

C2

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação

Modelagem e Controle Preditivo utilizando Multimodelos

Autor: Jeremias Barbosa Machado

Orientador: Prof. Dr. Wagner Caradori do Amaral

**Co-orientador: Prof. Dr. Ricardo José Gabrielli Barreto
Campello**

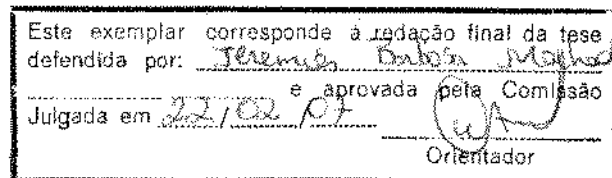
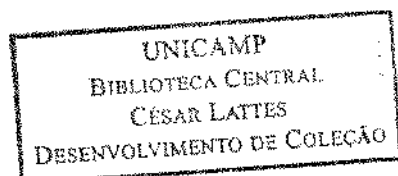
Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: **Automação**.

Banca Examinadora

Carlos Alberto Murari Pinheiro, Dr. IESTI/UNIFEI
Fernando Antônio Campos Gomide, Ph.D. DCA/FEEC/Unicamp

Campinas, SP

Fevereiro/2007



UNIDADE BC
Nº CHAMADA: _____
T/UNICAMP M18m
V. _____ EX. _____
TOMBO BCCL 73613
PROC 16.8.045.07
C _____
PREÇO 11.00
DATA 2/8/07
BIB-ID 423422

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

M18m Machado, Jeremias Barbosa
Modelagem e controle preditivo utilizando multimodelos
/ Jeremias Barbosa Machado. – Campinas, SP: [s.n.], 2007.

Orientadores: Wagner Caradori do Amaral; Ricardo José Gabrielli Barreto Campello.

Dissertação (Mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Sistemas de controle. 2. Controle preditivo. 3. Identificação de sistemas. 4. Lógica difusa. 5. Sistemas não Lineares. I. Amaral, Wagner Caradori do. II. Campello, Ricardo José Gabrielli Barreto. III Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. IV. Título.

Título em Inglês: Modeling and Predictive Control using Multi-models

Palavras-chave em Inglês: Nonlinear predictive control, Orthonormal base functions, Takagi-Sugeno fuzzy model, System identification and control, Nonlinear systems

Área de concentração: Automação

Titulação: Mestre em Engenharia Elétrica

Banca examinadora: Carlos Alberto Murari Pinheiro, Fernando Antônio Campos Gomide

Data da defesa: 22/02/2007

Programa de Pós-Graduação: Engenharia Elétrica

Aos meus pais Mário e Terezinha

000734558

Agradecimentos

A Deus pela vida e pela iluminação na execução deste trabalho.

Ao Professor Wagner Caradori do Amaral pela orientação, amizade e confiança demonstrados nestes dois anos de mestrado.

Ao Professor Ricardo J. G. B. Campello por coorientar meus estudos e pela amizade e dedicação neste período.

Aos meus pais Mário e Terezinha, e meu irmão Juliano pelo apoio, amor e incentivo durante esta jornada.

À minha namorada Andrezza pelo amor, carinho e compreensão durante estes dois anos.

Ao professor Carlos Alberto M. Pinheiro pela indicação e ajuda para que eu chegasse aqui.

Aos amigos Luis Henrique e Filipe pela amizade e por me acolherem nesta até então desconhecida cidade.

Aos amigos Adler, Leandro, Renato e tantos outros que fiz no LCA e na FEEC neste período.

A todos aqueles que de maneira direta ou indireta me ajudaram, ensinaram e participaram da minha vida na caminhada até aqui.

Ao CNPq, pelo apoio financeiro, com o qual foi possível a dedicação integral na elaboração deste trabalho.

“Toda a nossa ciência, comparada com a realidade, é primitiva e infantil - e, no entanto, é a coisa mais preciosa que temos.”

Albert Einstein

Resumo

O interesse na utilização de algoritmos de controle sofisticados cresce no meio industrial devido à necessidade de melhor qualidade dos produtos produzidos. Uma abordagem que vem ganhando destaque é a utilização de sistemas de controle não-linear que modelam os sistemas por meio de multimodelos lineares. Neste contexto, este trabalho apresenta a modelagem e controle de sistemas não-lineares através de controladores preditivos não-lineares que utilizam multimodelos lineares.

Os controladores preditivos baseados em modelos (*MBPC - Model Based Predictive Controllers*) são controladores cuja principal característica é a utilização de um modelo na determinação de um conjunto de previsões de saída, e a lei de controle é calculada em função destas previsões minimizando-se uma função de custo. O desempenho deste controlador depende da qualidade do modelo utilizado para predição dos sinais de saída. A proposta do trabalho é modelar as não-linearidades do processo sob controle através de modelos *fuzzy* Takagi-Sugeno - TS com funções de base ortonormal - FBO nos consequentes das regras. As FBO's apresentam diversas características conceituais e estruturais de interesse na elaboração dos modelos utilizados nos controladores preditivos, como a ausência de realimentação de saída, o que evita a propagação de erro, além de outras que serão discutidas ao longo deste trabalho.

Os parâmetros de um modelo *fuzzy* TS a serem determinados são os antecedentes das regras, com suas funções de pertinência, e as funções nos consequentes das regras, que neste trabalho dar-se-ão de forma automática, sendo os antecedentes das regras obtidos através de agrupamento *fuzzy* (*fuzzy clustering*) das amostras de entrada e saída. Para esta tarefa será utilizado o algoritmo de Gustafson-Kessel. A fim de determinar o número de grupos que irão compor o modelo e, por consequência, definir o número de regras e modelos locais, utilizar-se-ão critérios que avaliam a qualidade dos agrupamentos *fuzzy*, como *Fuzzy Silhouette*, *Fuzzy Hipervolume*, *Average Partition Density* e *Average Within-Cluster Distance*, sendo proposta a combinação dos resultados obtidos em cada um dos critérios.

O controle é feito de forma que, para cada modelo local, presente no modelo *fuzzy* TS-FBO, tem-se um controlador atuando sobre este. As ações de controle locais são combinadas conforme a ativação de cada regra do respectivo modelo local, e a ação de controle global resultante dessa combinação é aplicada ao processo a ser controlado.

A abordagem proposta apresenta vantagens estruturais na modelagem e controle de processos não-lineares, quando comparado a outras metodologias de modelagem (como modelos polinomiais NARMAX) e controle, uma vez que esta abordagem é composta de uma estrutura simples com modelos locais lineares (ou afins) formados por FBO's. Para ilustrar o que foi desenvolvido, são apresentadas, no final deste trabalho, implementações na modelagem e controle de processos não-lineares.

Palavras-chave: Controle Preditivo Não-Linear, Funções de Base Ortonormais, Modelos *Fuzzy* Takagi-Sugeno, Identificação e Controle de Sistemas Dinâmicos, Sistemas Não-lineares.

Abstract

The use of advanced control strategies has been increased in the last years due to the needs of more accurate quality on products. An approach that seems attractive on control and modeling of the nonlinear processes is the use of multiple linear models. In this context, this work presents an alternative approach for modeling and controlling nonlinear processes through nonlinear predictive control (NMBPC) using multi-models.

The main characteristic of the Model Based Predictive Controllers is the use of a model for the determination of the output predictions. The control law is derived based on these output predictions, minimizing a specified cost function. Its performance is directly related to the quality of the model predictor. Therefore, in this work, the process is modeling through Takagi-Sugeno-TS fuzzy models with orthonormal base functions - OBF - on the rules consequents. OBF's models present several conceptual and structural characteristics of interest on the elaboration of models predictors, such as, absence of output recursion and feedback of prediction errors, often leading to superior performances over long-range horizon predictions and natural decoupling between multiple outputs; there is no need for previous knowledge about the relevant past terms of the system signals; the representation of a stable system is assuredly stable; tolerance to unmodeled dynamics; ability to deal with time delays. The antecedents of the TS fuzzy models are obtained through fuzzy clustering of the input and output measures. The algorithm of Gustafson-Kessel is used to perform this task. In order to determine the number of the local models, clustering validity criteria such as Fuzzy Silhouette, Fuzzy Hipervolume, Average Partition Density e Average Within-Cluster Distance are used.

A predictive controller is derived for local model and the global control law is obtained by combining each local control law, using the degree of activation of every rule of the respective local model. The proposed approach presents structural advantages in the modeling and controlling nonlinear process, when compared to other modeling (like polynomial models-NARMAX) and controlling strategies, as this approach is constituted of a simple structure with linear local models using OBF's. The performance of the proposed strategies is illustrated using some simulated examples.

Keywords: Nonlinear Predictive Control, Orthonormal Base Functions, Takagi-Sugeno Fuzzy Model, System Identification and Control, Nonlinear Systems.

Sumário

Lista de Figuras	xv
Lista de Tabelas	xvii
Lista de Siglas e Abreviações	xix
1 Introdução	1
2 Modelagem e Identificação de Sistemas Dinâmicos	7
2.1 Modelagem Matemática	7
2.1.1 Equações Diferenciais	8
2.1.2 Funções de Transferência	9
2.1.3 Representação Discreta Geral	11
2.1.4 Espaço de Estados	14
2.1.5 Funções de Base Ortonormal	16
2.1.6 Modelos Não-lineares	19
2.1.7 Identificação de sistemas dinâmicos	23
2.2 Resumo	24
3 Sistemas <i>Fuzzy</i>	25
3.1 Introdução aos conjuntos <i>Fuzzy</i>	25
3.2 Modelo Relacional	28
3.3 Modelo Mamdani	29
3.3.1 Modelos <i>Singleton</i>	30
3.4 Modelo Takagi-Sugeno	30
3.4.1 Inferência em Modelos Takagi-Sugeno	32
3.5 Agrupamento	35
3.5.1 Algoritmo de Gustafson-Kessel	36
3.5.2 Métodos de determinação do número de grupos	37
3.6 Construção dos modelos <i>Fuzzy</i>	41
3.6.1 Determinação dos antecedentes das Regras	41
3.6.2 Determinação dos consequentes das Regras com FBO's	43
3.6.3 Exemplos	46
3.7 Resumo	55

4	Controle Preditivo Não-linear	57
4.1	O Controlador Preditivo	57
4.2	Modelagem do Processo e Previsão da Saída	59
4.3	Lei de Controle Preditivo Linear	62
4.3.1	Controle Preditivo Linear com FBO	64
4.4	Controlador Preditivo Não-linear	65
4.5	Resumo	67
5	Resultados e Simulações	69
5.1	Modelagem de um levitador magnético	69
5.2	Controle de Sistemas Não-lineares	76
5.2.1	Controle de sistema CSTR	76
5.3	Resumo	86
6	Conclusões	87
6.1	Contribuições	88
6.2	Trabalhos Futuros	89
	Referências bibliográficas	90
A	Método de Mínimos Quadrados	97
B	Teoria de Conjuntos <i>Fuzzy</i>	99
B.1	Conjuntos <i>Fuzzy</i>	99
B.1.1	Representação do conjunto <i>fuzzy</i>	101
B.1.2	Operações em conjuntos <i>Fuzzy</i>	101
B.2	Raciocínio Aproximado	103
B.2.1	Variáveis Lingüísticas	103
B.2.2	Proposição <i>Fuzzy</i>	104
B.2.3	Regras Lingüísticas	105

Lista de Figuras

2.1	Sistema a ser modelado.	9
2.2	(a) - Plano s ; (b) - Plano z	11
2.3	Representação Geral de um Sistema Linear.	12
2.4	Diagrama de blocos do modelo FBO	17
3.1	Funções de pertinência trapezoidais	26
3.2	Modelo relacional <i>fuzzy</i>	28
3.3	Intersecção de dois conjuntos <i>fuzzy</i> A_1 e A_2 (em escuro).	32
3.4	Inferência <i>fuzzy</i> TS com operador prod. para conjunção.	33
3.5	Agrupamento dos dados por GK.	42
3.6	Funções de pertinência geradas a partir das projeções.	43
3.7	Conjunto de amostras com 9 grupos.	46
3.8	Crítérios de avaliação do número ideal de grupos (não normalizados)	47
3.9	Resultado global do número de grupos via combinação por produto e média.	47
3.10	Grupos obtidos com o algoritmo de GK.	48
3.11	Conjunto de amostras com 3 grupos.	48
3.12	Métodos de avaliação do número de grupos (normalizados)	49
3.13	Média e Produto dos critérios de avaliação.	49
3.14	Grupos determinados pelo algoritmo de GK.	50
3.15	Conjunto de amostras com 3 grupos formando um "A".	50
3.16	Média e Produto dos critérios de avaliação do núm. de grupos.	51
3.17	Resultado do agrupamento dado pelo algoritmo de GK.	51
3.18	Dados de identificação contaminados com ruído.	52
3.19	Resultado dado pela combinação dos critérios de análise de agrupamento.	52
3.20	Resultado do agrupamento dos objetos por GK.	53
3.21	Funções de pertinência dos antecedentes das regras.	53
3.22	Comparação entre os valores simulados e os estimados.	54
4.1	Diagrama de blocos do Controlador Preditivo	58
4.2	Sinais do Controlador Preditivo	62
4.3	Controlador preditivo proposto.	66
5.1	Levitador Magnético; e - emissor laser; s - sensor laser.	70
5.2	Dados de identificação do sistema de levitação.	71
5.3	Respostas medida do sistema e estimada pelo modelo único linear.	72

5.4	Média e Produto dos critérios de validade de agrupamento p/ dados do sistema de levitação magnética agrupados pelo algoritmo de GK.	73
5.5	Funções de pertinência presentes nos antecedentes das regras do modelo TS - levitador magnético.	73
5.6	Resposta medida do sistema e resposta estimada do modelo.	74
5.7	Resposta em malha aberta do sistema CSTR.	77
5.8	Resultado global da estimação do número de grupos (modelos locais) via combinação de critérios de validade por média e produto - sistema CSTR.	78
5.9	Funções de pertinência presentes nos antecedentes das regras do modelo TS - sistema CSTR.	78
5.10	Respostas do sistema e do modelo do CSTR (dados de validação).	79
5.11	Resposta CSTR sob ação do controlador.	80
5.12	Sinal de entrada e Resposta em malha aberta do sistema CSTR.	81
5.13	Resultado do produto e média dos critérios de avaliação do número de grupos - CSTR 1.	82
5.14	Funções de pertinência presentes nos antecedentes das regras do modelo TS-FBO que modela o CSTR.	83
5.15	Respostas do sistema e do modelo do CSTR (dados de validação).	83
5.16	Resposta do sistema CSTR sob ação do controlador proposto com <i>set point</i> de 0,08 [mol/l] - Preditivo <i>fuzzy</i> (-), PID (-), GPC (.), Preditivo TS-FBO (-).	85
5.17	Resposta do sistema CSTR sob ação do controlador proposto com <i>set point</i> de 0,12 [mol/l] - Preditivo <i>fuzzy</i> (-), PID (-), GPC (.), Preditivo TS-FBO (-).	85
B.1	Função Característica para o conj. $A = \{x x \in E^+ \text{ e } x > 3\}$	100
B.2	Funções de pertinência: a) Função Triangular b) Função Trapezoidal c) Função Gaussiana	101
B.3	Operações com conjuntos <i>fuzzy</i> : a) Complemento b) Intersecção c) União	103
B.4	Conjunto T (velocidade) com suas respectivas funções de pertinência	104

Lista de Tabelas

5.1	Comparação entre os EQM's para modelos TS-FBO do levitador com 1 e 2 variáveis nas premissas das regras <i>fuzzy</i> .	72
5.2	Comparação entre os EQM's para modelos TS-FBO do levitador com diferentes quantidades de estados de Kautz nos modelos locais	74
5.3	Comparação entre os EQM's para modelos TS-FBO do levitador com diferentes quantidades de modelos locais	75
5.4	Parâmetros do modelo CSTR	81
5.5	Comparação entre os EQM's para modelos TS-FBO do CSTR com diferentes pólos de Laguerre	84

Lista de Siglas e Abreviações

APD	- <i>Average Partition Density</i>
ARMA	- <i>AutoRegressive Moving Average</i>
ARMAX	- <i>AutoRegressive Moving Average with eXogenous inputs</i>
ARX	- <i>AutoRegressive with eXogenous inputs</i>
AWCD	- <i>Average Within-Cluster Distance</i>
CSTR	- <i>Continuous Stirred Tank Reactor</i>
DMC	- <i>Dynamic Matrix Control</i>
EQM	- Erro Quadrático Médio
FBO	- Função de Base Ortonormal
FHV	- <i>Fuzzy Hypervolume</i>
FIR	- <i>Finite Impulse Response</i>
FS	- <i>Fuzzy Silhouette</i>
FT	- Função de Transferência
GK	- Gustafson-Kessel
GPC	- <i>Generalized Predictive Control</i>
IIR	- <i>Infinite Impulse Response</i>
MBPC	- <i>Model Based Predictive Control</i>
MLP	- <i>MultiLayer Perceptron</i>
MQ	- Mínimos Quadrados
NARX	- <i>Nonlinear AutoRegressive with eXogenous inputs</i>
NARMAX	- <i>Nonlinear AutoRegressive Moving Average with eXogenous inputs</i>
NFIR	- <i>Nonlinear Finite Impulse Response</i>
NOBF	- <i>Nonlinear Orthonormal Base Function</i>
NOE	- <i>Nonlinear Output Error</i>
PID	- Proporcional+Integral+Derivativo
RBF	- <i>Radial Base Function</i>
RNA	- Rede Neural Artificial
TS	- Takagi-Sugeno

Lista de Símbolos

K	-	Número de regras <i>fuzzy</i> - modelos locais
d	-	Número de variáveis nos antecedentes das regras <i>fuzzy</i>
μ_i	-	Valor de ativação de uma regra <i>fuzzy</i> i
$\mu_{A_{i,j}}$	-	Valor de ativação de uma dada função de pertinência $A_{i,j}$
$\bar{\mu}_i$	-	Valor normalizado de Ativação de uma regra <i>fuzzy</i> i
$\mu_{i,j}$	-	Valor de ativação de um objeto i num dado grupo <i>fuzzy</i> j
N	-	Número de objetos utilizados para Identificação
F_i	-	Matriz de covariância do grupo i
σ	-	Desvio padrão
N_1	-	Horizonte inicial de previsão de saída
N_y	-	Horizonte final de previsão de saída
N_u	-	Horizonte de controle
$\hat{y}$	-	Saída estimada
y	-	Saída medida ou simulada
ω	-	Sinal de referência
Δu	-	Variação do sinal de controle
u	-	Sinal de controle
l_i	-	Estado i da representação em espaço de estados da FBO

Capítulo 1

Introdução

Nas últimas décadas, com os avanços obtidos na área de computação e eletrônica, tem sido possível suprir à demanda por sistemas de controle digital mais sofisticados que atendam às necessidades cada vez maiores dos meios de produção na busca por qualidade e maior eficiência dos processos produtivos. Muitas aplicações, onde se necessita fazer o controle do processo, podem ser atendidas por sistemas de controle clássico como controladores PID (Proporcional+Integral+Derivativo) de ganhos fixos ou controladores clássicos de realimentação de estados, dentre outros. Já algumas aplicações, como processo multivariáveis, não-lineares, variantes no tempo, de ordem elevada ou com especificações de projeto muito rígidas, exigem sistemas de controle digital com algoritmos complexos para atender às necessidades do processo e possibilitar um controle eficiente.

Sendo assim, dentre estes novos algoritmos para sistemas de controle mais elaborados, surgem no final da década de 70 os controladores preditivos baseados em modelo (*Model Based Predictive Control*- MBPC). Seu conceito inicial foi introduzido nos trabalhos de Richalet, Rault, Testud e Papon (1978) com o *Identification Command* (IDCOM) e de Cutler e Ramaker (1980) com o *Dynamic Matrix Controller* (DMC). Estes algoritmos estabeleceram a metodologia dos MBPC, partindo de uma representação do processo feita através de modelos de resposta ao impulso. Desde então, outras estratégias de controle preditivo, utilizando diferentes formas para a modelagem de processo, foram proposta na literatura (Clarke, Mohtadi e S., 1987; Clarke, 1994; Oliveira, 1997). No trabalho Oliveira (1997) são apresentadas propostas de controladores preditivos com modelos dados por funções de base ortonormal (FBO), os quais serão utilizados neste trabalho.

Em sua essência os controladores preditivos refletem o comportamento humano uma vez que é característica humana atuar de maneira antecipativa no controle de um dado processo que esteja controlando, como, por exemplo, quando está ao volante de um veículo. Para atuar de maneira antecipativa, o controlador preditivo se baseia na resposta de um modelo interno do processo em questão, onde as ações de controle são atualizadas constantemente conforme novas observações estão

disponíveis. Desta maneira, a lei de controle para este controlador terá os seguintes componentes:

1. Um modelo do processo - O modelo é utilizado para prever o comportamento do processo em um horizonte de tempo futuro, denominado horizonte de previsão;
2. Um critério de custo - O desempenho do sistema em malha fechada durante o horizonte de previsão é especificado através de um critério de custo, definido a partir da saída prevista, do sinal de referência e do esforço de controle;
3. Um processo de otimização do critério de custo - A ação de controle atual é determinada a partir da otimização do critério de custo definido;

Os controladores preditivos são capazes de controlar uma grande variedade de processos, desde aqueles com dinâmicas relativamente simples até mais complexos, incluindo sistemas com grandes atrasos de transporte, sistemas de fase não-mínima, instáveis ou sistemas multivariáveis. Eles compensam intrinsecamente os atrasos de transporte apresentando fácil implementação e custo computacional reduzido.

Nos algoritmos preditivos, a lei de controle e, conseqüentemente, o desempenho do sistema em malha fechada, está diretamente ligada à qualidade do modelo. Assim, um sistema de controle com modelo de baixa qualidade irá produzir um controlador ruim. Neste trabalho serão utilizadas funções de base ortonormal - FBO nos controladores preditivos para modelar os processos sob controle (Oliveira, 1997).

Os modelos FBO têm sido aplicados com sucesso nas tarefas de modelagem e controle de sistemas dinâmicos (Wiener, 1956; Wahlberg, 1991; Wahlberg, 1994; Campello e Amaral, 2002). Esses modelos apresentam características que os tornam atraentes na modelagem de sistemas dinâmicos:

- Ausência de recursão de saída, o que conseqüentemente leva a modelos mais precisos e evita a propagação de erro na predição de sinais;
- Não há necessidade de se conhecer previamente a estrutura exata do vetor de regressão (ordens e atrasos de transporte), podendo assim tratar atrasos de transporte de maneira explícita ou implícita;
- Se necessário, é possível aumentar a capacidade de representação do modelo aumentando o número de funções ortonormais utilizadas;
- Desacoplamento natural das saídas em sistemas multivariáveis;
- Tolerância a dinâmicas não modeladas;

- É garantido que a representação de um sistema estável é também estável.

Estas propriedades fazem da representação dinâmica em FBO's interessante quando se deseja modelar processos em sistemas de controles preditivos. Como não apresentam realimentação de sinal de saída no modelo, não propagam erro na predição gerando modelos mais precisos para serem aplicados em MBPC's.

Este trabalho procura desenvolver algoritmos para o controle preditivo de sistemas não-lineares através da utilização de uma abordagem de multimodelos locais. Para a modelagem destes sistemas utiliza-se modelos *fuzzy* Takagi-Sugeno com estruturas FBO nos conseqüentes das regras.

Os conceitos de conjuntos *fuzzy*, ou conjuntos nebulosos, foram originalmente propostos por Zadeh (1965) e com os trabalhos de Takagi e Sugeno (1985) foi introduzida uma nova metodologia de modelagem baseada em regras, munida de conjuntos *fuzzy*, cujos modelos são capazes de representar o comportamento de sistemas não-lineares graças a propriedade de serem aproximadores universais de funções em um espaço compacto (Wang e Mendel, 1992; Wang, 1998; Kosko, 1994). A estes modelos dá-se a denominação de modelos *fuzzy* Takagi-Sugeno (TS). Estes modelos consistem de interpolações de múltiplos modelos locais (usualmente lineares ou afins), o que favorece amplamente a interpretabilidade matemática e o projeto de sistemas de controle uma vez que localmente se trabalha com sistemas lineares.

Os modelos *fuzzy* TS utilizados nessa dissertação apresentam FBO's nos conseqüentes das regras. Portanto, esses modelos, denominados TS-FBO (Campello e Amaral, 2002; Campello, 2002; Medeiros, 2006; Medeiros, Amaral e Campello, 2006), dispõem de todas as vantagens destacadas anteriormente para as estruturas dinâmicas FBO. Outras propostas de aproximadores universais de funções utilizando FBO's são encontradas na literatura (envolvendo redes neurais artificiais, funções de base radial, etc); contudo, tais modelos não são matematicamente interpretáveis e, além disso, não são constituídos por modelos locais, como os modelos *fuzzy* TS.

Como o modelo TS é composto de regras, faz-se necessário determinar os antecedentes e os conseqüentes destas regras. A determinação dos antecedentes pode ser feita com conhecimento *a priori* do sistema, como a partir do conhecimento de um especialista, ou então de forma automática, a partir de medidas de entrada e saída do sistema. Neste trabalho a determinação dos antecedentes das regras dar-se-á de forma automática através de agrupamento *fuzzy* (*fuzzy clustering*) das amostras de entrada e saída. Para esta tarefa será utilizado o algoritmo de Gustafson-Kessel (Gustafson e Kessel, 1979). A determinação do número de grupos que melhor agrupa os dados se dá através de critérios de avaliação de agrupamentos *fuzzy* (Gath e Geva, 1989; Campello e Hruschka, 2006), dos quais alguns serão utilizados neste trabalho e terão resultados combinados a fim de determinar o número de modelos locais a compor o modelo TS-FBO. Uma vez determinados os antecedentes das regras, determina-se os parâmetros dos modelos locais através de estimação por mínimos quadrados

locais, que levam em consideração a ativação das regras de acordo com os dados de entrada.

Uma vez modelado o sistema através do modelo TS-FBO, a etapa de controle consiste em colocar controladores preditivos atuando localmente sobre os modelos locais. As ações de controle locais são combinadas de acordo com a ativação de cada regra do respectivo modelo local, e a ação de controle global resultante dessa combinação é aplicada ao processo a ser controlado.

Desta forma, no contexto apresentado nesta introdução, a metodologia de controle preditivo proposta se torna atrativa no controle de sistemas não-lineares graças a sua estrutura e às características apresentadas pelos modelos TS-FBO, uma vez que os mesmos são compostos por modelos lineares locais relativamente simples, com características importantes na modelagem de sistemas dinâmicos, além de interpretabilidade matemática. Tem-se ainda como contribuições, além da proposta do controlador, o algoritmo desenvolvido para determinar as funções de pertinência a partir da projeção da ativação das amostras no espaço de entrada, a modificação proposta para o critério de determinação do número de grupos *fuzzy* denominado *Average Within-Cluster Distance* - AWCD e a metodologia para síntese das avaliações dos diversos critérios em uma resposta global a fim de determinar o número ideal de grupos (modelos locais) que irá formar o modelo TS-FBO.

O conteúdo desta dissertação apresenta a seguinte estrutura:

Após esta introdução, o capítulo 2 apresenta os conceitos utilizados na modelagem de sistemas dinâmicos lineares e não-lineares, dando-se maior ênfase à representação de sistemas lineares por modelos de FBO. A seguir, serão apresentadas as abordagens mais utilizadas na modelagem de sistemas dinâmicos não-lineares e, posteriormente, o estimador de mínimos quadrados utilizado na identificação dos parâmetros do modelo. Maiores detalhes sobre os estimadores de Mínimos Quadrados podem ser obtidos no Apêndice A.

No capítulo 3 apresenta-se uma breve revisão da teoria de conjuntos e sistemas *fuzzy* (para maiores detalhes ver Apêndice B). Posteriormente, são apresentados os modelos mais utilizados, onde se apresenta de forma mais detalhada os modelos *fuzzy* Takagi-Sugeno - TS, e, em seguida, os modelos *fuzzy* TS-FBO, que são utilizados na implementação do sistema de modelagem e controle propostos. Também neste capítulo é introduzida a técnica de agrupamento proposto por Gustafson e Kessel (1979), utilizada aqui para se obter os antecedentes das regras do modelo TS. Ainda no contexto da determinação das regras, são apresentados alguns critérios de avaliação do melhor número de grupos para se representar o sistema sob análise, com contribuições nesta área. Além disso, são descritas duas variações do método de estimação de mínimos quadrados, estimação local e estimação global, utilizados na obtenção dos modelos locais presentes nos consequentes das regras.

O capítulo 4 inicialmente apresenta o controlador preditivo linear com o modelo de predição dado por funções de base ortonormal, o qual irá atuar de forma local sobre os modelos de cada regra. Por fim, é descrito como a ação de controle global, que atuará sobre o sistema não-linear, se dá a partir da

combinação *fuzzy* das ações de controle locais.

No capítulo 5 são apresentados exemplos de simulações que ilustram o desempenho do método proposto neste trabalho. Primeiramente, é apresentada a modelagem de um sistema de levitação magnética presente no laboratório de sistemas e controle desta instituição. Para ilustrar a aplicação do sistema de controle proposto foi feita uma aplicação (por simulação) no controle de uma planta de polimerização CSTR. Os resultados do modelo identificado e da resposta do sistema sob ação do controlador são apresentados e discutidos.

O Capítulo 6 traz as conclusões finais, uma revisão sobre as contribuições do trabalho e a perspectiva de trabalhos futuros como continuação deste trabalho.

Capítulo 2

Modelagem e Identificação de Sistemas Dinâmicos

Este capítulo tem como objetivo apresentar as formas mais utilizadas de se representar sistemas dinâmicos lineares e não-lineares através de modelos matemáticos. Outro tema abordado neste capítulo se refere às maneiras existentes para se determinar os modelos com seus parâmetros e estruturas, além de abordar as principais características dos métodos de identificação de sistemas dinâmicos.

Primeiramente serão apresentados modelos utilizados na modelagem de sistemas dinâmicos lineares contínuos e discretizados no tempo dando-se maior foco às Funções de Base Ortonormal que serão utilizadas como método de modelagem neste trabalho. Serão apresentados ainda métodos de modelagem não-lineares presentes na literatura e o estimador de mínimos quadrados, utilizado na identificação dos sistemas dinâmicos.

2.1 Modelagem Matemática

Modelos matemáticos são maneiras pelas quais se podem representar as características observadas de um dado sistema real. As maneiras de se representar um sistema real através de equações ou relações matemáticas são muitas, podendo haver algumas que melhor se adaptam em representar as características desejadas do sistema sob estudo. Os objetivos pelos quais se deseja obter tais modelos podem variar muito, desde tentar descrever o comportamento de um dado sistema social até modelar o funcionamento de uma complexa máquina. Nesta dissertação o objetivo principal é obter tais representações matemáticas com a finalidade de descrever o comportamento de sistemas dinâmicos e posteriormente utilizá-las no projeto de sistemas de controle.

Uma consideração importante a ser feita é que um modelo raramente consegue descrever de maneira exata o comportamento do sistema ao qual se refere o modelo. Tal característica dos modelos

matemáticos serem então representações aproximadas do sistema real leva à conclusão que muitas vezes não existe “o modelo” do sistema, mas sim conjuntos ou famílias de modelos que podem representar de maneira satisfatória as características desejadas. Ainda neste contexto, como geralmente não se consegue representar exatamente o comportamento do sistema, ou talvez não seja necessário representá-lo exatamente, pois exigiria a busca de modelos complexos, usa-se aproximações simplificadoras das características do sistema como, por exemplo, tratar o sistema sob estudo de forma mais linear possível, sendo tal linearidade em toda a área de operação do sistema ou somente em regiões restritas de operação. Assim, um sistema complexo e de comportamento não-linear pode ser representado por diversos modelos lineares onde cada um melhor representa o sistema em uma dada região de operação. Um sistema é dito linear quando obedece ao princípio de superposição (Aguirre, 2000). Embora as aproximações lineares possam ser utilizadas satisfatoriamente em uma ampla gama de aplicações práticas, muitos sistemas somente podem ser representados adequadamente através de modelos não-lineares. A seguir serão abordadas algumas das técnicas de modelagem mais utilizadas na área de controle.

2.1.1 Equações Diferenciais

O estudo de equações diferenciais tem atraído a atenção de muitos dos grandes matemáticos nos últimos quatro séculos. O estudo de equações diferenciais teve início com Isaac Newton¹ e Gottfried Wilhelm Leibniz² no século XVII. Muitos dos princípios e leis que descrevem o comportamento do mundo real são expressões ou relações que descrevem sistemas com comportamento que não respondem instantaneamente a perturbações. Quando essas relações são expressas em termos matemáticos, se tornam equações e as taxas com que respondem derivadas. Equações contendo tais derivadas são então chamadas de equações diferenciais. Para entender e investigar problemas envolvendo movimento de fluidos, fluxo de corrente elétrica em um circuito elétrico, entre outros, é necessário conhecer as equações diferenciais que regem o comportamento de tais processos, sendo essas equações os modelos matemáticos dos processos sob estudo.

¹Newton nasceu em 1642 e cresceu no interior da Inglaterra tendo ido estudar no Trinity College, Cambridge. Ele inicia suas descobertas do Cálculo e Leis Fundamentais da Mecânica a partir 1665, vindo a publicar seu famoso livro *Philosophiae Naturalis Principia Mathematica* em 1687. Ele foi nomeado Cavaleiro da Coroa Inglesa em 1705 e após sua morte em 1716 foi sepultado na Abadia de Westminster.

²Leibniz nasceu em Leipzig e completou seu doutorado em Filosofia aos 20 anos pela Universidade de Altdorf. Era auto-didata nos estudos de Matemática e chegou aos resultados fundamentais do cálculo independentemente dos estudos de Newton, vindo a publicá-los em 1684. Leibniz foi um grande conhecedor do poder de uma boa notação matemática e as notações para derivada, dy/dx , e o sinal de integral são devidos a ele. Veio a falecer assim como Newton no ano de 1716.

Uma equação diferencial de n -ésima ordem pode ser descrita da seguinte forma:

$$P_0(t) \frac{d^n y}{dt^n} + P_1(t) \frac{d^{n-1} y}{dt^{n-1}} + \dots + P_{n-1}(t) \frac{dy}{dt} + P_n y = G(t) \quad (2.1)$$

Assumindo que as funções $P_0, \dots, P_n$ e G são funções contínuas que assumem valor real no intervalo $I : \alpha < t < \beta$, e que P_0 é diferente de zero no intervalo, pode-se então, dividindo a equação (2.1) por P_0 e com o uso da notação $dy/dt = y'$, reescrever a equação (2.1) como:

$$y^{(n)} + p_1(t)y^{(n-1)} + \dots + p_{n-1}(t)y' + p_n(t)y = g(t) \quad (2.2)$$

A equação (2.2) envolve a n -ésima derivada de y com respeito a t . Assim essa equação é uma equação diferencial ordinária de ordem n . Se os parâmetros $p_i(t)$, com $i = 1, \dots, n$ apresentam valores constantes para todo t a equação (2.2) é chamada de equação diferencial linear invariante no tempo.

