

Laboratório de Modelagem, Análise e Controle de Sistemas Não-Lineares

Departamento de Engenharia Eletrônica

Universidade Federal de Minas Gerais

Av. Antônio Carlos 6627, 31270-901 Belo Horizonte, MG Brasil

Fone: +55 3499-4866 - Fax: +55 3499-4850

aguirre@cpdee.ufmg.br



Identificação Multiobjetivo de Sistemas Não-Lineares

Erivelton Geraldo Nepomuceno

Dissertação submetida à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, como parte dos requisitos necessários à obtenção do grau de Mestre em Engenharia Elétrica.

Orientador: Dr. Ricardo H. C. Takahashi

Co-orientador: Dr. Luis A. Aguirre

Belo Horizonte, 19 de Junho de 2002

Dedicatória

*Aos meus pais, pelo amor
Ao meu irmão, pela fraterna amizade
A minha noiva, por considerar-me mais importante
de que seus próprios sonhos*

Agradecimentos

Agradeço a Deus!

Agradeço a você que está lendo essa dissertação!

Agradeço aos colegas do CPDEE e CPH.

Agradeço a CAPES por financiar esse trabalho.

Agradeço aos colegas do LABSIS (FUNREI), onde dei meus primeiros passos na ciência.

Agradeço àqueles que me orientaram: Mário Neto, Élice Melo, Oriane M. Neto, Eduardo Mendes, João Antônio, Michael Lewis, Luís Aguirre e Ricardo Takahashi.

Agradeço àqueles que me ensinaram: Carlos, Maria do Carmo, Mário D'Ávila, Lane, Rodney, Silvino, Hani.

Agradeço àqueles que estudaram comigo: Edson Wander, Rodrigo J. R. Campos, Francisco Moreira, Marcelo Eduardo, Fábio Domingues, Elizabeth Silva.

Agradeço àqueles que foram companheiros: Gleison, Dair, Mara, Márcio, Kleber, Cristiano, Gutemberg, Rosilene, Rô (ela sabem quem é), Adriano, Marcelo Costa.

Agradeço àqueles com quem filosofei: Elton, Janilton, Gleison, Cristiano, Flávio Maia.

Agradeço aos amigos que fiz em Budapeste: David Alonso, Judith Creso, Marc Riemann, Tim Huerta, Alesh, Monika, Erik Schultes.

Agradeço àqueles com quem convivi: Aurélio, Alisson, Anderson, Cássio, William, Edmundo, Lázaro, Evando, Denilson, Luis.

Agradeço àqueles com quem aprendi: Jesus, Ghandi, Buda, Sartre, Nietzsche, Rubem Alves, Kant, Madre Tereza.

Agradeço a quem me mostrou o que é errado: FHC, Bush, Hitler, Collor, Maníaco do Parque, Juiz Nicolau, ACM.

Agradeço àqueles que me orientaram espiritualmente: Eduardo, Flávio, Emerson.

Agradeço àqueles que esqueci, e a eles peço o perdão.

Agradeço àqueles que me suportaram: Neuza Lombardi, Rogéria Cardoso, Tio Reis.

Agradeço àquela que acreditou: minha mãe (e também sofreu por acreditar...).

Agradeço ao meu pai: pelos cafés matinais, pelo amor discreto.

Agradeço ao “Manovisk”, meu querido irmão: Erismar. Permissão!!!

Agradeço a Maristela por ser Maristela!

Agradeço por agradecer...

Epígrafe

“Mineiro escorrega para cima.”

João Guimarães Rosa

“E que todo dia em que se não haja dançado, pelo menos uma vez, seja para nós perdido! E toda a verdade que não traga ao menos um riso nos pareça verdade falsa.”

Friedrich Nietzsche

“Buscai primeiro o reino de Deus, e a sua justiça, e todas estas coisas vos serão acrescentadas”

Mateus 6,33

Conteúdo

Resumo	xv
Abstract	xvii
Lista de Tabelas	xix
Lista de Figuras	xxiii
Lista de Símbolos	xxv
Lista de Abreviações	xxix
1 Introdução	1
1.1 Motivação	3
1.2 Objetivo	4
1.3 Justificativa epistemológica	5
1.4 Estrutura da dissertação	7
2 Identificação Caixa-Preta	11
2.1 Introdução	11
2.2 Testes dinâmicos e coletas de dados	12
2.2.1 Execução do teste	12
2.2.2 Sinais de excitação	12

2.2.3	Tempo de amostragem	13
2.2.4	Detecção de não-linearidades	14
2.3	Escolha da representação matemática	14
2.3.1	A série de Volterra	15
2.3.2	Modelos de Hammerstein e de Wiener	15
2.3.3	Modelos polinomiais contínuos	16
2.3.4	Funções de base radial	17
2.3.5	Redes neurais artificiais	18
2.3.6	Wavelets	18
2.3.7	NARMAX	19
2.4	Propriedades dos modelos NARMAX	21
2.4.1	Agrupamento de termos	22
2.4.2	Pontos fixos em sistemas autônomos	24
2.4.3	Pontos fixos em sistemas não-autônomos	24
2.4.4	Simetria de pontos fixos	26
2.4.5	Curva estática	27
2.5	Determinação de estrutura	28
2.5.1	Determinação de estrutura utilizando ERR	29
2.5.2	Critério de informação de Akaike	30
2.6	Estimação de parâmetros	31
2.6.1	Mínimos quadrados estendidos	32
2.7	Validação do modelo	34
2.8	Conclusões do capítulo	36
3	Identificação Multiobjetivo - Estado da Arte	37
3.1	Introdução	37
3.2	Motivação	39
3.3	Classificação da identificação	40

3.4	Comparação entre os tipos de identificação	43
3.5	Incorporação de informação auxiliar	44
3.5.1	Seleção de estrutura	44
3.5.2	Estimação de parâmetros	46
3.6	Estudos de casos	49
3.7	Alguns trabalhos relacionados	51
3.7.1	(Johansen, 1996)	51
3.7.2	(Johansen, 2000)	52
3.7.3	(Aguirre et al., 2000)	54
3.8	Conclusões do capítulo	55
4	Metodologia para Identificação Multiobjetivo	57
4.1	Introdução	57
4.2	Formulação do problema	58
4.3	Identificação mono e multiobjetivo	59
4.3.1	Formulação mono-objetivo	60
4.3.2	Formulação multiobjetivo	61
4.4	Formulação computacional	62
4.5	Incorporação de informação auxiliar	65
4.5.1	Pontos fixos	65
4.5.2	Curva estática	66
4.6	O Problema da decisão	67
4.6.1	Equilíbrio entre polarização e variância	69
4.6.2	Norma mínima dos objetivos normalizados	70
4.7	Conclusões do capítulo	71
5	Uso da Informação dos Pontos Fixos	73
5.1	Introdução	73

5.2	Sistemas caóticos	75
5.3	Circuito de Chua	76
5.3.1	Descrição do circuito de Chua	76
5.3.2	Metodologia	80
5.3.3	Caso 1 - sem ruído	82
5.3.4	Caso 2 - com ruído	85
5.3.5	Problema de decisão	89
5.3.6	Discussão	91
5.4	Mapa senoidal	93
5.4.1	Descrição de mapas unidimensionais	93
5.4.2	Metodologia e resultados	93
5.4.3	Problema de decisão	96
5.4.4	Discussão	97
5.5	Conclusões do capítulo	99
6	Uso da Informação da Curva Estática	101
6.1	Introdução	101
6.2	Aquecedor elétrico	102
6.2.1	Descrição do sistema	102
6.2.2	Teste dinâmico	103
6.2.3	Teste estático	105
6.2.4	Resultados	105
6.2.5	Problema de decisão	108
6.2.6	Discussão	109
6.3	Conversor buck	111
6.3.1	Descrição do sistema	111
6.3.2	Teste dinâmico	112
6.3.3	Característica estática	114

6.3.4	Resultados	115
6.3.5	Problema de decisão	115
6.3.6	Discussão	116
6.4	Conclusões do capítulo	117
7	Identificação Multiobjetivo de um TCSC	121
7.1	Introdução	121
7.2	O TCSC	122
7.2.1	Relevância do TCSC	122
7.2.2	Descrição do TCSC	123
7.3	Metodologia	125
7.4	Resultados e discussão	131
7.5	Conclusões do capítulo	134
8	Conclusão e Pesquisas Futuras	135
8.1	Discussões finais	135
8.2	Resumo das contribuições	136
8.2.1	Elucidação de conceitos	137
8.2.2	Formulação multiobjetivo	137
8.2.3	Formulação computacional	138
8.2.4	Informação auxiliar: pontos fixos e curva estática . . .	138
8.2.5	Problema de decisão	139
8.3	Conclusões	140
8.4	Pesquisas futuras	140
	Bibliografia	156
A	Rotinas Computacionais	157
A.1	genterms	157

A.2	mcand	158
A.3	akaike	159
A.4	myhouse	159
A.5	orthreg	160
A.6	simodeld	161
A.7	build_pr	161
A.8	build_no	162
A.9	fpoints	163
A.10	fixedp	163
A.11	matrizs	164
A.12	lyapunov	165
A.13	mapa1	165
A.14	im	166
A.15	decisor	167

Resumo

A Identificação de Sistemas estuda como modelar e analisar sistemas a partir de dados. Em situações práticas é comum coletar um conjunto limitado de dados, corrompido por ruído e de caráter local. Nessas situações, se forem considerados apenas os dados coletados, dificilmente um modelo não-linear adequado será obtido. As primeiras tentativas de resolver esse problema usam uma metodologia mono-objetivo, em que o erro de predição é o objetivo a ser minimizado e a *informação auxiliar* é incorporada na forma de restrição, usando técnicas de otimização mono-objetivo. Tal abordagem, contudo, não coloca em perspectiva a determinação de um conjunto de soluções dentro do qual se verifica o *compromisso* entre os diversos objetivos. O conjunto dessas soluções é chamado *Pareto-ótimo*. Este trabalho emprega técnicas multiobjetivo na identificação de sistemas não-lineares e apresenta uma estrutura sistemática de incorporação de informação auxiliar. Tal procedimento recebeu o nome de *identificação multiobjetivo*. A representação NARMAX (*Nonlinear AutoRegressive Moving Average model with eXogenous input*) foi escolhida por permitir a incorporação de informação auxiliar, particularmente a respeito de pontos fixos e da curva estática. A metodologia proposta foi aplicada na modelagem de dois sistemas caóticos (o circuito de Chua e o mapa senoidal), um aquecedor elétrico, um conversor buck e um compensador de reativos (o TCSC – *Thyristor Controlled Series Capacitor*, ou capacitor série controlado a tiristor). As soluções finais dos diversos problemas, que pertencem ao *Pareto-ótimo* em cada caso, foram escolhidas através da análise de preferência *a priori* e *a posteriori* e de dois decisores: um baseado no equilíbrio de polarização e variância e outro, desenvolvido nesse trabalho, baseado na norma mínima dos objetivos normalizados. Verificou-se que a identificação multiobjetivo de sistemas não-lineares é realizável. A determinação de soluções pertencentes ao *Pareto-ótimo* permite o estudo, nesse conjunto, da maneira como variam propriedades relevantes dos modelos e, com base nessa análise, a escolha dos modelos identificados mais adequados às necessidades específicas de cada situação. A possibilidade dessa análise é a principal vantagem da metodologia multiobjetivo.

Abstract

System Identification studies how to model and analyse systems from data. In practical situations it is common to acquire a limited set of data, corrupted by noise and of local character. In these situations, if only the collected data are considered, a suitable model would hardly be yielded. The first attempts to solve this problem used a mono-objective methodology, where the prediction error was the objective to be minimized and the *auxiliary information* was incorporated in the form of constraint, using mono-objective optimization techniques. This mono-objective approach, however, does not put in perspective the search for a set of solutions in which the trade-off between solutions is verified. This set of solutions is called *Pareto-optimal*. This work applies multiobjective techniques in the identification of nonlinear systems and presents a systematic approach to incorporate auxiliary information. Such procedure is referred to as *multiobjective identification*. NARMAX (Nonlinear AutoRegressive Moving Average model with eXogenous input) models have been chosen because in this context it is possible to use auxiliary information, particularly, about fixed points and static curve. The proposed methodology was applied in the modelling of two chaotic systems (Chua's circuit and the sine-map), an electric heater, a buck converter and a reactance compensator (the TCSC - Thyristor Controlled Series Capacitor). The final solutions, which belong to the Pareto-optimal set in each case, were chosen by means of preference analysis *a priori* and *a posteriori* and two decisors: one based on the bias-variance tradeoff and other, which was developed in this work, based on the minimum norm of the normalized objectives. It was verified that multiobjective identification of nonlinear systems is feasible. The determination of solutions belonging to the *Pareto-optimal* allows the study, in this set, of the variation of the relevant properties of the models and, on the basis of this analysis, the choice of the identified models that better fit to the specific needs in each situation. The possibility of this analysis is the main advantage of the multiobjective methodology.

Lista de Tabelas

2.1	Agrupamento de termos necessários para simetria	27
5.1	Propriedades de modelos do circuito de Chua - sem ruído . . .	83
5.2	Propriedades de modelos do circuito de Chua - com ruído . . .	86
5.3	Comparação de parâmetros de modelos do circuito de Chua. . .	88
5.4	Propriedades estáticas e dinâmicas para o mapa senoidal . . .	95
6.1	RMSE de modelos do aquecedor elétrico	107
6.2	RMSE de modelos do conversor buck	115

Lista de Figuras

1.1	Representação em diagramas de blocos de um sistema.	2
3.1	Esquema ilustrativo dos tipos de identificação	39
3.2	Identificação caixa-branca	42
3.3	Identificação caixa-preta	42
3.4	Identificação com o uso de informação auxiliar	42
4.1	Conjunto pareto hipotético de um problema bi-objetivo	58
4.2	Conjunto pareto hipotético e curva de validação	70
5.1	Circuito de Chua	77
5.2	Característica corrente-tensão para o diodo de Chua	77
5.3	Série temporal do circuito de Chua	79
5.4	Atrator dupla volta do circuito de Chua	79
5.5	Atrator dupla-volta reconstruído - circuito de Chua	81
5.6	Mapa de primeiro retorno - circuito de Chua	81
5.7	Pareto-ótimo de modelos do circuito de Chua - sem ruído	82
5.8	Atratores de modelos do circuito de Chua - sem ruído	84
5.9	Mapa de primeiro retorno do circuito de Chua - sem ruído	84
5.10	Pareto-ótimo de modelos do circuito de Chua - com ruído	85
5.11	Atratores de modelos do circuito de Chua - com ruído	87

5.12	Mapa de primeiro retorno do circuito de Chua - com ruído . . .	87
5.13	Pareto-ótimo e curvas de validação	89
5.14	Pareto-ótimo e curvas de validação - circuito de Chua	90
5.15	Atrator da solução ótima - circuito de Chua	91
5.16	Erro de predição com dados de treinamento e validação	92
5.17	Mapa de primeiro retorno - mapa senoidal	94
5.18	Pareto-ótimo - mapa senoidal	94
5.19	Mapa de primeiro retorno de seis modelos do mapa senoidal . . .	96
5.20	Histograma de dados gerados por modelos do mapa senoidal . . .	97
5.21	Diagrama de bifurcação de modelo do mapa senoidal	98
5.22	Variação do expoente de Lyapunov - mapa senoidal	98
6.1	Diagrama em blocos funcionais do aquecedor elétrico.	102
6.2	Dados de identificação - aquecedor elétrico	104
6.3	Dados de validação - aquecedor elétrico	104
6.4	Dados estáticos - aquecedor elétrico	105
6.5	Pareto-ótimo - aquecedor elétrico	106
6.6	Validação dinâmica - aquecedor elétrico	107
6.7	Validação estática - aquecedor estático	108
6.8	Validação dinâmica do melhor modelo - aquecedor elétrico . . .	109
6.9	Validação estática do melhor modelo - aquecedor elétrico . . .	110
6.10	Conversor buck	111
6.11	Dados de identificação - conversor buck	113
6.12	Dados de validação - conversor buck	113
6.13	Dados estáticos - conversor buck	114
6.14	Pareto-ótimo - conversor buck	116
6.15	Validação dinâmica - conversor buck	117
6.16	Validação estática - conversor buck	118

6.17	Validação dinâmica do modelo s^* - conversor buck	119
6.18	Validação estática do modelo s^* - conversor buck	119
7.1	Diagrama <i>on-line</i> do TCSC	123
7.2	Variação de impedância em um TCSC	124
7.3	Impedância de um TCSC para a frequência fundamental . . .	124
7.4	Relação típica entre impedância do TCSC e corrente de linha .	126
7.5	Sistema de potência com um TCSC modelado no Simulink . .	126
7.6	Simulação do sistema-teste com um TCSC	127
7.7	Simulação do sistema-teste com um TCSC - 6 s a 8 s	128
7.8	Simplificação do sistema com um TCSC	129
7.9	Reatância em regime permanente do TCSC	130
7.10	Validação dinâmica para o TCSC	132
7.11	Validação estática para o TCSC	133

Lista de Símbolos

$\Omega_{y^p u^{m-p}}$	Agrupamento de termos;
α	Ângulo de disparo do tiristor;
τ_m	Atraso da função de autocorrelação;
τ_d	Atraso puro de tempo;
$r_y(\tau)$	Autocorrelação linear;
$r_{y^2}(\tau)$	Autocorrelação não-linear;
$\Sigma_{y^p u^{m-p}}$	Coefficiente de agrupamento e termos;
$\hat{\Theta}^*$	Conjunto de <i>soluções eficientes</i> ;
Φ_{MQ}	Conjunto de coeficientes de agrupamento estimados por MQ;
Σ	Conjunto de coeficientes de agrupamento normalizados;
\mathbf{D}	Conjunto factível;
Φ	Conjunto genérico de coeficientes de agrupamento;
P^*	Conjunto Pareto-ótimo;
\dot{x}	Derivada primeira de x ;
$\det[\cdot]$	Determinante;
z	Domínio da frequência complexo;
s	Domínio da frequência contínuo;
$\xi(k)$	Erro de modelagem;

\mathbb{Z}	Espaço dos números inteiros;
\mathbb{N}	Espaço dos números naturais;
\mathbb{R}	Espaço dos números reais;
\square	Fim de definição, teorema e nota;
\blacksquare	Fim de demonstração;
Δ	Fator de decimação;
$b(\cdot)$	Função polinomial qualquer;
$a(\cdot)$	Função polinomial qualquer;
Υ	Função <i>wavelet</i> ;
J_{CE}	Função custo de erro da curva estática;
J_{PF}	Função custo de erro da localização do ponto-fixa;
J_{MQ}	Função custo de erro quadrático de predição;
J	Função custo genérica;
\hat{J}	Função custo normalizada;
F^ℓ	Função genérica com grau de não-linearidade ℓ ;
$f(\cdot)$	Função matemática;
K	Ganho estático;
$e(k)$	Erro no instante k , pode ou não ser branco;
$-$	Indica valores estáticos;
τ_y	Instante de primeiro mínimo da autocorrelação linear;
τ_{y^2}	Instante de primeiro mínimo da autocorrelação não-linear;
h_j	<i>Kernels</i> da série de Volterra;
S	Mapeamento linear;
S^*	Mapeamento linear estendido;
$\mathbf{0}_{\text{MA}}$	Matriz de zeros (com número de linhas de S e n_ξ colunas);

\mathbf{B}	Matriz de entradas no espaço de estados discreto;
Ψ	Matriz de regressores;
Q	Matriz de regressores estáticos;
Ψ^*	Matriz de regressores estendida;
n_u	Máximo atraso do regressor de entrada;
n_e	Máximo atraso do regressor de ruído;
n_y	Máximo atraso do regressor de saída;
e_V	Média do somatório dos erros quadráticos para validação;
$H(s)$	Modelo dinâmico não-linear;
$\ \cdot \ $	Norma euclidiana;
n_p	Número de termos de processo;
n_ξ	Número de termos de ruído;
$E[\cdot]$	Operador esperança matemática;
μ_b	Parâmetro de bifurcação;
c_i	Parâmetro de ruído em um modelo NARMAX;
θ_{di}	Parâmetro do denominador do modelo NARMAX racional;
θ_i	Parâmetro do modelo NARMAX polinomial;
θ_{nj}	Parâmetro do numerador do modelo NARMAX racional;
g_i	Parâmetros ortogonais;
T_i	Passo de integração;
w	Peso para o problema de ponderação;
P_w	Problema de ponderação;
$u(k-j)$	Regressor de entrada;
$\nu(k-i)$	Regressor de ruído em um modelo NARMAX;
$y(k-i)$	Regressor de saída;

ν	Ruído branco;
$y(k)$	Sinal de trabalho;
x_i	Sinal do sistema
$y^*(k)$	Sinal observado;
p^*	Solução ótima;
$\hat{\theta}^*$	Soluções eficientes;
T_s	Tempo de amostragem;
k	Tempo discreto;
T_b	Tempo mínimo de chaveamento;
$\hat{\cdot}$	Indica o valor estimado;
$\min(\cdot)$	Valor mínimo de um vetor;
$\max(\cdot)$	Valor máximo de um vetor;
$\text{var}[\cdot]$	Variância;
\mathbf{e}	Vetor de erro, pode ou não ser branco;
\mathbf{J}	Vetor de funções objetivo;
\mathbf{g}	Vetor de parâmetros ortogonais;
$\boldsymbol{\theta}$	Vetor de parâmetros a estimar;
$\boldsymbol{\theta}^*$	Vetor de parâmetros estendido;
$\hat{\boldsymbol{\theta}}$	Vetor de parâmetros estimado;
$\hat{\boldsymbol{\theta}}_{\text{MQ}}$	Vetor de parâmetros estimados por MQ;
$\hat{\boldsymbol{\theta}}_{\text{MQE}}$	Vetor de parâmetros estimados por MQE;
$\psi(k-1)$	Vetor de regressores do NARMAX polinomial.

Lista de Abreviações

AC	Corrente alternada;
AIC	Critério de informação de Akaike;
CC	Corrente contínua;
CLS	Mínimos quadrados grampeados;
ERR	Taxa de redução de erro;
FACTS	Sistema flexível de transmissão AC;
ICB	Identificação caixa-branca;
ICC	Identificação caixa-cinza;
ICP	Identificação caixa-preta;
IM	Identificação multiobjetivo;
MIMO	Múltiplas entradas e múltiplas saídas (<i>Multiple-Input and Multiple-Output</i>);
MA	Média móvel;
ML	Maior expoente de Lyapunov;
MQ	Mínimos quadrados convencionais;
MQE	Mínimos quadrados estendidos;
MQR	Mínimos quadrados restritos;

NARMAX	Modelos não-lineares auto-regressivos com média móvel e entrada exógena (<i>nonlinear autoregressive moving average model with exogenous input</i>);
NARMA	Modelos não-lineares auto-regressivos com média móvel (<i>nonlinear autoregressive moving average model</i>);
NARX	Modelos não-lineares auto-regressivos com entradas exógenas (<i>nonlinear autoregressive model with exogenous input</i>);
NAR	Modelos não-lineares auto-regressivos (<i>nonlinear autoregressive model</i>);
p.u.	Valor por unidade;
PCA	Análise em componentes principais;
PRBS	Sinal binário pseudo-aleatório;
PWM	Modulação por largura de pulso;
QP	Programação quadrática;
RMSE	Raíz do erro médio quadrático;
SISO	Uma entrada e uma saída (<i>Single-Input and Single-Output</i>);
SNR	Relação sinal-ruído;
SQEP	Somatório do quadrado dos erros de predição;
SQPF	Somatório do quadrado dos erros da localização dos pontos fixos;
SVD	Decomposição em valores singulares;
TCSC	Capacitor série controlado a tiristor.

Introdução

“No princípio criou Deus os céus e a terra. E a terra era sem forma e vazia; e havia trevas sobre a face do abismo; e o Espírito de Deus se movia sobre a face das águas. E disse Deus: Haja luz; e houve luz. ”

Gênesis, 1,1-3

A Identificação de Sistemas é a área do conhecimento que estuda maneiras de modelar e analisar sistemas a partir de observações, ou seja, de dados (Aguirre, 2000a, p. 39). É uma das atividades mais antigas e relevantes da ciência, como bem expressa um dos mais importantes pesquisadores nesse campo, o cientista sueco Lennart Ljung, em seu livro *System Identification: theory for the user*:

“Inferir modelos a partir de observações e estudar suas propriedades é o que realmente faz a ciência. Os modelos (“hipóteses”, “leis da natureza”, “paradigmas”, etc.) podem apresentar um caráter mais ou menos formal, mas todos eles possuem uma mesma característica básica, que é a tentativa de encontrar algum padrão em observações.”¹

Para melhor compreensão do que se trata a Identificação de Sistemas, os conceitos de *sistema* e *modelo* são apresentados. Segundo (Haykin e Van Veen, 2001) um *sistema* é formalmente definido como uma entidade que manipula um ou mais sinais para realizar uma função, produzindo assim, novos sinais. A interação entre um sistema e seus sinais associados é ilustrada esquematicamente na Figura 1.1.

¹(Ljung, 1987, p. 1)



Figura 1.1: Representação em diagramas de blocos de um sistema.

O termo *modelo* significa o conjunto de hipóteses sobre a estrutura ou o comportamento de um sistema físico pelo qual se procuram explicar ou prever, dentro de uma teoria científica, as propriedades do sistema (Holanda, 1988). Do ponto de vista matemático, um modelo nada mais é do que uma abstração, comumente expressa através de equações, de um sistema real (Seborg et al., 1989). Os modelos podem ser classificados principalmente como:

- *Modelos estáticos*: relacionam variáveis sem quantificar sua dependência temporal, ou seja, a saída no instante t_1 depende apenas da entrada no instante t_1 ;
- *Modelos dinâmicos*: são usados para representar sistemas cujo estado atual depende de seus estados em outros instantes de tempo;
- *Modelos autônomos*: são modelos que não contém sinais de entrada de forma explícita;
- *Modelos não-autônomos*: são modelos que apresentam pelo menos uma entrada;
- *Modelos contínuos*: são representados por equações diferenciais e representam a evolução do sistema continuamente no tempo;
- *Modelos discretos*: representam a evolução do sistema em instantes discretos e são descritos por equações de diferença.
- *Modelos lineares*: são representados por equações que satisfazem o *princípio de superposição*².
- *Modelos não-lineares*: são representados por equações que não satisfazem o *princípio de superposição*;

²Uma equação $y = f(x)$ satisfaz o princípio de superposição, e portanto é linear, se e somente se $y_1 = f(x_1)$, $y_2 = f(x_2)$ e $ay_1 + by_2 = f(ax_1 + bx_2)$, para quaisquer constantes a e b .

Dentre os tipos de modelos listados acima, os modelos não-lineares são de considerável importância e vêm sendo largamente estudados. A motivação para essa especial atenção é discutida a seguir.

1.1 Motivação

Devido ao fato de que a maioria dos sistemas no mundo real ser de natureza não-linear, faz-se necessário o desenvolvimento de identificação de sistemas, processamento de sinais, técnicas de controle e análise de sistemas que sejam aplicáveis a esses sistemas (Billings, 1980; Corrêa, 2001). Significantes estudos incluem, recentemente, redes neurais aplicados a estes problemas (Braga et al., 2000). Embora, redes neurais apresentem excelentes propriedades, não se pode extrair facilmente informações analíticas do modelo encontrado (Corrêa, 2001). Isso não é importante em predição pura e reconhecimento de padrões, mas é uma considerável desvantagem em Engenharia e aplicações científicas, em que é importante entender, analisar e simular os mecanismos que produzem a relação entre causa e efeito das entradas e saídas do sistema (Eykhoff, 1981). Entretanto, esforços têm sido canalizados no sentido de entender melhor como informações do sistema são representadas em redes neurais (Amaral, 2001).

A estrutura de modelo conhecida como NARMAX (*Nonlinear Auto Regressive Moving Average model with exogenous inputs*, ou modelo não-linear auto-regressivo de média móvel com entradas exógenas) (Leontaritis e Billings, 1985a,b; Chen e Billings, 1989) pode levar em conta as preocupações acima citadas. Em particular, modelos polinomiais NARMAX permitem, com relativa facilidade, a obtenção analiticamente de algumas informações sobre a dinâmica do sistema (Aguirre et al., 2000).

Nas últimas duas décadas houve um grande volume de trabalhos a respeito de identificação de sistemas não-lineares (Aguirre, 2000a). E apesar de grandes conquistas, a identificação de sistemas não-lineares possui grandes dificuldades. Um dos desafios desse campo é a escolha de uma estrutura adequada e compacta para o modelo (Lindskog, 1996; Aguirre et al., 2000). Uma solução geral para esse problema está ainda longe de ser encontrada.

No início da década de noventa, alguns trabalhos começaram a utilizar conhecimento *a priori* (Bohlin, 1991; Tulleken, 1993; Eskinat et al., 1993). Aguirre e colegas (2000) sugerem que o uso de informação *a priori* pode ser usado na escolha da estrutura do modelo. O termo *conhecimento a priori*

é largamente usado na literatura para descrever toda informação disponível antes de iniciar o processo de identificação. Entretanto, Eskinat e colegas (1993) cunharam o termo *informação auxiliar* para designar qualquer informação sobre o sistema adicional aos dados de entrada e saída. Tal nomenclatura parece ser mais consistente, pois não é raro, adquirir conhecimento sobre o sistema durante o processo de identificação e incorporar esse conhecimento em etapas subseqüentes.

Inúmeros artigos continuaram a ser produzidos durante a década de noventa (Duran e White, 1994; Sjöberg et al., 1995; Bohlin e Graebe, 1995; Johansen, 1997; Alessandri e Parisini, 1997) e vários outros já foram publicados nos últimos anos (Pearson e Pottmann, 2000; Murakami e Seborg, 2000; Aguirre, 2000b; Barroso, 2001; Nepomuceno et al., 2002b). Apesar disso, poucos trabalhos mostraram alguma metodologia sistematizada para incorporar a informação auxiliar no processo de identificação de sistemas não-lineares. Alguns trabalhos como (Tulleken, 1993; Johansen, 1997; Pearson e Pottmann, 2000) incorporam a informação auxiliar utilizando uma *metodologia mono-objetivo*.

Um dos primeiros trabalhos que abordam o uso de informação auxiliar em uma *metodologia multiobjetivo* é (Johansen, 2000). Tal método, diferentemente da abordagem mono-objetivo, é capaz de encontrar soluções intermediárias que levam em conta o compromisso entre os objetivos, e não apenas soluções individuais, sendo que o conjunto dessas soluções é chamado *Pareto-ótimo*. Porém, o sistema utilizado em (Johansen, 2000) é simulado e os modelos produzidos são lineares. De fato, essa dissertação juntamente com o trabalho de Nepomuceno e colegas (2002a), parece ser a primeira tentativa de identificar sistemas não-lineares usando informação auxiliar através de uma abordagem multiobjetivo.

