

IDENTIFICAÇÃO DE SISTEMAS DINÂMICOS NÃO LINEARES E PROBLEMAS INVERSOS

*Ubiratan S. Freitas*¹ and *Elbert E. N. Macau*²

¹Lab. Ass. de Computação e Matemática Aplicada, INPE, S. José dos Campos, Brazil ubiratan@lac.inpe.br
²elbert@lac.inpe.br

Abstract: Neste trabalho, a identificação de sistemas dinâmicos não lineares é analisada. Diversos métodos de identificação são colocados num arcabouço comum e avaliados no escopo de um problema inverso. Essa abordagem permite identificar um problema fundamental com os métodos mais usados e apontar uma possível solução. Essas idéias são ilustradas através um exemplo numérico.

Keywords: Identificação, sistemas não lineares, problemas inversos.

1. INTRODUÇÃO

Os modelos matemáticos são uma das mais importantes ferramentas das ciências exatas. Eles permitem que se realizem diversas análises quantitativas dos sistemas que eles representam, como observar o comportamento no caso de uma variação paramétrica ou até mesmo prever comportamentos futuros. Outra grande utilidade dos modelos é que eles viabilizam a *simulação* dos mesmos sistemas, sob as mais diversas condições desejadas.

Essas utilidades fazem dos modelos matemáticos uma ferramenta tecnológica importante e muitas vezes essencial em várias atividades. Numa indústria, por exemplo, podem-se usar modelos para auxiliar o projeto de controladores automáticos e assim automatizar processos. Outro exemplo é o de uma usina hidrelétrica que, para ser construída, precisa, entre outras coisas, de um modelo de vazão para o rio onde será instalada.

Uma das principais dificuldades da análise de sistemas baseada em modelos é a obtenção dos últimos, também chamada *modelagem*. Há duas formas de se obter um modelo. Na primeira, chamada modelagem por primeiros princípios ou pela física do processo [16], divide-se o sistema em estudo em subsistemas cujos modelos são conhecidos e constrói-se o modelo global a partir das interrelações dos subsistemas. Essa abordagem tem a vantagem de gerar modelos gerais, que são válidos para toda uma classe de sistemas. No entanto, ela requer o conhecimento dos mecanismos intrínsecos por trás do sistema. Um exemplo é o modelo para um circuito elétrico obtido a partir das leis de Kirchoff e das características dos componentes. Quando o conhecimento sobre

esses mecanismos não está disponível ou essa abordagem leva a modelos excessivamente complexos, emprega-se a segunda forma de modelagem: a *empírica*. Nesta abordagem, os modelos são obtidos a partir de dados coletados no sistema sob estudo. Os modelos assim obtidos são válidos apenas para o sistema do qual os dados foram coletados, porém nenhum conhecimento *a priori* sobre o sistema é necessário.

A identificação de sistemas é a área do conhecimento que trata da modelagem de sistemas dinâmicos a partir de dados experimentais [1, 11, 13, 18]. Embora variados, os diversos métodos de identificação possuem uma estrutura comum, que pode ser descrita de uma forma unificada, pelo menos no que tange aos métodos mais usados.

O objetivo deste artigo é descrever essa estrutura comum e apontar seus principais problemas, bem como indicar possíveis soluções. De particular interesse nesse sentido é a *formulação do problema de identificação como um problema inverso*.

Quanto à organização do artigo, inicialmente a estrutura básica dos métodos mais usados de identificação é apresentada na Seção 2. A seguir, a Seção 3 descreve a identificação de sistemas como um problema inverso. Os problemas do método tradicional de identificação e possíveis soluções são mostrados na Seção 4 e, finalmente, a Seção 5 conclui o trabalho.

2. PROCEDIMENTO DE IDENTIFICAÇÃO

A identificação de sistemas consiste no procedimento que permite obter, a partir de dados experimentais, modelos matemáticos para sistemas quaisquer [18]. O produto da identificação é, pois, um modelo matemático que representa (bem ou mal) algumas características de um determinado sistema [1, 16].

