

UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA ELÉTRICA

DEPARTAMENTO DE ENGENHARIA DE COMPUTAÇÃO E AUTOMAÇÃO INDUSTRIAL

## ESTIMAÇÃO PARAMÉTRICA ROBUSTA ATRAVÉS DE REDES NEURAIIS ARTIFICIAIS

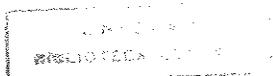
*Por:* Ivan Nunes da Silva

*Orientação:* Dra. Lúcia Valéria Ramos de Arruda

*Co-Orientação:* Prof. Dr. Wagner Caradori do Amaral

Dissertação apresentada à Faculdade de Engenharia Elétrica da Universidade Estadual de Campinas, como parte dos requisitos exigidos para obtenção do Título de Mestre em Engenharia Elétrica.

JANEIRO DE 1995



*Este trabalho contou com o apoio financeiro do CNPq .*

## Agradecimentos

Aos meus orientadores Dra. Lúcia Valéria Ramos de Arruda e Prof. Dr. Wagner Caradori do Amaral pela segurança e competência demonstradas na orientação deste trabalho, além do apoio e incentivos dedicados.

Aos meus familiares pela presença sempre constante.

A secretaria do Departamento de Engenharia de Computação e Automação Industrial (DCA) e a seção de pós-graduação, da Faculdade de Engenharia Elétrica da Unicamp pela atenção e serviços prestados.

Por último, gostaria de mostrar a minha gratidão a todos os amigos que fiz na Unicamp, em especial aos amigos do LCA.

Porque *dele* e por *ele*, e para *ele*, são todas  
as coisas; glória, pois a *ele* eternamente.  
Amém.

Romanos 11,36.

Dedico este trabalho para os meus pais *Gerson e Hosana* pelo apoio em todas as horas, e para a minha noiva *Sileide* pelo carinho e compreensão em todos os momentos.

## **ABSTRACT**

Artificial Neural Networks can achieve high computation rates by employing a massive number of simple processing elements with a high degree of connectivity between these elements. Neural networks with feedback connections provide a computing model to solve a rich class of optimization problems. This dissertation presents an application of Hopfield's Neural Networks in Robust Parametric Estimation with unknown-but-bounded disturbance. The Discret Hopfield's Network is used to calculate a parameter uncertainty set for model parameters. Any element in this set can be considered a good estimator for the real parameters. A Modified Hopfield's Network has also been described and it is useful for getting efficient and reliable sets. Comparative analysis with others robust estimation approaches are included.

The Valid-Subspace technique is used to obtain the internal parameters of the Hopfield's Neural Network. These parameters are explicitly computed, based upon problem specifications, to assure the network convergence. In this case, the equilibrium point represents a solution to robust estimation problem with unknown-but-bounded error.

## SUMÁRIO

Redes Neurais Artificiais atingem altas taxas de computação através de um número massivo de elementos processadores simples com um alto grau de conectividade entre estes elementos. Redes neurais com conexões realimentadas podem ser utilizadas para resolver problemas de otimização. Nesta dissertação utilizam-se Redes Neurais Artificiais do tipo Hopfield, na solução de problemas de Estimção Paramétrica Robusta com perturbação desconhecida-mas-limitada. A rede de Hopfield discreta é usada no cálculo de regiões de incerteza para os parâmetros do modelo. Qualquer elemento pertencente a estas regiões são considerados um bom estimador dos parâmetros reais do processo. Uma Rede de Hopfield Modificada também é descrita e é utilizada para assegurar maior eficiência e confiabilidade na obtenção de resultados. Análises são realizadas comparando os resultados obtidos pela rede em relação a outras abordagens tradicionais de cálculo de incertezas paramétricas.

Os parâmetros internos das Redes para o problema são derivados utilizando uma técnica denominada subespaço-válido de soluções. Estes parâmetros são explicitamente calculados, baseado nas especificações do problema, e devem assegurar a convergência da rede para um ponto de equilíbrio que representa uma solução para o problema de estimção robusta de modelos com perturbações desconhecidas-mas-limitadas.

# ÍNDICE

## *Capítulo 1 - Introdução*

1.1-) Motivação e Relevância do Trabalho .....	01
1.2-) Proposta da Dissertação .....	03
1.3-) Organização da Dissertação.....	04
Bibliografia .....	05

## *Capítulo 2 - Métodos Clássicos Para Estimação Robusta de Regiões Paramétricas*

2.1-) Introdução.....	06
2.2-) Formulação do Problema.....	09
2.3-) Métodos (UBBE) Para Estimação de Modelos Lineares .....	10
2.3.1-) Aproximação do Politopo $S$ Por uma Região Elipsoidal .....	12
2.3.2-) Aprox. da Região Politópica por uma Região Ortotópica.....	18
2.3.3-) Descrição Exata da Região Politópica.....	19
2.4-) O Problema dos Modelos Não Lineares .....	26
2.5-) Conclusão .....	27
Bibliografia .....	28

## *Capítulo 3 - Fundamentos de Redes Neurais Artificiais*

3.1-) Introdução.....	31
3.2-) Topologias de Redes .....	34
3.3-) Treinamento de Redes Neurais Artificiais .....	35
3.4-) Redes de Hopfield .....	36
3.4.1-) A Estabilidade da Rede de Hopfield.....	39
3.4.2-) Rede de Hopfield Discreta .....	41
3.5-) Conclusão .....	42
Bibliografia .....	42

## **Capítulo 4 - Estimação Paramétrica Robusta Através de Redes Neurais**

4.1-) Introdução.....	45
4.2-) Subespaço-Válido e a Rede de Hopfield .....	46
4.2.1-) Derivando os Parâmetros de $T$ e $i^b$ Utilizando-se o Subespaço-Válido...	49
4.3-) Rede de Hopfield Modificada .....	50
4.3.1-) Rede Modificada e o Confinamento no Subespaço-Válido .....	53
4.4-) Mapeando Problemas de Estimação Paramétrica Robusta Através da Rede de Hopfield .....	54
4.4.1-) Subespaço-Válido Para o Problema .....	55
4.4.2-) Cálculo dos Vértices do Politopo.....	58
4.5-) Conclusão .....	65
Bibliografia .....	66

## **Capítulo 5 - Simulações de Problemas de Estimação Paramétrica Robusta Através de Redes Neurais Artificiais**

5.1-) Introdução.....	67
5.2-) Detalhes Sobre os Resultados das Simulações.....	67
5.3-) Problema 1 - Modelo AR de Segunda Ordem .....	69
5.4-) Problema 2 - Modelo MA de Terceira Ordem.....	72
5.5-) Problema 3 - Modelo ARMAX de Segunda Ordem .....	76
5.6-) Problema 4 - Modelo ARMAX de Terceira Ordem.....	79
5.7-) Problema 5 - Modelo ARMAX de Quarta Ordem .....	82
5.8-) Conclusão .....	86
Bibliografia .....	88

## **Capítulo 6 - Linhas Futuras de Pesquisas e Conclusões**

6.1-) Sumário da Dissertação .....	89
6.2-) Conclusões.....	90
6.3-) Linhas Futuras de Pesquisas .....	92

### **Apêndice A**

### **Apêndice B**

### **Apêndice C**

### **Apêndice D**

## Lista de Notação

### Convenções Gerais

A notação de matrizes é feita geralmente utilizando letras itálicas maiúsculas em negrito, e vetores são denotados usando letras itálicas minúsculas em negrito. Os símbolos superscript em letra itálica maiúscula indicam a dimensão de uma matriz ou vetor, e símbolos superscript em letra itálica minúscula denotam o rótulo para um vetor ou matriz. Símbolos subscript em letra itálica minúscula são usados geralmente para distinguir os elementos de um vetor ou matriz. Letras itálicas maiúsculas em formato simples denotam conjuntos.

### Definições Gerais

---

$a_g$ e $b_g$	Delimitadores das funções de ativação sigmoidal e "rampa-simétrica"
$g_r$	Função de ativação "rampa-simétrica"
$E$	Energia total de uma rede de Hopfield
$E^{conf}$	Termo de energia que confina várias restrições ao mesmo tempo
$E^{ot}$	Termo de energia ótimo
$E^{rest}$	Termo de energia que confina uma restrição
$F(\theta)$	Família de Modelos que podem caracterizar o sistema a ser modelado
$M^{eq}$	Número de restrições de igualdade
$M^{in}$	Numero de restrições de desigualdade
$M^{med}$	Número de medidas realizadas sobre o sistema a ser modelado
$N$	Número de parâmetros a ser estimado ou dimensão do espaço paramétrico
$S$	Região de Pertinência Paramétrica
$V$	Conjunto (lista) de vértices de um politopo
$\Omega$	Conjunto Paramétrico admissível à priori
$\nabla_v$	Operador Gradiente em relação ao vetor $v$

---

---

## Vetores

---

$e$	Vetor erro de saída do sistema a ser modelado
$i^b$	Vetor de entrada externa da rede
$i^{ot}$	Vetor de entrada externa do termo de energia $E^{ot}$
$i^{conf}$	Vetor de entrada externa do termo de energia $E^{conf}$
$s$	Vetor de offset do subespaço-válido
$u$	Vetor de estado da rede de Hopfield
$v$	Vetor de saída da rede de Hopfield
$v^{val}$	Vetor que pertence ao subespaço-válido de soluções
$v^{inv}$	Vetor que pertence ao subespaço-inválido (ortogonal a $v^{val}$ )
$y$	Vetor de medidas realizadas sobre o sistema a ser modelado
$\theta$	Vetor de parâmetros estimados

---



---

## Matrizes

---

$I$	Matriz identidade
$\Sigma$	Matriz que especifica a orientação e o tamanho de um elipsóide
$T$	Matriz de conexão da rede de Hopfield
$T^{val}$	Matriz-projeção do subespaço-válido
$T^{inv}$	Matriz-projeção do subespaço-inválido
$T^{ina}$	Matriz-projeção do subespaço-inativo
$T^{ot}$	Matriz de conexão do termo de energia $E^{ot}$
$T^{conf}$	Matriz de conexão do termo de energia $E^{conf}$

---



---

## Rótulos Superscript

---

$b$	refere-se ao termo de entrada externa da rede de Hopfield
$conf$	refere-se ao termo de energia de confinamento de restrições
$ina$	refere-se ao subespaço-inativo
$inv$	refere-se ao subespaço-inválido
$ot$	refere-se ao termo de energia ótimo
$val$	refere-se ao subespaço-válido

---

## **Abreviações**

---

EB	Ellipsoidal-Bounding
EOB	Ellipsoidal Outer-Bounded
EIB	Ellipsoidal Inner-Bounded
IA	Inteligência Artificial
PRBS	Pseudorandom Binary Sequence
RNA	Redes Neurais Artificiais
UBBE	Unknown-But-Bounded Error

---

## Lista de Figuras

2.1 - Processo de Identificação de Sistemas.....	07
2.2 - Região de Pertinência Paramétrica Para Modelos Lineares .....	11
2.3 - Interseção de um Elipsóide Com um Par de Hiperplanos Paralelos .....	13
2.4 - Região Elipsoidal Externa e os Intervalos de Incerteza Associados.....	16
2.5 - Região Elipsoidal Interna e os Intervalos de Incerteza Associados.....	17
2.6 - Ortotopo Externo Alinhado Com os Eixos Paramétricos .....	18
2.7 - Ortotopo Interno Alinhado Com os Eixos Paramétricos .....	19
2.8 - Alteração de Vértices de um Politopo .....	21
2.9 - Introdução de Novas Faces .....	23
2.10 - Interpolação de um Novo Vértice do Politopo .....	23
2.11 - Região de Pertinência Paramétrica Para Modelos Não Lineares.....	26
3.1 - Modelo de um Neurônio Artificial.....	32
3.2 - Funções de Ativação do Neurônio Artificial .....	33
3.3 - Topologias de Redes Neurais.....	34
3.4 - Rede Neural Tipo Hopfield .....	37
3.5 - Diagrama de Estado da Rede de Hopfield .....	38
3.6 - Rede de Hopfield Discreta .....	41
4.1 - Diagrama Esquemático da Rede de Hopfield Contínua .....	46
4.2 - Diagrama Esquemático da Rede de Hopfield Modificada .....	51
4.3 - Função de Ativação "Rampa Simétrica" .....	52
4.4 - Perspectivas de Seleção dos Vetores $i^{of}$ .....	64
4.5 - Região de Pertinência Pelo Método do Politopo Exato- Ex. Ilustrativo .....	64
4.6 - Região de Pertinência Pelo Método de Redes Neurais - Ex. Ilustrativo.....	64
5.1 - Gráfico Ilustrando a Saída do Modelo - Problema 3 .....	79
5.2 - Gráfico Ilustrando a Saída do Modelo - Problema 4 .....	82
5.3 - Gráfico Ilustrando a Saída do Modelo - Problema 5 .....	86
A.1 - Hardware Analógico da Rede de Hopfield.....	93
D.1 - PRBS Através de Registrador de Deslocamento.....	100

## Lista de Tabelas

4.1 - Medidas Simuladas Para o Exemplo Ilustrativo .....	60
4.2 - Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados - Ex. Ilustrativo .....	62
4.3 - Características dos Intervalos de Incerteza Paramétrica - Exemplo Ilustrativo .....	62
4.4 - Vértices da Região de Pertinência Paramétrica - Exemplo Ilustrativo .....	63
4.5 - Estimadores Paramétricos - Exemplo Ilustrativo .....	63
5.1 - Medidas Simuladas Para o Problema 1 .....	69
5.2 - Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados - Problema 1 .....	71
5.3 - Intervalos de Incerteza Paramétrica - Problema 1 .....	71
5.4 - Vértices da Região de Pertinência Paramétrica - Problema 1 .....	71
5.5 - Estimadores Paramétricos - Problema 1 .....	71
5.6 - Valores de Predição Multipasso - Problema 1 .....	72
5.7 - Medidas Simuladas Para o Problema 2 .....	73
5.8 - Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados - Problema 2 .....	74
5.9 - Intervalos de Incerteza Paramétrica - Problema 2 .....	74
5.10 - Vértices da Região de Pertinência Paramétrica - Problema 2 .....	75
5.11 - Estimadores Paramétricos - Problema 2 .....	75
5.12 - Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados - Problema 3 .....	77
5.13 - Intervalos de Incerteza Paramétrica - Problema 3 .....	78
5.14 - Vértices da Região de Pertinência Paramétrica - Problema 3 .....	78
5.15 - Estimadores Paramétricos - Problema 3 .....	78
5.16 - Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados - Problema 4 .....	81
5.17 - Intervalos de Incerteza Paramétrica - Problema 4 .....	81
5.18 - Vértices da Região de Pertinência Paramétrica - Problema 4 .....	81
5.19 - Estimadores Paramétricos - Problema 4 .....	82
5.20 - Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados - Problema 5 .....	84
5.21 - Intervalos de Incerteza Paramétrica - Problema 5 .....	85
5.22 - Estimadores Paramétricos - Problema 5 .....	85
D.1 - Adição em Complemento de Dois .....	101
D.2 - Valores dos Coeficientes de Retroalimentação da PRBS .....	101

# *Capítulo 1*

## **Introdução**

### **1.1-) MOTIVAÇÃO E RELEVÂNCIA DO TRABALHO**

O crescimento do interesse em Redes Neurais Artificiais (RNA) nos últimos anos certamente será visto pelos historiadores da Ciência da Computação como um fenômeno notável dos anos de 1980, do mesmo modo que as pesquisas em Inteligência Artificial (IA) foram marcantes nos anos de 1970. Existe entretanto, uma diferença primordial. Enquanto as pesquisas em IA atraíram majoritariamente os cientistas de computação, RNA também despertou o interesse de pesquisadores de diversas áreas: físicos, matemáticos, engenheiros, psicólogos, neuro-fisiologistas e biólogos.

Enquanto os pesquisadores de IA baseiam-se no paradigma de que, se um conhecimento é adquirido então ele pode ser expresso por um conjunto de regras possível de ser executado num computador, os pesquisadores de RNA acreditam que o cérebro, constituído de neurônios, constrói suas próprias estratégias de processamento de informações, através do que é usualmente chamado de "experiência". Assim, o foco deste novo paradigma é baseado no reconhecimento, por esta comunidade diversa, de que o cérebro processa informações de uma forma diferente do computador convencional [1].

A justificativa principal para a utilização de RNA na solução de problemas complexos e variados é inspirada no próprio cérebro humano, que tem capacidade de resolver, de forma rápida e satisfatória, problemas mal definidos e que exigem um enorme esforço computacional como, por exemplo, reconhecimento de imagens visuais, reconhecimento de voz, etc. A principal semelhança entre o sistema neural biológico e uma RNA é que ambos atingem altas taxas de computação por intermédio de um grande número de processadores (elementos de processamento simples) com alto grau de conectividade entre eles [10].

Algumas das discussões atuais em RNA foram iniciadas cerca de 10 anos antes da invenção do computador digital. O modelo de *neurônio artificial* proposto por McCulloch e Pitts [4] em 1943, é ainda o elemento básico da maioria dos modelos atuais de RNA. Por outro lado, os anos de 1960 foram muito produtivos nesta área, destacando-se neste período por exemplo, os trabalhos de Rosenblatt [7], Widrow e Holf [11], que propuseram as redes *Perceptron* e *Adaline* (Adaptative Linear Element), respectivamente.

Minsky e Papert [5] em 1969, provocaram um esfriamento da pesquisa nesta área quando publicaram o livro *Perceptrons*, onde prova-se que as estruturas utilizadas nos dispositivos da época não eram capazes de aprender regras lógicas tão simples quanto a do *ou-exclusivo*. No entanto, alguns pesquisadores continuaram as pesquisas sobre o assunto, entre eles destacaram-se Teuvo Kohonen (Finlândia), Edoardo Caianiello (Itália), Stephen Grossberg, Bernard Widrow e James Anderson (E.U.A) e Kunihiko Fukushima (Japão).

O ressurgimento do interesse na área durante os anos de 1980, deve-se principalmente a dois eventos independentes. Um dos eventos foi o trabalho desenvolvido por Hopfield [3] sobre o projeto de memórias associativas, e mais tarde sobre a solução de problemas de otimização utilizando um tipo especial de redes recorrentes. O outro foi o desenvolvimento de um novo algoritmo de aprendizado (*backpropagation*) explicitado por Rumelhart, Hinton e Williams [8] para as redes do tipo Perceptron com estrutura multicamadas.

As RNA têm sido utilizadas em várias aplicações, tais como: Reconhecimento de Padrões, Processamento de Sinais, Processamento de Fala, Robótica, etc. Vale ressaltar que trabalhos importantes na área de Controle e Identificação de Sistemas [2,6,9] têm sido também desenvolvidos nos últimos anos.

Esta dissertação é voltada ao estudo de aplicações de técnicas de redes neurais dentro da perspectiva de teoria de sistemas, particularmente da teoria de identificação de sistemas. Avanços na Teoria de Sistemas vêm sendo obtidos através da combinação da matemática, modelagem, computação e experimentação. Estes elementos são também essenciais ao estudo de redes neurais artificiais. Através deles, as diferentes áreas da pesquisa atual em Teoria de Sistemas, incluindo sistemas adaptativos e de aprendizagem, sistemas não lineares, sistemas estocásticos, sistemas hierárquicos-multiníveis, tornam-se diretamente relacionados ao estudo de sistemas neurais.

Este trabalho está relacionado principalmente com o uso de redes neurais na solução de problemas de Estimção Paramétrica Robusta para sistemas lineares. Embora avanços expressivos na área de estimação tenham sido testemunhados nos últimos anos, é necessária a pesquisa de métodos alternativos que explorem arquiteturas de processamento inerentemente paralelas e adaptativas. Além disso, problemas de estimação são frequentemente resolvidos através da aplicação de técnicas de otimização, os quais são essenciais ao desenvolvimento de novos mecanismos de aprendizagem e adaptação de sistemas em geral, e de redes neurais em particular.

## **1.2-) PROPOSTA DA DISSERTAÇÃO**

Neste trabalho propõe-se utilizar Redes Neurais do tipo Hopfield visando solucionar problemas de Estimção Paramétrica Robusta com erro desconhecido-mas-limitado. Os parâmetros internos das Redes para este problema são derivados utilizando a técnica denominada subespaço-válido de soluções.

A solução de um problema de estimação robusta com erro desconhecido-mas-limitado consiste em determinar a região do espaço paramétrico onde todos os parâmetros são consistentes com as estruturas do modelo e da perturbação que são assumidas representar o sistema em estudo. Esta região, no caso de sistemas lineares, resulta da interseção de vários hiperplanos gerados pelas medidas e que por sua vez, formam uma região paramétrica convexa que pode ser expressa pelo conjunto de seus vértices. Portanto, o objetivo da Rede é localizar estes vértices e obter o melhor estimador do sistema contido nesta região, segundo algum critério geométrico relacionado com a região delimitada por estes vértices.

Neste trabalho, os vértices da região paramétrica procurada estão relacionados aos pontos de equilíbrio estáveis na Rede de Hopfield. Assim o estudo das condições de existência e de técnicas para obtenção destes pontos é de suma importância na aplicação das redes de Hopfield para solução do problema de estimação robusta. Dentro deste contexto, descreve-se também uma rede de Hopfield modificada com o intuito de fornecer maior eficiência e confiabilidade na obtenção dos vértices da região estimada.

Além de fornecer um novo método de solução destes problemas, este trabalho (como será visto posteriormente) também fornece duas perspectivas de busca de soluções que são associadas com a rede de Hopfield a ser proposta. Estas perspectivas aplicadas a uma mesma rede substituem dois métodos considerados clássicos na obtenção de solução destes problemas, ou seja, é possível simular o método do ortotopo-externo e do politopo-exato através de uma única rede.

### 1.3-) ORGANIZAÇÃO DA DISSERTAÇÃO

A organização desta dissertação é feita como segue:

**No Capítulo 2**, formula-se o problema de estimação paramétrica robusta sob hipótese de erro desconhecido-mas-limitado. Faz-se uma breve revisão dos métodos existentes na literatura para esta abordagem, destacando-se os métodos baseados no cálculo de regiões elipsoidais, ortotópicas e do politopo-exato.

**No Capítulo 3**, são introduzidos os conceitos básicos relativos a RNA com enfoque principalmente sobre a concepção da Rede de Hopfield.

**No Capítulo 4**, é introduzido o conceito de subespaço-válido de soluções que será usado na derivação dos parâmetros internos da rede. Em seguida faz-se o mapeamento de problemas de Estimação Paramétrica Robusta através de uma Rede de Hopfield

**No Capítulo 5**, a rede proposta anteriormente é usada na solução de vários problemas de identificação de sistemas. A partir de exemplos simulados, é realizado também, uma análise comparativa entre os resultados fornecidos pela rede e os resultados obtidos com os métodos descritos no capítulo 2.

**No Capítulo 6**, faz-se a conclusão do trabalho apresentado.

**BIBLIOGRAFIA**

- [1] Aleksander, L. - **Neural Computing Architectures: The Design of Brain Like Machines**, The MIT Press, Cambridge, Massachusetts, 1989.
- [2] Antsaklis, P. J. - "Neural Networks in Control Systems", *IEEE Control Systems Magazine*, 12 (1992) 8-10.
- [3] Hopfield, J. J. - "Neurons With a Graded Response Have Collective Computational Properties Like Those of Two-State Neurons", *Proceedings of the National Academy of Science, USA*, 81 (1984) 3088-3092.
- [4] McCulloch, W. S. & Pitts, W. - "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics*, 5 (1943) 115-133.
- [5] Minsky, M. & Papert, S. - **Perceptrons: An Introduction to Computational Geometry**, The MIT Press, Cambridge, Massachusetts, 1969.
- [6] Narendra, K. S. & Parthasarathy, K. - "Identification and Control of Dynamical Systems Using Neural Networks", *IEEE Transactions on Neural Networks*, 1 (1990) 4-27.
- [7] Rosenblatt, F. - **Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms**, Spartan Books, Washington - DC, 1959.
- [8] Rumelhart, D. E. & Hinton, G. E. & Williams, R. J. - **Learning Internal Representations by Error Propagation**, Chap 8 (D. E. Rumelhart and J. L. McClelland, Eds), The MIT Press, Cambridge, Massachusetts, 1986.
- [9] Sartori, M. A. & Antsaklis, P. J. - "Implementations of Learning Control Systems Using Neural Networks", *IEEE Control Systems Magazine*, 12 (1992) 49-57.
- [10] Tagliarini, G. A. & Christ, J. F. & Page, E. W. - "Optimization Using Neural Networks", *IEEE Trans. on Computers*, 12 (1991) 1347-1358.
- [11] Widrow, B. & Holf, M. E. - **Adaptative Switching Circuits**, IRE WESCON Con. Record, Part 4, 96, 1960.

## *Capítulo 2*

# **Métodos Clássicos Para Estimação Robusta de Regiões Paramétricas**

### **2.1-) INTRODUÇÃO**

A modelagem matemática de sistemas físicos a partir de informações observadas sobre este sistema, constitui um típico problema de *identificação de sistemas*. Esta modelagem do sistema pode ser realizada através de três passos principais:

- i-) Determinar uma estrutura de modelo que seja adequada ao tipo de sistema físico observado,
- ii-) Calcular os parâmetros desconhecidos deste modelo,
- iii-) Testar e validar o modelo.

A determinação da estrutura de modelo adequada ao sistema observado é realizada a partir de um conjunto de modelos candidatos com base em algum conhecimento a priori sobre o sistema a ser identificado. Quando nenhum conhecimento sobre o sistema é disponível, técnicas específicas podem ser utilizadas para escolher o melhor modelo entre os possíveis modelos candidatos [17].

O cálculo dos parâmetros desconhecidos do modelo constitui o objetivo do processo de *estimação paramétrica*. Em suma, o processo de identificação de sistemas pode ser ilustrado como mostra o fluxograma na *figura 2.1*.

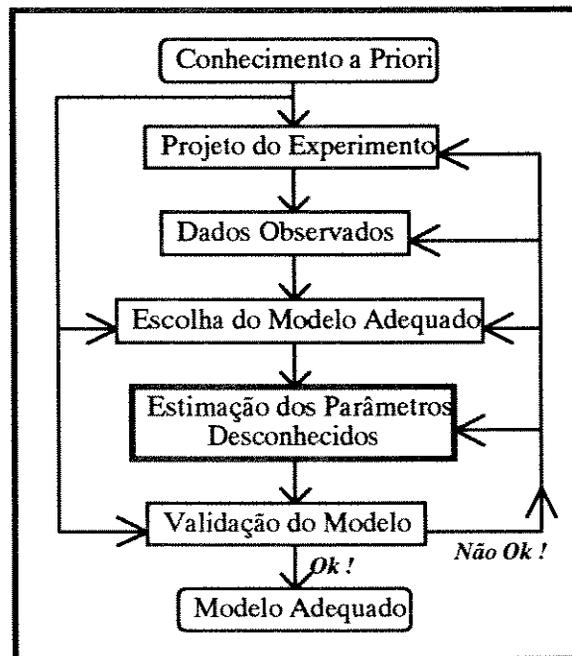


Figura 2.1 - Processo de Identificação de Sistemas

As informações coletadas (medidas) para estimar os parâmetros de um modelo que caracteriza o comportamento de um sistema físico são frequentemente corrompidas pela presença de *incertezas* (ruídos) que representam algum tipo de perturbação associada com o sistema observado. O conhecimento do tipo de incerteza associada a essas informações é de suma importância para se fazer uma avaliação de como estas incertezas afetam as variáveis estimadas.

De um modo geral, as incertezas associadas com as informações coletadas são tipicamente caracterizadas através de um ruído aleatório aditivo com função densidade de probabilidade (fdp) conhecida como por exemplo: um ruído branco gaussiano. Dentro deste contexto, vários métodos podem ser utilizados para estimar os parâmetros de modelos que possuem uma caracterização estatística das incertezas, entre estes destacam-se o estimador de máxima-verossimilhança, mínimos quadrados e o estimador de Bayes [17,25].

Entretanto, esta abordagem que considera incertezas estocásticas possui alguns contra-pontos, entre estes pode-se citar:

i-) o número de medidas disponíveis deve ser suficientemente grande para validar o comportamento assintótico do modelo,

ii-) a função densidade de probabilidade assumida para a incerteza, nem sempre é baseada em informações a priori válidas,

iii-) incertezas de natureza determinística e que correspondem a erros estruturais não são adequadamente descritas por uma função densidade de probabilidade [14].

Além disso, em muitas situações reais, o sistema que gera os dados é muito complexo, possuindo não-linearidades, parâmetros variantes no tempo e mesmo aberrações (outliers). Nestes casos, torna-se necessário o uso de modelos simplificados onde as incertezas são caracterizadas de forma mais simples.

Uma alternativa atrativa para a descrição estocástica de erros é caracterizá-los apenas através de *limites inferiores e superiores*. Uma motivação para esta abordagem, denominada UBBE ( Unknown-But-Bounded Error Approach), é o fato que em muitos casos práticos a descrição UBBE é mais realística e menos exigente que a abordagem estatística. A abordagem de erro desconhecido-mas-limitado foi introduzida no domínio de estimação de estado por Witsenhausen [31] e Schweppe [25] em 1968, e atualmente vários trabalhos já foram publicados na área [5,15,16,26]. Dentro do contexto de estimação paramétrica, também já existem na literatura, trabalhos que supõem que a perturbação atuando no processo é limitada [2,3,7,18,21,22,27,28].

Ao contrário da abordagem estocástica onde os métodos de estimação fornecem um vetor de parâmetros únicos para o sistema, um método de estimação robusta no contexto UBBE calcula uma *região de pertinência* (membership set) para os parâmetros reais, que seja compatível com as medidas e com os limites supostamente conhecidos das perturbações. Isto tem significado prático para aplicações de controle, quando a informação disponível sobre a perturbação é fornecida como limites sobre seu valor absoluto; em aplicações onde os dados medidos não são suficientes para garantir resultados assintóticos confiáveis, e para diagnósticos de falhas, neste caso as perturbações correspondem a flutuações dos valores dos componentes em torno de seus valores nominais [1].

Esta abordagem UBBE está começando a ser utilizada em diversas aplicações [1]: farmacocinética, previsão de séries temporais, química, biologia, processamento de sinais, controle robusto e adaptativo, restaurações de imagens, etc.

