

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/293419148>

Identifying non-linear systems using NARMAX polynomial models – a review of novel results

Article · May 1998

CITATIONS

2

READS

331

3 authors, including:



Luis Antonio Aguirre

Federal University of Minas Gerais

369 PUBLICATIONS 4,913 CITATIONS

[SEE PROFILE](#)



Giovani G Rodrigues

Centro Federal de Educação Tecnológica de Minas Gerais

9 PUBLICATIONS 101 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Applications [View project](#)



Modelling tumor growth [View project](#)

IDENTIFICAÇÃO DE SISTEMAS NÃO-LINEARES UTILIZANDO MODELOS NARMAX POLINOMIAIS — UMA REVISÃO E NOVOS RESULTADOS

LUIS A. Aguirre†

Centro de Pesquisa e Desenvolvimento em Engenharia Elétrica

† Departamento de Engenharia Eletrônica
Universidade Federal de Minas Gerais
Av. Antônio Carlos 6627, 31270-901,
Belo Horizonte, MG

Fax: (031) 499-5480 E-mail:
aguirre@cpdee.ufmg.br

Giovani G. Rodrigues

Departamento de Engenharia Elétrica,
CEFET-MG

Av. Amazonas 7675, Belo Horizonte, MG
E-mail: giovani@cpdee.ufmg.br

Cristiano R.F. Jácome

Federação das Indústrias do Estado de
Minas Gerais (FIEMG)

Av. Contorno 4520, 5º Andar, Belo
Horizonte, MG
Fax: (031) 229-6645

Resumo Este artigo tem dois objetivos principais. Em primeiro lugar, ele aborda a identificação de sistemas não-lineares de forma introdutória, provendo literatura específica para um estudo mais detalhado. O segundo objetivo é apresentar alguns novos resultados úteis na seleção da estrutura de modelos NARMAX polinomiais (os quais são lineares nos parâmetros). Especificamente, discute-se como certo tipo de conhecimento prévio sobre o sistema pode ser utilizado na determinação de uma representação adequada e na análise de modelos identificados. O artigo inclui dois exemplos (um com dados simulados e outro com dados reais) para ilustrar o uso de ferramentas e os novos resultados.

Abstract. The aim of this paper is twofold. First, the fundamentals of nonlinear system identification are briefly reviewed and a list of specific references are provided for further reading. Secondly, this paper presents some novel results concerning structure selection of NARMAX polynomial models. It is shown how a certain type of prior knowledge about the system can be used to determine a proper model structure. The paper also discusses how some new results can be used to analyze identified nonlinear models. Two numerical examples are included which use both simulated and real data. Such examples illustrate the use of standard identification tools and also the use of new results.

1 INTRODUÇÃO

Um dos grandes desafios na história da ciência tem sido obter sistemas análogos aos processos e fenômenos observados no universo. Por sistema análogo entende-se um sistema capaz de reproduzir algumas características do fenômeno observado, assim como uma maquete reproduz as escalas, proporções, cores, etc. de uma construção real. Quando o análogo é um sistema matemático, ele constitui um modelo matemático do fenômeno observado.

Com a crescente disponibilidade de computadores, o uso de modelos matemáticos tem aumentado em praticamente todas as áreas do conhecimento humano. Além disso, tem-se observado uma mudança no tipo de modelos utilizados. Mais especificamente, tem havido um crescente interesse por representações não-lineares para caracterizar sistemas e fenômenos reais. Na medida em que as representações lineares são substituídas em algumas aplicações por seus correspondentes não-lineares, torna-se possível analisar e reproduzir certos fenômenos e comportamentos dinâmicos mais complexos (Aguirre, 1996). Por outro lado, a obtenção de modelos não-lineares é significativamente mais trabalhosa do que para modelos lineares e, além disso, há algumas etapas na modelagem de sistemas não-lineares que ainda não estão bem estabelecidas.

De forma geral, é possível agrupar as técnicas de modelagem em duas grandes categorias, a saber: modelagem pela física do processo e modelagem a partir de testes. A segunda categoria é normalmente conhecida por identificação de sistemas (Ljung, 1987). As diferenças entre as duas abordagens são muitas, bem

Artigo submetido em 23/04/97

Revisado em 30/06/97

Aceito sob recomendação do Ed. Cons. Prof.Dr. Paulo Sérgio P. da Silva

como as diferenças no tipo de modelos obtidos em cada caso. Entretanto, o que provavelmente mais diferencia essas duas abordagens é a quantidade de conhecimento sobre o processo real utilizado na obtenção dos modelos. Na modelagem pela física do processo, o modelo é desenvolvido a partir de toda informação disponível sobre o processo. Por outro lado, métodos de identificação normalmente não pressupõem qualquer conhecimento prévio do sistema, justificando o nome "identificação caixa-preta" (Sjöberg et al., 1996).

Ultimamente, tem havido algum interesse em desenvolver métodos que permitam incorporar alguma informação que se tenha sobre o sistema durante a sua identificação (Lindskog e Ljung, 1994). Procedimentos com esta característica são denominados métodos de "identificação caixa-cinza" e são especialmente interessantes porque não exigem do usuário um profundo conhecimento "a priori" do processo, mas permitem a utilização de conhecimento prévio. Isso normalmente resulta em modelos melhores e, principalmente, modelos fisicamente mais significativos.

A escolha de qual representação utilizar na modelagem de sistemas não-lineares ainda é uma questão que parece não ter uma resposta definitiva. A opinião dos autores sobre este assunto é que o volume e o tipo de informação "a priori" disponível sobre o sistema poderá vir a ser um elemento determinante na escolha da representação. Assim, representações que permitirem incorporar informação prévia com maior facilidade serão preferidas na identificação do tipo "caixa-cinza", enquanto outras representações continuarão a ser usadas em problemas de identificação do tipo "caixa-preta".

O objetivo do presente trabalho é rever alguns dos pontos mais importantes na identificação de sistemas não-lineares utilizando modelos polinomiais. Textos em identificação de sistemas são escassos e normalmente cobrem apenas os sistemas lineares (Norton, 1986; Ljung, 1987; Söderström e Stoica, 1989). Apesar de não se pretender aqui produzir um texto autocontido sobre identificação de sistemas não-lineares, espera-se prover o leitor com um conjunto básico de ferramentas matemáticas e referências bibliográficas para trabalhar na área. Um segundo objetivo é apresentar alguns novos resultados referentes ao uso de conhecimento "a priori" na identificação de modelos não-lineares polinomiais. As principais idéias são ilustradas com dois exemplos numéricos.

O artigo está organizado da seguinte forma: na seção 2 é feita uma revisão sobre técnicas básicas de identificação de modelos não-lineares polinomiais. A seção 3 faz uma revisão do conceito de agrupamento de termos e coeficiente de agrupamentos. Estes conceitos são úteis na determinação da estrutura de modelos não-lineares polinomiais, assunto esse discutido na seção 4. Alguns novos resultados sobre a estrutura de modelos polinomiais na presença de constante de tempo e ganho variáveis são apresentados na seção 5. Dois exemplos são discutidos na seção 6 e, finalmente, algumas observações finais são apresentadas na seção 7.

2 IDENTIFICAÇÃO DE SISTEMAS NÃO-LINEARES

O problema de identificação de sistemas pode ser dividido em cinco etapas principais (Ljung, 1987): (i) obtenção de dados de experimentação do sistema que se deseja modelar; (ii) aplicação de testes aos dados obtidos para detecção de não-linearidades; (iii) escolha da estrutura que será utilizada para

representar o modelo; (iv) estimação dos parâmetros do modelo; (v) validação do modelo obtido.

O procedimento descrito acima é empregado na identificação tanto de sistemas lineares quanto sistemas não-lineares. As principais diferenças se devem à maneira como cada passo é implementado, conforme será descrito nas subseções seguintes.

O objetivo primordial dessa seção é apresentar uma revisão dos principais tópicos relacionados à identificação de sistemas não-lineares utilizando estruturas conhecidas como modelos NARMAX polinomiais.

2.1 Experimentação do Sistema e Detecção de Não-Linearidades

Nessa etapa do procedimento de identificação, o sistema deve ser experimentado através da aplicação de entradas adequadas e da observação das saídas correspondentes (e/ou das variáveis de estado observáveis). Os dados de identificação assim obtidos serão utilizados na detecção de não-linearidades e no ajuste dos parâmetros do modelo escolhido.

Os dados utilizados na identificação devem conter informações sobre o sistema que será modelado. Assim, o sinal de excitação deve apresentar espectro suficientemente amplo em frequência e amplitude de tal forma que excursions o sistema pelos regimes dinâmicos de interesse. No caso de sistemas não-lineares, isto requer que os efeitos não-lineares sejam excitados por tais sinais e estejam presentes nos dados. A escolha da excitação adequada para a identificação de sistemas não-lineares é abordada formalmente em (Leontaritis e Billings, 1987).

O procedimento de identificação geralmente deriva modelos dinâmicos discretos a partir de dados amostrados. Existem vários métodos práticos para a determinação do período adequado para a amostragem de um sinal. Aguirre (1995) apresenta um método de seleção de período de amostragem que utiliza a auto-correlação e uma correlação não-linear para detectar o tipo de interações presentes naquele sinal. A motivação básica do método é simples: se os dados tiverem sido amostrados a uma taxa superior do que a necessária, haverá redundância entre amostras vizinhas. A fim de quantificar tal 'redundância' utilizam-se funções de auto-correlação. A novidade em (Aguirre, 1995) reside no fato de que é mostrado que se os dados tiverem informação não linear presente, a função de autocorrelação linear em alguns casos é insuficiente para quantificar a 'redundância' mencionada acima e consequentemente torna-se inadequada na escolha do tempo de amostragem. A fim de contornar o problema, foi sugerido usar, além da função de auto-correlação linear, a seguinte função

$$\Phi_{u^2, \xi}(\tau) = E\{(u^2(t) - E\{u^2(t)\})\xi(t - \tau)\} = 0 \quad \forall \tau, \quad (2.1)$$

onde $y(t)$ é o sinal de saída do sistema analisado e E indica esperança matemática. O procedimento completo pode ser enunciado como se segue: i) calcular a função de auto-correlação linear e a função (2.1), ii) determinar os atrasos, τ , para os quais tais funções passam pelos primeiros mínimos locais, iii) define-se τ_m como o menor valor entre os valores calculados em ii), iv) finalmente, o período de amostragem T_s deve ser escolhido de forma a satisfazer:

$$\frac{\tau_m}{20} \leq T_s \leq \frac{\tau_m}{10}. \quad (2.2)$$

Os algoritmos de detecção de não-linearidades são utilizados para quantificar o nível de interações não-lineares encontradas nos dados de identificação (Billings e Voon, 1983; Haber, 1985; Haber e Unbehauen, 1990).

Billings e Voon (1983, 1986) mostraram que a relação

$$\Phi_{y^2, y^2}(\tau) = E\{(y(t) - E\{y(t)\})(y^2(t - \tau) - E\{y^2(t)\})\} = 0 \quad \forall \tau \quad (2.3)$$

é válida se e somente se o sistema original for linear. Os limites de um intervalo de confiança de 95% são: $\pm 1,96/\sqrt{N}$, onde N é o comprimento do registro de dados disponíveis. A função de correlação (2.3) pode ser estimada utilizando os dados de identificação disponíveis. O sistema que gerou os dados deverá ser representado por um modelo não-linear quando a correlação calculada não permanece dentro do intervalo de confiança. Nesse caso, as interações não-lineares nos dados de identificação são consideráveis e devem ser modeladas.

2.2 Representações Matemáticas para Sistemas Não-Lineares

Na modelagem, uma importante questão é a escolha da estrutura que deverá representar o comportamento de um sistema dinâmico. Algumas representações utilizadas na modelagem de sistemas não-lineares são: (i) redes neurais (Elsner, 1992; Masri et al., 1993); (ii) funções de base radial, RBF (Casdagli, 1989); (iii) séries de Volterra (Billings, 1980); (iv) "wavelets" (Strang, 1989); (v) funções polinomiais e racionais (Chen e Billings, 1989; Haber e Unbehauen, 1990; Foss e Johansen, 1992; Noshiro et al., 1993; Jang e Kim, 1994); (vi) equações diferenciais polinomiais (Gouesbet e Letellier, 1994). Destaca-se que os conhecidos modelos bilineares constituem uma classe especial dos modelos polinomiais não-lineares (Ljung, 1987; Chen e Billings, 1989).

