
MÉTODOS DE VARIÁVEL INSTRUMENTAL *FUZZY* PARA IDENTIFICAÇÃO DE SISTEMAS

Ginalber L.O. Serra*

ginalber@dmcsi.fee.unicamp.br

Celso P. Bottura*

bottura@dmcsi.fee.unicamp.br

*Laboratório de Controle e Sistemas Inteligentes
UNICAMP - FEEC - DMCSI
Campinas, SP, Brasil

ABSTRACT

This paper presents a theoretical proposal of fuzzy instrumental variable (FIV) methods for nonlinear discrete time systems identification and parameter estimation in noisy environment based on the fuzzy instrumental variable concept. The theoretical analysis is presented using a suitable formulation associated to Takagi-Sugeno (TS) fuzzy model. The complexity of the algorithm is quite low and the statistical properties show that the asymptotic error of the parameter estimates go to zero as the data series length increases.

KEYWORDS: System identification, Fuzzy systems, Fuzzy instrumental variable.

RESUMO

Este artigo apresenta uma proposta teórica de métodos de variável instrumental *fuzzy* para identificação de sistemas não-lineares no tempo discreto e estimação paramétrica em ambientes ruidosos a partir do conceito de variável instrumental *fuzzy*. A análise teórica é apresentada a partir de uma formulação adequada associada ao modelo *fuzzy*

Takagi-Sugeno. A complexidade do algoritmo é baixa e as propriedades estatísticas mostram que o erro assintótico das estimações paramétricas tende a zero quando o conjunto de dados aumenta.

PALAVRAS-CHAVE: Identificação de sistemas, Sistemas *Fuzzy*, Variável instrumental *fuzzy*.

1 INTRODUÇÃO

A área de identificação de sistemas, em particular a estimação paramétrica em modelos dinâmicos, é de grande importância em muitos campos tais como engenharia de controle, análise de séries temporais, economia e processamento de sinais. Vários métodos de identificação têm sido propostos, cada um com suas vantagens e restrições, e estão em uso na comunidade científica ou até mesmo na indústria (Ljung, 1999). Atualmente existe um grande interesse em métodos paramétricos, isto é, métodos de identificação cuja parte crucial inclui a estimação paramétrica a partir de dados experimentais. As metodologias de identificação inteligente empregam técnicas motivadas por sistemas biológicos e inteligência humana e tem sido introduzidas explorando esquemas de representações alternativas usando linguagem natural, regras, redes semânticas ou modelos qualitativos (Bottura and Serra, 2004; Kasabov and Song, 2002; Serra, 2005; Serra and Bottura, 2006). Identificação *fuzzy* é um exemplo típico de técnicas que fazem uso de conhecimento humano e processo dedutivo. Para identificação de sistemas não-lineares, modelos *fuzzy* Takagi-Sugeno (TS) são

ARTIGO CONVIDADO:

Versão completa e revisada de artigo apresentado no SBAI-2005

Artigo submetido em 23/05/2006

1a. Revisão em 11/09/2006

2a. Revisão em 17/12/2006

3a. Revisão em 12/04/2007

Aceito sob recomendação do Editor Convidado

Prof. Osvaldo Ronald Saavedra Mendez

grandemente investigados, visto que eles são aproximadores universais e fornecem boas características de interpolação e generalização (Papadakis and Theocaris, 2002; Takagi and Sugeno, 1985). Este artigo apresenta, como contribuição original, uma proposta teórica de métodos de variável instrumental *fuzzy* para identificação de sistemas não-lineares no tempo discreto e estimação paramétrica em ambientes ruidosos a partir do conceito de variável instrumental *fuzzy* proposto em (a) Bottura and Serra, 2005; (b) Bottura and Serra, 2005). A análise teórica é apresentada a partir de uma formulação adequada associada ao modelo *fuzzy* Takagi-Sugeno. A complexidade do algoritmo é baixa e as propriedades estatísticas mostram que o erro assintótico das estimações paramétricas tende a zero quando o conjunto de dados aumenta.

2 MODELO FUZZY TAKAGI-SUGENO

O sistema de inferência *fuzzy* Mandani, representa uma planta ou controlador por meio de regras linguísticas SE-ENTÃO com proposições *fuzzy* no antecedente e no conseqüente. O sistema de inferência *fuzzy* Takagi-Sugeno (Takagi and Sugeno, 1985) usa no conseqüente uma expressão funcional das variáveis linguísticas do antecedente. A i -ésima regra TS é da seguinte forma:

$$R^i : \text{SE } \tilde{x}_1 \text{ é } F_1^i \text{ E } \dots \text{ E } \tilde{x}_n \text{ é } F_n^i \\ \text{ENTÃO } \tilde{y}_i = f_i(\tilde{\mathbf{x}}) \quad |^{i=1,2,\dots,l} \quad (1)$$

onde l é o número de regras. O vetor $\tilde{\mathbf{x}} \in \mathbb{R}^n$ contém as variáveis linguísticas do antecedente. Cada variável linguística tem seu próprio universo de discurso $\mathcal{U}_{\tilde{x}_1}, \dots, \mathcal{U}_{\tilde{x}_n}$ particionados pelos conjuntos *fuzzy* representando os termos linguísticos correspondentes. A variável \tilde{x}_j pertence a um conjunto *fuzzy* F_j^i com um valor de verdade dado por uma função de pertinência $\mu_{F_j^i} : \mathbb{R} \rightarrow [0, 1]$. Esta notação refere-se ao fato de que $F_j^i \in \{F_{j1}^i, F_{j2}^i, \dots, F_{jp_j}^i\}$ e $\mu_{F_j^i} \in \{\mu_{F_{j1}^i}, \mu_{F_{j2}^i}, \dots, \mu_{F_{jp_j}^i}\}$, onde $p_{\tilde{x}_j}$ é o número de partições do universo de discurso associado à variável linguística \tilde{x}_j . O grau de ativação h_i para a regra i é dado por:

$$h_i(\tilde{\mathbf{x}}) = \mu_{F_1^i}(\tilde{x}_1^*) \star \mu_{F_2^i}(\tilde{x}_2^*) \star \dots \star \mu_{F_n^i}(\tilde{x}_n^*) \quad (2)$$

onde \tilde{x}_i^* é algum ponto em $\mathcal{U}_{\tilde{x}_i}$ e \star é uma norma-t (em nosso caso, o produto algébrico). O grau de ativação normalizado para a regra i é definido como:

$$\gamma_i(\tilde{\mathbf{x}}) = \frac{h_i(\tilde{\mathbf{x}})}{\sum_{r=1}^l h_r(\tilde{\mathbf{x}})} \quad (3)$$

Esta normalização implica em

$$\sum_{i=1}^l \gamma_i(\tilde{\mathbf{x}}) = 1 \quad (4)$$

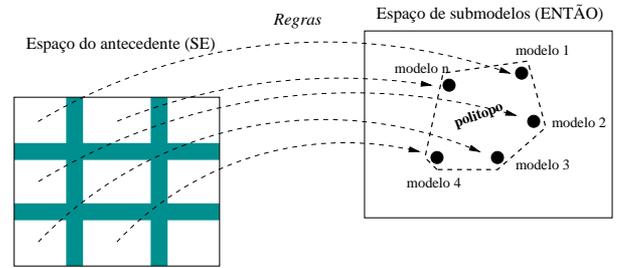


Figura 1: Mapeamento *fuzzy* do espaço de entrada no espaço de submodelos locais.

A resposta do modelo TS é uma soma ponderada dos conseqüentes, isto é, uma combinação convexa das funções locais (modelos) f_i :

$$y = \sum_{i=1}^l \gamma_i(\tilde{\mathbf{x}}) f_i(\tilde{\mathbf{x}}) \quad (5)$$

Tal modelo pode ser visto como um sistema linear variante nos parâmetros LPV (*Linear Parameter Varying*). Neste sentido, um modelo TS pode ser considerado como um mapeamento do espaço do antecedente (entrada) à uma região convexa (polígono) no espaço dos submodelos locais definidos pelos parâmetros do conseqüente (Bottura and Serra, 2004; Kuma, Stoll and Stoll, 2006; Skrjanc, Blazic and Agamennoni, 2005; Kukolj and Levi, 2004), conforme mostrado na Fig. 1. Esta propriedade simplifica a análise de modelos TS num contexto de sistema linear robusto tanto para identificação como para projeto de controladores com características de malha-fechada desejadas ou análise de estabilidade.