2.1.2 Funções de Transferência

Uma das mais utilizadas e importantes representações matemáticas de sistemas dinâmicos lineares e invariantes no tempo é a função de transferência (FT). Estas funções modelam o comportamento dinâmico entra/saída (Fig. 2.1) de um sistema dinâmico linear invariante no tempo

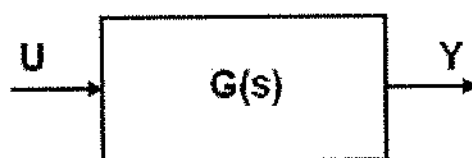


Fig. 2.1: Sistema a ser modelado.

A FT de um sistema é, por definição, a transformada de Laplace da resposta do sistema a uma entrada do tipo impulso sob condições iniciais nulas. A FT pode ser obtida dividindo-se a transformada de Laplace da saída pela transformada de Laplace da entrada, tornando assim uma FT representada pela razão de dois polinômios em s ($s = j\omega$, sendo ω uma dada frequência em rad/s), como pode ser observado na equação abaixo:

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_0 + b_1 s + \dots + b_n s^n}{a_0 + a_1 s + \dots + a_q s^p}$$

Os zeros de $G(s)$ são valores de s para os quais a função $G(s)$ se anula, sendo portanto as raízes do polinômio $Y(s)$. Assim como os zeros da FT os valores de s que anulam o polinômio $U(s)$ são conhecidos como os pólos do sistema. Para o sistema ser estável a parte real dos pólos devem apresentar somente valores negativos (Fig. 2.2 - (a)) (Chen, 1999). FT com pólos na origem representam sistemas com integradores puros, sendo estes marginalmente estáveis (Phillips e Harbor, 1999).

Muitas vezes se torna interessante ou até mesmo necessário que não se utilize modelos contínuos no tempo, mas sim sistemas discretizados com período de amostragem bem definido (definido em z , onde z é uma variável complexa utilizada na transformada discreta). A partir de uma FT contínua (em s) é possível se obter o equivalente discreto através de transformações, como por exemplo, a transformação Bilinear ou Tustin (Phillips e Nagle, 1995):

$$s = \frac{2}{T_s} \frac{1 - z^{-1}}{1 + z^{-1}} \quad (2.3)$$

Resolvendo para z a equação (2.3) tem-se:

$$z = \frac{(2/T_s) + s}{(2/T_s) - s} \quad (2.4)$$

onde T_s é o tempo de amostragem, o qual deve obedecer ao critério de Nyquist (Phillips e Nagle, 1995).

Desta forma a representação de uma FT discretizada será como a apresentada na equação (2.5).

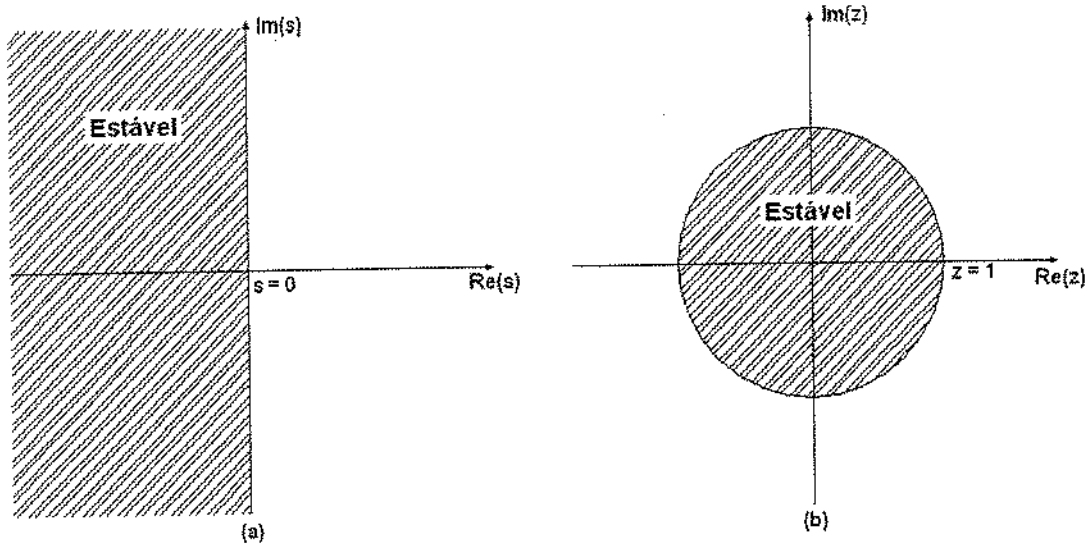
$$G(z) = \frac{Y(z)}{U(z)} = \frac{r_0 + r_1 z + \dots + r_{o_n} z^{o_n}}{d_0 + d_1 z + \dots + d_{o_d} z^{o_d}} \quad (2.5)$$

Esta representação é muito utilizada quando se deseja trabalhar com modelos em computadores ou sistemas microprocessados. Assim como para os sistemas contínuos, em que os sistemas estáveis são aqueles que têm pólos na região à esquerda do eixo das ordenadas (eixo $\text{Re}(s)$; Fig. 2.2 - (a)), os sistemas discretizados são estáveis quando possuem pólos dentro do círculo central unitário, $\|z\| \leq 1$. (Fig. 2.2 - (b))

Dividindo o denominador e o numerador da equação (2.5) por z^{o_d} e fazendo $z^{-i} = q^{-i}$ com $i = 1, \dots, o_d$, pode-se reescrever a equação (2.5) como sendo:

$$G(q) = \frac{Y(q)}{U(q)} = \frac{r_0 q^{-o_d} + r_1 q^{-(o_d+1)} + \dots + r_{o_n} q^{-(o_d-o_n)}}{d_0 q^{-o_d} + d_1 q^{-(o_d+1)} + \dots + d_{o_d}} \quad (2.6)$$

Para a utilização da representação discretizada em sistemas microprocessados, muitas vezes se mostra útil obter a representação no tempo discreto. Para realizar tal transformação deve-se substituir

Fig. 2.2: (a) - Plano s ; (b) - Plano z .

o operador q^{-i} por um atraso de tempo de ordem i , com $i = 1, \dots, o_d$. Assim, a equação (2.5) será reescrita sob a forma de uma equação a diferenças como indica (2.7).

$$d_{o_d}y(k) + \dots + d_1y(k - o_d + 1) + d_0y(k - o_d) = r_{o_n}u(k) + \dots + r_1u(k - (o_d + 1)) + r_0u(k - (o_d - o_n)) \quad (2.7)$$

2.1.3 Representação Discreta Geral

Como já foi abordado anteriormente, muitas são as possibilidades existentes para se representar um sistema dinâmico através de modelos matemáticos. Dentre estas representações existem algumas que são especialmente adequadas à identificação de sistemas onde se deseja conhecer os parâmetros do processo a ser modelado.

Seja o processo linear invariante no tempo apresentado na Fig. 2.3, onde a perturbação $X(z)$ pode ser descrita por um processo estocástico com densidade espectral racional:

Este processo apresentado na Fig. 2.3 é estocástico estacionário no sentido amplo, podendo ser modelado pela transformada z da saída, dada por (Astrom, 1970):

$$Y(z) = G(z)U(z) + H(z)W(z) \quad (2.8)$$

onde $G(z)$ é a FT da parte determinística e $H(z)$ a FT da parte estocástica do sistema e ambas são funções racionais e estáveis, $U(z)$ o sinal de excitação do sistema e $W(z)$ o ruído branco. Além disso

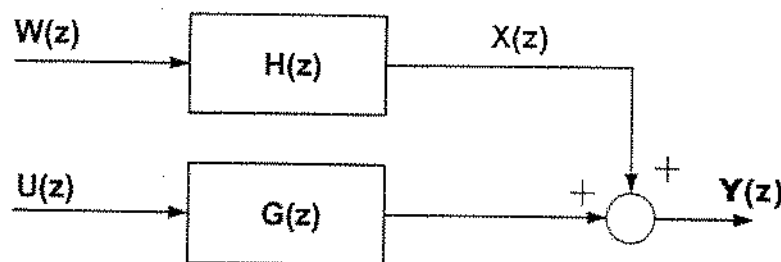


Fig. 2.3: Representação Geral de um Sistema Linear.

a FT $H(Z)$ apresenta as seguintes propriedades adicionais (Astrom, 1970):

- i. $H(z)^{-1}$ é estável;
- ii. $\lim_{z \rightarrow \infty} H(z) = 1$.

onde a primeira garante ao modelo a estabilidade e a propriedade de reconstruir o sinal a partir da saída; a segunda garante que o primeiro termo da representação discreta é sempre 1.

Supondo que, no caso geral, as FT da parte determinística e estocástica podem ter alguns pólos em comum, a equação (2.8) pode ser reescrita de maneira geral como sendo:

$$Y(z) = \frac{z^{-d}B(z)}{F(z)A(z)}U(z) + \frac{C(z)}{D(z)A(z)}W(z) \quad (2.9)$$

onde $A(z)$, $B(z)$, $C(z)$, $D(z)$, $F(z)$ são polinômios em z , cujas raízes são os pólos e zeros da parte determinística e estocástica do sistema, e d é o atraso de transporte do sistema. Aplicando a anti-transformada z e utilizando o operador de atraso q^{-1} ($q^{-d}y(k) = y(k-d)$) resulta no modelo discreto no tempo:

$$A(q^{-1})y(k) = q^{-d} \frac{B(q^{-1})}{F(q^{-1})}u(k) + \frac{C(q^{-1})}{D(q^{-1})}w(k) \quad (2.10)$$

onde $y(k)$ é a saída, $u(k)$ a entrada e $w(k)$ o ruído, todos no instante k .

A representação feita na equação (2.10) possibilita diferentes tipos de modelos. Alguns desses modelos são apresentados abaixo:

• Modelo de Resposta Finita ao Impulso FIR³

Um sistema linear causal pode ser representado pelo somatório de convolução entre a resposta ao impulso, $h(k)$, de um sistema e um sinal de entrada, $u(k)$, resultando no seguinte modelo (modelo IIR⁴):

$$y(k) = \sum_{j=0}^{\infty} h(j)u(k-j) + w(k) \quad (2.11)$$

³FIR - Finite Impulse Response

⁴IIR - Infinite Impulse Response

onde $w(k)$ é a perturbação branca atuando no sistema. Se o sistema for BIBO⁵ estável, então existe um $M < \infty$ tal que $|h(k)| < \epsilon, \forall k > M$. Portanto, truncando a equação (2.11), obtém-se o modelo de resposta finita ao impulso, FIR:

$$y(k) = \sum_{j=0}^M h(j)u(k-j) + w(k) \quad (2.12)$$

onde M é o número de elementos da resposta ao impulso. O modelo FIR pode ser obtido através da equação (2.10) tomando $A(q^{-1}) = C(q^{-1}) = D(q^{-1}) = F(q^{-1}) = 1$ e $B(q^{-1})$ um polinômio arbitrário de ordem M . Nesse caso, o modelo FIR (2.12) pode ser escrito como:

$$y(k) = q^{-d}B(q^{-1})u(k) + w(k)$$

• Modelo ARX⁶

Além da resposta ao impulso finita, outra forma simples de representar uma relação de entrada/saída em um sistema é através da seguinte equação a diferenças:

$$y(k) + a_1y(k-1) + \dots + a_{n_a}y(k-n_a) = b_1u(k-\tau_d) + \dots + b_{n_b}u(k-n_b) + w(k) \quad (2.13)$$

Como o termo de ruído $w(k)$ aparece na equação a diferenças (2.13), esse modelo é classificado como um modelo de erro na equação, ao contrário do modelo **FIR**, onde o erro é adicionado diretamente à saída $y(k)$.

Ao se introduzir a notação:

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_{n_a}q^{-n_a}$$

e

$$B(q^{-1}) = b_1q^{-1} + \dots + b_{n_b}q^{-n_b}$$

obtém-se o modelo autoregressivo com entradas externas (ou **ARX**):

$$A(q^{-1})y(k) = q^{-\tau_d}B(q^{-1})u(k) + w(k)$$

A exemplo do modelo **FIR**, o modelo **ARX** também pode ser reescrito na forma a partir da

⁵BIBO - Bounded Input/ Bounded Output

⁶ARX - Autoregressive with exogenous inputs

equação (2.10):

$$y(k) = \frac{q^{-\tau_d} B(q^{-1})}{A(q^{-1})} u(k) + \frac{1}{A(q^{-1})} w(k) \quad (2.14)$$

Neste caso, o ruído adicionado à saída na equação (2.14), $e(k) = w(k)/A(q^{-1})$, não é branco, ou seja, neste modelo o ruído é modelado como ruído branco filtrado por um modelo auto-regressivo.

• Modelo ARMAX⁷

O modelo ARMAX, a partir do modelo geral representado pela equação (2.10), é obtido tomando-se $D(q^{-1}) = F(q^{-1}) = 1$ e $A(q^{-1})$, $B(q^{-1})$ e $C(q^{-1})$ como polinômios arbitrários, resultando em:

$$y(k) = \frac{q^{-\tau_d} B(q^{-1})}{A(q^{-1})} u(k) + \frac{C(q^{-1})}{A(q^{-1})} w(k) \quad (2.15)$$

Assim como o modelo ARX, o modelo ARMAX também é classificado como um modelo de erro na equação. Neste caso, a perturbação (não-branca) é modelada como a saída de um filtro ARMA com ruído branco na entrada. O modelo ARMAX é considerado uma estrutura padrão usada em modelagem e controle de sistemas. Outras estruturas que são formas particulares de representação do modelo (2.10) podem ainda ser encontradas na literatura (Aguirre, 2000).

2.1.4 Espaço de Estados

Uma outra alternativa para se modelar matematicamente um sistema dinâmico é a representação através de espaço de estados. A representação em espaço de estados traz, além da relação dinâmica entre entrada e saída de um sistema como uma FT, informações relativas às condições internas do sistema através de suas variáveis de estado. Esta informação é importante em muitos sistemas de controle moderno, que baseiam suas leis de controle também nas condições internas do sistema, além de sua saída.

O estado de um sistema dinâmico é o menor conjunto de valores de variáveis (chamadas variáveis de estado) de modo que o conhecimento destes valores em $t = t_0$ junto com o conhecimento dos valores do sinal de entrada para $t \geq t_0$ determina completamente o comportamento do sistema em qualquer instante $t \geq t_0$. Portanto, as variáveis de estado de um sistema dinâmico são as grandezas cujo conjunto de valores determina o estado do sistema. Considerando-se que as variáveis de estado do modelo são $x_1, \dots, x_n$, o espaço n -dimensional cujos eixos coordenados consistem nos eixos $x_1, \dots, x_n$ é chamado espaço de estados (Ogata, 2003).

⁷ARMAX - Autoregressive Moving Average with exogenous inputs

Considere um sistema dinâmico linear invariante no tempo a ser modelado, o qual possui como entradas $u_1, \dots, u_q$, onde q é o número de entradas e $x_1, \dots, x_n$ os estados do sistema dinâmico, sendo n o número de estados do sistema. O conjunto de equações abaixo estabelece a relação dinâmica entre as entradas e os estados do sistema:

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + b_{11}u_1 + b_{12}u_2 + \dots + b_{1q}u_q \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + b_{21}u_1 + b_{22}u_2 + \dots + b_{2q}u_q \\ &\vdots \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n + b_{n1}u_1 + b_{n2}u_2 + \dots + b_{nq}u_q \end{aligned} \quad (2.16)$$

Já a(s) relação(ões) entre a(s) saída(s) e os estados e as entradas é (são) dada(s) por:

$$\begin{aligned} y_1 &= c_{11}x_1 + c_{12}x_2 + \dots + c_{1n}x_n + d_{11}u_1 + d_{12}u_2 + \dots + d_{1q}u_q \\ y_2 &= c_{21}x_1 + c_{22}x_2 + \dots + c_{2n}x_n + d_{21}u_1 + d_{22}u_2 + \dots + d_{2q}u_q \\ &\vdots \\ y_m &= c_{m1}x_1 + c_{m2}x_2 + \dots + c_{mn}x_n + d_{m1}u_1 + d_{m2}u_2 + \dots + d_{mq}u_q \end{aligned} \quad (2.17)$$

onde os termos a_{ij} , b_{ij} , c_{ij} , d_{ij} em sistemas lineares invariantes no tempo são valores constantes que dependem do sistema, sendo que os termos a_{ij} e b_{ij} modelam a parte dinâmica do sistema e os termos c_{ij} e d_{ij} a parte estática. As equações (2.16) e (2.17) podem ser reescritas sob a forma matricial:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \quad (2.18)$$

Assim como na representação através de FT, o modelo em espaço de estados possui uma representação discretizada. A representação discretizada em espaço de estado é equivalente à equação (2.18), onde x deve ser substituído por $x(k)$, u por $u(k)$, y por $y(k)$ e a derivada dos estados $\dot{x}$ é substituída por $x(k+1)$.

$$\begin{aligned} x(k+1) &= A_d x(k) + B_d u(k) \\ y(k) &= C_d x(k) + D_d u(k) \end{aligned} \quad (2.19)$$

Transformações de similaridade podem ser utilizadas a fim de se obter, a partir da representação de estados, a função de transferência que representa o mesmo sistema (Chen, 1999):

$$G(s) = C(sI - A)^{-1}B + D \quad (2.20)$$

Da mesma forma é possível a partir da representação em FT obter a representação canônica matricial da representação em espaço de estados (Chen, 1999).

2.1.5 Funções de Base Ortonormal

As funções de base ortonormal (FBO) vêm sendo aplicadas na identificação de sistemas dinâmicos há várias décadas. É de conhecimento que um sistema linear estável pode ser completamente caracterizado por sua resposta ao impulso $h(k)$. Admitindo que o processo seja causal e que $h(k)$ pertença ao espaço das funções quadraticamente somáveis (Espaço de Lebesgue $L_2[0, \infty)$ no caso contínuo), então o processo pode ser modelado, através de $h(k)$, por uma série de funções ortonormais:

$$h(k) = \sum_{i=1}^{\infty} c_i \phi_i(k) \quad (2.21)$$

onde $\{\phi_i(k)\}_{i=1}^{\infty}$ é uma base de funções ortonormais e c_i são os coeficientes associados ao desenvolvimento em série da resposta ao impulso $h(k)$ através dessa base. Esses coeficientes são escritos a partir de (2.21) e da propriedade de ortonormalidade da seguinte forma:

$$c_i = \sum_{k=0}^{\infty} h(k) \phi_i(k) \quad (2.22)$$

Duas FBO's utilizadas com muita frequência e de especial interesse neste trabalho são as FBO's de Laguerre⁸ e Kautz pois elas formam uma base completa no espaço das funções quadraticamente somáveis no intervalo $[0, \infty)$. Isto significa que qualquer função descrita neste espaço pode ser modelada com precisão arbitrária através de um truncamento finito da combinação linear descrita pela equação (2.21). A base de funções ortonormais de Laguerre é caracterizada pela utilização de funções de transferência com apenas um pólo real (Wahlberg, 1991):

$$\Phi_{laguerre,i}(q^{-1}) = \sqrt{1-p^2} \frac{q^{-1}(q^{-1}-p)^{i-1}}{(1-pq^{-1})^i} \quad (2.23)$$

onde $p = \{p : p \in \mathbb{R} \text{ e } |p| < 1\}$ é o pólo que parametriza as funções ortonormais e q^{-1} é o operador de atraso de tempo.

⁸Edmond Nicolas Laguerre - Nasceu 1834 em Bar-le-Duc, França. Graduiu-se pela École Polytechnique em Paris em 1854. Publicou seus trabalhos a respeito das funções polinomiais de Laguerre em 1879 e mostrou que uma função arbitrária pode ser expandida em série de funções polinomiais de Laguerre. Faleceu em 1886 em Bar-le-Duc, França.

Já as funções de Kautz (Wahlberg, 1994) são parametrizadas pelo par de pólos complexos conjugados $\beta = \alpha \pm j\omega$ segundo a equação:

$$\begin{aligned}\Phi_{2i-1}(q^{-1}) &= \frac{\sqrt{(1-c^2)}(z-b)}{z^2 + b(c-1)z - c} \left(\frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c} \right)^{i-1} \\ \Phi_{2i}(q^{-1}) &= \frac{\sqrt{(1-b^2)(1-c^2)}}{z^2 + b(c-1)z - c} \left(\frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c} \right)^{i-1}\end{aligned}\quad (2.24)$$

onde $b = (\beta + \beta^*)/(1 + \beta\beta^*)$ e $c = -\beta\beta^*$, sendo $-1 < b < 1$ e $-1 < c < 1$. Se os pólos das representações (2.23) e (2.24) são iguais a zero a base FBO se torna um modelo FIR. A base de Laguerre é mais apropriada para representar sistemas com pólos puramente reais ou com parte imaginária de valor reduzido, já a base de Kautz necessita de um número menor de funções para representar sistemas com dinâmica oscilatória, por ser parametrizada por pólos complexos conjugados.

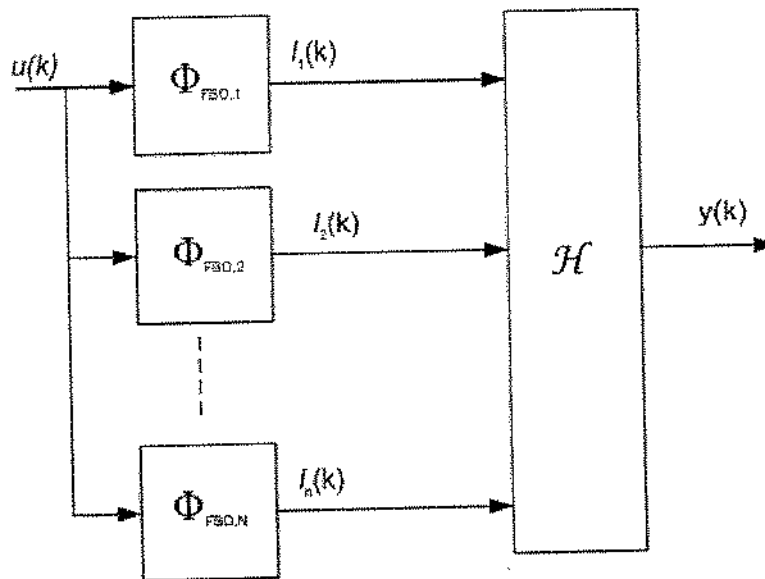


Fig. 2.4: Diagrama de blocos do modelo FBO

As funções ortonormais podem ser representadas através de equações de estado (Zervos e Dumont, 1988) devido à propriedade de serem recursivas, isto é, a i -ésima função pode ser escrita em função da $(i-1)$ -ésima. No caso, o modelo da Fig. 2.4 pode ser representado no espaço de estados como:

$$\begin{aligned}l(k+1) &= A_f l(k) + B_f u(k) \\ y(k) &= C_f^T l(k)\end{aligned}\quad (2.25)$$

onde $l(k) = [l_1(k) \dots l_d(k)]^T$ é o vetor de estados ortonormais (denominação usada por conveniência já que a ortonormalidade é uma propriedade da função e não dos estados) de ordem d . Este modelo pode ainda incluir uma parcela na parte estática da representação que modela o nível contínuo do sistema sob estudo. É também de interesse neste trabalho (para a modelagem para o controle preditivo como será visto na seção 4.2) a representação dos sistemas por modelos variacionais em espaço de estado. Como esta representação modela a variação nos estados e saídas entre dois instantes, dispensa a parcela que modela o nível contínuo do sistema:

$$\begin{aligned}\Delta l(k+1) &= A_f \Delta l(k) + B_f \Delta u(k) \\ y(k) &= y(k-1) + C_f^T \Delta l(k)\end{aligned}\quad (2.26)$$

Nas representações em espaço de estado as matrizes A_f e B_f dependem apenas da base de funções ortonormais. No caso da base de Laguerre tem-se (Campello, 2002):

$$A_f = \begin{bmatrix} p & 0 & 0 & \dots & 0 \\ 1-p^2 & p & 0 & \dots & 0 \\ (-p)(1-p^2) & 1-p^2 & p & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (-p)^{d-2}(1-p^2) & (-p)^{d-3}(1-p^2) & \dots & p \end{bmatrix}$$

$$B_f = \sqrt{1-p^2} \begin{bmatrix} 1 & -p & (-p)^2 & \dots & (-p)^{d-1} \end{bmatrix}^T$$

onde p é o pólo de Laguerre e d é o número de filtros de Laguerre. Assim como a base de Laguerre, a base Kautz também possui uma representação matricial (Wahlberg, 1994).

Uma vantagem desta abordagem é que o modelo é linear com relação aos coeficientes da matriz C_f . Este, portanto, pode ser estimado através de método de identificação clássico como mínimos quadrados, que é descrito no Apêndice A. Além disso, o sucesso do emprego de modelos dinâmicos baseados em funções de base ortonormais na modelagem de sistemas lineares e não-lineares, quando comparados a outros modelos como os da equação (2.10), se deve às seguintes características:

- Ausência de realimentação da saída, evitando realimentação do erro de predição;
- Não há a necessidade de determinar previamente a ordem do modelo nem o atraso de transporte, eliminando assim a etapa (geralmente árdua) de determinação dos regressores;
- É possível aumentar a capacidade de aproximação dos modelos simplesmente aumentando o número de funções;

- Os modelos FBO apresentam grande tolerância às dinâmicas não modeladas e sensibilidade reduzida aos parâmetros estimados;
- Capacidade de lidar de forma robusta com atrasos de transporte.

2.1.6 Modelos Não-lineares

Até o presente momento nesta dissertação foram apresentadas somente representações de modelos lineares. Embora a maioria dos sistemas encontrados na prática sejam não-lineares, muitas vezes representações lineares são suficientes para representar tais sistemas, mas em alguns casos se torna necessário o uso de modelos não-lineares, o que aumenta a complexidade de tais modelos e na forma de sua identificação. A utilização de tais modelos não-lineares se justifica quando é necessário representar certos comportamentos dinâmicos que os modelos lineares não conseguem reproduzir. A seguir serão apresentadas algumas abordagens de modelos não-lineares utilizados na modelagem de sistemas dinâmicos (Aguirre, 2000; Aguirre, Rodrigues e Jácome, 1998):

- **NARX - NARX⁹ e NARMAX¹⁰**

Os modelos não-lineares NARX e NARMAX (N de *nonlinear*) são extensões de seus equivalentes lineares. O modelo NARX descreve a saída do sistema como uma função não-linear dos termos passados da entrada e da saída, ou seja:

$$y(k) = \mathcal{F}[y(k-1), \dots, y(k-n_a), u(k-1), \dots, u(k-n_b)] \quad (2.27)$$

O modelo NARMAX se diferencia do modelo NARX pela inserção da perturbação e seus termos passados na dinâmica do modelo:

$$y(k) = \mathcal{F}[y(k-1), \dots, y(k-n_a), u(k-1), \dots, u(k-n_b), e(k), \dots, e(k-n_c)] \quad (2.28)$$

Nas equações (2.27) e (2.28), n_a , n_b e n_c representam os maiores atrasos da saída, da entrada e da perturbação (ruído), respectivamente, sendo $\mathcal{F}$ uma função não-linear. Algumas realizações especiais da função $\mathcal{F}[\cdot]$ para os modelos NARX permitem a obtenção de modelos lineares nos parâmetros, mesmo com regressores não-lineares, o que se torna vantajoso na identificação de tais parâmetros. Contudo, persiste um grande problema que é a definição da estrutura do modelo, ou seja, a escolha de uma função $\mathcal{F}$ apropriada e dos termos n_a , n_b e n_c (Aguirre, 2000). Outra dificuldades para a identificação de modelos NARX ou NARMAX é que o número de

⁹Nonlinear Autoregressive with exogenous inputs

¹⁰NARMAX - Nonlinear Autoregressive Moving Average with exogenous inputs

parâmetros cresce rapidamente com o aumento da ordem polinomial do modelo ou com o aumento dos atrasos, o que torna essencial a determinação dos termos relevantes nos regressores, acrescido da necessidade de uma estimação numericamente robusta de parâmetros o que torna a obtenção de tais modelos muitas vezes uma tarefa onerosa (Aguirre et al., 1998). Destaca-se também que modelos não-lineares polinomiais sobreparametrizados podem apresentar regimes dinâmicos espúrios (Aguirre e Billings, 1995).

• Modelos de Hammerstein e de Wiener

Há ainda dois modelos bastante populares na modelagem de sistemas dinâmicos não-lineares, os modelos de Hammerstein e de Wiener. Estes modelos são a composição de um modelo dinâmico linear $H(s)$ em cascata com uma função estática não-linear $f(\cdot)$, sendo que no modelo de Hammerstein, a não-linearidade estática precede o modelo dinâmico linear:

$$U^*(s) = f(U(s)); \text{ e } Y(s) = H(s)U^*(s) \quad (2.29)$$

Para o modelo de Wiener, quem precede na representação é o modelo dinâmico linear, sendo sucedido pela não-linearidade estática:

$$Y^*(s) = H(s)U(s); \text{ e } Y(s) = f(Y^*(s)) \quad (2.30)$$

Os modelos apresentados nas equações (2.29) e (2.30) podem ser definidos no tempo discreto. Estas representações têm como inconveniente o fato de a função $f(\cdot)$ ser usualmente desconhecida.

• Modelos NOBF¹¹

Genericamente, um modelo não-linear do tipo entrada/saída (por simplicidade SISO¹² ou MISO¹³) pode ser escrito como:

$$y(k) = \mathcal{H}(\lambda(k))$$

onde $\lambda(k)$ é o vetor de regressão, que pode conter entradas, saídas e erros de previsão em diferentes instantes de tempo, e $\mathcal{H}$ é um mapeamento estático genérico. Os modelos NOBF são generalizações não-lineares dos modelos lineares baseados em funções ortonormais. No caso

¹¹NOBF - Nonlinear Orthonormal Base Function

¹²SISO - Single-Input, Single-Output

¹³MISO - Multi-Input, Single-Output

SISO o vetor de regressão baseado na Fig. 2.4 é dado por:

$$\lambda(k) = \begin{bmatrix} l_1(k) & l_2(k) & \dots & l_d(k) \end{bmatrix}$$

onde $\Phi_{FBO,i}$ é o i -ésimo filtro ortonormal.

Este modelo apresenta vantagens com relação àqueles com recursão de saída, como acontece com os modelos FBO lineares, apresentando ainda a vantagem adicional de garantia de estabilidade na representação de sistemas estáveis. Evidentemente trata-se de uma estrutura do tipo Wiener, onde o mapeamento estático $\mathcal{H}$ pode ser realizado, por exemplo, através de séries de Volterra, sendo que nesse caso o operador $\mathcal{H}$ é um polinômio multidimensional. Também este mapeamento pode ser uma Rede Neural ou ainda um mapeamento *fuzzy* (Campello, 2002).

• Série de Volterra

Outra forma de se representar um modelo de um sistema não-linear é através das séries de Volterra, cujos estudos se iniciaram em 1887 com V. Volterra. A saída $y(t)$ de um sistema não-linear com entrada $u(t)$ pode ser representada por uma série de Volterra como (Aguirre et al., 1998):

$$y(t) = \sum_{j=1}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h_j(\tau_1, \dots, \tau_j) \prod_{i=1}^j u(t - \tau_i) d\tau_i \quad (2.31)$$

onde $h_j(\tau_1, \dots, \tau_j)$ são funções denominadas kernels (núcleos) de Volterra, que nada mais são que generalizações não-lineares da resposta ao impulso de um sistema linear. De fato, para um sistema linear ($j = 1$), a equação (2.31) se reduz à integral de convolução.

O modelo discreto de Volterra é obtido substituindo-se as integrais de convolução pelos somatórios de convolução. Em uma notação compacta, tem-se (Aguirre et al., 1998):

$$y(k) = \sum_{j=1}^{\infty} \sum_{\tau_1=0}^{\infty} \dots \sum_{\tau_j=0}^{\infty} h_j(\tau_1, \dots, \tau_j) \prod_{i=1}^j u(k - \tau_i) \quad (2.32)$$

As séries de Volterra não usam informação de recorrência (valores anteriores de saída) em seu modelo e isso, embora traga benefícios, também acarreta problemas, porque mesmo para sistemas com não-linearidades, pode haver um grande número de parâmetros a se determinar. Para casos onde trunca-se a série, como o modelo não usa recorrência, o mesmo passa a ser um modelo NFIR¹⁴, ou seja:

$$y(k) = \mathcal{F}(u(k-1), \dots, u(k-n_b))$$

¹⁴NFIR - Nonlinear Finite Impulse Response

com $\mathcal{F}$ polinomial.

- **Abordagens com Inteligência Computacional**

O uso de técnicas de inteligência computacional - IC na identificação e modelagem de sistemas dinâmicos é cada vez mais difundida. Muitas dessas técnicas apresentam características interessantes à modelagem de sistemas dinâmicos não-lineares pois permitem a construção de modelos como aproximadores universais de funções.

Uma das técnicas de IC empregadas na modelagem de sistemas dinâmicos são as Redes Neurais Artificiais - RNA. O conceito de RNA's foi introduzido por McCulloch e Pitts (McCulloch e Pitts, 1943) num trabalho em que propunham um modelo matemático para representar um neurônio artificial. Atualmente as RNA's reúnem diversos modelos matemáticos distintos com aplicação nos mais diferentes campos de aplicação dentro das áreas da engenharia e de outras ciências. As redes mais utilizadas para aproximações de mapeamentos não-lineares genéricos, as quais são utilizadas em aplicações como modelagem e controle de sistemas dinâmicos, são as redes *feedforward* como as redes MLP¹⁵ e RBF¹⁶ (Haykin, 1999; Kosko, 1992). Apesar das abordagens serem bastante distintas, há problemas inerentes a cada uma que são comuns, como a determinação da estrutura dos modelos não-lineares, que no caso de redes neurais irá implicar na determinação da topologia da rede (onde se deve escolher o número de neurônios, número de camadas, tipo de função de ativação, quais entradas usar, dentre outros parâmetros de projeto).

Outra técnica muito utilizada na aproximação de mapeamentos não-lineares é a aproximação por Modelos *Fuzzy* (Hellendoorn e Driankov, 1997; Yager e Filev, 1994). As idéias originais de teoria de conjuntos *Fuzzy* na modelagem de sistemas foram introduzidas por Zadeh (1973) para representação de sistemas complexos através de um conjunto de declarações condicionais denominado algoritmo comportamental *fuzzy*. Os modelos *fuzzy* têm crescido de importância, dentre diversos outros fatores, devido à sua capacidade de aproximação universal de funções, ou seja, podem aproximar com precisão arbitrária qualquer mapeamento não-linear contínuo definido numa região compacta (limitada e fechada) (Wang e Mendel, 1992; Kosko, 1994; Wang, 1998). Outra característica que torna esses modelos muito atrativos é sua habilidade de tratar tanto informações de natureza qualitativa como quantitativa, podendo a modelagem ser feita através de informações obtidas a partir do conhecimento de especialistas do processo e/ou da extração de informações obtidas diretamente dos dados do processo (Wang e Mendel, 1992; Zeng e Singh, 1994; Zeng e Singh, 1995). Além das aplicações em modelagem, os conceitos

¹⁵MLP - Multi-Layer Perceptron

¹⁶RBF - Radial Base Function

de sistemas *Fuzzy* podem ser usados no controle de sistemas dinâmicos (Lee, 1990). A Técnica de Conjuntos *Fuzzy* será melhor abordada no Capítulo 3.

