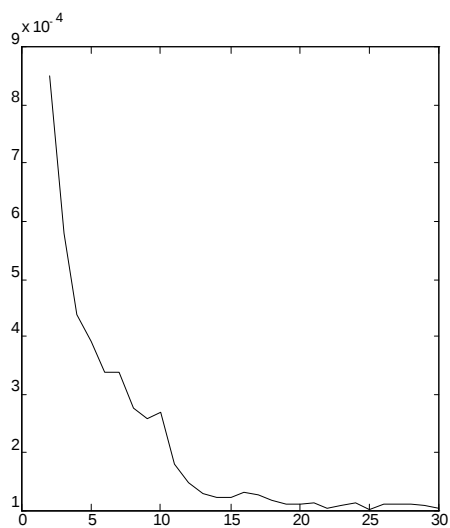


Figura 7.6 - Variância dos resíduos.

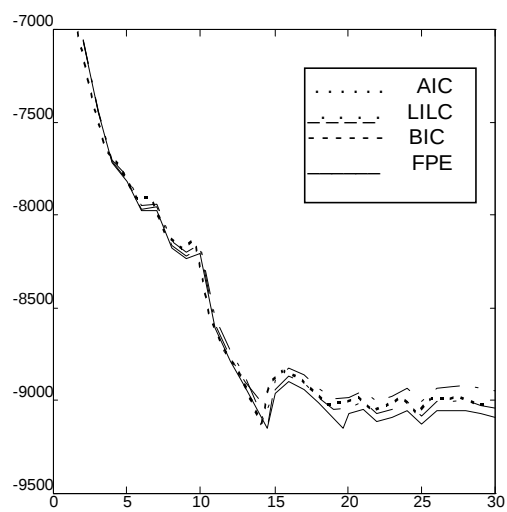


Figura 7.7 - Critérios de informação.

Através das figuras 7.6 e 7.7 percebe-se que tais critérios indicam um modelo de 15 termos. Pelos coeficientes de agrupamentos da figura 7.5 também percebe-se que a partir de 15 termos no modelo, os coeficientes tendem a permanecer em torno de um mesmo valor. Com isto escolheu-se um modelo com 15 termos, que é mostrado a seguir:

$$\begin{aligned} y(k) = & + 0,02428 - 0,20522 y(k-1) + 1,08220 y(k-2) - 0,02922 u(k-2)u(k-2)u(k-2) \\ & - 3,80950 u(k-3)u(k-3)y(k-1) + 6,46510 u(k-3)y(k-1) - \\ & - 5,57840 u(k-3)y(k-2) + 4,08710 u(k-3)u(k-3)y(k-2) - \\ & - 0,35280 u(k-3)u(k-3)u(k-2) + 2,83290 u(k-2)y(k-3) - \\ & - 1,05690 u(k-3)u(k-2)y(k-3) - 0,33528 y(k-3) + \\ & + 0,56637 u(k-3)u(k-3)u(k-3) - 1,37440 u(k-3)y(k-3) - \\ & - 1,31640 u(k-2)y(k-1) \end{aligned} \quad (7.1)$$

A figura 7. 8 mostra o modelo com 15 termos e os *dados de identificação*. Na figura 7.9 são apresentados os testes de validação estatística.

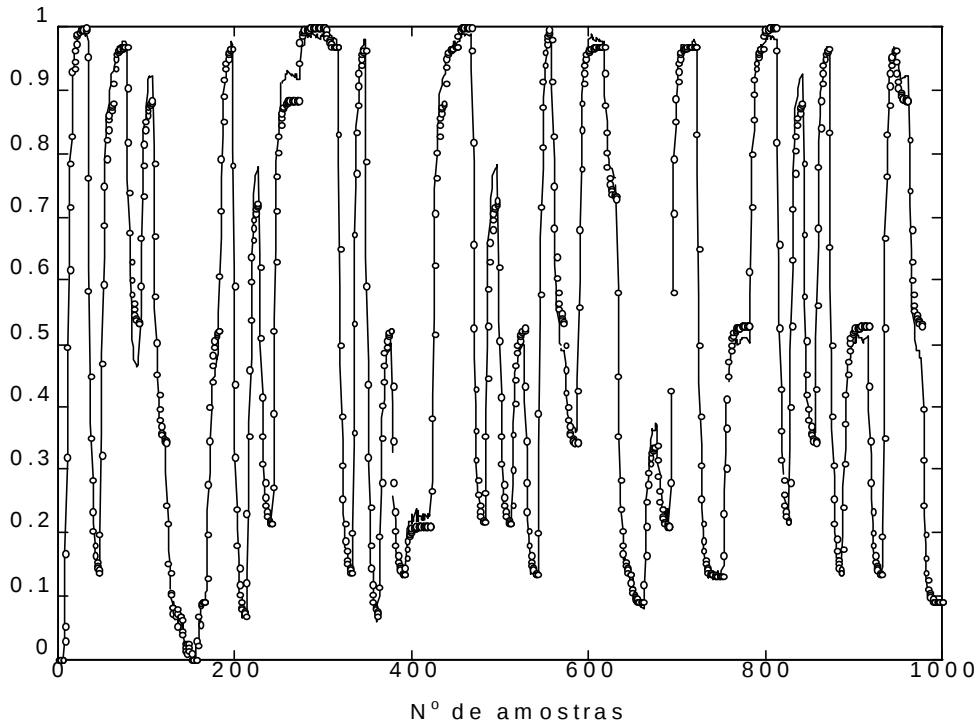


Figura 7.8 - (—) Dados reais de identificação e (. . .) modelo identificado.

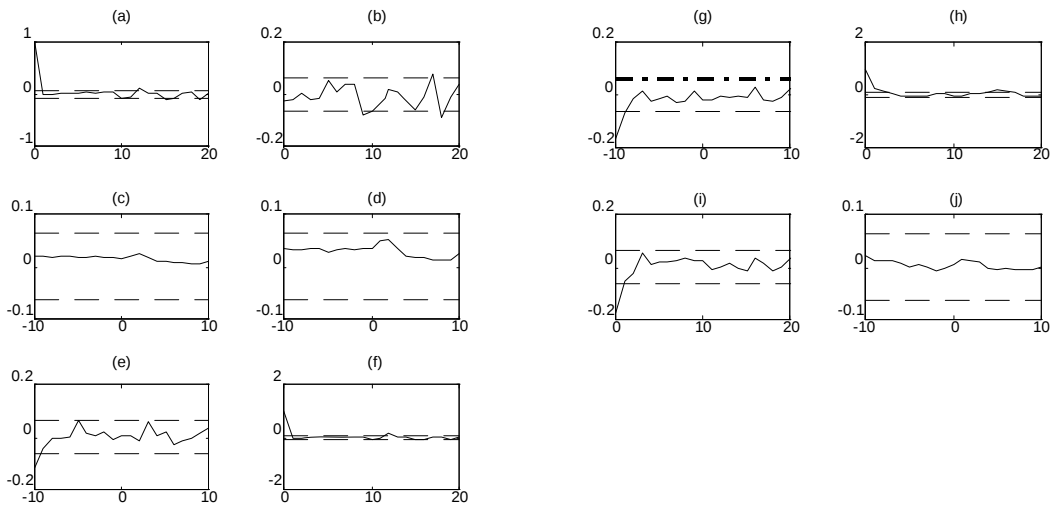


Figura 7.9 - Testes de validação estatística para o modelo final. Correlações calculadas:
 (a) $\xi(t) \times \xi(t)$ (b) $u(t) \times \xi(t)$ (c) $\xi(t) \times \xi(t) \times u(t)$ (d) $u^2(t) \times \xi(t)$ (e) $u^2(t) \times \xi^2(t)$;
 (f) $\xi'(t) \times \xi'(t)$ (g) $\xi'(t) \times \xi^2(t)$ (h) $\xi^2(t) \times \xi^2(t)$ (i) $(y(t) \times e(t))' \times \xi^2(t)$;
 (j) $(y(t) \times e(t))' \times u^2(t)$

A figura 7.10 apresenta o modelo com 15 termos utilizado na predição de uma *massa de dados diferente dos dados de identificação*.

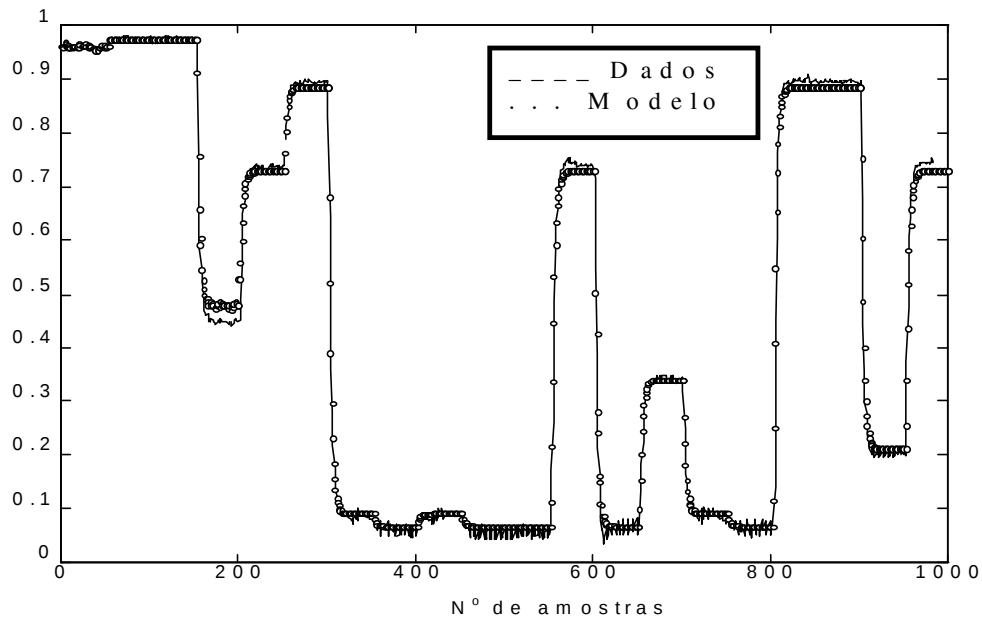


Figura 7.10 - predição do modelo de 15 termos.

Pelas figuras 7.24 e 7.25, do ponto de vista visual, percebe-se que o modelo com 15 termos é um modelo satisfatório. Os critérios estatísticos de validação estão bem dentro da faixa de confiança. Resta saber se através de tal modelo pode-se recuperar a característica estática da válvula, esta tentativa será feita a seguir.

Tomando-se a equação (7.1) e considerando regime estático tem-se:

$$0,4583Y = 0,02428 + 0,1842 U^3 + 1,029YU - 0,7792 YU^2$$

onde os coeficientes de agrupamentos presentes são:

$$1 - \Sigma_y = 0,4583, \Sigma_0 = 0,02428, \Sigma_{u^3} = 0,1842, \Sigma_{yu} = 1,029 \text{ e } \Sigma_{yu^2} = -0,7792.$$

A equação anterior pode ser reescrita, isolando-se o termo em Y do lado esquerdo desta equação, da seguinte forma:

$$Y = \frac{(0,02428 + 0,1842U^3)}{(0,4583 - 1,029U + 0,7792U^2)} \quad (7.2)$$

A equação (7.2) encontra-se na forma $y = f(u)$, de acordo com as suposições iniciais. Assim, pode-se recuperar a característica estática da válvula de uma maneira analítica direta, ao contrário da obtenção através de simulação ou pelas raízes. Portanto, substituindo valores da faixa de entrada, com $0 \leq U \leq 1$, tem-se a característica estática obtida através do modelo. Esta característica é mostrada na figura 7.11, sobreposta à característica estática medida da válvula.

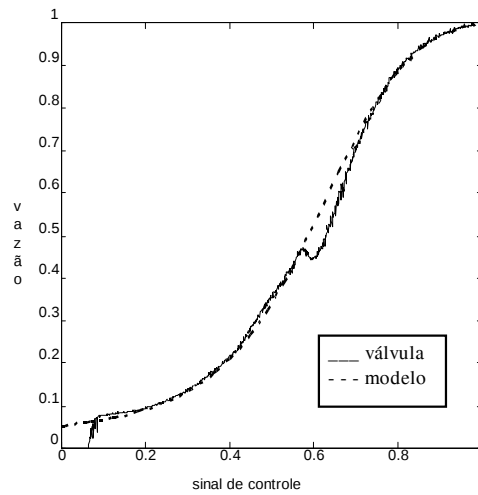


Figura 7.11 - Características estática real e característica estática recuperada do modelo.

A seguir encontrar-se-á a relação estática saída-entrada. Partindo-se da equação (7.2), e dividindo-se ambos os lados por U tem-se:

$$\frac{Y}{U} = \frac{0,02428 + 0,1842U^2}{(0,4583 - 1,029U + 0,7792U^2)U}$$

A equação anterior pode ser reescrita da seguinte maneira:

$$K_{re} = \frac{Y}{U} = \frac{0,02428}{0,4583U - 1,029U^2 + 0,7792U^3} + \frac{0,1842U^2}{0,4583 - 1,029U + 0,7792U^2} \quad (7.3)$$

A equação (7.3) representa a variação da relação estática, ao longo da característica da figura 7.1a.