## 2.2-) FORMULAÇÃO DO PROBLEMA

Admitindo-se que um sistema real possa ser representado por um modelo com  $N$  parâmetros desconhecidos e que sejam realizadas  $M^{med}$  medidas da saída deste sistema, então o modelo para o sistema real pode ser escrito matematicamente em termos de uma relação linear ou não, entre um vetor paramétrico real  $\theta^* = [\theta_1^* \ \theta_2^* \ \dots \ \theta_N^*]^T$  no espaço  $N$ -dimensional e um vetor de medidas da saída do sistema  $y = [y_1 \ y_2 \ \dots \ y_{M^{med}}]^T$  no espaço  $M^{med}$ -dimensional. Além disso, este modelo é corrompido aditivamente por um vetor de incertezas que pode representar os erros de modelagem, erros estruturais, erros de medidas, ruídos externos, etc. Portanto, tem-se:

$$y(k) = F(\theta^*) + e(k) \quad k = 1, \dots, M^{med} \quad (2.1)$$

onde  $k$  representa a  $k$ -ésima medida realizada,  $F(\theta^*)$  é uma família de modelos condicionados por  $\theta$ , e capaz de representar os dados medidos no sistema real.

Na abordagem UBBE, o erro  $e(k) \in \Re^{M^{med}}$  deve satisfazer o seguinte critério:

$$e(k) \in [e^{min}(k), e^{max}(k)] \quad (2.2)$$

onde  $e^{min}$  e  $e^{max}$  são limites supostamente conhecidos para o vetor de incerteza.

Dado  $\Omega$  ser o conjunto admissível (domínio) a priori dos parâmetros, um método de estimação paramétrica robusta sob hipótese de erro desconhecido-mas-limitado calcula o conjunto  $S$  de todos os valores de  $\theta$  admissíveis que sejam consistentes com (2.1) e (2.2).  $S$  é então definido como a interseção entre  $\Omega$  e o conjunto  $R(k)$  das soluções para  $\theta$  das  $2 \cdot M^{med}$  desigualdades:

$$R(k) = \{ \theta \mid y(k) - e^{max}(k) \leq F(\theta) \leq y(k) - e^{min}(k) \}, \quad k = 1, \dots, M^{med} \quad (2.3)$$

$$S = \Omega \cap R(k), \quad k = 1, \dots, M^{med} \quad (2.4)$$

Quando nenhum conhecimento a priori for disponível sobre os parâmetros a serem estimados, o conjunto  $\Omega$  será o próprio conjunto  $\mathbf{R}^N$ , e o conjunto  $S$  em (2.4) passa a ser definido como  $S = \cap R(k)$ .

## 2.3-) MÉTODOS (UBBE) PARA ESTIMAÇÃO DE MODELOS LINEARES

Considerando-se que o sistema real possa ser representado por uma relação linear, então  $F(\theta^*)$  pode ser escrito como um modelo de regressão linear:

$$F(\theta^*) = \phi^T(k) \cdot \theta^* \quad (2.5)$$

onde  $\phi^T$  é uma matriz de regressão contendo medidas do sistema real com dimensão  $M^{med} \times N$  ( $\phi^T(k)$  é um vetor coluna de  $\phi$ ), portanto, a equação (2.1) torna-se:

$$y(k) = \phi^T(k) \cdot \theta^* + e(k) \quad (2.6)$$

Substituindo (2.5) em (2.3), tem-se:

$$R(k) = \{\theta \mid y(k) - e^{max}(k) \leq \phi^T(k) \cdot \theta \leq y(k) - e^{min}(k)\}, \quad k = 1, \dots, M^{med} \quad (2.7)$$

Todos os parâmetros  $\theta$  que são consistentes com a estrutura do modelo (2.6), com os limites a priori dos erros (2.2) e com as medidas  $y(k)$  pertencem a um conjunto que pode ser visto como a região do espaço paramétrico que é delimitada pelos  $k$  pares de hiperplanos  $H_1(k)$  e  $H_2(k)$ , definidos por:

$$H_1(k) = \{\theta \mid \phi^T(k) \cdot \theta = y(k) - e^{min}(k)\} \quad (2.8)$$

$$H_2(k) = \{\theta \mid \phi^T(k) \cdot \theta = y(k) - e^{max}(k)\} \quad (2.9)$$

Cada hiperplano  $H_i(k)$ ,  $i = 1, 2$ , divide o espaço paramétrico em dois semi-espacos  $H_i^+(k)$  e  $H_i^-(k)$ , tal que :

$$H_1^+(k) = \{\theta \mid \phi^T(k) \cdot \theta \leq y(k) - e^{min}(k)\} \quad (2.10)$$

$$H_1^-(k) = \{\theta \mid \phi^T(k) \cdot \theta > y(k) - e^{min}(k)\} \quad (2.11)$$

e

$$H_2^+(k) = \{\theta \mid \phi^T(k) \cdot \theta \geq y(k) - e^{max}(k)\} \quad (2.12)$$

$$H_2^-(k) = \{\theta \mid \phi^T(k) \cdot \theta < y(k) - e^{max}(k)\} \quad (2.13)$$

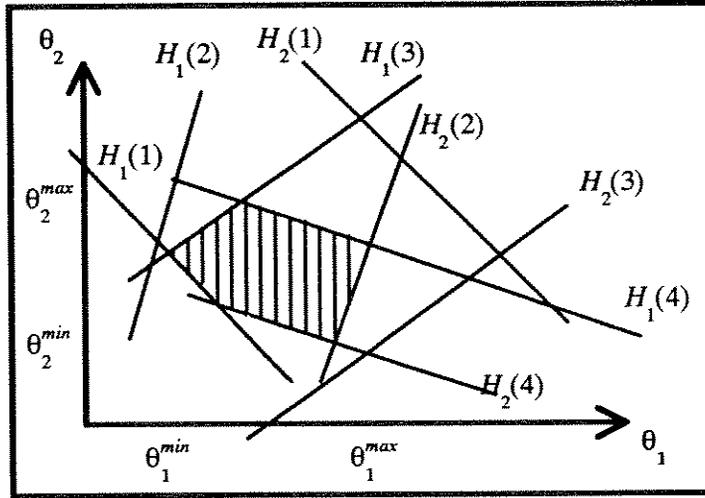


Figura 2.2 - A Região de Pertinência Paramétrica Para Modelos Lineares

Então, o conjunto  $S$  será dado por:

$$S = \Omega \cap \left( \bigcap_{k=1}^{M^{med}} R(k) \right), \quad \text{onde } R(k) = H_1^+(k) \cap H_2^+(k) \tag{2.14}$$

A equação (2.14) representa uma sequência monótona não crescente de regiões de forma convexa, como ilustrado na **figura 2.2** para o caso de  $N = 2$ ,  $M^{med} = 4$  e  $\Omega = \mathbf{R}^2$ . Esta região é denominada **Região de Pertinência Paramétrica** ou **Região Politépica** no caso de modelos lineares. Qualquer vetor de parâmetros  $\theta$  pertencente ao conjunto  $S$  é uma estimação válida de  $\theta^*$ . Na prática, o centro de  $S$  em algum sentido geométrico é escolhido como o "melhor" estimador dos parâmetros reais. De fato, a região de pertinência paramétrica  $S$  pode ser considerada como o conjunto de todos os estimadores verdadeiros e admissíveis para o modelo que representa o sistema real.

Uma outra definição importante relacionada a forma da região  $S$  é o conceito de **Intervalos de Incerteza Paramétrica**. Estes intervalos correspondem (veja **figura 2.2**) aos valores mínimo ( $\theta_i^{min}$ ) e máximo ( $\theta_i^{max}$ ) associados a cada parâmetro dentro da região de pertinência paramétrica  $S$ . Na abordagem UBBE, os intervalos de incerteza paramétrica são usados para caracterizar o grau de confiança associado aos estimadores: quanto menor o intervalo de incerteza associado a um parâmetro, menor será o erro de estimação cometido na substituição do parâmetro real por seu valor estimado.

A determinação da região politópica  $S$  torna-se complicada quando o número de medidas aumenta e conseqüentemente aumenta o número de vértices de  $S$ . Baseado nisto, uma aproximação da Região  $S$  por uma região de forma mais simples, muitas vezes torna-se necessário. Para se ter uma idéia da necessidade de uma aproximação mais simples para o politopo  $S$ , considere por exemplo, que qualquer interseção de hiperplanos linearmente independentes possa ser um vértice de  $S$ , então o número máximo de vértices ( $nv_{max}$ ) é dado por:

$$nv_{max} = 2 \cdot N \cdot \binom{M^{med}}{N} \tag{2.15}$$

onde:  $\binom{i}{j} = \frac{i!}{j! \cdot (i-j)!}$

No caso de modelos com 4 parâmetros e 10 observações o número máximo de vértices será de 1680, para 5 parâmetros e 15 observações este número ( $nv_{max}$ ) corresponde a 30030 vértices, e com 6 parâmetros e 20 observações obtém-se 465120 vértices.

Em geral, as regiões de forma mais simples usadas nas aproximações da região paramétrica  $S$  são *elipsóides* ou *ortotopos* (hipercubos). Estas regiões podem *conter* (outer-bound) ou estar *contidas* na região paramétrica  $S$ . O elipsóide/ortotopo com *volume mínimo* contendo a região de pertinência paramétrica, ou o elipsóide/ortotopo com *volume máximo* e que está contido na região de pertinência paramétrica são geralmente considerados escolhas "ótimas" [19,27].

A aproximação ou a determinação da região de pertinência  $S$  pode então ser dividida em 3 famílias de algoritmos segundo a forma geométrica considerada: elipsoidal (aproxima  $S$  por elipsóides), ortotópica (aproxima  $S$  por hipercubos) e politópica-exata (determinação exata da região  $S$ ). Estas famílias de algoritmos são apresentadas nas seções seguintes deste capítulo.

### **2.3.1-) Aproximação do Politopo $S$ Por uma Região Elipsoidal**

Nesta abordagem considera-se que à região politópica  $S$  é aproximada por uma elipsóide. Esta aproximação foi introduzida por Schweppe e outros [12,24,25,26] no contexto de estimação de estado e, alguns trabalhos foram desenvolvidos no caso específico de estimação paramétrica [1,3,10,11,22]. Os métodos que fazem a aproximação da região

politópica  $S$  por elipsóides podem ser divididos em dois grupos distintos: métodos que calculam o elipsóide externo e métodos que calculam o elipsóide interno à região  $S$ .

**MÉTODOS DE APROXIMAÇÃO POR ELIPSÓIDES EXTERNO**

O elipsóide calculado delimita uma região paramétrica externa à região politópica  $S$ . Os métodos utilizados para calcular este elipsóide caem dentro de uma mesma estrutura geral e diferem um do outro pelo critério utilizado na minimização do elipsóide. Estes métodos são recursivos e funcionam do seguinte modo: Seja  $E(k-1)$  uma região elipsoidal no espaço paramétrico obtido das  $(k-1)$  medidas anteriores, então o novo elipsóide  $E(k)$  que é menor que  $E(k-1)$  segundo algum critério geométrico, contém a interseção de  $E(k-1)$  com a região do espaço paramétrico delimitado pelos dois hiperplanos paralelos formando  $R(k)$ , e gerados pela medida atual  $y(k)$  em (2.7), como ilustrado na *figura 2.3*.

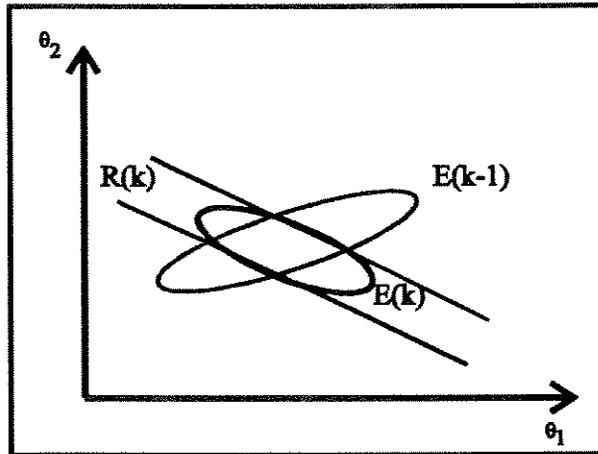


Figura 2.3 - Interseção de  $E(k-1)$  e  $R(k)$  gerando  $E(k)$

Os métodos que aproximam a região de pertinência  $S$  por uma região elipsoidal externa são conhecidos na literatura por "Algoritmos EOB (Ellipsoidal-Outer-Bounding)" e possuem a característica principal de minimizar o tamanho geométrico do elipsóide.

Os algoritmos da abordagem EOB consistem em determinar recursivamente a partir de um elipsóide inicial  $E(0)$ , uma sequência de elipsóides  $E(k)$  que podem ser concisamente descritos como:

$$E(k) = \left\{ \theta \mid [\theta - \theta_k]^T \cdot \Sigma_k^{-1} \cdot [\theta - \theta_k] \leq 1 \right\}, \tag{2.16}$$

onde:  $\theta_k$  é o centro do elipsóide  $E(k)$ , e

$\Sigma_k^{-1}$  é uma matriz que especifica o tamanho geométrico do elipsóide e sua orientação.

A determinação de  $\theta_0$  e  $\Sigma_0^{-1}$  definindo o elipsóide inicial  $E(0)$  é baseado no conhecimento a priori sobre o sistema a ser identificado. Se nenhum conhecimento for disponível, o elipsóide  $E(0)$  poderá ser inicializado com:

$$\theta_0 = \mathbf{0} \quad (2.17)$$

$$\Sigma_0 = \gamma \cdot I^N \quad (2.18)$$

onde  $\gamma$  é um número suficientemente grande tal que  $E(0)$  contenha toda região politópica  $S$ . Nesta abordagem, o centro  $\theta_k$  do elipsóide resultante de  $k$ -medidas realizadas sobre o sistema geralmente é escolhido como o melhor estimador de  $\theta^*$ .

No contexto de estimação paramétrica, os dois algoritmos apresentados por Fogel e Huang em [11] são os mais utilizados.

O primeiro algoritmo procura minimizar o volume do elipsóide (2.16), enquanto o segundo procura minimizar a soma dos quadrados dos eixos deste mesmo elipsóide. Dado que  $\Sigma$  é uma matriz que reflete o tamanho geométrico do elipsóide, o primeiro método minimiza o *determinante* de  $\Sigma$ , enquanto o segundo método minimiza o *traço* da matriz  $\Sigma$ . Estes algoritmos são chamados respectivamente de algoritmo de *volume mínimo* e algoritmo de *traço mínimo* e são inicializados tal que  $E(0)$  seja um elipsóide grande o suficiente para garantir que toda a região politópica  $S$  esteja contida nela.

Quando apenas um dos hiperplanos que limita  $R(k)$  intersecciona  $E(k-1)$ , o elipsóide  $E(k)$  calculado com o algoritmo de volume mínimo não é ótimo. O volume deste elipsóide pode ser reduzido pela substituição do hiperplano que não intersecciona  $E(k-1)$  por um hiperplano paralelo e que seja tangente a  $E(k-1)$ . Este algoritmo é denominado de "Algoritmo de Volume Mínimo Melhorado" [3,23].

Assim, como exemplo de métodos elipsoidais, apresenta-se a seguir o algoritmo do volume mínimo melhorado. Supõe-se, por simplicidade de notação que  $e^{max}(k) = e^{min}(k) = e(k)$ .

Resumo do Algoritmo de Volume Mínimo Melhorado:

- (1) Inicializar  $\theta_0$  e  $\Sigma_0^{-1}$
- (2) Obter a Medida  $y(k)$  e Formar o Vetor  $\phi(k)$
- (3) Calcular os Dois Indicadores

$$(3.1) \alpha_k^- = \frac{\phi^T(k) \cdot \theta_{k-1} - y(k) - e(k)}{\sqrt{\phi^T(k) \cdot \Sigma_{k-1} \cdot \phi(k)}}$$

$$(3.2) \alpha_k^+ = \frac{y(k) - \phi^T(k) \cdot \theta_{k-1} - e(k)}{\sqrt{\phi^T(k) \cdot \Sigma_{k-1} \cdot \phi(k)}}$$

- (4) IF  $\alpha_k^- > 1$  ou  $\alpha_k^+ > 1$   
 THEN  $E(k) = 0$  ; GO TO (8)  
 ELSE (4.1)  $\alpha_k^- = \max(\alpha_k^-, -1)$   
 (4.2)  $\alpha_k^+ = \max(\alpha_k^+, -1)$
- (5) IF  $\alpha_k^- \cdot \alpha_k^+ \geq 1/N$   
 THEN  $E(k) = E(k-1)$  ; GO TO (7)

- (6) IF  $\alpha_k^+ = \alpha_k^-$

THEN (6.1)  $\theta_k = \theta_{k-1}$

$$(6.2) \Sigma_k = \frac{N(1-\alpha_k^2)}{N-1} \left( \Sigma_{k-1} - \frac{1-N\alpha_k^2}{(1-\alpha_k^2) \cdot \phi^T(k) \cdot \Sigma_{k-1} \cdot \phi(k)} \Sigma_{k-1} \cdot \phi(k) \cdot \phi^T(k) \cdot \Sigma_{k-1} \right)$$

ELSE (6.3) Calcular as Variáveis

$$(6.3.1) \rho_k = \sqrt{4(1-(\alpha_k^-)^2)(1-(\alpha_k^+)^2) + N^2((\alpha_k^+)^2 - (\alpha_k^-)^2)^2}$$

$$(6.3.2) \delta_k = \frac{N^2}{(N^2-1)} \left( 1 - \frac{(\alpha_k^+)^2 + (\alpha_k^-)^2 - \rho_k/N}{2} \right)$$

$$(6.3.3) \alpha_k = \frac{1}{N+1} \left( N + \frac{2}{(\alpha_k^+ - \alpha_k^-)^2} (1 - \alpha_k^+ \cdot \alpha_k^- - \rho_k/2) \right)$$

$$(6.4) \theta_k = \theta_{k-1} + \frac{\alpha_k(\alpha_k^+ - \alpha_k^-)}{2 \cdot \sqrt{\phi^T(k) \cdot \Sigma_{k-1} \cdot \phi(k)}} + \Sigma_{k-1} \cdot \phi(k)$$

$$(6.5) \Sigma_k = \delta_k \left( \Sigma_{k-1} - \frac{\alpha_k}{\phi^T(k) \cdot \Sigma_{k-1} \cdot \phi(k)} \Sigma_{k-1} \cdot \phi(k) \cdot \phi^T(k) \cdot \Sigma_{k-1} \right)$$

- (7) IF Existem Mais Medidas  
 THEN Incrementar  $k$  ; GO TO (2)
- (8) STOP

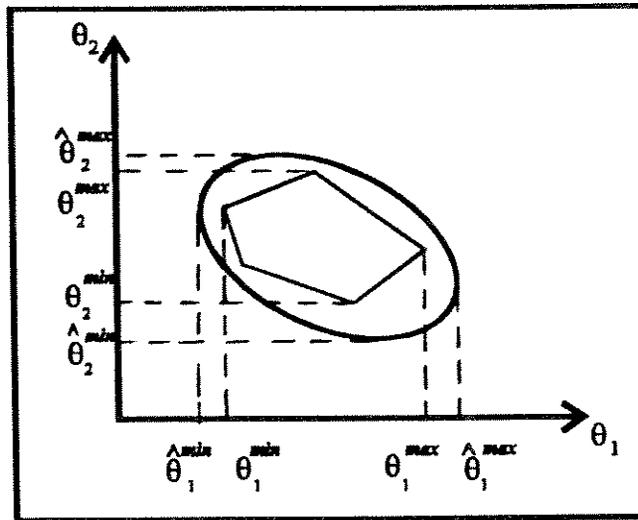


Figura 2.4 - Região Elipsoidal Externa e os Intervalos de Incerteza

As principais vantagens dos algoritmos EOB são:

- i-) Uma representação matemática simples da região calculada.
- ii-) Utilização para aplicações em tempo real devido ao caráter recursivo dos algoritmos.
- iii-) Uma forte semelhança com o conhecido algoritmo dos Mínimos Quadrados Recursivo, o que facilita a análise das propriedades de convergência e estabilidade dos algoritmos EOB.

A principal desvantagem dos algoritmos EOB é que o elipsóide resultante não é o menor elipsóide que contém a região  $S$ . A cada medida  $y(k)$  realizada, existe um erro de aproximação no cálculo do elipsóide  $E(k)$ , visto que na obtenção do novo elipsóide o hiperplano delimitado por  $R(k)$  (figura 2.3) é substituído por um elipsóide degenerado. Este erro de aproximação é propagado e incrementado de iteração para iteração, portanto o elipsóide resultante frequentemente conduz a uma representação pessimista de  $S$  como mostrado na figura 2.4.

Uma outra desvantagem dos algoritmos EOB, como também mostra a figura 2.4, é o fato de que nem todos os estimadores paramétricos contidos na região elipsoidal resultante pode ser considerado como uma estimação válida de  $\theta^*$ . Além disso, vale

salientar que os intervalos de incerteza paramétrica são maiores comparados com os intervalos da região politópica  $S$  (figura 2.4), ou seja:

$$\hat{\theta}_i^{max} > \theta_i^{max} \quad e \quad \hat{\theta}_i^{min} < \theta_i^{min}$$

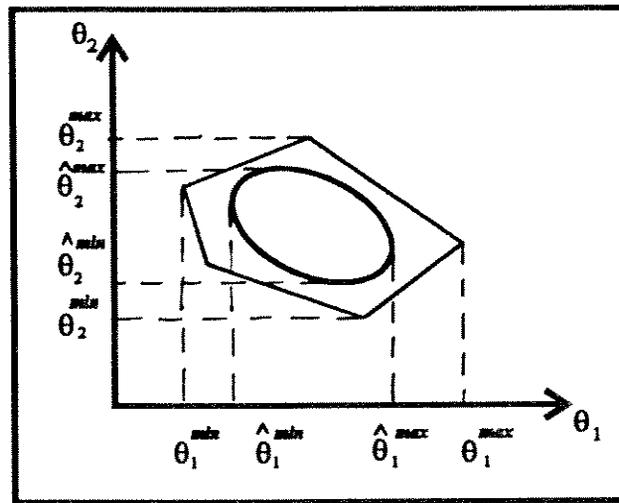


Figura 2.5 - Região Elipsoidal Interna e os Intervalos de Incerteza

A seguir, apresenta-se o segundo grupo de métodos elipsoidais.

### MÉTODOS DE APROXIMAÇÃO POR ELIPSÓIDES INTERNOS

O elipsóide calculado delimita uma região paramétrica interna a região politópica  $S$  como ilustrado na figura 2.5. Os métodos deste grupo são conhecidos como "Algoritmos EIB (Ellipsoidal-Inner-Bouding)".

Estes métodos são similares aos algoritmos EOB , ou seja, a partir de um elipsóide  $E(k-1)$  obtido das  $(k-1)$  medidas anteriores, calcula-se um novo elipsóide  $E(k)$  de volume máximo e que está contido na interseção de  $E(k-1)$  com  $R(k)$ .

Algoritmos EIB possuem a vantagem de serem simples e recursivos, mas dependendo do número de medidas, há uma tendência do elipsóide resultante encolher-se e até mesmo desaparecer, fato este que determina uma menor utilização dos algoritmos EIB. Por outro lado, ao contrário dos algoritmos EOB, todos os vetores paramétricos contidos na região elipsoidal delimitada pelos algoritmos EIB são uma estimativa válida de  $\theta^*$ , por conseguinte, os intervalos de incerteza paramétrica são menores se comparados com os

intervalos reais calculados a partir da região  $S$  (figura 2.5). Uma descrição completa de um dos algoritmos EIB pode ser encontrado em [22].

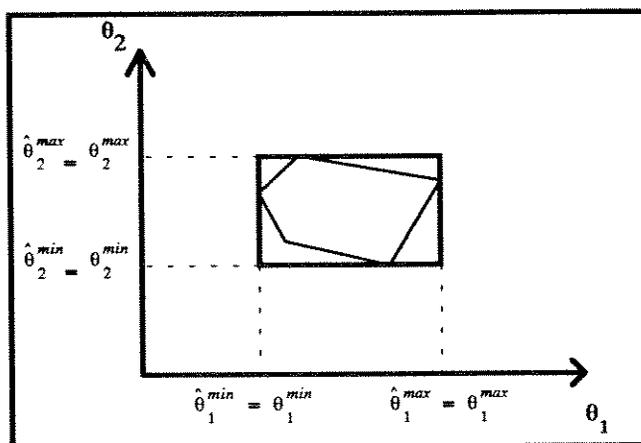


Figura 2.6 - Ortotopo Externo Alinhado Com os Eixos

### 2.3.2-) Aproximação da Região Politópica por uma Região Ortotópica

Nesta abordagem considera-se que a região politópica  $S$  é aproximada por um hipercubo alinhado com os eixos paramétricos (ortotopo). Os métodos desta abordagem também podem ser divididos em dois grupos distintos: Métodos que calculam o ortotopo externo e métodos que calculam o ortotopo interno à região politópica  $S$ .

Os métodos pertencentes ao primeiro grupo calculam o ortotopo de volume mínimo que contém a região politópica  $S$ . Cada limite do ortotopo externo pode ser obtido a partir da solução de  $2.N$  problemas de programação linear, cada um com  $N$  variáveis e  $2.M^{med}$  desigualdades lineares. Vale ressaltar que estes métodos conduzem ao cálculo de intervalos de incerteza idênticos ao da região  $S$  (figura 2.6).

Os métodos do segundo grupo calculam o ortotopo de volume máximo e que está contido na região politópica  $S$  (figura 2.7). O ortotopo de volume máximo também é calculado resolvendo um problema de programação linear com  $(N+1)$  variáveis e  $(2.M^{med} + 1)$  desigualdades. Como no caso EIB, os intervalos de incerteza calculados são menores que os intervalos fornecidos pela região  $S$  e alguns possíveis estimadores paramétricos não estão contidos no ortotopo calculado.

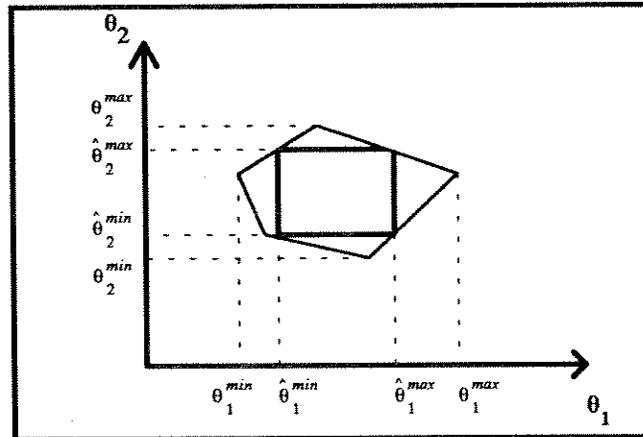


Figura 2.7 - Ortotopo Interno Alinhado Com os Eixos

A principal desvantagem da abordagem ortotópica é que os métodos são não recursivos, dificultando assim, implementações em tempo real. Nesta abordagem, os estimador central expressado por Milanese e Tempo em [19], geralmente é escolhido como o melhor estimador de  $\theta^*$ . O estimador central é definido por:

$$\hat{\theta}_i^{central} = \frac{\hat{\theta}_i^{max} + \hat{\theta}_i^{min}}{2}, \quad i = 1, \dots, N \tag{2.19}$$

Métodos pertencentes ao primeiro grupo podem ser encontrados em [10] e [18], enquanto os do segundo grupo são discutidos em [27].

### 2.3.3-) Descrição Exata da Região Politópica

A representação exata da Região politópica  $S$  pode ser feita através de dois métodos principais. No primeiro método, apresentado por Broman e Shensa em [6] e [7], o politopo  $S$  é representado por uma lista de faces de dimensão  $(N-1)$ . Cada face  $j$ -dimensional é composta por uma lista de faces de dimensão  $(j-1)$ . Assim, o politopo  $S$  é descrito por um conjunto de todas suas faces de dimensão  $j(j=0, \dots, N-1)$ , onde as faces ( $f$ ) são definidas como triplas, ou seja:

$$f = (j, L, T) \tag{2.20}$$

onde:  $j$  é a dimensão de  $f$ ,

$L$  é um conjunto de inteiros que rotulam as subfaces que compõem  $f$ ,

$T$  define a natureza de  $f$ : se  $f$  é um vértice ( $j = 0$ ),  $T$  contém suas coordenadas, se  $f$  é um hiperplano,  $T$  exprime sua equação.

No segundo método, apresentado por Mo e Norton em [21], a região politópica  $S$  é representada por dois conjuntos de índices associados a cada vértice ( $v$ ), ou seja:

$$v \overset{\Delta}{=} (v-v, v-f) \tag{2.21}$$

onde:  $v-v$  é a lista de todos os vértices adjacentes ao vértice  $v$ ,

$v-f$  é a lista dos hiperplanos que cortam  $v$ .

Vale ressaltar que estes métodos fornecem uma caracterização exata de  $S$ , são recursivos, e conduzem a intervalos de incerteza paramétrica exatos. Nesta abordagem, o centro de massa de  $S$ , geralmente é escolhido como o melhor estimador de  $\theta^*$ .

Já que o primeiro método necessita de mais operações lógicas e produz uma representação muito mais complexa de  $S$  se comparado com o método proposto por Mo e Norton, utiliza-se este segundo no capítulo 5, onde seu desempenho será comparado com o do algoritmo proposto nesta tese. Antes de se descrever o algoritmo do segundo método, apresenta-se em seguida algumas definições matemáticas essenciais para a compreensão do funcionamento do algoritmo de Mo e Norton.

### **DEFINIÇÃO E PROPRIEDADES DOS POLITOPOS**

**Definição 1 :** Um conjunto poliédrico  $S \in \mathbf{R}^N$  é a interseção de uma família finita de semi-espacos do  $\mathbf{R}^N$  [13].

**Definição 2 :** Dados dois pontos pertencentes a um conjunto  $S$ , este conjunto é considerado *convexo* se todos os pontos que compõem o segmento de reta entre os dois pontos pertencem ao mesmo conjunto, ou matematicamente:

$$(1-\lambda).\theta_j + \lambda\theta_g \in S, \text{ tal que } 0 \leq \lambda \leq 1 \text{ e } \theta_j, \theta_g \in S \tag{2.22}$$

**Definição 3 :** Um *politopo*  $P$  é a envoltória convexa (convex hull) de seu conjunto de vértices [21] :

$$P = \text{conv}(V), \tag{2.23}$$

onde  $V = \{v_1 v_2 \dots v_x\}$  com  $v_i \in \mathbf{R}^N$ ,  $i = 1, 2, \dots, x$ .

