

Universidade Estadual de Campinas  
Faculdade de Engenharia Elétrica e de Computação  
Departamento de Computação e Automação Industrial

# Uma Nova Metodologia para Identificação Adaptativa de Modelos Relacionais Nebulosos

**Autor:** Ricardo José Gabrielli Barreto Campello  
**Orientador:** Prof. Dr. Wagner Caradori do Amaral †

Este exemplar corresponde a redação final da tese  
defendida por Ricardo José Gabrielli  
Barreto Campello e aprovada pela Comissão  
Julgada em 13 / 06 / 97  
Wagner  
Orientador

Dissertação apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas - FEEC - UNICAMP, como parte dos requisitos exigidos para a obtenção do título de MESTRE EM ENGENHARIA ELÉTRICA.

Campinas, Maio de 1997.

DE	BC
AMADA:	unicamp
53-ne	
Ex.	
BC/	31521
	281197
<input type="checkbox"/>	D <input checked="" type="checkbox"/>
RB	11.00
	30102197
PD	

CM-00099917-0

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

C153n      Campello, Ricardo José Gabrielli Barreto  
Uma nova metodologia para identificação adaptativa de  
modelos relacionais nebulosos / Ricardo José Gabrielli  
Barreto Campello.--Campinas, SP: [s.n.], 1997.

Orientador: Wagner Caradori do Amaral.  
Dissertação (mestrado) - Universidade Estadual de  
Campinas, Faculdade de Engenharia Elétrica e de  
Computação.

1. Identificação de sistemas.\* 2. Sistemas difusos.\* 3.  
Modelos matemáticos.\* I. Amaral, Wagner Caradori do.  
II. Universidade Estadual de Campinas. Faculdade de  
Engenharia Elétrica e de Computação. III. Título.

---

*À minha esposa*  
*Adriana*

# Agradecimentos

Agradeço inicialmente a Deus, pela graça de nossa existência.

Agradeço de maneira especial à minha querida esposa Adriana por todo amor e companheirismo que me sustentaram nessa difícil jornada e que sempre me sustentarão ao longo de nossas vidas.

Agradeço aos meus pais e irmãos por minha educação e pelo incentivo, e aos meus tios Cecília e Francisco, minha segunda família.

Agradeço ao Prof. Dr. Wagner Caradori do Amaral pela amizade e orientação que me engrandeceram pessoal e profissionalmente.

Aos amigos Rubén, Gustavo, Jussara, Gonzaga, Ivan, Samuel, Fábio e em especial ao Maurício Carvalho, agradeço pela amizade e pelo apoio moral e técnico.

Agradeço também ao CNPq pelo apoio financeiro e ao FAEP pelo patrocínio à divulgação desse trabalho no exterior.

Por fim, agradeço a todos aqueles professores e funcionários da UNICAMP, entre outros, que direta ou indiretamente contribuíram para tornar possível a realização desse trabalho.

# Índice

<b>Resumo</b>	<b>iii</b>
<b>1 Introdução à Modelagem Nebulosa de Sistemas</b>	<b>1</b>
1.1 Introdução . . . . .	1
1.2 Modelos nebulosos . . . . .	3
1.2.1 Modelos baseados em regras . . . . .	3
1.2.2 Modelos funcionais . . . . .	4
1.2.3 Modelos baseados em relação . . . . .	5
1.3 Sumário . . . . .	8
<b>2 Modelos Relacionais Nebulosos</b>	<b>9</b>
2.1 Introdução . . . . .	9
2.2 Discretização nebulosa . . . . .	11
2.3 Identificação de modelos relacionais nebulosos . . . . .	15
2.4 Sumário . . . . .	16
<b>3 Interfaces de Entrada e Saída</b>	<b>18</b>
3.1 Introdução . . . . .	18
3.2 Estrutura de cognição . . . . .	19
3.3 Otimização de interfaces . . . . .	20
3.3.1 Critério de equivalência de informações . . . . .	21
3.3.2 Algoritmo $\Sigma$ -PAFIO . . . . .	22
3.4 Exemplo numérico . . . . .	27
3.5 Sumário . . . . .	36
<b>4 Algoritmo de Identificação Adaptativa</b>	<b>37</b>
4.1 Introdução . . . . .	37
4.2 O modelo nebuloso . . . . .	38

4.3	O problema de otimização . . . . .	39
4.3.1	O ganho $G$ da função de custo . . . . .	40
4.3.2	Solução do problema de otimização . . . . .	41
4.4	Algoritmo de identificação . . . . .	45
4.5	Topologias de realimentação . . . . .	46
4.6	O modelo incremental . . . . .	47
4.7	Sumário . . . . .	49
<b>5</b>	<b>Exemplos Numéricos de Identificação Adaptativa</b>	<b>50</b>
5.1	Introdução . . . . .	50
5.2	Exemplo 1 . . . . .	51
5.3	Exemplo 2 . . . . .	54
5.4	Exemplo 3 . . . . .	59
5.5	Sumário . . . . .	61
<b>6</b>	<b>Conclusões e Perspectivas</b>	<b>63</b>
	<b>Apêndice A: Algoritmo de Aproximação Linguística</b>	<b>65</b>
	<b>Apêndice B: Algoritmo <math>\Sigma</math>-PAFIO “On-line”</b>	<b>67</b>
	<b>Apêndice C: Definições Básicas</b>	<b>71</b>
	<b>Bibliografia</b>	<b>74</b>

# Resumo

Modelos nebulosos constituem uma importante ferramenta matemática na identificação de processos complexos onde técnicas convencionais normalmente não são eficientes. Uma classe particular é formada pelos modelos relacionais nebulosos, onde o mapeamento entre variáveis de entrada e saída é realizado através de equações relacionais nebulosas. Os principais procedimentos para modelagem de sistemas utilizando esses modelos são a determinação da estrutura, a estimação dos parâmetros da relação nebulosa e a construção das interfaces de entrada e saída quando são necessárias.

Neste trabalho é proposto um algoritmo para estimação dos parâmetros da relação nebulosa em modelos relacionais de estrutura conhecida. A principal vantagem deste método, em comparação aos algoritmos existentes na literatura, é sua maior capacidade de adaptação. Essa característica evita a dependência de uma boa estimativa da relação nebulosa e torna o algoritmo adequado à modelagem de sistemas variantes no tempo. Outra característica é a utilização de um método de otimização simples que permite a identificação de sistemas nebulosos e não-nebulosos. Para construção das interfaces de entrada e saída utiliza-se um algoritmo baseado em otimização que reduz a influência das interfaces no modelo através da minimização dos erros nos mapeamentos entre grandezas numéricas e linguísticas, mantendo a integridade semântica dos conjuntos nebulosos. A otimização das interfaces de entrada e saída e a estimação dos parâmetros da relação nebulosa constituem a metodologia de identificação.

Vários processos foram identificados utilizando a metodologia proposta. Características como não-linearidades, perturbações e variações temporais foram analisadas e os modelos obtidos foram avaliados através da geração de séries sintéticas e previsões da saída um passo à frente.

# Capítulo 1

## Introdução à Modelagem Nebulosa de Sistemas

Esse capítulo apresenta uma visão geral de identificação de sistemas utilizando Lógica Nebulosa. São descritas as principais propostas encontradas na literatura e discutidos os aspectos relevantes de cada uma delas, assim como suas vantagens e desvantagens.

### 1.1 Introdução

A crescente necessidade de identificar processos complexos que frequentemente não podem ser bem representados por modelos matemáticos determinísticos ou estocásticos convencionais [Ljung, 1987, Söderström e Stoica, 1989] torna atrativa a abordagem através de Lógica Nebulosa para a identificação de sistemas. A aplicação de Lógica Nebulosa na análise e representação desse tipo de processo torna-se interessante por sua capacidade de manipular incertezas e tratar informações imprecisas de caráter linguístico [Gomide e Gudwin, 1994, Tanscheit, 1992]. Sua importância pode ser mais bem compreendida através do *princípio da incompatibilidade* estabelecido por Zadeh [Zadeh, 1973]:

*“A nossa opinião é que as técnicas quantitativas convencionais de análise de sistemas são intrinsecamente inadequadas para tratar sistemas humanísticos ou qualquer sistema cuja complexidade é comparável à dos sistemas humanísticos. Essa opinião está baseada no que poderia se denominar princípio da incompatibilidade. Definida informalmente, a essência desse princípio é que conforme*



*aumenta a complexidade de um sistema, nossa habilidade de tornar ao mesmo tempo precisas e significantes as características do seu funcionamento diminui até um nível além do qual precisão e significância (ou relevância) tornam-se características quase mutuamente exclusivas.*”

Os modelos nebulosos são, em geral, constituídos por três blocos como ilustrado na Figura (1.1). O bloco de processamento realiza o mapea-

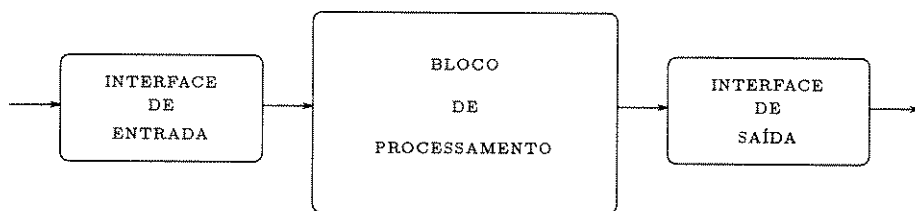


Figura 1.1: Diagrama de blocos genérico para um modelo nebuloso.

mento entre valores nebulosos de entrada e saída através dos mecanismos de implicação e inferência e as interfaces realizam as transformações entre valores numéricos (*crisp*) e linguísticos (*fuzzy*) [Pedrycz, 1995]. Nesses modelos, problemas clássicos como não-linearidades, perturbações e incertezas paramétricas são considerados genericamente com o auxílio das teorias de conjuntos nebulosos e relações nebulosas [Zadeh, 1965, Zadeh, 1973, Kaufmann, 1975, Pedrycz, 1993]. Modelos não-lineares convencionais demandam grande conhecimento do processo para uma adequada escolha da sua estrutura. Isso não ocorre com os modelos nebulosos que são flexíveis e necessitam apenas de valores aproximados para a estrutura tais como ordem e atraso de transporte do processo [Tong, 1978, Sugeno e Kang, 1988]. Esses modelos têm sido aplicados com sucesso em sistemas com estrutura desconhecida como, por exemplo, na identificação de uma fornalha de gás [Pedrycz, 1984a], de um incinerador multicamadas [Sugeno e Kang, 1986] e na previsão do fluxo de água na barragem de um rio [Sugeno e Tanaka, 1991].

Em aplicações de controle, a modelagem nebulosa é tão importante para a obtenção de um modelo da planta usado no projeto do controlador, como

também para a obtenção do próprio controlador através da identificação das ações de controle de um especialista na operação do processo. Nesses casos, o controle realizado por um operador humano em um determinado processo pode ser identificado, como foi feito, por exemplo, para uma planta química [Sugeno e Yasukawa, 1993] e para um processo de limpeza de água [Takagi e Sugeno, 1985]. Nesses controladores é possível a incorporação de restrições linguísticas de origem humana extremamente úteis em aplicações práticas.

## 1.2 Modelos nebulosos

Os modelos nebulosos são divididos em três classes principais:

- Modelos baseados em regras.
- Modelos funcionais (*Functional models*).
- Modelos baseados em relação.

### 1.2.1 Modelos baseados em regras

Os modelos baseados em regras ou modelos linguísticos, como o próprio nome sugere, são formados por um conjunto de regras linguísticas do tipo

$$\text{SE } x \text{ é } X_i \text{ ENTÃO } y \text{ é } Y_i$$

onde  $x$  é a entrada do modelo (antecedente),  $y$  é a saída (consequente) e  $X_i$  e  $Y_i$  são rótulos linguísticos (pequeno, grande, etc) associados a conjuntos nebulosos nos universos de discurso das respectivas variáveis. Esses modelos, pioneiros no uso de Lógica Nebulosa em identificação de sistemas, foram inicialmente propostos por Tong [Tong, 1978]. Nesse trabalho inicial, Tong utilizou um método “off-line” de investigação lógica (*logical examination*) para determinação da base de regras que constitui o modelo. Esse método analisa um conjunto de dados de entrada e saída, selecionando apenas aqueles dados com valor de pertinência unitário para todas as variáveis do sistema. Como a probabilidade de ocorrência dessa combinação diminui com o aumento do número de variáveis, o método necessita de uma enorme quantidade de dados para análise no caso de modelos com estrutura complexa. Após essa análise,

os conjuntos nebulosos associados ao valor de pertinência unitário referentes a cada conjunto de dados de entrada e saída selecionado constituem uma regra. Esse procedimento permite a existência de regras conflitantes, que indicam saídas diferentes para uma mesma combinação de entradas. Também pode ocorrer a inexistência de regras, quando nenhuma saída é indicada para determinada combinação de entradas, implicando em um modelo incompleto. Para contornar esses problemas e gerar um modelo final consistente, Tong utilizou a combinação de métodos estatísticos, heurísticos e de tentativa-e-erro. Posteriormente, Graham e Newell [Graham e Newell, 1989] propuseram uma generalização do método de Tong, baseada no conceito de conjuntos nebulosos de referência [Pedrycz, 1984a], agregando certa flexibilidade aos modelos através da inserção automática de regras associada à ponderação das regras conflitantes.

Os modelos baseados em regras são de difícil aplicação em sistemas multivariáveis devido ao aumento da complexidade da inferência no conjunto de regras em função do aumento da complexidade do modelo. Esses modelos tornam-se atrativos quando se deseja conhecimento qualitativo sobre o processo mais que conhecimento quantitativo.

### 1.2.2 Modelos funcionais

Os modelos nebulosos funcionais [Takagi e Sugeno, 1985, Sugeno e Tanaka, 1991, Kroll, 1996], frequentemente denominados de modelos do tipo de Sugeno, são estruturalmente constituídos por uma base de regras e baseados em um tipo diferente de implicação como segue:

$$\text{SE } x_1 \text{ é } X_1^i \text{ e } x_2 \text{ é } X_2^i \text{ e } \dots \text{ e } x_k \text{ é } X_k^i \text{ ENTÃO } y = f^i(x_1, x_2, \dots, x_k)$$

onde  $y$  é a variável consequente cujo valor é inferido,  $x_1, \dots, x_k$  são as variáveis antecedentes que também ocorrem em parte do consequente,  $X_1^i, \dots, X_k^i$  são conjuntos nebulosos e  $f^i$  é uma função que realiza o mapeamento dos valores de  $x_1, \dots, x_k$  em  $y$ . As funções  $f^i$  são normalmente lineares em seus argumentos, ou seja,

$$f^i(x_1, x_2, \dots, x_k) = p_i(0) + \sum_{j=1}^k p_i(j)x_j$$

Os parâmetros  $p_i$  podem ser estimados por métodos convencionais como Mínimos Quadrados ou Filtro de Kalman [Ljung, 1987,

Söderström e Stoica, 1989]. A identificação de modelos funcionais possui as dificuldades inerentes a esses métodos além das particularidades referentes aos conjuntos nebulosos e às regras, o que a torna complexa. Além disso, embora sejam precisos, os modelos funcionais praticamente não fornecem conhecimento qualitativo sobre o processo e não são adequados à identificação de sistemas nebulosos por manipularem variáveis consequentes numéricas.

### 1.2.3 Modelos baseados em relação

Os modelos baseados em relação ou modelos relacionais são constituídos por uma equação relacional nebulosa [Pedrycz, 1993] e não por um conjunto de regras. Essa equação realiza o mapeamento da saída do modelo a partir da composição entre as entradas e a relação nebulosa. Para universos de discurso finitos, ou seja, definidos como vetores contendo um número finito de elementos, essa relação é representada por uma matriz denominada matriz relacional. A principal vantagem dos modelos relacionais é que, tendo sido definida sua estrutura, o problema de identificação restringe-se à estimação dos elementos dessa matriz. Outra vantagem é a simplicidade da composição nebulosa e, em consequência, a facilidade de se obter a saída a partir de um determinado conjunto de entradas. Os modelos relacionais baseados em conjuntos nebulosos de referência também fornecem um elevado conhecimento qualitativo sobre o processo, sendo que cada elemento da matriz relacional representa uma medida de possibilidade [Zadeh, 1978] da relação entre os conjuntos nebulosos de entrada e saída [Pedrycz, 1984a, Graham e Newell, 1989].

O trabalho pioneiro na identificação de modelos relacionais [Czogala e Pedrycz, 1981] propôs a utilização de métodos analíticos para resolver a equação relacional nebulosa que descreve o sistema a partir de um conjunto de dados de entrada e saída. A partir desse trabalho vários outros seguiram utilizando essa técnica [Pedrycz, 1981, Czogala e Pedrycz, 1982, Pedrycz et al., 1984, Pedrycz, 1984b, Higashi e Klir, 1984a]. Um conjunto de métodos analíticos para solução de equações relacionais é dado na literatura [Sanchez, 1976, Higashi e Klir, 1984b, Pedrycz, 1985b, Pedrycz, 1985a, Di Nola et al., 1991, Pedrycz, 1993]. Esses métodos tem como hipótese básica a existência de uma família de soluções não-vazia. Como essa condição não pode ser verificada *a priori*, a sua utilização pode gerar relações nebulosas inconsistentes mas, em geral, fornece uma condição inicial aproximada para

métodos mais refinados. Essa aproximação é denominada aproximação de “força-bruta” [Pedrycz, 1995].

Devido à dificuldade de existência de soluções exatas para as equações relacionais na presença de um número elevado de dados, um grupo de métodos foi desenvolvido para obtenção “off-line” de soluções aproximadas para essas equações através da seleção e/ou alteração desses dados. Pedrycz propôs a ponderação dos dados utilizando técnicas de “clustering” [Pedrycz, 1984b]. Através dessas técnicas, os conjuntos nebulosos são agrupados e substituídos em cada grupo (*cluster*) por um conjunto nebuloso resultante chamado conjunto probabilístico. Esse conjunto é caracterizado por uma função de pertinência calculada como a média aritmética, ponto-a-ponto, das funções de pertinência pertencentes ao grupo. A quantidade de dados é reduzida e a equação relacional pode ser mais facilmente solucionada. A solução gerada satisfaz na média os dados originais. Posteriormente, Gottwald e Pedrycz [Gottwald e Pedrycz, 1986] utilizaram uma medida de solucionabilidade para equações relacionais nebulosas. Essa medida é diretamente proporcional à similaridade entre a saída da equação para a melhor solução possível e o respectivo conjunto nebuloso real de saída. Essa similaridade é avaliada por um índice de igualdade entre conjuntos nebulosos. Nessa técnica, selecionam-se os dados que proporcionam maior medida de solucionabilidade para a equação ou ainda alteram-se os conjuntos nebulosos para aumentar essa medida. Pedrycz propôs ainda uma seleção dos dados através da análise da consistência desses dados também avaliada através de um índice de igualdade entre conjuntos nebulosos [Pedrycz, 1988]. Essa análise é realizada considerando que a existência de saídas diferentes para um mesmo conjunto de entradas representa uma inconsistência. Utiliza-se a medida de similaridade de conjuntos nebulosos de entrada e saída para representar o grau de factibilidade (*degree of feasibility*) das equações relacionais e eliminar conjuntos de dados inconsistentes. Pedrycz também utilizou uma medida representativa da capacidade de solução da equação relacional para cada par de dados de entrada e saída. Essa medida é baseada na condição básica de existência de soluções para as equações relacionais nebulosas [Sanchez, 1976, Pedrycz, 1988, Pedrycz, 1993] e é representada por um vetor de probabilidades. Qualquer solução é infactível quando esse vetor é nulo. Nesse caso, pode-se eliminar o par de dados em questão, aumentando assim a solucionabilidade da equação.

Outra abordagem para determinação da matriz relacional, desenvolvida segundo a teoria clássica de Lógica Nebulosa, é conhecida como

aproximação linguística [Tong, 1978, Pedrycz, 1984a, Graham e Newell, 1988]. Nessa técnica, a matriz relacional é obtida a partir da agregação dos Produtos Cartesianos (Apêndice C) dos conjuntos nebulosos de entrada e saída. Essa abordagem é extremamente simples mas, como os modelos linguísticos baseados em regras, é intrinsecamente imprecisa e está sujeita ao acúmulo de relações conflitantes ou inconsistentes em aplicações “on-line”, onde não é possível a análise prévia do conjunto de dados. A aproximação linguística fornece uma condição inicial aproximada para métodos mais refinados e, em função disso, um algoritmo é apresentado no Apêndice A.

Recentemente, diversas aproximações numéricas foram propostas para refinar modelos relacionais previamente obtidos por outros métodos como aproximação linguística ou aproximação de força-bruta. Pedrycz utilizou o método numérico de Newton para a obtenção “off-line” de uma aproximação da matriz relacional nebulosa [Pedrycz, 1983] a partir de um conjunto de dados. Em relação a aplicações “on-line”, Graham e Newell [Graham e Newell, 1989] e Xu e Lu [Xu e Lu, 1987, Xu, 1989] desenvolveram algoritmos com características adaptativas a partir de fatores de ponderação utilizados na atualização da matriz relacional. Contudo, esses fatores são de difícil ajuste. Nesses algoritmos, a cada iteração os elementos predominantes da composição nebulosa são selecionados e os respectivos elementos da matriz relacional são calculados e ponderados com os valores existentes na iteração anterior. Posteriormente, Lee *et al.* [Lee et al., 1994] propuseram um algoritmo recursivo baseado na minimização de um critério de otimização que implica na minimização aproximada do erro entre as saídas do sistema real e do modelo. O desempenho desse e dos outros métodos numéricos adaptativos depende de uma boa estimativa da relação nebulosa e pode ser degradado se ocorrer uma variação abrupta no sistema. Essa dependência também existe na aproximação proposta por Oliveira e Lemos [Oliveira e Lemos, 1993, Oliveira e Lemos, 1995] com respeito à escolha dos elementos da matriz relacional a serem otimizados. No Capítulo 4 propõe-se uma solução para esse problema que utiliza uma nova metodologia baseada em otimização para a identificação de modelos relacionais. Essa metodologia utiliza conhecimento heurístico para resolver o problema de otimização quando a matriz relacional está polarizada, ou seja, quando uma boa estimativa da relação nebulosa não é disponível.

Por constituírem a principal motivação desse trabalho os modelos relacionais nebulosos são detalhados no Capítulo 2. No Capítulo 3, são discutidas as interfaces de entrada e saída e o método de otimização adotado para essas

interfaces. No Capítulo 4, é apresentada a nova proposta para identificação de modelos relacionais [Campello et al., 1997]. Simulações utilizando o algoritmo proposto são providas no Capítulo 5 e as conclusões e perspectivas são apresentadas no Capítulo 6.