Cabe ressaltar que, da equação (7.2), tem-se que $\Omega_{x_k u} = \Omega_{u^3}$ relativo à variação do ganho, e $\Omega_{x_{\tau} y} = \Omega_{y u^2} \cup \Omega_{y u}$ relativamente à variação da constante de tempo. Logo, substituindo-se os valores dos coeficientes de agrupamento na equação (5.6), com $\Sigma_u = 0$, obtém-se a segunda parcela da equação (7.3). Note-se que a equação (5.6) não considera o termo constante no numerador, correspondente à primeira parcela da equação (7.3), por isto K_{re} não pode ser totalmente recuperada pela equação (5.6).

A curva $K_{re} \times u$, obtida a partir da equação (7.3), é apresentada na figura 7.12.

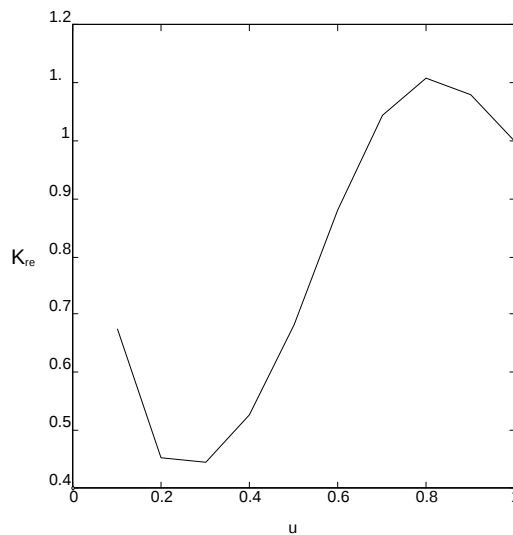


Figura 7.12 - Relação estática $\times$ sinal de controle.

Pela equação (7.3), percebe-se que K_{re} e a constante de tempo variável estão diretamente relacionados, atestando que a variação do ganho é influenciada pela constante de tempo, pelo fato de ambas variarem com o sinal de controle na entrada. Nestas condições, tem-se que $\Sigma_{x,y} \neq 0$, como dito nos comentários finais do capítulo 5.

Uma análise do funcionamento do sistema real permite verificar que realmente, o ganho e a constante de tempo estão relacionadas. Isto porque a aplicação de um sinal de controle na válvula implica na modificação do ponto de operação, deslocando o sistema ao longo da característica estática da figura 7.1a. Entretanto, a mudança do ponto de operação do sistema, demora um tempo determinado pela constante de tempo deste, ou seja, do tempo em que a válvula se movimenta da posição atual até a posição final. O tempo de deslocamento da válvula, por sua vez, depende da amplitude do sinal de controle aplicado, estabelecendo-se assim a relação de dependência entre o ganho e a constante de tempo variável.

Cabe ressaltar que não foi possível recuperar a equação para a variação da constante de tempo, pelo fato do modelo obtido não ser racional, de acordo com o lema 4.1.

7.3.3 Identificação do Tipo “caixa-preta”

Nesta seção serão mostrados os resultados do processo de identificação “caixa-preta” da válvula de controle. Ou seja, a identificação será realizada sem a utilização de qualquer conhecimento a respeito do sistema, sendo baseada puramente nas técnicas de agrupamentos de termos e coeficientes de agrupamentos, com um procedimento de identificação semelhante ao do capítulo 6.

Os dados utilizados nesta seção, tanto de identificação, quanto de validação, são os mesmos da seção anterior.

Portanto, estimou-se um conjunto de modelos variando-se o número de termos de 2 a 40, e mantendo-se 22 termos de ruído a fim de evitar a polarização das estimativas, como dito no capítulo anterior. Para cada modelo calculou-se os valores coeficientes de agrupamentos e agrupamentos de termos presentes. Os critérios estatísticos de informação também foram utilizados no auxílio à escolha do número de termos adequados para o modelo. Os resultados são apresentados nas figuras 7.13 a 7.15.

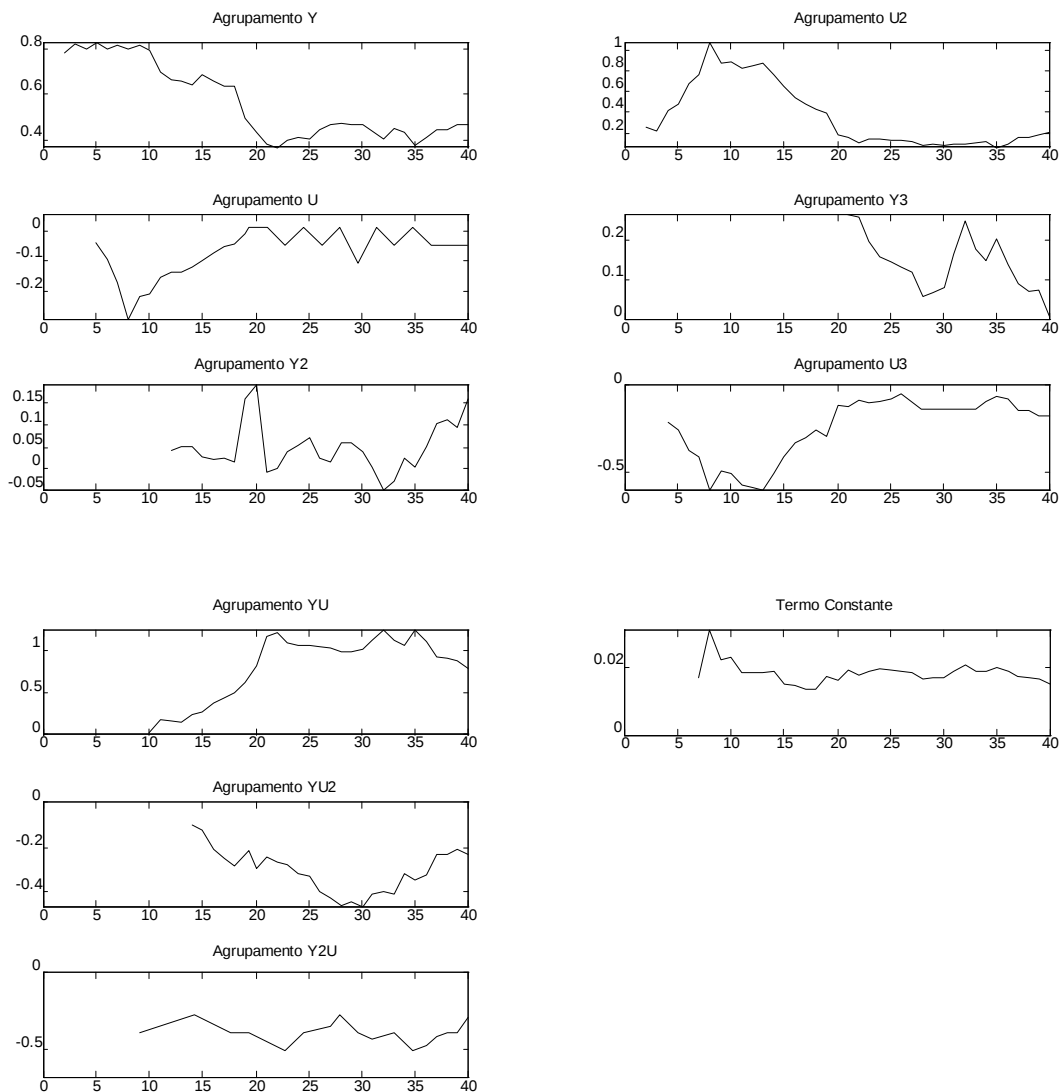


Figura 7.13 - Coeficientes de agrupamentos em função do número de termos no modelo.

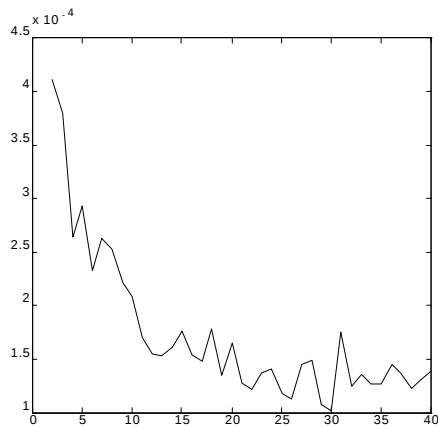


Figura 7.14 - Variância dos resíduos.

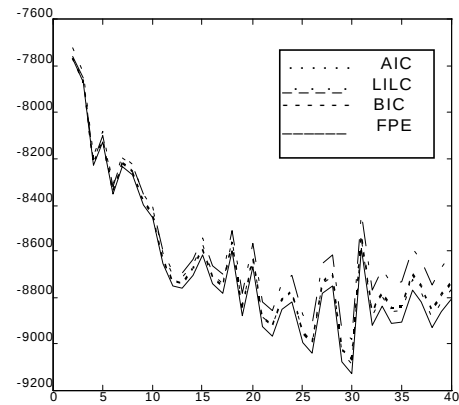


Figura 7.15 - Critérios de informação.

Pelas figuras 7.14 e 7.15, percebe-se que os critérios estatísticos sugerem um modelo com 30 termos de processo. A figura 7.16 apresenta os dados de identificação e o modelo identificado com 30 termos. Na figura 7.17 são mostrados os testes de validação estatística de tal modelo.

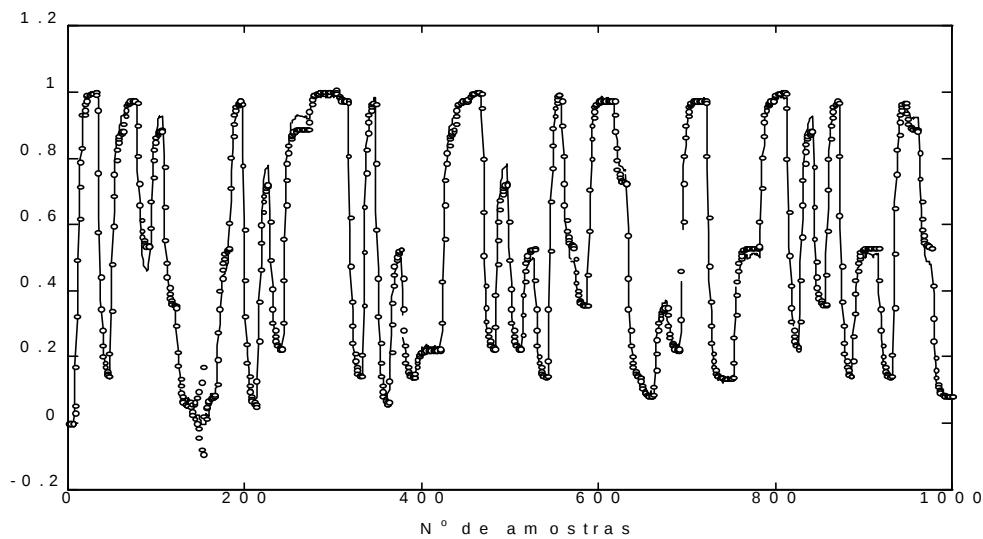


Figura 7.16 - (____) Dados de identificação e (.....) modelo identificado.

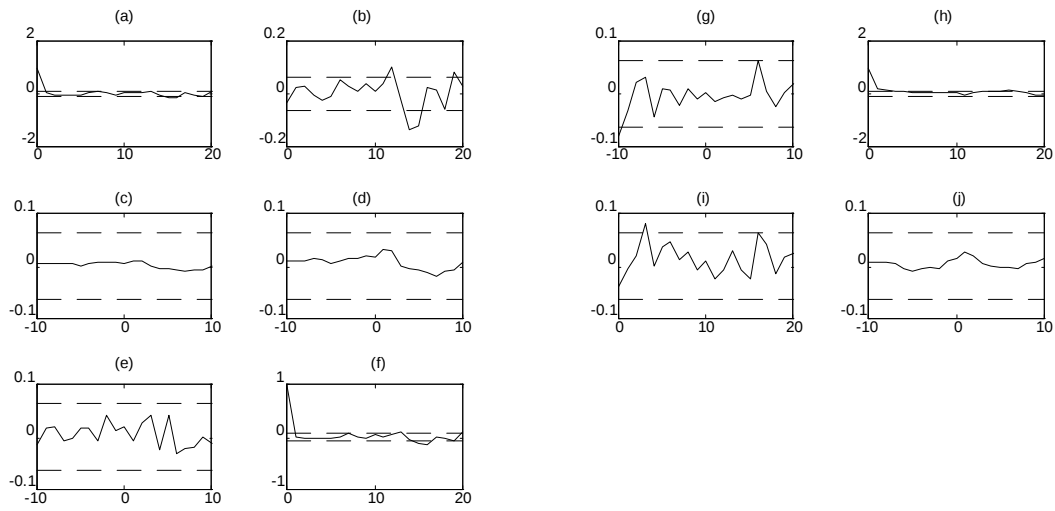


Figura 7.17 - Testes de validação estatística. Correlações calculadas: (a) $\xi(t) \times \xi(t)$; (b) $u(t) \times \xi(t)$; (c) $\xi(t) \times \xi(t) \times u(t)$; (d) $u^2(t) \times \xi(t)$; (e) $u^2(t) \times \xi^2(t)$; (f) $\xi'(t) \times \xi'(t)$; (g) $\xi'(t) \times \xi^2(t)$; (h) $\xi^2(t) \times \xi^2(t)$; (i) $(y(t) \times e(t))' \times \xi^2(t)$; (j) $(y(t) \times e(t))' \times u^2(t)$.

