**INSTITUTO MILITAR DE ENGENHARIA**

**BIANCA BOREM FERREIRA**

# APLICAÇÃO DE FERRAMENTAS DE LÓGICA NEBULOSA À PREDIÇÃO DE SÉRIES TEMPORAIS

Dissertação de Mestrado apresentada ao Curso de Mestrado em Engenharia Mecânica do Instituto Militar de Engenharia, como requisito parcial para a obtenção do título de Mestre em Ciências em Engenharia Mecânica.

Orientador:
Prof. Jorge Audrin Morgado de Gois, Dr. – Ing.

**Rio de Janeiro**
**2008**

INSTITUTO MILITAR DE ENGENHARIA
Praça General Tibúrcio, 80 – Praia Vermelha
Rio de Janeiro – RJ CEP: 22290-270



F383a Ferreira, B. B.
Aplicação de Ferramentas de Lógica Nebulosa à Predição de Séries Temporais / Bianca Borem Ferreira. - Rio de Janeiro: Instituto Militar de Engenharia, 2008.
124 p.: il.

Dissertação (mestrado) - Instituto Militar de Engenharia, - Rio de Janeiro, 2008.

1. Sistemas Não-Lineares. 2. Séries Temporais. 3. Lógica Nebulosa. 4. Algoritmo Genético. 5. Agrupamento de Dados. I. Título II. Instituto Militar de Engenharia

CDD 003.75

**INSTITUTO MILITAR DE ENGENHARIA**

**BIANCA BOREM FERREIRA**

# APLICAÇÃO DE FERRAMENTAS DE LÓGICA NEBULOSA À PREDIÇÃO DE SÉRIES TEMPORAIS

Dissertação de Mestrado apresentada ao Curso de Mestrado em Engenharia Mecânica do Instituto Militar de Engenharia como requisito parcial para a obtenção do título de Mestre em Ciências em Engenharia Mecânica.

Orientador: Prof. Jorge Audrin Morgado de Gois, Dr. – Ing.

Aprovada em 26 de Setembro de 2008 pela seguinte banca examinadora:

_____

Prof. Jorge Audrin Morgado de Gois, Dr. – Ing. – Presidente

_____

Prof. Marcelo Amorim Savi, Ph. D. da UFRJ

_____

Prof. Fernando Ribeiro da Silva, D. Sc. do IME

Rio de Janeiro

2008

À minha mãe, Deloina de Almeida Borem, alicerce da minha vida e fonte de amor incondicional.

## AGRADECIMENTOS

Agradeço primeiramente a Deus por ter me dado força e sabedoria para subir mais esse degrau na grande escada do conhecimento.

Agradeço a CAPES pela ajuda financeira através da concessão da bolsa de estudos.

Agradeço aos meus pais, Dervan da Silva Ferreira e Deloina de Almeida Borem cujo apoio e carinho foram imprescindíveis no decorrer de toda minha existência.

Agradeço ao professor Jorge Audrin Morgado de Gois pela paciência, pela amizade, pela orientação em todos os passos desse trabalho e por ter confiado que eu poderia realizar o mesmo.

Agradeço à Tálita S. P. Sono e ao Fernando Pereira por terem compartilhado seus conhecimentos de programação e pela paciência e carinho nas horas mais desesperadoras.

Agradeço aos amigos que fiz no IME, principalmente ao Maurício O. Brito, Ten. Arantes, Mônica M. Gomes, Wellington Bitencourt, e Leandro Aguiar, pela compreensão e apoio nas horas difíceis, pela ajuda nas dificuldades e pelo companheirismo nos momentos de descontração.

Agradeço a todos os professores e funcionários do IME os quais possibilitaram que eu me tornasse Mestre.

Agradeço a todos que de alguma forma me ajudaram para a conclusão deste trabalho.

“O caos freqüentemente alimenta a vida, enquanto a ordem alimenta o hábito.”

HENRY ADAMS

**SUMÁRIO**

LISTA DE ILUSTRAÇÕES.....10
LISTA DE TABELAS.....14
LISTA DE ABREVIATURAS E SÍMBOLOS.....15

**1 INTRODUÇÃO.....19**
1.1 Revisão bibliográfica.....20
1.1.1 Aprendizagem genética de sistemas nebulosos.....20
1.1.2 Identificação, predição e estimação dos parâmetros dos sistemas não-lineares.....21
1.1.3 Reconstrução do espaço de estados de séries temporais.....25
1.1.4 Experimentos e aplicações.....27
1.2 Objetivo do trabalho.....28
1.3 Organização da dissertação.....31

**2 SÉRIES TEMPORAIS.....33**
2.1 O que são séries temporais.....35
2.2 Dimensão de imersão.....38
2.3 Passo de reconstrução.....39
2.3.1 Método da informação mútua.....41
2.3.2 Método da função de auto-correlação.....43
2.4 Reconstrução do espaço de estado.....43
2.5 Predição.....46
2.5.1 Predição por média móvel.....47
2.5.2 Predição por auto-regressão.....48
2.5.3 Predição por média móvel auto-regressiva.....48

**3 LÓGICA NEBULOSA.....50**
3.1 Introdução a conjuntos nebulosos.....52
3.2 Funções de pertinência.....53
3.2.1 Função triangular.....54
3.2.2 Função gaussiana.....55
3.2.3 Função trapezoidal.....55
3.2.4 Função sigmoidal.....56
3.2.5 Função sino.....57
3.2.6 Conjunto unitário (*singleton*).....58
3.3 Definições básicas em conjuntos nebulosos.....58
3.3.1 Corte $\alpha$ .....59
3.3.2 Suporte.....59
3.3.3 Núcleo.....59
3.3.4 Altura.....59
3.3.5 Normalização.....60
3.4 Operações básicas com conjuntos nebulosos.....60
3.4.1 União.....60
3.4.2 Interseção.....61
3.4.3 Complemento.....61
3.5 Variável lingüística.....61
3.6 Regras nebulosas.....63
3.7 Raciocínio aproximativo.....64
3.8 Regra composicional de inferência.....65
3.9 Sistemas nebulosos.....66
3.9.1 Modelos de sistemas nebulosos.....68
3.9.2 Sistemas nebulosos Takagi-Sugeno-Kang.....69

**4 AGRUPAMENTO DE DADOS.....71**
4.1 Agrupamento de dados nebuloso.....74
4.2 Agrupamento de dados subtrativo.....79

**5 ALGORITMO GENÉTICO.....82**
5.1 Métodos de seleção.....85
5.1.1 Seleção pela roleta.....86
5.1.2 Seleção por amostragem universal estocástica.....86
5.1.3 Seleção por torneio.....87
5.1.4 Seleção por truncamento.....87
5.1.5 Seleção elitista ou elitismo.....88
5.1.6 Outros métodos.....88
5.2 Operadores genéticos.....89
5.2.1 Cruzamento.....89
5.2.1.1 Cruzamento mono-ponto.....90
5.2.1.2 Cruzamento de n-pontos.....90
5.2.1.3 Cruzamento uniforme.....91
5.2.2 Mutação.....92
5.2.2.1 Mutação simples.....93
5.3 Parâmetros genéticos.....93

**6 SIMULAÇÕES E RESULTADOS.....95**
6.1 Série temporal de Mackey-Glass.....98
6.2 Sistema de Lorenz.....103
6.3 Sistema de Rössler.....107
6.4 Sistema de Duffing.....110
6.3 Sistema de Van der Pol.....114

**7 CONCLUSÕES E COMENTÁRIOS.....119**

**8 REFERÊNCIAS BILBIOGRÁFICAS.....121**

## LISTA DE ILUSTRAÇÕES


FIG 1.1 Algoritmo para definição da estrutura do sistema nebuloso.....30

FIG 2.1 Representação gráfica de uma série temporal. No eixo horizontal $x$ estão representados os 30 dias do mês de junho e no eixo vertical $y$ estão às unidades vendidas do produto $A$ .....38

FIG 3.1 Função triangular.....54

FIG 3.2 Função gaussiana.....55

FIG 3.3 Função trapezoidal.....56

FIG 3.4 Função sigmoidal.....57

FIG 3.5 Função sino.....57

FIG 3.6 Conjunto unitário.....58

FIG 3.7 Estrutura básica de um sistema nebuloso.....67

FIG 4.1 Algoritmo do agrupamento de dados nebuloso *c-means*.....75

FIG 4.2 Algoritmo do agrupamento de dados subtrativo.....80

FIG 5.1 Estrutura básica do algoritmo genético.....84

FIG 5.2 Exemplo de uma roleta de seleção.....86

FIG 5.3 Seleção por amostragem universal estocástica.....87

FIG 5.4 Cruzamento mono-ponto.....90

FIG 5.5 Cruzamento de três-pontos.....91

FIG 5.6 Cruzamento uniforme.....93

FIG 5.7 Mutação simples.....92

FIG 6.1 Exemplo de uma subcadeia que compõe o cromossomo.....98

FIG 6.2 Exemplo da formação de uma regra do sistema de inferência nebuloso.....98

FIG 6.3 Atrator de Mackey-Glass.....99

FIG 6.4 (a) Resultado obtido através do método da informação mútua aplicado aos dados da série temporal de Mackey-Glass e (b) Localização do primeiro mínimo local no gráfico da informação mútua.....100

FIG 6.5 Curvas de convergência do agrupamento de dados nebuloso *c-means* para os vetores do espaço de estados reconstruído da série temporal de Mackey-Glass com $\tau = 5$ e $m = 2 - 10$.....101

FIG 6.6 Comparação entre a saída da série temporal de Mackey-Glass real (verde) e a estimada através sistema de inferência nebuloso proposto (azul).....102

FIG 6.7 Atrator de Lorenz.....103

FIG 6.8 (a) Resultado obtido através do método da informação mútua aplicado aos dados da saída $x$ do atrator de Lorenz e (b) Localização do primeiro mínimo local no gráfico da informação mútua.....104

FIG 6.9 Curvas de convergência do agrupamento de dados nebuloso *c-means* para os vetores do espaço de estados reconstruído da saída $x$ do atrator de Lorenz com $\tau = 12$ e $m = 2-10$ .....105

FIG 6.10 Comparação entre a saída $x$ do atrator Lorenz de real (verde) e a estimada através sistema de inferência nebuloso proposto (azul).....106

FIG 6.11 Atrator de Rössler.....107

FIG 6.12 (a) Resultado obtido através do método da informação mútua aplicado aos dados da saída $x$ do atrator de Rössler e (b) Localização do primeiro mínimo local no gráfico da informação mútua.....108

FIG 6.13 Curvas de convergência do agrupamento de dados nebuloso *c-means* para os vetores do espaço de estados reconstruído da saída $x$ do atrator de Rössler com $\tau = 17$ e $m = 2-10$ .....109

FIG 6.14 Comparação entre a saída $x$ do atrator Rössler real (verde) e a estimada através sistema de inferência nebuloso proposto (azul).....110

FIG 6.15 Trajetória obtida através do sistema de Duffing.....111

FIG 6.16 (a) Resultado obtido através do método da informação mútua aplicado aos dados do sistema de Duffing e (b) Localização do primeiro mínimo local no gráfico da informação mútua.....112

FIG 6.17 Curvas de convergência do agrupamento de dados nebuloso *c-means* para os vetores do espaço de estados reconstruído da do sistema de Duffing com $\tau = 11$ e $m = 2-10$ .....113

FIG 6.18 Comparação entre a saída do sistema de Duffing real (verde) e a estimada através sistema de inferência nebuloso proposto (azul).....114

FIG 6.19 Trajetória obtida através do sistema de Van der Pol.....115

FIG 6.20 (a) Resultado obtido através do método da informação mútua aplicado aos dados do sistema de Van der Pol e (b) Localização do primeiro mínimo local no gráfico da informação mútua.....116

FIG 6.21 Curvas de convergência do agrupamento de dados nebuloso *c-means* para os vetores do espaço de estados reconstruído do sistema de Van der Pol com $\tau = 23$ e $m = 2 - 10$.....117

FIG 6.22 Comparação entre do sistema de Van der Pol real (verde) e a estimada através sistema de inferência nebuloso proposto (azul).....118

## LISTA DE TABELAS

TAB 3.1 Características, vantagens e desvantagens da lógica nebulosa.....51

TAB 6.1 Resultados do método do agrupamento de dados nebuloso *c-means* proposto para a série temporal de Mackey-Glass.....102

TAB 6.2 Resultados do método do agrupamento de dados nebuloso *c-means* proposto para a saída $x$ do atrator de Lorenz.....106

TAB 6.3 Resultados do método do agrupamento de dados nebuloso *c-means* proposto para a saída $x$ do atrator de Rössler.....109

TAB 6.4 Resultados do método do agrupamento de dados nebuloso *c-means* proposto para o sistema de Duffing.....113

TAB 6.5 Resultados do método do agrupamento de dados nebuloso *c-means* proposto para o sistema de Van der Pol.....117

## LISTA DE ABREVIATURAS E SÍMBOLOS

### ABREVIATURAS

ANFIS - Analytical Neuro Fuzzy Inference System
AR - Auto-Regressive
ARMA - Auto Regressive Moving Average
FCM - Fuzzy C-Means
HCM - Hard C-Means
MA - Moving Average
mGA - Modified Genetic Algorithm
PID - Proporcional Integral Derivativo
RMSE - Root Mean Squared Error
SVD - Singular Value Decomposition
TSK - Takagi Sugeno Kang

### SÍMBOLOS

$m$ - dimensão de imersão
$m_0$ - dimensão de imersão mínima
$D_0$ - dimensão real do atrator
$\tau$ - passo de reconstrução
$y_j$ - ponto da série temporal
$y(i)$ - série temporal
$\varepsilon(k)$ - componente aleatória ou ruído
$P_{AB}$ - densidade de probabilidade da medida de $A$ e de $B$ em resultar nos valores $a$ e $b$
$P_A(a)$ - densidade de probabilidade individual medida em $A$

$r_k$ - função de auto-correlação

$\Gamma_j^i$ - função de pertinência triangular

$f_\beta$ - função objetivo

$\mu_A(x)$ - grau de pertinência de $x$ em $A$

$I(a_i, b_j)$ - informação mútua entre uma medida $a_i$ e a medida $b_j$

$Z$ - matriz que contém as coordenadas dos centros das classes

$A_{ij}$ - matriz que contém o grau de pertinência do vetor do espaço de estados $i$ na classe $j$

$Nmáx$ - número máximo de iterações

$Na$ - número máximo de pontos do espaço de estados

$P_i$ - potencial ou medida de possibilidade que um ponto tem de ser centro de uma classe ou grupo

$c$ - quantidade de centros

$r_a$ - raio do grupo ou raio da vizinhança

$R^i$ - regra de inferência nebulosa

$R(x, y)$ - relação nebulosa

$\hat{y}_k$ - saída estimada pelo sistema de inferência nebuloso

$yk$ - saída real do sistema estudado

$\varepsilon$ - tolerância

$U_{x_j}$ - universo de discussão

$w^i$ - valor de ativação da regra de inferência nebulosa

$y^i$ - variável de saída da regra de inferência nebulosa

## RESUMO

Modelar uma série temporal e prever um dado futuro é útil, pois torna possível a tomada de decisões e ações antecipadamente. Paralelamente, a identificação dos parâmetros de um modelo nebuloso para sistemas não-lineares é um problema complexo, sendo comumente resolvido por tentativa e erro. O foco deste trabalho é o estudo de fundamentos teóricos para análise, desenvolvimento e implementação de ferramentas utilizadas na modelagem de dados de sistemas dinâmicos complexos e séries temporais com a finalidade de reprodução de suas saídas e predição. Nesse sentido foi desenvolvido um modelo nebuloso Takagi-Sugeno, onde sua estrutura mínima e parâmetros ótimos foram obtidos via métodos de agrupamento de dados e algoritmo genético, respectivamente. O algoritmo proposto foi testado e apresentou bons resultados em cinco casos distintos: série temporal de Mackey-Glass, sistema de Lorenz, sistema de Rössler, sistema de Duffing e sistema de Van der Pol.

## ABSTRACT

Time series modeling and prediction has many different applications, because it enables decision making. Besides, the parametric identification of a fuzzy model of a non-linear system is very complex, therefore usually solved by trial. The focus of this work is to study the theoretical base behind the analysis, development and implementation of modeling tools used on reproduction or prediction of time series or data generated by complex dynamical systems. In this sense, it was implemented a Takagi-Sugeno fuzzy model, where its minimal structure and optimal set of parameters were obtained through clustering and genetic algorithm, respectively. The proposed algorithm was tested in five different systems: the Mackey-Glass, Lorenz, Rössler, Duffing and Van der Pol systems, good results were obtained in all the cases.

# 1 INTRODUÇÃO

Na década de 90, Zadeh introduziu o conceito de computação flexível, a qual representa uma combinação de técnicas de inteligência computacional, como por exemplo, lógica nebulosa, algoritmo genético e agrupamento de dados (entre outras técnicas que formam a computação flexível, mas aqui serão consideradas somente essas três como principais componentes). A utilização cooperativa destas técnicas oferece formas de raciocínio e busca para a solução de problemas complexos do mundo real que apresentam situações indeterminadas (Pires, 2004). Um aspecto essencial da computação flexível é o fato de que as metodologias que a constituem serem complementares e simbióticas, ao invés de competitivas e exclusivas.

A lógica nebulosa introduzida por Zadeh nos dá uma linguagem com sintaxe e semântica capaz de traduzir o conhecimento do domínio do problema em sentenças lingüísticas de fácil compreensão para o ser humano, sendo que a maior característica da lógica nebulosa é a robustez do seu mecanismo de inferência no tratamento das informações representadas por estas sentenças. Os algoritmos genéticos propostos por John Holland (Pires, 2004) são algoritmos de otimização e busca baseados nos mecanismos de seleção natural e genética, capazes de executar uma procura global em um espaço de solução complexo e irregular. O agrupamento de dados nebuloso *c-means* introduzido por Jim Bezdek (Cardoso, 2003) é uma técnica na qual há o particionamento de um conjunto de dados em subconjuntos (*clusters*) de modo que cada ponto tenha um grau de pertinência aos *clusters*, sendo que sua maior vantagem é a minimização das variações *inter-cluster*.

Considerando as características apresentadas de cada técnica e utilizando-as de uma forma onde as vantagens de uma se sobrepõem às desvantagens de outra,

é possível construir sistemas híbridos cada vez mais robustos para resolução de problemas por demais complexos.

## 1.1 REVISÃO BIBLIOGRÁFICA

### 1.1.1 APRENDIZAGEM GENÉTICA DE SISTEMAS NEBULOSOS

A combinação de sistemas nebulosos com algoritmos genéticos, conhecida também como sistemas genéticos nebulosos, tem grande aceitação na comunidade científica, uma vez que estes sistemas são robustos e capazes de encontrar soluções em espaços complexos e irregulares (Pires, 2004). Além de ter um bom desempenho em termos de acuidade e interpretabilidade, essa abordagem aumenta a autonomia do projeto ao minimizar a intervenção do usuário.

O principal foco do trabalho de Pires é a investigação das abordagens de modelagem automática de sistemas nebulosos aplicados a problemas de classificação de padrões, através de algoritmo genético para a definição e sintonia dos conjuntos nebulosos que compõem as partições nebulosas dos domínios envolvidos. O aprendizado genético é empregado somente na base de dados do sistema nebuloso, isto é, as funções de pertinência (Pires, 2004).

Para solucionar o problema de projeto automático de sistemas nebulosos, Delgado propôs uma abordagem co-evolutiva (Delgado, 2002). A co-evolução permitiu que relações de hierarquia e cooperação fossem estabelecidas entre indivíduos representando diferentes parâmetros dos sistemas nebulosos. A abordagem proposta recorreu a diferentes espécies que codificaram soluções parciais do problema de projeto automático de sistemas nebulosos e estavam organizadas em quatro níveis hierárquicos. Cada nível hierárquico codificou as funções de pertinência, as regras individuais, as bases de regras e os sistemas nebulosos, respectivamente. Restrições e objetivos locais foram observados em todos os níveis, de modo a garantir a ocorrência de indivíduos caracterizados pela

simplicidade das regras nebulosas, compactação e consistência da base de regras e visibilidade na partição do universo.

Devido ao uso de múltiplas populações, com informações significativamente diferentes, o ajuste dos parâmetros evolutivos do sistema proposto por Delgado se torna um problema complexo. Sendo assim, Maruo propôs uma abordagem auto-adaptativa, na qual o próprio algoritmo genético se encarrega de selecionar um bom conjunto de parâmetros, liberando o usuário do processo de definição manual dos parâmetros evolutivos mantendo o bom desempenho do sistema nebuloso (Maruo, 2006). A utilização do algoritmo evolutivo, não apenas para encontrar a solução do problema, mas também para ajustar uma série de parâmetros do próprio algoritmo, se constitui em uma das principais contribuições da pesquisa de Maruo. O desempenho do mecanismo de auto-adaptação de parâmetros evolutivos é avaliado em duas fases: inicialmente, a auto-adaptação é testada, utilizando problemas de otimização contínua e combinatória; depois, a auto-adaptação é aplicada para resolver o problema do projeto automático de sistemas baseados em regras nebulosas, e para isto, o sistema genético-nebuloso resultante é usado na aproximação de funções.

### 1.1.2 IDENTIFICAÇÃO, PREDIÇÃO E ESTIMAÇÃO DOS PARÂMETROS DE SISTEMAS NÃO-LINEARES

A pesquisa na área de sistemas dinâmicos não-lineares tem despertado a atenção crescente de diversas áreas, tais como: engenharia, física, matemática e biomedicina. Várias dificuldades encontradas no desenvolvimento desse assunto têm mostrado uma necessidade real em se usar alguns tipos de aproximações inteligentes (Lee et al., 2006). Essas dificuldades aparecem porque ao lidar com informações, como no caso de modelos matemáticos ou de qualquer outra natureza para representação de fenômenos ou sistemas físicos, a incerteza e a imprecisão estão ligados entre si.

Várias metodologias, baseadas, na maioria das vezes, em lógica nebulosa e algoritmo genético, têm sido aplicadas nas áreas de controle, identificação, predição e estimação dos parâmetros de sistemas, no sentido de suplantar essas dificuldades.

Demonstrando a capacidade superior de predição das aproximações baseadas nas redes neurais nebulosas quando comparadas às que utilizam redes neurais convencionais, Jang aplicou 16 regras de aprendizagem híbridas SE-ENTÃO (mesmas regras de aprendizagem utilizadas nas redes neurais artificiais) à arquitetura ANFIS (Analytical Neuro-Fuzzy Inference System - Sistema de Inferência Nebuloso baseado em Redes Neurais) empregadas na predição de series temporais, comparando seus resultados com aproximações obtidas anteriormente, tais como: regressões lineares e redes neurais convencionais (Jang et al., 1993). Manguire também trabalhou com uma arquitetura *neuro-fuzzy* alternativa aplicado-a a predição de séries temporais caóticas (Manguire et al., 1998). A arquitetura apresentada por ele propõe uma aproximação para o sistema de inferência nebuloso a fim de reduzir consideravelmente as dimensões da rede se comparada às aproximações similares.

Pisarenko propõe uma discussão sobre a possibilidade de aplicar alguns métodos estatísticos padrão (método dos mínimos quadrados, método da máxima verossimilhança e o método de momentos estatísticos para estimar parâmetros) a um sistema dinâmico com baixa dimensionalidade e deterministicamente caótico (o mapa logístico), contendo um ruído observacional (Pisarenko et al., 2004).

Utilizando técnicas baseadas na programação genética, Zhang propõe uma forma para modelar séries temporais caóticas (Zhang et al., 2004). Primeiramente, utilizou um algoritmo com técnicas baseadas na programação genética para encontrar estruturas do modelo apropriadas localizadas no espaço da função. Depois, introduziu um algoritmo de otimização de partículas *Swarm* para estimar os parâmetros não-lineares das estruturas do modelo dinâmico. Finalmente, os resultados da análise da série temporal são integrados ao algoritmo baseado na

programação genética para melhorar a qualidade da modelagem e os critérios de estabilidade do modelo.

Goldberg e Deb propuseram um algoritmo genético modificado (Modified Genetic Algorithm - mGA) que assegurava uma convergência para um ótimo global (Goldberg e Deb, 1989). O mGA, primeiramente, descobre e enfatiza os blocos de construção bons para soluções ótimas ou quase ótimas, que é chamada de fase de seleção primordial. Em seguida, os operadores de corte, encaixe e a fase de seleção justaposicional recombinam os blocos de construção bons formando pontos ótimos com probabilidades altas.

Lee, a partir da pesquisa desenvolvida por Goldberg e Deb e de seu trabalho de análise e desenvolvimento do projeto de um controlador nebuloso robusto para sistemas Takagi-Sugeno-Kang aplicados a sistemas não-lineares com parâmetros incertos, buscando desenvolver um modelo nebuloso bem sucedido para identificação e predição de sistemas não-lineares (Lee et al., 2001), propõe um método para identificação automática da estrutura e dos parâmetros de um sistema de inferência nebuloso Takagi-Sugeno-Kang de ordem zero com um algoritmo híbrido utilizando o mGA e, posteriormente, utiliza o método do gradiente para fazer o ajuste fino dos parâmetros obtidos (Lee et al., 2006).

Chang desenvolveu um algoritmo genético com codificação real e multi cruzamentos para estimar parâmetros de uma classe de sistemas não-lineares, mesmo que esses parâmetros tenham termos de atraso no tempo ou apresentem não-linearidades (Chang, 2006). Dando continuidade à pesquisa (Chang, 2007) foi proposto um algoritmo genético, desta vez, com codificação real, aplicando-o no controle e identificação de sistemas não-lineares. Primeiramente, Chang utilizou um algoritmo genético de codificação real para identificar sistemas desconhecidos nos quais as estruturas são supostamente conhecidas. Em seguida, aplicou o modelo estimado de um controlador PID offline, resolvendo otimamente o problema utilizando algoritmo genético com codificação real.

Visando apresentar um método de modelagem baseado nos conjuntos nebulosos aplicado na predição de sistemas complexos e com características não lineares, Pucciarelli utilizou o modelo nebuloso Takagi-Sugeno-Kang aplicado na modelagem de séries temporais onde os conjuntos nebulosos do antecedente e os parâmetros do conseqüente são estimados via métodos de agrupamentos e identificação paramétrica, respectivamente (Pucciarelli, 2005).

Wang e outros pesquisadores (Wang et al., 2005) se basearam na aprendizagem competitiva nebulosa, para confirmar o espaço nebuloso das variáveis de entrada, e nos mínimos quadrados recursivo, empregando o método de decomposição do valor singular (SVD), para confirmar os parâmetros conseqüentes do modelo nebuloso. Alguns anos depois, Wang e Gu continuam a pesquisa feita anteriormente, utilizando, para esse trabalho, um método *fuzzy clustering* para confirmar o espaço de entrada do modelo nebuloso (Gu e Wang, 2007).

Para construir um modelo nebuloso ótimo ajustado para sistemas não-lineares, Eftekhari e Katebi apresentaram um procedimento constituído de dois estágios, utilizando algoritmo genético (Eftekhari e Katebi, 2007). O primeiro estágio consistiu em uma otimização estrutural, atribuindo uma aptidão apropriada para cada membro individual da população. No segundo estágio, foi aplicada filtragem para otimizar as funções de pertinência de entrada e saída. A aproximação híbrida proposta explora vantagens e utiliza características desejáveis para a extração de modelos nebulosos exatos e compactos.

Revendo algumas modelagens de preditores baseados em funções de base radial e nos mínimos quadrados, Lau e Wu propuseram um preditor local baseado na regressão do vetor suporte para melhorar a predição do espaço de fases de séries temporais caóticas combinando a resistência da regressão do vetor suporte e as propriedades de reconstrução da dinâmica caótica (Lau e Wu, 2008).

### 1.1.3 RECONSTRUÇÃO DO ESPAÇO DE ESTADOS A PARTIR DE SÉRIES TEMPORAIS

Considerando a reconstrução do espaço estado, a análise no domínio da freqüência, a determinação de invariantes dinâmicas, dos expoentes de Lyapunov e da dimensão do atrator, em (Franca e Savi, 2001 **(A)**) é discutida a análise experimental de um pêndulo não-linear e apresentado um procedimento para reconstruir o mapa de Poincaré do sinal. Também são feitas análises dos movimentos periódicos e caóticos a fim estabelecer a diferença entre eles. No mesmo ano, (Franca e Savi, 2001 **(B)**) aplicam a investigação feita sobre a dimensão de atratores caóticos ao pêndulo não-linear, onde os sinais são gerados a partir da integração numérica do modelo matemático, selecionando uma variável do sistema como uma série temporal. A reconstrução do espaço estado e a determinação das dimensões do atrator são obtidas considerando sinais caóticos e periódicos. Os resultados foram comparados com os valores de referência obtidos através da análise do modelo matemático.

Caracterizado o comportamento de uma série temporal, o espaço de estado é reconstruído a partir da série temporal observada exigindo a seleção de sua dimensão de imersão. Em (Jiang e Adeli, 2003) são investigados o método do fator de preenchimento, método da deformação local integral média e o método dos falsos vizinhos para determinar a dimensão de imersão usando três exemplos diferentes com equações analíticas disponíveis, onde o valor exato da dimensão de imersão ótima era conhecida. Em seguida, uma aproximação por meio de agrupamento nebuloso *c-means* é proposta para encontrar a dimensão de imersão ótima. A aproximação proposta retorna a resposta exata para todos os exemplos. Além disso, não requer a seleção por tentativa e erro ou arbitrária dos parâmetros sendo computacionalmente eficiente.

Com intuito de selecionar a dimensão de imersão de um sistema dinâmico, em (Abonyi et al., 2004) também é proposto um algoritmo baseado em agrupamentos para esta finalidade, sendo essa uma etapa importante na análise e predição de

séries temporais caóticas não-lineares. O agrupamento foi aplicado no espaço reconstruído definido pelas variáveis de saída defasadas. A dimensão intrínseca do espaço reconstruído foi então estimada baseando-se na análise dos autovalores das matrizes de covariância do agrupamento nebuloso, enquanto a dimensão de imersão correta foi inferida através do desempenho de predição dos modelos locais dos grupamentos. A maior vantagem da solução proposta seria que 3 etapas foram simultaneamente resolvidas durante o agrupamento: seleção da dimensão de imersão, estimação da dimensão intrínseca e identificação do modelo que pode ser usado para a predição.

Na mesma época, em (Feil et al., 2004) é também proposto um algoritmo baseado em agrupamentos, mas para aumentar a eficiência do modelo de estimação de sistemas lineares e não-lineares baseado no algoritmo dos falsos vizinhos. O agrupamento foi aplicado no espaço gerado pelo produto cartesiano das variáveis de entrada e saída e a estrutura do modelo, da mesma forma que o trabalho descrito acima, estimado com base nos autovalores das matrizes de covariância do grupamento nebuloso. A vantagem principal da solução proposta reside no fato de seu modelo ser livre. Isto significa que nenhum modelo particular precisou ser construído para selecionar a ordem da modelagem. Isto conservou o esforço computacional e evitou uma polarização possível devido ao método particular de construção utilizado.

Uma nova estrutura de dados de séries temporais para reconstrução do espaço de fase foi proposta em (Feng e Huang, 2005) para identificar padrões temporais que são característicos e preditivos de eventos significativos nas séries temporais complexas. Essa nova estrutura utiliza conjuntos nebulosos com funções de pertinência gaussianas para definir os padrões temporais no espaço de imersão.

Além disso, foram utilizados os métodos dos falsos vizinhos e da informação mútua para estimar a defasagem no tempo e a dimensão do espaço de fase.

### 1.1.4 EXPERIMENTOS E APLICAÇÕES

No sentido de desenvolver uma estrutura efetiva e sistemática para integração de modelagem e projeto de controle digital de sistemas complexos, incluindo sistemas caóticos, Joo e outros pesquisadores (Joo et al. 1999) propuseram uma metodologia para projetar controladores baseados em modelos nebulosos de espaço de estados híbrido com uma taxa dupla de amostragem. Para isso, o projeto é desenvolvido da seguinte forma: primeiramente, foi construído um modelo de espaço de estado com tempo discreto e taxa fixa equivalente do sistema com tempo contínuo para ser utilizado no sistema de inferência nebuloso Takagi-Sugeno-Kang. Para obter um ganho ótimo de realimentação do estado de tempo continuo, é construído um sistema nebuloso de tempo discreto para ser convertido em um sistema de tempo continuo. A lei de controle ótimo de tempo continuo desenvolvida é, finalmente, convertida em uma lei de controle digital de tempo lento equivalente.