### **1.3 Sumário**

Modelos nebulosos constituem uma importante ferramenta matemática para identificação de sistemas complexos, onde técnicas convencionais podem não ser eficientes. As três classes principais de modelos nebulosos são os modelos baseados em regras, os modelos funcionais e os modelos relacionais nebulosos. Os modelos baseados em regras são de difícil aplicação em sistemas multi-variáveis devido à complexidade da inferência no conjunto de regras e são intrinsecamente imprecisos devido à sua construção puramente heurística. Contudo podem ser utilizados quando se deseja ter conhecimento qualitativo sobre o sistema. Os modelos funcionais possuem não apenas as dificuldades inerentes aos métodos convencionais de identificação, mas também as particularidades referentes aos conjuntos nebulosos e às regras, o que torna a modelagem complexa. Embora sejam precisos, esses modelos fornecem pouco conhecimento qualitativo e não são adequados à identificação de sistemas não nebulosos em função de manipularem variáveis consequentes numéricas, ao invés de variáveis linguísticas. Os modelos relacionais fornecem grande conhecimento qualitativo quando implementados com base no conceito de conjuntos nebulosos de referência. Nessa abordagem, definida a estrutura do modelo, o problema de identificação restringe-se à estimação dos parâmetros da matriz relacional. A inferência da saída é simples, calculada através da composição nebulosa entre as entradas e a matriz. Portanto, sua complexidade independe da estrutura do modelo. Os modelos relacionais podem ser identificados através da solução analítica de equações relacionais nebulosas, através de aproximação linguística ou através de métodos numéricos. Entre essas técnicas, os métodos numéricos adaptativos são de particular interesse para aplicações “on-line”. O principal problema desses métodos é a dependência de uma estimativa adequada da matriz relacional, que não é disponível quando o sistema é variante no tempo. No Capítulo 4 apresenta-se uma proposta para solução desse problema.

# Capítulo 2

## Modelos Relacionais Nebulosos

Esse capítulo apresenta uma descrição de modelos relacionais nebulosos no contexto que serão utilizados ao longo desse trabalho, e introduz a discussão sobre a identificação desses modelos.

### 2.1 Introdução

Modelos relacionais nebulosos constituem uma representação matemática de sistemas através de equações relacionais nebulosas. Um sistema discreto, por exemplo com múltiplas entradas e saída única (MISO), pode ser descrito por um modelo nebuloso através da seguinte equação relacional:

$$\begin{aligned} Y(k) = & Y(k-1) \bullet Y(k-2) \bullet \dots \bullet Y(k-p) \\ & \bullet U_1(k-t_1-1) \bullet \dots \bullet U_1(k-t_1-q_1) \\ & \bullet U_2(k-t_2-1) \bullet \dots \bullet U_2(k-t_2-q_2) \\ & \vdots \\ & \bullet U_v(k-t_v-1) \bullet \dots \bullet U_v(k-t_v-q_v) \bullet R \end{aligned} \quad (2.1)$$

onde  $p$  e  $q_i$  para  $i = 1, \dots, v$  estão relacionados às ordens do modelo,  $t_i$  são os atrasos de transporte,  $U_i(k)$  e  $Y(k)$  são as representações linguísticas das entradas  $u_i(k)$  e saída  $y(k)$  numéricas no instante de amostragem  $k$ , respectivamente,  $R$  é a relação nebulosa e “ $\bullet$ ” denota o operador de composição nebulosa. Este operador e os parâmetros  $p$ ,  $q_i$  e  $t_i$  definem a estrutura do



modelo. A extensão para sistemas com múltiplas entradas e múltiplas saídas (MIMO) é direta. Nesse caso, um conjunto de equações como em (2.1) pode representar o conjunto de saídas do sistema, sendo que cada saída ocorre como entrada também em cada uma das outras equações.

A Figura (2.1) apresenta um diagrama para o modelo relacional descrito pela Equação (2.1). Os blocos  $N/L$  (Numérico/Linguístico) e  $L/N$

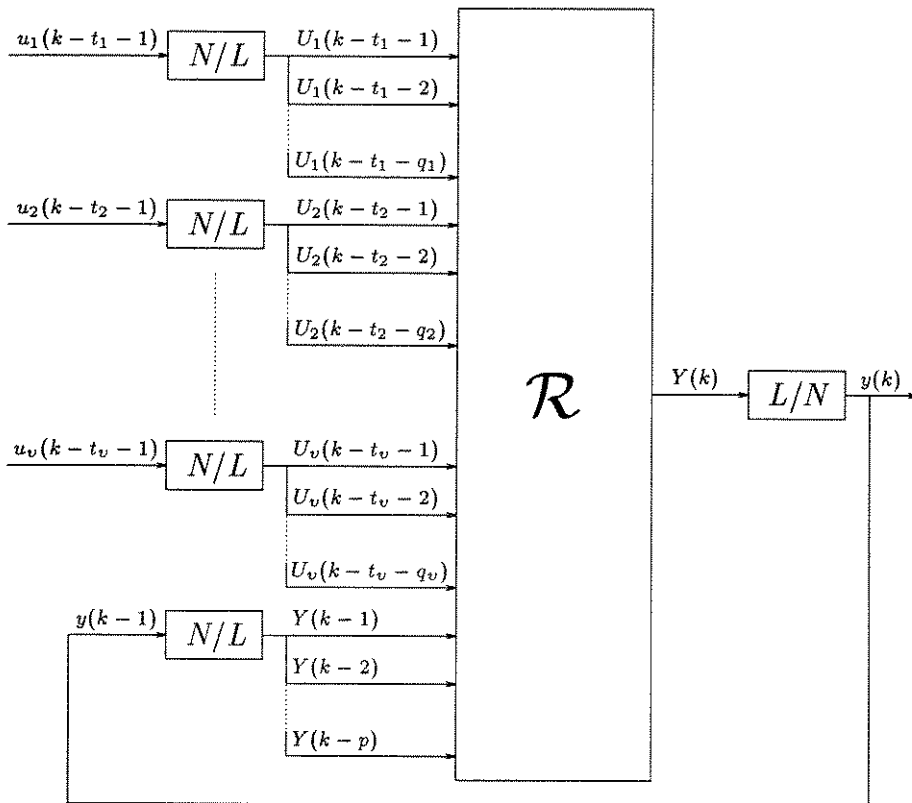


Figura 2.1: Diagrama de um modelo relacional nebuloso.

(Linguístico/Numérico) representam as interfaces de entrada (fuzzificação) e saída (defuzzificação), respectivamente, e o bloco  $\mathcal{R}$  representa a inferência da saída do modelo através da composição entre as variáveis de entrada e a

relação nebulosa (ver Seção 1.1). Pode-se observar, através da Figura (2.1), que os atrasos são considerados após os procedimentos de fuzzificação, acarretando a necessidade de apenas  $v + 1$  interfaces de entrada no modelo. Pode-se observar também que esse diagrama apresenta a topologia de realimentação numérica [Oliveira e Lemos, 1995], onde a variável de saída é realimentada após o procedimento de defuzzificação. O motivo da utilização dessa topologia e não da topologia de realimentação linguística (onde a variável de saída nebulosa é realimentada diretamente) é analisado no Capítulo 4.

Por simplicidade, a Equação (2.1) pode ser reescrita como

$$Y(k) = X_1(k) \bullet X_2(k) \bullet \cdots \bullet X_n(k) \bullet R \quad (2.2)$$

onde  $X_1(k) = Y(k-1)$ ,  $\dots$ ,  $X_p(k) = Y(k-p)$ ,  $X_{p+1}(k) = U_1(k-t_1-1)$ ,  $\dots$ ,  $X_{p+q_1}(k) = U_1(k-t_1-q_1)$ ,  $\dots$ ,  $X_n(k) = U_v(k-t_v-q_v)$ ,  $n = p+q_1+q_2+\dots+q_v$  e as variáveis nebulosas de entrada  $X_i(k)$  ( $i = 1, \dots, n$ ) e saída  $Y(k)$  são definidas sobre os universos de discurso  $\mathbf{X}_i$  e  $\mathbf{Y}$ , respectivamente. Como foi descrito no capítulo anterior, se esses universos são finitos (discretos e limitados) a relação nebulosa é representada por uma matriz relacional nebulosa definida no espaço dos Produtos Cartesianos ( $\times$ ) desses universos como segue:

$$R : \mathbf{X}_1 \times \mathbf{X}_2 \times \cdots \times \mathbf{X}_n \times \mathbf{Y} \rightarrow [0, 1]$$

Na seção seguinte será apresentada a abordagem de discretização nebulosa (*fuzzy discretization*) [Pedrycz, 1984a] que representa a relação nebulosa como uma matriz relacional independentemente da natureza finita ou não dos universos de discurso das variáveis do modelo. Essa abordagem apresenta diversas vantagens em relação à abordagem clássica do cálculo relacional, como será descrito a seguir, e será utilizada ao longo desse trabalho.

## 2.2 Discretização nebulosa

Seja um modelo relacional nebuloso descrito pela seguinte equação relacional com uma entrada e uma saída:

$$Y = U \bullet R$$

Sejam também  $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_n$  e  $\mathcal{Y}_1, \mathcal{Y}_2, \dots, \mathcal{Y}_\beta$  conjuntos nebulosos de referência definidos nos universos de discurso  $\mathbf{U}$  e  $\mathbf{Y}$ , respectivamente,

$$\begin{aligned} \mathcal{U}_j : \mathbf{U} &\rightarrow [0, 1], & j = 1, \dots, \eta \\ \mathcal{Y}_l : \mathbf{Y} &\rightarrow [0, 1], & l = 1, \dots, \beta \end{aligned}$$

de modo a satisfazer a condição de integralidade (*completeness*):

$$\begin{aligned} \forall u \in \mathbf{U} : \exists j : 1 \leq j \leq \eta : \mathcal{U}_j(u) > 0 \\ \forall y \in \mathbf{Y} : \exists l : 1 \leq l \leq \beta : \mathcal{Y}_l(y) > 0 \end{aligned}$$

De acordo com o conceito de discretização nebulosa, cada conjunto nebuloso  $U$  em  $\mathbf{U}$  e  $Y$  em  $\mathbf{Y}$  pode ser descrito como um vetor de possibilidades em relação aos conjuntos nebulosos de referência como segue:

$$\begin{aligned} Pu &= [Pu_1 \ Pu_2 \ \dots \ Pu_\eta] \\ Py &= [Py_1 \ Py_2 \ \dots \ Py_\beta] \end{aligned} \tag{2.3}$$

Cada elemento desses vetores representa uma medida de possibilidade [Zadeh, 1978] do conjunto nebuloso em relação a um conjunto nebuloso de referência específico, ou seja,

$$\begin{aligned} Pu_j &= \text{Poss}(U|\mathcal{U}_j) \triangleq \sup_{u \in \mathbf{U}} [U(u) \ t \ \mathcal{U}_j(u)], & j = 1, \dots, \eta \\ Py_l &= \text{Poss}(Y|\mathcal{Y}_l) \triangleq \sup_{y \in \mathbf{Y}} [Y(y) \ t \ \mathcal{Y}_l(y)], & l = 1, \dots, \beta \end{aligned} \tag{2.4}$$

onde  $t$  é uma norma triangular [Pedrycz, 1985b, Di Nola et al., 1991, Pedrycz, 1993] (ver Apêndice C). Nesse caso pode-se reescrever o modelo como

$$Py = Pu \bullet R' \tag{2.5}$$

onde  $R' : \mathbf{Pu} \times \mathbf{Py} \rightarrow [0, 1]$  é uma matriz relacional com  $\eta \times \beta$  elementos onde cada um dos elementos,  $r'_{ji}$ , representa uma implicação do tipo

SE entrada é  $\mathcal{U}_j$  ENTÃO saída é  $\mathcal{Y}_l$  com possibilidade  $r'_{jl}$  (2.6)

Nesse trabalho o Singleton nebuloso (Apêndice C) é utilizado no procedimento de fuzzificação. Nesse caso, um valor numérico  $u_0 \in \mathbf{U}$  é linguisticamente representado por

$$U(u) = \begin{cases} 1, & u = u_0 \\ 0, & \text{caso contrário} \end{cases} \quad (2.7)$$

Utilizando (2.7) a Equação (2.4) para  $Pu_j$  torna-se

$$Pu_j = \mathcal{U}_j(u_0), \quad j = 1, \dots, \eta \quad (2.8)$$

e substituindo (2.8) em (2.3) obtém-se

$$Pu = [\mathcal{U}_1(u_0) \ \mathcal{U}_2(u_0) \ \dots \ \mathcal{U}_\eta(u_0)] \quad (2.9)$$

que é a equação geral de fuzzificação para a associação de discretização nebulosa com o Singleton nebuloso. Um exemplo gráfico da geração do vetor de possibilidades  $Pu$  a partir da Equação (2.8) para um valor numérico  $u_0$  qualquer é apresentado na Figura (2.2). Para esse vetor  $Pu$ , um vetor  $Py$

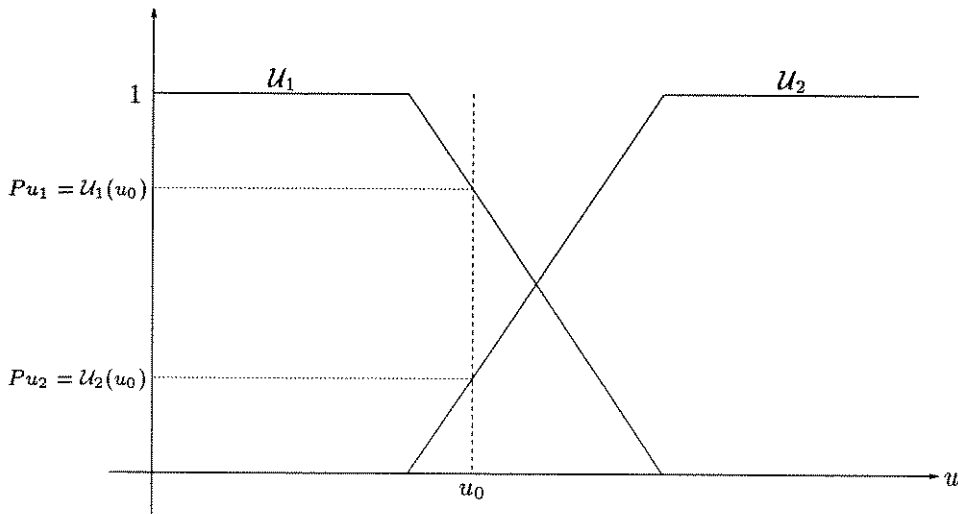


Figura 2.2: Geração de um vetor de possibilidades a partir da fuzzificação de um valor numérico utilizando discretização nebulosa.

hipotético pode ser obtido através da Equação (2.5). A Figura (2.3) apresenta um exemplo da interpretação geométrica de Pedrycz [Pedrycz, 1984a] para esse vetor de possibilidades de saída do modelo, onde  $Y^*$  é um conjunto nebuloso representativo, dado pela seguinte equação:

$$Y^*(y) = \max_{i=1}^{\beta} [\min(\mathcal{Y}_i(y), Py_i)]$$

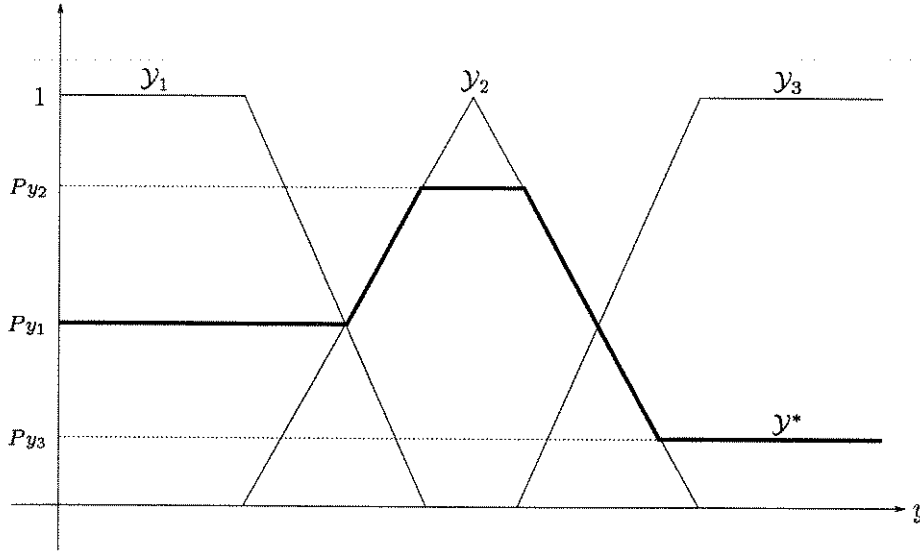


Figura 2.3: Interpretação geométrica para um vetor de possibilidades de saída do modelo.

Existem vantagens significativas na utilização da discretização nebulosa frente à abordagem clássica do cálculo relacional. A mais importante é a redução do esforço computacional (memória e processamento) em função da dimensão da matriz relacional. Como exemplo, pode-se tomar um modelo com uma variável de entrada e uma de saída, ambas definidas em universos de discurso limitados entre os valores  $-10$  e  $10$  e com uma discretização de  $0,1$ . Nesse caso, cada universo é constituído por um vetor de  $200$  elementos e a matriz relacional possui portanto  $40.000$  ( $200 \times 200$ ) elementos. Supondo  $5$  conjuntos nebulosos de referência em cada universo de discurso, através da discretização nebulosa obtém-se uma matriz com apenas  $25$  ( $5 \times 5$ ) elementos, ou seja,  $1.600$  vezes menor. Cada elemento da matriz também possui um significado claro dado pela implicação (2.6). Outra vantagem da discretização nebulosa é a possibilidade da utilização de universos de discurso discretos ou contínuos, com conjuntos nebulosos de referência definidos como vetores ou funções contínuas, uma vez que os vetores de possibilidades não possuem informação sobre esses universos.

## 2.3 Identificação de modelos relacionais nebulosos

Pedrycz propôs um algoritmo com os procedimentos básicos para identificação de modelos relacionais nebulosos [Pedrycz, 1984a]. Esse algoritmo consiste nos seguintes passos:

1. Coletar dados de entrada e saída do sistema e definir todos os universos de discurso necessários.
2. Construir os conjuntos nebulosos de referência.
3. Descrever os dados em termos dos conjuntos nebulosos de referência.
4. Fixar a estrutura do modelo.
5. Calcular a relação nebulosa do modelo.
6. Testar o modelo obtido através de dados coletados.

Nesse trabalho, um método baseado em otimização é adotado para construção dos conjuntos nebulosos de referência. Como é descrito no Capítulo 3, esse método utiliza um critério que minimiza os erros nos mapeamentos realizados pelas interfaces, melhorando o desempenho dos modelos. A descrição dos dados em termos desses conjuntos nebulosos de referência é realizada utilizando o conceito de discretização nebulosa em função das vantagens que essa abordagem apresenta e que foram discutidas na seção anterior. A determinação da estrutura do modelo não será abordada nesse trabalho. O cálculo da relação nebulosa em modelos de estrutura conhecida é descrito no Capítulo 4. Nesse capítulo propõe-se uma solução para o problema, introduzido na Seção 1.2.3, da dependência de boas estimativas da relação nebulosa na identificação de modelos relacionais. Os testes dos modelos obtidos através dos dados coletados são feitos no Capítulo 5 com previsões da saída um passo à frente e geração de séries sintéticas.

Os aspectos principais referentes aos modelos relacionais nebulosos e a identificação desses modelos são:

- **Dimensão:** Sendo a matriz relacional definida no Produto Cartesiano dos universos de discurso das variáveis do modelo, a dimensão dessa matriz é diretamente proporcional à discretização dos universos. Como foi

descrito na seção anterior, com a utilização da discretização nebulosa os universos de discurso são discretizados em função da quantidade de conjuntos nebulosos existentes nas interfaces. Nesse caso, a dimensão da matriz relacional depende não apenas da quantidade de variáveis nebulosas do modelo mas também da quantidade de conjuntos nebulosos associados a cada uma dessas variáveis. Como o esforço computacional de identificação e o tempo de convergência crescem com o aumento da dimensão da matriz, existe um compromisso entre essas grandezas e a precisão do modelo que está diretamente relacionada com o número de conjuntos nebulosos nas interfaces.

- **Tempo de convergência:** O tempo de convergência do modelo depende dos valores iniciais e da dimensão da matriz relacional. Inicializando adequadamente essa matriz, o tempo de convergência pode ser significativamente reduzido. Para essa finalidade pode-se obter uma matriz relacional aproximada através de uma pré-identificação utilizando, por exemplo, o algoritmo de aproximação linguística apresentado no Apêndice A.
- **Excitação:** Ao contrário dos modelos convencionais que representam relações no plano frequencial e devem ser excitados nas frequências do sistema, os modelos nebulosos constituem um mapeamento direto no espaço de amplitudes. Para que todas as relações do modelo sejam excitadas, é necessário gerar em amplitude valores que representam todas as possíveis combinações linguísticas de entrada e saída. A abordagem usual de utilizar excitações do tipo ruído branco permite gerar, após um período de tempo, as combinações possíveis de entrada. No entanto, as saídas em termos de amplitude dependem da constante de tempo do sistema. Assim, deve-se gerar os sinais de entrada de modo que a saída excursiona sobre todo seu universo.

## 2.4 Sumário

Modelos relacionais nebulosos constituem uma representação matemática de sistemas através de equações relacionais nebulosas. Quando implementados com base nos conceitos de discretização nebulosa e conjuntos nebulosos de referência, esses modelos podem ter suas dimensões significativamente reduzidas. Nesse caso, cada entrada e cada saída nebulosa é dada por um vetor

de possibilidades, onde cada elemento representa uma medida de possibilidade em relação a um determinado conjunto de referência, e cada elemento da relação nebulosa representa uma medida de possibilidade de uma implicação que relaciona conjuntos específicos de entrada e saída.

Nos modelos relacionais nebulosos existe um compromisso entre precisão, esforço computacional e tempo de convergência, porque todas essas grandezas dependem da estrutura do modelo e conseqüentemente da dimensão da matriz relacional. O tempo de convergência pode ser reduzido através da inicialização do procedimento de identificação utilizando uma matriz aproximada. A identificação deve ser realizada utilizando uma excitação de forma a ativar todas as relações do modelo, permitindo que o sistema responda adequadamente em amplitude.