A figura 7.18 mostra o modelo com 30 termos utilizado na predição de uma massa de dados *diferente* dos dados de identificação

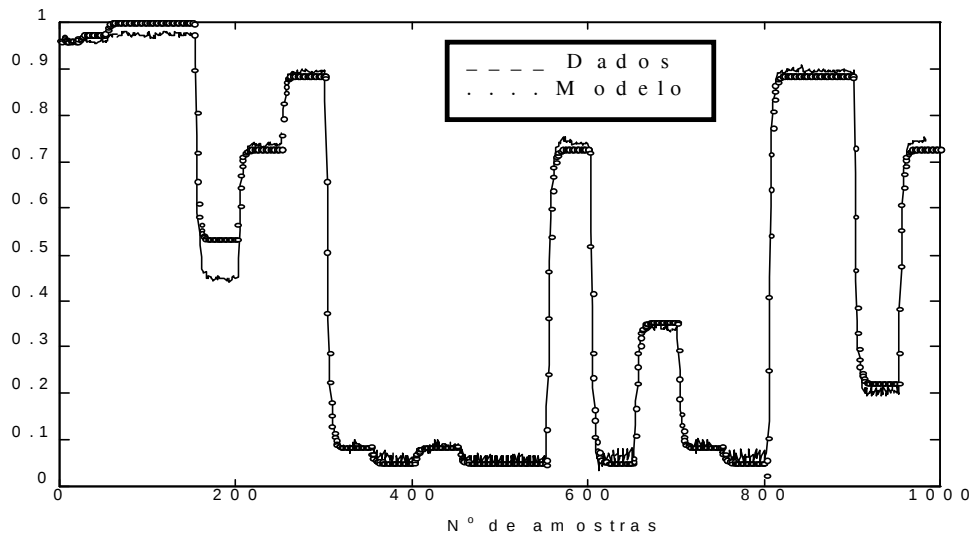


Figura 7.18 -Predição do Modelo com 30 termos.

Observando-se a figura 7.18, nota-se que o modelo obtido com 30 termos em certos pontos oscila mais que o sistema real, evidenciando uma possível sobreparametrização

do modelo. Além disso alguns testes de validação estatística da figura 7.17, estão fora da faixa.

Pelos agrupamentos presentes no modelo, de acordo com a figura 7.13, percebe-se que a característica estática só poderá ser recuperada via simulação do modelo, ou cálculo das raízes ou pontos fixos (ver seção 6.4), isto porque não é possível obter-se uma equação na forma $y = f(u)$ que permita uma recuperação analítica direta da mesma. Nas figuras 7.19, 7.20a e 7.20b são apresentados resultados da característica estática do modelo com 30 termos obtida via simulação e por raízes, respectivamente.

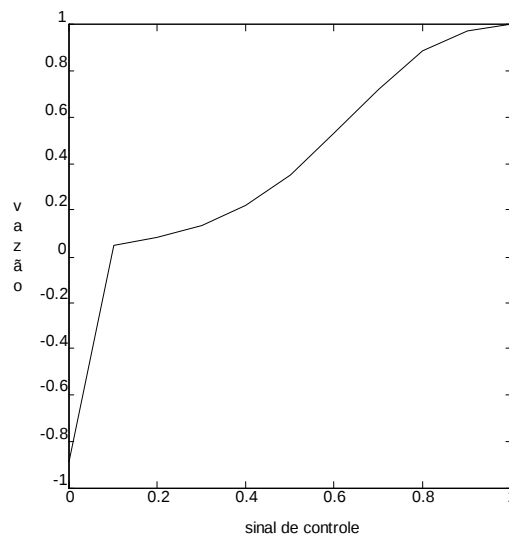


Figura 7.19 - Característica estática obtida via simulação.

As raízes do modelo com 30 termos são mostradas na figura 7.20a, e a raiz estável é mostrada na figura 7.20b, abaixo.

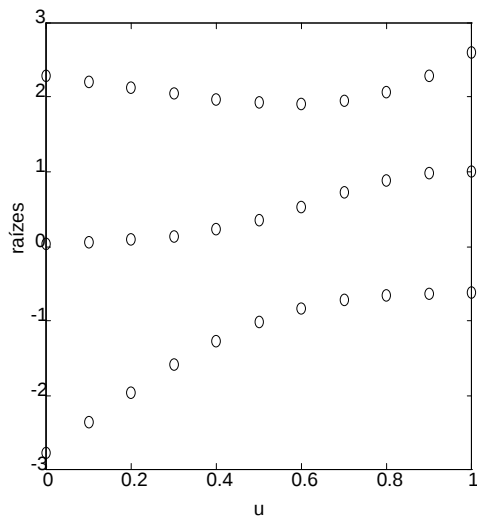


Figura 7.20a - Raízes do modelo (30 termos)

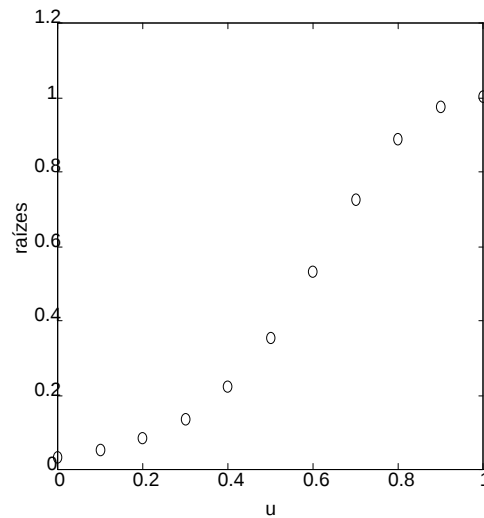


Figura 7.20b - Raíz estável do modelo

Pela figura 7.19, percebe-se que o modelo com 30 termos contém alguma informação a respeito da característica estática. Cabe ressaltar que características estáticas recuperadas através de simulação do modelo obtido seria atualmente inviável para aplicações on-line, devido ao dispêndio de grande tempo computacional, o que tornaria esta informação praticamente inacessível para o controle. Através da figura 7.20b percebe-se que a característica estática da válvula só pode ser recuperada mediante o cálculo das raízes (ou pontos fixos).

A seguir tentar-se-á melhorar o modelo, de forma que permaneçam somente os agrupamentos de termos efetivos, e que eventualmente possibilitem uma representação direta da característica estática.

Analisando-se a variação dos coeficientes de agrupamentos da figura 7.4, percebe-se que o agrupamento Ω_{y^3} passa a fazer parte do modelo, somente quando este apresenta 20 termos. Além do mais, tal coeficiente apresenta-se oscilando continuamente. Estas características sugerem que este termo é espúrio, podendo-se retirá-lo então, do conjunto de termos candidatos ao modelo.

O coeficiente Σ_{y^2} além de oscilar muito, apresenta algumas trocas de sinais, evidenciados pelas oscilações em torno do zero, podendo também ser retirado do conjunto de termos candidatos ao modelo.

O coeficiente do agrupamento Ω_{u^2} percorre uma faixa muito grande de valores, em comparação aos demais coeficientes, tendendo posteriormente a um valor próximo de zero, sendo também descartado. O coeficiente do agrupamento Ω_u , oscila constantemente em torno do zero, também sugerindo que tais termos sejam retirados do conjunto de termos candidatos ao modelo. Os demais agrupamentos de termos serão mantidos para a próxima etapa da identificação.

Retirados os agrupamentos considerados não necessários, reestimou-se outro conjunto de modelos, variando-se desta vez o número de termos de 2 a 30. Esta redução deve-se à redução do conjunto de termos candidatos ao modelo. A variação dos coeficientes de agrupamentos e os critérios de informação são mostrados nas figuras 7.21 a 7.23.

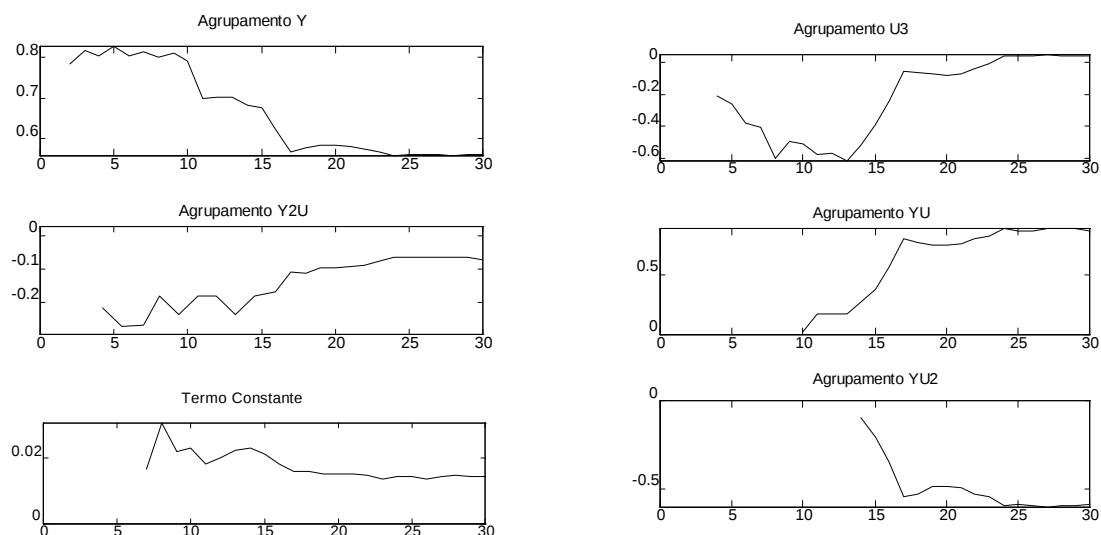


Figura 7.21 - Coeficientes de agrupamentos em função do número de termos no modelo.

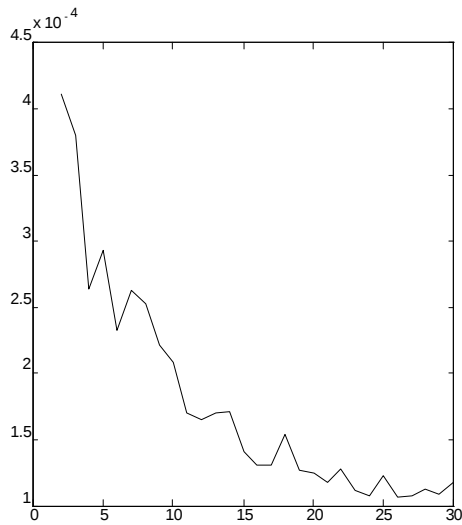


Figura 7.22 - Variância dos resíduos.

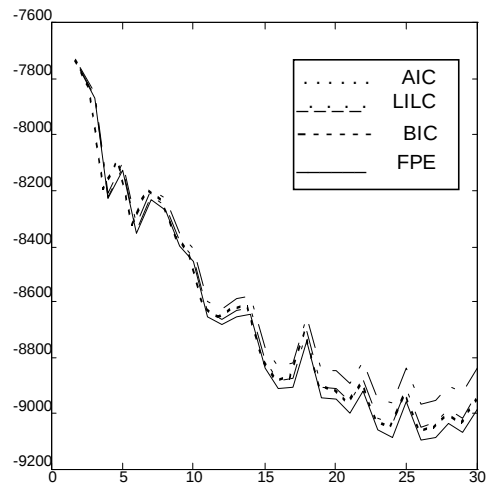


Figura 7.23 - Critérios de informação

Os critérios das figuras 7.22 e 7.23 sugerem um modelo com 24 termos, que é mostrado na equação 7.4:

$$\begin{aligned}
 y(k) = & 0.02295 - 0.2574 y(k-1) + 1.18 y(k-2) - 0.4231 u(k-2)u(k-2)u(k-2) + \\
 & +0.5209 u(k-3)y(k-1)y(k-1) + 1.618 u(k-2)y(k-3) - 1.302 u(k-2)u(k-2)y(k-3) + \\
 & +3.108 u(k-2)y(k-1) - 0.4876 u(k-2)u(k-2)y(k-2) - 0.5729 u(k-2)y(k-1)y(k-1) - \\
 & -3.775 u(k-3)u(k-2)y(k-1) + 2.472 u(k-3)u(k-3)u(k-3) - 2.676 u(k-3)y(k-2) + \\
 & +0.3763 u(k-3)u(k-3)y(k-2) + 2.185 u(k-3)y(k-1) + 2.656 u(k-3)u(k-2)u(k-2) - \\
 & -0.381 y(k-3) - 2.386 u(k-3)u(k-3)y(k-3) + 0.009485 u(k-1)u(k-1)u(k-1) - \\
 & -4.561 u(k-3)u(k-3)u(k-2) + 4.297 u(k-3)u(k-2)y(k-2) - 3.191 u(k-2)y(k-2) + \\
 & +2.576 u(k-3)u(k-2)y(k-3) - 0.007693 u(k-1)y(k-3)y(k-2)
 \end{aligned} \tag{7.4}$$

Os resultados da identificação do modelo da equação (7.4), bem como a validação estatística de tal modelo, são mostrados na figura 7.24 e 7.25 respectivamente.