É possível verificar que um dos problemas enfrentados por muitas empresas jornalísticas é o de determinar a quantidade de jornais que devem ser impressos e distribuídos entre os numerosos pontos de vendas. A quantidade certa que deve ser reposta depende de vários fatores, especialmente da demanda de cada ponto de venda a qual, por sua vez, depende de sua localização. Atualmente, a previsão da demanda de cada ponto de venda é baseada em taxas de reposição observadas no passado e por um especialista da área. Cardoso propôs o uso de *Knowledge Discovery in Databases* como uma técnica de previsão de reposição (Cardoso, 2003). O objetivo era prever a quantidade de jornais que deveriam ser repostos diariamente em cada uma das bancas de jornal. O modelo de previsão proposto utilizou agrupamento nebuloso para a exploração dos dados e regras nebulosas para a previsão. Os resultados experimentais obtidos com uma base de dados real mostraram a eficácia do modelo, especialmente quando comparados com outras

metodologias e com os resultados proporcionados por métodos de previsão baseados em reposição e em redes neurais.

Investigando a hipótese de modelos não-lineares de previsão de séries temporais serem capazes de fornecer uma previsão fora da amostra mais acurada que modelos lineares tradicionais, Santos analisou modelos não-lineares, tais como: redes neurais *perceptron* multicamadas, redes neurais com funções de base radial e sistemas nebulosos Takagi-Sugeno-Kang, e modelos lineares, tais como: auto-regressivos, média móvel, auto-regressivos de média móvel e auto-regressivos de média móvel considerando a heteroscedasticidade condicional auto-regressiva dos resíduos, aplicados na avaliação do *root mean squared error* (RMSE), do índice de desigualdade U-Theil, do percentual de sinais corretamente previstos e da estatística de falha de previsão Pesaran-Timmermann (Santos, 2005). Além disso, foi avaliado também o retorno e o risco de uma estratégia de negociação estabelecida com base nas previsões geradas pelos modelos

## 1.2 OBJETIVO DO TRABALHO

Estudos baseados em séries temporais vêm ganhando uma grande importância no decorrer dos últimos anos. Isso está acontecendo porque vários fenômenos têm seus modelos ou seus mecanismos de funcionamento muito complexos, o que impossibilita o desenvolvimento de modelos matemáticos que possam descrevê-los, mostrando uma necessidade real em buscar aproximações inteligentes.

A identificação de sistemas dinâmicos não-lineares através de dados de séries temporais tem despertado a atenção crescente de diversas áreas. Além disso, a possibilidade de prever um dado futuro através de uma série temporal é muito útil porque torna possível a tomada de decisões e ações antecipadamente.

Nesse sentido, esse trabalho visa à reprodução e predição séries temporais não-lineares utilizando modelos baseados na lógica nebulosa e como ferramentas de

apoio o algoritmo genético e as técnicas de agrupamento de dados também conhecido como *clustering*.

O sistema de inferência nebuloso funcionará como preditor, de modo que dado um ponto da série temporal para a qual ele foi treinado, pretende-se estimar um valor da série a frente no tempo. Para tanto será necessária informação sobre a dinâmica do sistema.

A fim de se obter tal informação, a predição será feita a partir dos estados do sistema correspondente à série temporal e, portanto, faz-se necessária a reconstrução de seu espaço de estados. A saída do sistema nebuloso será a série predita, ou seja, com um avanço no tempo.

Utilizando-se o teorema de Takens para a reconstrução, onde os estados são obtidos pela introdução de retardos fixos na série (passo de reconstrução), o sistema nebuloso será treinado por um conjunto de séries defasadas, onde uma será a saída (defasada em avanço) e as outras serão as entradas correspondentes (espaço reconstruído – retardadas).

O funcionamento do sistema nebuloso proposto depende da determinação de alguns parâmetros essenciais:

- **As entradas:** As entradas do sistema de inferência nebuloso proposto são dadas pelos espaços de estado reconstruídos, determinados através do teorema de Takens. Sendo assim, fez-se necessário o cálculo do passo de reconstrução (*time delay*) e da dimensão de imersão

  - **Passo de reconstrução:** O passo de reconstrução é dado pelo primeiro mínimo local encontrado através da análise do gráfico com os resultados da informação mútua.

  - **Dimensão de imersão:** A dimensão de imersão ótima é calculada através do algoritmo de agrupamento de dado nebuloso *c-means*.

- **A quantidade de particionamentos do espaço de entrada:** A quantidade de particionamentos do espaço de entrada, ou seja, a quantidade de funções de pertinência que vão compor o antecedente de cada regra do sistema de inferência nebuloso proposto vai será igual à dimensão de imersão encontrada através do algoritmo de agrupamento de dados nebuloso *c-means*.

- **A quantidade de regras:** A quantidade de regras que vão compor o sistema de inferência nebuloso proposto é igual à quantidade de centros calculada através do algoritmo de agrupamento de dados subtrativo.

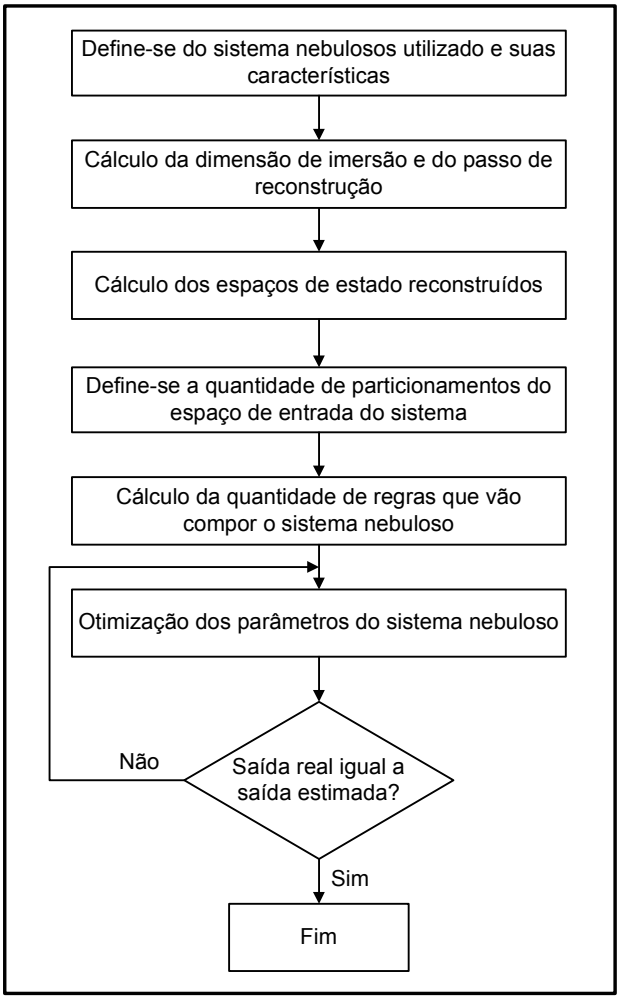


**FIG. 1.1 – Algoritmo para definição da estrutura do sistema nebuloso**

Além da determinação desses parâmetros, foi feita a otimização dos parâmetros das funções que compõem o antecedente e o conseqüente das regras de inferência

nebulosas através do algoritmo genético. O processo de otimização é finalizado quando a saída real for igual à saída predita estimada pelo sistema de inferência nebuloso proposto. A definição do sistema de inferência nebuloso proposto está ilustrado na FIG. 1.1.

Nos capítulos seguintes são apresentadas revisões sobre todos os conceitos e ferramentas usadas no desenvolvimento desse trabalho e no capítulo 6 é explicado como foram determinados os parâmetros e como é feita a otimização do sistema proposto através do ferramental estudado.

## 1.3 ORGANIZAÇÃO DA DISSERTAÇÃO

Este trabalho está dividido em sete capítulos.

O Capítulo 1 apresenta uma revisão bibliográfica focada na aprendizagem genética de sistemas nebulosos; na identificação, predição e estimação dos parâmetros dos sistemas não-lineares; na reconstrução dos espaços de estados de séries temporais; e em alguns experimentos e aplicações. Além disso, apresenta o objetivo do trabalho, contendo um escopo do sistema nebuloso proposto e do seu funcionamento.

O Capítulo 2 apresenta uma revisão dos principais conceitos de séries temporais utilizados no desenvolvimento desta dissertação. Isto inclui definições de séries temporais, dimensão de imersão, passo de reconstrução, reconstrução de espaço de estados e predição.

O Capítulo 3 apresenta os principais conceitos relacionados à teoria da lógica nebulosa abordados nesta dissertação, tais como: uma breve introdução, definições e operações básicas com conjuntos nebulosos; funções de pertinência; variável lingüística; regras nebulosas e regra composicional de inferência; e sistemas nebulosos, tendo um enfoque especial no sistema Takagi-Sugeno-Kang.

O Capítulo 4 apresenta os dois métodos de agrupamentos de dados utilizados nesta dissertação: agrupamento de dados nebuloso *c-means*, utilizado para encontrar a dimensão de imersão ótima e, conseqüentemente, a quantidade de particionamentos do espaço de entrada do antecedente de cada regra, e o agrupamento de dados subtrativo, usado para determinar a quantidade de regras que irão compor o sistema nebuloso proposto.

O Capítulo 5 apresenta alguns conceitos básicos relacionados à teoria dos algoritmos genéticos, a serem utilizados na otimização dos parâmetros do sistema nebuloso proposto, tais como: métodos de seleção, operadores e parâmetros genéticos.

O Capítulo 6 apresenta algumas simulações e resultados obtidos através do sistema nebuloso Takagi-Sugeno-Kang proposto, juntamente com a utilização das ferramentas do algoritmo genético e do agrupamento de dados.

Finalmente, o Capítulo 7 apresenta as conclusões sobre os resultados obtidos e propostas para pesquisas futuras utilizando a metodologia proposta nesta dissertação.

## 2 SÉRIES TEMPORAIS

Apesar da dinâmica ter grande aplicabilidade em vários ramos da ciência, ela foi originalmente restrita à física. O inicio foi por volta do meio do século XVII, quando Newton desenvolveu as equações diferenciais, descobriu as leis de movimento, da gravitação universal e combinou-as para explicar as leis de Kepler sobre o movimento dos planetas, resolvendo o problema da interação entre dois corpos.

Gerações subseqüentes de matemáticos e físicos tentaram expandir a teoria de Newton para o problema de três corpos, por exemplo: Sol, Terra e Lua, mas esse problema tornou-se muito mais difícil de se resolver. Depois de décadas de esforço em vão, muitos já consideravam esse problema impossível de ser resolvido no sentido de obter equações explícitas que explicassem os movimentos.

Em fins do século XIX, Poincaré introduziu um novo ponto de vista na análise do problema, ao invés de tentar determinar as posições exatas dos planetas em qualquer instante ele tentaria responder se o sistema solar seria estável para sempre ou não através de uma análise qualitativa. Poincaré desenvolveu uma ferramenta geométrica poderosa para analisar essas questões, resultando na teoria moderna de sistemas dinâmicos. Ele foi também o primeiro a perceber a possibilidade do caos, um comportamento aperiódico que depende sensivelmente das condições iniciais, tornando a previsão em longo prazo impossível.

Lorenz, em 1963, descobriu um comportamento caótico em um modelo simplificado de convecção atmosférica. Ele notou que as soluções de suas equações nunca convergiam para o equilíbrio ou para um estado periódico. Além disso, para duas condições iniciais ligeiramente diferentes, o comportamento do sistema era totalmente diferente. Lorenz representou graficamente a solução de suas equações construindo um espaço de estado, mostrando que há uma certa estrutura no caos.

A idéia de que um experimento real possa ser governado por um conjunto de equações é uma ficção. Um conjunto de equações diferenciais ou um mapa pode modelar um sistema apenas da forma suficiente para fornecer resultados úteis. Nesse contexto, torna-se importante analisar sistemas dinâmicos sem que se conheçam detalhes sobre a sua dinâmica, não possuindo, portanto, um modelo matemático estabelecido (Savi, 2006). Para esta análise são utilizadas, freqüentemente, as séries temporais obtidas de experimentos.

Dados de séries temporais surgem nos mais variados campos do conhecimento como economia (preços diários de ações, taxa mensal de desemprego, produção industrial), mercadologia (vendas semanais, gastos semanais com propaganda), medicina (eletrocardiograma, eletroencefalograma, diagnóstico e comportamento de pacientes), engenharia (monitoramento baseado em sensores), epidemiologia (número mensal de novos casos de meningite), demografia (população anual, nascimentos e mortes mensais), meteorologia (precipitação pluviométrica, temperatura diária, velocidade do vento), oceanografia (maré horária), sociologia (criminalidade mensal, greves anuais), entre outros.

Os primeiros trabalhos sobre recuperação de séries temporais consideravam “casamento exato” de séries. No entanto, atualmente existe um consenso de que a recuperação de dados de uma série temporal deve ser baseada em “similaridade”. Critérios de similaridade são usados tanto em recuperação quanto em várias tarefas de mineração de dados em séries temporais (Sanches, 2006). Em cada aplicação da técnica de mineração de dados existe um conjunto de métodos e algoritmos que são os candidatos potenciais para a extração de relações relevantes implícitas em base de dados. Entre eles incluem-se métodos e algoritmos de análise de seqüências, agrupamento de dados, classificação, estimativas, regras de associação e, mais recentemente, técnicas que utilizam a teoria de conjuntos nebulosos e algoritmos genéticos. Cada um destes candidatos pode ser utilizado nos diferentes tipos de problemas relacionados com a aplicação em mente (Cardoso, 2003).

Existe, porém, uma diferença significativa entre a mineração de dados e os outros mecanismos de análise de dados, justamente na maneira como cada um

deles explora as relações existentes entre os dados que estão sendo analisados. Os diversos mecanismos de análise dispõem de métodos baseados na verificação, isto é, o usuário constrói hipóteses sobre relações específicas e as verifica com o auxílio do próprio sistema. Esse modelo torna-se dependente da intuição e habilidade do analista em propor hipóteses interessantes, em manipular a complexidade do espaço de atributos, e em refinar a análise baseando-se em resultados de consultas à base de dados potencialmente complexas. No processo de mineração de dados, ele mesmo é responsável pela geração de hipóteses, garantindo maior rapidez, precisão e integralidade aos resultados.

Verifica-se que a mineração de dados é uma metodologia que objetiva encontrar uma descrição lógica ou matemática, eventualmente de natureza complexa, de padrões e regularidades em um determinado conjunto de dados (Cardoso, 2003).

## 2.1 O QUE SÃO SÉRIES TEMPORAIS

Uma série temporal é uma coleção de observações, feitas em diferentes instantes de tempo, não necessariamente igualmente espaçadas, e sujeitas a variações aleatórias. A inevitável presença de ruído em sinais experimentais torna o seu estudo ainda mais difícil, podendo acarretar interpretações incorretas dos resultados. Desta forma, o uso de técnicas adequadas é extremamente importante na análise de séries temporais (Savi, 2006). Um modelo adequado que forneça uma visão do funcionamento do mecanismo gerador dos dados pode ser usado, por exemplo, para prever valores futuros da série.

Algumas características são particulares deste tipo de dados, por exemplo:

- É preciso levar em conta a ordem temporal das observações;

- Fatores complicadores como presença de tendências e variação sazonal ou cíclica podem ser difíceis de estimar ou remover;

- A seleção de modelos pode ser bastante complicada e as ferramentas podem ser de difícil interpretação;

- É mais difícil lidar com observações perdidas e dados discrepantes, devido à natureza seqüencial.

A análise de séries temporais possui dois caminhos: a análise no domínio do tempo e a análise no domínio da freqüência. A análise no domínio do tempo concentra-se em descrever a magnitude de eventos que ocorrem em determinados instantes e na relação entre observações em diferentes instantes de tempo. A análise no domínio da freqüência analisa a freqüência de certos eventos que ocorrem em determinados períodos de tempo. As duas formas de análise de séries temporais se complementam, pois cada uma captura os diferentes aspectos existentes em uma série (Cardoso, 2003). As duas formas de análise não são alternativas, mas sim, complementares, cada uma mostrando diferentes aspectos da natureza da série temporal.

São objetivos dos estudos de séries temporais: a análise da dinâmica do sistema que gerou a série, permitindo reproduzí-lo ou mesmo predizer valores futuros da série.

As séries temporais são denotadas como $\{y(1), y(2),..., y(N)\}$, sendo, no caso de uma taxa de amostragem periódica, $y(k)$ a observação da série correspondente ao instante $t = kT$, $k = 1,2,...N$, onde $T$ é o período de amostragem. Além disso, as séries podem ser decompostas em várias componentes (funções) como: constante, tendência linear, variação cíclica, impulso, função degrau e rampa. Desta maneira, existem muitos modelos que se pode utilizar para representar uma série temporal. No caso de uma série aleatória a representação seria:

$$y(k) = b + \varepsilon(k) = bf(k) + \varepsilon(k), \qquad 2.1$$

onde $b$ é a média (constante), $\varepsilon(k)$ é a componente aleatória (também chamada de ruído) e $f(k) = 1$.

Em outros casos, onde a série apresenta um comportamento de tendência linear, pode-se reescrever a EQ. 2.1, como:

$$y(k) = b_1 + b_2 k + \varepsilon(k) = b_1 f_1(k) + b_2 f_2(k) + \varepsilon(k), \qquad 2.2$$

onde $b_1$ e $b_2$ são constantes, $f_1(k) = 1$ e $f_2(k) = k$ e $\varepsilon(k)$ é o ruído.

Observando os casos anteriores, podemos generalizar que a série temporal é representada da seguinte maneira:

$$y(k) = b_1 f(k-1) + b_2 f(k-2) + ... + b_q f(k-q) + b_{q+1} + \varepsilon(k), \qquad 2.3$$

onde $b_i$ são os parâmetros, $f_i(k)$ funções de $k$ e $q$ é o número de funções.

Por exemplo, considerando intervalos eqüidistantes, a série

$$y(30) = \{47,30,42,25,55,28,33,42,30,23,28,37,23,23,28, \qquad 2.4$$
$$42,23,56,14,24,35,33,45,42,10,30,28,17,36,25\}$$

pode representar o registro das vendas diárias de um produto $A$ numa loja no mês de junho de um determinado ano. Neste caso temos $N = 30$, pois são trinta valores de venda de produto, um valor por unidade de tempo (dia). A série temporal descrita na EQ. 2.4 foi representada graficamente na FIG. 2.1 e, embora estejamos considerando unidades discretas de tempo, a série será representada por uma curva.

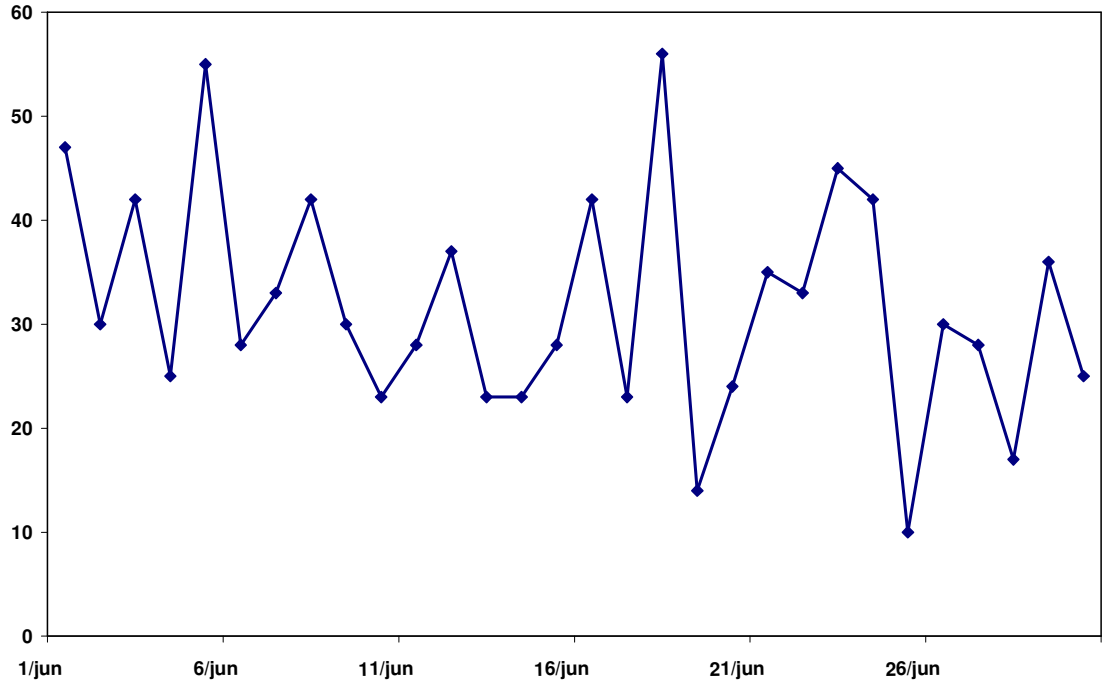


**FIG. 2.1 – Representação gráfica de uma série temporal. No eixo horizontal $x$ estão representados os 30 dias do mês de junho e no eixo vertical $y$ estão as unidades vendidas do produto $A$.**

Os gráficos são formas úteis para a visualização e interpretação de dados e muitas vezes são utilizados como ferramentas auxiliares à análise de dados. No entanto, para dados com dimensões elevadas, a análise visual não é uma tarefa simples. Além disso, devido à dimensão tempo, a série sempre apresentará no mínimo igual a dois.

## 2.2 DIMENSÃO DE IMERSÃO

A dimensão de imersão é a dimensão mínima de uma estrutura de características dinâmicas equivalentes às do sistema de interesse. A utilização da técnica de imersão procura gerar a estrutura mais próxima da real do espaço de estado, e também como base para outros cálculos, como a dimensão de correlação e o cálculo do expoente de Lyapunov. Existem, basicamente, três métodos utilizados na determinação da dimensão mínima de imersão:

- **Método da Saturação de algum Invariante do Sistema:** Neste método varia-se a dimensão de imersão para calcular um invariante. Assim, quando o invariante geométrico calculado convergir para um determinado valor, à medida que a dimensão de imersão aumenta, a dimensão correspondente é a dimensão mínima de imersão (Savi, 2006).

- **Método das Falsas Vizinhanças ou dos Falsos Vizinhos Próximos:** A idéia desse método é bastante intuitiva. Admite-se, a priori, que a dimensão de imersão mínima para uma determinada série temporal $y(k)$ é $m_0$. Isto significa que, em um espaço de dimensão $m_0$ é possível reconstruir o atrator. A ausência do cruzamento na órbita é testada verificando o número de pontos vizinhos da série. Um determinado ponto da série, sobre uma órbita, tem vizinhos que constroem o atrator, se sua dimensão for menor que o valor de imersão mínimo $m_0$, haverá cruzamento na órbita e surgirão os falsos vizinhos. Na ausência de falsos vizinhos, as propriedades topológicas do atrator são preservadas, pois ele está imerso num espaço de dimensão adequada. Em conclusão, a dimensão de imersão mínima $m_0$ é obtida quando, para dimensões crescentes, o número de falsos vizinhos é zero pela primeira vez.

- **Método da Decomposição dos Valores Singulares:** Este método é utilizado na diagonalização da matriz de covariância, identificando seus autovalores ou valores singulares. O número de autovalores não nulos é um valor estimado da dimensão mínima de imersão (Savi, 2006).

## 2.3 PASSO DE RECONSTRUÇÃO

O passo de reconstrução é a defasagem aplicada sobre a série temporal para construir o mapa de retorno dinamicamente equivalente ao original, ou seja, é a defasagem necessária à reconstrução dos atratores. Para reconstruir um atrator no espaço de estados devemos usar um atraso na série temporal com intuito de gerar uma outra coordenada.

O teorema da imersão de Takens (Takens, 1981) prevê as condições sob as quais um atrator pode ser reconstruído a partir dos dados de uma série temporal. Ele provou que, no espaço de fase formado por $\{y(t), y(t-\tau), y(t-2\tau),..., y(t-(m-1)\tau)\}$, o atrator reconstruído é topologicamente equivalente ao atrator "real", sobre o qual conhece-se apenas a evolução em tempo discreto da variável $y$. Na sua prova, Takens assumiu que a série era formada por infinitos pontos $y_j$ e que não há ruído. Com essas condições satisfeitas pode-se garantir que as propriedades topológicas do atrator reconstruído são preservadas, tomando-se:

$$m \geq 2D_0+1, \qquad 2.5$$

onde $D_0$ é a dimensão real do atrator, $m$ é a dimensão de imersão e $\tau$ é o passo de reconstrução. Chamou-se de espaço de imersão o espaço no qual se realizou a reconstrução.

A partir do teorema de imersão de Takens verifica-se a importância da escolha correta do passo de reconstrução. Dependendo do tipo de estrutura que queremos explorar, temos que escolher um passo de reconstrução que seja adequado. Obviamente, unidades de defasagem para dados de fluxo com uma grande amostragem produzirão vetores que são todos concentrados ao redor da diagonal no espaço de imersão e assim toda a estrutura perpendicular à diagonal torna-se quase imperceptível. A questão é que passos de reconstrução pequenos produzem elementos de vetor fortemente correlacionados, enquanto que passos de reconstrução grandes levam a vetores cujas componentes são (quase) não correlacionados e os dados são assim (aparentemente) aleatoriamente distribuídos no espaço de imersão.

Algumas ferramentas quantitativas estão disponíveis para guiar a escolha do passo de reconstrução. A função de auto-correlação e a informação mútua provêem uma informação razoável do passo de reconstrução.

### 2.3.1 MÉTODO DA INFORMAÇÃO MÚTUA

O passo de reconstrução obtido através do método da informação mútua foi sugerido por Fraser e Swinney (Fraser e Swinney, 1986) e, ao contrário da função de auto-correlação, também leva em conta as correlações não lineares. Portanto, esta técnica é uma boa ferramenta para se obter o passo de reconstrução.

A informação mútua entre uma medida $a_i$, pertencente ao conjunto $A = \{a_i\}$, e a medida de $b_j$, pertencente ao conjunto $B = \{b_j\}$, é a soma lida da medida de $a_i$ sobre a medida de $b_j$ em bits, dada por:

$$I(a_i, b_j) = \log_2 \left[ \frac{P_{AB}(a_i, b_j)}{P_A(a_i) P_B(b_j)} \right], \qquad 2.6$$

onde $P_{AB}$ é a densidade de probabilidade da medida de $A$ e de $B$ em resultar nos valores $a$ e $b$. $P_A(a)$ e $P_B(b)$ são densidades de probabilidades individuais medidas em $A$ e $B$. Em um sistema determinístico calculamos estas probabilidades através de um histograma das variações de $a_i$ e $b_j$ a serem medidos.

Se a medida de $A$ resultando em $a_i$ é totalmente independente da medida de $B$ resultando em $b_j$ então $P_{AB}(a,b)$ fatora da forma: $P_{AB}(a,b) = P_A(a) P_B(b)$ e a soma da informação entre as medidas, a informação mútua, é zero. O cálculo de todas as medidas desta informação estatística é chamado de informação mútua média entre a medida de $A$ e a medida de $B$ e é dada por:

$$I_{AB} = \sum_{a_i, b_j} P_{AB}(a_i, b_j) \log_2 \left[ \frac{P_{AB}(a_i, b_j)}{P_A(a_i) P_B(b_j)} \right]. \qquad 2.7$$

A quantidade de $I_{AB}$ não está expressa de forma a satisfazer as regras de evolução linear ou não linear das quantidades medidas. A idéia é colocar o conjunto de medidas como conectadas entre si, estabilizando-se assim, o critério de suas mútuas dependências da informação. Para isso definimos uma medida $y(t)$ no tempo fazendo uma ligação entre a informação teórica da media $y(t+\tau)$ num tempo $t+\tau$.

Tomamos o conjunto de medias de $A$ com valores observáveis $y(n)$ e para as medidas de $B$ os valores de $y(n+\tau)$. Então a informação mútua entre essas duas medidas é:

$$I(t) = \sum_{y(n), y(n+\tau)} P(y(n), s(n+\tau)) \log_2 \left[ \frac{P(y(n), y(n+\tau))}{P(y(n)) P(y(n+\tau))} \right]. \qquad 2.8$$

Em geral $I(\tau) \geq 0$. Quando $\tau$ se torna grande, o comportamento caótico do sinal torna as medidas de $y(n)$ e $y(n+\tau)$ independentes e $I(\tau)$ irá tender a zero.

A sugestão dada por Fraser e Swinney é tomar $\tau$ onde ocorre o primeiro mínimo da informação mútua $I(\tau)$ que será o melhor valor do passo de reconstrução para se reconstruir as componentes no espaço de estados. A sugestão mais requerida na generalização da noção da informação mútua para alta dimensionalidade no espaço de estados é que $y(n)$ seja substituído por vetores $m$-dimensionais e estes permitam uma boa reconstrução do atrator. Como “boa reconstrução” queremos dizer que as grandezas de interesse como dimensão fractal e expoente de Lyapunov são bem próximas das do atrator no espaço de estados original. A escolha do primeiro mínimo na informação mútua é equivalente à escolha do primeiro zero na função de auto-correlação linear.

### 2.3.2 MÉTODO DA FUNÇÃO DE AUTO-CORRELAÇÃO

A função de auto-correlação tenta estimar a medida da dependência linear entre um mesmo sinal defasado. O passo de reconstrução é encontrado quando a função de auto-correlação atinge o ponto zero pela primeira vez, caracterizando uma independência linear entre $y(t)$ e $y(t+\tau)$ (Savi, 2006). A função de auto-correlação estima a correlação entre observações defasadas no tempo que é dada por:

$$r_1 = \frac{\sum_{t=1}^{n-1}(y_t - \bar{y}_1)(y_{t+1} - \bar{y}_2)}{\sqrt{\sum_{t=1}^{n-1}(y_t - \bar{y}_1)^2 (y_{t+1} - \bar{y}_2)^2}}, \qquad 2.9$$

onde as médias amostrais são: $\bar{y}_1 = \sum_{i=1}^{n-1} y_t \big/ (n-1)$ e $\bar{y}_2 = \sum_{i=2}^{n} y_t \big/ (n-1)$. Já que $\bar{y}_1 \approx \bar{y}_2$ e assumindo variância constante, é possível reescrever a EQ. 2.9 da seguinte forma:

$$r_1 = \frac{\sum_{t=1}^{n-1}(y_t - \bar{y})(y_{t+1} - \bar{y})}{(n-1)\sum_{t=1}^{n-1}(y_t - \bar{y})^2 \big/ n}. \qquad 2.10$$

Generalizando a EQ. 2.10 para $k$ períodos de defasagem no tempo, temos:

$$r_k = \frac{\sum_{t=1}^{n-k}(y_t - \bar{y})(y_{t+k} - \bar{y})}{(n-1)\sum_{t=1}^{n-1}(y_t - \bar{y})^2}. \qquad 2.11$$

## 2.4 RECONSTRUÇÃO DO ESPAÇO DE ESTADO

Normalmente, um experimento não mede todas as variáveis de estado do sistema, geralmente tem-se somente a evolução temporal de uma variável,

representada pela série temporal $y(t)$. A técnica de reconstrução do espaço de estado usa essa série para extrair informações sobre a dinâmica do sistema, ou seja, utiliza-se das informações sobre variáveis de estado, contidas na história temporal do sinal de saída do sistema, para estimar suas características.

A técnica baseia-se no teorema da imersão de Takens (Takens, 1981). Conforme visto anteriormente, este teorema permite reconstruir um espaço de estado $m$-dimensional similar ao espaço de estado original, $D_0$-dimensional, a partir de uma única variável de estado, a variável medida (Savi, 2006). A partir dela, pode-se reconstruir um espaço de estados preservando invariantes geométricos do sistema, tais como a dimensão do atrator e os expoentes de Lyapunov.

Existem diferentes métodos para a reconstrução do espaço de estado, mas três métodos são mais utilizados: o método das derivadas, a decomposição em valores singulares (SVD) e o método das coordenadas defasadas.

O método das derivadas ou coordenadas derivativas corresponde ao primeiro método utilizado na reconstrução do espaço de estado. Neste método, as coordenadas são aproximações numéricas das derivadas de ordem sucessivamente superiores de uma variável medida, ou seja, (Savi, 2006)

$$\dot{y}(t) \approx \frac{y[(n+1)\Delta t] - y(n\Delta t)}{\Delta t}, \qquad 2.12$$

onde $y(t)$ é uma coordenada de um sistema físico em função do tempo, $\dot{y}(t)$ é a derivada no tempo da coordenada $y(t)$ e $n$ é um número inteiro.

A vantagem deste método é o seu significado físico palpável, pois a derivada de uma coordenada gera uma outra coordenada com significado físico. A desvantagem, por outro lado, é a sensibilidade ao ruído.