3 FORMULAÇÃO DO PROBLEMA

Um modelo não-linear estabelece uma relação não-linear entre a coleção de dados de entrada u e saída y passados e a saída predita

$$\hat{y}(k+1) = \tilde{F}(y(k), \dots, y(k-n_y+1), u(k), \dots, u(k-n_u+1)) \quad (6)$$

onde k é o instante de amostragem, n_y e n_u são inteiros relacionados à ordem do sistema. Neste trabalho, será estabelecida a estrutura de modelo autorregressivo com entrada exógena não-linear NARX (*Nonlinear AutoRegressive with eXogenous input*), usada na maioria dos métodos de identificação não-linear tais como redes neurais, funções de base radial, CMAC (*Cerebellar Model Articulation Controller*), e também modelos *fuzzy* (Brown and Harris, 1994; Tsoukalas and Uhrig, 1997). Com uma

formulação baseada no modelo *fuzzy* Takagi-Sugeno, tem-se

$$R^i : \text{SE } y(k) \text{ é } F_1^i \text{ E } \dots \text{ E } y(k - n_y + 1) \text{ é}$$

$$F_{n_y}^i \text{ E } u(k) \text{ é } G_1^i \text{ E } \dots \text{ E } u(k - n_u + 1)$$

$$\text{é } G_{n_u}^i \text{ ENTÃO } \hat{y}_i(k+1) = \sum_{j=1}^{n_y} a_{i,j} y(k-j+1) + \sum_{j=1}^{n_u} b_{i,j} u(k-j+1) + c_i \quad (7)$$

onde $a_{i,j}$, $b_{i,j}$ e c_i são os parâmetros do conseqüente a serem determinados. A fórmula de inferência é dada por

$$\hat{y}(k+1) = \frac{\sum_{i=1}^l h_i(\mathbf{x}_k) \hat{y}_i(k+1)}{\sum_{i=1}^l h_i(\mathbf{x}_k)} \quad (8)$$

ou

$$\hat{y}(k+1) = \sum_{i=1}^l \gamma_i(\mathbf{x}_k) \hat{y}_i(k+1) \quad (9)$$

com

$$\mathbf{x}_k = (y(k), \dots, y(k - n_y + 1), u(k), \dots, \dots, u(k - n_u + 1)) \quad (10)$$

Este modelo NARX representa sistemas de múltiplas entradas e única saída MISO (*Multiple Input Single Output*) diretamente e sistemas de múltiplas entradas e múltiplas saídas MIMO (*Multiple Input Multiple Output*) em uma forma decomposta como um conjunto de modelos MISO acoplados.

3.1 Problema de estimação dos parâmetros do conseqüente

A fórmula de inferência do modelo *fuzzy* em (9) pode ser expressa como

$$\begin{aligned} \hat{y}(k+1) = & \gamma_1(\mathbf{x}_k)[a_{1,1}y(k) + \dots + a_{1,n_y}y(k - n_y \\ & + 1) + b_{1,1}u(k) + \dots + b_{1,n_u}u(k - n_u + 1) + c_1] + \\ & + \gamma_2(\mathbf{x}_k)[a_{2,1}y(k) + \dots + a_{2,n_y}y(k - n_y + 1) + \\ & + b_{2,1}u(k) + \dots + b_{2,n_u}u(k - n_u + 1) + c_2] + \dots \end{aligned}$$

$$\begin{aligned} & + \gamma_l(\mathbf{x}_k)[a_{l,1}y(k) + \dots + a_{l,n_y}y(k - n_y + 1) + \\ & + b_{l,1}u(k) + \dots + b_{l,n_u}u(k - n_u + 1) + c_l] \quad (11) \end{aligned}$$

que é linear nos parâmetros do conseqüente, \mathbf{a} , \mathbf{b} e \mathbf{c} . Para um conjunto de N pares de dados de entrada-saída $\{(\mathbf{u}_k, y_k) | i = 1, 2, \dots, N\}$ disponível, a seguinte forma vetorial é obtida

$$\mathbf{Y} = [\psi_1 \mathbf{X}, \psi_2 \mathbf{X}, \dots, \psi_l \mathbf{X}] \Theta + \Xi \quad (12)$$

onde $\psi_i = \text{diag}(\gamma_i(\mathbf{x}_k)) \in \mathfrak{R}^{N \times N}$, $\mathbf{X} = [y_k, \dots, y_{k-n_y+1}, \mathbf{u}_k, \dots, \mathbf{u}_{k-n_u+1}, \mathbf{1}] \in \mathfrak{R}^{N \times (n_y+n_u+1)}$, $\mathbf{Y} \in \mathfrak{R}^{N \times 1}$, $\Xi \in \mathfrak{R}^{N \times 1}$ e $\Theta \in \mathfrak{R}^{l(n_y+n_u+1) \times 1}$ são a matriz de grau de pertinência normalizada, a matriz de dados, o vetor de saída, o vetor de ruído desconhecido e o vetor de parâmetros estimados, respectivamente.

Diferentes técnicas podem ser usadas quando as variáveis associadas com os parâmetros desconhecidos são *exatamente conhecidas*. Na prática, e no presente contexto, as relações básicas entre os parâmetros ainda estão na forma apresentada em (12), mas os elementos de \mathbf{X} não são exatamente conhecidos e podem ser observados com erro de maneira que o valor observado é expresso por

$$y_k = \chi_k^T \theta + \eta_k \quad (13)$$

onde, no k -ésimo instante de amostragem, $\chi_k^T = [\gamma_k^1(\mathbf{x}_k + \xi_k), \dots, \gamma_k^l(\mathbf{x}_k + \xi_k)]$ é o vetor de dados com erro nas variáveis, $\mathbf{x}_k = [y_{k-1}, \dots, y_{k-n_y}, u_{k-1}, \dots, u_{k-n_u}, 1]^T$ é o vetor de dados com quantidades exatamente conhecidas, isto é, dados de entrada-saída livres de ruído, ξ_k é um vetor de ruído associado com a observação de \mathbf{x}_k , e η_k é um ruído de saída.

As equações normais são formuladas por

$$\left[\sum_{j=1}^k \chi_j \chi_j^T \right] \hat{\theta}_k = \sum_{j=1}^k \chi_j y_j \quad (14)$$

e multiplicando por $\frac{1}{k}$ resulta

$$\begin{aligned} & \left\{ \frac{1}{k} \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)] [\gamma_j^1(\mathbf{x}_j + \xi_j), \right. \\ & \left. \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \right\} \hat{\theta}_k = \frac{1}{k} \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \\ & \gamma_j^l(\mathbf{x}_j + \xi_j)] y_j \quad (15) \end{aligned}$$

Visto que $y_j = \chi_j^T \theta + \eta_j$,

$$\left\{ \frac{1}{k} \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)] [\gamma_j^1(\mathbf{x}_j + \xi_j), \right.$$

$$\begin{aligned} \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \hat{\theta}_k &= \frac{1}{k} \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \\ \gamma_j^l(\mathbf{x}_j + \xi_j)] [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \theta &+ \\ + \frac{1}{k} \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)] \eta_j &\quad (16) \end{aligned}$$

e

$$\begin{aligned} \tilde{\theta}_k &= \left[\frac{1}{k} \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)] [\gamma_j^1(\mathbf{x}_j + \right. \\ &+ \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \right]^{-1} \frac{1}{k} \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \\ &\dots, \gamma_j^l(\mathbf{x}_j + \xi_j)] \eta_j \quad (17) \end{aligned}$$

onde $\tilde{\theta}_k = \hat{\theta}_k - \theta$ é o erro paramétrico. Tomando-se a probabilidade no limite, quando $k \rightarrow \infty$,

$$p. \lim_{k \rightarrow \infty} \tilde{\theta}_k = p. \lim_{k \rightarrow \infty} \left\{ \frac{1}{k} \mathbf{C}_k^{-1} \frac{1}{k} \mathbf{b}_k \right\} \quad (18)$$

com

$$\begin{aligned} \mathbf{C}_k &= \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)] [\gamma_j^1(\mathbf{x}_j + \\ &+ \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \\ \mathbf{b}_k &= \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)] \eta_j \end{aligned}$$

Aplicando-se o teorema de Slutsky (Goodwin and Payne, 1977) e supondo que os elementos de $\frac{1}{k} \mathbf{C}_k$ e $\frac{1}{k} \mathbf{b}_k$ convergem em probabilidade, temos

$$p. \lim_{k \rightarrow \infty} \tilde{\theta}_k = p. \lim_{k \rightarrow \infty} \frac{1}{k} \mathbf{C}_k^{-1} p. \lim_{k \rightarrow \infty} \frac{1}{k} \mathbf{b}_k \quad (19)$$