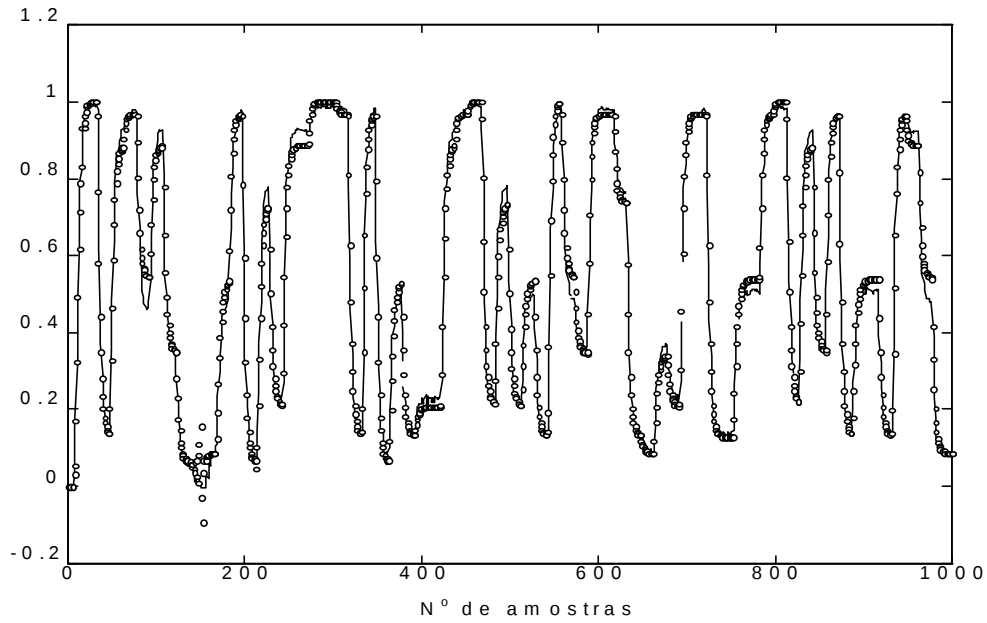


Figura 7.24 - (____) Dados de identificação e (. . . .) predição do modelo.

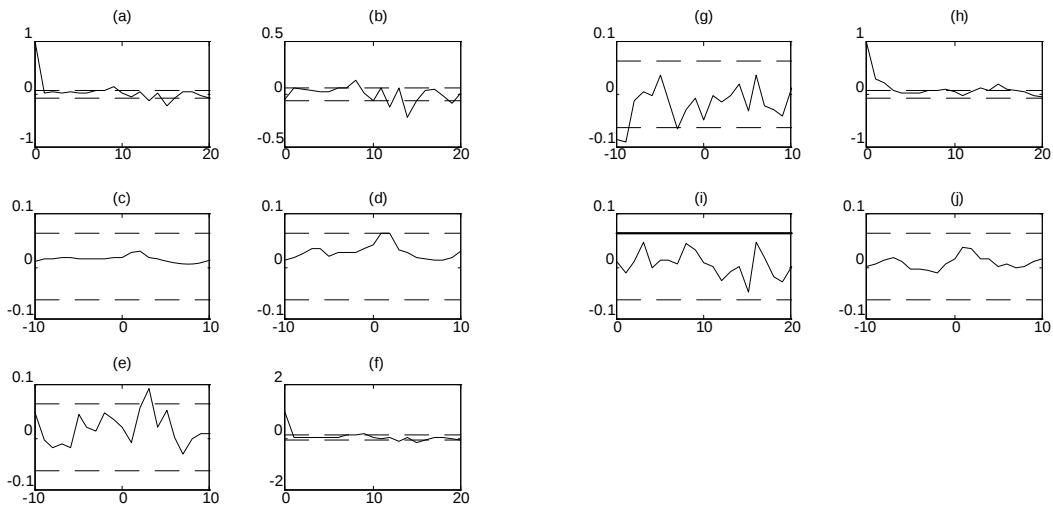


Figura 7.25 - Validação estatística do modelo com 24 termos. Correlações calculadas:
 (a) $\xi(t) \times \xi(t)$; (b) $u(t) \times \xi(t)$; (c) $\xi(t) \times \xi(t) \times u(t)$; (d) $u^{2\cdot}(t) \times \xi(t)$; (e) $u^{2\cdot}(t) \times \xi^2(t)$;
 (f) $\xi'(t) \times \xi'(t)$; (g) $\xi'(t) \times \xi^{2\cdot}(t)$; (h) $\xi^{2\cdot}(t) \times \xi^{2\cdot}(t)$; (i) $(y(t) \times e(t))' \times \xi^{2\cdot}(t)$;
 (j) $(y(t) \times e(t))' \times u^{2\cdot}(t)$.

O modelo de 24 termos utilizado na predição é mostrado na figura 7.26 abaixo.

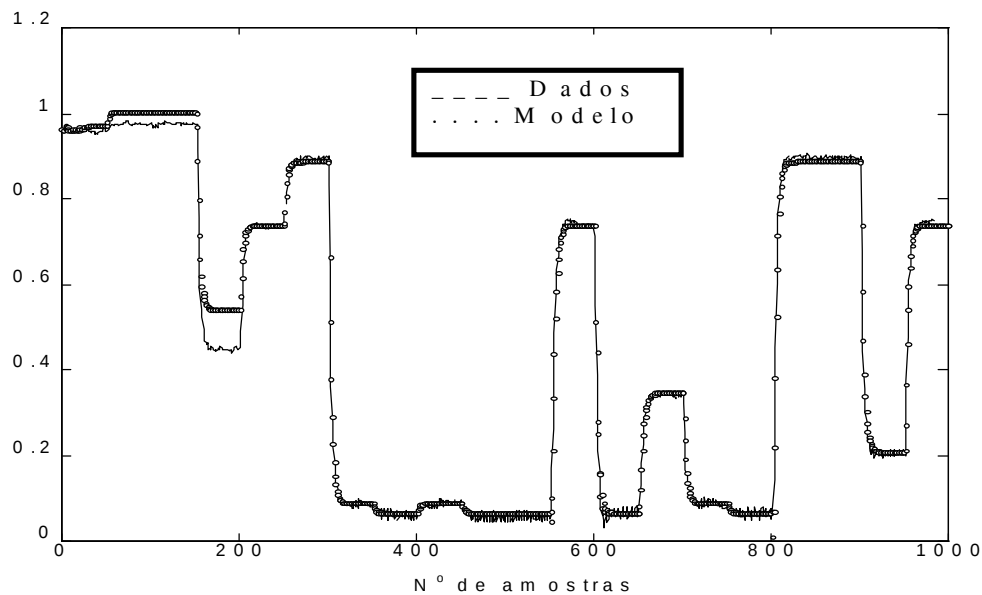


Figura 7.26 - Predição do modelo de 24 termos.

Como no caso do modelo de 30 termos, a característica estática do modelo de 24 termos só pode ser recuperada mediante simulação ou análise das raízes. Estes resultados são mostrados nas figuras 7.27, 7.28a e 7.28b respectivamente.

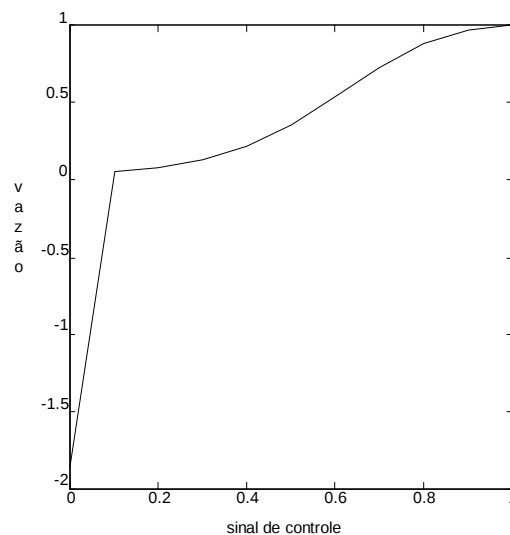


Figura 7.27 - Característica estática do modelo de 24 termos obtida via simulação.

As raízes do modelo da equação (7.4) são mostradas na figura 7.28a. Na figura 7.28b é mostrada somente a raiz estável deste modelo, nos diversos pontos de operação.

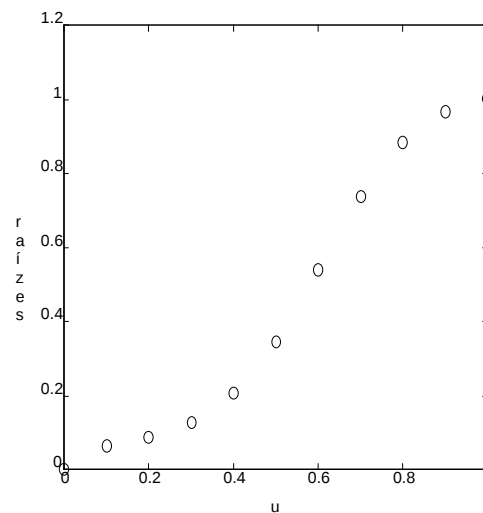
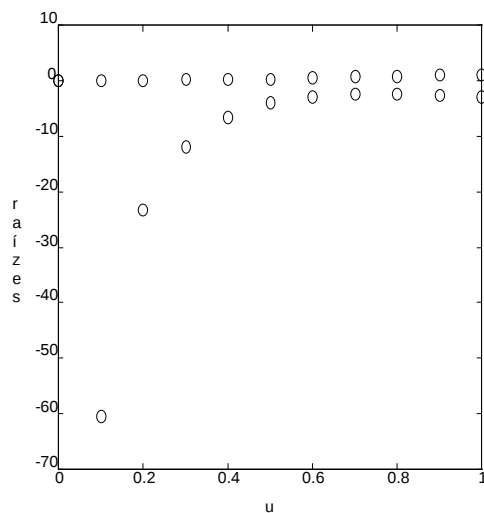


Figura 7.28a - Raízes do modelo (24 termos). Figura 7.28b - Raíz estável do modelo.

Pelas figuras 7.27, 7.28a e 7.28b percebe-se que a característica estática só poderá ser bem recuperada mediante o cálculo das raízes do modelo.

A partir do modelo de 24 termos, os critérios usuais de identificação não permitem melhorar mais o modelo, sendo que este último então seria considerado o modelo final.

7.4 Comentários Finais

Um sistema de controle de um processo que apresenta uma válvula ajustada para operar em um ponto de sua característica estática pode não estar bem ajustado quando muda-se tal ponto de operação, consequentemente podendo trazer problemas para o controle do processo. O interesse em obter-se um modelo dinâmico para a válvula de controle, através do qual pode-se recuperar de maneira direta informação a respeito de sua característica estática, visa aliviar este problema, através da implementação de um sistema controle ‘inteligente’ capaz de processar tal informação.

Neste capítulo, conhecimentos *a priori* do sistema auxiliaram na redução do conjunto de termos candidatos ao modelo, encontrando-se um modelo conciso e dinamicamente válido, e que permitiu a recuperação analítica da característica estática da válvula.

Note que os testes de validação estatística estão mais dentro da faixa para o modelo baseado em informações a priori do que os modelos obtidos através dos critérios usuais de identificação.

Neste capítulo percebeu-se a interação entre o ganho e a constante de tempo, ambas variando com o ponto de operação do sistema, como dito anteriormente, além do ganho variável. Este fato faz com que, em regime estático $\Sigma_{x,y} \neq 0$, que como dito no capítulo 5, ocorre quando o ganho e a constante de tempo relacionam-se entre si.

Na seção 7.3.3, chegou-se a um modelo final (eq. 7.4) que não permitia uma boa recuperação da característica estática da válvula. Neste modelo, a utilização do conhecimento *a priori* do sistema permitiria que o agrupamento $\Omega_{y^2_u}$ fosse retirado do modelo. Com as técnicas usuais de identificação, tal agrupamento não poderia ser retirado. Isto é crítico na medida em que a inclusão de tal agrupamento no modelo é um impedimento para que a curva característica do mesmo seja obtido analiticamente.