Existem diversos procedimentos disponíveis para um eventual usuário que queira identificar um modelo a partir de dados coletados. Apesar dessa multiplicidade, os métodos mais usados partilham a mesma essência e podem ser apresentados em uma forma geral comum.

Nesta Seção, esta forma geral é apresentada, com o objetivo mostrar o mecanismo de identificação e, principalmente, apontar os problemas que esses métodos de identificação compartilham. O tipo de modelo abordado

aqui é o de tempo discreto e não linear.

O processo de identificação pode ser dividido em 5 etapas principais [1]:

1. coleta de dados;
2. escolha da representação matemática a ser usada;
3. determinação da estrutura do modelo;
4. estimação de parâmetros;
5. validação do modelo.

A coleta de dados constitui o ponto de partida do processo de identificação. Esses dados são grandezas que devem ser amostradas no sistema sob estudo e que quantificam características de interesse do mesmo.

Quanto a forma de obtenção dos dados, diversas situações podem acontecer, dependendo do controle que o usuário de identificação tenha sobre o processo de coleta de dados. Num extremo, o usuário pode projetar completamente o ensaio de coleta de dados, determinando questões como formas de onda das entradas e período de amostragem adequado. No outro extremo, o usuário pode ter que partir apenas dos dados, sobre os quais ele não teve nenhuma influência no processo de obtenção, como acontece com frequência na análise de séries temporais.

A importância da coleta de dados reside no fato de que toda a informação que o método de identificação precisa sobre o sistema deve estar nos dados¹. Em outras palavras, se não houver informação sobre uma determinada característica do sistema nos dados, não há como garantir que o modelo preserve essa característica. Logo, o ensaio de coleta de dados deve procurar garantir que haja informação suficiente sobre as características de interesse nos dados e que essa informação esteja disponível para o processo de identificação.

Um dos principais parâmetros da coleta de dados é o tempo ou período de amostragem T_s . Este deve ser compatível com as escalas de tempo características do sistema. Um período de amostragem grande demais não será capaz de acomodar informação sobre os comportamentos dinâmicos mais rápidos do sistema. Por outro lado, razões práticas como capacidade de armazenamento de dados e limitações do sistema de coleta de dados impedem que se utilize tempos de amostragem arbitrariamente curtos. Deve-se, portanto, empregar um T_s intermediário. Maiores detalhes sobre como determinar T_s podem ser obtidos em [10] e em [1].

Outro ponto muito importante para o ensaio de coleta de dados é o sinal de excitação ou de entrada. Nos casos em que o sistema a ser identificado possui entradas, muitas vezes é possível, e mesmo desejável, que sinais de excitação especialmente projetados para a identificação sejam utilizados. O objetivo é garantir que o máximo de informação possível sobre o sistema esteja

¹Na verdade, existem outras formas de incluir informação sobre um sistema no processo de identificação, além dos dados [2]. No entanto, os dados são a principal fonte de informação.

presente nos dados a serem coletados. Na prática, na identificação de sistemas não lineares, isso se traduz em sinais de entrada de amplo espectro de frequências e de amplitudes variadas. Estratégias para a escolha de sinais de entrada podem ser vistas em [1].

Após a coleta dos dados, é possível realizar um pré-processamento nos mesmos antes da identificação. Um pré-processamento muito comum consiste em realizar uma *reconstrução de espaço de estados*, no qual, a partir de séries temporais de baixa dimensão, geram-se séries de dimensão mais alta. Dentre as diversas técnicas existentes para isso, a mais utilizada é a reconstrução por coordenadas de atraso. Nela, parte-se de uma série temporal $y(k)$, sendo $k = 1, 2, \dots, N$ os instantes discretos de tempo, e geram-se vetores $\mathbf{y}(k)$ dados por

$$\mathbf{y}(k) = [y(k), y(k - \tau), \dots, y(k - (d_e - 1)\tau)]^T, \quad (1)$$

sendo τ o atraso de reconstrução e d_e a dimensão do espaço de reconstrução. Tal procedimento é muito importante na identificação de sistemas de dimensão mais alta do que a dos dados coletados.