**Definição 4 :** Se  $H$  é um hiperplano que contribui para a formação de um conjunto convexo fechado, a interseção de  $H$  com  $P$  determina uma face de  $P$ .

A definição 3 será relevante quando discutir-se-á o algoritmo de alteração dos vértices do politopo  $S(k-1)$  para inclusão da nova face gerada pela medida  $y(k)$ , mas deve-se analisar antes disso, como os limites (2.5) geram um politopo.

Escreve-se  $R(k)$ , dado em (2.7) como:

$$R(k) = \{ \theta \mid c^+(k) \geq \phi^T(k) \cdot \theta \geq c^-(k) \}, \quad \theta \text{ e } \phi(k) \in \mathbb{R}^N, \quad (2.24)$$

onde:  $c^+(k) = y(k) - e^{\min}(k)$  e  $c^-(k) = y(k) - e^{\max}(k)$ .

Pela definição 1,  $S = \bigcap_{k=1}^{M^{med}} R(k)$  é um conjunto poliédrico. Todo conjunto poliédrico limitado gera um politopo [13], então  $S$  é um politopo se e somente se  $M^{med} \geq N$  e  $N$  elementos de  $\{\phi(k)\}$  são linearmente independentes. Por outro lado, o conjunto  $S$  é também convexo, visto que a interseção de uma família finita de conjuntos convexos é convexo [13], e pela definição 2, cada  $R(k)$  é convexo.

### Método de Mo e Norton [21]

O método recursivo para cálculo da região de pertinência paramétrica proposto por Mo e Norton, consiste em calcular a cada iteração o politopo  $S(k)$  resultante da interseção entre o politopo  $S(k-1)$ , obtido com as  $(k-1)$  medidas anteriores, e o hiperplano  $R(k) = H_1^+(k) \cap H_2^+(k)$ , gerado pela nova medida.

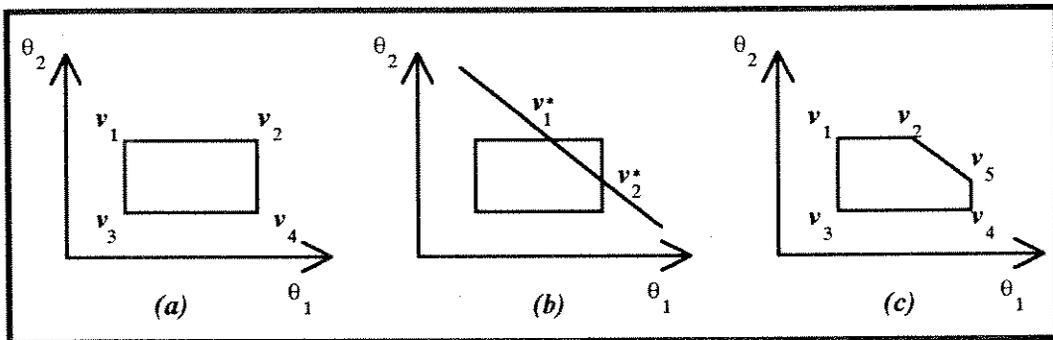


Figura 2.8 - Alterando os Vértices do Politopo Exato

Como exposto anteriormente, este método descreve um politopo através do conjunto de seus vértices (2.21). Assim, a lista de vértices do politopo  $S(k)$  é formada por todos os vértices de  $S(k-1)$  contidos na região  $S(k-1) \cap R(k)$  e pelos novos vértices obtidos nos pontos onde o hiperplano  $R(k)$  intercepta as faces de  $S(k-1)$ . Este procedimento é ilustrado na **figura 2.8**, onde  $N=2$  e  $M^{med} = 5$ . Nesta figura, para alterar o conjunto de vértices  $V$  pertencentes a  $P$  (**figura 2.8-a**) quando uma nova observação é feita, os vértices que serão rejeitados por qualquer nova face devem ser identificados (**figura 2.8-b**), e novos vértices serão formados onde o hiperplano corta as faces existentes (**figura 2.8-c**). Portanto, os vértices adjacentes a cada novo vértice  $v$  devem ser listados, e também as faces que interseptam este. O algoritmo que é descrito a seguir guarda a lista de todos os vértices adjacentes a cada vértice (listas  $v-v$ ) do politopo  $P$  e a lista dos hiperplanos (faces) a que cada vértice pertence (listas  $v-f$ ).

### INTRODUÇÃO DE NOVAS FACES

Dado  $H_1^+(k)$  e  $H_2^+(k)$  serem os semi-espacos formados pela última observação  $y(k)$ :

$$\begin{aligned} H_1^+(k) &= \{\theta \mid \phi^T(k) \cdot \theta \leq y(k) - e^{\min}(k) = c^+(k)\} \\ H_2^+(k) &= \{\theta \mid \phi^T(k) \cdot \theta \leq y(k) - e^{\max}(k) = c^-(k)\} \end{aligned} \quad (2.25)$$

denota-se suas fronteiras por:

$$\begin{aligned} H_1(k) &= \{\theta \mid \phi^T(k) \cdot \theta = c^+(k)\} \\ H_2(k) &= \{\theta \mid \phi^T(k) \cdot \theta = c^-(k)\} \end{aligned} \quad (2.26)$$

As posições de todos os vértices em  $V$  com relação a  $H_1$  são indicados por  $\alpha^+$ , onde:

$$\alpha^+ = [\alpha_1^+ \ \alpha_2^+ \ \dots \ \alpha_x^+]^T, \quad \alpha_i^+ = c^+(k) - \phi^T(k) \cdot v_i, \quad i = 1, 2, \dots, x \quad (2.27)$$

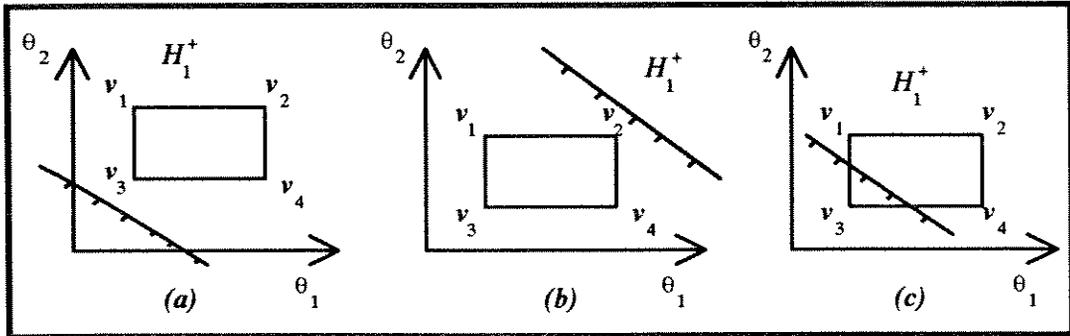


Figura 2.9 - Introdução de Novas Faces

Se  $\alpha_i^+$  é não negativo,  $v_i$  está em  $H_1^+(k)$ ; se negativo, fora de  $H_1^+(k)$ ; se zero, em  $H_1$ . O conjunto  $S(k-1)$  não é modificado por  $H_1^+(k)$  se  $\alpha^+$  é completamente não negativo (figura 2.9-a),  $S(k)$  será vazio se  $\alpha^+$  é todo negativo (figura 2.9-b), e  $H_1^+(k)$  intersecciona e modifica  $S(k-1)$  se alguns mas não todos elementos de  $\alpha^+$  são negativos (figura 2.9-c). Similarmente é feito o mesmo procedimento para  $H_2^+(k)$  com mudanças apropriadas de sinais. Daqui por diante para simplificar, utiliza-se apenas  $H_1^+(k)$ .

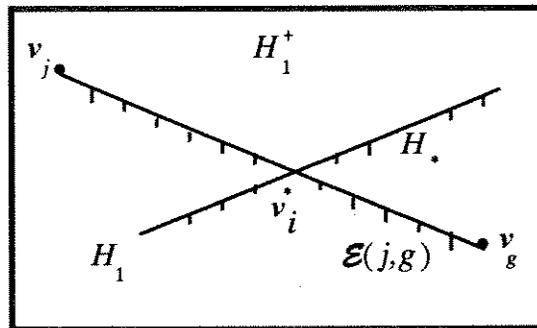


Figura 2.10 - Interpolação de Um Novo Vértice

O politopo é alterado da seguinte forma. O hiperplano  $H_1(k)$  corta  $S(k-1)$ , formando uma nova face  $H_2(k)$  de  $S(k)$ . A face  $H_2(k)$  pela definição 3 pode ser considerada a envoltória convexa de seu conjunto de vértices  $V^*$ , em que cada vértice é um vértice de  $S(k)$  mas não de  $S(k-1)$ . Cada aresta  $\mathcal{E}(j,g)$  de  $S(k-1)$  é uma combinação convexa de dois vértices adjacentes  $v_j$  e  $v_g$  :

$$\mathcal{E}(j,g) = \left\{ \theta \mid \theta = (1-\lambda)v_j + \lambda v_g \right\}, \quad 0 \leq \lambda \leq 1, \quad v_j, v_g \in V_{k-1} \quad (2.28)$$

e deve ser encontrado na lista  $v-v$  de  $V(k-1)$ . Se uma aresta é cortada por  $H_1(k)$ , a interseção é um vértice em  $V^*$ . Para se determinar se  $H_1(k)$  corta  $\mathcal{E}(j,g)$ , examina-se as posições dos vértices  $v_j$  e  $v_g$  em relação a  $H_1(k)$ :  $H_1(k)$  corta  $\mathcal{E}(j,g)$ , se e somente se,  $v_j \in H_1^+(k)$  e  $v_g \notin H_1^+(k)$ , isto é, se e somente se,  $\alpha_j^+ \geq 0$  e  $\alpha_g^+ < 0$ , como mostrado na **figura 2.10**.

O cálculo do novo vértice  $v_i^* \in \mathcal{E}(j,g)$  é obtido a partir de (2.28):

$$v_i^* = (1-\lambda).v_g + \lambda.v_j, \quad 0 \leq \lambda \leq 1, \quad \alpha_j^+ \geq 0 \text{ e } \alpha_g^+ < 0 \tag{2.29-a}$$

além disso,  $v_i^*$  dado por (2.29-a) deve satisfazer:

$$\phi^T(k).v_i^* = c^+(k) \tag{2.29-b}$$

isto é,  $v_i^* \in H_1$ .

Substituindo (2.29-a) em (2.29-b) obtém-se:

$$\lambda = \frac{\phi^T(k).v_g - c^+(k)}{\phi^T(k).(v_g - v_j)} \tag{2.30}$$

Portanto, o conjunto  $V^*$  dos novos vértices é encontrado pela aplicação de (2.30) a todos pares de vértices adjacentes de  $V(k-1)$  que passam no teste em (2.29-a). Os elementos negativos de  $\alpha^+$  mostram quais vértices de  $V(k-1)$  poderão ser eliminados.

**ALTERAÇÃO DA LISTA DE VÉRTICE-VÉRTICE ADJACENTE (Lista  $v-v$ )**

A adjacência  $v_j \leftrightarrow v_g$  de vértices onde  $v_j \in H_1^+$  e  $v_g \notin H_1^+$  é trocada por  $v_j \leftrightarrow v_i^*$  quando  $S$  é alterado. O novo vértice  $v_i^*$  é também adjacente a outros novos vértices pertencentes a  $V^*$  que devem ser identificados. A adjacência de dois vértices é detectada checando se eles possuem  $N-1$  faces (hiperplanos) em comum.

**ALTERAÇÃO DA LISTA DE VÉRTICE-FACE (Lista  $v-f$ )**

A nova lista  $v-f$  de  $S(k)$  é formada por dois tipos de faces. As primeiras pertencem ao hiperplano  $H_1^+$ , já que todo hiperplano limitante  $H_1$  de  $H_1^+$ , interseccionando  $S(k-1)$ , aparecerá na lista  $v-f$  dos vertices  $v_i^*$  em  $V^*$  uma vez que  $V^* \subset H_1$ . O segundo tipo de face pertence a aresta  $\mathcal{E}(j,g)$  que é a aresta ligando os vértices adjacentes  $v_j$  e  $v_g$  de  $S(k-1)$ . Como

$\mathcal{E}(j,g)$  é a interseção das  $(N-1)$  faces comuns as listas  $v\text{-}f$  de  $v_j$  e  $v_g$ , e como também  $v_i^* \in \mathcal{E}(j,g)$ , estas são as faces de  $S(k-1)$  que continuam na lista  $v\text{-}f$  de  $V^*$ .

### INICIALIZAÇÃO DO ALGORITMO

Como no caso elipsoidal, a determinação de  $S(0)$  é baseada no conhecimento a priori sobre o sistema a ser identificado. Se nenhum conhecimento for disponível, a região politópica  $S(0)$  pode ser inicializada como um hipercubo  $N$ -dimensional, suficientemente grande tal que  $S(k) \subset S(0)$  para todo  $k$ .

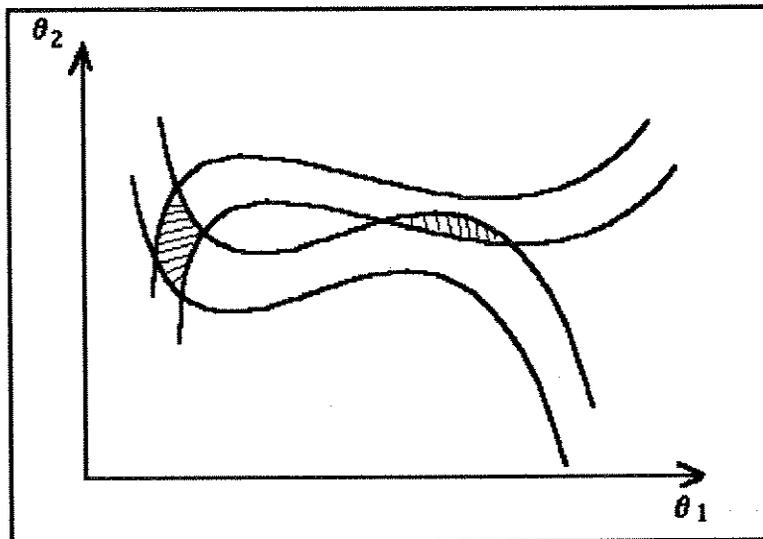
### RESUMO DO ALGORITMO

- (1) Formar  $S_0$  e alterar  $k$  para 1
- (2) Obter a medida  $y(k)$  e formar vetor  $\phi(k)$
- (3) FOR  $H_1^+(k)$  e  $H_2^+(k)$  DO
  - (3.1) FOR  $i = 1$  TO  $x$  DO { onde  $x =$  nro. de vértices de  $S_{k-1}$  }
    - IF  $H_1^+(k)$  THEN  $\alpha_i^+ = c^+(k) - \phi^T(k).v_i$
    - ELSE  $\alpha_i^- = \phi^T(k).v_i - c^-(k)$
  - (3.2) Formar  $\alpha = [ \text{IF } H_1^+(k) \text{ THEN } \alpha^+ \text{ ELSE } \alpha^- ]$
  - (3.3) IF todos elementos de  $\alpha$  são não negativos
    - THEN  $S_k = S_{k-1}$ ; GO TO (3.4)
    - ELSE IF todos elementos de  $\alpha$  são negativos
      - THEN  $S_k = 0$ ; GO TO (5)
      - ELSE (3.3.1) Encontrar  $V^*$ 
        - (3.3.2) Formar listas  $v\text{-}f$  para  $V^*$
        - (3.3.3) Formar listas  $v\text{-}v$  para  $V^*$
        - (3.3.4) Alterar  $S_{k-1}$ 
          - Acrescentar  $V^*$  a  $V$
          - Eliminar vértices  $\notin H_1^+(k)$
- (4) IF existem mais medidas
  - THEN incrementar  $k$  e voltar ao passo (2)
- (5) STOP

Uma das vantagens da abordagem politópica, assim como os algoritmos EOB, refere-se ao caráter recursivo de seus métodos, que facilitam implementações em tempo real. Além disso, estes algoritmos conduzem ao cálculo de intervalos de incerteza paramétrica idênticos aos valores reais da região  $S$ . A principal desvantagem desta abordagem é o fato da complexidade de descrição dos vértices que cresce com o número de medidas.

## 2.4-) O PROBLEMA DOS MODELOS NÃO LINEARES

Quando se utilizam modelos não-lineares, a região de pertinência paramétrica não possui forma politópica como no caso linear (*figura 2.2*). Isto é devido ao fato de que a região produzida pela equação (2.3) não é mais limitada por hiperplanos e sim por hipersuperfícies complexas, logo a aproximação desta região por elipsóides ou ortotopos torna-se complicada. Além disso, a região de pertinência paramétrica total pode ser não convexa, e também ser composta por sub-regiões como ilustrado na *figura 2.11*.



*Figura 2.11* - A Região de Pertinência Para Modelos Não Lineares

Os intervalos de incerteza paramétrica para o caso de modelos não lineares também são desconexos permitindo apenas o cálculo de intervalos de incerteza locais (*figura 2.11*).

Os métodos usados para estimação paramétrica de modelos não lineares utilizando a abordagem UBBE podem ser divididos em 3 grupos principais. O primeiro grupo destes métodos consiste em procedimentos que realizam múltiplas linearizações sobre o modelo não linear, o segundo grupo consiste na caracterização das fronteiras da região de pertinência paramétrica associada com cada medida, o terceiro grupo de métodos utiliza algoritmos de programação não linear. Este métodos são discutidos com maior abrangência em [4,20,29,30].

Nesta tese, o enfoque é dirigido principalmente aos modelos lineares e aos modelos não lineares que permitam ser linearizados.

## **2.5-) CONCLUSÃO**

A maioria das técnicas para estimação paramétrica assumem que os dados são corrompidos por ruídos randômicos, dos quais a função densidade de probabilidade é conhecida. Existem muitas situações onde esta hipótese não é baseada em qualquer informação a priori digna de confiança, e onde o número de medidas é muito pequeno para permitir sua validação a posteriori. Poucos dados também proíbem o uso de teorias assintóticas a fim de se obter intervalos de incerteza paramétrica. Nestas circunstâncias, uma caracterização mais geral da incerteza sobre as medidas torna-se necessária, e a imposição de limites sobre os erros (abordagem UBBE) fornece uma alternativa interessante.

Um método de estimação robusta na abordagem UBBE calcula uma região do espaço paramétrico, denominada região de incerteza paramétrica, que é compatível com a estrutura do modelo, com os limites conhecidos do erro e com todos os dados medidos. Qualquer vetor de parâmetros pertencente a esta região pode ser considerado um bom estimador dos parâmetros reais do sistema que se deseja identificar. A caracterização das incertezas para estes estimadores é dada por um intervalo (intervalo de incerteza paramétrica) de valores possíveis para cada parâmetro estimado. Este intervalo é calculado a partir dos valores extremos assumidos por cada estimador na região de pertinência paramétrica.

Quando o modelo é considerado linear nos parâmetros, é possível empregar três tipos de algoritmos para cálculo de regiões de incerteza paramétricas. O primeiro produz regiões elipsoidais internas ou externas à  $S$ . Os cálculos envolvidos são recursivos, limitados

e semelhantes àqueles requeridos pelo algoritmo mínimos quadrados recursivo (MQR) [30]. Contudo, os intervalos de incerteza paramétrica obtidos são freqüentemente muito pessimistas. O segundo tipo de algoritmo produz regiões ortotópicas internas ou externas à  $S$ . Os cálculos envolvidos não são recursivos e são mais complexos, mas os intervalos de incerteza paramétrica podem ser obtidos exatamente (método do ortotopo externo). O terceiro tipo de algoritmo produz uma descrição exata de  $S$ , dos quais os intervalos de incerteza paramétrica podem ser obtidos. Os cálculos são recursivos e bastante simples, mas a complexidade da descrição do conjunto  $S$  cresce linearmente com o número de vértices, geralmente muito menor que o número de medidas.

## **BIBLIOGRAFIA**

- [1] Arruda, L. V. R. - "**Etude d'algorithmes d'estimation Robuste et Développement d'un Système à Base de Connaissance pour l'Identification**", tese de doutorado, Université de Nice - Sophia Antipolis, França, 1992.
- [2] Barmish, B. R. & Sankaran, J. - "**The Propagation of Parametric Uncertainty Via Polytopes**", *IEEE Transaction on Automatic Control*, AC-24 (1979) 346-349.
- [3] Belforte, G & Bona, B - "**An Improved Parameter Identification Algorithm For Signals With Unknown-But-Bounded Errors**", in: *Prep. 7th IFAC/IFORS Symp. on Identification and System Parameter Estimation*, York (1985) 1507-1512.
- [4] Belforte, G. & Bona, B. & Cerone, V. - "**Parameter Estimation With Set Membership Uncertainty: Nonlinear Families of Models**", in: *Proc. 8th IFAC/IFORS Symp. on Identification and System Parameter Estimation*, Beijing (1988) 399-404.
- [5] Bertsekas, D. P. & Rhodes, I. B - "**Recursive State Estimation For a Set-Membership Description of Uncertainty**", *IEEE Trans. on Automatic Control*, AC-16 (1971) 117-128.
- [6] Broman, V & Shensa, M. J. - "**Polytopes, a Novel Approach to Tracking**", in: *Proc. 25th IEEE Conference on Decision and Control*, Athens (1986) 1749-1752.
- [7] Broman, V. & Shensa, M. J. - "**A Compact Algorithm to Compute Exact Polytope Parameter Bounds**", *Math. Comput. Simulation*, 32 (5&6) (1990) 469-480.

- [8] Clement, T. & Gentil, S. - "Reformulation of Parameter Identification With Unknown-But-Bounded Errors", *Math. Comput. Simulation*, 30 (1988) 257-270.
- [9] Favier, G & Arruda, L. V. R. - "A Review and a Comparison of Ellipsoidal Bounding Algorithms", in "Bounding Approaches to System Identification", Plenum Publishing Company, London, 1994.
- [10] Fogel, E. - "System Identification Via Membership Set Constraints With Energy Constrained Noise", *IEEE Trans. Automat. Control*, AC-24 (1979) 752-758.
- [11] Fogel, E. & Huang, Y. F. - "On the Value of Information in System Identification-Bounded Noise Case", *Automatica*, 18 (1982) 229-238.
- [12] Glover, J. D. & Schweppe, F. C. - "Control of Linear Dynamic Systems With Set Constrained Disturbances", *IEEE Trans. Automat. Control*, AC-16 (1971) 411-423.
- [13] Grunbaum, B. - **Convex Polytopes**, John Wiley and Sons, London, 1967.
- [14] Hadaegh, F. Y. & Bekey, G. A. - "Near-Identifiability of Dynamical Systems", *Math. Biosci.*, 77 (1985) 325-340.
- [15] Kurzhanski, A. B. & Hapalov, A. Yu. - "On the State Estimation Problem For Distributed System", in: A. Bensoussan and J. L. Lions, eds., *Analysis and Optimization of Systems*, Lecture Notes on Control and Information Sciences, 83 (Springer, Berlin, 1986) 102-113.
- [16] Lin, J. N. - "Determination of Reachable Set For a Linear Discrete System", *IEEE Trans. Automat. Control.*, AC-15 (1970) 339-342.
- [17] Ljung, L. - **System Identification: Theory For the User**, Prentice-Hall, Englewood Cliffs, NJ, 1987.
- [18] Milanese, M. & Belforte, G. - "Estimation Theory and Uncertainty Intervals Evaluation in Presence of Unknown-But-Bounded Errors: Linear Families of Models and Estimators", *IEEE Trans. Automat. Control*, AC-27 (1982) 408-414.
- [19] Milanese, M. & Tempo, R. - "Optimal Algorithms Theory For Robust Estimation and Prediction", *IEEE Trans. Automat. Control*, AC-30 (1985) 730-738.
- [20] Milanese, M. & Vicino, A. - "Estimation Theory For Nonlinear Models and Set Membership Uncertainty", *Automatica*, 27 (1991) 403-408.
- [21] Mo, S. H. & Norton, J. P. - "Fast and Robust Algorithm to Compute Exact Polytope Parameter Bounds", *Math. Comput. Simulation*, 32 (5&6) (1990) 481-493.

- [22] Norton, J. P. - "Recursive Computation of Inner Bounds For the Parameters of Linear Models", *Internat. J. Control*, 50 (1989) 2423-2430.
- [23] Pronzato, L. & Walter, E. & Piet-Lahanier, H. - "Mathematical Equivalence of Two Ellipsoidal Algorithms For Bounded-Error Estimation", in: Proc 28th IEEE Conference on Decision and Control, Tamp (1989) 1952-1955.
- [24] Schlaepfer, F. M. & Schweppe, F.C - "Continuous-Time State Estimation Under Disturbances Bounded By Convex Sets", *IEEE Trans. Automat. Control*, AC-17 (1972) 197-205.
- [25] Schweppe, F. C. - "Recursive State Estimation: Unknown-But-Bounded Errors and System Inputs", *IEEE Trans. Automat. Control*, AC-13 (1968) 22-28.
- [26] Schweppe, F. C. - **Uncertain Dynamic Systems**, Prentice-Hall, Englewood Cliffs, NJ, 1973.
- [27] Vicino, A. & Milanese, M. - "Optimal Inner Bounds of Feasible Parameter Sets in Linear Estimation With Bounded Noise", in: Proc 28th IEEE Conference on Decision and Control, Tampa (1989) 2576-2580.
- [28] Walter, E. & Piet-Lahanier, W. & Happel, J. - "Estimation of Non-Uniquely Identifiable Parameters Via Exhaustive Modeling and Membership Set Theory", *Math. Comput. Simulation*, 28 (1986) 479-490.
- [29] Walter, E. & Piet-Lahanier, H. - "Robust Nonlinear Estimation in the Bounded Noise Case", in: Proc 25th IEEE Conference on Decision and Control, Athens (1986) 1037-1042.
- [30] Walter, E. & Piet-Lahanier, H. - "Estimation of Parameter Bounds From Bounded-Error Data: A Survey", *Mathematics and Computers in Simulation*, 32 (1990) 449-468.
- [31] Witsenhausen, H. S. - "Sets of Possible States of Linear Systems Given Perturbed Observations", *IEEE Trans. Automat. Control*, AC-13 (1968) 556-558.

## *Capítulo 3*

# **Fundamentos de Redes Neurais Artificiais**

### **3.1-) INTRODUÇÃO**

O quarto tipo de algoritmo para cálculo de regiões de incerteza paramétricas, que será introduzido nesta tese, utiliza técnicas de redes neurais artificiais para calcular a região de pertinência paramétrica e os intervalos de incerteza paramétrica, esta abordagem é discutida no capítulo 4.

Neste capítulo, apresenta-se inicialmente alguns conceitos básicos usado para descrever e analisar sistemas através de redes neurais. Em seguida, descreve-se a rede de Hopfield que será utilizada na solução do problema de estimação robusta, e estabelece-se condições para a convergência e estabilidade desta rede.

Modelos de redes neurais artificiais ou simplesmente "redes neurais" podem ser tratados como: modelos conexionistas, processamento distribuído paralelo e sistemas neuromórficos [15,18]. Todos estes modelos tentam obter bom desempenho por intermédio da *alta densidade* de interconexões de elementos computacionais simples. Ao invés de um programa com instruções seqüenciais como em um computador de von Neumann, as redes neurais artificiais exploram simultaneamente várias possibilidades concorrentes, usando

redes *altamente paralelas*, compostas de vários elementos computacionais (nós) ligados entre si por conexões com pesos variáveis, e que são tipicamente adaptados no processo de "aprendizagem". Esta fase de aprendizagem, onde os pesos são adaptados para fornecer uma solução desejada é denominada de "treinamento" da rede. Assim um modelo baseado em uma rede neural artificial é unicamente especificado pela topologia da rede, pelas características dos nós e pelas regras de treinamento ou aprendizagem.

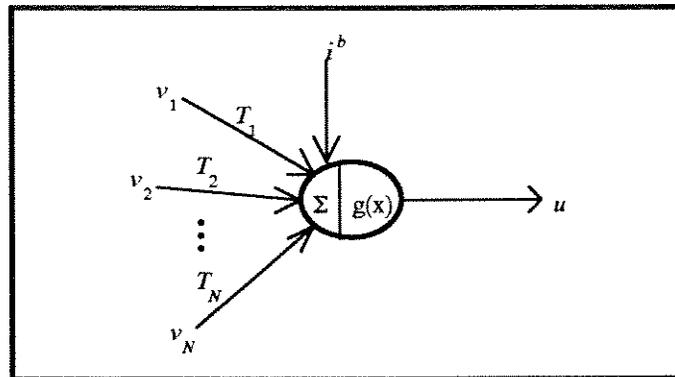


Figura 3.1 - Modelo de um Neurônio Artificial

A estrutura de uma rede neural é baseada no conhecimento atual de sistemas nervosos biológicos e sobre o próprio cérebro. Os elementos computacionais ou unidades processadoras, aqui chamados simplesmente de *neurônios*, correspondem aos nós da rede e são modelos simplificados dos neurônios biológicos. Estes modelos foram obtidos a partir de estudos realizados sobre a geração e propagação de impulsos elétricos pela membrana celular dos neurônios [8]. Os neurônios artificiais usados nos modelos de redes neurais são não lineares, tipicamente analógicos e realizam funções muito simples, como a de coletar os sinais existentes em suas entradas, agregá-los de acordo com a sua função de entrada e produzir sua saída de acordo com a função de ativação inerente.