A idéia principal da decomposição em valores singulares é a de introduzir uma base de vetores ortonormais para o espaço de estados de forma que projeções em um certo número de direções preservem a fração máxima da variância dos vetores originais. Em outras palavras, o erro nas projeções é minimizado para um certo número de direções através de um problema de autovalores. As direções principais são aquelas para as quais obtêm-se uma matriz de covariância com os maiores autovalores. Essa matriz de covariância pode ser obtida de diferentes formas que correspondem às variações do método. Este método exige muito poder computacional e é, portanto, mais utilizado para sistemas de remoção de ruído do que para encontrar os vetores do espaço de estado.

O método das coordenadas defasadas ou método do *delay* é uma simplificação da EQ. 2.5. Este método foi proposto primeiramente por Ruelle (1979) e Packard (1980) e, logo depois, por Takens (1981) e Sauer (1991). A idéia básica da técnica é traçar $y(t)$ versus $y(t+\tau)$, onde $\tau$ é o passo de reconstrução. Este procedimento é motivado pelo fato de que a trajetória representada no espaço de fase reconstruído possui propriedades similares ao espaço de fase original, sendo topologicamente equivalentes (Savi, 2006). A sua principal característica é ser mais imune a ruído do que o método das derivadas. Isto se dá porque não dividimos por $\Delta t$. Não existem evidências claras sobre qual dos métodos é o melhor, mas a reconstrução do espaço de estado por meio de coordenadas defasadas é o mais explorado na literatura (Savi, 2006).

Um sistema dinâmico pode conter muitas variáveis, mas, muitas vezes, fazemos a medida de apenas uma delas. Em geral, sistemas dinâmicos apresentam-se na forma de equações dependentes, o que permite reconstruir a trajetória do sistema no tempo e no espaço de estados utilizando, na maior parte das vezes, um dos métodos de reconstrução do espaço de estados descrito acima.

## 2.5 PREDIÇÃO

A possibilidade de se prever o comportamento dos fenômenos de forma a poder utilizá-los com maior eficiência e eficácia é de grande interesse para a sociedade, constituindo-se em um ponto estratégico para um sistema de apoio a decisão. Desta forma, é de suma importância o desenvolvimento de métodos que possam garantir o entendimento, com a maior precisão e o menor custo possível, dos fenômenos de interesse no desenvolvimento das entidades contidas nos meios não-lineares, dinâmicos e complexos.

De forma genérica, um meio bastante eficiente e seguro para a tomada de uma decisão é o conhecimento dos fatos futuros, das tendências e dos possíveis cenários que estão por acontecer através de um sistema de análise e apoio à tomada de decisão que, de alguma forma, deve ser capaz de realizar uma previsão dos acontecimentos relevantes, sendo esta previsão um elemento chave para o seu sucesso ou o fracasso.

Predizer séries temporais significa representar as características de um determinado processo através de um modelo matemático e/ou computacional que possa estender tais características ao futuro, o que para tal é requerido que este modelo seja uma boa representação das observações em qualquer segmento de tempo próximo ao presente.

Desta forma, a predição de séries temporais consiste na estimativa dos parâmetros desconhecidos dos modelos apropriados para sua descrição, com o intuito de projetar tais modelos no futuro. É a estimativa, a partir de uma série temporal escalar, dos seus valores futuros, sem que se conheçam as equações de governo do fenômeno físico associado ao sistema.

O principal objetivo da utilização de técnicas de previsão é a identificação de determinados padrões presentes no conjunto de dados, determinando assim, um modelo que seja capaz de reconstruir os próximos padrões temporais. De uma

maneira geral, a predição consiste em ajustar um modelo aos dados de uma série (Savi, 2006).

Diversos modelos matemáticos e estatísticos têm sido utilizados determinar os acontecimentos futuros. Esses modelos de predição podem ser classificados em lineares e não-lineares. As técnicas lineares são aquelas em que o modelo de predição satisfaz às condições de linearidade, as demais técnicas podem ser classificadas como não-lineares (Savi, 2006). Modelos não-lineares são bastante complexos, tanto matematicamente quanto computacionalmente, tornado a eficiência prática destas técnicas equivalentes à eficiência de previsão dos modelos lineares.

Dentre as técnicas de predição lineares podemos destacar a predição por média móvel, predição por auto-regressão e predição por média móvel auto-regressiva. Já as técnicas não-lineares mais utilizadas são a lógica nebulosa, algoritmo genético e redes neurais. Neste trabalho, entretanto, dentre as técnicas não-lineares de predição, serão utilizadas a lógica nebulosa e algoritmo genético, estudadas posteriormente.

Tomando-se uma trajetória, definida sobre um espaço de estados, pode-se, a princípio, considerá-la como uma série temporal de dimensão superior, ou um conjunto de séries temporais unidimensionais, onde cada uma corresponde a uma componente. Deste modo, os mecanismos de predição usuais para séries temporais podem ser utilizados sobre espaços de estado quaisquer.

### 2.5.1 PREDIÇÃO POR MÉDIA MÓVEL

Dada uma série temporal qualquer y(t), o processo de média móvel de ordem $q$ ($MA(t,q)$ – Moving Average) pode ser descrito da seguinte forma:

$$MA(t,q) = \varepsilon(t) + \alpha_1\varepsilon(t-1) + \alpha_2\varepsilon(t-2) + ... + \alpha_q\varepsilon(t-q), \qquad 2.13$$

onde $\varepsilon(t)$ representa uma seqüência com distribuição normal, com média nula, variância constante e descorrelacionada, representando uma parcela não controlável da série. Assim o processo de média móvel representa em $MA(t,q)$ uma média ponderada dos termos $y(t-k)$ para $k = 0,...,q$.

### 2.5.2 PREDIÇÃO POR AUTO-REGRESSÃO

Podemos descrever um processo de auto-regressão de ordem $p$ ($AR(t,p)$ – Auto-Regressive) da seguinte forma:

$$AR(t,p) = \alpha_1 y(t-1) + \alpha_2 y(t-2) + ... + \alpha_p y(t-p) + \varepsilon(t), \qquad 2.14$$

onde $\varepsilon(t)$ representa uma seqüência com distribuição normal, com média nula, variância constante e descorrelacionada, representando uma parcela não controlável da série. Esse processo estima o valor esperado para a variável de estudo $AR(t,p)$ em função de seu próprio passado, atribuindo pesos a cada período ocorrido no passado. Por uma questão de estabilidade faz-se necessária a descrição das condições limitantes para os coeficientes $\alpha_k$, no caso de uma auto-regressão, por exemplo, $|\alpha_1| < 1$.

### 2.5.3 PREDIÇÃO POR MÉDIA MÓVEL AUTO-REGRESSIVA

É possível combinar um processo auto-regressivo com um processo de média móvel, chamamos de ARMA - Auto Regressive Moving Average, obtendo um modelo com a seguinte forma:

$$ARMA(p,q) = a_0 + \sum_{i=1}^{p} a_i AR(t-i) + \sum_{i=1}^{q} b_i MA(t-i). \qquad 2.15$$

Para estimarmos um modelo ARMA é necessário descrever um processo estável. Sendo assim, haverá restrições quanto aos coeficientes.

## 3. LÓGICA NEBULOSA

Aristóteles, filósofo grego que viveu entre 384-322 a.C. na Macedônia, foi o fundador da lógica. Ele vivia em busca de instrumentos para a compreensão de um mundo real e verdadeiro. Na obra Organon, reunião de seus trabalhos e de seus discípulos, Aristóteles estabeleceu um conjunto de regras rígidas para que as conclusões pudessem ser aceitas como logicamente válidas. O emprego da lógica de Aristóteles é sintetizado em uma linha de raciocínio baseada em premissas e conclusões. Dentro desse raciocínio, a lógica clássica tem sido binária, isto é, uma declaração é falsa ou verdadeira, não podendo ser ao mesmo tempo parcialmente verdadeira e parcialmente falsa.

As lógicas não clássicas violam justamente estas suposições binárias que não admitem ambigüidades e contradições. Por outro lado, o conceito de dualidade, estabelecendo que algo pode e deve coexistir com o seu oposto, faz as aplicações das lógicas não clássicas parecerem naturais e, até mesmo, inevitáveis. De acordo com o que foi dito por Zadeh (Camargos, 2002),

> "meu artigo de 1965 sobre conjuntos nebulosos foi motivado em grande escala pela convicção de que os métodos tradicionais de análise de sistemas não serviam para lidar com sistemas em que relações entre variáveis não prestavam para representação em termos de diferenciação ou equações diferenciais. Tais sistemas são o padrão em biologia, sociologia, economia e, usualmente, nos campos em que os sistemas são humanistas, ao invés de maquinistas, em sua natureza".

As lógicas multivaloradas surgem, a partir daí, como uma opção para o tratamento de situações que não sejam tão determinísticas (somente verdadeiro ou falso). Na lógica multivalorada, o valor verdade é visto como graus de verdade pertencentes ao intervalo unitário $[0,1]$. Ela possibilita os tratamentos dos valores indeterminados, que não são totalmente verdadeiros ou totalmente falsos (Pires, 2004).

A lógica nebulosa, também conhecida como lógica *fuzzy*, foi introduzida por Lofti A. Zadeh em seu artigo em 1965. Neste artigo, Zadeh introduziu uma teoria em que os objetos – conjuntos nebulosos (*fuzzy sets*) – são conjuntos com limites não precisos. A pertinência em um conjunto nebuloso não é uma questão de afirmação ou negação, mas uma questão de grau (Pires, 2004). O significado do artigo de Zadeh não foi apenas um desafio à teoria das probabilidades como única forma para a representação de incertezas, mas também pela necessidade de um método capaz de expressar de maneira sistemática quantidades imprecisas, vagas, mal definidas (Shaw e Simões, 1999), ou seja, é uma ferramenta capaz de capturar informações vagas, em geral descritas em uma linguagem natural e as converte para um formato numérico, de fácil manipulação.

Na TAB. 3.1, estão sintetizadas as principais características, vantagens e desvantagens da lógica nebulosa:

**TAB. 3.1 – Características, vantagens e desvantagens da lógica nebulosa**

| CARACTERÍSTICAS | VANTAGENS | DESVANTAGENS |
|---|---|---|
| A lógica nebulosa está baseada em palavras e não em números, ou seja, os valores verdades são expressos lingüisticamente. Por exemplo: quente, frio, longe, perto, etc. | O uso de variáveis lingüísticas nos deixa mais perto do pensamento humano. | São necessárias mais simulações e testes. |
| Possui vários modificadores de predicado, tais como: muito, pouco, mais, menos, entre outros. | São necessárias poucas regras, valores e decisões. | Não aprendem facilmente. |
| Possui um amplo conjunto de qualificadores, tais como: pouco, vários, em torno de, usualmente, etc. | Simplifica a solução de problemas e a aquisição da base do conhecimento. | Há uma certa dificuldade em estabelecer as regras corretamente. |
| Utiliza-se de probabilidades lingüísticas que são interpretadas como números nebulosos e manipuladas pela sua aritmética. | Mais variáveis observáveis podem ser valoradas. | Não há uma definição matemática precisa. |
| Manuseia todos os valores entre 0 e 1, tomando estes apenas como um limite. | É mais fácil de entender, manter e testar. | |
| | É robusta. Opera com falta de regras ou com regras defeituosas. | |
| | Acumula evidências contra e a favor. | |
| | Proporciona um rápido protótipo do sistema. | |

## 3.1 INTRODUÇÃO A CONJUNTOS NEBULOSOS

Na teoria clássica dos conjuntos, os conjuntos são ditos *"crisp"*, de tal forma que um dado elemento do universo de discurso, domínio ou espaço pertence ou não pertence ao referido conjunto. Na teoria dos conjuntos nebulosos existe um grau de compatibilidade de cada elemento com as propriedades ou características distintas de um determinado conjunto. Um conjunto nebuloso é um agrupamento impreciso e indefinido onde a transição de não pertinência para pertinência é gradual, não abrupta (Shaw e Simões, 1999).

Um conjunto nebuloso é caracterizado por uma função de pertinência que mapeia os elementos do domínio, $X$ no intervalo $[0,1]$. Isto é,

$$A : X \rightarrow [0,1]. \qquad 3.1$$

Utilizando uma convenção algébrica, o conjunto nebuloso $A$ em $X$ pode ser representado como um conjunto de pares ordenados de um elemento genérico $x \in X$,

$$A = \{(\mu_A(x)/x) \mid x \in X\}, \qquad 3.2$$

onde $\mu_A(x)$ representa do grau de $x$ em $A$, ou seja, o conjunto nebuloso $A$ pode ser escrito em uma forma em que cada argumento do conjunto mostra o elemento com seu respectivo valor de pertinência.

Em um conjunto com argumentos de pares ordenados, conforme representado na EQ. 3.2, o argumento consiste de dois elementos em uma ordem preestabelecida, onde seu primeiro elemento representa o grau de pertinência de $x$ em $A$ e o segundo elemento denota o valor de $x$. O conceito de função de pertinência desempenha um papel fundamental na teoria dos conjuntos nebulosos e será discutido na próxima seção.

## 3.2 FUNÇÕES DE PERTINÊNCIA

As funções de pertinência *fuzzy* representam os aspectos fundamentais de todas as ações teóricas e práticas de sistemas *fuzzy*. Uma função de pertinência é uma função numérica gráfica ou tabulada que atribui valores de pertinência *fuzzy* para valores discretos de uma variável, em seu universo de discurso (Shaw e Simões, 1999).

Os gráficos das funções podem ter diferentes formas e podem ter algumas propriedades específicas. Na prática, devido a sua formulação simples e eficiência computacional, as funções do tipo triangular e trapezoidal são amplamente utilizadas, especialmente em implementações com requisitos de funcionamento em tempo real. Contudo, o fato dessas funções de pertinência serem constituídas por segmentos de reta faz com que a derivada de primeira ordem da função não seja contínua. Isto é importante se for utilizado algum método de otimização baseado no gradiente. A ausência de continuidade nas transições entre segmentos de reta pode acarretar na divergência do método. Mesmo em métodos que independem da diferenciabilidade das funções de pertinência, o uso de funções não lineares pode ser recomendável por permitir transições mais suaves no valor de pertinência (Maruo, 2006). Além dos formatos tradicionais existe uma forma bastante utilizada em aplicações práticas: o conjunto unitário (*singleton*) (Pires, 2004).

A especificação de um formato para as funções de pertinência nem sempre é obvia, podendo inclusive não estar ao alcance do conhecimento de um especialista para a aplicação desejada. No entanto, existem sistemas nebulosos cujos parâmetros das funções de pertinência são completamente definidos pelo especialista. Neste caso, a escolha de funções triangulares e trapezoidais é mais comum porque a idéia de definição de regiões de pertinência total, média e nula é mais intuitiva (Maruo, 2006). Existem ainda algumas classes de métodos experimentais para a determinação de funções de pertinência como a abordagem horizontal, abordagem vertical, comparação em pares, inferência baseada na especificação do problema, estimação paramétrica e agrupamento nebuloso. A utilização da classe adequada depende da aplicação, em particular, da forma como a incerteza se manifesta e é observada durante o experimento (Maruo, 2006).

Entretanto, em trabalhos mais recentes existem tendências a projetos nos quais os parâmetros das funções de pertinência são ajustados de tal forma que a saída do sistema nebuloso descreva adequadamente o comportamento do sistema.

A seguir serão descrito os tipos mais utilizados de funções de pertinência e o conjunto unitário (*singleton*).

### 3.2.1 FUNÇÃO TRIANGULAR

A função triangular é definida da seguinte forma:

$$A(x) = \begin{cases} 0, & se \quad x \le a \\ \dfrac{x-a}{m-a}, & se\ x \in [a,m] \\ \dfrac{b-x}{b-m}, & se\ x \in [m,b] \\ 0, & se \quad x \ge b \end{cases}, \qquad 3.3$$

onde $m$ é um valor modal e $a$ e $b$ são os limites inferior e superior, respectivamente, para valores não nulos de $A(x)$. Na FIG. 3.1 é ilustrado o gráfico de uma função triangular, sendo $a = 1$, $b = 2.5$ e $c = 4$.

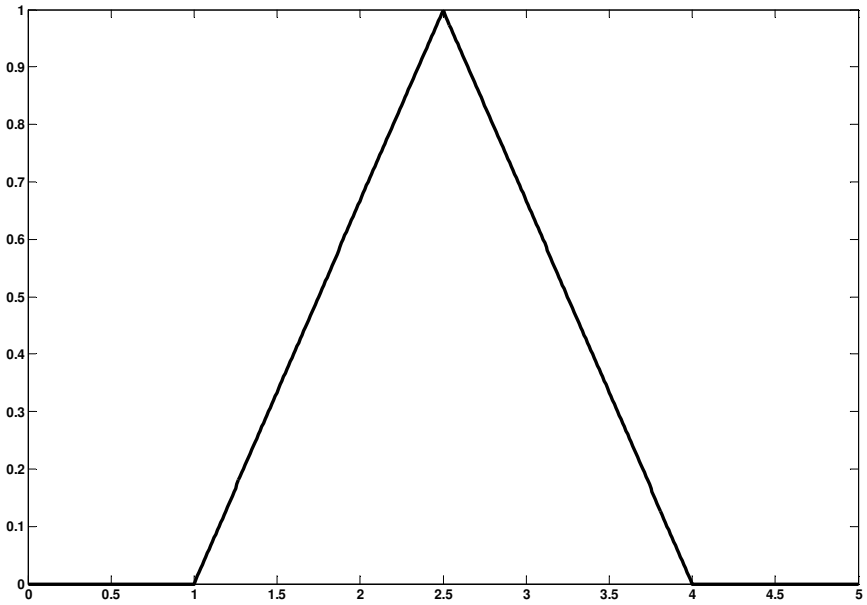


**FIG. 3.1 – Função triangular**

### 3.2.2 FUNÇÃO GAUSSIANA

A função gaussiana é definida por:

$$A(x) = e^{-x^2/\sigma^2}, \qquad 3.4$$

onde $\sigma$ é a meia largura a uma altura de $1/e$. Na FIG 3.2 é ilustrada forma de uma função gaussiana com $\sigma = 1$.

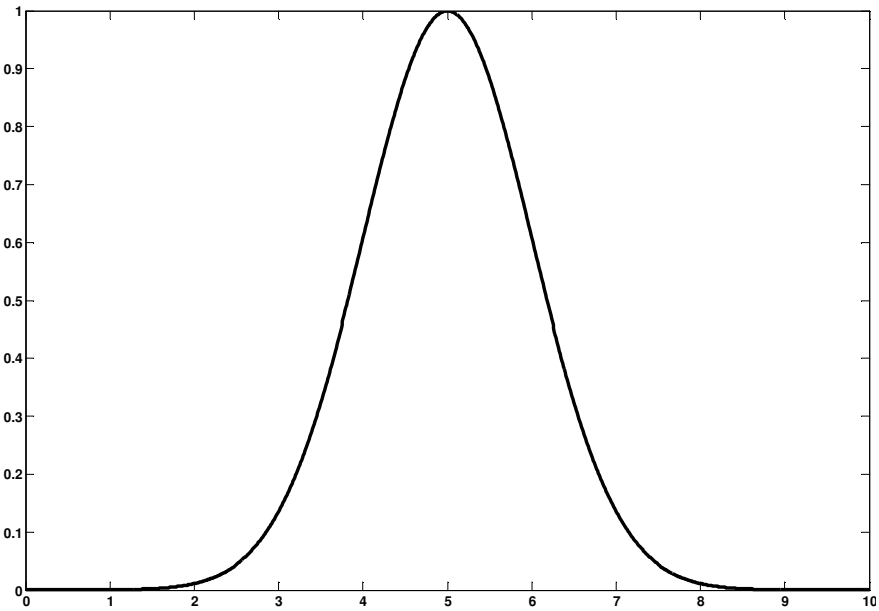


**FIG. 3.2 – Função gaussiana**

### 3.2.3 FUNÇÃO TRAPEZOIDAL

A função trapezoidal é definida por:

$$A(x) = \begin{cases} 0, & se \quad x < a \\ \dfrac{x-a}{b-a}, & se \quad a < x \le b \\ 1, & se \quad b < x \le c \\ \dfrac{d-x}{d-c}, & se \quad c < x \le d \\ 0, & se \quad x > d \end{cases} \qquad 3.5$$

e especificada por quatro parâmetros $\{a,b,c,d\}$, conforme EQ. 3.5. Graficamente, pode-se interpretar o conjunto $\{a,b,c,d\}$ com $a \le b \le c \le d$ como as coordenadas dos quatro vértices do trapézio formado pela função de pertinência. Para o caso especial em que $b = c$ a função trapezoidal se reduz a uma função triangular. E se $a = b$ e $c = d$, tem-se um conjunto clássico. Na FIG. 3.3 é ilustrado o gráfico de uma função trapezoidal, onde $a = 1$, $b = 4$, $c = 5$ e $d = 8$.

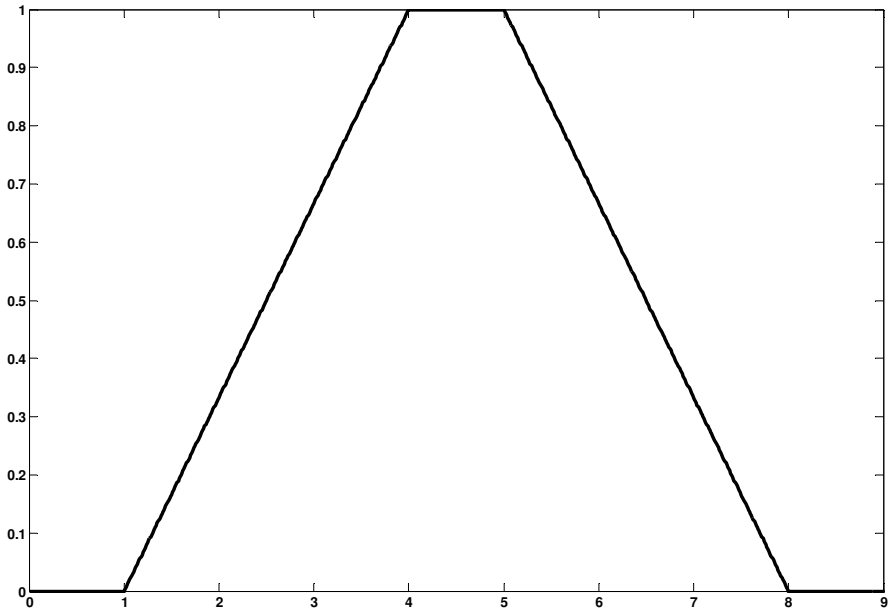


**FIG. 3.3 – Função trapezoidal**

### 3.2.4 FUNÇÃO SIGMOIDAL

As funções sigmoidais são definidas como funções monotônicas crescentes que apresentam propriedades assintóticas e de suavidade. Um exemplo de função sigmoidal é a função logística definida por:

$$A(x) = \frac{1}{1 + e^{-ax}}, \qquad 3.6$$

onde $a$ é o parâmetro que define a inclinação da função. Na FIG. 3.4 é ilustrado o gráfico de uma função sigmoidal.

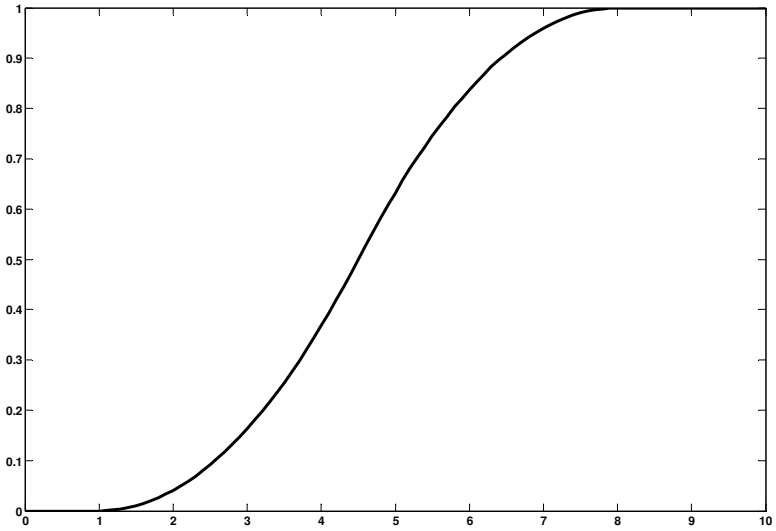


**FIG. 3.4 – Função sigmoidal**

### 3.2.5 FUNÇÃO SINO

A função sino é definida por:

$$A(x) = \frac{1}{1 + \left| \frac{x - c}{a} \right|^{2b}} \qquad 3.7$$

e especificada por três parâmetros $\{a, b, c\}$, conforme EQ. 3.7 Sendo que $b$ é usualmente positivo e $c$ indica o centro da curva.

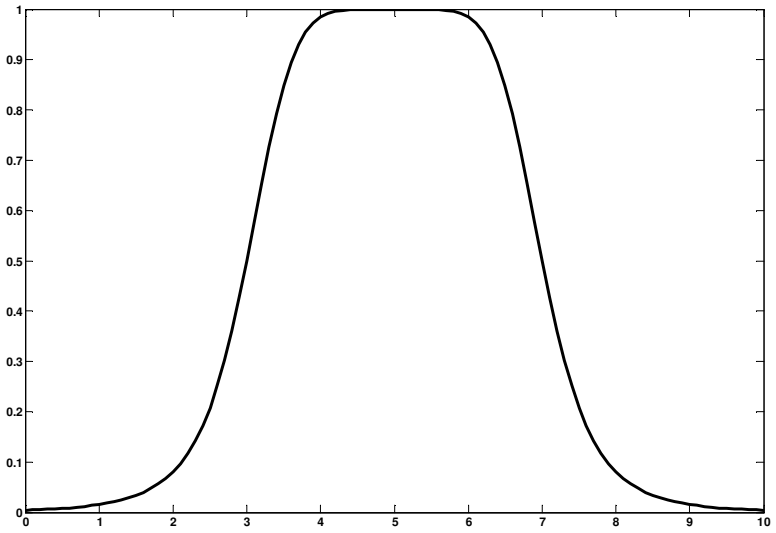


**FIG. 3.5 – Função sino**

Na FIG. 3.5 é ilustrada a forma de uma função gaussiana com $a = 2$, $b = 3$ e $c = 5$.

### 3.2.6 CONJUNTO UNITÁRIO (SINGLETON)

Um conjunto unitário (singleton) é uma função delta de Kronecker com altura controlada pelo parâmetro $h \in \Re \mid h < 1$ (Maruo, 2006). Ela é definida por:

$$A(x) = \begin{cases} h, & se \quad x = m \\ 0, & caso\ contrário \end{cases}. \qquad 3.8$$

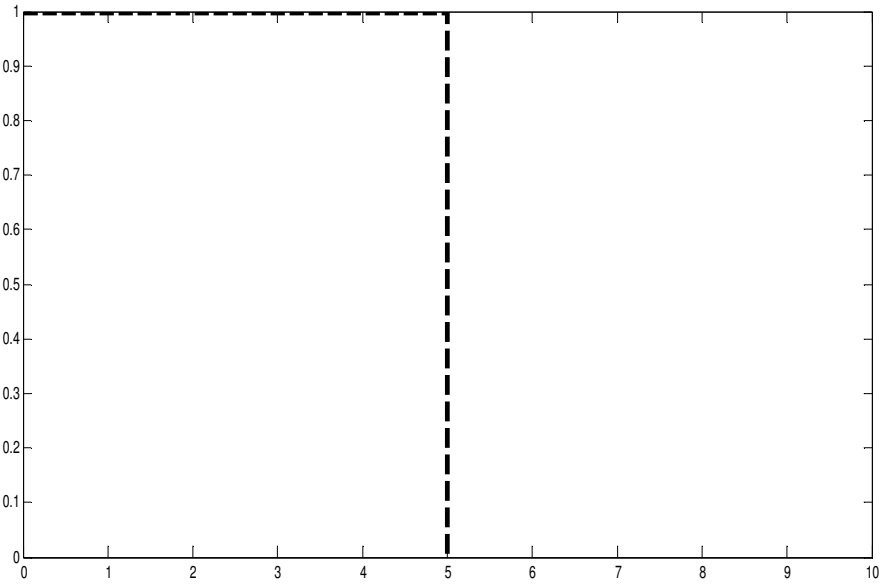


**FIG. 3.6 – Conjunto unitário**

O formato de um conjunto unitário é dado por um ponto. Na FIG. 3.6 observa-se que este ponto é dado pelo par ordenado (5, 1).

## 3.3 DEFINIÇÕES BÁSICAS EM CONJUNTOS NEBULOSOS

Alguns conceitos básicos estão associados aos conjuntos nebulosos. Entre eles encontram-se: corte $\alpha$, suporte, comprimento, núcleo, altura e normalização. A

seguir esses conceitos serão apresentados de forma sucinta, supondo que $A$ é um conjunto nebuloso sobre o conjunto base.

### 3.3.1 CORTE $\alpha$

O corte $\alpha$ é o conjunto de todos os elementos do universo $X$ com grau de pertinência em $A$ maior ou igual a $\alpha$.

$$A_\alpha = \{x \in X / \mu_A(x) \geq \alpha\}. \qquad 3.9$$

### 3.3.2 SUPORTE

O suporte é o conjunto que contém todos os elementos de $X$ que possuem grau de pertinência em $A$ diferentes de zero.

$$\sup orte(A) = \{x \in X / \mu_A(x) > 0\}. \qquad 3.10$$

### 3.3.3 NÚCLEO

O núcleo de $A$ compreende a todos os elementos de $X$ com grau de pertinência igual a um, ou seja, todos os elementos que são completamente compatíveis com o conceito expresso por $A$.

$$nucleo(A) = \{x \in X / \mu_A(x) = 1\}. \qquad 3.11$$

### 3.3.4 ALTURA

A altura de $A$ representa o maior grau de pertinência dos elementos de $X$, ou seja, o supremo da função de pertinência de $A$, conforme expresso na EQ. 3.12.

$$altura(A) = \sup_{x \in X}\{\mu_A(x)\}. \qquad 3.12$$

### 3.3.5 NORMALIZAÇÃO

Um conjunto nebuloso é dito normal quando sua altura é igual a um, ou seja, pelo menos um grau de pertinência dos elementos do conjunto atinge valor máximo ou se o seu núcleo for não vazio, enquanto que os conjuntos que não possuem altura igual a um são chamados subnormais.

Caso um conjunto nebuloso possua apenas um elemento com grau de pertinência igual a um, este elemento é denominado protótipo do conjunto.

## 3.4 OPERAÇÕES BÁSICAS COM CONJUNTOS NEBULOSOS

De forma semelhante à teoria dos conjuntos clássicos, as operações de união, interseção e o complemento possuem operações similares quando os operandos são conjuntos nebulosos.

Sejam $A$ e $B$ dois conjuntos nebulosos definidos no universo de discurso $E$, com suas funções de pertinência $\mu_A$ e $\mu_B$ respectivamente, tal que $x \in E$. Na teoria dos conjuntos nebulosos, a interseção é implementada por uma família de operadores denominados T-normas e a união são implementadas por uma família de operadores denominados T-conormas (Pires, 2004).

### 3.4.1 UNIÃO

Sejam os conjuntos $A \subset E$ e $B \subset E$. A união $A \cup B$ é o menor subconjunto do universo de discurso $E$, que inclui ambos os conjuntos nebulosos $A$ e $B$. A união é o contorno que inclui ambos os conjuntos nebulosos $A$ e $B$ e, portanto, é sempre

maior que qualquer um dos conjuntos individuais $A$ e $B$. Por essa razão, o vetor de pertinência de união, ou seja, os componentes do vetor de pertinência são calculados dos valores individuais de $A$ e $B$ (Shaw e Simões, 1999), conforme segue:

$$\mu_{A\cup B}(x) = \max[\mu_A(x), \mu_B(x)] \text{ ou } \mu_A \cup \mu_B(x) = \mu_A(x) \tilde{*} \mu_B(x) \qquad 3.13$$

### 3.4.2 INTERSEÇÃO

Sejam os conjuntos $A \subset B$ e $A \subset E$, então a interseção $A \cap B$ é o maior subconjunto do universo de discurso $E$, o qual é ao mesmo tempo parte de $A$ e também parte de $B$. A interseção é a parte comum dos conjuntos $A$ e $B$ e, como resultado, é sempre menor que qualquer um dos conjuntos individuais $A$ e $B$. Por essa razão, o vetor de pertinência da interseção $A \cap B$ é calculado dos vetores individuais de $A$ e $B$ (Shaw e Simões, 1999), conforme segue:

$$\mu_{A\cap B}(x) = \min[\mu_A(x), \mu_B(x)] \text{ ou } \mu_A \cap \mu_B(x) = \mu_A(x) \hat{*} \mu_B(x). \qquad 3.14$$

### 3.4.3 COMPLEMENTO

Seja o conjunto $A \subset E$, o complemento de $A$ em relação à $E$ é denominado $A'$ composto por todos os elementos $x \in E$ que não são membros de $A$ (Shaw e Simões, 1999). O vetor de pertinência do complemento é calculado como segue:

$$\mu_{A'}(x) = 1 - \mu_A(x). \qquad 3.15$$

## 3.5 VARIÁVEL LINGÜÍSTICA

A variável lingüística é a estruturação primária de qualquer sistema baseado na lógica nebulosa, onde múltiplas categorias subjetivas que descrevem o mesmo

contexto são combinadas. Uma variável lingüística pode ser definida, de uma maneira informal, como uma variável cujos valores são conjuntos de termos, terminologias, nomes ou rótulos, ao invés de números.