Conclusão

A obtenção de modelos matemáticos que reproduzam a dinâmica e características de sistemas em estudo é uma tarefa de vital importância em engenharia. Uma abordagem possível para a construção de modelos matemáticos é a identificação de sistemas. A identificação de sistemas é a técnica que permite desenvolver modelos matemáticos a partir de dados medidos. Esse trabalho investigou um procedimento de aplicação de técnicas para identificação de sistemas dinâmicos não-lineares, com ganho e constante de tempo variáveis, utilizando modelos polinomiais NARMAX, usando um conhecimento *a priori* do mesmo.

A identificação de sistemas dinâmicos não-lineares é uma área de pesquisa relativamente recente. Os fundamentos teóricos de identificação de modelos não-lineares, tal como a utilização de conhecimentos *a priori* do sistema no auxílio a detecção da estrutura do modelo, ainda estão sendo estabelecidos, conforme pode ser verificado no grande número de trabalhos publicados em revistas especializadas (ver capítulo 2). Atualmente, o procedimento de identificação necessita da experiência, conhecimentos do sistema e capacidade de julgamento do usuário para determinar modelos não-lineares para os sistemas em estudo, fato refletido pelos resultados obtidos nesta dissertação. Assim, esta área de pesquisa ainda está em fase de consolidação.

Este trabalho analisou a utilização conjunta de técnicas de agrupamentos de termos e coeficientes de agrupamentos, e um conhecimento *a priori* do sistema, na obtenção de estruturas concisas de modelos, onde fizessem parte somente os agrupamentos supostamente efetivos.

Inicialmente analisaram-se modelos discretizados de sistemas lineares até terceira ordem, onde observou-se os agrupamentos presentes. Posteriormente foram incluídas nos modelos variações do ganho e de uma constante de tempo, mediante a utilização de uma equação genérica de variação. Através disto investigou-se a inclusão de novos agrupamentos de termos no modelo.

As observações feitas para os modelos discretizados, mediante certas considerações, foram extendidas e comparadas a certos modelos polinomiais NARMAX. Com isto tentou-se recuperar as equações representativas da relação estática e constante de tempo variáveis presentes nos modelos NARMAX, sendo resumidas em equações genéricas em função dos agrupamentos de termos e coeficientes de agrupamentos.

Os resultados obtidos até então foram utilizados na identificação de dois modelos simulados de Wiener e Hammerstein, supondo-se ganho variável mediante a utilização de uma característica estática não-linear em série com um modelo linear dinâmico. Através destas aplicações verificou-se a obtenção de modelos dinamicamente válidos e concisos que permitiam de alguma forma mais rápida que simulação, recuperar-se as características estáticas dos sistemas, através de raízes ou analiticamente.

O procedimento desenvolvido ao longo da dissertação, foi aplicado então com sucesso, à uma válvula de controle real, presente em uma planta de tanques interativos, no laboratório de Controle de Processos Industriais (LCPI) da Universidade Federal de Minas Gerais. Através deste procedimento, obteve-se um modelo dinâmico para esta válvula, que considerando-se em regime estático transformava-se numa equação algébrica, através da qual obtém-se diretamente a característica estática da válvula. Tal informação pode ser muito útil na implementação de sistemas de controle ‘inteligente’.

O modelo da válvula obtido através de conhecimento *a priori* pôde ser comparado com o modelo obtido à partir de técnicas usuais de identificação, diferenciando-se assim a qualidade de ambos.

Portanto, o uso das técnicas de agrupamento de termos e coeficientes de agrupamentos em conjunto com um conhecimento *a priori* do sistema, no auxílio a detecção de estrutura de modelos polinomiais NARMAX é válida e importante. É certo que nem sempre este conhecimento *a priori* está disponível, mas sempre que possível pode e deve ser utilizado no processo de identificação, como visto neste trabalho.

É interessante notar que as técnicas de agrupamentos de termos e coeficientes de agrupamentos, que são técnicas novas, além de auxiliarem na detecção de estruturas podem vir a auxiliar também na escolha do número adequado de termos para o modelo, como visto em diversos pontos neste trabalho. Cabe ressaltar também, que, um agrupamento cujo coeficiente esteja muito próximo de zero nem sempre é um agrupamento espúrio, como dito em certos pontos durante a dissertação

A presença de ruído nos dados reais pode dificultar o processo de identificação, entretanto, as técnicas utilizadas neste trabalho demonstraram ser bastante robustas.

Algumas propostas para a continuação desse trabalho são:

- análise mais profunda do papel dos agrupamentos de termos decorrentes da variação das constantes de tempo no modelo;
- projetos de controladores para sistemas não-lineares utilizando modelos NARMAX polinomiais identificados;
- projetos de controladores ‘inteligentes’ para sistemas não-lineares utilizando modelos polinomiais NARMAX, que façam uso das informações contidas no modelo;
- utilização do modelo NARMAX identificado para a detecção de falhas de operação do sistema original.

APÊNDICE A

Neste apêndice proceder-se-á ao raciocínio da seção 4.5, agora para sistemas de segunda e terceira ordem. Será visto como o ganho e a constante de tempo variável influem no aparecimento de termos no modelo discretizado.

A.1 Modelo Discretizado de Segunda Ordem

Da mesma forma que na seção 4.5.1, considere o sistema de segunda ordem com ganho e uma das constantes de tempo variáveis de acordo com as equações (4.1) e (4.2). No desenvolvimento a seguir, τ_2 é considerado constante.

Para facilitar a visualização e a manipulação das expressões subsequentes, omitir-se-ão as funções de dependência da constante de tempo e do ganho, passando a reescrevê-los da seguinte maneira:

$$\tau_1 = \tau_{c1} + \tau_{v1} x_{\tau_1} \quad (\text{A.1a})$$

$$K = K_0 + K_1 x_K \quad (\text{A.1b})$$

Tomando-se as equações (A.1a) e (A.1b) e substituindo-as na equação (4.14) e procedendo-se de forma análoga á da seção 4.5.1, chega-se a:

$$\begin{aligned} \left(\frac{\tau_{c1}\tau_2}{T_s^2} + \frac{\tau_{c1} + \tau_2}{T_s} \right) y(k) = & \left(\frac{\tau_{c1}\tau_2}{T_s^2} + \frac{\tau_{c1} + \tau_2}{T_s} \right) y(k-1) - \frac{\tau_{v1}\tau_2}{T_s^2} x_{\tau_1} [y(k) - y(k-1)] - \\ & - \frac{\tau_{v1}}{T_s} x_{\tau_1} [y(k) - y(k-1)] + \frac{\tau_{v1}\tau_2}{T_s^2} x_{\tau_1} [y(k-1) - y(k-2)] + \\ & + \frac{\tau_{c1}\tau_2}{T_s^2} [y(k-1) - y(k-2)] - y(k-2) + K_0 u(k-2) + K_1 x_K u(k-2) \end{aligned} \quad (\text{A.2})$$

Da equação (A.2), além dos agrupamentos de termos lineares, para uma dependência da constante de tempo, observa-se a presença do agrupamento $\Omega_{x_{\tau_1}y}$. Nesta situação,

também é válido notar que os coeficientes dos termos do agrupamento de termos $\Omega_{x_{\tau_1}y}$, é o mesmo, porém com sinais contrários, ou seja $\Sigma_{x_{\tau_1}y} = 0$, sugerindo que tais termos poderiam ser espúrios. Para uma dependência do ganho, o agrupamento de termos presente é o mesmo que para o sistema de primeira ordem, ou seja $\Omega_{x_k u}$.

A.2 Modelo Discretizado de Terceira Ordem

Da mesma maneira que para o sistema de segunda ordem, o sistema de terceira ordem serão considerados τ_2 e τ_3 constantes. O ganho K e τ_1 variam de acordo com as equações (4.1) e (4.2), respectivamente.

Então, pode-se reescrever a equação (4.15) da seguinte forma :

$$\begin{aligned} & \frac{(\tau_{c1}\tau_{c2}\tau_{c3} + \tau_{c3}\tau_{c2}\tau_{v1}x_{\tau_1})}{T_s^3} \Delta y(k) + \frac{(\tau_{c1}\tau_{c2} + \tau_{c2}\tau_{v1}x_{\tau_1} + \tau_{c1}\tau_{c3} + \tau_{c3}\tau_{v1}x_{\tau_1} + \tau_{c2}\tau_{c3})}{T_s^2} \nabla y(k) + \\ & + \frac{(\tau_{c1} + \tau_{v1}x_{\tau_1} + \tau_{c2} + \tau_{c3})}{T_s} \Delta y(k) + y(k-3) = K_0 u(k-3) + K_1 x_K u(k-3) \end{aligned} \quad (A.3)$$

Para um sistema discretizado de terceira ordem, da mesma maneira que o de segunda ordem, pode-se dizer que, além dos agrupamentos de termos lineares, o agrupamento $\Omega_{x_{\tau_1}y}$, estará presente para uma dependência da constante de tempo. O mesmo efeito de cancelamento visto para primeira e segunda ordem, pode ser visto aqui, ou seja $\Sigma_{x_{\tau_1}y} = 0$.

O agrupamento de termos para uma dependência do ganho é o mesmo para sistemas de primeira e segunda ordem.

APÊNDICE B

Neste apêndice serão feitas as generalizações da seção 4.6 para sistemas de segunda e terceira ordem.

B.1 Sistema de Segunda Ordem

$$\Sigma_y = 2 \frac{\tau_{c1}\tau_{c2}}{T_s^2} - 2 \frac{\tau_{c1}\tau_{c2}}{T_s^2} + \frac{(\tau_{c1} + \tau_{c2})}{T_s} - \frac{(\tau_{c1} + \tau_{c2})}{T_s} + 1$$

$$\Sigma_{x_{\tau y}} = 2 \frac{\tau_{v1}\tau_{c2}}{T_s^2} - 2 \frac{\tau_{v1}\tau_{c2}}{T_s^2} + \frac{\tau_{v1}}{T_s} - \frac{\tau_{v1}}{T_s}$$

$$\Sigma_u = K_0$$

$$\Sigma_{x_{ku}} = K_1$$

B.2 Sistema de Terceira Ordem

$$\begin{aligned} \Sigma_y = & \frac{4\tau_{c1}\tau_{c2}\tau_{c3} - 4\tau_{c1}\tau_{c2}\tau_{c3}}{T_s^3} + \frac{2(\tau_{c1}\tau_{c2} + \tau_{c2}\tau_{c3} + \tau_{c1}\tau_{c3})}{T_s^2} - \frac{2(\tau_{c1}\tau_{c2} + \tau_{c2}\tau_{c3} + \tau_{c1}\tau_{c3})}{T_s^2} \\ & + \frac{(\tau_{c1} + \tau_{c2} + \tau_{c3})}{T_s} - \frac{(\tau_{c1} + \tau_{c2} + \tau_{c3})}{T_s} + 1 \end{aligned}$$

$$\Sigma_{x_{\tau y}} = \frac{4\tau_{v1}\tau_{c2}\tau_{c3} - 4\tau_{v1}\tau_{c2}\tau_{c3}}{T_s^3} + \frac{2(\tau_{v1}\tau_{c2} + \tau_{v1}\tau_{c3})}{T_s^2} - \frac{2(\tau_{v1}\tau_{c2} + \tau_{v1}\tau_{c3})}{T_s^2} + \frac{\tau_{v1}}{T_s} - \frac{\tau_{v1}}{T_s}$$

$$\Sigma_u = K_0$$

$$\Sigma_{x_{ku}} = K_1$$

APÊNDICE C

Neste apêndice serão feitas generalizações como a da seção 5.2, para modelos NARX de segunda e terceira ordem, com parâmetros variáveis.

C.1 Generalização para um Sistema de Segunda Ordem

Seja o seguinte modelo de segunda ordem, com uma constante de tempo e ganho variáveis:

$$\begin{aligned} y(k) = & Ay(k-1) + By(k-2) + Cu(k-1) + Du(k-2) + E_{x_\tau}(k-1)y(k-1) + \\ & + F_{x_\tau}(k-1)y(k-2) + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2) \end{aligned} \quad (C.1)$$

A seguir, vai-se manipular esta equação para que fique na forma da equação discretizada expressa pela fórmula genérica da equação (4.11 b), sendo então comparada com esta.