A escolha da representação matemática é uma etapa crucial. Essa escolha determina como serão as próximas etapas. Nesse ponto, o usuário de identificação deve optar entre diversas possibilidades. Por exemplo, ele pode escolher entre modelos de tempo discreto ou de tempo contínuo e entre modelos lineares e não lineares, entre muitas outras decisões.

Entre as representações para modelos dinâmicos não lineares e de tempo discreto pode-se citar as redes MLP (*Multilayer Perceptron*), [9], redes RBF's (*Radial Basis Function*), [7], lógica nebulosa, [19], modelos NARMAX polinomiais, [14, 15], entre muitas outras.

As etapas de detecção de estrutura e estimação de parâmetros, devido a sua importância, são descritas em detalhe a seguir.

A validação é a etapa na qual a qualidade dos modelos identificados é aferida. Nela são descartados os modelos que não foram capazes de representar adequadamente as características desejadas do sistema original. É importante observar que *a qualidade do modelo é um conceito relativo, pois depende do objetivo para o qual o modelo se destina*. Assim, durante a identificação, deve-se ter em mente o uso a que se destina o modelo [16].

Um emprego muito freqüente para os modelos é a predição, ou seja, usar o modelo para saber como o sistema se comportaria frente a determinadas condições. Nesse caso, deseja-se que o modelo tenha um comportamento temporal o mais próximo possível daquele do sistema original.

2.1. Determinação da estrutura e estimação de parâmetros

Qualquer representação matemática útil para identificação de sistemas deve possuir uma flexibilidade que permita se moldar a um sistema qualquer. Num exemplo didático, seja o problema de se aproximar uma função qualquer $f : \mathbb{R} \rightarrow \mathbb{R}$ por um polinômio $p(x) =$

$\sum_{i=0}^n a_i x^i$. Para se determinar um polinômio, deve-se escolher, ou ajustar, tanto o seu grau n , quanto os seus coeficientes a_i . Essas duas “escolhas” são distintas em sua natureza.

A determinação do grau n deve ser feita num espaço discreto. Não existem polinômios de grau fracionário. A escolha do grau está relacionada com a flexibilidade do polinômio. Quanto maior o grau, maior a flexibilidade. Também, o grau determina o número de coeficientes que precisarão ser ajustados.

Já o ajuste dos coeficientes é uma busca num espaço contínuo. É ajustando ou estimando os coeficientes que o polinômio genérico $p(x)$ toma uma forma específica e pode ser usado no lugar da função que ele aproxima.

A escolha do grau na aproximação polinomial é o análogo da escolha da estrutura para a identificação de sistemas. Esta última é feita num espaço discreto e influencia a flexibilidade dos modelos a serem gerados. Essa escolha depende da representação utilizada. O ajuste dos coeficientes na aproximação polinomial é o análogo da estimação de parâmetros na identificação. Esta é uma busca num espaço contínuo e é responsável por colocar o modelo numa forma específica capaz de representar o sistema.

A escolha da estrutura e a estimação de parâmetros são o cerne das técnicas de identificação, pois são esses procedimentos que geram os modelos para os sistemas.

Um modelo não linear dinâmico de tempo discreto pode ser descrito por

$$\mathbf{y}(k) = f(\mathbf{y}(k-1), \mathbf{u}(k-1)) , \quad (2)$$

sendo $\mathbf{y}(k) \in \mathbb{R}^q$ o vetor de estados no instante de tempo discreto k , $\mathbf{u}(k) \in \mathbb{R}^m$ o vetor de entradas no instante k e $f: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^q$ uma função não linear qualquer. Os vetores $\mathbf{y}(k)$ e $\mathbf{u}(k)$ podem representar tanto vetores observáveis no sistema, quanto vetores de um espaço vetorial reconstruído, vide equação (1). A equação (2) permite obter, a partir do conhecimento do estado e das entradas no instante $k-1$, o estado no instante k . Ou seja, o modelo permite avançar no tempo o conhecimento do estado. Repetindo-se o processo para o instante $k+1$, usando o valor de $\mathbf{y}(k)$ obtido no passo anterior, pode-se conhecer o valor de $\mathbf{y}(k+1)$. Assim, por iteração, é possível obter a evolução temporal dos estados. A equação (2) é, junto com o processo de iteração, um modelo não linear de tempo discreto para um sistema dinâmico.