O modelo de neurônio mais simples e que engloba as principais características de uma rede neural biológica, paralelismo e alta conectividade, foi proposto por McCulloch e Pitts [16]. Este modelo efetua a soma algébrica ponderada das entradas de um neurônio e passa este resultado através de uma função não linear, sendo o neurônio caracterizado ainda, por um limiar interno ou offset ( $i^b$ ). A *figura 3.1* mostra um neurônio com  $N$  entradas, e saída descrita pela equação (3.1):

$$u = g\left(\sum_{i=1}^N v_i \cdot T_i + i^b\right) \tag{3.1}$$

onde  $v_i$  corresponde a uma das entradas,  $T_i$  corresponde ao peso dado a esta entrada,  $i^b$  ao limiar do neurônio, e  $g(.)$  é a função de ativação do neurônio geralmente não linear, que define a característica digital/analógica (discreta/contínua) do neurônio. Por exemplo, para neurônios digitais usa-se, em geral, a função *sinal*:

$$g(x) = \text{sinal}(x) = \begin{cases} -1, & \text{se } x < 0 \\ +1, & \text{se } x \geq 0 \end{cases} \tag{3.2}$$

e para implementações analógicas, a função do tipo *sigmóide* :

$$g(x) = \text{sigmóide}(x) = \frac{a_g}{1 + \exp(-\beta x)} - b_g, \quad g(x) \in [-b_g, a_g - b_g] \tag{3.3}$$

onde  $a_g$  e  $b_g$  são os termos que delimitam a saída da função sigmóide e  $\beta$  é o termo de ganho de inclinação da função sigmóide. Se for utilizado os valores ( $a_g=2$  e  $b_g=1$ ) a saída será simétrica em (-1; 1), para valores ( $a_g=1$  e  $b_g=0.5$ ) a saída será simétrica em (-0.5; 0.5), e para ( $a_g=1$  e  $b_g=0$ ) a saída será assimétrica em (0; 1). As funções sinal e sigmóide estão ilustradas na **figura 3.2**. A função tangente hiperbólica também pode ser utilizada para se obter um comportamento de saída de forma equivalente ao da função sigmóide. Como as funções sigmóide, dada pela equação 3.3, e tangente hiperbólica são ambas diferenciáveis, elas foram historicamente usadas na derivação de um algoritmo de aprendizagem, baseado no gradiente de descida do erro de saída, para redes neurais de mais de uma camada [14]. Estas funções são até hoje as mais utilizadas em RNA.

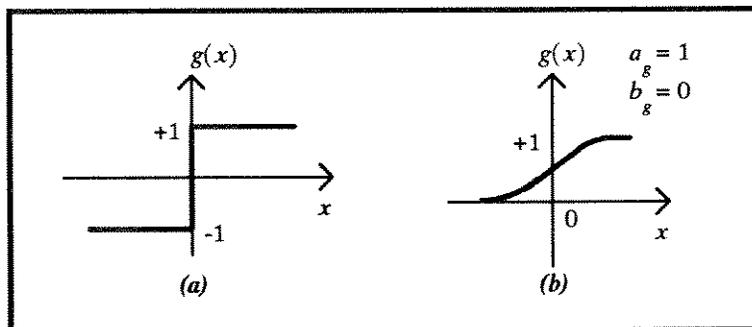


Figura 3.2 - Funções de Ativação do Neurônio, (a) Sinal, (b) Sigmóide

Uma rede neural é caracterizada pelo tipo de neurônio presente, por sua topologia e pelo algoritmo de aprendizagem utilizado. Este algoritmo especifica as condições iniciais da rede (valores dos pesos e offset) e como os pesos são ajustados durante o treinamento para atingir-se o desempenho desejado. As topologias básicas das redes e os diversos tipos de algoritmos de aprendizado utilizados serão discutidos a seguir.

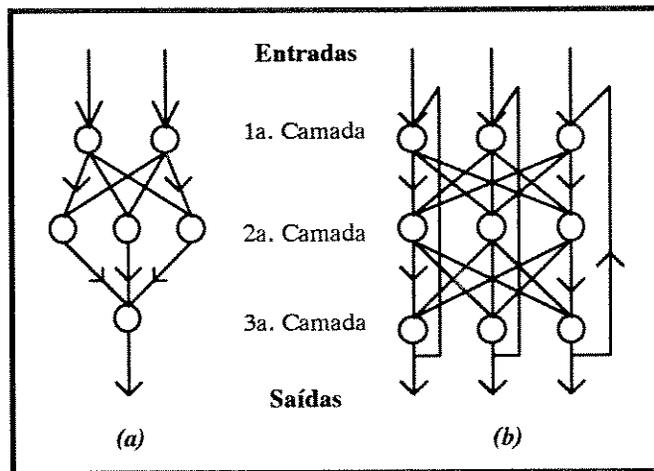


Figura 3.3 - Topologias de Redes, (a) Camadas, (b) Recorrente

As redes neurais possuem outras características importantes além da alta capacidade computacional devido ao *denso* paralelismo. Redes neurais são geralmente mais tolerantes a falhas, se comparada aos computadores seqüenciais. Esta robustez é devido à grande quantidade de nós processadores, cada um com seu conjunto primário de conexões independentes: defeitos em alguns dos nós ou nas conexões não irão afetar significativamente o desempenho global da rede, uma vez que os pesos nos outros nós serão ajustados de modo a compensar os erros de mau-funcionamento [15].

### 3.2-) TOPOLOGIAS DE REDES

As topologias básicas de redes neurais são relacionadas com os padrões de conexões entre as unidades processadoras e com a direção do fluxo de propagação de dados [14].

Segundo o tipo de conexão, as redes neurais podem ser classificadas em:

- *Redes em Camadas (Feed-Forward)*: onde o fluxo de dados das entradas para as saídas segue estritamente em uma única direção, isto é, existem apenas ligações entre as saídas dos nós de um nível inferior para as entradas dos nós de um nível superior, não existindo conexões entre os elementos de um mesmo nível (*figura 3.3-a*). Redes em camadas são utilizadas em várias aplicações, tais como: reconhecimento de padrões, aproximação de funções, simulação de funções lógicas, etc.

- *Redes Recorrentes (Feed-Back)*: este tipo de rede tem a mesma estrutura que a rede em camadas, no entanto são permitidas conexões entre as saídas de nós de um nível superior e as entradas de nós (realimentação) de um nível inferior (*figura 3.3-b*). Redes com realimentação são apropriadas para aplicações de identificação de sistemas (modelagem), controle, filtragem, predição não linear, etc.

Redes em camadas não têm memória; sua saída é unicamente determinada pela entrada atual e pelos valores dos pesos armazenados. Redes recorrentes realimentam as entradas com as saídas anteriores, portanto, sua saída atual é determinada por sua entrada atual e sua saída no passo anterior, o que fornece uma característica dinâmica à rede.

Exemplos clássicos de redes em camadas são o *Perceptron* e *Adaline*. Exemplos de redes recorrentes têm sido apresentados por Anderson [3], Kohonen [12] e Hopfield [9,10] a qual será objeto de estudo nas seções posteriores.

### 3.3-) TREINAMENTO DE REDES NEURAI ARTIFICIAIS

Uma rede neural deve ter seus pesos ajustados tal que a aplicação de um conjunto de entradas dado produza (ou por meio "direto" ou via um processo de relaxação) um conjunto de saídas desejado. Vários métodos para alteração dos pesos das conexões são disponíveis [11]. Alguns modificam os pesos explicitamente, utilizando um conhecimento a priori do sistema a ser modelado, outros propõem uma fase de treinamento da rede neural através do fornecimento de padrões de ensinamentos com a finalidade de alterar os pesos da rede de acordo com alguma regra de aprendizagem.

As situações de aprendizagem podem ser categorizadas em duas naturezas distintas:

- *Supervisionado, Associativo ou Auto-Supervisionado*: a rede é treinada pelo fornecimento de padrões de saídas associados às suas respectivas entradas. Estes pares entradas-saídas podem ser fornecidos por um meio externo, ou pelo sistema que contiver a rede. Quando um vetor de entrada é aplicado, a saída da rede é calculada e comparada com o respectivo padrão de saída, e a diferença (erro) é propagada em sentido inverso ao fluxo de informações da rede (das saídas para as entradas) e os pesos são alterados de acordo com algum algoritmo que tende a minimizar o erro. Os vetores de entrada e de padrão de saída do conjunto de treinamento são aplicados seqüencialmente, e os erros são calculados e os pesos ajustados para cada vetor, até que o erro para o conjunto de treinamento seja aceitável.

- *Não supervisionado, Autônomo ou Auto Organizado*: uma unidade (de saída) é treinada a fim de fornecer grupos de padrões dentre as entradas. Os vetores do conjunto de treinamento consistem unicamente de vetores de entrada. O algoritmo de treinamento modifica os pesos da rede para produzir vetores de saída que são consistentes, isto é, vetores do conjunto de treinamento que são similares entre si produzirão o mesmo padrão de saída. Neste tipo de aprendizagem espera-se que o sistema deva estatisticamente descobrir características marcantes da população de entrada. Ao contrário do aprendizado supervisionado, onde não existe um conjunto à priori de categorias dentro do qual os padrões irão ser classificados, o sistema é quem deve desenvolver sua própria representação do estímulo de entrada.

Exemplos de redes que são treinadas pelo primeiro tipo é o Perceptron [17] e a rede de Hopfield, e com aprendizado não supervisionado pode-se citar a de Kohonen [13] e redes Counterpropagation [22].

### 3.4-) REDES DE HOPFIELD

A rede de Hopfield é provavelmente o melhor exemplo conhecido de rede *recorrente* [11]. Esta rede apresenta geralmente uma única camada com conexões realimentadoras entre os nós. Na maioria dos casos os nós (neurônios) são completamente interconectados, isto é todos os nós da rede são conectados a todos os outros e a si próprio (*figura 3.4*), ou seja, todas as saídas da rede alimentam todas as suas entradas. Portanto,

alguns ou todos os nós de saídas podem ser vistos também como entradas da rede. Além disso, alguns ou todos os nós da rede podem receber entradas externas.

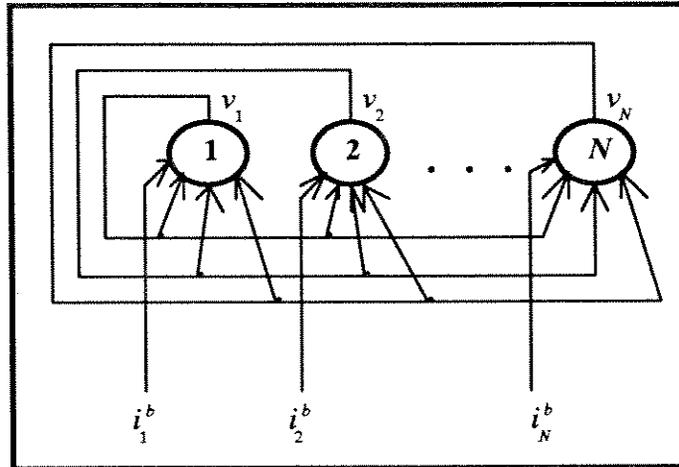


Figura 3.4 - Rede de Hopfield Típica

As respostas de tais redes é *dinâmica*, isto é, depois de aplicar um conjunto de novas entradas, as saídas são calculadas e retro-alimentadas com a finalidade de modificar as entradas do instante anterior. A saída é então recalculada e o processo é repetido iterativamente. Para uma rede estável, estas iterações sucessivas produzem mudanças nas saídas cada vez menores até que todas as saídas tornam-se constantes. Para alguns tipos de redes, este processo pode se repetir indefinidamente, e tais redes são ditas serem instáveis. Redes instáveis possuem características interessantes e têm sido estudadas como exemplo de sistemas caóticos.

Um amplo domínio de sistemas não-lineares de comportamento complexo pode ser representado através de redes de Hopfield. Dependendo de como os parâmetros da rede são escolhidos, esta deve funcionar como um sistema *estável*, ou como um oscilador, ou então como um sistema totalmente caótico [2, 5, 7, 19]. A maioria das aplicações da rede de Hopfield requerem que a rede se comporte como um sistema estável com múltiplos pontos de equilíbrio estáveis assintoticamente [1,6,11]. As condições que garantem este tipo de comportamento serão descritas na próxima seção.

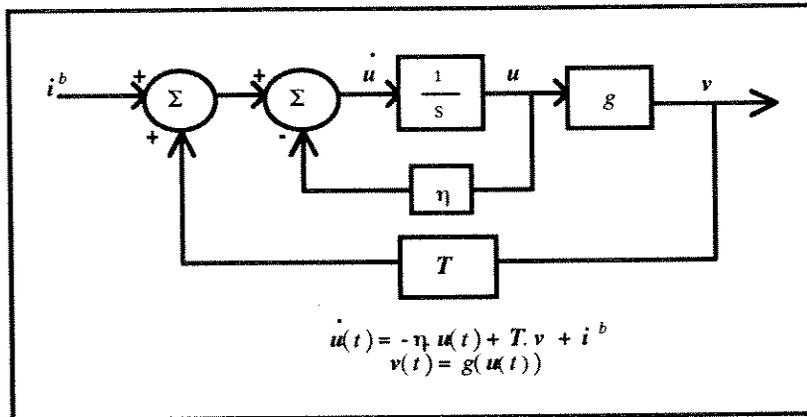
A rede de Hopfield contínua no tempo modela um sistema dinâmico não linear através de um vetor de entrada  $i$ , um vetor de estado  $u(t)$ , e um vetor de saída  $v(t)$  como

mostrado na *figura 3.5*. Estes vetores são relacionados através das equações dos nós (neurônios) dadas por:

$$\dot{u}_i(t) = -\eta u_i(t) + \sum_{j=1}^N T_{ij} \cdot v_j(t) + i_i^b \tag{3.4}$$

$$v_i(t) = g(u_i(t))$$

onde  $u_i(t)$  é o estado interno do  $i$ -ésimo neurônio,  $v_i(t)$  é a ativação de saída (ou estado de saída) do  $i$ -ésimo neurônio,  $T_{ij}$  é o peso conectando o  $j$ -ésimo neurônio com o  $i$ -ésimo neurônio,  $i_i^b$  é a entrada do  $i$ -ésimo neurônio, e  $-\eta u(t)$  é um termo de decaimento passivo. Utilizando a função de ativação do tipo sigmoidal (3.3), o vetor de saída estará no interior de um hipercubo  $N$ -dimensional, isto é,  $v(t) \in [-b_g, a_g - b_g]^N$ . Este fato será importante no mapeamento do problema de estimação robusta através da Rede de Hopfield. Uma outra vantagem da rede de Hopfield é que esta pode ser implementada diretamente em Hardware (**Apêndice A**), facilitando aplicações em tempo real.



*Figura 3.5* - Diagrama de Estado da Rede de Hopfield

Vale ressaltar que os valores dos pesos  $T_{ij}$  das conexões da rede de Hopfield para resolver um problema específico, geralmente são obtidos por meios explícitos (diretos) que utilizam um conhecimento a priori do sistema, dispensando-se portanto a fase de "treinamento" da rede. Neste trabalho, utiliza-se a técnica de subespaço-válido de soluções para derivar a matriz  $T$  de maneira explícita. Esta técnica será definida no capítulo 4.

### 3.4.1-) A Estabilidade da Rede de Hopfield

A estabilidade assintótica de uma rede de Hopfield pode ser mostrada utilizando o segundo método de Liapunov [21], isto é, mostrando que o sistema dissipa energia com o passar do tempo. Para isso, dada uma função de Liapunov (*ou função de energia*) associada à rede, mostra-se que suas derivadas temporais são não crescentes.

Uma função de Liapunov para a rede de Hopfield foi definida por [10, 20] como:

$$E(t) = -\frac{1}{2} \mathbf{v}(t)^T \cdot \mathbf{T} \cdot \mathbf{v}(t) - \mathbf{v}(t)^T \cdot \mathbf{i}^b + \eta \sum_{i=1}^N \int_0^{v_i(t)} g^{-1}(\alpha) d\alpha \quad (3.5)$$

onde  $g(\cdot)$  é a função de ativação do tipo sigmoidal (3.3). Embora a equação (3.5) não se comporte como uma função de Liapunov típica pois pode assumir valores negativos, é fácil mostrar que esta função é limitada, e pode portanto tornar-se positiva através da adição de uma constante apropriada.

Obtendo as derivadas temporais da Equação (3.5), tem-se:

$$\dot{E}(t) = \frac{dE(t)}{dt} = (\nabla_v E(t))^T \cdot \dot{\mathbf{v}}(t) \quad (3.6)$$

onde  $\nabla_v$  é o operador gradiente com relação a  $\mathbf{v}$ .

Para  $\mathbf{T}$  simétrica, tem-se

$$\nabla_v E(t) = -\mathbf{T} \cdot \mathbf{v}(t) - \mathbf{i}^b + \eta \cdot \mathbf{u}(t), \quad (3.7)$$

A partir da equação (3.4) para os neurônios da rede, conclui-se que:

$$\nabla_v E(t) = -\dot{\mathbf{u}}(t) \quad (3.8)$$

Com isto,  $\dot{E}(t)$  na equação (3.6) torna-se:

$$\begin{aligned} \dot{E}(t) &= -\dot{\mathbf{u}}(t)^T \cdot \dot{\mathbf{v}}(t) = \\ &= -\sum_{i=1}^N \dot{u}_i(t) \cdot \dot{v}_i(t) = -\sum_{i=1}^N \dot{u}_i(t) \frac{\partial v_i}{\partial u_i} \cdot \frac{\partial u_i}{\partial t} \\ &= -\sum_{i=1}^N (\dot{u}_i(t))^2 \cdot \frac{\partial v_i}{\partial u_i} \end{aligned} \quad (3.9)$$

Como  $(\dot{u}_i(t))^2$  é sempre positivo, precisa-se verificar apenas que  $\partial v_i / \partial u_i \geq 0$  para se mostrar que  $\dot{E}(t) \leq 0$ . Isto é trivialmente verdadeiro desde que a função  $g(\cdot)$  seja monotonicamente crescente como é o caso da função sigmóide na equação (3.3).

A partir da equação (3.9), nota-se também que  $\dot{E}(t) = 0$  apenas quando  $\dot{\mathbf{u}}(t) = 0$ . Isto completa a prova de estabilidade assintótica para a rede de Hopfield. Portanto, considerando a condição que  $T$  é simétrica, a rede eventualmente alcança um ponto de equilíbrio fixo. Devido a sua natureza não linear, a rede de Hopfield pode possuir múltiplos pontos de equilíbrio, os quais podem ser localizados calculando  $\dot{\mathbf{u}}(t) = 0$  na equação (3.4). Além disso, pela equação (3.8) estes pontos são também pontos extremos de  $E(t)$ .

Os resultados anteriores estabelecem que dado qualquer conjunto de condições iniciais  $\mathbf{v}(0)$ , a rede de Hopfield com  $T$  simétrica convergirá para um ponto de equilíbrio fixo, isto é, para um ponto onde  $\dot{\mathbf{v}}(t) = 0$ . Este ponto de equilíbrio  $\mathbf{v}_f$ , é um ponto fixo no hipercubo  $[-b_g, a_g - b_g]^N$  que torna-se limitado em virtude da função sigmoidal. Como a rede é determinística, a localização deste ponto é unicamente determinada pelas condições iniciais. Para todas condições iniciais que caem dentro da região de atração de um ponto de equilíbrio, a rede convergirá assintoticamente para aquele ponto.

A quantidade exata de pontos de equilíbrio, e suas localizações, é determinada pelos parâmetros  $T, \mathbf{i}^b$  da rede (equação 3.4) e pelo parâmetro  $\beta$  da função sigmóide (equação 3.3). Para ganhos pequenos ( $\beta$  pequeno) o número de pontos de equilíbrio é pequeno (possivelmente tão pequeno quanto 1), e estes pontos estão localizados no interior do hipercubo. Aumentando-se o ganho, o número de pontos de equilíbrio cresce bastante e suas localizações movem-se em direção aos vértices do hipercubo [10].

Para ganhos muito altos ( $\beta$  tendendo ao infinito) os pontos de equilíbrio realmente alcançarão os vértices do hipercubo e serão máximos em quantidade. Pode ser mostrado que quando  $\beta$  aproxima-se do infinito, o terceiro termo da equação (3.5) tende a zero [10], tal que a função de energia para a rede é simplificada para:

$$E(t) = -\frac{1}{2} \mathbf{v}(t)^T \cdot T \cdot \mathbf{v}(t) - \mathbf{v}(t)^T \cdot \mathbf{i}^b \quad (3.10)$$

Esta característica da rede de Hopfield ( $\beta \rightarrow \infty$ ) é de grande interesse para problemas que requerem soluções binárias; isto é, deseja-se que os pontos de equilíbrio  $v_f$  sejam vetores binários.

Uma vez que o ganho é fixo, as localizações dos pontos de equilíbrio são determinadas pela matriz  $T$  e  $i^b$ . Se a rede é usada para resolver diversos tipos de problemas, uma metodologia usual consiste em determinar  $T$  e  $i^b$  tal que os pontos de equilíbrio da rede correspondam às soluções do problema.

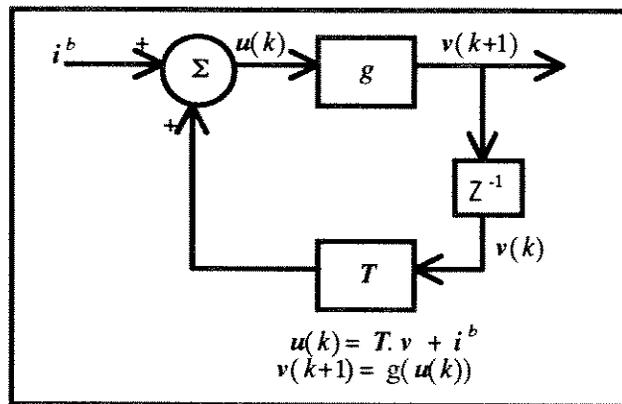


Figura 3.6 - Rede de Hopfield em Tempo-Discreto

### 3.4.2-) Rede de Hopfield Discreta

Uma versão em tempo-discreto da rede de Hopfield é descrita pelas seguintes equações:

$$u_i(k) = \sum_{j=1}^N T_{ij} \cdot v_j(k) + i_i^b \tag{3.11}$$

$$v_i(k+1) = g(u_i(k)) \tag{3.12}$$

onde  $k$  é o incremento de tempo. As equações (3.11) e (3.12) são as aproximações em tempo-discreto das equações (3.4). Assim como no caso contínuo, dado qualquer conjunto de condições iniciais  $v(0)$ , e restrições apropriadas sobre os pesos (para garantir estabilidade), esta rede convergirá para um ponto de equilíbrio fixo. O diagrama de estado do sistema para esta rede está mostrado na *figura 3.6*.

A estabilidade e propriedades de convergência da rede de Hopfield discreta são discutidos em [4]. As condições suficientes para a estabilidade desta rede são em geral que a matriz  $T$  seja simétrica e definida positiva. Além disso, estas condições dependem do modo que será realizado as alterações dos nós da rede. Quando são realizadas alterações *síncronas*, as quais implicam que todos os nós sejam alterados simultaneamente como sugerido na *figura 3.6*, então  $T$  deve ser definida positiva; mas quando utilizam-se alterações *assíncronas* (um nó por vez) é suficiente que os elementos da diagonal principal de  $T$  sejam não negativos, isto é,  $T_{ii} \geq 0$  [9]. Estas condições são mais exigentes que àquelas para a rede de Hopfield em tempo-contínuo.

### 3.5-) CONCLUSÃO

As estruturas de Redes Neurais Artificiais têm sido inspiradas por redes biológicas de neurônios encontradas no cérebro humano. Apesar das similaridades entre as redes artificiais e biológicas serem ainda pequenas, esta emulação modesta do cérebro através de RNA tem produzido resultados encorajadores. RNA exibem características que são semelhantes a do cérebro humano, entre estas: a habilidade de aprender através da experiência, generalização do conhecimento, execução de abstração, etc.

Nesta tese, utiliza-se a rede de Hopfield para solucionar problemas de Estimção Paramétrica Robusta. Portanto, vale ressaltar que o conceito de estabilidade, assim como as propriedades de convergência desta rede é de suma importância em relação a sua utilização. Além disso, vale ressaltar que os parâmetros internos desta rede serão explicitados no capítulo seguinte através de uma estrutura matemática denominada subespaço-válido de soluções.

### BIBLIOGRAFIA

- [1] Abu-Mostafa, Y. S. & St. Jaques, J. M. - "Information Capacity of the Hopfield Model", *IEEE Transactions on Information Theory*, 31(1985) 461-464.

- [2] Aihara, K & Takabe T. & Toyoda, M. - "Chaotic Neural Networks", *Physics Letters*, 144 (1990) 333-340.
- [3] Anderson, J.A - "Neural Models With Cognitive Implications", in *Basic Processes in Reading Perception and Comprehension Models*, D. LaBerge & S. J. Samuels, eds., Erlbaum, Hillsdale, NJ, 1977, 27-90.
- [4] Bruck, J. & Goodman, J. W. - "A generalized Convergence Theorem for Neural Networks", *IEEE Transaction on Information Theory*, 34 (1988) 1089-1092.
- [5] Choi, M. Y. & Huberman, B. A. - "Dynamic Behavior of Nonlinear Neural Networks", *Physical Review, A*, 28 (1983) 1204-1206.
- [6] Farrell, J. A. & Michel, A. N. - "A Syntesis Procedure for Hopfield's Continuous-Time Associative Memory.", *IEEE Transaction on Circuits and Systems*, 37 (1990) 877-884.
- [7] Freeman, W. J. & Yao, Y. - "Model of Biological Pattern Recognition With Spatially Chaotic Dynamics", *Neural Networks*, 3 (1990) 153-170.
- [8] Hodgkin, A. L. & Huxley L. F. - "A Quantitative Description of Membrane Current and Its Applications to Conduction and Excitation in Nerve", *Journal of Physiology*, 117 (1952) 500-544.
- [9] Hopfield, J. J. - "Neural Networks and Physical Systems With Emergent Colletive Computational Abilities", *Proceedings of the National Academy of Science USA*, 79 (1982) 2554-2558.
- [10] Hopfield, J. J. - "Neurons With a Graded Response Have Collective Computational Properties Like Those of Two-State Neurons", *Proceedings of the National Academy of Science, USA*, 81 (1984) 3088-3092.
- [11] Hush, D. R. & Horne B. G. - "Progress in Supervised Neural Networks", *IEEE Signal Processing Magazine*", january (1993) 8-39
- [12] Kohonen, T. - *Associative Memory: A System Theoretical Aproach*, Springer-Verlag, 1977.
- [13] Kohonen, T. - *Self-Organization and Associative Memory*, Springer-Verlag, 3rd. edition, 1989.

- [14] Kröse, B. J. A. & Smagt, V. & Patrick, P - **An Introduction to Neural Networks**, University of Amsterdam, 1993.
- [15] Lippmann, R. P. - "An Introduction to Computing With Neural Networks", *IEEE ASSP Magazine*, April (1987) 4-22.
- [16] McCulloch, W. & Pitts W. - "A Logical Calculus of the Ideas Immanent in Nervous Activity", *Bulletin of Mathematical Biophysics*, 5 (1943) 115-133.
- [17] Minsky, M. & Papert, S. - **Perceptrons: An Introduction to Computational Geometry**, The MIT Press, Cambridge, 1969.
- [18] Rumelhart, D. E. & McClelland, J. L. - **Parallel Distributed Processing**, Volume I e II, The MIT Press, Cambridge, Massachusetts, 1986.
- [19] Sandler, Y. M. - "Model of Neural Networks With Selective Memorization and Chaotic Behavior", *Physics Letters, A*, 144 (1990) 462-466.
- [20] Sherman, S. M. & Cohen, D. H. - "The Nervous System and Its Components", in R. M. Berne and M. N. Levy, editors, *Physiology*, pages 69-76. C. V. Mosby, St. Louis, MO, 1983.
- [21] Vidyasagar, M. - **Nonlinear Systems Analysis**, Prentice-Hall, Englewood Cliffs, NJ, 1978.
- [22] Wasserman, P. D. - **Neural Computing, Theory and Practice**, Van Nostrand Reinhold, New York, 1989.

## *Capítulo 4*

# **Estimação Paramétrica Robusta Através de Redes Neurais**

### **4.1-) INTRODUÇÃO**

Muitos problemas que modelam situações do mundo real podem ser formulados como problemas de otimização, e portanto qualitativamente descrito como um método de busca pela melhor solução. Estes problemas são encontrados normalmente em áreas de engenharia e comércio, e em cada um destes problemas de otimização o uso de um modelo matemático específico é útil para buscar a melhor solução.

No caso de estimação paramétrica robusta como apresentado no capítulo 2, o que se procura é o conjunto de soluções admissíveis (estimadores) que satisfaçam as várias restrições impostas pelo modelo.

Neste capítulo, introduz-se uma nova abordagem que utiliza os princípios de Redes Neurais Artificiais para cálculo destes estimadores. Nesta abordagem o que se propõe é um método que utiliza a rede de Hopfield, levando em consideração o conceito de *subespaço-válido* de soluções, com o objetivo de identificar os vértices do poliedro ou politopo convexo que é gerado por um conjunto de desigualdades lineares. No caso de estimação robusta, esta abordagem permitirá a identificação da *região de pertinência paramétrica*,

bem como dos *intervalos de incerteza paramétrica* para o problema apresentado na seção 2.2.

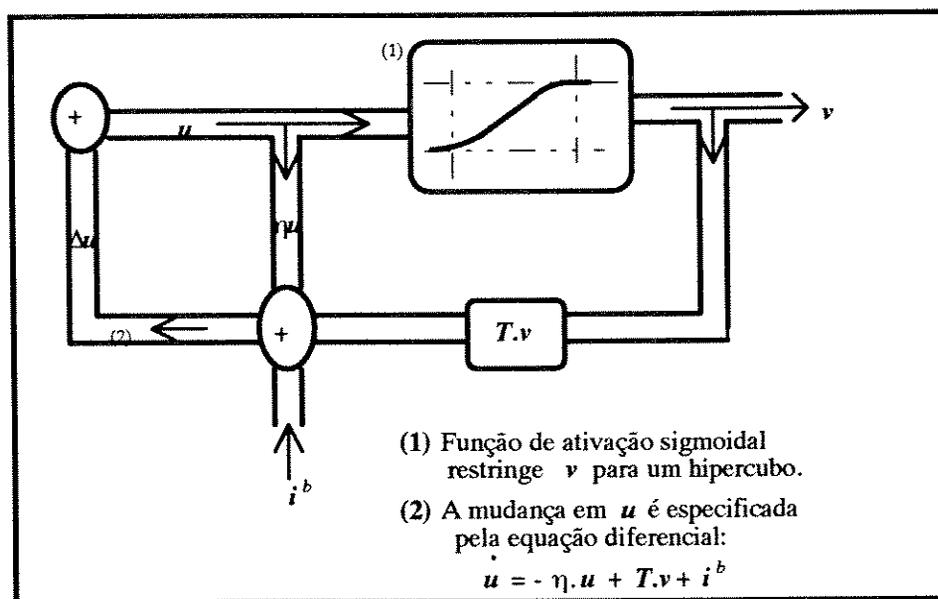


Figura 4.1 - Diagrama Esquemático da Rede de Hopfield Contínua

Uma das vantagens de utilizar este novo método é englobar dois métodos clássicos que foram apresentados no capítulo 2, isto é, ortotopo externo e politopo-exato. Neste capítulo, descreve-se como mapear estes problemas com o objetivo de garantir o confinamento do vetor de estado da rede dentro de um politopo. Este mapeamento utiliza a abordagem de *subespaço-válido* apresentado em [1,2]. Esta técnica estabelece que todos os pontos no espaço de saída da rede que correspondem às soluções válidas do problema pertencem a um subespaço-válido de soluções, que pode ser calculado com a finalidade de fornecer os parâmetros  $T$  e  $i^b$  da rede.