Conforme apontado por Zadeh, técnicas convencionais para a análise de sistemas são essencialmente inadequadas para o tratamento de sistemas no conhecimento humano, cujo comportamento é influenciado pela percepção, julgamento e emoções (Maruo, 2006). Sendo assim, Zadeh propôs o conceito de variáveis lingüísticas como sendo uma variável onde os valores são palavras ou sentenças na linguagem natural ou artificial, utilizadas como alternativa na modelagem do pensamento humano em que a informação é processada através de conjuntos nebulosos.

Caracteriza-se uma variável lingüística por uma quíntupla denotada por

$$\langle v, T, X, g, m \rangle, \qquad 3.16$$

onde:

- $v$ é o nome da variável;
- $T$ é o conjunto de termos lingüísticos de $v$ que faz referencia a uma base que pertence à regra superior num conjunto universal. Cada elemento de $T(v)$ representa um rótulo $L$ dos termos que a variável $v$ pode assumir;
- $X$ é o universo de discurso da variável $v$;
- $g$ é uma regra sintática para gerar termos ou rótulos lingüísticos;
- $m$ é uma regra semântica que associa a cada rótulo $L$, um conjunto nebuloso no universo de discurso, significando $m(L)$.

A gramática define como os termos primários {baixa, média, alta} serão associados aos modificadores {muito, pouco, maior, menor, não} para formar os nomes dos termos não primários. Os conjuntos nebulosos associados aos termos não primários (muito alta, pouco alta, não alta) podem ser derivados do uso de modificadores pré especificados (Pires, 2004).

A quantidade de valores lingüísticos define a granularidade, isto é, a especificação e distribuição dos termos lingüísticos e, por conseguinte, a partição nebulosa do universo de discurso correspondente. Um número pequeno de termos lingüísticos define uma partição esparsa ou grossa, ao passo que um número maior resulta numa partição fina.

A partição do universo de discurso também pode ser vista como uma forma de compreensão nebulosa de dados. Utilizando o agrupamento nebuloso de informações de natureza similar (*fuzzy clusters*), pode-se desprezar parte da informação inútil, indesejada ou redundante. Assim, a granularização da partição pode ser usada para direcionar a análise nos aspectos de interesse permitindo maior ênfase em áreas especificas dos universos das variáveis de entrada (Maruo, 2006).

Não existe uma metodologia consistente para a determinação da partição nebulosa ideal. Em geral essa tarefa é realizada manualmente através da intervenção de um especialista ou utilizando um método de particionamento automático a partir de dados de treinamento.

## 3.6 REGRAS NEBULOSAS

As regras nebulosas, também conhecidas como implicações nebulosas ou declarações condicionais nebulosas, permitem uma maneira formal de representação de diretivas e estratégias. As regras nebulosas são muito apropriadas quando o conhecimento do domínio resulta de associações empíricas e experiências do operador humano, ou quando se deseja uma representação lingüística do conhecimento adquirido (Delgado, 2002).

Em geral, as regras nebulosas assumem a forma:

Se $x$ é $A$ então $y$ é $B$ 3.17

onde $A$ e $B$ são rótulos lingüísticos definidos nos universos $X$ e $Y$ respectivamente. Freqüentemente, $x$ é $A$ é denominado de antecedente ou premissa, enquanto $y$ é $B$ é denominado de conseqüente ou conclusão (Maruo, 2006). Os antecedentes descrevem uma condição, enquanto a parte conseqüente descreve uma conclusão ou uma ação que pode ser esboçada quando as premissas se verificam. A expressão "se a *velocidade* é *alta* então a *pressão* é *baixa*" é um exemplo de uma regra nebulosa que relaciona as variáveis lingüísticas *velocidade* e *pressão*, combinando os conjuntos nebulosos associados aos termos lingüísticos *alta* e *baixa*.

## 3.7 RACIOCÍNIO APROXIMATIVO

Todo método de raciocínio pode ser definido como um processo de inferência que produz conclusões a partir de um conjunto formado por uma ou mais regras e fatos conhecidos (Delgado, 2002).

Na lógica tradicional de dois valores, as principais ferramentas de raciocínio são as tautogias, como por exemplo, o *modus ponens*:

Premissa: *A* é verdade
Implicação: Se *A* = *B*. 3.18
Conclusão: *B* é verdade

Como exemplo, considere o fato de que a "banana é amarela" e a regra "se a banana é amarela então ela está madura". Pode-se concluir que "a banana está madura".

Duas óbvias generalizações do *modus ponens* são: (1) permitir que sentenças fossem caracterizadas por conjuntos nebulosos e (2) relaxar levemente a exigência de igualdade dos *B*'s da implicação e da conclusão (Pires, 2004).

Esta versão generalizada do *modus ponens* foi então chamada de "*modus ponens* generalizado". A inferência do *modus ponens* generalizado ocorre da seguinte forma:

Premissa: *x* é *A'*
Implicação: Se *x* é *A* então *y* é *B*. 3.19
Conclusão: *y* é *B'*

Neste caso tem-se o fato de que a "banana é mais ou menos amarela", podendo concluir-se que "a banana está mais ou menos madura".

Quando $A$, $A'$, $B$ e $B'$ são conjuntos nebulosos, o método de raciocínio é chamado de raciocínio nebuloso. É sugerida a regra de inferência composicional para o tipo de inferência apresentada pelo *modus ponens* generalizado (Pires, 2004). A seguir será descrito como obter uma conclusão a partir de premissas e implicações nebulosas, considerando a regra de inferência composicional de Zadeh.

## 3.8 REGRA COMPOSICIONAL DE INFERÊNCIA

O procedimento conhecido como regra composicional de inferência é uma generalização do processo de avaliação de uma função em um dado ponto (Maruo, 2006).

Assumindo que um conjunto nebuloso $A \subset E_1$ induz outro conjunto nebuloso $B \subset E_2$ cuja função de pertinência é dada por $\mu_B(y/x)$. Se o parâmetro $x$ da função de pertinência pertencer ao conjunto $A(x) \subset E_1$ e adicionalmente $B(y) \subset E_2$ e a

relação nebulosa $R(x, y)$ em $E_1 \times E_2$. Dessa forma, podemos descrever a regra composicional como:

$$B(y) = A(x) \circ R(x, y) \qquad 3.20$$

onde $\circ$ é o operador composicional que indica uma operação generalizada similar a uma T-norma ou T-conorma (Shaw e Simões, 1999).

Para fins práticos, a regra composicional de inferência pode ser escrita em termos das funções de pertinência dos respectivos conjuntos. Usando os operadores composicionais mais comuns, os valores de pertinência do vetor de pertinências $B(y)$ podem ser calculados da seguinte forma:

• **Composição máx-min:** $\mu_B(y) = \max\{\min[\mu_A(x), \mu_R(x, y)]\}$ 3.21

• **Composição máx-produto:** $\mu_B(y) = \max\{\mu_A(x) \cdot \mu_R(x, y)\}$ 3.22

A regra composicional de inferência diz que a relação nebulosa que representa um sistema $R(x, y)$ é conhecida, então a resposta do sistema $B(y)$, pode ser determinada a partir de uma excitação conhecida $A(x)$ (Shaw e Simões, 1999).

## 3.9 SISTEMAS NEBULOSOS

Os sistemas de inferência nebulosos, criados por Zadeh nos anos 60, podem ser aceitos como a melhor ferramenta para modelar o raciocínio humano, que é aproximado e parcial em sua essência. Ainda existem muitas dificuldades em aplicar métodos de modelagem existentes a muitos sistemas reais complexos, com características não lineares ou variantes no tempo. As técnicas de sistemas nebulosos são especialmente utilizadas nos casos onde não existem modelos matemáticos capazes de descrever precisamente o processo proposto. Estas

técnicas fornecem uma estrutura poderosa para manipular informações aproximadas.

O modelo nebuloso de um sistema consiste em um grupo finito de implicações nebulosas que juntas formam um algoritmo para determinar as saídas do processo, baseado em um número finito de entradas e saídas passadas. Os sistemas de controle resultantes proporcionam um resultado mais acurado além de um desempenho estável e robusto.

Um sistema de decisão nebuloso ou controlador nebuloso é então formado por um conjunto de regras de controle, às quais se associam implicações nebulosas utilizando as operações precedentes.

A estrutura básica de um sistema nebuloso possui três componentes conceituais: uma base de regras, que contém o conjunto de regras nebulosas; uma base de dados, que define as funções de pertinência usadas nas regras nebulosas; e um mecanismo de raciocínio, que realiza um procedimento de inferência (raciocínio nebuloso) para obter a saída ou conclusão, baseado nas regras e fatos conhecidos (Delgado, 2002).

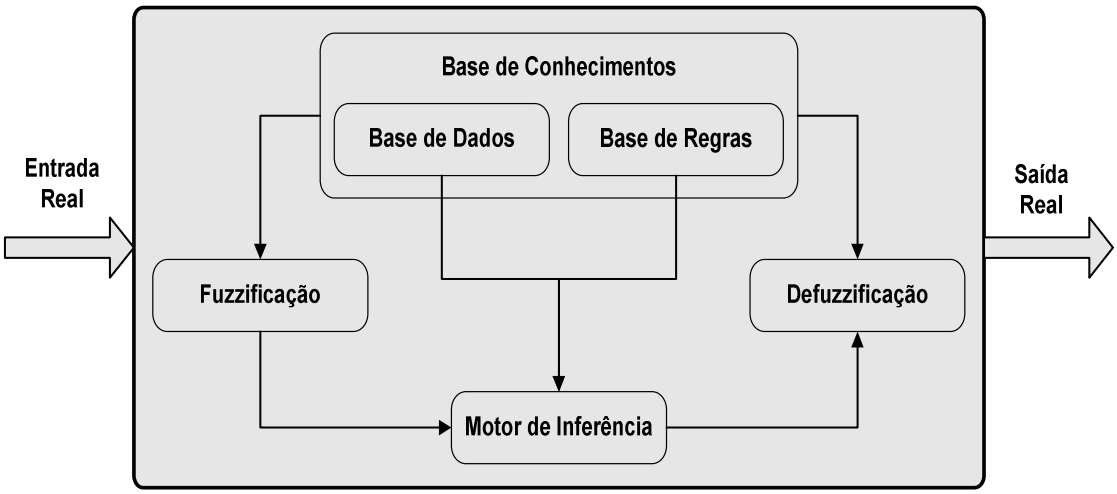


**FIG. 3.7 – Estrutura básica de um sistema nebuloso**

Na FIG. 3.7 está ilustrada a estrutura básica de um sistema nebuloso, (Pires, 2004) constituída por: um módulo de *fuzzificação*, que tem como função converter os

valores de entrada do sistema para termos lingüísticos representados por conjuntos nebulosos; uma base de conhecimento, onde todo o conhecimento sobre o domínio do problema em questão é armazenado. A base de conhecimento é formada por uma base de regras, que contém o conjunto de regras nebulosas e por uma base de dados, que define as funções de pertinência usadas nas regras nebulosas; um motor de inferência, que é responsável pelo desenvolvimento do raciocínio nebuloso baseado no conhecimento representado na base de conhecimento; e por um módulo de *defuzzificação*, que realiza uma transformação da resposta do sistema nebuloso, a qual está representada por um conjunto nebuloso, em uma resposta não nebulosa.

Devido à simplicidade de implementação e ajuste e da não necessidade do modelo matemático preciso do processo, os sistemas nebulosos vêm sendo mais e mais aplicados em vários campos, qualquer que seja a complexidade do sistema.

No entanto, a falta de metodologia para projeto e ajustes tornam essas tarefas muito dependentes da sensibilidade que o projetista possui a respeito do processo. Se este for muito complexo, as variáveis utilizadas no controlador podem dificultar o ajuste.

### 3.9.1 MODELOS DE SISTEMAS NEBULOSOS

Existem vários tipos de sistemas nebulosos. Na maioria dos casos, o antecedente é formado por proposições lingüísticas e a distinção entre os modelos se dá no conseqüente das regras nebulosas. Entre os modelos mais conhecidos pode-se destacar (Pires, 2004 e Maruo, 2006):

- o modelo **Lingüístico** ou **Mandani** que utiliza conjuntos nebulosos também nos conseqüentes das regras nebulosas. A saída final é representada por um conjunto nebuloso resultante da agregação da saída inferida de cada regra. Para se obter

uma saída final não nebulosa adota-se um dos métodos de transformação da saída nebulosa em não nebulosa.

- o modelo **Interpolativo** ou **Takagi-Sugeno-Kang**, no qual o conseqüente é representado por uma função das variáveis de entrada. A saída final é obtida pela média ponderada das saídas inferidas de cada regra. Os coeficientes da ponderação são dados pelos graus de ativação das respectivas regras.

- o modelo de **Tsukamoto**, que utiliza funções de pertinência monotônicas no conseqüente. Assim como no modelo Takagi-Sugeno-Kang, é inferido um valor não nebuloso induzido pelo nível de ativação da regra. A saída final é obtida por média ponderada das saídas inferidas de cada regra.

### 3.9.2 SISTEMA NEBULOSO TAKAGI-SUGENO-KANG

O modelo de Takagi-Sugeno-Kang, também conhecido como TSK ou simplesmente TS, é uma abordagem sistemática para a geração de regras nebulosas a partir de um conjunto de dados de entrada e saída. O modelo TSK foi proposto como resultado de um esforço para se desenvolver, de forma sistemática, uma abordagem para a geração de regras nebulosas a partir de dados de entrada e saída.

O modelo TSK pode ser visto como uma combinação entre conhecimento lingüístico (parte antecedente) e regressão matemática (parte conseqüente), de tal forma que os antecedentes descrevem regiões nebulosas no espaço de entrada nas quais as funções conseqüentes são válidas. Uma regra típica de um sistema com duas variáveis de entrada utilizando o sistema TSK tem a forma:

Se $x_1$ é $A$ e $x_2$ é $B$ então $y^i = g(w, x)$, 3.23

onde $A$ e $B$ são conjuntos nebulosos no antecedente e $g(w,x)$ é uma função das variáveis de entrada $x=[x_1,x_2]$ com parâmetros $w$, também conhecida como conseqüente (Maruo, 2006).

Geralmente $g(w,x)$ é uma função polinomial de $x$, mas pode ser qualquer função que descreva apropriadamente o comportamento do sistema em uma região determinada pelo antecedente da regra. Geralmente, a função que mapeia a entrada e saída para cada regra é uma combinação linear das entradas, isto é, $z=px_1+qx_2+r$. No caso em $p=q=0$, temos $z=r$, chamado modelo Takagi-Sugeno de ordem zero, que pode ser visto como um caso especial de um modelo de Mandani no qual o conseqüente é especificado por um conjunto unitário (singleton).

A saída de um modelo nebuloso TSK de ordem zero é uma função suave de suas variáveis de entrada enquanto as funções de pertinência vizinhas apresentam um grau de sobreposição suficiente, ou seja, a sobreposição entre as funções de pertinência no antecedente de um modelo TSK de ordem zero é condição necessária para a suavidade do mapeamento de entrada e saída (Maruo, 2006).

Como cada regra possui uma saída convencional, a saída global é obtida através da média ponderada. Na prática, o operador de média ponderada é algumas vezes substituído pela soma ponderada visando reduzir o custo computacional. Esta simplificação, embora seja interessante, por exemplo, em aplicações de controle com tempo real, pode levar a uma perda do significado lingüístico associado aos sistemas nebulosos. Nos últimos tempos, a possibilidade de obtenção dos parâmetros dos conseqüentes através de métodos de otimização tem reavivado o interesse por estes modelos, uma vez que os resultados obtidos justificam, em muitas aplicações, a perda de interpretabilidade lingüística resultante do uso de conseqüentes não lingüísticos (Delgado, 2002).

Nesse trabalho somente os sistemas nebulosos TSK foram considerados.

# 4 AGRUPAMENTO DE DADOS

Para o ser humano é natural realizar a tarefa de agrupar objetos por suas similaridades, como, por exemplo, aves, peixes, plantas, etc. É dessa maneira que aprende a distinguir o que há a sua volta. De forma similar, o processo artificial de agrupamento visa separar os dados em grupos que contenham atributos similares, facilitando a compreensão das informações que guardam.

Classificar ou agrupar objetos ou dados em categorias é atividade bastante comum e vem sendo intensificada devido ao número elevado de informações que estão disponíveis atualmente. Para realizar esta tarefa utiliza-se um mecanismo denominado análise de *cluster* ou clusterização. A clusterização é um método que utiliza o aprendizado não supervisionado ou auto-organizável, ou seja, não há um "professor" ou "crítico" que lhe indique o que cada padrão representa.

A idéia básica do processo de agrupamento de dados (*clustering*) é uma tarefa em que se busca identificar um número finito de categorias ou grupos (*clusters*) para descrever um determinado conjunto de dados, tanto minimizando a homogeneidade de cada grupo como a heterogeneidade entre grupos distintos (Pucciarelli, 2005).

Um bom algoritmo de agrupamento caracteriza-se pela produção de classes de alta qualidade, nas quais a similaridade intra-classe é alta e a inter-classes é baixa. A qualidade do resultado depende do método utilizado para medir esta similaridade, normalmente baseado em análise de distância entre pares de objetos, e da habilidade em descobrir algum ou todos os padrões escondidos.

Os algoritmos de agrupamento podem ser organizados de acordo com a categoria de similaridade ou dissimilhança, dentro desta categoria existem cinco categorias principais:

- **Métodos de particionamento:** dados $n$ objetos, o método constrói $k$ partições, onde cada partição representa um grupo, e $k \leq n$. O conjunto dos grupos satisfaz os seguintes requisitos: (1) cada grupo tem ao menos um objeto e (2) cada objeto pertence somente a um grupo. Este tipo de algoritmo utiliza normalmente uma dessas duas técnicas: *Kmeans*, onde o grupo é representado por uma medida média de todos seus objetos, e a *Kmedoids*, onde o grupo é representado por um objeto próximo ao centro deste;

- **Métodos hierárquicos:** este método decompõe os dados do conjunto em uma hierarquia, que pode ser formada de forma aglomerativa ou divisiva. Na aglomerativa, cada objeto inicia como sendo um grupo, e a cada iteração é ligado a um grupo similar. Na divisiva, todos começam em um grande grupo, que vai sendo repartido em cada iteração. Os principais tratamentos para melhorar a qualidade dos métodos hierárquicos são: (1) realizar uma análise cuidadosa das conexões dos objetos em cada partição hierárquica ou (2) utilizar um algoritmo de hierarquização aglomerativo e aplicar em seguida uma relocação iterativa para refinar os resultados;

- **Métodos baseados em densidade:** este método baseia-se na idéia de densidade, ou seja, continuar aumentando um dado grupo até que a densidade (número de objetos) na vizinhança exceda algum limite estabelecido. Assim, para cada objeto em um grupo, a vizinhança deste, dentro de um raio dado, deve conter ao menos um número mínimo de outros objetos;

- **Métodos baseados em grade:** estes métodos quantificam os objetos em um número finito de células, que formam uma estrutura de grade. As operações de agrupamento são executadas nesta estrutura de grade. A principal vantagem é a velocidade de processamento, que depende apenas do número de células da grade;

- **Métodos baseados em modelo:** neste método, um modelo hipotético é criado para cada grupo e encontra-se o melhor ajuste do objeto ao modelo dado. Podem

localizar os grupos através de uma função densidade que reflete a distribuição espacial dos objetos.

Também podem ser organizados de acordo com a categoria dos conceitos e fundamentos. Dessa maneira, os algoritmos podem ser:

- **Nebulosos:** utilizam conjuntos nebulosos para classificar dados e consideram que um ponto pode ser classificado em mais de um grupo, mas com diferentes graus de associação. Estes tipos de algoritmos levam para um esquema de agrupamento que são compatíveis com experiências da vida cotidiana, pois tratam as incertezas dos dados reais. O mais representativo algoritmo de agrupamento nebuloso é o *Fuzzy C-Means* (FCM);

- **Rígidos:** não consideram a sobreposição de grupos aos quais um ponto de X pertence. O algoritmo deverá resultar em um agrupamento com valores na matriz de pertinência restritos ao conjunto {0, 1};

- **Neurais:** adotam abordagens conexionistas para o agrupamento. Em geral, utilizam redes neurais artificiais com aprendizagem não supervisionada. Um exemplo típico é o método dos mapas organizáveis de Kohonen;

- **Estatísticos:** os conceitos lógico-matemáticos empregados na análise estatística para agrupamento de dados são principalmente probabilísticos. Os métodos estatísticos são conhecidos também como métodos de máxima verossimilhança (*maximum-likelihood*), sendo que a aproximação mais aplicada para o cálculo é a abordagem Bayesiana.

Os algoritmos de agrupamento nem sempre são facilmente classificados dentro de apenas uma das categorias citadas, pois utilizam vários métodos na sua construção. A aplicação das técnicas e métodos disponíveis depende do problema a ser resolvido e da quantidade de dados a ser trabalhada.

Atualmente, algoritmos de agrupamento de dados estão sendo utilizados em várias áreas de conhecimento, principalmente em economia, biologia, medicina, geografia, classificação de documentos, entre outras. Um exemplo clássico é o de classificação demográfica, que serve de início para a determinação das características de um grupo social, visando desde hábitos de compras até utilização de meios de transporte e de comunicação (Cardoso, 2003).

Neste capítulo apresentamos as duas técnicas de agrupamento de dados utilizadas nesta dissertação: agrupamento nebuloso *c-means*, utilizado para encontrar exatamente a dimensão de imersão ótima e, conseqüentemente, a quantidade de funções de pertinência que compõem o antecedente de cada regra do sistema nebuloso proposto; e agrupamento subtrativo, utilizado para determinar a quantidade de regras que vão compor o sistema nebuloso proposto.

## 4.1 AGRUPAMENTO DE DADOS NEBULOSO

O primeiro algoritmo de agrupamento de dados nebuloso, desenvolvido em 1969 por Ruspini, é uma extensão do algoritmo *c-means* rígido (HCM), chamado de ISODATA proposto por Ball e Hall em1965. O HCM é um dos mais populares métodos de agrupamento. Ruspini, em 1969, introduziu a partição nebulosa para descrever estruturas de grupos de um conjunto de dados e sugeriu um algoritmo computacional que otimiza esta partição nebulosa. Dunn, em 1973, generalizou o procedimento de agrupamento de variância mínima para a técnica de agrupamento nebuloso do HCM. Bezdek, em 1981, generalizou a aproximação de Dunn criando assim o algoritmo de agrupamento de dados nebuloso *c-means* (FCM) (Pucciarelli, 2005).

O algoritmo de agrupamento nebuloso *c-means* consiste num método onde os elementos pertencem a todos os grupos com graus de pertinência diferentes, obtidos a partir da distância entre o elemento e os centros dos grupos. O algoritmo executa um particionamento nebuloso de um conjunto de dados em classes.

Existem, em contrapartida, vetores de dados e atribuições para classes distintas empregadas nas técnicas de agrupamento estatístico convencionais. A maioria das técnicas de agrupamento de dados tem propriedades similares e produzem resultados comparáveis, mas as aproximações utilizando lógica nebulosa têm a vantagem da representação efetiva da imprecisão (Jiang e Adeli, 2003).

Os passos básicos do algoritmo FCM, fornecido por Bezdek, são os seguintes:

**1)** Escolher o número de grupos $c$, $2 \le c \le n$, o parâmetro $m > 1$, o critério de parada $\varepsilon > 0$, o número máximo de iterações $lmáx$.

**2)** Inicializar $U^0$ aleatoriamente e o contador $l = 1$.

**3)** Calcular os $c$ centros das classes $v_i^l$ para $i = 1,...,c$ usando $U^l$ com a equação abaixo:

$$v_i^l = \frac{\sum_{k=1}^{n} \left[u_{ki}^{l-1}\right]^m x_k}{\sum_{k=1}^{n} \left[u_{ki}^{l-1}\right]^m}, \quad i = 1,...,c .$$

**4)** Atualizar a matriz de pertinência:

4.1) Para $1 \le i \le c$ e $1 \le k \le n$

4.1.1) Se $d_{ki} > 0$, então:

$$u_{ki}^l = \left[ \sum_{j=1}^{c} \left( \frac{d_{k1}}{d_{kj}} \right)^{\frac{1}{m-1}} \right]^{-1}$$

4.1.2) Senão

Se $d_{ki} = 0$ para $i \in I \le c$

Então definir $u_{ki}^l$ para $i \in I$ com um número real positivo que satisfaça a condição $\sum u_{ki}^l = 1$, deste modo $u_{ki}^l = 1 - \sum u_{ki}^l$.

Definir $u_{ki}^l = 0$ para $i \in c - I$

**5)** Calcular $\Delta = \left\| U^l - U^{l-1} \right\| = máx_{ji} \left| u_{ji}^l - u_{ji}^{l-1} \right|$, $j = 1,...,n$ e $i = 1,...,c$.

**6)** Se $\Delta > \varepsilon$ e $l < lmáx$, $l = l + 1$ e retornar para o passo 2. Senão parar.

**FIG. 4.1 – Algoritmo do agrupamento de dados nebuloso *c-means***

Nesta dissertação um método de agrupamento de dados nebuloso *c-means* foi empregado para encontrar exatamente a dimensão de imersão ótima (Jiang e Adeli, 2003) dos sistemas estudados e, conseqüentemente, a quantidade de partições do espaço de entrada, ou seja, a quantidade de funções de pertinência que compõem o antecedente das regras de inferência nebulosas.

Para agrupar uma dada série temporal $y(i)$, uma matriz com padrão dos dados,

$$Y = \{y(1), y(2),..., y(Na)\}^T , \quad 4.1$$

de dimensão $Na \times m$, foi criada para representar o espaço de estados reconstruído, onde $y(n)$ é o vetor de espaço de estados definido pela equação:

$$y(n) = [y(n), y(n-\tau), y(n-2\tau),..., y(n-(m-1)\tau)]^T , \quad 4.2$$

onde $Na$ é o número máximo de pontos do espaço de estados definido por:

$$n < Na - \tau(m-1), \; n \in Z \quad 4.3$$

e $m$ é a variável da dimensão de imersão. Foi criada uma matriz, conforme descrito abaixo:

$$Z = \{z_1, z_2,..., z_c\}, \quad 4.4$$

de dimensão $m \times c$, onde $c$, ao contrário das aproximações por tentativa e erro comumente utilizadas na seleção do número de agrupamentos, neste trabalho corresponde ao passo de reconstrução, $\tau$, encontrado através do método da informação mútua. Os vetores $z_i$ representam as classes ou aglomerações em $Y$.

A técnica de agrupamento nebuloso pode ser definida como uma técnica de otimização forçada, ou seja, minimiza-se

$$f_\beta(z) = \sum_{i=1}^{Na} \sum_{j=1}^{c} A_{ij}^{\beta} \left\| y_i - z_j \right\|^2 \qquad 4.5$$

com $\sum_{j=1}^{c} A_{ij} = 1,\ 1 \le i \le Na$ 4.6

$A_{ij} \ge 0,\ 1 \le i \le Na$ e $1 \le j \le c$, 4.7

onde $f_\beta(z)$ é a função objetivo, $A_{ij}$ é o grau de pertinência do vetor do espaço de estados $i$ na classe $j$, $\|.\|$ denota a norma euclideana e o parâmetro $\beta$ representa o grau de nebulosidade (*fuzziness*) dos dados. Freqüentemente é escolhido um valor de $\beta$ dentro do limite $1 < \beta \le 2$, onde os maiores valores são selecionados quando os dados são nebulosos (Jiang e Adeli, 2003).

A formulação descrita acima pelas EQ. 4.5, EQ. 4.6 e EQ. 4.7 remete a um problema de otimização não convexa, que não produz sempre uma solução ótima. Para solucionar esse problema foi utilizado um procedimento iterativo, onde o grau de pertinência do vetor do espaço de estados $y_i$ foi expresso em termos da distância euclideana dos centros das classes, conforme podemos verificar a seguir:

$$A_{ij}^{t+1} = \left[ \sum_{k=1}^{c} \left( \frac{\left\| y_i - z_j^t \right\|^2}{\left\| y_i - z_k^t \right\|^2} \right)^{\frac{1}{\beta-1}} \right]^{-1},\ 1 \le i \le Na \text{ e } 1 \le j \le c, \qquad 4.8$$

onde $t$ sobrescrito denota o número da iteração.

Os passos básicos do procedimento iterativo de agrupamento nebuloso c-means proposto, são os seguintes:

**1)** Inicializar os vetores $z$ aleatoriamente.

**2)** Encontrar a matriz de pertinências $A_{ij}$ através da EQ. 4.8.

**3)** Calcular a função objetivo utilizando a EQ. 4.5.

**4)** Atualizar os centros das classes para a matriz de padrão dos dados $Y$ utilizando a EQ. 5.9 abaixo:

$$z_j^t = \frac{\sum_{i=1}^{n} A_{ij}^m y_i}{\sum_{i=1}^{n} A_{ij}^m}. \qquad 4.9$$

**5)** Atualizar a função objetivo substituindo a EQ. 4.9 na EQ. 4.5.

**6)** Verificar o critério de convergência utilizando uma tolerância dada por $\varepsilon$, conforme descrito abaixo na EQ. 4.10:

$$\left| f_\beta\left(z^t\right) - f_\beta\left(z^{t-1}\right) \right| < \varepsilon,\ 1 \le j \le c \qquad 4.10$$

e um dado número máximo de iterações $Nmáx$.

**7)** Se os critérios de convergência não são satisfeitos, retornar a etapa 2.

Segundo Jiang e Adeli, as etapas a seguir descrevem o método de agrupamento de dados nebuloso *c-means* proposto para identificar a dimensão de imersão ótima (Jiang e Adeli, 2003):

**1)** Utilizar o método da informação mútua para determinar o passo de reconstrução.

**2)** Assumir um valor inicial para a dimensão de imersão $m$.

**3)** Reconstruir o espaço de estados utilizando a EQ. 4.1.

**4)** Agrupar os vetores do espaço de estados reconstruído utilizando o processo iterativo de agrupamento de dados nebuloso *c-means* descrito acima.

**5)** Traçar a curva de convergência, verificando o valor mínimo da função objetivo do método de agrupamento de dados nebuloso *c-means* e o número de iterações correspondente.

**6)** Aumentar a dimensão de imersão e repetir as etapas 3, 4 e 5.

**7)** A curva de convergência com o número mínimo de iterações, ou seja, a curva de convergência mais rápida indica a dimensão de imersão ótima.

## 4.2 AGRUPAMENTO DE DADOS SUBTRATIVO

O agrupamento de dados subtrativo é, essencialmente, uma forma modificada do método da montanha (Eftekhari e Katebi, 2007). No método da montanha os centros iniciais são formados pelas interseções do *grid*, normalmente igualmente espaçado, realizado no espaço de entrada. Define-se então a função de montanha associada a cada grupo e qual o ponto com maior grau de pertinência para ser o primeiro centro. A partir daí a função de montanha associada ao centro $c_1$ é "eliminada" e inicia-se a busca por um novo centro $c_2$. Já no agrupamento subtrativo os *clusters* iniciais são obtidos entre os próprios pontos.

Quando não se conhece "a priori" quantos agrupamentos deve haver para um determinado conjunto de dados, o agrupamento subtrativo é um algoritmo rápido e robusto para saber este número. Esta técnica permite também a localização dos centros de agrupamentos de um conjunto de dados, que é o centro da função de pertinência e, a partir destas, é possível obter as regras nebulosas.

Nesta dissertação o método de agrupamento de dados subtrativo foi proposto para determinar a quantidade de regras que vão compor o sistema de inferência nebuloso proposto. O fluxograma da FIG. 4.2 ilustra o funcionamento do algoritmo do agrupamento de dados subtrativo.