Este trabalho mostra a aplicação de estruturas "não-lineares auto-regressivas com média móvel e entrada exógena" NARMAX ("non-linear auto-regressive with moving average and exogenous inputs"). Propriedades gerais deste tipo de modelos, que constituem uma representação natural para uma grande classe de sistemas não-lineares (Chen e Billings, 1989), podem ser encontradas em (Leontaritis e Billings, 1985a e 1985b). A estrutura de um modelo NARMAX monovariável com período de amostragem normalizado é:

$$y(t) = F^l(y(t-1), y(t-2), \dots, y(t-n_y), u(t-d), u(t-d-1), \dots, u(t-d-n_u+1), e(t-1), e(t-2), \dots, e(t-n_e)) + e(t), \quad (2.4)$$

onde $t=1, \dots, N$. F^l é uma função não-linear qualquer. $y(t)$, $u(t)$, $e(t)$ são saída, entrada e ruído aditivo do sistema, cujos atrasos máximos são representados por n_y , n_u , n_e . d representa o retardo ou tempo morto do sistema.

A forma da função F^l normalmente não é conhecida "a priori". Assim, a dinâmica do sistema deve ser reconstruída utilizando-

se uma aproximação para representar F^l . Possíveis aproximações para esta função são os modelos polinomiais e racionais (Chen e Billings, 1989). A aproximação polinomial de grau l para o modelo (2.4) apresenta a seguinte estrutura (Chen e Billings, 1989):

$$y(t) = \theta_0 + \sum_{i_1=1}^n \theta_{i_1} x_{i_1}(t) + \sum_{i_1=1}^n \sum_{i_2=i_1}^n \theta_{i_1 i_2} x_{i_1}(t) \cdot x_{i_2}(t) + \dots + \sum_{i_1=1}^n \dots \sum_{i_l=i_{l-1}}^n \theta_{i_1 \dots i_l} x_{i_1}(t) \dots x_{i_l}(t) + e(t), \quad (2.5)$$

onde:

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

As constantes θ_i são parâmetros que devem ser estimados para ajustar a estrutura escolhida aos dados de identificação.

Os modelos polinomiais apresentam algumas vantagens sobre as demais representações para dinâmicas não-lineares. Geralmente, é possível obter modelos NARMAX polinomiais que ajustem dados com boa exatidão desde que estes dados não apresentem variações abruptas. Outra vantagem da representação polinomial é a facilidade com que a informação analítica sobre a dinâmica do modelo pode ser obtida (Aguirre e Mendes, 1996). Por fim, os polinômios não-lineares são funções lineares nos parâmetros, o que permite a utilização de algoritmos de estimação de parâmetros para modelos lineares (Davis e Vinter, 1985; Korenberg et al., 1988; Chen et al., 1989).

Por outro lado, modelos lineares nos parâmetros (muitas vezes chamados de modelos com não linearidade fraca) requerem mais parâmetros do que modelos que são não lineares nos parâmetros (chamados de modelos com não linearidade forte). Entretanto, o que vai determinar o número de parâmetros necessários num modelo linear nos parâmetros é o tipo de regressores incluídos no modelo. Se regressores adequados forem usados, mesmo um modelo linear nos parâmetros com poucos coeficientes poderá modelar a dinâmica não linear satisfatoriamente. Determinar os 'regressores adequados' entretanto é um problema não trivial e que está longe de ser satisfatoriamente resolvido. Acredita-se que alguns dos resultados descritos no presente artigo são uma importante contribuição ao estudo desse problema. Uma clara limitação de modelos polinomiais lineares nos parâmetros é na modelagem de sistemas com não linearidade estática que não possa ser adequadamente aproximada por polinômios de baixa ordem (Aguirre, 1997a).

A identificação de modelos NARMAX multivariáveis foi analisada em (Billings et al., 1989), enquanto que diversas representações não-lineares para séries temporais (incluindo redes neurais e modelos NARMA) foram comparadas por Çinar (1995).

2.3 Detecção de Estrutura de Modelos

NARMAX Polinomiais

O número de termos possíveis em modelos polinomiais cresce bastante com o aumento do grau de não-linearidade l e dos atrasos máximos n_y, n_u, n_e . De fato, esse número pode ser determinado para modelos monovariáveis através da seguinte expressão (Korenberg et al., 1988):

$$n_\theta = M + 1, \quad (2.6)$$

onde n_θ é o número total de termos no modelo e

$$M = \sum_{i=1}^l n_i, \quad n_i = \frac{n_{i-1}(n_y + n_u + n_e + i - 1)}{i},$$

$$n_0 = 1.$$

A união de todos os termos possíveis em um modelo polinomial é denominado conjunto de termos candidatos, o qual possui normalmente muitos elementos. Entretanto, representações polinomiais concisas podem ser obtidas para uma grande diversidade de sistemas não-lineares garantindo-se que os termos importantes no modelo possam ser corretamente encontrados. O procedimento de seleção dos termos a serem incluídos em um modelo é denominado detecção de estrutura.

Çinar (1995) comenta que a utilização de uma estrutura não compatível com os tipos de não-linearidades existentes nos dados tem um efeito significativo sobre o esforço de identificação e a qualidade do modelo gerado. Assim, o processo de detecção da estrutura de um modelo dinâmico para sistemas não-lineares deve receber uma atenção especial durante o procedimento de identificação. Destaca-se também que modelos não-lineares sobreparametrizados podem apresentar regimes dinâmicos espúrios (Aguirre e Billings, 1995a).

O uso de algoritmos genéticos para detecção de termos em modelos não-lineares foi investigado por Fonseca et al. (1993). Outro método utilizado para a detecção de estrutura de modelos não-lineares é o "zeroing-and-refitting" (Kadtke et al., 1993). Neste método, os termos cujos parâmetros apresentam valores absolutos reduzidos em relação aos demais são eliminados do conjunto de termos candidatos e um novo modelo é reestimado. Thouvez e Jezequel (1996) propuseram um procedimento de seleção de estrutura de modelos NARMAX baseado em análise modal, enquanto que Wang e Cluett (1996) utilizaram uma ferramenta estatística (os resíduos PRESS) para avaliar a qualidade de uma estrutura não-linear.

A taxa de redução do erro ("error reduction ratio" ou ERR) é um critério utilizado na detecção de estrutura de modelos NARMAX polinomiais (Korenberg et al., 1988; Billings et al., 1989; Chen et al., 1989). O ERR de cada termo candidato é um número que indica a melhoria obtida na representação do sistema através da sua inclusão no modelo. O critério do ERR será definido formalmente na seção seguinte.

2.4 Estimação de Parâmetros

Determinada a estrutura do modelo, deve-se estimar seus parâmetros para aproximar o comportamento dinâmico apresentado pelo sistema original. Conforme mencionado anteriormente, os modelos NARMAX polinomiais são estruturas lineares nos parâmetros. Assim, estes parâmetros podem ser estimados através do algoritmo de mínimos quadrados lineares (Billings e Voon, 1984; Chen et al., 1989; Zhu e Billings, 1996).

A estrutura mostrada em (2.5) pode ser representada na forma do erro de predição:

$$y(t) = \sum_{i=1}^{n_\theta} p_i(t) \hat{\theta}_i + \xi(t), \quad (2.7)$$

onde os regressores $p_i(t)$ do modelo correspondem aos diferentes termos no polinômio e os θ_i são os respectivos parâmetros. O símbolo " $\hat{\cdot}$ " sobre variáveis indica valores estimados e o resíduo de identificação $\xi(t, \theta)$ é definido como:

$$\xi(t, \theta) = y(t) - \hat{y}(t, \theta), \quad (2.8)$$

$$\hat{y}(t, \hat{\theta}) = \sum_{i=1}^{n_\theta} p_i(t) \hat{\theta}_i. \quad (2.9)$$

O vetor de resíduos $\{\xi(t), t=1, \dots, N\}$ representa os erros de modelagem, o ruído aditivo do sistema e incertezas de ordem qualquer. A equação (2.9) é denominada preditor de "um-passo-a-frente" e $\hat{y}(t)$ é a predição de "um-passo-a-frente" de $y(t)$.

Os parâmetros θ_i do modelo podem ser escolhidos, dentro de um espaço de busca, de modo a minimizar a função de custo do algoritmo de mínimos quadrados:

$$J_N(\theta) = \frac{1}{N} \sum_{t=1}^N \xi^T(t, \theta) \xi(t, \theta), \quad (2.10)$$

O vetor de parâmetros estimados será omitido na representação dos resíduos de identificação e da saída predita do modelo para simplificação de notação. A equação (2.7) pode ser apresentada em notação matricial:

$$Y = P \Theta + \Xi, \quad (2.11)$$

onde:

$$Y = [y(1) \ y(2) \ \dots \ y(N)]^T, \quad \Xi = [\xi(1) \ \xi(2) \ \dots \ \xi(N)]^T, \quad P = \begin{bmatrix} p_1(1) & p_2(1) & \dots & p_{n_\theta}(1) \\ p_1(2) & p_2(2) & \dots & p_{n_\theta}(2) \\ \dots & \dots & \dots & \dots \\ p_1(N) & p_2(N) & \dots & p_{n_\theta}(N) \end{bmatrix},$$

e T indica transposição da matriz ou vetor. A matriz P é denominada matriz de regressores do modelo e Θ indica o vetor de parâmetros nominal. A solução em batelada do problema de mínimos quadrados é dada por (Golub e Van Loan, 1989):

$$\hat{\theta}_{LS} = (P^T P)^{-1} . P^T Y . \quad (2.12)$$

A matriz $P^T P$ é denominada matriz de informação ou matriz normal. Esta matriz é simétrica e positiva definida quando P tem posto pleno de colunas e positiva semidefinida em caso contrário. Entretanto, ela pode perder essa propriedade em casos de mal-condicionamento numérico grave. A equação (2.12) é denominada equação normal. A solução $\hat{\theta}_{LS}$ da equação normal existe e é única desde que $P^T P$ seja não-singular.

A estimativa obtida é dita não-polarizada ($E\{\hat{\theta}_{LS}\} = \theta$) se os resíduos $\{\xi(t), t=1, \dots, N\}$ forem brancos e não apresentarem correlação com os regressores (Davis e Vinter, 1985). Quando isso não é verificado, os resíduos apresentam alguma dinâmica que não foi devidamente explicada pelo modelo. Nesse caso, novos termos devem ser incluídos neste modelo para que as estimativas se tornem não-polarizadas e toda a dinâmica dos dados seja absorvida pelo modelo.

A formação de $P^T P$ torna-se sujeita a problemas numéricos quando P é mal-condicionada. Estes problemas podem afetar a estabilidade do algoritmo de mínimos quadrados e inviabilizar a sua solução através da equação (2.12) (Chen et al., 1989). Uma alternativa para aliviar tais problemas é a ortogonalização da matriz P . Nesta situação, as colunas de P serão não-correlacionadas e formarão uma base ortogonal para o espaço imagem de P . Após a ortogonalização da matriz de regressores, o problema de mínimos quadrados pode ser solucionado através do procedimento de Gram-Schmidt (clássico ou modificado) ou do método da transformação de Householder. Estes métodos foram desenvolvidos e comparados no contexto de sistemas não-lineares por Chen et al. (1989). Neste trabalho, os autores apresentam ainda um algoritmo de estimação ortogonal baseado na decomposição em valores singulares da matriz de regressores P . Este algoritmo pode ser utilizado para solucionar o problema de mínimos quadrados em casos onde a matriz de regressores P não tem posto pleno de colunas e a matriz $P^T P$ é singular.