## 4.2-) SUBESPAÇO-VÁLIDO E A REDE DE HOPFIELD

Como visto no capítulo anterior, as equações dinâmicas da rede de Hopfield são as seguintes :

$$\dot{\mathbf{u}} = -\eta \cdot \mathbf{u} + \mathbf{T} \cdot \mathbf{v} + \mathbf{i}^b \tag{4.1}$$

$$v_i = g(u_i)$$

Se  $\mathbf{T}$  é simétrica, e  $\eta = 0$  ou o ganho da função de ativação é alto, uma possível função de Liapunov, para este sistema é :

$$E(\mathbf{v}) = -\frac{1}{2} \mathbf{v}^T \cdot \mathbf{T} \cdot \mathbf{v} - \mathbf{v}^T \cdot \mathbf{i}^b \tag{4.2}$$

Portanto, a rede de Hopfield (*figura 4.1*) é um método de minimizar funções da forma descrita pela equação (4.2) com  $\mathbf{T}$  simétrica. Para resolver um problema específico é necessário determinar o conjunto dos parâmetros  $\mathbf{T}$  e  $\mathbf{i}^b$  da rede, para que a minimização da função de Liapunov (4.2) coincida com a minimização da função objetivo do problema, e também forçar que as restrições do problema sejam satisfeitas. Este processo é denominado de "*mapeamento*" do problema dentro da rede. É também possível usar a rede de Hopfield para casos onde  $\mathbf{T}$  não é simétrica [6] modificando apenas a função (4.2) para :

$$E(\mathbf{v}) = -\frac{1}{2} \mathbf{v}^T \cdot \left[ \frac{1}{2} \cdot (\mathbf{T} + \mathbf{T}^T) \right] \cdot \mathbf{v} - \mathbf{v}^T \cdot \mathbf{i}^b \tag{4.3}$$

onde a matriz  $\frac{1}{2}(\mathbf{T} + \mathbf{T}^T)$  é simétrica.

A dificuldade em mapear problemas de otimização com restrições através de uma rede de Hopfield está em satisfazer as várias restrições que são impostas pelo problema como um todo. A rede atua com o objetivo de minimizar simultaneamente uma função de Energia simples ( $E^{ot}$ ) e as funções que regem as várias restrições ( $E^{rest}$ ): se qualquer uma destas restrições for violada, então a solução será dita "inválida". Uma técnica simples de mapeamento codifica as restrições válidas como termos na função de Energia que são minimizados quando as restrições são satisfeitas, ou seja:

$$E = E^{ot} + c_1 \cdot E_1^{rest} + c_2 \cdot E_2^{rest} + \dots \tag{4.4}$$

No processo de otimização com uma rede de Hopfield, a multiplicidade de termos da função (4.4) tende a frustrar uma ou várias restrições simultaneamente, portanto as

soluoes obtidas no final da otimizao podem ser inactivveis [7,8]. Por outro lado, o desempenho da rede  sensvel aos valores dos parmetros  $c_i$  em (4.4).

Para contornar estes problemas, Aiyer e outros em [1] fizeram uma anlise dos autovetores do subespao de comportamento da rede (isto , analisaram como a rede comporta-se em sua convergncia para soluoes vlidas) que revela como  $E^{ot}$  e os termos de restrioes  $E^{rest}$  na equao (4.4) podem ser efetivamente separados dentro de subespaos diferentes tal que eles no mais frustrem um ao outro.

Para uma grande variedade de problemas que podem ser resolvidos pela rede de Hopfield, foi verificado que as soluoes vlidas (pontos de equilbrio) pertenciam todas a um mesmo subespao-comum. Este subespao foi denominado subespao-vlido de soluoes com equao definida por:

$$v = T^{val} \cdot v + s \tag{4.5}$$

e que possui as seguintes propriedades:

1-)  $T^{val}$   uma *matriz projeco* ( isto :  $T^{val} \cdot T^{val} = T^{val}$  ) que projeta o vetor  $v$  dentro do subespao-vlido ( $v^{val} = T^{val} \cdot v$  ).

2-)  $T^{val} \cdot s = 0$ , onde o vetor  $s$  est num subespao ortogonal ao subespao-vlido (subespao-invlido).

Uma anlise matemtica detalhada do subespao-vlido pode ser encontrada em [4]. Assim, todos os termos de restrioes na equao (4.4) podem ser confinados dentro de um subespao nico, onde todas as restrioes imposta pelo problema so agrupadas em um nico termo denotado  $E^{conf}$  :

$$E = E^{ot} + E^{conf} \tag{4.6}$$

A partir do conceito de subespao-vlido, uma rede de Hopfield ideal, deve operar como segue:

- ( i ) - A rede  inicializada para um estado aleatrio tal que  $v$  est dentro do hipercubo limitado pela funo de ativao sigmide (3.3), e tambm sobre o subespao- vlido.

( ii ) - A minimizao de  $E$  move  $v$  em direo a um ponto de equilbrio da rede que satisfaz as restries impostas pelo problema, isto , que pertence ao subespao-vlido.

( iii ) - Quando  $v$  aproxima-se do ponto de equilbrio, a interao entre a minimizao de  $E$  e a funo sigmide fora  $v$  para o ponto de equilbrio que  uma soluo vlida e tambm uma soluo de "custo" timo para o problema, visto que  $E^{ot}$  est sendo otimizada. Portanto, a funo sigmide  $g(u)$  interage com a minimizao de  $E$  atravs da variao do vetor de estado interno da rede ( $u$ ), haja visto que

$$\nabla_v E(t) = - \dot{u}(t).$$

#### 4.2.1-) Derivando os Parmetros de $T$ e $i^b$ Utilizando-se o Subespao-Vlido

O prximo passo agora  obter as expresses analticas de  $T$  e  $i^b$ , utilizando o conceito de subespao-vlido tal que estas expresses permitam a minimizao da funo de Energia ( $E = E^{ot} + E^{conf}$ ) atravs dos passos ( i ), ( ii ) e ( iii ).

Dado que  $E$  foi dividido em um termo de confinamento ( $E^{conf}$ ) e em um termo de otimizao ( $E^{ot}$ ), tal que a minimizao de  $E^{conf}$  da rede confina  $v$  dentro de um subespao-vlido, e a minimizao de  $E^{ot}$  da rede move  $v$  em direo a uma soluo de "custo" timo, ento os parmetros  $T$  e  $i^b$  tambm podem ser escritos como:

$$\begin{aligned} T &= T^{conf} + T^{ot} \\ i^b &= i^{conf} + i^{ot} \end{aligned} \tag{4.7}$$

Portanto, a partir da definio (4.2) os termos da funo de Energia em (4.6) podem ser escritos como:

$$E^{conf} = -\frac{1}{2} v^T . T^{conf} . v - v^T . i^{conf} \tag{4.8}$$

$$E^{ot} = -\frac{1}{2} v^T . T^{ot} . v - v^T . i^{ot} \tag{4.9}$$

**TEOREMA:** Se  $\nu$  minimiza  $E^{conf}$ , então  $\nu$  satisfaz a equação do subespaço-válido, se e somente se:

$T^{conf} = c_0 \cdot (T^{val} - I) \quad (4.10)$
$e \quad i^{conf} = c_0 \cdot s \quad (4.11)$

onde  $c_0$  é um parâmetro livre e  $s$  é um vetor constante do supespaço ortogonal ao subespaço-válido. Este teorema está mostrado no (apêndice B).

Os parâmetros livres  $T$  e  $i^b$  da rede de Hopfield utilizando a abordagem de subespaço-válido são dados por:

$T = T^{ot} + c_0 \cdot (T^{val} - I)$
$i^b = i^{ot} + c_0 \cdot s \quad (4.12)$

A necessidade de assegurar que  $E^{conf}$  seja dominante sobre  $E^{ot}$  implica em atribuir um valor elevado para o termo  $c_0$  em (4.12), isto significa que a maior parte da mudança em  $\nu$  causada por  $\Delta \nu$  (por meio de  $\Delta u$ ) está relacionada com a minimização de  $E^{conf}$  e muito pouco com a minimização de  $E^{ot}$  [5]. Isto faz a simulação muito ineficiente, desde que a maioria do tempo será gasto forçando o confinamento de  $\nu$  no subespaço-válido. Com a finalidade de evitar este problema, analisa-se a seguir uma rede de Hopfield modificada [4], que diretamente força a validade da restrições.

### 4.3-) REDE DE HOPFIELD MODIFICADA

Se  $T^{conf}$  e  $i^{conf}$  são alterados de acordo com as equações (4.10) e (4.11), pode ser mostrado que a rede de Hopfield, segue aproximadamente os passos ( i ), ( ii ) e ( iii ), já que a minimização de  $E^{conf}$  assegura que  $\nu$  sempre pertence ao subespaço-válido, enquanto a minimização de  $E^{ot}$  move  $\nu$  em direção a um ponto de equilíbrio que é de "custo" mínimo.

Entretanto, uma análise mais detalhada de como a rede percorre os passos ( i )...( iii ) revela que isto é feito de um modo ineficiente e não confiável. Isto porque a minimização de  $E^{ot}$  interfere com a minimização de  $E^{conf}$ . Conseqüentemente  $c_0$  em (4.10) e (4.11) tem que ser grande o bastante para assegurar que  $E^{conf}$  seja suficientemente dominante para forçar o confinamento de  $\nu$  no subespaço-válido.

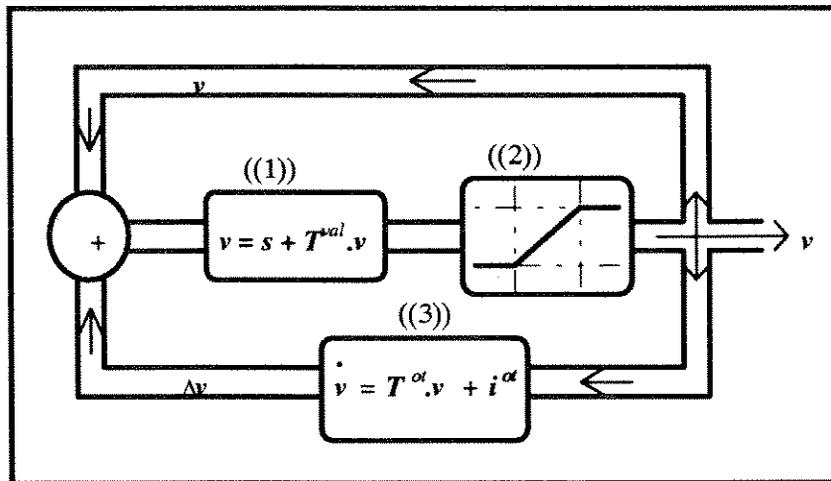


Figura 4.2 - Diagrama Esquemático da Rede de Hopfield Modificada

Em simulações por computador da rede de Hopfield, uma aproximação discreta para a equação diferencial contínua (4.1) é utilizada, em que:

$$\Delta v = g(u + \Delta u) - g(u)$$

onde:  $\Delta u = -\Delta t \cdot \nabla_v E = -\Delta t \cdot \nabla_v (E^{conf} + E^{ot})$  (4.13)

Portanto, a rede de Hopfield modificada ilustrada na *figura 4.2* e proposta por Fallside e Aiyer em [4] funciona através dos seguintes passos :

((1)) Projção de  $v$  dentro do subespaço-válido:

$$v = T^{val} \cdot v + s$$

((2)) Função de ativação "rampa simétrica" restringindo  $v$  dentro de um hipercubo

((3)) Alteração em  $v$  dado pelo gradiente do termo de energia de otimização  $E^{ot}$ , com:

$$\Delta v = \Delta t \cdot \dot{v} = \Delta t \cdot (T^{ot} v + i^{ot})$$

A vantagem principal da rede modificada, é que a rede confina  $v$  para o subespaço-válido diretamente pela operação de projeção (( 1 )), o que implica em minimizar  $E^{conf}$  diretamente. A operação não-linear (( 2 )) assegura então que  $v$  esteja dentro de

um hipercubo, visto que a função de ativação "rampa simétrica" (figura 4.3) é aplicada a todos elementos de  $v$ , isto é:

$$v_i \rightarrow g_r(v_i)$$

$$\text{onde, } g_r(v_i) = \begin{cases} a_g - b_g & \text{se } v_i > a_g - b_g \\ v_i & \text{se } -b_g \leq v_i \leq a_g - b_g \\ -b_g & \text{se } v_i < -b_g \end{cases} \quad (4.14)$$

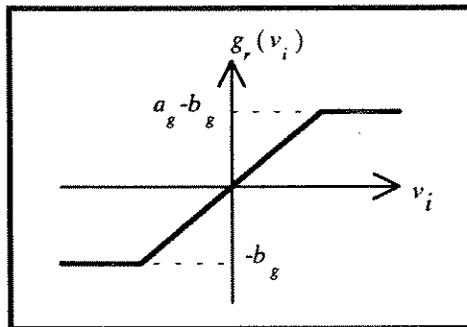


Figura 4.3 - Função de Ativação "Rampa Simétrica"

A minimização de  $E^{ot}$  é realizada pela operação (( 3 )) que atualiza  $v$  da seguinte forma:

$$\Delta v = \Delta t \cdot \dot{v} \quad \text{onde: } \dot{v} = T^{ot} \cdot v + i^{ot} \quad (4.15)$$

Da equação (4.9) tem-se que:

$$E^{ot} = -\frac{1}{2} v^T \cdot T^{ot} \cdot v - v^T \cdot i^{ot}$$

Portanto, pode ser visto que:

$$\Delta v = -\Delta t \cdot \nabla E^{ot} \quad (4.16)$$

$$\Delta v \cdot \nabla E^{ot} \leq 0 \quad (4.17)$$

Por conseguinte, a rede altera a direção de  $v$  em um sentido que minimize  $E^{ot}$ . Ao contrário da rede de Hopfield original (seção anterior),  $\Delta v$  está exclusivamente relacionado com a minimização de  $E^{ot}$ , e conseqüentemente, simulações por computador da rede de

Hopfield modificada converge (utilizando a técnica de subespaço-válido) em um número menor de iterações do que a rede de Hopfield original.

A rede de Hopfield modificada possui uma vantagem importante sobre a rede de Hopfield original que é a de ter somente um parâmetro a ser definido,  $\Delta t$ , enquanto a rede de Hopfield original possui dois parâmetros,  $c_0$  e  $\Delta t$ , como discutido na seção anterior.

Finalmente, as operações básicas da rede modificada, ou seja, multiplicação por uma matriz seguida por adição de um vetor fixo são idênticas àquelas da rede de Hopfield original.

### 4.3.1-) Rede Modificada e o Confinamento no Subespaço-Válido

É preciso provar que o vetor  $v$  calculado com a rede modificada permanece confinado no subespaço-válido. A suposição que a aplicação sucessiva das operações ((1)) e ((2)) confinará o vetor  $v$  calculado em ((3)) tanto no subespaço-válido quanto no hipercubo gerado pela função de ativação "rampa simétrica" ( $g_r$ ) é crucial para a viabilidade da rede de Hopfield modificada. Esta subseção é relacionada com a prova desta suposição.

Assume-se inicialmente que  $v$  no instante  $t$  satisfaz esta suposição. O valor de  $v$  no instante  $t + \Delta t$  é dado por  $v + \Delta v$ , onde  $\Delta v$  é dado pela equação dinâmica (4.16).

Se  $v + \Delta v$  permanece tanto dentro do subespaço-válido quanto do hipercubo gerado por  $g_r$ , então  $\Delta v$  deve satisfazer as seguintes condições:

- (c1)  $\Delta v$  é tal que  $v + \Delta v$  continua dentro do hipercubo gerado por  $g_r$ .
- (c2)  $\Delta v$  está completamente no subespaço-válido.

De certa forma, é muito improvável que  $\Delta v$  satisfaça ambas as condições acima, contudo, existe sempre uma componente de  $\Delta v$ , denominada  $\hat{\Delta v}$ , que satisfaz ambas condições (c1) e (c2). É mostrado no (Apêndice C) que existe uma matriz projeção  $\hat{T}$  tal que:

$$\hat{\Delta v} = \hat{T} \cdot \Delta v$$

e que a aplicação sucessiva das operações ((1)) e ((2)) sobre  $v + \Delta v$ , é equivalente a multiplicação de  $\Delta v$  por  $\hat{T}$ , isto é:

$$v + \Delta v \xrightarrow{((1))(2)) \dots ((1))(2))} v + \hat{T} \cdot \Delta v = v + \hat{\Delta v}$$

Desde que  $v + \Delta \hat{v}$  satisfaça (c1) e (c2), então a aplicação repetida das operações ((1)) e ((2)) realmente confina  $v$  tanto no subespaço-válido quanto no hiper-cubo gerado por  $g_r$ , apesar da mudança em  $v$  ser realizada por  $\Delta v$ .

#### 4.4) MAPEANDO PROBLEMAS DE ESTIMAÇÃO PARAMÉTRICA ROBUSTA ATRAVÉS DA REDE DE HOPFIELD

O problema de estimação paramétrica robusta para modelos lineares com erro desconhecido-mas-limitado (UBBE), como descrito no capítulo 2, consiste em estimar uma região do espaço paramétrico que pode ser expressa da seguinte forma:

$$R(k) = \{ \theta \mid y(k) - e^{max}(k) \leq \phi^T(k) \cdot \theta \leq y(k) - e^{min}(k) \} , \quad k=1, \dots, M^{med} \quad (4.18)$$

Neste caso, dado um vetor de medidas  $y$ , o conjunto de todos valores admissíveis de  $\theta$  que sejam consistentes com a estrutura do modelo, e tal que o erro de modelagem obedeça  $e^{min}$  e  $e^{max}$  é um poliedro convexo ou politopo. Este politopo é denominado de região de pertinência paramétrica.

O politopo acima pode ser mapeado de uma forma geral por uma rede de Hopfield que resolve o seguinte problema de otimização:

$$\text{minimizar } E^{ot}(\theta) = -\frac{1}{2} \theta^T \cdot T^{ot} \cdot \theta - \theta^T \cdot i^{ot} \quad (4.19)$$

$$\text{sujeito a } (\phi^{eq})^T \cdot \theta = b^{eq} \quad (4.20)$$

$$\text{e } (\phi^{in})^T \cdot \theta \leq b^{in} \quad (4.21)$$

$$\text{e } c_i^{min} \leq \theta_i \leq c_i^{max} \quad i \in \{1, \dots, N\} \quad (4.22)$$

onde:  $\theta \in \mathfrak{R}^N$  é o conjunto de vértices do politopo.

$\phi^{eq} \in \mathfrak{R}^{N \times M^{eq}}$  e  $\phi^{in} \in \mathfrak{R}^{N \times M^{in}}$  são matrizes contendo variáveis conhecidas inerentes ao sistema.

$b^{in} \in \mathfrak{R}^{M^{in}}$  é um vetor de medidas da saída do sistema levando em conta os limites conhecidos para o erro de modelagem.

$b^{eq} \in \mathfrak{R}^{M^{eq}}$  é um vetor de medidas da saída do sistema, obtido sob condições estabelecidas a priori.

$c_i^{min}$  e  $c_i^{max}$  são valores limites, definindo um intervalo de incerteza paramétrica inicial para cada estimador  $\theta_i$ .

A equação (4.19) refere-se a alguma condição que se deseja minimizar a fim de que algum critério a respeito do sistema seja ótimo.

A equação (4.20) refere-se às restrições de igualdade descrevendo alguma condição do problema estabelecida a priori .

A inequação (4.21) diz respeito às restrições de desigualdade obtidas a partir da equação (4.18) e também de alguma outra condição inerente ao problema que não possa ser descrita pela equação (4.18) .

A condição (4.22) refere-se aos limites mínimo e máximo que são considerados admissíveis a priori para cada parâmetro  $\theta_i$  . Estes limites determinam os intervalos de incerteza paramétrica inicial para cada estimador  $\theta_i$  .

O sistema (4.19)-(4.22) descreve um problema de otimização em que se deve minimizar uma função objetivo quadrática (4.19) sujeita a um conjunto de igualdades lineares (4.20) e a um conjunto de restrições de desigualdade (4.21). A condição final (4.22) assegura que  $\theta_i \in [c_i^{min}, c_i^{max}]$

De um ponto de vista geométrico, nota-se que as condições (4.20)-(4.22), se admissíveis, definem um poliedro convexo limitado ou politopo como na equação (4.18) , dentro do qual o vetor  $\theta$  deve estar caso represente uma solução válida. Este politopo será denotado pelo símbolo  $P$ . Portanto, estas condições definem a região de pertinência paramétrica para os estimadores. Para resolver este problema com redes de Hopfield (original ou modificada) utilizando a idéia de subespaço-válido, é necessário calcular o mapeamento em que o vetor  $\theta$  permanece dentro de  $P$  enquanto se otimiza a função objetivo descrita pela equação (4.19). Portanto, deseja-se obter um subespaço-válido (como descrito na seção 4.2) que assegure que as condições (4.20) e (4.21) sejam satisfeitas, já que a condição (4.22) pode ser implementada através da função de ativação dos neurônios da rede, ou seja, função sigmoidal (3.3) para a rede de Hopfield original ou a função "rampa simétrica" (4.14) para a rede modificada.

#### 4.4.1-) Subespaço-Válido Para o Problema

Primeiro, considera-se o caso em que exista apenas restrições de igualdade na forma (4.20). As restrições de igualdade constituem um sistema de equações lineares que se solucionadas simultaneamente determinam um subespaço-afim de soluções. Se as linhas de

$(\phi^{eq})^T$  são linearmente independentes, este subespaço pode ser representado da seguinte forma:

$$\theta = \phi^{eq} \cdot \left( (\phi^{eq})^T \cdot \phi^{eq} \right)^{-1} \cdot b^{eq} \quad (4.23)$$

neste caso, o vetor solução  $\theta$  é único (visto que o sistema de equações é possível e determinado), logo a expressão do subespaço-válido ( $\theta = T^{val} \cdot \theta + s$ ) deve possuir  $T^{val} \cdot \theta = 0$ , portanto o vetor  $s$  (que é um termo independente) torna-se igual ao vetor  $\theta$  e o subespaço-válido reconhecerá o vetor solução. Substituindo a equação (4.23) no lugar de  $s$  na equação do subespaço-válido, tem-se:

$$\begin{aligned} \theta &= T^{val} \cdot \theta + \phi^{eq} \cdot \left( (\phi^{eq})^T \cdot \phi^{eq} \right)^{-1} \cdot b^{eq} \\ T^{val} \cdot \theta &= \theta - \phi^{eq} \cdot \left( (\phi^{eq})^T \cdot \phi^{eq} \right)^{-1} \cdot b^{eq} \end{aligned} \quad (4.24)$$

Utilizando o valor de  $b^{eq}$  descrito pela equação (4.20), obtém-se a partir de (4.24) a expressão para  $T^{val}$ :

$$\begin{aligned} T^{val} \cdot \theta &= \theta - \phi^{eq} \cdot \left( (\phi^{eq})^T \cdot \phi^{eq} \right)^{-1} \cdot (\phi^{eq})^T \cdot \theta \\ T^{val} \cdot \theta &= \left( I - \phi^{eq} \cdot \left( (\phi^{eq})^T \cdot \phi^{eq} \right)^{-1} \cdot (\phi^{eq})^T \right) \cdot \theta \end{aligned}$$

Portanto,  $T^{val}$  e  $s$  para o subespaço-válido, são dados por:

$$T^{val} = I - \phi^{eq} \cdot \left( (\phi^{eq})^T \cdot \phi^{eq} \right)^{-1} \cdot (\phi^{eq})^T \quad (4.25)$$

$$s = \phi^{eq} \cdot \left( (\phi^{eq})^T \cdot \phi^{eq} \right)^{-1} \cdot b^{eq} \quad (4.26)$$

Neste caso, as restrições são combinadas em um único termo na função de Energia da rede. Além disso, este termo não interfere com a função objetivo em qualquer ponto dentro do politopo  $P$ .

Agora resta destacar que a função de ativação dos neurônios ( (3.3) ou (4.14) ) para a rede de Hopfield naturalmente força que a condição (4.22) seja satisfeita e, por conseguinte, o confinamento de  $\theta$  dentro de  $P$  é garantido. É de suma importância verificar que as equações (4.25) e (4.26) obedecem às propriedades do subespaço-válido, ou seja,  $T^{val} \cdot T^{val} = T^{val}$  e  $T^{val} \cdot s = 0$ .

A técnica descrita anteriormente para as restrições de igualdade pode ser estendida para as restrições de desigualdade da forma (4.21). Neste caso, convertem-se todas as restrições de desigualdade em restrições de igualdade introduzindo-se variáveis auxiliares para cada restrição de desigualdade [3]. Por exemplo, uma restrição de desigualdade típica dada por:

$$[\phi^{in}]_{i1}^T \cdot \theta_1 + [\phi^{in}]_{i2}^T \cdot \theta_2 + \dots + [\phi^{in}]_{iN}^T \cdot \theta_N \leq b_i^{in} \quad (4.27)$$

torna-se:

$$[\phi^{in}]_{i1}^T \cdot \theta_1 + [\phi^{in}]_{i2}^T \cdot \theta_2 + \dots + [\phi^{in}]_{iN}^T \cdot \theta_N + \sum_{j=1}^{M^{in}} q_j \cdot w_j = b_i^{in}, \quad w_j \geq 0 \quad (4.28)$$

onde  $w_j$  são as variáveis auxiliares que podem agora ser tratadas como componentes do vetor  $\theta$ , e  $q_j$  são constantes dadas por:

$$q_j = \begin{cases} 1, & \text{para } j = i \\ 0, & \text{para todo } j \neq i \end{cases} \quad (4.29)$$

Após a transformação do conjunto de restrições de desigualdade em um conjunto de restrições de igualdade, que é representado por um conjunto de variáveis estendidas  $(\theta^+)^T = [\theta^T \ w^T]$ , onde  $w$  é o vetor de variáveis auxiliares  $[w_1 \ w_2 \ \dots \ w_{M^{in}}]^T$ , o tamanho de  $\theta^+$  é  $N^+ = N + M^{in}$  e  $(\phi^+)^T$  é de dimensão  $(N+M^{in}) \times (M^{eq} + M^{in})$ , o problema de estimação paramétrica robusta pode agora ser expresso como:

$$\text{minimizar } E^{ot}(\theta) = -\frac{1}{2}(\theta^+)^T \cdot T^{ot} \cdot \theta^+ - (\theta^+)^T \cdot i^{ot} \quad (4.30)$$

$$\text{sujeito a } (\phi^+)^T \cdot \theta^+ = b^+ \quad (4.31)$$

$$\text{e } c_i^{min} \leq \theta_i^+ \leq c_i^{max}, \quad i \in \{1, \dots, N\} \quad (4.32-a)$$

$$\text{e } 0 \leq \theta_i^+ \leq c_i^{max}, \quad i \in \{N+1, \dots, N^+\} \quad (4.32-b)$$

onde  $\theta^+ \in \mathfrak{R}^{N^+}$  e  $b^+ \in \mathfrak{R}^{(M^{eq} + M^{in})}$ . Vale a pena notar que  $E^{ot}$  não depende das variáveis auxiliares  $w$ , e a restrição de igualdade estendida (4.31) incorpora as restrições de igualdade (4.20) e as restrições de desigualdade (4.21) originais. Além disso, a condição (4.22) fica dividida em duas partes: a condição (4.32-a) que estabelece um intervalo de incerteza paramétrica inicial para os parâmetros a serem estimados (originais), e a condição (4.32-b) que estabelece um limite superior para os parâmetros que representam as variáveis estendidas do vetor  $w$ .

Portanto a matriz de projeção  $T^{val}$  em (4.25) e o vetor  $s$  em (4.26) tornam-se:

$$T^{val} = I - \phi^+ \cdot \left( (\phi^+)^T \cdot \phi^+ \right)^{-1} \cdot (\phi^+)^T \quad (4.33)$$

$$s = \phi^+ \cdot \left( (\phi^+)^T \cdot \phi^+ \right)^{-1} \cdot b^+ \quad (4.34)$$

onde  $T^{val}$  será de dimensão  $N^+ \times N^+$  e o vetor  $s$  será de dimensão  $N^+$ .

#### 4.4.2-) Cálculo dos Vértices do Politopo

Como mostrado na seção anterior, o problema de estimação paramétrica robusta para modelos lineares com erro desconhecido-mas-limitado pode ser solucionado através de uma rede de Hopfield cuja dinâmica é baseada no gradiente de descida de  $E^{ot}$ . Para se obter os vértices do politopo  $P$ , é necessário alterar a matriz  $T^{ot}$  e o vetor  $i^{ot}$  para que o vetor paramétrico procurado (vetor  $\theta$ ) varie na direção contrária ao gradiente da função  $E^{ot}$ . Vale salientar que o termo  $E^{conf}$ , agrupando todas as restrições impostas pelo problema, é minimizado através da técnica de cálculo do subespaço-válido definido pela equações (4.33) e (4.34).

Assumindo-se que a matriz  $T^{ot}$  seja nula, o termo quadrático da função  $E^{ot}$  em (4.19) é eliminado, portanto  $E^{ot}$  depende apenas do vetor  $i^{ot}$ , ou seja:

$$E^{ot}(\theta) = -\theta^T \cdot i^{ot} \quad (4.35)$$

Baseado nisto, o cálculo dos vértices pertencentes a um politopo é realizada através de dois passos principais:

i-) fixa-se um vetor  $i^{ot}$  arbitrário,

ii-) utiliza-se a rede de Hopfield para calcular o ponto de equilíbrio da rede em relação ao vetor  $i^{ot}$  escolhido. Este ponto de equilíbrio representa o vértice do politopo.

A utilização de valores arbitrários para o vetor  $i^{ot}$  torna a busca dos vértices do politopo muito deficiente, visto que nenhum critério para a seleção dos vetores  $i^{ot}$  foi utilizado.