Na identificação de sistemas, escolhe-se para $f(\cdot)$ uma representação paramétrica qualquer, como por exemplo as já citadas redes *perceptron* multicamadas. Nessa representação, $f(\cdot)$ toma uma forma que depende de uma estrutura $s \in S$ e de um vetor de parâmetros $\boldsymbol{\theta} \in \mathbb{R}^{p(s)}$, sendo S o conjunto de todas as estruturas possíveis. O número de parâmetros p depende da estrutura escolhida, enquanto o significado da estrutura e dos parâmetros depende da representação.

Por exemplo, seja o caso de um modelo com $q = 1$ e sem entradas. Uma possível representação para $f(\cdot)$ é a

polinomial. Nesse caso, a estrutura s seria um conjunto de termos de um polinômio, e S o conjunto de todos os conjuntos de termos possíveis. O vetor de parâmetros seria formado então pelos coeficientes desses termos.

Se se emprega o *perceptron* multicamadas como representação, o conjunto S é o conjunto de todas as topologias de redes com uma entrada e uma saída. Uma estrutura particular s para o modelo é uma topologia específica, determinando quantas camadas escondidas, quantos neurônios em cada camada e quais as ligações entre eles. O vetor de parâmetros é o conjunto de todos os pesos e termos de polarização (*bias*).

Assim, escolher a estrutura é escolher um s particular de S e estimar os parâmetros é tomar um vetor particular $\boldsymbol{\theta}$. Ambas ações devem ter por objetivo gerar um modelo que capture as características desejadas de um determinado sistema. Para isso lança-se mão dos dados coletados.

O procedimento adotado na maior parte das técnicas de identificação de sistemas é o de se definir uma função custo que dependa do modelo e dos dados e de se utilizar algum método de otimização.

Normalmente, a função de custo empregada em identificação é definida sobre o erro de predição de um passo à frente

$$\boldsymbol{\xi}(k) = \mathbf{y}(k) - f(\mathbf{y}(k-1), \mathbf{u}(k-1), s, \boldsymbol{\theta}) , \quad (3)$$

sendo que a dependência de $f(\cdot)$ com relação a estrutura s e ao vetor de parâmetros $\boldsymbol{\theta}$ foi tornada explícita. De agora em diante, supõe-se que $\mathbf{y}(k)$ e $\mathbf{u}(k)$, para $k = 1, 2, \dots, N$ são formados pelos dados coletados, sejam vetores observados diretamente, sejam vetores oriundos de algum procedimento de reconstrução de espaço de estados.

A função de custo normalmente utilizada é o somatório dos erros quadráticos

$$J = \sum_{k=1}^N \|\boldsymbol{\xi}(k)\|_2^2 . \quad (4)$$

Essa função depende da estrutura s e do vetor de parâmetros $\boldsymbol{\theta}$, bem como dos dados coletados. Dessa forma, os problemas de detecção de estrutura e estimação de parâmetros se resumem em escolher s e $\boldsymbol{\theta}$ de modo que J seja mínimo. Esses problemas são normalmente resolvidos separadamente, escolhendo-se primeiro uma estrutura e depois estimando os parâmetros.

Existem técnicas automáticas de detecção de estrutura para diversas representações, como modelos NARMAX polinomiais [6], racionais [5, 8], redes neurais [17], entre outras. Tais técnicas permitem escolher a estrutura de um modelo baseada nos dados de identificação. No entanto, elas apenas se aproximam da estrutura ideal. Encontrar a estrutura ótima é, em geral, um problema muito difícil.

Uma estrutura pode também ser escolhida de forma arbitrária. Dessa forma, espera-se que apenas a estimação de parâmetros seja suficiente para gerar um modelo adequado.

Uma vez obtida a estrutura, seja arbitrariamente ou utilizando algum critério, utiliza-se alguma técnica de otimização para estimar os parâmetros. Dependendo da representação escolhida, o erro de predição um passo à frente pode ser linear ou não linear nos parâmetros. No primeiro caso, a função custo é convexa em θ e o problema de otimização é mais simples de se resolver do que quando o erro de predição é não linear nos parâmetros. Neste último caso, há a possibilidade de existência de mínimos locais.