No procedimento de Gram-Schmidt, uma matriz P com posto pleno de colunas é decomposta em duas submatrizes:

$$P = WA, \quad (2.13)$$

onde $A (n_\theta \times n_\theta)$ é uma matriz triangular superior com a diagonal unitária e W é uma matriz ortogonal ($W^T W = D$, sendo D uma matriz diagonal de dimensão $N \times n_\theta$). Define-se então:

$$g = A\theta, \quad (2.14)$$

Aplicando (2.13) e (2.14) na equação (2.11):

$$Y = Wg + \Xi, \quad (2.15)$$

onde as colunas de W constituem os novos regressores ortogonais do problema e g é o vetor de parâmetros para este conjunto de regressores. A solução do problema de mínimos quadrados ortogonais (2.15) é (Korenberg et al., 1988; Billings et al., 1989; Chen et al., 1989):

$$\hat{g} = D^{-1} W^T Y, \quad \hat{\theta}_{LS} = A^{-1} \hat{g}. \quad (2.16)$$

O critério do ERR foi mencionado na seção anterior e pode ser formalizado através da estrutura do modelo ortogonal (2.15). A variância dos resíduos de identificação $\xi(t)$ é igual ao erro quadrático médio da saída $y(t)$ quando nenhum termo é incluído no modelo (2.15). A cada novo termo colocado neste modelo, a variância de $\xi(t)$ é decrescida de um fator $1/N (g_i^2 w_i^T w_i)$, onde w_i indica o termo incluído e g_i o seu respectivo parâmetro. Assim, o ERR de cada termo é definido formalmente como:

$$[ERR]_i = \frac{g_i^2 w_i^T w_i}{Y^T Y}, \quad 1 \leq i \leq n_\theta. \quad (2.17)$$

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 não-lineares polinomiais. Escolhe-se o número de termos desejados para o modelo e considera-se aqueles que possuem os maiores valores de ERR (Korenberg et al., 1988; Billings et al., 1989).

Nos algoritmos mencionados nessa seção, a estimação de parâmetros é feita em batelada (ou seja, de maneira "off-line"). Algoritmos de estimação recursiva ("on line") também podem ser utilizados. Por fim, existem algoritmos para detecção de estrutura e estimação de parâmetros "on-line", também baseados na ortogonalização da matriz de regressores do modelo (Luo et al., 1994). No entanto, tais algoritmos são sensivelmente mais complexos do que aqueles disponíveis para a detecção "off-line" e sua utilidade prática ainda precisa ser estabelecida. Na prática, a detecção de estrutura de modelos não-lineares é feita de maneira "off-line", enquanto que a estimação de parâmetros pode ser feita "on-line" (Pröl e Karim, 1994).

O procedimento descrito nesta seção pode ser resumido da seguinte forma (Mendes e Aguirre, 1995)¹

1. Formar a equação matricial (2.11) com os M possíveis termos candidatos.
2. Supondo que se deseje um modelo com n_θ termos, para $k = 1$ e para $i = 1, \dots, M$ faça
 - $w_1^i = p_i$. Tome o i -ésimo regressor original para compor o 1º regressor ortogonal.
 - $g_1^i = \frac{(w_1^i)^T Y}{(w_1^i)^T (w_1^i)}$. Estime por mínimos quadrados o respectivo coeficiente.
 - $[ERR]_i = \frac{g_i^2 w_i^T w_i}{Y^T Y}$, $1 \leq i \leq n_\theta$.. Usando a definição (2.17) determine o ERR de cada possível regressor, candidato ao 1º regressor ortogonal. Então faça:

¹ No procedimento abaixo, após as fórmulas decreve-se em palavras aquilo que é feito no respectivo passo.

$$3. [ERR]_i = \frac{g_i^2 w_i^T w_i}{Y^T Y}, \quad 1 \leq i \leq n_\theta \dots \text{Escolha para ser o}$$

1º regressor ortogonal, aquele com maior ERR. O índice de tal regressor é h_1 . Portanto,

$$4. w_1 = w_1^{h_1} = p_{h_1}. \text{ Chame de } w_1 \text{ o } 1^\circ \text{ regressor ortogonal, que é a primeira coluna de } W.$$

5. Agora, para $k = 2, \dots, n_\theta$ e para $i = 1, \dots, M, i \neq h_1, \dots, i \neq h_{k-1}$ (ou seja, até completar o número de termos desejado no modelo, *para todos os regressores que ainda não foram escolhidos*) faça

$$\bullet \quad \alpha_{jk}^i = \frac{(w_j^i)^T p_i}{(w_j^i)^T (w_j^i)}, \quad 1 \leq j < k. \text{ Estime por mínimos}$$

quadrados os coeficientes (que são os elementos da matriz A em (2.13)) dos regressores ortogonais escolhidos até a presente iteração (k).

$$\bullet \quad w_k^i = p_i - \sum_{j=1}^{k-1} \alpha_{jk}^i w_j. \text{ Determine o próximo}$$

regressor ortogonal (candidato) eliminando de um regressor original o efeito dos $k-1$ regressores ortogonais escolhidos até o presente.

$$\bullet \quad g_k^i = \frac{(w_k^i)^T Y}{(w_k^i)^T (w_k^i)}. \text{ Calcule o coeficiente do}$$

regressor ortogonal determinado no passo anterior. E, finalmente

$$\bullet \quad [ERR]_i = \frac{g_i^2 w_i^T w_i}{Y^T Y}, \quad 1 \leq i \leq n_\theta \dots \text{Determine o}$$

valor do respectivo ERR.

$$\bullet \quad [ERR]_i = \frac{g_i^2 w_i^T w_i}{Y^T Y}, \quad 1 \leq i \leq n_\theta \dots \text{Escolha para ser}$$

o k -ésimo regressor ortogonal aquele regressor ortogonal, dentre os restantes, com maior ERR. Finalmente,

$$\bullet \quad w_k = w_k^{h_k} = p_{h_k} - \sum_{j=1}^{k-1} \alpha_{jk}^{h_k} w_j. \text{ Chame de } w_k \text{ o } k\text{-}$$

ésimo regressor ortogonal, que é a k -ésima coluna de W .

É interessante notar que os índices $h_1, h_2, \dots, h_{n_\theta}$ indicam quais os regressores, ou seja, quais as colunas de P que devem ser incluídas nos modelos. Também, de acordo com a equação (2.14), pode-se obter o vetor Θ a partir do vetor de parâmetros estimados $g = [g_1^{h_1} \ g_2^{h_2} \ \dots \ g_{n_\theta}^{h_{n_\theta}}]^T$. Portanto, o modelo da equação (2.7) está totalmente determinado.

2.5 Critérios de Informação

No âmbito da identificação de sistemas, existem diversos procedimentos que permitem estimar a ordem de modelos dinâmicos a partir de dados medidos. Entre tais procedimentos, destacam-se os critérios de informação de Akaike ou AIC

(Akaike, 1974), Bayes ou BIC (Kashyap, 1977), Schwarz (Crutchfield e McNamara, 1987), "entropia do modelo" (Mees, 1993) e o critério de Rissanen (Rissanen, 1978). O trabalho (Gooijer et al., 1985) apresenta alguns métodos estatísticos para determinar a ordem de modelos lineares auto-regressivos.

O critério de informação de Akaike (1974) é um dos métodos mais utilizados para estimar o número de termos em modelos dinâmicos. De acordo com o critério de Akaike, o número ótimo de termos deve minimizar a seguinte função de custo:

$$J = N \log(\text{Var}\{\xi(t)\}) + 2 n_p, \quad (2.18)$$

onde N é o comprimento dos registros de dados e n_p é o número de termos de processo no modelo. A função de custo de Akaike estabelece um compromisso entre a qualidade do ajuste aos dados de identificação (quantificada pelo primeiro termo de (2.18)) e a procura por representações parcimoniosas (quantificada pelo segundo termo de (2.18)).

A utilização dos critérios de informação no contexto de sistemas não-lineares foi discutida em (Aguirre, 1994b). É importante ressaltar que o uso de critérios de informação pressupõe que o conjunto de termos candidatos esteja pré-ordenado segundo uma certa hierarquia. No caso de modelos não-lineares polinomiais, esta hierarquia entre os termos possíveis pode ser definida através do critério do ERR.

2.6 Validação de Modelos

Uma vertente da validação de modelos dinâmicos utiliza funções de correlação para detectar possíveis dinâmicas não-modeladas nos resíduos de identificação. Tratando-se de sistemas lineares, a validação deve verificar se os resíduos $\xi(t)$ são brancos e não-correlacionados com a entrada. Esta verificação pode ser feita calculando-se as funções auto-correlação dos resíduos e correlação cruzada dos resíduos com a entrada (Box e Jenkins, 1976; Ljung, 1987). Tratando-se de sistemas não-lineares, esses testes não são suficientes pois eles falham em detectar certos tipos de dinâmica não-modelada. Nesse caso, é importante verificar se os resíduos de um determinado modelo são (linear e não linearmente) brancos com respeito à entrada, à saída e não autocorrelacionados. A fim de se fazer isso é necessário avaliar se existem correlações estatisticamente significativas nos resíduos. As seguintes funções de correlação foram projetadas com esse fim. Portanto, se tais identidades podem ser verificadas então aceita-se a hipótese de que os resíduos de modelagem são brancos. Os referidos testes são (Billings e Voon, 1986; Billings e Tao, 1991):

$$\Phi_{\xi\xi}(\tau) = E\{\xi(t)\xi(t-\tau)\} = \delta(\tau), \quad (2.19)$$

$$\Phi_{\xi u}(\tau) = E\{\xi(t)u(t-\tau)\} = 0 \quad \forall \tau, \quad (2.20)$$

$$\Phi_{\xi\xi u}(\tau) = E\{\xi(t)\xi(t-\tau)u(t-\tau)\} = 0 \quad \forall \tau, \quad (2.21)$$

$$\Phi_{u^2\xi}(\tau) = E\{(u^2(t) - E\{u^2(t)\})\xi(t-\tau)\} = 0 \quad \forall \tau, \quad (2.22)$$

$$\Phi_{u^2\xi^2}(\tau) = E\{(u^2(t) - E\{u^2(t)\})\xi^2(t-\tau)\} = 0 \quad \forall \tau, \quad (2.23)$$

$$\Phi_{\xi\xi u}(\tau) = E\{\xi(t)\xi(t-\tau)u(t-\tau)\} = 0 \quad \forall \tau, \quad (2.24)$$

$$\Phi_{\xi\xi u}(\tau) = E\{\xi(t)\xi(t-\tau)u(t-\tau)\} = 0 \quad \forall \tau, \quad (2.25)$$

$$\Phi_{\xi\xi u}(\tau) = E\{\xi(t)\xi(t-\tau)u(t-\tau)\} = 0 \quad \forall \tau, \quad (2.26)$$

$$\Phi_{\xi\xi u}(\tau) = E\{\xi(t)\xi(t-\tau)u(t-\tau)\} = 0 \quad \forall \tau, \quad (2.27)$$

$$\Phi_{(y\xi)u^2}(\tau_1) = E\{(y(t)\xi(t) - E\{y(t)\xi(t)\})(u^2(t-\tau_1) - E\{u^2(t)\})\} = 0, \quad \forall \tau_1 \quad (2.28)$$

onde $\delta(\tau_1)$ é a função delta de Dirac e o apóstrofe indica que a média foi subtraída dos sinais. Essas equações procuram determinar possíveis correlações não lineares existentes nos resíduos. Se tais correlações não forem detectadas, então diz-se que os resíduos são (linear e não linearmente) brancos.

O procedimento descrito acima é denominado validação estatística de um modelo dinâmico. Esta validação garante apenas que não existem correlações não-modeladas nos resíduos de identificação. Um modelo estatisticamente válido pode não reproduzir uma ou mais propriedades dinâmicas do sistema original (Aguirre e Billings, 1994; Aguirre e Billings, 1995a). Por esse motivo, um procedimento de validação dinâmica deve estar sempre associado com a validação estatística. Esta validação dinâmica deverá verificar se o modelo identificado reproduz as principais características dinâmicas originais. Algumas propriedades dinâmicas que podem ser utilizadas para validar modelos são (Aguirre e Billings, 1994; Fiedler-Ferrara e Prado, 1994; Letellier e Gouesbet, 1995): (i) expoentes de Lyapunov; (ii) mapas e seções de Poincaré; (iii) dimensão de correlação; (iv) diagramas de bifurcação; (v) características topológicas de atratores reconstruídos; (vi) constantes de tempo, ganhos e características estáticas conhecidas "a priori".