Assim,

$$\begin{aligned} p. \lim_{k \rightarrow \infty} \frac{1}{k} \mathbf{C}_k &= p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \\ &+ \xi_j)] [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \\ p. \lim_{k \rightarrow \infty} \frac{1}{k} \mathbf{C}_k &= p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k (\gamma_j^1)^2 (\mathbf{x}_j + \xi_j) (\mathbf{x}_j + \xi_j)^T \\ &+ \dots + p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k (\gamma_j^l)^2 (\mathbf{x}_j + \xi_j) (\mathbf{x}_j + \xi_j)^T \end{aligned}$$

Supondo \mathbf{x}_j e ξ_j estatisticamente independentes,

$$\begin{aligned} p. \lim_{k \rightarrow \infty} \frac{1}{k} \mathbf{C}_k &= p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k (\gamma_j^1)^2 [\mathbf{x}_j \mathbf{x}_j^T + \xi_j \xi_j^T] + \dots + \\ &+ p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k (\gamma_j^l)^2 [\mathbf{x}_j \mathbf{x}_j^T + \xi_j \xi_j^T] \\ p. \lim_{k \rightarrow \infty} \frac{1}{k} \mathbf{C}_k &= p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \mathbf{x}_j \mathbf{x}_j^T [(\gamma_j^1)^2 + \dots + (\gamma_j^l)^2] \\ &+ p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \xi_j \xi_j^T [(\gamma_j^1)^2 + \dots + (\gamma_j^l)^2] \quad (20) \end{aligned}$$

com $\sum_{i=1}^l \gamma_j^i = 1$. Logo, a análise assintótica da estimação dos parâmetros do consequente associada a modelos fuzzy TS é baseada na soma ponderada das matrizes de covariância de \mathbf{x} e ξ . Similarmente,

$$\begin{aligned} p. \lim_{k \rightarrow \infty} \frac{1}{k} \mathbf{b}_k &= p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \\ &+ \xi_j)] \eta_j \\ p. \lim_{k \rightarrow \infty} \frac{1}{k} \mathbf{b}_k &= p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\gamma_j^1 \xi_j \eta_j, \dots, \gamma_j^l \xi_j \eta_j] \quad (21) \end{aligned}$$

Substituindo-se (20) e (21) em (19), resulta

$$\begin{aligned} p. \lim_{k \rightarrow \infty} \tilde{\theta}_k &= \left\{ p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \mathbf{x}_j \mathbf{x}_j^T [(\gamma_j^1)^2 + \dots + (\gamma_j^l)^2] + \right. \\ &+ p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \xi_j \xi_j^T [(\gamma_j^1)^2 + \dots + (\gamma_j^l)^2] \left. \right\}^{-1} p. \lim_{k \rightarrow \infty} \frac{1}{k} \\ &\sum_{j=1}^k [\gamma_j^1 \xi_j \eta_j, \dots, \gamma_j^l \xi_j \eta_j] \quad (22) \end{aligned}$$

com $\sum_{i=1}^l \gamma_j^i = 1$. Para o caso de apenas uma regra ($l = 1$), a análise é simplificada ao modelo linear, com $\gamma_j^i \Big|_{i=1, \dots, k} = 1$. Visto que a entrada u_k continua a excitar o processo e, ao mesmo tempo, os coeficientes nos submodelos do consequente não são todos nulos, então a saída y_k existirá para todo intervalo de observação k . Conseqüentemente, a matriz de covariância $\mathbf{x}_j \mathbf{x}_j^T$ também será não-singular e a sua inversa existirá. Assim, o único modo no qual o erro paramétrico será nulo é para $\xi_j \eta_j$ zero. Mas, em geral, ξ_j e η_j estão correlacionados e o erro assintótico não será zero e a estimação de mínimos quadrados será polarizada definida pela taxa relativa da variância sinal/ruído. Em outras palavras, a estimação de mínimos quadrados não é apropriada para estimar os parâmetros do consequente pois a estimação será inconsistente e a polarização permanecerá independentemente do número de dados usados na estimação.

4 VARIÁVEL INSTRUMENTAL FUZZY

Em identificação de sistemas usando modelos lineares, para baixos níveis de ruído, o método de mínimos quadrados apresenta desempenho satisfatório na estimação dos parâmetros do modelo. Na presença de ruído colorido, o método dos mínimos quadrados fornece uma estimativa polarizada, ou seja, ao se fazer muitas observações a estimativa não se aproxima do valor verdadeiro dos parâmetros. Para níveis maiores de ruído, ou na presença de ruído colorido, métodos tais como GLS (*Generalized Least Squares*) (^aSöderström and Stoica, 1989), ELS (*Extended Least Squares*) (Ljung and Söderström, 1987) e PE (*Prediction Error*) (Ljung, 1999; Young, 1984) podem ser usados, embora sejam inevitavelmente dependentes da precisão do modelo do ruído. Para superar o problema de polarização e inconsistência da estimação paramétrica, sem a necessidade de modelar o ruído, o método IV (*Instrumental Variable*) pode ser usado (^bSöderström and Stoica, 1983). Neste método, um vetor estatisticamente independente dos ruídos e correlacionado com o vetor de dados do sistema \mathbf{x}_j se faz necessário. Se isto é possível, então será fácil ver que a escolha deste vetor é eficiente para remover a polarização assintótica da estimação dos parâmetros do consequente. A estimação fuzzy no sentido dos mínimos quadrados é dada por:

$$\left\{ \frac{1}{k} \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)] [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \right\} \hat{\theta}_k = \frac{1}{k} \sum_{j=1}^k [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)] \{ [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \theta + \eta_j \}$$

Usando-se um novo vetor de variáveis linguísticas $[\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j]$, a última equação fica da forma

$$\left\{ \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \right\} \hat{\theta}_k = \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] \{ [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \theta + \eta_j \} \quad (23)$$

onde \mathbf{z}_j é um vetor com a ordem de \mathbf{x}_j , associado ao comportamento dinâmico do sistema, e $\beta_j^i \mid^{i=1, \dots, l}$ é o grau de ativação normalizado, como em (3), associado a \mathbf{z}_j . Para análise de convergência, são propostos os seguintes lemas e teorema:

Lema 1 *Seja \mathbf{z}_j um vetor com a ordem de \mathbf{x}_j , associado ao comportamento dinâmico do sistema e independente da*

entrada de ruído ξ_j , e $\beta_j^i \mid^{i=1, \dots, l}$ é o grau de ativação normalizado, como em (3), associado a \mathbf{z}_j . Então, no limite

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] \xi_j^T = \mathbf{0} \quad (24)$$

Prova: Desenvolvendo-se o lado esquerdo de (24), resulta

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] \xi_j^T = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j \xi_j^T, \dots, \beta_j^l \mathbf{z}_j \xi_j^T]$$

Como $\beta_j^i \mid^{i=1, \dots, l}$ é um escalar, e, por definição, o vetor de variáveis escolhido é estatisticamente independente das entradas ruidosas, o produto interno entre \mathbf{z}_j e ξ_j será zero. Assim, tomando-se o limite, resulta

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j \xi_j^T, \dots, \beta_j^l \mathbf{z}_j \xi_j^T] = \mathbf{0}$$

□

Lema 2 *Seja \mathbf{z}_j um vetor com a ordem de \mathbf{x}_j , associado ao comportamento dinâmico do sistema e independente do ruído η_j , e $\beta_j^i \mid^{i=1, \dots, l}$ o grau de ativação normalizado, como em (3), associado a \mathbf{z}_j . Então, no limite*

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] \eta_j = \mathbf{0} \quad (25)$$

Prova: Desenvolvendo o lado esquerdo de (25), resulta

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] \eta_j = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j \eta_j, \dots, \beta_j^l \mathbf{z}_j \eta_j]$$

Visto que o vetor de variáveis escolhido é estatisticamente independente do ruído de observação, o produto entre \mathbf{z}_j e η_j será zero no limite. Logo,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j \eta_j, \dots, \beta_j^l \mathbf{z}_j \eta_j] = \mathbf{0}$$

□

Lema 3 *Seja \mathbf{z}_j um vetor com a ordem de \mathbf{x}_j , associado ao comportamento dinâmico do sistema, e $\beta_j^i \mid^{i=1, \dots, l}$ é o grau de ativação normalizado, como em (3), associado a \mathbf{z}_j . Então, de acordo com (23), no limite*

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T = \theta$$

$$+\xi_j)]^T = \mathbf{C}_{\mathbf{z}\mathbf{x}} \neq 0 \quad (26)$$

Prova: Desenvolvendo o lado esquerdo de (26), resulta

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T &= \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \gamma_j^1(\mathbf{z}_j \mathbf{x}_j^T + \mathbf{z}_j \xi_j^T) + \dots + \\ &+ \beta_j^l \gamma_j^l(\mathbf{z}_j \mathbf{x}_j^T + \mathbf{z}_j \xi_j^T)] \end{aligned}$$