Existem ainda outras formulações para modelagem de sistemas dinâmicos não-lineares que podem ser encontradas na literatura como através de equações ou sistemas de equações diferenciais, entre outros.

2.1.7 Identificação de sistemas dinâmicos

Ao se realizar a modelagem de um sistema dinâmico, além de se determinar a estrutura a ser utilizada para a modelagem e representação do sistema, a determinação dos parâmetros do modelo é uma questão muito relevante, pois a qualidade de bons modelos depende diretamente dos seus parâmetros. A estimativa dos parâmetros necessita de ferramentas específicas para ser resolvida, algumas até mesmo empíricas, outras, ferramentas matemáticas que utilizam os dados disponíveis do sistema, como medições do sistema em operação ou informações de elementos que compõem o sistema.

Para processos modelados através de equações ou sistemas de equações diferenciais muitas são as ferramentas apresentadas na literatura para determinação dos parâmetros do modelo.

As funções de transferência de resposta ao impulso contínuas no tempo, a forma mais utilizada em sistemas de controle clássico para se representar sistemas dinâmicos, apresentam muitas técnicas para se obter seus parâmetros. Algumas dessas técnicas possibilitam a obtenção dos parâmetros pela inspeção da resposta ao degrau (sistemas de primeira e segunda ordem) (Aguirre, 2000). Outras técnicas utilizam ainda informações de resposta na frequência destes mesmos sistemas (Ljung, 1999). Para sistemas que têm suas FT's discretizadas no tempo (domínio z) os parâmetros podem ser obtidos através do Estimador de Mínimos Quadrados (MQ). Este mesmo método também é utilizado para identificação de representações em espaço de estados discretizadas como apresentadas nas equações (2.19) e (2.25).

O estimador de mínimos quadrados minimiza de forma quadrática o erro entre as medidas obtidas do sistema e o valor estimado pelo modelo, e através deste processo de minimização fornece os parâmetros do modelo (para maiores detalhes vide Apêndice A). O estimador de mínimos quadrados apresenta as seguintes propriedades quando a perturbação agindo no processo é branca e aditiva na saída do processo (em malha aberta, de forma que a saída seja estatisticamente independente da perturbação) (Ljung, 1999):

- Não polarizado, ou seja, o valor esperado (média) do erro de estimação é nulo;
- Melhor estimador linear não polarizado BLUE¹⁷, implicando que ele minimiza a variância do

¹⁷BLUE - Best Linear Unbiased Estimator

erro de estimação;

- Consistente - o que significa que a variância do erro de estimação tende a zero na medida em que o número N de medições tende a infinito.

Estas propriedades fazem do estimador de mínimos quadrados atrativo para a estimação de parâmetros de sistemas.

2.2 Resumo

O objetivo deste capítulo foi apresentar de maneira introdutória a teoria de modelagem e identificação de sistemas dinâmicos. Inicialmente foram apresentados os modelos mais utilizados na modelagem de sistemas dinâmicos lineares e invariantes no tempo LTI¹⁸. Foi discutido como os sistemas são modelados através de equações diferenciais, funções de transferência e representações em espaço de estados, sendo estas duas últimas formas nos domínio contínuos e discretos (s e z respectivamente). Foram introduzidas as Funções de Base Ortonormal com suas principais propriedades, sendo que esta estrutura dinâmica será a de maior interesse nesta dissertação. Posteriormente, foram abordados algumas realizações específicas de modelos dinâmicos não-lineares, apresentando entre outras, realizações através de inteligência computacional, Modelos *Fuzzy*, os quais serão apresentados mais detalhadamente no próximo capítulo.

Para a identificação de sistemas dinâmicos onde se torna necessário determinar os parâmetros do modelo sob estudo, o estimador de mínimos quadrados é o método de maior interesse no presente trabalho como poderá ser verificado no Capítulo 3.

¹⁸LTI - Linear Time-invariant System

Capítulo 3

Sistemas *Fuzzy*

Este capítulo tem como objetivo introduzir os conceitos elementares de teoria de conjuntos *fuzzy* e apresentar alguns fundamentos e aplicações que serão empregados neste trabalho, tanto na modelagem de sistemas como no controle de sistemas lineares e não-lineares.

Os conceitos de conjuntos *fuzzy*, também conhecida como teoria de conjuntos nebulosos, foram originalmente propostos por Zadeh (1965) como uma generalização da idéia de um conjunto ordinário ou *crisp* (não *fuzzy*). Um conjunto *fuzzy* pode ser visto como um conjunto em que o valor do predicado pode pertencer ao intervalo $I = [0, 1]$, ao contrário dos conjuntos ordinários em que o valor pertence ao conjunto $\{0, 1\}$, ou seja, para conjuntos ordinários um dado elemento $\in$ ou $\notin$ a um dado conjunto enquanto que nos conjuntos *fuzzy* um elemento pertence a um conjunto com valor de verdade entre 0 e 1. Assim, os conjuntos *fuzzy* estão sujeitos a uma lógica com abordagem multivalorada. Os modelos *fuzzy* apresentam a característica, além de outras que serão apresentadas a seguir, de descrever sistemas de forma aproximada, porém com precisão arbitrária, ou seja, como aproximadores universais de funções numa região compacta, apresentando a capacidade de mapear de forma eficiente relações entre entradas e saídas de sistemas em geral (Wang e Mendel, 1992; Kosko, 1994; Wang, 1998).

Neste capítulo discute-se um algoritmo iterativo para a determinação das funções de pertinência trapezoidais utilizadas nos antecedentes das regras *fuzzy* a partir das projeções dos objetos presentes em grupos *fuzzy*. Discute-se ainda uma metodologia para a determinação do número de grupos e, por consequência, o número de regras no modelo *fuzzy* Takagi Sugeno.

3.1 Introdução aos conjuntos *Fuzzy*

A teoria de lógica e conjuntos *fuzzy* permite o tratamento de implicações lógicas seguindo regras naturais de raciocínio nas quais, através da análise de condições, estipulam-se consequências baseadas em tais condições. Um conjunto *fuzzy* é definido como uma coleção de objetos cujo grau de

pertinência ao conjunto varia no intervalo contínuo entre zero e um. Formalmente, um conjunto *fuzzy* é caracterizado por uma função de pertinência que mapeia os elementos do universo de discurso U (domínio) no intervalo contínuo $[0, 1]$, ou seja:

$$A : U \rightarrow [0, 1] \quad (3.1)$$

Desta forma, o valor $A(x)$ indica com que grau o elemento x pertence ao conjunto A (Pedrycz e Gomide, 1998). Um grau de pertinência 1 equivale ao clássico símbolo de pertinência $\in$ enquanto um grau de pertinência 0 equivale ao clássico símbolo $\notin$.

Um conceito fundamental quando se trabalha com conjuntos *fuzzy* é o de variável lingüística. Entende-se por variável um identificador que pode assumir um dentre vários valores e, deste modo, uma variável lingüística pode assumir um valor lingüístico dentre vários outros em um conjunto de termos lingüísticos. Cada termo lingüístico é caracterizado por uma função de pertinência. É permitido ainda o uso de modificadores qualificadores, como por exemplo: “pouco pequeno positivo”, “muito grande negativo”, “muito aproximadamente zero”, etc. Por definição as funções de pertinência podem ser qualquer função que produza valores entre 0 e 1. As mais utilizadas na literatura são as funções triangulares, trapezoidais, sigmóides ou gaussianas, podendo também as funções de pertinência serem do tipo *singleton*. A Fig. 3.1 mostra funções de pertinência trapezoidais, as quais são de interesse neste trabalho, como será discutido na seção 3.6.1. Uma definição satisfatória da quantidade e a determinação dos parâmetros das funções de pertinência no universo de discurso é fundamental para determinar a qualidade de um modelo *fuzzy* aplicado a um dado sistema.

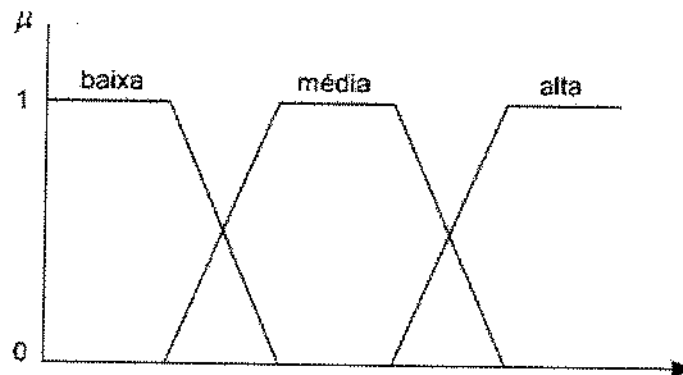


Fig. 3.1: Funções de pertinência trapezoidais

Em um sistema *fuzzy* (seja um modelo *fuzzy* de um processo ou um controlador *fuzzy*) as funções de pertinência são responsáveis pela conversão das grandezas do domínio do mundo real, captadas por sensores, dispositivos computadorizados ou mesmo provenientes de outros segmentos do processo de controle, para números *fuzzy*. Essa conversão, conhecida como “*fuzzificação*”, é essencial para a atuação da máquina de inferência *fuzzy*.

A etapa de inferência consiste em se avaliar um conjunto de regras (a base de conhecimento) do tipo “se ... então ...” que descrevem a dependência entre as variáveis lingüísticas de entrada e as de saída. Essas regras seguem o paradigma *modus ponens*, um mecanismo de inferência progressiva (ao contrário do *modus tollens*, regressivo, mais utilizado em sistemas especialistas), para fazer o mapeamento do conceito de implicação lógica.

Os antecedentes das regras em conjuntos *fuzzy* possuem a mesma estrutura geral, qual seja:

$$\begin{aligned} &\text{Se } x_1 \text{ é } A_{1,1} \text{ e } x_2 \text{ é } A_{1,2} \text{ e } \dots \text{ e } x_d \text{ é } A_{1,d} \\ &\text{Se } x_1 \text{ é } A_{2,1} \text{ e } x_2 \text{ é } A_{2,2} \text{ e } \dots \text{ e } x_d \text{ é } A_{2,d} \\ &\quad \vdots \\ &\text{Se } x_1 \text{ é } A_{K,1} \text{ e } x_2 \text{ é } A_{K,2} \text{ e } \dots \text{ e } x_d \text{ é } A_{K,d} \end{aligned}$$

sendo $x_1, \dots, x_d$ as entradas do sistema e $A_{1,1}, \dots, A_{K,d}$ os termos lingüísticos dessas variáveis, definidos como funções de pertinência que cobrem todo o universo de discurso. A inferência consiste de dois passos: avaliação da premissa de cada regra (conjunção), através dos operadores *t*-norma, e em seguida a etapa de agregação, unindo (disjunção) as diferentes conclusões das regras ativas sob o operador *s*-norma (Yager e Filev, 1994; Lee, 1990). Os operadores de *t*-norma e *s*-norma são respectivamente normas e co-normas triangulares que fornecem métodos genéricos para as operações de intersecção e união em conjuntos *fuzzy*. Alguns exemplos de *t*-normas são o produto e o operador de mínimo. Para as *s*-normas têm-se como exemplos o operador de máximo ou a soma limitada (Pedrycz e Gomide, 1998).

Um sistema *fuzzy* contém um conjunto dessas regras, todas ativadas em paralelo. Assim, o sistema *fuzzy* trabalha com inferência associativa paralela, ou seja, quando uma entrada é fornecida, o controlador dispara as regras paralelamente, com diferentes graus de ativação, para inferir um resultado ou saída. Em algumas arquiteturas *fuzzy*, após a inferência da ação a ser tomada, necessita-se de uma tradução do valor lingüístico para a variável numérica de saída, sendo este passo conhecido como “*defuzzificação*”. Como pode acontecer de saídas distintas serem acionadas num mesmo momento, com diferentes graus de ativação, deve-se encontrar o valor que melhor corresponda à combinação das saídas dos conjuntos *fuzzy* do modelo. Dentre os métodos de *defuzzificação* mais utilizados pode-se citar o método de centro de área ou a abordagem de centro dos máximos que equivale a calcular a média ponderada das saídas pelos valores de ativação das respectivas regras (Lee, 1990; Dubois e Prade, 1984; Gupta, Kandel, Bandler e Kiszka, 1985).

O modo como são formadas as regras, relativas ao tipo de antecedentes e conseqüentes, possibilita uma classificação do tipo de estrutura utilizada. A seção seguinte apresenta três tipos estruturas de modelos *fuzzy* sendo a primeira, a estrutura dos modelos relacionais *fuzzy*, em seqüência a estrutura para o modelo *fuzzy* Mamdani e por fim a estrutura utilizada para a construção de modelos *fuzzy* TS.

Maiores detalhes dos conceitos de lógica e conjuntos *fuzzy* podem ser obtidos no Apêndice B.

3.2 Modelo Relacional

Modelos relacionais representam o mapeamento entre os conjuntos *fuzzy* de entrada A_i e de saída B_i através de uma relação *fuzzy* (Pedrycz e Gomide, 1998; Babuska e Verbruggen, 1996). Como exemplo, sejam x a entrada de um sistema, onde $x \in \mathcal{X}$ e y a saída do sistema com $y \in \mathcal{Y}$. Denotemos por $\mathbf{A}$ uma coleção de M_a termos lingüísticos (conjuntos *fuzzy*) definidos no domínio de $\mathcal{X}$ e $\mathbf{B}$ uma coleção de N_a conjuntos *fuzzy* definidos em $\mathcal{Y}$:

$$\mathbf{A} = \{A_1, A_2, \dots, A_{M_a}\}$$

$$\mathbf{B} = \{B_1, B_2, \dots, B_{N_a}\}$$

Uma relação *fuzzy* $R = [r_{ij}]_{M \times N}$ define o mapeamento a partir dos conjuntos de entrada $\mathbf{A}$ para os conjuntos de saída $\mathbf{B}$, $R : \mathbf{A} \rightarrow \mathbf{B}$, como mostra a Fig. 3.2:

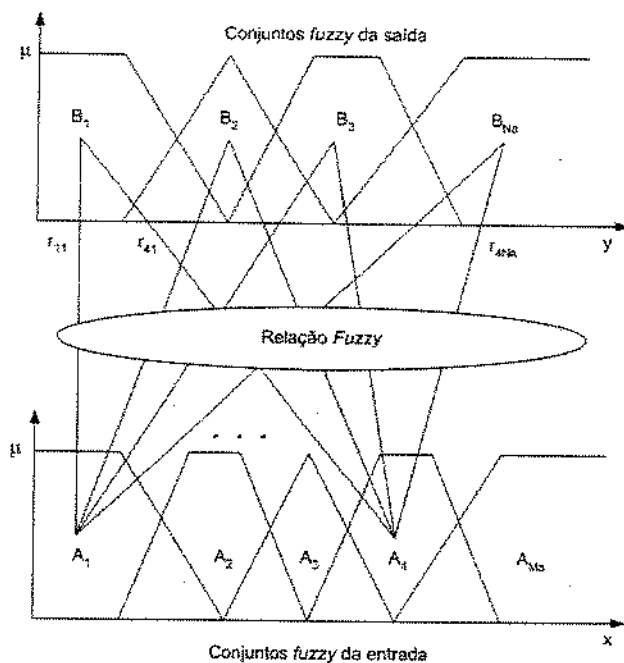


Fig. 3.2: Modelo relacional *fuzzy*

Para um dada entrada x , o conjunto *fuzzy* $X = \{\mu_{A_1}, \mu_{A_2}, \dots, \mu_{A_{M_a}}\}$ representa o grau de ativação dos respectivos termos lingüísticos. O correspondente conjunto *fuzzy* da saída $Y = \{\mu_1, \mu_2, \dots,$

$\dots, \mu_{N_q}\}$ são obtidos utilizando a composição relacional:

$$Y = X \circ R$$

Se um valor crisp é requerido no modelo de saída, o conjunto *fuzzy* Y deve ser *defuzzificado*. Usualmente, a *defuzzificação* pela média é aplicada aos centróides dos conjuntos B_i

$$b_i = \frac{\sum_{q=1}^{N_q} \mu_{B_i}(y_q) y_q}{\sum_{q=1}^{N_q} \mu_{B_i}(y_q)} \quad i = 1, 2, \dots, N;$$

onde N_q é o número de níveis de discretização, onde, qualquer outro método poderia ser utilizado para calcular b_i 's (como centro de área ou média da máxima). A saída do modelo relacional *fuzzy* y_0 é calculado como a média ponderada dos b_i 's:

$$y_0 = \frac{\sum_{i=1}^N \mu_i b_i}{\sum_{i=1}^N \mu_i}$$

3.3 Modelo Mamdani

Um modelo do tipo Mamdani para um sistema SISO consiste de um conjunto de regras, como por exemplo:

R_1 : Se x é A_1 então y é B_1

R_2 : Se x é A_2 então y é B_2

R_3 : Se x é A_3 então y é B_1

R_4 : Se x é A_4 então y é B_2

onde no conjunto de regras acima, x representa a variável de entrada que tem seu valor no domínio de um universo de discurso U enquanto A_1, A_2, A_3 e A_4 são as variáveis lingüísticas *fuzzy* com suas respectivas funções de pertinência $\mu_{A_i}(x)$. A variável de saída do modelo é y , sendo B_1 e B_2 variáveis lingüísticas definidas no universo de discurso dessa variável através das respectivas funções de pertinência $\mu_{B_i}(y)$.

Os modelos de Mamdani podem ainda modelar sistemas MISO ou MIMO onde os antecedentes e os conseqüentes serão compostos. Para cada regra definida no modelo pode ser relacionada uma região no espaço entrada/saída, podendo ser isoladas ou superpostas conforme necessário (Yager e Filev, 1994).

As estruturas de Mamdani necessitam de um processo de "*defuzzificação*" após a inferência dos dados. Os processos de *defuzzificação* mais utilizados na literatura, como citado anteriormente, são os métodos de centro de gravidade e a abordagem de centro dos máximos (Lee, 1990; Dubois e

Prade, 1984; Gupta et al., 1985).

3.3.1 Modelos *Singleton*

Um caso especial de modelo lingüístico *fuzzy* é obtido quando os conseqüentes B_i das regras são *singletons*, ou seja, representados por números reais b_i como mostra a regra a seguir:

$$R_i : \text{Se } x \text{ é } A_i \text{ então } y \text{ é } b_i, \quad i = 1, 2, \dots, K. \quad (3.2)$$

Este modelo é denominado *singleton*. Um método de inferência (com *defuzzificação* integrada) muito utilizado com esse modelo é:

$$\hat{y} = \frac{\sum_{i=1}^K \mu_i b_i}{\sum_{i=1}^K \mu_i} \quad (3.3)$$

que corresponde ao método de centro de gravidade, onde μ_i é o valor de ativação da regra i , definido pelo valor de ativação da função de pertinência da variável lingüística A_i . O modelo *fuzzy singleton* pertence a uma classe de aproximadores gerais de funções, chamadas também de expansão de funções base (Friedman, 1991), seguindo a seguinte forma:

$$\hat{y} = \sum_{i=1}^K \phi_i(x) b_i \quad (3.4)$$

Muitas estruturas usadas em identificação de sistemas não-lineares, tais como redes neurais de funções de base radial ou *splines*, pertencem a essa classe de sistemas.

A equação (3.3) é exatamente a mesma que descreve os modelos Takagi-Sugeno (como será descrito na seção 3.4) quando os conseqüentes das regras desses são constantes, ao invés de funções das variáveis das premissas.

3.4 Modelo Takagi-Sugeno

Um dos modelos baseados em regras mais usado para aproximação de sistemas não-lineares foi introduzido por Takagi e Sugeno (1985) e, na literatura referente à modelagem *fuzzy*, é citado como modelo Takagi-Sugeno (TS). Tais modelos consistem essencialmente de uma interpolação de múltiplos modelos locais (usualmente lineares ou afins), o que é favorável ao projeto de controladores e de aproximadores universais de funções. A forma mais comum de um modelo TS baseado em regras é dada por:

$$\text{Se } \underbrace{x \text{ é } A_i}_{\text{antecedente}} \text{ então } \underbrace{y_i = f_i(x)}_{\text{conseqüente}} \quad (3.5)$$

A primeira parte da regra, chamada de antecedente, é definida como uma proposição *fuzzy* “ x é A_i ” onde x é um vetor real exato (não-*fuzzy*) e A_i é um conjunto *fuzzy* definido pela função de pertinência

$$\mu_{A_i}(x) : \mathbf{R}^d \rightarrow [0, 1]$$

O indexador $i = 1, \dots, K$ denota a i -ésima regra, onde K é número de regras na base. O grau de desempenho (grau de verdade) μ_i dos antecedentes para um dado valor do vetor de regressão x é avaliado como o grau de ativação da função de pertinência no conjunto A_i : $\mu_i = \mu_{A_i}(x)$. O antecedente é usualmente expresso pela combinação dos conjuntos *fuzzy* dos antecedentes em cada sub-espaco univariável definido para cada componente de x . Os antecedentes são combinados usando conjunção, isto é, conectivos lógicos E. A i -ésima regra formulada então com um conjunto multidimensional na entrada é dada por:

$$\text{Se } x_1 \text{ é } A_{i,1} \text{ e } x_2 \text{ é } A_{i,2}, \dots, x_d \text{ é } A_{i,d} \text{ então } y_i = f_i(x) \quad (3.6)$$

sendo o grau de ativação da i -ésima regra, $\mu_i(x)$, calculado como:

$$\mu_i(x) = \mu_{A_{i,1}}(x_1) \wedge \mu_{A_{i,2}}(x_2) \wedge \dots \wedge \mu_{A_{i,d}}(x_d) \quad (3.7)$$

onde o operador ($\wedge$) é usado para conjunção. Para o operador produto nós obtemos:

$$\mu_i(x) = \mu_{A_{i,1}}(x_1) \cdot \mu_{A_{i,2}}(x_2) \cdot \dots \cdot \mu_{A_{i,d}}(x_d) \quad (3.8)$$

Note que $\mu_i(x)$ em (3.7) e (3.8) é uma função de pertinência *fuzzy* com entrada multivariável constituída pela intersecção dos conjuntos univariáveis *fuzzy* das entradas x . O formato e a característica do conjunto *fuzzy* resultante depende do operador de intersecção utilizado. A Fig. 3.3 mostra um exemplo de intersecção realizado com uso do operador mínimo.

A segunda parte da regra (3.5), chamada de conseqüente, é uma função das variáveis da premissa. As regras devem em geral se diferenciar pelos conjuntos *fuzzy* dos antecedentes e também pelas funções f_i dos conseqüentes. No entanto, geralmente na prática, f_i são funções parametrizadas, cuja estrutura não muda independentemente da regra, variando somente os parâmetros. Uma estrutura simples e prática muito utilizada é a “função afim”, que resulta na seguinte regra:

$$\text{Se } x \text{ é } A_i \text{ então } y_i = a_i^T x + b_i \quad (3.9)$$

onde a_i é um vetor de parâmetros e b_i é o *offset*. Este modelo é denominado TS afim. Quando $b_i = 0$, $i = 1, \dots, K$, o modelo é TS linear. Os antecedentes definem em que região do espaço de regressão o

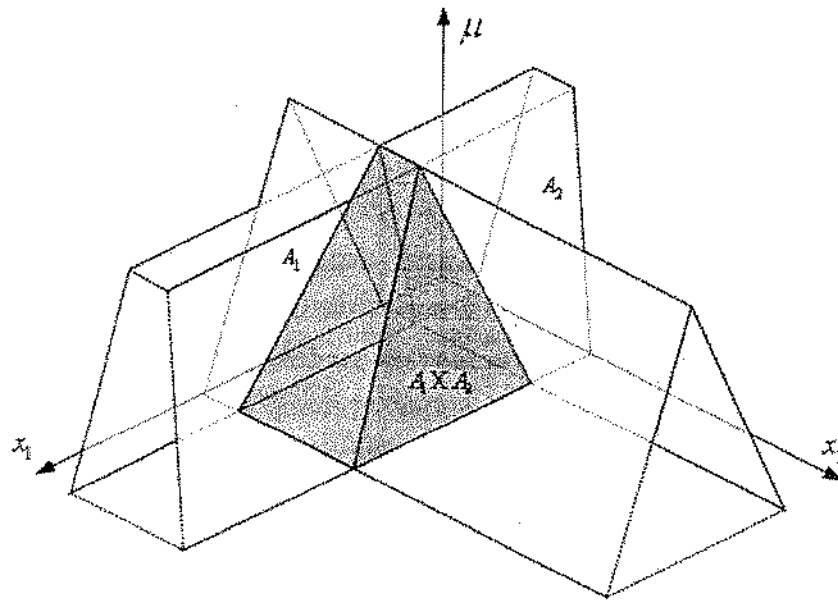


Fig. 3.3: Intersecção de dois conjuntos fuzzy A_1 e A_2 (em escuro).

hiperplano do conseqüente do modelo é válido. Neste sentido, os modelos TS afins podem aproximar a região de “interpolação” por hiperplanos. Um caso especial de (3.9) existe para funções constantes $y_i = b_i$, $i = 1, \dots, K$. Este modelo é chamado de um modelo singleton, conforme apresentado na seção 3.3.1.

Para a descrição anterior do modelo foram utilizadas funções lineares e afins estáticas. Porém, o modelo fuzzy TS pode ter qualquer tipo de função no conseqüente, como funções não-lineares e mesmo relações dinâmicas como equações diferenciais, equações à diferenças, etc. Neste trabalho as funções nos conseqüentes serão FBO's, as quais já foram apresentadas na seção 2.1.5.

3.4.1 Inferência em Modelos Takagi-Sugeno

Inferência de sistemas fuzzy baseados em regras é um processo de determinação da saída, dadas as regras e conhecidas as entradas, no qual se usa uma composição das regras de inferência (Hellendoorn e Driankov, 1997). A Fig. 3.4 mostra o processo de *fuzzificação* dos valores de entrada, a ativação das regras e os valores das saídas nas funções presentes nos conseqüentes das regras para o caso particular de modelos TS.

Como os conseqüentes em um modelo TS não são conjuntos fuzzy, mas funções crisp, a inferência e *defuzzificação* são reduzidas à média ponderada das saídas pela ativação das respectivas regras, dada por:

$$\hat{y} = \frac{\sum_{i=1}^K \mu_i(x) y_i}{\sum_{i=1}^K \mu_i(x)} \quad (3.10)$$

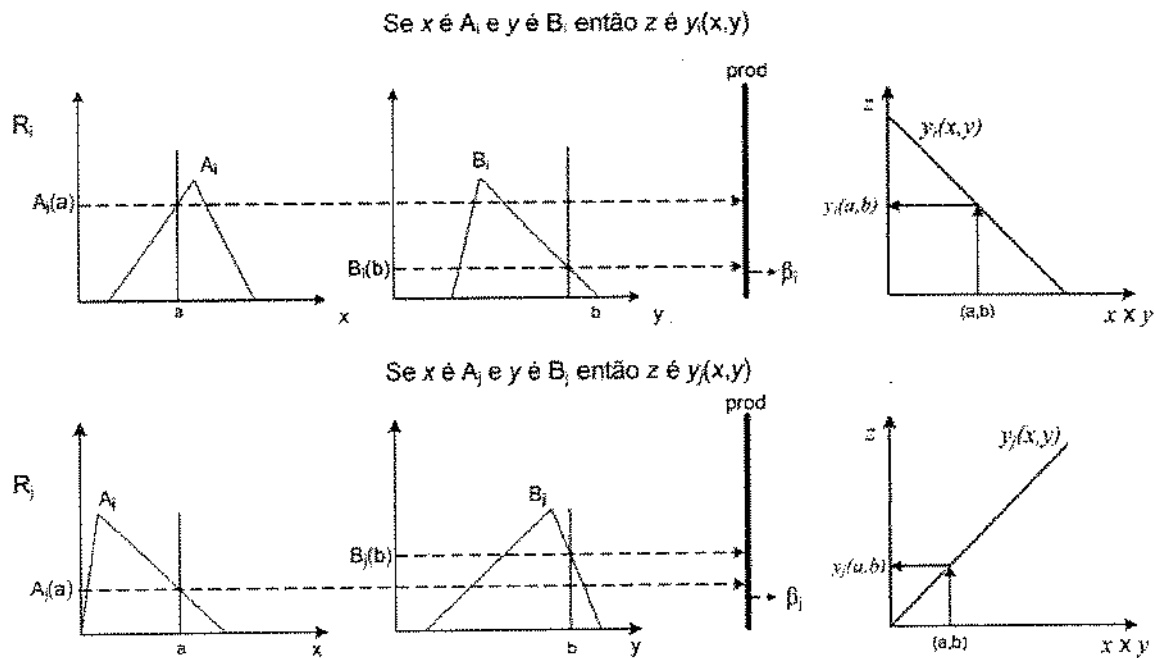


Fig. 3.4: Inferência *fuzzy* TS com operador prod. para conjunção.

Na equação (3.10), $\mu_i(x)$ é o grau de ativação do antecedente da i -ésima regra e y_i é a saída do modelo local (consequente) da i -ésima regra. Denotando a normalização do grau de ativação por:

$$\bar{\mu}_i(x) = \frac{\mu_i(x)}{\sum_{j=1}^K \mu_j(x)}$$

o modelo global afim pode ser escrito como:

$$\hat{y} = \left(\sum_{i=1}^K \bar{\mu}_i(x) a_i^T \right) x + \sum_{i=1}^K \bar{\mu}_i(x) b_i \quad (3.11)$$

que pode ser escrito como um modelo com parâmetros dependentes da entrada:

$$\hat{y} = a(x)^T x + b(x) \quad (3.12)$$

onde os parâmetros $a(x)$, $b(x)$ são obtidos através da combinação linear convexa dos parâmetros dos consequentes a_i e b_i , isto é:

$$a(x) = \sum_{i=1}^K \bar{\mu}_i(x) a_i, \quad b(x) = \sum_{i=1}^K \bar{\mu}_i(x) b_i$$

Contudo, neste trabalho é de interesse se construir modelos TS com FBO nos consequentes devido as vantagens desta modelagem já apresentadas na seção 2.1.5. Estas vantagens aliadas as ca-

racterísticas dos modelos TS, possibilitam aos modelos TS-FBO a capacidade de modelar processos com comportamento não-linear através de modelos simples e com vantagens estruturais como a simplicidade do modelo resultante, onde a determinação da ordem dos modelos locais não é um problema crítico. Outra vantagem da modelagem TS-FBO é a interpretabilidade matemática, muito útil no projeto de sistemas de controle. Para o modelo TS com FBO no conseqüente a função $y = f(x)$ em (3.5) deve ser substituída por um modelo local FBO como aquele em (2.25). Para casos em que se utiliza uma mesma base FBO para todos os modelos locais, faz-se necessário somente o cálculo da parte estática da formulação para cada modelo uma vez que a parte dinâmica é a mesma para todos os modelos (Oliveira, Campello e Amaral, 1999; Campello e Amaral, 2002; Campello, Meleiro e Amaral, 2004; Campello, 2002; Medeiros et al., 2006; Medeiros, 2006):

$$\begin{aligned} \text{Se } l_1(k) \text{ é } A_{1,1} \text{ e } l_2(k) \text{ é } A_{1,2} \text{ e } \dots \text{ e } l_d(k) \text{ é } A_{1,d} & \text{ então } y_1 = C_1^T l(k) \\ \text{Se } l_1(k) \text{ é } A_{2,1} \text{ e } l_2(k) \text{ é } A_{2,2} \text{ e } \dots \text{ e } l_d(k) \text{ é } A_{2,d} & \text{ então } y_2 = C_2^T l(k) \\ & \vdots \\ \text{Se } l_1(k) \text{ é } A_{K,1} \text{ e } l_2(k) \text{ é } A_{K,2} \text{ e } \dots \text{ e } l_d(k) \text{ é } A_{K,d} & \text{ então } y_K = C_K^T l(k) \end{aligned}$$

onde $l(k)^T = [l_1(k) \ l_2(k) \ \dots \ l_d(k) \ 1]$, d é o número de estados da base FBO e K é o número de regras no modelo TS. O valor unitário no final do vetor representa o *offset* do modelo afim.

O valor de ativação de cada regra para cada instante k é então dado por:

$$\mu_j(k) = \mu_{A_{j,1}}(l_1(k)) \cdot \mu_{A_{j,2}}(l_2(k)) \cdot \dots \cdot \mu_{A_{j,d}}(l_d(k)), \quad j = 1, \dots, K. \quad (3.13)$$

A saída do modelo geral, no instante k , é dada pela inferência *fuzzy* das saídas dos modelos locais segundo a equação (3.10):

$$\hat{y} = \frac{\sum_{i=1}^K (\mu_i(k) C_i l(k))}{\sum_{i=1}^K \mu_i(k)} \quad (3.14)$$

Sendo assim, é conveniente montar um vetor que atribui os relativos “pesos” a cada saída local na composição da resposta global do modelo de TS num dado instante k :

$$\Phi(k) = [\bar{\mu}_1(k) \ \bar{\mu}_2(k) \ \dots \ \bar{\mu}_K(k)] \quad (3.15)$$

onde:

$$\bar{\mu}_i(k) = \frac{\mu_i(k)}{\sum_{j=1}^K \mu_j(k)} \quad (3.16)$$

A partir desta nova representação matricial pode-se então reescrever a equação (3.14) como:

$$\hat{y}(k) = \Phi(k) \cdot C \cdot l(k) \quad (3.17)$$

sendo $C^T = [C_1 \ C_2 \ \dots \ C_K]$.

3.5 Agrupamento

Como já foi citado anteriormente, um conjunto de regras *fuzzy* pode ser obtido através de informações fornecidas por um especialista ou então diretamente extraído de dados obtidos do processo, como, por exemplo, a resposta do processo a determinado sinal de excitação. Esta última abordagem, a qual se trata de uma abordagem *bottom-up*, será a utilizada neste trabalho e será apresentada a seguir.