Partindo-se da equação (C.1), e passando todos os termos em y para o lado esquerdo da equação:

$$\begin{aligned} y(k) - Ay(k-1) - By(k-2) - E_{x_\tau}(k-1)y(k-1) - F_{x_\tau}(k-1)y(k-2) = & Cu(k-1) + \\ & + Du(k-2) + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2) \end{aligned}$$

Uma técnica para se obter o primeiro $\Delta y(k)$, é somar e subtrair simultaneamente $y(k-1)$ do lado esquerdo da equação :

$$\begin{aligned} y(k) + y(k-1) - y(k-1) - Ay(k-1) - By(k-2) - E_{x_\tau}(k-1)y(k-1) - F_{x_\tau}(k-1)y(k-2) = & \\ Cu(k-1) + Du(k-2) + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2) \end{aligned}$$

Chamando $\Delta y(k) = y(k) - y(k-1)$ e somando os termos em $y(k-1)$ e $y(k-2)$:

$$\Delta y(k) + (1 - A)y(k-1) - By(k-2) - E_{x_\tau}(k-1)y(k-1) - F_{x_\tau}(k-1)y(k-2) = Cu(k-1) + Du(k-2) + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2)$$

Fazendo-se o mesmo procedimento para o termo em $y(k-1)$, completando-se com um termo de uma ordem de superioridade:

$$\Delta y(k) + (1 - A)y(k-1) - (1 - A)y(k-2) + (1 - A)y(k-2) - By(k-2) - E_{x_\tau}(k-1)y(k-1) - F_{x_\tau}(k-1)y(k-2) = Cu(k-1) + Du(k-2) + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2)$$

Logo,

$$\Delta y(k) + (1 - A)\Delta y(k-1) + (1 - A - B)y(k-2) - E_{x_\tau}(k-1)y(k-1) - F_{x_\tau}(k-1)y(k-2) = Cu(k-1) + Du(k-2) + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2)$$

Assumindo-se que $\Delta y(k) \approx y(k) - y(k-1) \approx y(k-1) - y(k-2)$:

$$(2-A)\Delta y(k) + (1 - A - B)y(k-2) - E_{x_\tau}(k-1)y(k-1) - F_{x_\tau}(k-1)y(k-2) = Cu(k-1) + Du(k-2) + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2)$$

Com os termos do agrupamento $\Omega_{x_\tau y}$, ao contrário do agrupamento Ω_y , começar-se-á a completar os termos a partir do termo de maior ordem, interagindo-se com um termo de uma ordem abaixo. Assim:

$$(2-A)\Delta y(k) + (1 - A - B)y(k-2) - E_{x_\tau}(k-1)y(k-1) - \mathbf{F}_{x_\tau}(\mathbf{k-1})\mathbf{y(k)} + \mathbf{F}_{x_\tau}(\mathbf{k-1})\mathbf{y(k-1)} - F_{x_\tau}(k-1)y(k-2) = Cu(k-1) + Du(k-2) + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2)$$

$$(2-A)\Delta y(k) + (1-A-B)y(k-2) - (E+F)x_\tau(k-1)y(k-1) + F\Delta y(k) = Cu(k-1) + Du(k-2) + Gx_k(k)u(k-1) + Hx_k(k)u(k-2)$$

$$(2-A)\Delta y(k) + (1-A-B)y(k-2) - (E+F)x_\tau(k-1)y(k-1) + (E+F)x_\tau(k-1)y(k) - (E+F)x_\tau(k-1)y(k-1) + F\Delta y(k) = Cu(k-1) + Du(k-2) + Gx_k(k)u(k-1) + Hx_k(k)u(k-2)$$

$$(2-A)\Delta y(k) + (1-A-B)y(k-2) - (E+F)x_\tau(k-1)y(k) + (E+F)x_\tau(k-1)\Delta y(k) + Fx_\tau(k-1)\Delta y(k) = Cu(k-1) + Du(k-2) + Gx_k(k)u(k-1) + Hx_k(k)u(k-2)$$

$$(2-A+(E+2F)x_\tau(k-1))\Delta y(k) = -(1-A-B)y(k-2) + (E+F)x_\tau(k)y(k) + Cu(k-1) + Du(k-2) + Gx_k(k)u(k-1) + Hx_k(k)u(k-2)$$

Explicitando-se o termo $y(k-2)$:

$$\Delta y(k) = \frac{(1-A-B)}{(2-A+(E+2F)x_\tau(k-1))} \left[\frac{(Cu(k-1) + Du(k-2) + x_k[Gu(k-1) + Hu(k-2)] + (F+E)x_\tau(k-1)y(k))}{(1-A-B)} - y(k-2) \right] \quad (C.2)$$

Portanto, comparando-se a equação (4.11.b) a (C.2), chega-se em:

$$\frac{T_s}{(\tau_1 + \tau_2)} = \frac{(1-A-B)}{(2-A+(E+2F)x_\tau(k-1))} \quad (C.3)$$

logo,

$$K_{re}u(k-2) = \hat{y} = \left[\frac{(Cu(k-1) + Du(k-2) + x_k[Gu(k-1) + Hu(k-2)] + (F+E)x_\tau(k-1)y(k))}{(1-A-B)} \right] \quad (C.4)$$

Considerando o sistema em regime estático $y(k) = y(k-1) = y(k-2) = Y$, $u(k) = u(k-1) = u(k-2) = U$, $x_\tau = X_\tau$ e $x_k = X_k$, e aplicando-se estas considerações nas equações (C.3) e (C.4), expressando-as em termos de agrupamentos de termos e coeficientes de agrupamento, obtém-se:

$$\frac{T_s}{\tau_1 + \tau_2} = \frac{(1 - \Sigma_y)}{(2 - A + (E + 2F)X_\tau)}$$

e

$$K_{re}U = \hat{y} = \frac{(\Sigma_u U + \Sigma_{x_k u} X_k U + \Sigma_{x_\tau y} X_\tau Y)}{(1 - \Sigma_y)}$$

Que ainda podem ser reescritas da seguinte forma:

$$\frac{\tau_1 + \tau_2}{T_s} = \frac{(2 - A)}{(1 - \Sigma_y)} + \frac{(E + 2F)}{(1 - \Sigma_y)} X_\tau \quad (C.5)$$

e

$$K_{re} = \frac{\Sigma_u}{(1 - \Sigma_y) - (\Sigma_{x_\tau y})X_\tau} + \frac{(\Sigma_{x_k u})X_k}{(1 - \Sigma_y) - (\Sigma_{x_\tau y})} \quad (C.6)$$

C.2 Generalização para um Sistema de Terceira Ordem

Seja o modelo de terceira ordem com uma constante de tempo e ganho variáveis:

$$\begin{aligned}
 y(k) = & Ay(k-1) + By(k-2) + Cu(k-1) + Du(k-2) + E_{x_\tau}(k-1)y(k-1) + \\
 & + F_{x_\tau}(k-1)y(k-2) + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2) + Iy(k-3) + Ju(k-3) + \\
 & + K_{x_\tau}(k-1)y(k-3) + L_{x_k}(k)u(k-3)
 \end{aligned} \tag{C.7}$$

A seguir será feito o mesmo procedimento que foi feito para um sistema de segunda ordem.

Tomando-se a equação (C.7), e passando-se todos os termos em y para o lado esquerdo da equação:

$$\begin{aligned}
 y(k) - Ay(k-1) - By(k-2) - Iy(k-3) - E_{x_\tau}(k-1)y(k-1) - F_{x_\tau}(k-1)y(k-2) - \\
 - K_{x_\tau}(k-1)y(k-3) = Cu(k-1) + Du(k-2) + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2) + \\
 + Ju(k-3) + L_{x_k}(k)u(k-3)
 \end{aligned}$$

Agora continua-se o procedimento usual, como para o sistema de segunda ordem:

$$\begin{aligned}
 y(k) - y(k-1) + y(k-1) - Ay(k-1) - By(k-2) - Iy(k-3) - E_{x_\tau}(k-1)y(k-1) - \\
 - F_{x_\tau}(k-1)y(k-2) - K_{x_\tau}(k-1)y(k-3) = Cu(k-1) + Du(k-2) + G_{x_k}(k)u(k-1) + \\
 + H_{x_k}(k)u(k-2) + Ju(k-3) + L_{x_k}(k)u(k-3)
 \end{aligned}$$

$$\begin{aligned}
 y(k) - y(k-1) + y(k-1) - Ay(k-1) - By(k-2) - Iy(k-3) - E_{x_\tau}(k-1)y(k-1) - \\
 - F_{x_\tau}(k-1)y(k-2) - K_{x_\tau}(k-1)y(k-3) = Cu(k-1) + Du(k-2) + G_{x_k}(k)u(k-1) + \\
 + H_{x_k}(k)u(k-2) + Ju(k-3) + L_{x_k}(k)u(k-3)
 \end{aligned}$$

Prosseguindo-se no raciocínio, como na seção C.1:

$$\begin{aligned}
 \Delta y(k) + (1-A)y(k-1) - (1-A)y(k-2) + (1-A)y(k-2) - E_{x_\tau}(k-1)y(k-1) - By(k-2) - \\
 - Iy(k-3) - F_{x_\tau}(k-1)y(k-2) - K_{x_\tau}(k-1)y(k-3) = Cu(k-1) + Du(k-2) + \\
 + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2) + Ju(k-3) + L_{x_k}(k)u(k-3)
 \end{aligned}$$

Considerando-se, $\Delta y(k) \approx y(k) - y(k-1) \approx y(k-1) - y(k-2) \approx y(k-2) - y(k-3)$:

$$\begin{aligned} &\Delta y(k) + (1 - A)\Delta y(k) + (1 - A - B)y(k-2) - E_{x_\tau}(k-1)y(k-1) - Iy(k-3) - \\ &- F_{x_\tau}(k-1)y(k-2) - K_{x_\tau}(k-1)y(k-3) = Cu(k-1) + Du(k-2) + G_{x_k}(k)u(k-1) + \\ &+ H_{x_k}(k)u(k-2) + Ju(k-3) + L_{x_k}(k)u(k-3) \end{aligned}$$

$$\begin{aligned} &(2 - A)\Delta y(k) + (1 - A - B)y(k-2) - (1 - A - B)y(k-3) + (1 - A - B)y(k-3) - \\ &- E_{x_\tau}(k-1)y(k-1) - Iy(k-3) - F_{x_\tau}(k-1)y(k-2) - K_{x_\tau}(k-1)y(k-3) = Cu(k-1) + \\ &+ Du(k-2) + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2) + Ju(k-3) + L_{x_k}(k)u(k-3) \end{aligned}$$

$$\begin{aligned} &(2 - A)\Delta y(k) + (1 - A - B)y(k-2) - (1 - A - B)y(k-3) + (1 - A - B - I)y(k-3) - \\ &- E_{x_\tau}(k-1)y(k-1) - F_{x_\tau}(k-1)y(k-2) - K_{x_\tau}(k-1)y(k-3) = Cu(k-1) + Du(k-2) + \\ &+ G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2) + Ju(k-3) + L_{x_k}(k)u(k-3) \end{aligned}$$

$$\begin{aligned} &(2 - A)\Delta y(k) + (1 - A - B)\Delta y(k) + (1 - A - B - I)y(k-3) - E_{x_\tau}(k-1)y(k-1) - \\ &- F_{x_\tau}(k-1)y(k-2) - K_{x_\tau}(k-1)y(k-3) = Cu(k-1) + Du(k-2) + G_{x_k}(k)u(k-1) + \\ &+ H_{x_k}(k)u(k-2) + Ju(k-3) + L_{x_k}(k)u(k-3) \end{aligned}$$

$$\begin{aligned} &(3 - 2A - B)\Delta y(k) + (1 - A - B - I)y(k-3) - E_{x_\tau}(k-1)y(k-1) - F_{x_\tau}(k-1)y(k-2) - \\ &- K_{x_\tau}(k-1)y(k-3) = Cu(k-1) + Du(k-2) + G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2) + \\ &+ Ju(k-3) + L_{x_k}(k)u(k-3) \end{aligned}$$