1.2 Objetivo

Em situações práticas é comum coletar um conjunto limitado de dados, corrompido por ruído e de caráter local. Nessas situações, se forem considerados apenas os dados coletados, dificilmente um modelo não-linear adequado será obtido. As primeiras tentativas de resolver esse problema usaram uma metodologia mono-objetivo, em que o erro de predição é o objetivo a ser minimizado e a *informação auxiliar* é incorporada na forma de restrição, usando técnicas de otimização mono-objetivo. Tal abordagem, contudo, não

coloca em perspectiva a determinação de um conjunto de soluções dentro do qual se verifica o *compromisso* entre os diversos objetivos. O conjunto dessas soluções é chamado *Pareto-ótimo*.

O objetivo deste trabalho é empregar técnicas multiobjetivo na identificação de sistemas não-lineares através de uma estrutura sistemática de incorporação de informação auxiliar. Tal procedimento recebeu o nome de *identificação multiobjetivo (IM)*. A representação NARMAX foi escolhida por permitir a incorporação de informação auxiliar com relativa facilidade, particularmente, de pontos fixos e da curva estática.

1.3 Justificativa epistemológica

A epistemologia³, também chamada teoria do conhecimento, é o ramo da filosofia interessado na investigação da natureza, fontes e validade do conhecimento. Entre as questões principais que ela tenta responder estão as seguintes. O que é o conhecimento? Como o alcançamos? Podemos conseguir meios para defendê-lo contra o desafio cético? Essas questões são, implicitamente, tão velhas quanto a filosofia, embora seu primeiro tratamento explícito seja o encontrado em Platão (427-347 AC), em particular no *Theaetetus*. Mas primordialmente na era moderna, a partir do século XVII, com o resultado do trabalho de Descartes (1596-1650) e Locke (1632-1704) e em associação com a emergência da ciência moderna é que a epistemologia tem ocupado um plano central na filosofia (Grayling, 1996).

Um passo óbvio na direção de responder à primeira questão é tentar uma definição. A definição padrão, preliminarmente, é a de que o conhecimento é *crença verdadeira justificada*. Esta definição parece plausível porque, ao menos, ela dá a impressão de que para conhecer algo, alguém deve acreditar nisso, que a crença deve ser verdadeira, e que a razão de alguém para acreditar deve ser satisfatória à luz de algum critério, pois alguém não poderia dizer conhecer algo se sua razão para acreditar fosse arbitrária ou aleatória. Assim, cada uma das três partes da definição parece expressar uma condição necessária para o conhecimento, e a reivindicação é a de que, tomadas em

³[Do gr. epistéme, 'ciência'; 'conhecimento', + -o- + -logia.] S. f. 1. Conjunto de conhecimentos que têm por objeto o conhecimento científico, visando a explicar os seus condicionamentos (sejam eles técnicos, históricos, ou sociais, sejam lógicos, matemáticos, ou lingüísticos), sistematizar as suas relações, esclarecer os seus vínculos, e avaliar os seus resultados e aplicações.

conjunto, elas são suficientes.

Desde os primórdios da humanidade uma das estratégias mais utilizadas para auxiliar o homem na busca de conhecimento é a criação de modelos. Sejam gráficos, fluxogramas, descrição textual de processos ou representações matemáticas. As representações matemáticas tiveram um grande avanço com o advento do cálculo diferencial. Sistemas dinâmicos passaram, então, a ser modelados. Uma outra etapa importante foi o ajuste de modelos matemáticos a dados observados. Um dos exemplos mais importantes na história da ciência é o trabalho de Johann Kleper (1571-1630). A partir de dados coletados por Tycho Brahe (1546-1601), Kleper formulou um modelo matemático capaz de descrever o movimento dos corpos celestes. Esse modelo, posto à prova, era capaz de prever, com considerável precisão, a posição dos corpos celestes. Era a teoria submetida ao teste da observação. Nesse sentido, um modelo matemático pode ser entendido como um objeto intelectual, no qual o homem deposita a sua fé, uma *crença verdadeira*; e que através de critérios de validação, normalmente associados à observação do fenômeno, tem a sua crença *justificada*.

Entretanto, a observação nem sempre corrobora os intuítos humanos: um cientista pode ser levado a considerar válida ou falsa uma teoria, contrariando suas expectativas *a priori*. A ciência, sob uma perspectiva fenomenológica, procura observar fenômenos de seu interesse, descrevê-los, classificá-los e encontrar ordem. Vez por outra, o fenômeno, ao invés de corroborar uma teoria, revela a sua falsidade. É o que aconteceu com Max Plank (1858-1947) ao formular a teoria do *quanta* de energia. Como atestado em sua autobiografia, tal evidência era surpreendente e, em princípio, deixou-o desesperado:

“... posso qualificar todo o feito como um ato de desespero. Pois por natureza eu sou pacífico e avesso a aventuras. Mas, uma interpretação teórica tinha que ser encontrada a qualquer preço, mesmo que tão alto. As duas primeiras leis da Termodinâmica me parecem a única coisa que em todas as circunstâncias tinha que ser mantida. De resto, estava pronto para qualquer sacrifício em minhas convicções físicas anteriores.”⁴

Embora ainda esteja por ser formulada uma teoria do desenvolvimento científico nas ciências da Engenharia (ciências que investigam a possibilidade

⁴(Martins, 1992, p. 298)

de se construir mecanismos não existentes na natureza), pode-se extrapolar a sistematização filosófica que já foi elaborada para as ciências da natureza (Takahashi, 1998). A Engenharia, então, também se beneficia da modelagem de sistemas. E além de ampliar o conhecimento acerca de um fenômeno, um modelo matemático, permite que a Engenharia desenvolva novos processos, otimize sistemas, crie novos dispositivos não existentes na natureza. É interessante observar a relação entre a teoria que estuda como obter modelos e seu emprego. De certa forma, essas duas linhas de atuação, podem ser enquadradas no campo da ciência pura e da tecnologia, respectivamente. Um equilíbrio entre essas duas linhas é de fundamental importância, como atesta Simon Schwartzman:

“A nova política de C&T deve implementar tarefas aparentemente contraditórias: estimular a liberdade, a iniciativa e a criatividade do pesquisador e, ao mesmo tempo, estabelecer um forte vínculo entre o que eles fazem e as necessidades da economia, do sistema educacional e da sociedade como um todo.”⁵

Dentro do contexto da Engenharia e particularmente dentro da teoria de controle, a modelagem matemática pode ser atingida através de técnicas de *Identificação de Sistemas*. Ljung (1987) afirma que “*Identificação de Sistemas* lida com a construção de modelos matemáticos de sistemas dinâmicos baseado em dados observados dos sistemas físicos.” A sua importância, sob a perspectiva do que acima foi mencionado, é crucial para o desenvolvimento da ciência e da Engenharia.

A utilização de informação auxiliar no processo de identificação é justificável do ponto de vista epistemológico, pois utiliza toda evidência disponível. Tal argumento parece tautológico, porém há correntes de pesquisadores que seguem o princípio enunciado por Fisher: “Deixem os dados falarem por si mesmos”. Apesar de Fisher ter o intuito de tornar imparciais os resultados obtidos ao analisar uma massa de dados, em muitos casos características importantes do sistema não estão presentes nos dados.

1.4 Estrutura da dissertação

Essa dissertação está organizada em oito capítulos da seguinte forma:

⁵(Schwartzman, 2001, p. xxii)

No Capítulo 1 a Identificação de Sistemas é introduzida. Os conceitos de sistema e modelo, bem como algumas classificações de modelos são abordados. É também apresentada a motivação e a justificativa epistemológica para o assunto em questão. O objetivo da dissertação é definido e sua estrutura apresentada.

No Capítulo 2 são abordadas questões relativas à *identificação caixa-preta*. As principais etapas de identificação (testes dinâmicos e coleta de dados; escolha da representação matemática, determinação de estrutura; estimação de parâmetros; validação do modelo) são discutidas. São apresentadas algumas das mais importantes representações matemáticas presentes na literatura, sendo dada particular atenção aos modelos NARMAX.

O Capítulo 3 traz uma revisão bibliográfica de metodologias utilizadas para a incorporação de informação auxiliar na identificação de sistemas. Os conceitos de informação auxiliar, informação (ou conhecimento) *a priori*, identificação caixa-preta, caixa-branca e caixa-cinza são investigados. Também são mostradas algumas vantagens e desvantagens desses tipos de identificação.

No Capítulo 4 é apresentada a metodologia para identificação multiobjetivo de sistemas não-lineares. A formulação de um problema multiobjetivo é mostrada e o conceito de *conjunto Pareto-ótimo* é destacado. É mostrada uma formulação computacional que permite estimar parâmetros usando determinados tipos de informação auxiliar. Também é mostrado como conhecimento a respeito de pontos fixos e da curva estática pode ser incorporado na identificação de modelos NARMAX. Em seguida, o problema da decisão, ou seja, a sistemática para escolher a solução ótima pertencente ao conjunto *Pareto-ótimo* é discutido.

No Capítulo 5 o uso da informação acerca de pontos fixos é empregado para a identificação de dois sistemas caóticos: o circuito de Chua e o Mapa senoidal com linearidades cúbicas. Análise qualitativa e o equilíbrio entre polarização e variância são empregadas para a escolha da solução ótima.

No Capítulo 6 o uso da informação acerca da curva estática é aplicado na identificação de dois sistemas: um aquecedor elétrico e um conversor buck.

No Capítulo 7 é apresentada a identificação multiobjetivo de um TCSC (*Thyristor Controlled Series Capacitor*, ou capacitor série controlado a tirisitor) usando informação auxiliar a respeito da curva estática. A solução ótima é escolhida com base em propriedades qualitativas.

O Capítulo 8 discute sobre as características da identificação multiob-

jetivo de sistemas não-lineares. Conclusões sobre o método proposto são formuladas e as propostas para pesquisas futuras são apresentadas.

Um Apêndice traz as principais rotinas computacionais usadas na dissertação.

Identificação Caixa-Preta

“Let the data speak for themselves.”

Sir Ronald Aylmer Fisher¹

2.1 Introdução

A Identificação de Sistemas envolve, em linhas gerais, as seguintes etapas (Ljung, 1987; Aguirre, 2000a, p. 66):

1. Testes dinâmicos e coleta de dados;
2. Escolha da representação matemática a ser usada;
3. Determinação da estrutura do modelo;
4. Estimação de parâmetros;
5. Validação do modelo.

Essas etapas são comuns tanto para sistemas que apresentam comportamento aproximadamente linear quanto para sistemas não-lineares. Nas situações em que só se utilizam os dados de entrada e saída a identificação é normalmente chamada de *identificação caixa-preta (ICP)*. As seções seguintes comentam brevemente essas etapas. As etapas de determinação de estrutura e estimação de parâmetros utilizaram a representação NARMAX. A Seção 2.4 fornece algumas propriedades dos modelos NARMAX.

¹Deixe os dados falarem por si mesmos (Jaynes, 1996, p. 641).

2.2 Testes dinâmicos e coletas de dados

Uma vez que a identificação se propõe a obter modelos a partir de dados, é necessário coletar tais dados. Em muitos casos, os dados disponíveis são obtidos do sistema em *operação normal*. Em outras situações, entretanto, será possível e desejável efetuar testes de forma a extrair informação dinâmica do sistema. Costuma-se chamar *dados de identificação* o conjunto de dados destinados às etapas de escolha da representação matemática, determinação de estrutura e estimação de parâmetros; como *dados de validação*, o conjunto de dados destinados a validação do modelo. Problemas importantes relacionados a esta etapa são a execução de testes, escolha dos sinais de excitação, a escolha do tempo de amostragem e a detecção de não-linearidades.

2.2.1 Execução do teste

A *execução do teste* envolve aspectos de medição e instrumentação até a escolha das variáveis que serão usadas como entrada e saída. A medição deve procurar minimizar os efeitos de ruído nos sinais coletados. Em muitas situações, a escolha das variáveis é simples de ser feita. Contudo, para sistemas em que as malhas de controle não são óbvias, como sistemas biológicos e econômicos, não é uma tarefa óbvia escolher quais e quantas variáveis devem ser usadas para compor o modelo. Nesses casos, procedimentos como decomposição em valores singulares (SVD) (Chen, 1999, p. 76) ou a análise em componentes principais (PCA) (Hotelling, 1933; Jolliffe, 1986) podem ser adotadas.

2.2.2 Sinais de excitação

É desejável que os *sinais de excitação* do sistema tenham espectros de frequência que lhes permitam excitar persistentemente a dinâmica de interesse do sistema. No caso de sistemas não-lineares, isso requer que os efeitos não-lineares sejam excitados por esses sinais e assim estejam presentes nos dados de identificação (Aguirre, 2000a, p. 362).

Intuitivamente, um sinal persistentemente excitante de ordem n é um sinal que tem potência espectral em n ou mais frequências distintas. Uma condição para estimar n parâmetros é que o sinal de excitação deve ser persistentemente excitante de pelo menos ordem n . Um sinal geralmente usado é o sinal binário pseudo-aleatório (PRBS) do inglês *pseudo random binary*

signal. Uma descrição desse sinal pode ser encontrada em (Aguirre, 2000a, p. 170). Em (Aguirre et al., 2000) o sinal PRBS é utilizado para identificar um conversor buck.

2.2.3 Tempo de amostragem

Grande parte dos sistemas reais são processos contínuos. Em várias aplicações científicas e tecnológicas é necessário, entretanto, amostrar o sinal de interesse. O período entre amostras é chamado de período ou tempo de amostragem, T_s . Para que o sinal amostrado possua as características fundamentais do sinal original, é necessário que o tempo de amostragem seja suficientemente curto. De acordo com o Teorema de Shannon (Shannon, 1949), o sinal deve ser amostrado em pelo menos duas vezes o valor da maior componente de frequência desejada. Nas situações práticas aqui consideradas, costuma-se amostrar em 5 a 10 vezes tal frequência.

Quando se deseja utilizar um sinal para identificação, costuma-se superamostrar o sinal, ou seja, amostrá-lo com T_s bastante reduzido. A questão passa a ser a definição de uma taxa pela qual o sinal observado $y^*(k)$ será decimado de forma a gerar o sinal de trabalho $y(k)$, devidamente amostrado. Um procedimento prático é a análise da autocorrelação linear $r_y(\tau)$ (2.1) e não-linear $r_{y^2}(\tau)$ (2.2) do sinal de trabalho (Aguirre, 1995)

$$r_y(\tau) = E[(y(k) - \overline{y(k)})(y(k - \tau) - \overline{y(k)})], \quad (2.1)$$

$$r_{y^2}(\tau) = E[(y^2(k) - \overline{y^2(k)})(y^2(k - \tau) - \overline{y^2(k)})], \quad (2.2)$$

sendo que $E[\cdot]$ indica a esperança matemática. O sinal $y(k)$ é considerado ergódico, permitindo, então, substituir a esperança matemática pela média temporal. Com base nas funções de autocorrelação descritas acima pode-se chegar à seguinte constante:

$$\tau_m = \min\{\tau_y, \tau_{y^2}\}, \quad (2.3)$$

em que τ_y é o instante do primeiro mínimo de $r_y(\tau)$ e τ_{y^2} é o instante do primeiro mínimo de $r_{y^2}(\tau)$. Finalmente, escolhe-se o fator de decimação, Δ , de forma que as funções (2.1) e (2.2) do sinal decimado $y(k) = y^*(\Delta k)$ satisfaçam

$$10 \leq \tau_m \leq 20, \quad (2.4)$$

sendo que os limites inferior e superior de (2.4) podem ser relaxados para 5 e 25, respectivamente (Aguirre, 2000a, p. 368).

Um exemplo de aplicação dessa técnica pode ser encontrado em (Aguirre, 2000a, p. 371). Uma outra abordagem pode ser encontrada em (Rosenstein et al., 1994) que utiliza conceitos da geometria do sistema em estudo. Recentemente, Tseng (2002) propôs um critério específico para amostragem de sistemas não-lineares, que permite, em alguns casos, amostrar o sinal com frequência abaixo da frequência de Nyquist.

2.2.4 Detecção de não-linearidades

Por fim, os dados de identificação devem passar por testes de *detecção de não-linearidades*. Esses testes verificam, dentro de um limite de confiança pré-determinado, se o sistema possui algumas características próprias dos sistemas lineares. Caso não sejam verificadas essas propriedades, é necessária a utilização de modelos não-lineares para aproximar as características do mesmo.

Em (Haber, 1985) podem-se encontrar alguns algoritmos simples para verificação se há ou não indícios de não-linearidades nos dados. Dentre tais algoritmos, o mais simples consiste em realizar dois testes independentes no sistema a ser modelado operando em condições diferentes. Em seguida, verifica-se se o princípio da superposição é aplicável.

Um outro procedimento consiste em calcular

$$r_{y'y'^2}(\tau) = E[y'(k + \tau)y'^2(k)], \quad (2.5)$$

sendo $y'^2 = (y(k) - E[y(k)])^2$ e o sinal de apóstrofe significa que a média foi removida do sinal. Portanto, $r_{y'y'^2}(\tau) = 0, \forall \tau$ se o sistema for linear (Billings e Voon, 1983).

2.3 Escolha da representação matemática

A escolha da representação não tem sido uma etapa cumprida criteriosamente na identificação de sistemas dinâmicos não-lineares. Tal ausência pode ser em parte justificada pela falta de estudos mais consistentes a respeito das características de cada representação (Corrêa, 2001).

Há uma grande variedade de representações não-lineares que, pelo menos

em princípio, podem ser utilizadas na identificação de sistemas. As seções seguintes apresentam de forma muito breve algumas representações estudadas nos últimos anos.

2.3.1 A série de Volterra

A saída $y(t)$ de um sistema não-linear com entrada $u(t)$ pode ser representada pela chamada série de Volterra definida como

$$y(t) = \sum_{j=1}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_j(\tau_1, \dots, \tau_j) \prod_{i=1}^j u(t - \tau_i) d\tau_i, \quad (2.6)$$

sendo que as funções h_j são denominadas *kernels* e claramente são generalizações não-lineares da resposta ao impulso $h_1(t)$.

Uma dificuldade prática da aplicação da série de Volterra na identificação de sistemas é que, mesmo para sistemas pouco não-lineares, o número de parâmetros a determinar é muito grande. A grande necessidade de parâmetros é consequência da série de Volterra tentar explicar a saída $y(t)$ apenas em função da entrada. Uma forma de reduzir o número de parâmetros é utilizar valores da própria saída, além de valores da entrada, para determinar $y(t)$, ou seja, utilizar recorrência ou auto-regressão da saída. Uma boa revisão a respeito da série de Volterra pode ser encontrada em (Billings, 1980). O uso das séries de Volterra para mapeamento no domínio da frequência usando a técnica explorativa pode ser encontrado em (Bedrosian e Rice, 1971), e, em (Bui et al., 2001) uma recente aplicação na descrição de distorções não-lineares em ressonância magnética.

2.3.2 Modelos de Hammerstein e de Wiener

Os modelos de Hammerstein e de Wiener foram bastante populares até duas ou três décadas atrás. Recentemente o uso desses modelos tem atraído atenção de vários pesquisadores (Wigren, 1993; Greblicki, 1996; Pearson e Pottmann, 2000).

Os modelos de Hammerstein e de Wiener são uma composição de um modelo dinâmico linear² $H(s)$ em cascata com uma função estática não-

²Um modelo dinâmico nas representações de Hammerstein e de Wiener é normalmente descrito no domínio da frequência s . $H(s)$ é a transformada de Laplace da resposta ao impulso $h(t)$.

linear $f(\cdot)$. No caso do modelo de Hammerstein, a não-linearidade estática precede o modelo dinâmico linear, ou seja,

$$U^*(s) = f(U(s)); \text{ e } Y(s) = H(s)U^*(s). \quad (2.7)$$

No caso do modelo de Wiener, o modelo dinâmico linear precede a não-linearidade estática, isto é,

$$Y^*(s) = H(s)U(s); \text{ e } Y(s) = f(Y^*(s)). \quad (2.8)$$

Recentes trabalhos (Coelho et al., 2001; Coelho, 2002) apresentam uma metodologia de obter modelos Wiener e Hammerstein a partir de modelos NARX (*Nonlinear AutoRegressive Model with eXogenous input*, ou modelo não-linear auto-regressivo de média móvel com entradas exógenas), contornando assim, o problema de obtê-los a partir de dados de entrada e saída.

2.3.3 Modelos polinomiais contínuos

Um modelo polinomial contínuo pode ser formado utilizando um sinal e suas derivadas como base. No caso de um polinômio de ordem três, o modelo pode ser representado da seguinte forma (Gouesbet e Letellier, 1994)

$$\begin{aligned} \dot{X} &= \dot{y}(t) \\ \dot{Y} &= \ddot{y}(t) \\ \dot{Z} &= \sum_{l=1}^{n_\theta} \theta_l \psi^l, \end{aligned} \quad (2.9)$$

sendo $\psi^l = X^i Y^j Z^k$ e $i, j, k \in \mathbb{N}$. Nesse caso os regressores (X, Y, Z) são o sinal $y(t)$ e suas derivadas primeira e segunda. A grande dificuldade dessa representação é estimar as derivadas do sinal medido. Trabalhos recentes são (Aguirre et al., 2001; Freitas, 2001), em que os autores aplicaram técnicas de seleção de estruturas para a obtenção de modelos contínuos.

2.3.4 Funções de base radial

As funções de base radial (RBF), do inglês *radial basis functions* (Bromhead e Lowe, 1988) são mapeamentos do tipo

$$f(\mathbf{y}) = w_0 + \sum_i w_i \phi(\|\mathbf{y} - \mathbf{c}_i\|), \quad (2.10)$$

sendo que $\mathbf{y} \in \mathbb{R}^n$, $\|\cdot\|$ é a norma euclidiana, $w_i \in \mathbb{R}$ são pesos, $\mathbf{c}_i \in \mathbb{R}^n$ são os centros e $\phi(\cdot) : \mathbb{R}^+ \rightarrow \mathbb{R}$ é uma função, normalmente escolhida a priori, como por exemplo:

$$\phi(\|\mathbf{y} - \mathbf{c}_i\|) = \exp\left(-\frac{\|\mathbf{y} - \mathbf{c}_i\|}{\sigma_i^2}\right), \quad (2.11)$$

sendo σ_i constante. A função (2.11) é chamada de gaussiana. Várias outras funções são utilizadas, tais como *linear*, *cúbica*, *multiquadrática* (Haykin, 1994; Braga et al., 2000).

Essa representação pode ser interpretada como uma técnica de interpolação global com boas propriedades locais. As funções de base radial são casos particulares de redes neurais, mas são lineares nos parâmetros w_i , ao contrário de redes neurais.

É comum, no contexto de Identificação de Sistemas, acrescentar termos auto-regressivos lineares e termos de entrada, à função (2.10), resultando em

$$\begin{aligned} y(k) = & w_0 + \sum_i w_i \phi(\|\mathbf{y}(k-1) - \mathbf{c}_i\|) + \\ & + \sum_{i=1}^{n_y} n_y a_i y(k-1) + \sum_{i=1}^{n_u} n_u a_i u(k-i) + e(k), \end{aligned} \quad (2.12)$$

sendo $\mathbf{y}(k-1) = [y(k-1) \dots y(k-n_y) u(k-1) \dots u(k-n_u)]^T$ e $e(k)$ o erro. No caso do sistema possuir atraso puro de tempo, τ_d , os termos lineares de entrada começariam com $u(k - \tau_d - 1)$.

2.3.5 Redes neurais artificiais

As redes neurais artificiais são compostas por camadas de “neurônios” interconectados. A saída de um único neurônio com n entradas é do tipo

$$y(\mathbf{x}) = f\left(\sum_{j=1}^n w_j x_j + b\right), \quad (2.13)$$

sendo que b (*bias*)³ e w_j são constantes e f é chamada de função de ativação. Há vários tipos de função de ativação sendo que uma das mais usadas é a sigmóide

$$f(z) = \frac{1}{1 + e^{-z}}. \quad (2.14)$$

Deve ser notado que, como é normalmente o caso nas redes neurais multicamadas, a saída de um neurônio do tipo mostrado na Equação (2.13) é conectada à entrada de um outro neurônio. Assim a saída será uma composição das funções de cada camada. Modelos que levam em conta termos auto-regressores e termos de entrada podem ser obtidos de maneira semelhante aos modelos descritos em (2.10) e (2.12).

A fim de estimar os parâmetros de uma rede neural multicamada, são utilizadas algoritmos de estimação não-linear, tais como o de *retropropagação* do erro (do inglês *backpropagation*).

Há inúmeros trabalhos sobre modelos que usam RBF ou redes neurais. Uma introdução geral a respeito do assunto pode ser encontrada em (Haykin, 1994; Braga et al., 2000).

2.3.6 Wavelets

Wavelets são funções matemáticas que ampliam intervalos de dados, separando-os em diferentes componentes de frequência, permitindo a análise de cada componente em sua escala correspondente. A transformada *wavelet* permite representar uma série temporal ou uma função no tempo. Além disso, essa transformada é um poderoso e genérico método usado para detectar comportamentos singulares em sinais. Em geral, a transformada *wavelet* revela determinadas características que outros métodos não conseguem obter

³Apesar de possuir um termo em português, *polarização*, o termo *bias* é consideravelmente usado na literatura científica brasileira.

(Combastel et al., 2002). Teoricamente, a aplicação da transformada *wavelet* a um sinal é equivalente à aplicação de um filtro.

A idéia fundamental da transformada *wavelet* é representar simultaneamente uma função nos domínios do tempo e da frequência. Para se ter uma breve idéia da transformada *wavelet*, considere a clássica transformada de Fourier

$$F(w) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt, \quad (2.15)$$

que nada mais é do que a soma ao longo do tempo do sinal $f(t)$ multiplicado por $e^{-j\omega t}$. Os resultados dessa transformação são os *coeficientes de Fourier* $F(w)$, que quando multiplicados por senóides de frequência w apropriada, produzem as componentes do sinal original.

Similarmente, a transformada *wavelet* é definida como o somatório ao longo do tempo de um sinal multiplicado por escalonadas e deslocadas versões da função *wavelet* Υ :

$$C(\text{escala}, \text{posição}) = \int_{-\infty}^{\infty} f(t)\Upsilon(\text{escala}, \text{posição}, t) dt. \quad (2.16)$$

O resultado de (2.16) são vários *coeficientes*, que são função da escala e da posição (MathWorks, 1997).

O mundo científico passou a dar atenção a *wavelets* após o trabalho de Daubechies (1988). Em trabalhos recentes a representação *wavelet* tem sido usada também na identificação de sistemas não-lineares (Cao et al., 1995; Billings e Coca, 1999a). Combastel et al. (2002) utilizam modelos baseados em *wavelets* para detectar falhas em motores de indução.

2.3.7 NARMAX

A representação matemática NARMAX (*Nonlinear AutoRegressive Moving Average with eXogenous input*) (Leontaritis e Billings, 1985b; Chen e Billings, 1989) que, traduzindo, seria modelo não-linear auto-regressivo de média móvel com entradas exógenas, vem sendo largamente empregada nas últimas duas décadas.

Modelos NARMAX podem ser descritos pela equação:

$$y(k) = F^\ell \begin{bmatrix} y(k-1), \dots, y(k-n_y), \\ u(k-d), \dots, u(k-d-n_u+1), \\ e(k), \dots, e(k-n_e) \end{bmatrix}, \quad (2.17)$$

sendo $k = 1, \dots, N$. F^ℓ é uma função não-linear qualquer, $y(k)$ e $u(k)$ são, respectivamente, saída e entrada do sistema, que têm seus atrasos representados por n_y , n_u respectivamente. d representa o tempo de retardo do sistema, $e(k)$ representa incertezas e n_e , o atraso de $e(k)$.

O presente trabalho utilizará a representação NARMAX polinomial, ou seja, quando F é um polinômio de grau $\ell \in \mathbb{N}$. Apesar de ser quase impossível definir a melhor representação em termos gerais, podem-se destacar algumas vantagens dos modelos polinomiais sobre outras representações. É possível obter modelos NARMAX polinomiais que representem sistemas não-lineares com boa exatidão, contanto que em tais sistemas não haja uma variação muito brusca do sinal de interesse no tempo, evitando uma taxa de variação muito elevada. Além disso, o modelo NARMAX polinomial pode ser transformado em uma representação linear fixando-se o ponto de operação do sistema, ou seja, obtendo-se uma linearização do modelo. Outra vantagem é a facilidade de se obter informações analíticas sobre a dinâmica e as características em estado estacionário do modelo (Jácome, 1996).

As funções não-lineares polinomiais são lineares nos parâmetros⁴, o que permite a utilização de algoritmos de estimação de parâmetros para sistemas lineares (Billings e Voon, 1984). Esses algoritmos de estimação são fáceis de implementar, convergem rapidamente e já foram estudados em um vasto número de trabalhos (Korenberg et al., 1988; Chen et al., 1989; Aström e Witternmark, 1990).

É importante citar que a representação *NARMAX racional* pode ser considerada como uma extensão natural do modelo polinomial não-linear. Um

⁴A rigor, os modelos NARMAX são não-lineares nos parâmetros, pois o vetor estendido de regressores depende dos parâmetros e não apenas das medições (Aguirre, 2000a, p. 251). A estimação é feita através de algoritmos iterativos (ver Seção 2.6.1). Modelos com esta propriedade são às vezes chamados de modelos *pseudolineares*.

modelo racional estocástico é definido como (Billings e Chen, 1989)

$$y(k) = \frac{a \left(\begin{array}{c} y(k-1), \dots, y(k-n_y), u(k-1), \dots, \\ u(k-n_u), e(k-1), \dots, e(k-n_e) \end{array} \right)}{b \left(\begin{array}{c} y(k-1), \dots, y(k-n_y), u(k-1), \dots, \\ u(k-n_u), e(k-1), \dots, e(k-n_e) \end{array} \right)} + e(k) \quad (2.18)$$

sendo que $a(\cdot)$ e $b(\cdot)$ são funções polinomiais não-lineares quaisquer.