Uma forma de se obter as regras a partir dos dados do sistema é utilizando algoritmos de agrupamento *fuzzy* (*fuzzy clustering*). Agrupamento de dados numéricos forma a base de muitos algoritmos de classificação e modelagem. O intuito do agrupamento é identificar grupos naturais de dados a partir de uma grande massa de dados para produzir uma representação concisa do comportamento do sistema. Os principais algoritmos de agrupamento *fuzzy* podem ser classificados nos três seguintes grupos (Hellendoorn e Driankov, 1997):

- i. Algoritmos que utilizam uma medida de distância adaptativa, tais como Gustafson-Kessel (Gustafson e Kessel, 1979) ou algoritmos de estimação de máxima verossimilhança *fuzzy* (Gath e Geva, 1989).
- ii. Algoritmos baseados em protótipos hiperplanares, também conhecidos como variedade linear *fuzzy* e *c*-elipsóides *fuzzy* (Bezdek, 1981b; Bezdek, 1981c; Bezdek, 1981a).
- iii. Algoritmos de *c*-regressão, os quais utilizam protótipos definidos por funções de regressão (Hathaway e Bezdek, 1993).

Todas as três classes de algoritmos podem ser vistas como uma extensão do algoritmo *c-means* (Bezdek, 1981a), que minimiza a variância dos dados para os protótipos (centros) dos grupos, sendo assim algoritmos de otimização não-linear como será descrito na seção 3.5.1. A principal diferença entre estes algoritmos está na definição de distância utilizada para estabelecer os grupos, o que influencia na sua estrutura e formato. Neste trabalho dar-se-á maior atenção ao algoritmo de Gustafson-Kessel, que será apresentado em detalhes na seção a seguir. As informações obtidas a partir dos grupos serão, então, utilizadas para obter modelos *fuzzy* TS que representem de maneira satisfatória o sistema sob estudo, com o menor número de grupos (i.e. regras/ modelos locais) possível.

3.5.1 Algoritmo de Gustafson-Kessel

Seja um conjunto de dados de entrada-saída medidos do sistema (x_j, y_j) , $j = 1, \dots, N$, onde $x \in \mathbb{R}^d$ são as d -variáveis nas premissas, y a saída do sistema e N o número de medidas. Denotando $z_j = [x_j, y_j]^T$, pode-se então escrever o conjunto de dados de forma matricial:

$$Z = [z_1, \dots, z_N], \quad z_j \in \mathbb{R}^{d+1} \quad (3.18)$$

Os vetores z_j serão particionados em K grupos com seus protótipos (centros) dados por:

$$v_i = [v_{i,1}, \dots, v_{i,d+1}]^T \in \mathbb{R}^{d+1}, \quad i = 1, \dots, K. \quad (3.19)$$

Ao conjunto dos protótipos dos K grupos pode-se dar a denotação: $V = [v_1, \dots, v_K]$. O particionamento dos dados é definido pela matriz de particionamento *fuzzy* dada por:

$$U = [\mu_{i,j}]_{K \times N}$$

onde $\mu_{i,j} \in [0, 1]$ representa o grau de pertinência da j -ésima medida z_j ao grupo (*fuzzy cluster*) com protótipo v_i . O algoritmo de Gustafson-Kessel (GK) é um dentre aqueles que encontram a matriz de particionamento e os protótipos dos grupos pela minimização da seguinte função objetivo (Bezdek, 1981a; Babuska, Van der Veen e Kaymak, 2002):

$$J(Z, V, U) = \sum_{i=1}^K \sum_{j=1}^N \mu_{i,j}^m d^2(z_j, v_i) \quad (3.20)$$

sujeito a

$$\sum_{i=1}^K \mu_{i,j} = 1, \quad j = 1, \dots, N, \quad (3.21)$$

$$0 < \sum_{j=1}^N \mu_{i,j} < N, \quad i = 1, \dots, K. \quad (3.22)$$

onde $m > 1$ é um parâmetro que controla a *fuzzificação* do grupo, onde valores maiores de m acarretam uma maior sobreposição dos grupos. Tipicamente o valor $m = 2$ é utilizado (Babuska e Verbruggen, 1997). A função $d(z_j, v_i)$ é a distância do vetor z_j com relação ao protótipo v_i do grupo. A restrição (3.21) impede uma solução trivial $U = 0$ e a restrição (3.22) garante que os grupos não sejam vazios e nem contenham todas as medidas com grau 1. Esta é uma definição padrão de agrupamento com uma função de objetivo *fuzzy* em que a minimização da função objetivo deve ser feita através de um processo iterativo (Bezdek, 1981a). O formato geométrico do grupo é determinado pelo tipo de

função $d(z_j, v_i)$ utilizada. Gustafson e Kessel (1979) generalizaram o algoritmo de agrupamento para um método de distância adaptativa:

$$d^2(z_j, v_i) = (z_j - v_i)^T M_i (z_j - v_i) \quad (3.23)$$

onde M_i é uma matriz definida positiva adaptada de acordo com a forma atual do grupo, que pode ser descrita em função da matriz de covariância do grupo, F_i :

$$F_i = \frac{\sum_{j=1}^N \mu_{i,j}^m (z_j - v_i)(z_j - v_i)^T}{\sum_{j=1}^N \mu_{i,j}^m} \quad (3.24)$$

A matriz M_i utilizada em (3.23) é calculada como a matriz inversa normalizada da matriz de covariância do grupo, dada por:

$$M_i = \det(F_i)^{\frac{1}{d+1}} F_i^{-1} \quad (3.25)$$

A normalização dada pelo determinante de F_i é utilizada para impedir uma solução trivial de (3.20) para $M_i = 0$.

Algorithm 1 Algoritmo de Gustafson-Kessel

Dado um conjunto de dados Z , o número de clusters K e uma matriz de particionamento U inicialmente aleatória, os protótipos dos clusters V e a matriz de particionamento final U são encontrados através da repetição do seguinte algoritmo:

- 1: Calcule os protótipos dos grupos: $v_i = \frac{\sum_{j=1}^N \mu_{i,j}^m z_j}{\sum_{j=1}^N \mu_{i,j}^m}$
- 2: Calcule as matrizes de covariância F_i usando (3.24)
- 3: Obtenha as matrizes M_i usando (3.25)
- 4: Calcule as distâncias quadráticas $d^2(z_j, v_i)$ como dado em (3.23)
- 5: Atualize as matrizes de particionamento U : $\mu_{i,j} = \frac{d^2(z_j, v_i)^{\frac{-1}{m-1}}}{\sum_{l=1}^K d^2(z_j, v_l)^{\frac{-1}{m-1}}}$

Se $d^2(z_j, v_i) = 0$ para algum $i = k$, faça $\mu_{k,j} = 1$ e $\mu_{i,j} = 0, \forall i \neq k$

Repita o procedimento acima até que a convergência seja obtida, isto é, $\|U_l - U_{l-1}\| < \epsilon$ onde $\|\cdot\|$ é uma norma matricial, l é o passo de iteração e ϵ é uma tolerância adotada para o término da otimização, tipicamente $\epsilon = 0,01$ ou $\epsilon = 0,001$ (Babuska e Verbruggen, 1997).

Após a convergência do algoritmo 1, a matriz de particionamento U , os protótipos v_i dos grupos e as matrizes de covariância $F_i, i = 1, \dots, K$ são obtidos.

3.5.2 Métodos de determinação do número de grupos

Além de como determinar os grupos *fuzzy*, outro fator de grande importância é determinar o número ideal de grupos para que o menor número possível possa produzir o melhor agrupamento

dos dados. Na literatura existem diversos critérios para se determinar a quantidade ideal de grupos. Alguns deles, utilizados neste trabalho, são apresentados a seguir.

Inspirado no conceito de que um bom agrupamento é aquele que não apresenta grupos muito dispersos, Gath e Geva (1989) propuseram medidas de validação baseadas no critério de hipervolume ou densidade do agrupamento *fuzzy*. Mais especificamente, o critério de avaliação conhecido como *Fuzzy Hypervolume* (FHV) é dado por:

$$FHV = \sum_{i=1}^K [\det(F_i)]^{-1/2} \quad (3.26)$$

onde F_i é a matriz de covariância do i -ésimo grupo *fuzzy*, dada por:

$$F_i = \frac{\sum_{j=1}^N h(i|z_j)(z_j - v_i)(z_j - v_i)^T}{\sum_{j=1}^N h(i|z_j)} \quad (3.27)$$

Sob a perspectiva de uma estimação de máxima verossimilhança, $h(i|z_j)$ é a probabilidade de seleção do i -ésimo grupo dada a j -ésima amostra z_j . Quando o valor mais utilizado do parâmetro *fuzzificador* é escolhido, $m = 2$, esta probabilidade aproxima-se do grau de ativação da j -ésima amostra ao i -ésimo grupo (Gath e Geva, 1989). Portanto, a matriz de covariância em (3.27) pode ser reescrita como:

$$F_i = \frac{\sum_{j=1}^N \mu_{i,j}(z_j - v_i)(z_j - v_i)^T}{\sum_{j=1}^N \mu_{i,j}} \quad (3.28)$$

Assim, como os autovalores de F_i estão diretamente relacionados à variância do i -ésimo grupo, o determinante em (3.26), dado pelo produto dos autovalores de F_i , fornece a medida da dispersão n -dimensional (hipervolume) do grupo. Bons agrupamentos *fuzzy* são aqueles que apresentam menor valor de FHV, ou seja, aqueles que apresentam grupos *fuzzy* mais compactos.

A Densidade média de particionamento (APD¹) é outro critério de qualidade de agrupamento *fuzzy* que pode ser utilizado para determinação do número ideal de grupos. Este critério é definido através da seguinte equação:

$$APD = \frac{1}{K} \sum_{i=1}^K \frac{R_i}{[\det(F_i)]^{1/2}} \quad (3.29)$$

onde K é o número de grupos e R_i é a "soma dos números centrais", definida por:

$$R_i = \sum_j \mu_{i,j}, \forall j \text{ tal que } (z_j - v_i)^T F_i^{-1} (z_j - v_i) < 1 \quad (3.30)$$

que é a soma dos valores de pertinência das amostras j localizados dentro de um raio de proximidade

¹APD - Average Partition Density

do grupo i . Como para grupos mais compactos o valor do determinante será menor e R_i será maior, se conclui pela análise de (3.29) que os melhores agrupamentos serão aqueles que apresentarem maior APD.

Os critérios FHV e APD são critérios que de maneira geral não apresentam resultados satisfatórios quando se trabalha com muitos atributos (variáveis e/ou dimensões). Este fato acontece porque estes critérios exigem o cálculo do determinante e/ou inversa (ver (3.26), (3.29) e (3.30)) da matriz de covariância F_i . Os cálculos do determinante e da inversão demandam um grande esforço computacional ($O(n_1^3)$ com n_1 sendo o número de parâmetros) e um grande número de parâmetros na matriz F_i pode causar problemas numéricos no cálculo de tal determinante e/ou inversa (Campello e Hruschka, 2006).

Outro critério que pode ser utilizado para avaliação do número ideal de grupos *fuzzy* é o que faz a medida de distância média interna de cada grupo (AWCD²), dada pela equação:

$$AWCD = \frac{1}{K} \sum_{i=1}^K \frac{\sum_{j=1}^N \mu_{i,j}^m \|x_j - v_i\|^2}{\sum_{j=1}^N \mu_{i,j}^m} \quad (3.31)$$

onde $\|\cdot\|$ e m se referem à mesma norma e ao parâmetro *fuzzificador* utilizados no algoritmo de agrupamento, respectivamente. Este método utiliza como avaliação o valor médio das distâncias internas dos grupos. A distância interna de um determinado grupo é dada pela média ponderada das distâncias entre todas as amostras e o protótipo (centro) do grupo com cada distância ponderada pelo grau de pertinência da amostra ao grupo (μ_{ij}). Os valores de AWCD decrescem monotonicamente, no caso ideal, conforme o número de grupos aumenta, estando o valor ideal de grupos no ponto onde o gráfico referente aos valores de AWCD descreve um “joelho” mais acentuado. Na seção 3.6.3 será apresentado um exemplo que ilustra esta condição.

Com a finalidade de se determinar de maneira automática onde se encontra o “joelho” da curva no gráfico gerado pelo critério AWCD foi desenvolvida uma técnica levando em consideração as informações de primeira ordem da curva. São obtidas as derivadas das retas que interligam os pontos no gráfico e calculada a variação entre as derivadas de duas retas adjacentes no ponto. Assim, o ponto em que esta variação apresentar maior valor indicará que há uma maior mudança na inclinação da curva que interliga os pontos, ou seja, neste ponto é onde se obtém o melhor agrupamento dos dados. Com os valores das variações é traçado um novo gráfico onde agora não mais se deve analisar o “joelho” da curva e sim o ponto onde há a maior variação da informação de primeira ordem do método originalmente proposto.

O quarto critério de determinação do número ideal de grupos considerado neste trabalho é conhecido como *Fuzzy Silhouette* - FS (Campello e Hruschka, 2006) o qual é uma generalização *Fuzzy* do

²AWCD - *Average Within-Cluster Distance* - AWCD

critério *Average Silhouette Width Criterion* ou *Crisp Silhouette* (Kaufman e Rousseeuw, 1990; Everitt, Landau e Leese, 2001). Para definir este critério, considere as amostras $j \in \{1, 2, \dots, N\}$ pertencentes ao grupo $p \in \{1, \dots, K\}$. Sob o ponto de vista do contexto de agrupamentos *crisp*, isto significa que a amostra j está mais próximo ao protótipo do grupo p do que do protótipo de qualquer outro grupo. No entanto, sob o ponto de vista do agrupamento *fuzzy*, isso significa que a amostra j pertence ao grupo p com grau $\mu_{p,j}$ maior do que pertence a outro grupo *fuzzy*, ou seja, $\mu_{p,j} > \mu_{q,j} \forall q \in \{1, \dots, K\}, q \neq p$.

Tomemos a distância média da amostra j para todos as outras amostras pertencentes ao grupo p sendo denotada por a_{pj} . Também, tomemos a distância média desta amostra para todos as amostras pertencentes a um outro grupo q , com $q \neq p$, denotada por d_{qj} . Finalmente, seja b_{pj} o menor d_{qj} calculado sobre os grupos $q = 1, \dots, K, q \neq p$, que representa a dissimilaridade da amostra j para o grupo vizinho mais próximo do seu grupo. Então, a silhueta da amostra j é definida como:

$$s_j = \frac{b_{pj} - a_{pj}}{\max(a_{pj}, b_{pj})} \quad (3.32)$$

Levando-se em consideração o grau de ativação da amostra j dentro dos grupos, pode-se construir um critério generalizado chamado *fuzzy silhouette*, ou silhueta *fuzzy*, definido por:

$$FS = \frac{\sum_{j=1}^N (\mu_{p,j} - \mu_{q,j})^\alpha s_j}{\sum_{j=1}^N (\mu_{p,j} - \mu_{q,j})^\alpha} \quad (3.33)$$

onde s_j é a silhueta da amostra j de acordo com a equação (3.32), $\mu_{p,j}$ e $\mu_{q,j}$ são respectivamente o primeiro e o segundo maior elementos da j -ésima coluna da matriz de partição *fuzzy* e $\alpha \geq 0$ é um coeficiente que dá maior ou menor peso à diferença $(\mu_{p,j} - \mu_{q,j})$. Quando $\alpha = 0$ a equação (3.33) passa a ser uma média aritmética dos termos s_j , não levando em consideração as informações do agrupamento *fuzzy*, o que corresponde ao critério da silhueta *crisp* original. O valor típico de α para o caso *fuzzy* é 1 (Campello e Hruschka, 2006).

A diferença $(\mu_{p,j} - \mu_{q,j})$, que corresponde à diferença entre o primeiro e o segundo maiores valores de pertinência da amostra j com relação aos grupos, tem o objetivo de dar maior peso aos dados pertencentes a determinado grupo com maior proximidade do protótipo do grupo do que àqueles em regiões de mais acentuada sobreposição entre os grupos. O valor ideal de grupos é sugerido pelo maior valor de FS.

A análise isolada dos critérios de determinação do número de grupos não satisfaz o objetivo deste trabalho em determinar o número de grupos (por consequência o número de regras e de modelos locais) uma vez que cada critério pode fornecer um valor diferente como ideal (Gath e Geva, 1989; Campello e Hruschka, 2006) deixando em dúvida de qual valor adotar. Para a determinação do

número ideal de grupos foi então desenvolvido um método que concilia os critérios acima citados, de acordo com suas características, quando aplicados aos problemas sob estudo. O método desenvolvido consiste primeiramente em executar o algoritmo GK de agrupamento para cada valor de grupos especificado em um dado intervalo de número de grupos aceitável pré-especificado pelo usuário. Ao final de cada execução do algoritmo de GK realiza-se uma análise com os critérios apresentados anteriormente e o desempenho de cada um é então armazenado. Os resultados de cada critério são então normalizados e, especificamente para o critério FHV, os valores dos resultados são subtraídos de 1 para que este método também tenha seu valor ideal de grupos dado pelo maior valor. Com os critérios compreendidos no intervalo $[0, 1]$, calcula-se então a média aritmética e a multiplicação entre os valores de cada critério para um dado número de grupos. Analisando a média e o produto destes resultados, obtém-se informações para a decisão do valor adequado do número de grupos. Utiliza-se a multiplicação dos critérios porque ele permite excluir agrupamentos que apresentam avaliação com valores baixos, embora se para algum agrupamento houver algum critério com avaliação errônea, apresentando valor muito baixo ou nulo, este agrupamento será penalizado na avaliação global. Neste sentido a média dos critérios permite completar a avaliação, uma vez que ela não exclui agrupamentos com algum(ns) valor(es) de avaliação baixo(s) e valoriza os agrupamentos que apresentam avaliações com maiores valores.

A seguir será discutido o procedimento adotado para se obter as funções de pertinência dos antecedentes de modelos *fuzzy* a partir dos agrupamentos realizados.

3.6 Construção dos modelos *Fuzzy*

Uma vez determinados os grupos, é então necessário se obter as regras do modelo TS que determinarão os modelos locais. Assim deve-se determinar os antecedentes e os conseqüentes das regras. Neste trabalho as funções de pertinência nos antecedentes das regras serão funções trapezoidais obtidas através das projeções dos grupos *fuzzy*, determinados na etapa de agrupamento, como será visto na seção 3.6.1. Para os conseqüentes é necessário se determinar os parâmetros da função que foi escolhida para estar no conseqüente dos respectivos modelos locais em uma abordagem do tipo TS. Para este trabalho, como será discutido na seção 3.6.2, serão utilizadas FBO's nos conseqüentes e dois métodos serão apresentados para a determinação dos parâmetros.

3.6.1 Determinação dos antecedentes das Regras

As medidas de distância (3.23) definem os grupos como hiperelipsóides cujo formato é descrito através das matrizes F_i . Os autovetores de F_i definem a orientação dos eixos da hiperelipsóide e

o comprimento dos eixos é dado pelos autovalores associados aos autovetores. O agrupamento dos dados ao redor da superfície de regressão resulta em hiperelipsóides aproximadamente planas, que podem ser vistas localmente como hiperplanos. O autovetor $\Phi_i = [\phi_{i,1} \dots \phi_{i,d+1}]$, associado ao menor autovalor da i -ésima matriz de covariância F_i , é normal ao hiperplano correspondente, como ilustrado na Fig. 3.5³.

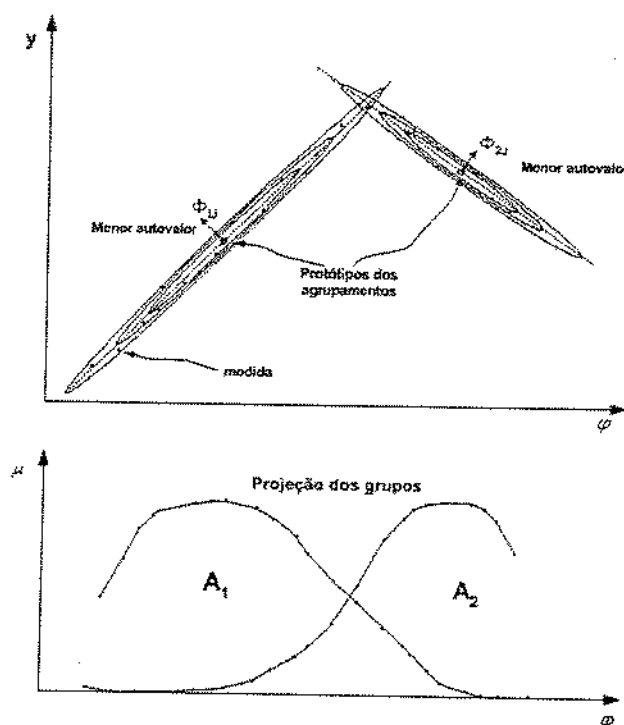


Fig. 3.5: Agrupamento dos dados por GK.

Os antecedentes A_i das regras *fuzzy* são obtidos a partir da projeção dos valores de ativação dos objetos com relação aos grupos que definem as regras em cada dimensão do espaço de entrada (i -ésima coluna da matriz de particionamento U , $\mu_{i,j} : (\mathcal{D} \times \mathcal{Y}) \rightarrow [0, 1]$) como mostra o gráfico na parte inferior da Fig. 3.5. Estes antecedentes serão compostos pelas funções de pertinência que aproximam a casca convexa das projeções por funções convexas, podendo estas, serem funções trapezoidais, triangulares, gaussianas, etc. Na literatura há diversas referências quanto aos métodos de se obter e otimizar tais funções, como o método do gradiente descendente (Espinosa, Vandewalle e Wertz, 2004), através da dispersão da projeção dos valores de ativação (Hellendoorn e Driankov, 1997) ou outros métodos de otimização não-linear (Campello, 2002).

Neste trabalho é de interesse a aproximação da projeção dos objetos por funções de pertinência trapezoidais, como mostra a Fig. 3.6. Este interesse se deve ao fato das funções trapezoidais, ao contrário das triangulares e gaussianas, possuírem intervalos de pertinência unitária (*core não singleton*) que permite a um único modelo local (regra) representar sozinho uma determinada região.

³Esta figura é baseada na figura encontrada em (Babuska e Verbruggen, 1997) página 86.

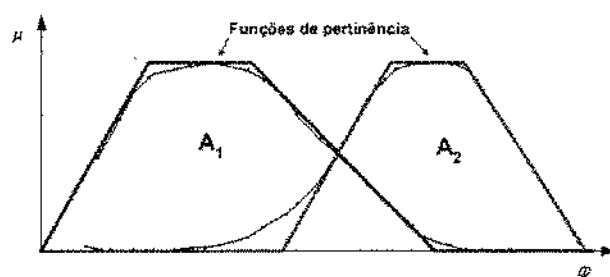


Fig. 3.6: Funções de pertinência geradas a partir das projeções.

Para obtenção das funções de pertinência foi utilizado um método iterativo que utiliza as projeções dos valores de ativação dos objetos com relação aos grupos *fuzzy* no(s) eixo(s) do espaço de entrada. Ao final das iterações é obtida uma casca convexa para os dados de projeção e a partir de um processo de otimização por mínimos quadrados é obtida a função de pertinência trapezoidal.

3.6.2 Determinação dos consequentes das Regras com FBO's

Como citado na seção 3.4, o consequente de um modelo de TS pode ser composto por diversos tipos de função, como funções lineares, afins, equações diferenciais, etc (Yager e Filev, 1994). Neste trabalho é de interesse se construir os modelos de TS com FBO's lineares ou afins, com representação em espaço de estados nos consequentes das regras. A adoção de modelos FBO lineares ou afins nos consequentes das regras dos modelos *fuzzy* TS-FBO, além de representar com precisão arbitrária uma ampla classe de sistemas não-lineares, apresenta a característica de linearidade nos parâmetros das funções locais além de manter as vantagens estruturais apresentadas para as FBO's.

Para se determinar o valor dos consequentes é necessário estabelecer um pólo (ou par de pólos complexos) que são determinados como o(s) pólo(s) dominante(s). Estando definido o pólo é então aplicado ao modelo FBO, sob a forma matricial (equação (2.25)), a mesma entrada que é aplicada ao sistema real afim de se determinar os estados da base de funções ortonormais nos N instantes de simulação. Dessa maneira estarão disponíveis as saídas do sistema real $y(k)$ e os estados da base ortonormal $l^T(k) = [l_1(k) \ l_2(k) \ \dots \ l_d(k)]$.

Para se determinar os parâmetros da parte estática das representações matriciais das FBO's presentes nos consequentes (a parte dinâmica não precisa ser estimada como descrito em 2.1.5) podem ser utilizadas duas abordagens que empregam MQ, uma em que os parâmetros são estimados globalmente, ou seja, de uma só vez para todos os modelos locais, e outra em que são estimados localmente, onde os parâmetros têm seus valores estimados separadamente para cada modelo local, ou seja, cada regra tem sua função no consequente estimada localmente e independentemente (Ljung, 1999; Abonyi e Babuska, 2000; Delgado, Von Zuben e Gomide, 2001; Johansen, Shorten e Murray-Smith, 2000).

- **Estimação Global** - A partir do erro de estimação $\xi(k) = y(k) - \hat{y}(k)$ pode-se usar a seguinte

função de custo para otimização dos parâmetros:

$$J = \frac{1}{2N} \sum_{k=1}^N \xi(k)^T \xi(k) \quad (3.34)$$

onde N é o número de amostras do sinal e y e $\hat{y}$ são respectivamente as saídas medida e estimadas do sistema.

A função de custo (3.34) pode então ser reescrita sob a forma vetorial:

$$J = \frac{1}{2N} \|Y - \Psi C\|^2 \quad (3.35)$$

onde $Y^T = [y(1) \ y(2) \ \dots \ y(N)]$ é o vetor de saídas reais do sistema, $C^T = [C_1 \ C_2 \ \dots \ C_K]$ é um vetor coluna com os parâmetros dos conseqüentes de cada regra e a matriz é:

$$\begin{aligned} \Psi &= [\psi_{ij}]_{i=1,2,\dots,N; j=1,2,\dots,K} \\ \psi_{kj} &= l(k) \mu_j(k) \end{aligned} \quad (3.36)$$

onde $\mu_j(k)$ é dado em (3.13).

O mínimo para a equação (3.35), conforme demonstrado no Apêndice A, é obtido por:

$$\hat{C} = [\Psi^T \Psi]^{-1} \Psi^T Y \quad (3.37)$$

A inversão em (3.37) representa um problema de complexidade $O((d.K)^3)$ (Murray-Smith, 1994), cuja solução poderá representar um esforço computacional elevado quando o número n de regras é elevado.

- **Estimação Local** - Os parâmetros lineares podem também ser determinados de maneira local utilizando-se K estimadores de mínimos quadrados independentes. Isso pode ser feito porque localmente os modelos têm pouca interação, devido à estrutura local das regiões de validade. Pode-se observar na Fig. 3.6 que em regiões do espaço de entradas onde $\mu_{j_0} \approx 1$, tem-se que $\mu_{j \neq j_0}$ são aproximadamente zero.

A partir do erro de estimação em relação ao j -ésimo modelo linear local, $\xi_j(k) = y(k) - C_j^T l(k)$, pode-se definir as funções de custo dos modelos locais, sendo para cada região de validade dado por:

$$J = \frac{1}{2N} \sum_{k=1}^N \xi_j(k)^T \mu_j(k) \xi_j(k) \quad (3.38)$$

Os termos $\mu_j(k)$, $k = 1, \dots, N$, são introduzidos como termos de ponderação na otimização do critério (3.38). Desta maneira, vetores de entrada (estados da FBO) $l(k)$ que estejam mais próximos dos protótipos dos grupos são ponderados com maior peso nos cálculos dos parâmetros da função no conseqüente da regra correspondente.

Os parâmetros lineares da j -ésima regra, considerando o critério (3.38), são obtidos por (Ljung, 1999):

$$\hat{C}_j = [\Psi^T \Lambda_j \Psi]^{-1} \Psi^T \Lambda_j Y \quad (3.39)$$

onde:

$$\Psi = \begin{bmatrix} l(1)^T \\ l(2)^T \\ \vdots \\ l(N)^T \end{bmatrix}$$

e

$$\Lambda_j = \text{diag} \left(\mu_j(1), \mu_j(2), \dots, \mu_j(N) \right)$$

A inversa em (3.39) representa um problema de complexidade $O(d^3)$, pois a matriz $\Psi^T \Lambda_j \Psi$ é de dimensão $d \times d$. Mas como é necessário fazer K vezes (uma para cada modelo local), tem-se para o processo total uma complexidade $O(K.d^3)$, que é menor que no caso de otimização dos modelos lineares locais com um critério de otimização global ($O(K.d^3)$). A otimização local dos parâmetros lineares apresenta melhores resultados quando os antecedentes das regras estão sobreparametrizados (número grande de regiões de validade), devido a problemas numéricos na resolução do problema de mínimos quadrados com o critério global (Murray-Smith, 1994). Para parâmetros lineares determinados com critério local, os modelos lineares locais associados a cada região de validade são forçados a ajustar localmente os dados, o que não ocorre para a otimização dos parâmetros lineares com um critério global. Sendo assim, os modelos estimados pelo critério local apresentam importantes características de interpretabilidade, importante para o projeto de controladores locais em sistemas de controle multimodelo, os quais serão abordados neste trabalho. Nos trabalhos de Johansen et al. (2000) e Shorten, Murray-Smith, Bjorgan e Gollee (1999) discute-se a interpretação dos modelos locais como aproximações locais na modelagem de Takagi-Sugeno e propõem-se condições na identificação para garantir uma aproximação local adequada. A otimização local é superior nesse contexto.

A próxima seção traz exemplos que mostram como é feito o agrupamento, a determinação do número de grupos e a definição das regras dos modelos locais de TS.

3.6.3 Exemplos

Alguns exemplos serão apresentados nesta seção para de ilustrar a aplicação do método de GK no agrupamento de dados, o método de determinação do número de grupos e a modelagem de um sistema através da técnica proposta neste trabalho.

No primeiro exemplo deseja-se determinar os grupos que melhor agrupam um conjunto de dados apresentados na Fig. 3.7:

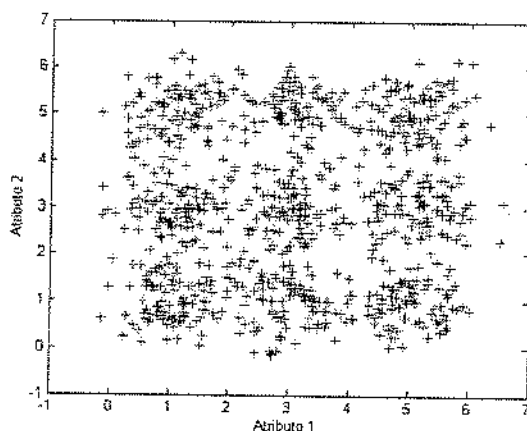


Fig. 3.7: Conjunto de amostras com 9 grupos.

Neste conjunto de dados foram utilizados 900 objetos (amostras) contendo 9 grupos sobrepostos com 100 objetos cada um. Tais objetos foram gerados aleatoriamente a partir de 9 gaussianas bidimensionais, com desvio padrão igual a 0,5, distribuídas numa dada região $\mathcal{R}^2$ (Hruschka, Campello e de Castro, 2004). Desta forma o número ideal de grupos já é conhecido e assim pode-se verificar se o método proposto para determinação do número de grupos funciona corretamente.

Os grupos resultantes do processo de agrupamento foram gerados de maneira automática analisando-se o número de grupos entre 2 e 12. A Fig. 3.8 mostra os resultados da avaliação dos critérios apresentados na seção 3.5.2, onde o critério AWCD já se apresenta com a modificação proposta neste trabalho.

Como se pode constatar através de análise da Fig. 3.8, para todos os critérios de avaliação o resultado apresentado como melhor número de grupos foi 9, sendo indicado pelos maiores valores para os critérios FS e APD, para o menor valor em FHV e o maior valor em AWCD-modificado, que corresponde ao “joelho” da curva para o método AWCD descrito habitualmente na literatura. A Fig. 3.9 mostra o resultado dos métodos propostos de avaliação global através da média e produto entre os resultados normalizados dos critérios e fornece o número ideal de grupos.

Por fim, a Fig. 3.10 mostra os objetos representados pelo símbolo + e as elipses que indicam regiões de fronteira entre os grupos gerados pelo algoritmo de GK. O resultado mostra a eficácia do

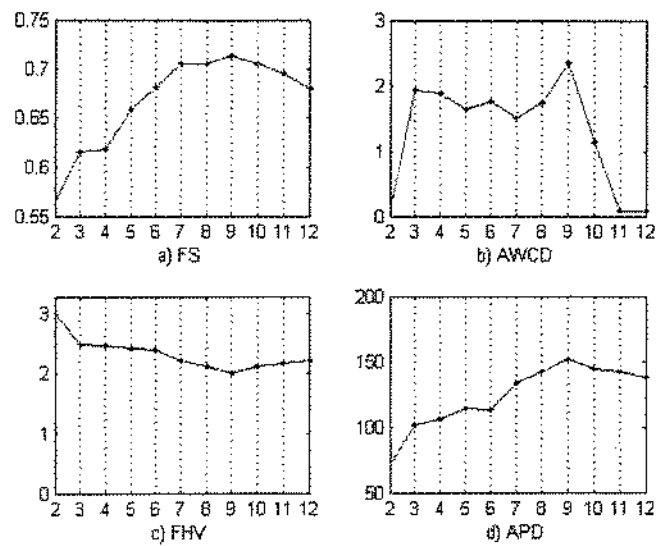


Fig. 3.8: Critérios de avaliação do número ideal de grupos (não normalizados)

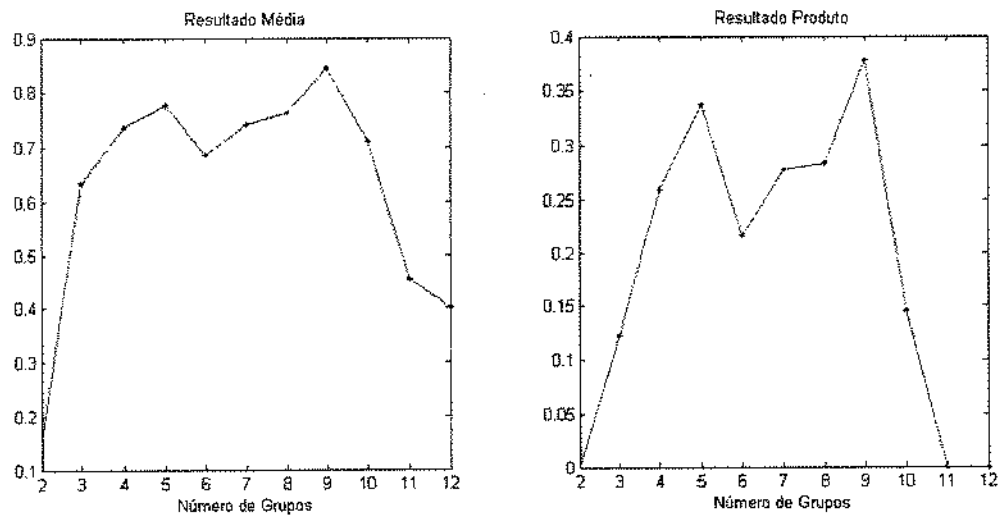


Fig. 3.9: Resultado global do número de grupos via combinação por produto e média.

método proposto e dos critérios de avaliação na determinação do número de grupos.