O método aqui apresentado é a forma usual de se fazer identificação de sistemas. Essa formulação tem problemas intrínsecos que dificultam a obtenção de bons modelos, principalmente para sistemas não lineares. A seguir, os principais problemas dessa abordagem são explorados. Mas, antes, é importante formular a identificação de sistemas como um problema inverso.

3. IDENTIFICAÇÃO DE SISTEMAS COMO UM PROBLEMA INVERSO

A identificação de sistemas, ou mais precisamente, as etapas de detecção de estrutura e estimação de parâmetros, consiste em gerar modelos a partir de comportamentos temporais conhecidos de sistemas dinâmicos. Esse é o problema inverso ao de se determinar, a partir do conhecimento do modelo, o comportamento temporal do mesmo.

Um problema inverso pode ser visto como o problema de se determinar as causas a partir de seus efeitos, enquanto o problema direto seria determinar os efeitos a partir do conhecimento das suas causas [4]. Matematicamente, sejam W e Z espaços normados e $K : W \rightarrow Z$ uma função. O problema direto é, dado as causas $w \in W$, determinar os seus efeitos $z \in Z$ correspondentes tal que $z = K(w)$. O problema inverso é, a partir de um dado efeito $z \in Z$ encontrar suas causas $w \in W$ tal que $K(w) = z$.

No caso da identificação de sistemas, w representa o modelo, ou, mais especificamente, uma estrutura s particular junto com um vetor de parâmetros θ específico, supondo uma representação já escolhida, enquanto z representa o conjunto de dados e K é a função que toma essa estrutura e parâmetros como entrada e gera a evolução temporal do modelo. A Figura 1 ilustra esse problema direto. O modelo, na Figura, é um mecanismo que produz o próximo estado a partir das entradas e do estado atual. Esse mecanismo depende de uma estrutura particular e de um vetor de parâmetros. O modelo é então usado em um processo iterativo que produz a evolução temporal. Na identificação de sistemas, a estrutura s e o vetor de parâmetros θ são as *causas* (w) do problema direto, enquanto a evolução temporal representa os *efeitos* (z) e está associada aos dados coletados no sistema a ser identificado.

O procedimento de identificação é o problema inverso de, a partir dos dados, obter a estrutura s e os parâmetros θ . É importante notar que existe, no problema direto, um processo de iteração que deve ser levado em

Tabela 1 – Causas e efeitos do PDC e do PDR

	PDC	PDR
Causas (x)	Estrutura (s) Parâmetros (θ) Condição inicial $\mathbf{u}(k), \forall k$	Estrutura (s) Parâmetros (θ) Estado atual ($\mathbf{y}(k-1)$) Entrada atual ($\mathbf{u}(k-1)$)
Efeitos	$\mathbf{y}(k), \forall k$	Próximo estado ($\mathbf{y}(k)$)

conta no momento de se atacar o problema inverso. Um modelo, tal qual o descrito na equação (2), apenas relaciona o próximo estado com o estado atual e as possíveis entradas. Para que esse modelo seja capaz de gerar sinais que evoluem no tempo, é necessário empregar o processo de iteração.

Um modelo da forma (2) representa um outro problema direto, no qual as causas são o estado atual, as entradas, a estrutura e os parâmetros. O efeito é o próximo estado. Esse outro problema direto é referido, daqui por diante, por *Problema Direto Reduzido*, ou PDR. A sua contra-partida, o problema direto de, a partir da estrutura e dos parâmetros, determinar *toda* a evolução temporal, é referido como *Problema Direto Completo*, ou PDC. Os problemas inversos associados são chamados *Problema Inverso Reduzido*, ou PIR, e *Problema Inverso Completo*, ou PIC, respectivamente. As características do PDC e do PDR estão listadas na Tabela 1.

Um problema inverso pode ser resolvido como um problema de otimização. Seja a função de custo dada por

$$J_{\text{pi}}(w) = \|z - K(w)\|, \quad (5)$$

sendo $\|\cdot\|$ uma norma em Z . Pode-se definir como solução do problema inverso o w^* para o qual $J_{\text{pi}}(w)$ seja mínimo.