Agora, será feito o mesmo procedimento para os termos $x_\tau y$:

$$\begin{aligned} &(3 - 2A - B)\Delta y(k) + (1 - A - B - I)y(k-3) - E_{x_\tau}(k-1)y(k-1) - F_{x_\tau}(k-1)y(k-2) - \\ &- K_{x_\tau}(k-1)y(k-3) + K_{x_\tau}(k-1)y(k-2) - K_{x_\tau}(k-1)y(k-2) = Cu(k-1) + Du(k-2) + \\ &+ G_{x_k}(k)u(k-1) + H_{x_k}(k)u(k-2) + Ju(k-3) + L_{x_k}(k)u(k-3) \end{aligned}$$

$$(3 - 2A - B)\Delta y(k) + (1 - A - B - I)y(k-3) - E x_{\tau}(k-1) y(k-1) - \\ -(F + K) x_{\tau}(k-1) y(k-2) + \mathbf{K} x_{\tau}(\mathbf{k-1})\Delta \mathbf{y(k)} = Cu(k-1) + Du(k-2) + \\ + G x_k(k)u(k-1) + H x_k(k)u(k-2) + Ju(k-3) + L x_k(k)u(k-3)$$

$$(3 - 2A - B + K x_{\tau}(k-1))\Delta y(k) + (1 - A - B - I)y(k-3) - E x_{\tau}(k-1) y(k-1) - \\ (F + K) x_{\tau}(k-1) y(k-2) + (\mathbf{F + K}) x_{\tau}(\mathbf{k-1})y(\mathbf{k-1}) - (\mathbf{F + K}) x_{\tau}(\mathbf{k-1})y(\mathbf{k-1}) = \\ Cu(k-1) + Du(k-2) + G x_k(k)u(k-1) + H x_k(k)u(k-2) + Ju(k-3) + L x_k(k)u(k-3)$$

$$(3 - 2A - B + K x_{\tau}(k-1))\Delta y(k) + (1 - A - B - I)y(k-3) - \\ -(E + F + K) x_{\tau}(k-1) y(k-1) - (\mathbf{F + K}) x_{\tau}(\mathbf{k-1})y(\mathbf{k-2}) + (\mathbf{F + K}) x_{\tau}(\mathbf{k-1})y(\mathbf{k-1}) = \\ +Cu(k-1) + Du(k-2) + G x_k(k)u(k-1) + H x_k(k)u(k-2) + Ju(k-3) + L x_k(k)u(k-3)$$

$$(3 - 2A - B + K x_{\tau}(k-1))\Delta y(k) + (1 - A - B - I)y(k-3) - \\ -(E + F + K) x_{\tau}(k-1) y(k-1) + (\mathbf{F + K}) x_{\tau}(\mathbf{k-1})\Delta \mathbf{y(k)} = Cu(k-1) + Du(k-2) + \\ +G x_k(k)u(k-1) + H x_k(k)u(k-2) + Ju(k-3) + L x_k(k)u(k-3)$$

$$(3 - 2A - B + (F + 2K) x_{\tau}(k-1))\Delta y(k) + (1 - A - B - I)y(k-3) - \\ -(E + F + K) x_{\tau}(k-1) y(k-1) + (\mathbf{E + F + K}) x_{\tau}(\mathbf{k-1})y(\mathbf{k}) - \\ -(\mathbf{E + F + K}) x_{\tau}(\mathbf{k-1})y(\mathbf{k}) = Cu(k-1) + Du(k-2) + G x_k(k)u(k-1) + H x_k(k)u(k-2) + \\ +Ju(k-3) + L x_k(k)u(k-3)$$

$$(3 - 2A - B + (F + 2K) x_{\tau}(k-1))\Delta y(k) + (1 - A - B - I)y(k-3) - \\ (\mathbf{E + F + K}) x_{\tau}(\mathbf{k-1})y(\mathbf{k-1}) + (\mathbf{E + F + K}) x_{\tau}(\mathbf{k-1})y(\mathbf{k}) - (E + F + K) x_{\tau}(k-1) y(k) = \\ Cu(k-1) + Du(k-2) + G x_k(k)u(k-1) + H x_k(k)u(k-2) + Ju(k-3) + L x_k(k)u(k-3)$$

$$(3 - 2A - B + (F + 2K) x_{\tau}(k-1))\Delta y(k) + (1 - A - B - I)y(k-3) -$$

$$(\mathbf{E} + \mathbf{F} + \mathbf{K}) \mathbf{x}_\tau(\mathbf{k}-1) \Delta \mathbf{y}(\mathbf{k}) - (\mathbf{E} + \mathbf{F} + \mathbf{K}) \mathbf{x}_\tau(\mathbf{k}-1) \mathbf{y}(\mathbf{k}) = \mathbf{C} \mathbf{u}(\mathbf{k}-1) + \mathbf{D} \mathbf{u}(\mathbf{k}-2) + \\ + \mathbf{G} \mathbf{x}_k(\mathbf{k}) \mathbf{u}(\mathbf{k}-1) + \mathbf{H} \mathbf{x}_k(\mathbf{k}) \mathbf{u}(\mathbf{k}-2) + \mathbf{J} \mathbf{u}(\mathbf{k}-3) + \mathbf{L} \mathbf{x}_k(\mathbf{k}) \mathbf{u}(\mathbf{k}-3)$$

Portanto:

$$(3 - 2\mathbf{A} - \mathbf{B} + (\mathbf{E} + 2\mathbf{F} + 3\mathbf{K}) \mathbf{x}_\tau(\mathbf{k}-1)) \Delta \mathbf{y}(\mathbf{k}) + (1 - \mathbf{A} - \mathbf{B} - \mathbf{I}) \mathbf{y}(\mathbf{k}-3) - \\ - (\mathbf{E} + \mathbf{F} + \mathbf{K}) \mathbf{x}_\tau(\mathbf{k}-1) \mathbf{y}(\mathbf{k}) = \mathbf{C} \mathbf{u}(\mathbf{k}-1) + \mathbf{D} \mathbf{u}(\mathbf{k}-2) + \mathbf{G} \mathbf{x}_k(\mathbf{k}) \mathbf{u}(\mathbf{k}-1) + \mathbf{H} \mathbf{x}_k(\mathbf{k}) \mathbf{u}(\mathbf{k}-2) \\ + \mathbf{J} \mathbf{u}(\mathbf{k}-3) + \mathbf{L} \mathbf{x}_k(\mathbf{k}) \mathbf{u}(\mathbf{k}-3)$$

Isolando-se o termo $\Delta \mathbf{y}(\mathbf{k})$ tem-se:

$$(3 - 2\mathbf{A} - \mathbf{B} + (\mathbf{E} + 2\mathbf{F} + 3\mathbf{K}) \mathbf{x}_\tau(\mathbf{k}-1)) \Delta \mathbf{y}(\mathbf{k}) = - (1 - \mathbf{A} - \mathbf{B} - \mathbf{I}) \mathbf{y}(\mathbf{k}-3) + \\ + (\mathbf{E} + \mathbf{F} + \mathbf{K}) \mathbf{x}_\tau(\mathbf{k}-1) \mathbf{y}(\mathbf{k}) + \mathbf{C} \mathbf{u}(\mathbf{k}-1) + \mathbf{D} \mathbf{u}(\mathbf{k}-2) + \mathbf{G} \mathbf{x}_k(\mathbf{k}) \mathbf{u}(\mathbf{k}-1) + \mathbf{H} \mathbf{x}_k(\mathbf{k}) \mathbf{u}(\mathbf{k}-2) \\ + \mathbf{J} \mathbf{u}(\mathbf{k}-3) + \mathbf{L} \mathbf{x}_k(\mathbf{k}) \mathbf{u}(\mathbf{k}-3)$$

A equação anterior pode ser reescrita na forma:

$$\Delta \mathbf{y}(\mathbf{k}) = \frac{(1 - (\mathbf{A} + \mathbf{I} + \mathbf{B}))}{(3 - 2\mathbf{A} - \mathbf{B} + (\mathbf{E} + 2\mathbf{F} + 3\mathbf{K}) \mathbf{x}_\tau(\mathbf{k}-1))} \left[\begin{array}{l} \mathbf{C} \mathbf{u}(\mathbf{k}-1) + \mathbf{D} \mathbf{u}(\mathbf{k}-2) + \\ + \mathbf{J} \mathbf{u}(\mathbf{k}-3) + \\ + \mathbf{G} \mathbf{x}_k(\mathbf{k}) \mathbf{u}(\mathbf{k}-1) + \\ + \mathbf{H} \mathbf{x}_k(\mathbf{k}) \mathbf{u}(\mathbf{k}-2) + \\ + \mathbf{L} \mathbf{x}_k(\mathbf{k}) \mathbf{u}(\mathbf{k}-3) + \\ + (\mathbf{K} + \mathbf{E} + \mathbf{F}) \mathbf{x}_\tau(\mathbf{k}-1) \mathbf{y}(\mathbf{k}) \end{array} \right] - \mathbf{y}(\mathbf{k}-3) \quad (\text{C.8})$$

Comparando-se a equação (C.8) com a equação discretizada (4.12.b), tem-se:

$$\frac{T_s}{\tau_1 + \tau_2 + \tau_3} = \frac{(1 - (A + I + B))}{(3 - 2A - B + (E + 2F + 3K)x_\tau(k - 1))} \quad (C.9)$$

e

$$K_{re} \frac{u(k) + u(k - 1)}{2} = \hat{y} = \frac{Cu(k - 1) + Du(k - 2) + Ju(k - 3) + Gx_k(k)u(k - 1) + Hx_k(k)u(k - 2) + Lx_k(k)u(k - 3) + (K + E + F)x_\tau(k - 1)y(k)}{(1 - (A + B + I))} \quad (C.10)$$

Considerando o sistema em regime estático, as equações (C.9) e (C.10) podem ser reescritas em termos de agrupamento de termos e coeficiente de agrupamento como segue:

$$\frac{\tau_1 + \tau_2 + \tau_3}{T_s} = \frac{(3 - 2A - B)}{(1 - \Sigma_y)} + \frac{(E + 2F + 3K)}{(1 - \Sigma_y)} X_\tau \quad (C.11)$$

e

$$K_{re} = \frac{Y}{U} = \frac{(\Sigma_u)}{(1 - \Sigma_y) - (\Sigma_{x_\tau y})X_\tau} + \frac{(\Sigma_{x_k u})X_k}{(1 - \Sigma_y) - (\Sigma_{x_\tau y})X_\tau} \quad (C.12)$$

Referências Bibliográficas

- Aguirre, L. A. [1994a] “Term Clustering and the Order Selection of Linear Continuous Systems”. *Journal of the Franklin Institute*, 331B(4):403-415.
- Aguirre, L. A. [1994b] “Some Remarks on Structure Selection for Nonlinear Models” *International Journal of Bifurcation and Chaos*, 4 (6):1707-1714.
- Aguirre, L. A. [1995] “A Nonlinear Correlation Function for Selecting the Delay Time in Dynamical Reconstructions” *Physics Letters A*, 203:88-94.
- Aguirre, L. A. and Aguirre, A [1995] “A Tutorial Introduction to Nonlinear Dynamics in Economics”, *CEDEPLAR Universidade Federal de Minas Gerais*, Working Paper # 88.
- Aguirre, L. A. and Billings, S. A. [1994a] “ Validating Identified Nonlinear Models with Chaotic Dynamics” *International Journal of Bifurcation and Chaos* 4 (4):109-125.
- Aguirre, L. A. and Billings, S. A. [1994b] “Discrete Reconstruction of Strange Attractors in Chua’s Circuit” *International Journal of Bifurcation and Chaos*, 4(4):853-864.
- Aguirre, L. A. and Billings, S. A. [1995a] “Improved Structure Selection for Nonlinear Models Based on Term Clustering” *International Journal of Control*, 62 (3):569-587.
- Aguirre, L. A. and Billings, S. A. [1995b] “Dynamical Effects of Overparametrization in Nonlinear Models” *Physica D.*, 80(1,2):26-40.
- Aguirre, L. A. and Billings, S. A. [1995c] “Retrieving Dynamical Invariants from Chaotic Data Using NARMAX Models” *International Journal of Bifurcation and Chaos*, 5(2):449-474.
- Aguirre, L. A. and Mendes, E. M. A. M. [1996] “Global Nonlinear Polynomial Models: Structure, Terms Clusters and Fixed Points” *International Journal of Bifurcation and Chaos*, 6(2):279-294.
- Akaike, H. [1970] “Statistical Predictor Identification” *Ann. Inst. Stat. Math.* 22:203-217.
- Akaike, H. [1974] “A New Look at the Statistical Model Identification” *IEEE Transactions on Automatic Control*, 19(6):716-723.
- Åström, K. J. and Eykhoff, P. [1971] “System Identification - a Survey” *Automática* 7:123-167.
- Billings, S. A. [1980] “Identification of Nonlinear Systems- a Survey” *IEE Proceedings* 127(6):272-285.
- Billings, S. A. and Aguirre, L. A. [1995] “Effects of the Sampling Time and Identification of Nonlinear Models” *International Journal of Bifurcation and Chaos*, 5(6):1541-1556.
- Billings, S. A. and Chen, S. [1989a] “Representation of Non-linear Systems: The NARMAX Model” *International Journal of Control*, 49(3):1013-1032.