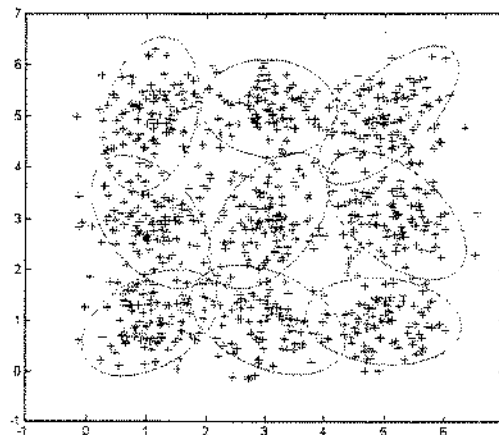


Fig. 3.10: Grupos obtidos com o algoritmo de GK.

Assim como no exemplo anterior, também deseja-se conhecer o melhor número de grupos para um conjunto de dados. Este conjunto é composto por 350 objetos contendo 3 grupos formados por estes objetos. Os objetos são gerados aleatoriamente usando gaussianas 2D distribuídas com desvio padrão 3 (200 objetos), 2 (100 objetos), e 1 (50 objetos) como mostra a Fig. 3.11 (Campello e Hruschka, 2006):

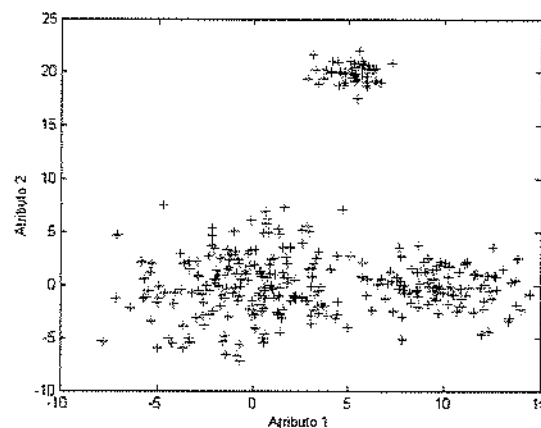


Fig. 3.11: Conjunto de amostras com 3 grupos.

A Fig. 3.12 mostra as respostas individuais dos critérios utilizados na análise dos agrupamentos. Desta vez, ao contrário do que foi apresentado na Fig. 3.8, os gráficos são apresentados normalizados conforme será utilizado nos métodos propostos para se conciliar esses critérios de análise.

Como se pode verificar os critérios FS, AWCD e APD indicam como 3 o melhor número de grupos para agrupar os dados. Já o critério FHV indica como 4 o número ideal de grupos. Como

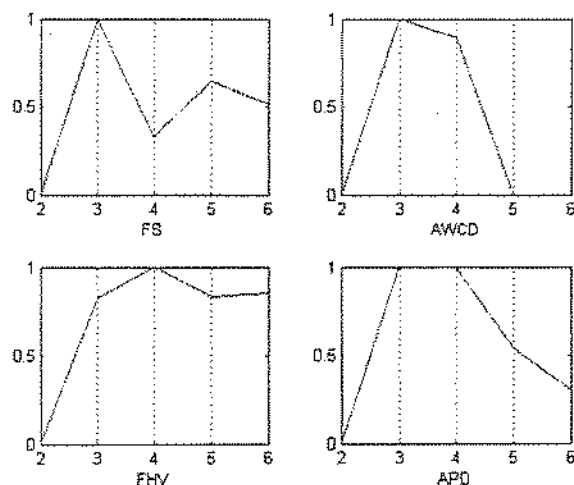


Fig. 3.12: Métodos de avaliação do número de grupos (normalizados)

métodos de se conciliar esses resultados e fornecer uma resposta global no sentido de determinar o melhor número de grupos no agrupamento dos dados, a Fig. 3.13 apresenta a média e o produto dos resultados obtidos pontualmente pelos critérios. Em ambas as avaliações verifica-se que 3 é o número ideal de grupos para agrupar este conjunto de dados.

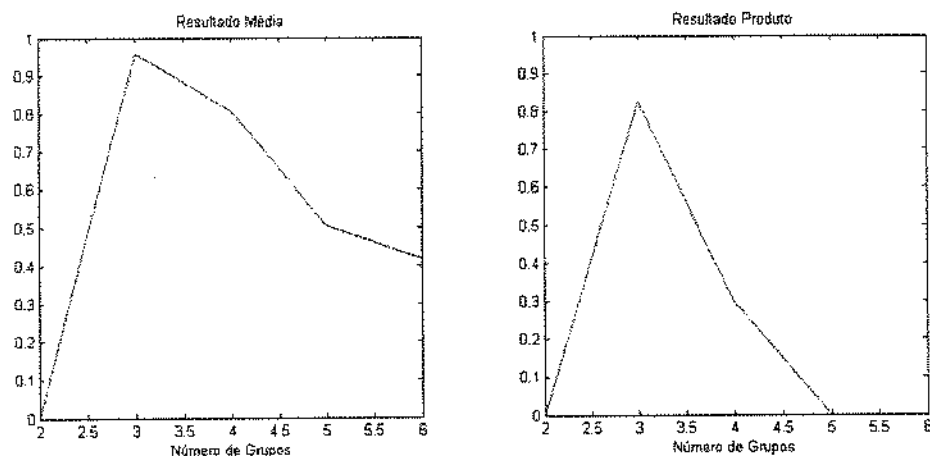


Fig. 3.13: Média e Produto dos critérios de avaliação.

Uma vez determinado que 3 é o número de grupos que melhor agrupa os objetos, os grupos são então recuperados a partir das informações que foram armazenadas durante a execução do algoritmo de GK. A Fig. 3.14 mostra os objetos representados pelo símbolo + e as elipses que indicam regiões de fronteira entre os grupos. Os métodos propostos, que levam em consideração as respostas dos critérios de avaliação dos grupos, funcionaram adequadamente, indicando de maneira esperada, uma vez que os dados foram gerados a partir de 3 grupos, o valor correto de grupos para o agrupamento

dos objetos.

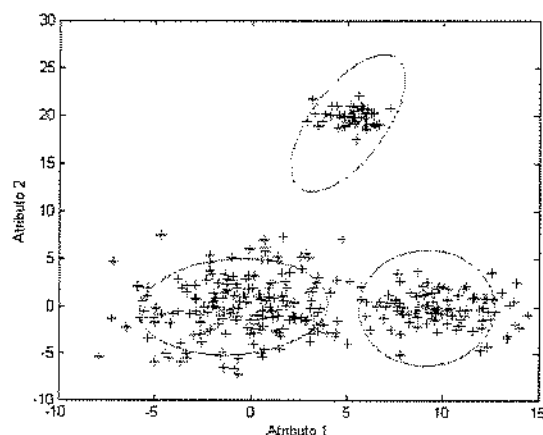


Fig. 3.14: Grupos determinados pelo algoritmo de GK.

O próximo exemplo é inspirado no trabalho de Krishnapuram e Freg (1992), onde o algoritmo de agrupamento *fuzzy* de GK é aplicado no problema de reconhecimento de caracteres. A Fig. 3.15 mostra os objetos utilizados neste exemplo. Estes objetos formam 3 grupos elipsoidais que compõem uma letra “A” com 100 objetos em cada elipsóide. Os objetos foram gerados utilizando-se distribuições gaussianas 2D com desvio padrão na horizontal e na vertical de $(0,5; 0,02)$ para dois grupos e $(1; 0,01)$ para o terceiro. Os dois primeiros grupos tiveram seus eixos rotacionados de 45° e -45° . (Campello e Hruschka, 2006)

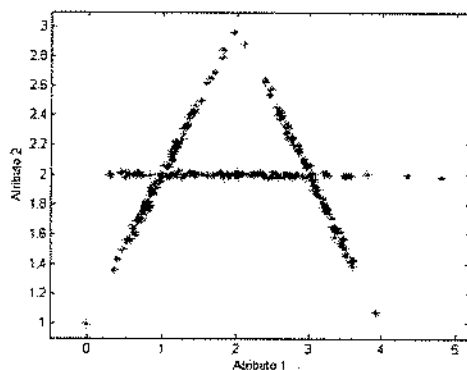


Fig. 3.15: Conjunto de amostras com 3 grupos formando um “A”.

Neste exemplo, o algoritmo de GK foi sequencialmente aplicado aos objetos da Fig. 3.15 analisando o valor de grupos que melhor agrupa os dados entre 1 e 6 grupos. A Fig. 3.16 mostra a média e o produto dos critérios de avaliação do melhor número de grupos para se agrupar os objetos em questão.

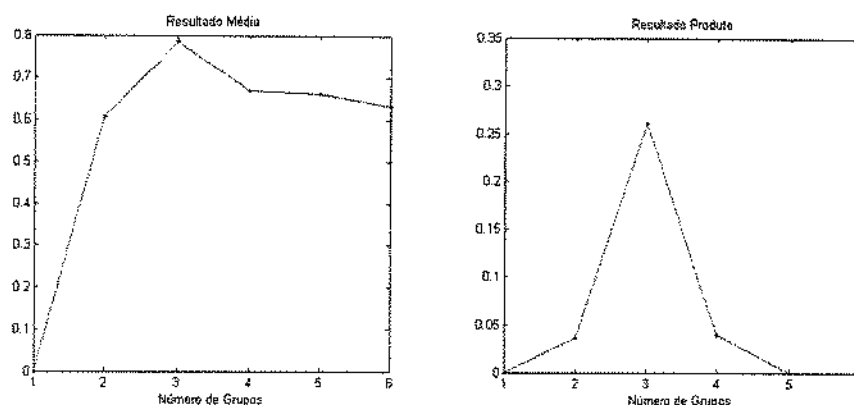


Fig. 3.16: Média e Produto dos critérios de avaliação do núm. de grupos.

Como pode ser observado, tanto na média quanto no produto tem-se que o melhor número de grupos para os objetos é de 3 grupos, exatamente o que se esperava como resultado. A Fig. 3.17 mostra o resultado do agrupamento dos dados onde as elipses indicam os grupos. Portanto, para o agrupamento de análise de objetos formando grupos planares o método proposto também funciona corretamente.

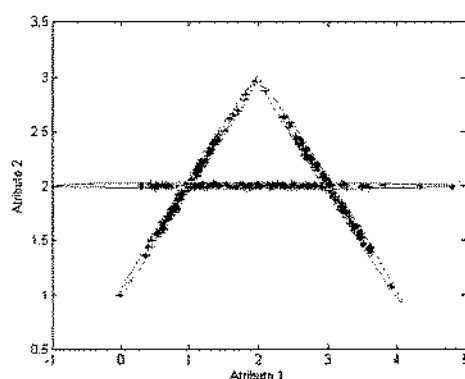


Fig. 3.17: Resultado do agrupamento dado pelo algoritmo de GK.

Para finalizar será apresentado um exemplo em que se realiza a construção de um modelo *fuzzy* TS que irá modelar uma dada função não-linear, baseado em (Setnes, Babuska e Verbruggen, 1998). Considere função dada pela equação (3.40):

$$y(x) = 3e^{-x^2} \cdot \text{sen}(\pi x) + \eta \quad (3.40)$$

onde η é um ruído gaussiano com média zero e desvio padrão dado $\sigma = 0,1$. Como entrada, foram gerados 300 números aleatórios x distribuídos uniformemente no intervalo $[-3, 3]$. A partir da equação (3.40), com as entradas geradas, foram obtidas as saídas $y(x)$ como indicado na Fig.

3.18.

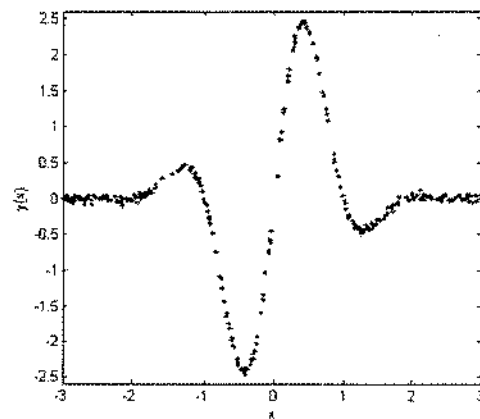


Fig. 3.18: Dados de identificação contaminados com ruído.

Os objetos foram agrupados utilizando-se o algoritmo de GK e, analisando-se entre um número de 2 a 10 grupos para o agrupamento dos objetos, pode-se verificar que 7 grupos é o valor de grupos que melhor agrupa os dados, conforme indicado tanto na análise pela média dos critérios de avaliação de agrupamento quando pelo produto dos mesmos, conforme indicado na Fig. 3.19.

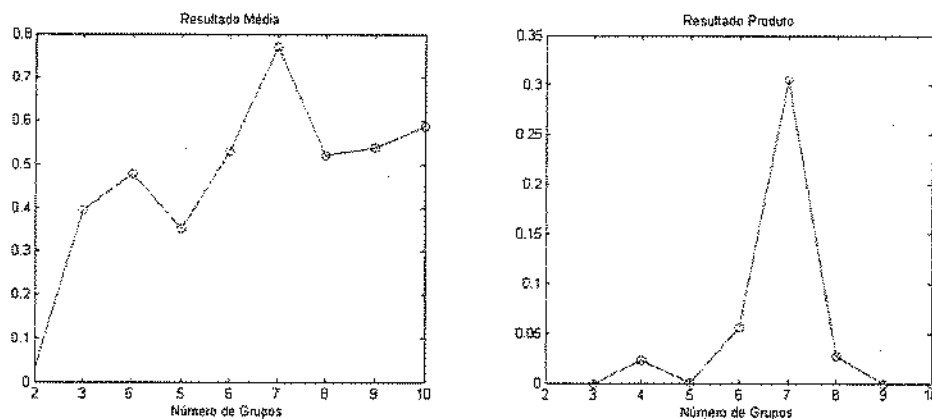


Fig. 3.19: Resultado dado pela combinação dos critérios de análise de agrupamento.

Como pode ser verificado na Fig. 3.20 o algoritmo de agrupamento conseguiu agrupar os dados de forma satisfatória, mesmo os dados apresentando ruídos.

A Fig. 3.21 mostra as funções de pertinência para os antecedentes das regras obtidas a partir das projeções dos valores de ativação dos objetos, nos respectivos grupos, no espaço de entrada. As funções de pertinência cobrem todo o universo de discurso, garantindo que em todo o universo de discurso haverá pelo menos uma regra ativada.

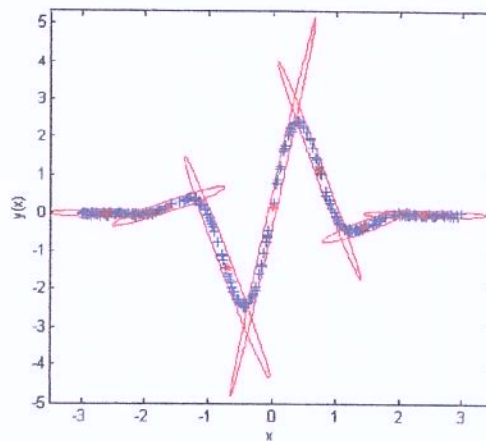


Fig. 3.20: Resultado do agrupamento dos objetos por GK.

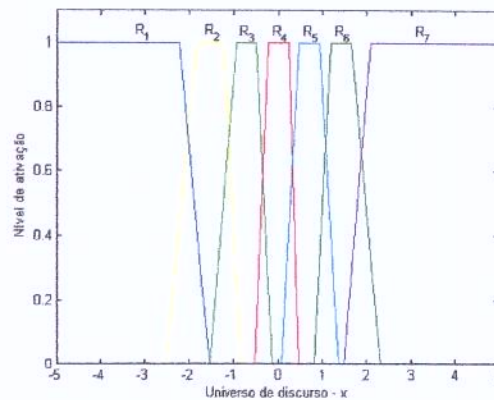


Fig. 3.21: Funções de pertinência dos antecedentes das regras.

Uma vez determinados os antecedentes das regras o próximo passo então consiste em se determinar os conseqüentes das mesmas regras. Para este modelo TS serão utilizadas funções afins do tipo $y(x) = a.x + b$, que irão aproximar a função não-linear localmente. Determinadas as funções dos conseqüentes das regras, via mínimos quadrados, foi então feita a estimação da saída do modelo, como apresentado na Fig. 3.22.

Como pode ser verificado visualmente no gráfico da Fig. 3.22, obteve-se um bom modelo para a função que se desejava mapear. O erro quadrático médio - EQM entre o valor estimado pelo modelo e o valor calculado através da função analítica dada em (3.40) é de 0,0047.

Estes exemplos ilustram que, os métodos propostos apresentaram resultados satisfatórios, tanto no agrupamento *fuzzy* dos dados quanto na determinação dos modelos *fuzzy* TS. Posteriormente, no capítulo 5, serão apresentados outros exemplos da aplicação destes conceitos na modelagem e no

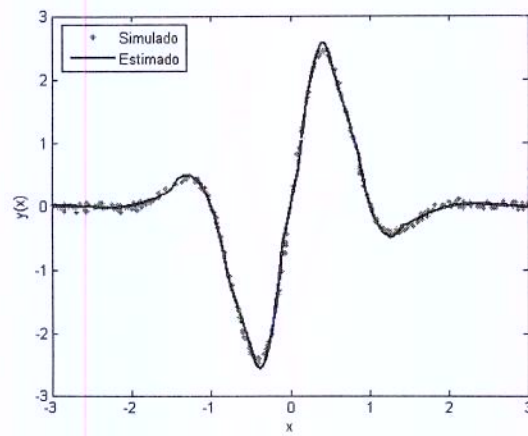


Fig. 3.22: Comparação entre os valores simulados e os estimados.

controle preditivo de sistemas dinâmicos não-lineares.

3.7 Resumo

Este capítulo teve como principal objetivo apresentar conceitos fundamentais da análise e projeto de sistemas *fuzzy*. Foi feita uma breve introdução da teoria de conjuntos fuzzy e apresentados os modelos *fuzzy* de Mamdani e Takagi-Sugeno, sendo dada maior atenção a este último, o qual é empregado no desenvolvimento deste trabalho, conforme já abordado neste capítulo e segundo o que será ainda discutido no capítulo 4.

Ainda neste capítulo foi apresentada a técnica de agrupamento utilizada para obtenção dos modelos *fuzzy* TS para os sistemas sob estudo. Foram abordados os critérios mais utilizados na determinação de um número apropriado de grupos (regras) a ser utilizado na obtenção de um modelo. Apresentou-se o método desenvolvido que concilia os resultados destes critérios com a finalidade de determinar o número de regras utilizadas no modelo *fuzzy* TS. Também desenvolveu-se um algoritmo iterativo para gerar as funções de pertinência trapezoidais presentes nos antecedentes das regras. Outro tópico abordado refere-se aos métodos local e global de obtenção dos coeficientes das FBO's utilizadas nos consequentes das regras do modelo. Este modelo proposto apresenta as vantagens estruturais de um modelo dado por FBO (ver seção 2.1.5), além de apresentar as características de aproximador universal de funções inerentes aos modelos TS, se tornando mais eficaz na modelagem de sistemas dinâmicos uma vez que é um modelo mais simples estruturalmente quando comparado a metodologias de modelagem de sistemas dinâmicos não-lineares hoje utilizadas, como os modelos polinomiais NARX e NARMAX.

Por fim, foram apresentados exemplos ilustrativos de agrupamento de dados e da geração do conjunto de regras do modelo a partir de dados de entrada e saída do sistema. Através destes exemplos verificou-se a eficácia dos métodos propostos para a determinação automática do número de grupos e obtenção das funções de pertinência presentes no antecedente das regras, como também na obtenção dos consequentes das regras do modelo *fuzzy* TS. Exemplos com maiores detalhes serão abordados no Capítulo 5, onde, além da modelagem, serão discutidas aplicações ao controle preditivo de sistemas dinâmicos.

Capítulo 4

Controle Preditivo Não-linear

Este capítulo tem como objetivo apresentar a técnica de controle preditivo (MBPC¹) que será utilizada na implementação do sistema de controle proposto neste trabalho. Tal técnica de controle, em certo aspecto, se aproxima do comportamento humano na execução de tarefas, uma vez que é característica do ser humano tomar decisões antecipativas ao comportamento de um dado sistema baseado no conhecimento prévio do desempenho do sistema em um horizonte limitado e no objetivo final a ser atingido, como por exemplo, ao guiar um veículo.

Nas primeiras seções deste capítulo serão discutidos e apresentados as principais características e fundamentos referentes aos controladores preditivos lineares. A proposta central desta dissertação será então apresentada na seção 4.4 que consiste na elaboração de um controlador preditivo não-linear que utiliza um modelo *fuzzy* TS-FBO e controladores preditivos lineares que atuam junto aos modelos locais nos conseqüentes das regras do modelo.

4.1 O Controlador Preditivo

Os controladores preditivos pertencem à classe de algoritmos de controle que utilizam o modelo do processo para obter a lei de controle através de um critério de otimização. O conceito de controle preditivo foi introduzido nos trabalhos de Richalet et al. (1978) e de Cutler e Ramaker (1980), com este último apresentando o controlador DMC². O trabalho de Clarke et al. (1987) apresentou o controlador preditivo generalizado (GPC³), o qual difundiu na literatura a metodologia de controle MBPC, tornando-se o GPC ao lado do DMC os métodos de controle preditivo mais utilizados na prática.

Tais controladores preditivos podem ser caracterizados por quatro etapas principais, como apre-

¹MBPC - *Model Based Predictive Control*

²DMC - *Dynamic Matrix Controller*

³GPC - *Generalised Predictive Control*

sentadas abaixo:

- i. Modelagem de processo - O modelo do sistema é utilizado no controlador para prever o comportamento do processo em um horizonte de tempo futuro, denominado horizonte de previsão, utilizando os dados reais de entrada e saída do sistema.
- ii. Definição do critério de custo - O desempenho do sistema em malha fechada durante o horizonte de previsão é especificado através de um critério de custo, definido a partir da saída prevista, do sinal de referência e do esforço de controle.
- iii. Otimização do critério de custo - O critério de custo é otimizado em relação ao conjunto de sinais de controle a serem aplicados no processo futuramente durante o horizonte de previsão.
- iv. Atuação no sinal de controle - Somente o primeiro sinal de controle resultante da otimização do critério de custo é aplicado na entrada do processo real, e a cada instante de amostragem o processo é repetido.

A Fig. 4.1 sintetiza em um diagrama de bloco estas quatro etapas descritas anteriormente, que compõem um controlador preditivo e descrevem seu funcionamento.

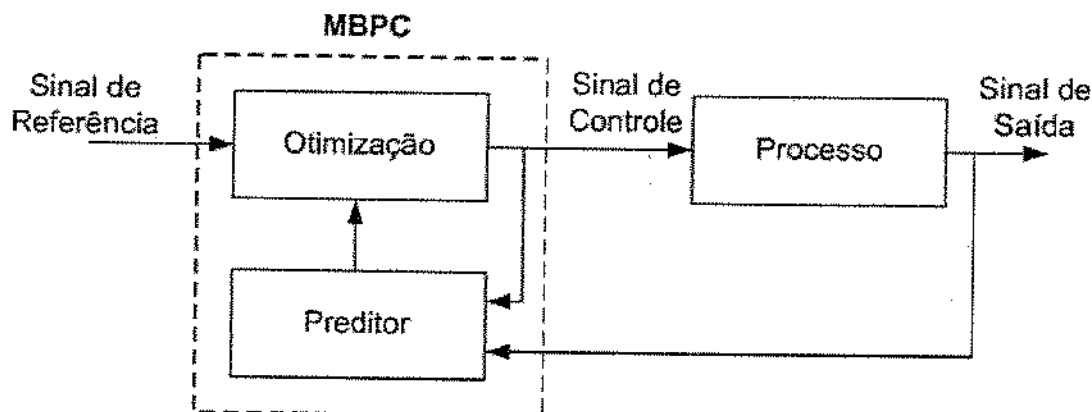


Fig. 4.1: Diagrama de blocos do Controlador Preditivo

Os controladores preditivos apresentam as seguintes características quando comparados a outras técnicas de controle (Camacho e Bordons, 1998):

- É particularmente atrativo para lidar quando se tem um conhecimento limitado de controle, uma vez que seus conceitos são muito intuitivos e ao mesmo tempo a sintonia é relativamente fácil;
- Podem ser usados para controlar uma grande variedade de processos desde aqueles com dinâmicas relativamente simples até mais complexos, incluindo sistemas com grandes atrasos de transporte, sistemas de fase não-mínima ou instáveis;

- Sistemas multi-variáveis podem ser controlados por esta abordagem;
- Compensa intrinsecamente os atrasos de transporte;
- O controlador resultante é de fácil implementação e exige um pequeno esforço computacional para processos sem restrições;
- Sua extensão para o tratamento de restrições é conceitualmente simples podem ser sistematicamente incluídas durante o processo de projeto;
- É muito útil quando se conhece os sinais de referência durante toda a operação do processo.
- É uma metodologia totalmente aberta baseado em princípios básicos os quais permitem futuras extensões;

Estas características justificam a escolha de controladores preditivos uma vez que os mesmos apresentam diversas características de interesse no controle de sistemas dinâmicos. A seguir, serão abordados com maiores detalhes os fundamentos que estão presentes em um controlador preditivo, como a modelagem do processo, a otimização da função de custo e a obtenção do sinal de controle.

4.2 Modelagem do Processo e Previsão da Saída

Como foi abordado anteriormente, uma das etapas do controle preditivo é a modelagem do processo a ser controlado. Nos algoritmos preditivos, a lei de controle e, conseqüentemente, o desempenho do sistema em malha fechada, está diretamente ligada à formulação do modelo, sendo assim a modelagem um dos fatores que mais influenciam um sistema de controle (Rossiter, 2003). A modelagem pode ser feita pelo emprego de uma das arquiteturas apresentadas na seção 2.1, sendo necessária a utilização de modelos discretizados no tempo de acordo com a taxa de amostragem do controlador. Neste trabalho a metodologia de modelagem utilizada serão as funções de base ortonormal - FBO (seção 2.1.5) (Oliveira, 1997). Para o sistema de controle preditivo em questão, o modelo de FBO será representado sob a forma variacional, dado pela equação (2.26), escrita com a adição de uma perturbação $\xi(k)$ ao sinal:

$$\begin{aligned}\Delta l(k+1) &= A_f \Delta l(k) + B_f \Delta u(k) \\ \hat{y}(k) &= \hat{y}(k-1) + C_f^T \Delta l(k) + \Delta \xi(k)\end{aligned}\tag{4.1}$$

A previsão de saída do modelo j passos a frente é dada por:

$$\hat{y}(k+j/k) = \mathcal{E}\{y(k+j)\}\tag{4.2}$$

ou

$$\hat{y}(k+j/k) = \mathcal{E}\left\{\frac{1}{\Delta}C^T\Delta l(k+j) + \frac{1}{\Delta}\Delta\xi(k+j)\right\} \quad (4.3)$$

Assumindo que o sinal de perturbação $\xi(k+j)$ tem valor médio igual a zero, a equação (4.3) pode ser reescrita como:

$$\hat{y}(k+j/k) = \mathcal{E}\left\{\frac{1}{\Delta}C^T\Delta l(k+j)\right\} \quad (4.4)$$

ou equivalentemente:

$$\hat{y}(k+j/k) = \hat{y}(k+j-1/k) + C^T\Delta l(k+j) \quad (4.5)$$

Substituindo $\Delta l(k+j)$ de (4.1) em (4.5) obtém-se:

$$\hat{y}(k+j/k) = \hat{y}(k+j-1/k) + C^T\{A\Delta l(k+j-1) + B\Delta u(k+j-1/k)\} \quad (4.6)$$

onde $\Delta u(k+j-1/k)$ é a variação do sinal de controle no instante $k+j-1$, calculada a partir da informação disponível no instante k .

Assim como foi substituída a parcela $\Delta l(\cdot)$ em (4.5), fazendo-se sucessivas substituições e supondo que $\Delta u(k+j/k) = 0 \forall j \geq N_u$ obtém-se:

$$\hat{y}(k+j/k) = \hat{y}(k+j-1/k) + C^T A^j \Delta l(k) + C^T \sum_{i=1}^{N_u} A^{j-i} B \Delta u(k+i-1/k) \quad (4.7)$$

onde $A^j = 0$ para $j < 0$, $l(k) = 0$ para $k \leq 0$ e N_u é um parâmetro de projeto denominado horizonte de controle, comum em algoritmos clássicos MBPC (ver seção 4.3).

Substituindo $\hat{y}(k+j-1/k)$ na equação (4.7), obtém-se:

$$\hat{y}(k+j/k) = \hat{y}(k+j-2/k) + C^T(A^{j-1} + A^j)\Delta l(k) + C^T \sum_{i=1}^{N_u} (A^{j-i-1} + A^{j-i})B \Delta u(k+i-1/k) \quad (4.8)$$

Repetindo-se esta substituição sucessivamente, irá se obter como resultado:

$$\hat{y}(k+j/k) = y(k) + C^T(K_j - I)\Delta l(k) + C^T \sum_{i=1}^{N_u} K_{j-i}B \Delta u(k+i-1/k) \quad (4.9)$$

onde $K_j = \sum_{i=0}^j A^i$, com $K_j = 0$ para $j < 0$ e I é a matriz identidade de ordem m , sendo m a dimensão do vetor $l(k)$, isto é, o número de estados ortonormais do modelo.

Na equação (4.9), a qual representa a previsão de saída do processo, pode-se distinguir dois termos: um termo é função de informações passadas do processo, conhecido como previsão de saída livre, e o outro termo é função das informações futuras a respeito do processo, denominado previsão

de saída forçada. Sendo assim, pode-se reescrever a equação (4.9) da seguinte forma:

$$\hat{y}(k + j/k) = \hat{y}_f(k + j/k) + \hat{y}_l(k + j/k) \quad (4.10)$$

onde:

$$\hat{y}_l(k + j/k) = y(k) + C^T(K_j - I)\Delta l(k) \quad (4.11)$$

e

$$\hat{y}_f(k + j/k) = C^T \sum_{i=1}^{N_u} K_{j-i} b \Delta u(k + i - 1/k) \quad (4.12)$$

sendo $\hat{y}_l$ é a previsão da saída livre e $\hat{y}_f$ é a previsão da saída forçada.

As previsões de saída, calculadas através da equação (4.10) para valores de j entre N_1 e N_y , que determinam o início e o final do horizonte de previsão, respectivamente, podem ser escritas na forma matricial mostrada abaixo:

$$\hat{y} = G\Delta u + \hat{y}_l \quad (4.13)$$

onde:

$$\hat{y} = \begin{bmatrix} \hat{y}(k + N_1/k) & \hat{y}(k + N_1 + 1/k) & \dots & \hat{y}(k + N_y/k) \end{bmatrix} \quad (4.14)$$

$$G = \begin{bmatrix} g_{N_1} & g_{N_1-1} & \dots & g_{N_1-N_u+1} \\ g_{N_1+1} & g_{N_1} & \dots & g_{N_1-N_u+2} \\ \vdots & \vdots & \ddots & \vdots \\ g_{N_y} & g_{N_y-1} & \dots & g_{N_y-N_u+1} \end{bmatrix} \quad (4.15)$$

$$\Delta u = \begin{bmatrix} \Delta u(k/k) & \Delta u(k + 1/k) & \dots & \Delta u(k + N_u - 1/k) \end{bmatrix}^T \quad (4.16)$$

$$\hat{y}_l = \begin{bmatrix} \hat{y}_l(k + N_1/k) & \hat{y}_l(k + N_1 + 1/k) & \dots & \hat{y}_l(k + N_y/k) \end{bmatrix} \quad (4.17)$$

com:

$$g_i = C^T K_{i-1} b \quad (4.18)$$

A seguir será apresentada a estrutura utilizada na formulação e implementação de um controlador preditivo.

4.3 Lei de Controle Preditivo Linear

A lei de controle dos MBPC é obtida otimizando-se um critério de custo, que geralmente é definido em função dos erros da previsão de saída do processo em relação a um sinal de referência especificado e da variação do sinal de controle a ser aplicado ao processo, durante um horizonte de tempo futuro denominado horizonte de previsão, como citado anteriormente. A Fig. 4.2 apresenta, através das áreas destacadas, os sinais considerados na função de custo e que devem ser minimizados. Outros critérios de custo podem ser definidos, como, por exemplo, quando se têm múltiplos modelos de referência ou para referência no sinal de controle (Irwing, Falinower e Fonte, 1986), ou ainda incluindo um termo de ponderação no sinal de controle presente na função de custo (Lopez, 1995).

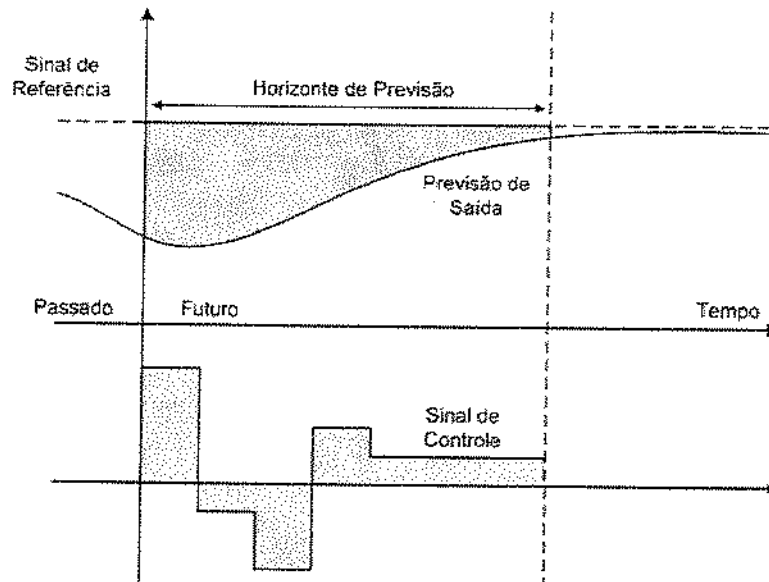


Fig. 4.2: Sinais do Controlador Preditivo

Para todas estas funções de custo, define-se o critério de custo utilizando-se normas espaciais e temporais para os sinais envolvidos. As normas espaciais relacionam os sinais em um mesmo passo de previsão, descrevendo a relação entre os sinais $\hat{y}(k+j/k)$, $\omega(k+j)$ e $\Delta u(k+j/k)$ em um dado passo j de previsão. As normas temporais relacionam os sinais considerando todo o horizonte de previsão, ou seja, realizam o cálculo para j variando ao longo do horizonte de previsão. Os principais tipos de normas utilizadas em controle preditivo são:

- i. Norma 1: $\|x\|_1 = \sum_i |x_i|$
- ii. Norma 2: $\|x\|_2 = \sqrt{\sum_i x_i^2}$
- iii. Norma infinita: $\|x\|_\infty = \max_i \{x_i\}$

Muitas podem ser as maneiras pelas quais são descritos os critérios de custo, sendo o mais utilizado em aplicações o critério de custo quadrático com norma temporal 1 e norma espacial 2. Tal critério é o utilizado nos algoritmos GPC, DMC, entre outros. Assim, minimizando-se o critério de custo em relação às variações futuras do sinal de controle obtém-se a seqüência ótima de sinais de controle a ser aplicado no processo. Dessa forma, a lei de controle geral para os controladores do tipo MBPC é dada, para a abordagem incremental, por:

$$\begin{aligned} \Delta u &= \arg \min \{ F(\hat{y}(k + j/k) - \omega(k + j), \Delta u(k + j - 1/k), j = N_1, \dots, N_y) \} \\ s. a \quad & \Delta u(k + j/k) = 0 \quad \forall j \geq N_u \end{aligned} \quad (4.19)$$

onde:

$F(\cdot)$ é o critério de custo do controlador, definido de acordo com a escolha das normas espacial e temporal;

N_1 é o horizonte inicial de previsão de saída;

N_y é o horizonte final de previsão de saída;

N_u é o horizonte de controle;

$\hat{y}(k + j/k)$ é a previsão de saída j passos a frente;

$\omega(k)$ é o sinal de referência;

$\Delta u(k + j/k)$ é a variação do sinal de controle no instante $k + j$, calculada em k .