A fim de simplificar, os polinômios do numerador e denominador da Equação (2.18) são definidos, respectivamente, como

$$a(k) = \sum_{j=1}^{\text{num}} p_{nj}(k) \theta_{nj} \quad (2.19)$$

$$b(k) = \sum_{j=1}^{\text{den}} p_{dj}(k) \theta_{dj} \quad (2.20)$$

sendo p_{nj} e p_{dj} os regressores dados por $y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u), e(k-1), \dots, e(k-n_e)$ ou combinações lineares dos mesmos; θ_{nj} e θ_{dj} os parâmetros, todos se referindo, respectivamente, ao numerador e denominador. num é o número de termos no numerador e den o número de termos no denominador. Assim, o número de termos a ser estimado é dado por: $\text{num} + \text{den}$.

2.4 Propriedades dos modelos NARMAX

Como já foi dito, uma das grandes vantagens dos modelos NARMAX é a facilidade com que certos tipos de conhecimento podem ser extraídos e incorporados. As próximas seções mostram os tipos de informação, a saber, pontos fixos e curva estática, e a sua extração a partir de modelos NARMAX. Grande parte das propriedades descritas estão em mais detalhes em (Corrêa, 2001). Os conceitos abordados nessa seção serão usados nesse capítulo e no Capítulo 4.

2.4.1 Agrupamento de termos

A parte determinística do modelo (2.17) definido na seção (2.3.7), isto é, o modelo NARX, pode ser escrita como (Peyton-Jones e Billings, 1989)

$$y(k) = \sum_{m=0}^{\ell} \sum_{p=0}^m \sum_{n_1, n_m}^{n_y, n_u} c_{p, m-p}(n_1, \dots, n_m) \prod_{i=1}^p y(k - n_i) \prod_{i=p+1}^m u(k - n_i), \quad (2.21)$$

sendo que

$$\sum_{n_1, n_m}^{n_y, n_u} \equiv \sum_{n_1=1}^{n_y} \cdots \sum_{n_m=1}^{n_u}. \quad (2.22)$$

Os monômios da Equação (2.21) são agrupados de acordo com sua ordem m ($0 \leq m \leq \ell$). Cada termo de ordem m contém p fatores multiplicativos em $y(k - i)$ e $(m - p)$ fatores multiplicativos em $u(k - j)$. Os parâmetros destes termos são representados por $c_{p, m-p}(n_1, \dots, n_m)$, nos quais (n_1, \dots, n_m) indicam os atrasos de cada fator constituinte do monômio considerado. Para efeito de ilustração, um polinômio de grau dois, $F^2[\cdot]$, em (2.21), torna-se:

$$\begin{aligned} y(k) = & c_{0,0} + \sum_{n_1=1}^{n_y} c_{1,0}(n_1)y(k - n_1) + \sum_{n_1=1}^{n_u} c_{0,1}(n_1)u(k - n_1) \\ & + \sum_{n_1=1}^{n_y} \sum_{n_2=1}^{n_y} c_{2,0}(n_1, n_2)y(k - n_1)y(k - n_2) \\ & + \sum_{n_1=1}^{n_y} \sum_{n_2=1}^{n_u} c_{1,1}(n_1, n_2)y(k - n_1)u(k - n_2) \\ & + \sum_{n_1=1}^{n_u} \sum_{n_2=1}^{n_u} c_{0,2}(n_1, n_2)u(k - n_1)u(k - n_2). \end{aligned} \quad (2.23)$$

O primeiro somatório da Equação (2.21) faz referência aos monômios da Equação (2.17), separando-os de acordo com sua ordem. O segundo somatório faz referência ao número de fatores em $y(k - n_i)$ no termo considerado. Dentro do conjunto de termos de ordem m , um termo qualquer pode ser acessado através do ajuste do valor de p adequado. Por fim, o último somatório permite que seja feita a distinção entre os termos de (2.17), através do ajuste dos atrasos de cada um dos fatores constituintes do termo. Analisando-se o

modelo em regime permanente para entradas constantes, tem-se

$$\begin{aligned} y(k-1) &= y(k-2) = \dots = y(k-n_y) \\ u(k-1) &= u(k-2) = \dots = u(k-n_u), \end{aligned} \quad (2.24)$$

aplicando (2.24) na Equação (2.21), chega-se a

$$y(k) = \sum_{n_1, n_m}^{n_y, n_u} c_{p, m-p}(n_1, \dots, n_m) \sum_{m=0}^{\ell} \sum_{p=0}^m y(k-1)^p u(k-1)^{m-p} \quad (2.25)$$

e a Definição 2.4.1 pode ser apresentada (Aguirre e Billings, 1995b).

Definição 2.4.1 $\sum_{n_1, n_m}^{n_y, n_u} c_{p, m-p}(n_1, \dots, n_m)$ na Equação (2.25) são os coeficientes dos agrupamentos de termos $\Omega_{y^p u^{m-p}}$, que contêm termos da forma $y(k-1)^p u(k-1)^{m-p}$ para $m = 0, \dots, \ell$ e $p = 0, \dots, m$. Tais coeficientes são chamados coeficientes de agrupamento e são representado por $\sum_{y^p u^{m-p}}$.

□

Os modelos NARX polinomiais são bastante sensíveis à sobreparametrização de sua estrutura (Mendes e Billings, 1998; Aguirre e Billings, 1995a). Assim, um modelo que contenha termos que não estejam dentro dos agrupamentos efetivos pode apresentar regimes dinâmicos espúrios, ou seja, regimes dinâmicos que não são exibidos pelo sistema real.

A importância de um agrupamento de termos pode ser quantificada pelos seus coeficientes (Aguirre e Billings, 1995b). Agrupamentos de termos que possuem coeficientes nulos podem indicar, em princípio, que aqueles agrupamentos não contribuem em nada para o modelo. Ou, se o coeficiente de algum agrupamento tiver o seu valor muito menor que dos agrupamentos efetivos, esse provavelmente pode ser descartado. Os agrupamentos não efetivos ou apresentam variação de sinal dos seus coeficientes em relação ao número de termos incluídos no modelo ou sua amplitude é insignificante comparada a outros agrupamentos. Embora essas análises sejam relevantes, nem sempre seus resultados são determinantes para se descartar um agrupamento de termos. Além disso, foi mostrado que o fato de certos agrupamentos de termos praticamente se anularem, não significa necessariamente que são espúrios (Aguirre e Jácome, 1998). O modelo polinomial, identificado a partir da análise dos agrupamentos de termos efetivos, tem melhores chances de reproduzir as dinâmicas dos sistemas que se deseja estudar.

2.4.2 Pontos fixos em sistemas autônomos

Pontos fixos de um sistema discreto autônomo, são os pontos de operação que apresentam a seguinte característica:

$$y(k) = y(k + i), \quad \forall i \in \mathbb{Z}^+. \quad (2.26)$$

Sistemas dinâmicos lineares apresentam apenas um ponto fixo trivial. Em sistemas não-lineares polinomiais o número de pontos fixos vai depender do grau de não-linearidade ℓ do sistema. Por exemplo, um sistema que apresenta não-linearidade cúbica em y (variável de saída) pode apresentar três pontos de equilíbrio, ou seja, três pontos de operação na saída do sistema.

Os pontos fixos podem ser obtidos a partir de modelos NARX polinomiais através da Definição 2.4.1 de agrupamentos de termos e coeficientes de agrupamentos, resolvendo a Equação (2.27) polinomial em que o argumento k foi omitido:

$$\Sigma_{y^\ell} y^\ell + \dots + \Sigma_{y^2} y^2 + (\Sigma_y - 1) y + \Sigma_0 = 0, \quad (2.27)$$

sendo que Σ_0 é o termo constante do modelo. O modelo poderá apresentar até ℓ pontos de operação na saída se o termo $\Sigma_{y^\ell} \neq 0$. Os pontos fixos do modelo são os valores que zeram o lado esquerdo da Equação (2.27), isto é, os pontos fixos são as soluções de (2.27).

2.4.3 Pontos fixos em sistemas não-autônomos

Um modelo NARX não-autônomo analisado em estado estacionário para entrada constante pode ser escrito como segue:

$$y(k) = \sum_{n_1, n_{m\ell}}^{n_y, n_u} c_{p, m\ell-p}(n_\ell, \dots, n_{m\ell}) \sum_{m\ell=0}^{\ell} \sum_{p=0}^{m\ell} y(k)^p u(k)^{m\ell-p}, \quad (2.28)$$

sendo que $m\ell$ corresponde ao grau de não linearidade de cada termo e está na faixa $1 \leq m\ell \leq \ell$. Cada termo de grau $m\ell$ pode conter um fator dos termos $y(k)$ de ordem p , um fator em $u(k)$ de ordem $(m\ell - p)$ e um coeficiente $c_{p, m\ell-p}(n_\ell, \dots, n_{m\ell})$.

Para melhor compreensão, a Equação (2.28) pode ser reescrita como:

$$\begin{aligned}
\Sigma_y y^\ell &+ \sum_{m\ell=\ell-1}^{\ell} [\Sigma_{u^{m\ell-(\ell-1)}y^{\ell-1}} u^{m\ell-(\ell-1)}] y^{\ell-1} &+ \dots \\
&+ \sum_{m\ell=p}^{\ell} [\Sigma_{u^{m\ell-p}y^p} u^{m\ell-p}] y^p &+ \dots \\
&+ \sum_{m\ell=1}^{\ell} [\Sigma_{u^{m\ell-1}y} u^{m\ell-1} - 1] y &+ \\
&+ \sum_{m\ell=1}^{\ell} \Sigma_{u^{m\ell}} u^{m\ell} + \Sigma_0 = 0.
\end{aligned} \tag{2.29}$$

Para a determinação do número de pontos fixos em sistemas não-autônomos o procedimento é idêntico ao adotado para sistemas autônomos, mas a localização apresenta diferença. Como exemplo, será mostrada a seguir a forma para modelos lineares e modelos quadráticos.

Para um modelo linear representado em sua forma de agrupamento de termos

$$(\Sigma_y - 1)y + \Sigma_u u = 0, \tag{2.30}$$

o ponto fixo é dado por:

$$\bar{y} = \frac{\Sigma_u u}{1 - \Sigma_y}. \tag{2.31}$$

Para modelos quadráticos escritos sob a mesma forma de (2.30) tem-se

$$(\Sigma_{y^2})y^2 + (\Sigma_y - 1 + \Sigma_{uy}u)y + \Sigma_{u^2}u^2 + \Sigma_u u + \Sigma_0 = 0, \tag{2.32}$$

sendo que os pontos fixos são dados por:

$$\bar{y}_{1,2} = \frac{1 - \Sigma_y - \Sigma_{uy}u \pm \sqrt{\Delta}}{2\Sigma_{y^2}}, \tag{2.33}$$

sendo $\Delta = (\Sigma_y - 1 + \Sigma_{uy}u)^2 - 4\Sigma_{y^2}(\Sigma_{u^2}u^2 + \Sigma_u u + \Sigma_0)$.

Modelos quadráticos também podem apresentar apenas um ponto fixo, bastando para isto que o coeficiente do agrupamento Ω_{y^2} seja igual a zero ($\Sigma_{y^2} = 0$). Então o ponto fixo será determinado por:

$$\bar{y} = \frac{\Sigma_0 + \Sigma_u u + \Sigma_{u^2}u^2}{1 - \Sigma_y - \Sigma_{uy}u}. \tag{2.34}$$

Pode-se notar que a localização dos pontos fixos depende dos valores de entrada (constante) do sistema, ou seja, a localização dos pontos fixos varia com o patamar do sinal de excitação. Essas equações passam a ter a função de um mapeamento que leva \bar{u} a um valor determinado para \bar{y} . Os valores gerados por estes mapeamentos caracterizam uma curva estática para o sistema.

Então dessa forma, a função estática de sistemas não-autônomos está relacionada com seus modelos dinâmicos. Nesse caso os conceitos de agrupamentos de termos e seus coeficientes são úteis para representar o procedimento de forma compacta.

2.4.4 Simetria de pontos fixos

Considere a Equação (2.27) com $\ell = 2$:

$$\Sigma_{y^2} y^2 + (\Sigma_y - 1)y + \Sigma_0 = 0. \quad (2.35)$$

Ao se garantir $\Sigma_0 \neq 0$ e $\Sigma_y = 1$, o modelo apresenta dois pontos fixos não-triviais (fora da origem) que são simétricos com respeito à origem, e nesse caso os pontos fixos são dados por:

$$\bar{y} = \pm \sqrt{\frac{\Sigma_0}{\Sigma_{y^2}}}. \quad (2.36)$$

Utilizando a Equação (2.28), é possível determinar quais agrupamentos de termos devem ser considerados no modelo para que este apresente pontos fixos não-triviais simétricos. Em (Aguirre e Mendes, 1996) foi apresentada uma tabela indicando quais os agrupamentos de termos que um modelo NAR polinomial deve possuir em função da localização dos pontos fixos. Parte desses dados é reproduzida na Tabela 2.1.

Tabela 2.1: Agrupamentos de termos que devem ser considerados em modelos polinomiais para que este apresente pontos fixos não-triviais simétricos e pontos triviais na quantidade indicada na segunda coluna (Aguirre e Mendes, 1996), de acordo com o grau de não-linearidade (primeira coluna).

ℓ	Nº de pontos fixos triviais	Agrupamentos de termos
1	1	Ω_y
2	0	Ω_{y^2}, Ω_0
3	1	Ω_{y^3}, Ω_y
4	0	$\Omega_{y^4}, \Omega_{y^2}, \Omega_0$
4	2	$\Omega_{y^4}, \Omega_{y^2}$
5	1	$\Omega_{y^5}, \Omega_{y^3}, \Omega_y$
5	3	$\Omega_{y^5}, \Omega_{y^3}$
6	0	$\Omega_{y^6}, \Omega_{y^4}, \Omega_{y^2}, \Omega_0$
6	2	$\Omega_{y^6}, \Omega_{y^4}, \Omega_{y^2}$
6	4	$\Omega_{y^6}, \Omega_{y^4}$

A última coluna da Tabela 2.1 mostra que, sempre que for desejado ter pontos fixos não-triviais e simétricos em relação a um número ímpar de pontos fixos triviais, não pode haver no modelo termos pertencentes aos agrupamentos Ω_0 e Ω_{y^2} (Aguirre, 2000a, p. 318).

2.4.5 Curva estática

Nesta seção será utilizado o conceito de agrupamento de termos e de coeficiente de agrupamento para derivar uma equação para o ganho estático. Esta aproximação é equivalente à utilização do teorema do valor final para a resposta ao degrau com amplitude final \bar{u} . Conforme visto na Seção (2.4.3), o modelo NARX polinomial pode ser escrito da seguinte forma:

$$\hat{y} = \Sigma_0 + \Sigma_y \bar{y} + \Sigma_u \bar{u} + \sum_{m=1}^{l-1} \sum_{p=1}^{\ell-m} \Sigma_{y^p u^m} \bar{y}^p \bar{u}^m + \sum_{p=2}^{\ell} \Sigma_{y^p} \bar{y}^p + \sum_{m=2}^{\ell} \Sigma_{u^m} \bar{u}^m, \quad (2.37)$$

sendo que os termos de processo e seus respectivos parâmetros foram agrupados da seguinte forma:

- termo constante : Σ_0

- termos lineares em y : $\Sigma_y \bar{y}$
- termos lineares em u : $\Sigma_u \bar{u}$
- termos cruzados : $\sum_{m=1}^{\ell-1} \sum_{p=1}^{\ell-m} \Sigma_{y^p u^m} \bar{y}^p \bar{u}^m$
- termos não-lineares em y : $\sum_{i=1}^{\ell} \Sigma_{y^i} \bar{y}^i$
- termos não-lineares em u : $\sum_{i=1}^{\ell} \Sigma_{u^i} \bar{u}^i$.

Dessa maneira a curva estática pode ser calculada por:

$$\hat{K}(\bar{y}, \bar{u}) = \frac{\bar{y}}{\bar{u}} = \frac{\frac{\Sigma_0}{\bar{u}} + \Sigma_u + \sum_{m=2}^{\ell} \Sigma_{u^m} \bar{u}^{m-1}}{1 - \Sigma_y - \sum_{m=1}^{\ell-1} \sum_{p=1}^{\ell-m} \Sigma_{y^p u^m} \bar{y}^{(p-1)} \bar{u}^m - \sum_{p=2}^{\ell} \Sigma_{y^p} \bar{y}^{(p-1)}}. \quad (2.38)$$

A função estática de um modelo descreve a relação entre a saída em estado estacionário \bar{y} e a respectiva entrada, também em estado estacionário, \bar{u} . Esta função pode ser entendida também como a função de pontos fixos para um sistema não-autônomo.

Uma vez obtida a expressão para ganho estático de um modelo, Equação (2.38), a expressão para a curva estática pode ser determinada por

$$\bar{y} = \frac{\Sigma_0 + \Sigma_u \bar{u} + \sum_{m=2}^{\ell} \Sigma_{u^m} \bar{u}^m}{1 - \Sigma_y - \sum_{m=1}^{\ell-1} \sum_{p=1}^{\ell-m} \Sigma_{y^p u^m} \bar{y}^{(p-1)} \bar{u}^m - \sum_{p=2}^{\ell} \Sigma_{y^p} \bar{y}^{(p-1)}}. \quad (2.39)$$

2.5 Determinação de estrutura

No caso de modelos lineares, a escolha da sua estrutura se restringe, basicamente, à escolha do número de pólos e de zeros, bem como à determinação do atraso puro de tempo. Em representações não-lineares há inúmeras técnicas para cada tipo de representação. Neste trabalho as considerações a respeito da determinação de estrutura limitar-se-ão à representação NAR-MAX polinomial.

O grau de não-linearidade é um dos fatores que vai definir o número de termos candidatos, ou seja, termos possíveis nos modelos polinomiais NAR-MAX. Os aumentos no grau de não-linearidade ℓ e dos máximos atrasos n_y e n_u , provocam um aumento significativo no número de termos candidatos no modelo polinomial. O número de termos em um modelo polinomial pode ser determinado através da seguinte expressão (Korenberg et al., 1988):

$$n_\theta = M + 1, \quad (2.40)$$

em que n_θ é o número de termos (de processo e de ruído) no modelo e

$$M = \sum_{i=1}^{\ell} n_i \quad (2.41)$$

$$n_i = n_{i-1} \frac{(n_y + n_u + n_e + i - 1)}{i}, n_0 = 1. \quad (2.42)$$

O número de termos rapidamente se torna demasiadamente grande para modelos polinomiais. Mas o problema não é tão crítico como em outras representações. Nas séries de Volterra, por exemplo, o número de termos pode facilmente chegar a 10^{10} para sistemas relativamente simples (Billings, 1980). Embora o número de termos do modelo polinomial seja muito grande, em muitos casos um pequeno número desses termos é suficiente para aproximar a dinâmica do processo. É desejável a obtenção de uma representação parcimoniosa, garantindo que os termos importantes sejam levados em conta e descartando só os termos que não contribuem para a dinâmica do sistema. O procedimento para a escolha dos termos a serem incluídos no modelo é chamado de determinação de estrutura.

Várias técnicas podem ser usada para fazer a determinação de estrutura como por exemplo algoritmos genéticos (Fonseca et al., 1993) e *zeroing-and-refitting* (Kadtke et al., 1993).

2.5.1 Determinação de estrutura utilizando ERR

A taxa de redução de erro (*error reduction ratio*) ou ERR (Billings et al., 1989) associa a cada termo candidato um índice correspondente à contribuição deste na explicação da variância dos dados de saída.

Para ver isso de forma matemática, primeiro define-se a variância do erro

de modelagem $\xi(k)$ como sendo:

$$\text{var}\{\xi(k)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \left[\mathbf{y}^T \mathbf{y} - \sum_{i=1}^N g_i^2 w_i^T w_i \right], \quad (2.43)$$

em que g_i indica os elementos do vetor de parâmetros \mathbf{g} e w_i indica regressores ortogonais e \mathbf{y} é o vetor contendo os dados de saída.⁵

Supondo que nenhum termo fosse acrescentado ao modelo, a variância de $\xi(k)$ seria igual ao erro quadrático da saída $y(k)$. A cada termo acrescentado, a variância de $\xi(k)$ decresce de um fator igual a $\frac{1}{N}(g_i^2 w_i^T w_i)$, em que w_i indica o termo incluído e g_i o seu respectivo parâmetro. A redução no valor da variância pode ser normalizada com relação ao erro quadrático médio do sinal de saída. Assim, o ERR de cada termo é definido como sendo:

$$[ERR] = \frac{(g_i^2 w_i^T w_i)}{\mathbf{y}^T \mathbf{y}}. \quad (2.44)$$

O ERR pode ser utilizado na determinação de estrutura de modelos polinomiais. Escolhe-se o número de termos desejado através de uma ferramenta auxiliar, o critério de informação de Akaike por exemplo, e consideram-se aqueles que possuem os maiores valores de ERR.

2.5.2 Critério de informação de Akaike

Um procedimento para a determinação do número de termos de um modelo é o critério de informação. Um método utilizado para estimar o número de termos que devem ser incluídos no modelo é o critério de informação de Akaike (AIC). De acordo com este método, o número de termos de um modelo deve minimizar a função custo J , que se apresenta da seguinte maneira

$$J = N \log(\text{var}\{\xi(k)\}) + 2n_p, \quad (2.45)$$

sendo N o comprimento do registro de dados e n_p o número de termos de processo no modelo. Esse critério estabelece um compromisso entre a qualidade dos ajustes de identificação, contida no primeiro termo da Equação (2.45) e a procura por representações parcimoniosas, revelada pelo segundo

⁵Em Aguirre (2000a), o autor descreve um algoritmo baseado na transformação de Householder para a ortogonalização dos regressores (Chen e Billings, 1989).

termo.

O número de termos determinado a partir do AIC minimiza a variância dos resíduos de identificação partindo de uma estrutura previamente ajustada por um critério de seleção de estrutura. Contudo, não se pode afirmar que o número de termos selecionados torne o modelo capaz de reproduzir as propriedades dinâmicas do sistema original (Aguirre e Billings, 1994). O resultado obtido através do AIC pode ser visto como um indicativo na procura do número *ideal* de termos do modelo.

2.6 Estimação de parâmetros

Essa etapa começa com a escolha do algoritmo a ser utilizado. Serão descritos os procedimentos para estimação de parâmetros apenas para modelos NARMAX polinomiais.

Uma vez escolhida a estrutura de um modelo NARMAX polinomial, devem-se estimar seus parâmetros para que o modelo possa se aproximar do comportamento dinâmico do sistema original. Isso é normalmente feito em modelos polinomiais aplicando-se técnicas de mínimos quadrados (MQ) aos dados de identificação.

Para estimar os parâmetros de um mapa polinomial $F^\ell[\cdot]$ com grau de não-linearidade $\ell \in \mathbb{Z}^+$, a Equação (2.17) tem de ser expressa na forma de regressão linear

$$y(k) = \psi^T(k-1)\hat{\boldsymbol{\theta}} + \xi(k), \quad (2.46)$$

em que $\psi(k-1)$ é o vetor de regressores que contém todas as combinações lineares e não-lineares, dos termos de entrada, saída e ruído até o instante $k-1$ inclusive. Os parâmetros correspondentes a cada termo nessas matrizes são os elementos do vetor $\hat{\boldsymbol{\theta}}$. $\xi(k)$ são os resíduos ou erros de predição que são definidos como a diferença entre os dados medidos $y(k)$ e a predição um passo à frente $\psi^T(k-1)\hat{\boldsymbol{\theta}}$.

O modelo dinâmico (2.46), tomado sobre uma massa de dados, gera restrições que podem se representadas pela Equação matricial:

$$\mathbf{y} = \Psi\hat{\boldsymbol{\theta}} + \boldsymbol{\xi}, \quad (2.47)$$

em que Ψ é chamada de *matriz de regressores*. O vetor de parâmetros $\boldsymbol{\theta}$ pode

ser estimado por técnicas de mínimos quadrados ao minimizar a função custo

$$J_{\text{MQ}}(\hat{\boldsymbol{\theta}}) = (\mathbf{y} - \Psi\hat{\boldsymbol{\theta}})^T(\mathbf{y} - \Psi\hat{\boldsymbol{\theta}}). \quad (2.48)$$

A solução do problema de minimização do erro quadrático (MQ) expresso em (2.48) é dada pela seguinte expressão:

$$\hat{\boldsymbol{\theta}}_{\text{MQ}} = [\Psi^T\Psi]^{-1}\Psi^T\mathbf{y}, \quad (2.49)$$

em que $[\Psi^T\Psi]^{-1}\Psi^T$ é chamado de matriz *pseudo inversa* e $\Psi^T\Psi$ de matriz de *informação*.

A estimativa $\hat{\boldsymbol{\theta}}_{\text{MQ}}$ é dita não-polarizada se

$$E[\hat{\boldsymbol{\theta}}_{\text{MQ}}] = \boldsymbol{\theta}, \quad (2.50)$$

sendo $\boldsymbol{\theta}$ o vetor de parâmetros “reais”. Isso não acontecendo, os parâmetros são polarizados e os resíduos apresentam alguma dinâmica que não foi devidamente explicada pelo modelo. Uma alternativa para se evitar a polarização é a inclusão ou troca dos regressores do modelo, na tentativa de melhor explicar a dinâmica não modelada anteriormente.

Existem métodos de estimação de parâmetros que garantem a obtenção de estimativas não-polarizadas mesmo quando o ruído $e(k)$ não for branco ou quando o sistema apresentar ruído branco aditivo na saída do sistema. O método utilizado neste trabalho é o denominado Mínimos Quadrados Estendidos (MQE) (Aguirre, 2000a, p. 244).

2.6.1 Mínimos quadrados estendidos

A polarização no estimador MQ ocorre quando o erro que aparece na equação de regressão, $\mathbf{y} = \Psi\boldsymbol{\theta} + \mathbf{e}$, for autocorrelacionado e ao mesmo tempo o modelo incluir regressores da saída. Se os dados puderem ser modelados com regressores de entrada apenas, o estimador MQ não será polarizado, mesmo na presença de erros do tipo ruído colorido.

Se os resíduos de identificação forem modelados como um processo de média móvel, da seguinte maneira,

$$e(k) = c_i\nu(k-i) + \nu(k), \quad (2.51)$$

sendo $\nu(k)$ ruído branco, os termos $\nu(k-i)$ podem ser incorporados à matriz

de regressores e os seus respectivos parâmetros ao vetor de parâmetros do modelo da seguinte forma:

$$\mathbf{y}^* = \Psi^* \boldsymbol{\theta}^* + \mathbf{e}^*, \quad (2.52)$$

sendo que $\mathbf{y}^* = \mathbf{y}$, $\mathbf{e}^* = [\nu(k) \cdots \nu(k + N - 1)]^T$,

$$\Psi^* = \begin{bmatrix} \vdots & \nu(k-1) \\ \vdots & \nu(k) \\ \Psi & \nu(k+1) \\ \vdots & \vdots \\ \vdots & \nu(k+N-2) \end{bmatrix} \quad (2.53)$$

e $\boldsymbol{\theta}^* = [\boldsymbol{\theta} : c_i]^T$. A parte parametrizada por c_i da Equação (2.51) foi incorporada à matriz de regressores. Por este motivo, pode-se notar que \mathbf{e}^* é ruído branco, de forma que em (2.52) o termo \mathbf{e}^* é não-correlacionado com Ψ^* . Logo a estimativa por (MQ)

$$\hat{\boldsymbol{\theta}}^* = [\Psi^{*T} \Psi^*]^{-1} \Psi^{*T} \mathbf{y} \quad (2.54)$$

não apresentaria polarização, ou $E[\hat{\boldsymbol{\theta}}^*] = [\boldsymbol{\theta} : \mathbf{c}]^T$. Os elementos entre colchetes da Equação (2.54), correspondem aos mesmos termos da Equação (2.49) na Pág. 32.

Um cuidado deve ser tomado com os termos de ruído no modelo: de maneira prática, o ruído é modelado apenas com o intuito de se evitar a polarização do modelo NARX polinomial. O modelo final contém normalmente apenas termos de processo. Apenas em alguns casos, como em controle adaptativo, a parte estocástica também é usada.

Evidentemente, (2.54) não pode ser calculado como mostrado pois Ψ^* inclui regressores do tipo $\nu(k-i)$, que não são medidos. A fim de resolver este problema, tais regressores precisam ser estimados. Quando isso é feito, o resultado é um processo iterativo que é equivalente a resolver (2.54). Em Aguirre (2000a) o autor descreve um algoritmo iterativo para a estimação de $\nu(k)$ da seguinte forma:

1. A partir da equação de regressão $y(k) = \Psi^T(k-1)\boldsymbol{\theta} + e(k)$ e dos dados disponíveis, monte a Equação matricial $\mathbf{y} = \Psi\boldsymbol{\theta} + \mathbf{e}$, como no método

- de mínimos quadrados, e determine $\hat{\boldsymbol{\theta}}_{\text{MQ}} = [\Psi^T \Psi]^{-1} \Psi^T \mathbf{y}$;
2. Calcule o vetor de resíduos $\xi_i = \mathbf{y} - \Psi \hat{\boldsymbol{\theta}}_{\text{MQ}}$;
 3. Faça $i = 2$ (i indica o número de iterações);
 4. Com ξ_{i-1} , monte a matriz estendida de regressores, Ψ_i^* , e estime $\hat{\boldsymbol{\theta}}_{\text{MQE}}^* = [\Psi_i^{*T} \Psi_i^*]^{-1} \Psi_i^{*T} \mathbf{y}$;
 5. Determine o vetor de resíduos $\xi_i = \mathbf{y} - \Psi_i^* \hat{\boldsymbol{\theta}}_{\text{MQE}_i}^*$;
 6. Faça $i = i + 1$ e volte ao passo 4. Repita até convergir.

Para verificar a convergência, pode-se verificar a variância dos resíduos ou dos parâmetros estimados em cada iteração.

Para a simulação do modelo, devem-se desprezar os termos de ruído do modelo, conforme já foi dito.

2.7 Validação do modelo

Tendo obtido uma família de modelos, é necessário verificar se eles incorporam ou não as características de interesse do sistema original. Além disso, é interessante poder comparar os modelos entre si e decidir se há algum candidato significativamente melhor que os demais. Essa etapa é certamente subjetiva, sendo que o resultado da validação dependerá da aplicação pretendida para o modelo e da quantidade de informação disponível sobre o sistema original.

A ausência de algum termo importante no modelo pode provocar polarização dos parâmetros e o aparecimento de dinâmicas espúrias na simulação do modelo. Por essa razão, faz-se necessária a utilização de um critério para determinar se o modelo incorpora as características que lhe são exigidas. Neste trabalho foram considerados os seguintes aspectos:

1. predição infinitos passos à frente (simulação livre);
2. capacidade de representar o comportamento estático do sistema;
3. erro quadrático médio (RMSE - do inglês *root mean squared error*);
4. reconstrução do espaço de estados;

5. funções de correlação de resíduo.