A partir do **Lema 1**, esta expressão é simplificada para

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T &= \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \gamma_j^1 \mathbf{z}_j \mathbf{x}_j^T + \dots + \beta_j^l \gamma_j^l \mathbf{z}_j \mathbf{x}_j^T] \end{aligned}$$

e esta matriz de covariância tem a seguinte propriedade:

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \gamma_j^1 \mathbf{z}_j \mathbf{x}_j^T + \dots + \beta_j^l \gamma_j^l \mathbf{z}_j \mathbf{x}_j^T] \neq 0 \quad (27)$$

e

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \gamma_j^1 \mathbf{z}_j \mathbf{x}_j^T + \dots + \beta_j^l \gamma_j^l \mathbf{z}_j \mathbf{x}_j^T] = \mathbf{C}_{\mathbf{z}\mathbf{x}} \neq 0$$

Teorema 1 *Seja \mathbf{z}_j um vetor com a ordem de \mathbf{x}_j , associado ao comportamento dinâmico do sistema, e $\beta_j^i \mid i=1, \dots, l$ é o grau de ativação normalizado, como em (3), associado a \mathbf{z}_j . Então, no limite*

$$p. \lim_{k \rightarrow \infty} \tilde{\theta} = 0 \quad (28)$$

Prova: A partir do novo vetor de variáveis linguísticas $[\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j]$, a estimação dos parâmetros do conseqüente, no sentido dos mínimos quadrados, pode ser modificada como segue:

$$\begin{aligned} \left\{ \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \right\} \hat{\theta}_k &= \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] \{ [\gamma_j^1(\mathbf{x}_j + \xi_j), \\ &\dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \theta + \eta_j \} \end{aligned}$$

a qual pode ser expressa como

$$\left\{ \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \right\} (\hat{\theta}_k - \theta) = \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] \eta_j$$

e

$$\left\{ \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \right\} \tilde{\theta} = \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] \eta_j$$

Tomando-se a probabilidade no limite com $k \rightarrow \infty$, e aplicando-se o teorema de Slutsky, temos

$$p. \lim_{k \rightarrow \infty} \tilde{\theta}_k = \{ p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] [\gamma_j^1(\mathbf{x}_j + \xi_j), \dots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T \}^{-1} \{ p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] \eta_j \}$$

De acordo com o **Lema 1** e o **Lema 3**, resulta

$$p. \lim_{k \rightarrow \infty} \tilde{\theta}_k = \{ p. \lim_{k \rightarrow \infty} \mathbf{C}_{\mathbf{z}\mathbf{x}} \}^{-1} \{ p. \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] \eta_j \}$$

onde $\mathbf{C}_{\mathbf{z}\mathbf{x}}$ é não-singular e, conseqüentemente, a inversa existe. A partir do **Lema 2**, temos

$$p. \lim_{k \rightarrow \infty} \tilde{\theta}_k = \{ p. \lim_{k \rightarrow \infty} \mathbf{C}_{\mathbf{z}\mathbf{x}} \}^{-1} \mathbf{0}$$

Assim, o valor do erro paramétrico, no limite em probabilidade, é

$$p. \lim_{k \rightarrow \infty} \tilde{\theta} = 0 \quad (29)$$

e a estimação é assintoticamente não-polarizada, como desejado. \square

Em conseqüência desta análise, a definição do vetor $[\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j]$ como *vetor de variáveis instrumentais fuzzy* ou simplesmente *variáveis instrumentais fuzzy* (VIF) é proposta. Claramente, com o uso do vetor VIF na forma sugerida, é possível eliminar a polarização assintótica e preservar a existência de uma solução. Porém, a eficiência estatística da solução é dependente do grau de correlação entre $[\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j]$ e $[\gamma_j^1 \mathbf{x}_j, \dots, \gamma_j^l \mathbf{x}_j]$. Em particular, a menor variância da estimação ocorrerá quando $\mathbf{z}_j = \mathbf{x}_j$, isto é, quando \mathbf{z}_j são iguais às variáveis do sistema livres de ruído, os quais estão indisponíveis na prática. Uma escolha eficiente de VIF seria baseada na seqüência de entrada atrasada

$$\mathbf{z}_j = [u_{k-\tau}, \dots, u_{k-\tau-n}, u_k, \dots, u_{k-n}]^T$$

onde τ é escolhido de modo que os elementos da matriz de covariância $\mathbf{C}_{\mathbf{z}\mathbf{x}}$ são maximizados. Neste caso, o sinal de entrada é considerado excitante persistentemente, isto é, ele deve perturbar continuamente o sistema. Uma outra escolha de VIF seria baseada na seqüência de entrada-saída atrasada

$$\mathbf{z}_j = [y_{k-1-dl}, \dots, y_{k-n_y-dl}, u_{k-1-dl}, \dots, u_{k-n_u-dl}]^T$$

onde dl é o atraso aplicado. Outra escolha de VIF seria baseada na seqüência de entrada-saída de um “modelo *fuzzy* auxiliar” com a mesma estrutura do modelo utilizado para identificar o sistema dinâmico original. Assim,

$$\mathbf{z}_j = [\hat{y}_{k-1}, \dots, \hat{y}_{k-n_y}, u_{k-1}, \dots, u_{k-n_u}]^T$$

onde \hat{y} é a saída do modelo *fuzzy* auxiliar, e u é a entrada do sistema dinâmico. A fórmula de inferência deste modelo auxiliar é dada por

$$\begin{aligned} \hat{y}(k+1) = & \beta_1(\mathbf{z}_k)[\alpha_{1,1}y(k) + \dots + \alpha_{1,n_y}y(k-n_y \\ & +1) + \rho_{1,1}u(k) + \dots + \rho_{1,n_u}u(k-n_u+1) + \delta_1] + \\ & + \beta_2(\mathbf{z}_k)[\alpha_{2,1}\hat{y}(k) + \dots + \alpha_{2,n_y}y(k-n_y+1) + \\ & + \rho_{2,1}u(k) + \dots + \rho_{2,n_u}u(k-n_u+1) + \delta_2] + \dots \\ & + \beta_l(\mathbf{z}_k)[\alpha_{l,1}\hat{y}(k) + \dots + \alpha_{l,n_y}y(k-n_y+1) + \\ & + \rho_{l,1}u(k) + \dots + \rho_{l,n_u}u(k-n_u+1) + \delta_l] \end{aligned} \quad (30)$$

a qual é também linear nos parâmetros do conseqüente: α , ρ e δ . Quanto mais próximos dos parâmetros reais, mas desconhecidos, (a, b, c) como em (11), mais correlacionados \mathbf{z}_k e \mathbf{x}_k , e a estimação via VIF será próxima da ótima.

4.1 Algoritmo de variável instrumental *fuzzy*

Nesta seção é apresentada a formulação geral, *off-line* e *online*, do algoritmo de variável instrumental *fuzzy*, independente dos três métodos aqui apresentados em função de \mathbf{z}_j : seqüência de entrada atrasada, seqüência de entrada-saída atrasada, seqüência de entrada-saída de um modelo *fuzzy* auxiliar. Resultados computacionais bem como um estudo de caso à identificação de uma viga de alumínio (uma planta complexa, não-linear e variante com o tempo) cujo estudo serve como fundamento para aplicações em estruturas mecânicas de aeronaves e/ou veículos aeroespaciais podem ser vistos em (^bBottura and Serra, 2005).

4.2 Esquema *Off-line*

As equações normais VIF são formuladas como

$$\sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] [\gamma_j^1 (\mathbf{x}_j + \xi_j), \dots, \gamma_j^l (\mathbf{x}_j + \xi_j)]^T \hat{\theta}_k - \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] y_j = 0 \quad (31)$$

ou, com $\zeta_j = [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j]$,

$$\left[\sum_{j=1}^k \zeta_j \chi_j^T \right] \hat{\theta}_k - \sum_{j=1}^k \zeta_j y_j = 0 \quad (32)$$

de modo que a estimação VIF é obtida por

$$\hat{\theta}_k = \left\{ \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] [\gamma_j^1 (\mathbf{x}_j + \xi_j), \dots, \gamma_j^l (\mathbf{x}_j + \xi_j)]^T \right\}^{-1} \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] y_j \quad (33)$$

e, na forma vetorial, o problema de interesse pode ser definido por:

$$\hat{\theta} = (\mathbf{\Gamma}^T \mathbf{\Sigma})^{-1} \mathbf{\Gamma}^T \mathbf{Y} \quad (34)$$

onde $\mathbf{\Gamma}^T \in \mathfrak{R}^{l(n_y+n_u+1) \times N}$ é a matriz de variável instrumental *fuzzy* estendida com linhas dadas por ζ_j , $\mathbf{\Sigma} \in \mathfrak{R}^{N \times l(n_y+n_u+1)}$ é a matriz de dados estendida com linhas dadas por χ_j e $\mathbf{Y} \in \mathfrak{R}^{N \times 1}$ é o vetor de saída e $\hat{\theta} \in \mathfrak{R}^{l(n_y+n_u+1) \times 1}$ é o vetor de parâmetros.