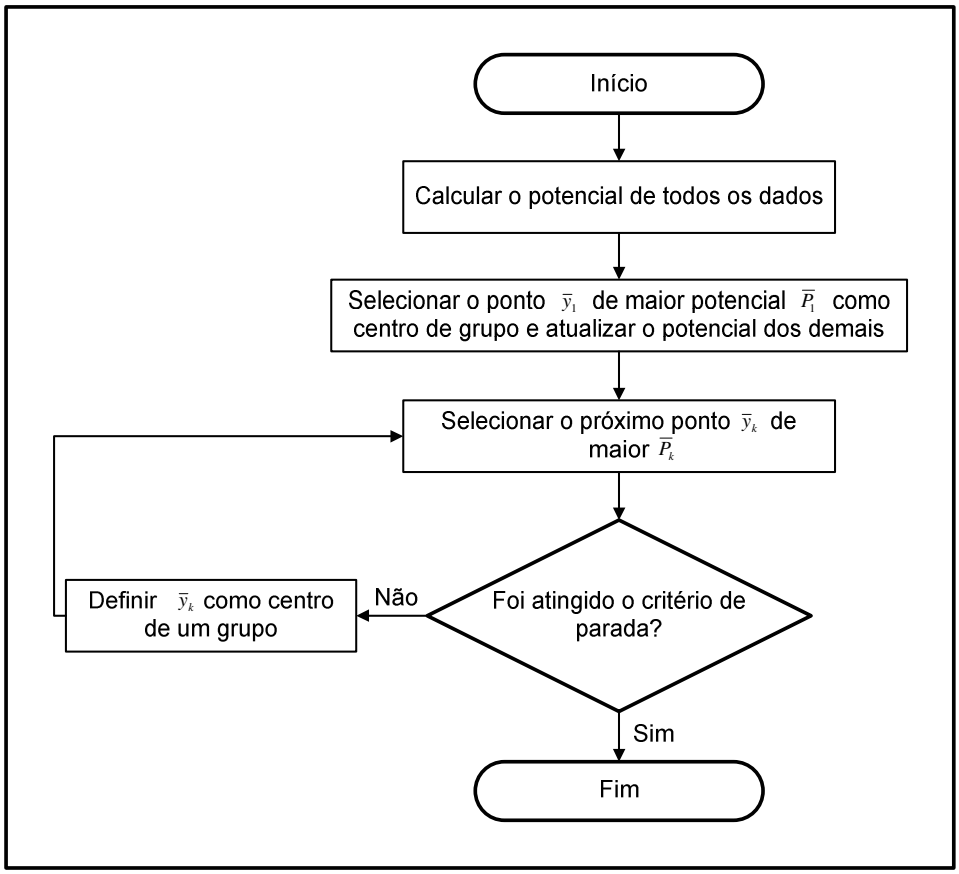


**FIG. 4.2 – Algoritmo do agrupamento de dados subtrativo**

Inicialmente cada ponto $y_i$ é considerado como sendo um possível centro de um grupo e tem seu potencial (medida da possibilidade dele ser centro) calculado em relação aos demais $y_j$ por meio da EQ. 4.11:

$$P_i = \sum_{j=1}^{n} e^{-\alpha \left\| y_i - y_j \right\|^2}, \qquad 4.11$$

onde $n$ é o número de pontos e com:

$$\alpha = 4 / r_a^2 \text{, com } r_a > 0. \qquad 4.12$$

O parâmetro $r_a$ é considerado o raio do grupo. Como o potencial de um ponto é função da sua distância a todos os demais pontos, aquele que apresentar o maior

número de vizinhos dentro do raio $r_a$ terá o maior potencial e será considerado efetivamente o centro do primeiro grupo, passando a ser denominado de $\bar{y}_1$ e seu potencial de $\overline{P}_1$. Neste momento, o potencial do próximo ponto é subtraído com base na sua distância do primeiro centro por meio da EQ. 4.13:

$$P_i = P_i - \overline{P}_1 e^{-\beta \|y_i - \bar{y}_1\|^2}, \qquad 4.13$$

$$\beta = 4 / r_b^2 \text{, com } r_b > 0. \qquad 4.14$$

O parâmetro $r_b$ é uma constante positiva. Somente os pontos dentro da área definida por este parâmetro é que sofrerão redução no seu potencial com relação ao primeiro centro do grupo. O valor de $r_b$ deve ser igual ou maior que o valor de $r_a$ para se evitar centros de grupos muito próximos. Em (Eftekhari e Katebi, 2007), é recomendada a escolha de $r_b = 1.25 r_a$.

Após o potencial de todos os pontos ser sido revisado com o uso da EQ. 4.13, o ponto que possuir o maior potencial é definido como sendo o centro do segundo grupo. Então são atualizados novamente os potenciais dos pontos que se encontram dentro do raio do segundo grupo com a mesma EQ. 4.13. Estes procedimentos são realizados iterativamente até que o critério de parada seja satisfeito.

# 5 ALGORITMO GENÉTICO

A primeira teoria sobre evolução das espécies foi proposta em 1809, pelo naturalista francês Jean Baptiste Pierre Antoine de Monet, conhecido como Lamarck. Para Lamarck as características as características que um animal adquire durante sua vida podem ser transmitidas hereditariamente, este estudo ficou conhecido pela ciência como a “Lei do Uso e Desuso”.

Charles Darwin, em 1858, vem debater a teoria de Lamack de forma agressiva, tentando de forma científica explicar como as espécies evoluem. De acordo com a teoria de Darwin, o princípio de seleção privilegia os indivíduos mais aptos com maior longevidade e, portanto, com maior probabilidade de reprodução. Segundo Darwin, isso ocorre porque indivíduos com mais descendentes têm mais chance de perpetuarem seus códigos genéticos nas próximas gerações.

A técnica do algoritmo genético foi proposta por John H. Holland, seus alunos e colegas da Universidade de Michigan, em meados das décadas de 60 e 70, inspirada nas teorias darwinianas. Seu objetivo ao desenvolver os algoritmos genéticos foi estudar formalmente o fenômeno de adaptação como ocorre na natureza e desenvolver modelos em que os mecanismos da adaptação natural pudessem ser importados para os sistemas computacionais.

Os algoritmos genéticos são algoritmos de busca e otimização baseados na analogia com os princípios de seleção natural, teoria evolutiva e genética. Eles utilizam uma forma simplificada do princípio da sobrevivência do mais apto e uma estratégia de busca paralela e estruturada, porém aleatória, entre soluções candidatas, que é voltada em direção ao reforço em busca de pontos de alta aptidão, ou seja, pontos nos quais a função a ser minimizada (ou maximizada) tem valores relativamente baixos (ou altos). É freqüentemente descrito como um método de busca global, não utilizando gradiente de informação e podendo ser combinado

com outros métodos para refinamento de buscas quando há aproximação de um máximo ou mínimo local. Um aspecto essencial a ser levado em conta quando se propõem soluções de problemas complexos com base em algoritmos genéticos trata da característica de propósito geral associada a esta área da computação. Outro aspecto importante é a eficiência no processo de busca propiciada pelo paralelismo inerente às buscas baseadas em algoritmo genético (Delgado, 2002). O algoritmo genético é considerado um método robusto aplicado em problemas complexos de otimização, tais como problemas com diversos parâmetros ou características que precisam ser combinadas em busca da melhor solução, problemas com muitas restrições ou condições que não podem ser representadas matematicamente e problemas com grandes espaços de busca.

A técnica do algoritmo genético consiste na simulação da evolução de estruturas individuais, via processo de seleção e os operadores de busca, também chamados de operados genéticos, tais como mutação e cruzamento. Todo o processo depende do grau de adaptação, ou seja, da aptidão do indivíduo frente ao ambiente. A seleção, inspirada na seleção natural das espécies, preconiza que os indivíduos mais aptos ou com melhor grau de adaptação ao meio terão maiores chances de repassas seu material genético as próximas gerações.

Segundo (Pires, 2004), para implementação de um algoritmo genético deve-se primeiramente definir alguns aspectos importantes:

- uma representação genética para as soluções do problema. A representação binária é tradicionalmente usada em algoritmos genéticos uma vez que é de fácil utilização e manipulação, além de simples de analisar teoricamente. Contudo, apresenta algumas desvantagens quando aplicada a problemas multidimensionais e a problemas numéricos de alta precisão;

- uma forma de criar uma população inicial das soluções. A escolha de uma população inicial maior que a população a ser utilizada nas gerações subseqüentes pode melhorar a representação do estado de busca. A população inicial pode ser gerada de várias maneiras. Se uma população inicial pequena for gerada

aleatoriamente, provavelmente algumas regiões do espaço de busca não serão representadas;

- uma função de avaliação que desempenha o papel do ambiente, ou seja, avalia as soluções em termos de *fitness*;

- operadores genéticos que alteram a composição dos filhos. O operador de cruzamento tem o intuito de trocar informações entre diferentes soluções em potencial. Já a aplicabilidade do operador de mutação, tem por objetivo introduzir uma variabilidade extra na população;

- valores para os parâmetros utilizados pelo algoritmo genético. A influência de cada parâmetro no desempenho do algoritmo depende da classe de problemas que se está tratando. Assim, a determinação de um conjunto de valores otimizado para estes parâmetros dependerá da realização de um grande número de experimentos e testes. Existem estudos que utilizam o algoritmo genético como método de otimização para escolha dos parâmetros de outro algoritmo genético.

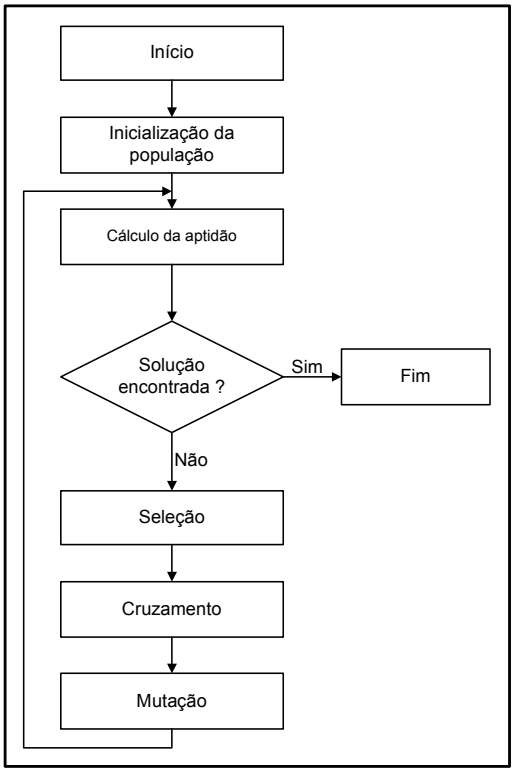


**FIG. 5.1 – Estrutura básica do algoritmo genético**

O funcionamento básico de um algoritmo genético clássico pode ser descrito pelo diagrama da FIG. 5.1. Observando esse diagrama é possível verificar que cada iteração do algoritmo genético corresponde, basicamente, à aplicação do conjunto de quatro operações: cálculo de aptidão, seleção, cruzamento e mutação. Ao fim destas operações uma nova população é criada, esperando que esta represente uma melhor aproximação da solução do problema que a população anterior. A população inicial é gerada atribuindo aleatoriamente valores aos genes de cada cromossomo.

## 5.1 MÉTODOS DE SELEÇÃO

O princípio básico do funcionamento dos algoritmos genéticos é que um critério de seleção vai fazer com que, depois de muitas gerações, o conjunto inicial de indivíduos gere indivíduos mais aptos. A idéia principal da seleção é oferecer aos melhores indivíduos da população corrente preferência para o processo de reprodução, permitindo que estes indivíduos possam passar as suas características às próximas gerações. Isto funciona como na natureza onde os indivíduos altamente adaptados ao seu ambiente possuem naturalmente mais oportunidades para reproduzir do que aqueles indivíduos considerados mais fracos. Métodos de seleção muito fortes tendem a gerar super-indivíduos (indivíduos com medida de desempenho muito superior aos demais), reduzindo a diversidade necessária para alterações e progressos futuros. A geração de super-indivíduos pode levar a uma convergência prematura do processo evolutivo. Por outro lado, métodos de seleção excessivamente fracos (pouca pressão seletiva) tendem a produzir progressos muito lentos na evolução (Delgado, 2002). A maioria dos métodos de seleção são projetados para escolher preferencialmente indivíduos com maiores notas de aptidão, embora não exclusivamente, a fim de manter a diversidade da população. A seguir veremos alguns destes métodos, com base em (Pires, 2004).

### 5.1.1 SELEÇÃO PELA ROLETA

O método de seleção da roleta é um método de seleção muito utilizado, onde os indivíduos de uma geração são escolhidos para fazer parte da população intermediária através de um sorteio de roleta. Neste método cada indivíduo da população é representado na roleta proporcionalmente ao seu índice de aptidão, como é mostrado na FIG. 5.2. Assim, aos indivíduos com alta aptidão é dada uma porção maior da roleta, enquanto aos de aptidão mais baixa é dada uma porção relativamente menor da roleta. Finalmente, a roleta é girada um determinado número de vezes dependendo do tamanho da população e são escolhidos como indivíduos que participarão da população intermediária aqueles sorteados na roleta.

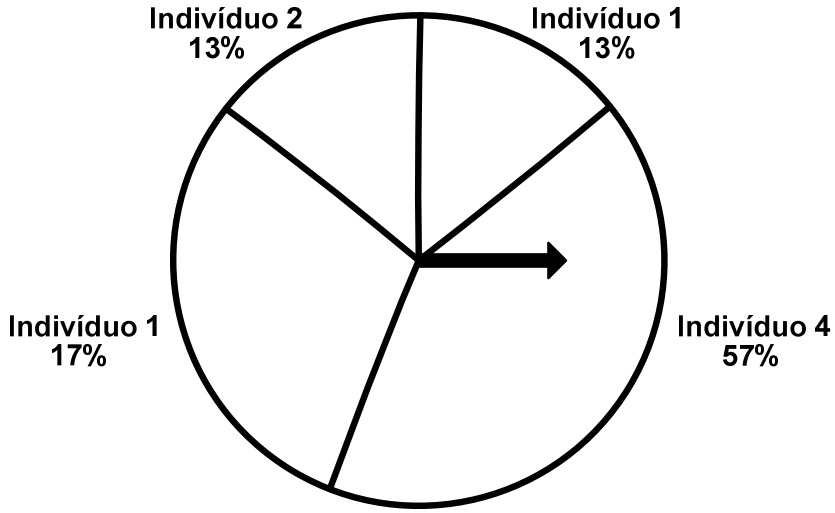


**FIG. 5.2 – Exemplo de uma roleta de seleção**

### 5.1.2 SELEÇÃO POR AMOSTRAGEM UNIVERSAL ESTOCÁSTICA

Para visualizar esse método, considere um círculo em $n$ regiões (tamanho da população), onde a área de cada região é proporcional à aptidão do indivíduo (FIG. 5.2). Coloca-se sobre este círculo uma "roleta" com $n$ cursores, igualmente espaçados. Após um giro da roleta a posição dos cursores indica os indivíduos selecionados. Evidentemente, os indivíduos cujas regiões possuem maior área, terão maior probabilidade de serem selecionados várias vezes. Como conseqüência,

a seleção de indivíduos pode conter várias cópias de um mesmo indivíduo enquanto outros podem desaparecer. Este tipo de seleção é mais rápido que a roleta, pois em uma só rodada já são selecionados todos os indivíduos.

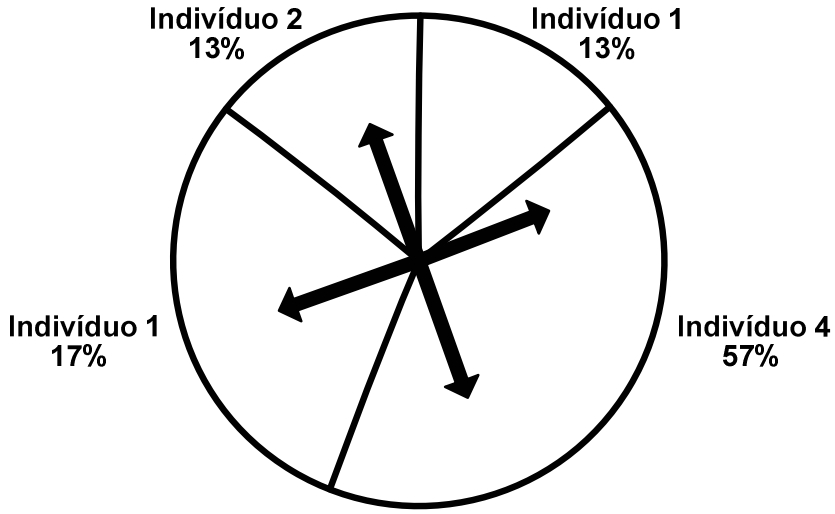


**FIG. 5.3 – Seleção por amostragem universal estocástica**

### 5.1.3 SELEÇÃO POR TORNEIO

No método de seleção por torneio são escolhidos aleatoriamente (com probabilidades iguais) $n$ indivíduos da população e o indivíduo que possuir maior aptidão é selecionado para a população intermediária. O processo repete-se até completar a população intermediária.

Existem variações deste algoritmo nas quais a escolha dentre os indivíduos candidatos a possuir maior aptidão é determinística (escolhe-se sempre o melhor) (Delgado, 2002).

### 5.1.4 SELEÇÃO POR TRUNCAMENTO

O método de seleção por truncamento tem como base um valor $T$ entre zero e um. A seleção é feita aleatoriamente entre os $T$ melhores indivíduos. Por exemplo,

$T = 0{,}4$, então a seleção é feita entre os 40% melhores indivíduos e os outros 60% são descartados.

### 5.1.5 SELEÇÃO ELITISTA OU ELITISMO

A seleção elitista ou elitismo, introduzido por Kenneth De Jong em 1975, é um método de seleção que força o algoritmo genético a reter os $r$ melhores indivíduos de uma geração para geração seguinte. Estes indivíduos poderiam ser perdidos caso não fossem reproduzidos de forma determinística para a próxima geração, ou se sofressem a ação dos operadores de cruzamento e mutação. Geralmente, estratégias elitistas associadas aos métodos de seleção melhoram o desempenho de um algoritmo genético.

### 5.1.6 OUTROS MÉTODOS

Segundo (Delgado, 2002), há vários outros métodos (determinísticos ou não) propostos para a implementação do processo de seleção como, por exemplo, a seleção puramente aleatória, na qual todos os indivíduos têm a mesma probabilidade de serem escolhidos. Normalmente, este método é utilizado em conjunto com estratégias elitistas. Outro exemplo é o método de seleção *steady-state*, que mantém a população original de uma geração para outra, com exceção de poucos indivíduos que são substituídos por descendentes do melhor, obtidos por mutação ou por cruzamento. Neste método, a escolha da população inicial tem papel muito importante, pois a evolução é mais lenta e obtida com pequenas modificações a cada geração. Na direção contrária, aparece a seleção por diversidade, em que parte da população é obtida escolhendo-se os indivíduos mais diversos, a partir do melhor indivíduo, à custa do uso de recursos computacionais adicionais. O conceito de mais diverso é baseado em algum critério de distância previamente definido.

## 5.2 OPERADORES GENÉTICOS

O princípio básico dos operadores genéticos é transformar a população através de sucessivas gerações, estendendo a busca até chegar a um resultado satisfatório. Os operadores genéticos são necessários para que a população se diversifique e mantenha características de adaptação adquiridas pelas gerações anteriores.

Os principais operadores genéticos são a mutação e o cruzamento (inspirado da biologia evolutiva) ou recombinação. Uma distinção entre estes dois operadores pode ser feita de acordo com o número de indivíduos utilizados como entradas do operador. O operador de mutação utiliza como operando um único indivíduo e produz um único indivíduo em sua saída. O operador de cruzamento, em contrapartida, atua em mais de um indivíduo e produz, geralmente, mais de um indivíduo em sua saída. Em geral, operadores de cruzamento utilizam dois indivíduos como “pais” (Maruo, 2006).

A escolha dos operadores genéticos está intimamente ligada à codificação adotada para a representação genética (Delgado, 2002), ou seja, existe uma variação nos comportamentos dos operadores, os quais estão relacionados à codificação empregada nos cromossomos. Em outras palavras, há operadores para o uso com codificação binária e operadores genéticos para o uso com codificação real ou inteira (Pires, 2004).

Algumas abordagens para problemas com restrições apresentadas na literatura sugerem o uso de algoritmos reparadores associados aos operadores genéticos. Estes algoritmos modificam a codificação genética resultante do *crossover* ou da mutação de forma que as novas soluções obtidas não violem as restrições impostas às soluções do problema (Delgado, 2002).

Nesse trabalho, somente operadores para uso com codificação binária foram considerados.

### 5.2.1 CRUZAMENTO

O cruzamento é o operador responsável pela recombinação de características dos pais durante a reprodução, permitindo que as próximas gerações herdem essas

características. Esta fase é marcada pela troca de segmentos entre "casais" de cromossomos selecionados para dar origem a novos indivíduos que formarão a população da próxima geração. A idéia central do cruzamento é a propagação das características dos indivíduos mais aptos da população por meio de troca de segmentos informações entre os mesmos, o que dará origem a novos indivíduos.

A seguir estão descritos os três tipos de cruzamento para cromossomos com codificação binária, com base em (Pires, 2004).

#### 5.2.1.1 CRUZAMENTO MONO-PONTO

O cruzamento mono-ponto é a forma mais simples deste operador. É aplicado a um par de cromossomos retirados da população intermediaria, gerando dois cromossomos filhos. Cada um dos cromossomos pais é cortado em uma posição aleatória produzindo duas cabeças e duas caudas. As caudas são trocadas gerando dois novos cromossomos. A FIG. 5.4 ilustra o comportamento deste operador.

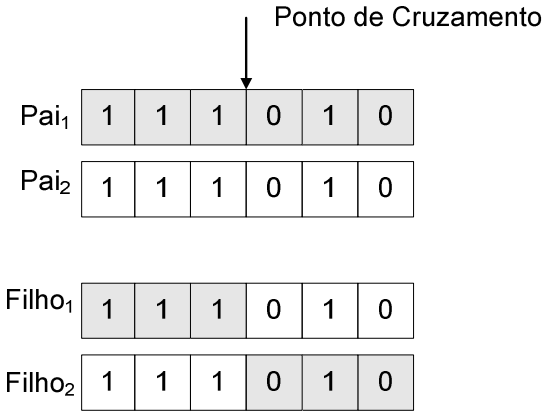


**FIG. 5.4 – Cruzamento mono-ponto**

#### 5.2.1.2 CRUZAMENTO DE N-PONTOS

O cruzamento de n-pontos funciona da mesma forma que o cruzamento mono-ponto, porém, neste há vários pontos de corte nos cromossomos pais. Na FIG. 5.5 é exemplificado o cruzamento de três-pontos.

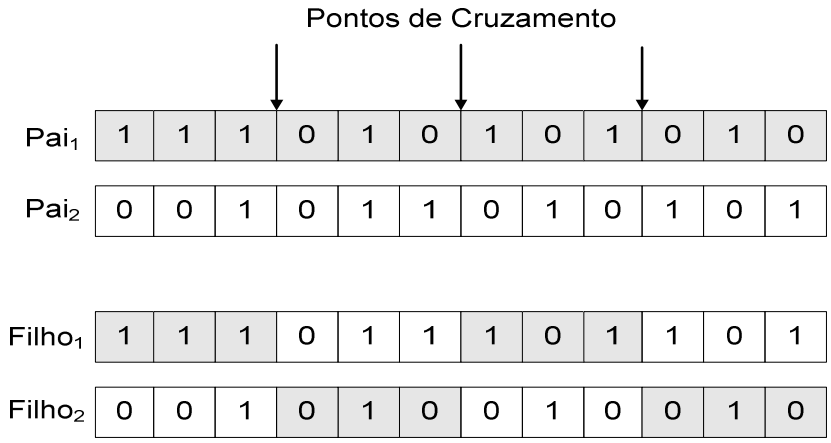


**FIG. 5.5 – Cruzamento de três-pontos**

5.2.1.3 CRUZAMENTO UNIFORME

No cruzamento uniforme cada gene do cromossomo filho é criado copiando o gene correspondente de um dos pais, escolhido de acordo com uma máscara de cruzamento gerada aleatoriamente. Onde houver uma máscara de cruzamento, o gene correspondente será copiado do primeiro pai e onde houver zero será copiado do segundo. O processo é repetido com os pais trocados para produzir o segundo descendente. Uma nova máscara de cruzamento é criada para cada par de pais. O número não é fixo, mas em geral é usado $L/2$ (onde $L$ é o comprimento do cromossomo). Na FIG. 5.6 o processo é ilustrado graficamente.

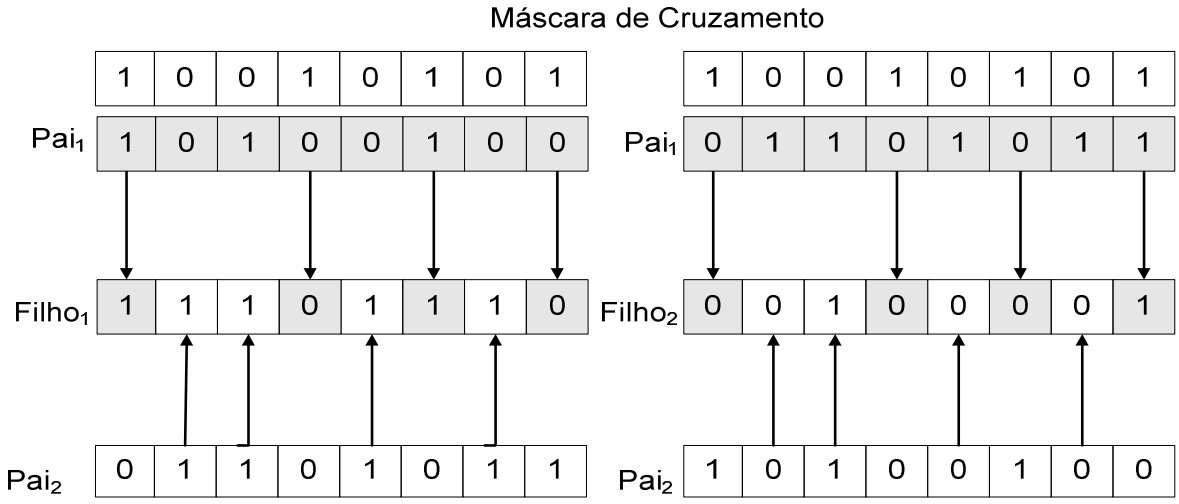


**FIG. 5.6 – Cruzamento uniforme**

Para uma combinação de esquemas com variáveis relacionadas próximas no genótipo, é provável que um cruzamento de múltiplos pontos apresente um desempenho superior. Entretanto, na ausência de conhecimento sobre a dependência entre as variáveis do problema, o uso de esquemas com grande distância entre os genes apresenta uma probabilidade maior de quebra devido ao cruzamento levando a um desempenho sub-ótimo. Um cruzamento uniforme costuma ser mais recomendável nesses casos (Maruo, 2006). Não existe nenhum impedimento quanto ao uso destes operadores junto a outros tipos de codificação. Mas, para problemas com restrições, existem operadores de cruzamento específicos, que garantem a factibilidade dos descendentes originários de indivíduos factíveis (Delgado, 2002).

### 5.2.2 MUTAÇÃO

A mutação é um processo de busca aleatória e é necessária porque ocasionalmente a seleção natural e o cruzamento podem eliminar material genético potencialmente promissor da população antes que ele seja avaliado adequadamente. Nos algoritmos genéticos a mutação evita este tipo de perda através de uma alteração aleatória do valor codificado em um determinado *locus gênico* (Goldberg et al., 1989).

A mutação é necessária para a introdução e manutenção da diversidade genética da população, alterando arbitrariamente um ou mais indivíduos, fornecendo assim, meios para introdução de novos indivíduos na população. Desta forma, a mutação assegura que a probabilidade de se chegar a qualquer ponto do espaço de busca nunca será zero, além de contornar o problema de mínimos locais, pois com este mecanismo, altera-se levemente a direção da busca. Geralmente se utiliza uma taxa de mutação pequena, pois se trata de um operador genético secundário. O operador de mutação é aplicado aos indivíduos com uma probabilidade dada pela taxa de mutação.

A seguir está descrito o processo de mutação utilizado para cromossomos com codificação binária, com base em (Pires, 2004).

#### 5.2.2.1 MUTAÇÃO SIMPLES

O operador de mutação simples é aplicado aleatoriamente em alguns genes do cromossomo que sofrerá a mutação. Este processo inverte os valores dos genes, ou seja, muda o valor de um dado gene de ‘1’ para ‘0’ ou de ‘0’ para ‘1’. Na FIG. 5.7 é ilustrado o comportamento deste operador.

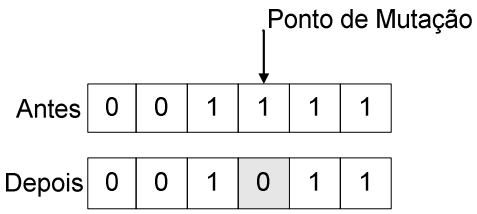


**FIG. 5.7 – Mutação simples**

## 5.3. PARÂMETROS GENÉTICOS

Além da forma como o cromossomo é codificado, existem vários parâmetros do algoritmo genético que podem ser escolhidos para melhorar o seu desempenho, adaptando-o às características particulares de determinadas classes de problemas. Entre eles os mais importantes são:

- **tamanho da população:** o tamanho da população afeta o desempenho global e a eficiência do algoritmo genético. Com uma população pequena o desempenho pode cair, pois deste modo a população fornece uma pequena cobertura do espaço de busca do problema e o algoritmo genético terá poucas possibilidades de realizar cruzamentos. Uma grande população geralmente fornece uma cobertura representativa do domínio do problema, além de prevenir convergências prematuras para soluções locais ao invés de globais. No entanto, para se trabalhar

com grandes populações são necessário maiores recursos computacionais ou que o algoritmo trabalhe por um período de tempo muito maior;

- **intervalo de gerações:** controla a porcentagem da população que será substituída durante a próxima geração. Com um valor alto, a maior parte da população será substituída, mas com valores muito altos pode ocorrer perda de estruturas de alta aptidão. Com um valor baixo, o algoritmo pode tornar-se muito lento;

- **taxa de cruzamento:** quanto maior for esta taxa, mais rapidamente novas estruturas serão introduzidas na população. Entretanto, isto pode gerar um efeito indesejado, pois a maior parte da população será substituída podendo ocorrer perda de estruturas de alta aptidão;

- **taxa de mutação:** uma baixa taxa de mutação previne que uma dada posição fique estagnada em um valor, além de possibilitar que se chegue a qualquer ponto do espaço de busca. A mutação não deve ocorrer com muita freqüência senão a busca se torna essencialmente aleatória.

A influência de cada parâmetro no desempenho do algoritmo depende da classe de problemas que se está tratando.

A forma tradicional de ajuste de parâmetros altera os valores de somente um parâmetro por vez, podendo acarretar em escolhas sub-ótimas. Isso ocorre devido à forma complexa com que os parâmetros estratégicos interagem. Em contrapartida, o ajuste de mais parâmetros ao mesmo tempo demanda uma grande quantidade de experimentos (Maruo, 2006).

## 6 EXPERIMENTOS E RESULTADOS

Este capítulo apresenta e discute exemplos de predição com dados de séries temporais obtidas a partir de sistemas dinâmicos não-lineares, onde foram aplicadas ferramentas baseadas na lógica nebulosa, tendo como técnicas de apoio o algoritmo genético e o agrupamento de dados.

O sistema nebuloso escolhido para a modelagem das séries temporais foi o sistema Takagi-Sugeno-Kang de ordem zero, sendo aqui o antecedente e o conseqüente de cada regra de inferência nebulosa compostos, respectivamente, por funções de pertinência triangulares e por funções polinomiais de ordem zero, no formato de uma função impulso unitário. A regra de ordem $i$ é descrita por

$R^i$: SE $x_1$ é aproximadamente $\Gamma_1^i$ ou ... ou $x_p$ é aproximadamente $\Gamma_p^i$

ENTÃO $y^i = w^i$, 6.1

onde $R^i$ com $i \in I_R = \{1,2,...,r\}$ são as $i$ regras de inferência nebulosas; $\Gamma_j^i$ são os conjuntos nebulosos com $j \in I_p = \{1,2,...,p\}$; $x_j$ é a variável de entrada composta pelos espaços de estado reconstruídos, encontrados através do teorema de Takens; $y^i$ é a variável de saída; e $w^i$ é o valor de ativação da regra. Para calcular o valor da pertinência de $x_j$ a $\Gamma_j^i$ utiliza-se funções de pertinência triangulares $\Gamma_j^i : U_{x_j} \subset \Re_{[a_j^i, c_i^j]} \rightarrow \Re_{[0,1]}$, modeladas matematicamente por

$$\Gamma_j^i(x_j) = \max\left\{\min\left\{\frac{x_j - (b_j^i i - a_j^i)}{a_j^i}, \frac{b_j^i + c_j^i - x_j}{c_j^i}\right\}, 0\right\}, \quad 6.2$$

onde $U_{x_j}$ é o universo de discussão de $x_j$ e $\{a_j^i, b_j^i, c_j^i\}$ com $(i, j) \in \mathrm{I}_R \times \mathrm{I}_p$ são os parâmetros da esquerda, central e da direita da EQ. 6.2. O método do centro de gravidade foi utilizado para agregação dos valores encontrados. Neste método calcula-se o centróide da área composta que representa o termo de saída nebuloso, esse termo de saída é composto pela união de todas as contribuições das regras (Shaw et al., 1999). O método é deduzido pela EQ. 6.3 abaixo:

$$\hat{y} = \frac{\sum_i^r \theta_i(x) w^i}{\sum_{i=1}^r \theta_i(x)}, \qquad 6.3$$

onde $\theta_i(x) = \sup[\Gamma_j^i(x_j)]$.