4. PROBLEMAS DA IDENTIFICAÇÃO DE SISTEMAS TRADICIONAL

As técnicas de identificação de sistemas, em geral, recaem em um problema de otimização. É possível supor, então, que tais técnicas estão, implicitamente, resolvendo um problema inverso, *embora sem postulá-lo*. Observando a definição do erro de predição um passo à frente (3) e a função de custo J (4), nota-se que, embora J dependa de s e de θ , não há referência a nenhum processo de iteração. Em outras palavras, a função de custo J depende do modelo que fornece o próximo estado a partir do atual, mas não depende do processo que usa esse modelo para gerar a evolução temporal característica dos sistemas dinâmicos.

O resultado disso é que o problema inverso que as técnicas apresentadas na Seção 2 resolvem não é aquele de determinar qual modelo que gerou os dados coletados, pois elas não fazem referência ao processo iterativo. O problema que elas na verdade tratam é determinar qual modelo (estrutura e parâmetros) retorna o próximo estado a partir do estado atual e entradas contidos nos dados. Em resumo, elas resolvem o PIR ao invés do PIC, ou seja, elas *resolvem o problema errado*.

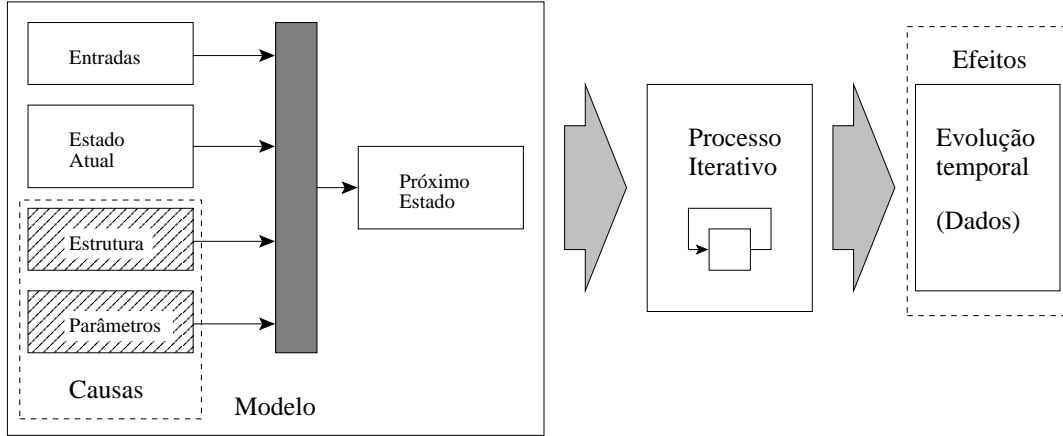


Figura 1 – Problema direto na identificação de sistemas. As causas são associadas à estrutura e aos parâmetros do modelo. Os efeitos são a evolução temporal do sistema, ou seja, os dados.

Embora a diferença entre o PIC e o PIR possa parecer sutil a princípio, ela pode comprometer seriamente o desempenho da identificação, pois a solução para o PIR não é, necessariamente, uma boa solução para o PIC, principalmente quando se trata de identificar modelos para sistemas não lineares. Por exemplo, um modelo solução para o PIR (um preditor um passo à frente para os dados) pode, quando colocado em um regime iterativo, apresentar um comportamento dinâmico completamente diferente daquele presente nos dados, podendo ser até mesmo instável. Tal modelo dificilmente poderia ser considerado como uma solução adequada para o PIC. Essa situação ocorre com frequência na identificação de sistemas não lineares [3, 10].

Um exemplo simples pode ilustrar esse problema. Suponha que o sistema a ser identificado é dado pela equação logística, um clássico exemplo de sistema caótico discreto

$$y(k) = \lambda y(k-1) [1 - y(k-1)] , \quad (6)$$

sendo λ o seu parâmetro. O valor de λ determina o comportamento temporal do sistema.