4.3 Esquema *online*

A estimação VIF em (33) pode ser da forma

$$\hat{\theta}_k = \mathbf{P}_k \mathbf{b}_k \quad (35)$$

onde

$$\mathbf{P}_k = \left\{ \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] [\gamma_j^1 (\mathbf{x}_j + \xi_j), \dots, \gamma_j^l (\mathbf{x}_j + \xi_j)]^T \right\}^{-1}$$

e

$$\mathbf{b}_k = \sum_{j=1}^k [\beta_j^1 \mathbf{z}_j, \dots, \beta_j^l \mathbf{z}_j] y_j$$

a qual pode ser expressa como

$$\mathbf{P}_k^{-1} = \mathbf{P}_{k-1}^{-1} + [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k] [\gamma_k^1 (\mathbf{x}_k + \xi_k), \dots,$$

$$\dots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \quad (36)$$

e

$$\mathbf{b}_k = \mathbf{b}_{k-1} + [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k] y_k \quad (37)$$

respectivamente. Pré-multiplicando (36) por \mathbf{P}_k e pós-multiplicando por \mathbf{P}_{k-1} resulta

$$\begin{aligned} \mathbf{P}_{k-1} = \mathbf{P}_k + \mathbf{P}_k [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k] [\gamma_k^1(\mathbf{x}_k + \xi_k), \dots, \\ \dots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1} \end{aligned} \quad (38)$$

e pós-multiplicando pelo vetor VIF, resulta

$$\begin{aligned} \mathbf{P}_{k-1} [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k] = \mathbf{P}_k [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k] \{1 + \\ [\gamma_k^1(\mathbf{x}_k + \xi_k), \dots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1} [\beta_k^1 \mathbf{z}_k, \dots, \\ \dots, \beta_k^l \mathbf{z}_k]\} \end{aligned}$$

Então, pós-multiplicando por $\{1 + [\gamma_k^1(\mathbf{x}_k + \xi_k), \dots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1} [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k]\}^{-1} [\gamma_k^1(\mathbf{x}_k + \xi_k), \dots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1}$, tem-se

$$\begin{aligned} \mathbf{P}_{k-1} [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k] \{1 + [\gamma_k^1(\mathbf{x}_k + \xi_k), \dots, \gamma_k^l(\mathbf{x}_k + \\ \xi_k)]^T \mathbf{P}_{k-1} [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k]\}^{-1} [\gamma_k^1(\mathbf{x}_k + \xi_k), \dots, \\ \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1} = \mathbf{P}_k [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k] [\gamma_k^1(\mathbf{x}_k + \\ \xi_k), \dots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1} \end{aligned} \quad (39)$$

Substituindo (38) em (39), resulta

$$\begin{aligned} \mathbf{P}_k = \mathbf{P}_{k-1} - \mathbf{P}_{k-1} [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k] \{1 + [\gamma_k^1(\mathbf{x}_k + \\ + \xi_k), \dots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1} [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k]\}^{-1} \\ [\gamma_k^1(\mathbf{x}_k + \xi_k), \dots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1} \end{aligned} \quad (40)$$

Substituindo (40) e (37) em (35), a equação recursiva para os parâmetros do conseqüente será:

$$\begin{aligned} \hat{\theta}_k = \{\mathbf{P}_{k-1} - \mathbf{P}_{k-1} [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k] \{1 + [\gamma_k^1(\mathbf{x}_k + \\ + \xi_k), \dots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1} [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k]\}^{-1} \\ [\gamma_k^1(\mathbf{x}_k + \xi_k), \dots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1}\} \{\mathbf{b}_{k-1} + \\ [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k] y_k\} \end{aligned}$$

de maneira que, finalmente

$$\begin{aligned} \hat{\theta}_k = \hat{\theta}_{k-1} - \mathbf{K}_k \{[\gamma_j^1(\mathbf{x}_k + \xi_k), \dots, \gamma_k^l(\mathbf{x}_k + \\ + \xi_k)]^T \hat{\theta}_{k-1} - y_k\} \end{aligned} \quad (41)$$

onde

$$\begin{aligned} \mathbf{K}_k = \mathbf{P}_{k-1} [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k] \{1 + [\gamma_k^1(\mathbf{x}_k + \xi_k), \dots, \\ \dots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1} [\beta_k^1 \mathbf{z}_k, \dots, \beta_k^l \mathbf{z}_k]\}^{-1} \end{aligned} \quad (42)$$

5 RESULTADOS COMPUTACIONAIS

5.1 Exemplo: planta não-linear

A planta a ser identificada consiste em um sistema altamente não-linear de segunda ordem

$$y(k+1) = \frac{y(k)y(k-1)(y(k)+2.5)}{1+y(k)^2+y(k-1)^2} + u(k) + e(k) \quad (43)$$

que é uma referência muito utilizada em modelagem *fuzzy*, onde $y(k)$ é a saída e $u(k) = \sin(\frac{2\pi k}{25})$ é a entrada aplicada. Neste caso $e(k)$ é um ruído branco com média zero e variância σ^2 . O modelo TS tem duas entradas $y(k)$ e $y(k-1)$ e uma saída $y(k+1)$, e os parâmetros do antecedente na base de regras do modelo *fuzzy* são definidos pelo método de agrupamento ECM (*Evolving Clustering Method*) proposto em (Kasabov and Song, 2002). O modelo é composto de regras da forma:

$$R^i : \text{IF } y(k) \text{ is } F_1^i \text{ AND } y(k-1) \text{ is } F_2^i \text{ THEN}$$

$$\hat{y}_i(k+1) = a_{i,1}y(k) + a_{i,2}y(k-1) + b_{i,1}u(k) + c_i \quad (44)$$

onde $F_{1,2}^i$ são conjuntos *fuzzy* com funções gaussianas. Conjuntos de dados experimentais de N pontos cada são criados a partir de (43), com $\sigma^2 \in [0, 0.23]$. Isto quer dizer que o ruído aplicado tem valores entre 0 e $\pm 30\%$ do valor da saída nominal, que é uma percentagem de ruído aceitável na prática. Estes conjuntos de dados são apresentados ao algoritmo proposto, para a obtenção de um modelo *fuzzy* IV, e a um algoritmo baseado em LS, para a obtenção de um modelo *fuzzy* LS. Os modelos são obtidos pelas seguintes abordagens:

- *Abordagem global* : Nesta abordagem, todos os parâmetros do conseqüente são estimados simultaneamente, minimizando o critério:

$$\hat{\Theta} = \arg \min \| \mathbf{S}\Theta - \mathbf{b} \|^2 \quad (45)$$

com $\mathbf{S} = \mathbf{\Gamma}^T \mathbf{W} \mathbf{\Sigma}$ e $\mathbf{b} = \mathbf{\Gamma}^T \mathbf{W} \mathbf{Y}$ para o algoritmo proposto, onde $\mathbf{\Gamma}$ é a matriz de variável instrumental estendida e $\mathbf{W} = \text{diag}(\lambda^{n-1}, \lambda^{n-2}, \dots, 1) \in \mathfrak{R}^{N \times N}$, com $0 < \lambda < 1$; e $\mathbf{S} = \mathbf{\Sigma}^T \mathbf{\Sigma}$ e $\mathbf{b} = \mathbf{\Sigma}^T \mathbf{Y}$ para o algoritmo baseado no método de mínimos quadrados, onde $\mathbf{\Sigma}$ é a matriz de dados estendida. O escalar λ , *fator de esquecimento*, é usado para atribuir menos peso aos dados passados.

- *Abordagem local* : Nesta abordagem os parâmetros do conseqüente são estimados para cada regra i , independentemente uns dos outros, minimizando o conjunto de critérios locais ponderados ($i = 1, 2, \dots, l$):

$$\hat{\Theta}_i = \arg \min \| \mathbf{S}_i \Theta_i - \mathbf{b}_i \|^2 \quad (46)$$

com $\mathbf{S}_i = \mathbf{Z}^T \Psi_i \mathbf{X}$ e $\mathbf{b}_i = \mathbf{Z}^T \Psi_i \mathbf{Y}$ para o algoritmo proposto, onde \mathbf{Z} é a matriz de variável instrumental e Ψ_i é a matriz de graus de ativação normalizados de cada regra *fuzzy*; e $\mathbf{S}_i = \mathbf{X}^T \Psi_i \mathbf{X}$ e $\mathbf{b}_i = \mathbf{X}^T \Psi_i \mathbf{Y}$ para o algoritmo baseado no método de mínimos quadrados, onde \mathbf{X} é a matriz de dados.