# Capítulo 3

## Interfaces de Entrada e Saída

Esse capítulo discute os aspectos da construção de interfaces para sistemas nebulosos e apresenta o método de otimização que foi adotado nesse trabalho para otimizar o desempenho geral dos modelos relacionais.

### 3.1 Introdução

Em sistemas nebulosos, os valores assumidos pelas variáveis linguísticas e numéricas estão diretamente relacionados com a estrutura das interfaces de entrada e saída. Os seguintes aspectos devem ser abordados para uma adequada construção dessas interfaces:

- A integridade semântica dos conjuntos nebulosos como representação matemática de rótulos linguísticos que devem ser compreensíveis ao ser humano.
- A relação entre as formas linguística e numérica de informação.
- A adaptabilidade a mudanças no sistema e conseqüentemente nas variáveis envolvidas.
- A representação de incertezas no caso de tratamento de perturbações diretamente através das interfaces [Pedrycz, 1995].

O primeiro item é discutido na seção seguinte através da abordagem de uma série de aspectos que constituem o conceito de estrutura de cognição (*frame*

of cognition) [Pedrycz, 1995]. Os dois itens seguintes são abordados no algoritmo de otimização descrito na Seção 3.3.2. As incertezas não são tratadas diretamente nas interfaces mas explicitamente no algoritmo de identificação proposto no Capítulo 4 e também através da própria natureza da lógica e dos conjuntos nebulosos.

## 3.2 Estrutura de cognição

A estrutura de cognição é definida como uma coleção  $A$  de conjuntos nebulosos  $A_1, A_2, \dots, A_m$  definidos no mesmo universo de discurso  $A$  satisfazendo as condições essenciais para uma descrição semanticamente compreensível de um sistema como segue:

- **Distinção:** Os rótulos linguísticos devem ter um sentido semântico claro. Os conjuntos nebulosos correspondentes devem ser suficientemente disjuntos para representar claramente uma região específica do universo de discurso. A distinção em uma estrutura de cognição está relacionada com a especificidade dos conjuntos nebulosos que a compõe. Esta característica pode ser avaliada através de uma medida de especificidade (*specificity measure*) [Pedrycz, 1993]. Em uma mesma distribuição dos conjuntos nebulosos ao longo do universo de discurso, quanto mais específicos são esses conjuntos, mais disjunta é a estrutura de cognição.
- **Convexidade:** A convexidade é uma característica desejada uma vez que é bem conhecido que conjuntos nebulosos convexos [Zadeh, 1978] (Apêndice C) possuem uma interpretabilidade linguística maior que os não-convexos. Deve-se ressaltar que em alguns casos específicos essa característica pode não ser verificada, como no caso de conjuntos gerados a partir da aplicação, em outros conjuntos, do transformador linguístico NÃO (*Hedge NOT*) [Zadeh, 1973, Tanscheit, 1992].
- **Normalização:** Os conjuntos nebulosos devem ser normais (Apêndice C), ou seja, para cada conjunto, pelo menos um elemento do universo de discurso deve ter uma correspondência máxima com o rótulo linguístico associado a ele.
- **Número justificável de elementos:** O número de conjuntos nebulosos deve ser compatível com a capacidade de compreensão humana. Estudos

apontam os bem conhecidos valores de  $7 \pm 2$  conceitos linguísticos como um limite adequado.

- **Cobertura do universo de discurso:** Os universos de discurso devem estar “cobertos” por completo pelos conjuntos nebulosos definidos sobre ele. Formalmente, pode-se definir a cobertura do universo de discurso  $\mathbf{A}$  como

$$\forall a \in \mathbf{A} : \exists i : 1 \leq i \leq m : A_i(a) > 0 \quad (3.1)$$

Com isso, qualquer valor numérico do universo de discurso tem um grau de pertinência diferente de zero em pelo menos um conjunto nebuloso, ou seja, qualquer  $a \in \mathbf{A}$  possui uma correspondência não-nula com pelo menos um rótulo linguístico. Mais rigidamente pode-se redefinir (3.1) como

$$\forall a \in \mathbf{A} : \exists i : 1 \leq i \leq m : A_i(a) > \xi$$

onde  $\xi \in [0, 1)$  é definido como um nível de cobertura.

Embora existam diversos métodos para construção de interfaces de entrada e saída em sistemas nebulosos como por exemplo “clustering” e métodos subjetivos (heurísticos) [Pedrycz, 1984a, Pedrycz, 1993, Pedrycz, 1995, Kroll, 1996], nem todos eles consideram as características da estrutura de cognição detalhadas acima. A seção seguinte apresenta uma metodologia baseada em otimização que considera essas características. Através dessa metodologia, a estrutura de cognição é garantida matematicamente com uma implementação computacional simples. As interfaces são construídas de acordo com um critério que otimiza o desempenho dos sistemas nebulosos. Esse critério é de particular interesse para o método de identificação de modelos relacionais proposto nesse trabalho, como será apresentado no Capítulo 4.

### 3.3 Otimização de interfaces

Utiliza-se o algoritmo  $\Sigma$ -PAFIO (*Sigma-Count Penalty Adaptive Fuzzy Interface Optimizer*) proposto por Oliveira [Oliveira, 1995] para otimização de interfaces de entrada e saída em sistemas nebulosos. Ele foi desenvolvido

segundo o conceito de discretização nebulosa descrito na Seção 2.2 e baseia-se em um critério de equivalência de informações [Pedrycz e Oliveira, 1993, Pedrycz e Oliveira, 1996], descrito a seguir.

### 3.3.1 Critério de equivalência de informações

Seja  $\mathcal{L}$  o mapeamento realizado por uma interface de entrada ( $N/L$ ) com  $m$  conjuntos nebulosos de referência definidos em um universo de discurso  $\mathbf{A}$ , ou seja,

$$\mathcal{L} : \mathbf{A} \rightarrow [0, 1]^m$$

Analogamente, seja  $\mathcal{N}$  o mapeamento realizado por uma interface de saída ( $L/N$ ) com a mesma estrutura da interface de entrada definida acima, ou seja,

$$\mathcal{N} : [0, 1]^m \rightarrow \mathbf{A}$$

A otimalidade segundo o critério de equivalência de informações é definida como

$$\forall a \in \mathbf{A} \quad \mathcal{N}(\mathcal{L}(a)) = a \quad (3.2)$$

e pode ser observada na Figura (3.1) considerando  $\tilde{a} = a$ . Essa otimalidade

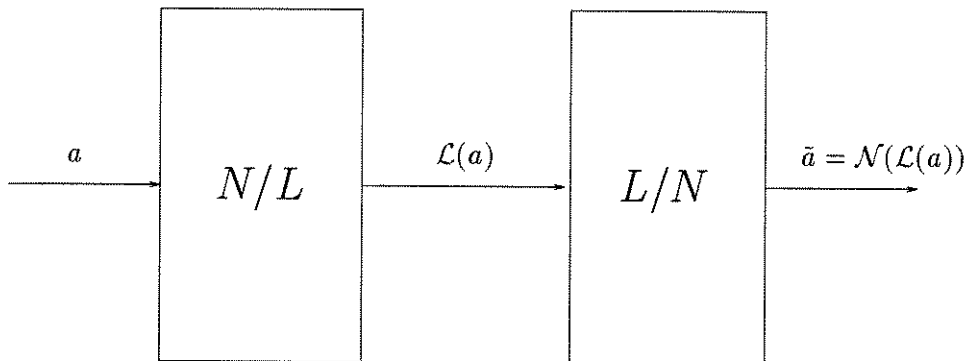


Figura 3.1: Sequência de fuzzificação e defuzzificação como representação do critério de equivalência de informações.

significa que a informação linguística é igualmente equivalente às informações

numéricas de entrada e saída. Logo, os erros inseridos em um sistema nebuloso através dos mapeamentos realizados nas interfaces são anulados, uma vez que a informação de entrada é completamente recuperada na saída após esses mapeamentos. Uma característica demonstrada por Oliveira [Oliveira, 1995] é que a verificação da otimalidade (3.2) garante a cobertura do universo de discurso definida em (3.1).

### 3.3.2 Algoritmo $\Sigma$ -PAFIO

O algoritmo  $\Sigma$ -PAFIO é uma extensão dos algoritmos FIO (*Fuzzy Interface Optimizer*) e PAFIO (*Penalty Adaptive Fuzzy Interface Optimizer*) [Oliveira, 1993] também baseados no critério de equivalência de informações. Ele considera explicitamente a característica de distinção entre conjuntos nebulosos de referência apresentada na Seção 3.3, o que o torna mais genérico. Esse algoritmo realiza a otimização paramétrica das interfaces de um sistema nebuloso utilizando funções de pertinência parametrizadas para a definição dos conjuntos nebulosos nessas interfaces. Por simplicidade, nesse trabalho são utilizadas apenas funções Gaussianas descritas como segue:

$$f_i(a, \theta_i) = \exp\left(\frac{-(a - \theta_{1_i})^2}{(\theta_{2_i})^2}\right) \quad (3.3)$$

onde  $a \in \mathbf{A}$ ,  $\theta_i = [\theta_{1_i}, \theta_{2_i}]$  com  $\theta_{1_i}$  e  $\theta_{2_i}$  sendo respectivamente o centro (que nesse caso coincide com o valor modal definido no Apêndice C) e a abertura do  $i$ -ésimo conjunto nebuloso de referência com função de pertinência  $f_i$ . A utilização de funções Gaussianas é atraente não apenas por elas serem parametrizadas em centro e abertura, mas também por garantirem a cobertura do universo de discurso definida em (3.1) devido à sua característica assintótica. Essas funções definem conjuntos nebulosos convexos e através de um ganho de amplitude unitário também definem conjuntos nebulosos normais. Portanto, a maior parte das especificações da estrutura de cognição são automaticamente satisfeitas.

A distinção entre os conjuntos nebulosos de referência é equacionada a partir da medida “sigma-count” definida como segue:

$$M_\rho(\mathcal{L}(a)) = \sqrt[\rho]{\sum_{i=1}^m f_i^\rho(a, \theta_i)} \quad (3.4)$$

onde  $\mathcal{L}(a) = [f_1(a, \theta_1) \cdots f_m(a, \theta_m)]$  é um conjunto nebuloso (vetor de possibilidades) gerado pelo mapeamento de entrada  $\mathcal{L}$  de um valor  $a \in \mathbf{A}$  (Seção 3.3.1),  $f_i(a, \theta_i)$  é o valor da função de pertinência, no ponto  $a$ , do  $i$ -ésimo conjunto nebuloso de referência com centro e abertura dados por  $\theta_i$  e  $\rho$  é um escalar positivo denominado norma para medida “sigma-count”. Em seus trabalhos, Oliveira inicialmente relacionou a medida “sigma-count” para  $\rho = 1$  com a distinção entre conjuntos nebulosos de referência em uma interface. Nesse caso, uma distribuição uniforme dos conjuntos ao longo do universo de discurso, satisfazendo as características de cobertura e distinção da estrutura de cognição (e.g. ver Figura (3.2-a)), representa uma boa aproximação em relação ao critério de equivalência de informações e, em geral, satisfaz a seguinte condição de desigualdade [Oliveira e Lemos, 1995, Oliveira, 1995]:

$$\forall a \in \mathbf{A} \quad M_1(\mathcal{L}(a)) \leq 1 \quad (3.5)$$

Para a interface ilustrada na Figura (3.2-b), que não apresenta adequada distinção entre os conjuntos nebulosos, essa condição não é satisfeita.

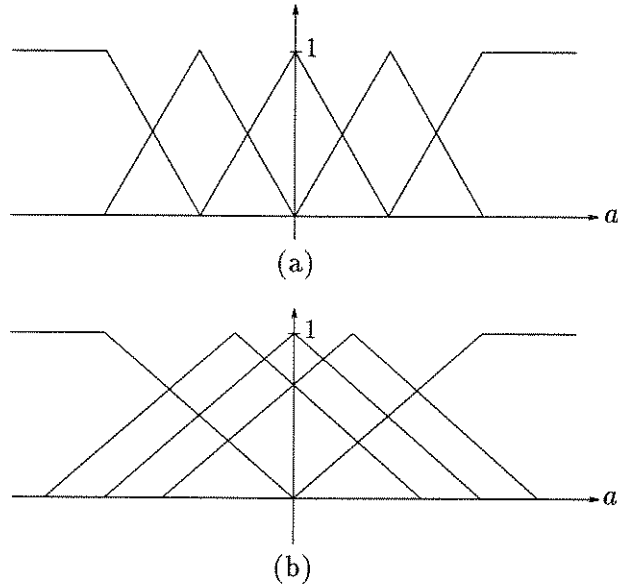


Figura 3.2: Interfaces com (a) conjuntos nebulosos distintos e homoganeamente distribuídos e (b) conjuntos nebulosos pouco distintos e não-homoganeamente distribuídos.

Genericamente pode-se reescrever (3.5) como

$$\forall a \in \mathbf{A} \quad M_\rho(\mathcal{L}(a)) \leq 1 \quad (3.6)$$

Utilizando valores de  $\rho$  maiores que 1 a condição de desigualdade (3.6) torna-se menos rígida que a condição (3.5), permitindo um grau de liberdade a mais em relação à especificação para otimização das interfaces.

Objetivando a otimalidade segundo o critério de equivalência de informações dado por (3.2), o algoritmo  $\Sigma$ -PAFIO baseia-se na minimização, sobre  $\theta_i$  ( $i = 1, \dots, m$ ) definido em (3.3), da seguinte função de custo

$$J = \frac{1}{2} \sum_{j=1}^N (a_j - \tilde{a}_j)^2 \quad (3.7)$$

sujeita à restrição dada por (3.6), onde  $a_j$  ( $j = 1, \dots, N$ ) é o  $j$ -ésimo valor de entrada na interface a ser otimizada e  $\tilde{a}_j = \mathcal{N}(\mathcal{L}(a_j))$  é o valor correspondente obtido após os mapeamentos de entrada e saída (sequência de fuzzificação e defuzzificação) ilustrados na Figura (3.1). Como os valores  $a_j$  ( $j = 1, \dots, N$ ) não necessariamente representam todo o universo de discurso  $\mathbf{A}$ , a restrição (3.6) é considerada apenas para os valores disponíveis e o problema de otimização torna-se

$$\begin{aligned} \min_{\theta_1, \dots, \theta_m} \quad & J = \frac{1}{2} \sum_{j=1}^N (a_j - \tilde{a}_j)^2 \\ \text{sujeito a} \quad & M_\rho(\mathcal{L}(a_1)) \leq 1 \\ & \vdots \\ & M_\rho(\mathcal{L}(a_N)) \leq 1 \end{aligned} \quad (3.8)$$

Esse problema é resolvido tornando-o irrestrito através do método da função de penalidade. Esse método permite considerar a minimização de uma função de custo  $P(\Omega)$  sujeita a uma restrição genérica  $g(\Omega) \leq 0$  no domínio  $\Omega$ , como a minimização sem restrições de uma nova função de custo  $\tilde{P}(\Omega)$  dada por

$$\tilde{P}(\Omega) = P(\Omega) + K g^2(\Omega) \mathcal{F}(g(\Omega))$$

onde  $K$  é um escalar positivo denominado constante de distinção e  $\mathcal{F}$  é uma função definida como

$$\mathcal{F}(x) = \begin{cases} 1, & \text{se } x > 0 \\ 0, & \text{se } x \leq 0 \end{cases}$$

Através da função de penalidade o Problema (3.8) pode ser reescrito como

$$\min_{\theta_1, \dots, \theta_m} \tilde{J} = J + J_1 \quad (3.9)$$

onde  $J$  é dado pela Equação (3.7) e  $J_1$  é dado por

$$J_1 = \frac{K}{2} \sum_{j=1}^N (M_\rho(\mathcal{L}(a_j)) - 1)^2 \mathcal{F}(M_\rho(\mathcal{L}(a_j)) - 1)$$

onde cada parcela da somatória refere-se a cada uma das  $N$  restrições em (3.8).

Da solução de (3.9) através do gradiente da função de custo  $\tilde{J}$  obtém-se, para  $i = 1, \dots, m$ ,

$$\theta_i(k+1) = \theta_i(k) - \eta(k) \frac{\partial \tilde{J}}{\partial \theta_i}$$

onde  $k$  refere-se à iteração e  $\eta(\cdot) \in (0, 1)$  é a taxa de aprendizado. O gradiente de  $\tilde{J}$  é dado por

$$\frac{\partial \tilde{J}}{\partial \theta_i} = \frac{\partial J}{\partial \theta_i} + \frac{\partial J_1}{\partial \theta_i} \quad (3.10)$$

Desenvolvendo a primeira parcela de (3.10) obtém-se

$$\frac{\partial J}{\partial \theta_i} = - \sum_{j=1}^N (a_j - \tilde{a}_j) \frac{\partial \tilde{a}_j}{\partial \theta_i}$$

onde  $\frac{\partial \tilde{a}_j}{\partial \theta_i} = \left[ \frac{\partial \tilde{a}_j}{\partial \theta_{1i}} \quad \frac{\partial \tilde{a}_j}{\partial \theta_{2i}} \right]$ . A relação entre  $\tilde{a}_j$  e  $\theta_i$  é determinada através do mapeamento de saída  $\mathcal{N}$ , ou mais precisamente através do método de defuzzificação associado a ele. Nesse trabalho utiliza-se o método do centro de gravidade devido à sua eficiência e facilidade de implementação aliadas à inexistência de variações abruptas na saída [Pedrycz, 1995]. A equação desse método é dada por



$$\tilde{a}_j = \frac{\sum_{i=1}^m f_i(a_j, \theta_i) \theta_{1i}}{\sum_{i=1}^m f_i(a_j, \theta_i)} \quad (3.11)$$

A partir de (3.11) e após algumas manipulações algébricas obtém-se

$$\frac{\partial \tilde{a}_j}{\partial \theta_{1i}} = \frac{f_i(a_j, \theta_i) + (\theta_{1i} - \tilde{a}_j) \frac{\partial f_i(a_j, \theta_i)}{\partial \theta_{1i}}}{\sum_{l=1}^m f_l(a_j, \theta_l)}$$

e

$$\frac{\partial \tilde{a}_j}{\partial \theta_{2i}} = \frac{(\theta_{1i} - \tilde{a}_j) \frac{\partial f_i(a_j, \theta_i)}{\partial \theta_{2i}}}{\sum_{l=1}^m f_l(a_j, \theta_l)}$$

Considerando  $M_\rho(\mathcal{L}(a_j)) > 1$ , ou seja,  $M_\rho(\mathcal{L}(a_j)) - 1 > 0$  que implica em  $\mathcal{F}(M_\rho(\mathcal{L}(a_j)) - 1) = 1$ , a segunda parcela da Equação (3.10) pode ser desenvolvida resultando em

$$\frac{\partial J_1}{\partial \theta_i} = K \sum_{j=1}^N (M_\rho(\mathcal{L}(a_j)) - 1) \left( \sum_{l=1}^m f_l^\rho(a_j, \theta_l) \right)^{\frac{1}{\rho}-1} f_i^{\rho-1}(a_j, \theta_i) \frac{\partial f_i(a_j, \theta_i)}{\partial \theta_i}$$

onde  $\frac{\partial f_i(a_j, \theta_i)}{\partial \theta_i} = \left[ \frac{\partial f_i(a_j, \theta_i)}{\partial \theta_{1i}} \frac{\partial f_i(a_j, \theta_i)}{\partial \theta_{2i}} \right]$ . Por outro lado, se  $M_\rho(\mathcal{L}(a_j)) \leq 1$  tem-se  $\mathcal{F}(M_\rho(\mathcal{L}(a_j)) - 1) = 0$  e portanto  $\frac{\partial J_1}{\partial \theta_i} = 0$ . Assim, para o caso geral tem-se

$$\frac{\partial J_1}{\partial \theta_i} = K \sum_{j=1}^N \mathcal{F}(M_\rho(\mathcal{L}(a_j)) - 1) (M_\rho(\mathcal{L}(a_j)) - 1) \left( \sum_{l=1}^m f_l^\rho(a_j, \theta_l) \right)^{\frac{1}{\rho}-1} \cdot f_i^{\rho-1}(a_j, \theta_i) \frac{\partial f_i(a_j, \theta_i)}{\partial \theta_i}$$

onde o termo  $M_\rho(\mathcal{L}(a_j))$  é calculado a partir da Equação (3.4).

Para evitar oscilações em torno do ponto de ótimo, propõe-se nesse trabalho a utilização de um critério de parada para o procedimento de otimização. Esse critério é apresentado no Apêndice B, junto a uma versão do algoritmo  $\Sigma$ -PAFIO para aplicações “on-line”, que é atraente por permitir a adaptação das interfaces a mudanças nos sinais de entrada e saída dos sistemas nebulosos. Com a utilização do critério de parada, o algoritmo é interrompido se o resultado de uma filtragem dos erros de otimização é inferior a um valor preestabelecido, denominado nível de desligamento.

O processo de otimização de interfaces descrito nessa seção é analisado brevemente uma vez que o objetivo desse trabalho não envolve a apresentação de novas propostas para esse problema. Inicialmente, deve-se enfatizar que o Problema (3.8) não é convexo. Em função disso, alguns cuidados devem ser adotados ao se utilizar a função de penalidade e o método do gradiente. Deve-se inicializar a otimização em um ponto da região factível, para que, com  $K$  suficientemente grande, o algoritmo tenda a manter a solução dentro dessa região. Deve-se também inicializar o algoritmo próximo a um ponto de solução ótima, para que, mesmo dentro da região factível, ótimos locais indesejados possam ser evitados. Como foi afirmado anteriormente, esses dois requerimentos são em geral satisfeitos heurísticamente através da distribuição uniforme dos conjuntos nebulosos de referência sugerida por Oliveira [Oliveira e Lemos, 1995, Oliveira, 1995]. Por fim, também em função da não-convexidade da função de custo  $\tilde{J}$ , uma taxa de aprendizado  $\eta$  relativamente baixa torna-se adequada, embora valores mais altos não sejam necessariamente críticos como pode ser visto no exemplo da seção seguinte.

### 3.4 Exemplo numérico

O exemplo a seguir apresenta uma aplicação do algoritmo  $\Sigma$ -PAFIO e ilustra os principais aspectos da otimização de interfaces discutidos nesse capítulo, como a minimização dos erros nos mapeamentos de entrada e saída e a capacidade de adaptação dessas interfaces.