Os espaços de estado reconstruídos, usados como entrada sistema de inferência nebuloso, servirão para treinar o sistema proposto. O passo de reconstrução é dado pelo primeiro mínimo local encontrado através da análise do gráfico com os resultados da informação mútua. Os dados para construção desse gráfico são calculados através do pacote TISEAN – *Nonlinear Time Series Analysis*, versão 3.0.0. Já a dimensão de imersão ótima é calculada através do algoritmo de agrupamento de dados nebuloso *c-means*, conforme descrito no capítulo 4.

A quantidade de particionamentos do espaço de entrada, ou seja, a quantidade de funções de pertinência que irão compor o antecedente de cada regra do sistema de inferência nebuloso proposto será igual à dimensão de imersão.

A quantidade de regras que vão compor o sistema de inferência nebuloso proposto será igual à quantidade de centros encontrados através da aplicação

do algoritmo de agrupamento de dados subtrativo, descrito no capitulo 4, implementado através da função *subclust.m* do Matlab®.

Para otimizar os parâmetros do sistema de inferência nebuloso foi utilizado o algoritmo genético. Primeiramente foi definida a estrutura da cadeia de cromossomos que melhor represente o sistema de inferência nebuloso proposto. Na FIG. 6.1 é possível visualizar uma subcadeia, composta por uma regra, que integra a cadeia de cromossomos. Pode-se visualizar os parâmetros que compõem essa subcadeia através da FIG. 6.2.

No algoritmo genético proposto neste trabalho foi utilizada uma população composta por dez indivíduos, sendo que, cada parâmetro que compõem a cadeia de cromossomos contém cinco bits.

A função objetivo e a função de aptidão foram minimizadas, ao invés de maximizadas, ao contrário do que é adotado na maioria dos algoritmos genéticos convencionais. A função objetivo utilizada foi o erro médio quadrático, conforme descrito pela EQ 6.4 abaixo:

$$f_{objetivo} = \frac{1}{n}\sum_{k=1}^{n}(\hat{y}_k - y_k)^2 , \qquad 6.4$$

onde $y_k$ é a saída real e $\hat{y}_k$ é a saída estimada pelo sistema de inferência nebuloso. A saída estimada pelo sistema de inferência nebuloso proposto vai ser igual à saída predita um passo à frente, ou seja, $y(t+\tau)$. O processo de otimização é finalizado quando a saída real for igual à saída predita estimada pelo sistema proposto.

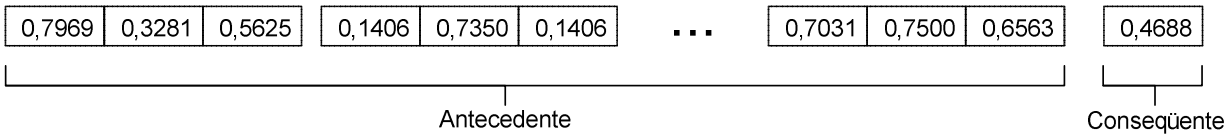


**FIG. 6.1 – Exemplo de uma subcadeia que compõe o cromossomo**

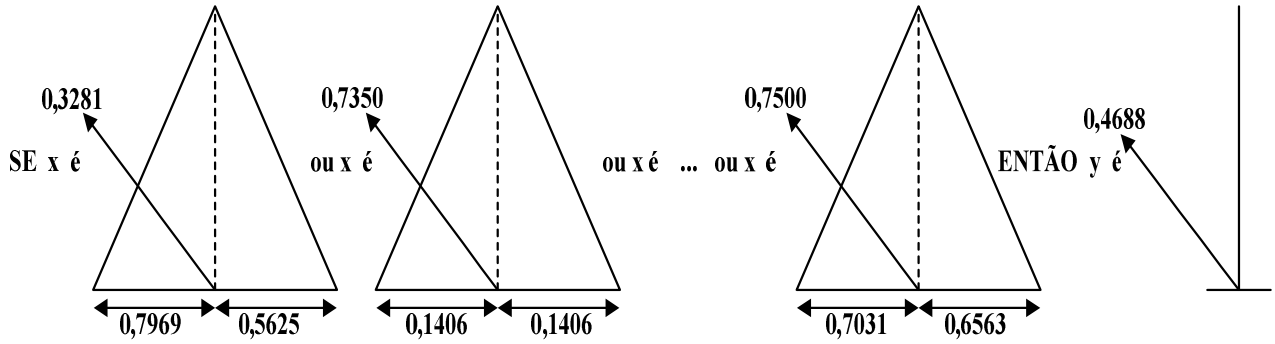


**FIG. 6.2 – Exemplo da formação de uma regra do sistema de inferência nebuloso**

## 6.1 SÉRIE TEMPORAL DE MACKEY GLASS

Um dos casos de estudo mais utilizados na identificação de sistemas consiste na predição da série temporal de Mackey-Glass. A equação de Mackey-Glass é uma equação diferencial com atraso no tempo que descreve um sistema de controle fisiológico (Jang e Sun, 1993, Manguire et al., 1998, Wang et al., 2005, Lee et al., 2006, Gu e Wang, 2007). Ela foi proposta primeiramente como um modelo de produção de glóbulos brancos do corpo humano porque taxas de proliferação da célula envolvem um atraso no tempo, dinâmica periódica e caos podem ser obtidos. Certamente, Mackey e Glass sugeriram que as flutuações de longo prazo nas contagens das células observadas em determinados formulários de leucemia fossem evidenciadas para estes comportamentos (Lee et al., 2006). A série

temporal de Mackey-Glass é gerada através da integração da equação diferencial abaixo:

$$\frac{dy}{dt} = \frac{Ay(t-\xi)}{1+y(t-\xi)^C} - By(t), \qquad 6.5$$

onde $y(t) \in \Re$ e as constantes são comumente escolhidas como sendo $A = 0.2$, $B = 0.1$ e $C = 10$. O parâmetro de atraso $\xi$ determina o comportamento da EQ. 6.5: para $\xi < 4.53$ há um atrator de ponto fixo estável; para $4.53 \le \xi < 13.3$ há um atrator de período limitado estável, período duplicado começa em $\xi = 13.3$ e continua até $\xi = 16.8$; para $\xi > 16.8$ o caos se desenvolve (Lee et al., 2006). A FIG. 6.3 mostra a órbita espacial de $(y(t), y(t-\xi), y(t-2\xi))$, onde $\xi = 17$ e $y(t) = 1.2$ para $t \in \Re_{\le 0}$.

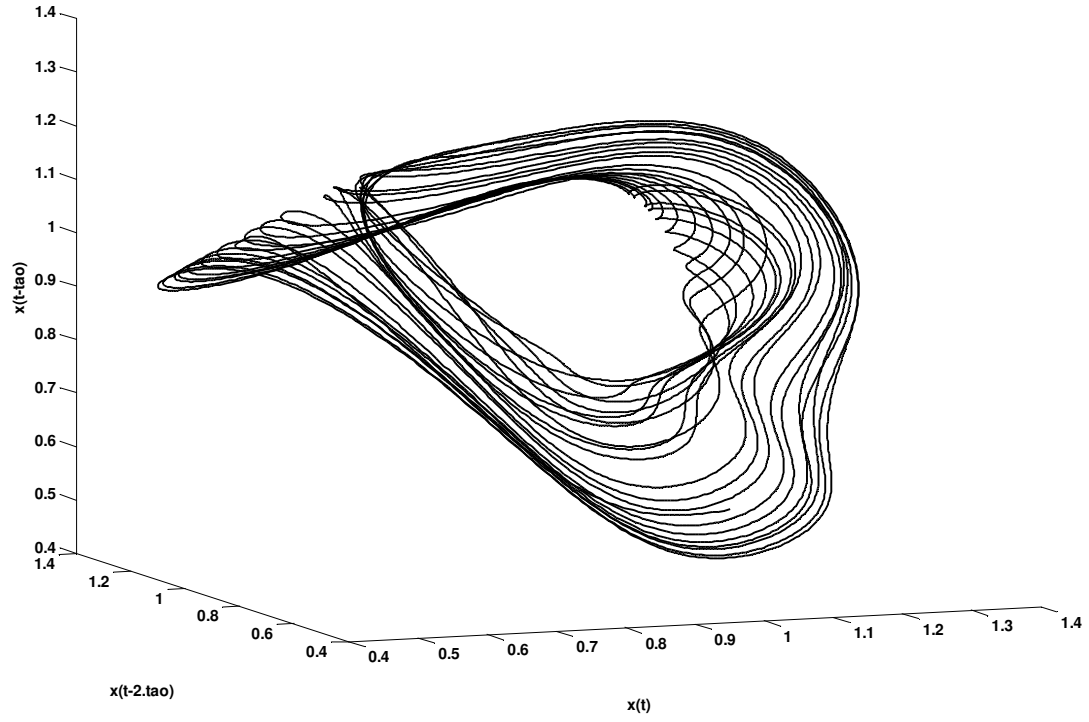


**FIG. 6.3 – Atrator de Mackey-Glass**

As soluções numéricas da forma pontual são geradas com base no algoritmo de Runge-Kutta modificado, implementado através da função *dde23.m* do Matlab® com condição inicial $y(0)=1.2$, função histórico $y(t)=1.2$ para $t \in \Re_{\leq 0}$ e atraso no de tempo $\xi = 17$. Por conveniência todos os dados gerados foram normalizados para estarem entre 0 e 1.

Na FIG. 6.4 (a) está ilustrado o resultado da informação mútua e na FIG. 6.4 (b) está ampliada a região onde se encontra o primeiro mínimo local do gráfico da informação mútua, calculado a partir dos dados da série temporal de Mackey-Glass. Através da análise desses dados, verifica-se que o passo de reconstrução a ser utilizado é 5.

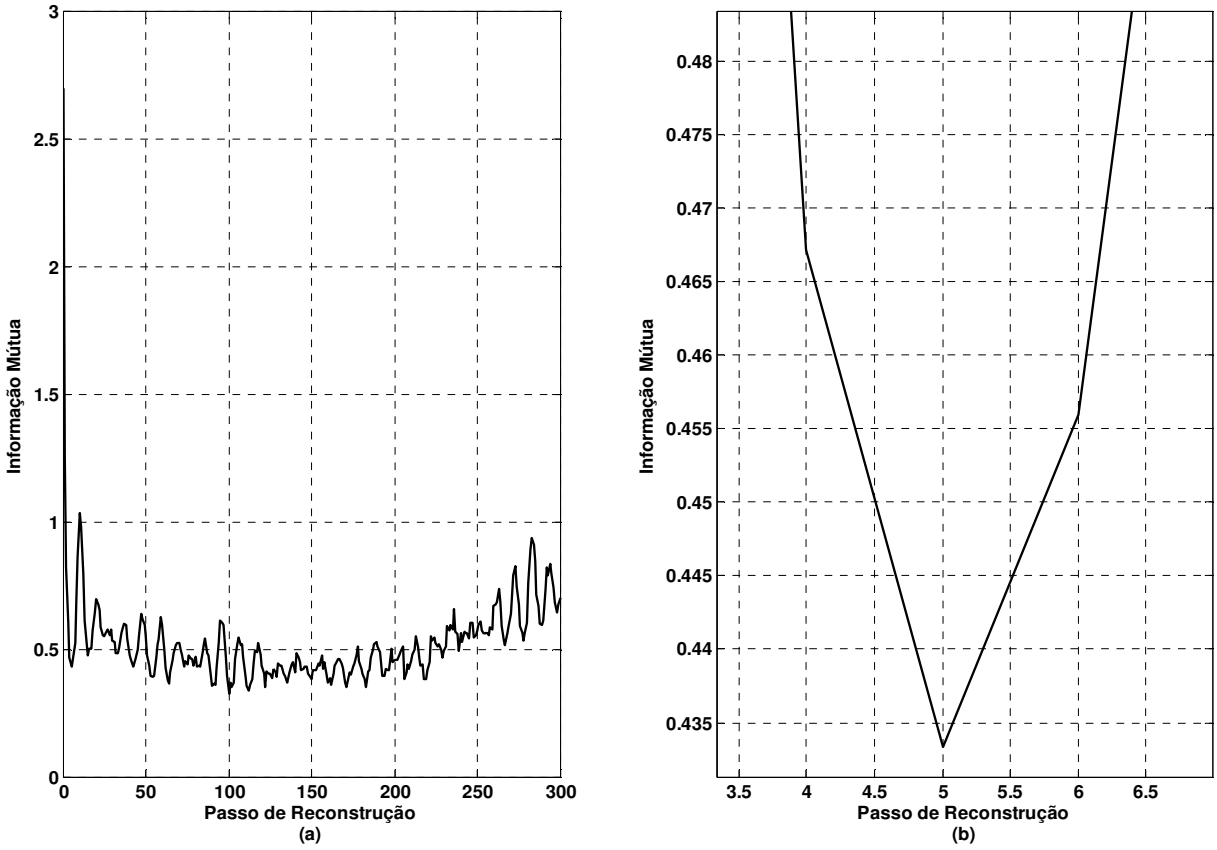


**FIG. 6.4 – (a) Resultado obtido através do método da informação mútua aplicado aos dados da série temporal de Mackey-Glass e (b) Localização do primeiro mínimo local no gráfico da informação mútua**

A FIG. 6.5 ilustra as curvas de convergência encontradas através do agrupamento de dados nebuloso *c-means* e a TAB. 6.1 relaciona a quantidade de iterações necessárias para atingir o critério de convergência estabelecido no capítulo 4, aplicando uma tolerância $\varepsilon = 0.00001$ e um $Nmáx = 600$. O algoritmo de agrupamento nebuloso foi testando, variando as dimensões de imersão de 2 a 10. Analisando a FIG. 6.5 e a TAB. 6.1, verifica-se que a dimensão de imersão para a qual foram necessários o menor número de iterações para se atingir os critérios de convergência estabelecidos foi $m = 3$, ou seja, serão utilizadas 3 funções de pertinência no antecedente de cada regra do sistema de inferência nebuloso proposto.

Aplicando-se os resultados obtidos à rotina *subclust.m* do Matlab®, encontra-se o número de centros do agrupamento de dados subtrativo igual a 7, ou seja, a quantidade de regras utilizadas no sistema de inferência nebulosos proposto é 7.

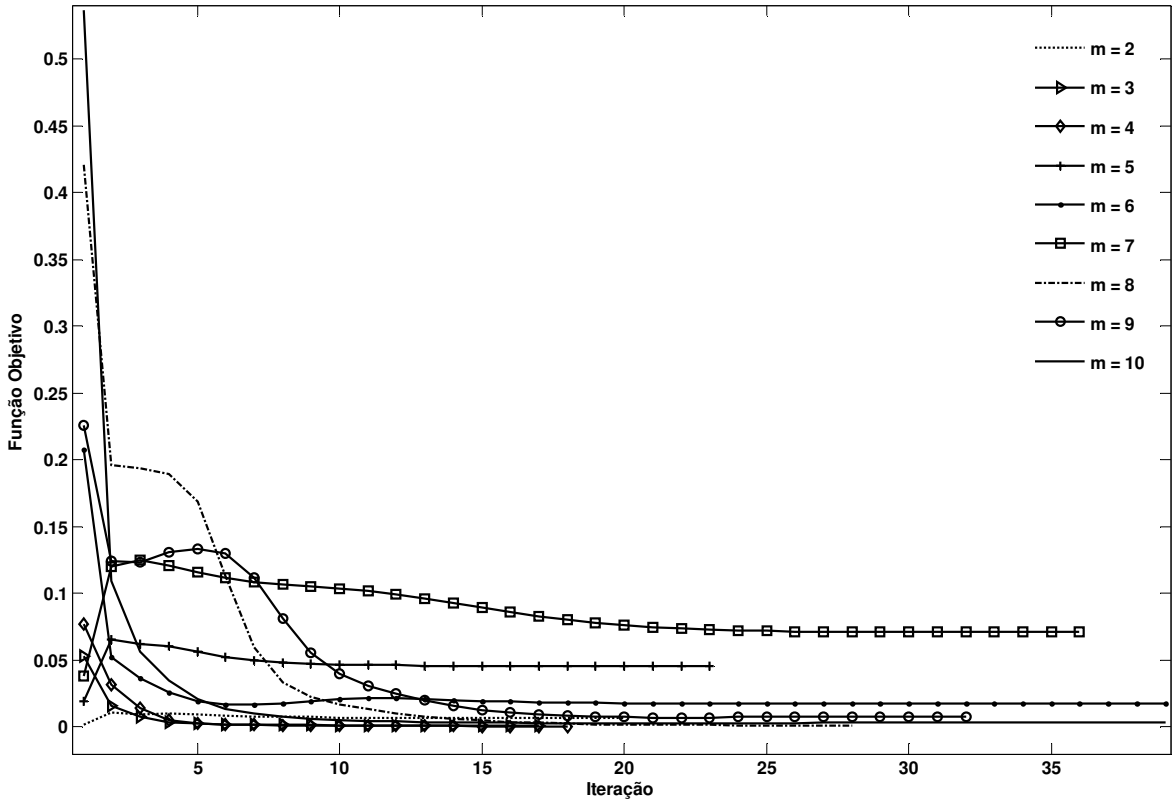


**FIG. 6.5 – Curvas de convergência do agrupamento de dados nebuloso *c-means* para os vetores do espaço de estados reconstruído da série temporal de Mackey-Glass com $\tau = 5$ e $m = 2-10$**

**TAB. 6.1 – Resultados do método do agrupamento de dados nebuloso *c-means* proposto para a série temporal de Mackey-Glass**

| Dimensão de imersão | 2 | **3** | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Valor Médio de Iterações | 18,85 | **16,75** | 17,35 | 21,75 | 46,10 | 30,45 | 25,65 | 34,25 | 39,40 |
| Valor mínimo do agrupamento nebuloso c-means | 0,002 | 0,0012 | 0,0005 | 0,0195 | 0,0168 | 0,0381 | 0,001 | 0,0072 | 0,0029 |

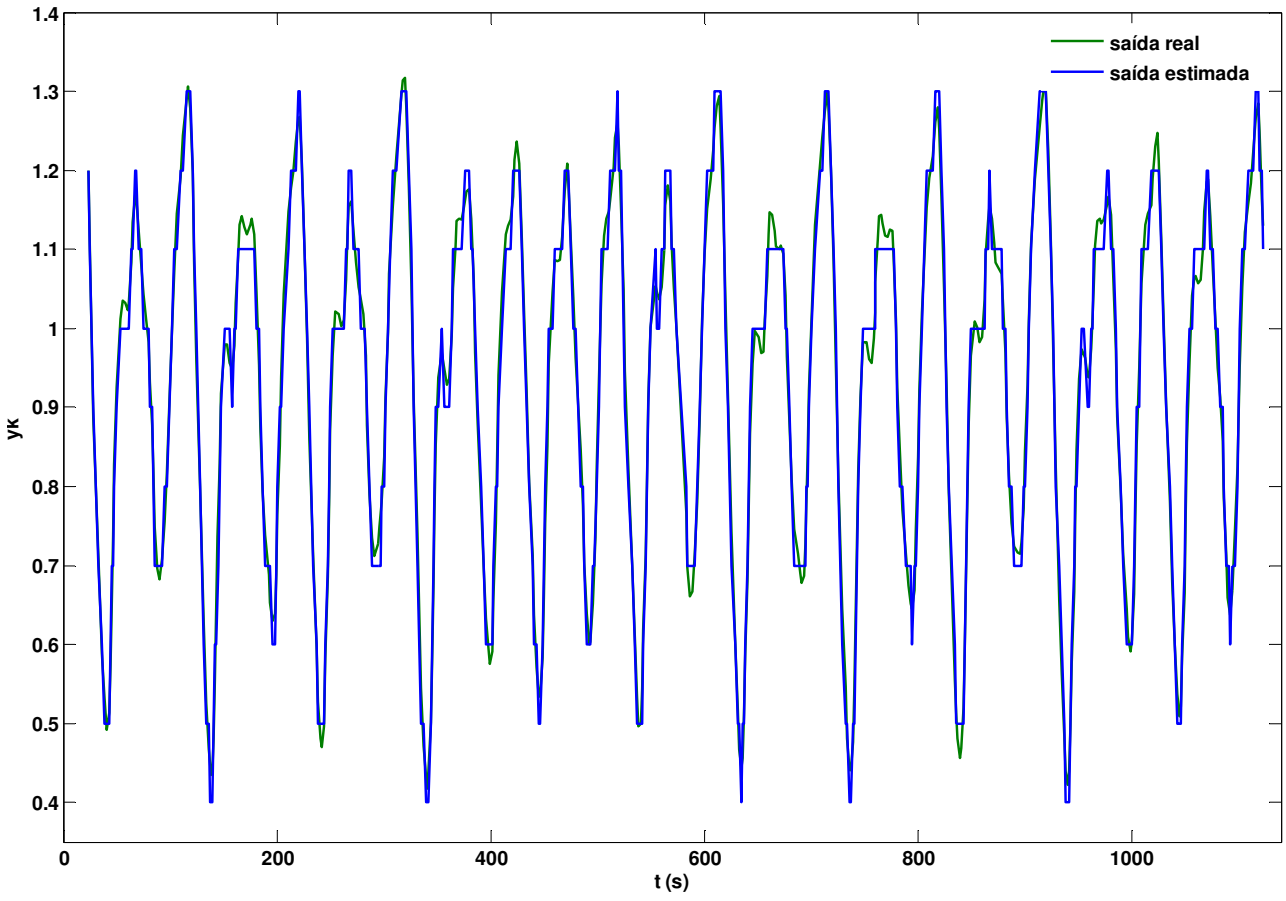


**FIG. 6.6 – Comparação entre a saída da série temporal de Mackey-Glass real (verde) e a estimada através sistema de inferência nebuloso proposto (azul)**

## 6.2 SISTEMA DE LORENZ

As equações de Lorenz foram introduzidas em 1963 como um modelo simples do movimento convectivo nas camadas superiores da atmosfera. Lorenz descobriu que para certos valores dos parâmetros $\sigma$, $\beta$ e $\rho$, da EQ. 6.8, o sistema nunca tende para um comportamento previsível a longo prazo e que, por essa razão, não é possível também fazer previsões do tempo meteorológico a longo prazo. Além disso, as trajetórias deste sistema nunca acabam num ponto fixo nem num ciclo limite estável e, contudo, nunca divergem para o infinito Como apresenta uma grande sensibilidade as condições iniciais tornam-se extremamente imprevisível apesar de ser determinístico.

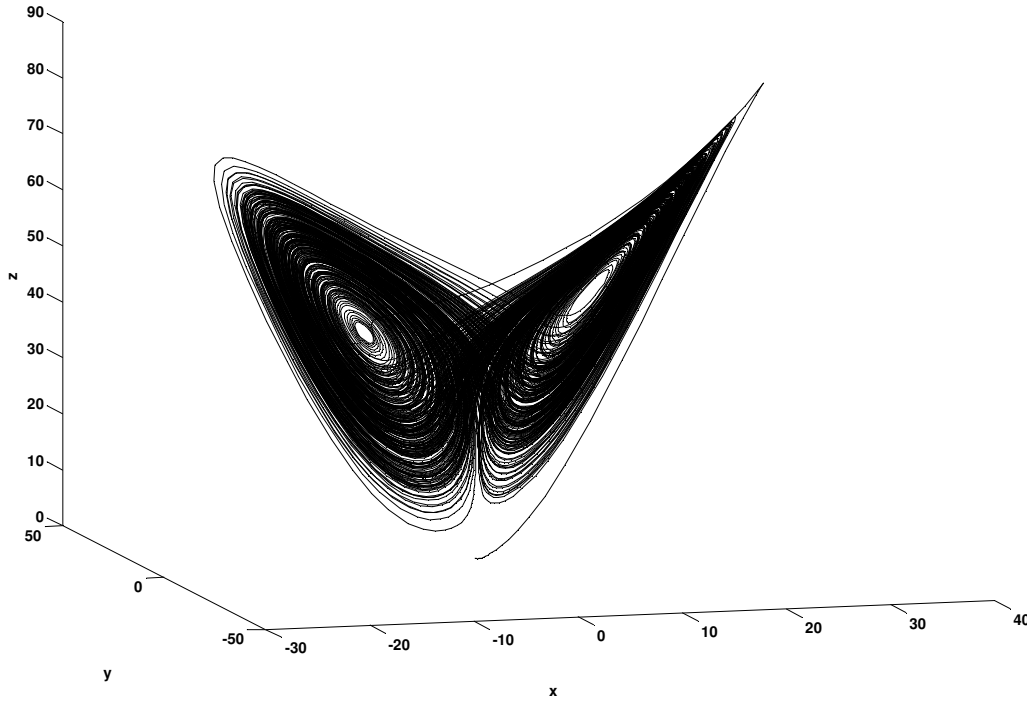


**FIG. 6.7 – Atrator de Lorenz**

O atrator de Lorenz é encontrado através da solução das três equações diferenciais não lineares a seguir:

$$\begin{cases} \dot{x} = \sigma(y - x) \\ \dot{y} = \rho x - y - xz \\ \dot{z} = xy - \beta z \end{cases}, \qquad 6.8$$

onde os parâmetros adotados foram $\sigma = 16$, $\rho = 45.92$ e $\beta = 4$ (Jiang e Adeli, 2003). As soluções numéricas da forma pontual são geradas a partir do algoritmo de Runge-Kutta de 4ª ordem, implementado através da função *ode45.m* do Matlab® com condição inicial $[0 \quad 1 \quad 1]$. Na FIG. 6.7 é possível visualizar o atrator de Lorenz encontrado utilizando os parâmetros e as condições iniciais relacionadas acima.

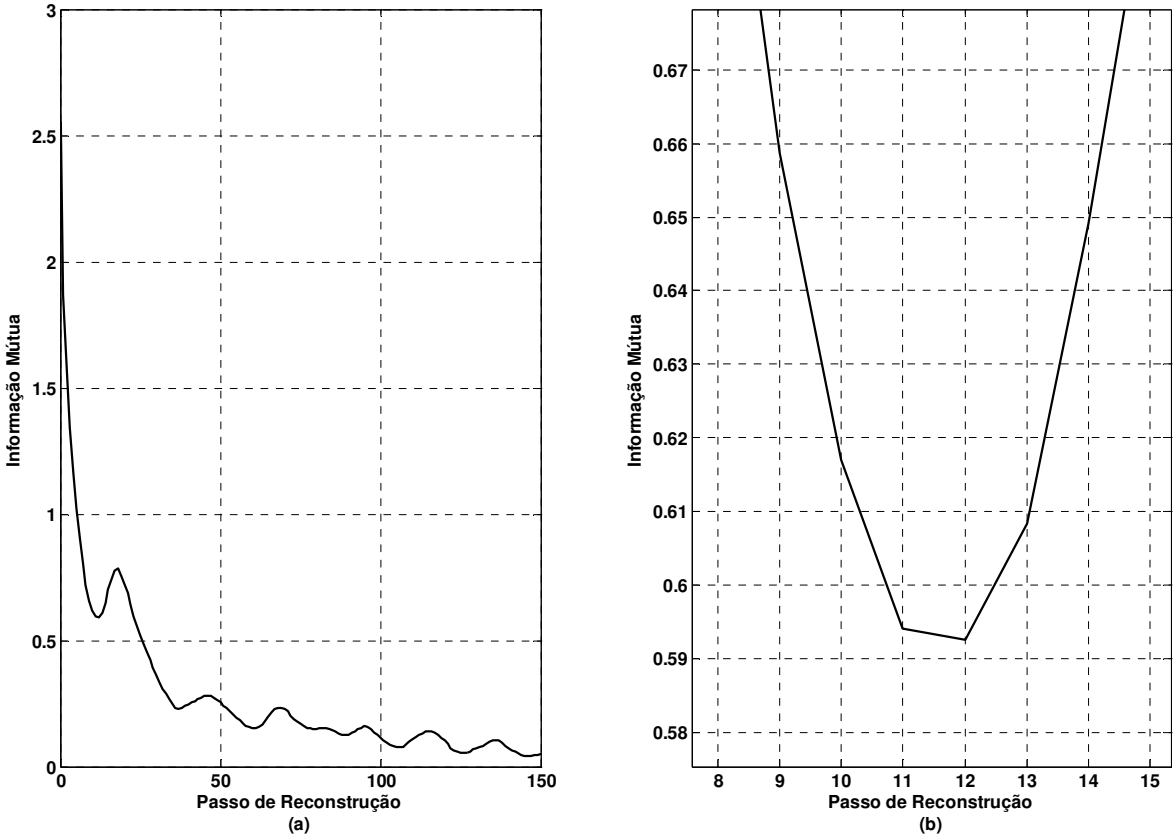


**FIG. 6.8 – (a) Resultado obtido através do método da informação mútua aplicado aos dados da saída $x$ do atrator de Lorenz e (b) Localização do primeiro mínimo local no gráfico da informação mútua**

Na FIG. 6.8 (a) está ilustrado o resultado da informação mútua e na FIG. 6.8 (b) está ampliada a região onde se encontra o primeiro mínimo local do gráfico da informação mútua, calculado a partir dos dados do sistema de Lorenz. Através da análise desses dados, verifica-se que o passo de reconstrução a ser utilizado é 12.

A FIG. 6.9 ilustra as curvas de convergência encontradas através do agrupamento de dados nebuloso *c-means* e a TAB. 6.2 relaciona a quantidade de iterações necessárias para atingir o critério de convergência estabelecido no capítulo 4, aplicando uma tolerância $\varepsilon = 0.00001$ e um $Nmáx = 600$. O algoritmo de agrupamento nebuloso foi testando, variando as dimensões de imersão de 2 a 10. Analisando a FIG. 6.9 e a TAB. 6.2, verifica-se que a dimensão de imersão para a qual foram necessários o menor número de iterações para se atingir os critérios de convergência estabelecidos foi $m = 3$, ou seja, serão utilizadas 3 funções de pertinência no antecedente de cada regra do sistema de inferência nebuloso proposto.

Aplicando-se os resultados obtidos à rotina *subclust.m* do Matlab®, encontra-se o número de centros do agrupamento de dados subtrativo igual a 12, ou seja, a quantidade de regras utilizadas no sistema de inferência nebulosos proposto é 12.