O sistema da equação logística foi simulado com $\lambda = 3,97$ gerando 500 pontos. A estes, foi adicionado ruído gaussiano i.i.d. com desvio padrão de 0,02. A identificação foi então conduzida empregando uma representação polinomial de grau 2 e um modelo de primeira ordem, ou seja, a função $f(\cdot)$ em (2) foi tomada como sendo um polinômio

$$f(y(k-1)) = \theta_0 + \theta_1 y(k-1) + \theta_2 y^2(k-1) . \quad (7)$$

Dessa forma, a estrutura foi fixada e possui um vetor de três parâmetros $\theta = [\theta_0 \ \theta_1 \ \theta_2]^T$. Esses parâmetros foram então estimados usando duas metodologias: a tradicional, via minimização do erro de predição um passo à frente (4), que resolve o PIR e uma que emprega a minimização dos erros até n passos à frente, que se aproxima de resolver o PIC. O erro de predição n passos à frente é dado por

$$\xi_n(k) = \mathbf{y}(k+n) - f^{[n]}(\mathbf{y}(k)) , \quad (8)$$

Tabela 2 – Parâmetros estimados e expoente de Lyapunov dos modelos obtidos usando diferentes estratégias e do sistema original.

	θ_0	θ_1	θ_2	Expo. de Lyapunov
Sistema	0	3,97	-3,97	0,59116
Tradicional	$1,44 \times 10^{-2}$	3,8760	-3,8662	0,51312
Mínimo J_2	$-9,70 \times 10^{-3}$	4,0046	-3,9839	0,60886
Mínimo J_3	$9,83 \times 10^{-4}$	3,9694	-3,9698	0,59432
Mínimo J_4	$3,84 \times 10^{-3}$	3,9627	-3,9692	0,59998
Mínimo J_5	$1,58 \times 10^{-3}$	3,9682	-3,9669	0,60437
Mínimo J_6	$1,33 \times 10^{-3}$	3,9670	-3,9641	0,59289
Mínimo J_7	$-3,00 \times 10^{-3}$	3,9811	-3,9725	0,60558
Mínimo J_8	$-1,53 \times 10^{-3}$	3,9772	-3,9671	0,60022
Mínimo J_9	$4,77 \times 10^{-4}$	3,9720	-3,9668	0,60466
Mínimo J_{10}	$-1,93 \times 10^{-3}$	3,9803	-3,9739	0,61156

sendo $f^{[n]}(\cdot) = f(f^{[n-1]}(\cdot))$, com $f^{[1]}(\cdot) = f(\cdot)$, a n -ésima iteração de $f(\cdot)$. A função de custo adotada é o somatório dos erros até o passo n [12]

$$J_n = \sum_{k=1}^N \sum_{i=1}^n \|\xi_i(k)\|_2^2 . \quad (9)$$

Se $n = N$, ou seja, o erro envolver todos os dados coletados, a função de custo (9) está associada ao PIC. Na prática, $n \ll N$, mas qualquer $n > 1$ deveria ser melhor do que o erro de predição um passo à frente.

Como estratégias alternativas à identificação tradicional, foram estimados modelos minimizando J_i para $i = 2, \dots, 10$. Os modelos obtidos foram então simulados. Os parâmetros estimados usando a estratégia convencional (erro um passo à frente) e as funções de custo (9) são mostrados na Tabela 2, juntamente com os parâmetros do sistema original.

Pode-se notar que todas as estratégias alternativas apresentam coeficientes mais próximos do sistema original que os do método convencional, sendo o modelo de mínimo J_3 o que mais se aproxima do sistema. No entanto, mais importante que os valores dos parâmetros

em si é o comportamento temporal do modelo identificado. Para mensurar esse comportamento temporal, foram computados os expoentes de Lyapunov do sistema original e dos modelos identificados. Os resultados também se encontram na Tabela 2.

O melhor modelo identificado, segundo o critério do expoente de Lyapunov, é aquele que minimiza J_6 , enquanto o modelo de “melhores” coeficientes é o de mínimo J_3 . Isso é uma indicação de que, sob o ponto de vista dinâmico, os coeficientes em si talvez não sejam tão importantes. Outro ponto interessante é que todos os modelos que minimizam de J_2 a J_{10} possuem os expoentes de Lyapunov mais próximos dos do sistema original.