Para casos de sistemas caóticos podem-se também citar:

6. expoentes de Lyapunov⁶;

7. mapas de primeiro retorno⁷;

8. diagramas de bifurcação⁸.

O índice RMSE é dado por:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - \hat{y}(k)]^2}, \quad (2.55)$$

que, normalizado, é escrito como:

$$\text{RMSE} = \frac{\sqrt{\sum_{k=1}^N [y(k) - \hat{y}(k)]^2}}{\sqrt{\sum_{k=1}^N [y(k) - \bar{y}]^2}}, \quad (2.56)$$

sendo que $\hat{y}(k)$ é a *simulação livre* do sinal e \bar{y} é o valor médio do sinal medido $y(k)$.

Esse índice mede o erro, em uma unidade de medida coerente com os dados reais, da simulação livre do modelo em relação aos dados reais. Neste trabalho serão considerados bons modelos aqueles que apresentarem o RMSE normalizado (2.56) menor que a unidade. Quanto menor for seu valor, melhor será considerado o modelo.

Outra maneira de verificar a validade dinâmica do modelo é o uso de funções de correlação. Essas funções podem ser aplicadas aos resíduos para a

⁶Os *expoentes de Lyapunov* medem a divergência exponencial de trajetórias vizinhas em certas direções do espaço de estados de um sistema dinâmico não-linear. Os expoentes de Lyapunov constituem uma generalização do conceito de autovalores para sistemas não-lineares (Fiedler-Ferrara e Prado, 1994, p. 139).

⁷O *mapa de primeiro retorno* (ou mapa de Poincaré) é a seqüência de pontos em que uma órbita do sistema intercepta a seção de Poincaré (Fiedler-Ferrara e Prado, 1994, p. 66).

⁸Uma *bifurcação* é uma mudança qualitativa na dinâmica de um sistema em decorrência da variação de um parâmetro de bifurcação μ_b . O *diagrama de bifurcações* mostra a dinâmica de um sistema em função de μ_b . Este diagrama apresenta de maneira concisa todas as bifurcações de comportamento do sistema analisado (Fiedler-Ferrara e Prado, 1994, p. 34).

verificação de correlação linear e não-linear. Se os resíduos possuírem alguma correlação, o modelo será considerado polarizado, uma vez que existem evidências de dinâmicas não modeladas contidas nos resíduos. Outros testes são as correlações cruzadas, lineares e não lineares, entre os resíduos e os dados de entrada e saída do sistema (Ljung, 1987; Billings e Voon, 1983, 1986).

2.8 Conclusões do capítulo

O presente capítulo teve como objetivo familiarizar o leitor com a identificação caixa-preta (ICP). As cinco principais etapas da identificação foram brevemente abordadas e algumas representações foram ilustradas.

Essas técnicas são a base de todo um conjunto de novas técnicas que estão sendo desenvolvidas para tornarem os modelos cada vez mais representativos dos sistemas reais.

O próximo passo desse trabalho é fazer um apanhado geral sobre novas técnicas em Identificação de Sistemas que tenham como ponto de partida as etapas descritas neste capítulo e a perspectiva de incorporação de *informação auxiliar*.

Identificação Multiobjetivo - Estado da Arte

“Johann Kepler was one of the most speculative astronomers of any age. He was forever theorizing, but such was the peculiar quality of his mind that his theories never satisfied him for long unless he could put them to the test of observation.”

Henry Smith Williams¹

3.1 Introdução

O uso de vários tipos de informação em identificação de sistemas não-lineares é uma área de recente e promissora pesquisa (Tulleken, 1993; Herbert e Tulleken, 1993; Sjöberg et al., 1995; Bohlin e Graebe, 1995; Johansen, 1997; Aguirre, 2000b; Corrêa, 2001; Barroso, 2001; Nepomuceno et al., 2002b,a). Tal área tem sido comumente denotada por *Identificação Caixa-Cinza (ICC)*. Entretanto, tal nomenclatura está confusa e com significado divergente em vários trabalhos. Exemplificando, Sjöberg et al. (1995) considera *caixa-cinza* a identificação que utiliza um conhecimento *a priori*² estrutural do sistema,

¹Johann Kepler foi um dos mais especulativos astrônomos de todas as épocas. Ele estava sempre teorizando, mas a peculiar qualidade de sua mente era tal que suas teorias nunca o satisfaziam, a menos que ele pudesse submetê-las ao teste da observação (Williams, 1904, p. 69).

² A informação *a priori* é aquela sabida anteriormente ao início do procedimento de identificação. Normalmente, na literatura, bem como nesse texto, refere-se ao conheci-

ou seja, as leis físicas são conhecidas parcial ou totalmente. Do mesmo modo, Aguirre et al. (2000) utilizam o conhecimento *a priori* da curva estática de um conversor buck, adquirido a partir de equações do circuito eletrônico, para identificar um modelo global, considerando corretamente tal procedimento como ICC. Por outro lado, (Corrêa, 2001; Barroso, 2001) mostram a identificação caixa-cinza de um aquecedor elétrico, em que se usa a informação da curva estática, adicionalmente ao critério convencional de *mínimos quadrados* (MQ). A informação da curva estática é obtida experimentalmente, não constituindo assim, uma informação estrutural acerca do sistema. O conceito é também utilizado de maneira ambígua por Pearson e Pottmann (2000). Nesse trabalho, os autores afirmam que o uso da informação do comportamento de regime permanente será um exemplo de ICC, sendo esse conhecimento adquirido pela física do processo ou por dados coletados empiricamente. Além disso, há trabalhos que utilizam o procedimento de agregação de informação ao MQ, denotando tal procedimento simplesmente de *uso de informação auxiliar na identificação de sistemas* (Eskinat et al., 1993); e outros, que utilizam diferente nomenclaturas, tais como *modelagem híbrida* (Cubillos et al., 1996; Forssell e Lindskog, 1997; Thompson e Kramer, 1994) e *modelagem semi-mecânica* (Braake et al., 1999) para o uso de informação auxiliar na identificação.

Essa *confusão* é de certo modo compreensível, pois a razão para a nomenclatura *caixa-cinza* vem do fato de que esse procedimento costuma usar algum tipo de técnica para estimação de parâmetros, normalmente o critério convencional de MQ – que isoladamente caracteriza a ICP – concomitantemente com algum tipo de informação adicional. Entretanto, o que parece mais evidente é que se essa informação adicional for relativa a algum conhecimento físico do sistema, esse procedimento deverá ser chamado de ICC. De um modo mais genérico, a identificação que utilizar um ou mais tipos de informação auxiliar de qualquer procedência, associada a técnicas de estimação de parâmetros, MQ ou outras, será chamada *Identificação Multiobjetivo (IM)*. A Figura 3.1 ilustra o posicionamento dos tipos de identificação. Note que a IM pode situar-se dentro das três faixas de identificação.

O objetivo desse trabalho é fazer uma breve revisão da literatura a res-

mento estrutural do sistema. Entretanto, pode-se também possuir um conhecimento *a priori* advindo de testes empíricos, e dessa forma, não se referindo a estrutura do sistema. O termo *informação auxiliar*, por sua vez, é mais genérico, e denota qualquer informação utilizada na identificação adicionalmente ao critério dos MQ.

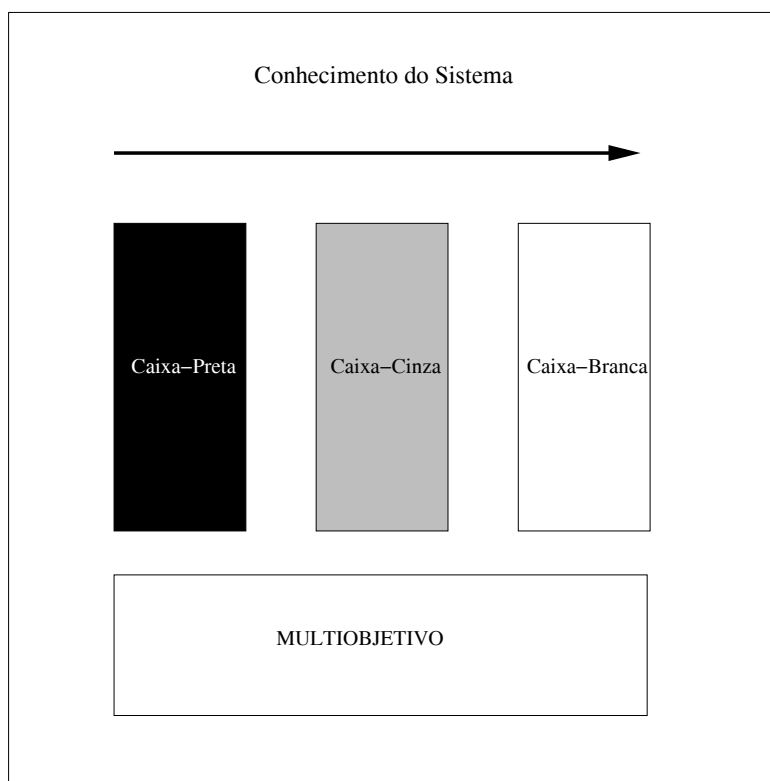


Figura 3.1: Esquema ilustrativo dos tipos de identificação

Esquema ilustrativo dos tipos de identificação. Note que a Identificação Multiobjetivo (IM) pode situar-se dentro das três faixas de identificação.

peito da IM. Inicialmente, na Seção 3.2 é apresentada a motivação para o uso de IM. Em seguida, algumas definições apresentadas na literatura para os tipos de identificação serão apresentadas na Seção 3.3. Na Seção 3.4, são feitas comparações, destacando vantagens e desvantagens, para os vários tipos de identificação. A Seção 3.5 discute algumas metodologias para a incorporação de informação auxiliar e na Seção 3.6 são apresentados vários estudos de casos. Três trabalhos relacionados com a dissertação são destacados na Seção 3.7. E por fim, a Seção 3.8 traz as conclusões do capítulo.

3.2 Motivação

O emprego da IM pode ser assim justificado:

- A minimização do erro de predição é realizada sobre uma série finita de dados, fazendo que haja sempre a presença de incerteza no modelo, que pode ser diminuída com a agregação de mais informação;
- Os dados na série podem apresentar algum tipo de erro sistemático (polarização, por exemplo), que pode ser minimizado com o uso de informação auxiliar;
- A série temporal pode conter poucas amostras, ou mesmo nenhuma amostra, em determinadas regiões de operação do sistema, que conteriam alguma característica particular que deveria ser modelada. A informação auxiliar pode ser usada para incorporar conhecimento na identificação a respeito desse comportamento particular que não esteja presente nos dados.

Além disso, a utilização de informação auxiliar é justificável do ponto de vista epistemológico, pois assegura que toda evidência disponível será utilizada. Tal argumento parece tautológico, porém há correntes de pesquisadores que seguem o princípio atestado por Fisher: “Deixem os dados falarem por si mesmos”. O princípio de Fisher tem o intuito de tornar imparciais os resultados obtidos ao analisar uma massa de dados.

3.3 Classificação da identificação

A utilização de informação auxiliar originou uma classificação nos métodos de identificação. De acordo com o nível de informação incorporada, os métodos de modelagem podem ser classificados como (Herbert e Tulleken, 1993; Sjöberg et al., 1995; Bohlin e Graebe, 1995):

- *Modelagem caixa-branca*: consiste em procedimentos em que a estrutura do modelo é totalmente conhecida, geralmente determinada pelas equações físicas ou químicas que regem o comportamento estático e dinâmico do sistema (Garcia, 1997). No inglês esse termo é escrito como *white-box*. Mas há trabalhos, como (Tan et al., 1997), em que o termo *clear-box* é usado com o mesmo significado;
- *Modelagem caixa-preta*: não se tem nenhuma informação *a priori* do sistema. Os parâmetros do modelo, geralmente, não possuem nenhum

significado físico. O processo de modelagem, chamado de identificação, se baseia única e exclusivamente em dados de entrada e saída medidos do sistema. A estrutura é escolhida de tal forma a ser flexível e usar da experiência do cientista/engenheiro (Sjöberg et al., 1995);

- *Modelagem caixa-cinza*: é a construção de modelos que incorporam informação *a priori* do sistema, com um certo grau de incerteza na seleção da estrutura da representação (Jorgensen e Hangos, 1995). Sjöberg e outros (1995), subdividem a modelagem caixa-cinza em dois subgrupos:
 - a) *Modelagem física*: toda a estrutura é determinada por conhecimentos físicos do sistema, e apenas os parâmetros, ou um certo número desses são estimados a partir dos dados;
 - b) *Modelagem semifísica*: informações do sistema são usadas para sugerir combinações não-lineares entre os sinais medidos, utilizando tais informações da estrutura do modelo.

A ICP é uma modelagem fenomenológica, pois utiliza apenas as informações disponíveis nos dados observados dos fenômenos em estudo. A identificação caixa-cinza, por sua vez, é a abordagem que faz uso da praticidade da ICP com a preocupação de compreender o fenômeno em estudo da *Identificação Caixa-Branca (ICB)*.

As Figuras 3.2 a 3.4 esquematizam os tipos de identificação de acordo com o uso de informação auxiliar.

Jorgensen e Hangos (1995) afirmam que a modelagem caixa-cinza surgiu a partir de duas disciplinas:

1. *Construção estatística de modelos*. Para projeto de controladores a informação *a priori* sobre o fenômeno físico, adicionalmente aos dados de entrada e saída, é usualmente levada em consideração pela imposição através de restrições no modelo caixa-preta;
2. *Inteligência Artificial*. Inclui todas as abordagens na física qualitativa que considera a descrição do fenômeno físico. Um exemplo dessa abordagem pode ser visto em (Bohlin, 1994).

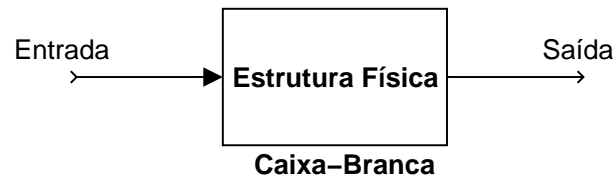


Figura 3.2: Identificação caixa-branca

Esquema ilustrativo da identificação caixa-branca (ICB)



Figura 3.3: Identificação caixa-preta

Esquema ilustrativo da identificação caixa-preta (ICP)

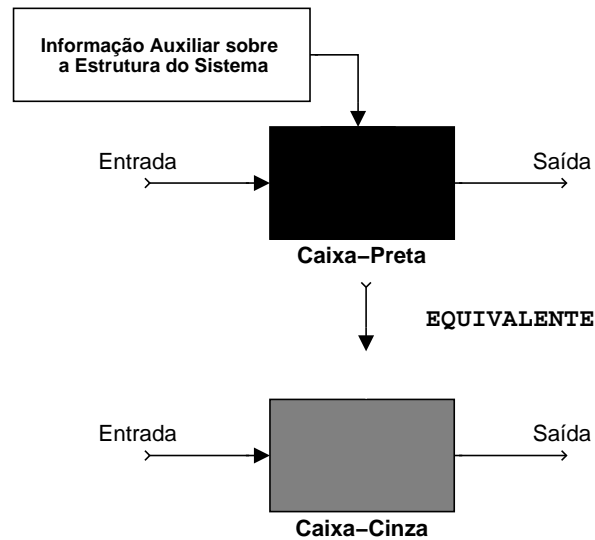


Figura 3.4: Identificação com o uso de informação auxiliar

Identificação de sistemas não-lineares utilizando informação auxiliar. Se a informação auxiliar for de qualquer natureza, a identificação será chamada de Identificação Multiobjetivo (IM). Caso a informação auxiliar refira-se a algum conhecimento físico estrutural do processo, a identificação será chamada Identificação Caixa-Cinza (ICC).

3.4 Comparação entre os tipos de identificação

Nessa seção, as abordagens caixa-branca, caixa-preta e caixa-cinza são comparadas e suas vantagens e desvantagens são destacadas.

A vantagem da modelagem caixa-branca é o fato de que os parâmetros e a estrutura do modelo possuem significado físico e são determinados a partir de informação *a priori* do sistema (Corrêa, 2001). Dessa forma, o processo de modelagem não utiliza dados de entrada e saída medidos no sistema. Em (Pottmann e Pearson, 1998) são apresentadas duas desvantagens. A primeira refere-se à maior complexidade que, em geral, tem-se na estrutura das equações, normalmente envolvendo equações diferenciais ordinárias, equações diferenciais parciais e/ou integrais. A segunda é que os modelos são geralmente constituídos por equações contínuas no tempo, enquanto os processos de medição e controle são em geral discretos.

A modelagem caixa-preta, por sua vez, possui a flexibilidade da escolha de estrutura, facilitando a formulação e a resolução de problemas de controle (Pottmann e Pearson, 1998). Como desvantagem, apresenta um alto grau de liberdade na seleção de estrutura do modelo, tornando o espaço de procura dos modelos normalmente muito extenso. Além disso, a ICP apresentará maior dificuldade na obtenção do modelo do que a ICC, em casos que apresentem número limitado de observações (Tikhonov e Arsenin, 1977).

Entre os extremos: modelagem caixa-branca e identificação caixa-preta, situa-se a identificação *caixa-cinza*, que utiliza informação auxiliar de natureza estrutural do sistema em conjunto com dados de entrada e saída medidos (Corrêa, 2001). A IM constitui-se de fato uma metodologia para a incorporação de informação auxiliar que pode ser ou não de natureza estrutural (ver Figura 3.1). Inúmeras vantagens têm sido apresentadas para a modelagem caixa-cinza. Billings (1980) afirma: *A escolha entre as várias abordagens para a identificação de sistemas não-lineares será freqüentemente conduzida pelo processo, a quantidade de informação a priori e o objetivo da identificação*. Tulleken (1993) cita que os modelos caixa-cinza podem trazer benefícios ao projeto de controladores que requerem descrições adequadas do processo. Já em (Johansen, 1996), o uso de informação *a priori* é enfatizado pois reduz no caso estudado (que não é a situação geral) efetivamente o número de parâmetros a serem determinados e torna o problema de identificação melhor condicionado. Como resultado, os modelos gerados são mais robustos,

mesmo com conjunto de dados escassos ou incompletos. Aguirre et al. (2000) reafirma Billings (1980) e ressalta a utilização de informação *a priori* na seleção de uma estrutura adequada, mostrando que deve haver um compromisso entre validação e precisão na predição. Em outras palavras, a modelagem semi-física, ou simplesmente *identificação caixa-cinza*, combina características desejáveis da modelagem caixa-branca e da identificação caixa-preta (Corrêa, 2001). Enfim, como descrito em (Kárný et al., 1995b): “modelos caixa-cinza são balanceados pois geram modelos simples, adequados para tomadas de decisão, predição ou projeto de controladores, enquanto exploram as fontes de informações disponíveis.”

No contexto de redes neurais (Cubillos et al., 1996; Cubillos e Lima, 1997) citam que “... modelagem de processos baseados em redes neurais requerem um grande número de parâmetros (pesos), resultando em dois problemas principais:

- pesada carga computacional no treinamento da rede;
- a possibilidade de sobreparametrização.

Esses dois problemas podem ser minimizados usando algum tipo de informação *a priori* na construção do modelo”. Essa última afirmação é também encontrada em (Forssell e Lindskog, 1997; Lindskog e Ljung, 1995). Kárný et al. (1995a) alerta, entretanto, que se houver extração incorreta de informação das várias fontes disponíveis, há um grande risco de ocorrer sobreparametrização.

3.5 Incorporação de informação auxiliar

A seguir será feito um levantamento dos principais tipos de informação auxiliar e como essa informação é incorporada na seleção de estrutura e estimação de parâmetros.

3.5.1 Seleção de estrutura

Será apresentada uma breve lista de trabalhos que utilizam informação auxiliar na seleção de estrutura em identificação de sistemas.

Aguirre (1994) discute as abordagens construtivas (ERR) e eliminativas. Também comenta como *zeroing-and-refitting* (Kadtke et al., 1993) pode ser

utilizada para ajuste fino de modelos. Além disso, destaca que para modelos que apresentem dinâmica não-linear a estrutura deve ser necessariamente não-linear. E por fim o autor destaca a influência da taxa de amostragem e afirma a validade do critério de Akaike para determinar o número de termos de um modelo.

Em (Bohlin, 1994) o programa *IdKit* é descrito para identificar um processo de lavagem de barras de aço de maneira interativa com o usuário. A lógica empregada é continuar o processo de interação enquanto houver falseamento das hipóteses e do modelo.³ Esse processo foi escolhido por ser não-linear, por parte do processo ser conhecido e por estar sujeito a incertezas. Além disso, dados experimentais são disponíveis. Em princípio esse procedimento determina a estrutura mais adequada baseada numa seqüência de hipóteses e falsificações. É feita uma comparação entre as abordagens caixa-branca, caixa-preta e caixa-cinza, sendo que essa última apresentou melhores resultados. Essa abordagem requer um alto grau de conhecimento do usuário.

No trabalho de (Lindskog e Ljung, 1995) a modelagem semifísica é estudada. A seleção de estrutura ocorre a partir do conhecimento físico do sistema sendo representado por equações de diferença. Os parâmetros são estimados através de redes neurais. Em um outro trabalho, Aguirre (1997) mostra como funções de correlação e espectro de ordem superior podem ser úteis na determinação da estrutura.

Em (Alessandri e Parisini, 1997) a estrutura é reduzida pois se aplica um procedimento híbrido de identificação. Assim partes do processo, no caso, uma planta de sistemas de potência conhecida, são modeladas segundo a física do processo: caixa-branca. Enquanto que outras partes, dados foram coletados e o modelo foi completado pelo uso de redes neurais.

Petrick e Wigdorowitz (1997) sugerem a utilização de expoentes de Lyapunov, mapas de Poincaré e técnicas de dimensão para a seleção de estrutura de modelos. O circuito de Chua é utilizado como estudo de caso. Apesar de parecer promissoras, as técnicas são desenvolvidas de maneira subjetiva tendo mais características de validação do que propriamente de estimação de parâmetros ou seleção de estrutura.

³Essa lógica é a idéia central presente em (Popper, 1974). O conceito de refutação e corroboração de teorias e hipóteses está presente em quase todos os campos da ciência. Em Identificação de Sistemas, essa metodologia é muito útil em problemas de validação do modelo.

Em (Lorito, 1998), o autor usa a modelagem física e estima os parâmetros através de estatística Bayesiana. O autor afirma que essa abordagem geralmente fornece modelos precisos e simples.

Braake et al. (1999) utiliza redes neurais e modelos *fuzzy* preliminares a fim de gerar novos dados. Em seguida, faz uso de identificação caixa-preta na tentativa de obter modelos mais simples. A grave pressuposição de que o modelo preliminar seja bom ou passível de melhora não é investigada.

Aguirre et al. (2000) utiliza o conhecimento da curva estática de um conversor buck CC-CC para determinar que um modelo polinomial NARMAX deva ser não-linear e não possuir termos cruzados⁴. Os autores sugerem que se informação *a priori* está disponível, então tal informação pode ser usada para restringir a estrutura do modelo. Ainda nesse trabalho os autores mostram que modelos obtidos usando técnicas caixa-preta são modelos com características locais, apesar de serem não-lineares, enquanto que “o uso de informação *a priori* força o modelo a um comportamento mais global, aumentando a faixa sobre a qual o modelo é dinamicamente válido, com uma pequena perda de precisão”. Em alguns casos, como relatado em (Barroso, 2001), a perda de precisão pode ser considerada insignificante.

Em (Aguirre, 2000b) é apresentada uma síntese das vantagens das abordagens caixa-branca, caixa-cinza e caixa-preta, além das dificuldades que serviram como motivação para o desenvolvimento de novas técnicas. Também se encontra uma lista de referências bibliográficas sobre o assunto.

Em um trabalho mais recente, Barany (2001) analisa a estimação de parâmetros em sistemas que apresentam simetria. O autor afirma que a simetria de uma solução restringe quais combinações de parâmetros um modelo linear nos parâmetros deverá possuir.

3.5.2 Estimação de parâmetros

O número de trabalhos que levam em conta algum tipo de informação auxiliar na estimação de parâmetros é significativo. Antes de listar os trabalhos que relatam estimação de parâmetros com restrição, será aqui enunciada uma formulação formal para a incorporação de conhecimento apresentada em (Tulleken, 1993). Considere-se o modelo ARMAX pertencente à classe MIMO

⁴Termos cruzados são termos de um modelo NARMAX polinomial que apresentam uma multiplicação de regressores de entrada e saída. São termos na forma $y(k-i)^p u(k-j)^m$ com $m \neq 0$ e $p \neq 0$.

(*Multiple-Input and Multiple-Output* ou múltiplas entradas e múltiplas saídas), linear nos parâmetros, caixa-preta:

$$\mathcal{M}^{\text{cp}} := \{y_k = \theta z_k + \xi_k \mid \theta \in \mathbb{R}^{\nu \times \rho}; k \in \mathbb{Z}\}, \quad (3.1)$$

em que y_k é vetor ν -dimensional de saída do processo, θ é a matriz de parâmetros desconhecida, z_k é um vetor de ρ regressores conhecidos e ξ_k são resíduos⁵. A incorporação de informação *a priori* gerará um subconjunto, Ω , que atenda a restrições de igualdade ou desigualdade:

$$\Omega := \{\theta \in \mathbb{R}^{\nu \times \rho} \mid f(\theta) = 0; g(\theta) \geq 0\}. \quad (3.2)$$

Dessa forma a estimação de parâmetros com restrição resultará no seguinte modelo:

$$\mathcal{M}^{\text{cc}} := \{y_k = \theta z_k + \xi_k \mid \theta \in \mathbb{R}^{\nu \times \rho}; k \in \mathbb{Z}; \theta \in \Omega\} \subset \mathcal{M}^{\text{cp}}. \quad (3.3)$$

A formulação acima pode ser facilmente estendida para modelos NARMAX polinomiais. Fica evidente a necessidade de se conhecer o conjunto de regressores, que nada mais é do que o resultado da determinação da estrutura do modelo.

Com ligeiras alterações, a lista de trabalhos a ser apresentada utiliza a formulação acima.

Bai e Sastry (1986) comentam que a utilização de informação *a priori* na etapa de estimação de parâmetros de modelos lineares leva a uma diminuição do erro de estimação com um menor custo computacional.

Em (Tulleken, 1993) mostra-se uma abordagem para a estimação de parâmetros que é consistente com o conhecimento *a priori* do processo. Utilizando como conhecimento *a priori* a estabilidade do processo e o ganho em estado estacionário, para selecionar o melhor estimador é utilizada estatística Bayesiana. Esse trabalho mostra aplicações em projeto de controladores para controle avançado.

Em (Eskinat et al., 1993) usa-se o conhecimento da resposta em frequência do sistema, particularmente da frequência de ressonância, para estimar os parâmetros de um modelo ARX. Também é apontada a importância da validade de modelos na região de frequência de ressonância para o projeto de controladores.

⁵A notação usada na Equação (3.1) é a usada em (Tulleken, 1993).

Em (Duran e White, 1994) encontra-se a seguinte afirmação: “O método Bayesiano é o procedimento estatístico que permite a incorporação sistemática e a ponderação apropriada de dados experimentais, conhecimento *a priori* do modelo e parâmetros e de modelos probabilísticos das origens de erro experimental.” O autor também afirma que a técnica Bayesiana permite a melhor estimação de parâmetros a um menor custo computacional.

Johansen (1996) discute três pontos em identificação caixa-cinza utilizando otimização mono-objetivo, a saber:

1. quais os tipos de conhecimento *a priori* podem ser incluídos como restrições e como isso influencia a obtenção de um modelo ótimo;
2. um procedimento numérico prático para a identificação de um modelo semi-paramétrico;
3. a seleção de estrutura através da escolha de pesos e penalidades com base nos dados empíricos.

Nesse trabalho o autor utilizou um tanque de neutralização de pH como exemplo para ilustrar os efeitos de imprecisão, de adequados e não adequados tipos de conhecimento *a priori*. Entre os tipos de conhecimento *a priori* utilizados estão:

- seqüência de dados para uma região limitada de operação;
- dados em regime permanente de cinco pontos de operação;
- modelo simplificado de balanço de massa;
- faixa de variação de pH (0 a 14);
- suavidade do comportamento do sistema;
- estabilidade em malha aberta do sistema.

Johansen (1996) mostra que a inclusão gradativa de conhecimento *a priori* pode melhorar a qualidade do modelo identificado, e, conclui que “se a informação *a priori* for correta, o modelo obtido, em geral, será melhor e mais robusto mesmo quando o conjunto de dados for deficiente ou incompleto”. Johansen (1997) diz que “...Essa abordagem pode ser útil tanto para determinação da estrutura do modelo quanto para estimação e regularização dos parâmetros”.

Alessandri e Parisini (1997) utilizam filtro de Kalman estendido para estimar parâmetros através de medições e uso de informação *a priori* sobre a física do processo. Em (Omlin e Reichert, 1999) é feita uma comparação entre os métodos clássicos e Bayesiano.

Em Johansen (2000) desenvolve-se uma abordagem multiobjetivo para modelos de resposta ao impulso finito (FIR). Segundo o autor, a estimação de parâmetros pode ser vista como um problema de múltiplos objetivos e restrições derivadas de dados empíricos (dinâmicos e estáticos), modelos físicos e do bom senso, propriedades desejadas do modelo. Particular atenção é dedicada à análise e seleção do equilíbrio entre objetivos e restrições conflitantes.

Pearson e Pottmann (2000) descrevem a identificação caixa-cinza de modelos não-lineares do tipo: Hammerstein, Wiener e modelo realimentado orientado por blocos. O conhecimento *a priori* utilizado em todos os casos é a curva estática. É sugerido nesse trabalho o uso de MQR (Mínimos Quadrados Restritos)⁶ para garantir ganho unitário em modelos dinâmicos lineares.

Em (Murakami e Seborg, 2000) os parâmetros são restritos a uma faixa de operação de acordo com o conhecimento da curva estática e de constante temporal. Dois métodos para estimação dos parâmetros são sugeridos: mínimos quadrados grampeados (CLS – do inglês *clipped least squares*) e programação quadrática (QP). O desempenho desses dois métodos é comparado ao método convencional dos mínimos quadrados.

3.6 Estudos de casos

A seguir, é apresentada uma breve lista de estudos de casos e aplicações que utilizaram algum tipo de informação auxiliar em alguma etapa do processo de identificação.