Na literatura, existem técnicas para a validação de modelos que não se encaixam em nenhum dos dois procedimentos de validação mencionados acima. Por exemplo, Brown e outros sugeriram a utilização do princípio da sincronização para validar modelos reconstruídos (Brown et al., 1994). De acordo com este princípio, um modelo válido é capaz de entrar em sincronia com uma série temporal gerada pelo sistema original. Em outro trabalho, Parameswaran e Raol derivaram modelos para os resíduos da identificação na tentativa de validar uma dada representação (Parameswaran e Raol, 1994).

3 AGRUPAMENTOS DE TERMOS E COEFICIENTES DE AGRUPAMENTOS

O modelo NARMAX polinomial foi definido na equação (2.5). A parte determinística (que não envolve termos contendo o ruído $\xi(t)$) do modelo polinomial pode ser reescrita como (Peyton-Jones e Billings, 1989):

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

onde:

$$\sum_{n_1, n_m}^{n_y, n_u} \equiv \sum_{n_1=1}^{n_y} \dots \sum_{n_m=1}^{n_u}.$$

Os monômios da equação (2.5) são agrupados de acordo com sua ordem m ($0 \leq m \leq l$, onde l é o grau de não-linearidade do modelo). Cada termo de ordem m contém p fatores multiplicativos em $y(t-i)$ e $m-p$ fatores multiplicativos em $u(t-j)$. Os parâmetros destes termos são representados pelas constantes $c_{p, m-p}(n_1, \dots, n_m)$, onde n_1, \dots, n_m indicam os atrasos de cada fator constituinte do monômio considerado. O período de amostragem T_s foi omitido na expressão (3.1) por questões de coesão e simplicidade.

O primeiro somatório da equação (3.1) faz referência aos monômios de (2.5), separando-os de acordo com sua ordem. O segundo somatório referencia o número de fatores em $y(t-i)$ no termo considerado. Dentro do conjunto de termos de ordem m , um termo qualquer pode ser acessado através do ajuste do valor de p adequado. Por fim, o último somatório permite que seja feita a distinção entre os termos de (2.5) através do ajuste dos atrasos de cada um dos fatores constituinte do termo.

Se o período de amostragem T_s é escolhido suficientemente pequeno ($T_s \rightarrow 0$):

$$\begin{aligned} y(t-1) &\approx y(t-2) \approx \dots \approx y(t-n_y), \\ u(t-1) &\approx u(t-2) \approx \dots \approx u(t-n_u). \end{aligned} \quad (3.2)$$

Aplicando (3.2) na equação (3.1):

$$y(t) \approx \sum_{n_1, n_m}^{n_y, n_u} c_{p, m-p}(n_1, \dots, n_m) \sum_{m=0}^l \sum_{p=0}^m y(t-1)^p u(t-1)^{m-p} \quad (3.3)$$

O conjunto de termos da forma $y(t-i)^p u(t-j)^{m-p}$ é denominado agrupamento de termos ("term cluster" em Aguirre e Billings, 1995b). Os agrupamentos de termos serão representados por $\Omega_{y^p u^{m-p}}$ ($m=0, \dots, l$ e $p=0, \dots, m$). A constante $\sum_{n_1, n_m}^{n_y, n_u} c_{p, m-p}(n_1, \dots, n_m)$ é o coeficiente do agrupamento de termos $\Omega_{y^p u^{m-p}}$ ("cluster coefficient" em Aguirre e Billings, 1995b) e será representada por $\sum_{y^p u^{m-p}}$. Todos os termos pertencentes a um dado agrupamento de termos explicam o mesmo tipo de não-linearidade no modelo.

As aproximações na equação (3.2) foram utilizadas para efeito de exposição. Na prática, as definições de agrupamentos de termos e coeficientes de agrupamento continuam válidas para os valores usuais de T_s e n_y (Aguirre e Billings, 1995b; Aguirre e Mendes, 1996). A seguir, será apresentado um exemplo que ilustra as definições acima.

Exemplo 3.1 (Rodrigues, 1996):

Seja o modelo NARMAX polinomial:

$$\begin{aligned} y(t) = & 0,44551y(t-1) + 0,57773y(t-2) - 0,83997u(t-2) + 0,64166u(t-1) \\ & - 0,00867y(t-1)^2 u(t-1) - 1,31720u(t-3)u(t-1)^2 - 0,16796y(t-3)^3 \\ & + 0,78831y(t-2)u(t-2) + 0,00056y(t-3)^2 u(t-3). \end{aligned} \quad (3.4)$$

O modelo acima foi obtido com $l = 3$, $n_y = n_u = 3$. Os nove termos do modelo foram escolhidos pelo critério do ERR dentre um conjunto de oitenta e quatro termos candidatos.

Expressando o modelo (3.4) como em (3.1):

$$y(t) = \sum_{m=0}^3 \sum_{p=0}^m \sum_{n_1, \dots, n_m}^{3,3} c_{p,m-p}(n_1, \dots, n_m) \prod_{i=1}^p y(t-n_i) \prod_{i=p+1}^m u(t-n_i), \quad (3.5)$$

onde:

$$\begin{aligned} c_{1,0}(1) &= 0,44551, & c_{0,3}(3,1,1) &= -1,31720 \\ c_{1,0}(2) &= 0,57773, & c_{3,0}(3,3,3) &= -0,16796 \\ c_{0,1}(2) &= -0,83997, & c_{1,1}(2,2) &= 0,78831 \\ c_{0,1}(1) &= 0,64166, & c_{2,1}(3,3,3) &= 0,00056 \\ c_{2,1}(1,1,1) &= -0,00867, & \text{demais } c_{p,m-p}(\cdot) &= 0. \end{aligned} \quad \sum_{n_1, n_m}^{3,3} \equiv \sum_{n_1=1}^3 \cdots \sum_{n_m=1}^3.$$

O modelo NARMAX polinomial (3.4) possui seis agrupamentos de termos distintos. Os agrupamentos do modelo e seus coeficientes são:

Agrupamento	Coeficiente
Ω_y	$\sum y = c_{1,0}(1) + c_{1,0}(2) = 1,00324$
Ω_u	$\sum u = c_{0,1}(1) + c_{0,1}(2) = -0,19831$
Ω_{yu}	$\sum yu = c_{1,1}(2,2) = 0,78831$
Ω_{y^2u}	$\sum y^2u = c_{2,1}(1,1,1) + c_{2,1}(3,3,3) = -0,0081$
Ω_{y^3}	$\sum y^3 = c_{3,0}(3,3,3) = -0,16796$
Ω_{u^3}	$\sum u^3 = c_{0,3}(3,1,1) = -1,31720$

(3.6)

4 SELEÇÃO DE ESTRUTURA DE MODELOS NÃO-LINEARES

Conforme descrito na seção 2.4, a seleção da estrutura de modelos não-lineares polinomiais pode ser implementada utilizando o critério do ERR. Entretanto, o ERR é um critério estatístico e não apresenta relações claras com aspectos dinâmicos do sistema a ser modelado. Além disso, o desempenho do ERR cai com o aumento do ruído nos dados de identificação.

O conceito de agrupamentos de termos (seção 3) possui características que permitem relacionar a estrutura de um modelo com algumas de suas características dinâmicas. Assim, este conceito pode ser utilizado para incrementar o desempenho do ERR na seleção da estrutura de modelos NARMAX. Um algoritmo de seleção de estrutura que utilize a idéia de agrupamentos de termos deverá ser capaz de determinar quais são os agrupamentos que devem compor um modelo para reproduzir um dado comportamento dinâmico. Após a determinação dos agrupamentos importantes, os termos do modelo poderão ser escolhidos através do critério do ERR.

Um agrupamento é dito espúrio quando os seus termos não são necessários para representar a dinâmica do sistema original. Por outro lado, um agrupamento de termos efetivo é essencial

para modelar aquela dinâmica. A importância de um agrupamento de termos em um modelo NARMAX pode ser quantificada pelo seu coeficiente de agrupamento. Aguirre e Billings (1995b) mostraram que os coeficientes dos agrupamentos de termos espúrios normalmente possuem valores reduzidos em relação aos coeficientes dos agrupamentos de termos efetivos. Além disso, o coeficiente de um agrupamento espúrio é muito sensível ao número de termos do modelo. Destaca-se ainda que existe uma situação na qual um pequeno valor de um coeficiente de agrupamento não indica que tal agrupamento é espúrio (Jácome, 1996).

Os modelos NARMAX polinomiais são bastante sensíveis à sobreparametrização de sua estrutura (Aguirre e Billings, 1995a). Um modelo com termos pertencentes a agrupamentos espúrios pode apresentar regimes dinâmicos não refletidos nos dados de identificação. Portanto, os modelos identificados a partir do conjunto de agrupamentos efetivos têm melhores chances de reproduzir as propriedades dinâmicas do sistema original.

4.1 Análise de Propriedades Dinâmicas Utilizando Agrupamentos de Termos

Os agrupamentos de termos permitem relacionar estrutura e características dinâmicas de modelos NARMAX polinomiais. A seguir serão mencionadas algumas relações verificadas entre agrupamentos de termos e propriedades dinâmicas de um sistema não-linear.

Agrupamentos e pontos fixos: A análise de um modelo dinâmico não-linear geralmente começa com a determinação da localização, estabilidade e simetria de seus pontos fixos (Guckenheimer e Holmes, 1983). Tais propriedades dos pontos fixos de um modelo não-linear podem ser determinadas a partir dos agrupamentos de termos que compõem a sua estrutura (Aguirre e Mendes, 1996). Pode ser demonstrado que o número de pontos fixos de um modelo poderia ser diferente do número original se todos os agrupamentos possíveis fossem incluídos na estrutura. Todo conhecimento "a priori" sobre o número e a simetria dos pontos fixos da dinâmica de um sistema pode ser utilizado para definir uma estrutura adequada para representá-la (Aguirre et al., 1997). Vale a pena ressaltar que existem algoritmos que permitem estimar localização de pontos fixos a partir de uma massa de dados (Glover e Mees, 1992).

Agrupamentos e características estáticas de mapas: O conceito de agrupamentos de termos permite derivar uma representação de análise para um modelo NARX qualquer na forma $y(k) = f(y(k-1))$. Se o polinômio f fosse agrupado (usando-se a teoria da seção 3), o resultado é uma equação algébrica que descreve a não linearidade estática do mapa em questão. Nesse caso, cada agrupamento de termos do modelo original gera um monômio em f cujo parâmetro é dado pelo coeficiente do agrupamento correspondente. Essa formulação permite observar que os agrupamentos de termos efetivos num modelo são aqueles necessários para reproduzir a característica estática original (Aguirre, 1997a) e consequentemente qualquer conhecimento *a priori* da não linearidade estática pode ser usada para definir quais agrupamentos de termos considerar na etapa de determinação da estrutura do modelo. Uma extensão desses resultados para o caso de modelos não autônomos é apresentada na próxima seção.

Agrupamentos e análise espectral de ordem elevada: Observou-se a existência de uma relação clara entre o

agrupamento de termos Ω_{y^2} e o biespectro bem como entre o agrupamento Ω_{y^3} e o triespectro (Aguirre, 1997b). Algumas relações semelhantes foram estabelecidas entre espectros cruzados e agrupamentos mistos de termos do processo. O conhecimento destas relações pode simplificar a seleção "a priori" dos agrupamentos de termos efetivos na modelagem de um dado sistema não-linear.

5 USO DE CONHECIMENTO "A PRIORI" NA SELEÇÃO DE ESTRUTURAS

A seleção da estrutura dos modelos é uma das etapas mais importantes do procedimento de identificação de sistemas não-lineares. Por esse motivo, seria interessante que qualquer conhecimento prévio do sistema pudesse ser utilizado para reduzir o conjunto de termos candidatos na identificação. Esta eliminação de termos aumentaria as chances de obter modelos NARMAX polinomiais concisos e dinamicamente válidos. Esta seção apresenta um procedimento que permite utilizar conhecimento prévio na seleção da estrutura de modelos dinâmicos quando as não-linearidades presentes surgem em função da variação do ganho e da constante de tempo do sistema (Jácome, 1996).