Além disso, é importante notar que as equações (4.35), (4.31), (4.32-a) e (4.32-b) permitem calcular uma região de pertinência paramétrica (convexa) a partir de uma combinação de qualquer número de pontos de equilíbrio (vértices) obtidos pela rede. O número de vértices é proporcional ao número de vetores  $i^{ot}$  que serão selecionados

previamente, com a finalidade de resolver algum problema específico. Para esta tese, define-se duas perspectivas de seleção dos vetores  $i^{ot}$  denominadas de perspectiva *eixos-direcionados* e perspectiva *eixos-hipercúbicos*.

Na *perspectiva eixos-direcionados*, a lista dos vetores  $i^{ot}$  escolhidos são os respectivos vetores unitários que estão na direção direta e inversa a cada eixo paramétrico (*figura 4.4-a*). Sendo  $N$  a dimensão do espaço paramétrico (ou quantidade de parâmetros a serem estimados), o número de vetores ( $nv_{ed}$ )  $i^{ot}$  a ser selecionados nesta perspectiva será dado por:

$$nv_{ed} = 2.N \tag{4.36}$$

A utilização desta perspectiva leva ao cálculo imediato dos intervalos de incerteza paramétrica exatos para os estimadores. Isto é, a rede de Hopfield que minimiza  $E^{ot}$  através de uma perspectiva eixos-direcionados é um método de estimação robusta que aproxima a região de pertinência paramétrica por uma região ortotópica externa.

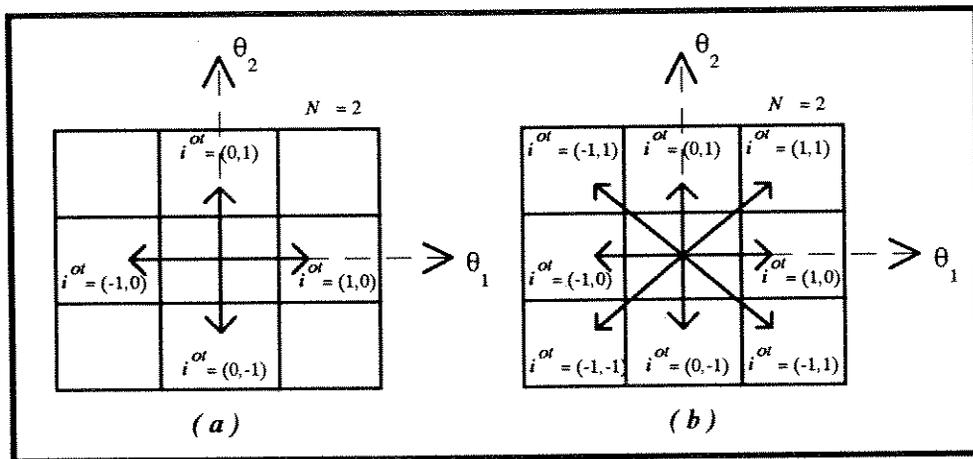


Figura 4.4 - Perspectivas de Seleção de  $i^{ot}$  : (a) Eixos-Direcionados, (b) Eixos-Hipercúbicos

Na *perspectiva eixos-hipercúbicos*, a lista de vetores  $i^{ot}$  selecionados são dados pelos  $2^N$  vetores unitários que ligam o centro de um hipercubo (alinhado com os eixos paramétricos) centrado na origem aos seus próprios vértices e também aos centros de suas faces (*figura 4.4-b*). Portanto, o número de vetores ( $nv_{eh}$ ) a ser selecionados nesta perspectiva é dado por:

$$nv_{eh} = 2.N + 2^N \tag{4.37}$$

Esta perspectiva será utilizada no cálculo dos vértices de uma região de pertinência paramétrica por intermédio do uso de redes neurais artificiais. Neste contexto, a rede de Hopfield que minimiza  $E^{ot}$  através de uma perspectiva eixos-hipercúbicos, assim como o método de Mo e Norton (seção 2.3.3), faz uma descrição politópica da região de pertinência paramétrica.

O exemplo a seguir ilustra o uso destas perspectivas no cálculo de soluções de problemas de Estimação Paramétrica Robusta.

**EXEMPLO ILUSTRATIVO**

Admitindo-se um sistema físico representado por um modelo linear seja dado por:

$$y(k) = a_1(k) \cdot \theta_1 + a_2(k) \cdot \theta_2 + e(k) \equiv \phi^T(k) \cdot \theta + e(k) , k = 1, \dots, 10 \quad (4.38)$$

onde:  $y(k)$  é a medida realizada na saída do sistema,

$a_1$  e  $a_2$  são variáveis conhecidas inerentes ao sistema,

$\theta_1$  e  $\theta_2$  são os parâmetros desconhecidos do modelo e que serão estimados,

$e(k)$  representa uma perturbação que possui característica desconhecida-mas-limitada.

A **tabela 4.1** fornece 10 medidas simuladas para o sistema com perturbação dada por:

$$|e(k)| \leq 1, \text{ ou seja, } e^{max}(k) = -e^{min}(k) = 1.0 \quad (4.39)$$

$k$	1	2	3	4	5	6	7	8	9	10
$a_1(k)$	0.5454	0.3721	-0.3809	0.1342	0.1110	0.3332	0.2000	-0.0667	-0.1250	-0.2857
$a_2(k)$	0.0909	0.4651	0.1904	0.5369	0.3332	0.2222	0.4000	0.3332	0.2500	0.1429
$y(k)$	1.909	1.5172	-0.6190	1.8859	1.3332	1.6667	1.8000	0.6667	0.5000	-0.4286

**Tabela 4.1** - Medidas Simuladas Para o Exemplo Ilustrativo

Portanto, os pares de desigualdades que serão formados a partir da equação (4.18), com base no modelo (4.38) e nos limites a priori do erro (4.39) serão dados por:

$$R(k) = \{ \theta / y(k) - 1 \leq \phi^T(k) \cdot \theta \leq y(k) + 1 \} \quad k = 1, \dots, 10 \quad (4.40)$$

onde:  $\phi^T(k) = [a_1(k) \ a_2(k)]$

$$\theta^T = [\theta_1 \ \theta_2]$$

Os pares de desigualdades dados por (4.40) formam um politopo em um espaço de duas dimenses ( $N=2$ ) com eixos paramétricos  $\theta_1$  e  $\theta_2$ . Estes pares de desigualdades podem ser colocados na forma da equaço (4.21) com  $M^n = 20$ . O exemplo no apresenta nenhuma restrico de igualdade que possa ser colocada na forma da equaço (4.20), logo  $M^e = 0$ .

Convertendo-se o conjunto de restrices de desigualdade na forma de (4.21) em um conjunto de restrices de igualdade na forma de (4.28), pode-se derivar os vetores de variáveis estendidas que sero utilizados tanto na rede de Hopfield original quanto na rede de Hopfield modificada. Estes vetores so dados por:

$$(\theta^+)^T = [ \theta_1 \ \theta_2 \ w_1 \ w_2 \ \dots \ w_{20} ]$$

$$(b^+) = [ b_1 \ b_2 \ b_3 \ \dots \ b_{20} ]$$

$$(\phi^+(i))^T = [ \pm a_1 \ \pm a_2 \ q_1 \ q_2 \ q_3 \ \dots \ q_{20} ] \quad , \quad i = 1, \dots, 20$$

onde os termos  $q_j$  so calculados pela equaço (4.29).

Assim, o problema de Estimaco Paramétrica Robusta do exemplo ilustrativo, torna-se:

$$\text{Minimizar } E^{ot}(\theta) = -(\theta^+)^T \cdot i^{ot}$$

$$\text{sujeito a: } (\phi^+)^T \cdot \theta^+ = b^+$$

onde a lista dos vetores  $i^{ot}$  sero dada ou pela perspectiva eixos-direcionados ou pela perspectiva eixos-hipercúbicos.

O cálculo da matriz  $T^{val}$  e do vetor  $s$  so feitos utilizando respectivamente as equaçes (4.33) e (4.34), e a rede de Hopfield original ou modificada sero utilizada para encontrar os pontos de equilíbrio da rede, os quais correspondem aos vértices do politopo gerado a partir de  $R(k)$ .

A **tabela 4.2** mostra os vértices calculados pela rede utilizando a perspectiva eixos direcionados, faz-se tambem uma comparaco com os valores calculados pelo método do ortotopo externo apresentado na seço 2.3.2.

Vértices	vetor $i^{or}$	Rede Neural		Ortotopo Externo	
		$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$
vértice 1	( 1 , 0)	5.0814	1.6490	5.0630	1.6255
vértice 2	(-1 , 0)	0.9980	4.0009	1.0000	4.0011
vértice 3	( 0 , 1)	1.5000	5.0098	1.4995	5.0003
vértice 4	( 0 , -1)	4.5278	0.5173	4.5116	0.5223

Tabela 4.2 - Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados

A *tabela 4.3* ilustra as características dos Intervalos de Incerteza Paramétrica com base nos resultados obtidos na *tabela 4.2*.

Intervalos de Incerteza	Rede Neural			Ortotopo Externo		
	[min , max]	largura	centro	[min , max]	largura	centro
$\theta_1$	[0.9980 , 5.0814]	4.0834	3.0397	[1.0000 , 5.0630]	4.0630	3.0315
$\theta_2$	[0.5173 , 5.0098]	4.4925	2.7636	[0.5223 , 5.0003]	4.4780	2.7613

Tabela 4.3 - Características dos Intervalos de Incerteza Paramétrica - Exemplo Ilustrativo

A *tabela 4.4* mostra os vértices da Região de Pertinência Paramétrica calculados pela rede utilizando a perspectiva eixos-hipercúbicos, faz-se também uma comparação com os valores calculados pelo método do politopo exato apresentado na seção 2.3.3.

Nesta tabela mostra-se também, que o vértice 7 não foi identificado pela rede utilizando a perspectivas eixos-hipercúbicos. A razão disto é que este vértice mantém um grau de colinearidade bastante forte com os vértices 1 e 6 que compõem a região de pertinência paramétrica exata (*figura 4.5*), logo a região de pertinência paramétrica pelo método de redes neurais utilizando a perspectiva eixos-hipercúbicos descarta o vértice 7 (*figura 4.6*).

Vértices	Rede Neural		Politopo-Exato	
	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$
vértice 1	1.5000	5.0098	1.4995	5.0003
vértice 2	1.4510	1.2855	1.4524	1.2870
vértice 3	4.5278	0.5173	4.5116	0.5223
vértice 4	0.9980	4.0009	1.0000	4.0011
vértice 5	5.0814	1.6490	5.0630	1.6255
vértice 6	4.7497	3.5377	4.7507	3.4994
vértice 7	-	-	3.4995	4.5004

**Tabela 4.4** - Vértices da Região de Pertinência Paramétrica - Exemplo Ilustrativo

Calcula-se a seguir o estimador levando em consideração o critério centro de massa do politopo que define a Região de Pertinência Paramétrica composta pelos vértices da *tabela 4.4*. A *tabela 4.5* compara os valores dos estimadores com relação aos 6 vértices obtidos pela rede utilizando a perspectiva eixos-hipercúbicos e aos 7 vértices obtidos pelo método politopo-exato.

Estimador	Rede Neural	Politopo-Exato
$\theta_1$	3.0513	3.1109
$\theta_2$	2.6667	2.9194

**Tabela 4.5** - Estimadores Paramétricos - Exemplo Ilustrativo

A *figura 4.6* ilustra a região de pertinência paramétrica obtida pela rede neural para o exemplo ilustrativo baseada na *tabela 4.4* enquanto a *figura 4.5* ilustra a região de pertinência paramétrica obtida pelo método do politopo-exato.

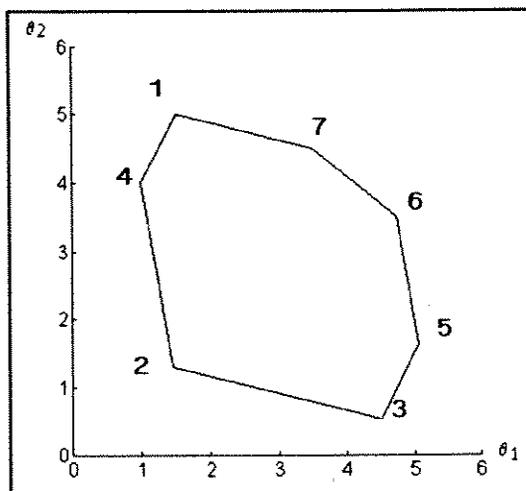


Figura 4.5 - A Região de Pertinência Pelo Método Politopo-Exato

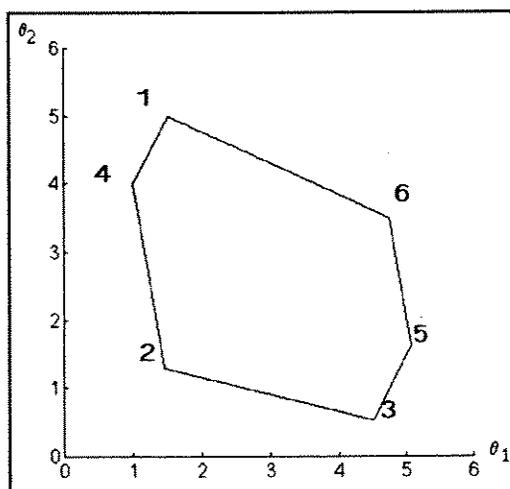


Figura 4.6 - A Região de Pertinência Pelo Método de Redes Neurais

## 4.5-) CONCLUSÃO

Neste capítulo ilustrou-se como é possível resolver problemas de Estimação Paramétrica Robusta através de Redes Neurais Artificiais. Vale ressaltar que técnicas utilizando RNA estão começando a ser introduzidas na área de Estimação Paramétrica Robusta.

Duas perspectivas de seleção dos vetores  $i^{ot}$  foram definidas. A perspectiva eixos-direcionados é utilizada para caracterizar os intervalos de incerteza paramétrica associados com cada parâmetro a ser estimado. A perspectiva eixos-hipercúbicos é utilizada quando é necessário obter-se uma representação mais fiel da região de pertinência paramétrica.

O método apresentado neste capítulo produz regiões de pertinência paramétrica mais otimista se comparado com o método elipsoidal ou ortotópico. Além disso, os intervalos de incerteza paramétrica associados com cada parâmetro são praticamente idênticos aos valores exatos calculados pelo método ortotopo externo.

Uma outra característica importante da rede de Hopfield apresentada é que esta permite tratar os métodos ortotopo-externo e politopo-exato de um modo unificado, ou seja: o método ortotopo-externo é implementado de maneira implícita pela rede quando se utiliza a perspectiva eixos-direcionados, enquanto o método politopo-exato é implementado quando se utiliza a perspectiva eixos-hipercúbicos.

A utilização de funções de ativação do tipo sigmoidal e "rampa simétrica" garantem que os estimadores calculados (através dos vértices obtidos pela rede) estão sempre confinados dentro de um hipercubo inicial. Além disso, devido as características de estabilidade e convergência desta rede, a utilização da técnica de subespaço-válido de soluções, garante também o confinamento do vetor de estado da rede no politopo  $P$ , portanto a possibilidade de encontrar soluções inválidas é eliminada.

A principal desvantagem desta rede está relacionada com a sua característica não recursiva, que dificulta implementações em tempo real.

## BIBLIOGRAFIA

- [1] Aiyer, S. V. B. & Niranja, M. & Fallside, F - "A Theoretical Investigation Into the Performance of the Hopfield Model", *IEEE Transaction on Neural Networks*, 1 (1990) 53-60.
- [2] Aiyer, S.V.B. - "Solving Combinatorial Optimization Problems Using Neural Networks", *Technical Report CUED/F-INFENG/TR 89*, Cambridge University Engineering Department, 1991.
- [3] Bazaraa, M. S. & Jarvis, J. J. - **Linear Programming and Network Flows**, John Wiley & Sons, Toronto, 1977.
- [4] Fallside, F. & Aiyer, S.V.B. - "A Subspace Approach to Solving Combinatorial Optimization Problem With Hopfield Networks", *Technical Report CUED/F-INFENG/TR 55*, Cambridge University Engineering Department, 1990.
- [5] Gee, A. H. & Aiyer, S. V. & Prager, R. V. - "Neural Networks and Combinatorial Optimization Problems - The Key to Successful Mapping", *Technical Report CUED/F-INFENG/TR 77*, Cambridge University Engineering Department, 1991.
- [6] Gee, A. H. & Prager, R. V. - "Polyhedral Combinatorics and Neural Networks", *Technical Report CUED/F-INFENG/TR 100*, Cambridge University Engineering Department, 1992.
- [7] Kamgar-Parsi, B. & Kamgar-Parsi, B - "On Problem Solving With Hopfield Neural Networks", *Biological Cybernetics*, 62(1990)415-423.
- [8] Wilson, V. & Pawley, G. S. - "On the Stability of the TSP Problem Algorithm of Hopfield and Tank", *Biological Cybernetics*, 58(1988) 63-70.

## *Capítulo 5*

# **Simulações de Problemas de Estimação Paramétrica Robusta Através de Redes Neurais Artificiais**

### **5.1-) INTRODUÇÃO**

Neste capítulo faz-se a simulação de problemas de Estimação Paramétrica Robusta utilizando a rede de Hopfield modificada como apresentada no capítulo anterior. Faz-se também uma comparação dos valores obtidos pela rede com outros métodos clássicos que foram apresentados no capítulo 2. A implementação desta rede foi realizada em computador digital e através de programas escritos para o ambiente interpretado do software matemático MATLAB [6]. Os métodos clássicos ortotopo-externo e politopo-exato foram escritos na linguagem Turbo Pascal.

### **5.2-) DETALHES SOBRE OS RESULTADOS DAS SIMULAÇÕES**

A maioria dos resultados das simulações realizadas são apresentados na forma de tabelas e gráficos. Portanto, para simplificar, descreve-se a seguir o conteúdo de cada tabela apresentada nos problemas subsequentes. Existem 4 tipos de tabelas intituladas por:

**a-) "Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados"**

Estas tabelas mostram os vértices calculados pela rede de Hopfield modificada utilizando a perspectiva eixos-direcionados. Faz-se também nestas tabelas uma comparação com os valores calculados pelo método do ortotopo-externo apresentado na seção 2.3.2.

**b-) "Intervalos de Incerteza Paramétrica"**

Estas tabelas ilustram as características dos Intervalos de Incerteza Paramétrica calculados com base nos resultados obtidos pela tabela intitulada "Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados". Vale ressaltar que o centro de cada intervalo paramétrico corresponde ao estimador central definido no capítulo 2 pela equação (2.19).

**c-) "Vértices da Região de Pertinência Paramétrica"**

Estas tabelas mostram os vértices da Região de Pertinência paramétrica calculados pela rede utilizando a perspectiva eixos-hipercúbicos. Faz-se também nestas tabelas uma comparação com os valores calculados pelo método do politopo-exato apresentado na seção 2.3.3.

**d-) "Estimadores Paramétricos"**

Estas tabelas fornecem os estimadores paramétricos levando em consideração o critério centro de massa do politopo que define a Região de Pertinência Paramétrica composta pelos vértices da tabela intitulada "Vértices da Região de Pertinência Paramétrica".

A coluna { Erro R-P } destas tabelas correspondente ao erro relativo (em porcentagem) entre o estimadores obtidos pela rede e os estimadores obtidos pelo método do politopo-exato, ou seja:

$$\text{Erro R-P} = \frac{(\theta_i^{\text{politopo}} - \theta_i^{\text{rede}}).100}{\theta_i^{\text{politopo}}} \% , i = 1, \dots, n$$

A coluna { Erro R-N } destas tabelas correspondente ao erro relativo (em porcentagem) entre o estimadores obtidos pela rede e os valores nominais dos parâmetros, ou seja:

$$\text{Erro R-N} = \frac{(\theta_i^{nominal} - \theta_i^{rede}) \cdot 100}{\theta_i^{nominal}} \% , i = 1, \dots, n$$

A coluna { Erro P-N } destas tabelas correspondente ao erro relativo (em porcentagem) entre o estimadores obtidos pelo método politopo-exato e os valores nominais dos parâmetros, ou seja:

$$\text{Erro P-N} = \frac{(\theta_i^{nominal} - \theta_i^{politopo}) \cdot 100}{\theta_i^{nominal}} \% , i = 1, \dots, n$$

### 5.3-) PROBLEMA 1 - Modelo AR (AutoRegressive) de Segunda Ordem.

Seja o modelo AR de segunda ordem, dado por (Problema proposto originalmente por Milanese e Vicino em [3]):

$$y(k) = a_1 \cdot y(k-1) + a_2 \cdot y(k-2) + e(k) \equiv \phi^T(k) \cdot \theta + e(k) , k = 1, \dots, 10 \quad (5.1)$$

onde:  $y(k)$  é a medida realizada na saída do sistema.

$a_1$  e  $a_2$  são os parâmetros desconhecidos do modelo e que serão estimados.

$e(k)$  representa a perturbação que possui característica desconhecida-mas-limitada.

A *tabela 5.1* fornece as 12 medidas simuladas para o sistema. Estes dados foram gerados a partir de (5.1) com valores nominais para os parâmetros dados por:

$$a_1^{nominal} = 0.3 , a_2^{nominal} = -0.69$$

e assumindo-se que  $e(k)$  é uniformemente distribuído no intervalo  $[-0.5, 0.5]$ , ou seja:

$$|e(k)| \leq 0.5, \quad \text{portanto: } e^{max}(k) = -e^{min}(k) = 0.5 \quad (5.2)$$

$k$	-1	0	1	2	3	4	5	6	7	8	9	10
$y(k)$	0.19	-0.72	-0.82	-0.22	0.88	0.80	-0.20	-0.88	0.31	0.32	-0.33	-0.63

**Tabela 5.1** - Medidas Simuladas Para o Problema 1

Os pares de desigualdades, formados a partir da equação (4.18), com base no modelo (5.1) e nos limites a priori do erro (5.2) são dados por:

$$R(k) = \{ \theta / y(k) - 0.5 \leq \phi^T(k) \cdot \theta \leq y(k) + 0.5 \} \quad k = 1, \dots, 10 \quad (5.3)$$

onde:  $\phi^T(k) = [y(k-1) \ y(k-2)]$

$$\theta^T = [a_1 \ a_2]$$

Os pares de desigualdades dados por (5.3) formam um politopo em um espaço de duas dimensões ( $N=2$ ) com eixos paramétricos  $a_1$  e  $a_2$ . Estes pares de desigualdades podem ser colocados na forma da equação (4.21) com  $M^{in} = 20$ . O exemplo não apresenta nenhuma restrição de igualdade que possa ser colocada na forma da equação (4.20), logo  $M^{eq} = 0$ .

Como nenhuma informação é disponível sobre o domínio a priori dos parâmetros a serem estimados para o modelo, admite-se que:

$$\theta_i \in [c_i^{min}, c_i^{max}] \ , \quad i = 1, 2 \quad (5.4)$$

onde:  $c_i^{min} = -100$  e  $c_i^{max} = 100$ .

Convertendo-se o conjunto de restrições de desigualdade da forma (4.21) em um conjunto de restrições de igualdade na forma (4.28), pode-se derivar os vetores de variáveis estendidas que serão utilizados na rede de Hopfield modificada. Estes vetores são dados por:

$$(\theta^+)^T = [a_1 \ a_2 \ w_1 \ w_2 \ \dots \ w_{20}]$$

$$(b^+) = [b_1 \ b_2 \ b_3 \ \dots \ b_{20}]$$

$$(\phi^+(i))^T = [\pm y(k-1) \ \pm y(k-2) \ q_1 \ q_2 \ q_3 \ \dots \ q_{20}] \ , \quad i = 1, \dots, 20$$

onde os termos  $q_i$  são fornecidos pela equação (4.29).

Utilizando-se a rede de Hopfield modificada, os parâmetros de inicialização desta rede são dados por:

$$\Delta t = 0.001$$

$$a_g = 200 \quad \text{e} \quad b_g = 100 \quad \text{para} \quad \theta_1^+ \quad \text{e} \quad \theta_2^+$$

onde os valores para  $a_g$  e  $b_g$  garantem que a saída da rede será confinada no hiper-cubo definido pelos valores admissíveis de  $c_i^{min}$  e  $c_i^{max}$  dados para cada parâmetro  $\theta_i$  na equação (5.4).

As tabelas de resultados para o problema 1 obtidos pela rede são apresentadas a seguir:

Vértices	vetor $i^{ot}$	Rede Neural		Ortotope Externo	
		$a_1$	$a_2$	$a_1$	$a_2$
vértice 1	( 1 , 0)	0.2606	-0.6972	0.2628	-0.6882
vértice 2	(-1 , 0)	0.3981	-0.7922	0.3960	-0.7923
vértice 3	( 0 , 1)	0.3645	-0.8054	0.3641	-0.8036
vértice 4	( 0 , -1)	0.3007	-0.5417	0.3009	-0.5441

Tabela 5.2 - Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados - Problema 1

Intervalo de Incerteza	Rede Neural			Ortotope Externo		
	[min , max]	largura	centro	[min , max]	largura	centro
$a_1$	[0.2606 , 0.3981]	0.1375	0.3294	[0.2628 , 0.3960]	0.1332	0.3294
$a_2$	[-0.8054 , -0.5417]	0.2637	-0.6736	[-0.8036 , -0.5441]	0.2595	-0.6739

Tabela 5.3 - Intervalos de Incerteza Paramétrica - Problema 1

Vértices	Rede Neural		Politopo-Exato	
	$a_1$	$a_2$	$a_1$	$a_2$
vértice 1	0.3645	-0.8054	0.3641	-0.8036
vértice 2	0.3981	-0.7922	0.3960	-0.7923
vértice 3	0.2606	-0.6872	0.2628	-0.6882
vértice 4	0.3007	-0.5417	0.3009	-0.5441
vértice 5	0.3433	-0.5538	0.3421	-0.5552

Tabela 5.4 - Vértices da Região de Pertinência Paramétrica - Problema 1

Estimador	Rede Neural	Politopo-Exato	Erro R-P	Erro R-N	Erro P-N
$a_1$	0.3334	0.3332	0.060 %	11.133 %	11.067 %
$a_2$	-0.6761	-0.6767	0.088 %	2.014 %	1.928 %

Tabela 5.5 - Estimadores Paramétricos - Problema 1

A tabela 5.6 fornece os valores previstos 4 passos a frente para  $y$ , com base nos valores estimados para os parâmetros conforme a tabela 5.5.

	Predição-Nominal	Predição-Rede	Predição-Polítipo
$y_{11}$	0.1387	0.1131	0.1134
$y_{12}$	0.6763	0.6637	0.6641
$y_{13}$	0.4072	0.4448	0.4445
$y_{14}$	0.0555	0.0996	0.0987

Tabela 5.6 - Valores de Predição Multipasso - Problema 1

Neste exemplo é importante notar que mesmo para poucas medidas realizadas, a abordagem UBBE fornece resultados atraentes, tanto na obtenção dos Intervalos de Incerteza Paramétrica como na obtenção da Região de Pertinência Paramétrica e dos Estimadores envolvidos.

### 5.4-) PROBLEMA 2 - Modelo MA (Moving Average) de Terceira Ordem.

Seja o modelo MA de terceira ordem, dado por (Problema proposto originalmente por Vicino e Milanese em [7]):

$$y(k) = \lambda_1.u(k) + \lambda_2.u(k-1) + \lambda_3.u(k-1) + e(k) \equiv \phi^T(k).\theta + e(k) , k = 1, \dots, 10 \quad (5.5)$$

onde:  $y(k)$  é a medida realizada na saída do sistema.

$u(k)$  são as medidas realizadas inerentemente ao sistema.

$\lambda_1, \lambda_2$  e  $\lambda_3$  são os parâmetros desconhecidos do modelo e que serão estimados.

$e(k)$  representa a perturbação que possui característica desconhecida-mas-limitada.

A *tabela 5.7* fornece as 12 medidas simuladas para o sistema. Estes dados foram gerados a partir de (5.5) com valores nominais para os parâmetros dado por:

$$\lambda_1^{nominal} = 1.25 , \lambda_2^{nominal} = 2.35 , \lambda_3^{nominal} = 0.50$$

e assumindo-se que  $e(k)$  é uniformemente distribuído no intervalo  $[-0.5, 0.5]$ , ou seja:

$$|e(k)| \leq 0.5, \quad \text{portanto: } e^{max}(k) = -e^{min}(k) = 0.5 \quad (5.6)$$

$k$	-1	0	1	2	3	4	5	6	7	8	9	10
$u(k)$	1.5	2.53	0.08	-0.03	0.43	0.05	0.35	-0.51	-0.71	0.57	-0.02	-0.42
$y(k)$	-	-	1.25	1.00	0.21	0.82	1.12	0.06	-1.72	-0.94	1.04	-0.51

Tabela 5.7 - Medidas Simuladas Para o Problema 2

Os pares de desigualdades, formados a partir da equação (4.18), com base no modelo (5.5) e nos limites a priori do erro (5.6), são dados por:

$$R(k) = \{ \theta / y(k) - 0.5 \leq \phi^T(k) \cdot \theta \leq y(k) + 0.5 \} \quad k = 1, \dots, 10 \quad (5.7)$$

onde:  $\phi^T(k) = [u(k) \quad u(k-1) \quad u(k-2)]$

$$\theta^T = [\lambda_1 \quad \lambda_2 \quad \lambda_3]$$

Os pares de desigualdades, dados por (5.7) formam um politopo em um espaço de três dimensões ( $N=3$ ) com eixos paramétricos  $\lambda_1$ ,  $\lambda_2$  e  $\lambda_3$ . Estes pares de desigualdades podem ser colocados na forma da equação (4.21) com  $M^{in} = 20$ . O exemplo não apresenta nenhuma restrição de igualdade que possa ser colocada na forma da equação (4.20), logo  $M^{eq} = 0$ .

Como nenhuma informação é disponível sobre o domínio a priori dos parâmetros a serem estimados para o modelo, admite-se que em (5.4) os valores de  $c_i^{min}$  e  $c_i^{max}$  sejam respectivamente 0 e 10 para cada parâmetro  $\theta_i$ , isto é:

$$\theta_i \in [0, 10] \quad , \quad i = 1, 2, 3 \quad (5.8)$$

Os vetores de variáveis estendidas para o problema 2 que serão utilizados pela rede de Hopfield modificada são dados por:

$$(\theta^+)^T = [\lambda_1 \quad \lambda_2 \quad \lambda_3 \quad w_1 \quad w_2 \quad \dots \quad w_{20}]$$

$$(b^+) = [b_1 \quad b_2 \quad b_3 \quad \dots \quad b_{20}]$$

$$(\phi^+(i))^T = [\pm u(k) \quad \pm u(k-1) \quad \pm u(k-2) \quad q_1 \quad q_2 \quad q_3 \quad \dots \quad q_{20}] \quad , \quad i = 1, \dots, 20$$

onde os termos  $q_i$  são fornecidos pela equação (4.29).