Seja o processo com uma entrada  $u$  e uma saída  $y$  descrito pela seguinte equação a diferenças:

$$y(k) = 0,7y(k-1) + 0,6u(k-1)$$

Seja também um modelo nebuloso desse processo com as variáveis  $Y(k-1)$ ,  $U(k-1)$  e  $Y(k)$  respectivamente como representações linguísticas de  $y(k-1)$ ,

$u(k-1)$  e  $y(k)$  (ver Equação (2.1)). Como as saídas  $y(k-1)$  e  $y(k)$  representam a mesma grandeza defasada no tempo, pode-se utilizar a mesma estrutura de interface para ambas.

Utiliza-se como entrada  $u$  uma sequência uniformemente distribuída entre os valores  $-5$  e  $5$  com 300 amostras. Essa entrada e a saída  $y$  do processo são apresentadas na Figura (3.3).

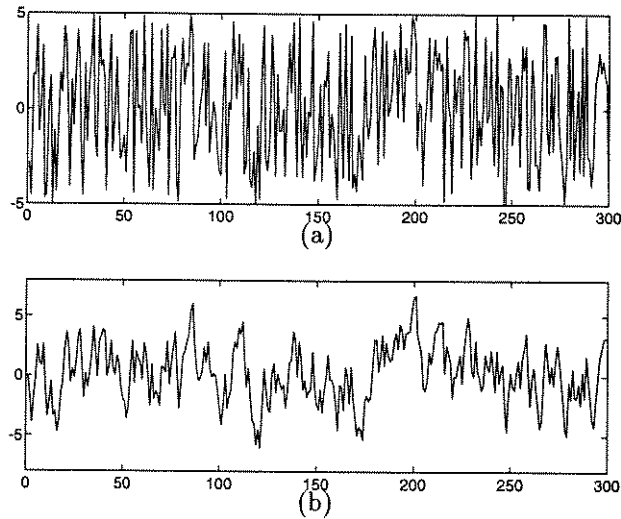


Figura 3.3: Sequência dos sinais (a) de entrada  $u$  e (b) de saída  $y$  para otimização das interfaces do modelo.

Durante os 300 instantes relativos a cada uma das amostras, a versão “on-line” do algoritmo  $\Sigma$ -PAFIO apresentada no Apêndice B é executada para ambas as interfaces de entrada e saída com 5 conjuntos nebulosos de referência, taxa de aprendizado  $\eta = 0,8$ , constante de distinção  $K = 5$ , norma para medida “sigma-count”  $\rho = 4$ , nível de desligamento do algoritmo  $\lambda = 0,02$ , janela da sequência de erros  $H = 10$  e pólo do filtro  $\gamma = 0,9$ . Os parâmetros  $\lambda$ ,  $H$  e  $\gamma$  estão relacionados ao critério de parada do algoritmo definido no Apêndice B. Os valores utilizados para  $\gamma$ ,  $H$  e  $\rho$  são os sugeridos neste Apêndice. O valor de  $\lambda$  corresponde ao erro médio aceitável para otimização, nesse caso escolhido como aproximadamente 0,4% dos universos de discurso, que foram tomados como o intervalo entre os valores mínimo e máximo de excursão dos sinais de entrada e saída apresentados na Figura (3.3). Selecionou-se um va-

lor alto para  $\eta$ , para permitir uma convergência rápida das interfaces, e um valor para  $K$  suficientemente grande (em relação aos valores assumidos pela função de custo) para garantir a distinção entre os conjuntos nebulosos de referência. Esses conjuntos foram inicialmente distribuídos uniformemente nos universos de discurso e, posteriormente, os conjuntos extremos tiveram seus centros e aberturas acrescidos de 10% em função da bem conhecida dificuldade do método do centro de gravidade em defuzzificar valores extremos dos universos [Tanscheit, 1992].

A Figura (3.4) apresenta o comportamento do erro com relação ao critério de equivalência de informações ao longo dos 300 instantes de otimização para as interfaces (a) de entrada e (b) de saída do modelo. Deve-se enfatizar dois aspectos que podem ser notados através dessa figura. Primeiro, os erros são reduzidos a níveis quase nulos como já era esperado em função da otimização. Segundo, os erros iniciais são pequenos (relativamente às dimensões dos universos de discurso), demonstrando que a distribuição uniforme inicial dos conjuntos nebulosos é realmente adequada como foi descrito na seção anterior.

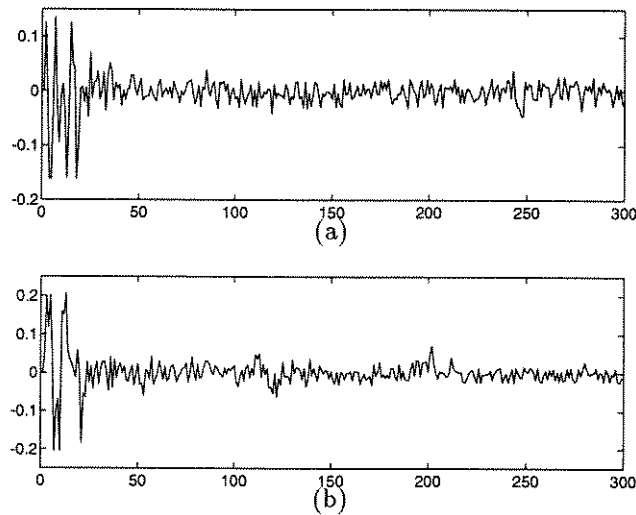


Figura 3.4: Sequência de erros ( $a - \tilde{a}$ ) para as interfaces (a) de entrada e (b) de saída do modelo.

O desempenho do algoritmo  $\Sigma$ -PAFIO pode ser mais bem observado na Figura (3.5) que apresenta o comportamento da função de custo durante a

otimização. Nota-se nessa figura que essa função é praticamente zerada, o que representa o objetivo do procedimento de otimização levando em conta a existência do critério de parada imposto ao algoritmo. Com relação a esse critério, a filtragem  $E$  da sequência de erros (Apêndice B), que é comparada ao nível de desligamento  $\lambda$ , é mostrada na Figura (3.6).

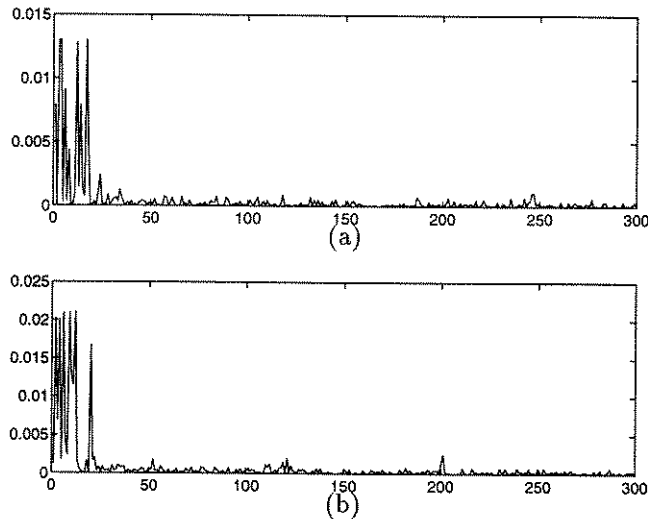


Figura 3.5: Função de custo  $\tilde{J}$  para as interfaces (a) de entrada e (b) de saída do modelo.

As Figuras (3.7) e (3.8) ilustram respectivamente a evolução dos centros e aberturas dos conjuntos nebulosos de referência para a interface de entrada do modelo ao longo do procedimento de otimização. O mesmo é feito para a interface de saída através das Figuras (3.9) e (3.10). Pode-se notar nessas figuras que os parâmetros sofrem variações muito pequenas, o que confirma mais uma vez a adequação da distribuição inicial dos conjuntos que foi utilizada. Pode-se notar também uma característica convergente desses parâmetros até níveis constantes serem obtidos durante os períodos de desligamento do algoritmo, ou seja, nos instantes de tempo tais que  $E < \lambda$ .

As estruturas iniciais ( $k = 1$ ) e finais ( $k = 300$ ) da interface de entrada do modelo são apresentadas na Figura (3.11). O mesmo é feito na Figura (3.12) para a interface de saída do modelo. Nota-se que a variação entre as interfaces iniciais e finais é praticamente imperceptível em função da variação pequena

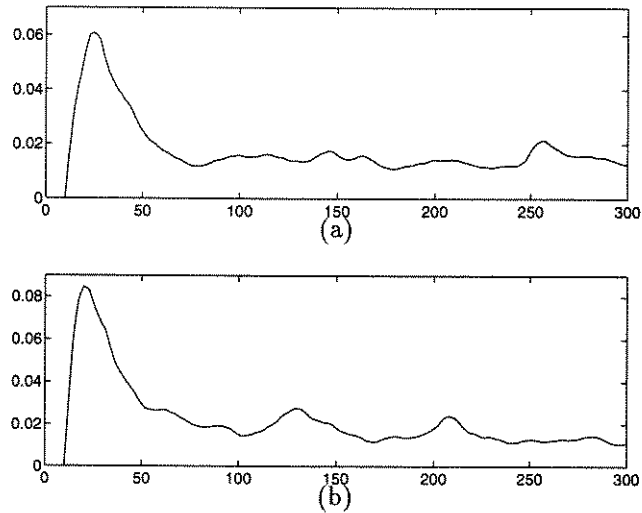


Figura 3.6: Filtragem  $E$  da seqüência de erros para as interfaces (a) de entrada e (b) de saída do modelo.

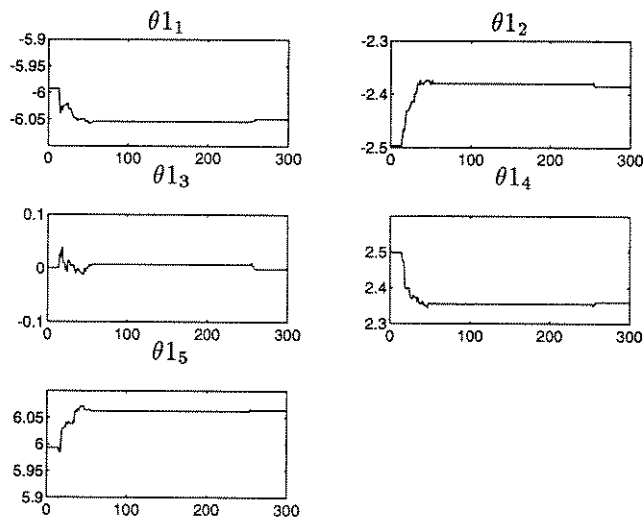


Figura 3.7: Evolução dos centros  $\theta_{1_i}$  dos conjuntos nebulosos de referência para a interface de entrada do modelo.

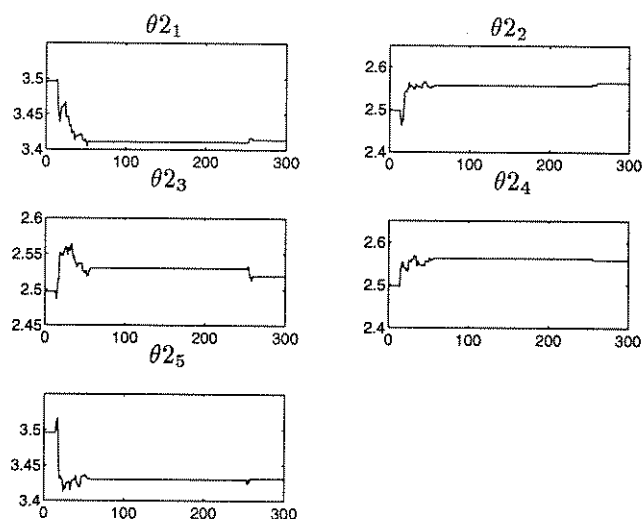


Figura 3.8: Evolução das aberturas  $\theta_{2_i}$  dos conjuntos nebulosos de referência para a interface de entrada do modelo.

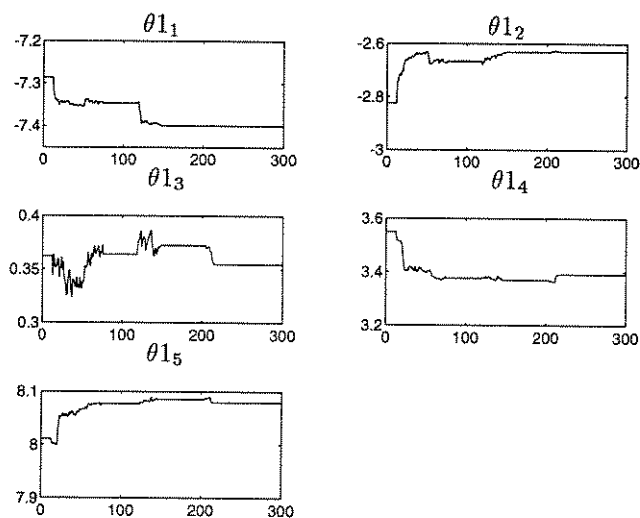


Figura 3.9: Evolução dos centros  $\theta_{1_i}$  dos conjuntos nebulosos de referência para a interface de saída do modelo.

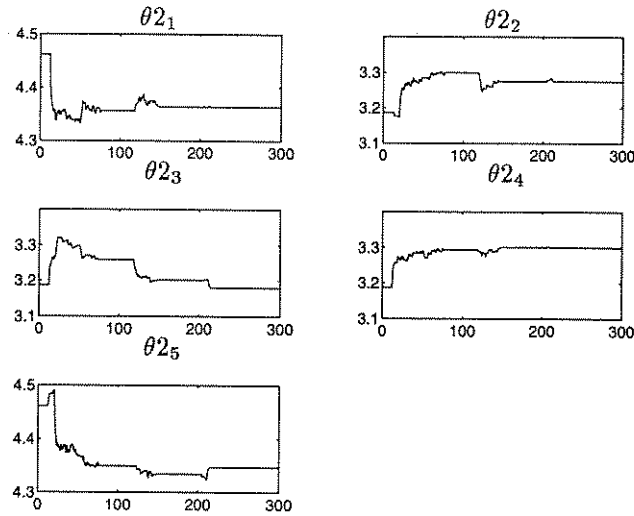


Figura 3.10: Evolução das aberturas  $\theta_{2i}$  dos conjuntos nebulosos de referência para a interface de saída do modelo.

ocorrida nos parâmetros.

Para ilustrar a capacidade de adaptação das interfaces com a utilização do algoritmo  $\Sigma$ -PAFIO, uma segunda otimização com 400 amostras dos sinais de entrada  $u$  e saída  $y$  foi realizada com os conjuntos nebulosos iniciais concentrados em uma pequena região em torno dos centros dos universos de discurso. Os erros relativos ao critério de equivalência de informações são apresentados na Figura (3.13). Nota-se que os erros iniciais são consideravelmente maiores do que aqueles apresentados na Figura (3.4), correspondente a uma distribuição inicial uniforme dos conjuntos nebulosos nas interfaces. Os picos obtidos na sequência de erros da interface de saída são devido aos picos no sinal de saída  $y$  (eventos raros) que chegam a valores do universo de discurso até então ainda não alcançados, e portanto não cobertos pelos conjuntos nebulosos de referência. Esses picos e os erros consequentes causam a expansão dos conjuntos nebulosos ao longo dos universos de discurso. Esse efeito, que demonstra a capacidade de adaptação das interfaces, pode ser observado nas Figuras (3.14) e (3.15) que mostram as estruturas iniciais ( $k = 1$ ) e finais ( $k = 400$ ) das interfaces de entrada e saída do modelo, respectivamente.



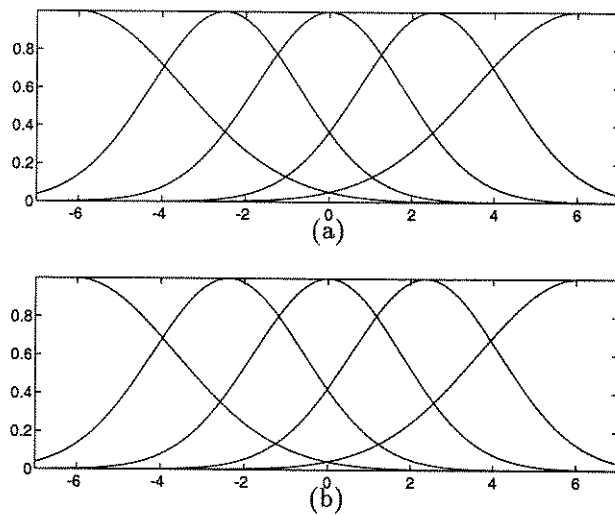


Figura 3.11: Estruturas (a) inicial e (b) final da interface de entrada do modelo.

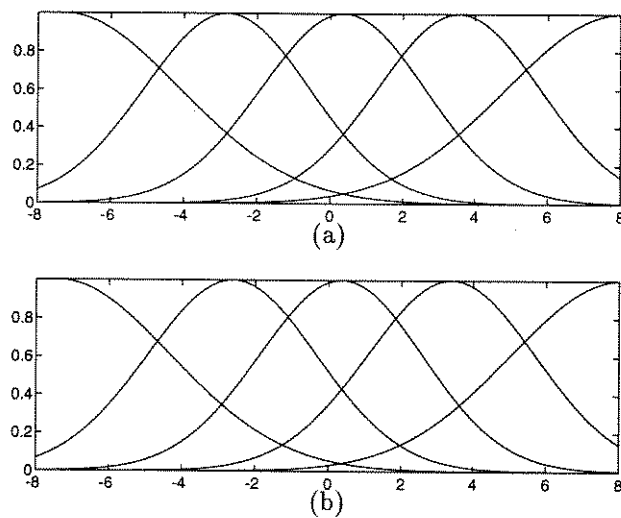


Figura 3.12: Estruturas (a) inicial e (b) final da interface de saída do modelo.

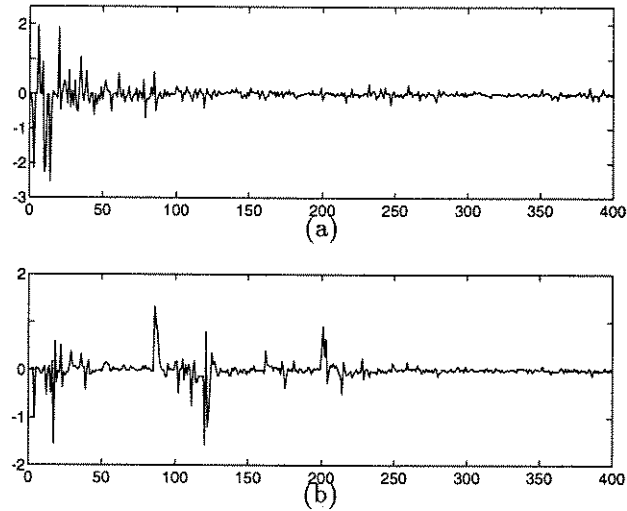


Figura 3.13: Sequência de erros ( $a - \tilde{a}$ ) para as interfaces (a) de entrada e (b) de saída do modelo.

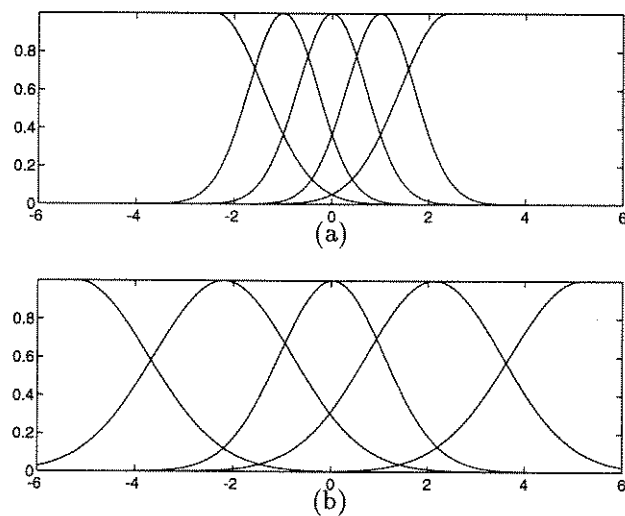


Figura 3.14: Estruturas (a) inicial e (b) final da interface de entrada do modelo.

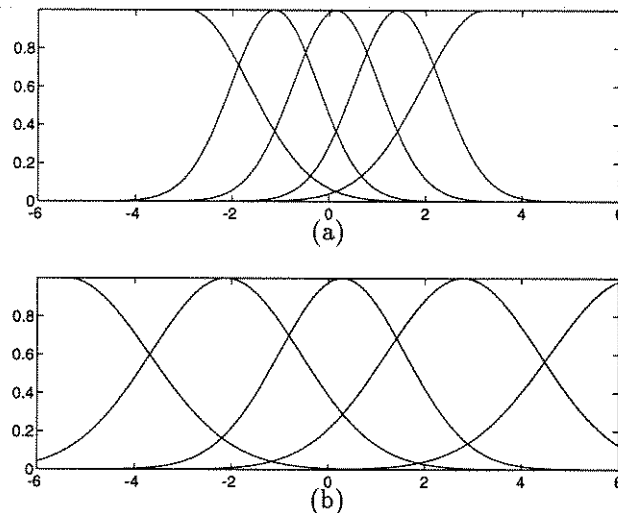


Figura 3.15: Estruturas (a) inicial e (b) final da interface de saída do modelo.

### 3.5 Sumário

As interfaces de entrada e saída tem uma influência significativa nos sistemas nebulosos uma vez que esses sistemas constituem um mapeamento entre conjuntos nebulosos de entrada e saída. A construção dessas interfaces deve ser realizada de forma a preservar a integridade semântica dos conjuntos nebulosos. Portanto, o método adotado deve satisfazer as condições da estrutura de cognição que garantem uma descrição semanticamente compreensível das relações entre esses conjuntos. O algoritmo  $\Sigma$ -PAFIO satisfaz essas condições otimizando parametricamente as interfaces segundo o critério de equivalência de informações, cuja otimalidade implica o cancelamento dos erros nos mapeamentos de fuzzificação e defuzzificação. Além dessa otimalidade, a versão "on-line" do algoritmo  $\Sigma$ -PAFIO provê uma característica adaptativa às interfaces, permitindo que os conjuntos nebulosos adaptem-se à mudanças nos sinais do sistema. Contudo, a minimização do critério de equivalência de informações constitui um problema não-convexo. Portanto, uma escolha criteriosa dos parâmetros do algoritmo torna-se necessária, bem como uma condição inicial aproximada, que pode ser obtida através de uma distribuição uniforme dos conjuntos nebulosos sobre os universos de discurso.

## Capítulo 4

# Algoritmo de Identificação Adaptativa

Esse capítulo apresenta um novo algoritmo de identificação adaptativa de sistemas utilizando modelos relacionais nebulosos [Campello et al., 1997].