5.1 Análise de Modelos Discretizados

Seja um sistema representado por um modelo de primeira ordem no domínio de Laplace:

$$\frac{Y(s)}{U(s)} = \frac{K}{(\tau s + 1)}, \quad (5.1)$$

onde τ é a constante de tempo do sistema. O objetivo da presente seção é verificar quais agrupamentos de termos aparecem na estrutura de um modelo discreto correspondente ao modelo (5.1) quando este é discretizado e quando seu ganho e/ou constante de tempo são variáveis, isto é, dependem de alguma outra variável do sistema. Para atingir este fim, assume-se que o ganho e a constante de tempo do sistema dependem de variáveis do processo de forma conhecida (ver equações (5.2) e (5.3) abaixo) e tais

Nessa análise, considerar-se-á que a constante de tempo e o ganho estático do modelo (5.1) variam de modo que:

$$K_{re}(x_{k_1}, \dots, x_{k_{n_t}}) = K_0 + \sum_{i=1}^{n_t} K_{1i} x_{k_i}, \quad (5.2)$$

$$\tau(x_{\tau_1}, \dots, x_{\tau_{n_\tau}}) = \tau_0 + \sum_{i=1}^{n_\tau} \tau_{1i} x_{\tau_i}, \quad (5.3)$$

onde $K_0, K_{1i}, \tau_0, \tau_{1i}$ são constantes e x_{k_i} e x_{τ_i} são variáveis que afetam o ganho e a constante de tempo do sistema original ($i = 1, 2, \dots, n_t$). Estas variáveis podem ser expressas na forma de polinômios com n_t termos relativos a sinais do processo. Destaca-se que estas duas variáveis não precisam ser iguais.

A equação (5.1) pode ser discretizada pela aplicação da transformação de Euler:

$$\tau(y(t) - y(t-1)) + T_d y(t-1) = K T_d u(t-1), \quad (5.4)$$

onde T_d é o tempo de discretização utilizado. Substituindo as relações (5.2) e (5.3) em (5.4), obtém-se:

$$\begin{aligned} (\tau_0 + \tau_{1i} x_{\tau_i}(t)) y(t) = \\ (\tau_0 + \tau_{1i} x_{\tau_i}(t)) y(t-1) - T_d y(t-1) + (K_0 + K_{1i} x_{k_i}(t)) T_d u(t-1). \end{aligned} \quad (5.5)$$

As equações (5.4) e (5.5) podem ser rescritas, respectivamente, como:

$$\Delta y(t) = y(t) - y(t-1) = \frac{T_d}{\tau} [K u(t-1) - y(t-n)], \quad (5.6)$$

onde n é a ordem do modelo, e

$$\begin{aligned} y(t) = - \left(\frac{T_d}{\tau_0} - 1 \right) y(t-1) - \frac{\sum \tau_{1i} x_{\tau_i}}{\tau_0} [y(t) - y(t-1)] + \\ K_0 \frac{T_d}{\tau_0} u(t-1) + T_d \frac{\sum K_{1i} x_{k_i}}{\tau_0} u(t-1). \end{aligned} \quad (5.7)$$

Apesar da última equação não estar na forma preditiva, ela serve para ilustrar como é que os diversos agrupamentos de termos se formam num modelo NARMAX como consequência das não linearidades do sistema. Por exemplo, além dos termos lineares de entrada e de saída, a última equação possui termos pertencentes aos agrupamentos $\Omega_{x_{\tau_i} y}$ (induzidos em função da variação da constante de tempo) e os agrupamentos $\Omega_{x_{k_i} u}$ (consequentes da variação e do ganho do modelo). A tabela abaixo mostra os agrupamentos de termos que podem estar presentes em modelos não-lineares para refletir variações de ganho e constante de tempo, de acordo com a equação (5.7).

Tabela 5.1. Agrupamento de termos em modelos não-lineares.

Tipo de dependência	Agrupamento de termos possíveis	Agrupamento de termos indesejados
Constante de Tempo	$\Omega_0, \Omega_y, \Omega_u, \Omega_{x_{\tau_i} y}$	$\Omega_{x_{\tau_i} u}$
Ganho	$\Omega_0, \Omega_y, \Omega_u, \Omega_{x_{k_i} u}$	$\Omega_{x_{k_i} y}$

A análise de modelos de ordens superiores feita por Jácome (1996) confirma a generalidade dos resultados apresentados na tabela 5.1. A última afirmativa pode ser verificada notando-se que as svariáveis x_{k_i} e x_{τ_i} podem ser (e normalmente são) termos de ordem mais elevada. Por exemplo, supondo que $x_{k_i} = y(t-1)y(t-2)u(t-3)^2$, então $\Omega_{x_{k_i} u} = \Omega_{y^2 u^3}$. Se no presente caso o modelo for de quarta ordem em vez de segunda e, digamos, $x_{k_i} = y(t-1)y(t-4)u(t-3)^2$, o agrupamento de termos $\Omega_{x_{k_i} u}$ ainda é $\Omega_{y^2 u^3}$. Este simples exemplo revela que a tabela 5.1 independe da ordem do modelo.

5.2 Análise de Modelos NARX com Ganho e Constante de Tempo Variáveis

A fim de generalizar os resultados da seção anterior adotar-se-á o seguinte procedimento. O ponto de partida será um modelo de terceira ordem (equação (5.8)). Tal modelo pode ser reescrito no mesmo formato da equação (5.6). Por analogia, então, o que

estiver fora do colchetes pode ser igualado a $\frac{T_d}{\tau}$. Por outro

lado, com à exceção do termo $y(t-n)$ dentro do colchetes, o restante pode ser relacionado com a parcela $Ku(t-1)$ na equação (5.6) de onde pode ser obtida uma expressão para o ganho K . Procedimento semelhante foi proposto e utilizado anteriormente na literatura (Haber e Keviczky, 1985; Haber e Unbehauen, 1990).

A equação seguinte representa um modelo NARX polinomial de terceira ordem:

$$y(t) = C_0 + Ay(t-1) + By(t-2) + Cu(t-1) + Du(t-2) + Ex_{\tau_i}(t)y(t-1) + Fx_{\tau_i}(t)y(t-2) + Gx_{k_i}(t)u(t-1) + Hx_{k_i}(t)u(t-2) + Iy(t-3) + Ju(t-3) + Mx_{\tau_i}(t)y(t-3) + Lx_{k_i}(t)u(t-3). \quad (5.8)$$

Assume-se que a constante de tempo τ_1 do modelo varia de acordo com a equação (5.3), enquanto que τ_2 e τ_3 são constantes. Por sua vez, o ganho K varia de acordo com a equação (5.2).

O modelo (5.8) é bastante genérico pois os termos $x_{\tau_i}(t)y(t-1)$ e $x_{k_i}(t)u(t-1)$ podem ser quadráticos, cúbicos ou de ordens mais elevadas. Este modelo pode ser manipulado e reescrito de forma semelhante à equação (5.6) (Jácome, 1996):

$$\Delta y(t) = \frac{1 - (A + I + B)}{3 - 2A - B + (E + 2F + 3M)x_{\tau_i}(t)} \left[\begin{array}{l} \frac{C_0 + Cu(t-1) + Du(t-2) + Ju(t-3)}{1 - (A + B + I)} \\ + \frac{Gx_{k_i}(t)u(t-1) + Hx_{k_i}(t)u(t-2)}{1 - (A + B + I)} \\ + \frac{Lx_{k_i}(t)u(t-3) + (M + E + F)x_{\tau_i}(t)y(t)}{1 - (A + B + I)} \\ - y(t-3) \end{array} \right] \quad (5.9)$$

Ao comparar as equações (5.9) em regime estático e (5.6), obtêm-se as relações abaixo.

(i) para a variação da constante de tempo:

$$\frac{\tau_1 + \tau_2 + \tau_3}{T_d} = \frac{3 - 2A - B}{1 - \Sigma_y} + \frac{\sum_i ((E + 2F + 3M)X_{\tau_i})}{1 - \Sigma_y},$$

(ii) para a relação estática saída/entrada:

$$K_{re} = \frac{Y}{U} = \frac{\Sigma_0}{(1 - \Sigma_y - \sum_i (\Sigma_{x_{\tau_i}y} X_{\tau_i}))U} + \frac{\Sigma_u}{1 - \Sigma_y - \sum_i (\Sigma_{x_{\tau_i}y} X_{\tau_i})} + \frac{\sum_i (\Sigma_{x_{k_i}u} X_{k_i})}{1 - \Sigma_y - \sum_i (\Sigma_{x_{\tau_i}y} X_{\tau_i})},$$

onde X_{τ_i} e X_{k_i} são os valores estáticos (em regime) das variáveis x_{k_i} e x_{τ_i} .

O procedimento descrito pode ser generalizado na forma de agrupamentos de termos e coeficientes de agrupamentos para modelos de ordem $n \leq 3$, derivando as seguintes relações:

(i) para a constante de tempo:

$$\frac{\sum_{j=1}^n \tau_j}{T_s} = \frac{n - (n-1)_+ A - (n-2)_+ B}{1 - \Sigma_y} + \frac{\sum_i ((E + (2)_+ F + (3)_+ M) X_{\tau_i})}{1 - \Sigma_y}, \quad (5.10)$$

onde n é a ordem do modelo e

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

(ii) para a relação estática, utiliza-se a equação (5.11)

(a qual já está na forma genérica desejada).

Os resultados apresentados nessa seção podem ser úteis na recuperação de informações relativas à variação do ganho e de constantes de tempo em modelos NARX. A próxima seção apresentará um exemplo no qual estes resultados serão utilizados para recuperar uma dada característica estática não-linear.

6 EXEMPLOS DE APLICAÇÃO

Essa seção apresenta alguns exemplos que empregam as técnicas de identificação apresentadas anteriormente para derivar modelos matemáticos para sistemas não-lineares conhecidos.

6.1 Recuperação de Características Estáticas Não-Lineares

Esse exemplo pretende recuperar características estáticas de modelos não-lineares polinomiais utilizando alguns dos novos resultados descritos na seção 5. Nesse contexto, modelos NARX polinomiais serão identificados a partir de dados gerados por um modelo de Hammerstein (Billings e Fakhouri, 1982; Greblicki, 1996). Em seguida, tais modelos serão processados para tentar recuperar analiticamente a característica estática embutida na representação de Hammerstein.

O exemplo será dividido em duas etapas. Primeiro, a característica estática original deverá ser recuperada a partir de modelos identificados utilizando algum conhecimento "a priori" sobre o sistema. Em seguida, o mesmo sistema será modelado utilizando técnicas usuais de identificação "caixa-preta". Os modelos assim derivados serão empregados na recuperação da característica estática original.

Uma representação possível para sistemas dinâmicos não-lineares é formada por um modelo dinâmico linear em série com uma característica estática não-linear (Billings e Fakhouri, 1982). A representação analisada nesse exemplo é constituída por uma característica estática especial seguida de um modelo dinâmico linear de primeira ordem com ganho e constante de tempo unitários ($K = 1, \tau = 1$). Esta representação é denominada modelo de Hammerstein e apresentada na figura 6.1. A característica estática a ser recuperada é mostrada na figura 6.2.

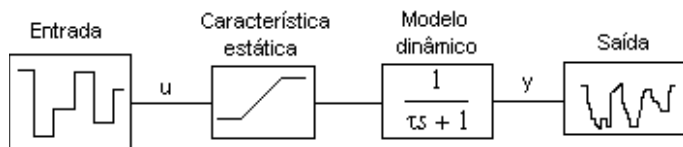


Figura 6.1. Representação de Hammerstein.

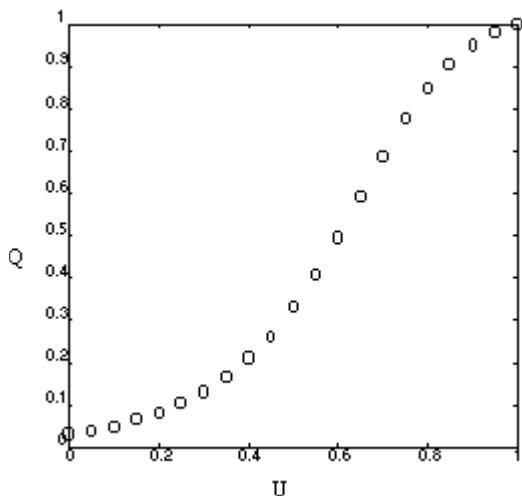


Figura 6.2. Característica estática não-linear de um modelo de Hammerstein.