O horizonte de controle N_u restringe os graus de liberdade na determinação dos valores de $\Delta u(k + j/k)$ durante o horizonte de previsão, pois $\Delta u(k + j/k) = 0$ para $j \geq N_u$, e por isso representa um parâmetro de projeto de controlador.

O controlador preditivo tem a característica de tratar de forma simples a presença de atraso de transporte no sistema a ser controlado, tratando isto de forma direta, através de considerações explícitas no próprio modelo do sistema, ou de forma indireta na determinação dos parâmetros do horizonte de previsão N_1 e N_y . Para a forma indireta, garantindo-se que N_y seja maior que o atraso de transporte e suas eventuais variações, não se faz necessário o conhecimento do atraso do sistema. Uma outra maneira para o caso indireto é adotar um valor de $N_1 > 1$, reduzindo assim o esforço computacional tendo em vista que as previsões não dependem, devido ao tempo morto, da determinação dos sinais de controle.

A solução do problema de otimização (4.19) tem como resultado um vetor Δu composto de N_u variações no sinal de controle dado por:

$$\Delta u = \begin{bmatrix} \Delta u(k/k) & \Delta u(k + 1/k) & \dots & \Delta u(k + N_u - 1/k) \end{bmatrix}^T \quad (4.20)$$

Somente o primeiro elemento de Δu é aplicado no processo, isto é, $u(k) = \Delta u(k/k) + u(k-1)$, e no próximo instante de amostragem todo este procedimento é repetido. A definição do critério de custo do algoritmo de controle preditivo com modelos de funções de base ortonormal (Oliveira, 1997), o qual será utilizado neste trabalho, será apresentado, na próxima seção.

4.3.1 Controle Preditivo Linear com FBO

Nesta seção será apresentada a lei de controle para um controlador preditivo com modelo dado por funções de base ortonormal (CP-FBO). Este controlador apresenta uma função de custo do tipo quadrática. O critério de custo empregado utiliza o quadrado da norma 2 como norma espacial e como norma temporal. Tal tipo de critério é o utilizado com maior frequência em algoritmos de controle preditivo devido à possibilidade de se obter uma solução analítica para a lei de controle quando não existem restrições nas variáveis do processo. Assim, o critério de custo quadrático é dado por:

$$J_q(\Delta u) = \sum_{j=N_1}^{N_y} (\hat{y}(k+j/k) - \omega(k+j))^2 + \sum_{j=1}^{N_u} \lambda \Delta u^2(k+j-1) \quad (4.21)$$

onde λ ($\lambda \geq 0$) é um fator de ponderação para os futuros incrementos do sinal de controle.

A equação (4.21) pode ser reescrita sob a forma vetorial como apresentado a seguir:

$$J_q(\Delta u) = \Delta u^T Q \Delta u + f^T \Delta u + d \quad (4.22)$$

onde:

$$Q = G^T G + \lambda I, \quad I \text{ é a matriz identidade} \quad (4.23)$$

$$f = 2G^T(\hat{y}_l - \omega) \quad (4.24)$$

$$d = (\hat{y}_l - \omega)^T(\hat{y}_l - \omega) \quad (4.25)$$

e

$$\omega = \begin{bmatrix} \omega(k+N_1) & \omega(k+N_1+1) & \dots & \omega(k+N_y) \end{bmatrix}^T \quad (4.26)$$

sendo G , Δu e $\hat{y}_l$ dados por (4.15), (4.16) e (4.17), respectivamente.

A lei de controle então consiste em minimizar esta função de custo (4.22) em relação a Δu , como apresentado abaixo:

$$\min_{\Delta u} \Delta u^T Q \Delta u + f^T \Delta u \quad (4.27)$$

A formulação apresentada em (4.27) se trata de um problema de programação quadrática irrestrita (Bazaraa e Shetty, 1979) onde a solução factível é um mínimo global, não admitindo assim mínimos

locais. Este problema admite solução analítica dada por:

$$\Delta u = -(1/2)Q^{-1}f \quad (4.28)$$

Se houver restrições nos sinais de entrada e/ou saída, a solução é obtida numericamente através de algoritmos de programação quadrática com restrições. A solução permanece ótima global se as restrições forem lineares.

4.4 Controlador Preditivo Não-linear

Controladores preditivos não-lineares têm emergido como uma técnica promissora para o controle de processos altamente não-lineares. Um desses controladores é o controlador preditivo de múltiplos modelos (MMAC⁴) que é uma estratégia de controle baseada em modelo que incorpora um conjunto de pares modelo/controlador que operam em diferentes regiões de operação do sistema e apresentam melhor desempenho do que um único modelo e controlador atuando em todas as condições de operação do sistema (Schott e Bequette, 1997). Além disso, a decomposição do problema de controle em um conjunto de subproblemas menores pode simplificar a busca por uma solução ótima (Yeung e Bekey, 1993).

A proposta deste trabalho consiste então na elaboração de um sistema de controle preditivo para sistemas não-lineares com um controlador preditivo de múltiplos modelos sendo utilizados modelos *fuzzy* com estrutura TS com funções de base ortonormal nos consequentes das regras. A escolha de modelos *fuzzy* para a modelagem do processo se deve ao fato destes modelos serem aproximadores universais de funções num dado espaço compacto (Wang e Mendel, 1992; Kosko, 1994; Wang, 1998). Em particular, a escolha de modelos TS se deve ao fato que estes permitem a composição do modelo global a partir de múltiplos modelos locais o que favorece a interpretabilidade e projeto de controladores. A utilização de FBO's nos consequentes das regras do modelo TS se deve ao fato destas apresentarem a característica já citadas na seção 2.1.5.

A etapa de projeto do controlador é realizada de maneira *offline*. Primeiramente, é necessário identificar o processo a ser controlado. Para tanto é feita uma excitação do sistema em malha aberta, armazenando, ao final da excitação, os dados de entrada e saída. O processo de modelagem do sistema a ser controlado segue o algoritmo apresentado nas seções 3.5 e 3.6, ou seja, os dados coletados na saída do sistema, junto aos dados gerados pelos estados da FBO, são agrupados pelo algoritmo de GK tendo seu valor ótimo de grupos determinado pelo processo que envolve os critérios apresentados na seção 3.5.2. Após a determinação dos grupos a próxima fase consiste em se formar o conjunto

⁴MMAC - Multiple model adaptative control

de regras que irá modelar o sistema, tendo as funções de pertinência dos antecedentes estabelecidas através do processo descrito na seção 3.6.1 e posteriormente determinando os parâmetros das FBO's nos consequentes das regras. A determinação dos parâmetros dos consequentes pode ser realizada de maneira global ou de maneira local como descrito na seção 3.6.2. Para as aplicações em controle com controlador preditivo de múltiplos modelos a estimação local de parâmetros se mostra mais adequada, uma vez que os modelos locais apresentam comportamento e interpretabilidade locais, o que não ocorreria com a estimação global dos parâmetros. Como poderá ser verificado no capítulo 5, que traz exemplos de implementações, a estimação local apresenta resultados adequados na modelagem dos sistemas sob controle. Neste trabalho foram utilizadas FBO's com pólo único ou com par conjugado de pólos comum a todas as funções presentes nos modelos locais.

Uma vez determinadas as regras *fuzzy* do modelo TS e os modelos locais FBO, pode-se então implementar o controlador preditivo não-linear. A implementação do controlador não-linear consiste em aplicar um controlador preditivo linear (conforme apresentado em 4.3.1) para cada modelo local definido na etapa de modelagem. A Fig. 4.3 exemplifica o sistema de controle em questão.

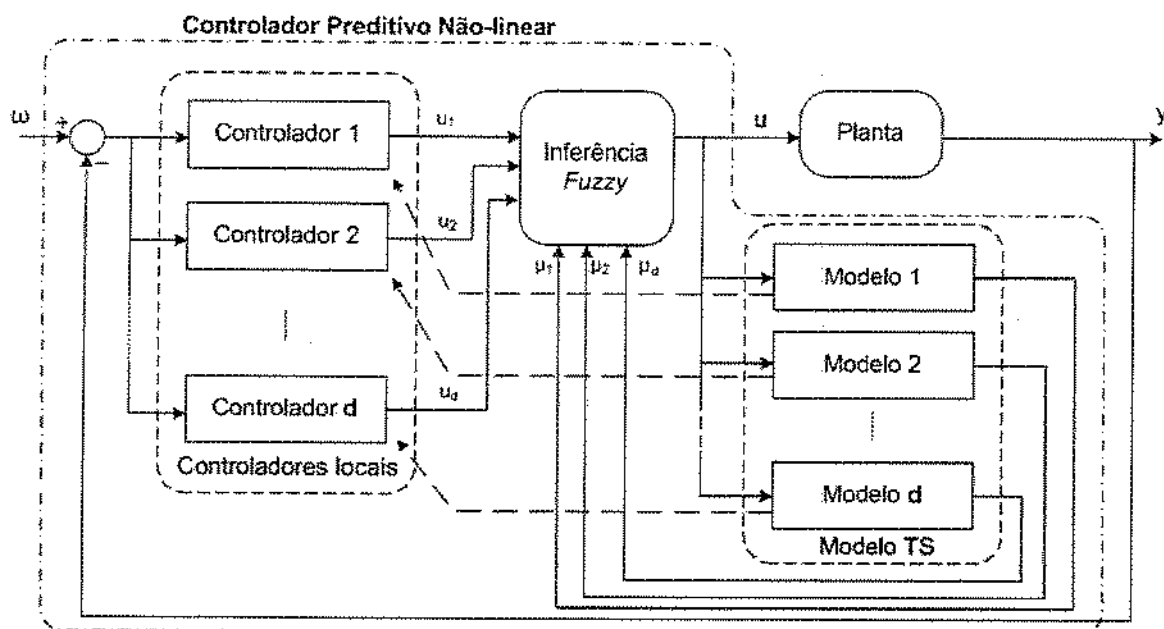


Fig. 4.3: Controlador preditivo proposto.

A ação de controle é obtida pela combinação ponderada, pelo modelo *fuzzy* TS, das ações de controle locais. O mesmo sinal que é fornecido à planta também é fornecido aos modelos locais que por sua vez fornecerão informação para os controladores. Baseado na resposta e previsão dos modelos locais os controladores lineares “atuam” sobre estes modelos de maneira a corrigir a resposta de cada um deles. O sinal de entrada aplicado a cada modelo local também tem a função de gerar os estados das FBO's e estes, por sua vez, a função de ativarem as regras do modelo TS. Com o valor de ativação

das regras dos respectivos modelos locais, é feita a inferência sobre as ações de controle locais que irão compor a ação de controle global “ u ” que será aplicada à planta. A cada período de amostragem se repete esse ciclo no controlador e a ação de controle é atualizada na entrada da planta. Sendo assim, a ação de controle é dada por:

$$u(k) = \frac{\sum_{i=1}^N (\mu_i(k) u_i(k))}{\sum_{i=1}^N \mu_i(k)} \quad (4.29)$$

onde $\mu_i(k)$ é dado por (3.13) e $u_i(k)$ é a ação de controle determinado pelo i -ésimo controlador.

O próximo capítulo traz os resultados obtidos em simulações de modelagem a fim de ilustrar o desempenho do modelo utilizado no controlador preditivo como também os resultados do sistema de controle como um todo aplicado a algumas configurações de plantas não-lineares.

4.5 Resumo

Este capítulo teve como objetivo apresentar uma proposta para implementação de um controlador preditivo não-linear de multimodelos locais utilizando estruturas *fuzzy* TS para inferência das ações de controle. Além disso, foram discutidos os conceitos centrais da abordagem de controle preditivo, cada vez mais utilizada na indústria de processos com o avanço dos recursos computacionais.

A qualidade do modelo do processo sob controle é determinante para o desempenho do controlador, mostrando assim a importância de se definir modelos adequados para o referido processo. O controle de sistemas não-lineares pode se tornar crítico por esta questão, tendo em vista que é necessária a identificação de modelos não-lineares e controladores associados com um bom desempenho. A vantagem do controlador proposto está justamente no fato de que este controlador trabalha somente com modelos e controladores locais lineares, o que simplifica muito a solução do problema global uma vez que a identificação do sistema e o projeto de controle podem ser decompostos de forma mais simples, menos onerosa computacionalmente e com bons resultados.

No próximo capítulo serão apresentadas algumas aplicações deste controlador no contexto do trabalho proposto e os resultados obtidos nas implementações realizadas.

Capítulo 5

Resultados e Simulações

Este capítulo tem como objetivo mostrar aplicações e resultados através de simulações das propostas apresentadas neste trabalho, que são a identificação de sistemas dinâmicos não-lineares através de modelos *fuzzy* TS como funções de base ortonormal nos consequentes das regras e um controlador preditivo não-linear que utiliza multimodelos lineares para a modelagem do processo sob controle.

Primeiramente será apresentada a modelagem de um sistema de levitação magnética. Esse sistema não-linear não permite a implementação do controlador proposto em sua malha de controle devido às características da planta, presente nos laboratórios desta instituição. Posteriormente, serão então apresentados dois caso de controle em malha fechada cujos processos, plantas de polimerização *CSTR*¹, são utilizados como *benchmark* no controle preditivo de processos não-lineares.

5.1 Modelagem de um levitador magnético

Para o perfeito funcionamento de um controlador preditivo baseado em modelo (MBPC) é de suma importância que o modelo do sistema, sob ação do controlador, seja de qualidade e que possa de fato representar as características e a dinâmica do sistema. A fim de exemplificar a eficácia do processo de identificação de sistemas não-lineares proposto neste trabalho, será realizada a modelagem de um sistema de levitação magnética presente no laboratório de controle desta instituição. A Fig. 5.1 traz uma representação gráfica da estrutura deste levitador.

O sistema é composto de duas bobinas que, ao serem submetidas a corrente elétrica, criam campos magnéticos que vão interagir com os campos dos respectivos disco magnéticos permanentes. O sentido da corrente em cada bobina, bem como a polaridade de cada magneto, são estabelecidos de forma a criar uma força resultante sobre cada disco magnético que seja:

¹CSTR - *Continuous Stirred Tank Reactor*

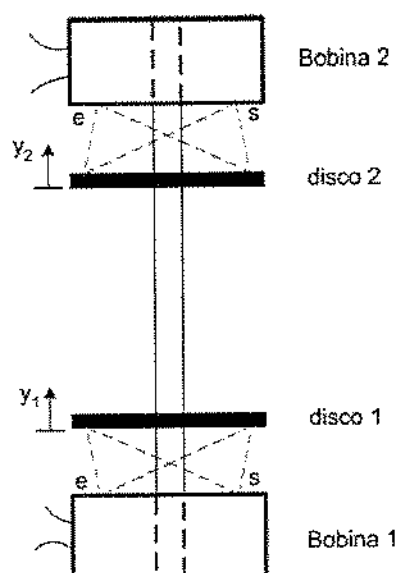


Fig. 5.1: Levitador Magnético; e - emissor laser; s - sensor laser.

- a. Repulsiva para o magneto inferior acionado pela bobina inferior,
- b. Atrativa para o magneto superior acionado pela bobina superior.

Utiliza-se um emissor de laser e um sensor correspondente, fixados na superfície de cada uma das bobinas, a fim de fazer as medidas de posição dos discos de forma independente e não intrusiva. Iremos considerar neste trabalho o movimento do disco magnético inferior (disco 1). Logo, a entrada do sistema será a corrente aplicada na bobina superior (com corrente na bobina inferior mantida constante) e a saída será dada pela altura do disco 1 y_1 em relação à bobina inferior. Portanto, deseja-se estimar, através de medidas sucessivas da resposta de posicionamento do disco magnético, um modelo que descreva o comportamento dinâmico do sistema de levitação magnética. Restrições quanto ao sistema de aquisição de dados e acionamento do sistema impedem que seja feito o controle do sistema através do controlador proposto. Por isso, para esse sistema, ficar-se-á restrito à modelagem do mesmo.

Para a identificação do sistema será utilizada um conjunto de dados coletados diretamente do processo descrito, por meio de um sistema de aquisição de dados. A Fig. 5.2 traz os dados utilizados no processo de identificação, sendo os dados amostrados com período de 0,1 segundos em um tempo total de 102 segundos. Os sinais de entrada e saída serão normalizados para evitar erros numéricos na determinação dos modelos. O sinal de entrada, conforme explicado na seção 3.4.1, será utilizado na obtenção dos estados da dinâmica FBO que estão presentes nos antecedentes e conseqüentes das regras fuzzy. Para estimação dos estados e identificação do sistema de levitação foram utilizados como pólos da dinâmica FBO o par de pólos complexos $p = 0.799 \pm 0.268i$, extraídos de Medeiros

(2006), em cujo trabalho foram determinados como pólos ótimos para modelagem deste sistema por FBO's através de um processo de otimização por algoritmos genéticos. Além dos pólos, no trabalho de Medeiros (2006) foi otimizado o número de filtros de Kautz a serem utilizados, cujo valor ideal é dado por oito filtros de Kautz e que serão utilizados na modelagem (esta análise sobre o número de filtros será novamente abordado nesta seção). Como o sistema de levitação apresenta comportamento oscilatório quando submetido a uma variação no sinal de entrada, é justificável a utilização de funções de Kautz, uma vez que estas são mais adequadas a sistemas com comportamento oscilatório subamortecido. Caso não se tivesse nenhuma informação *a priori* da dinâmica dominante do sistema poderia também ser utilizada alguma técnica de determinação dos pólos apresentada na literatura (Campello, 2002; Rosa, 2005).

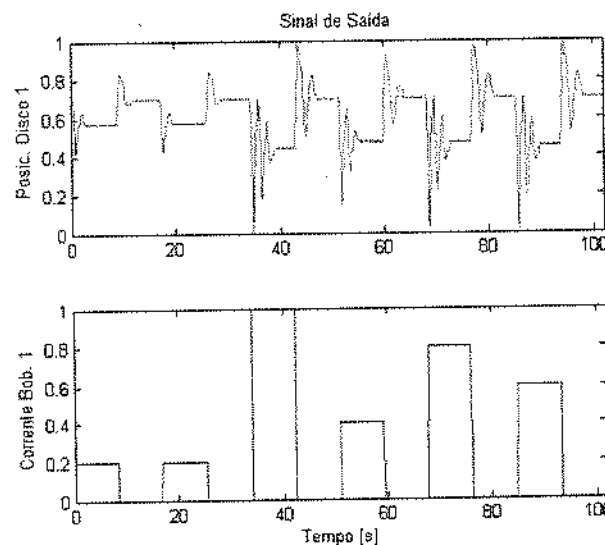


Fig. 5.2: Dados de identificação do sistema de levitação.

Empregando um único modelo linear com 8 estados de Kautz na modelagem do processo foi obtida uma resposta com erro quadrático médio entre saída real e saída prevista para os dados de validação - EQM = 0,0069. A resposta pode ser observada na Fig. 5.3.

Para a estimação do modelo utilizando o método proposto neste trabalho os estados da representação em espaço de estados FBO e os dados de saída do sistema foram reunidos em $Z_i = [l_i^T \ y_i]$, onde l_i^T são os estados de Kautz no instante i e y_i a saída medida de posicionamento do disco magnético inferior no mesmo instante. Para gerar modelos com regras mais simples diminuindo o tamanho da base de regras será usada somente o primeiro estado, l_1 , de Kautz. Além de gerar regras mais simples, a utilização de um menor número de variáveis na identificação possibilita uma diminuição na complexidade do processo de identificação. A diminuição do número de variáveis não acarreta perdas na qualidade do modelo, como pôde ser verificado nos resultados apresentados, uma vez que modelos

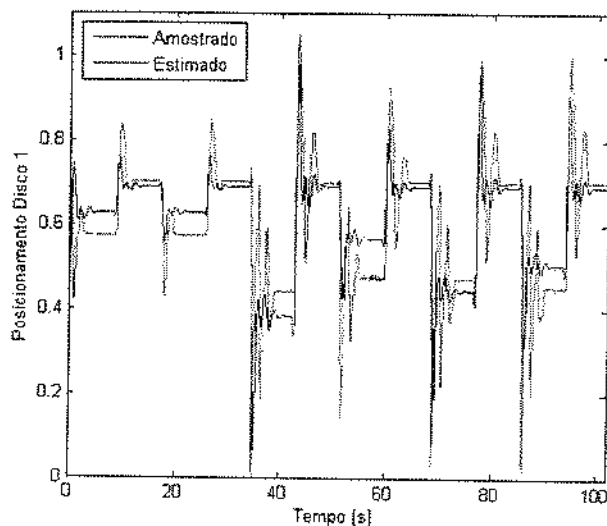


Fig. 5.3: Respostas medida do sistema e estimada pelo modelo único linear.

com mais estados nas premissas das regras apresentaram o mesmo EQM conforme indicado na Tab. 5.1 (resultado válido para este exemplo).

Tab. 5.1: Comparação entre os EQM's para modelos TS-FBO do levitador com 1 e 2 variáveis nas premissas das regras *fuzzy*.

Número Variáveis nas premissas	EQM	Número de modelos
1	0.0015	7
2	0.0015	7

Ao conjunto de objetos Z_i é aplicado o algoritmo de Gustafson-Kessel de agrupamento para se agrupar as medidas no espaço formado pela(s) entrada(s) (estado(s) de Kautz) e saída do sistema real. Durante o processo de modelagem, incluindo o agrupamento, a obtenção dos antecedentes das regras e dos conseqüentes serão descartados os primeiros 60 valores dos estados de Kautz com a finalidade de garantir que o transitório inicial na obtenção dos estados seja descartado e não influa negativamente no modelo final. Para se determinar o número de grupos que melhor representa os dados foi analisado o intervalo entre 2 e 10 grupos através dos critérios apresentados na seção 3.5.2. A Fig. 5.4 mostra que o melhor número de grupos para se compor o modelo TS é 7. Sendo assim, o modelo *fuzzy* para o levitador apresentará sete modelos locais com 1 estado (filtro) de Kautz nas premissas de cada regra e modelos locais com 8 estados de Kautz nos conseqüentes.

Determinados o número de grupos e os parâmetros do melhor agrupamento, foi realizada a modelagem das funções de pertinência, presentes nas regras do modelo *fuzzy*. A Fig. 5.5 apresenta as funções de pertinência, obtidas pelo método proposto em 3.6.1, que compõem os antecedentes das

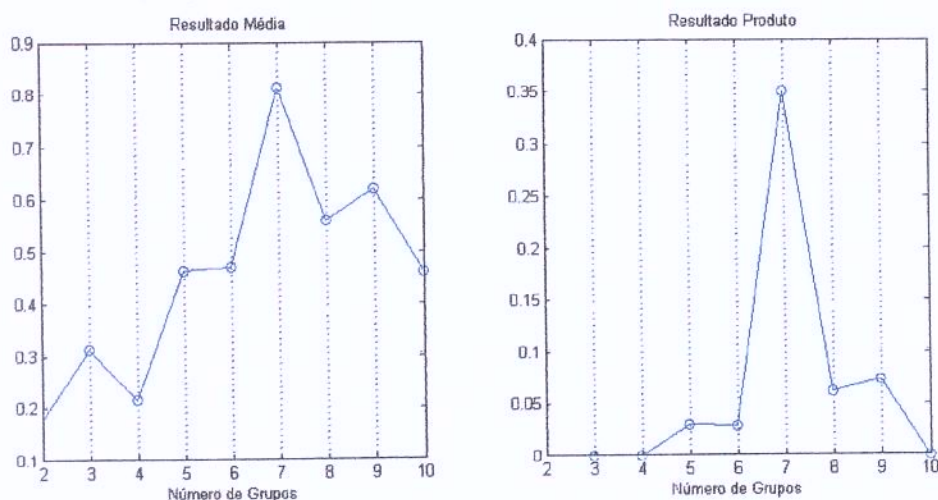


Fig. 5.4: Média e Produto dos critérios de validade de agrupamento p/ dados do sistema de levitação magnética agrupados pelo algoritmo de GK.

regras obtidas a partir das projeções dos valores de ativação dos objetos em cada grupo. Pode-se verificar que as funções de pertinência cobrem todo o universo de discurso, garantindo que em todo ele haverá pelo menos uma regra ativada, não apresentando regiões em que o modelo não é válido dentro do universo de discurso.

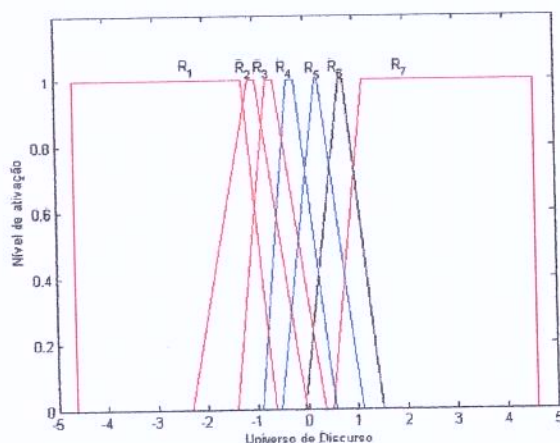


Fig. 5.5: Funções de pertinência presentes nos antecedentes das regras do modelo TS - levitador magnético.

Uma vez determinados os antecedentes das regras, é possível realizar a identificação dos coeficientes das FBO's presentes nos consequentes das regras, as quais definem os modelos locais dados pelos respectivos grupos. A estimação dos coeficientes dos modelos locais é feita através de esti-

mação local do estimador de mínimos quadrados, como apresentado na seção 3.6.2. A Fig. 5.6 mostra, através de um outro conjunto de dados utilizados exclusivamente para validação, a resposta amostrada diretamente do sistema comparada com a resposta estimada através do modelo identificado. O EQM apresentado para a resposta estimada é de 0,0015.

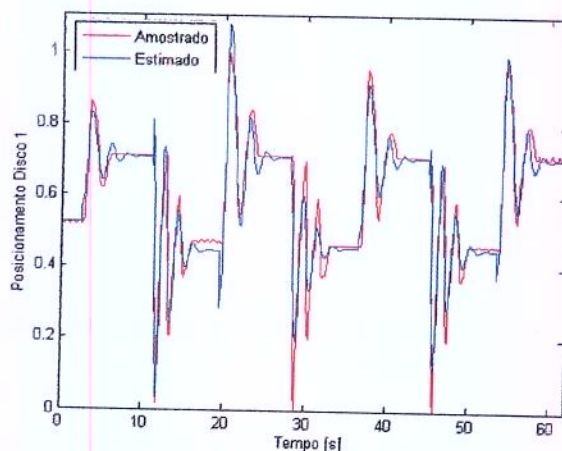


Fig. 5.6: Resposta medida do sistema e resposta estimada do modelo.

Para verificar se o modelo com 8 estados de Kautz era de fato o melhor modelo, foram feitas simulações variando-se o número de estados de Kautz e os resultados são apresentados na Tab. 5.2. Os resultados mostram que diminuindo o número de estados aumenta o EQM, desta maneira, reduzindo a precisão do modelo, o que mostra que de fato que modelos com 8 estados de Kautz se apresentam como uma boa escolha.

Tab. 5.2: Comparação entre os EQM's para modelos TS-FBO do levitador com diferentes quantidades de estados de Kautz nos modelos locais

Número Estados de Kautz	EQM	Número de modelos
5	0.0027	7
6	0.0019	7
7	0.0016	7
8	0.0015	7

Como pode ser observado na Fig. 5.6 o modelo apresentou um bom resultado com um EQM correspondente a 22% do EQM para o modelo linear (1 só modelo local modelando o processo, como apresentado na Fig. 5.3), ilustrando o bom desempenho do modelo proposto neste trabalho. O resultado obtido, com o método proposto, na modelagem deste processo apresenta um melhor resultado do que o apresentado por Medeiros (2006) quando se utiliza um único pólo no modelo

TS. No presente trabalho utiliza-se 7 modelos locais enquanto que na referência citada utiliza-se 24 modelos locais também com 8 estados de Kautz, apresentando um EQM de 0,0019, com modelo TS-FBO otimizado através de algoritmo genético.

Para ilustrar que a escolha do número de modelos locais apresentada pelos critérios de avaliação está adequada, variou-se o número de modelos locais com modelos dados FBO's na representação matricial com 8 estados de Kautz. A Tab. 5.3 traz um comparativo entre os EQM's entre configurações dos modelos TS com o valor de 1, 6, 7 e 8 modelos locais (grupos), onde pode-se observar que o menor valor de EQM ocorre para o modelo TS-FBO com 7 modelos locais.

Tab. 5.3: Comparação entre os EQM's para modelos TS-FBO do levitador com diferentes quantidades de modelos locais

Número de Modelos Locais	EQM
1	0.0069
6	0.0019
7	0.0015
8	0.0018

5.2 Controle de Sistemas Não-lineares

Esta seção exemplifica a implementação do controlador proposto no controle de sistemas não-lineares.

5.2.1 Controle de sistema CSTR

Considere um reator de polimerização do tipo CSTR em que ocorre uma reação exotérmica, sendo seu modelo de estados dado por:

$$\begin{aligned}
 \dot{x}_1(t) &= 60 - x_1(t)(2,4568\sqrt{x_2(t)} + 10) \\
 \dot{x}_2(t) &= 80u(t) - 10,1022x_2(t) \\
 \dot{x}_3(t) &= 0,002412x_1(t)\sqrt{x_2(t)} + 0,11218x_2(t) - 10x_3(t) \\
 \dot{x}_4(t) &= 245,9811x_1(t)\sqrt{x_2(t)} - 10x_4(t) \\
 y(t) &= x_4(t)/x_3(t)
 \end{aligned} \tag{5.1}$$

onde a saída $y(t)$ é o número médio do peso molecular (*Number Average Molecular Weight - NAMW*) [$kg/kmol$] do polímero resultante da reação que é controlada pela manipulação da taxa de fluxo $u(t)[m^3/h]$ da substância iniciadora. Mais detalhes sobre esse processo são apresentados em (Maner, Doyle, Ogunnaike e PearsonMcCulloch, 1996). Este sistema será utilizado para ilustrar a aplicação do controlador proposto neste trabalho.

A fim de realizar a identificação foi feita a simulação do sistema e foram armazenados os dados de entrada e saída. A simulação do conjunto de equações diferenciais do modelo (eq. (5.1)) foi feita utilizando-se o método Runge-Kutta de quarta ordem com passo fixo de 0,03h (1,8 min) e condições iniciais dadas por $x_1(0) = 5,50677$, $x_2(0) = 0,132906$, $x_3(0) = 0,0019752$, $x_4(0) = 49,3818$, $u(0) = 0,016783$ e $y(0) = 25000,5$. O período de simulação se deu de $t = 0h$ até $t = 24h$ como o sinal de entrada $u(t)$ sendo dado por uma sequência de degraus com período de uma 1h e amplitude aleatória uniformemente distribuída no intervalo $[0,002; 0,02]$. O conjunto de dados resultante é dividido em duas partes: as primeiras 12h são destinadas à identificação do modelo TS-FBO e a outra metade para a validação do modelo, sendo os dados normalizado entre $[0, 1]$ para se evitar a ocorrência de problemas numéricos durante a fase de estimação dos parâmetros do modelo. A primeira hora de cada conjunto de dados é utilizada apenas para calcular os estados de Laguerre, já que os estados de Laguerre tem início nulo e seu transitório inicial poderia exercer influências indesejáveis na estimação dos parâmetros do modelo. A Fig 5.7 mostra os sinais entrada e saída em malha aberta, utilizados na identificação do CSTR.

A Fig. 5.7 mostra que este sistema apresenta uma dinâmica menos complexa que a do sistema

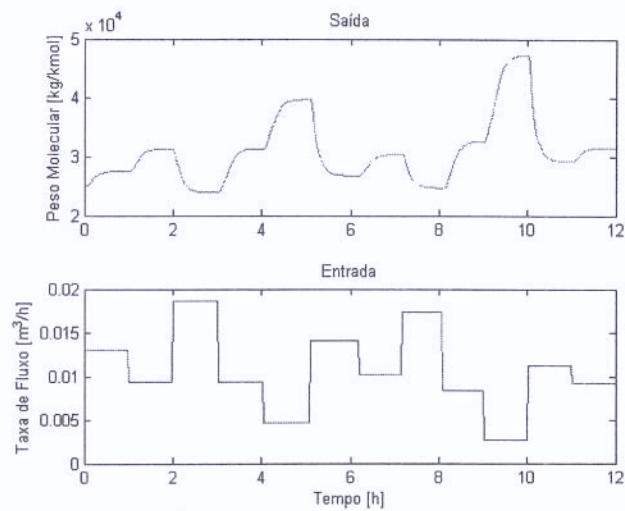


Fig. 5.7: Resposta em malha aberta do sistema CSTR.

de levitação magnética, apresentando neste caso um ganho altamente não-linear. O sistema CSTR apresenta uma resposta temporal sobreamortecida, sem oscilações, onde, um modelo com pólo real é suficiente para modelar esta dinâmica. Para a modelagem do processo pelo método proposto neste trabalho é necessário, além da entrada e saída, obter os estados da FBO em espaço de estados escolhida. Para este sistema, baseado nos resultados de Campello (2002), sabe-se que um modelo com funções de Laguerre e pólo $p = 0,75$ modela adequadamente o sistema. Foram utilizados, na elaboração dos modelos locais, FBO's em representação em espaço de estados com três estados (três filtros). Para a determinação do modelo *fuzzy* TS, o algoritmo de GK foi sequencialmente aplicado aos objetos da Fig. 5.7 analisando o valor de grupos que melhor agrupa os dados entre 2 e 10 grupos. A Fig. 5.8 mostra os resultados para a média e o produto das respostas dos critérios de avaliação, os quais indicam que 9 grupos são necessários para se modelar o sistema de maneira eficiente. Vale salientar que para determinação das regras do modelo TS foi utilizado somente o primeiro estado da representação em espaço de estados da FBO e a saída do sistema simulado. A diminuição da dimensionalidade do espaço de entrada simplifica o modelo e reduz a complexibilidade computacional do agrupamento dos dados, não diminuindo para este caso a eficiência do modelo como poderá ser visto.