### 4.1 Introdução

Como foi descrito na Seção 1.2.3, na literatura foram propostos diferentes algoritmos adaptativos para refinar modelos relacionais obtidos através de métodos aproximados. O principal problema desses algoritmos é a dependência com uma estimativa adequada da matriz relacional, que se torna crítica quando o sistema a ser identificado é variante no tempo. Nesses casos, uma matriz pré-determinada pode não ser adequada e o desempenho do algoritmo pode ser degradado.

Nesse trabalho, propõe-se uma heurística para resolver o problema de identificação quando a matriz relacional está polarizada, ou seja, quando essa matriz não representa uma estimativa adequada da relação nebulosa que se deseja determinar. Essa heurística, aliada à capacidade de adaptação das interfaces de entrada e saída, torna o algoritmo proposto adequado a aplicações “on-line” e de particular interesse na identificação de sistemas variantes no tempo. O algoritmo é obtido otimizando um critério baseado nas saídas nebulosas do modelo e do sistema, ao invés dos critérios tradicionais baseados nos valores não-nebulosos dessas saídas. Esse critério permite o tratamento de informações linguísticas e numéricas e a utilização de um método de oti-

mização simples. Como a saída nebulosa do modelo independe do método de defuzzificação adotado, evita-se as eventuais não-linearidades e/ou descontinuidades desse método.

## 4.2 O modelo nebuloso

A Figura (4.1) apresenta um esquema básico para identificação de um sistema através de um modelo relacional nebuloso. Por simplicidade, utiliza-se um sistema monovariável de primeira ordem. Nessa figura,  $x_1$  é a entrada do sistema que também excita o modelo,  $y$  é a saída medida do sistema,  $\tilde{y}$  é

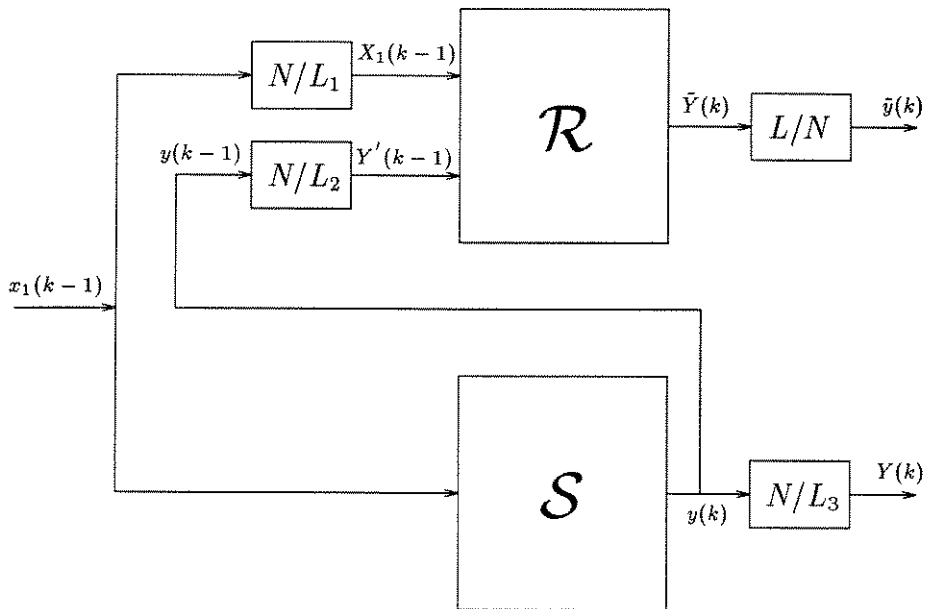


Figura 4.1: Esquema de identificação de um sistema ( $S$ ) utilizando um modelo relacional nebuloso.

a saída estimada do modelo e  $Y/Y'$  e  $\tilde{Y}$  são as representações linguísticas de  $y$  e  $\tilde{y}$ , respectivamente. Para o caso descrito na Seção 2.1 que considera sistemas MISO, pode-se reescrever para a saída  $\tilde{Y}$  a equação relacional (2.2) que descreve o modelo como

$$\tilde{Y}(k) = X_1(k) \bullet X_2(k) \bullet \dots \bullet X_n(k) \bullet R \quad (4.1)$$

Conforme definido naquela seção,  $X_i$  para  $i = 1, \dots, n$  são as  $n$  entradas nebulosas do modelo que pode ser considerado, sem perda de generalidade, como tendo uma interface de fuzzificação para o mapeamento de cada uma dessas entradas. Nesse caso, considerando a existência de um número  $c_i$  de conjuntos nebulosos de referência na  $i$ -ésima interface de entrada e um número  $c_0$  na interface de saída, de acordo com a discretização nebulosa, a cada conjunto de dados numéricos de entrada e saída  $x_i \in \mathbf{X}_i$  e  $\tilde{y} \in \mathbf{Y}$  associam-se as respectivas representações linguísticas  $X_i \in [0, 1]^{c_i}$  e  $\tilde{Y} \in [0, 1]^{c_0}$ .

### 4.3 O problema de otimização

Baseando-se no exemplo da Figura (4.1), se as estruturas das interfaces de fuzzificação e defuzzificação utilizadas nas saídas do sistema e do modelo ( $N/L_3$  e  $L/N$ , respectivamente) são iguais, tem-se que  $Y \in [0, 1]^{c_0}$ . Também tem-se que, de acordo com o critério de otimização de interfaces do algoritmo  $\Sigma$ -PAFIO (Seção 3.3.1), a igualdade  $\tilde{Y} = Y$  implica que  $\tilde{y} = y$  se a função de custo desse algoritmo é nula, ou que  $\tilde{y} \approx y$  se essa função é apenas aproximadamente nula. Portanto, uma aproximação para a modelagem nebulosa poderia ser obtida minimizando, em cada instante de amostragem, o seguinte conjunto de funções de custo:

$$J_j = \frac{1}{2}(Y_j - \tilde{Y}_j)^2, \quad j = 1, \dots, c_0 \quad (4.2)$$

onde  $\tilde{Y}_j$  e  $Y_j$  são os  $j$ -ésimos elementos de  $\tilde{Y}$  e  $Y$ , respectivamente. No entanto, de acordo com a Equação (4.1) para a composição nebulosa max-t, que será utilizada nesse trabalho (Apêndice C), para os elementos de  $R$  no intervalo  $[0, 1]$  a saída  $\tilde{Y}$  é limitada superiormente pelas entradas  $X_1, \dots, X_n$ . Essa limitação pode impedir que as saídas nebulosas do modelo e do sistema se tornem iguais, ou seja, que o conjunto de funções (4.2) se torne nulo. Para solucionar esse problema, utiliza-se um ganho  $G \in (0, 1)$  na saída nebulosa do sistema para permitir numericamente que as funções  $J_j$  ( $j = 1, \dots, c_0$ ) sejam iguais a zero para qualquer combinação possível de entradas não-nebulosas  $x_i \in \mathbf{X}_i$  ( $i = 1, \dots, n$ ). Assim reescreve-se (4.2) como

$$J_j = \frac{1}{2}(GY_j - \tilde{Y}_j)^2, \quad j = 1, \dots, c_0 \quad (4.3)$$

Esse ganho não influencia o procedimento de defuzzificação se o método do centro de gravidade for utilizado. Assim, a igualdade  $\tilde{Y} = GY$  ( $J_j = 0$ ,  $j = 1, \dots, c_0$ ) é equivalente a  $\tilde{Y} = Y$  com respeito a saída não-nebulosa  $\tilde{y}$ , como segue:

$$\tilde{y} = \frac{\sum_{j=1}^{c_0} GY_j \theta_j}{\sum_{j=1}^{c_0} GY_j} = \frac{\sum_{j=1}^{c_0} Y_j \theta_j}{\sum_{j=1}^{c_0} Y_j}$$

onde  $\theta_j$  é o valor modal do  $j$ -ésimo conjunto nebuloso de referência da interface de saída. No entanto, se algum outro método for utilizado, o ganho inverso  $G^{-1}$  deve ser aplicado antes da defuzzificação de  $\tilde{Y}$ .

A seção seguinte discute a influência do ganho  $G$  no problema de otimização e descreve como esse ganho é calculado.

### 4.3.1 O ganho $G$ da função de custo

Seja  $X \in [0, 1]^{c_1 \dots c_n}$  o vetor dos Produtos Cartesianos das entradas nebulosas, como segue:

$$X = X_1 \times X_2 \times \dots \times X_n \quad (4.4)$$

A solução ótima para a minimização de (4.3) é  $\tilde{Y} = GY$ . Nesse caso, a Equação (4.1) pode ser escrita como

$$\tilde{Y} = X \bullet R = GY \quad (4.5)$$

Para o pior caso em relação às entradas  $x_i$ , o máximo elemento de  $X$  assume seu valor mínimo, que é igual a  $G$ . Em outras palavras,  $\max(X)$  é igual ou maior que  $G$  para todo  $x_i \in X_i$  ( $i = 1, \dots, n$ ). Como  $Y$  pertence a  $[0, 1]^{c_0}$  tem-se que  $GY$  pertence a  $[0, G]^{c_0}$ . Logo,

$$\forall x_i, j: \max(X) \geq GY_j$$

e verifica-se a condição para a existência de soluções da Equação (4.5) em relação a  $R$  e para a composição max-t [Pedrycz, 1988, Pedrycz, 1993]. Assim, é sempre possível encontrar uma matriz relacional  $R$  tal que a saída ótima do modelo ( $\tilde{Y} = GY$ ) seja obtida, isso é,  $J_j = 0$  para  $j = 1, \dots, c_0$  em (4.3).

Seja  $\psi$  a norma triangular associada ao Produto Cartesiano na Equação (4.4). O ganho  $G$  é definido como o mínimo valor possível para o elemento máximo de  $X$ . Em função da monotonicidade das normas triangulares, esse ganho é igual ao produto da aplicação da norma-t  $\psi$ , individualmente, entre os mínimos valores possíveis dos máximos elementos das entradas  $X_i$  ( $i = 1, \dots, n$ ). Denominando cada um desses valores como  $g_i$ , tem-se

$$G = g_1 \psi g_2 \psi \dots \psi g_n$$

De acordo com a lei de fuzzificação definida pela Equação (2.9), para um conjunto específico de entradas  $x_i \in \mathbf{X}_i$  ( $i = 1, \dots, n$ ) tem-se as entradas nebulosas como vetores  $X_i = [f_{i_1}(x_i) \dots f_{i_{c_i}}(x_i)]$ , onde  $f_{i_\phi}$  é a função de pertinência do  $\phi$ -ésimo conjunto nebuloso pertencendo à  $i$ -ésima interface de entrada. Para obter os valores  $g_i$  deve-se calcular os mínimos valores possíveis dos máximos elementos desses vetores com relação a  $x_i$ , ou seja,

$$g_i = \text{Inf}_{x_i \in \mathbf{X}_i} \max(f_{i_1}(x_i), f_{i_2}(x_i), \dots, f_{i_{c_i}}(x_i))$$

A Figura (4.2) ilustra o cálculo de um valor  $g_i$  para uma variável de entrada nebulosa definida em uma interface com dois conjuntos nebulosos.

É importante notar que em função da condição limitante inferior da norma-t, qualquer valor  $g_i$  nulo implica em um ganho  $G$  também nulo. Para evitar essa situação claramente indesejada para o problema de otimização, os conjuntos nebulosos das extremidades direita e esquerda devem ser normais à direita e à esquerda (Apêndice C), respectivamente, como pode ser verificado na Figura (4.2). Outra situação indesejada é a ocorrência de um ganho  $G$  muito pequeno, porque isso pode ocasionar problemas numéricos. Em função da norma-t  $\psi$  escolhida, por exemplo o produto, essa situação pode ocorrer na presença de um grande número de entradas  $n$  no modelo. Nesse caso é importante uma adequada escolha de  $\psi$ , como por exemplo o operador min (Apêndice C).

### 4.3.2 Solução do problema de otimização

A partir da Equação (4.1) para a composição max-t,  $\tilde{Y}_j$  é escrito como segue:

$$\tilde{Y}_j = \bigvee_{\substack{l_i = 1 \\ i = 1, \dots, n}}^{c_i} (X_{1_{l_1}} \psi \dots \psi X_{n_{l_n}}) \tau R_{l_1, \dots, l_n, j}, \quad j = 1, \dots, c_0 \quad (4.6)$$



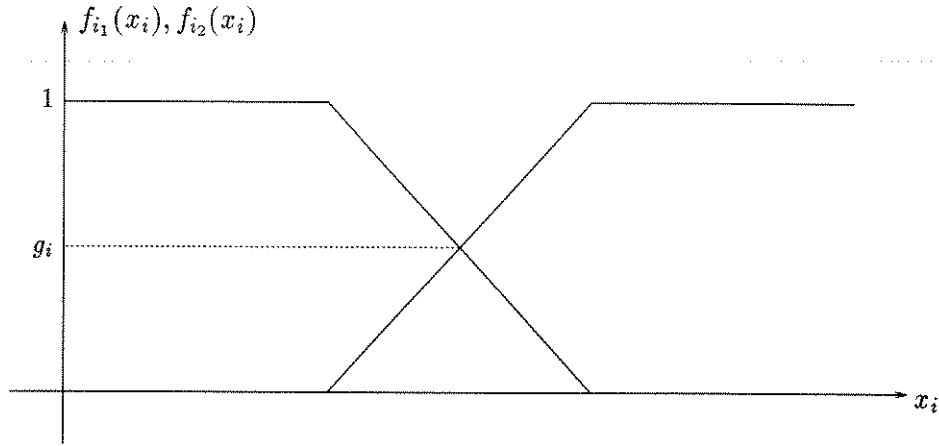


Figura 4.2: Valor mínimo  $g_i$  para a união de dois conjuntos nebulosos de referência.

onde  $\tau$  é a norma-t referente à composição,  $X_{m_{l_m}}$  é o  $l_m$ -ésimo elemento da  $m$ -ésima entrada nebulosa e  $R_{l_1, \dots, l_n, j}$  é um elemento específico da matriz relacional  $R$ . Por simplicidade, seja cada elemento do vetor de entrada  $X$ , definido em (4.4), denominado  $X_{l_1, \dots, l_n}$ . Então, substituindo (4.6) em (4.3) obtém-se

$$J_j = \frac{1}{2} \left( GY_j - \bigvee_{\substack{l_i = 1 \\ i = 1, \dots, n}}^{c_i} X_{l_1, \dots, l_n} \tau R_{l_1, \dots, l_n, j} \right)^2, \quad j = 1, \dots, c_0 \quad (4.7)$$

e o problema de otimização é dado por

$$\begin{aligned} \min_{R_{(\cdot), j}} J_j & \quad j = 1, \dots, c_0 \\ \text{sujeito a} & \quad 0 \leq R_{(\cdot), j} \leq 1 \end{aligned} \quad (4.8)$$

Desde que a informação correta sobre o elemento máximo  $X_{(\cdot)} \tau R_{(\cdot), j}$  não está disponível antes da identificação do modelo, ou seja, antes da determinação da matriz  $R$ , o operador  $\max (\vee)$  deve ser considerado como uma parte ativa de (4.8). Assim, para cada elemento  $R_{l_1, \dots, l_n, j}$  ( $j = 1, \dots, c_0$ ;  $l_i = 1, \dots, c_i$ ;  $i = 1, \dots, n$ ) define-se um problema de otimização como segue:

$$\min_{R_{l_1, \dots, l_n, j}} J_{l_1, \dots, l_n, j} = \frac{1}{2} (GY_j - X_{l_1, \dots, l_n} \tau R_{l_1, \dots, l_n, j})^2$$

sujeito a  $\Gamma_{l_1, \dots, l_n, j} \leq R_{l_1, \dots, l_n, j} \leq 1$  (4.9)

onde  $\Gamma_{l_1, \dots, l_n, j}$  é o menor valor de  $R_{l_1, \dots, l_n, j}$  para o qual  $X_{l_1, \dots, l_n} \tau R_{l_1, \dots, l_n, j}$  permanece máximo. Em outras palavras, para cada elemento de cada coluna  $j$  da matriz otimiza-se uma função de custo individual sobre o domínio restrito dos valores que maximizam a operação  $\tau$  entre esse elemento e o respectivo elemento do vetor de entrada em relação às demais operações na coluna.

As restrições  $\Gamma_{(\cdot), j}$  são calculadas como segue. Desconsiderando  $X_{l_1, \dots, l_n} \tau R_{l_1, \dots, l_n, j}$ , que se deseja manter máximo, toma-se a máxima operação  $\tau$  na coluna  $j$ , denominada  $X_{\ell_1, \dots, \ell_n} \tau R_{\ell_1, \dots, \ell_n, j}$ , entre aquelas restantes. Deseja-se

$$X_{l_1, \dots, l_n} \tau R_{l_1, \dots, l_n, j} \geq X_{\ell_1, \dots, \ell_n} \tau R_{\ell_1, \dots, \ell_n, j}$$

Devido à monotonicidade de  $\tau$ , decrescendo  $R_{l_1, \dots, l_n, j}$  a partir de um valor que satisfaça a condição acima, obtém-se o limite dessa condição, ou seja, a igualdade. Nessa situação, tem-se o valor de  $R_{l_1, \dots, l_n, j}$  que define  $\Gamma_{l_1, \dots, l_n, j}$ , que é dado por

$$\Gamma_{l_1, \dots, l_n, j} = \frac{X_{\ell_1, \dots, \ell_n} \tau R_{\ell_1, \dots, \ell_n, j}}{X_{l_1, \dots, l_n}} \quad (4.10)$$

Se a norma- $\tau$  for o produto, o Problema (4.9) é convexo e contínuo. Nesse caso, se  $\Gamma_{l_1, \dots, l_n, j} \leq 1$  a solução  $R_{l_1, \dots, l_n, j}^*$  é obtida diretamente por

$$\frac{\partial J_{l_1, \dots, l_n, j}}{\partial R_{l_1, \dots, l_n, j}} = X_{l_1, \dots, l_n} R_{l_1, \dots, l_n, j}^2 - GY_j R_{l_1, \dots, l_n, j} = 0$$

$$R_{l_1, \dots, l_n, j}^* = \begin{cases} \Gamma_{l_1, \dots, l_n, j}, & \text{se } \frac{GY_j}{X_{l_1, \dots, l_n}} \leq \Gamma_{l_1, \dots, l_n, j} \\ 1, & \text{se } \frac{GY_j}{X_{l_1, \dots, l_n}} \geq 1 \\ \frac{GY_j}{X_{l_1, \dots, l_n}}, & \text{caso contrário} \end{cases} \quad (4.11)$$

A partir de  $R_{l_1, \dots, l_n, j}^*$  obtém-se o mínimo valor da função de custo denominado  $J_{l_1, \dots, l_n, j}^*$ . Assim o Problema (4.8) torna-se

$$\min_{\substack{l_i = 1, \dots, c_i \\ i = 1, \dots, n}} J_{l_1, \dots, l_n, j}^*, \quad j = 1, \dots, c_0 \quad (4.12)$$

A solução de (4.8) é  $R_{l_1^*, \dots, l_n^*, j}^*$ , que é a solução ótima apresentada em (4.11) para o elemento da matriz referente aos índices ótimos  $l_1^*, \dots, l_n^*$  obtidos em (4.12). Quando ocorrem múltiplas soluções em (4.12), nesse trabalho escolhe-se aquela que envolve os índices relacionados ao máximo produto  $X_{(\cdot)} R_{(\cdot), j}$  baseado em  $R$  existente no instante de amostragem anterior. Isso faz com que após a convergência da matriz relacional para um modelo aproximado do sistema, a estrutura básica dessa matriz tenda a ser mantida.

Pode-se notar que se  $\Gamma_{l_1, \dots, l_n, j} > 1$ , o Problema (4.9) é infactível. Nesse caso, sendo que qualquer  $J_{(\cdot), j}^*$  pertence ao intervalo  $[0, \frac{1}{2}]$ , um valor elevado para  $J_{l_1, \dots, l_n, j}^*$  ( $> \frac{1}{2}$ ) deve ser imposto para que os índices  $l_1, \dots, l_n$  referentes a ele sejam descartados em (4.12).

Para sistemas invariantes no tempo ou lentamente variantes no tempo, os procedimentos acima podem ser simplificados através da inicialização do algoritmo de identificação com uma adequada estimativa da matriz  $R$  ou após a convergência dessa matriz para um modelo aproximado. Nessa situação supõe-se que a matriz possui informações corretas para a determinação dos elementos máximos em (4.6) e em consequência no Problema (4.8). Assim, em cada instante de amostragem é apenas necessário resolver, para cada  $j = 1, \dots, c_0$ , aquele Problema (4.9) referente aos índices  $l_1, \dots, l_n$  que maximizam (4.6).

### A heurística de otimização

Da Equação (4.10) pode-se verificar que as restrições de otimização  $\Gamma_{(\cdot), j}$  dependem diretamente da estrutura da matriz relacional  $R$ . Se essa matriz está polarizada, não correspondendo a um modelo adequado do sistema, as restrições são em geral incorretas. A seguir, propõe-se uma heurística para determinar e solucionar essa polarização. Supondo que se dispõe de um modelo exato do sistema, para cada  $j = 1, \dots, c_0$  existe um elemento  $R_{(\cdot), j}$  tal que  $J_j = 0$  em (4.7). Logo, a matriz  $R$  está polarizada se para cada  $j = 1, \dots, c_0$  não for possível determinar, devido às restrições  $\Gamma_{(\cdot), j}$ , pelo menos um elemento  $R_{(\cdot), j}$  tal que  $J_{(\cdot), j} \leq \varepsilon$  em (4.9), para  $\varepsilon \in [0, \frac{1}{2})$ . O parâmetro  $\varepsilon$  representa imprecisões numéricas, perturbações e dinâmicas não-modeladas. Se essa determinação não for possível, a solução para o problema de otimização é obtida ignorando as restrições  $\Gamma_{(\cdot), j}$ , isto é, utilizando o domínio  $[0, 1]$  ao invés de  $[\Gamma_{(\cdot), j}, 1]$  na solução de cada problema (4.9), dada pela Equação (4.11).

