



Edgard Ubaldo Guillén Salas

**Aplicações de Técnicas Baseadas no SVD à Análise e
Previsão de Dados**

Tese de Doutorado

Tese apresentada como requisito parcial para
obtenção do título de Doutor pelo Programa de Pós-
Graduação em Engenharia Elétrica da PUC-Rio.

Orientador: Jacques Szczupack

Rio de Janeiro, setembro de 2004



Edgard Ubaldo Guillén Salas

Aplicações de Técnicas Baseadas no SVD à Análise e Previsão de Dados

Tese apresentada como requisito parcial para obtenção do título de Doutor pelo Programa de Pós-Graduação em Engenharia Elétrica da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Jacques Szczupack

Orientador
PUC-Rio

Prof. Jacques Szczupak

Departamento de Engenharia Elétrica PUC-Rio

Prof. Marco Aurélio Cavalcanti Pacheco

Departamento de Engenharia Elétrica PUC-Rio

Prof. Carlos Eduardo Pedreira

Departamento de Engenharia Elétrica PUC-Rio

Prof. Luiz Pereira Calôba

UFRJ

Prof. Alexandre Pinto Alves da Silva

UFRJ

Prof. Bruno Cosenza de Carvalho

IPD

Prof. José Eugenio Leal

Coordenador(a) Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 9 de setembro de 2004

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Edgard Ubaldo Guillén Salas

Graduou-se em Engenharia Eletrônica na Universidade Nacional de San Agustín de Arequipa (Peru) em 1993. Obteve o título de Mestre em Engenharia Elétrica pela Pontifícia Universidade Católica do Rio de Janeiro em 2000, tendo como área de concentração: Processamento de Sinais e Controle e como linha de pesquisa: Teoria de Controle

Ficha Catalográfica

Guillén Salas, Edgard Ubaldo

Aplicações de técnicas baseadas no SVD à análise e previsão de dados / Edgard Ubaldo Guillén Salas ; orientador: Jacques Szczupak. – Rio de Janeiro : PUC, Departamento de Engenharia Elétrica, 2004.

160 f. : il. ; 30 cm

Tese (doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica.

Incluí referências bibliográficas.

1. Engenharia elétrica – Teses. 2. Algoritmo de cooperação. 3. Redes neurais. 4. Modelos neuro-fuzzy. 5. Decomposição espectral. I. Szczupak, Jacques. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

Para meus pais, Juan e Alejandrina, pelo apoio e confiança

Agradecimentos

Ao meu orientador Professor Jacques Szczupak pelo estímulo e parceria para a realização deste trabalho.

Ao CNPq pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Ao meus amigos, Jhony, Juan Otalora, Larissa, Sibebe, Adriana, Flavia, Diana e Carol por todo apoio, paciência e compreensão.

Aos meus pais, pela educação, atenção e carinho de todas as horas.

Aos meus colegas da PUC-Rio.

Aos professores que participaram da Comissão examinadora.

A todos os professores e funcionários do Departamento pelos ensinamentos e pela ajuda.

A todos os amigos e familiares que de uma forma ou de outra me estimularam ou me ajudaram.

Resumo

Guillén Salas, Edgard Ubaldo. **Aplicações de Técnicas Baseadas no SVD à Análise e Previsão de Dados**. Rio de Janeiro, 2004. 160p. Tese de Doutorado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

O objetivo do presente trabalho é desenvolver uma técnica para a modelagem de sistemas, capaz de se adaptar a uma larga classe de problemas. Como aspecto inovador esta a forma como é orientada a modelagem do sinal, feita segundo a análise dos espaços dos sinais de entrada e saída, destes análises são feitas partições iterativamente em tais espaços até atingir o erro de modelagem desejado. A técnica proposta aqui foi desenvolvida usando redes neurais RBF e modelos Neuro-fuzzy, ajudando-se mutuamente com o objetivo de gerar uma estimativa mais próxima do ideal, esta ajuda mutua é feita pela combinação lineal dos autovetores e autovalores, de forma tal, a gerar novos autovetores e autovalores mais próximos dos ideais. Um objetivo extra, associado ao processo de identificação de sistemas, é a incorporação de facilidades de identificação nas relações entrada-saída por meio de técnicas de decomposição espectral. Desta forma, por um lado pretende-se reduzir o tempo de treinamento e análise para a identificação, eliminando testes a priori julgados desnecessários. Por outro lado, esta técnica sinalaria caminhos para soluções mais viáveis ao processo.