Uma vez obtidos os grupos, é possível a partir destes se determinar as funções de pertinência dos antecedentes das regras através da projeção do valor de ativação dos objetos nos subespaços de entrada, conforme descrito na seção 3.6.1. A Fig. 5.9 mostra as funções de pertinência para os antecedentes das regras obtidas a partir das tais projeções. Como pode ser verificado, as funções de pertinência cobrem todo o universo de discurso, garantindo que em todo ele haverá pelo menos uma regra ativada.

Determinados os antecedentes das regras é possível se obter os coeficientes dos modelos locais

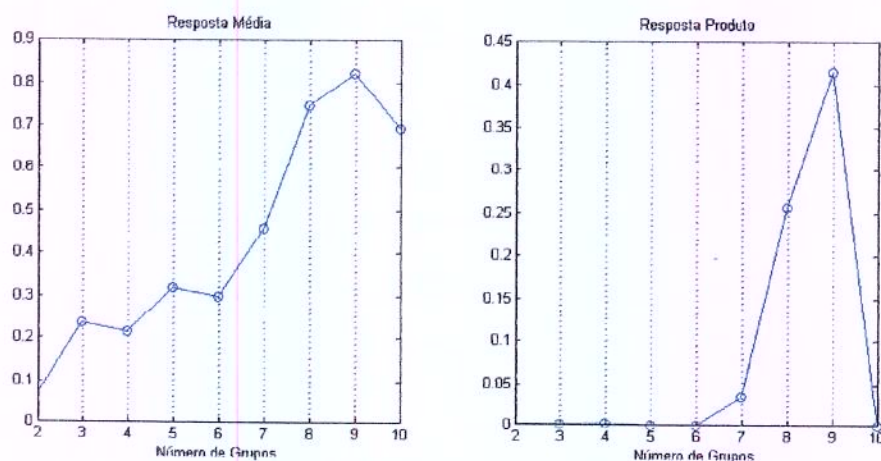


Fig. 5.8: Resultado global da estimação do número de grupos (modelos locais) via combinação de critérios de validade por média e produto - sistema CSTR.

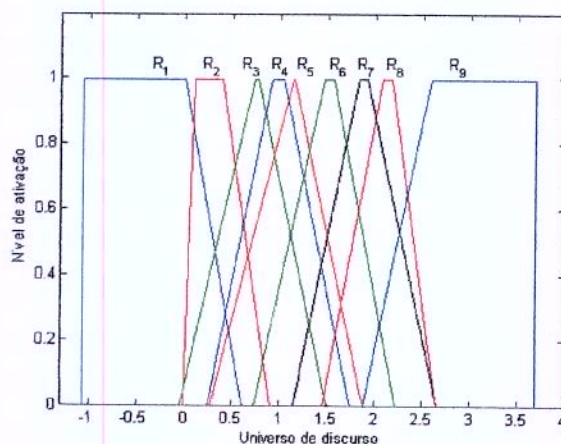


Fig. 5.9: Funções de pertinência presentes nos antecedentes das regras do modelo TS - sistema CSTR.

presentes nos conseqüentes das regras através de estimação local, como discutido na seção 3.6.2. Desta forma é concluído o processo de identificação estando os modelos locais então determinados e, conseqüentemente, o modelo *fuzzy* TS-FBO. A fim de averiguar se o modelo é de boa qualidade, a Fig. 5.10 traz a comparação entre a resposta do modelo e a resposta do sistema submetidos ao um sinal de excitação utilizado para validação, que assim como no caso da identificação tem sua primeira uma hora de simulação descartados uma vez que os estados de Laguerre estão no transitório inicial. Como pode ser verificado, a resposta do modelo está coerente com a resposta do sistema, apresentando um $EQM = 1,4341 \cdot 10^4$.

Depois de verificado que o modelo identificado é satisfatório, o próximo passo consta em utilizar

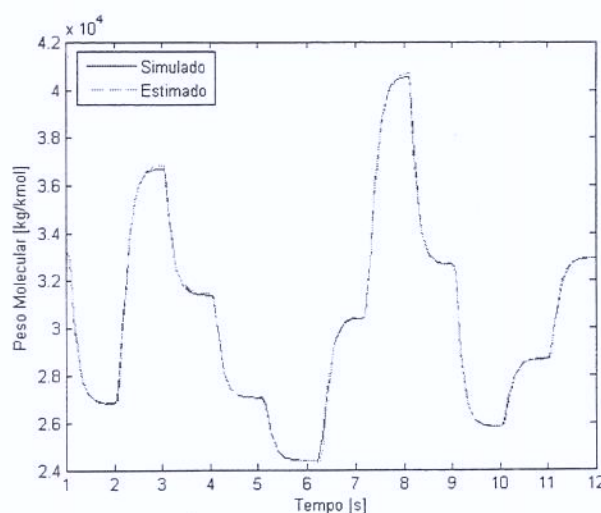


Fig. 5.10: Respostas do sistema e do modelo do CSTR (dados de validação).

o algoritmo de controle proposto para controlar o peso molecular dentro do reator de polimerização. Os controladores preditivos lineares atuam localmente sobre cada modelo local do modelo *fuzzy* TS e as ações destes controladores locais são combinadas conforme a ativação das respectivas regras, como apresentado na seção 4.4. Os parâmetros dos controladores foram sintonizados experimentalmente como $N_1 = 1$ (horizonte inicial), $N_u = 2$ (horizonte de controle), $N_y = 3$ (horizonte de previsão) e $\lambda = 0,001$ (fator de ponderação do sinal de controle). Habilita-se o controle em malha fechada após um período transitório de 1h, necessário para convergência dos estados de Laguerre (que inicialmente são nulos), e que ao longo do qual mantém-se constante a variável manipulada u (ou seja, neste período não há ação do sistema de controle). Neste exemplo, os controladores locais apresentam a mesma parametrização. A Fig. 5.11 mostra a resposta do sistema CSTR em malha fechada sob a ação do sistema de controle proposto neste trabalho.

Como pode ser verificado na Fig. 5.11 o sistema de controle funcionou adequadamente e a concentração manteve o valor desejado, seguindo a referência estipulada. Em (Oliveira e Amaral, 2000) foi utilizado o mesmo modelo de polimerização para avaliar o desempenho de controladores preditivos baseados em modelos com estrutura Volterra-FBO. O desempenho do controlador proposto neste trabalho e ilustrado na Fig. 5.11 apresenta um desempenho superior ao da referência citada, uma vez que apresenta máximo *overshoot* e tempos de subida e acomodação inferiores. Em (Campello, 2002) também foi utilizado o modelo CSTR para avaliação de um controlador GPC baseado na linearização de modelos TS-FBO. O desempenho do controlador proposto nesta dissertação quando comparado ao proposto na citada referência, apresenta resultados de igual qualidade com *overshoot* máximo e tempo de subida com valores semelhantes. Contudo, o controlador proposto neste trabalho apresenta vantagens do ponto de vista estrutural uma vez que trabalha com a metodologia de sistema de

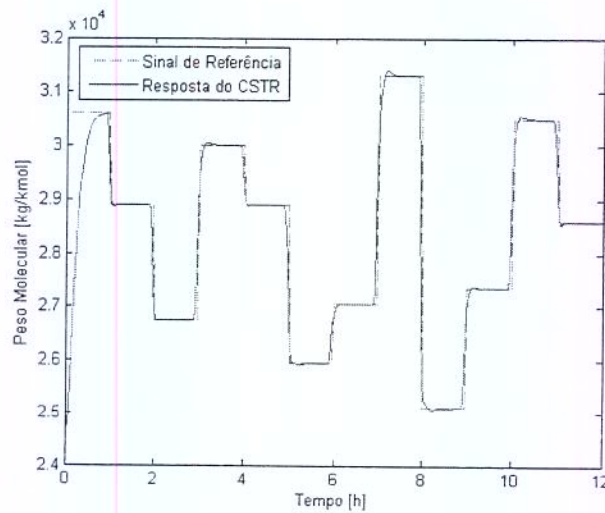


Fig. 5.11: Resposta CSTR sob ação do controlador.

controle distribuído dando maior interpretabilidade ao controlador aplicado a sistemas não-lineares possibilitando sintonias individuais dos controladores locais que atuam em determinadas regiões de operação do sistema, ao contrário do apresentado na referência que é sintonizado globalmente. Estes resultados mostram que o controlador aqui proposto funciona tão bem quanto os apresentados, com bons resultados no controle de sistemas não-lineares.

Considere agora um segundo modelo de reator de polimerização do tipo CSTR onde, diferentemente do exemplo anterior em que a dinâmica do sistema era superamortecida em todo intervalo de operação, este sistema apresenta uma resposta com dinâmica variando de sobreamortecida, para entradas de menor valor, para resposta oscilatória subamortecida para sinais de entrada de maiores amplitudes. Este sistema CSTR se diferencia do anterior devido à características inerentes a este processo, sendo seu modelo diferencial dado por:

$$\dot{C}_A = \frac{q}{V}(C_0 - C_A) - k_0 \exp\left(-\frac{E}{RT}\right)C_A$$

$$\dot{T} = \frac{q}{V}(T_0 - T) - \frac{(-\Delta H)k_0 C_A}{\rho C_p} \exp\left(-\frac{E}{RT}\right)C_A + \frac{\rho_c C_c}{\rho C_p V} q_c \left(1 - \exp\left(\frac{-hA}{q_c \rho_c C_c}\right)\right)(T_C - T)$$

onde C_A é a concentração do polímero, T é a temperatura do reator, T_C a temperatura do refrigerante, q é a taxa de fluxo de alimentação e C_0 é a concentração da alimentação. O objetivo do controle é manter a concentração da reação, que é obtida através da manipulação do fluxo de elemento refrigerante no tanque (Morningred, Paden, Seborg e Mellicamp, 1992). Os valores nominais dos parâmetros deste modelo estão presentes na Tab. 5.4:

A Fig. 5.12 mostra a resposta da concentração C_a no CSTR em malha aberta (Moor, 1998). Como

Tab. 5.4: Parâmetros do modelo CSTR

Concentração inicial do produto	C_{A0}	0 mol/l
Temperatura de reação	T	$438,54 \text{ K}$
Taxa de fluxo do refrigerante	q_{c0}	$103,41 \text{ l.min}^{-1}$
Taxa de fluxo do processo	q	100 l.min^{-1}
Concentração de Alimentação	C_0	1 mol/l
Temperatura de Alimentação	T_0	350 K
Temperatura do refrigerante na alimentação	T_C	350 K
Volume do CSTR	V	100 l
Coeficiente de transferência de calor	hA	$7 \cdot 10^5 \text{ cal min}^{-1} \text{ K}^{-1}$
Constante de reação	k_0	$7,2 \cdot 10^{10} \text{ min}^{-1}$
Termo de energia de ativação	E/R	$1 \cdot 10^4 \text{ K}$
Calor de reação	ΔH	$-2 \cdot 10^5 \text{ cal/mol}$
Densidade dos líquidos	ρ, ρ_C	$1 \cdot 10^3 \text{ g/l}$
Calor específico	C_p, C_C	$1 \text{ cal g}^{-1} \text{ K}^{-1}$

pode se verificar através da característica da resposta do sistema aos diversos níveis de entrada, trata-se claramente de um sistema não-linear, uma vez o ganho e a dinâmica do sistema variam conforme a região de operação do sistema.

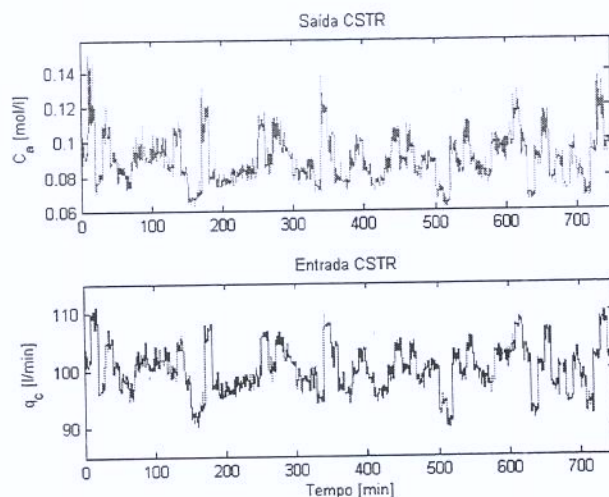


Fig. 5.12: Sinal de entrada e Resposta em malha aberta do sistema CSTR.

Os dados apresentados na Fig. 5.12 serão utilizados no processo de estimação e validação do modelo, sendo os primeiros 5000 pontos utilizados na estimação e os 2500 restantes utilizados na validação do modelo. Estes dados foram amostrados com período de 0,1 min num total de 750 minutos. Como pode ser observado o sistema foi excitado com um sinal em que a taxa de fluxo do elemento refrigerante varia entre 89 a 111 l/min. Para a identificação os dados de entrada e saída

foram normalizados e as 60 primeiras amostras excluídas do processo de identificação para garantir que os estados de Laguerre converjam para valores adequados (são inicialmente nulos) e não interfiram negativamente na obtenção do modelo. Como já descrito anteriormente, para a modelagem do processo pelo método proposto neste trabalho é necessário, além da entrada e saída, obter os estados da FBO em espaço de estados escolhida. Para este sistema, através de análise da resposta temporal da saída verificou-se que o pólo real $p = 0,7$ é capaz de modelar a dinâmica dominante (uma análise sobre a escolha do pólo será abordada a seguir nesta seção). Foram utilizados, na elaboração do modelo, uma FBO em representação em espaço de estados com oito estados (oito filtros) e utilizou-se o primeiro estado de Laguerre para determinar os antecedentes das regras. Aplicando o algoritmo de GK para o intervalo de 2 a 10 grupos, conforme descrito na seção 3.5.1, foram obtidos os resultados para a média e o produto das respostas dos critérios de avaliação conforme apresentado na Fig. 5.13. Em ambos os casos indicam que 8 grupos são necessários para se modelar o sistema de maneira eficiente.

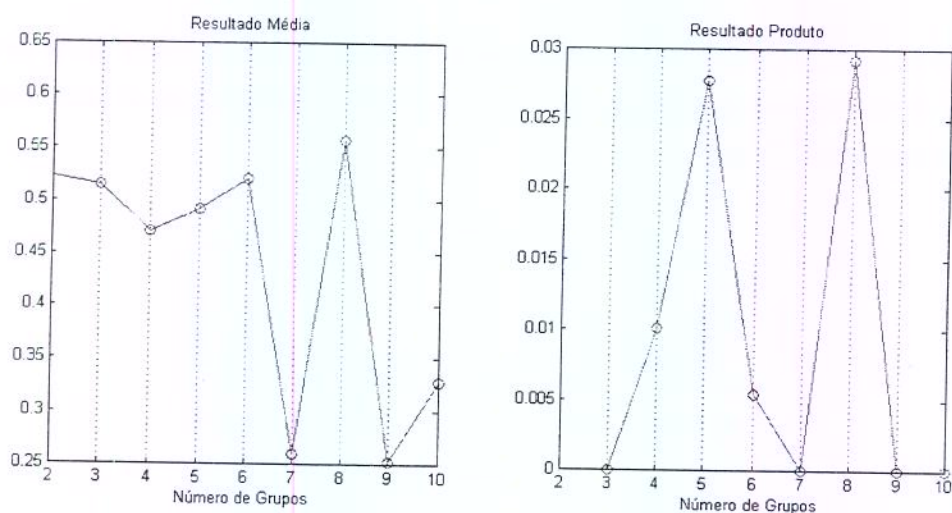


Fig. 5.13: Resultado do produto e média dos critérios de avaliação do número de grupos - CSTR 1.

Obtidos os grupos, determina-se as funções de pertinência dos antecedentes das regras, utilizando-se os valores de ativação dos objetos nos grupos, projetados nos subespaços de entrada conforme descrito na seção 3.6.1. A Fig. 5.15 mostra as funções de pertinência obtidas para os respectivos antecedentes das regras. Estas cobrem todo o universo de discurso normalizado, garantindo assim que o modelo TS tem validade em todo o universo de discurso.

Uma vez determinados os antecedentes das regras, obtém-se então os coeficientes dos modelos locais presentes nos consequentes das regras, sendo este processo realizado através do método de estimação local descrito na seção 3.6.2. Desta maneira o processo de identificação do modelo TS-FBO está concluído. Para verificar se o modelo é de boa qualidade realiza-se uma simulação com o

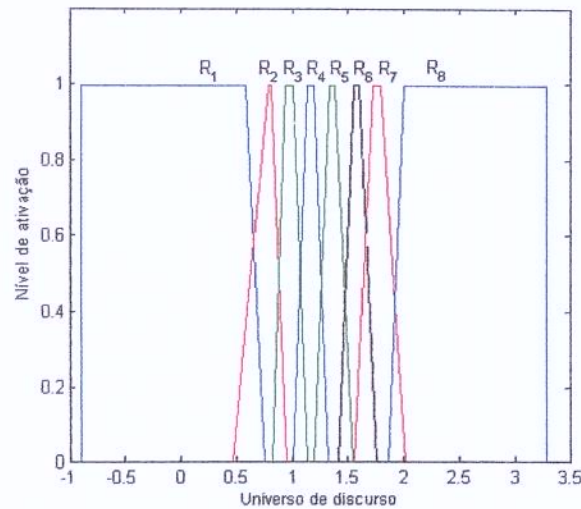


Fig. 5.14: Funções de pertinência presentes nos antecedentes das regras do modelo TS-FBO que modela o CSTR.

modelo obtido tendo como entrada os dados de validação já citados. A Fig. 5.15 traz a comparação entre os sinais estimados e os destinados à validação. Como pode ser verificado, o modelo obtido apresenta comportamento dinâmico correspondente ao sistema modelado, apresentando um $EQM = 5,8263 \cdot 10^{-5}$. Para fins de comparação, o modelo FBO com um modelo somente e pólo de Laguerre $p = 0,7$ apresenta um $EQM = 67,390 \cdot 10^{-5}$, ilustrando a qualidade do modelo TS-FBO obtido.

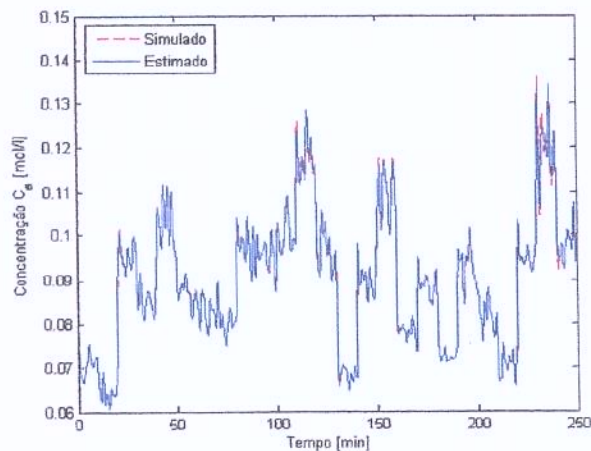


Fig. 5.15: Respostas do sistema e do modelo do CSTR (dados de validação).

A Tab. 5.5 mostra um comparativo entre o EQM de diversos modelos TS-FBO com diferentes valores de pólos utilizados na identificação, onde o objetivo é verificar se de fato o pólo $p = 0,7$ é

uma boa escolha.

Tab. 5.5: Comparação entre os EQM's para modelos TS-FBO do CSTR com diferentes pólos de Laguerre

Pólo utilizado	EQM
0,68	$10,1039.10^{-5}$
0,70	$5,8263.10^{-5}$
0,72	$7,4282.10^{-5}$
0,75	$7,4398.10^{-5}$

Após verificado que o modelo identificado apresenta um desempenho satisfatório, utiliza-se o algoritmo de controle proposto para controlar a concentração dentro do reator de polimerização. Assim como no exemplo anterior, os controladores preditivos lineares atuam localmente sobre cada modelo local do modelo *fuzzy* TS e suas ações de controle locais são combinadas conforme a ativação das respectivas regras, como apresentado na seção 4.4. Os parâmetros dos controladores foram sintonizados experimentalmente como $N_1 = 1$ (horizonte inicial), $N_u = 3$ (horizonte de controle), $N_y = 10$ (horizonte de previsão) e $\lambda = 0.2$ (fator de ponderação do sinal de controle). O controle em malha fechada é habilitado após um período transitório de 1 hora, período este necessário para convergência dos estados de Laguerre (que inicialmente são nulos) e que ao longo do qual mantém-se constante a variável de entrada q_c . Neste exemplo, todos os controladores locais apresentam a mesma parametrização. As Fig. 5.16 e 5.17² mostram a resposta do sistema CSTR em malha fechada sob a ação do sistema de controle proposto neste trabalho em diferentes pontos de operação.

Como pode ser verificado nas Fig. 5.16 e 5.17 o sistema de controle funcionou adequadamente fazendo com que a concentração seguisse os valores de referência desejados. Nestas figuras é comparado o desempenho do controlador proposto com o resultado de três sistemas de controle apresentados em (Espinosa et al., 2004), sendo dois deles sistemas de controle linear muito utilizados na indústria e um sistema de controle não-linear. Os sistemas de controle linear são um PID e um Controlador Preditivo Generalizado (GPC) e o sistema de controle não-linear é um controlador preditivo baseado em modelo *fuzzy*, cujo modelo identificado é um modelo NOE (nonlinear output error) com 81 regras, sendo considerados neste os seguintes regressores $C_a(k)$, $C_a(k-1)$, $C_a(k-2)$, $q_c(k-1)$. As Fig. 5.16 e 5.17 ilustram o desempenho destes controladores em duas distintas regiões de operação do sistema CSTR, sendo para a Fig. 5.16 com concentração em torno de 0,085 mol/l e para a Fig. 5.17 em torno de 0,12 mol/l. Como pode ser verificado, o desempenho do controlador proposto neste trabalho e apresentado na Fig. 5.16 se apresenta superior com relação ao desempenho do controlador *fuzzy* preditivo e dos demais controladores apresentados na referência citada, já que possui

²Estes gráficos se baseiam nos gráficos que aparecem na página 165 de (Espinosa et al., 2004)

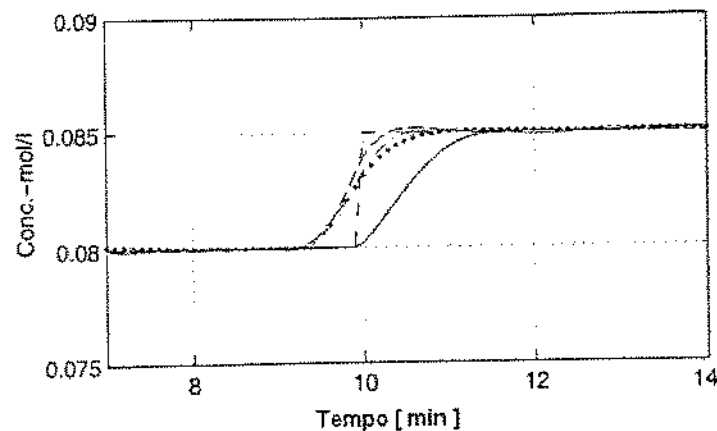


Fig. 5.16: Resposta do sistema CSTR sob ação do controlador proposto com *set point* de 0,08 [mol/l]
- Preditivo *fuzzy*(.-), PID (-), GPC (.), Preditivo TS-FBO (- -).

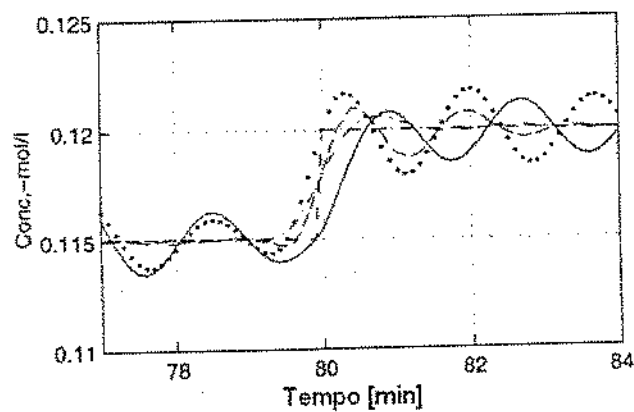


Fig. 5.17: Resposta do sistema CSTR sob ação do controlador proposto com *set point* de 0,12 [mol/l]
- Preditivo *fuzzy*(.-), PID (-), GPC (.), Preditivo TS-FBO (- -).

menor tempo de subida com o mesmo tempo de acomodação da resposta de saída em malha fechada. Com relação a Fig. 5.17, onde a região de operação se encontra em torno de 0,12 mol/l, pode-se observar que o desempenho do algoritmo de controle proposto em relação ao controlador preditivo *fuzzy* apresentado em (Espinosa et al., 2004) apresenta um menor tempo de subida, com *overshoot* máximo superior e resposta de saída em malha fechada com um tempo de acomodação também um pouco superior. Este desempenho deve-se ao fato de que a dinâmica do sistema é bastante oscilatória nessa região de operação, dificultando a modelagem do processo por FBO's com um pólo real. Uma possível melhora no controlador proposto poderá ser obtida com a utilização de modelos locais com FBO's apresentando pólos distintos, o que deve ser avaliado em trabalhos futuros.

Com estes exemplos pôde-se ilustrar o desempenho das estruturas de identificação e controle proposta neste trabalho, mostrando que os mesmos apresentam um bom desempenho na modelagem e controle de sistemas não-lineares.

5.3 Resumo

Este capítulo teve como objetivo apresentar implementações e resultados que exemplifiquem a aplicação das técnicas de modelagem e controle investigados neste trabalho.

Na primeira seção mostrou-se a aplicação da técnica de identificação proposta fazendo-se a modelagem de um sistema não-linear, exemplificado por uma planta de levitação magnética presente no laboratório de sistemas e controle desta instituição. Como se pôde verificar, utilizando a metodologia proposta, com um modelo TS contando com poucos modelos locais FBO foi possível obter um modelo não-linear com precisão satisfatória, reproduzindo o comportamento dinâmico deste sistema que é não-linear, apresenta saturação de ganho e é variante no tempo devido ao atrito seco entre o disco magnético e seu guia.

Já para exemplificar a aplicação do sistema de controle proposto, a seção seguinte trouxe a aplicação do controlador no controle de sistemas não-lineares, sendo utilizada duas plantas CSTR nos exemplos. Este tipo planta é muito utilizada como *benchmark* em aplicações de controladores preditivos não-lineares (Maner et al., 1996; Espinosa et al., 2004). Assim como no caso da planta de levitação, os resultados na modelagem foram satisfatórios, e além disso a atuação do controlador sobre os sistemas garantiu que a resposta do sistema em malha fechada seguisse o sinal de referência como desejado.

Desta maneira, fica exemplificada a eficiência dos métodos propostos de modelagem e controle de sistemas não-lineares.

Capítulo 6

Conclusões

Neste trabalho apresentou-se uma abordagem para a modelagem e controle de sistemas não-lineares. A identificação é feita através de modelos *fuzzy* Takagi-Sugeno com funções de base ortonormal nos conseqüentes das regras do modelo. Esta abordagem de modelos *fuzzy* TS-FBO já foi apresentada em outros trabalhos (Oliveira et al., 1999; Campello, 2002; Campello e Amaral, 2002; Campello et al., 2004; Medeiros et al., 2006) e vem demonstrando sua eficácia na modelagem de sistemas dinâmicos não-lineares, o que em parte motivou sua utilização neste trabalho.

Como citado no capítulo 2, os modelos construídos a partir de FBO's apresentam uma série de vantagens se comparados a outras possíveis abordagens utilizadas na identificação de sistemas dinâmicos não-lineares. A utilização de uma topologia dinâmica do tipo FBO nos modelos TS se torna vantajosa uma vez que os modelos baseados em funções de base ortonormal apresentam características desejáveis como a ausência de realimentação da saída, o que leva também à ausência eventuais erros de predição, necessidade de um menor número de parâmetros para alcançar uma dada precisão, escolha arbitrária dos pólos que definem as funções da base ortonormal que irão modelar o sistema, representação de atrasos de transporte e tolerância a dinâmicas não modeladas ou a pequenas diferenças na ordem dos vetores de regressão. Além disso, os modelos *fuzzy* TS baseados em funções de base ortonormais, além das características inerentes à modelagem por dinâmica FBO, são aproximadores universais de funções em um espaço compacto e agregam as propriedades de interpretabilidade, úteis no projeto de sistemas de controle.

Um dos objetivos desta dissertação foi apresentar uma outra abordagem na obtenção destes modelos e na implementação do sistema de controle. Para determinação das regras do modelo *fuzzy* foi utilizada uma técnica de agrupamento *fuzzy* (*fuzzy clustering*) proposta por Gustafson e Kessel (1979), a partir da qual foram definidos os antecedentes e conseqüentes das regras. Utilizando critérios de avaliação dos agrupamentos foi possível determinar o número de modelos locais a compor o modelo TS-FBO.

Uma vez determinados os modelos locais, implementou-se a etapa de controle, que consiste na aplicação de controladores preditivos (Oliveira, 1997) atuando localmente sobre os modelos FBO determinados previamente. A partir da combinação da ação destes controladores através da máquina de inferência *fuzzy* se obtém a ação de controle global que será aplicada ao processo a fim de conseguir que o sistema siga o sinal de referência.

O Capítulo 5 apresentou alguns exemplos que exemplificam a aplicação tanto na modelagem quanto no controle de processos não-lineares. A modelagem do sistema de levitação (seção 5.1) mostrou a eficácia do método proposto quando aplicado a um sistema não-linear real, apresentando resultados, para modelagem com somente um pólo nos modelos locais, com menor EQM que nos resultados apresentados em (Medeiros, 2006). A implementação do sistema de controle aplicado a uma planta de polimerização CSTR ilustrou a eficiência do controlador quando aplicado no controle de sistemas não-lineares.

Assim pode-se concluir que as propostas tanto para a identificação quanto no controle funcionam satisfatoriamente, fornecendo uma alternativa na modelagem e controle de sistemas dinâmicos não-lineares, sendo um atrativo nessa modelagem o fato de se trabalhar somente com modelos lineares, que por se tratarem de modelos FBO, gozam de todas as vantagens já descritas anteriormente, e assim possibilitam uma abordagem mais simples e com menor esforço computacional na modelagem de sistemas não-lineares.

6.1 Contribuições

Neste trabalho algumas contribuições foram dadas no sentido tornar possível as implementações propostas ou o aperfeiçoamento de técnicas já utilizadas. Algumas destas contribuições podem ser destacadas:

- Alteração no critério de avaliação AWCD utilizado na determinação do melhor número de grupos no agrupamento *fuzzy*. Com as contribuições apresentadas em 3.5.2 pôde-se melhorar a avaliação do melhor número de grupo indicado pelo critério AWCD, tornando esta análise mais objetiva uma vez que a mesma até então dependia exclusivamente da percepção da pessoa ao avaliar os resultados apresentados tendo em vista que o melhor número de grupos era indicado pelo ponto onde havia um “joelho” na curva de resposta gerada pelo critério.
- Ainda na seção 3.5.2 é proposta uma metodologia para se combinar as respostas dos critérios de avaliação do agrupamento dos dados. Essa combinação permite, através da análise de múltiplos critérios diferentes inferir, qual o melhor número de grupos a se utilizar na determinação de um eficiente modelo *fuzzy* TS.

- Na seção 3.6.1 foi apresentada uma metodologia desenvolvida a fim de determinar as funções de pertinência dos antecedentes das regras *fuzzy* a partir de um método iterativo que utiliza a projeção dos objetos presentes nos grupos, obtidos pelo algoritmo de Gustafson-Kessel, nos subespaços de entrada (formados pelos estados FBO).
- A proposta do controlador preditivo não-linear que tem sua ação de controle dada pela combinação das ações dos controladores preditivos que atuam localmente sobre os modelos locais das regras do modelo *fuzzy* TS-FBO.

6.2 Trabalhos Futuros

Como continuidade dos trabalhos aqui apresentados ficam algumas propostas a seguir que irão complementar o que foi desenvolvido até aqui:

- Simplificação das bases de regras do modelo TS via medidas de similaridade (Setnes, 1998).
- Otimização dos pólos e funções de pertinência - Uma proposta a ser implementada como continuação deste trabalho será uma otimização dos pólos e funções de pertinência a fim de gerar um ajuste fino nos modelos locais.
- Estudo de robustez e estabilidade do sistema de controle - Deseja-se futuramente se realizar estudos de robustez e estabilidade para o sistema de controle proposto a fim de que o mesmo tenha tais garantias.
- Implementação de modelagem e sistemas de controle com modelos locais com pólos distintos - Uma das propostas de continuação deste trabalho consiste em utilizar em controle modelos *fuzzy* TS-FBO que utilizem não somente um pólo em todos os modelos locais, mas que utilizem modelos com pólos distintos (Campello, 2002; Campello e Amaral, 2002; Medeiros, 2006; Medeiros et al., 2006) o que pode melhorar a qualidade dos modelos locais, exigindo menos FBO's que se adequem melhor a determinada dinâmica presente no comportamento não-linear do sistema a ser modelado e controlado.
- Aplicação no controle de sistemas não-lineares reais - Em trabalhos futuros, o controlador apresentado neste trabalho será aplicado para realizar o controle de processos não-lineares reais a fim de comprovar sua eficácia também com resultados obtidos em sistemas reais além dos resultados já obtidos com simulações.

Bibliografia

- Abonyi, J. e Babuska, R. (2000). Local and global identification and interpretation of parameters in takagi-sugeno fuzzy models, San Antonio, USA, pp. 835–840.
- Aguirre, L. A. (2000). *Introdução à Identificação de Sistemas*, Editora UFMG, Belo Horizonte.
- Aguirre, L. e Billings, S. A. (1995). Dynamical effects of overparametrization in nonlinear models, *Physica D*. **80**: 26–40.
- Aguirre, L., Rodrigues, G. e Jácome, C. (1998). Identificação de sistemas não lineares utilizando modelos narmax polinomiais - uma revisão e novos resultados, *Controle e Automação* **9**(2): 90–106.
- Astrom, K. J. (1970). *Introduction to Stochastic Control Theory*, Academic Press, New York.
- Babuska, R., Van der Veen, P. e Kaymak, U. (2002). Improved covariance estimation for gustafson-kessel clustering, USA, pp. 1081 – 1085.
- Babuska, R. e Verbruggen, H. B. (1996). An overview of fuzzy modeling for control, *Control Engineering Practice* **4**(11): 1593–1606.
- Babuska, R. e Verbruggen, H. B. (1997). Fuzzy set methods for local modelling and identification, in R. Murray-Smith e T. A. Johansen (eds), *Multiple Model Approaches to Modelling and Control*, Taylor and Francis, London, chapter 2.
- Bazaraa, M. S. e Shetty, C. M. (1979). *Nonlinear Programming*, John Wiley and Sons.
- Bezdek, J. (1981a). *Pattern Recognition with Fuzzy Objective Function*, 1st edn, Plenum Press, New York.
- Bezdek, J. C.; Coray, C. G. R. W. J. (1981b). Detection and characterization of cluster substructure, i. linear structure: Fuzzy c-lines., *SIAM Journal on Applied Mathematics* **40**(2): 339–357.