A influência do ruído é analisada de acordo com a diferença entre as saídas dos modelos *fuzzy* obtidos a partir dos dados ruidosos, e a saída da planta sem ruído. Os parâmetros do antecedente e a estrutura dos modelos *fuzzy* são as mesmas nos experimentos, enquanto os parâmetros do conseqüente são obtidos pelo método proposto e pelo método LS. Assim, os resultados obtidos são devidos a estes algoritmos e conclusões precisas serão obtidas sobre o desempenho do algoritmo proposto na presença de ruído. Dois critérios são aplicados para validação dos modelos *fuzzy* obtidos : VAF (*Variance Accounted For*)

$$\text{VAF}(\%) = 100 \times \left[1 - \frac{\text{var}(\mathbf{Y} - \hat{\mathbf{Y}})}{\text{var}(\mathbf{Y})} \right] \quad (47)$$

onde \mathbf{Y} é a saída nominal da planta, $\hat{\mathbf{Y}}$ é a saída do modelo *fuzzy* e var é a variância do sinal, e MSE (*Mean Square Error*)

$$\text{MSE} = \frac{1}{N} \sum_{k=1}^N (y(k) - \hat{y}(k))^2 \quad (48)$$

onde $y(k)$ é a saída nominal da planta, $\hat{y}(k)$ é a saída do modelo *fuzzy*, N é o número de pontos. Uma vez obtidos estes valores, uma análise comparativa é estabelecida entre o algoritmo proposto, baseado em IV, e o algoritmo baseado em LS, de acordo com as abordagens apresentadas acima. A Fig. 2 mostra os desempenhos dos modelos obtidos *off-line* de acordo com (45) e (46). Para estes resultados, o número de pontos foi 500, o algoritmo proposto usou um fator de esquecimento λ igual a 0.99; o número de regras foi 4, a estrutura foi apresentada em (44) e os parâmetros do antecedente foram obtidos pelo método ECM para ambos os algoritmos. O fator de esquecimento bem como o número de regras foram escolhidos levando-se em consideração a velocidade de estimação paramétrica bem como a simplicidade do modelo *fuzzy*. A partir da Fig. 2 pode-se ver, claramente, que o algoritmo proposto apresenta melhor desempenho que o algoritmo baseado em LS nas duas abordagens, sendo mais robusto ao ruído. Isto é devido à matriz de variável instrumental escolhida baseada na seqüência de entrada e saída atrasada, com $ld = 1$, satisfazendo as condições dos Lemas e Teorema propostos tão bem quanto possível. Na abordagem global, para baixa variância do ruído, ambos algoritmos apresentaram desempenhos similares com VAF e MSE de 99.50% e 0.0071 para o algoritmo proposto e de 99.56% e 0.0027 para o

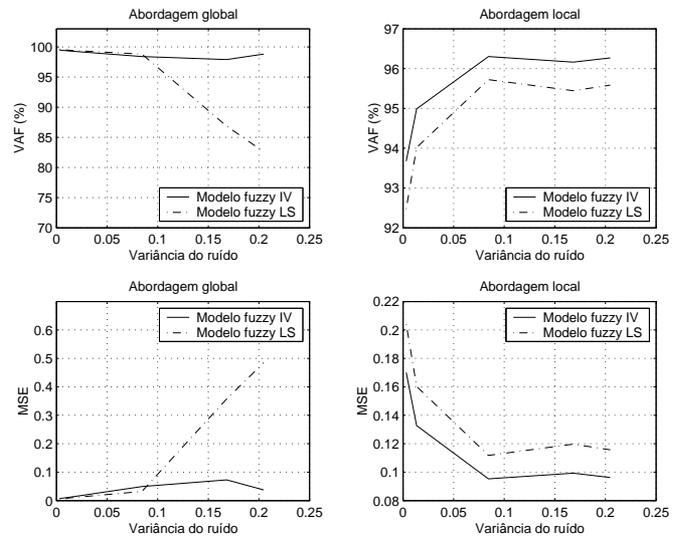


Figura 2: Análise comparativa

algoritmo baseado em LS, respectivamente. Porém, quando a variância do ruído aumenta, a matriz de variável instrumental escolhida satisfaz os Lemas e Teorema propostos, o que, conseqüentemente, torna o algoritmo proposto mais robusto ao ruído com VAF e MSE de 98.81% e 0.0375, enquanto que o algoritmo baseado em LS apresentou VAF e MSE de 82.61% e 0.4847, respectivamente. Uma análise similar pode ser feita com a abordagem local: aumentando a variância do ruído, ambos algoritmos apresentam bom desempenho onde os valores de VAF e MSE aumentam também. Isto é devido à propriedade politópica (ver seção 2), onde os modelos obtidos podem representar aproximações locais com maior flexibilidade. O algoritmo proposto apresentou valores de VAF e MSE de 93.70% e 0.1701 para o pior caso e de 96.3% e 0.0962 para o melhor caso. O algoritmo baseado em LS apresentou valores de VAF e MSE de 92.4% e 0.2042 para o pior caso e de 95.5% e 0.1157 para o melhor caso. A Fig. 3 mostra parte dos dados de saída ruidosos para, $\sigma^2 = 0.24$, bem como as saídas dos modelos *fuzzy* para ambos algoritmos e a saída da planta livre de ruído para ambas abordagens objetivando a validação dos modelos. Claramente, pode ser observada a precisão da estimação da saída para o método proposto para ambas abordagens. Para a abordagem global, os valores de VAF e MSE foram de 96.5% e 0.09 para o método proposto e de 81.4% e 0.52 para o algoritmo baseado em LS, respectivamente. Para a abordagem local, os valores de VAF e MSE foram de 96.0% e 0.109 para o método proposto e de 95.5% e 0.1187 para o algoritmo baseado em LS, respectivamente. Para a versão recursiva implementada de acordo com a seção 4.3, a análise de desempenho do algoritmo proposto é feita para a seguinte

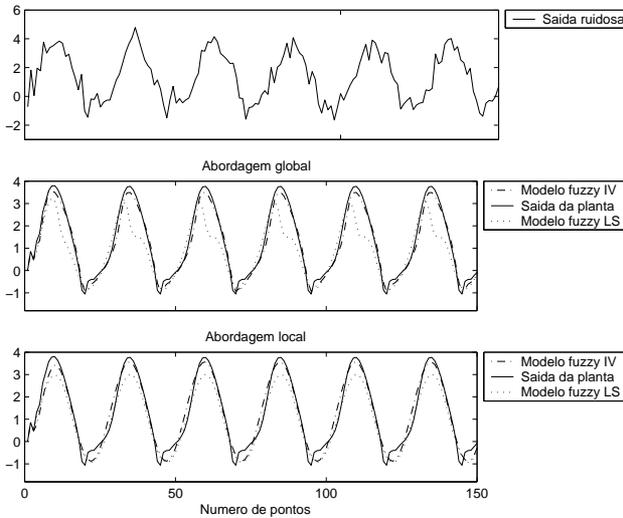


Figura 3: Validação dos modelos

planta não-linear

$$y(k+1) = \frac{y(k)y(k-1)(y(k) + prmt(k))}{1 + y(k)^2 + y(k-1)^2} + u(k) + e(k) \quad (49)$$

onde $prmt(k) = 2.5 + 0.1 \sin(\frac{2\pi k}{45})$. A Fig. 4 mostra o rastreamento da saída da planta não-linear variante no tempo em um ambiente ruidoso com $\sigma^2 = 0.01$. Nesta aplicação, 25 pontos foram usados para fazer a estimação inicial e $\lambda = 0.99$. Como esperado, o método proposto fornece uma estimação da saída com alta precisão, uma estimação consistente e não-polarizada dos parâmetros *fuzzy* e, conseqüentemente, alta velocidade de convergência

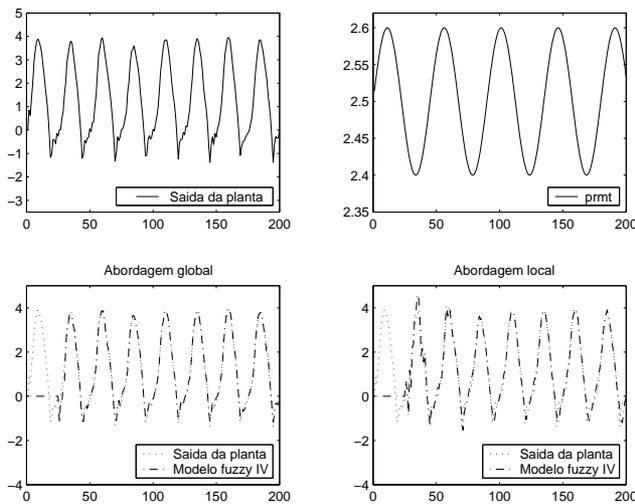


Figura 4: Rastreamento da saída da planta não-linear variante no tempo

dos modelos *fuzzy* para seguir o comportamento variante no tempo da planta em um ambiente ruidoso. Estas características são muito importantes em aplicações ao projeto de controle adaptativo.