A curva da figura 6.2 pode ser descrita pela seguinte equação:

$$q = q(u) = \frac{1}{\sqrt{1 + \left(\frac{1}{R^{2(u-1)}} - 1 \right)^p}}, \quad (6.1)$$

onde q é a saída, u é a entrada, enquanto que R e p são parâmetros ajustados como $R = 124$ e $p = 0,066$ (Jácome e Aguirre, 1996). Destaca-se que esta equação representa a característica estática não-linear de diversos atuadores presentes em processos industriais tais como válvulas de controle (Shinsky, 1988). Por fim, a dinâmica da representação de Hammerstein mostrada na figura 6.1 pode ser modelada pela função de transferência:

$$\frac{Y(s)}{U(s)} = \frac{q(u)}{s+1}. \quad (6.2)$$

A figura 6.3 mostra um conjunto de dados entrada-saída gerados pelo modelo de Hammerstein com a característica estática descrita pela equação (6.1). O sinal de entrada é aleatório e normalizado entre $0 \leq u \leq 1$, possuindo largura de pulso igual a 40 amostras (figura 6.3(a)). Uma sequência de ruído branco foi adicionada ao sinal de saída do modelo gerando um novo sinal com relação sinal/ruído de 80 dB (figura 6.3(b)).

O conjunto de dados apresentado na figura 6.3 será utilizado na seleção de estrutura e estimação dos parâmetros de modelos NARX polinomiais para a representação (6.2).

Na primeira etapa do exemplo, assume-se conhecida a forma geral da característica estática original. O ganho do modelo (6.2) é claramente variável com a entrada pois a inclinação da reta tangente a sua característica estática varia com o ponto de operação considerado. Observa-se que a característica estática original possui um ponto de inflexão e, portanto, um modelo NARX deverá possuir termos cúbicos para reproduzir a sua

forma. Como a característica estática está relacionada diretamente com o ganho do sistema, o agrupamento de termos $\Omega_{x_k u}$, relativo à variação do ganho deve ter grau cúbico (conforme tabela 5.1).

A equação (5.2) mostra que X_k é responsável pelas variações de ganho estático em um dado sistema. Assim, a variação do ganho do sistema (6.2) com o sinal de entrada ($K = f(u)$) implica que X_k também é função deste sinal; isto é $X_k = g(u)$. Observando que o agrupamento de termos $\Omega_{x_k u}$ deve ter grau cúbico, percebe-se que $X_k = U^2$. Essa consideração não impede que X_k dependa também de termos com grau inferior a 2 e assim $X_k = g(c_0, U, U^2)$, onde c_0 é uma constante. Em resumo, o agrupamento Ω_{u^3} deve fazer parte dos modelo NARX para representar a variação do ganho, enquanto que os agrupamentos Ω_u e Ω_{u^2} podem integrar o modelo mas não são indispensáveis. A presença do agrupamento constante Ω_0 nos modelos também é necessária pois a característica estática considerada não parte da origem (vide figura 6.2). Por outro lado, a presença dos agrupamentos Ω_{y^2} , Ω_{y^3} , Ω_{yu} , $\Omega_{y^2 u}$ e Ω_{yu^2} é indesejada pois o ganho do sistema não depende do sinal de saída e a constante de tempo é constante.

O conjunto de termos candidatos foi limitado pela análise da forma da característica estática original. Em seguida, utilizaram-se vários critérios de informação para estimar o número ideal de termos em um modelo do sistema (6.2). Os resultados obtidos são apresentados abaixo.

A figura 6.4 mostra que os critérios de informação sugerem um número ideal de treze termos no modelo. Portanto, um modelo NARX com treze termos de processo foi identificado a partir do conjunto de termos candidatos reduzido. O modelo obtido é (Jácome, 1996):

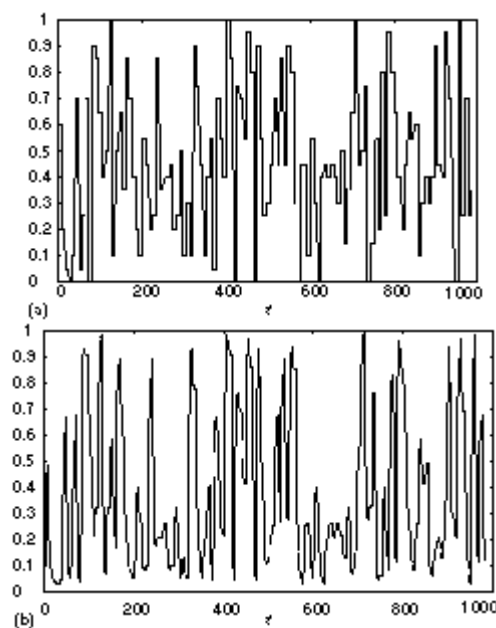


Figura 6.3. Dados de identificação gerados pela simulação de um modelo de Hammerstein (a) sinal de entrada (b) sinal de saída.

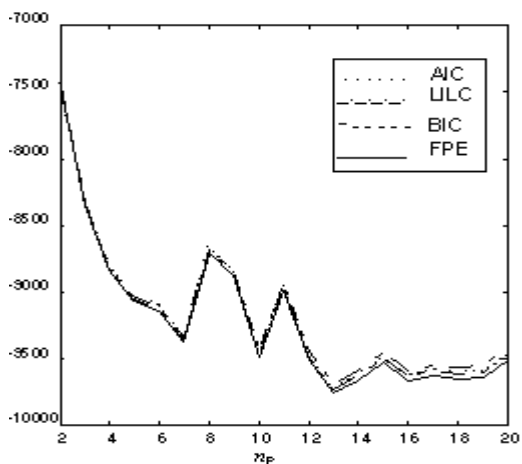


Figura 6.4. Critérios de informação na estimação do número de termos ideal em modelos NARX.

$$y(t) = 0,0091 + 0,7784y(t-1) + 1,3740u(t-1)^2 - 0,7221u(t-1)^3 - 0,2622u(t-1) + 0,1282u(t-3) + 0,0314u(t-2) - 0,2700y(t-3) - 0,2189u(t-2)^2 + 0,1150u(t-2)^3 - 0,6422u(t-3)^2 + 0,3447y(t-2) + 0,3423u(t-3)^3. \quad (6.3)$$

A figura seguinte mostra a validação dinâmica do modelo identificado. Na figura, t indica o número de amostras coletadas.

A figura 6.5 mostra que o modelo de treze termos prediz os dados de validação com boa exatidão. Portanto, este modelo será utilizado para recuperar a característica estática não-linear do modelo de Hammerstein original. Os coeficientes dos agrupamentos de termos do modelo (6.3) são:

$$\Sigma_0 = 0,0091 \quad \Sigma_y = 0,8531 \quad \Sigma_u = -0,1026 \quad \Sigma_{u^2} = 0,5128 \quad \Sigma_{u^3} = -0,2644 \\ \Omega_{x_{k_1}u} = \Omega_{u^2} \cup \Omega_{u^3} \quad \Sigma_{x_{k_1}u} = \Sigma_{u^2} = 0,5128 \quad \Sigma_{x_{k_2}u} = \Sigma_{u^3} = -0,2644.$$

Substituindo estes coeficientes na equação (5.11), obtém-se a relação estática

$$Y = 0,0618 - 0,6984U + 3,4908U^2 - 1,8033U^3. \quad (6.4)$$

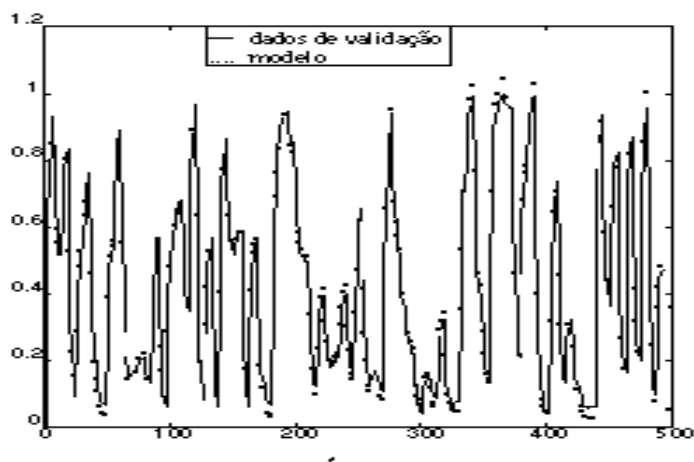


Figura 6.5. Predição de comportamento de modelo NARX identificado.

Assim, a característica estática do modelo (6.3) pode ser recuperada de forma analítica e direta através da atribuição de valores para U na equação (6.4) obedecendo à faixa $0 \leq U \leq 1$.

A figura 6.6 mostra a característica estática recuperada dessa maneira e também a característica estática original.

Em uma segunda etapa do exemplo, os modelos NARX para o sistema (6.2) foram derivados através das técnicas de identificação "caixa-preta". Observou-se que os melhores modelos "caixa-preta" identificados não podem ser colocados na forma $y = f(u)$ pois eles possuem termos não-lineares em $y(t)$. Em outras palavras, eles possuem termos de agrupamentos indesejados (vide tabela 5.1). Assim, a característica estática destes modelos só pode ser recuperada por simulação ou pelo cálculo dos seus pontos fixos (Jácome, 1996). Além disso, as características estáticas recuperadas a partir dos modelos não reproduziram a curva original com exatidão aceitável.

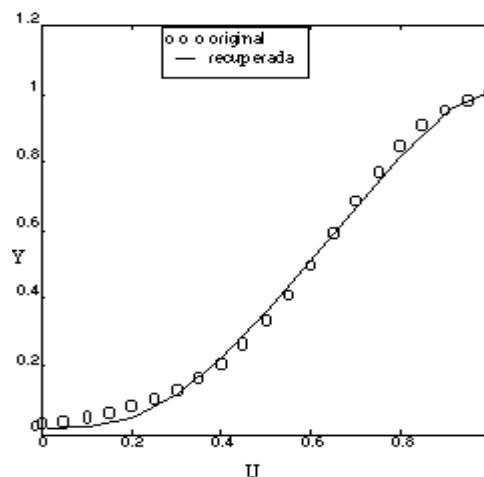


Figura 6.6. Característica estática recuperada a partir do modelo (6.3).

Esse exemplo mostrou que apenas um modelo identificado utilizando conhecimento prévio permitiu recuperar analiticamente a característica estática mostrada na figura 6.2. Nesse caso, a utilização de informações prévias sobre o sistema original permitiu selecionar a estrutura adequada para representar a sua dinâmica. Essa observação reitera a importância de uma estrutura concisa e escolhida a partir de um conjunto de termos candidatos efetivos.

6.2 Modelagem de um Forno Elétrico Real

Este exemplo mostra a modelagem de um forno elétrico real utilizando modelos NARMAX polinomiais (Rodrigues, 1996). O forno modelado é uma caixa metálica com um elemento de aquecimento interno de 200W. O forno não é isolado termicamente, de modo que variações na temperatura ambiente afetam as constantes de tempo envolvidas.

Devido ao mecanismo interno de transferência de calor, o processo é altamente não-linear e dependente da condição de operação. Além disso, ele apresenta diferentes constantes de tempo de aquecimento e resfriamento, caracterizando uma dinâmica bilinear (Abreu, 1993). Estes fenômenos não-lineares não podem ser reproduzidos por modelos lineares convencionais.

A figura 6.7 apresenta dados de entrada e de saída do forno. A variável de saída é a temperatura normalizada da superfície externa do forno e a variável de entrada é o sinal de tensão aplicado ao elemento de potência. O período de amostragem foi determinado a partir do método mencionado na seção 2.1.