## 4.4 Algoritmo de identificação

Obtém-se a identificação “on-line” de um modelo relacional nebuloso e a otimização de suas interfaces através do seguinte algoritmo básico:

- (a) Selecionar a estrutura do modelo (ver Equação (2.1)): As ordens  $p$  e  $q_i$  ( $i = 1, \dots, v$ ), os atrasos de transporte  $t_i$ , a norma-t  $\psi$  e os valores  $c_0, c_1, \dots, c_n$  dos conjuntos nebulosos de referência.
- (b) Selecionar um valor para  $\varepsilon$ . Em geral um valor pequeno (tal como  $\varepsilon = 10^{-4}$ ) apenas para evitar imprecisões numéricas é adequado, especialmente para modelos de ordem completa e na ausência de perturbações.
- (c) Selecionar parâmetros adequados para o algoritmo  $\Sigma$ -PAFIO e para as interfaces iniciais conforme descrito no Capítulo 3 e no Apêndice B.
- (d) Fazer  $k = 0$  e inicializar  $R$ .
- (e) Fazer  $k = k + 1$ .
- (f) Atualizar o ganho  $G$  ou, por simplicidade, manter o valor existente no instante de tempo anterior se apenas variações desprezíveis ocorrerem nas interfaces.
- (g) Fuzzificar as entradas não-nebulosas  $x_i(k)$  ( $i = 1, \dots, n$ ) e a saída  $y(k)$  para obter  $X_i(k)$  e  $Y(k)$ , respectivamente. Calcular  $X$  a partir da Equação (4.4).
- (h) Para  $j = 1$  até  $c_0$  executar as operações até o Passo I.
  - (i) Para  $l_i = 1$  até  $c_i$  ( $i = 1, \dots, n$ ) executar as operações até o Passo k.
    - (j) Calcular  $\Gamma_{l_1, \dots, l_n, j}$  baseado em  $R$  existente no instante de tempo anterior. Se  $\Gamma_{l_1, \dots, l_n, j} > 1$ , fazer  $J_{l_1, \dots, l_n, j} > \frac{1}{2}$ , caso contrário calcular uma variável auxiliar  $\Lambda_{l_1, \dots, l_n, j}$  utilizando a Equação (4.11) e calcular  $J_{l_1, \dots, l_n, j} = \frac{1}{2}(GY_j - X_{l_1, \dots, l_n} \Lambda_{l_1, \dots, l_n, j})^2$ .
    - (k) Se  $J_{l_1, \dots, l_n, j} > \varepsilon$ , calcular uma variável auxiliar  $\Omega_{l_1, \dots, l_n, j}$  utilizando a Equação (4.11) para  $\Gamma_{l_1, \dots, l_n, j} = 0$  e calcular  $Q_{l_1, \dots, l_n, j} = \frac{1}{2}(GY_j - X_{l_1, \dots, l_n} \Omega_{l_1, \dots, l_n, j})^2$  como uma função de custo auxiliar.

(l) Obter os índices  $l_1, \dots, l_n$  relacionados ao mínimo  $J_{l_1, \dots, l_n, j}$ . Se existirem múltiplas soluções, selecionar aquela cujo produto  $X_{(\cdot)} R_{(\cdot), j}$  é máximo baseado em  $R$  existente no instante de tempo anterior. Se  $J_{l_1, \dots, l_n, j} \leq \varepsilon$  então  $R_{l_1, \dots, l_n, j} = \Lambda_{l_1, \dots, l_n, j}$ , caso contrário obter os índices  $l_1, \dots, l_n$  referentes ao mínimo  $Q_{l_1, \dots, l_n, j}$ , e então  $R_{l_1, \dots, l_n, j} = \Omega_{l_1, \dots, l_n, j}$ . Nesse caso, o procedimento para múltiplas soluções é o mesmo.

(m) Executar uma iteração do algoritmo  $\Sigma$ -PAFIO se um critério de erro que avalia o desempenho desse algoritmo for maior que um valor específico (Apêndice B).

(n) Voltar para o Passo e.

Pode-se notar que após a convergência do algoritmo para um modelo aproximado do sistema, a simplificação proposta na Seção 4.3.2 permite, para identificação de sistemas invariantes no tempo ou lentamente variantes no tempo, a substituição dos passos **h**, **i**, **j**, **k** e **l** pelo seguinte passo:

( $\cdot$ ) Para  $j = 1$  até  $c_0$ : Para os índices  $l_1, \dots, l_n$  que maximizam (4.6), calcular  $\Gamma_{l_1, \dots, l_n, j}$  e o valor para  $R_{l_1, \dots, l_n, j}$  utilizando a Equação (4.11).

## 4.5 Topologias de realimentação

O esquema de identificação para um modelo relacional nebuloso apresentado na Figura (4.1) utiliza a topologia de realimentação numérica, onde a saída do sistema é realimentada numericamente como entrada no modelo. Essa topologia permite que a saída realimentada possa ser fuzzificada através de uma interface independente ( $N/L_2$ ). As variáveis envolvidas nas interfaces  $N/L_2$  e  $N/L_3$  (que possui a mesma estrutura da interface de saída do modelo  $L/N$ ) são as mesmas, apenas em instantes de amostragem diferentes. Logo, a interface  $N/L_2$  poderia ser eliminada através do uso da topologia de realimentação linguística, onde a variável nebulosa de saída é realimentada diretamente. Contudo, a realimentação numérica permite um grau de liberdade maior na especificação das interfaces. Através desse tipo de realimentação pode-se, por exemplo, utilizar um número menor de conjuntos nebulosos de referência em  $N/L_2$  com relação a  $N/L_3$  e  $L/N$ , diminuindo assim a dimensão da matriz relacional e mantendo a precisão na saída.

Outra vantagem da topologia de realimentação numérica diz respeito à identificação de processos auto-regressivos lentos. Para a identificação de processos auto-regressivos através da topologia de realimentação linguística, qualquer elemento particular  $Y_j$  é incluído em (4.9) diretamente como saída no instante de amostragem  $k$  e indiretamente como entrada em  $X_{(\cdot)}$ , no instante  $k - 1$ , para conjuntos específicos de índices  $l_1, \dots, l_n$ . Nesse caso, os elementos da relação nebulosa relacionados a esses índices praticamente independem de  $Y_j$  nos instantes  $k$  e  $k - 1$  para processos com constante de tempo lenta ou discretizados através de uma taxa de amostragem excessivamente rápida, porque nesses processos a saída  $Y$  no instante  $k$  é aproximadamente igual a essa saída no instante  $k - 1$ . Desse modo, os elementos  $Y_j$  nesses diferentes instantes são praticamente cancelados no numerador e denominador da Equação (4.11), acarretando problemas de convergência na matriz relacional  $R$ . A utilização da topologia de realimentação numérica com interfaces diferentes para fuzzificação da saída do sistema e para fuzzificação da entrada de realimentação no modelo evita esse problema. Nesse caso, mesmo as saídas numéricas nos instantes  $k$  e  $k - 1$  sendo semelhantes, suas representações linguísticas  $Y$  e  $Y'$  (ver Figura (4.1)) são distintas. Contudo, a necessidade de interfaces distintas restringe a especificação de estrutura de uma interface com relação à outra. Uma solução mais genérica é a utilização de um modelo incremental, descrito na seção seguinte, que possui variáveis nebulosas distintas de saída e realimentação da saída, independentemente da estrutura das interfaces.

## 4.6 O modelo incremental

O esquema de identificação de um modelo relacional nebuloso incremental é apresentado na Figura (4.3). Esse modelo utiliza a variação da saída  $\Delta\tilde{Y}$  e não a saída posicional  $\tilde{Y}$ . Nesse caso, o critério de otimização é dado por

$$J_j = \frac{1}{2}(G \Delta Y_j - \Delta\tilde{Y}_j)^2, \quad j = 1, \dots, c_0$$

A saída nebulosa do modelo permanece equacionada através de (4.6), porém para a variação, ou seja,

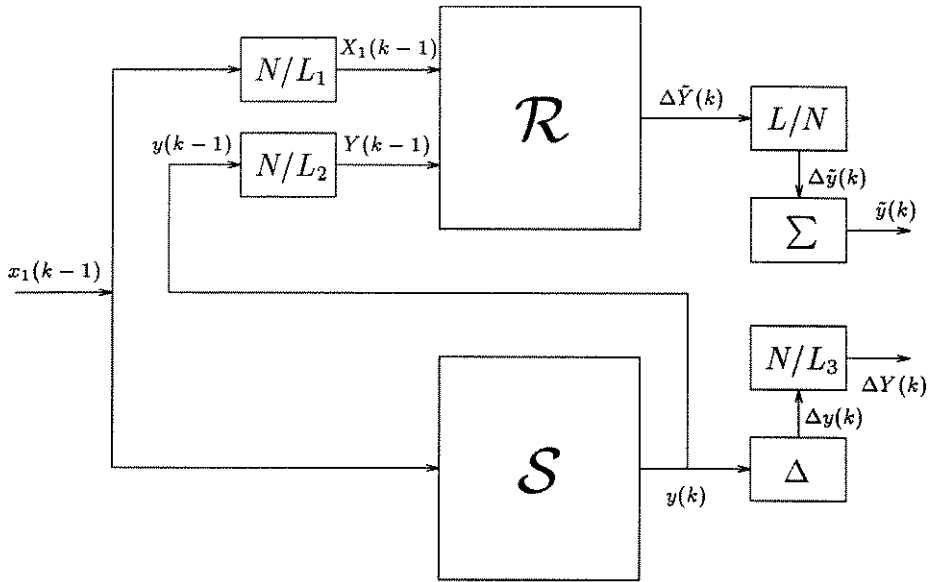


Figura 4.3: Esquema de identificação de um sistema ( $\mathcal{S}$ ) utilizando um modelo relacional nebuloso incremental.

$$\Delta \tilde{Y}_j = \prod_{\substack{i=1, \dots, n \\ l_i=1}}^{c_i} (X_{1_{l_1}} \psi \cdots \psi X_{n_{l_n}}) \tau R_{l_1, \dots, l_n, j}, \quad j = 1, \dots, c_0$$

Analogamente ao desenvolvimento da Seção 4.3.2 a solução para o problema de otimização é dada por

$$R_{l_1, \dots, l_n, j}^* = \begin{cases} \Gamma_{l_1, \dots, l_n, j}, & \text{se } \frac{G \Delta Y_j}{\bar{X}_{l_1, \dots, l_n}} \leq \Gamma_{l_1, \dots, l_n, j} \\ 1, & \text{se } \frac{G \Delta Y_j}{\bar{X}_{l_1, \dots, l_n}} \geq 1 \\ \frac{G \Delta Y_j}{\bar{X}_{l_1, \dots, l_n}}, & \text{caso contrário} \end{cases} \quad (4.13)$$

Pode-se observar na Figura (4.3) que a entrada do modelo relativa à realização do sistema permanece envolvendo a saída posicional  $y$ . Assim os elementos  $\Delta Y_j$  ocorrem apenas no numerador de (4.13) e o problema de cancelamento descrito na seção anterior não ocorre.

## 4.7 Sumário

Na literatura de sistemas nebulosos foram propostos diferentes algoritmos adaptativos para refinar modelos relacionais obtidos através de métodos aproximados. O principal problema desses algoritmos é a dependência de uma estimativa adequada da relação nebulosa que se torna crítica quando ocorrem variações no sistema. Esse capítulo apresentou um algoritmo de identificação que utiliza uma heurística para resolver o problema de otimização quando essa estimativa da relação não é disponível. Essa heurística consiste na verificação, em cada instante de tempo, da otimalidade da solução do problema de identificação e da realização de uma otimização irrestrita caso essa otimalidade não seja satisfeita. O algoritmo proposto substitui o operador max da composição nebulosa por restrições de otimização ao invés de suprimi-lo. O operador é suprimido apenas quando uma estimativa adequada da relação nebulosa é disponível. Essas características proporcionam uma elevada capacidade de adaptação ao algoritmo, que utiliza um critério de otimização que possibilita a identificação de sistemas nebulosos ou não-nebulosos e independe das características do método de defuzzificação adotado.

A identificação de modelos relacionais nebulosos pode ser realizada através das topologias de realimentação numérica ou linguística. No entanto, o algoritmo de identificação apresenta problemas de convergência na modelagem de sistemas auto-regressivos lentos através da topologia de realimentação linguística. A topologia de realimentação numérica evita esse problema e provê maior mobilidade ao modelo permitindo a utilização de interfaces diferentes para a saída e para a entrada realimentada. Como a utilização de interfaces diferentes restringe a especificação de estrutura dessas interfaces, um modelo nebuloso incremental pode ser utilizado. Esse modelo possui variáveis nebulosas distintas de saída e entrada realimentada independentemente da estrutura das interfaces, evitando os problemas de convergência do modelo relacional.



# Capítulo 5

## Exemplos Numéricos de Identificação Adaptativa

Esse capítulo apresenta três exemplos numéricos de identificação adaptativa de modelos relacionais nebulosos utilizando o algoritmo proposto no capítulo anterior.

### 5.1 Introdução

Para a análise da eficiência de um algoritmo de identificação, devem ser considerados aspectos tais como convergência, robustez e precisão. Esse capítulo apresenta três exemplos de identificação adaptativa abordando esses aspectos. Utilizam-se sistemas não-lineares e inicializa-se o algoritmo com matrizes relacionais nebulosas sem informação sobre o sistema, para analisar a capacidade de adaptação dos modelos. A avaliação do algoritmo e a validação dos modelos são realizadas através de previsões da saída um passo à frente e geração de séries sintéticas.

O primeiro exemplo numérico utiliza um sistema com um parâmetro variando lentamente no tempo. Essa variação e a inicialização do algoritmo através de uma matriz polarizada permitem testar a capacidade de adaptação do modelo, cuja convergência é ilustrada através da evolução da matriz relacional. O segundo exemplo utiliza um sistema com ruído branco aditivo e um modelo nebuloso de ordem reduzida, permitindo a análise de desempenho do algoritmo na presença de perturbações e dinâmicas não-modeladas. O terceiro exemplo realiza uma comparação entre os desempenhos de identificação de um

sistema através do algoritmo proposto no capítulo anterior para duas normas triangulares diferentes para o Produto Cartesiano e através do algoritmo de aproximação linguística apresentado no Apêndice A.

## 5.2 Exemplo 1

O processo não-linear e variante no tempo apresentado por Oliveira e Lemos [Oliveira e Lemos, 1995] é utilizado para ilustrar o desempenho do algoritmo proposto para identificação “on-line”. Esse processo é descrito pela seguinte equação a diferenças:

$$y(k) = 0,1y(k-1)\exp(-0,25u^2(k-1)) + 0,5u(k-1) + a(k-1)y(k-1)$$

onde  $a$  é um parâmetro variante no tempo, conforme mostrado na Figura (5.1), e a entrada  $u$  é uma sequência aleatória uniformemente distribuída entre  $-5,0$  e  $5,0$ .

A estrutura do modelo é dada por

$$\Delta Y(k) = X_1(k) \bullet X_2(k) \bullet R$$

onde  $X_1(k) = Y(k-1)$ ,  $X_2(k) = U(k-1)$  e “ $\bullet$ ” denota a composição max-produto.

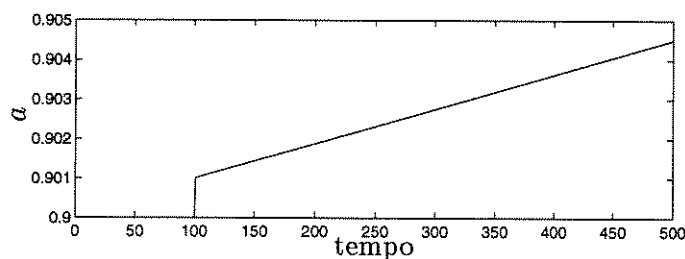


Figura 5.1: Parâmetro  $a$ .

Para ilustrar a adaptação do modelo, o procedimento de identificação é inicializado com uma matriz relacional nula e executado durante 500 instantes de tempo. A norma-t  $\psi$  utilizada é o produto. Como não há perturbações e o modelo é de ordem completa, seleciona-se  $\varepsilon = 10^{-4}$ , conforme descrito na Seção 4.4.

Simultaneamente à identificação “on-line”, as interfaces são otimizadas pelo algoritmo  $\Sigma$ -PAFIO com  $c_0 = c_1 = c_2 = 5$  conjuntos nebulosos de referência, taxa de aprendizado  $\eta = 0,8$ , constante de distinção  $K = 5$ , norma para medida “sigma-count”  $\rho = 4$ , nível de desligamento  $\lambda = 0,05$ , janela da sequência de erros  $H = 10$  e pólo do filtro  $\gamma = 0,9$ . A seleção desses parâmetros foi discutida na Seção 3.4. As funções de pertinência para o instante final de identificação  $k = 500$  são mostradas na Figura (5.2), onde  $UoD$  representa o universo de discurso.

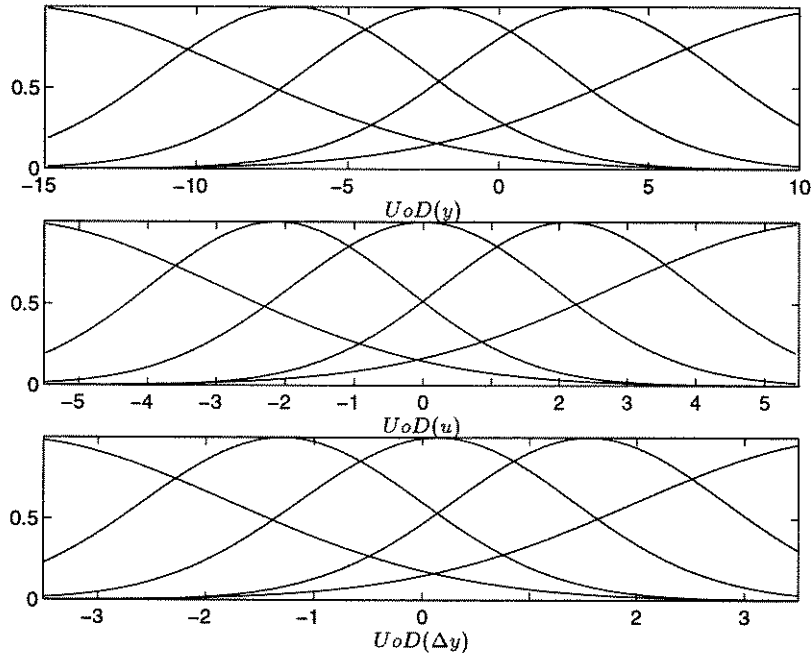


Figura 5.2: Funções de pertinência para  $k = 500$ .

Na Figura (5.3) apresenta-se (a) a variação da saída medida do processo, (b) a previsão um passo à frente “on-line” do modelo incremental e (c) uma ampliação das curvas nos últimos 50 instantes de tempo. Dessa figura pode-se verificar que o modelo prevê adequadamente a saída do sistema, ilustrando a eficiência do método de otimização em cada instante.

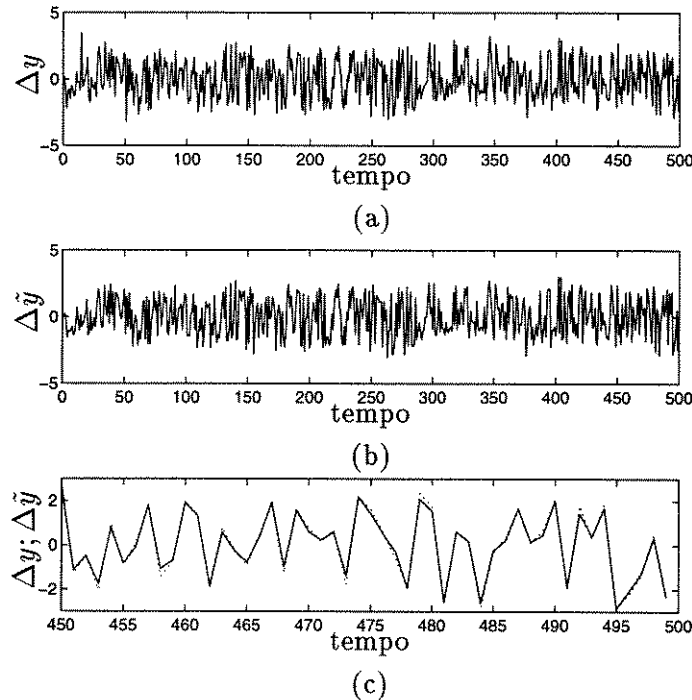


Figura 5.3: (a) Variação da saída do processo; (b) Previsão “on-line” do modelo; (c) Curvas sobrepostas: processo (linha cheia) e modelo (linha pontilhada).

As representações 3D da matriz relacional nos instantes de amostragem 20, 200, 350 e 500 durante o procedimento de identificação são mostrados na Figura (5.4), onde  $\times$  representa o espaço dos Produtos Cartesianos. Essa figura ilustra a convergência do modelo e sua capacidade de adaptação.

A eficiência do modelo incremental pode ser vista na Figura (5.5) que mostra a variação da saída do processo sobreposta à série sintética do modelo. Essa série sintética foi gerada utilizando a relação identificada e as interfaces ótimas para  $k = 500$ , e é importante para avaliar o modelo em aplicações de

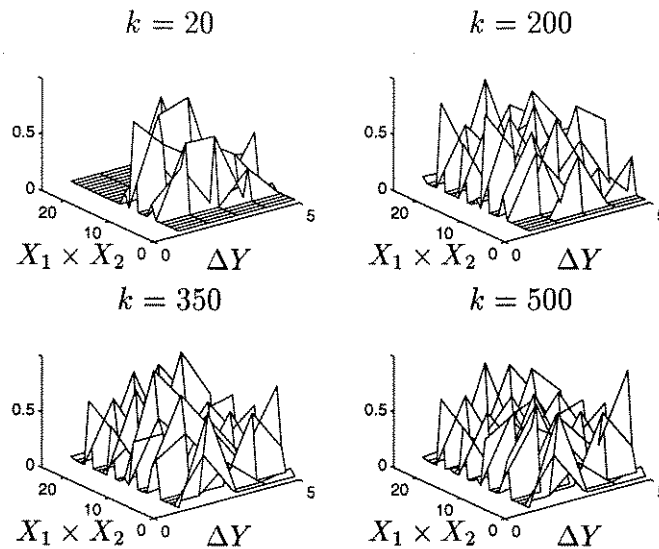


Figura 5.4: Matriz relacional em diferentes instantes de amostragem.

controle preditivo, onde são necessárias previsões da saída em vários instantes de tempo.

A Figura (5.6) apresenta a saída posicional do processo sobreposta à série sintética da saída integrada do modelo (a) com a mesma sequência de excitação utilizada na identificação e (b) com uma sequência diferente de excitação. A partir desses resultados pode-se notar que o modelo obtido apresenta um bom desempenho mesmo para outras sequências de excitação.