Palavras-chave

Algoritmo de Cooperação; Redes Neurais; Modelos Neuro-fuzzy; Decomposição espectral.

Abstract

Guillén Salas, Edgard Ubaldo. **Applications of Techniques Based on the SVD to the Analysis and Forecast of Data**. Rio de Janeiro, 2004. 160p. Thesis of Doutorado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

The objective of the present work is to develop one technique for the modeling of systems, capable of if adapting to a wide classroom of problems. As innovative aspect this the form as the modeling of the signal, made is guided according to analysis of the spaces of the entrance signals and exit, of these you analyze are made partitions iteratively in such spaces until reaching the desired error of modeling. The technique proposal was developed here using neural nets RBF and Neuro-fuzzy models, helping itself with the objective to generate a estimate next to the ideal, this aid lends is made by the lineal combination of the autovetores and autovalores, form such, to generate new autovetores and autovalores next to the ideals. An extra objective, associated to the process of identification of systems, is the incorporation of easinesses of identification in the relations enter-exit by means of techniques of spectral decomposition. Of this form, on the other hand it is intended to reduce the time of training and analysis for the identification, being eliminated tests a priori unnecessary judgeships. On the other hand, this technique would sinalaria ways for more viable solutions to the process.

Keywords

Algorithm of Cooperation; Neural Nets; Neuro-fuzzy Models; Spectral decomposition.

Sumário

1	Introdução	15
1.1.	Introdução	15
1.2.	Relevância	16
1.3.	Objetivos e Metodologia	16
1.4.	Conteúdo dos capítulos	17
2	Técnicas convencionais usadas na modelagem de sistemas não lineares	19
2.1.	Revisão Bibliográfica	19
2.1.1.	Mistura Adaptativa de Sistemas Especialistas	21
2.2.	Descrição das Técnicas Básicas Utilizadas	28
2.2.1.	Rede RBF em Identificação de Sistemas	28
2.2.1.1.	Utilização Proposta	29
2.2.2.	Lógica Fuzzy	30
3	Concepção Técnica	32
3.1.	Análise: Descrição do Sistema	32
3.1.1.	Técnica de Decomposição do Sinal: SVD	34
3.1.1.1.	SVD Modificado	35
3.1.2.	Partições	38
3.1.3.	Critério das Partições	38
3.1.3.1.	Obtenção dos autovetores e autovalores dos sinais de entrada e saída	38
3.1.3.2.	Re-ordenamento dos vetores característicos de cada atualização da matriz de dados	44
3.1.3.3.	Trajetória descrita pelos vetores característicos	48
3.1.3.4.	Realização das partições	49
3.1.3.5.	Mapeamento das partições dos sinais de entrada e saída	59
3.1.4.	Armazenamento e identificação dos vetores característicos representativos de cada partição	62

3.1.5. Algoritmo de Cooperação entre técnicas de redes	67
3.1.5.1. Modelagem das partições	67
3.1.5.2. Atribuição de pesos	69
3.1.5.3. Aplicação do processo inverso do EIG	74
3.1.5.3.1. Caso ilustrativo	76
3.1.5.4. Conferindo se o erro de modelagem foi satisfeito	92
3.1.5.5. Estrutura Piramidal	92
3.2. Síntese: Descrição do Sistema	96
3.2.1. Identificação da Partição	97
4 Testes e Comparações	109
5 Conclusões	144
5.1. Observações e Conclusões	144
5.1.1. Modificação da técnica	150
5.2. Sugestões	152
6 Referências Bibliográficas	153
7 Apêndice: Cooperação das técnicas nas modelagens das partições para o caso do exemplo 1	159