AGRADECIMENTOS

Os autores agradecem à FAPESP e ao CNPq pelo apoio.

5. CONCLUSÃO

Neste artigo, a estrutura básica dos métodos de identificação de sistemas tradicionais é apresentada. O problema da identificação de sistemas é formulado explicitamente como um problema inverso. Tal problema inverso consiste na determinação da estrutura e dos parâmetros de um modelo a partir de dados da evolução temporal de um sistema. Observa-se que as técnicas tradicionais de identificação não atacam esse problema, chamado aqui de Problema Inverso Completo, ou PIC, mas sim resolvem um problema diferente, o Problema Inverso Reduzido, ou PIR. Este equívoco causa uma piora no desempenho dos modelos identificados, que pode ser revertida com uma alteração da estratégia, levando em conta o PIC.

Referências

- [1] Aguirre, L. A. (2000a). *Introdução à Identificação de Sistemas: Técnicas Lineares e Não Lineares Aplicadas a Sistemas Reais*. Editora UFMG.
- [2] Aguirre, L. A. (2000b). A nonlinear dynamical approach to system identification. *IEEE Circuits Syst. Newsletter*, 11(2):10–23.
- [3] Aguirre, L. A., Freitas, U. S., Letellier, C., and Maquet, J. (2001). Structure selection techniques applied to continuous-time nonlinear models. *Physica D*, 158(2):1–18.
- [4] Alifanov, O. M. (1979). *Identification of Heat Transfer Process of Flying Vehicles (An Introduction to the Theory of Inverse Heat Transfer Problems)*. Mashinostroenie Publishing Academy, Moscou.
- [5] Billings, S. A. and Chen, S. (1989). Identification of nonlinear rational systems using a predictor-error estimation algorithm. *Int. J. Systems Sci.*, 20(3):467–494.
- [6] Billings, S. A., Chen, S., and Korenberg, M. J. (1989). Identification of MIMO nonlinear systems using a forward-regression orthogonal estimator. *Int. J. Control*, 49(6):2157–2189.
- [7] Broomhead, D. S. and Lowe, D. (1988). Multivariable functional interpolation and adaptive networks. *Complex Systems*, 2:321–355.
- [8] Corrêa, M. (1997). Identificação de sistemas dinâmicos não lineares utilizando modelos NARMAX racionais — aplicação a sistemas reais. Master’s thesis, PPGEE, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- [9] Elsner, J. B. (1992). Predicting time series using a neural network as a method of distinguishing chaos from noise. *J. Phys. A: Math. Gen.*, 25:843–850.
- [10] Freitas, U. S. (2001). Uso de técnicas de detecção de estrutura na identificação de modelos dinâmicos não lineares contínuos polinomiais. Master’s thesis, PPGEE, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- [11] Hsia, T. C. (1977). *System Identification*. Lexington Books, Lexington, Massachusetts.
- [12] Jaeger, L. and Kantz, H. (1996). Unbiased reconstruction of the dynamics underlying a noisy chaotic time series. *Chaos*, 6(3):440–450.
- [13] Johansson, R. (1993). *System Modeling and Identification*. Prentice-Hall, Englewood Cliffs.
- [14] Leontaritis, I. J. and Billings, S. A. (1985a). Input-output parametric models for nonlinear systems part I: deterministic nonlinear systems. *Int. J. Control*, 41(2):303–328.
- [15] Leontaritis, I. J. and Billings, S. A. (1985b). Input-output parametric models for nonlinear systems part II: stochastic nonlinear systems. *Int. J. Control*, 41(2):329–344.
- [16] Monteiro, L. H. A. (2002). *Sistemas Dinâmicos*. Editora Livraria da Física, São Paulo.
- [17] Reed, R. (1993). Pruning algorithms — a survey. *IEEE Trans. Neural Networks*, 4(5):740–746.
- [18] Söderström, T. and Stoica, P. (1989). *System Identification*. Prentice Hall International, London.
- [19] Takagi, T. and Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Syst., Man, Cybern.*, 15(1):116–132.