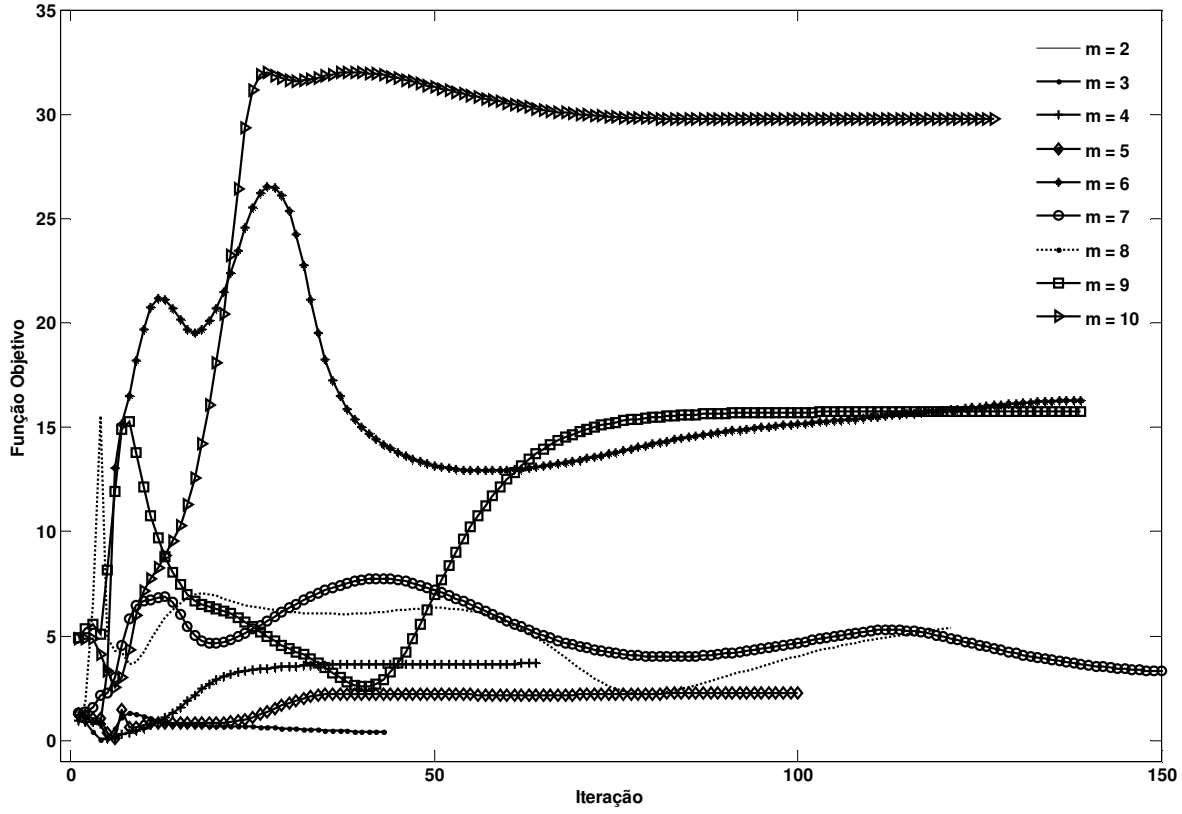


**FIG. 6.9 – Curvas de convergência do agrupamento de dados nebuloso *c-means* para os vetores do espaço de estados reconstruído da saída $x$ do atrator de Lorenz com $\tau = 12$ e $m = 2-10$**

**TAB. 6.2 – Resultados do método do agrupamento de dados nebuloso *c-means* proposto para a saída $x$ do atrator de Lorenz**

| Dimensão de imersão | 2 | **3** | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Valor Médio de Iterações | 64,20 | **43,95** | 64,65 | 100,10 | 139,82 | 412,25 | 121,75 | 139,75 | 127,30 |
| Valor mínimo do agrupamento nebuloso c-means | 0,0701 | 0,0202 | 0,0701 | 0,0642 | 0,8283 | 1,2971 | 1,2834 | 2,5796 | 2,5378 |

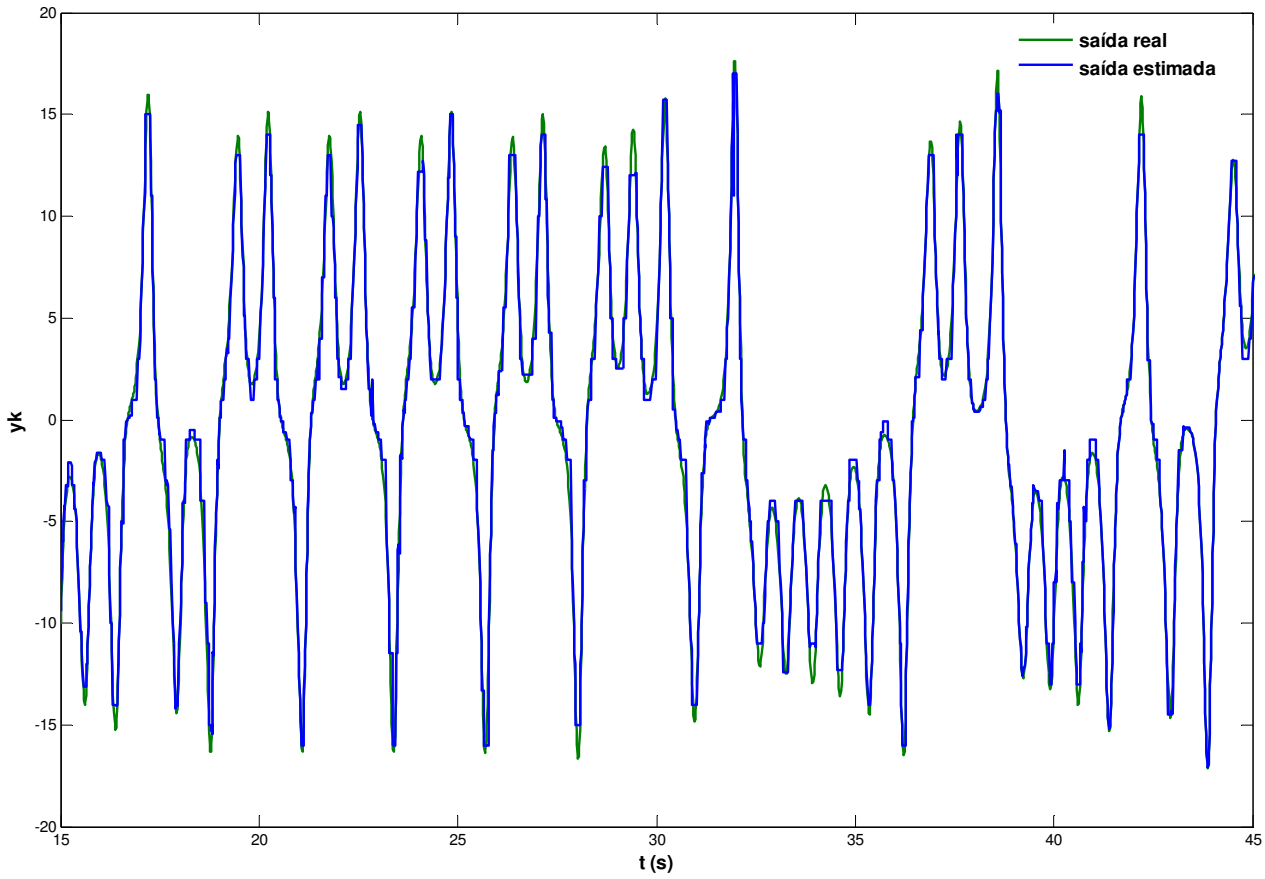


**FIG. 6.10 – Comparação entre a saída $x$ do atrator Lorenz de real (verde) e a estimada através sistema de inferência nebuloso proposto (azul)**

## 6.3 SISTEMA DE RÖSSLER

Baseado nos estudos numéricos desenvolvidos por Lorenz, em 1976, o matemático O. E. Rössler propôs um modelo tri-dimensional de equações diferenciais ordinárias, que, apesar de mais simples, apresentando somente um termo quadrático, também apresentava comportamento caótico, similar ao apresentado pelo sistema de Lorenz. De início apresentado como um sistema puramente teórico, o sistema de Rössler mostrou-se útil na modelagem e análise do equilíbrio de reações químicas. O sistema de Rössler é definido pelas três equações diferenciais descritas na EQ. 6.9 abaixo:

$$\begin{cases} \dot{x} = -(y+z) \\ \dot{y} = x + ay \\ \dot{z} = b + z(x-c) \end{cases}, \qquad 6.9$$

onde os parâmetros adotados foram $a = 0.38$, $b = 0.3$ e $c = 4.5$ (Jiang e Adeli, 2003). As soluções numéricas da forma pontual são geradas a partir do algoritmo de Runge-Kutta de 4ª ordem, implementado através da função *ode45.m* do Matlab® com condição inicial $[0 \quad 0 \quad 0]$. Na FIG. 6.11 é possível visualizar o atrator de Rössler encontrado utilizando os parâmetros e as condições iniciais relacionadas acima.

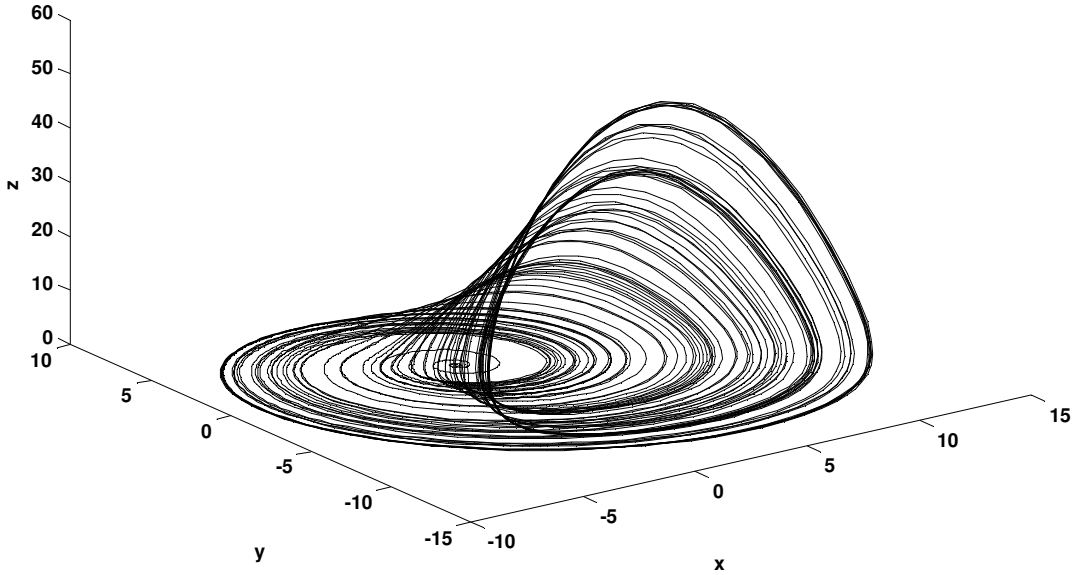


**FIG. 6.11 – Atrator de Rössler**

Na FIG. 6.12 (a) está ilustrado o resultado da informação mútua e na FIG. 6.12 (b) está ampliada a região onde se encontra o primeiro mínimo local do gráfico da informação mútua, calculado a partir dos dados do sistema de Rössler. Através da análise desses dados, verifica-se que o passo de reconstrução a ser utilizado é 17.

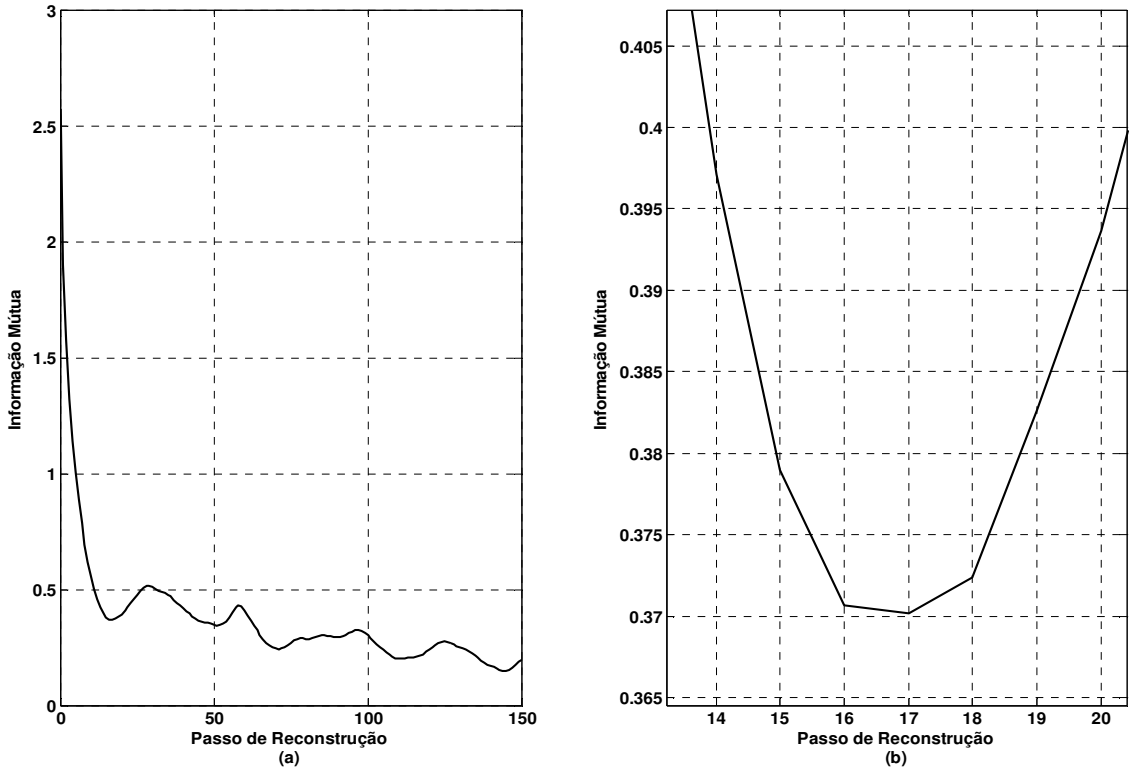


**FIG. 6.12 – (a) Resultado obtido através do método da informação mútua aplicado aos dados da saída $x$ do atrator de Rössler e (b) Localização do primeiro mínimo local no gráfico da informação mútua**

A FIG. 6.13 ilustra as curvas de convergência encontradas através do agrupamento de dados nebuloso *c-means* e a TAB. 6.3 relaciona a quantidade de iterações necessárias para atingir o critério de convergência estabelecido no capítulo 4, aplicando uma tolerância $\varepsilon = 0.00001$ e um $Nmáx = 600$. O algoritmo de agrupamento nebuloso foi testando, variando as dimensões de imersão de 2 a 10. Analisando a FIG. 6.13 e a TAB. 6.3, verifica-se que a dimensão de imersão para a qual foram necessários o menor número de iterações para se atingir os critérios de convergência estabelecidos foi $m = 3$, ou seja, serão utilizadas 3 funções de pertinência no antecedente de cada regra do sistema de inferência nebuloso proposto.

Aplicando-se os resultados obtidos à rotina *subclust.m* do Matlab®, encontra-se o número de centros do agrupamento de dados subtrativo igual a 8, ou seja, a quantidade de regras utilizadas no sistema de inferência nebulosos proposto é 8.

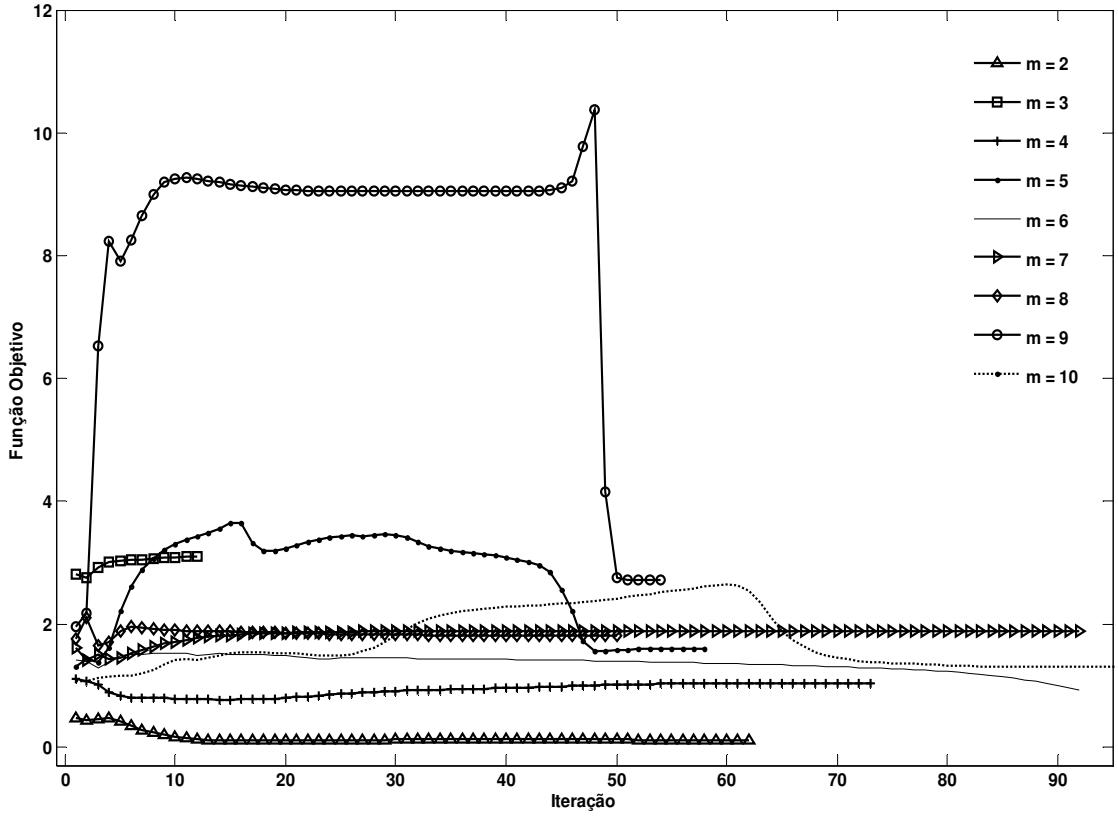


**FIG. 6.13 – Curvas de convergência do agrupamento de dados nebuloso *c-means* para os vetores do espaço de estados reconstruído da saída $x$ do atrator de Rössler com $\tau = 17$ e $m = 2-10$**

**TAB. 6.3 – Resultados do método do agrupamento de dados nebuloso *c-means* proposto para a saída $x$ do atrator de Rössler**

| Dimensão de imersão | 2 | **3** | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Valor Médio de Iterações | 63,30 | **18,50** | 77,10 | 58,80 | 107,80 | 93,60 | 41,00 | 44,00 | 132,90 |
| Valor mínimo do agrupamento nebuloso c-means | 0,1013 | 1,0739 | 0,769 | 1,3068 | 0,9212 | 1,4169 | 1,6549 | 1,9654 | 2,765 |

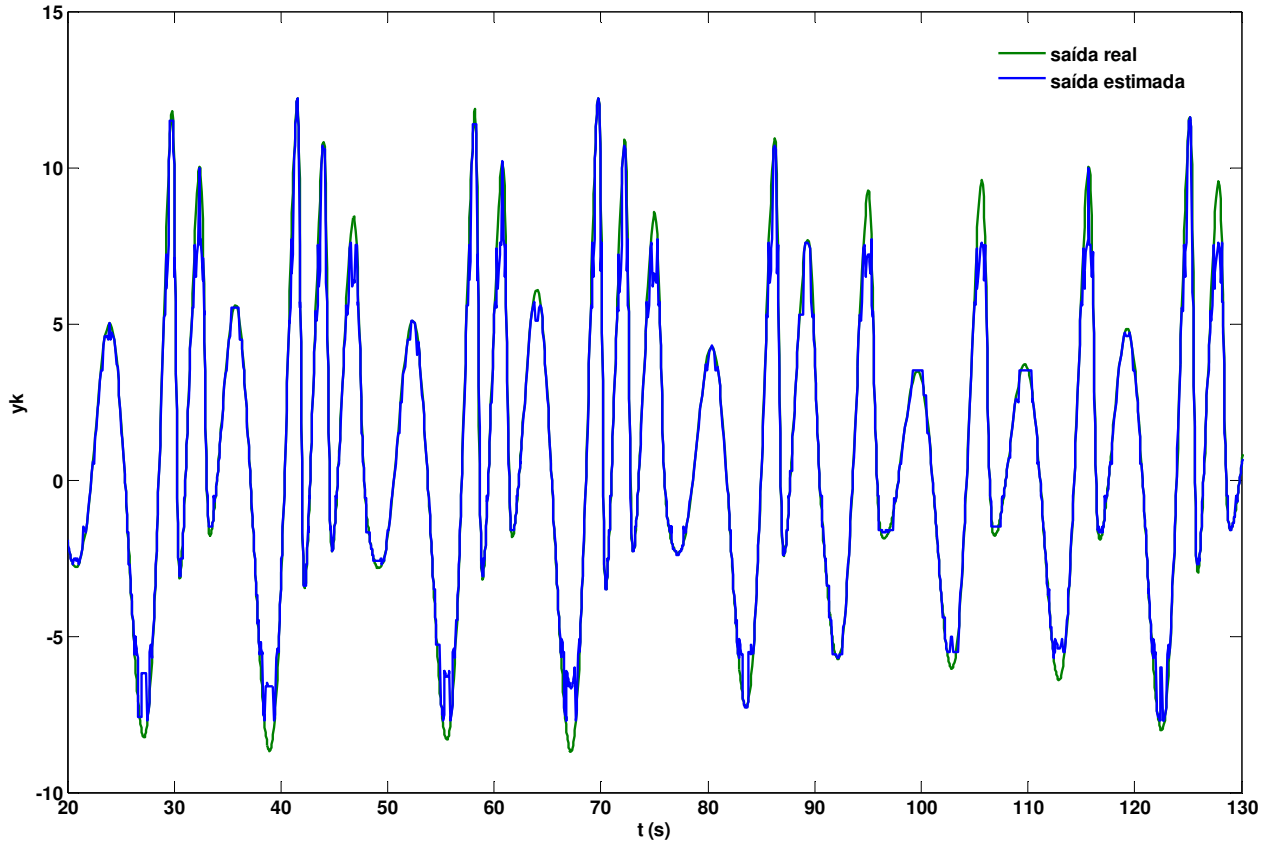


**FIG. 6.14 – Comparação entre a saída $x$ do atrator Rössler de real (verde) e a estimada através sistema de inferência nebuloso proposto (azul)**

## 6.4 SISTEMA DE DUFFING

A equação de Duffing é usada para descrever a dinâmica não linear de sistemas elétricos e mecânicos, e recebeu este nome em homenagem aos estudos de G. Duffing na década de 1930 (Savi, 2006). Do ponto de vista mecânico, a equação de Duffing descreve a oscilação forçada de uma partícula conectada a uma mola não-linear sob a influência de amortecimento viscoso. As características caóticas do sistema de Duffing dependem sensivelmente das condições iniciais e da determinação dos parâmetros. O movimento do sistema é descrito pela seguinte equação diferencial não-linear de segunda ordem:

$$\ddot{x} = -\alpha x - \beta x^3 - \xi \dot{x} + \gamma \cos(\omega t). \qquad 6.10$$

As soluções numéricas da forma pontual são geradas a partir do algoritmo de Runge-Kutta de 4ª ordem, implementado através da função *ode45.m* do Matlab® com os parâmetros $\alpha = 1$, $\beta = 1$, $\xi = 0$ , $\gamma = 12$ e $\omega = 1$, e a condição inicial $x(0) = [0 \quad 4]$. Na FIG. 6.15 é possível visualizar a trajetória obtida através da equação de Duffing encontrada utilizando os parâmetros e as condições iniciais relacionadas acima.

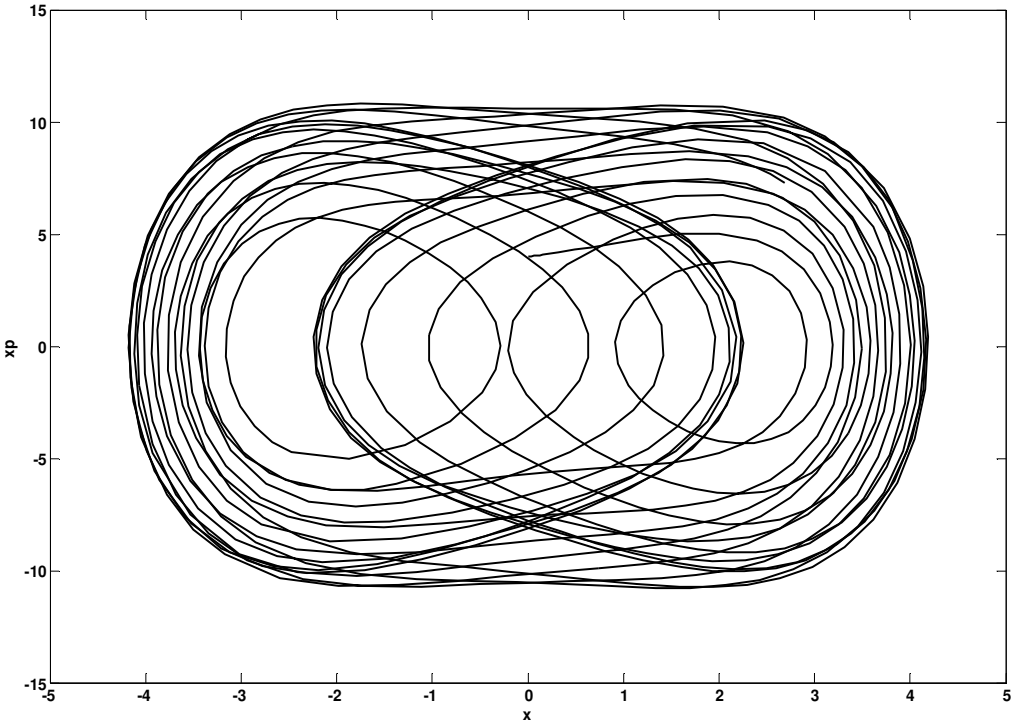


**FIG. 6.15 – Trajetória obtida através da equação Duffing**

Na FIG. 6.16 (a) está ilustrado o resultado da informação mútua e na FIG. 6.16 (b) está ampliada a região onde se encontra o primeiro mínimo local do gráfico da informação mútua, calculado a partir dos dados do sistema de Duffing. Através da análise desses dados, verifica-se que o passo de reconstrução a ser utilizado é 11.

A FIG. 6.17 ilustra as curvas de convergência encontradas através do agrupamento de dados nebuloso *c-means* e a TAB. 6.4 relaciona a quantidade de iterações necessárias para atingir o critério de convergência estabelecido no capítulo

4, aplicando uma tolerância $\varepsilon = 0.00001$ e um $Nmáx = 600$. O algoritmo de agrupamento nebuloso foi testando, variando as dimensões de imersão de 2 a 10. Analisando a FIG. 6.17 e a TAB. 6.4, verifica-se que a dimensão de imersão para a qual foram necessários o menor número de iterações para se atingir os critérios de convergência estabelecidos foi $m = 2$, ou seja, serão utilizadas 2 funções de pertinência no antecedente de cada regra do sistema de inferência nebuloso proposto.

Aplicando-se os resultados obtidos à rotina *subclust.m* do Matlab®, encontra-se o número de centros do agrupamento de dados subtrativo igual a 6, ou seja, a quantidade de regras utilizadas no sistema de inferência nebulosos proposto é 6.

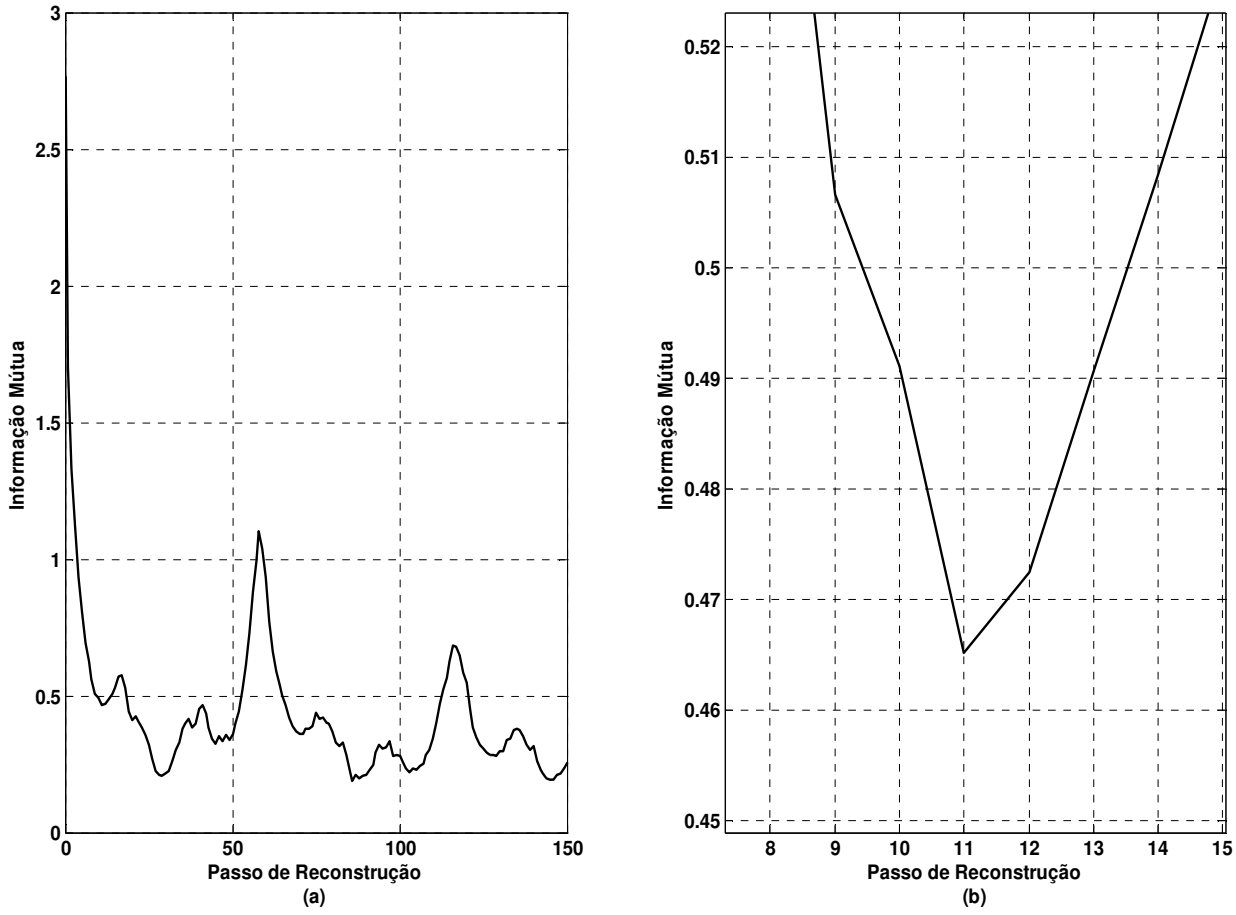


**FIG. 6.16 – (a) Resultado obtido através do método da informação mútua aplicado aos dados do sistema de Duffing e (b) Localização do primeiro mínimo local no gráfico da informação mútua**

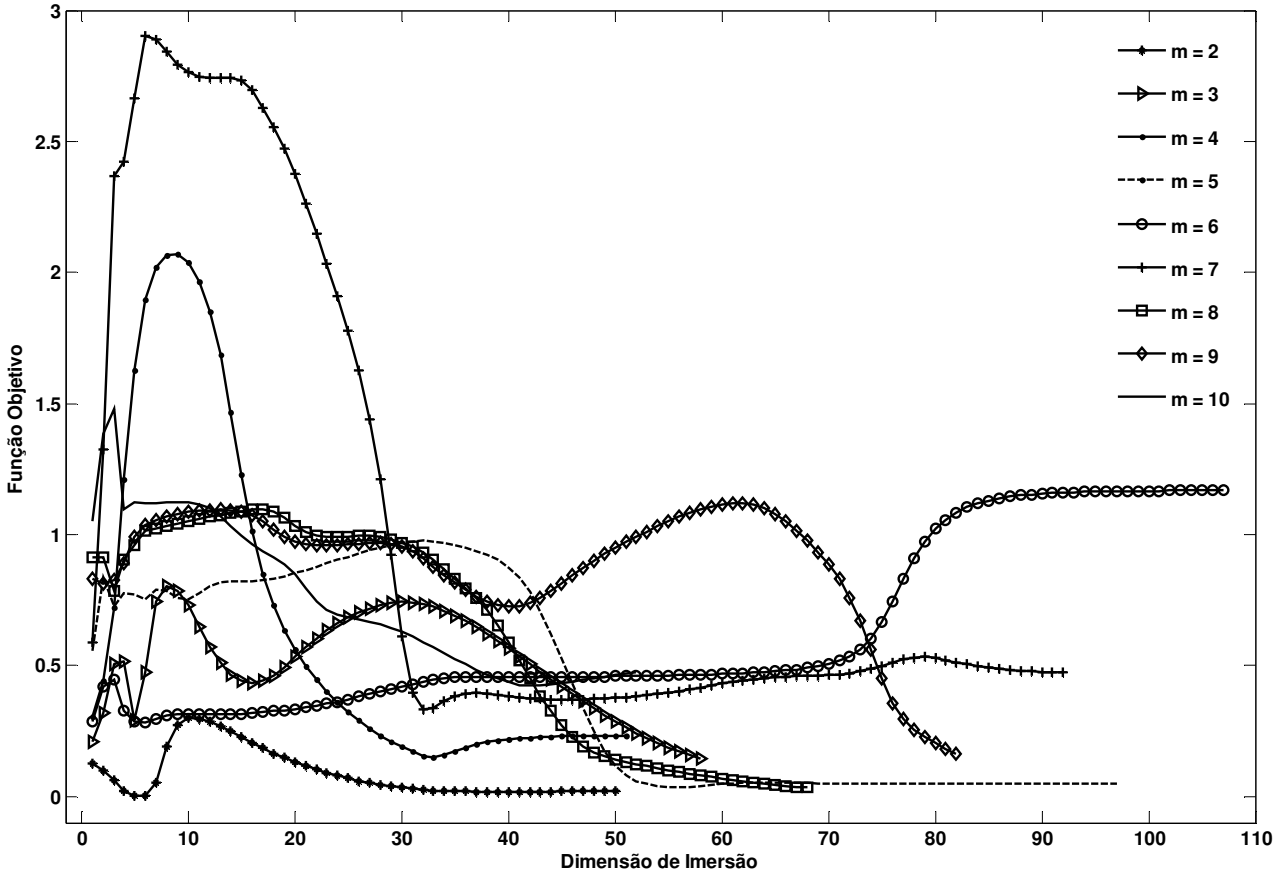


**FIG. 6.17 – Curvas de convergência do agrupamento de dados nebuloso *c-means* para os vetores do espaço de estados reconstruído do sistema de Duffing com $\tau = 11$ e $m = 2-10$**