Utilizando-se a rede de Hopfield modificada, os parâmetros de inicialização desta rede são dados por:

$$\Delta t = 0.001$$

$$a_g = 10 \text{ e } b_g = 0 \text{ para } \theta_1^+, \theta_2^+ \text{ e } \theta_3^+$$

onde os valores para  $a_g$  e  $b_g$  garantem que a saída da rede será confinada no hipercubo definido pelos valores admissíveis dados para cada parâmetros  $\theta_i$  na equação (5.8).

As tabelas de resultados obtidos pela rede para o problema 2 são apresentadas a seguir:

Vértices	vetor $i^{ot}$	Rede Neural			Ortotopo Externo		
		$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_1$	$\lambda_2$	$\lambda_3$
vértice 1	(1, 0, 0)	0.7859	2.2792	0.5303	0.7938	2.2847	0.5301
vértice 2	(-1, 0, 0)	1.7677	2.0002	0.1553	1.7623	2.0060	0.1551
vértice 3	(0, 1, 0)	1.2011	1.3644	0.3032	1.2032	1.3676	0.3035
vértice 4	(0, -1, 0)	1.1636	2.8475	0.1630	1.1639	2.8450	0.1637
vértice 5	(0, 0, 1)	1.2263	2.7322	0.1243	1.2266	2.7316	0.1258
vértice 6	(0, 0, -1)	1.6660	1.7020	0.5600	1.6659	1.7019	0.5588

Tabela 5.8 - Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados - Problema 2

Intervalo de Incerteza	Rede Neural			Ortotopo Externo		
	[min, max]	largura	centro	[min, max]	largura	centro
$\lambda_1$	[0.7859, 1.7677]	0.9818	1.2768	[0.7938, 1.7623]	0.9685	1.2781
$\lambda_2$	[1.3644, 2.8475]	1.4831	2.1060	[1.3676, 2.8450]	1.4774	2.1063
$\lambda_3$	[0.1246, 0.5600]	0.4354	0.3423	[0.1258, 0.5588]	0.4330	0.3423

Tabela 5.9 - Intervalos de Incerteza Paramétrica - Problema 2

Vértices	Rede Neural			Politopo Exato		
	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_1$	$\lambda_2$	$\lambda_3$
vértice 1	1.6760	1.6130	0.4893	1.6731	1.6129	0.4872
vértice 2	1.7566	1.9144	0.1573	1.7554	1.9156	0.1579
vértice 3	1.2011	1.3644	0.3032	1.2032	1.3676	0.3035
vértice 4	-	-	-	1.3445	1.5816	0.1636
vértice 5	1.6660	1.7020	0.5600	1.6659	1.7019	0.5588
vértice 6	-	-	-	0.7139	2.3371	0.5393
vértice 7	1.7677	2.0002	0.1553	1.7623	2.0060	0.1551
vértice 8	0.8523	1.6609	0.5511	0.8576	1.6631	0.5505
vértice 9	1.3944	2.7730	0.5225	1.3930	2.7717	0.5218
vértice 10	0.7859	2.2792	0.5303	0.7938	2.2847	0.5301
vértice 11	1.1636	2.8475	0.1630	1.1639	2.8450	0.1637
vértice 12	1.2263	2.7322	0.1246	1.2266	2.7316	0.1258

Tabela 5.10 - Vértices da Região de Pertinência Paramétrica - Problema 2

Estimador	Rede Neural	Politopo-Exato	Erro R-P	Erro R-N	Erro P-N
$\lambda_1$	1.3490	1.3794	2.2039 %	7.920 %	10.352 %
$\lambda_2$	2.0887	2.0682	0.9912 %	11.119 %	11.991 %
$\lambda_3$	0.3557	0.3548	0.2537 %	28.86 %	29.040 %

Tabela 5.11 - Estimadores Paramétricos - Problema 2

Na *tabela 5.10*, os vértices 4 e 6 não foram identificados pela rede utilizando a perspectiva eixos-hipercúbicos porque estes vértices devem manter um grau de colinearidade muito forte com alguns outros vértices obtidos pela rede. Além disso, vale notar na *tabela 5.11* que o erro dos estimadores calculados pelo método do politopo-exato é maior se comparado com os obtidos pela rede em relação aos valores nominais (Erro R-N e Erro P-N). Portanto, pode-se concluir a priori que a precisão dos estimadores calculados não depende de todos os vértices que compõem a região de pertinência paramétrica exata.

**5.5-) PROBLEMA 3 - Modelo ARMAX (AutoRegressive Moving Average with eXternal input) de Segunda Ordem.**

Seja o modelo ARMAX de segunda ordem, dado por (Problema proposto originalmente por Söderström e Stoica em [5]):

$$y(k) = a_1 \cdot y(k-1) + b_1 \cdot u(k-1) + e(k) \equiv \phi^T(k) \cdot \theta + e(k) , k = 1, \dots, 50 \quad (5.9)$$

onde:  $y(k)$  é a medida realizada na saída do sistema.

$u(k)$  é uma sequência binária pseudo-aleatória de amplitude  $\pm 1$  (Apêndice D) .

$a_1$  e  $b_1$  são os parâmetros desconhecidos do modelo e que serão estimados.

$e(k)$  representa a perturbação que possui característica desconhecida-mas-limitada .

Os dados foram gerados a partir de (5.9) com valores nominais para os parâmetros dado por:

$$a_1^{nominal} = 0.8 , b_1^{nominal} = 1.0$$

e assumindo-se que  $e(k)$  é uma distribuição Normal com média nula e variância unitária e que é limitada no intervalo  $[-1, 1]$  , ou seja:

$$|e(k)| \leq 1 , \quad \text{portanto: } e^{max}(k) = -e^{min}(k) = 1.0 \quad (5.10)$$

O domínio a priori dos parâmetros é dado por:

$$\Omega : \begin{cases} 0.70 \leq a_1 \leq 0.85 \\ 0.95 \leq b_1 \leq 1.50 \end{cases}$$

Portanto, os pares de desigualdades formados, a partir da equação (4.18), com base no modelo (5.9) e com os limites a priori do erro (5.10), são dados por:

$$R(k) = \{ \theta / y(k) - 1.0 \leq \phi^T(k) \cdot \theta \leq y(k) + 1.0 \} \quad k = 1, \dots, 50 \quad (5.11)$$

onde:  $\phi^T(k) = [y(k-1) \quad u(k-1)]$

$$\theta^T = [a_1 \quad b_1]$$

Os pares de desigualdades dado por (5.10) formam um politopo em um espaço de duas dimensões ( $N=2$ ) com eixos paramétricos  $a_1$  e  $b_1$  . Estes pares de desigualdades podem ser colocados na forma da equação (4.21) com  $M^{in} = 100$ . Como nos exemplos anteriores, não existe nenhuma restrição de igualdade, logo  $M^{eq} = 0$ .

Como o problema fornece informações sobre o domínio a priori dos parâmetros a serem estimados para o modelo, os valores de  $c_i^{min}$  e  $c_i^{max}$  para cada parâmetro  $\theta_i$  em (5.4) torna-se :

$$\begin{aligned} \theta_1 &\in [ 0.70 , 0.85 ] & (5.12) \\ \theta_2 &\in [ 0.95 , 1.50 ] \end{aligned}$$

Os vetores de variáveis estendidas para o problema 3 que serão utilizados pela rede de Hopfield modificada são dados por:

$$\begin{aligned} (\theta^+)^T &= [a_1 \ b_1 \ w_1 \ w_2 \ \dots \ w_{100}] \\ (b^+) &= [ b_1^+ \ b_2^+ \ b_3^+ \ \dots \ b_{100}^+ ] \\ (\phi^+(i))^T &= [\pm y(k-1) \ \pm u(k-1) \ q_1 \ q_2 \ q_3 \ \dots \ q_{100}] \ , \ i = 1, \dots, 100 \end{aligned}$$

onde os termos  $q_j$  são fornecidos pela equação (4.29), e os símbolos superscript nos elementos do vetor  $b^+$  são apenas para diferenciar do estimador de  $b_1$  no vetor  $\theta^+$ .

Utilizando-se a rede de Hopfield modificada, os parâmetros de inicialização desta rede são dados por:

$$\begin{aligned} \Delta t &= 0.001 \\ a_g &= 1.65 \text{ e } b_g = -0.70 \ , \text{ para } \theta_1^+ \\ a_g &= 2.05 \text{ e } b_g = -0.85 \ , \text{ para } \theta_2^+ \end{aligned}$$

onde os valores para  $a_g$  e  $b_g$  estão de acordo com os valores admissíveis para cada parâmetro  $\theta_i$ , delimitados pelo domínio a priori dos parâmetros dado por (5.12).

As tabelas de resultados obtidos pela rede para o problema 3 são apresentadas a seguir:

Vértices	vetor $i^{ot}$	Rede Neural		Ortotope Externo	
		$a_1$	$b_1$	$a_1$	$b_1$
vértice 1	( 1 , 0 )	0.7212	0.9861	0.7234	0.9866
vértice 2	(-1 , 0)	0.8414	0.9824	0.8385	0.9841
vértice 3	( 0 , 1 )	0.7924	0.9649	0.7931	0.9666
vértice 4	( 0 , -1 )	0.7719	1.0604	0.7714	1.0578

**Tabela 5.12 - Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados - Problema 3**

Intervalo de Incerteza	Rede Neural			Ortotope Externo		
	[min , max]	largura	centro	[min , max]	largura	centro
$a_1$	[0.7212 , 0.8414]	0.1202	0.7813	[0.7234 , 0.8385]	0.1151	0.7809
$b_2$	[0.9649 , 1.0604]	0.0955	1.0126	[0.9666 , 1.0578]	0.0912	1.0122

Tabela 5.13 - Intervalos de Incerteza Paramétrica - Problema 3

Vértices	Rede Neural		Politopo-Exato	
	$a_1$	$b_1$	$a_1$	$b_1$
vértice 1	0.8418	0.9812	0.8392	0.9824
vértice 2	0.7924	0.9649	0.7931	0.9666
vértice 3	0.8414	0.9824	0.8385	0.9841
vértice 4	0.7212	0.9861	0.7234	0.9866
vértice 5	0.7719	1.0604	0.7714	1.0578

Tabela 5.14 - Vértices da Região de Pertinência Paramétrica - Problema 3

Estimador	Rede Neural	Politopo-Exato	Erro R-P	Erro R-N	Erro P-N
$a_1$	0.7937	0.7931	0.076 %	0.788 %	0.863 %
$b_1$	0.9950	0.9955	0.050 %	0.500 %	0.450 %

Tabela 5.15 - Estimadores Paramétricos - Problema 3

Neste exemplo mostra-se como é feito o mapeamento de problemas de estimação paramétrica robusta quando utilizam-se limites mínimo e máximo para cada parâmetro a ser estimado. Vale salientar que o tempo de convergência para soluções válidas diminui somente se houver interferência dos parâmetros  $c_i^{min}$  e  $c_i^{max}$  na região de pertinência paramétrica exata. Entretanto, a precisão dos parâmetros estimados independe dos limites escolhidos a priori para cada parâmetro a ser estimado.

Na *figura 5.1*, a linha tracejada corresponde aos valores da saída do modelo (5.9) utilizando os parâmetros estimados pela rede (*tabela 5.15*) em função de um novo conjunto de medidas, enquanto a linha cheia corresponde aos valores da saída do modelo quando se utilizam os parâmetros nominais (as duas linhas se coincidem).

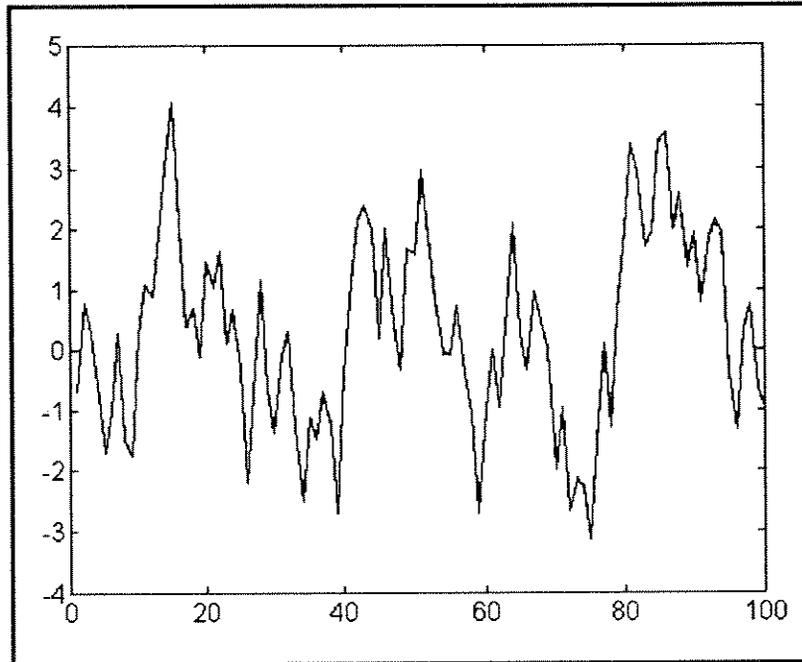


Figura 5.1 - Gráfico da Saída do Modelo em Função de Um Conjunto de Medidas - Problema 3

### 5.6-) PROBLEMA 4 - Modelo ARMAX de Terceira Ordem.

Seja o modelo ARMAX de terceira ordem, dado por:

$$y(k) = a_1 \cdot y(k-1) + b_1 \cdot u(k-1) + b_2 \cdot u(k-2) + e(k) \equiv \phi^T(k) \cdot \theta + e(k) , k = 1, \dots, 50 \quad (5.13)$$

onde:  $y(k)$  é a medida realizada na saída do sistema.

$u(k)$  é uma sequência binária pseudo-aleatória de amplitude  $\pm 1$ .

$a_1$ ,  $b_1$  e  $b_2$  são os parâmetros desconhecidos do modelo que serão estimados.

$e(k)$  representa a perturbação que possui característica desconhecida-mas-limitada

Os dados foram gerados a partir de (5.13) com valores nominais para os parâmetros dado por:

$$a_1^{nominal} = 0.8 , b_1^{nominal} = 1.0 , b_2^{nominal} = 0.3$$

e assumindo-se que  $e(k)$  é uma distribuição Normal com média nula e variância unitária e que é limitada no intervalo  $[-1, 1]$ , ou seja:

$$|e(k)| \leq 1, \quad \text{portanto: } e^{\max}(k) = -e^{\min}(k) = 1.0 \quad (5.14)$$

Portanto, os pares de desigualdades, que são formados a partir da equação (4.18), com base no modelo (5.13) e nos limites a priori do erro (5.14) serão dados por:

$$R(k) = \{ \theta / y(k) - 1.0 \leq \phi^T(k) \cdot \theta \leq y(k) + 1.0 \} \quad k = 1, \dots, 50 \quad (5.15)$$

onde:  $\phi^T(k) = [y(k-1) \quad u(k-1) \quad u(k-2)]$

$$\theta^T = [a_1 \quad b_1 \quad b_2]$$

Os pares de desigualdades dados por (5.15) formam um politopo em um espaço de três dimensões ( $N=3$ ) com eixos paramétricos  $a_1$ ,  $b_1$  e  $b_2$ . Estes pares de desigualdades podem ser colocados na forma da equação (4.21) com  $M^{in} = 100$ . Como nos exemplos anteriores, não existe nenhuma restrição de igualdade, logo  $M^{eq} = 0$ .

Como nenhuma informação é disponível sobre o domínio a priori dos parâmetros a serem estimados para o modelo, admite-se que em (5.4) os valores de  $c_i^{\min}$  e  $c_i^{\max}$  sejam respectivamente -10 e 10 para cada parâmetro  $\theta_i$ , isto é:

$$\theta_i \in [-10, 10], \quad i = 1, 2, 3 \quad (5.16)$$

Os vetores de variáveis estendidas para o problema 4 que serão utilizados pela rede de Hopfield modificada são dados por:

$$(\theta^+)^T = [a_1 \quad b_1 \quad b_2 \quad w_1 \quad w_2 \quad \dots \quad w_{100}]$$

$$(b^+) = [b_1^+ \quad b_2^+ \quad b_3^+ \quad \dots \quad b_{100}^+]$$

$$(\phi^+(i))^T = [\pm y(k-1) \quad \pm u(k-1) \quad \pm u(k-2) \quad q_1 \quad q_2 \quad q_3 \quad \dots \quad q_{100}] \quad , \quad i = 1, \dots, 100$$

onde os termos  $q_i$  são fornecidos pela equação (4.29), e os símbolos superscript nos elementos do vetor  $b^+$  são apenas para diferenciar dos estimadores  $b_1$  e  $b_2$  no vetor  $\theta^+$ .

Utilizando-se a rede de Hopfield modificada, os parâmetros de inicialização desta rede são dados por:

$$\Delta t = 0.001$$

$$a_g = 20 \quad \text{e} \quad b_g = 10 \quad \text{para} \quad \theta_1^+, \theta_2^+ \quad \text{e} \quad \theta_3^+$$

onde os valores para  $a_g$  e  $b_g$  garantem que a saída da rede está confinada no hiper-cubo definido pelos intervalos dados para cada parâmetro  $\theta_i$  na equação (5.16).

As tabelas de resultados obtidos pela rede para o problema 4 são apresentadas a seguir:

Vértices	vetor $i^{ot}$	Rede Neural			Ortopoto Externo		
		$a_1$	$b_1$	$b_2$	$a_1$	$b_1$	$b_2$
vértice 1	(1, 0, 0)	0.7042	0.9929	0.3309	0.7068	0.9781	0.3162
vértice 2	(-1, 0, 0)	0.8430	0.9803	0.2800	0.8398	0.9687	0.2933
vértice 3	(0, 1, 0)	0.7137	0.9604	0.3067	0.7132	0.9627	0.3076
vértice 4	(0, -1, 0)	0.8274	1.0726	0.2278	0.8275	1.0691	0.2283
vértice 5	(0, 0, 1)	0.8423	1.0622	0.1954	0.8398	1.0605	0.2015
vértice 6	(0, 0, -1)	0.7115	1.0384	0.3845	0.7113	1.0364	0.3813

Tabela 5.16 - Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados - Problema 4

Intervalo de Incerteza	Rede Neural			Ortopoto Externo		
	[min, max]	largura	centro	[min, max]	largura	centro
$a_1$	[0.7042, 0.8430]	0.1388	0.7736	[0.7068, 0.8398]	0.1330	0.7733
$b_1$	[0.9604, 1.0726]	0.1122	1.0165	[0.9627, 1.0691]	0.1064	1.0159
$b_2$	[0.1954, 0.3845]	0.1891	0.2900	[0.2015, 0.3813]	0.1798	0.2914

Tabela 5.17 - Intervalos de Incerteza Paramétrica - Problema 4

Vértices	Rede Neural			Politopo Exato		
	$a_1$	$b_1$	$b_2$	$a_1$	$b_1$	$b_2$
vértice 1	0.7115	1.0384	0.3845	0.7132	1.0364	0.3813
vértice 2	0.8430	0.9803	0.2800	0.8398	0.9687	0.2933
vértice 3	-	-	-	0.8275	0.9855	0.3120
vértice 4	0.7137	0.9604	0.3067	0.7132	0.9627	0.3076
vértice 5	0.7074	1.0452	0.3832	0.7068	1.0432	0.3813
vértice 6	0.7042	0.9929	0.3309	0.7068	0.9781	0.3162
vértice 7	0.8423	1.0622	0.1954	0.8398	1.0605	0.2015
vértice 8	0.8274	1.0726	0.2278	0.8275	1.0691	0.2283
vértice 9	0.7119	0.9616	0.3057	0.7132	0.9708	0.2994

Tabela 5.18 - Vértices da Região de Pertinência Paramétrica - Problema 4

Estimador	Rede Neural	Politopo-Exato	Erro R-P	Erro R-N	Erro P-N
$a_1$	0.7577	0.7601	0.316 %	5.288 %	4.988 %
$b_1$	1.0142	1.0038	1.036 %	1.420 %	0.380 %
$b_2$	0.3018	0.3028	0.330 %	0.600 %	0.933 %

Tabela 5.19 - Estimadores Paramétricos - Problema 4

Na *figura 5.2*, a linha tracejada corresponde aos valores da saída do modelo (5.13) utilizando os parâmetros estimados pela rede (*tabela 5.19*) em função de um novo conjunto de medidas, enquanto a linha cheia corresponde aos valores da saída do modelo quando se utilizam os parâmetros nominais.

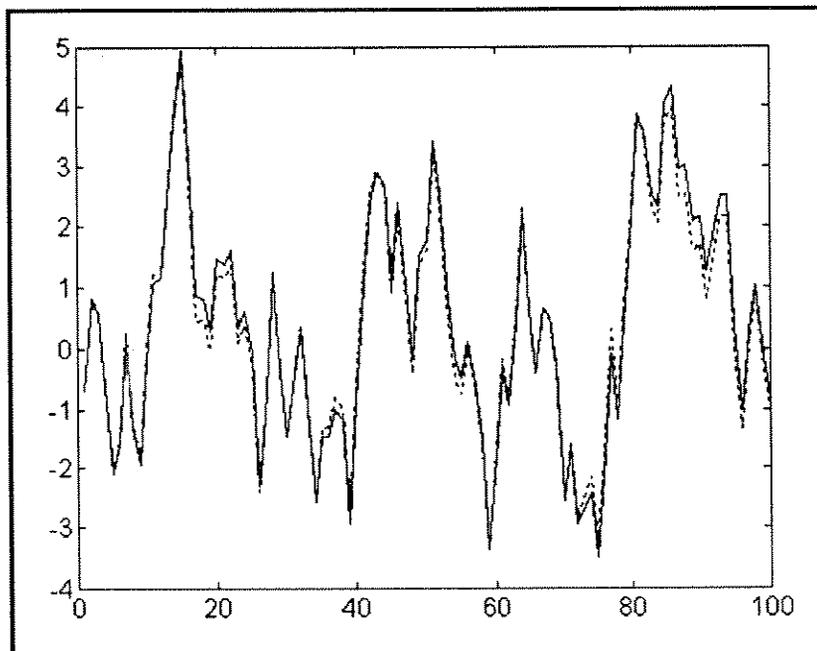


Figura 5.2 - Gráfico da Saída do Modelo em Função de Um Conjunto de Medidas - Problema 4

### 5.7-) PROBLEMA 5 - Modelo ARMAX de Quarta Ordem.

Seja o modelo ARMAX de quarta ordem, dado por (Problema proposto originalmente por Ljung em [2]):

$$y(k) = a_1 \cdot y(k-1) + a_2 \cdot y(k-2) + b_1 \cdot u(k-1) + b_2 \cdot u(k-2) + e(k) = \phi^T(k) \cdot \theta + e(k) , k = 1, \dots, 50 \quad (5.17)$$

onde:  $y(k)$  é a medida realizada na saída do sistema.

$u(k)$  é uma sequência binária pseudo-aleatória de amplitude  $\pm 1$ .

$a_1, a_2, b_1$  e  $b_2$  são os parâmetros desconhecidos do modelo que serão estimados.

$e(k)$  representa a perturbação que possui característica desconhecida-mas-limitada

Os dados foram gerados a partir de (5.17) com valores nominais para os parâmetros dado por:

$$a_1^{nominal} = 1.5 , a_2^{nominal} = -0.7 , b_1^{nominal} = 1.0 , b_2^{nominal} = 0.5$$

e assumindo-se que  $e(k)$  é uma distribuição Normal com média nula e variância unitária e que é limitada no intervalo  $[-1, 1]$ , ou seja:

$$|e(k)| \leq 1 , \quad \text{portanto: } e^{max}(k) = -e^{min}(k) = 1.0 \quad (5.18)$$

Os pares de desigualdades, que serão formados a partir da equação (4.18), com base no modelo (5.17) e nos limites a priori do erro (5.18) serão dados por:

$$R(k) = \{ \theta / y(k) - 1.0 \leq \phi^T(k) \cdot \theta \leq y(k) + 1.0 \} \quad k = 1, \dots, 50 \quad (5.19)$$

onde:  $\phi^T(k) = [y(k-1) \quad y(k-2) \quad u(k-1) \quad u(k-2)]$

$$\theta^T = [a_1 \quad a_2 \quad b_1 \quad b_2]$$

Os pares de desigualdades dados por (5.19) formam um politopo em um espaço de quatro dimensões ( $N=4$ ) com eixos paramétricos  $a_1, a_2, b_1$  e  $b_2$ . Estes pares de desigualdades podem ser colocados na forma da equação (4.21) com  $M^{in} = 100$ . Como nos exemplos anteriores, não existe nenhuma restrição de igualdade, logo  $M^{eq} = 0$ .

Como nenhuma informação é disponível sobre o domínio a priori dos parâmetros a serem estimados para o modelo, admite-se que em (5.4) os valores de  $c_i^{min}$  e  $c_i^{max}$  sejam respectivamente -100 e 100 para cada parâmetro  $\theta_i$ , isto é:

$$\theta_i \in [-100, 100] , \quad i = 1, 2, 3, 4 \quad (5.20)$$

Os vetores de variáveis estendidas para o problema 5 e que serão utilizados pela rede de Hopfield modificada são dados por:

$$(\theta^+)^T = [a_1 \ a_2 \ b_1 \ b_2 \ w_1 \ w_2 \ \dots \ w_{100}]$$

$$(b^+) = [b_1^+ \ b_2^+ \ b_3^+ \ \dots \ b_{100}^+]$$

$$(\phi^+(i))^T = [\pm y(k-1) \ \pm y(k-2) \ \pm u(k-1) \ \pm u(k-2) \ q_1 \ q_2 \ q_3 \ \dots \ q_{100}], \ i = 1, \dots, 100$$

onde os termos  $q_i$  são calculados pela equação (4.29), e os símbolos superscript nos elementos do vetor  $b^+$  são apenas para diferenciar dos estimadores  $b_1$  e  $b_2$  no vetor  $\theta^+$ .

Utilizando-se a rede de Hopfield modificada, os parâmetros de inicialização desta rede para o problema 5 são dados por:

$$\Delta t = 0.001$$

$$a_g = 200 \text{ e } b_g = 100 \text{ para } \theta_1^+, \theta_2^+, \theta_3^+ \text{ e } \theta_4^+$$

onde os valores para  $a_g$  e  $b_g$  garantem que a saída da rede está entre os respectivos valores admissíveis para cada parâmetro  $\theta_i$  como definido na equação (5.20).

As tabelas de resultados obtidos pela rede para o problema 5 são apresentadas a seguir:

Vértices	vetor $i^{ot}$	Rede Neural				Ortotopo Externo			
		$a_1$	$a_2$	$b_1$	$b_2$	$a_1$	$a_2$	$b_1$	$b_2$
vértice 1	(1, 0, 0, 0)	1.2964	-0.5975	0.9963	0.5521	1.3084	-0.6004	0.9927	0.5466
vértice 2	(-1, 0, 0, 0)	1.6415	-0.8170	1.0960	0.2397	1.6296	-0.8158	1.0962	0.2433
vértice 3	(0, 1, 0, 0)	1.6309	-0.8272	1.0939	0.2416	1.6296	-0.8158	1.0955	0.2433
vértice 4	(0, -1, 0, 0)	1.3146	-0.5596	1.1478	0.7995	1.3171	-0.5782	1.1246	0.7645
vértice 5	(0, 0, 1, 0)	1.4016	-0.6539	0.6837	0.5358	1.3994	-0.6522	0.7092	0.5405
vértice 6	(0, 0, -1, 0)	1.4021	-0.6259	1.3610	0.7388	1.4109	-0.6404	1.3302	0.6922
vértice 7	(0, 0, 0, 1)	1.6240	-0.7960	1.1354	0.1325	1.6221	-0.7959	1.1332	0.1642
vértice 8	(0, 0, 0, -1)	1.3136	-0.5536	1.2119	0.9055	1.3341	-0.5802	1.0543	0.7669

Tabela 5.20 - Vértices Calculados Utilizando a Perspectiva Eixos-Direcionados - Problema 5

Intervalo de Incerteza	Rede Neural			Ortotope Externo		
	[min , max]	largura	centro	[min , max]	largura	centro
$a_1$	[1.2964 , 1.6415]	0.3451	1.4690	[1.3084 , 1.6296]	0.3212	1.4544
$a_2$	[-0.8272 , -0.5596]	0.2676	-0.6934	[-0.8158 , -0.5782]	0.2376	-0.6967
$b_1$	[0.6837 , 1.3610]	0.6773	1.0223	[0.7092 , 1.3302]	0.6210	1.0354
$b_2$	[0.1325 , 0.9055]	0.7730	0.5090	[0.1642 , 0.7669]	0.6027	0.4432

Tabela 5.21 - Intervalos de Incerteza Paramétrica - Problema 5

A aplicação da rede de Hopfield modificada neste problema resultou na estimação de uma região de pertinência paramétrica com 24 vértices. Para este mesmo problema, o método do politopo-exato estimou uma região composta por 34 vértices. Os estimadores obtidos como o critério do centro de massa para estas duas regiões são fornecidos na tabela 5.22.

Estimador	Rede Neural	Politopo-Exato	Erro R-P	Erro R-N	Erro P-N
$a_1$	1.4656	1.4544	0.770 %	2.293 %	3.040 %
$a_2$	-0.6858	-0.6967	1.565 %	2.029 %	0.471 %
$b_1$	1.0645	1.0354	2.811 %	6.450 %	3.540 %
$b_2$	0.4768	0.4432	7.581 %	4.640 %	11.360 %

Tabela 5.22 - Estimadores Paramétricos - Problema 5

Na *tabela 5.22* deste exemplo nota-se também que os parâmetros  $a_1$  e  $b_2$  são melhores estimados pela rede (Erro R-N), enquanto os parâmetros  $a_2$  e  $b_1$  são melhores estimados pelo método do politopo-exato (Erro P-N). Este exemplo reforça a idéia de que a precisão dos estimadores não é função direta da quantidade de vértices encontrados.