- Bezdek, J. C.; Coray, C. G. R. W. J. (1981c). Detection and characterization of cluster substructure ii. fuzzy c- varieties and convex combinations thereof, *SIAM Journal on Applied Mathematics* 40(2): 358–372.
- Camacho, E. F. e Bordons, C. (1998). *Model Predictive Control*, 1st edn, Springer Verlag.
- Campello, R. J. G. B. (2002). *Arquiteturas e Metodologias para Modelagem e Controle de Sistemas Complexos Utilizando Ferramentas Clássicas e Modernas*, PhD thesis, Universidade Estadual de Campinas, Campinas.
- Campello, R. J. G. B. e Amaral, W. C. (2002). Takagi-sugeno fuzzy models within orthonormal basis function framework and their application to process control, *2002 IEEE International Conference on Fuzzy Systems*, pp. 1399 – 1404.
- Campello, R. J. G. B. e Hruschka, E. R. (2006). A fuzzy extension of the silhouette width criterion for cluster analysis, *Fuzzy Sets and Systems* 157(21): 2858–2875.
- Campello, R. J. G. B., Meleiro, L. A. C. e Amaral, W. C. (2004). Control of a bioprocess using orthonormal basis function fuzzy models, *2004 IEEE International Conference on Fuzzy Systems*, pp. 801 – 806.
- Chen, C. T. (1999). *Linear Systems Theory and Design*, 3rd edn, Oxford University Press.
- Clarke, D. W. (1994). *Advances in Model Based Predictive Control*, Oxford University Press.
- Clarke, D. W., Mohtadi, C. e S., T. P. (1987). Generalized predictive control - part i. the basic algorithm and part ii. extensions and interpretations, *Automatica* 23(2): 137–160.
- Cutler, C. R. e Ramaker, B. L. (1980). Dynamic matrix control - A computer control algorithm, *Joint Automatic Control Conf.*
- Delgado, M. R., Von Zuben, F. e Gomide, F. (2001). Local and global estimation of takagi-sugeno consequent parameters in genetic fuzzy systems, *Joint 9th IFSA World Congress and 20th NAFIPS International Conference*, Vancouver, Canada, pp. 1247–1252.
- Dubois, D. e Prade, H. (1984). Fuzzy logics and the generalized modus ponens revisited, *Cybernetics and Systems* 15(3-4): 293–332.
- Espinosa, J., Vandewalle, J. e Wertz, V. (2004). *Fuzzy Logic, Identification, And Predictive Control*, 1st edn, Springer Verlag.
- Everitt, B. S., Landau, S. e Leese, M. (2001). *Cluster analysis*, 4rd edn, Arnold.

- Friedman, J. H. (1991). Multivariate adaptive regression splines, *The Annals of Statistics* 19(1): 1–67.
- Gath, I. e Geva, A. B. (1989). Unsupervised optimal fuzzy clustering, *IEEE Transactions on pattern analysis and machine intelligence* 2(7): 773–781.
- Godjevac, J. (1997). *Neuro-Fuzzy Controllers Design and Application*, Presses Polytechniques et Universitaires Romandes, Lausanne, Switzerland.
- Gupta, M. M., Kandel, A., Bandler, W. e Kiszka, J. B. (1985). *The generalized modus ponens under sup-min composition - A theoretical study*, Eds. Amsterdam: North Holland.
- Gustafson, D. e Kessel, W. C. (1979). Fuzzy clustering with a fuzzy covariance matrix, *Proc. IEEE CDC* pp. 761–766.
- Hathaway, R. e Bezdek, J. (1993). Switching regression models and fuzzy clustering, *IEEE Transactions on Fuzzy Systems* 1(3): 195–204.
- Haykin, S. (1999). *Neural Networks, A Comprehensive Foundation*, 2nd edn, Prentice Hall.
- Hellendoorn, H. e Driankov, D. (1997). *Fuzzy Model Identification: Selected Approaches*, Springer-Verlag.
- Hruschka, E. R., Campello, R. J. G. B. e de Castro, L. N. (2004). Evolutionary search for optimal fuzzy c-means clustering, *13th IEEE Internat. Conf. on Fuzzy Systems*, Hungary, pp. 685–690.
- Irwing, E., Falinower, C. M. e Fonte, C. (1986). Adaptive generalized predictive control with multiple reference control, *2nd IFAC Workshop on Adaptive Control and Signal Processing* pp. 384–386.
- Johansen, T. A., Shorten, R. e Murray-Smith, R. (2000). On the interpretation and identification of dynamic takagi-sugeno fuzzy models, *IEEE Transactions on Fuzzy Systems* 8(3): 297–313.
- Kaufman, L. e Rousseeuw, P. J. (1990). *Finding Groups in Data: an Introduction to Cluster Analysis*, 1st edn, John Wiley and Sons.
- Kosko, B. (1992). *Neural Network and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*, 1st edn, Prentice Hall.
- Kosko, B. (1994). Fuzzy systems as universal approximators, *IEEE Transaction on Computers* 43(11): 1329–1333.
- Krishnapuram, R. e Freg, C.-P. (1992). Fitting an unknown number of lines and planes to image data through compatible cluster merging, *Pattern Recognition* 25: 385–400.

- Lee, C. C. (1990). Fuzzy logic in control systems: Fuzzy logic controller - parts 1 and 2, *IEEE Transaction on Systems, Man and Cybernetics* **20**(2): 404–435.
- Ljung, L. (1999). *System identification: Theory for the user*, 2nd edn, Prentice Hall.
- Lopez, J. F.; Oliveira, G. H. C. F. G. A. E. A. W. C. L. L. G. (1995). Multivariable constrained predictive control applied to the shell benchmark problem: A comparison, *Proc. of the 3rd European Control Conference*.
- Maner, B. R., Doyle, F. J., Ogunnaike, B. A. e PearsonMcCulloch, R. K. (1996). Nonlinear model predictive control of a simulated multivariable polymerization reactor using second-order volterra models, *Automatica* **32**(9): 1285–1301.
- McCulloch, W. S. e Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activities, *Bulletin of Mathematical Biophysics* **5**: 115–133.
- Medeiros, A., Amaral, W. e Campello, R. (2006). Ga optimization of generalized obf ts fuzzy models with global and local estimation approaches, *2006 IEEE International Conference on Fuzzy Systems*, Vol. 5, pp. 1835 – 1842.
- Medeiros, A. V. (2006). *Modelagem de sistemas dinâmicos não-lineares utilizando sistemas fuzzy, algoritmo genético e funções de base ortonormal*, Master's thesis, Universidade Estadual de Campinas, Campinas, Brasil.
- Moor, B. D. (1998). Daisy: Database for the identification of systems, *Continuous Stirred Tank Reactor*, Departament of Electrical Engeneering - ESAT, Leuven, Belgium. <http://www.esat.kuleuven.ac.be/sista/daisy>.
- Morningred, J. D., Paden, B. E., Seborg, D. E. e Mellicamp, D. A. (1992). An adaptive nonlinear predictive controller, *Chemical Engineering Science* **47**(4): 755–762.
- Murray-Smith, R. (1994). *A Local Model Network Approach to Nonlinear Modeling*, PhD thesis, University of Strathclyde, U.K.
- Ogata, K. (2003). *Engenharia de Controle Moderno*, 4th edn, Prentice-Hall.
- Oliveira, G., Campello, R. e Amaral, W. (1999). Fuzzy models within orthonormal basis function framework, *1999 IEEE International Conference on Fuzzy Systems*, pp. 957 – 962.
- Oliveira, G. H. C. (1997). *Controle Preditivo para Processos com incertezas estruturadas baseado em séries de funções ortonormais*, PhD thesis, Universidade Estadual de Campinas, Campinas.

- Oliveira, G. H. C. e Amaral, W. C. (2000). Identificação e controle preditivo de processos não lineares utilizando séries de Volterra e bases de funções ortonormais, *XIII Congresso Brasileiro de Automática* pp. 2174–2179.
- Pedrycz, W. e Gomide, F. A. C. (1998). *An Introduction to Fuzzy Sets: Analysis and Design (Complex Adaptive Systems)*, 1st edn, MIT Press.
- Phillips, L. C. e Harbor, R. D. (1999). *Feedback Control Systems*, 4th edn, Prentice Hall, Inc., New Jersey.
- Phillips, L. C. e Nagle, T. H. (1995). *Digital control system analysis and design*, 3rd edn, Prentice Hall, Inc., New Jersey.
- Richalet, J., Rault, A., Testud, J. L. e Papon, J. (1978). Model predictive heuristic control : Applications to industrial processes, *Automatica* 14(5): 413–428.
- Rosa, A. (2005). *Desenvolvimento de modelos discretos de volterra usando funções de kautz*, Master's thesis, Universidade Estadual de Campinas, Campinas, Brasil.
- Rossiter, J. A. (2003). *Model-Based Predictive Control*, 1st edn, CRC Press, New York.
- Schott, K. D. e Bequette, B. W. (1997). Multiple model adaptive control, in R. Murray-Smith e T. A. Johansen (eds), *Multiple Model Approaches to Modelling and Control*, Taylor and Francis, London, chapter 11.
- Setnes, M.; Babuska, R. K. U. V. N. L. H. R. (1998). Similarity measures in fuzzy rule base simplification, *IEEE Transactions on Systems, Man and Cybernetics - Part B* 28(3): 376–386.
- Setnes, M., Babuska, R. e Verbruggen, H. B. (1998). Rule-based modeling: Precision and transparency, *IEEE Transactions on Systems, Man and Cybernetics - Part C: Applications and Reviews* 28(1): 165–169.
- Shorten, R., Murray-Smith, R., Bjorgan, R. e Gollee, H. (1999). On the interpretation of local models in blended multiple model structures, *Int. Journal of Control* 72(7-8): 620–628.
- Sugeno, M. e Yasukawa, T. (1993). Fuzzy-logic-based approach to qualitative modeling, *IEEE Transactions on Fuzzy Systems* 1(1): 7–31.
- Takagi, T. e Sugeno, M. (1985). Fuzzy identification of systems and its application to modeling and control, *IEEE Transaction on Systems, Man and Cybernetics* 15(1): 116–132.

- Wahlberg, B. (1991). System identification using Laguerre models, *IEEE Transactions on Automatic Control* **36**(5): 551–562.
- Wahlberg, B. (1994). System identification using Kautz models, *IEEE Transactions on Automatic Control* **39**(6): 1276–1282.
- Wang, L. X. (1998). Universal approximation by hierarchical fuzzy systems, *Fuzzy Sets and Systems* **93**(2): 223–230.
- Wang, L.-X. e Mendel, J. M. (1992). Fuzzy basis functions, universal approximation, and orthogonal least-square learning, *IEEE Transaction on Neural Network* **43**(5): 807–814.
- Wiener, N. (1956). *The Theory of Prediction*, McGraw-Hill, New York.
- Yager, R. R. e Filev, D. P. (1994). *Essentials of Fuzzy Modeling and Control*, John Wiley and Sons.
- Yeung, D. Y. e Bekey, G. A. (1993). On reducing learning time in context-dependent mappings, *IEEE Transaction on Neural Networks* **4**(1): 31–42.
- Zadeh, L. A. (1965). Fuzzy sets, *Information and Control* **8**: 338–353.
- Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Transaction on Systems, Man and Cybernetics* **3**(2): 28–44.
- Zeng, X.-J. e Singh, M. G. (1994). Approximation theory of fuzzy system - SISO case, *IEEE Transaction on Fuzzy System* **2**(2): 162–176.
- Zeng, X.-J. e Singh, M. G. (1995). Approximation theory of fuzzy system - MIMO case, *IEEE Transaction on Fuzzy System* **3**(2): 219–235.
- Zervos, C. C. e Dumont, G. A. (1988). Deterministic adaptive control based on laguerre series representation, *Int. J. Control* **48**(6): 2333–2359.

Apêndice A

Método de Mínimos Quadrados

O método de mínimos quadrados foi originalmente proposto de forma independente por Gauss e Legendre. Gauss¹ durante suas observações astronômicas afirmou que:

"O valor mais provável de grandezas desconhecidas é o que minimiza a soma dos quadrados da diferença entre o valor medido e o valor calculado, ponderado pelo grau de precisão da medida".

Seja $\hat{\theta}$ o valor estimado de θ e $\hat{y}$ a saída estimada do sistema, dada por:

$$\hat{y} = U\hat{\theta} \quad (\text{A.1})$$

Sejam os erros entre os vetores de parâmetros desconhecidos e estimados e entre a saída do sistema e a saída do modelo definidos respectivamente por $\tilde{\theta} = \theta - \hat{\theta}$ e $\tilde{y} = y - \hat{y}$. Logo, tem-se que:

$$\tilde{y} = U\tilde{\theta} \quad (\text{A.2})$$

A equação (A.2) é denominada equação do erro. O estimador é obtido minimizando-se:

$$J(\hat{\theta}) = \omega(1)\tilde{y}_1^2 + \omega(2)\tilde{y}_2^2 + \dots + \omega(N)\tilde{y}_N^2$$

onde $\omega(i)$ é a ponderação em cada componente do erro, que é função da precisão da medida. Esta equação pode ser rescrita como segue:

$$J(\hat{\theta}) = \tilde{y}^T W \tilde{y}$$

¹ Johann Carl Friedrich Gauss (Braunschweig, 1777 - Göttingen, 1855) foi um famoso matemático, astrônomo e físico alemão. Era conhecido como o *príncipe dos matemáticos*. Muitos consideram Gauss o maior gênio da história da Matemática. Seu QI foi estimado em cerca de 240

onde $\tilde{y} = [\tilde{y}_1 \ \tilde{y}_2 \ \dots \ \tilde{y}_N]^T$ e $W = \text{diag}(\omega(1) \ \omega(2) \ \dots \ \omega(N))$. Deseja-se:

$$\hat{\theta}_M = \arg \min_{\hat{\theta}} J(\hat{\theta}) = \arg \min_{\hat{\theta}} \|y - U\hat{\theta}\|_W^2$$

sendo que a função $J(\hat{\theta})$ ainda pode ser rescrita como:

$$J(\hat{\theta}) = y^T W y - 2y^T W U \hat{\theta} + \hat{\theta}^T U^T W U \hat{\theta}$$

Sabe-se que $\hat{\theta}_M$ deve ser tal que $dJ/d\hat{\theta} = 0$, o que implica:

$$\frac{dJ}{d\hat{\theta}} = -2[y^T W U]^T + 2U^T W U \hat{\theta} = 0$$

$$[U^T W U] \hat{\theta} = U^T W y \quad (\text{A.3})$$

que é denominada Equação Normal e produz o seguinte estimador:

$$\hat{\theta}_M = [U^T W U]^{-1} U^T W y \quad (\text{A.4})$$

A matriz $[U^T W U]^{-1} U^T W$ é denominada matriz Pseudo-Inversa. O estimador (A.4) é denominado Estimador de Markov, Mínimos Quadrados Ponderados (*Weighted Least Squares - WLS*) ou ainda Mínimos Quadrados Generalizado. Quando $W = \sigma^2 I_N$, onde σ é um escalar e I_N é a matriz de identidade $N \times N$, obtém-se o estimador dos Mínimos Quadrados convencional, descrito por:

$$\hat{\theta}_{MQ} = [U^T U]^{-1} U^T y \quad (\text{A.5})$$

Neste caso tem-se que $\hat{\theta}_{MQ}$ é seguramente um ponto de mínimo pois a matriz Hessiana:

$$\frac{d^2 J}{d\hat{\theta}^2} = 2U^T W U$$

é definida positiva posto que $W = \sigma^2 I_N$ é definida positiva. Logo, $\hat{\theta}_{MQ}$ em (A.5) é uma solução ótima. Como é única, é uma solução ótima global. No caso geral do estimador em (A.4) a otimalidade requer apenas que a matriz de ponderação W seja definida positiva.

Apêndice B

Teoria de Conjuntos *Fuzzy*

B.1 Conjuntos *Fuzzy*

Um conceito fundamental em matemática é a noção de conjunto. Um conjunto é uma coleção de elementos específicos e discerníveis. Um conjunto pode ser infinito, finito contável ou incontável e pode ser descrito de três maneiras:

- Apresentando de forma explícita os elementos:

$$A = \{a, b, c, 10, \pi\} \Rightarrow a, b, \dots, \pi \in A; \quad \text{sendo que } k, \beta \notin A$$

- Enunciando uma propriedade comum a todos os elementos:

$$A = \{x | P(x)\}, \quad \text{onde } P \text{ define as propriedades dos elementos } x \text{ pertencentes a } A \\ (\text{p/ ex. } A = \{x | x > 3\})$$

- Pela definição de uma função característica para todos os elementos x do universo de discurso U que contém todos os objetos relacionados a um dado contexto. Assim, para um conjunto A pertencente a U , a função característica $\mu_A : U \rightarrow \{0, 1\}$ é definida como:

$$\mu_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases} \quad (\text{B.1})$$

A função (B.1) é também chamada de função de pertinência na teoria de conjuntos *fuzzy*, mas nesse caso é tal que $\mu_A : U \rightarrow [0, 1]$. Por exemplo, para a função (B.1) assumamos que U seja o conjunto de todos os positivos inteiros denotados como E^+ . Assim, para o conjunto de elementos de $A = \{x | x > 3\}$ a sua função característica pode ser definida como mostra a figura abaixo:

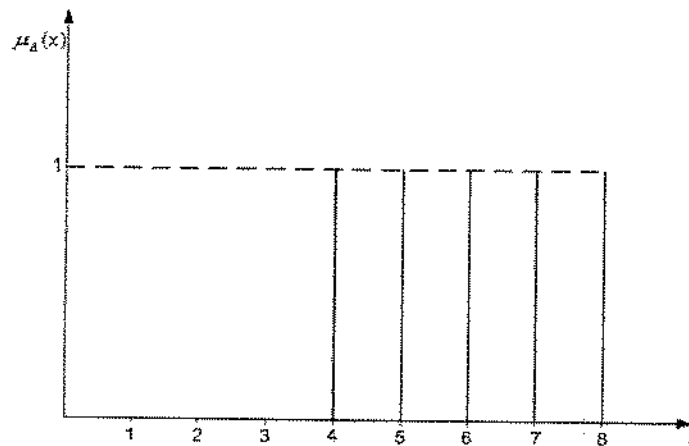


Fig. B.1: Função Característica para o conj. $A = \{x | x \in E^+ \text{ e } x > 3\}$

Os conjuntos que seguem a definição clássica de conjuntos como apresentada acima são também conhecidos como conjuntos *crisp* (não *fuzzy*). A seguir são definidas as principais operações com conjuntos. Sejam dois conjuntos A e B pertencentes ao mesmo universo de discurso U :

- Complemento: $\bar{A} = \{x | x \notin A\}$, $\mu_{\bar{A}}(x) = 1 - \mu_A(x)$
- Intersecção: $A \cap B = \{x | x \in A \text{ e } x \in B\}$, $\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$
- União: $A \cup B = \{x | x \in A \text{ ou } x \in B\}$, $\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$

Os símbolos “ $\wedge$ ” e “ $\vee$ ” também serão usados ao longo deste texto para descrever respectivamente intersecção e união. Os conjuntos, como caracterizados acima, ainda possuem as propriedades de comutatividade, associatividade e distributividade.

O conceito fundamental em teoria de conjuntos *fuzzy* é a idéia do conjunto *fuzzy* como: “um conjunto *fuzzy* é uma classe de objetos com uma função de pertinência com grau de ativação contínua no universo de discurso. Assim, um conjunto é caracterizado por uma função de pertinência que atribui a cada objeto um grau de ativação entre 0 e 1” (Zadeh, 1965). Um conjunto *fuzzy* F é inteiramente definido por um conjunto de par ordenado:

$$F = \{(x, \mu_F(x)) | x \in U\} \text{ e } \mu_F : U \rightarrow [0, 1] \quad (\text{B.2})$$

onde x é o elemento no universo de discurso U e μ_F é uma função de pertinência que atribui um grau de pertinência $\mu_F(x)$ para cada elemento x de F . A função de pertinência μ_F de um conjunto *fuzzy* F corresponde à função característica de um conjunto ordinário (*crisp*). No entanto, enquanto a função característica pode assumir somente valores 0 ou 1, a função de pertinência de um conjunto *fuzzy* pode assumir qualquer valor no intervalo $[0, 1]$. Assim o conjunto *fuzzy* é uma generalização do conjunto ordinário.

B.1.1 Representação do conjunto fuzzy

Igualmente aos conjuntos ordinários, os conjuntos *fuzzy* também podem ser finitos contáveis e terem seus elementos definidos explicitamente. Por exemplo, dado o universo de discurso $U = \{c, j, m, 5, \pi\}$, um conjunto *fuzzy* F de U pode ser definido como:

$$F = \{(c, 0.28), (j, 0.46), (m, 1), (5, 0.3), (\pi, 0)\}$$

onde para cada par ordenado deve-se ler que “ c ” pertence a F com grau de 0.28, “ j ” com 0.46, etc.

Quando o universo de discurso U é contínuo ou incontável, as funções de pertinência devem ser expressa analiticamente através de funções contínuas que mapeiam todo o universo de discurso. A figura abaixo mostra algumas das funções de pertinência mais utilizadas em aplicações com conjuntos *fuzzy*:

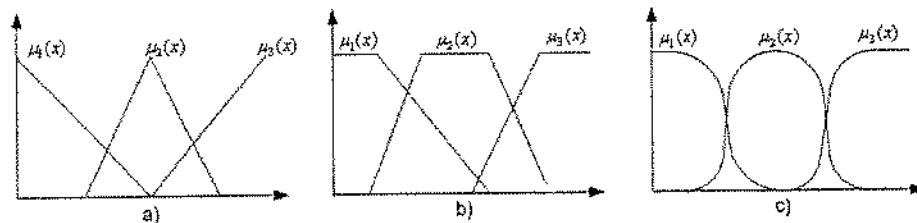


Fig. B.2: Funções de pertinência: a) Função Triangular b) Função Trapezoidal c) Função Gaussiana

B.1.2 Operações em conjuntos Fuzzy

As operações nos conjuntos *fuzzy* são definidas por meio de suas funções de pertinência para que propriedades de operações dadas para conjuntos ordinários possam ser generalizadas para conjuntos *fuzzy*. Sendo assim, se as funções de pertinência são tais que adotam somente valores 0 ou 1, os operadores *fuzzy* produzirão resultados semelhantes aos operadores para conjuntos ordinários. As operações mais importantes em conjuntos *fuzzy* serão definidas a seguir.

Complemento: dado um conjunto *fuzzy* A definido no universo de discurso U , seu complemento é dado pelo conjunto:

$$\bar{A} = \{(x, \mu_{\bar{A}}(x)) \mid x \in U\} \text{ e } \mu_{\bar{A}}(x) = \neg(\mu_A(x)) = 1 - \mu_A(x) \quad (\text{B.3})$$

onde “ $\neg$ ” denota a “negação *fuzzy*” da função de pertinência, e neste caso coincide com a operação correspondente para conjuntos ordinários. Há ainda outros operadores de complemento que geralmente são denominados Norma-C (Yager e Filev, 1994; Lee, 1990).

Intersecção: Sejam os conjuntos fuzzy A e B definidos em U com funções de pertinência μ_A e μ_B respectivamente. Sua intersecção é o conjunto fuzzy:

$$A \cap B = \{(x, \mu_{A \cap B}(x)) \mid x \in U\} \text{ com } \mu_{A \cap B}(x) = \mu_A \wedge \mu_B \quad (\text{B.4})$$

sendo " $\wedge$ " o operador que denota a intersecção de conjuntos fuzzy ou ainda a relação "e" fuzzy (conjunção).

O operador de intersecção pertence a classe também chamada de Norma Triangular ou Norma-T. Este tipo de operação está definida no intervalo $[0, 1]$ satisfazendo as seguintes propriedades:

- a. $T(0, 1) = T(1, 0) = T(0, 0) = 0, T(1, 1) = 1$; (condição de contorno)
- b. $T(x, 1) = x, \forall x \in [0, 1]$; (identidade)
- c. $T(x, y) = T(y, x), \forall x, y \in [0, 1]$; (comutatividade)
- d. $T(x_1, y) \leq T(x_2, y), \forall x_1, x_2, y \in [0, 1], x_1 \leq x_2$; (monotonicidade)

Alguns operadores de norma-T são:

- min : $\mu_{A \cap B}(x) = \mu_A(x) \wedge \mu_B(x) = \min(\mu_A(x), \mu_B(x))$
- produto algébrico: $\mu_{A \cap B}(x) = \mu_A(x) \wedge \mu_B(x) = \mu_A(x)\mu_B(x)$
- produto limitado: $\mu_{A \cap B}(x) = \mu_A(x) \wedge \mu_B(x) = \max(0, \mu_A(x) + \mu_B(x) - 1)$

União: dados os conjuntos fuzzy A e B , definidos em U com funções de pertinência μ_A e μ_B , então sua união é um conjunto fuzzy dado por:

$$A \cup B = \{(x, \mu_{A \cup B}(x)) \mid x \in U\} \text{ e } \mu_{A \cup B}(x) = \mu_A(x) \vee \mu_B(x) \quad (\text{B.5})$$

onde " $\vee$ " é o símbolo que denota a união fuzzy ou então "ou" fuzzy. (disjunção)

Os operadores de união pertencem a uma classe conhecida como Norma-S ou Co-norma Triangular, os quais devem satisfazer as seguintes propriedades:

- a. $S(0, 1) = S(1, 0) = S(1, 1) = 1, S(0, 0) = 0$; (cond. de contorno)
- b. $S(x, 0) = x, \forall x \in [0, 1]$; (identidade)
- c. $S(x, y) = S(y, x), \forall x, y \in [0, 1]$; (comutatividade)
- d. $S(x_1, y) \leq S(x_2, y), \forall x_1, x_2, y \in [0, 1], x_1 \leq x_2$; (monotonicidade)

Alguns operadores tem sido propostos como operadores de co-norma-T, sendo os mais usados deles:

- max : $\mu_{A \cup B}(x) = \mu_A(x) \vee \mu_B(x) = \max(\mu_A(x), \mu_B(x))$
- soma algébrica: $\mu_{A \cup B}(x) = \mu_A(x) \vee \mu_B(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)$
- soma limitada: $\mu_{A \cup B}(x) = \mu_A(x) \vee \mu_B(x) = \min(1, \mu_A(x) + \mu_B(x))$

A Fig. B.3 mostra uma representação gráfica para as três operações básicas apresentadas nessa seção.

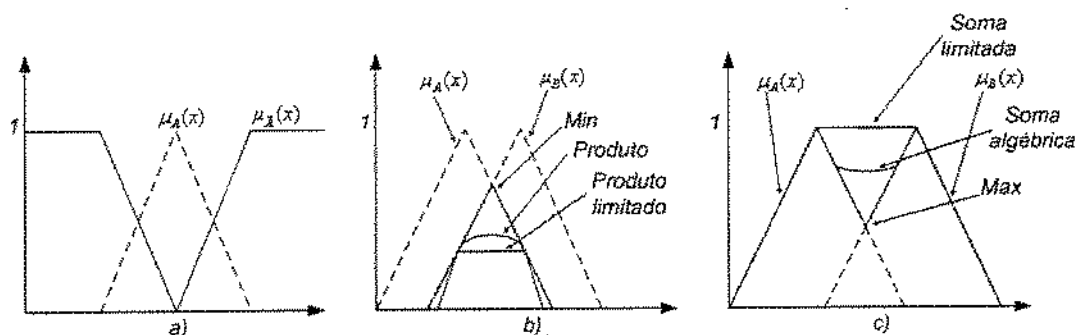


Fig. B.3: Operações com conjuntos *fuzzy*: a) Complemento b) Intersecção c) União

B.2 Raciocínio Aproximado

Raciocínio aproximado ou raciocínio *fuzzy* é um modo de raciocínio que se apóia em toda a teoria de conjuntos e lógica *fuzzy* e oferece um recurso para representar o raciocínio humano (Zadeh, 1965). O raciocínio aproximado é o fundamento para o sistema de inferência *fuzzy*. Para entender como se dá este tipo de raciocínio, três conceitos básicos devem se apresentados:

- Variáveis lingüísticas;
- Proposição *Fuzzy*;
- Regras Lingüísticas;

A seguir serão apresentados de forma resumida os conceitos citados anteriormente. Uma apresentação matemática mais rigorosa destes conceitos pode ser encontrada em (Godjevac, 1997) e (Yager e Filev, 1994).

B.2.1 Variáveis Lingüísticas

Uma variável lingüística é uma variável em que seus valores podem ser palavras ou sentenças em uma linguagem natural ou artificial como também pode ser numérico. Uma variável lingüística é caracterizada pelos seguintes termos:

- Seu nome: " x ";

- O conjunto de termos: " $T(x)$ ", que é o conjunto dos valores lingüísticos de " x ". Em outras palavras, as funções de pertinência para os valores lingüísticos de " x " estando definidos no domínio de " u ";

- O universo de discurso " U " associado como a variável base " u ".

A variável " x " não deve ser confundida com " u "; " x " é o nome da variável lingüística (por exemplo temperatura, ângulo, etc), enquanto " u " é a base da variável lingüística (por exemplo °C, radianos, etc). Em aplicações físicas, " x " deve adotar valores lingüísticos, isto é, pequeno, médio, grande, etc, enquanto " u " deve adotar valores numéricos, isto é, 100°C, 2.15 rad, etc. Por exemplo, se velocidade é definida como a variável lingüística, então o conjunto de termos $T(\text{velocidade})$ pode ser definido como:

$$T(\text{velocidade}) = \{\text{muito devagar, devagar, moderado, rápido, muito rápido}\}$$

onde cada termo em $T(\text{velocidade})$ é caracterizado por um conjunto *fuzzy* no universo de discurso " U ", conforme ilustrado na figura B.4.

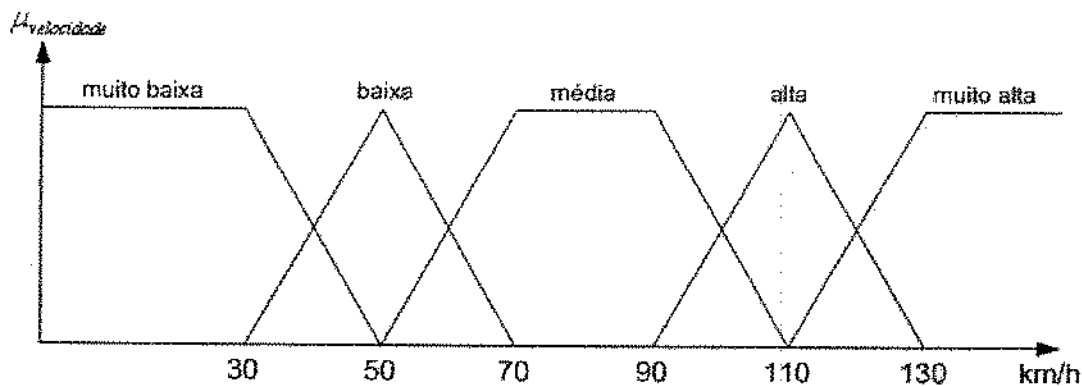


Fig. B.4: Conjunto $T(\text{velocidade})$ com suas respectivas funções de pertinência

B.2.2 Proposição Fuzzy

Uma proposição fuzzy é uma declaração expressa através de linguagem natural ou artificial. Ao contrário das proposições em lógica clássica, uma proposição *fuzzy* deve assumir valores verdade no intervalo $[0, 1]$. Para o exemplo apresentado na figura B.4, algumas das seguintes proposições são: a velocidade é baixa, a velocidade é muito alta, etc, onde o significado dessas proposições são determinados pela correspondente função de pertinência definida na figura. As proposições podem ainda envolver mais de uma variável, como, por exemplo, velocidade e aceleração relacionados através de

conectivos lingüísticos como “e”, “ou”, etc, e.g.: velocidade é baixa e aceleração alta, velocidade média e aceleração baixa, etc.

B.2.3 Regras Lingüísticas

As regras lingüísticas utilizadas em modelos *fuzzy* são do tipo *se ... então*, sendo formadas da seguinte forma:

$$\text{Se } \underbrace{\langle \text{proposição fuzzy} \rangle}_{\text{antecedente}} \text{ então } \underbrace{\langle \text{proposição fuzzy} \rangle}_{\text{consequente}} \quad (\text{B.6})$$

sendo as proposições do antecedente e consequente simples ou compostas, como discutido anteriormente. Em aplicações que envolvem tomadas de decisão, normalmente o antecedente (ou premissa) está relacionado à(s) informação(ões) de entrada e o consequente à saída. Tomando por exemplo o sistema de controle de velocidade de um veículo, uma possível regra seria:

Se velocidade *baixa* então aceleração *alta*

Esta regra relaciona a variável de entrada velocidade, com valor baixa, a uma decisão a ser tomada, a qual indica que a aceleração deve ser alta para a velocidade ser aumentada. Contudo, normalmente mais de uma regra poderá ser ativada, havendo assim mais de uma saída. Dessa forma se torna necessário usar algum processo de agregação, ponderando os diferentes valores nos consequentes das regras, utilizando alguma norma-S como operação a fim de obter uma saída global para o sistema *fuzzy*. Para casos de entradas multivariáveis, ou seja, regras com mais de uma premissa, algum operador de norma-T deve ser utilizado para realizar a composição, onde o operador mais utilizado é o operador de *mínimo*. Estes dois passos corresponde ao processo de inferência dos conjunto *fuzzy*.

Muitas vezes, dependendo da estrutura utilizada para o modelo *fuzzy*, deve haver um processo de *defuzzificação* do valor inferido, ou seja, uma “tradução” do valor lingüístico da inferência para valores numéricos que serão fornecidos ao sistema real. Os métodos de *defuzzificação*. Já a operação que acontece nos antecedentes das regras em que os valores de entrada são convertidos em grau de ativação das funções de pertinência é conhecida como *fuzzificação*.

O conjunto de regras utilizado no modelo *fuzzy* pode ser obtido através de informações obtidas através do mapeamento do conhecimento de especialistas, modelos que possam descrever o processo sob estudo ou então de forma automática a partir da obtenção das regras diretamente dos dados de entrada e saída do processo (Wang e Mendel, 1992; Zeng e Singh, 1994; Zeng e Singh, 1995). Um maneira automática de se obter tais regras é através da utilização de técnicas de agrupamento de dados (Sugeno e Yasukawa, 1993) como será descrito mais detalhadamente na seção 3.5.

Os modelos *fuzzy* podem se diferenciar em alguns critérios como o tipo de regras, tipo de funções

nos conseqüentes ou até mesmo no processo de inferência das regras, entre outros (Lee, 1990). O capítulo 3 traz uma descrição das estruturas mais utilizadas na literatura para se construir um modelo *fuzzy*, sendo elas as estruturas de Mandani e Takagi-Sugeno.