Modelos matemáticos e dados simulados são empregados nos seguintes trabalhos: (Eskinat et al., 1993; Johansen, 1996, 2000; Pearson e Pottmann, 2000; Barany, 2001), dados simulados do circuito de Chua (Petrick e Wigdorowitz, 1997).

Thompson e Kramer (1994) apresentam um método para a síntese de modelos de processos químicos que combina informação *a priori* e redes neurais

⁶Uma descrição mais detalhada do método dos MQR pode ser encontrada em (Draper e Smith, 1998). Barroso (2001) aplicou com sucesso o MQR como método de otimização mono-objetivo para identificação de sistemas dinâmicos não-lineares.

artificiais.

Em (Duran e White, 1994) apresenta-se uma abordagem Bayesiana de estimação para análise e estimação dos parâmetros de um modelo de transferência de calor.

Lindskog e Ljung (1995) aplicam a metodologia desenvolvida em seu trabalho para o aquecimento solar em residências.

Em (Alessandri e Parisini, 1997) uma seção de um sistema real de potência de 320 MW é identificada.

Aplicação em controle e otimização podem ser encontradas em (Simutis et al., 1997).

Em (Lorito, 1998) é proposto um procedimento para identificação de um modelo dinâmico não-linear para um transformador de corrente a partir de dados empíricos e de informação *a priori* da física do processo.

Em (Bechmann et al., 1999) o fluxo de deposição de poluentes em uma estação de esgotos é analisado e identificado.

Em (Braake et al., 1999) processos bioquímicos são identificados a fim de projetar controladores capazes de controlar temperatura, pH e fluxos de massa.

Em (Weyer et al., 2000) é utilizada uma modelagem caixa-cinza para a detecção de falhas de um trocador de calor.

Em (Aguirre, 2000a) um conversor buck CC-CC é modelado.

Barroso (2001) usa métodos de otimização mono-objetivo para identificação caixa-cinza de sistemas não-lineares. Com esse propósito, esse trabalho parte das definições gerais em identificação caixa-preta (sem informação auxiliar) e das características gerais da estrutura dos modelos NARX polinômiais, para definir uma metodologia que permita usar conhecimento *a priori*, ou outro tipo de informação auxiliar, nas etapas de escolha de estrutura e estimação de parâmetros de tais modelos utilizando-se ferramentas de otimização. Para avaliação da metodologia foram escolhidos dois sistemas: um conversor CC-CC buck, sendo que os dados de identificação estão contidos em uma estreita faixa de operação, o que torna a identificação de um modelo global difícil; e um sistema térmico, sendo que os dados estáticos estão contidos nos dados dinâmicos. Os resultados descritos sugerem que é possível identificar modelos dinâmicos na presença de restrições que forçam o modelo a ter certas características pré definidas em estado estacionário. Em geral, os modelos identificados fazendo-se uso de informação auxiliar, têm um comportamento global melhor de que seus análogos obtidos por técnicas

caixa-preta.

3.7 Alguns trabalhos relacionados

Nessa seção são estudados três trabalhos em maior profundidade. São os seguintes:

1. Johansen, T. (1996). Identification of non-linear systems using empirical data and prior knowledge - an optimization approach. *Automatica*, 32(3):337–356;
2. Johansen, T. A. (2000). Multi-objective identification of FIR models. In *Proceedings of 14th IFAC Symposium on System Identification 2000*, Santa Barbara, USA;
3. Aguirre, L. A., Donoso-Garcia, P. F., e Santos-Filho, R. (2000). Use of *a priori* information in the identification of global nonlinear model – a case study using a buck converter. *IEEE Transactions on Circuits and Systems Part I: Fundamental Theory and Applications*, 47(7):1081–1085.

3.7.1 (Johansen, 1996)

Nesse trabalho o critério de padrão de predição de erro é aumentado com termos de penalidade e restrições para incorporar o conhecimento *a priori*, incluindo penalidade de regularização que garante a suavidade do modelo. O projeto de parâmetros é discutido e a relevância do trabalho é ilustrada por um exemplo detalhado.

Logo na introdução, o autor mostra que a determinação de um modelo a partir de um amostra finita de observações sem nenhum conhecimento *a priori* é um problema mal-condicionado, no sentido de que não existe uma única solução. Assim, o problema de identificação é essencialmente que um conjunto finito de amostras não impõe restrições suficientes no conjunto de modelos. Para simplificar os estudos, o autor considerou que os sistemas apresentavam suavidade nos sinais (ou seja, seriam diferenciáveis até certo limite). Em seguida, afirma que a estrutura proposta nesse trabalho é uma formulação de otimização juntamente com o problema de identificação em

que os diferentes tipos de informação *a priori* são tratados como penalidades ou restrições.

O problema, sob uma abordagem de otimização, pode ser expresso da seguinte forma⁷.

$$\begin{aligned} \hat{J}(\theta) = & \frac{1}{l} \sum_{t=1}^l \xi^2(t; f(\cdot; \theta)) + \lambda \| \hat{S}(\theta) \|^2 \\ & + \gamma \mathcal{D}^2(M(f(\cdot; \theta)), M_a) + \beta \| \hat{O}(\theta) - q \|^2 \end{aligned} \quad (3.4)$$

sujeito a

$$\hat{H}(\theta) \leq 0; \hat{P}(\theta) = 0.$$

em que

$\hat{J}(\cdot)$: Função custo
λ, γ, β	: Constante não-negativas
ξ	: Resíduo de Predição
M	: Modelo
M_a	: Modelo ideal
$\mathcal{D}^2 M(f(\cdot; \theta))$: Distância entre M e M_a
S	: Penalidade de suavidade
O	: Restrições leves de igualdade

Observa-se que, além da minimização do erro quadrático, uma série de outras características podem ser incorporadas nessa metodologia. Um inconveniente presente é a sintonia das constantes λ, γ, β , ocorrendo situações em que procedimentos automáticos levarão a resultados espúrios.

A metodologia é empregada na simulação de um tanque de neutralização de pH. Nessa situação vários tipos de informação *a priori* (veja Seção 3.5.2) foram utilizados. A conclusão dos resultados foi de que a incorporação de gradual quantidade de informação *a priori* pôde melhorar a qualidade do modelo identificado. Além disso, uma sistemática para a sintonia das constantes λ, γ, β parece ser necessária.

3.7.2 (Johansen, 2000)

Johansen (2000) mostra que estimação dos parâmetros do modelo pode

⁷Notação fidedigna ao trabalho de Johansen (1996)

ser visto como um problema de múltiplos objetivos e restrições derivados de dados experimentais (dinâmicos e estáticos), modelos físicos, experiência do usuário, e propriedades desejadas do modelo. Nesse trabalho, uma abordagem multiobjetivo para identificação de sistemas é sugerida e formalizada para a identificação de modelos FIR lineares. Particular atenção é dada para a análise entre objetivos e restrições conflitantes.

Logo no princípio do trabalho, o autor identifica as razões para o uso de identificação multiobjetivo: (i) seqüência incompleta de dados; (ii) elevado grau de liberdade; (iii) informação adicional normalmente é disponível. O seu uso reduz a variância dos parâmetros do modelo; e (iv) propriedades, como suavidade, aumenta a aplicabilidade do modelo. É mostrada uma metodologia para identificação de sistemas usando otimização multiobjetivo aplicada para um modelo FIR

$$\begin{pmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{pmatrix} = \sum_{k=0}^T \begin{pmatrix} h_{11}^k & \cdots & h_{1r}^k \\ \vdots & & \vdots \\ h_{m1}^k & \cdots & h_{mr}^k(t) \end{pmatrix} \begin{pmatrix} u_1(t-k) \\ \vdots \\ u_r(t-k) \end{pmatrix}, \quad (3.5)$$

em que $y(k)$ é a saída do sistema, $u(k)$ a entrada e $h(k)$ é a resposta ao impulso.

Os objetivos e restrições usados são

- Série temporal;
- Ganho em estado estacionário;
- Regularidade (redução da variância do modelo);
- Limites dos parâmetros;
- Limite da incerteza de precisão do parâmetro.

A estratégia para solução do problema de Otimização Multiobjetivo adotada é P_w (soma ponderada de objetivos). Utiliza-se como objetivo adicional o ganho em estado estacionário e a análise geométrica permite verificar o quanto cada objetivo ou restrição está distante de seu mínimo. Além disso, limites dos parâmetros e incerteza inerente na estimação também são levados em conta.

A metodologia é empregada em um sistema dinâmico linear contínuo de 3ª ordem (que de uma certa forma pode ser considerado bem simples). O autor conclui que os resultados permitem avaliar a consistência do modelo e das técnicas de otimização e evita o cálculo de parâmetros, semelhante ao mostrado em (Johansen, 1997).

3.7.3 (Aguirre et al., 2000)

Esse artigo investiga algumas questões relacionadas à identificação caixa-cinza. É discutido como a informação *a priori* sobre a não-linearidade estática pode ser incorporada no modelo. Mostra-se também que, em algumas aplicações envolvendo sistemas não-lineares, há um compromisso entre a precisão do modelo da não-linearidade estática e a qualidade de predição. Esses conceitos foram aplicados na identificação de um conversor buck CC-CC operando em modo contínuo. Mostra-se que o modelo se torna estável em uma faixa mais ampla quando se incorpora informação *a priori*.

Os autores começam o artigo destacando a importância da modelagem de conversores de potência e afirmando que o procedimento de identificação *online* permite superar problemas como variação de valores de parâmetros em aplicações de controle. Logo em seguida, ressaltam que o grande desafio no campo da identificação é a seleção de uma adequada e compacta estrutura. Mas apontam que informação *a priori* pode ser utilizada para restringir a estrutura do modelo, conseqüentemente, facilitando a seleção.

Preliminarmente, o *conversor buck* é brevemente descrito. Tal dispositivo constitui-se em um conversor CC-CC controlado por técnica de largura de pulso modulado (PWM), cuja relação da tensão em regime permanente é dada por

$$V_o = (1 - D)V_d \quad (3.6)$$

em que V_o é a tensão na carga e D é a proporção de tempo que o dispositivo fica no estado “ligado” (*duty-cycle* ou ciclo de trabalho).

Para realizar os testes dinâmicos e aquisição de dados, foram considerados os seguintes aspectos:

- A entrada deve ser persistentemente excitante(foi usado PRBS);
- A entrada deve ser segurada por um tempo suficiente para que o sistema responda adequadamente;

- Sinal de saída: tensão de saída do conversor. Sinal de entrada: sinal CC usado para gerar o PWM.

Coletados os dados de entrada e saída, foi utilizada um modelo NARMAX polinomial. De acordo com (3.6) pode-se concluir em relação a seleção de estrutura:

- Um modelo linear seria inadequado para representar a relação estática entre entrada e saída;
- O modelo polinomial não pode conter termos cruzados de u e y nos regressores.

Como conclusão os autores destacaram que o conhecimento do formato geral da curva estática permite restringir a seleção da estrutura e que a identificação caixa-cinza para o caso estudado apresentou melhores resultados. Além disso, a identificação com a inclusão da informação da curva estática mostra aparentemente um compromisso de qualidade entre a característica dinâmica e estática.

3.8 Conclusões do capítulo

O presente capítulo procurou mostrar os recentes desenvolvimentos da *Identificação Multiobjetivo (IM)*, destacando o uso da informação *auxiliar* nas etapas de seleção de estrutura e estimação de parâmetros.

Inicialmente, os conceitos de identificação caixa-cinza, caixa-preta e caixa-branca e outras nomenclaturas presentes na literatura foram elucidadas. Foi adotada a nomenclatura *Identificação Multiobjetivo (IM)* como o procedimento que utiliza um ou mais tipos de informação auxiliar de qualquer natureza, associada a técnicas de estimação de parâmetros.

A seção seguinte destacou a motivação para o uso da IM. Foram mostradas as nomenclaturas presentes na literatura e um comparativo entre os tipos de identificação, particularmente a *identificação caixa-branca*, *identificação caixa-cinza* e *identificação caixa-preta*. Procurou-se deixar claro que a identificação caixa-cinza constitui-se em caso particular da IM, quando a informação auxiliar for relativa a um conhecimento estrutural do sistema.

Na Seção 3.5 foram apresentados vários trabalhos em que a informação auxiliar foi incorporada nas etapas de seleção de estrutura e estimação de

parâmetros. Para mostrar a relevância do assunto, seja no aspecto científico ou tecnológico, foram mostrados vários estudos de casos na seção seguinte, e posteriormente, três trabalhos mais diretamente relacionados com essa dissertação foram apresentados.

Por fim, pode-se concluir que o uso de informação auxiliar na identificação de sistemas não-lineares é uma área de grande concentração de trabalhos nos últimos anos. A formulação de conceitos claros para caracterizar as várias formas de incorporar informação auxiliar e a criação de um procedimento sistemático para incorporar essa informação são de grande relevância e constituem-se nas principais contribuições dessa dissertação.

Metodologia para Identificação Multiobjetivo

“Sempre que estiver em um dos extremos, estará fora da realidade.”

Buda

4.1 Introdução

A identificação *caixa-preta* (ICP) convencional de sistemas utiliza os dados de entrada/saída do sistema, que são obtidos através de testes dinâmicos, como única fonte de informação. Mas a observação das características dinâmicas não precisa ser o único critério utilizado para definir um bom modelo. Características em estado permanente como, por exemplo, a função estática, o seu ganho em regime permanente, o número e localização dos pontos de equilíbrio, entre outros, podem ser importantes ingredientes na obtenção de um bom modelo dinâmico.

O que se espera de modelos em geral é que sejam capazes de representar o sistema real tanto dinamicamente quanto em regime permanente. Por isso, características em regime permanente também devem ser usadas (se estiverem disponíveis) na validação de modelos.

Este capítulo descreve, de forma sucinta, a incorporação de informação auxiliar através de uma metodologia multiobjetivo. Em seguida, apresenta-se um caso particular em que determinados tipos de informação podem ser enquadrados e incorporados na representação NARMAX polinomial.

4.2 Formulação do problema

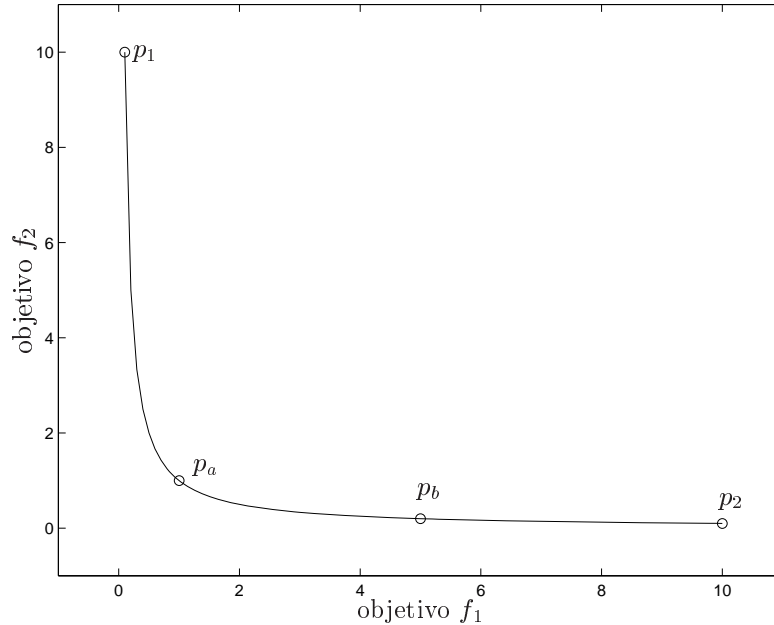


Figura 4.1: Conjunto Pareto hipotético de um problema bi-objetivo

A linha contínua ligando o ponto p_1 ao ponto p_2 é um hipotético conjunto Pareto-ótimo de um problema bi-objetivo de otimização. As soluções p_1 e p_2 denotam os ótimos individuais das funções objetivos f_1 e f_2 . p_a e p_b são soluções pertencentes ao conjunto Pareto-ótimo.

A Figura 4.1 representa o conjunto *soluções eficientes* ou *Pareto-ótimo* de um problema bi-objetivo de otimização. Isso significa que não existe nenhuma solução “abaixo” daquele conjunto, no gráfico da Figura 4.1, e qualquer método de otimização que minimize f_1 e/ou f_2 , no melhor caso, encontrará soluções pertencentes àquele conjunto.

As técnicas de identificação caixa-preta (que utilizam algoritmos de otimização mono-objetivo sem restrições) normalmente resultam em soluções como p_1 , pois levam em consideração apenas um dos objetivos, não possuindo controle nos outros objetivos (sejam eles explícitos ou não). A identificação caixa-cinza baseada em otimização mono-objetivo forneceria uma única solução que é provavelmente uma solução como p_b , já que não há razão para que esse procedimento encontre a solução p_a . A solução p_b é provavelmente inadequada, já que há outras melhores, que representariam uma pequena

degradação em um dos objetivos, com uma grande melhoria no outro.

A abordagem multiobjetivo abrange duas situações:

- Se os critérios a serem considerados na escolha final de uma solução são objetivos f_1 e f_2 apenas, a solução p_b é possivelmente inadequada, pois há melhores soluções, que representariam um pequeno decréscimo de qualidade em um objetivo e um grande aprimoramento para o outro objetivo. Nesse caso, a solução p_a seria possivelmente “a melhor”. Essa solução é próxima (no espaço dos objetivos) a ambos ótimos individuais dos objetivos f_1 e f_2 . Em outras palavras, a solução p_a , se modificada, levaria a uma pequena melhoria em um dos objetivos, enquanto que o outro objetivo sofreria uma grande depreciação em seu resultado. A abordagem multiobjetivo é um procedimento adequado para encontrar soluções como p_a .
- Se, entretanto, há outros critérios que não podem ser expressos em termos de f_1 e f_2 mas também são importantes na caracterização de soluções significativas, então todo o conjunto de *soluções eficientes* deverá ser examinado para encontrar a melhor solução. A solução, neste caso, é encontrada com alguma interação com o usuário.

Em ambas situações, quando algum tipo de análise de compromisso ocorre, a estrutura multiobjetivo será provavelmente mais útil do que técnicas mono-objetivo.

4.3 Identificação mono e multiobjetivo

O problema de identificação de sistemas pode ser considerado como um caso particular de otimização, que usualmente minimiza o erro de predição de um passo à frente associado ao modelo dinâmico, ou seja, ajustando a estrutura matemática aos dados. Em uma estrutura mono-objetivo, a incorporação do conhecimento *a priori* pode ser executada ao menos de duas maneiras básicas: aumentando a função objetivo ou introduzindo novas restrições, de tal forma a representar aquele conhecimento. Em (Johansen, 1996) o critério padrão de minimização do erro de predição é aumentado com termos de penalidade e restrições para incorporar o conhecimento *a priori*, incluindo

penalidade de regularização, que garante a suavidade do modelo. Em (Corrêa, 2001) o algoritmo elipsoidal é usado para estimar os parâmetros sujeitos a restrições que representam o conhecimento disponível. Barroso (2001) aprimorou as técnicas desenvolvidas por Corrêa (2001) utilizando algoritmo elipsoidal e algoritmo Quase-Newton BFGS.

Vários tipos de informação no procedimento de identificação podem ser representados como vários objetivos, que usualmente tendem a ser conflitantes (Aguirre et al., 2000). A fim de permitir uma análise do compromisso entre os diferentes tipos de objetivo, a *otimização multiobjetivo* se apresenta como uma conveniente alternativa (Chankong e Haimes, 1983).

A seguir será apresentada a formulação mono-objetivo e em seguida a formulação multiobjetivo.

4.3.1 Formulação mono-objetivo

A abordagem mono-objetivo para incorporação de informação auxiliar pode ser conduzida, em geral, como proposto em (Johansen, 1996). A Equação (4.1) representa um caso genérico

$$\begin{aligned} \min J(\hat{\theta}) &= J_{\text{MQ}}(\hat{\theta}) + \sum_{i=1}^m K_i J_{\text{PK}_i}(\hat{\theta}) \\ \text{sujeito a : } &\begin{cases} G_i(\hat{\theta}) \leq 0 & \forall i = 1, \dots, r \\ H_i(\hat{\theta}) = 0 & \forall i = 1, \dots, p \end{cases} \end{aligned} \quad (4.1)$$

em que $J_{\text{MQ}}(\hat{\theta})$ é o somatório do erro quadrático, $J_{\text{PK}}(\hat{\theta})$, $G(\hat{\theta})$ e $H(\hat{\theta})$ são a penalidade, restrições de desigualdade e restrições de igualdade, respectivamente. K_i são constantes não negativas. Obviamente, o conhecimento *a priori* e a escolha das constantes K_i terá uma grande influência no modelo. A escolha das constantes K_i não é uma tarefa trivial (Johansen, 1996).

De fato, a Equação (4.1) apresenta mais de um objetivo. Entretanto, sua formulação permite apenas que se extraia *uma solução* desse problema. Conforme já discutido na Seção 4.2, a geração do conjunto Pareto-ótimo é uma atraente estratégia para elencar soluções intermediárias e facilitar o processo de escolha da solução ótima.

4.3.2 Formulação multiobjetivo

Ao invés de tratar o problema de incorporação de informação auxiliar como em (4.1), a abordagem multiobjetivo considera um vetor de funções definido como

$$\mathbf{J}(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} J_1(\hat{\boldsymbol{\theta}}) \\ \vdots \\ J_n(\hat{\boldsymbol{\theta}}) \end{bmatrix}. \quad (4.2)$$

Cada uma das funções $J_1(\hat{\boldsymbol{\theta}}) \dots J_n(\hat{\boldsymbol{\theta}})$ representa o critério dos MQ e $(n - 1)$ tipos de informações auxiliares incorporadas no sistema. O problema multiobjetivo pode ser, então, denotado por

$$\begin{aligned} \min \mathbf{J}(\hat{\boldsymbol{\theta}}) \\ \text{sujeito a } \left\{ \hat{\boldsymbol{\theta}} \in \mathbf{D} \right. \end{aligned} \quad (4.3)$$

em que \mathbf{D} representa o conjunto factível do problema.

Geralmente, a minimização dos funcionais J_i possuem soluções competitivas. Isso é razoável, pois na maioria dos casos, a melhor solução não coincidirá com o ótimo individual de todos os funcionais. Uma metodologia adequada para caracterizar esse compromisso entre as soluções é produzida pela *teoria de programação multiobjetivo* (Chankong e Haimes, 1983). O problema é minimizar um vetor de funções $\mathbf{J}(\hat{\boldsymbol{\theta}})$. O conceito central na programação multiobjetivo é o conjunto *Pareto-ótimo* ou *soluções eficientes*. O conjunto $\hat{\Theta}^*$ de soluções eficientes $\hat{\boldsymbol{\theta}}^*$ pode ser caracterizado por

$$\hat{\boldsymbol{\theta}}^* \in \hat{\Theta}^* \Leftrightarrow \{ \nexists \hat{\boldsymbol{\theta}} \in \mathbf{D} : \mathbf{J}(\hat{\boldsymbol{\theta}}) \leq \mathbf{J}(\hat{\boldsymbol{\theta}}^*) \text{ e } \mathbf{J}(\hat{\boldsymbol{\theta}}) \neq \mathbf{J}(\hat{\boldsymbol{\theta}}^*) \}, \quad (4.4)$$

Uma solução eficiente deve pertencer a \mathbf{D} . Na Equação (4.4), as relações de igualdade, diferença e desigualdade são definidas como:

$$\mathbf{V} \leq \mathbf{Z} \Leftrightarrow v_i \leq z_i, \forall i \in 1, \dots, n \quad (4.5)$$

$$\mathbf{V} \neq \mathbf{Z} \Leftrightarrow \exists i \in 1, \dots, n \mid v_i \neq z_i$$

em que v_i e z_i são as componentes de um vetor n -dimensional \mathbf{V} e \mathbf{Z} . A solução é *eficiente* se e somente se não existir nenhuma outra solução que minimize um dos objetivos sem degradar ao menos um outro objetivo. Em

outras palavras, as soluções pertencentes ao Pareto-ótimo, são as melhores soluções entre as quais não existe um ordenamento (ou seja, não há como definir, a partir da avaliação dos funcionais objetivo, que uma solução é melhor que a outra).

4.4 Formulação computacional

Para determinar o conjunto de soluções eficientes, a estratégia mais comum é caracterizar esse conjunto em termos de soluções de vários problemas mono-objetivo. Uma das estratégias mais comuns é o chamado *problema de ponderação* P_w (Chankong e Haimes, 1983). Considere

$$W = \{\mathbf{w} \mid \mathbf{w} \in \mathbb{R}^n, w_j \geq 0 \text{ e } \sum_{j=1}^n w_j = 1\} \quad (4.6)$$

como sendo pesos não negativos. O problema de ponderação é definido como $\mathbf{w} \in W$ tal que:

$$\hat{\boldsymbol{\theta}}^* = \arg \min_{\hat{\boldsymbol{\theta}} \in \mathbf{D}} \sum_{i=1}^n w_i J_i(\hat{\boldsymbol{\theta}}) \quad (4.7)$$

em que J_i são todos os objetivos e restrições.

Esse método, muito simples, é capaz de achar completamente o conjunto de soluções eficientes para problemas convexos (Chankong e Haimes, 1983). O Teorema 4.4.1, desenvolvido nesse trabalho, mostra uma eficiente estratégia computacional para resolver 4.7, para casos em que os funcionais J_i puderem ser expressos pelo somatório dos erros quadráticos.

Teorema 4.4.1 *Seja $w_i \in \mathbb{R}^+$ e $J_i(\hat{\boldsymbol{\theta}})$ com $i = 1, \dots, n$ funcionais denotados por:*

$$J_i(\hat{\boldsymbol{\theta}}) = (\mathbf{v}_i - G_i \hat{\boldsymbol{\theta}})^T (\mathbf{v}_i - G_i \hat{\boldsymbol{\theta}}) \quad (4.8)$$

em que $\hat{\boldsymbol{\theta}} \in \mathbb{R}^{n_{\hat{\boldsymbol{\theta}}} \times 1}$, $\mathbf{v}_i \in \mathbb{R}^{n_v \times 1}$ e $G_i \in \mathbb{R}^{n_v \times n_{\hat{\boldsymbol{\theta}}}}$.

O problema:

$$\hat{\boldsymbol{\theta}}^* = \arg \min_{\hat{\boldsymbol{\theta}} \in \mathbf{D}} \mathbf{J}(\hat{\boldsymbol{\theta}}) = \arg \min_{\hat{\boldsymbol{\theta}} \in \mathbf{D}} \sum_{i=1}^n w_i J_i(\hat{\boldsymbol{\theta}}) \quad (4.9)$$

tem por solução:

$$\hat{\boldsymbol{\theta}}^* = \left[\sum_{i=1}^n w_i G_i^T G_i \right]^{-1} \left[\sum_{i=1}^n w_i G_i^T \mathbf{v}_i \right]. \quad (4.10)$$

□

DEMONSTRAÇÃO: Reescrevendo (4.8)

$$J_i(\hat{\boldsymbol{\theta}}) = \mathbf{v}_i^T \mathbf{v}_i - \mathbf{v}_i^T G_i \hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^T G_i^T \mathbf{v}_i + \hat{\boldsymbol{\theta}}^T G_i^T G_i \hat{\boldsymbol{\theta}}. \quad (4.11)$$

e substituindo (4.11) em (4.9)

$$\arg \min_{\hat{\boldsymbol{\theta}} \in \mathbf{D}} \mathbf{J}(\hat{\boldsymbol{\theta}}) = \arg \min_{\hat{\boldsymbol{\theta}} \in \mathbf{D}} \sum_{i=1}^n w_i (\mathbf{v}_i^T \mathbf{v}_i - \mathbf{v}_i^T G_i \hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^T G_i^T \mathbf{v}_i + \hat{\boldsymbol{\theta}}^T G_i^T G_i \hat{\boldsymbol{\theta}}). \quad (4.12)$$

A fim de minimizar (4.12) em relação a $\hat{\boldsymbol{\theta}}$, é necessário resolver $\frac{\partial \mathbf{J}(\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}}} = 0$.
Procedendo-se assim, tem-se

$$\begin{aligned} \frac{\partial \mathbf{J}(\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}}} &= 0 \\ \sum_{i=1}^n w_i (-G_i^T \mathbf{v}_i - G_i^T \mathbf{v}_i + 2G_i^T G_i \hat{\boldsymbol{\theta}}) &= 0 \\ \sum_{i=1}^n -2w_i G_i^T \mathbf{v}_i + \sum_{i=1}^n 2G_i^T G_i \hat{\boldsymbol{\theta}} &= 0 \\ \sum_{i=1}^n 2w_i G_i^T G_i \hat{\boldsymbol{\theta}} &= \sum_{i=1}^n 2w_i G_i^T \mathbf{v}_i \\ \left[\sum_{i=1}^n w_i G_i^T G_i \right] \hat{\boldsymbol{\theta}} &= \sum_{i=1}^n w_i G_i^T \mathbf{v}_i \\ \hat{\boldsymbol{\theta}} &= \left[\sum_{i=1}^n w_i G_i^T G_i \right]^{-1} \left[\sum_{i=1}^n w_i G_i^T \mathbf{v}_i \right] \end{aligned} \quad (4.13)$$

Para que $\hat{\boldsymbol{\theta}}$ seja o mínimo, é necessário verificar que

$$\frac{\partial^2 \mathbf{J}(\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}}^2} = \sum_{i=1}^n 2w_i G_i^T G_i > 0 \quad (4.14)$$

Como, por definição $w_i > 0$ e por construção $G_i^T G_i$ é positiva definida, (4.14) é satisfeito, terminando a demonstração. ■

Nota 4.4.1 O problema do Teorema 4.4.1 apresentado em (4.9) pode ser escrito na seguinte forma matricial:

$$\hat{\boldsymbol{\theta}}^* = \arg \min_{\hat{\boldsymbol{\theta}} \in \mathbf{D}} \mathbf{J}(\hat{\boldsymbol{\theta}}) = \arg \min_{\hat{\boldsymbol{\theta}} \in \mathbf{D}} [J_1(\hat{\boldsymbol{\theta}}) \dots J_n(\hat{\boldsymbol{\theta}})] \begin{bmatrix} w_1 & & 0 \\ & \ddots & \\ 0 & & w_n \end{bmatrix}, \quad (4.15)$$

cuja solução, também na forma matricial, pode se escrita como:

$$\hat{\boldsymbol{\theta}}^* = \begin{bmatrix} [G_1^T \dots G_n^T] \\ [G_1 \dots G_n] \end{bmatrix} \begin{bmatrix} w_1 & & 0 \\ & \ddots & \\ 0 & & w_n \end{bmatrix} \begin{bmatrix} G_1 \\ \vdots \\ G_n \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix}^{-1} \times \quad (4.16)$$

□

Apesar de particular, o Teorema 4.4.1 permite incorporar uma grande variedade de informações auxiliares que podem ser analiticamente expressas na representação NARMAX polinomial. Além disso, essa abordagem não necessita um ajuste de curva polinomial para uma $f(\cdot)$ de tal forma a relacionar $\bar{y} = f(\bar{u})$ como proposto em (Barroso, 2001, p. 46).