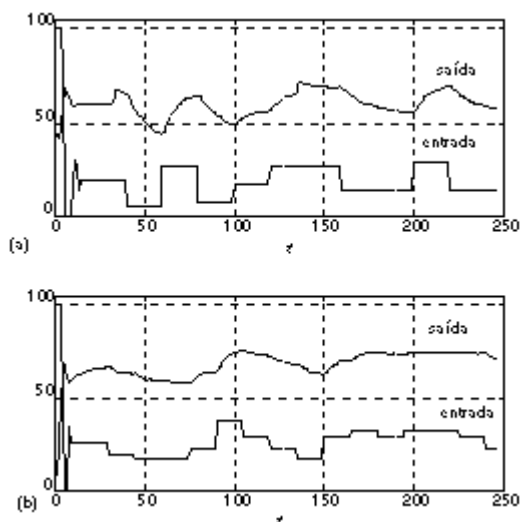


Figura 6.7. Dados de identificação. (a) Massa de dados *frq1* (b) Massa de dados *frq2*.

Os dados de identificação do forno elétrico foram obtidos escolhendo-se um a cada três pontos das massas de dados coletadas, o que resultou em um período de amostragem final de $T_s = 210s$.

A massa de dados *frq2* foi utilizada na detecção de estrutura dos modelos e na estimação de seus respectivos parâmetros. Por outro lado, os dados *frq1* foram utilizados na validação dos modelos obtidos. Os conjuntos de dados apresentados foram obtidos em dias distintos, durante ensaios totalmente independentes. Um ensaio para a aquisição de um conjunto de dados com 250 amostras tem duração aproximada de 4 horas e 50 minutos.

O resultado do teste de detecção de não-linearidades (equação (2.3)) mostrou um elevado nível de interações não-lineares presentes nos dados, o que justifica a utilização dos modelos não-lineares para descrever a dinâmica do sistema. O resultado deste teste é apresentado na figura 6.8.

Os termos dos modelos foram escolhidos dentro de um conjunto de 57 termos candidatos utilizando o critério do ERR (equação (2.17)). Os parâmetros dos modelos foram estimados através do algoritmo de mínimos quadrados ortogonal. Utilizou-se um modelo linear de ruído, com 17 termos, para minimizar a polarização dos parâmetros. Alguns dos melhores modelos identificados são apresentados a seguir. Outros modelos NARMAX polinomiais para o forno elétrico são apresentados em (Rodrigues, 1996; Rodrigues et al., 1996).

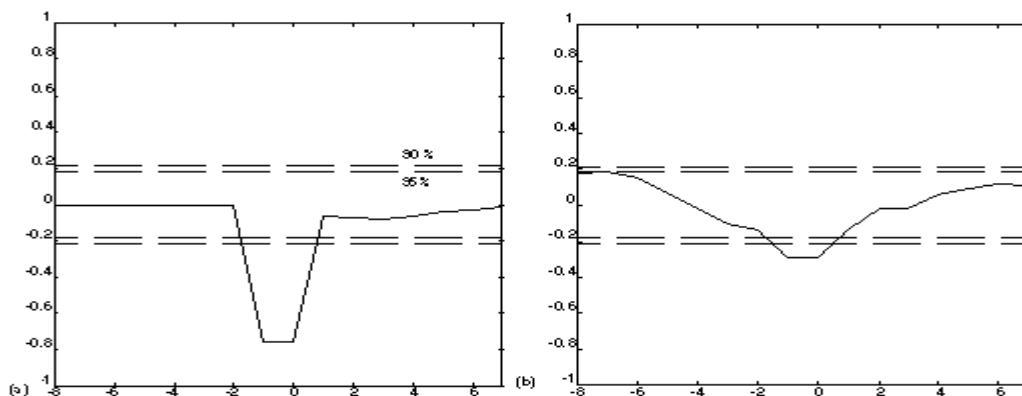


Figura 6.8. Detecção de não-linearidade nos dados do forno utilizando a função de correlação (2.3). (a) dados *frq1* (b) dados *frq2*.

(i) modelo não-linear *nl251*

$$y(t) = 0,672744y(t-1) + 0,234497y(t-2) - 0,038317u(t-2) + 0,515512u(t-1) - 0,000046y(t-1)^2u(t-1). \quad (6.5)$$

(ii) modelo linear *ml241*

$$y(t) = 1,6797y(t-1) - 0,67529y(t-2) - 0,24537u(t-2) + 0,23560u(t-1). \quad (6.6)$$

Os algoritmos de identificação utilizados mostraram-se bastante eficientes e robustos, pois estimaram bons modelos discretos para o forno elétrico empregando uma quantidade reduzida de dados (84 pontos). A relação sinal/ruído ($20 \log(\text{Var}\{y(t)\}/\text{Var}\{\xi(t)\})$) das massas de dados apresentadas variou entre 40 e 150 dB. As ordens dos modelos apresentados são estatisticamente adequadas segundo o critério de informação de Akaike (equação 2.18).

Os resultados dos teste de validação estatística (equações (2.19)– (2.28)) indicaram que os modelos do forno elétrico foram capazes de explicar adequadamente a dinâmica contida nos dados, pois não foram detectadas correlações significativas nos resíduos de identificação. No procedimento de validação dinâmica, os modelos foram simulados com o sinal de entrada da massa de dados *frq1*. A simulação foi implementada utilizando a estrutura identificada para gerar previsões do sinal de saída "infinitos-passos-a frente" onde valores preditos pelo modelo são reincorporados no modelo recursivamente para obter novas previsões. Esse procedimento é diferente das previsões de "um passo a frente" onde o modelo é sempre inicializado com dados medidos. Consequentemente, o uso de previsões de "um passo a frente" na validação de modelos é muito menos rigoroso do que o uso de previsões "infinitos-passos-a frente", como feito no presente artigo.

O modelo não-linear *nl251* foi capaz de reproduzir, com considerável exatidão, o sinal de saída da massa de dados *frq1* quando simulado com o correspondente sinal de entrada. As primeiras 17 amostras da massa de dados *frq2* foram utilizadas para inicializar o modelo.

Um modelo linear apresenta constantes de tempo de subida e de descida idênticas. De fato, o modelo *ml241* ajusta os mesmos valores para as constantes de tempo de subida e

descida, na resposta ao degrau, com base na constante de aquecimento do forno. Por sua vez, o modelo não-linear *nl251* reproduz duas constantes distintas para o aquecimento e o resfriamento do forno elétrico (conforme pode ser visto na figura 6.9). Logo, o modelo não-linear prediz melhor a dinâmica do forno que o modelo linear.

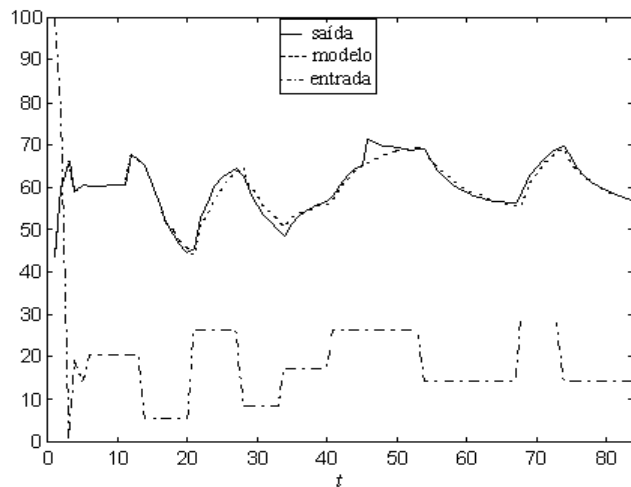


Figura 6.9. Predição de *frq1* utilizando *nl251*.

Conforme mencionado na seção 2.6, a validação estatística de um modelo dinâmico deve estar sempre conjugada com um procedimento de validação dinâmica. O modelo linear *ml241*, por exemplo, apresentou resíduos de identificação praticamente brancos e não correlacionados com a entrada e a saída do sistema. No entanto, este modelo falha em reproduzir adequadamente a dinâmica do forno elétrico, o que pode ser visto na figura 6.10.

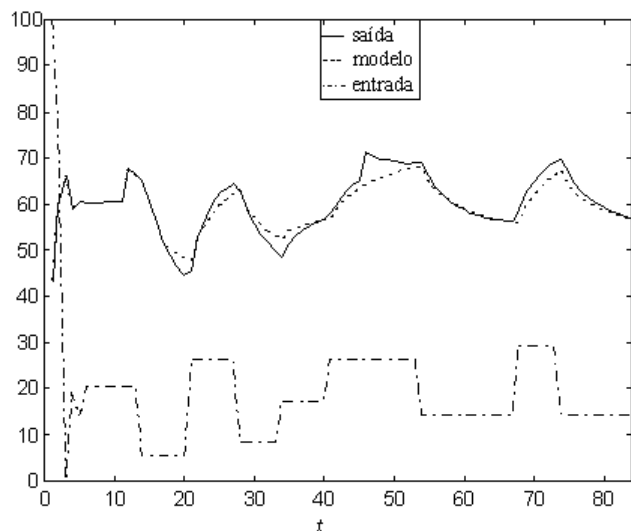


Figura 6.10. Predição de *frq1* utilizando *ml251*.

A figura 6.11 mostra a característica estática recuperada a partir do modelo não-linear *nl251*. Esta característica estática representa o ganho do modelo e só pode ser recuperada por simulação ou pelo cálculo dos pontos fixos deste modelo (vide seção 5). Essa curva é bastante parecida com a característica estática não-linear medida diretamente no forno elétrico (Abreu, 93). O modelo linear não reproduz a forma da característica original pois ele possui ganho constante (independente do ponto de operação considerado). Por outro lado, observou-se que o modelo não-linear reproduz a característica estática do forno com considerável exatidão. Esse

fato confirma a qualidade do modelo não-linear *nl251* identificado para representar a dinâmica do forno elétrico.

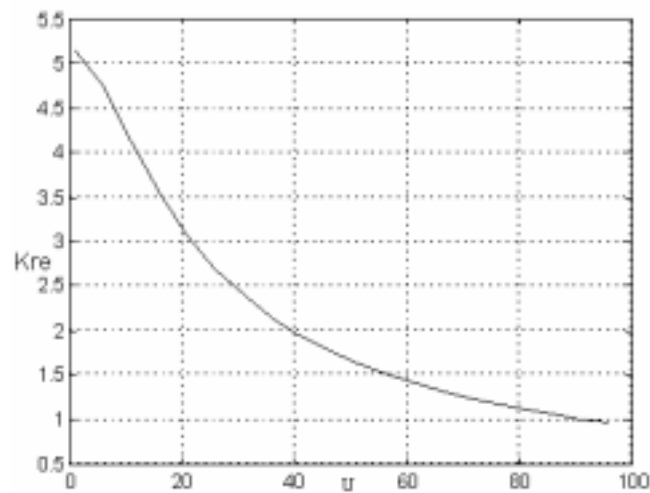


Figura 6.11. Características estáticas recuperadas a partir de modelo identificado.

7 CONCLUSÕES

Esse trabalho disponibiliza um resumo teórico e experimental sobre a identificação de sistemas dinâmicos não-lineares utilizando modelos NARMAX polinomiais. O texto apresenta alguns conceitos sobre a identificação e provê referências atuais para aqueles que queiram estudar o assunto mais a fundo. O trabalho também descreve um conjunto de ferramentas de cunho prático que permitem operacionalizar as diversas etapas do procedimento de identificação.

O artigo também apresenta alguns novos resultados na identificação dos modelos NARMAX polinomiais. A seção 5 descreve relações existentes entre a estrutura de um modelo e as características estáticas do sistema original. Observou-se que certos agrupamentos de termos são essenciais para representar uma dada característica, enquanto que outros agrupamentos devem ser eliminados do conjunto de termos candidatos para não distorcer a forma geral da mesma característica. Essa análise permite reduzir o conjunto de termos candidatos na identificação e facilita a obtenção de modelos polinomiais concisos e dinamicamente válidos.

Dois exemplos foram apresentados para ilustrar as técnicas descritas. O primeiro exemplo mostrou que a identificação tipo "caixa-preta" foi incapaz de gerar um modelo que reproduzisse a característica estática do sistema que gerou os dados. Esta característica estática só foi recuperada com exatidão a partir de um modelo identificado utilizando informação prévia. O segundo exemplo mostrou a modelagem NARMAX de um forno elétrico real. A constante de tempo do forno varia com o ponto de operação e também com a derivada do sinal de saída. Assim, a dinâmica do sistema não pode ser reproduzida por modelos lineares convencionais. Por outro lado, alguns modelos NARMAX polinomiais identificados reproduziram as características do forno elétrico com boa fidelidade.