### 5.3 Exemplo 2

Um processo não-linear é considerado como segue:

$$y(k) = 0,2y(k-1) - 0,1y(k-2) + 0,4\sqrt{|y(k-1)|} + 0,8u(k-1) + 0,1\xi$$

onde  $\xi$  é um ruído branco Gaussiano com média zero e variância  $\sigma^2 = 1,0$  e a entrada  $u$  é uma sequência aleatória uniformemente distribuída entre  $-2,0$  e  $2,0$ . O seguinte modelo nebuloso de ordem reduzida é utilizado para a identificação desse processo:

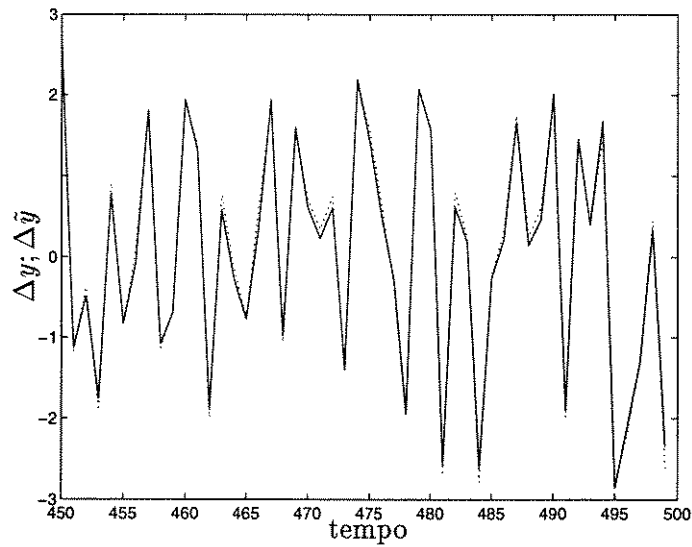
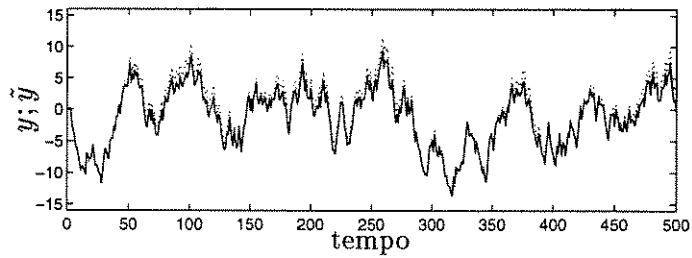
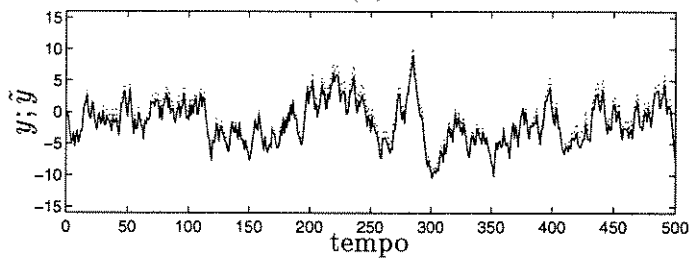


Figura 5.5: Variação da saída do processo (linha cheia) e série sintética do modelo incremental (linha pontilhada).



(a)



(b)

Figura 5.6: Saída do processo (linha cheia) e série sintética da saída integrada do modelo (linha pontilhada): (a) Para a mesma sequência de excitação da identificação; (b) Para uma sequência diferente de excitação.

$$Y(k) = X_1(k) \bullet X_2(k) \bullet R$$

onde  $X_1(k) = Y(k - 1)$ ,  $X_2(k) = U(k - 1)$  e “ $\bullet$ ” representa a composição max-produto.

Em função do modelo de ordem reduzida e da presença de perturbações no processo, seleciona-se  $\varepsilon = 5 \times 10^{-4}$ , que é um valor maior do que aquele utilizado no Exemplo 1. Selecionando os outros parâmetros com valores iguais aos definidos naquele exemplo, o procedimento de identificação é executado durante 400 instantes. Na Figura (5.7) apresenta-se (a) a saída medida do processo, (b) a previsão um passo à frente “on-line” do modelo e (c) uma ampliação das curvas nos últimos 100 instantes de tempo. Essa figura ilustra novamente que o modelo proposto representa adequadamente a saída do processo durante a identificação.

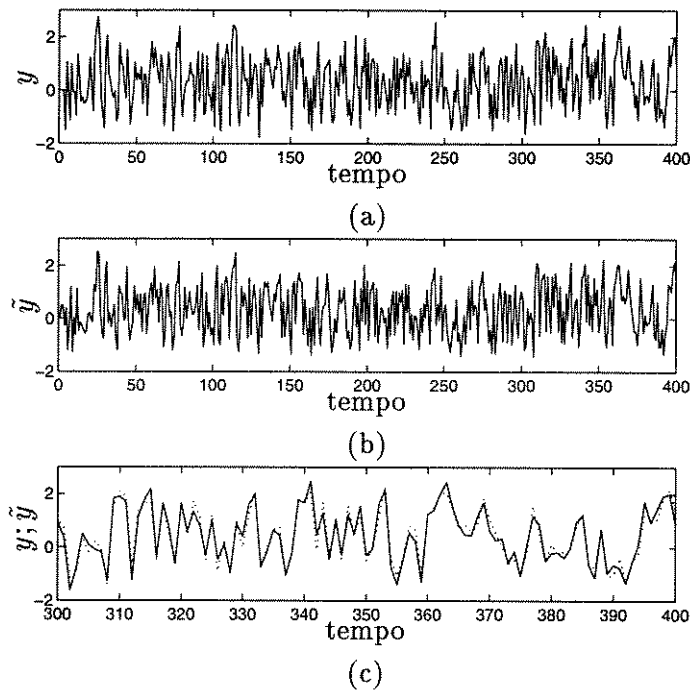


Figura 5.7: (a) Saída do processo; (b) Previsão “on-line” do modelo; (c) Curvas sobrepostas: processo (linha cheia) e modelo (linha pontilhada).

A Figura (5.8) mostra a saída do processo determinístico sobreposta à série sintética do modelo identificado, ilustrando a eficiência da modelagem proposta na previsão em vários instantes de tempo à frente.

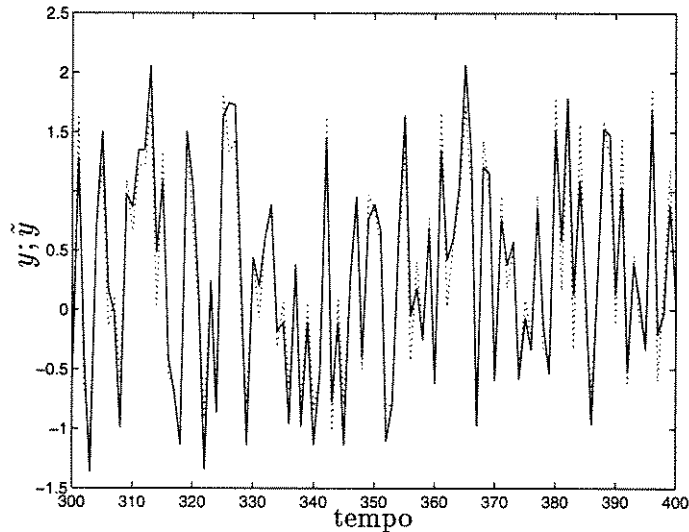


Figura 5.8: Saída do processo (linha cheia) e série sintética do modelo (linha pontilhada).

Embora o modelo nebuloso não tenha uma interpretação frequencial, uma vez que representa relações entre amplitudes de entrada e saída, os espectros das sequências de resposta desse modelo e do processo podem ser comparados. A Figura (5.9) mostra as densidades espectrais de potência da saída do processo e da série sintética do modelo. Pode-se observar que as curvas são bem semelhantes apesar da utilização de um modelo de ordem reduzida.

Para avaliar o comportamento do modelo em regime, a Figura (5.10-a) mostra a saída do processo e a série sintética do modelo para uma sequência de degraus de entrada apresentada na Figura (5.10-b). Pode-se notar um pequeno erro de regime nessa figura. Esse erro é causado por imprecisões em parte na relação nebulosa identificada e em parte na otimização das interfaces.



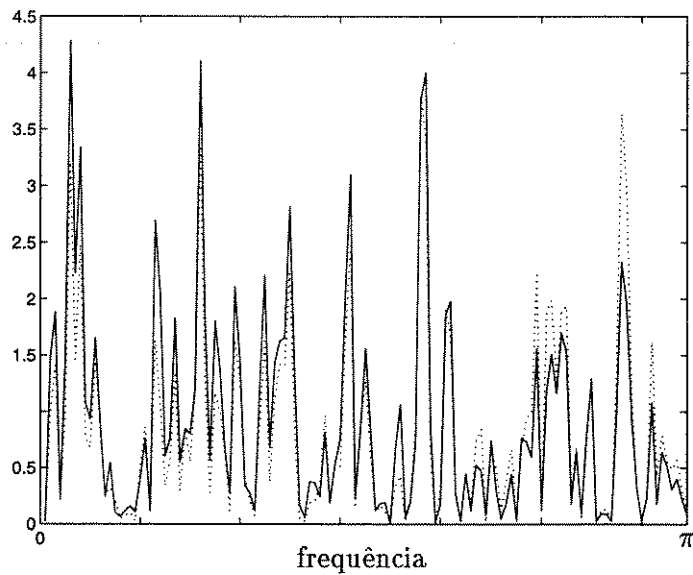


Figura 5.9: Densidade espectral de potência da saída do processo (linha cheia) e da série sintética do modelo (linha pontilhada).

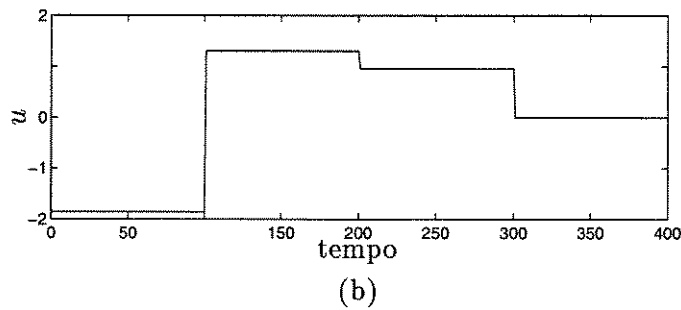
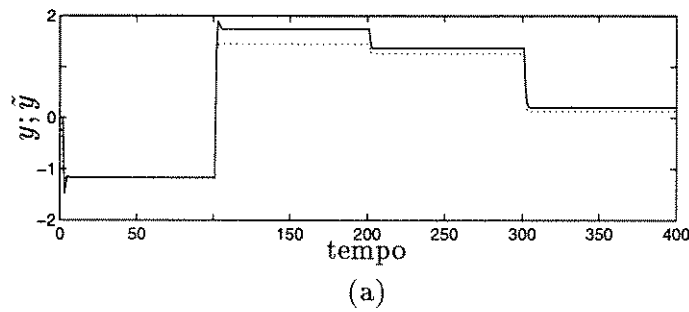


Figura 5.10: (a) Saída do processo (linha cheia) e série sintética do modelo (linha pontilhada); (b) Sequência de degraus de entrada.

## 5.4 Exemplo 3

Identifica-se a seguir um processo bi-linear dado por

$$y(k) = 0,3y(k-1) + 0,8u(k-1) \\ + 0,2u^2(k-1) + 0,1\xi$$

onde  $\xi$  é um ruído branco Gaussiano com média zero e variância  $\sigma^2 = 1,0$  e a entrada  $u$  é uma sequência aleatória uniformemente distribuída entre  $-2,0$  e  $2,0$ . O seguinte modelo nebuloso é utilizado para a identificação desse processo:

$$Y(k) = X_1(k) \bullet X_2(k) \bullet R$$

onde  $X_1(k) = Y(k-1)$ ,  $X_2(k) = U(k-1)$  e “ $\bullet$ ” denota a composição max-produto.

O procedimento de identificação é executado durante 500 instantes com  $\varepsilon = 10^{-3}$  e os outros parâmetros com valores iguais aos definidos nos exemplos anteriores. A utilização de um valor para  $\varepsilon$  superior aos utilizados anteriormente ilustra que a escolha desse parâmetro não é crítica. A Figura (5.11) apresenta a saída do processo determinístico sobreposta à série sintética do modelo entre os instantes de tempo  $k = 400$  e  $k = 500$ . Essa figura ilustra novamente um bom desempenho do modelo assim como a eficiência do algoritmo de identificação.

Para uma avaliação quantitativa de desempenho, a seguinte média temporal é tomada sobre  $N$  amostras:

$$\mu_\varphi(N) = \frac{1}{N} \sum_{k=1}^N |\varphi(k)|. \quad (5.1)$$

onde  $\varphi$  é uma sequência genérica de dados. Uma estimativa do valor absoluto médio da saída do processo, calculada utilizando a Equação (5.1) para  $\varphi = y$  e  $N = 500$  instantes de simulação, é  $\mu_y(500) = 0,8086$ . Analogamente, uma estimativa do erro absoluto médio entre a saída do processo e a série sintética do modelo, calculada utilizando a Equação (5.1) para  $\varphi = y - \tilde{y}$  e  $N = 500$ ,

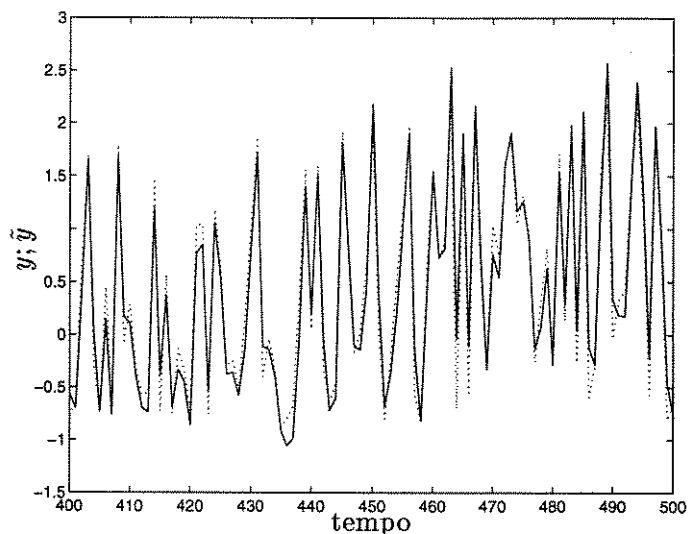


Figura 5.11: Saída do processo (linha cheia) e série sintética do modelo (linha pontilhada).

é  $\mu_{y-\hat{y}}(500) = 0,1768$ . O último valor ( $\mu_{y-\hat{y}}$ ) representa apenas aproximadamente 21,8% do primeiro ( $\mu_y$ ), o que é um resultado adequado especialmente considerando que utiliza-se 500 instantes de previsão e que os dados do processo real não são utilizados no modelo durante esse horizonte. Assim, apesar da acumulação dos erros durante a geração da série sintética, pode-se notar na Figura (5.11) que o modelo permanece como uma representação fiel do processo.

Para fins de comparação, o processo é identificado através de um modelo utilizando a norma-t  $\psi$  min e não o produto. A Figura (5.12) mostra a saída do processo real sobreposta à série sintética do modelo. Essa figura ilustra um desempenho equivalente àquele apresentado pelo modelo anterior. Nesse caso,  $\mu_{y-\hat{y}}(500) = 0,1829$ , que representa aproximadamente 22,6% de  $\mu_y(500)$ . Esse resultado demonstra que o desempenho do algoritmo proposto praticamente independe da norma-t utilizada para implementação do Produto Cartesiano.

Para comparar o desempenho dos dois modelos obtidos nesse exemplo através da modelagem proposta nesse trabalho, utiliza-se para identificação desse processo o algoritmo de aproximação linguística apresentado no Apêndice A para composição max-produto. Essa modelagem utiliza a mesma sequência de dados de entrada e saída do processo e as mesmas interfaces otimizadas

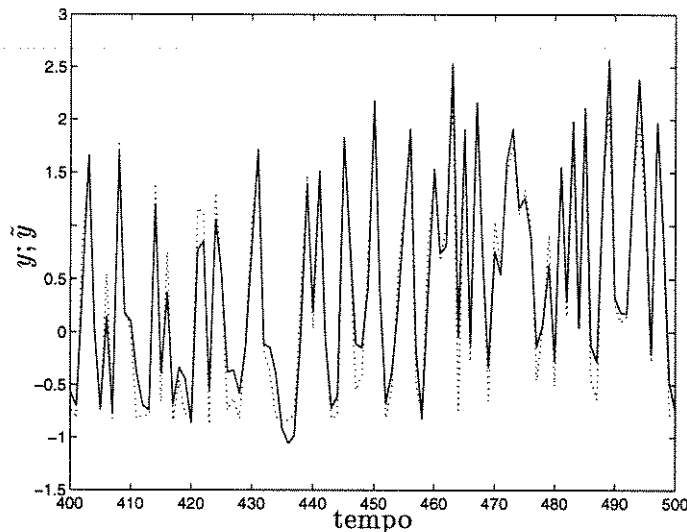


Figura 5.12: Saída do processo (linha cheia) e série sintética do modelo identificado com a norma-t  $\psi$  min (linha pontilhada).

para os dois modelos anteriores. A Figura (5.13) mostra a saída do processo determinístico sobreposta à série sintética do modelo. Nesse caso, tem-se  $\mu_{y-\hat{y}}(500) = 0,3540$ , que representa aproximadamente 43,8% de  $\mu_y(500)$  e é praticamente duas vezes maior que aquele valor obtido para os modelos anteriores. Esse resultado ilustra novamente a eficiência do algoritmo proposto.

## 5.5 Sumário

Esse capítulo apresentou três exemplos de identificação adaptativa de processos não-lineares utilizando modelos relacionais nebulosos. Nesses exemplos, o algoritmo de identificação foi inicializado com matrizes relacionais nebulosas sem qualquer informação sobre o processo. Essa inicialização possibilitou a verificação da elevada capacidade de adaptação dos modelos, que foi ilustrada no Exemplo 1 através da convergência da matriz relacional durante a modelagem de um sistema variante no tempo. A avaliação do algoritmo e a validação dos modelos foram realizadas através de previsões da saída um passo à frente e geração de séries sintéticas. Em todos os exemplos as respostas dos modelos representaram corretamente as saídas dos processos. Essas respostas

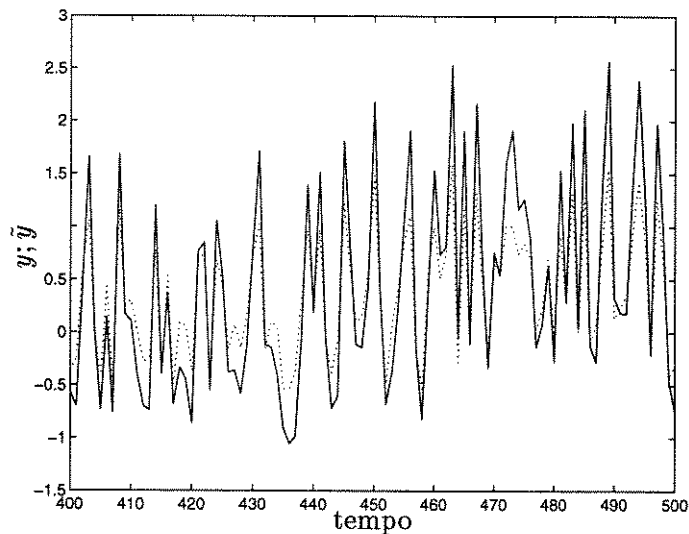


Figura 5.13: Saída do processo (linha cheia) e série sintética do modelo identificado por aproximação linguística (linha pontilhada).

ilustraram no Exemplo 2 a eficiência do algoritmo proposto na presença de perturbações e dinâmicas não-modeladas. Nesse exemplo, a comparação da densidade espectral de potência das sequências de saída do processo e do modelo comprovaram a semelhança desses sinais também no domínio da frequência. O comportamento desse modelo em regime foi analisado através da geração de série sintética para uma sequência de degraus de entrada. No Exemplo 3, o processo foi identificado por três modelos distintos. Os dois primeiros através do algoritmo proposto porém utilizando normas triangulares diferentes relacionadas ao Produto Cartesiano. O terceiro foi identificado através do algoritmo de aproximação linguística. Os resultados demonstraram um desempenho equivalente dos modelos identificados pelo algoritmo proposto, significativamente superior ao desempenho do modelo identificado por aproximação linguística.

## Capítulo 6

# Conclusões e Perspectivas

Esse trabalho apresentou uma nova metodologia para identificação adaptativa de modelos relacionais nebulosos. Foi obtido um algoritmo com elevada capacidade de adaptação que não necessita de uma estimativa adequada da matriz relacional. Esse algoritmo permite aliar a capacidade de tratamento de incertezas dos modelos nebulosos com um procedimento de otimização eficiente na modelagem de sistemas complexos.

A utilização do algoritmo  $\Sigma$ -PAFIO possibilitou a otimização das interfaces minimizando os erros nos mapeamentos de entrada e saída e mantendo a integridade semântica dos conjuntos nebulosos que as constituem. Além da otimalidade, o algoritmo  $\Sigma$ -PAFIO provê uma característica adaptativa às interfaces, permitindo aos conjuntos nebulosos adaptarem-se às mudanças nos sinais do sistema. A utilização simultânea desse algoritmo e do algoritmo de identificação adaptativa completam a metodologia, que se torna de particular interesse para aplicações “on-line”.

Três processos foram identificados utilizando a metodologia proposta. Os modelos obtidos foram avaliados através de séries sintéticas e previsões da saída um passo à frente. A inicialização do algoritmo com uma matriz relacional nula e a utilização de um modelo de ordem reduzida na modelagem de um processo com ruído aditivo ilustraram respectivamente a capacidade de adaptação e a eficiência do algoritmo de identificação. Através da identificação de um mesmo processo utilizando três modelos diferentes, verificou-se que a escolha da norma triangular referente ao Produto Cartesiano não é crítica. Verificou-se ainda que o desempenho do algoritmo proposto é significativamente superior ao desempenho do método de aproximação linguística. Embora isso também ocorra para

os algoritmos adaptativos propostos na literatura, nenhuma comparação com eles foi realizada em função de que as vantagens do algoritmo proposto são estruturais, especialmente no que se refere à independência de uma estimativa adequada da matriz relacional.