4.5 Incorporação de informação auxiliar

As seções seguintes mostram como as propriedades estruturais do modelo NARMAX polinomial podem ser utilizadas na identificação multiobjetivo de sistemas não-lineares, particularmente na etapa de estimação de parâmetros. Dois tipos de informação auxiliar: o conhecimento dos pontos fixos e da curva estática serão empregados.

4.5.1 Pontos fixos

A localização dos pontos fixos pode ser obtida a partir de (2.27), abaixo repetida para conveniência do leitor

$$\Sigma_{y^\ell} y^\ell + \dots + \Sigma_{y^2} y^2 + (\Sigma_y - 1) y + \Sigma_0 = 0,$$

ou a partir de dados medidos (Aguirre e Souza, 1998). Conhecendo os ℓ pontos fixos $[\alpha_1, \dots, \alpha_2, \alpha_\ell]$ é possível calcular o conjunto de coeficientes de agrupamento que definem tais pontos fixos. Ou seja, se um modelo tem a seguinte forma

$$\prod_{i=1}^{\ell} (\bar{y} - \alpha_i) = \Sigma_{y^\ell} \bar{y}^\ell + \dots + \Sigma_{y^2} \bar{y}^2 + (\Sigma_y - 1) \bar{y} + \Sigma_0 = 0, \quad (4.17)$$

então tal modelo tem os pontos fixos $y = [\alpha_1, \dots, \alpha_2, \alpha_\ell]$. Considere o conjunto Φ dos coeficientes de agrupamento do modelo dado pela Equação (4.18)

$$\Phi = [\Sigma_{y^\ell}, \dots, \Sigma_{y^2}, (\Sigma_y - 1), \Sigma_0]. \quad (4.18)$$

É importante ressaltar que (4.17) multiplicado por qualquer constante produzirá os mesmos pontos fixos. Então, para evitar uma grande variância nos parâmetros, o seguinte procedimento pode ser empregado. O conjunto de parâmetros obtidos por MQ é usado como referência. Através da Definição 2.4.1 um conjunto de coeficientes de agrupamento pode ser determinado

$$\Phi_{MQ} = [\Sigma_{MQy^\ell}, \dots, \Sigma_{MQy^2}, (\Sigma_{MQy} - 1), \Sigma_{MQ0}] \quad (4.19)$$

Seja a *norma euclidiana*, $\|\cdot\|$, do conjunto de parâmetros a raiz quadrada do somatório dos quadrados dos parâmetros. A norma euclidiana do conjunto

de parâmetros de qualquer modelo associado a determinados pontos fixos pode ser normalizada, a fim de igualar à norma euclidiana do conjunto Φ_{MQ} . Fazendo-se isso, tem-se o seguinte conjunto de coeficientes de agrupamento normalizado

$$\Sigma = \frac{\|\Phi_{\text{MQ}}\|}{\|\Phi\|} [\Sigma_{y^\ell}, \dots, \Sigma_{y^2}, (\Sigma_y - 1), \Sigma_0]. \quad (4.20)$$

Para minimizar o erro da localização dos pontos fixos, a seguinte função deve ser minimizada

$$J_{\text{PF}}(\hat{\boldsymbol{\theta}}) = (\Sigma - \hat{\Sigma})^T (\Sigma - \hat{\Sigma}) = (\Sigma - S\hat{\boldsymbol{\theta}})^T (\Sigma - S\hat{\boldsymbol{\theta}}). \quad (4.21)$$

em que S é um mapeamento linear tal que $\hat{\Sigma} = S\hat{\boldsymbol{\theta}}$.

A função (4.21) se enquadra no formato exigido pelo Teorema 4.4.1. É direto mostrar que para a situação em que há dois objetivos, o primeiro erro quadrático (2.48) e o segundo, localização dos pontos fixos (4.21), tem-se a estimativa da resposta estática

$$\hat{\boldsymbol{\theta}} = [w_1 \Psi^T \Psi + w_2 S^T S]^{-1} [w_1 \Psi^T \mathbf{y} + w_2 S^T \Sigma]. \quad (4.22)$$

Variando w_1 e w_2 de tal forma a $w_1 + w_2 = 1$ pode-se obter todo o conjunto Pareto-ótimo¹.

4.5.2 Curva estática

A curva estática de um sistema pode ser obtida através do teste estático que consiste em aplicar patamares de entrada fixos e esperar o sistema entrar em regime permanente para coletar a saída ou através da física do processo, como no caso do conversor buck (3.6).

Seja

$$Q = \begin{bmatrix} 1 & \bar{y} & \bar{u} & \sum_{m=1}^{\ell-1} \sum_{p=1}^{\ell-m} \bar{y}^p \bar{u}^m & \sum_{p=2}^{\ell} \bar{y}^p & \sum_{m=2}^{\ell} \bar{u}^m \end{bmatrix}, \quad (4.23)$$

em que Q está sendo chamada de *matriz de regressores estáticos*. Considere,

¹Para se empregar o estimador MQE (ver Pág. 32) deve-se estender também a matriz $S^* = [S \ ; \ \mathbf{0}_{\text{MA}}]$ em que $\mathbf{0}_{\text{MA}}$ é uma matriz de zeros com o número de linhas de S e o número de colunas iguais ao número de termos de ruído da parte média móvel (MA).

agora,

$$S\hat{\theta} = \left[\begin{array}{cccc} \Sigma_0 & \Sigma_y & \Sigma_u & \sum_{m=1}^{l-1} \sum_{p=1}^{\ell-m} \Sigma_{y^p u^m} \\ & & & \sum_{p=2}^{\ell} \Sigma_{y^p} \\ & & & \sum_{m=2}^{\ell} \Sigma_{u^m} \end{array} \right]^T, \quad (4.24)$$

que nada mais é do que uma re-organização dos coeficientes de agrupamento. $S\hat{\theta}$ é o mesmo mapeamento já descrito na Seção 4.5.1.

Substituindo (4.23) e (4.24) em (2.37), tem-se

$$\hat{y} = QS\hat{\theta}. \quad (4.25)$$

Suponha que os dados \bar{u} e \bar{y} são medidos, ou que se conheça *a priori* a relação entre eles. Pode-se expressar o erro quadrático entre a curva estática estimada e aquela medida ou conhecida *a priori*. Então é possível verificar o erro quadrático através da seguinte relação:

$$J_{CE}(\hat{\theta}) = (\bar{y} - \hat{y})^T (\bar{y} - \hat{y}) = (\bar{y} - QS\hat{\theta})^T (\bar{y} - QS\hat{\theta}). \quad (4.26)$$

E utilizando o Teorema 4.4.1 chega-se a um resultado idêntico a (4.22)

$$\hat{\theta} = [w_1 \Psi^T \Psi + w_2 (QS)^T (QS)]^{-1} [w_1 \Psi^T y + w_2 (QS)^T \bar{y}]. \quad (4.27)$$

Variando w_1 e w_2 de tal forma a $w_1 + w_2 = 1$ pode-se obter todo o conjunto Pareto-ótimo.

4.6 O Problema da decisão

Como se viu nas seções anteriores, a determinação das soluções eficientes pode ser feita de modo direto. O mesmo não acontece com a etapa em que se escolherá uma das soluções pertencentes ao Pareto-ótimo. Nessa etapa, chamada de *etapa de decisão*, deve em geral existir algum tipo de interação entre o *decisor* (usualmente um ser humano) e algum programa computacional, que irá utilizar a informação obtida do decisor para fazer a escolha convergir para uma única solução (Takahashi, 2001).

O conjunto de soluções eficientes, sendo pequeno, poderá ser apresentado ao decisor para que seja tomada uma decisão com base na avaliação simultânea. No entanto, normalmente existe uma quantidade muito grande de alternativas, que não podem ser simultaneamente processadas pelo decisor.

É necessária, portanto, uma sistemática de apresentação das alternativas que garanta que:

1. O número de consultas ao decisor será o menor possível;
2. Apresentação de um número de alternativas ao decisor seja inteligível;
3. A melhor solução não será perdida.

O encadeamento dessas etapas de decisão conduz a diferentes estruturas para o mecanismo de decisão. Basicamente, pode-se seguir uma das seguintes possibilidades descritas em (Takahashi, 2001):

- *Apresentação de preferências a posteriori*. Neste caso, será determinado um conjunto de soluções que seja representativo de todo o conjunto de soluções eficientes antes de iniciar qualquer interação com o decisor. A estrutura das consultas é formulada partindo da premissa de que as possíveis soluções já tiveram suas funções objetivo previamente avaliadas;
- *Apresentação de preferências progressivas*. Agora, a consulta ao decisor é feita concomitantemente com o processo de determinação de soluções eficientes. Cada consulta ao decisor é utilizada para determinar os parâmetros de uma nova busca de um ponto pertencente ao conjunto de soluções eficientes;
- *Apresentação de preferências a priori*. Neste esquema, o decisor é previamente consultado, e formula uma certa estrutura de preferências que “não depende” do conhecimento das alternativas concretas que vierem a se colocar. Esta situação não caracteriza propriamente um problema “multiobjetivo”, podendo ser entendida como um caso de otimização mono-objetivo. No entanto, esta é uma forma de agregar múltiplos objetivos em um problema de otimização.

Na Seção 4.6.1 é apresentada uma regra de decisão baseada no equilíbrio entre polarização e variância e na Seção 4.6.2 é apresentado um novo decisor baseado na norma mínima dos objetivos normalizados.

4.6.1 Equilíbrio entre polarização e variância

Em problemas de identificação que utilizam o somatório dos erros quadráticos como função custo, o erro pode ser decomposto em dois termos, um devido à escolha de um certo conjunto de treinamento (ou identificação) (*variância*) e outro devido à diferença entre o sistema e o modelo encontrado (*polarização*). Os efeitos devido à polarização e à variância somados formam o erro quadrático. Note que na medida em que o erro devido à polarização diminui, o erro devido à variância aumenta e vice-versa. Quanto mais complexo for um modelo, menor será o erro de polarização, enquanto o erro devido a variância aumentará. De uma certa forma, deve-se evitar modelos muito complexos que tendem a modelar o ruído presente nos dados. Quando isso ocorre, costuma-se dizer que os modelos estão sobreajustados (*overfitting*). Opostamente, quando o modelo é muito simples e o erro de polarização é maximizado, diz-se que os modelos estão sub-ajustados (*underfitting*). Esse equilíbrio entre variância e polarização também é conhecido como problema de *regularização* (Tikhonov e Arsenin, 1977). Um modelo que equilibre variância e polarização é desejável e a regularização pode, portanto, constituir em uma regra de decisão.

A regra de decisão pode ser expressa pela seguinte equação (Teixeira, 2001):

$$p^* = \arg \min_{p^* \in P^*} e_V, \quad (4.28)$$

em que p^* é a solução ótima, P^* é o conjunto Pareto-ótimo e e_V é a média do somatório dos erros quadráticos para padrões de validação, dada por :

$$e_V = \frac{1}{N_V} \sum_{i=1}^{N_V} [(d_{V_i} + \xi_i) - f(x_{V_i}; p)], \quad (4.29)$$

sendo N_V é o tamanho do conjunto de validação, $d_{V_i} + \xi_i$ é amostra do conjunto de validação somado a ruído gaussiano, $f(x_{V_i}; p)$ é o modelo estimado pertencente ao conjunto P^* . O conjunto P^* possui modelos sub-ajustados, sobre-ajustados, bem como modelos com ajuste adequado. A Figura 4.2 ilustra a solução ótima no conjunto Pareto-ótimo. As coordenadas da solução ótima, $p^* = (f_1^*, f_2^*)$ pertencentes ao Pareto-ótimo construído a partir dos dados de treinamento (identificação) são obtidas da seguinte forma. Primeiro,

encontra-se o valor mínimo de f_1 no Pareto (validação). A ordenada desse ponto será f_2^* . No Pareto (treinamento), a ordenada f_2^* determina um valor de f_1^* , formando o par ordenado que será a solução ótima.

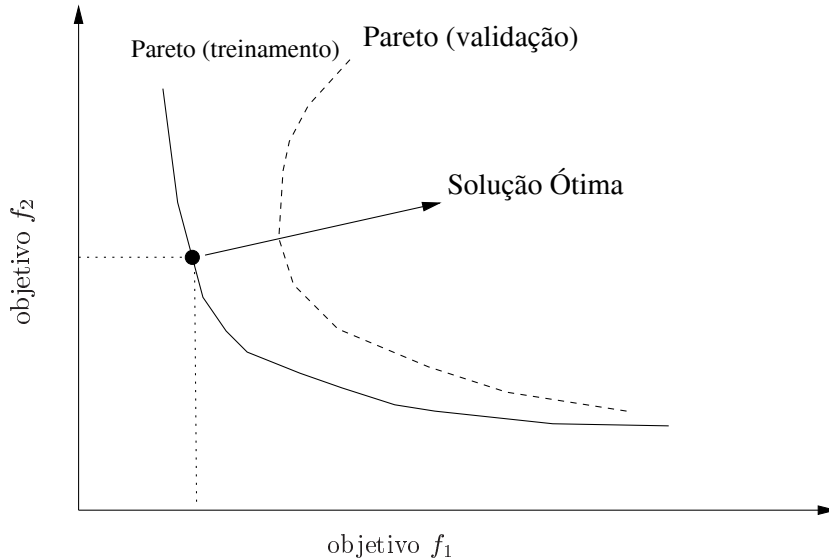


Figura 4.2: Conjunto pareto hipotético e curva de validação

Conjunto pareto hipotético e curva de validação. As coordenadas da solução ótima, pertencentes ao Pareto (treinamento) são obtidas da seguinte forma. Primeiro, encontra-se o valor mínimo de f_1 no Pareto (validação). A ordenada desse ponto será f_2^* . No Pareto (treinamento), a ordenada f_2^* determina um valor de f_1^* , formando o par ordenado que será a solução ótima.

4.6.2 Norma mínima dos objetivos normalizados

Uma outra alternativa aos procedimentos subjetivos pode ser procurar uma preferência *a priori*. Para tanto é proposto nesse trabalho uma formulação que permite encontrar uma solução que minimize a *norma euclidiana dos objetivos normalizados*. A normalização dos objetivos deverá ser realizada pois nem sempre as grandezas serão equivalentes. A normalização aqui empregada consiste em transformar os valores dos objetivos em uma escala de 0 a 1. Considere os seguintes funcionais $J_i(\hat{\theta})$ com $i = 1, \dots, n$, em que $\hat{\Theta}^*$ é o conjunto de soluções eficientes $\hat{\theta}^*$. A seguinte equação define o conjunto

de funcionais normalizados $\bar{J}_i(\hat{\boldsymbol{\theta}}^*)$

$$\bar{J}_i(\hat{\boldsymbol{\theta}}^*) = \frac{J_i(\hat{\boldsymbol{\theta}}^*) - \min(J_i(\hat{\boldsymbol{\theta}}^*))}{\max(J_i(\hat{\boldsymbol{\theta}}^*) - \min(J_i(\hat{\boldsymbol{\theta}}^*)))}, \quad (4.30)$$

em que $\min(\cdot)$ e $\max(\cdot)$ são o valor mínimo e valor máximo do vetor $J_i(\hat{\boldsymbol{\theta}}^*)$, respectivamente.

Seja $\hat{\boldsymbol{\theta}}^* = [\hat{\theta}_1^* \dots \hat{\theta}_m^*]$. A solução ótima p^* é

$$p^* = \mathbf{J}(\hat{\boldsymbol{\theta}}_k^*) \mid \min_{k \in \{1, \dots, m\}} \|[\bar{J}_1(\hat{\boldsymbol{\theta}}_k^*) \dots \bar{J}_n(\hat{\boldsymbol{\theta}}_k^*)] \|. \quad (4.31)$$

4.7 Conclusões do capítulo

Esse capítulo apresentou a incorporação de informação auxiliar através de uma abordagem multiobjetivo. O Teorema 4.4.1 permite o enquadramento de uma série de informações auxiliares. De fato, qualquer informação que possa relacionar os parâmetros de um modelo NARMAX polinomial na forma quadrática poderá ser incorporada.

Em seguida, mostrou-se dois tipos informação auxiliar que podem ser diretamente aplicados: *pontos fixos* e *curva estática*. Vale ressaltar que esses dois tipos de informação são relevantes em uma série de aplicações e que são facilmente obtidos a partir de modelos NARMAX polinomiais.

A metodologia apresentada no Teorema 4.4.1 obtém de maneira simples e sistemática o conjunto Pareto-ótimo. A seleção da melhor solução é feita na etapa de *decisão*. Mostrou-se brevemente uma estratégia de decisão baseada na procura do equilíbrio entre polarização e variância e outra que minimiza a norma euclidiana dos objetivos normalizados.

A metodologia apresentada apesar de particular revelou-se poderosa e os próximos capítulos mostrarão sua aplicação em vários estudos de casos.

Uso da Informação dos Pontos Fixos

“Mieux vaut prévoir sans certitude que de ne pas prévoir de tout.”

Henri Poincaré¹

5.1 Introdução²

A partir dos artigos (Packard et al., 1980; Takens, 1981), o campo da dinâmica não-linear testemunhou um grande interesse no problema de reconstrução do espaço de estado (ou fase) (Casdagli et al., 1991; Gibson et al., 1992) (Veja também Sauer et al. (1991) para um tratamento detalhado do assunto). Resumidamente, tal problema trata de, geralmente partindo de uma única série temporal, definir um espaço medido por variáveis que são funções da série temporal de tal maneira que o espaço reconstruído compartilhe de características dinâmicas e topológicas com o espaço original. A esse respeito, há um número de questões importantes para permitir a correta reconstrução em um ambiente não-ideal. Entre tais questões, podem-se mencionar: a escolha do tempo de atraso (Rosenstein et al., 1994), o comprimento dos dados (Kugiumtzis, 1996), a investigação dos efeitos das variáveis básicas (Casdagli

¹É muito melhor prever mesmo sem certeza, a nada prever.

²Os principais resultados apresentados nesse capítulo estão em um artigo aceito no *International Journal of Bifurcation and Chaos* (Nepomuceno et al., 2002b).

et al., 1991) e a própria escolha da variável a ser observada (Letellier et al., 1998).

Um problema diferente, embora completamente relacionado, é o da aproximação funcional ou simplesmente modelagem. Basicamente, este problema visa obter um modelo matemático (por exemplo uma ou mais equações dinâmicas) diretamente de um conjunto de dados disponíveis (normalmente, mas não necessariamente, uma única série temporal). Se tal modelo for válido, indicará a dinâmica relacionada com aquela presente nos dados. Uma grande variedade de procedimentos tem sido sugerida na literatura para obter modelos para dinâmicas não-lineares a partir de dados. Os exemplos incluem funções de base radial (Broomhead e Lowe, 1988), equações diferenciais (Gouesbet e Letellier, 1994; Brown et al., 1994; Lainscsek et al., 1998), redes neurais (Suykens e Vandewalle, 1995), equações de diferença (Aguirre et al., 1997), modelos mistos (Irving e Dewson, 1997), *wavelets* (Billings e Coca, 1999b) e modelos de Volterra (Floriani et al., 2000) para mencionar apenas alguns.

Uma característica comum a praticamente todos os métodos mencionados de reconstrução e de modelagem acima é obter um modelo utilizando apenas os dados de entrada e/ou saída. Ou seja, nenhuma informação auxiliar é utilizada. Supondo agora que algum conhecimento sobre o sistema esteja disponível antes da modelagem, seria possível usar eficientemente tal informação durante a construção do modelo? Como isso pode ser feito? Que tipo de informação pode ser usado? Acredita-se que tais perguntas abrem novas áreas de pesquisas em modelagem de dinâmica não-linear e que as primeiras respostas serão específicas para determinados tipos de representações matemáticas.

Algumas tentativas preliminares de usar a informação *a priori* na modelagem de dinâmicas não-lineares podem ser encontradas nos últimos anos. Brown e colaboradores modificaram a estrutura do modelo com base em características de simetria do fluxo modelado (Brown et al., 1994). Similarmente, a simetria dos pontos fixos (Aguirre et al., 1997) e o conhecimento do comportamento de estado estacionário foram usados para o caso não-autônomo (Aguirre et al., 2000) de tal forma a restringir o espaço factível das estruturas do modelo. Em todos esses artigos, entretanto, a informação auxiliar não foi usada diretamente no estágio de estimação de parâmetros, em que os parâmetros do modelo são estimados através de minimização do somatório do erro quadrático entre os dados e as predições do modelo. A

informação auxiliar na etapa de estimação de parâmetros foi usada em (Barroso, 2001; Aguirre e Corrêa, 2002).

Acredita-se que o presente trabalho seja pioneiro na literatura ao menos por duas razões. Primeiramente, informação auxiliar (localização de pontos fixos) será usada, no contexto de equações de diferença não-linear para diretamente atuar na estimação de parâmetros e não apenas para ajudar na restrição de classes da estrutura de modelos, como apresentando em (Aguirre et al., 2000). Segundo, o estágio de estimação de parâmetros será realizado através de uma abordagem multiobjetivo, ou seja, serão levadas em conta mais de uma função custo a ser minimizada (nesse trabalho, serão consideradas duas funções). O procedimento sugerido pode ser vantajoso em muitas situações. Esse procedimento não somente permite usar a informação disponível a respeito da localização dos pontos fixos, mas também permite que o usuário determine o peso com que tal informação deverá ser usada na modelagem. O resultado é um conjunto de modelos, o conjunto Pareto-ótimo, do qual um candidato, que melhor se adequar às necessidades do usuário, poderá ser escolhido.

No presente capítulo, a identificação multiobjetivo incorporando conhecimento *a priori* da localização de pontos fixos é aplicada em dois sistemas caóticos: circuito de Chua e o mapa senoidal. A metodologia utilizada está descrita no Capítulo 4 e também em (Nepomuceno et al., 2002b). As seções seguintes descrevem rapidamente os sistemas caóticos em estudo e discutem os resultados obtidos através da abordagem multiobjetivo.

5.2 Sistemas caóticos

Henri Poincaré constatou que sistemas dinâmicos podem exibir comportamentos de extrema complexidade no final do século XIX, durante a investigação das possíveis trajetórias apresentadas por três corpos no espaço sujeito às forças gravitacionais entre eles (Poincaré, 1957). O processo de investigação a respeito do caos foi acelerado mais de meio século depois das contribuições de Poincaré, principalmente com o trabalho de (Lorenz, 1963). A partir de então e motivado pelo advento de sistemas de computação mais poderosos, o estudo de sistemas caóticos penetrou profundamente nas áreas de física teórica e matemática computacional.

Um sistema caótico é caracterizado pela sua inerente sensibilidade a condições iniciais. Isso faz com que o sistema seja imprevisível a despeito do

conhecimento das equações diferenciais ou de diferença que descrevem o seu comportamento.

O significado usual da palavra *caos* como sendo um indicativo de desordem ou destruição diverge do significado empregado para descrever sistemas dinâmicos. Nesse caso, o regime caótico é de fato uma notável comprovação da hipótese determinística da natureza. De fato, o complexo comportamento caótico pode ser observado em sistemas descritos por equações extremamente simples. Isso tem motivado a busca por modelos matemáticos não-lineares capazes de prever o comportamento de fenômenos antes considerados puramente aleatórios (Gleick, 1987).

O estudo de sistemas caóticos tornou-se importante em várias áreas da ciência. Por exemplo, em biologia o desaparecimento de caos pode significar um fenômeno patológico (Codreanu, 2002). É importante para a mente humana e batimento cardíaco (Goldberger, 1991; Schiff et al., 1994). Em Mecânica, caos pode ser induzido para espalhar energia em pontos de ressonância (Georgiou e Scharz, 1999). Caos também é importante em sistemas de comunicação ótica (Kennedy, 2000), em processamento de informação (Codreanu, 2002), mistura de líquidos, entre outros.

Uma introdução ao estudo de sistemas caóticos pode ser encontrado em (Fiedler-Ferrara e Prado, 1994).³

5.3 Circuito de Chua

5.3.1 Descrição do circuito de Chua

O circuito de Chua (Chua e Hasler, 1993) é um sistema caótico amplamente conhecido e estudado (Tôrres, 2001). Vários trabalhos, como o controle e sincronismo de caos, utilizam o circuito de Chua como exemplo clássico de sistema caótico, dada a sua robustez, relativa facilidade de implementação (Tôrres e Aguirre, 2000) e grande variedade de regimes dinâmicos.

O circuito de Chua é um circuito eletrônico capaz de exibir oscilações periódicas ou caóticas (Kennedy, 1992), de acordo com a variação de um parâmetro do circuito. O circuito em si está mostrado na Figura 5.1 e é composto de componentes lineares, à exceção do *diodo de Chua*, que é um

³Também é possível encontrar vasto material na Internet, inclusive livros completos, tais como: *Classical and Quantum Chaos* (<http://www.nbi.dk/ChaosBook/>) e *Invitation to Dynamical Systems* (<http://www.mts.jhu.edu/~ers/invite.html>).

dipolo com a característica linear por partes mostrada na Figura 5.2. Esse diodo possui resistência negativa e é o elemento que confere energia ao sistema permitindo, assim, manter o circuito oscilando autonomamente. Fica claro que a presença de não-linearidade é devida à presença do diodo de Chua, que na prática pode ser implementado por meio de amplificadores operacionais e outros componentes simples (Kennedy, 1992).

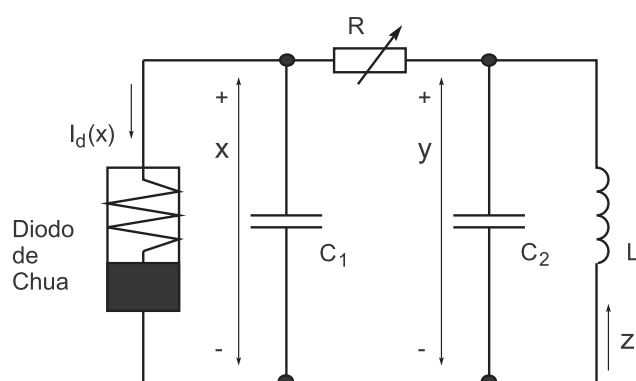


Figura 5.1: Circuito de Chua

Esquema do circuito de Chua. O diodo de Chua é o único componente não-linear.

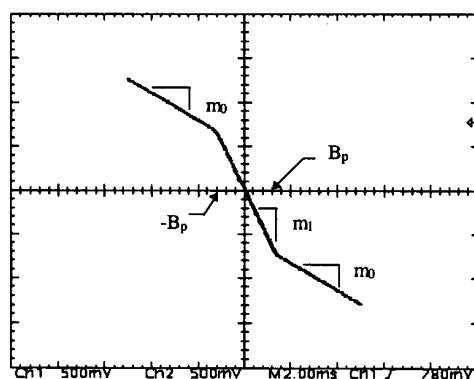


Figura 5.2: Característica corrente-tensão para o diodo de Chua

Característica corrente-tensão medida para o diodo de Chua implementado.

As Equações (5.1) descrevem o circuito de Chua:

$$\begin{aligned} C_1 \frac{dv_{C_1}}{dt} &= \frac{(v_{C_2} - v_{C_1})}{R} - i_d(v_1) \\ C_2 \frac{dv_{C_2}}{dt} &= \frac{(v_{C_1} - v_{C_2})}{R} + i_L \\ L \frac{di_L}{dt} &= -v_{C_2}, \end{aligned} \quad (5.1)$$

sendo que v_{C_i} é a tensão sobre o capacitor C_i , i_L é a corrente através do indutor e a corrente que passa pelo diodo de Chua é

$$i_d(v_1) = \begin{cases} m_0 v_1 + B_p(m_0 - m_1) & v_1 < -B_p \\ m_1 v_1 & |v_1| \leq B_p \\ m_0 v_1 + B_p(m_1 - m_0) & v_1 > B_p, \end{cases} \quad (5.2)$$

em que B_p , m_0 e m_1 são o ponto de quebra e as inclinações, respectivamente, da função linear por partes mostrada na Figura 5.2.

O seguinte conjunto de parâmetros pode ser utilizado para simular o circuito de forma a se aproximar do protótipo montado por Aguirre e colaboradores (Aguirre e Tôres, 1997): $C_1 = 10,0\text{nF}$, $C_2 = 100\text{nF}$, $L = 19,2\text{mH}$ e R é um *trimpot* de $2,0\text{ k}\Omega$. O diodo de Chua foi montado usando amplificadores operacionais (maiores detalhes em (Kennedy, 1992)), tendo sido medidos os seguintes valores para os parâmetros: $m_0 = -0,37$, $m_1 = -0,68$ e $B_p = 1,1$.

Vários comportamentos caóticos e regulares podem ser conseguidos através da variação do *trimpot* R . Nesse trabalho, fixou-se $R \approx 1800$ de tal forma a se obter o atrator dupla volta (Aguirre e Tôres, 1997). As Figuras 5.3 e 5.4 mostram a série temporal e o atrator dupla volta, respectivamente.

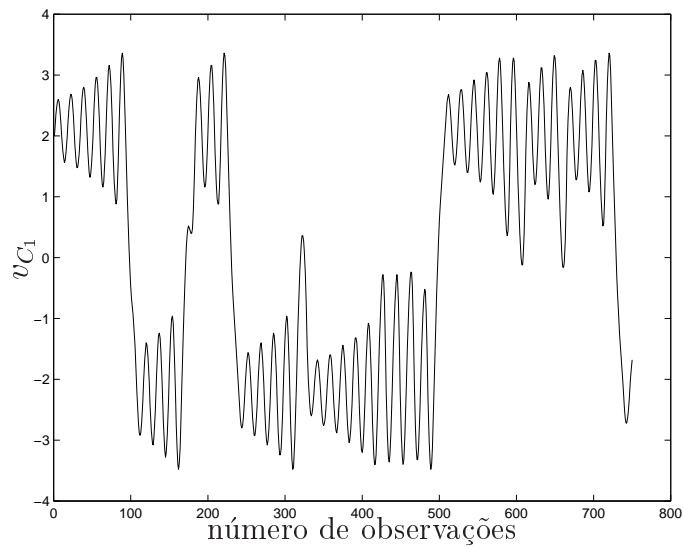


Figura 5.3: Série temporal do circuito de Chua

Série temporal do atrator dupla-volta do Circuito de Chua obtida a partir de um circuito implementado (Aguirre et al., 2000).