Agradecimentos

Os dados do forno elétrico descrito na seção 6.1 foram coletados no Laboratório de Controle de Processos Industriais

(LCPI). Agradecemos especialmente ao Professor Fábio Jota pela permissão de utilizarmos esse sistema. O presente trabalho foi parcialmente financiado por CNPq (351054/95-2), FAPEMIG (TEC 917/95), CAPES e PRPq/UFGM.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abreu, S. D. G. (1993). Avaliação de desempenho de controladores auto-ajustáveis. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- 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(1995):88-94.
- Aguirre, L.A. (1996). A tutorial introduction to nonlinear dynamics and chaos Part II: Modeling and control. *Controle & Automação*, 7(1):50-66.
- Aguirre, L.A. (1997a). Recovering map static nonlinearities from chaotic data using dynamical models. *Physica D*, 100(1,2):41-57.
- Aguirre, L.A. (1997b). On the structure of nonlinear polynomial models: Higher-order correlation functions, spectra and term-clusters. *IEEE Transactions on Circuits and Systems Part I*, 4(5):450-453.
- Aguirre, L. A. e Billings, S. A. (1994). Validating identified nonlinear models with chaotic dynamics. *International Journal of Bifurcation and Chaos*, 4(1):109-125.
- Aguirre, L. A. e Billings, S. A. (1995a). Dynamical effects of overparametrization in nonlinear models. *Physica D*, 80:26-40.
- Aguirre, L. A. e Billings, S. A. (1995b). Improved structure selection for nonlinear models based on term clustering. *International Journal of Control*, 62(3):569-587.
- Aguirre, L. A. e Mendes, E. M. A. M. (1996). Global nonlinear polynomial models: Structure, term clusters and fixed points. *International Journal of Bifurcation and Chaos*, 6(2):279-294.
- Aguirre, L. A., Rodrigues, G. G. e Mendes, E. M. A. M. (1997). Nonlinear identification and cluster analysis of chaotic attractors from a real implementation of Chua's circuit. *International Journal of Bifurcation and Chaos* (no prelo).
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716-723.
- Billings, S. A. (1980). Identification of nonlinear systems - A survey. *IEE Proceedings, Part D*, 127(6):272-285.
- Billings, S. A., Chen, S. e Korenberg, M. J. (1989). Identification of MIMO non-linear systems using a forward-regression orthogonal estimator. *International Journal of Control*, 49(6):2157-2189.
- Billings, S. A. e Fakhouri, S. Y. (1982). Identification of systems containing linear dynamic and static nonlinear elements. *Automatica*, 18(1):15-26.
- Billings, S. A. e Tao, Q. H. (1991). Model validation tests for nonlinear signal processing applications. *International Journal of Control*, 54:157-194.
- Billings, S. A. e Voon, W. S. F. (1983). Structure detection and model validity tests in the identification of nonlinear systems. *IEE Proceedings, Part D*, 130(4):193-199.
- Billings, S. A. e Voon, W. S. F. (1984). Least squares parameter estimation algorithms for non-linear systems. *International Journal of Systems and Science*, 15(6):601-615.
- Billings, S. A. e Voon, W. S. F. (1986). Correlation based model validity tests for non-linear models. *International Journal of Control*, 44(1):235-244.
- Box, G. E. P. e Jenkins, G. M. (1976). Time series analysis: Forecasting and control. Holden-Day Series in Time Series Analysis and Digital Processing, Holden-Day.
- Brown, R., Rulkov, N. e Tufillaro, N. (1994). Synchronization of chaotic systems: The effects of additive noise and drift in the dynamics of the driving. *Physical Review E*, 50(6):4488-4508.
- Casdagli, M. (1989). Nonlinear prediction of chaotic time series. *Physica D*, 35:335-356.
- Chen, S. e Billings, S. A. (1989). Representation of non-linear systems: the NARMAX model. *International Journal of Control*, 49(3):1013-1032.
- Chen, S., Billings, S. A. e Luo, W. (1989). Orthogonal least squares methods and their application to non-linear system identification. *International Journal of Control*, 50(5):1873-1896.
- Chen, G. (1995). Representation, approximation and identification. Em *CRC Handbook of Circuits and Filters*, páginas 973-1006. CRC Publishers, Boston.
- Crutchfield, J. P. e McNamara, B. S. (1987). Equations of motion from a data series. *Complex Systems*, 1:417-422.
- Çinar, A. (1995). Nonlinear time series models for multivariable dynamic processes. *Chemometrics and Intelligent Laboratory Systems*, 30:147-158.
- Davis, M. H. A., Vinter, R. B. (1985). Stochastic modeling and control. Chapman and Hall, London, U.K.
- Elsner, J. B. (1992). Predicting time series using a neural network as a method of distinguishing chaos from noise. *Journal of Physics A: Mathematical Gen.*, 25:843-850.
- Fiedler-Ferrara, N., Prado, C. P. C. (1994). Caos, Uma introdução. Editora Edgar Blücher Ltda, São Paulo, Brasil.
- Fonseca, C. M., Mendes, E. M., Fleming, P. J. e Billings, S. A. (1993). Non-linear model term selection with genetic algorithms. Em *IEEE/IEEE Workshop on Natural Algorithms in Signal Processing*, páginas 27/1-27/8.

- Foss, B. A. e Johansen, T. A. (1992). A NARMAX model representation for adaptive control based on local models. *Modeling, Identification and Control*, 13(1):25-39.
- Glover, J. e Mees, A. (1992). Reconstructing the dynamics of Chua's circuit. *Journal of Circuits, Systems and Computers*, 3(1):201-214.
- Golub, G. H. e Van Loan, C. F. (1989). Matrix computations. John Hopkins, London, U.K.
- Gouesbet, G. e Letellier, C. (1994). Global vector field reconstruction by using a multivariate polynomial L_2 approximation on nets. *Physical Review E*, 49(6):4955-4972.
- de Gooijer, J. G., Abraham, B., Gould, A., Robinson, L. (1985). Methods for determining the order of an autoregressive-moving average process: a survey. *International Statistical Review*, 53(3):301-329.
- Greblicki, W. (1996). Nonlinearity estimation in Hammerstein systems based on ordered observations. *IEEE Transactions on Signal Processing*, 44(5):1224-1233.
- Guckenheimer, J. e Holmes, P. (1983). Nonlinear oscillations, dynamical systems and bifurcation of vector fields. Springer-Verlag, New York, U.S.A.
- Haber, R. (1985). Nonlinearity tests for dynamic processes. Em IFAC Identification and System Parameter Estimation, páginas 409-413, York, U.K.
- Haber, R. e Keviczky, L. (1985). Identification of 'linear' systems having signal-dependent parameters. *Int. J. Systems Science*, 16(7):869-884.
- Haber, R. e Unbehauen, H. (1990). Structure identification of nonlinear dynamic systems - A survey on input/output approaches. *Automatica*, 26(4):651-677.
- Jácome, C. R. F. (1996). Uso de conhecimento prévio na identificação de modelos polinomiais NARMAX. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- Jácome, C.R.F. e Aguirre, L.A. (1996). Uso de técnicas de identificação não-linear na modelagem de válvulas para controle de processos. I Congresso Mineiro de Automação e IV Simpósio Regional de Instrumentação, Belo Horizonte, páginas 37-45.
- Jang, H. e Kim, K. (1994). Identification of loudspeaker nonlinearities using the NARMAX modeling technique. *Journal of Audio Engineering Society*, 42(1/2):50-59.
- Kadtke, J. B., Brush, J. e Holzfuss, J. (1993). Global dynamical equations and Lyapunov exponents from noisy chaotic time series. *International Journal of Bifurcation and Chaos*, 3(3):607-616.
- Kashyap, R. L. (1977). A bayesian comparison of different classes of dynamical models using empirical data. *IEEE Transactions on Automatic Control*, 22(5):715-727.
- Korenberg, M., Billings, S. A., Liu, Y. P. e McIlroy, P. J. (1988). Orthogonal parameter estimation algorithm for non-linear stochastic systems. *International Journal of Control*, 48(1):193-210.
- Leontaritis, I. J. e Billings, S. A. (1985a). Input-output parametric models for non-linear systems part I: deterministic non-linear systems. *International Journal of Control*, 41(2):303-328.
- Leontaritis, I. J. e Billings, S. A. (1985b). Input-output parametric models for non-linear systems part II: stochastic non-linear systems. *International Journal of Control*, 41(2):329-344.
- Leontaritis, I. J. e Billings, S. A. (1987). Experimental design and identifiability for nonlinear systems. *International Journal of Systems and Science*, 18(1):189-202.
- Letellier, C. e Gouesbet, G. (1995). Topological characterization of a system with high-order symmetries. *Physical Review E*, 52(5):4754-4761.
- Lindskog, P. and Ljung, L. (1994). Tools for semi-physical modeling. Report LiTH-ISY-R-1599, Department of Electrical Engineering, Linköping University.
- Ljung, L. (1987). System identification - Theory for the user. Prentice-Hall International, New Jersey, U.S.A.
- Luo, W., Billings, S. A., Tsang, K. M. (1994). On-line structures detection and parameter estimation with exponential windowing for nonlinear systems. Research Report número 503, Department of Automatic Control and Systems Engineering, University of Sheffield, U.K.
- Masri, S.F., Chassiakos, A.G. e Caughey, T. K. (1993). Identification of nonlinear dynamic systems using neural networks. *Transactions ASME - Journal of Applied Mechanics*, 60:123-133.
- Mees, A. (1993) Parsimonious dynamical reconstruction. *International Journal of Bifurcation and Chaos*, 3(3):669-675.
- Mendes, E.M.A.M. e Aguirre, L.A. (1995). Orthogonal estimation of nonlinear models. Research Report, UFMG, CPDEE/LCPI/01-95, January 1995.
- Norton, J. P. (1986). An introduction to identification. Academic Press.
- Noshiro, M., Furuya, M., Linkens, D. e Goode, K. (1993). Nonlinear identification of PCO_2 control system in man. *Computer Methods and Programs in Biomedicine*, 40:189-202.
- Parameswaran, V. e Raol, J. (1994). Estimation of model error for nonlinear system identification. *IEE Proceedings Part D*, 141(6):403-408.
- Peyton-Jones, J. C. e Billings, S. A. (1989). Recursive algorithm for computing the frequency response of a class of non-linear difference equation models. *International Journal of Control*, 50(5):1925-1940.
- Pröl, T. e Karim, M. N. (1994). Model-predictive PH control using real-time NARX approach. *AIChE Journal*, 40(2):269-282.

- Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14:467–471.
- Rodrigues, G. G. (1996). Identificação de sistemas dinâmicos não-lineares utilizando modelos NARMAX polinomiais - Aplicação a sistemas reais. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- Rodrigues, G.G., Aguirre, L.A., Freitas, U.S. e Pena, R.T. (1996). Identificação de um forno elétrico não-linear utilizando modelos NARMAX polinomiais. XI Congresso Brasileiro de Automática, São Paulo, páginas 953-958.
- Sjöberg, J., Zhang, Q., Ljung, L., Beneviste, A., Deylon, B., Glorennec, P.Y., Hjalmarsson, H. e Juditsky, A. (1996). Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, 31(12):1691-1724.
- Söderström, T. e Stoica, P. (1989). System identification. Series in Systems and Control Engineering. Prentice Hall International.
- Strang, G. (1989) Wavelets and dilation equations: a brief introduction. *SIAM Review*, 31(4):614-627.
- Shinskey, F. G. (1988). Process control systems. Chemical Engineering Series, McGraw-Hill International Editions.
- Thouverez, F. e Jezequel, L. (1996). Identification of NARMAX models on a modal base. *Journal of Sound and Vibration*, 189(2):193-213.
- Vallverdú, M., Korenberg, M. J. e Caminal, P. (1992). Model identification of the neural control of the cardiovascular system using NARMAX models. Em IEEE Conference, páginas 585-588.
- Wang, L. e Cluett, W. R. (1996). Use of PRESS residuals in dynamic system identification. *Automatica*, 32(5):781-784.
- Zhu, e Billings, S. A. (1996). Fast orthogonal identification of nonlinear stochastic models and radial basis functions. *International Journal of Control*, 64(5):871-886.