Lista de figuras

Figura 1	Sistema de redes especialistas e rede seletora: Cada especialista e a rede seletora é uma rede “feedforward”	24
Figura 2	Diagrama em blocos do algoritmo	32
Figura 3	Sinal	39
Figura 4	Arranjo de Vetores Característicos para cada nova amostra de dados	44
Figura 5	Autovetores e autovalores para cada atualização da matriz de dados em duas amostras novas	46
Figura 6	Ordenação dos autovetores e autovalores de cada atualização da matriz de dados segundo os autovetores mais parecidos.	47
Figura 7	Trajétória das distâncias dos autovetores para 6 amostras novas considerando só três dimensões	48
Figura 8	Trajétória descrita por um vetor característico no espaço de “n”dimensões	51
Figura 9	Gráfico “Gradientes por distâncias”	51
Figura 10	Curva “F” (Distâncias)	52
Figura 11	Curva “F” correspondente ao vetor característico da posição “j”	53
Figura 12	Sinal de Entrada	54
Figura 13	Gráfico “Gradientes x distâncias” do vetor característico 1 do sinal de entrada	54
Figura 14	Gráfico “Gradientes x distâncias” do vetor característico 2 do sinal de entrada	55
Figura 15	Gráfico “Gradientes x distâncias” do vetor característico 3 do sinal de entrada	55
Figura 16	Sinal de Saída	56
Figura 17	Gráfico “Gradientes x distâncias” do vetor característico 1 do sinal de saída	56
Figura 18	Gráfico “Gradientes x distâncias” do vetor característico 2 do	

sinal de saída	57
Figura 19 Gráfico “Gradientes x distâncias” do vetor característico 3 do sinal de saída	57
Gráficos: “Distâncias” entre três vetores característicos do sinal de saída	51
Figura 20 Distância “Default” assumida para realizar as partições.	58
Figura 21 Partições do sinal de Entrada e Saída.	59
Figura 22 Mapeamento das partições de entrada e saída	60
Figura 23 Superposição de partições do sinal de Entrada e Saída	61
Figura 24 Vetor Característico amostra 2xM-1	63
Figura 25 Vetor Característico, amostra 2xM	64
Figura 26 Vetor representativo da partição	65
Figura 27 Matrizes Vetores representativos	65
Figura 28 Matrizes de vetores representativos (Partições de um só elemento)	66
Figura 29 Comparação de vetores característicos Reais e Estimados.	68
Figura 30 Sinal de Entrada	76
Figura 31 Sinal de Saída	77
Figura 32 Saída estimada pela rede RBF	78
Figura 33 Saída estimada pelo Neuro-fuzzy	78
Figura 34 Comparação saídas estimadas	79
Figura 35 Erros quadráticos dos autovetores da saída estimada RBF com respeito aos autovetores reais	82
Figura 36 Erros quadráticos dos autovetores da saída estimada Neuro- fuzzy com respeito aos autovetores reais	82
Figura 37 Comparação autovalores	83
Figura 38 Novas saídas estimadas das recombinações	90
Figura 39 Comparação saídas estimadas	91
Figura 40 Esquema Piramidal	93
Figura 41 Processo iterativo de partições (estrutura piramidal)	94
Figura 42 Partições e sub-partições	95
Figura 43 Etapas análise e síntese	98
Figura 44 Preenchimento da matriz de dados com as “M-1” últimas amostras do	