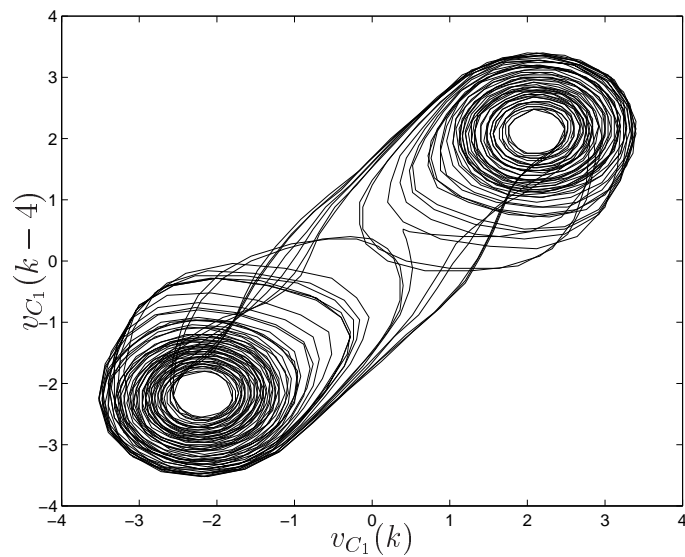


Figura 5.4: Atrator dupla volta do circuito de Chua

Atrator dupla volta para $R \approx 1800$, reconstruído a partir de 2000 observações da tensão sobre o capacitor C_1 do circuito de Chua.

5.3.2 Metodologia

Nesse exemplo usou-se um modelo NAR obtido da implementação real do circuito de Chua (Aguirre et al., 1997)

$$\begin{aligned}
y(k) = & +3,523y(k-1) - 4,2897y(k-2) \\
& -0,2588y(k-4) - 1,7784y(k-1)^3 \\
& +2,0652y(k-3) + 6,1761y(k-1)^2y(k-2) \\
& +0,1623y(k-1)y(k-2)y(k-4) - 2,7381y(k-1)^2y(k-3) \\
& -5,5369y(k-1)y(k-2)^2 + 0,1031y(k-2)^3 \\
& +0,4623y(k-4)^3 - 0,5247y(k-2)^2y(k-4) \\
& -1,8965y(k-1)y(k-3)^2 + 5,4255y(k-1)y(k-2)y(k-3) \\
& +0,7258y(k-2)y(k-4)^2 - 1,7684y(k-3)y(k-4)^2 \\
& +1,18y(k-3)^2y(k-4) + \xi(k).
\end{aligned} \tag{5.3}$$

O sistema é autônomo e não tem conseqüentemente nenhuma entrada *eXógena*, e a parte de média móvel foi estimada para reduzir a polarização, não sendo usada para a simulação. Esse exemplo usará dados gerados pelo modelo (5.3) a fim de evitar problemas na seleção de estrutura. Isto constitui a melhor situação possível para as técnicas mono-objetivo tradicionais. A técnica multiobjetivo proposta no Capítulo 4 será empregada neste caso, para diferentes níveis de ruído branco aditivo. Desta forma, a função da técnica multiobjetivo poderá ser avaliada.

Os pontos fixos do modelo estão aproximadamente em $\bar{y} = 0; \pm 2,2417$ e o mapa tem o *maior expoente de Lyapunov* (ML) estimado em $\lambda = 3,6250 \pm 0,075$ bits/s. A Figura 5.5 mostra o atrator (dupla-volta) e a Figura 5.6 mostra o mapa de primeiro retorno. Ambas as figuras foram obtidas simulando o modelo (5.3).

Este exemplo será estudado seguindo dois trajetos. Primeiramente, usando os dados obtidos do modelo (5.3) o sistema será identificado usando a mesma estrutura. Esse é um caso ideal, em que a estrutura é conhecida e os dados não são corrompidos por ruído. Em seguida, ruído branco com média zero e variância 0,0001 é adicionado aos dados. A relação sinal-ruído do sinal (SNR) é de aproximadamente 93,53 dB, comparado a 72,30 dB dos dados reais usados em (Aguirre et al., 1997). Apesar do fato que no caso atual a estrutura é supostamente conhecida, ver-se-á que a aproximação convencional não poderá identificar exatamente a dinâmica do sistema.

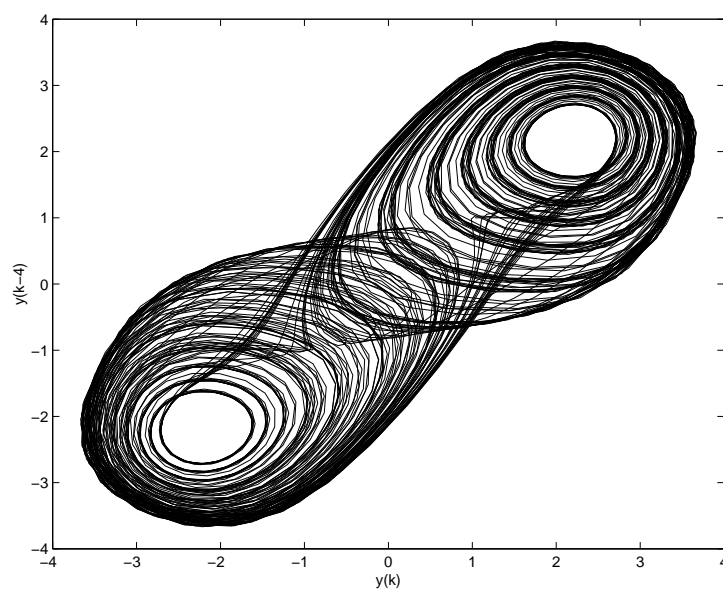


Figura 5.5: Atrator dupla-volta reconstruído - circuito de Chua

Atrator dupla-volta reconstruído a partir do modelo (5.3) do Circuito de Chua.

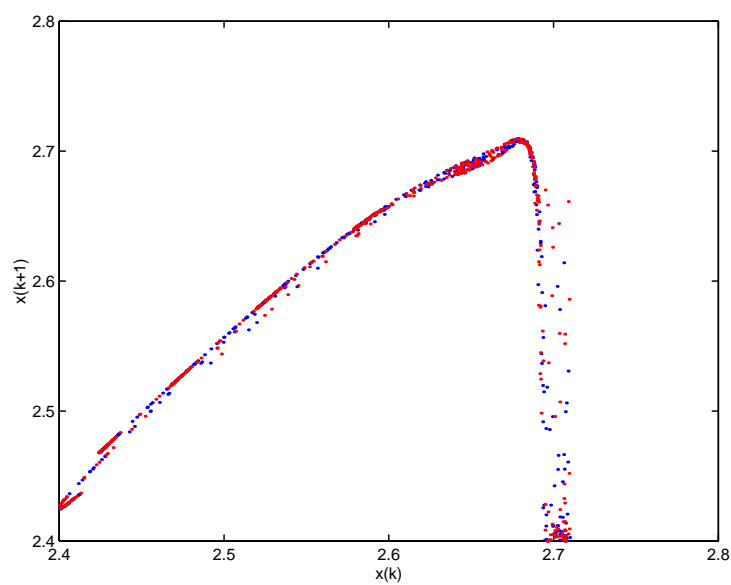


Figura 5.6: Mapa de primeiro retorno - circuito de Chua

Mapa de primeiro retorno do modelo (5.3) do circuito de Chua.

5.3.3 Caso 1 - sem ruído

O conhecimento dos pontos fixos foi incluído na estimação de parâmetros por meio de (4.22). Vinte modelos foram identificados. As soluções do *Pareto-ótimo* são mostradas na Figura 5.7, em que w foi variado de 0 a 1.

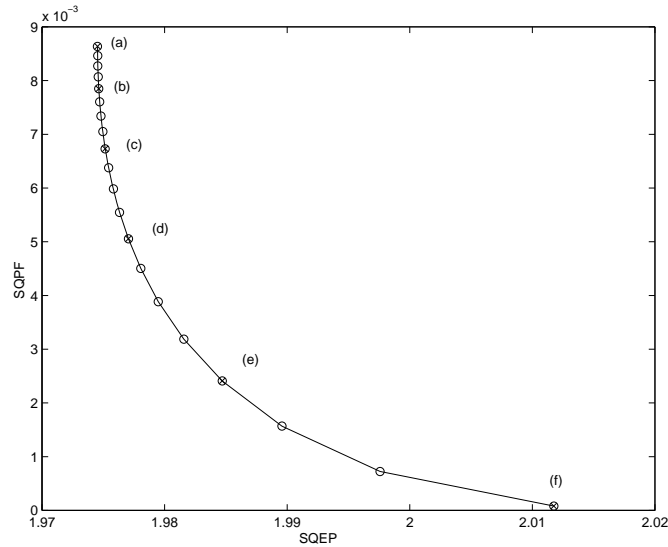


Figura 5.7: Propriedades de modelos do circuito de Chua - sem ruído

Soluções do *Pareto-ótimo*. Algumas propriedades dinâmicas e estáticas dos modelos assinalados pelas letras $(a-f)$, mostrados na Tab. 5.1. Caso sem ruído.

Por uma questão de simplicidade, apenas seis modelos são apresentados $(a-f)$, como mostrado na Figura 5.7, e serão analisados mais detalhadamente. A Tabela 5.1 mostra os seguintes parâmetros dos modelos $(a-f)$: (i) Peso usado no problema P_w ; (ii) somatório do quadrado dos erros de predição (SQEP); (iii) somatório do quadrado dos erros da localização dos pontos fixos (SQPF); (iv) a localização dos pontos fixos; (v) e o maior expoente de Lyapunov (ML).

Como esperado, quando o peso tende a 1 a soma dos erros entre os pontos fixos do modelo e os valores especificados vai a zero (ver Figura 5.7). Inversamente, enquanto o peso vai a zero, a solução aproxima-se gradualmente da solução obtida usando a técnica mono-objetivo. A Figura 5.8 mostra o atrator de projeção bi-dimensional para os modelos $(a-f)$. É interessante observar que enquanto o ponto fixo é forçado a seu valor original, a dinâmica

Tabela 5.1: Algumas propriedades estáticas e dinâmicas de seis modelos para o atrator dupla-volta do modelo original (5.3). Caso *sem* ruído. Legenda: (i) Peso usado no problem P_w ; (ii) somatório do quadrado dos erros de predição (SQEP); (iii) somatório do quadrado dos erros da localização dos pontos fixos (SQPF); (iv) a localização dos pontos fixos; (v) e o maior expoente de Lyapunov (ML).

Modelo	(a)	(b)	(c)
Peso	0,001	0,2111	0,4212
SQEP	1,9745	1,9746	1,9800
SQPF	0,0086	0,0078	0,0037
Ponto Fixo	0; $\pm 2,2466$	0; $\pm 2,2473$	0; $\pm 2,2606$
ML	$2,8263 \pm 0,0459$	$2,6291 \pm 0,1567$	$2,0530 \pm 0,0042$
Modelo	(d)	(e)	(f)
Peso	0,6313	0,8414	0,9990
SQEP	1,9918	2,0016	2,0197
SQPF	0,0013	0,0005	$2,4251 \times 10^{-7}$
Ponto Fixo	0; $\pm 2,2141$	0; $\pm 2,2292$	0; $\pm 2,2347$
ML	–	–	–

do modelo se deteriora e torna-se eventualmente instável (veja o *modelo f*). O algoritmo para cálculo dos expoentes de Lyapunov não convergiu para os modelos *d-e*, embora tais modelos sejam estáveis. Não se calculou o expoente de Lyapunov para o modelo *f*, pois este modelo é instável.

O mapa de primeiro retorno é apresentado na Figura 5.9. A forma do mapa muda consideravelmente do modelo *a* para o modelo *f*. Considerando o conjunto das características dos modelos apresentados na Tabela 5.1, é possível que o modelo *a* seja o melhor. Esse resultado, de certo modo, era esperado porque, devido ao fato de tratar-se de um caso simulado sem a presença de ruído, existe um modelo NAR que simultaneamente obteria o mínimo erro de predição, e também, o mínimo erro na localização dos pontos fixos para uma série temporal arbitrariamente longa: o modelo original (5.3). Nesse caso, a ICP, na verdade, produz um excelente resultado. Vale destacar que a variação do erro da localização do ponto fixo na Figura 5.7 é praticamente desprezível. No entanto, forçar os pontos fixos, ainda que pouco, produz fortes alterações na dinâmica global do modelo.

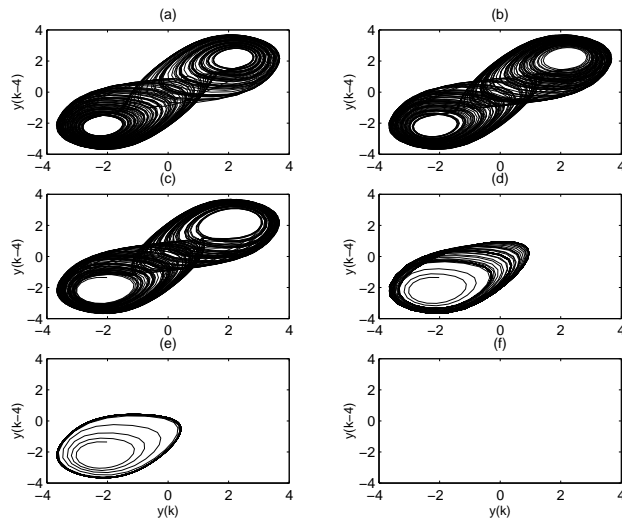


Figura 5.8: Atratores de modelos do circuito de Chua - sem ruído

Projeção dos atratores dupla-volta correspondentes aos modelos (a-f), veja Tabela 5.1 e Figura 5.7. Caso sem ruído.

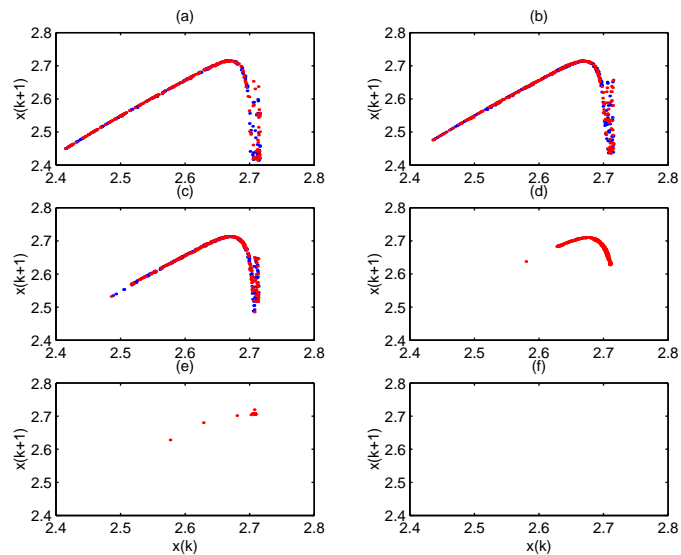


Figura 5.9: Mapa de primeiro retorno do circuito de Chua - sem ruído

Mapa de primeiro retorno dos modelos (a-f), veja Tabela 5.1 e Figura 5.7. Caso sem a presença de ruído.

5.3.4 Caso 2 - com ruído

No segundo caso, os dados da série temporal são corrompidos artificialmente com ruído aditivo. Nesse caso, deve-se esperar que o mínimo do erro de predição não esteja próximo ao mínimo do erro do ponto fixo, como estava para o caso sem ruído. Ou seja, deve haver uma diferença maior entre os modelos estimados pela minimização da soma de erros quadrados de predição e pela minimização da soma de erros quadrados da localização dos pontos fixos. Conseqüentemente, uma solução multiobjetivo seria altamente desejável.

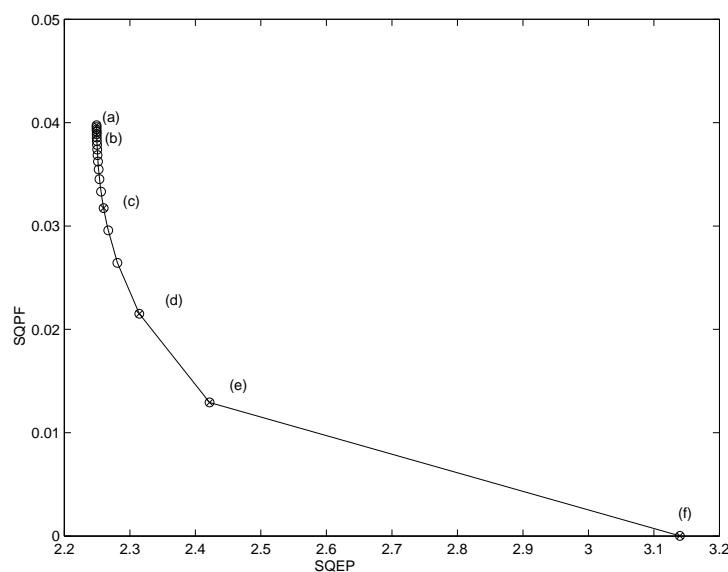


Figura 5.10: Pareto-ótimo de modelos do circuito de Chua - com ruído

Soluções do *Pareto-ótimo*. Algumas propriedades dinâmicas e estáticas dos modelos assinalados pelas letras *a-f* são mostradas na Tabela 5.2. Caso com ruído.

A Figura 5.10 mostra as soluções de Pareto-ótimo. A escala dos valores medidos pela soma de erros de ponto fixo ao quadrado é muito mais larga e os erros são dez vezes maiores do que no caso sem ruído. A Tabela 5.2 mostra algumas propriedades dinâmicas e estáticas de seis modelos para esse caso. A Figura 5.11 mostra as projeções dos atratores de dupla volta para os modelos (*a-f*) e a Figura 5.12 mostra o mapa de primeiro retorno para os modelos mencionados. Por contraste com o caso livre de ruído, o melhor modelo parece ser o modelo *c* que é uma solução multiobjetivo, isto é, ao estimar os parâmetros, dois objetivos diferentes foram considerados simultaneamente.

Tabela 5.2: Algumas propriedades estáticas e dinâmicas de seis modelos para o atrator dupla-volta do modelo original (5.3). Caso *com* ruído. Legenda: (i) Peso usado no problem P_w ; (ii) somatório do quadrado dos erros de predição (SQEP); (iii) somatório do quadrado dos erros da localização dos pontos fixos (SQPF); (iv) a localização dos pontos fixos; (v) e o maior expoente de Lyapunov (ML).

Modelo	(a)	(b)	(c)
Peso	0,001	0,2111	0,4212
SQEP	2,2494	2,2495	2,2599
SQPF	0,0397	0,0389	0,03174
Ponto Fixo	0; \pm 2,2333	0; \pm 2,2331	0; \pm 2,2308
ML	2,2122 \pm 0,0330	2,1475 \pm 0,0512	1,6107 \pm 0,0051
Modelo	(d)	(e)	(f)
Peso	0,6313	0,8414	0,9990
SQEP	2,3145	2,4216	3,1395
SQPF	0,02151	0,0129	2,0952 $\times 10^{-5}$
Ponto Fixo	0; \pm 2,2230	0; \pm 2,2279	0; \pm 2,2520
ML	–	–	–

Este modelo assimilou satisfatoriamente as propriedades dinâmicas (veja Figuras 5.11 e 5.12) e tem uma soma do erro de ponto fixo quadrado 25% menor do que o modelo obtido com a metodologia caixa-preta. Ou seja, com um aumento insignificante na soma do erro quadrado da predição, a técnica multiobjetivo sugerida no Capítulo 4 permite ao usuário obter sistematicamente modelos com desempenho melhor em relação à posição dos pontos fixos. Deve-se destacar que tais modelos, além de ter os pontos fixos que são mais precisos, têm freqüentemente a característica dinâmica mais próxima do sistema original, como por exemplo o *modelo c*.

Uma outra característica interessante de se observar é a variação dos parâmetros do modelo ao longo do Pareto. A Tabela 5.3 mostra a variação dos parâmetros dos modelos ao longo do Pareto. Na primeira coluna estão os termos, na segunda, estão os valores dos parâmetros retirados da Equação (5.3). As colunas de 3 a 8 mostram a variação dos parâmetros para os modelos (*a-f*). Note que mesmo para o modelo *a*, que está mais próximo da solução de MQ, os parâmetros não são semelhantes aos do modelo original. Fica também claro que apenas os primeiros termos, de cima para baixo, sofrem alterações pouco significativas ao longo do Pareto. Como essa Tabela está

respeitando a ordem decrescente do valor de ERR, pode-se conjecturar que termos com alto ERR sofrem menores alterações ao longo do Pareto.

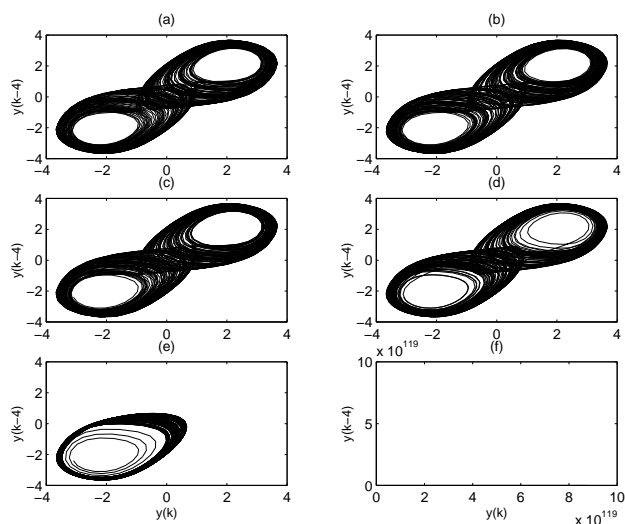


Figura 5.11: Atratores de modelos do circuito de Chua - com ruído

Projeção dos atratores dupla-volta correspondentes aos modelos (a-f), veja Tabela 5.2 e Figura 5.10. Caso com ruído.

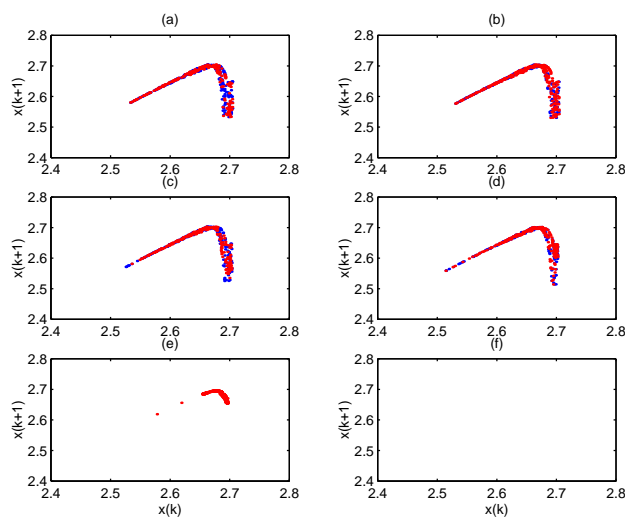


Figura 5.12: Mapa de primeiro retorno do circuito de Chua - com ruído

Mapa de primeiro retorno dos modelos (a-f), veja Tabela 5.2 e Figura 5.10. Caso com a presença de ruído.

Tabela 5.3: Comparação dos parâmetros do modelo original (5.3) com os seis modelos ($a-f$) do Pareto. Caso com ruído.

Termo \ Modelo	Original	(a)	(b)	(c)	(d)	(e)	(f)
$y(k-1)$	3,5230	2,3985	2,3978	2,3915	2,3791	2,3622	2,2315
$y(k-2)$	-4,2897	-0,8563	-0,8570	-0,8633	-0,8758	-0,8927	-1,0123
$y(k-4)^3$	-0,2588	0,9207	0,9198	0,9117	0,8957	0,8740	0,7119
$y(k-1)$	-1,7784	-0,6177	-0,6174	-0,6147	-0,6094	-0,6022	-0,5460
$y(k-3)$	2,0652	-1,3129	-1,3118	-1,3021	-1,2828	-1,2566	-1,0635
$y(k-1)^2 y(k-2)$	6,1761	0,7902	0,7973	0,8620	0,9907	1,1647	2,4643
$y(k-1)y(k-2)y(k-4)$	0,1623	-1,1704	-1,1617	-1,0822	-0,9243	-0,7107	0,8972
$y(k-2)^2 y(k-3)$	-2,7381	0,1008	0,0997	0,0900	0,0706	0,0444	-0,1517
$y(k-1)y(k-2)^2$	-5,5369	0,6808	0,6602	0,4731	0,1009	-0,4023	-4,1760
$y(k-2)^3$	0,1031	-1,1441	-1,1295	-0,9965	-0,7322	-0,3747	2,3084
$y(k-4)^3$	0,4623	-0,6120	-0,6101	-0,5924	-0,5573	-0,5099	-0,1560
$y(k-2)^2 y(k-4)$	-0,5247	1,3364	1,3231	1,2021	0,9614	0,6359	-1,8170
$y(k-1)y(k-3)^2$	-1,8965	-0,0471	-0,0488	-0,0643	-0,0952	-0,1369	-0,4549
$y(k-1)y(k-2)y(k-3)$	5,4255	0,3389	0,3411	0,3614	0,4019	0,4566	0,8734
$y(k-2)y(k-4)^2$	0,7258	-0,3270	-0,3227	-0,2833	-0,2049	-0,0988	0,7035
$y(k-3)y(k-4)^2$	-1,7684	1,4492	1,4432	1,3883	1,2792	1,1317	0,0246
$y(k-3)^2 y(k-4)$	1,1800	-0,8076	-0,8040	-0,7710	-0,7054	-0,6166	0,0502

5.3.5 Problema de decisão

Na seção anterior o melhor modelo foi escolhido segundo *avaliação a posteriori*. Questões qualitativas, tais como, o diagrama de bifurcação, espaço de fases e o maior expoente de Lyapunov serviram para apontar o modelo c , como o melhor modelo. Essa metodologia, mostra uma interação com um usuário, nesse caso, altamente especializado acerca do assunto. Isso nem sempre é possível. Dessa forma um outro método de decisão passa a ser desejável.

Para dar uma resposta a essas questões, nesta seção será aplicada a regra de decisão proposta por Teixeira (2001) para escolha do melhor modelo dentre aqueles mostrados no conjunto Pareto-ótimo, que consiste em uma abordagem para equilíbrio entre polarização e variância.

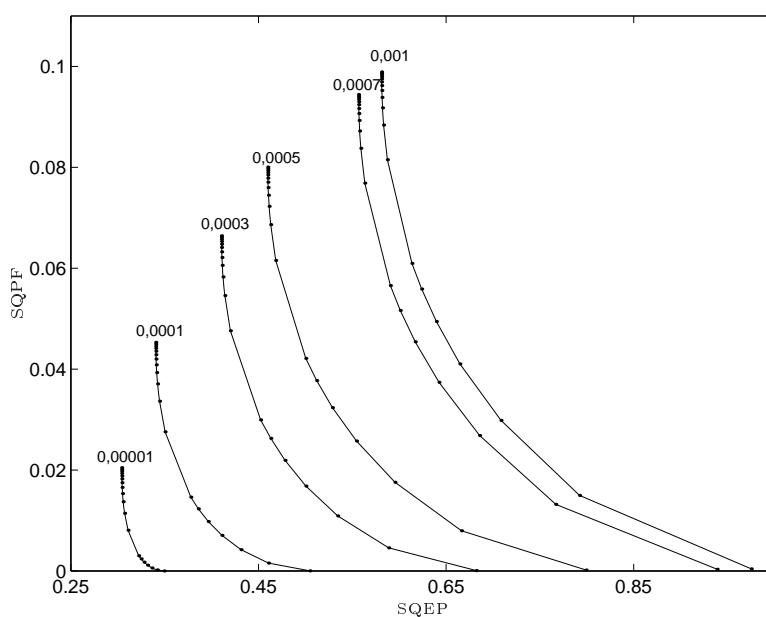


Figura 5.13: Pareto-ótimo e curvas de validação

Deslocamento do conjunto Pareto-ótimo em função do aumento da variância do ruído aditivo. SQEP: somatório do erro quadrático de predição (para melhor visualização, está apresentado o \log_{10} dos valores do eixo das abscissas SQPF: somatório do erro quadrático de localização dos pontos fixos.

A Figura 5.13 apresenta o deslocamento do conjunto Pareto-ótimo obtido

a partir dos dados de identificação, em função do aumento da variância do ruído aditivo. Fica evidente que quanto maior a variância do ruído, maior será a diferença entre os mínimos das funções objetivo, no caso, o erro quadrático de predição e o erro quadrático da localização dos pontos fixos. Para ilustração, foi escolhido o conjunto de validação com ruído aditivo de variância $\sigma^2 = 0,001$.

A Figura 5.14 apresenta a solução ótima do conjunto Pareto-ótimo. A Figura 5.14 (a) mostra o Pareto para os dados de treinamento e (b) apresenta as curvas de validação (V_1, V_2, V_3) com ruídos aditivos de variâncias (0,00001, 0,0001, 0,001), respectivamente. Observa-se que o ponto de mínimo na coordenada de SQEP de cada curva de validação em (b) não se altera em função da variância do ruído (Teixeira, 2001).

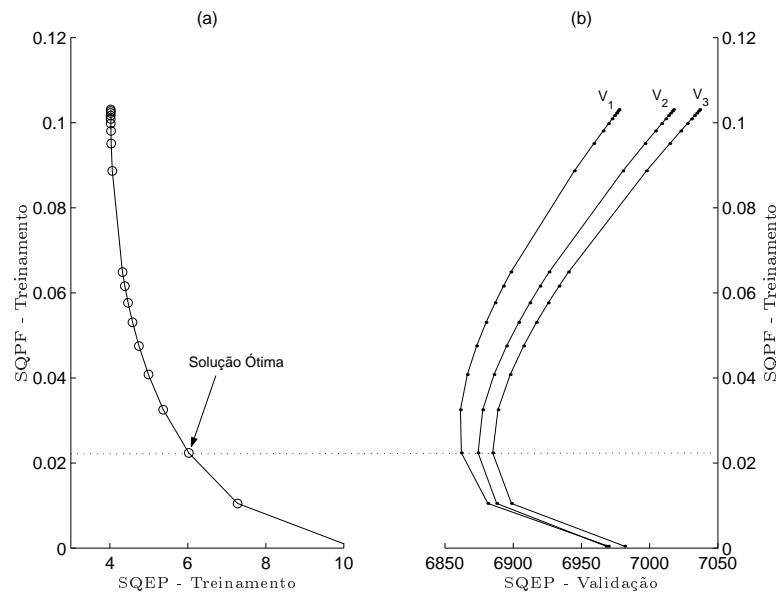


Figura 5.14: Pareto-ótimo e curvas de validação - circuito de Chua

(a) Conjunto Pareto-ótimo para o circuito de Chua com presença de ruído aditivo ($\bar{x} = 0$ e $\sigma^2 = 0,001$), utilizando dados de identificação. SQEP: Somatório do erro quadrático de predição (sendo \log_{10} para o eixo das abcissas). SQPF: Somatório do erro quadrático de localização dos pontos fixos. (b) apresenta as curvas de validação (V_1, V_2, V_3) com ruídos aditivos de variâncias (0,00001, 0,0001, 0,001), respectivamente.