5.2 Estudo de caso: Viga livre-livre

Esta seção estaremos dedicando à identificação de uma viga de alumínio, uma planta complexa, não-linear e variante com o tempo, cujas dimensões são: 1100x32x3 mm cujo estudo serve como fundamento para aplicações em estruturas mecânicas de aeronaves e/ou veículos aeroespaciais. A estrutura geral do experimento é do tipo livre-livre, ou seja, nenhuma das suas extremidades está fixa e a viga está suspensa por fios de *nilon* conforme mostra a Fig.5.

Esta viga faz parte do laboratório de Vibroacústica do Departamento de Mecânica Computacional na Faculdade de Engenharia Mecânica da UNICAMP. Vários estudos têm sido aplicados com o objetivo de se obter um modelo suficientemente preciso para fins de controle ativo de vibração. Foram utilizados 49152 pontos num intervalo de 49,152 s, que implica num intervalo de amostragem de 1 ms.

Um sinal senoidal com frequência variando de 1 a 125 Hz foi utilizado como entrada, em três repetições, e a vibração (leitura a partir de um micro-acelerômetro) da viga de alumínio como a saída ruidosa. O conjunto de dados de entrada-saída durante todo o experimento são mostrados na Fig. 6 (a) e (c), enquanto que os primeiros 20ms dos dados, para fins de ilustração, são apresentados na Fig. 6 (b) e (d). Este tipo de sinal de entrada apresenta boas propriedades estatísticas e é de grande interesse prático pois permite que os modos da planta, na varredura da frequência, sejam excitados e um modelo mais preciso seja obtido. Foram utilizados os primeiros 24576 pontos para estimação paramétrica e os próximos 24576, o que equivale à segunda metade do conjunto de dados, para a validação. Neste caso, as variáveis instrumentais *fuzzy* são as variáveis linguísticas y_k e y_{k-1} filtradas, geradas automaticamente a partir do modelo auxiliar *fuzzy*, e usada para criar as funções de

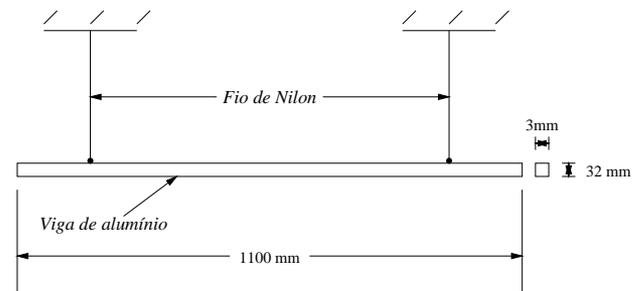


Figura 5: Viga de alumínio livre-livre

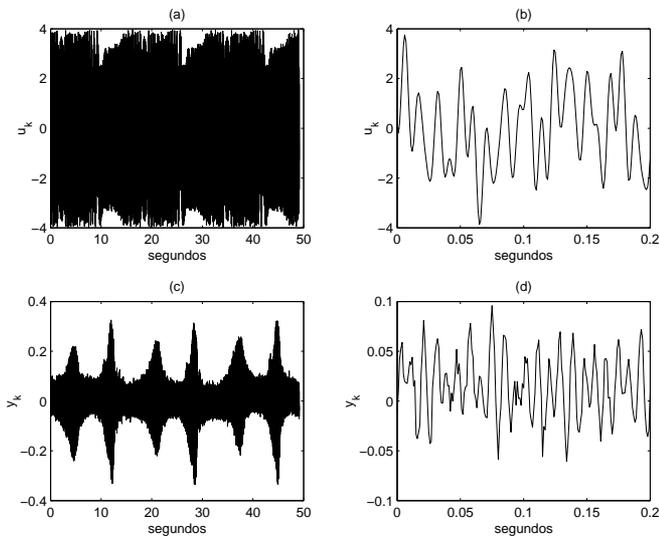


Figura 6: Teste dinâmico em uma viga livre-livre. Sinais de (a-b) entrada e (c-d) saída resultantes de um teste experimental. A entrada é o sinal que comanda a viga e a saída é a vibração.

pertinência bem como a matriz de variável instrumental *fuzzy* para estimação ótima dos parâmetros do consequente. A Fig. 7 mostra as partições das variáveis linguísticas do modelo *fuzzy* obtido. A Fig. 8 mostra os parâmetros estimados usando o algoritmo de variável instrumental *fuzzy* recursivo, e na Fig. 9 a matriz de covariância P_k na última recursão. Esta matriz apresenta valores elevados na diagonal principal, implicando num melhor ajuste dos parâmetros estimados a cada recursão; e é simétrica, o que representa a consistência e estabilidade do algoritmo proposto para estimação dos parâmetros do consequente. A base de regras do modelo *fuzzy* obtido na última recursão é da seguinte forma:

$$R^1 : SE \hat{y}_k \in F_1 \text{ E } \hat{y}_{k-1} \in G_1 \text{ ENTÃO}$$

$$\hat{y}_{k+1} = 6.3565\hat{y}_k - 3.4363\hat{y}_{k-1} - 0.0025u_k - 0.6552$$

$$R^2 : SE \hat{y}_k \in F_1 \text{ E } \hat{y}_{k-1} \in G_2 \text{ ENTÃO}$$

$$\hat{y}_{k+1} = 6.0951\hat{y}_k - 5.2626\hat{y}_{k-1} + 0.0067u_k - 2.0973$$

$$R^3 : SE \hat{y}_k \in F_2 \text{ E } \hat{y}_{k-1} \in G_1 \text{ ENTÃO}$$

$$\hat{y}_{k+1} = 7.3567\hat{y}_k - 5.4944\hat{y}_{k-1} + 0.0037u_k + 3.1799$$

$$R^4 : SE \hat{y}_k \in F_2 \text{ E } \hat{y}_{k-1} \in G_2 \text{ ENTÃO}$$

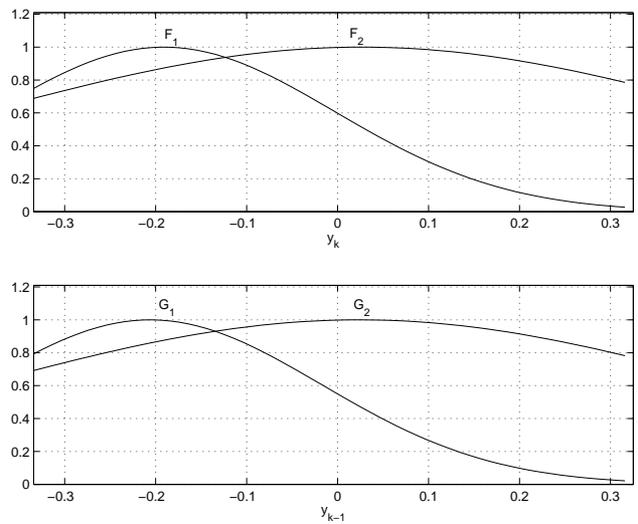


Figura 7: Partição das variáveis linguísticas do modelo *fuzzy* identificado.