**TAB. 6.4 – Resultados do método do agrupamento de dados nebuloso *c-means* proposto para o sistema de Duffing**

| Dimensão de imersão | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Valor médio de iterações | **49,78** | 57,67 | 51,86 | 97,19 | 106,05 | 92,58 | 67,37 | 81,00 | 52,88 |
| Valor mínimo do agrupamento nebuloso c-means | 0,002 | 0,145 | 0,150 | 0,034 | 0,283 | 0,331 | 0,034 | 0,163 | 0,422 |

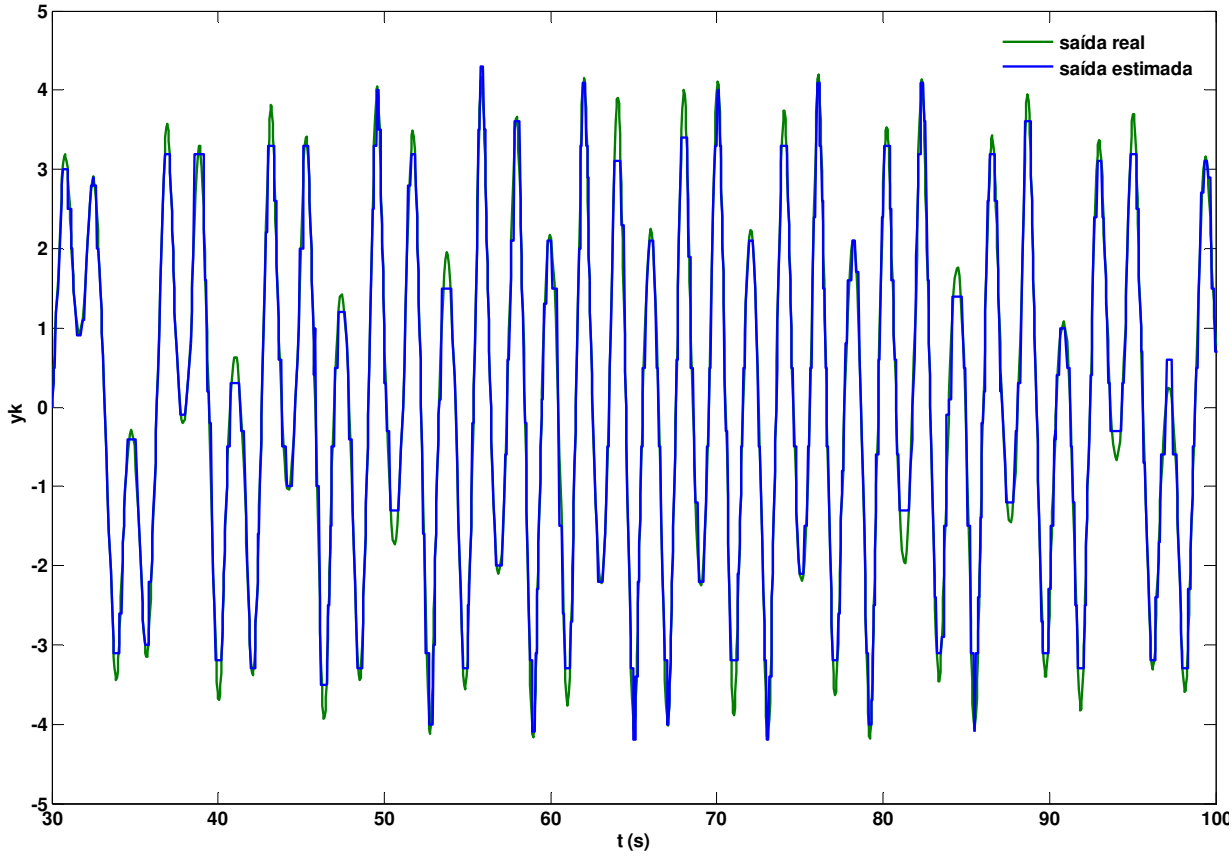


**FIG. 6.18 – Comparação entre a saída do sistema de Duffing real (verde) e a estimada através sistema de inferência nebuloso proposto (azul)**

## 6.5 SISTEMA DE VAN DER POL

A equação de Van der Pol foi introduzida por Rayleigh, em 1896, e estudada teoricamente e experimentalmente por Van der Pol, em 1927, em seus estudos sobre circuitos elétricos. Ela foi originalmente obtida para descrever o comportamento de um circuito elétrico com uma válvula triodo, uma espécie de "avô" do transistor, e uma resistência cujas propriedades variam de acordo com a corrente elétrica (Savi, 2006). O movimento do sistema é descrito pela seguinte equação diferencial não-linear de segunda ordem:

$$\ddot{x} = \mu\left(1 - x^2\right)\dot{x} - x\,. \qquad 6.11$$

As soluções numéricas da forma pontual são geradas a partir do algoritmo de Runge-Kutta de 4ª ordem, implementado através da função *ode45.m* do Matlab®, sendo $\mu = 1.7$ e a condição inicial $x(0) = [0 \quad 1]$. Na FIG. 6.19 é possível visualizar a trajetória obtida através da equação de Van der Pol encontrada utilizando os parâmetros e as condições iniciais relacionadas acima.

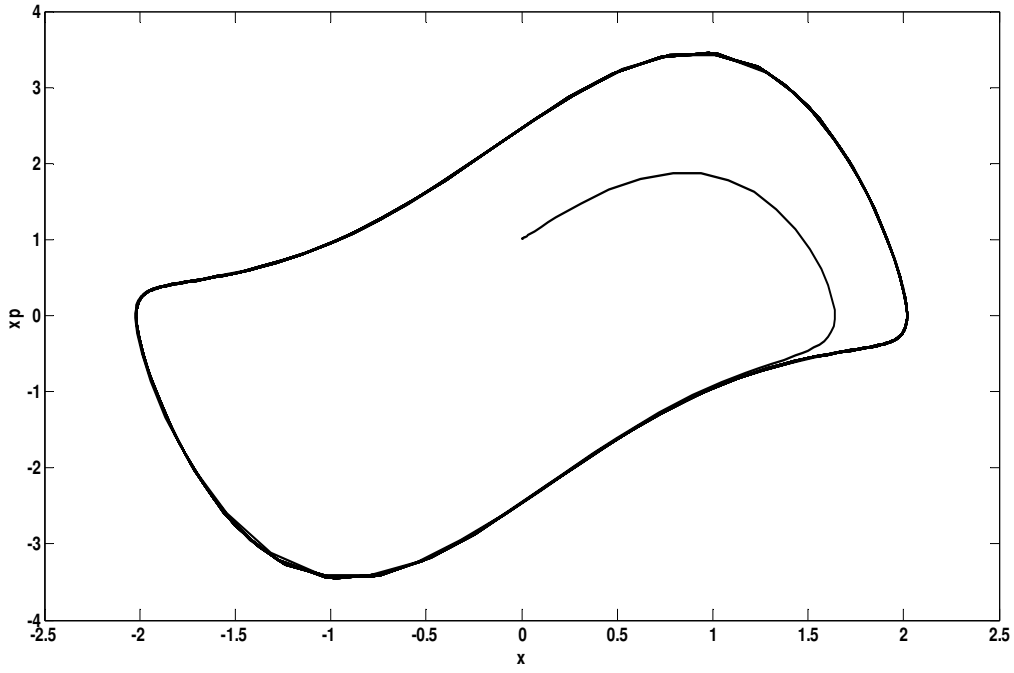


**FIG. 6.19 – Trajetória obtida através da equação Van der Pol**

Na FIG. 6.20 (a) está ilustrado o resultado da informação mútua e na FIG. 6.20 (b) está ampliada a região onde se encontra o primeiro mínimo local do gráfico da informação mútua, calculado a partir dos dados do sistema de Van der Pol. Através da análise desses dados, verifica-se que o passo de reconstrução a ser utilizado é 23.

A FIG. 6.21 ilustra as curvas de convergência encontradas através do agrupamento de dados nebuloso *c-means* e a TAB. 6.5 relaciona a quantidade de iterações necessárias para atingir o critério de convergência estabelecido no capítulo 4, aplicando uma tolerância $\varepsilon = 0.00001$ e um $Nmáx = 600$. O algoritmo de

agrupamento nebuloso foi testando, variando as dimensões de imersão de 2 a 10. Analisando a FIG. 6.21 e a TAB. 6.5, verifica-se que a dimensão de imersão para a qual foram necessários o menor número de iterações para se atingir os critérios de convergência estabelecidos foi $m = 2$, ou seja, serão utilizadas 2 funções de pertinência no antecedente de cada regra do sistema de inferência nebuloso proposto.

Aplicando-se os resultados obtidos à rotina *subclust.m* do Matlab®, encontra-se o número de centros do agrupamento de dados subtrativo igual a 9, ou seja, a quantidade de regras utilizadas no sistema de inferência nebulosos proposto é 9.

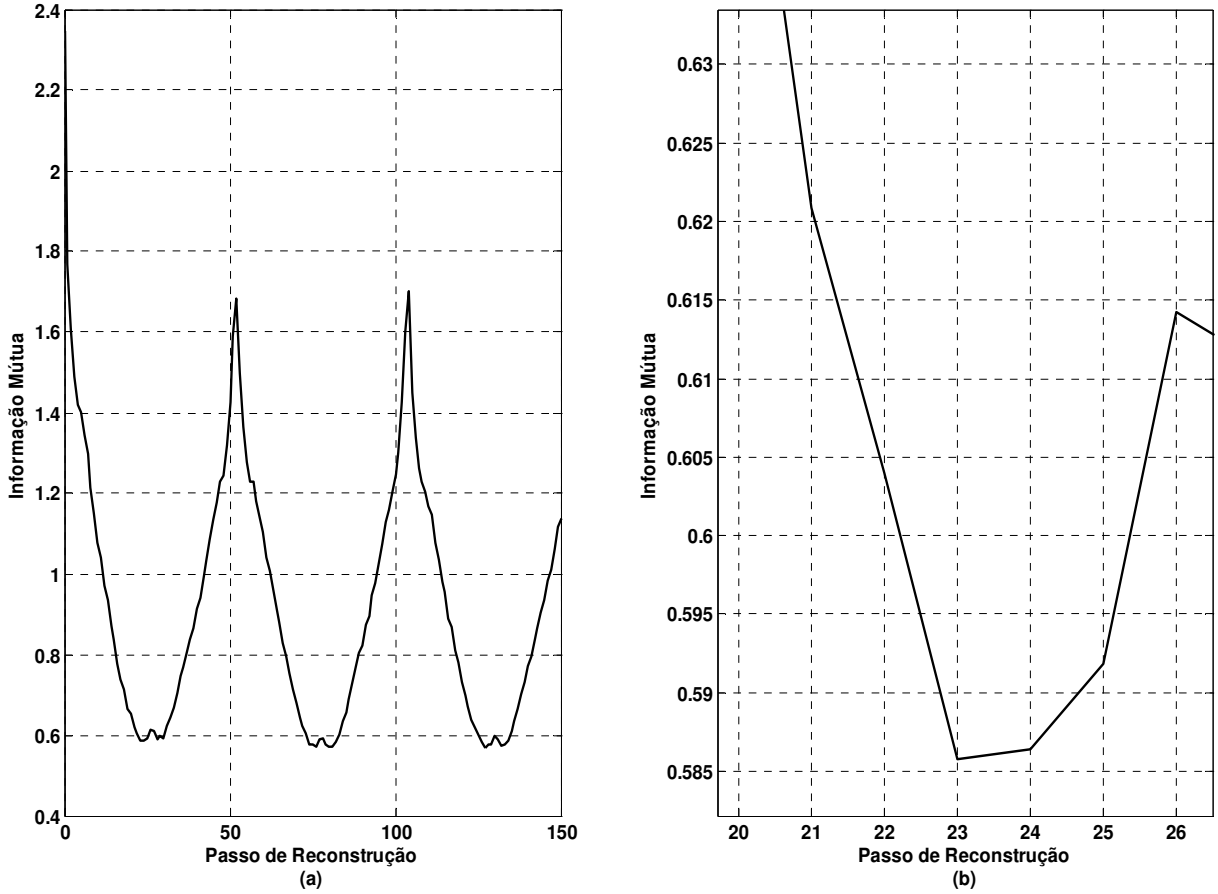


**FIG. 6.20 – (a) Resultado obtido através do método da informação mútua aplicado aos dados do sistema de Van der Pol e (b) Localização do primeiro mínimo local no gráfico da informação mútua**

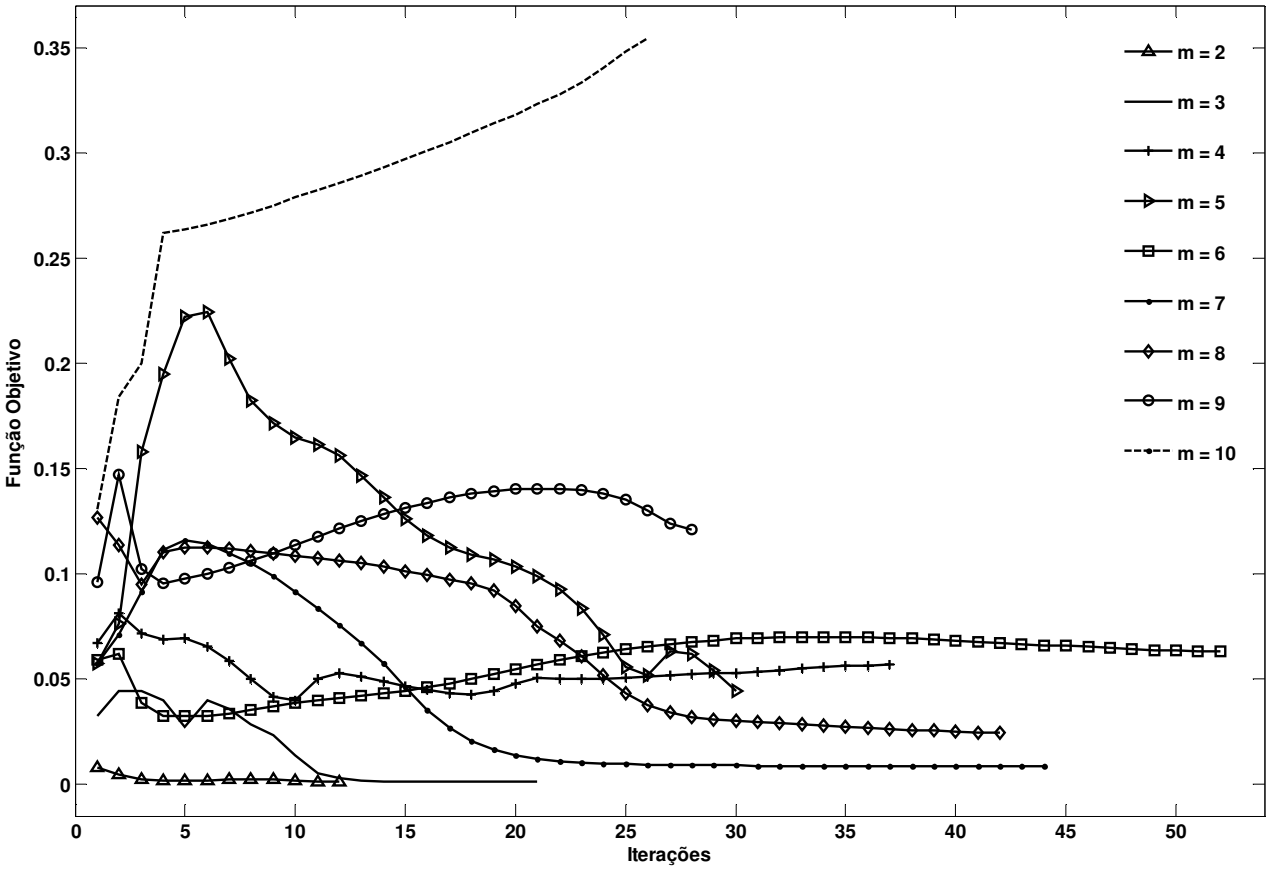


**FIG. 6.21 – Curvas de convergência do agrupamento de dados nebuloso *c-means* para os vetores do espaço de estados reconstruído do sistema de Van der Pol com $\tau = 23$ e $m = 2-10$**

**TAB. 6.5 – Resultados do método do agrupamento de dados nebuloso *c-means* proposto para o sistema de Van der Pol**

| Dimensão de imersão | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Valor médio de iterações | **12,12** | 20,98 | 36,94 | 30,41 | 52,27 | 45,85 | 40,82 | 28,50 | 28,97 |
| Valor mínimo do agrupamento nebuloso c-means | 0,0010 | 0,0011 | 0,0398 | 0,0443 | 0,0323 | 0,0084 | 0,0243 | 0,0954 | 0,1308 |

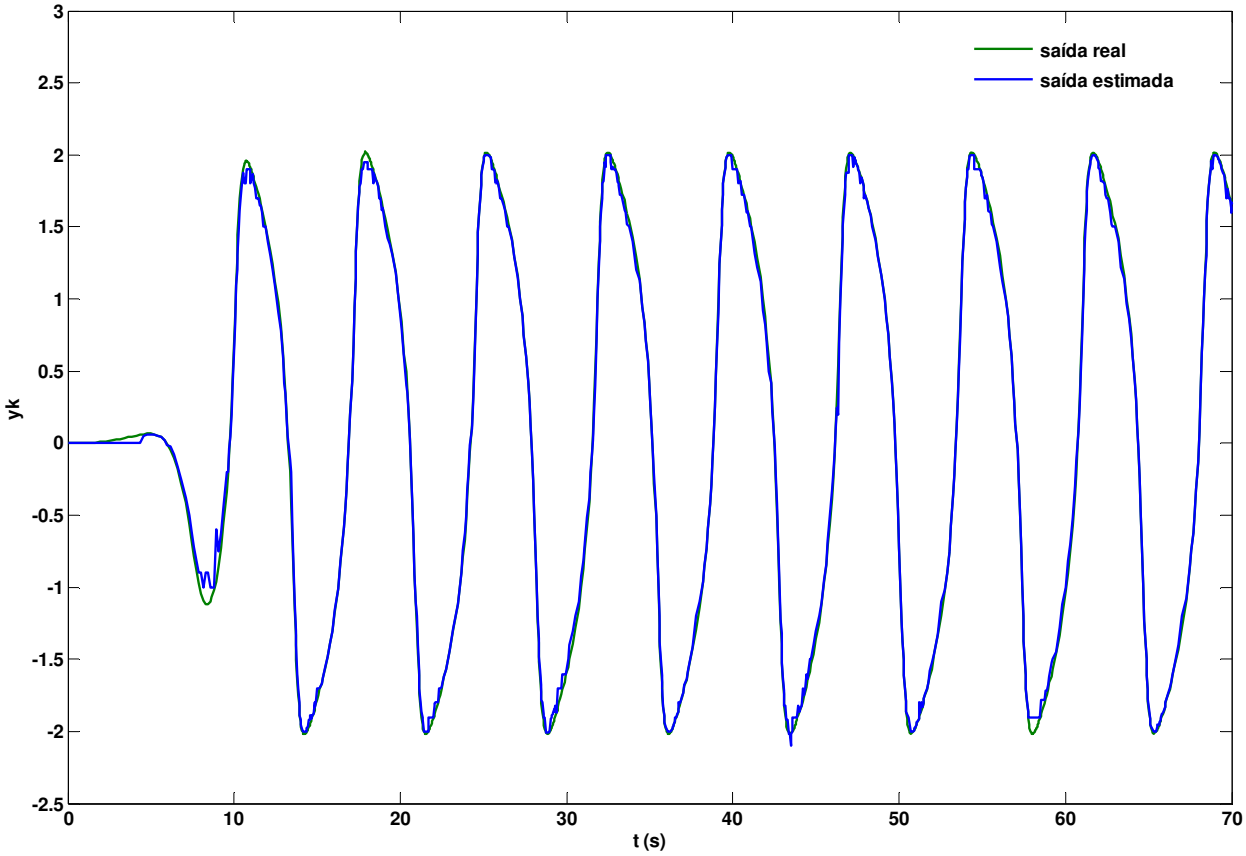


**FIG. 6.22 – Comparação entre a saída do sistema de Van der Pol real (verde) e a estimada através sistema de inferência nebuloso proposto (azul)**

# 7 CONCLUSÕES E COMENTÁRIOS

No capítulo 6, foi apresentado o modelo nebuloso Takagi-Sugeno proposto para a modelagem de séries temporais. O sistema de inferência nebuloso proposto e os métodos utilizados na estimação dos parâmetros do mesmo foram testados em cinco casos distintos: série temporal de Mackey-Glass, sistema de Rössler, sistema de Lorenz, sistema de Duffing e sistema de Van der Pol.

Os resultados obtidos que avaliam a modelagem estão ilustrados no gráfico comparativo entre a saída real do sistema estudado e a saída do sistema predito pelo modelo nebuloso proposto.

O algoritmo de agrupamento de dados nebuloso *c-means* tem a vantagem da facilidade na implementação e, em todos os casos testados, apresentando uma dimensão de imersão igual à dimensão real do sistema estudado. Além disso, ao invés de utilizar uma aproximação por tentativa e erro para definir o número de agrupamentos de dados a ser aplicado aos dados do sistema estudado, define a quantidade de agrupamentos como sendo igual ao passo de reconstrução.

O algoritmo de agrupamento de dados subtrativo é vantajoso porque propicia a determinação da quantidade de regras que irão compor o sistema de inferência nebuloso, o que possibilitou a modelagem o sistema de inferência nebuloso com estrutura mínima e com número mínimo de regras.

O algoritmo genético foi eficaz na otimização dos parâmetros das funções de pertinência que compõem o antecedente de cada regra e nos parâmetros da função que compõem o conseqüente de cada regra do modelo nebuloso, mas é uma ferramenta computacionalmente pesada, levando a uma demora demasiada na otimização desses parâmetros.

Os resultados apresentados mostram que os modelos nebulosos Takagi-Sugeno são ferramentas que fornecem um bom desempenho quando tratamos o problema de modelagem de séries temporais.

Para trabalhos futuros, a aplicação do método dos mínimos quadrados para estimar os parâmetros ótimos do modelo nebuloso Takagi-Sugeno, visando utilizar um ferramental computacionalmente mais leve que o algoritmo genético. Outra alternativa a ser estudada para diminuir o tempo computacional do algoritmo genético seria testar novos operadores genéticos.

Pode ser explorado também o estudo de redes neurais para predição de séries temporais. Além disso, o algoritmo genético também pode ser utilizado na determinação da estrutura da rede, no caso de um estudo com redes neurais.

Um caminho a ser explorado, seria a pesquisa de um algoritmo de agrupamento de dados para determinar o valor do passo de reconstrução, esperando-se assim encontrar um método mais robusto que o método da informação mútua.

E, finalmente, o algoritmo proposto nesta dissertação, bem como os trabalhos futuros propostos acima, podem ser implementados utilizando-se dados coletados de um sistema real como, por exemplo, um pêndulo não-linear.

## 8. REFERÊNCIAS BIBLIOGRÁFICAS


ABONYI, J., FEIL, B., NEMETH, S., ARVA, P. e BABUSKA, R. **State-Space Reconstruction and Prediction of Chaotic Time Series based on Fuzzy Clustering.** IEEE International Conference on, Systems, Man and Cybernetics, págs. 2374–2380, 2004.

CAMARGOS, Fernando Laudares. **Lógica Nebulosa:** Uma Abordagem Filosófica e Aplicada. Technical report, Universidade Federal de Santa Catarina, 2002.

CARDOSO, Giselle Cristina. **Modelos de Previsão baseado em Agrupamento e Base de Regras Nebulosas**. 2003. 84 p. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Estadual de Campinas, 2003.

CHANG, Wei-Der. **An Improved Real-Coded Genetic Algorithm for Parameters Estimation of Nonlinear Systems**. Mechanical Systems and Signal Processing 20, págs. 236–246, 2006.

CHANG, Wei-Der. **Nonlinear System Identification and Control using a Real-Coded Genetic Algorithm**. Applied Mathematical Modelling 31, págs. 541–550, 2007.

DELGADO, Myriam. Regattiere de Biase Silva. **Projeto Automático de Sistemas Nebulosos:** Uma Abordagem Co-Evolutiva. 2002. 186 p. Tese (Doutorado em Engenharia Elétrica), Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, 2002.

EFTEKHARI, M. e KATEBI, S. D. Extracting **Compact Fuzzy Rules for Nonlinear System Modeling using Subtractive Clustering, GA and Unscented Filter**. Applied. Mathematical Modelling, 2007.

FEIL, B., ABONYI, J. e SZEIFERT, F. **Model Order Selection of Nonlinear Imput-Output Models – A Clustering Based Approach**. Journal of Process Control 14, págs 593-602, 2004.

FENG, Xin e HUANG, Hai **A Fuzzy-Set-Based Reconstructed Phase Space Method for Identification of Temporal Patterns in Complex Time Series**. IEEE Transactions on Knowledge and Data Engineering, 17(5), págs. 601–613, 2005.

FRANCA, Luiz Fernando P. e SAVI, Marcelo Amorim. **Distinguishing Periodic and Chaotic Time Series Obtained from an Experimental Nonlinear Pendulum**. Nonlinear Dynamics 26, págs. 253–271, 2001. **(A)**

FRANCA, Luiz Fernando P. e SAVI, Marcelo Amorim. **Estimating Attractor Dimension on the Nonlinear Pendulum Time Series**. Journal of the Brazilian Society of Mechanical Sciences, 23(4), 2001. **(B)**

FRASER, A. M. e SWINNEY, H. L. **Independent Coordinates for Strange Attractors from Mutual Information**. Physical Review, 33(2), págs. 1134–1140, 1986.

GOLDBERG, David E. e DEB, Kalyanmoy. **Messy Genetic Algorithms:** Motivation, Analysis and First Results. Complex Systems 3, págs. 493–530, 1989.

GU, Hong e WANG, Hongwei. **Fuzzy Prediction of Chaotic Time Series based on Singular Value Decomposition**. Applied Mathematics and Computation 185, págs. 1171–1185, 2007.

JIANG, Xiaomo e ADELI, Hojjat. **Fuzzy Clustering Approach for Accurate Embedding Dimension Identification in Chaotic Time Series**. Integrated Computer-Aided Engineering (10), págs. 287-302, 2003.

JANG, Jyh- Shing R. e SUN, Chuen-Tsai. **Predicting Chaotic Time Series with Fuzzy If-Then Rules**. IEEE Second International Conference on Fuzzy Systems, 2, págs. 1079–1084, 1993.

JOO, Young Hoon, SHIEH, Leang-San e CHEN, Guanrong. **Hybrid State-Space Fuzzy Model-Based Controller with Dual-Rate Sampling for Digital Control of Chaotic Systems**. IEEE Transactions on Fuzzy Systems, 7(4), págs. 394–408, 1999.

LAU, K. W. e WU, H. **Local Prediction of Non-Linear Time Series using Support Vector Regression**. Pattern Recognition 41, págs. 1539–1547, 2008.

LEE, Ho Jae, PARK, Jin Bae e CHEN, Guanrong. **Robust Fuzzy Control of Nonlinear Systems with Parametric Uncertainties**. IEEE Transactions on Fuzzy Systems, 9(2), págs 369–379, 2001.

LEE, Ho Jae, PARK, Jin Bae e JOO, Young Hoon. **Fuzzy Model Identification using a Hybrid mGA Scheme with Application to Chaotic System Modeling**. Integration of Fuzzy Logic and Chaos Theory, págs. 81–97, 2006.

MANGUIRE, L. P., ROCHE, B., MCGINNITY, T. M. e MCDAID, L. J. **Predicting Chaotic Time Series using a Fuzzy Neural Network**. Information Sciences, 112(1), págs 125–136, 1998.

MARUO, Marcos Hideo. **Projeto Automático de Sistemas Nebulosos utilizando Algoritmos Genéticos Auto-Adaptativos**. 2006. 138 p. Dissertação (Mestrado em Informática Industrial) - Universidade Tecnológica Federal do Paraná – Faculdade de Engenharia Elétrica e Informática Industrial, 2006.

PIRES, Matheus Giovanni. **Aprendizado Genético de Funções de Pertinência na Modelagem Nebulosa**. 2004. 128 p. Tese (Doutorado em Ciências da Computação) - Universidade Federal de São Carlos, Faculdade de Ciências da Computação, 2004.

PISARENKO, V. F. e SORNETTE, D. **Statistical Methods of Parameter Estimation for Deterministically Chaotic Time Series**. Physical Review e 69, págs. 1–12, 2004.

PUCCIARELLI, Amilcar. José. **Modelagem de Séries Temporais Discretas utilizando Modelo Nebuloso Takagi-Sugeno**. 2005. 116 p. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e Computação, 2005.

SANCHES, André Rodrigo. **Redução de Dimensionalidade em Séries Temporais**. 2005. Dissertação (Mestrado em Ciências da Computação) - Universidade de São Paulo, Instituto de Matemática e Estatística, 2006.

SANTOS, André Alves Portela. **Previsão Não Linear da Taxa de Câmbio Real/Dólar utilizando Redes Neurais e Sistemas Nebulosos**. 2005. 85 p. Dissertação (Mestrado em Economia) - Universidade Federal de Santa Catarina, 2005.

SAVI, Marcelo Amorin. **Dinâmica Não-Linear e Caos**. Rio de Janeiro: E-papers Serviços Editoriais Ltda, 2006.

SHAW, Ian. S. e SIMÕES, Marcelo Godoy. **Controle e Modelagem Fuzzy**. Editora Edgard Blücher Ltda, 1999.

TAKENS, Floris. **Detecting Strange Attractors in Turbulence**, Lecture Notes in Mathematics. Vol. 898, Springer, New York, págs. 366-381, 1981.

WANG, Hong-Wei, GU, Hong e WANG, Zhe-Long. **Fuzzy Prediction of Chaotic Time Series based on SVD Matrix Decomposition**. Fourth International Conference on Machine Learning and Cybernetics, págs. 2493–2498, 2005.

ZHANG, Wei, YANG, Gen-Ke e WU, Zhi-Ming. **Genetic Programming-Based Modeling on Chaotic Time Series**. Third International Conference and Machine Learning and Cybernetics, págs. 2347–2352, 2004.

# Livros Grátis

( http://www.livrosgratis.com.br )

Milhares de Livros para Download:


Baixar livros de Administração
Baixar livros de Agronomia
Baixar livros de Arquitetura
Baixar livros de Artes
Baixar livros de Astronomia
Baixar livros de Biologia Geral
Baixar livros de Ciência da Computação
Baixar livros de Ciência da Informação
Baixar livros de Ciência Política
Baixar livros de Ciências da Saúde
Baixar livros de Comunicação
Baixar livros do Conselho Nacional de Educação - CNE
Baixar livros de Defesa civil
Baixar livros de Direito
Baixar livros de Direitos humanos
Baixar livros de Economia
Baixar livros de Economia Doméstica
Baixar livros de Educação
Baixar livros de Educação - Trânsito
Baixar livros de Educação Física
Baixar livros de Engenharia Aeroespacial
Baixar livros de Farmácia
Baixar livros de Filosofia
Baixar livros de Física
Baixar livros de Geociências
Baixar livros de Geografia
Baixar livros de História
Baixar livros de Línguas

Baixar livros de Literatura
Baixar livros de Literatura de Cordel
Baixar livros de Literatura Infantil
Baixar livros de Matemática
Baixar livros de Medicina
Baixar livros de Medicina Veterinária
Baixar livros de Meio Ambiente
Baixar livros de Meteorologia
Baixar Monografias e TCC
Baixar livros Multidisciplinar
Baixar livros de Música
Baixar livros de Psicologia
Baixar livros de Química
Baixar livros de Saúde Coletiva
Baixar livros de Serviço Social
Baixar livros de Sociologia
Baixar livros de Teologia
Baixar livros de Trabalho
Baixar livros de Turismo