- Billings, S. A. and Chen, S. [1989b] "The Identification of Linear and Non-linear Models of a Turbocharged Automotive Diesel Engine" *Mechanical Systems and Signal Processing*, 3(2):123-142.
- Billings, S. A. and Fakhouri, S. Y. [1978] "Identification of Nonlinear Systems Using Wiener Model" *Electronics Letters*, 13(17):502-504.
- Billings, S. A. and Fakhouri, S. Y. [1982] "Identification of Systems Containing Linear Dynamic and Static Nonlinear Elements" *Automatica*, 18(1):15-26.
- Billings, S. A. and Leontaritis, I. J. [1981] "Identification of Nonlinear Systems Using Parameter Estimation Techniques" *IEE Conference, Control and Its Applications*, Warwick University.
- Billings, S. A. and Fadzil, M.B. [1985] "The Practical Identification of Systems with Nonlinearities" In *Proceedings of the 7th IFAC, Symposium on Identification and System Parameter Estimation*, York, U.K., pp155-160.
- Billings, S. A. and Tao, Q. H. [1991] "Model Validation Tests for Nonlinear Signal Processing Applications" *International Journal of Control*, 54:157-194.
- Billings, S.A. and Voon, W.S.F. [1983] "Structure Detection and Model Validity Tests in the Identification of Nonlinear Systems" *IEE Proceedings*, 130(4):193-199.
- Billings, S. A. and Voon, W. S. F. [1984] "Least Squares Parameter Estimation Algorithms for Non-Linear Systems" *International Journal Systems Science*, 15(6):601-615.
- Billings, S. A. and Voon, W. S. F. [1986a] "A Prediction-error and Stepwise-regression Estimation Algorithm for Nonlinear Systems" *International Journal of Control*, 44(3):803-822.
- Billings, S.A. and Voon, W. S. F. [1986b] "Correlation Based Model Validity Tests for Nonlinear Models" *International Journal of Control*, 44(1):235-244.
- Billings, S. A. and Voon, W. S. F. [1987] "Piecewise Linear Identification of Non-linear Systems" *International Journal of Control* 46:215-235.
- Billings, S. A., Chen, S. and Korenberg, M. J. [1989] "Identification of MIMO Nonlinear Systems Using a Forward-regression Orthogonal Estimator" *International Journal of Control*, 49(6):2157-2189.
- Billings, S. A., Korenberg, M. J. and Chen, S. [1988] "Identification of Nonlinear Output Affine Systems Using an Orthogonal Least Squares Algorithm" *International Journal of Systems Science* 19:1559-1568.
- Box, G. E. P. and Jenkins, G. M. [1976] "*Time Series Analysis: Forecasting and Control*" Holden-Day Series in Time Series Analysis and digital Processing, Holden-day.
- Casdagli, M. [1989] "Nonlinear Prediction of Chotic Time Series" *Physica D*, (35):335-356.
- Chen, S. and Billings, S. A. and Luo, W. [1989] "Orthogonal Least Squares Methods and Their Application to Non-Linear System Identification" *International Journal of Control*, 50(5):1873-1896.

- Çinar, A. [1995] “Nonlinear Time Series for Multivariable Dynamic Processes” *Chemometrics and Intelligent Laboratory Systems*, 30:147-158.
- Diaz, H. and Desrochers, A. A. [1988] “Modeling of Nonlinear Discrete-time Systems from Input-Output Data” *Automatica*, 24(5):629-641.
- Doebelin, E. O. [1980] “*System Modeling and Response - Theoretical and Experimental Approaches*” John Wiley & Sons, Inc.
- Elsner, J. B. [1992] “Predicting Time Series Using a Neural Network as a Method of Distinguishing Chaos from Noise” *Journal of Physics A* (25):843-850.
- Eskinat, E., Johnson, S. H. and Luyben, W. L. [1991] “Use of Hammerstein Models in Identification of Nonlinear Systems” *AIChE Journal*, 37:255-268.
- Eykhoff, P. [1974] “*System Identification - Parameter and State Estimation*” John Wiley & Sons.
- Fonseca, C. M., Mendes, E. M., Fleming, P. J. and Billings, S. A. [1993] “Nonlinear Model Term Selection with Genetic Algorithms” *IEE Workshop on Natural Algorithms in Signal Processing*, Clemsford, Essex, U.K, pp 27/1-27/8.
- Genesio, R. and Tesi, A [1993] “Distortion Control Of Chaotic Systems: The Chua’s Circuit” *Journal of Circuits, Systems, and Computers*, 3(1):151-171.
- Godfrey, K. R. [1986] “Practical Problems In Identification” *Signal Processing For Control Conference, Lecture notes in Control and Information Sciences*, 358-386.
- Gray, G. J., Li, Y., Murray-Smith, D. J. and Sharman, K. C. [1996] “Structural System Identification Using Genetic Programming and a Block Diagram Oriented Simulation Tool” *Electronic Letters*, 32(15):1422-1424.
- Haber, R. [1985] “Nonlinearity Tests for Dynamic Process” *In IFAC Identification and System Parameter Estimation*, 409-413, York, U.K.
- Haber, R. and Keviczky, L. [1985] “Identification of ‘Linear’ Systems Having Signal-dependent Parameters” *International Journal Systems Science*, 16(7):869-884.
- Haber, R. and Unbehauen, H. [1990] “Structure Identification of Dynamic Systems - A Survey on Input/Output Approaches” *Automatica*, 26(4):651-677.
- Hannan, E. J. and Quin, B. G. [1979] “The Determination of the Order of an Autoregression” *J. R. Statist. Soc. B*(41):190-195.
- Hougen, J. O. [1972] “*Measurements and Control Applications for Practicing Engineers*” Cahnes Books, Massachussets, U.S.A.
- Jácome, C. R. F. e Aguirre, L. A. [1996a] “Uso de Técnicas de Identificação na Modelagem de Válvulas para Controle de Processos” *I Congresso Mineiro de Automação e IV Simpósio Regional de Instrumentação*, p.37-45, Belo horizonte - MG.
- Jácome, C. R. F. e Aguirre, L. A. [1996b] “Manual de Utilização de Rotinas MATLAB para Análise de Agrupamentos de Termos e Coeficientes de Agrupamentos de Modelos

Polinomiais NARMAX” *Relatório de Pesquisa, CPDEE/LCPI/08-96*, Centro de Pesquisa e Desenvolvimento em Engenharia Elétrica, Universidade Federal de Minas Gerais.

Jang, H. and Kim, K. [1994] “Identification of Loudspeaker Nonlinearities Using the NARMAX Modeling Technique” *J. Audio Eng. Soc.*, 42(1/2):50-59.

Johansen, T. A. and Foss, B. A. [1992] “A NARMAX Model Representation for Adaptive Control Based on Local Models” *Modeling, Identification and Control*, 13(1):25-39.

Jota, F. G.; Braga, A. P.; Polito, C. M. and Pena, R. T. [1995] “Development of an Interacting Tank System for the Study of Advanced Process Control Strategies”, *Proceedings of the 38th IEEE Midwest Symposium on Circuits and Systems*, pp 441-445, Rio de Janeiro.

Kalman, R. E. [1960] “A New Approach to Linear Filtering and Prediction” *Transactions ASME, Journal of Basic Engineering*, Series D, (82):342-345.

Kashyap, R. L. [1977] “A Bayesian Comparison of Different Classes of Dynamic Models Using Empirical Data” *IEEE Transactions on Automatic Control*, AC(22), 715-727.

Korenberg, M. J.; Billings, S. A.; Liu, Y. P., and McIlroy, P.J. [1988] “Orthogonal Parameter Estimation Algorithm for Nonlinear Stochastic Systems” *International Journal of Control*, 48(1):193-210.

Kortmann, M.; Janiszowski, K. and Unbehauen, H. [1988] “Application and Comparison of Different Identification Schemes Under Industrial Conditions” *International Journal of Control*, 48(6):2275-2296.

Kortmann, M. and Unbehauen, H. [1988] “Two Algorithms for Model Structure Determination of Nonlinear dynamic Systems with Applications to Industrial Processes” *IFAC Identification and Systems Parameter Estimation*, pp 649-656.

Leontaritis, I. J. and Billings, S. A. [1985a] “Input-Output Parametric Models for Nonlinear Systems Part I: Deterministic Nonlinear systems” *International Journal of Control*, 41(2):303-328.

Leontaritis, I. J. and Billings, S. A. [1985b] “Input-Output Parametric Models for Nonlinear Systems Part II: Stochastic Nonlinear systems” *International Journal of Control*, 41(2):329-344.

Leontaritis, I. J. and Billings, S. A. [1987] “Experimental Design and Identifiability for Nonlinear Systems” *International Journal of Systems and Science*, 18(1):189-202.

Lindskog, P. and Ljung, L. [1994] “Tools for Semi-Physical Modeling” Linköping University, Report *LiTH-ISY-R-1599*.

Ljung, L. [1987] “*System Identification Theory for the User*” Prentice Hall Inc.

Ljung, L. [1994] “System Identification in a MIC Perspective” *Modeling, Identification and Control*, 15(3):153-159.

Ljung, L. and Guo, L. [1995] “The Role of Model Validation for Assessing the Size of the Unmodeled Dynamics” Linköping University, Report *LiTH-ISY-R-1803*, ISSN 1400-3902.

Ljung, L. and Söderström, T. [1983] “Theory and Practice of Recursive Identification” *MIT Press*, Cambridge, Mass.

MathWorks, Inc. [1990a] PC-MATLAB™ for MS-DOS Personal Computers. Software Users Guide.

MathWorks, Inc. [1990b] Signal Processing Toolbox - for use with MATLAB™ Software Users Guide.

MathWorks, Inc. [1990c] System identification Toolbox - for use with MATLAB™ Software Users Guide.

MathWorks, Inc. [1990d] Control System Toolbox - for use with MATLAB™ Software Users Guide.

McMillan, G. K. [1994] “*Tuning and Control Loop Performance*” Instruments Society of America, Third Edition.

Nagy, P. A. J. and Ljung, L. [1992] “System Identification Using Boundgraphs” Linköping University, Report *LiTH-ISY-I-1337*.

Norton, J. P. [1986] “*An Introduction to Identification*” Academic Press.

Ogata, K. [1993] “*Engenharia de Controle Moderno*” Prentice Hall.

Parameswaran, V. and Raol, J. R. [1994] “Estimation of Model Error for Nonlinear Systems” *IEE Proc.-Control Theory Appl.*, 141 (6):403-408.

Pena, R. T. [1983] “Development of a Boiler Turbine Mathematical Model for Power Plant Operation Studies” The University of Texas, Austin, U.S.A. PhD dissertation, EE Dept.

Phillips, L. C. and Nagle, H. T. [1995] “*Digital Control System Analysis and Design*” Prentice Hall, International Editions.

Pottmann, M.; Unbehauen, H. and Seborg, D. E. [1993] “Application of a General Multi-model Approach for Identification Highly Nonlinear Processes-A Case Study” *International Journal of Control*, 57(1):97-120.

Pröll, T. and Karim, M. N. [1994] “Model-Predictive Ph Control Using Real-Time NARX Approach” *AIChE Journal*, 40(2):269-282.

Rodrigues, G. G. [1996] “Identificação de Sistemas Dinâmicos Não-lineares Utilizando Modelos Polinomiais NARMAX - Aplicação a Sistemas Reais” *Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica*, Universidade Federal de Minas Gerais.

Rodrigues, G. G. e Aguirre, L. A. [1996] “Manual de Utilização de rotinas MATLAB para a Identificação de Modelos NARMAX Polinomiais” *Relatório de Pesquisa, CPDEE/LCPI/05-96*, Centro de Pesquisa e Desenvolvimento em Engenharia Elétrica, Universidade Federal de Minas Gerais.

Rodrigues, G. G., Aguirre, L. A., Freitas, U. S. and Pena, R. T. [1996] “Identificação de um Forno Elétrico Não-linear Utilizando Modelos Polinomiais NARMAX” *XI Congresso Brasileiro de Automática*, p953-958, São Paulo - SP.

Schoukens, J.; Pintelon, R. and Van Hamme, H. [1994] “Identification of Linear Dynamic systems Using Piecewise Constant Excitations: Use, Misuse and Alternative” *Automatica*, 30(7):1153-1169.

Shinskey, F. G. [1988] “*Process Control Systems*” Chemical Engineering Series, McGraw-Hill International Editions.

Söderström, T. and Stoica, P. [1989] “*System Identification*” Series in Systems and Control Engineering, Prentice Hall.

Sorenson, H. W. [1970] “Least-squares Estimation from Gauss to Kalman” *IEE Spectrum*, 7:6368.

Stoica, P.; Eykhoff, P.; Janssen, P. and Söderström [1986] “Model Structure Selection by Cross-Validation” *International Journal of Control*, 43(6):1841-1878.

Thouverez, F. and Jezequel, L. [1996] “Identification of Narmax Models on a Modal Base” *Journal of Sound and Vibration*, 189(2):193-213.

Valverdú, M.; Korenberg, M. J. and Caminal, P. [1992] “Model Identification of the Neural Control of the Cardiovascular System Using NARMAX Models” *IEEE Conference*.

Wang, L. and Cluett, W. R. [1996] “Use of PRESS Residuals in Dynamic System Identification” *Automatica* 32(5):781-784.

Wiener, N. [1958] “Nonlinear Problems in Random Theory” *The Technology Press*, M.I.T. and John Wiley and Sons, New York, NY.

Wigren, T. [1990] “Recursive Identification Based on the Nonlinear Wiener Model” Doctoral Thesis at Uppsala University.

Zhao, M. W and Lu, Y. Z [1991] “Parameter Identification and Convergence Analysis Based on the Least-Squares Method for a Class of Non-Linear Systems” *International Journal Systems Science*, 22(1):33-48.