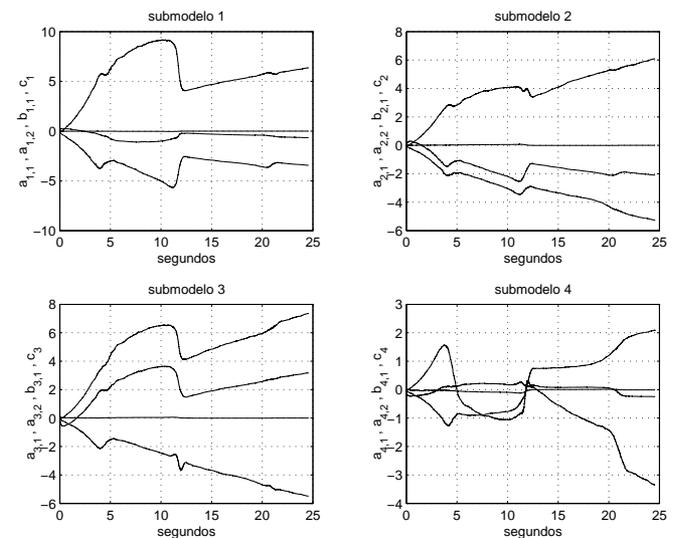


Figura 8: Estimação recursiva dos parâmetros do modelo *fuzzy* para a viga de alumínio.

$$\hat{y}_{k+1} = -3.3510\hat{y}_k + 2.0907\hat{y}_{k-1} - 0.0112u_k - 0.2468$$

O erro de estimação, o qual representa o ruído a partir dos componentes eletrônicos incluindo o atuador, sensor, sistema de aquisição de dados e amplificador de sinais, apresentou um valor quadrático médio MSE (*Mean Square Error*) de 4.2808×10^{-4} . A Fig. 10 mostra a validação do modelo *fuzzy* obtido de acordo com a curva da função de resposta em frequência (FRF) (Juang, 1994) da viga de alumínio. A complexidade consiste nas três frequências naturais da viga

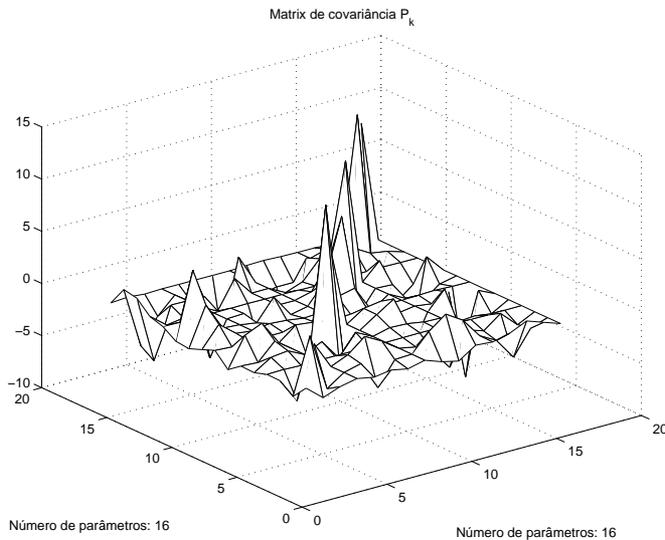


Figura 9: Matrix de covariância P_k .

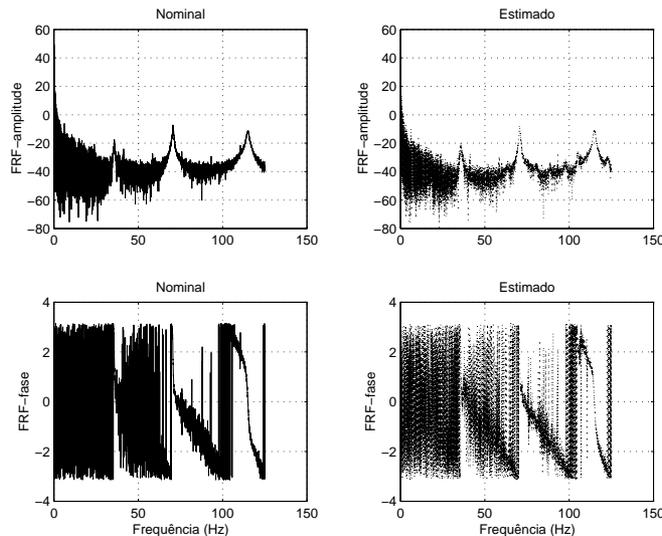


Figura 10: Validação do modelo.

de alumínio representadas pelos picos em 35 Hz, 70 Hz e 120 Hz, respectivamente. Claramente, o modelo *fuzzy* obtido é suficientemente preciso na faixa de interesse do experimento e pode ser usado eficientemente no projeto de controle ativo de vibração.

6 CONCLUSÃO

A partir do conceito de variável instrumental *fuzzy* proposto em (Bottura and Serra, 2005), métodos de variável instrumental *fuzzy* para identificação de sistemas em ambientes ruidosos são propostos teoricamente. Condições de convergência para identificação em ambiente ruidoso em

um contexto *fuzzy* foram estudadas. Um estudo de caso à identificação de uma viga de alumínio (uma planta complexa, não-linear e variante com o tempo) cujo estudo serve como fundamento para aplicações em estruturas mecânicas de aeronaves e/ou veículos aeroespaciais, foi apresentado. Resultados de simulação confirmaram a eficiência do método proposto que tem as seguintes características: a matriz de variável instrumental escolhida satisfaz os Lemas e Teoremas propostos, tornando o algoritmo proposto robusto ao ruído; estimação não-polarizada e consistente dos parâmetros *fuzzy*; convergência dos modelos *fuzzy* ao comportamento variante no tempo da planta não-linear; aplicabilidade ao projeto de controle adaptativo. Para trabalho futuro, a generalização do método proposto para plantas MIMO bem como sua aplicação ao projeto de controle adaptativo devem ser considerados.

AGRADECIMENTOS

Os autores gostariam de agradecer à CAPES e ao CNPq pelo apoio financeiro a este trabalho.

REFERÊNCIAS

- Bottura, C.P. and Serra, G.L.O. (2004). An Algorithm for Fuzzy Identification of Nonlinear Discrete-Time Systems, *43rd IEEE Conference on Decision and Control CDC'04*, Atlantis(Paradise Island)Bahamas, December, pp. 5421-5426.
- ^aBottura, C.P. and Serra, G.L.O. (2005). Fuzzy Instrumental Variable Concept and Identification Algorithm, *IEEE International Conference on Fuzzy Systems FUZZY-IEEE'05*, Nevada-EUA, May 22-25.
- ^bBottura, C.P. and Serra, G.L.O. (2005). Variável Instrumental *Fuzzy*, *IV Congresso Temático de Dinâmica e Controle DINCON'05*, Baurú-SP, Junho 6-10.
- Brown, M. and Harris, C. (1994). *Neurofuzzy Adaptive Modelling and Control*, Prentice Hall.
- Goodwin, G.C. and Payne, R.L. (1977). *Dynamic System Identification: Experiment Design and Data Analysis*, Academic Press.
- Juang, Jer-Nan. (1994). *Applied System Identification*, Prentice Hall.
- Kasabov, N.K. and Song, Q. (2002). DENFIS: Dynamic Evolving Neural-Fuzzy Inference System and Its Application for Time-Series Prediction, *IEEE Transactions on Fuzzy Systems*, Vol. 10, No 2, pp. 144-154.

- Kukolj, D. and Levi, E. (2004). Identification of Complex Systems Based on Neural and Takagi-Sugeno Fuzzy Model, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 34, No 1, pp. 272-282.
- Kuma, M., Stoll, R. and Stoll, N. (2006). A Robust Design Criterion for Interpretable Fuzzy Models with Uncertain Data, *IEEE Transactions on Fuzzy Systems*, Vol. 14, No 2, pp. 314-328.
- Ljung, L. (1999). *System Identification: Theory for user*, 2nd ed., Prentice Hall.
- Ljung, L. and Söderström, T. (1987). *Theory and Practice of Recursive Identification*, MIT PRESS.
- Papadakis, S.E. and Theocaris, J.B. (2002). A GA-based Fuzzy Modeling Approach for Generating TSK Models, *Fuzzy Sets and Systems*, Vol. 131, No 2, pp. 121-152.
- Serra, G.L.O. and Bottura, C.P. (2006). Multiobjective Evolution Based Fuzzy PI Controller Design For Nonlinear Systems, *IFAC Engineering Applications of Artificial Intelligence*, Vol. 19, No. 2, pp. 157-167.
- Serra, G.L.O. (2005). *Propostas de Metodologias para Identificação e Controle Inteligentes*, Tese de Doutorado, Laboratório de Controle e Sistemas Inteligentes, UNICAMP/FEEC/DMCSI, Campinas-SP.
- Skrjanc, L., Blazic, S. and Agamennoni, O. (2005). Interval Fuzzy Model Identification Using l_∞ -Norm, *IEEE Transactions on Fuzzy Systems*, Vol. 13, No 5, pp. 561-568.
- ^aSöderström, T. and Stoica, P.G. (1989). *System Identification*, Prentice Hall.
- ^bSöderström, T. and Stoica, P.G. (1983). *Instrumental Variable Methods for System Identification*, Lecture Notes in Control and Information Sciences, Springer-Verlag.
- Takagi, T. and Sugeno, M. (1985). Fuzzy Identification of Systems and Its Application to Modeling and Control, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 15, No. 1, pp. 116-132.
- Tsoukalas, L.H. and Uhrig, R.E. (1997). *Fuzzy and Neural Approaches in Engineering*, Wiley-Interscience.
- Young, P.C. (1984). *Recursive Estimation and Time-Series Analysis: An Introduction*, Springer-Verlag.