Na *figura 5.3*, a linha tracejada corresponde aos valores da saída do modelo (5.20) utilizando os parâmetros estimados pela rede (*tabela 5.22*) em função de um novo conjunto de medidas, enquanto a linha cheia corresponde aos valores da saída do modelo quando se utilizam os parâmetros nominais.

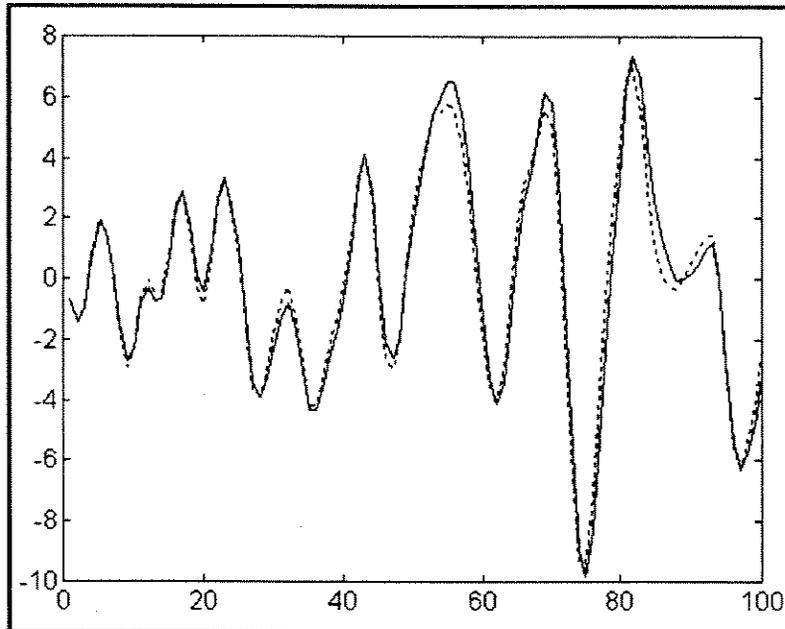


Figura 5.3 - Gráfico da Saída do Modelo em Função de Um Conjunto de Medidas - Problema 5

## 5.8-) CONCLUSÃO

O método para realizar Estimação Paramétrica Robusta utilizando redes neurais possui algumas peculiaridades em relação aos demais métodos clássicos utilizados nas simulações anteriores.

Um aspecto importante é que quando se utiliza a rede de Hopfield, os Intervalos de Incerteza Paramétrica podem ser obtidos diretamente pela aplicação da perspectiva eixos-direcionados, enquanto que a obtenção destes intervalos no método politopo-exato é realizada a partir do cálculo de todos os vértices do politopo. No método ortotopo-externo a obtenção destes intervalos também é feita de forma direta. Deve-se salientar que a forma original do algoritmo do método ortotopo-externo não se aplica quando os valores dos parâmetros a serem estimados são negativos, geralmente este problema é contornado com uma translação adequada dos eixos paramétricos [1]. Em suma, pode-se afirmar que os três métodos utilizados (rede, politopo-exato, ortotopo-externo) fornecem uma caracterização precisa dos Intervalos de Incerteza Paramétrica.

Com relação a Região de Pertinência Paramétrica, a rede de Hopfield produz uma caracterização bastante realística desta região se comparada com a descrição exata fornecida pelo método politopo-exato. Deve-se ressaltar que o método ortotopo-externo geralmente é utilizado no cálculo de intervalos de incerteza, haja visto que este método produz uma descrição muito pessimista da Região de Pertinência Paramétrica.

Foi observado também que o número de vértices produzidos para caracterizar a Região de Pertinência cresce em razão da dimensão do problema (ordem do problema), mas este número varia lentamente acima de uma certa quantidade de medidas realizadas (geralmente acima de 40 medidas). Este fato também foi confirmado por Mo e Norton em [4].

Em relação aos estimadores obtidos por estes métodos pode-se concluir que o estimador central obtido pela rede de Hopfield é praticamente idêntico ao obtido pelo método ortotopo-externo (*tabelas 5.3- 5.9- 5.13- 5.17- 5.21*). Conclui-se também que os estimadores obtidos pela rede a partir da perspectiva eixos-hipercúbicos e calculados segundo o critério do centro de massa do politopo diferem muito pouco dos estimadores obtidos pelo mesmo critério no método politopo-exato (erro relativo geralmente menor que 2% - *tabelas 5.5-5.11-5.15-5.19-5.22*). No entanto, a precisão dos estimadores obtidos pela rede com relação aos parâmetros nominais (Erro R-N) pode ser melhor ou pior do que a precisão dos estimadores obtidos pelo método do politopo-exato com relação aos parâmetros nominais (Erro P-N).

Vale ressaltar portanto, que a quantidade de vértices encontrados pela rede utilizando a perspectiva eixos-hipercúbicos não interfere na precisão dos estimadores em relação aos parâmetros nominais, visto que a perspectiva eixos-hipercúbicos fornece uma distribuição geometricamente equilibrada dos vetores de busca  $i^{oi}$  dentro do espaço paramétrico.

Nos problemas de estimação paramétrica robusta simulados pela rede de Hopfield modificada neste capítulo, adotou-se para o parâmetro  $\Delta t$  desta rede o valor de 0.001. Este valor foi obtido empiricamente levando em conta o tempo de convergência da rede e também a precisão das soluções obtidas pela rede. Vale salientar também que, quaisquer valores assumidos para os parâmetros  $a_g$  e  $b_g$  das funções de ativação "rampa simétrica" associadas aos problemas que não definem um domínio a priori dos parâmetros a serem estimados, não interferem no tempo de convergência da rede para a obtenção de soluções

válidas, haja visto que o hipercubo gerado pelos parâmetros  $a_g$  e  $b_g$  contém todo o politopo que delimita a região de pertinência paramétrica real para o problema.

## **BIBLIOGRAFIA**

- [1] Bazarara, M. S. & Jarvis, J. J. - **Linear Programming and Network Flows**, John Wiley & Sons, Toronto, 1977.
- [2] Ljung, L. - **System Identification: Theory For the User**, Prentice-Hall, Englewood Cliffs, NJ, 1989.
- [3] Milanese, M. & Vicino, A. - "Estimation Theory For Nonlinear Models and Set Membership Uncertainty", *Automatica*, 27 (1991) 403-408.
- [4] Mo, S. H. & Norton, J. P. - "Fast and Robust Algorithm to Compute Exact Polytope Parameter Bounds", *Math. Comput. Simulation*, 32 (5&6) (1990) 481-493.
- [5] Söderström, T. & Stoica, P. - **System Identification**, Prentice-Hall, Englewood Cliffs, NJ, 1989.
- [6] The Math Works Inc. - **The Student Edition of Matlab for MS-DOS Personal Computers: User's Guide**, Prentice-Hall, Englewood Cliffs, NJ, 1992.
- [7] Vicino, A. & Milanese, M. - "Optimal Inner Bounds of Feasible Parameter Sets in Linear Estimation With Bounded Noise", in: *Proc 28th IEEE Conference on Decision and Control*, Tampa (1989) 2576-2580.

## *Capítulo 6*

# **Linhas Futuras de Pesquisas e Conclusões**

### **6.1-) SUMÁRIO DA DISSERTAÇÃO**

Nesta dissertação, foi mostrado como problemas de Estimação Paramétrica Robusta podem ser formulados e resolvidos através de Redes Neurais Artificiais (em particular, a rede de Hopfield). Uma nova estrutura matemática foi utilizada na derivação dos parâmetros internos da rede de Hopfield com o objetivo de agregar de forma simples e compacta as várias restrições que são impostas por estes problemas quando se assume que suas perturbações possuem características desconhecidas-mas-limitadas (abordagem UBBE).

Esta nova estrutura matemática baseia-se na premissa de que todos os vetores de saídas da rede de Hopfield que correspondem às soluções válidas de um problema específico, pertencem a um mesmo subespaço de vetores. Este subespaço foi denominado subespaço-válido de soluções.

Com a finalidade de aperfeiçoar a eficiência de simulações por computador da rede de Hopfield proposta originalmente, uma rede, denominada rede de Hopfield modificada,

foi descrita. Esta rede confina suas saídas do subespaço-válido por projeção direta e foi mostrado que é funcionalmente equivalente à rede de Hopfield original.

Finalmente, como uma demonstração das aplicações práticas e das potencialidades desta rede, simulações de problemas de Estimção Paramétrica Robusta, em especial problemas caracterizados por modelos AR (AutoRegressive), MA (Moving Average) e ARMAX (AutoRegressive Moving Average eXternal input). Estas simulações foram realizadas utilizando as duas perspectivas propostas nesta tese para selecionar os vetores ótimos ( $i^{o}$ ) da rede de Hopfield proposta.

## **6.2-) CONCLUSÕES**

A pesquisa sobre a aplicação de Redes Neurais Artificiais para solucionar problemas de Estimção Paramétrica Robusta está ainda em fase inicial e já produz resultados no mínimo estimulantes como foi confirmado por esta tese.

Além de fornecer um novo método para solucionar problemas de Estimção Paramétrica Robusta, a rede de Hopfield apresentada permitiu tratar os métodos ortotopo externo e politopo exato em um modo unificado, mudando apenas o tipo de perspectiva utilizada na rede.

Como pontos positivos desta aplicação pode-se citar a caracterização quase exata dos Intervalos de Incerteza Paramétrica associados com cada problema, e também numa definição muito realística das Regiões de Pertinência Paramétrica produzidas por este tipo de rede. Entretanto, deve-se salientar que a eficiência desta aplicação depende fundamentalmente de uma escolha adequada dos parâmetros de convergência, isto é: o parâmetro  $\Delta t$  e  $c_0$  para a rede de Hopfield original e apenas o parâmetro  $\Delta t$  para a rede de Hopfield modificada.

Como principal desvantagem da utilização da rede de Hopfield em problemas envolvendo Estimção Paramétrica Robusta pode-se citar o aspecto de tempo de convergência desta rede para a obtenção de soluções válidas, e também o aspecto de limitação desta rede em resolver problemas em tempo-real, visto que esta rede não possui característica recursiva em relação a um conjunto de medidas realizadas anteriormente a uma medida corrente.

Em suma, devido aos avanços recentes nas teorias sobre as redes neurais e principalmente ao rápido crescimento das atividades relacionadas a elas, abriu-se um enorme campo de pesquisa e desenvolvimento das mesmas. Além disso, o fato das redes neurais serem modelos de sistemas nervosos biológicos, capazes de aprender com grande desenvoltura, contribui para torná-las uma matéria de estudos muito instigante.

Para concluir, apresenta-se a seguir uma breve comparação do método de estimação paramétrica robusta utilizando redes neurais com o método do politopo-exato:

***Pontos Positivos da Rede***

- A obtenção dos intervalos de incerteza paramétrica é realizada diretamente pela aplicação da perspectiva eixos-direcionados.
- O mapeamento dos problemas é feito de maneira simples para qualquer número de parâmetros a serem estimados.
- Maior abrangência no que se refere às restrições geradas pelos problemas de estimação paramétrica robusta (a rede incorpora restrições de igualdade e limites a priori dos parâmetros a serem estimados).

***Ponto Negativo da Rede***

- A rede não é recursiva.

***Pontos Positivos do Politopo-Exato***

- O método do politopo-exato é recursivo.
- Caracterização exata da região de pertinência paramétrica

***Pontos Negativos do Politopo-Exato***

- O mapeamento de problemas de estimação paramétrica robusta torna-se complicado para um número de parâmetros a serem estimados superior a 4.
- A obtenção dos intervalos de incerteza paramétrica não é realizada de forma direta.

***Ponto Comum***

- Os dois métodos fornecem estimadores com boa precisão.

### 6.3-) LINHAS FUTURAS DE PESQUISAS

Como fruto desta tese, sugerem-se ainda quatro linhas principais de pesquisas com extensões teóricas e experimentais.

A primeira consiste em estender a técnica de subespaço-válido de soluções para problemas de Estimção Paramétrica Robusta com características não lineares, haja visto que a linearização de modelos não lineares causam indefinição e imprecisão, tanto dos intervalos de incerteza paramétrica quanto da região de pertinência paramétrica.

A segunda linha de pesquisa diz respeito a novas perspectivas de seleção dos vetores de busca  $i^{ot}$  para problemas de Estimção Paramétrica Robusta com características lineares. Sugere-se como pesquisa futura, a seleção destes vetores por meios de métodos utilizando algoritmos de Busca Tabu (Tabu Search) ou algoritmos Genéticos (Genetic Algorithms).

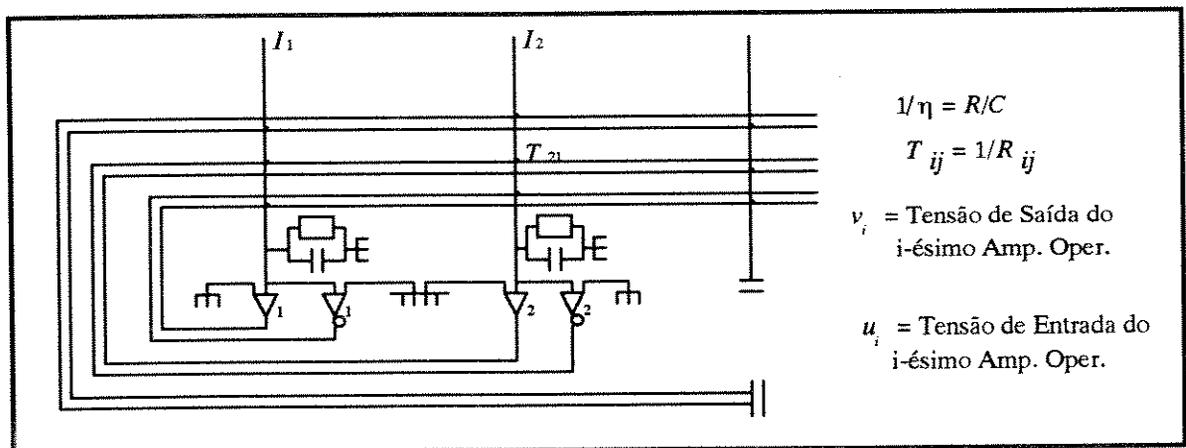
A terceira linha de pesquisa é analisar as modificações e implicações na implementação em hardware da rede de Hopfield modificada para resolver estes tipos de problemas.

A quarta linha de pesquisa é derivar outras estruturas de redes neurais artificiais, até mesmo com topologias híbridas, com a objetivo de solucionar problemas de Estimção Paramétrica Robusta nas aplicações em tempo real.

## APÊNDICE A

Uma das razões principais para a popularidade da rede de Hopfield é a possibilidade de implementá-la em *hardware*. Quando Hopfield e Tank propuseram a rede contínua, eles também demonstraram que existe um hardware analógico equivalente muito simples (*figura A.1*). Neste hardware, os neurônios são modelados através de amplificadores operacionais com relação entrada-saída determinada pela função sigmoideal (3.3).

Cada neurônio possui um resistor  $R_i^{in}$  e um capacitor  $C_i^{in}$  de entrada que definem uma constante de tempo dos neurônios e estabelece a soma analógica *integrativa* das correntes de entrada vindas de outros neurônios da rede. Uma conexão entre dois neurônios é definida por uma condutância  $T_{ij}$  que conecta uma das duas saídas do amplificador  $j$  a entrada do amplificador  $i$ . Esta conexão é feita com um resistor de valor  $R_{ij} = 1/|T_{ij}|$ . Se a sinapse (conexão) é excitatória ( $T_{ij} \geq 0$ ), este resistor é conectado a saída normal (+) do



*Figura A.1* - Hardware Analógico da Rede de Hopfield

amplificador  $j$ . Para uma sinapse inibitória ( $T_{ij} < 0$ ), é conectada a saída inversa (-) do amplificador  $j$ . A matriz  $T$  define então a conectividade entre os amplificadores operacionais (neurônios). A corrente de entrada em qualquer neurônio  $i$  é a soma das correntes fluindo através do conjunto de resistores que conecta suas entradas às saídas dos outros neurônios.

Como indicado na (*figura A.1*), este circuito também inclui uma corrente de entrada  $I_i$  fornecida externamente para cada neurônio.

As equações que descrevem a evolução deste circuito em relação ao tempo é definida por,

$$C_i \left( \frac{u_i}{dt} \right) = \sum_{j=1}^N T_{ij} \cdot V_j - \frac{u_i}{R_i} + I_i \quad (\text{A.1})$$

$$V_j = g_j(u_j)$$

Definimos o termo  $R_i$  como a associação em paralelo de  $R_i^{in}$  e os termos  $R_{ij}$ , ou seja,

$$\frac{1}{R_i} = \frac{1}{R_i^{in}} + \sum_{j=1}^N \left( \frac{1}{R_{ij}} \right) \quad (\text{A.2})$$

Assumindo que todos os neurônios possuem as mesmas configurações, ou seja,  $g_i = g$ ,  $R_i = R$  e  $C_i = C$ ; e dividindo o segundo membro da equação (A.1) por  $C$  e redefinindo  $T_{ij} / C$  e  $I_i / C$  como  $T_{ij}$  e  $I_i$  respectivamente, a equação (A.1) torna-se

$$\frac{du_i}{dt} = -\eta u_i + \sum_{j=1}^N T_{ij} \cdot V_j + I_i \quad (\text{A.3})$$

$\eta = 1 / RC$  e  $V_j = g(u_j)$ . Vale notar que a equação (A.3) é idêntica a equação (3.4).

Esta equação (A.3) fornece uma descrição completa da evolução temporal do estado do circuito. A integração desta equação em computador digital permite a simulação de qualquer rede hipotética.

## APÊNDICE B

Dado o subespaço ortogonal ao subespaço-válido, denominado subespaço-inválido (inv), então qualquer vetor pode ser decomposto em duas componentes pertencentes a cada um destes subespaços, isto é,

$$\mathbf{v} = \mathbf{v}^{inv} + \mathbf{v}^{val} \quad \text{com} \quad (\mathbf{v}^{inv})^T \cdot \mathbf{v}^{val} = \mathbf{0} \quad (\text{B.1})$$

O vetor  $\mathbf{s}$  está no subespaço-inválido, se

$$(\mathbf{v}^{val})^T \cdot \mathbf{s} = \mathbf{0}$$

Substituindo a decomposição de  $\mathbf{v}$  (B.1) em (4.8), e rescrevendo  $\mathbf{T}^{conf}$  como em (4.10) e  $\mathbf{i}^{conf}$  como em (4.11), vem:

$$\mathbf{E}^{conf} = -\frac{1}{2} c_0 (\mathbf{v}^{inv} + \mathbf{v}^{val})^T \cdot (\mathbf{T}^{val} - \mathbf{I}) \cdot (\mathbf{v}^{inv} + \mathbf{v}^{val}) - c_0 \cdot \mathbf{s}^T (\mathbf{v}^{inv} + \mathbf{v}^{val}) \quad (\text{B.2})$$

Como  $\mathbf{T}^{val}$  é uma matriz projeção ( $\mathbf{T}^{val} \cdot \mathbf{T}^{val} = \mathbf{T}^{val}$  e  $\mathbf{v}^{val} = \mathbf{T}^{val} \cdot \mathbf{v}$ ), portanto:

$$\begin{aligned} \mathbf{T}^{val} \cdot \mathbf{v}^{val} &= \mathbf{T}^{val} \cdot \mathbf{T}^{val} \cdot \mathbf{v} = \mathbf{v}^{val} \\ \text{e} \quad (\mathbf{v}^{val})^T \cdot \mathbf{T}^{val} &= \mathbf{v}^T \cdot \mathbf{T}^{val} \cdot \mathbf{T}^{val} = (\mathbf{v}^{val})^T \end{aligned}$$

também, tem-se que:

$$\begin{aligned} (\mathbf{v}^{inv})^T \cdot \mathbf{T}^{val} &= \mathbf{0}^T \\ \text{e} \quad \mathbf{T}^{val} \cdot \mathbf{v}^{inv} &= \mathbf{0} \end{aligned}$$

Portanto, a equação (B.2) torna-se:

$$\begin{aligned}
E^{conf} &= -\frac{1}{2}c_0 \left[ (\mathbf{v}^{val})^T \cdot \mathbf{v}^{val} - (\mathbf{v}^{val})^T \cdot \mathbf{v}^{val} - (\mathbf{v}^{inv})^T \cdot \mathbf{v}^{inv} \right] - c_0 (\mathbf{s}^T \cdot \mathbf{v}^{inv}) \\
&= \frac{1}{2}c_0 \left[ (\mathbf{v}^{inv})^T \cdot \mathbf{v}^{inv} - 2(\mathbf{s}^T \cdot \mathbf{v}^{inv}) \right] \\
&= \frac{1}{2}c_0 \left| \mathbf{v}^{inv} - \mathbf{s} \right|^2 - \frac{1}{2}c_0 \left| \mathbf{s} \right|^2
\end{aligned}$$

Desta forma, se  $\mathbf{s}$  é constante,  $E^{conf}$  é minimizado quando

$$\mathbf{v}^{inv} = \mathbf{s}$$

Substituindo  $\mathbf{v}^{inv}$  na equação (B.1), tem-se:

$$\mathbf{v} = \mathbf{s} + \mathbf{v}^{val} = \mathbf{s} + \mathbf{T}^{val} \cdot \mathbf{v}$$

o que mostra que  $\mathbf{v}$  satisfaz a equação do subespaço-válido.

## APÊNDICE C

Prova de equivalência entre o resultado da aplicação sucessiva das operações ((1)) e ((2)) da rede modificada e a multiplicação de  $\Delta v$  por  $\hat{T}$ . Adota-se por conveniência que a função de ativação da rede é limitada entre 0 e 1.

Dado  $A^0$  denotar o conjunto de elementos de  $v$  para o qual a restrição de desigualdade  $v_q \geq 0$  está ativa, isto é,  $v_q = 0$  e  $\Delta v_q < 0$ . O conjunto  $A^0$  é definido formalmente por,

$$A^0 = \{ q: v_q = 0 \text{ e } \Delta v_q < 0 \} \quad q \in \{1, \dots, N\} \quad (\text{C.1})$$

Dado  $A^1$  denotar o conjunto de elementos de  $v$  para o qual a restrição de desigualdade  $v_q \leq 1$  está ativa, isto é,  $v_q = 1$  e  $\Delta v_q > 0$ . O conjunto  $A^1$  é definido formalmente por,

$$A^1 = \{ q: v_q = 1 \text{ e } \Delta v_q > 0 \} \quad q \in \{1, \dots, N\} \quad (\text{C.2})$$

Dado  $A$  denotar o conjunto formado pela união de  $A^0$  e  $A^1$ , isto é,

$$A = A^0 \cup A^1 \quad (\text{C.3})$$

Então, pode ser visto que se,

$$\Delta \hat{v}_q = 0, \quad \forall q \in A \quad (\text{C.4})$$

e  $|\Delta \hat{v}|$  é pequeno suficiente para garantir que,

$$0 \leq v_p + \Delta \hat{v}_p \leq 1 \quad \text{para todo } p \notin A \quad (\text{C.5})$$

então  $\hat{\Delta v}$  satisfaz a condição (c1) da seção 4.3.1.

Dado  $T^{ina}$  ser uma matriz  $N \times N$  onde,

$$[T^{ina}]_{pq} = \begin{cases} 0 & \text{se } p \neq q \\ 0 & \text{se } p = q \text{ e } p \in A \\ 1 & \text{se } p = q \text{ e } p \notin A \end{cases} \quad (\text{C.6})$$

Então pode ser visto que  $T^{ina}$  é uma matriz-projeção desde que,

$$T^{ina} \cdot T^{ina} = T^{ina}$$

Além disso, se  $\hat{\Delta v}$  satisfaz (C.4), então

$$\hat{\Delta v} = T^{ina} \cdot \hat{\Delta v} \quad (\text{C.7})$$

Similarmente, se  $\hat{\Delta v}$  satisfaz a condição (c2) na sub-seção (4.3.1) (isto é,  $\hat{\Delta v}$  está no subespaço-válido) então,

$$\hat{\Delta v} = T^{val} \cdot \hat{\Delta v} \quad (\text{C.8})$$

Dado o subespaço estendido através de todos vetores  $u$  que satisfaz um  $u = T^{ina} \cdot u$  ser denominado o subespaço-inativo, então pode ser visto que  $\hat{T}$ , tal que  $\hat{\Delta v} = \hat{T} \cdot \Delta v$ , é simplesmente a matriz que projeta  $v$  dentro do subespaço que é a interseção do subespaço-válido e do subespaço-inativo.

Portanto  $\hat{T}$  satisfaz,

$$\begin{aligned} \hat{T} &= T^{val} \cdot \hat{T} \\ \hat{T} &= T^{ina} \cdot \hat{T} \end{aligned} \quad (\text{C.9})$$

Assumindo agora que  $\Delta v^{(0)} = \Delta v$ , dado que,  $\Delta v^{(n)} = \hat{\Delta v} + u^{(n)}$ , onde:

$$\begin{aligned}
\Delta v^{(n)} &= \begin{cases} T^{val} \cdot \Delta v^{(n-1)} & \text{se } n \text{ é ímpar} \\ T^{ina} \cdot \Delta v^{(n-1)} & \text{se } n \text{ é par} \end{cases} \\
&= \begin{cases} \hat{\Delta v} + T^{val} \cdot u^{(n-1)} & \text{se } n \text{ é ímpar} \\ \hat{\Delta v} + T^{ina} \cdot u^{(n-1)} & \text{se } n \text{ é par} \end{cases} \\
\Rightarrow u^{(n)} &= \begin{cases} T^{val} \cdot u^{(n-1)} & \text{se } n \text{ é ímpar} \\ T^{ina} \cdot u^{(n-1)} & \text{se } n \text{ é par} \end{cases} \tag{C.10}
\end{aligned}$$

Desde que  $T^{val}$  e  $T^{ina}$  são matrizes-projeção, então:

$$\begin{aligned}
|T^{val} \cdot u^{(n-1)}| &\leq |u^{(n-1)}| \\
\text{e } |T^{ina} \cdot u^{(n-1)}| &\leq |u^{(n-1)}|
\end{aligned}$$

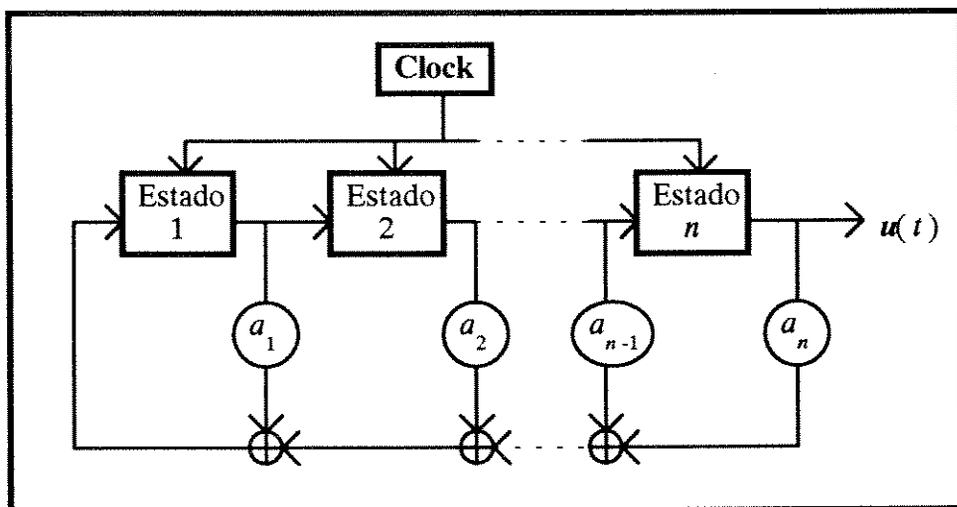
Portanto,

$$|u^{(n)}| \leq |u^{(n-1)}| \tag{C.11}$$

Multiplicar por  $T^{val}$  corresponde a operação ((1)) enquanto multiplicar por  $T^{ina}$  corresponde a operação ((2)). Da equação (C.11) tem-se que quando  $n \rightarrow \infty$ ,  $\Delta v^{(n)} \rightarrow \hat{\Delta v}$  em (C.10). Portanto, a aplicação das operações ((1)) e ((2)) é equivalente à projeção executada sobre  $\Delta v$  pela multiplicação por  $\hat{T}$ .

## APÊNDICE D

Uma seqüência binária pseudoaleatória ou PRBS (Pseudorandom Binary Sequence) pode ser gerada com a utilização de um registrador de deslocamento de ordem  $n$  como ilustrado na *figura D.1*. As variáveis de estado do registrador e os coeficientes de retroalimentação  $a_1, a_2, \dots, a_n$  assumem valores 0 ou 1. O registrador pode ser inicializado com qualquer vetor de estado que seja não nulo. O símbolo  $\oplus$  denota adição em complemento de dois e está definido na *tabela D.1*.



*Figura D.1* - PRBS Através de Registrador de Deslocamento

O espaço de estado da *figura D.1* pode ser representado por:

$$x(t+1) = \begin{pmatrix} a_1 & a_2 & \dots & a_{n-1} & a_n \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \cdot x(t)$$

$$u(t) = (0\ 0\ 0\ \dots\ 1) \cdot x(t)$$

onde todas as adições devem ser realizadas utilizando complemento de dois (*Tabela D.1*). Uma PRBS é um sinal com características puramente determinísticas, ou seja: dado um estado inicial do registrador, os estados futuros podem ser calculados exatamente.

$u_1$	$u_2$	$u_1 \oplus u_2$
0	0	0
1	0	1
0	1	1
1	1	0

*Tabela D.1* - Adição em Complemento de Dois

O período máximo ( $p_{max}$ ) da PRBS é determinada pela seguinte equação:

$$p_{max} = 2^n - 1$$

onde  $n$  é a ordem (número de estágios) do registrador. Este período máximo é obtido fazendo-se uma escolha adequada de cada valor binário do coeficientes de retroalimentação. A *tabela D.2* fornece os coeficientes que devem ter valores unitário para que o período da PRBS seja máximo, os demais coeficientes devem ter valores nulos.

$n$	Coefficientes Com Valor Unitário
3	$a_1$ e $a_3$
4	$a_1$ e $a_4$
5	$a_2$ e $a_5$
6	$a_1$ e $a_6$
7	$a_1$ e $a_7$
8	$a_1, a_2, a_7$ e $a_8$
9	$a_4$ e $a_9$
10	$a_7$ e $a_{10}$

*Tabela D.2* - Valores dos Coeficientes de Retroalimentação da PRBS