sinal de entrada	99
Figura 45 Vetor característico da matriz de dados das “M-1” últimas amostras do sinal de entrada	99
Figura 46 Partições identificadas nos dados de treinamento do sinal de entrada	100
Figura 47 Respectivas amostras no sinal de saída das amostras identificadas no sinal de entrada	101
Figura 48 Processo recursivo para identificar a partição no sinal de saída	102
Figura 49 Identificação de modelos	104
Figura 50 Recombinação de vetores característicos	105
Figura 51 Amostra estimada	106
Figura 52 “M-1” últimas amostras	106
Figura 53 Amostras interpoladas	107
Figura 54 Entrada e Saída sistema $y = x^3 + 10x$	110
Figura 55 Partições e sub-partições Sistema $y = x^3 + 10x$	111
Figura 56 Dados estimados pela técnica para o Sistema $y = x^3 + 10x$	112
Figura 57 Saída estimada Rede RBF	113
Figura 58 Saída estimada Neuro-fuzzy	113
Figura 59 Comparação de erros devido às várias técnicas empregadas	114
Figura 60 Sinais de entrada e saída teste 2	115
Figura 61 Partições	116
Figura 62 Amostras estimadas pela técnica	117
Figura 63 Amostras estimadas rede RBF	118
Figura 64 Amostras estimadas Neuro-fuzzy	118
Figura 65 Comparação erros	119
Figura 66 Sinais de entrada e saída (sistema retificador com memória)	120
Figura 67 Partições Sistema Retificador com memória	121
Figura 68 Saída estimada Sistema retificador com memória	122
Figura 69 Saída estimada pela rede RBF	123
Figura 70 Saída estimada pela rede RBF (ajustando os parâmetros)	124

Figura 71 Saída estimada pelo modelo Neuro-fuzzy	125
Figura 72 Comparação erros	125
Figura 73 Entrada e saída com ruído	127
Figura 74 Saída estimada (entrada com ruído)	128
Figura 75 Entrada e Saída	129
Figura 76 Partições	130
Figura 77 Amostras estimadas pela nova técnica	131
Figura 78 Amostras estimadas Rede RBF	132
Figura 79 Amostras estimadas Neuro-fuzzy	132
Figura 80 Comparação erros	133
Figura 81 Partições e sub-partições	134
Figura 82 Saída estimada (rede com função linear)	134
Figura 83 Partições e sub-partições	135
Figura 84 Saída estimada (rede com função linear)	136
Figura 85 Partições e sub-partições	137
Figura 86 Saída estimada (rede com função linear)	138
Figura 87 Vazão Normalizada de Itaipu, SST 1	140
Figura 88 Vazão Normalizada de Itaipu, SST 2	140
Figura 89 Vazões de Itaipu com o Trecho Sintetizado	142
Figura 90 Detalhe das Vazões de Itaipu e Sintetizadas	143
Figura 91 Erros produzidos nas sub-partições pelos resíduos	146
Figura 92 Identificação da partição no mapeamento	147
Figura 93 Curva de comportamento da partição identificada	148
Figura 94 Erro produzido pelos comportamentos de descontinuidades abruptas dos resíduos	149
Figura 95 Modificação da técnica	151

Lista de tabelas

Tabela 1 Erros quadráticos médios RBF e Neuro-fuzzy	79
Tabela 2 Erros quadráticos médios dos autovetores RBF e Neuro-fuzzy	81
Tabela 3 Erros quadráticos médios RBF e Neuro-fuzzy	83
Tabela 4 Erros quadráticos médios RBF, Neuro-fuzzy, combinação	90
Tabela 5 Erros quadráticos médios RBF, Neuro-fuzzy, combinação	91
Tabela 6 Tabela Código usado pelo programa para identificar os modelos das partições	103
Tabela 7 Erros quadráticos médios (exemplo 1)	114
Tabela 8 Erros quadráticos médios (exemplo 2)	119
Tabela 9 Erros quadráticos médios (exemplo 3)	126
Tabela 10 Erros quadráticos médios (exemplo 4)	133
Tabela 11 Erros quadráticos médios (exemplo 5)	135
Tabela 12 Erros quadráticos médios (exemplo 6)	137
Tabela 13 Erros quadráticos médios (exemplo 7)	139