Como perspectivas de continuação desse trabalho sugere-se:

- O desenvolvimento do algoritmo utilizando uma norma triangular referente à composição nebulosa diferente do produto, ou ainda, utilizando uma outra composição (como por exemplo a composição Inf-s).
- A utilização de outra técnica de fuzzificação diferente do singleton nebuloso, para propiciar inclusive o tratamento de incertezas diretamente nas interfaces de entrada.
- O desenvolvimento de um critério matemático para a convergência da matriz relacional nebulosa considerando a possibilidade de existência de múltiplas soluções para as equações relacionais.
- O desenvolvimento de um algoritmo para a identificação de modelos não-estruturados, ou seja, que não necessitam de um preestabelecimento da ordem e atraso de transporte do sistema.
- O estudo das relações entre as matrizes relacionais dos modelos incremental e posicional de um mesmo sistema.

# Apêndice A: Algoritmo de Aproximação Linguística

Um algoritmo de aproximação linguística baseado na agregação de Produtos Cartesianos para estimação da matriz relacional em modelos relacionais nebulosos é apresentado.

Sejam  $X_i(t)$  para  $i = 1, \dots, n$  e  $Y(t)$  respectivamente os  $n$  conjuntos nebulosos de entrada e o conjunto nebuloso de saída de um modelo relacional, com múltiplas entradas e saída única, no instante  $t$ . Então pode-se calcular a relação nebulosa  $R'$  desse modelo para os dados conjuntos nebulosos de entrada e saída através dos Produtos Cartesianos ( $\times$ ) desses conjuntos como segue:

$$R'(t) = X_1(t) \times X_2(t) \times \dots \times X_n(t) \times Y(t)$$

No caso da utilização da composição max-t, associa-se ao Produto Cartesiano uma norma triangular (norma-t). Sendo  $X_i(t) = [x_{i_1} \dots x_{i_{c_i}}]$ ,  $Y(t) = [y_1 \dots y_{c_0}]$  e  $\min (\wedge)$ , a norma triangular escolhida, tem-se a composição max-min onde cada elemento  $R'_{l_1, \dots, l_n, j}(t)$  da relação nebulosa ( $l_i = 1, \dots, c_i$  e  $j = 1, \dots, c_0$ ) é dado por:

$$R'_{l_1, \dots, l_n, j}(t) = \bigwedge (x_{1_{l_1}}(t), \dots, x_{n_{l_n}}(t), y_j(t))$$

Para uma seqüência de  $N$  conjuntos de dados de entrada e saída, a relação nebulosa resultante é obtida agregando-se, através de um operador de agregação, cada uma das relações individuais. Para a composição max-t, o operador união ( $\cup$ ) é utilizado obtendo-se

$$R(t) = \bigcup_{k=t-N+1}^t R'(k)$$



Como a operação de união de conjuntos e de relações nebulosas é realizada através do operador  $\max$  ( $\vee$ ), cada elemento de  $R$  é calculado como

$$R_{l_1, \dots, l_n, j}(t) = \bigvee_{k=t-N+1}^t R'_{l_1, \dots, l_n, j}(k)$$

Para aplicações “on-line” a seguinte forma recursiva pode ser utilizada:

$$R(t) = R'(t-1) \cup R'(t)$$

ou explicitamente

$$R_{l_1, \dots, l_n, j}(t) = \bigvee (R'_{l_1, \dots, l_n, j}(t-1), R'_{l_1, \dots, l_n, j}(t))$$

## Apêndice B: Algoritmo $\Sigma$ -PAFIO “On-line”

Uma versão “on-line” do algoritmo  $\Sigma$ -PAFIO para otimização de interfaces de entrada e saída de sistemas nebulosos é apresentada como segue. Inicialmente, deve-se executar os seguintes passos:

- 1) Selecionar a quantidade  $m$  de conjuntos nebulosos de referência (usualmente entre 3 e 9) em função da complexidade do sistema, e o tipo das funções de pertinência  $f_i$  (Gaussiana, Triangular, Senoidal, etc).
- 2) Selecionar a norma  $\rho$  para medida “sigma-count”. Um bom valor empírico sugerido por Oliveira é  $\rho = m - 1$  [Oliveira e Lemos, 1995].
- 3) Inicializar os centros e aberturas dos conjuntos nebulosos, ou seja,  $\theta_i = [\theta_{1i}, \theta_{2i}]$  para  $i = 1, \dots, m$ , segundo uma distribuição uniforme desses conjuntos nas interfaces, com centros equidistantes e aberturas equivalentes à distância entre dois centros consecutivos.
- 4) Escolher a taxa de aprendizado  $\eta \in (0, 1)$  para o método do gradiente e a constante de distinção  $K > 0$ .

Então, para cada iteração ou instante de amostragem  $k$ , executam-se os dois passos seguintes:

- 5) Verificar se a seguinte condição de parada do algoritmo é satisfeita:

$$E(k) < \lambda$$

onde  $\lambda$  é um escalar positivo e  $E(k)$  é uma filtragem da sequência dos erros de fuzzificação-defuzzificação (relativos ao critério de equivalência de informações) como segue:

$$E(k) = \gamma E(k-1) + (1-\gamma)e(k)$$

sendo  $\gamma \in [0, 1)$  o pólo do filtro e  $e(k)$  uma média temporal absoluta de uma janela  $H > 0$  da sequência de erros dada por

$$e(k) = \frac{1}{H} \sum_{t=k-H+1}^k |a(t) - \tilde{a}(t)|$$

A filtragem é utilizada para que um sinal com dinâmicas predominantemente lentas seja comparado ao valor de  $\lambda$ , evitando sucessivos acionamentos e interrupções do algoritmo. Embora a escolha dos parâmetros não seja crítica, bons valores empíricos são  $H = 10$  e  $\gamma = 0,9$ .

- 6) Caso a condição do passo anterior não seja verificada: Para o valor  $a(k)$  presente na entrada da interface (por simplicidade, denominado apenas  $a$ ), deve-se atualizar os parâmetros  $\theta_i$  ( $i = 1, \dots, m$ ) de acordo com a equação

$$\theta_i(k+1) = \theta_i(k) + \eta \Delta\theta_i$$

onde  $\Delta\theta_i = -\frac{\partial \tilde{J}}{\partial \theta_i}$  é o  $i$ -ésimo elemento do gradiente negativo da função de custo  $\tilde{J}$  que é dada por

$$\tilde{J} = J + J_1$$

onde

$$J = \frac{1}{2}(a - \tilde{a})^2$$

sendo  $\tilde{a} = \tilde{a}(k)$  obtido pelo método do centro de gravidade, ou seja,

$$\tilde{a} = \frac{\sum_{l=1}^m f_l(a, \theta_l) \theta_{1l}}{\sum_{l=1}^m f_l(a, \theta_l)}$$

e

$$J_1 = \frac{1}{2} K (M_\rho(\mathcal{L}(a)) - 1)^2 \mathcal{F} (M_\rho(\mathcal{L}(a)) - 1)$$

que representa a restrição de distinção  $M_\rho(\mathcal{L}(a)) \leq 1$  como função de penalidade (ver Seção 3.3.2). A medida "sigma-count" é dada por

$$M_\rho(\mathcal{L}(a)) = \sqrt{\rho \sum_{l=1}^m f_l^\rho(a, \theta_l)}$$

Para obter  $\Delta\theta_i$  ( $i = 1, \dots, m$ ) calcula-se

$$\frac{\partial \tilde{J}}{\partial \theta_i} = \frac{\partial J}{\partial \theta_i} + \frac{\partial J_1}{\partial \theta_i}$$

O desenvolvimento dessas duas parcelas é análogo ao realizado na Seção 3.3.2, porém para apenas uma amostra ( $N = 1$ ) de  $a$ . Assim tem-se

$$\frac{\partial J}{\partial \theta_i} = -(a - \tilde{a}) \frac{\partial \tilde{a}}{\partial \theta_i}$$

onde  $\frac{\partial \tilde{a}}{\partial \theta_i} = \left[ \frac{\partial \tilde{a}}{\partial \theta_{1i}} \quad \frac{\partial \tilde{a}}{\partial \theta_{2i}} \right]$ , sendo

$$\frac{\partial \tilde{a}}{\partial \theta_{1i}} = \frac{f_i(a, \theta_i) + (\theta_{1i} - \tilde{a}) \frac{\partial f_i(a, \theta_i)}{\partial \theta_{1i}}}{\sum_{l=1}^m f_l(a, \theta_l)}$$

e

$$\frac{\partial \tilde{a}}{\partial \theta_{2_i}} = \frac{(\theta_{1_i} - \tilde{a}) \frac{\partial f_i(a, \theta_i)}{\partial \theta_{2_i}}}{\sum_{l=1}^m f_l(a, \theta_l)}$$

e ainda

$$\begin{aligned} \frac{\partial J_1}{\partial \theta_i} = & K \mathcal{F}(M_\rho(\mathcal{L}(a)) - 1)(M_\rho(\mathcal{L}(a)) - 1) \left( \sum_{l=1}^m f_l^\rho(a, \theta_l) \right)^{\frac{1}{\rho} - 1} \\ & \cdot f_i^{\rho-1}(a, \theta_i) \frac{\partial f_i(a, \theta_i)}{\partial \theta_i} \end{aligned}$$

O termo  $\frac{\partial f_i(a, \theta_i)}{\partial \theta_i} = \left[ \frac{\partial f_i(a, \theta_i)}{\partial \theta_{1_i}} \frac{\partial f_i(a, \theta_i)}{\partial \theta_{2_i}} \right]$  pode ser calculado por derivação numérica, independentemente do tipo de função de pertinência utilizado.

## Apêndice C: Definições Básicas

Sejam  $A, A^1, A^2$  e  $B$ , conjuntos nebulosos definidos respectivamente nos universos de discurso  $\mathbf{A}$  e  $\mathbf{B}$ . Sejam também  $R, R^1$  e  $R^2$  relações nebulosas definidas no espaço do Produto Cartesiano de  $\mathbf{A}$  e  $\mathbf{B}$ . Alguns conceitos básicos em lógica e conjuntos nebulosos são apresentados como segue:

- **Suporte:** O suporte de um conjunto nebuloso  $A$  é o conjunto dos pontos  $a \in \mathbf{A}$  tais que a função de pertinência de  $A$  possui valor positivo ( $A(a) > 0$ ).
- **Singleton nebuloso:** O Singleton nebuloso é um conjunto nebuloso cujo suporte é um único ponto do universo de discurso.
- **Normas triangulares:** Uma norma triangular (norma-t) é uma função de dois argumentos  $t : [0, 1] \times [0, 1] \rightarrow [0, 1]$  monotônica, comutativa, associativa que satisfaz um conjunto de condições limitantes, isto é, sendo  $\alpha, \phi, \varphi$  e  $\xi \in [0, 1]$ ,
  1. Para  $\alpha \leq \varphi$  e  $\phi \leq \xi$ :  $\alpha t \phi \leq \varphi t \xi$ .
  2.  $\alpha t \phi = \phi t \alpha$ .
  3.  $\alpha t (\phi t \varphi) = (\alpha t \phi) t \varphi$ .
  4.  $\alpha t 0 = 0$  ;  $\alpha t 1 = \alpha$ .
- **Operadores min e max:** O mínimo e o máximo entre dois elementos  $a$  e  $b$  são definidos respectivamente através dos operadores min ( $\wedge$ ) e max ( $\vee$ ) como segue:

$$a \wedge b = \min(a, b) = \begin{cases} a & \text{se } a \leq b \\ b & \text{se } a > b \end{cases}$$
$$a \vee b = \max(a, b) = \begin{cases} a & \text{se } a \geq b \\ b & \text{se } a < b \end{cases}$$

Para um conjunto  $a_i$  ( $i = 1, \dots, N$ ) de elementos tem-se

$$\bigwedge_{i=1}^N a_i = a_1 \wedge a_2 \wedge \dots \wedge a_N = \min(a_1, a_2, \dots, a_N)$$

$$\bigvee_{i=1}^N a_i = a_1 \vee a_2 \vee \dots \vee a_N = \max(a_1, a_2, \dots, a_N)$$

O mínimo e o máximo entre dois conjuntos nebulosos  $A^1$  e  $A^2$  resultam em um terceiro conjunto nebuloso  $A$ , dado respectivamente por

$$A(a) = A^1(a) \wedge A^2(a) \quad \forall a \in \mathbf{A}$$

$$A(a) = A^1(a) \vee A^2(a) \quad \forall a \in \mathbf{A}$$

O mínimo e o máximo entre duas relações nebulosas  $R^1$  e  $R^2$  resultam em uma terceira relação nebulosa  $R$ , dada respectivamente por

$$R(a, b) = R^1(a, b) \wedge R^2(a, b) \quad \forall a \in \mathbf{A} \text{ e } b \in \mathbf{B}$$

$$R(a, b) = R^1(a, b) \vee R^2(a, b) \quad \forall a \in \mathbf{A} \text{ e } b \in \mathbf{B}$$

- **Valor modal:** O valor modal de um conjunto nebuloso  $A$  é o ponto  $a_0 \in \mathbf{A}$  em que a função de pertinência de  $A$  assume seu valor máximo.
- **Conjuntos nebulosos convexos:** Um conjunto nebuloso  $A$  é convexo se e somente se a seguinte condição for satisfeita:

$$A(\lambda a_1 + (1 - \lambda)a_2) \geq \min(A(a_1), A(a_2))$$

para todo  $a_1$  e  $a_2$  em  $\mathbf{A}$  e todo  $\lambda \in [0, 1]$ . Essa definição não implica que  $A(a)$  é uma função convexa.

- **Conjuntos nebulosos normais:** Um conjunto nebuloso  $A$  é normal se e somente se a seguinte condição for satisfeita:

$$\bigvee_{a \in \mathbf{A}} A(a) = 1$$

- **Conjuntos nebulosos normais à esquerda e à direita:** Um conjunto nebuloso  $A$  é normal à esquerda (direita) se e somente se existe  $a_0 \in \mathbf{A}$  tal que para todo  $a > a_0$  ( $a < a_0$ ) tem-se  $A(a) < 1$  e para todo  $a \leq a_0$  ( $a \geq a_0$ ) tem-se  $A(a) = 1$ .
- **Produto Cartesiano:** O Produto Cartesiano de dois conjuntos nebulosos  $A$  e  $B$  resulta em um conjunto nebuloso de duas dimensões  $R = A \times B$  dado por

$$R(a, b) = A(a) \, t \, B(b) \quad \forall a \in \mathbf{A} \text{ e } b \in \mathbf{B}$$

onde  $t$  é uma norma triangular.

- **Composição max-t:** A composição max-t entre um conjunto nebuloso  $A$  e uma relação nebulosa  $R$  resulta em um conjunto nebuloso  $B = A \bullet R$  dado por

$$B(b) = \bigvee_{a \in \mathbf{A}} A(a) \, t \, R(a, b)$$

onde  $t$  é uma norma triangular associada à composição.



# Bibliografia

- [Campello et al., 1997] Campello, R. J. G. B., Nazzetta, R. M., and Amaral, W. C. (1997). A new methodology for fuzzy model identification. In *Proc. 7th IFSA World Congress, Prague/Czech Republic* - (a ser publicado).
- [Czogala e Pedrycz, 1981] Czogala, E. and Pedrycz, W. (1981). On identification in fuzzy systems and its applications in control problems. *Fuzzy Sets and Systems*, 6:73–83.
- [Czogala e Pedrycz, 1982] Czogala, E. and Pedrycz, W. (1982). Control problems in fuzzy systems. *Fuzzy Sets and Systems*, 7:257–273.
- [Gomide e Gudwin, 1994] Gomide, F. A. C. e Gudwin, R. R. (1994). Modelagem, controle, sistemas e lógica fuzzy. *Controle & Automação*, 4:97–114.
- [Gottwald e Pedrycz, 1986] Gottwald, S. and Pedrycz, W. (1986). Solvability of fuzzy relational equations and manipulation of fuzzy data. *Fuzzy Sets and Systems*, 18:45–65.
- [Graham e Newell, 1988] Graham, B. P. and Newell, R. B. (1988). Fuzzy identification and control of a liquid level rig. *Fuzzy Sets and Systems*, 26:255–273.
- [Graham e Newell, 1989] Graham, B. P. and Newell, R. B. (1989). Fuzzy adaptive control of a first-order process. *Fuzzy Sets and Systems*, 31:47–65.
- [Higashi e Klir, 1984a] Higashi, M. and Klir, G. J. (1984a). Identification of fuzzy relation systems. *IEEE Trans. Systems, Man and Cybernetics*, SMC-14:349–355.
- [Higashi e Klir, 1984b] Higashi, M. and Klir, G. J. (1984b). Resolution of finite fuzzy relation equations. *Fuzzy Sets and Systems*, 13:65–82.

- [Kaufmann, 1975] Kaufmann, A. (1975). *Introduction to the Theory of Fuzzy Subsets*, volume 1. Academic Press.
- [Kroll, 1996] Kroll, A. (1996). Identification of functional fuzzy models using multidimensional reference fuzzy sets. *Fuzzy Sets and Systems*, 80:149–158.
- [Lee et al., 1994] Lee, Y. C., Hwang, C., and Shih, Y. P. (1994). A combined approach to fuzzy model identification. *IEEE Trans. Systems, Man and Cybernetics*, 24:736–743.
- [Ljung, 1987] Ljung, L. (1987). *System Identification: Theory for the user*. Prentice Hall.
- [Di Nola et al., 1991] Di Nola, A., Pedrycz, W., Sessa, S., and Sanchez, E. (1991). Fuzzy relation equations theory as a basis of fuzzy modelling: An overview. *Fuzzy Sets and Systems*, 40:415–429.
- [Oliveira, 1993] Oliveira, J. V. (1993). On optimal fuzzy systems I/O interfaces. In *Proc. 2nd IEEE Internat. Conference on Fuzzy Systems*, pages 851–856, San Francisco/USA.
- [Oliveira, 1995] Oliveira, J. V. (1995). A design methodology for fuzzy system interfaces. *IEEE Trans. Fuzzy Systems*, 3:404–414.
- [Oliveira e Lemos, 1993] Oliveira, J. V. and Lemos, J. M. (1993). System modelling and fuzzy relational identification. In *Proc. 2nd IEEE Internat. Conference on Fuzzy Systems*, pages 1074–1078, San Francisco/USA.
- [Oliveira e Lemos, 1995] Oliveira, J. V. and Lemos, J. M. (1995). Long-range predictive adaptive fuzzy relational control. *Fuzzy Sets and Systems*, 70:337–357.
- [Pedrycz, 1981] Pedrycz, W. (1981). An approach to the analysis of fuzzy systems. *Int. J. Control*, 34:403–421.
- [Pedrycz, 1983] Pedrycz, W. (1983). Numerical and applicational aspects of fuzzy relational equations. *Fuzzy Sets and Systems*, 11:1–18.
- [Pedrycz, 1984a] Pedrycz, W. (1984a). An identification algorithm in fuzzy relational systems. *Fuzzy Sets and Systems*, 13:153–167.

- [Pedrycz, 1984b] Pedrycz, W. (1984b). Identification in fuzzy systems. *IEEE Trans. Systems, Man and Cybernetics*, SMC-14:361–366.
- [Pedrycz, 1985a] Pedrycz, W. (1985a). Applications of fuzzy relational equations for methods of reasoning in presence of fuzzy data. *Fuzzy Sets and Systems*, 16:163–175.
- [Pedrycz, 1985b] Pedrycz, W. (1985b). On generalized fuzzy relational equations and their applications. *J. Mathematical Analysis and Applications*, 107:520–536.
- [Pedrycz, 1988] Pedrycz, W. (1988). Approximate solutions of fuzzy relational equations. *Fuzzy Sets and Systems*, 28:183–202.
- [Pedrycz, 1993] Pedrycz, W. (1993). *Fuzzy Control and Fuzzy Systems*. Research Studies Press/John Wiley & Sons, 2nd edition.
- [Pedrycz, 1995] Pedrycz, W. (1995). *Fuzzy Sets Engineering*. CRC Press.
- [Pedrycz et al., 1984] Pedrycz, W., Czogała, E., and Hirota, K. (1984). Some remarks on the identification problem in fuzzy systems. *Fuzzy Sets and Systems*, 12:185–189.
- [Pedrycz e Oliveira, 1993] Pedrycz, W. and Oliveira, J. V. (1993). Optimization of fuzzy relational models. In *Proc. 5th IFSA World Congress*, pages 1187–1190, Seoul/Korea.
- [Pedrycz e Oliveira, 1996] Pedrycz, W. and Oliveira, J. V. (1996). An algorithmic framework for development and optimization of fuzzy models. *Fuzzy Sets and Systems*, 80:37–55.
- [Sanchez, 1976] Sanchez, E. (1976). Resolution of composite fuzzy relation equations. *Information and Control*, 30:38–48.
- [Söderström e Stoica, 1989] Söderström, T. and Stoica, P. (1989). *System Identification*. Prentice Hall.
- [Sugeno e Kang, 1986] Sugeno, M. and Kang, G. T. (1986). Fuzzy modelling and control of multilayer incinerator. *Fuzzy Sets and Systems*, 18:329–346.
- [Sugeno e Kang, 1988] Sugeno, M. and Kang, G. T. (1988). Structure identification of fuzzy model. *Fuzzy Sets and Systems*, 28:15–33.

- [Sugeno e Tanaka, 1991] Sugeno, M. and Tanaka, K. (1991). Successive identification of a fuzzy model and its applications to prediction of a complex system. *Fuzzy Sets and Systems*, 42:315-334.
- [Sugeno e Yasukawa, 1993] Sugeno, M. and Yasukawa, T. (1993). A fuzzy-logic-based approach to qualitative modeling. *IEEE Trans. Fuzzy Systems*, 1:7-31.
- [Takagi e Sugeno, 1985] Takagi, T. and Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Systems, Man and Cybernetics*, SMC-15:116-132.
- [Tanscheit, 1992] Tanscheit, R. (1992). Controle nebuloso. 9º Congresso Brasileiro de Automática - Mini-Cursos, páginas 82-95, Vitória-ES/Brasil.
- [Tong, 1978] Tong, R. M. (1978). Synthesis of fuzzy models for industrial processes-some recent results. *Int. J. General Systems*, 4:143-162.
- [Xu, 1989] Xu, C. W. (1989). Fuzzy systems identification. *IEE Proceedings*, 136 - Pt. D:146-150.
- [Xu e Lu, 1987] Xu, C. W. and Lu, Y. Z. (1987). Fuzzy model identification and self-learning for dynamic systems. *IEEE Trans. Systems, Man and Cybernetics*, SMC-17:683-689.
- [Zadeh, 1965] Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8:338-353.
- [Zadeh, 1973] Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Systems, Man and Cybernetics*, SMC-3:28-44.
- [Zadeh, 1978] Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3-28.