A solução ótima encontrada atende aos requisitos de equilíbrio entre po-

larização e variância. Entretanto, as características dinâmicas estão longe de serem adequadas. O ML foi estimado em $\lambda = 1,0560 \pm 0,0048$ bits/s, originalmente o ML foi estimado em $\lambda = 3,6250 \pm 0,075$ bits/s. A Figura 5.15 mostra a projeção do atrator para a solução ótima, que não apresenta a dupla-volta e é praticamente periódica. Assim, é evidente que a regra de decisão empregada não foi capaz de levar em consideração propriedades dinâmicas importantes.

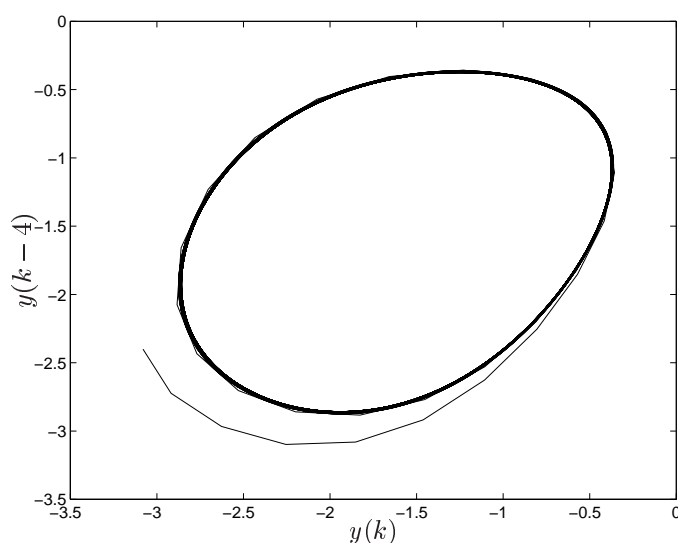


Figura 5.15: Atrator da solução ótima - circuito de Chua

Projecção do atrator (circuito de Chua) para a solução ótima obtida pela regra de decisão baseada no equilíbrio entre polarização e variância.

A Figura 5.16 ilustra o ponto de mínimo em relação ao erro de localização do ponto fixo. Esse ponto elucida que existe um modelo dentro do universo factível de modelos que equilibra a polarização e variância, independentemente da variância do ruído adicionado aos dados de validação.

5.3.6 Discussão

A incorporação de conhecimento *a priori* em um sistema caótico mostrou-se eficiente por uma série de motivos. Primeiramente, constituiu-se em uma forma de ajuste fino de modelos previamente identificados por metodologia ICP. Em segundo lugar, permitiu verificar que para situações com a presença

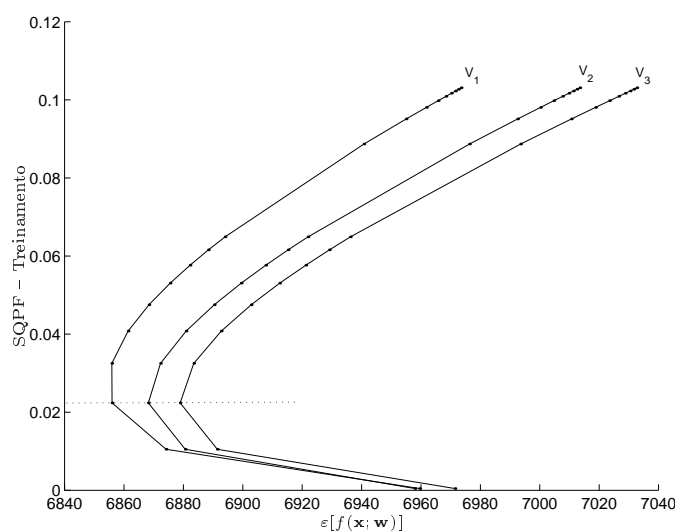


Figura 5.16: Diferença entre erro de predição com dados de treinamento e validação

Diferença entre o erro quadrático de predição para dados de treinamento e os três conjuntos de dados de validação.

de ruído - que de certo modo melhor explica situações práticas - a metodologia empregada passou ainda a ter mais relevância.

Nos casos estudados 1 e 2, o melhor modelo foi escolhido através da metodologia de *apresentação de preferência a posteriori*. O conjunto de soluções apresentadas permitiu ao usuário - que nesse caso, necessitou ser um especialista - a escolha do melhor modelo através de uma análise qualitativa. Também foi mostrado, que os termos de ERR mais elevado variam menos ao longo do Pareto do que os demais termos. Em seguida, foi mostrado um decisor baseado no equilíbrio entre a polarização e variância, fornecendo diretamente a melhor solução. A solução ótima encontrada atendeu ao equilíbrio de polarização e variância, mas não apresentou propriedades dinâmicas satisfatórias.

Não se pode esquecer que a qualidade da informação auxiliar vai, provavelmente, sofrer também com o aumento do ruído. Entretanto, foi mostrado em (Aguirre e Souza, 1998) que a estimação de pontos fixos é um procedimento bastante robusto na presença de ruído, tornando esse tipo de informação particularmente interessante.

5.4 Mapa senoidal

5.4.1 Descrição de mapas unidimensionais

Tem sido observado que mesmo os mais simples modelos - os mapas unidimensionais⁴ são capazes de explicar o comportamento de sistemas físicos relevantes (Artuso et al., 1993; Weibert e Main, 2001). Além disso, são modelos de fácil manipulação, comparados a sistemas de equações diferenciais, colaborando para o estudo de dinâmicas não-lineares.

Para este exemplo foi escolhido o mapa senoidal⁵ que apresenta não-linearidades cúbicas:

$$x(k) = \alpha \sin(x(k-1)), \quad (5.4)$$

sendo $\alpha = 1, 2\pi$. Para a condição inicial $x(0) \in [-\pi, \pi]$, a Equação 5.4 mapeia o intervalo $[-\pi, \pi]$ nele mesmo. Os pontos fixos são aproximadamente $\bar{y} = 0; \pm 2, 4383$ e o mapa tem o maior expoente de Lyapunov estimado em $\lambda = 1, 155 \pm 0, 009$ bits/s. A Figura 5.17 mostra o mapa de primeiro retorno que também é uma descrição de seu atrator.

5.4.2 Metodologia e resultados

O seguinte modelo de primeira ordem, usando técnicas mono-objetivo, foi estimado a partir de um conjunto de dados com 1000 pontos retirados do atrator mostrado na Figura 5.17

$$y(k) = 2, 6868y(k-1) - 0, 2462y(k-1)^3. \quad (5.5)$$

O modelo (5.5) tem os pontos fixos em $\bar{y} = 0; \pm 2, 6176$. O somatório

⁴Um mapa unidimensional é definido como $f : \mathbb{R} \rightarrow \mathbb{R}$, $x_{n+1} = f(x_n)$, sendo que o índice n indica que o sistema é discreto.

⁵Outros mapas unidimensionais comumente encontrados na literatura são:

- Mapa logístico:

$$x_{n+1} = ax_n(1 - x_n) \text{ para } 0 \leq a \leq 4$$

- Mapa de deslocamento:

$$x_{n+1} = 2x_n \text{ mod } 1$$

- Tent map:

$$x_{n+1} = \begin{cases} ax & x < 0,5 \\ a(1-x) & x > 0,5 \end{cases}$$

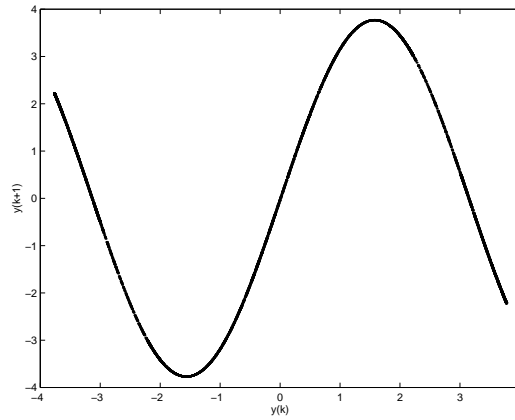


Figura 5.17: Mapa de primeiro retorno - mapa senoidal

Mapa de primeiro retorno para o mapa senoidal, Equação (5.4) com $\alpha = 1, 2\pi$.

do erro quadrático de predição é $0,2319$ e o ML é dado por $\lambda = 1,1458 \pm 3,4545 \times 10^{-5}$ bits/s.

Incorporando o conhecimento de pontos fixos por meio de (4.22) durante a estimação de parâmetros, vinte modelos foram identificados. As soluções do *Pareto-ótimo* são mostradas na Figura 5.18. Vale a pena mencionar que foi necessário variar muito finamente o parâmetro w na vizinhança de 1 a fim de produzir um conjunto significativo de soluções. Os mesmos índices usados anteriormente foram usados e são mostrados na Tabela 5.4.

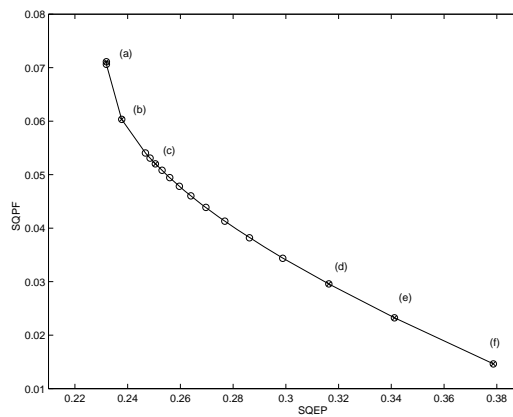


Figura 5.18: Pareto-ótimo - mapa senoidal

Soluções do Pareto-ótimo para o mapa senoidal com $\alpha = 1, 2\pi$.

Tabela 5.4: Algumas propriedades estáticas e dinâmicas de seis modelos para o mapa senoidal (5.4). Legenda: (i) Peso usado no problem P_w ; (ii) somatório do quadrado dos erros de predição (SQEP); (iii) somatório do quadrado dos erros da localização dos pontos fixos (SQPF); (iv) a localização dos pontos fixos; (v) e o maior expoente de Lyapunov (ML).

Modelo	(a)	(b)	(c)
Peso	0	0,99998207821415	0,99999141427971
SQEP	0,2319	0,2377	0,2506
SQPF	0,0711	0,0603	0,0520
Pontos Fixos	0; $\pm 2,6116$	0; $\pm 2,5854$	0; $\pm 2,5652$
ML	$1,0435 \pm 3,8190 \times 10^{-5}$	$-0,0249 \pm 6,1656 \times 10^{-5}$	$1,0225 \pm 1,0801 \times 10^{-5}$
Modelo	(d)	(e)	(f)
Peso	0,9999777856618	0,9999848570829	0,9999919285389
SQEP	0,3163	0,3412	0,3787
SQPF	0,0296	0,0232	0,0146
Pontos Fixos	0; $\pm 2,5103$	0; $\pm 2,4949$	0; $\pm 2,4739$
ML	$1,0264 \pm 9,1272 \times 10^{-6}$	$1,0223 \pm 2,5611 \times 10^{-5}$	$1,0096 \pm 1,4385 \times 10^{-5}$

5.4.3 Problema de decisão

A Figura 5.19 mostra os mapas de primeiro retorno para os modelos $(a-f)$. Observa-se que à medida em que os pontos fixos tendem àqueles do sistema original (indicado pelo cruzamento da linha $y = x$ com o mapa de primeiro retorno) o mapa de primeiro retorno dos modelos vão se aproximando do sistema original, particularmente para valores de $y(k)$ superiores a 2 e inferiores a -2 . Nesse caso, o benefício de se utilizar o conhecimento da localização dos pontos fixos torna-se óbvio.

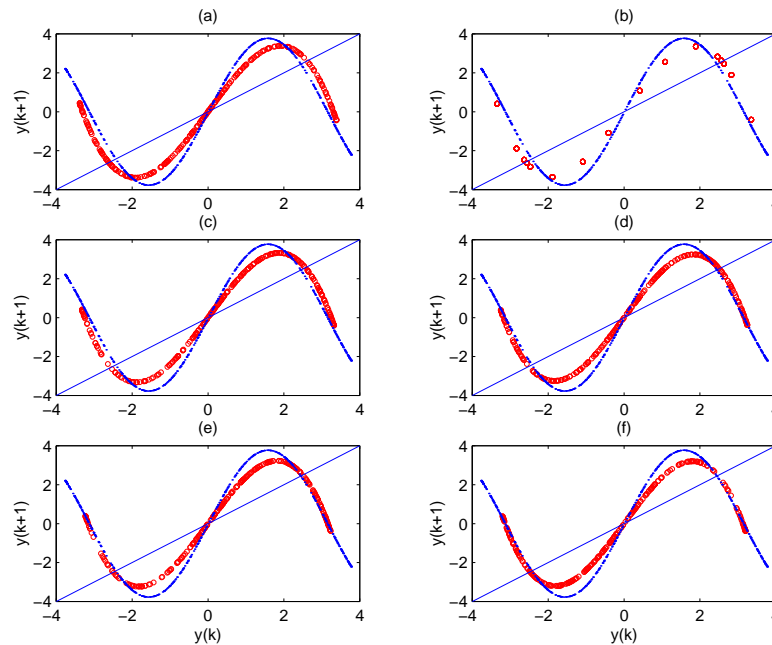


Figura 5.19: Mapa de primeiro retorno de seis modelos do mapa senoidal

Mapa de primeiro retorno de seis modelos $(a-f)$ para o mapa senoidal. (\dots) sistema original; (\circ) modelos, $(-)$ $y(k) = y(k+1)$. Veja Tabela 5.4. Os pontos fixos estão localizados nos cruzamentos entre os mapas de primeiro retorno e a linha $y = x$.

Uma avaliação qualitativa dos resultados pode apontar o modelo (d) como a solução ótima, pois apresenta um equilíbrio entre o erro da locação de pontos fixos e de predição, além de apresentar um mapa de primeiro retorno semelhante ao do sistema original.

5.4.4 Discussão

O modelo b apresenta um comportamento interessante. Embora tal modelo tenha os pontos fixos não demasiado distantes daqueles do sistema original, o atrator desse modelo revela dinâmica periódica ao invés de caos. A fim de investigar melhor essa característica peculiar, a Figura 5.20 mostra o histograma dos seis modelos. O modelo b não apresenta uma distribuição contínua e, como esperado, não tem um expoente de Lyapunov positivo. Assim torna-se evidente que uma *bifurcação* ocorreu ao longo do Pareto, à medida que o peso w foi variado. Na verdade, verificando a vizinhança do modelo b , considerando o erro da localização do ponto fixo como o parâmetro de bifurcação, encontra-se a suposta bifurcação, apresentada na Figura 5.21. O comportamento do ML é mostrado na Figura 5.22. De fato, o parâmetro de bifurcação é realmente o peso, w , mas há uma relação direta entre w e a soma de erros quadráticos dos pontos fixos, usada para traçar o diagrama da bifurcação.

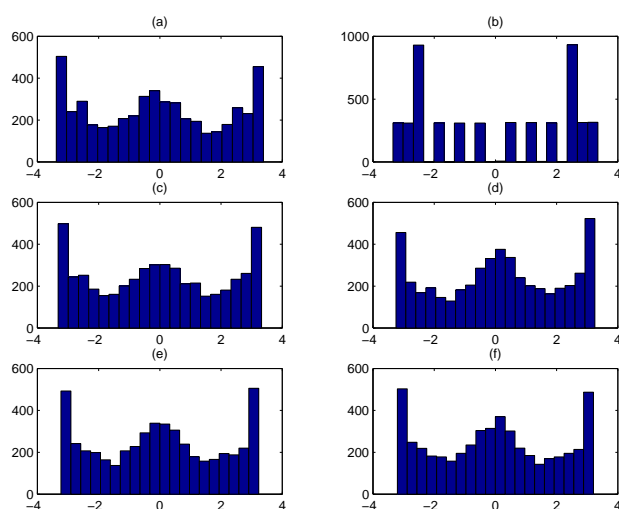


Figura 5.20: Histograma de dados gerados por modelos do mapa senoidal

Histograma que mostra a distribuição no espaço de dados produzidos pelos seis modelos ($a-f$) do mapa senoidal. Veja Tabela 5.4.

Assim, uma outra consideração importante pode ser destacada: o procedimento multiobjetivo define uma maneira sistemática para procurar soluções no espaço dos modelos. A abordagem mono-objetivo conduz a somente um ponto em tal espaço para cada estrutura do modelo (um conjunto finito de

soluções). A abordagem multiobjetivo permite uma busca contínua no espaço de soluções factíveis. Isto pode ser importante em casos quando o procedimento padrão de MQ não produzir modelos com propriedades desejadas em princípio, mas que produzam soluções na vizinhança de um modelo adequado (Aguirre et al., 2002).

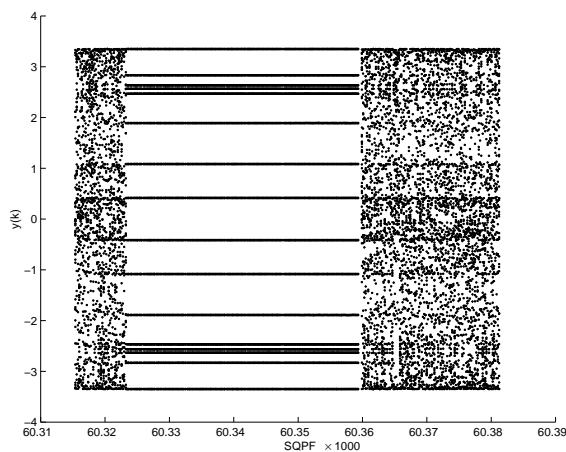


Figura 5.21: Diagrama de bifurcação de modelo do mapa senoidal

Diagrama de bifurcação na vizinhança do erro de localização do ponto fixo = 0,06035, usado como parâmetro de bifurcação.

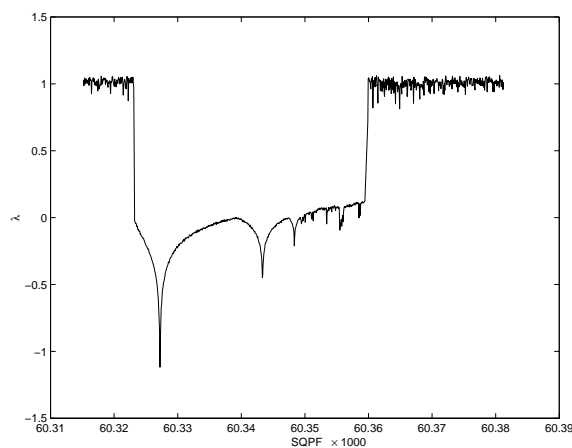


Figura 5.22: Variação do expoente de Lyapunov - mapa senoidal

Variação do maior expoente de Lyapunov (ML) na vizinhança do erro de ponto fixo = 0,06035.

5.5 Conclusões do capítulo

Esse capítulo discutiu o uso do conhecimento de pontos fixos na identificação multiobjetivo de sistemas não-lineares.

A metodologia apresentada no Capítulo 4 foi aplicada a dois sistemas caóticos: circuito de Chua e o mapa senoidal. No primeiro caso, muito pouco parece ser ganho no exemplo ideal de dados isentos de ruídos. Entretanto, quando ruído branco é adicionado aos dados de identificação, o uso de informação auxiliar resulta em modelos que reproduzem melhor a característica estática. Além disso, foi observado que a estimação dos pontos fixos é mais insensível ao ruído do que a estimação dos parâmetros do modelo (Aguirre e Souza, 1998). Uma explicação possível para isso é que ruído corrompe a informação (dinâmica e estática) presente nos dados. Isso pode ser compensado, ao menos em certa medida, por meio do uso de informação auxiliar. O melhor modelo foi escolhido seguindo duas estratégias: *apresentação de preferências a priori* e *apresentação de preferências a posteriori*. Para a primeira estratégia, necessitou-se de um conhecimento maior a respeito de características qualitativas do sistema, tais como, espaço de fases, expoentes de Lyapunov e diagramas de bifurcação. Para a segunda estratégia, apresentada em (Teixeira, 2001), a melhor solução é aquela que apresenta um equilíbrio entre polarização e variância. Contudo, a solução obtida não apresentou características dinâmicas satisfatórias, mostrando que um modelo que possui adequado equilíbrio entre polarização e variância não necessariamente possuirá boas características dinâmicas.

No exemplo do mapa senoidal, mesmo na ausência total de ruído, os modelos identificados não têm necessariamente um conjunto bom de pontos fixos, principalmente devido ao fato de que as estruturas do modelo usadas tinham não-linearidades cúbicas, enquanto o sistema original possui não-linearidade senoidal. O uso da informação da localização dos pontos fixos nesse exemplo resultou em modelos com propriedades dinâmicas melhores em termos do mapa de primeiro retorno e ligeiramente pior em relação ao maior expoente de Lyapunov. Nesse exemplo, tornou-se também claro que o peso, w , usado para produzir o conjunto Pareto pode ser considerado como um parâmetro de bifurcação, que poderia ser usado para procurar sistematicamente por modelos bons na vizinhança do modelo que não incorpora nenhuma espécie de informação auxiliar, superando, assim, problemas de outras abordagens. Isto tem conseqüências importantes na validação do modelo, pois são encontra-

dos freqüentemente modelos que, à primeira vista, não são dinamicamente válidos, mas que estão na vizinhança de modelos bons. Nessa situação, como podemos sistematicamente procurar por modelos bons que partem de um modelo inicial (Aguirre et al., 2002)? O procedimento apresentado aqui parece ser uma alternativa adequada.

Uso da Informação da Curva Estática

“O que ouço, esqueço. O que vejo, recordo. O que faço, compreendo.”

Confúcio

6.1 Introdução

Este capítulo tem como objetivo aplicar a metodologia desenvolvida no Capítulo 4, sendo utilizada a curva estática como informação auxiliar.

Dois sistemas serão estudados: um aquecedor elétrico (Cassini, 1999) e um conversor CC-CC buck (Aguirre et al., 2000).

Para o caso do aquecedor elétrico, os dados de identificação excursionam em uma ampla faixa de operação. Essa é uma situação mais favorável, mas importante para avaliar a metodologia.

Na seção seguinte o conversor CC-CC buck será identificado. Para esse sistema, os dados de identificação excursionam apenas em uma faixa de operação, tornando a identificação de um modelo global mais difícil.

Os sistemas estudados são descritos brevemente, apresentando os testes estáticos e dinâmicos para obtenção de identificação e validação.

6.2 Aquecedor elétrico

6.2.1 Descrição do sistema

O aquecedor elétrico em estudo nessa seção foi montado em conjunto com ventiladores que podem estar ligados ou desligados. Dessa forma, pode-se variar a dissipação desse aquecedor através dos ventiladores. A temperatura do aquecedor é medida por um termopar e a relação em estado estacionário entre potência elétrica na entrada e temperatura final depende do estado dos ventiladores (Cassini, 1999). A Figura 6.1 apresenta um diagrama em blocos funcionais do aquecedor com dissipação variável (Cassini, 1999).

A entrada e saída do sistema são a tensão elétrica no divisor de tensão e a tensão de saída do circuito amplificador, respectivamente. O divisor de tensão é construído de maneira que uma variação na faixa (0 a 127 V) corresponda a uma variação de (0 a 5 V) na entrada do sistema. Os dados de entrada e saída foram coletados através de uma placa de aquisição de dados (PCL 711s).

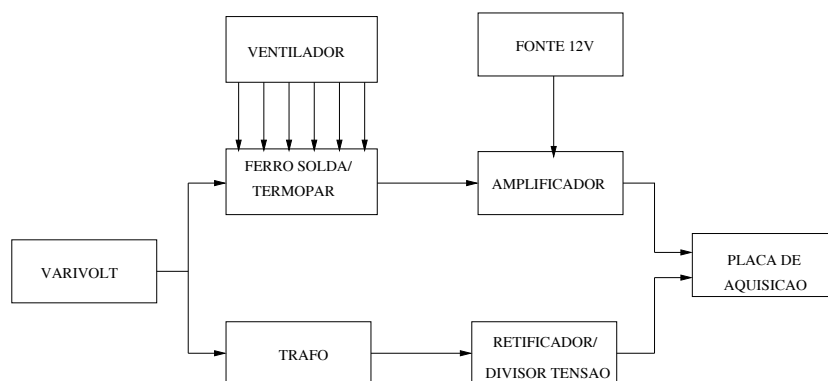


Figura 6.1: Diagrama em blocos funcionais do aquecedor elétrico

Diagrama em blocos funcionais do aquecedor com dissipação variável. O aquecedor com dissipação variável consiste em um pequeno aquecedor elétrico e um conjunto de ventiladores que podem estar ligados ou desligados.

A placa de aquisição de dados utilizada tem como limites de tensão, para coleta, as faixas (0 a 5V) e (0 a 10V). O conjunto funcional transformador-

retificador-divisor de tensão, tem como função garantir a manutenção dos limites de (0 a 5V).

O *varivolt* é utilizado para tornar possível a variação da tensão de entrada na faixa de (0 a 136V) da planta (ferro de solda). Esse sinal passa pelo transformador, passando para a faixa de (0 a 18V), e o sinal de saída é retificado.

As variações na temperatura do ferro de solda provocadas pela variação no *varivolt* são medidas por meio de um termopar acoplado ao mesmo. A tensão nos terminais do termopar são ampliadas por um amplificador de instrumentação para uma faixa de (0 a 4V) e coletadas pela placa de aquisição de dados. Essa planta piloto foi utilizada recentemente em estudos de Identificação de Sistemas (Corrêa, 2001; Barroso, 2001; Amaral, 2001; Coelho, 2002).

6.2.2 Teste dinâmico

Durante os testes que serão aqui considerados, o ventilador encontrava-se ligado em uma fonte de 127 V.

Esses testes foram realizados em uma temperatura ambiente de 24°C por um período de 2 h. Os dados foram coletados com uma taxa de amostragem de 6s, resultando num total de 2520 pontos para o teste com ventilador ligado. Cada patamar do sinal de excitação do sistema foi mantido por 1 min, gerando aproximadamente 10 amostras por patamar.

Os sinais utilizados nesse trabalho foram decimados por um fator de 10 para que os dados de identificação respeitassem os limites da Equação (2.4). A massa de dados foi dividida em duas, sendo a primeira metade utilizada para a identificação do sistema e a segunda metade para a validação do mesmo. Os dados utilizados na identificação podem ser vistos na Figura 6.2 e os dados utilizados na validação na Figura 6.3.

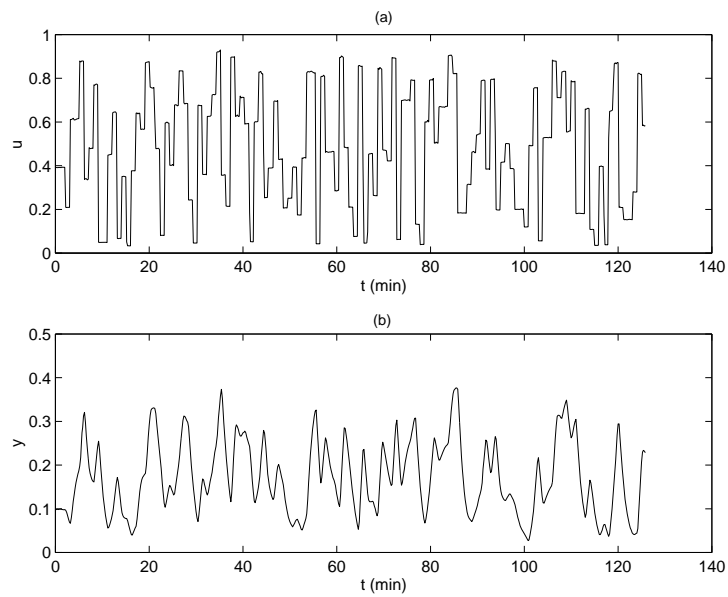


Figura 6.2: Dados de identificação - aquecedor elétrico

Dados de identificação do aquecedor elétrico (primeira metade), sendo: (a) sinal de entrada do sistema em p.u, (b) sinal de saída do sistema em p.u.

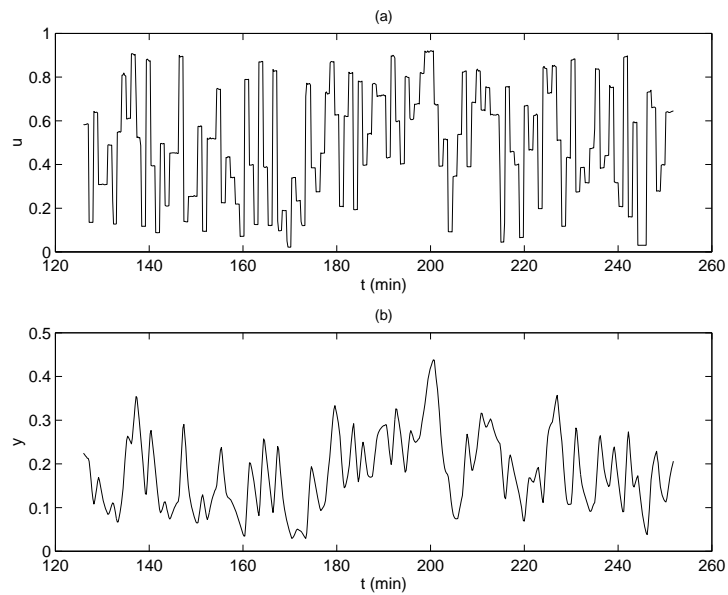


Figura 6.3: Dados de validação - aquecedor elétrico

Dados de validação do aquecedor elétrico (segunda metade), sendo: (a) sinal de entrada do sistema em p.u, (b) sinal de saída do sistema em p.u.

6.2.3 Teste estático

Os testes estáticos têm como finalidade o levantamento da curva estática do sistema. O teste utilizado nesse trabalho teve a duração de 3 h.

A Figura 6.4 mostra a curva estática do modelo. As grandezas estão indicadas em p.u., cuja base é 5V do sinal de entrada, visto pelo ferro de solda, e 998,51°C na temperatura do ferro de solda.

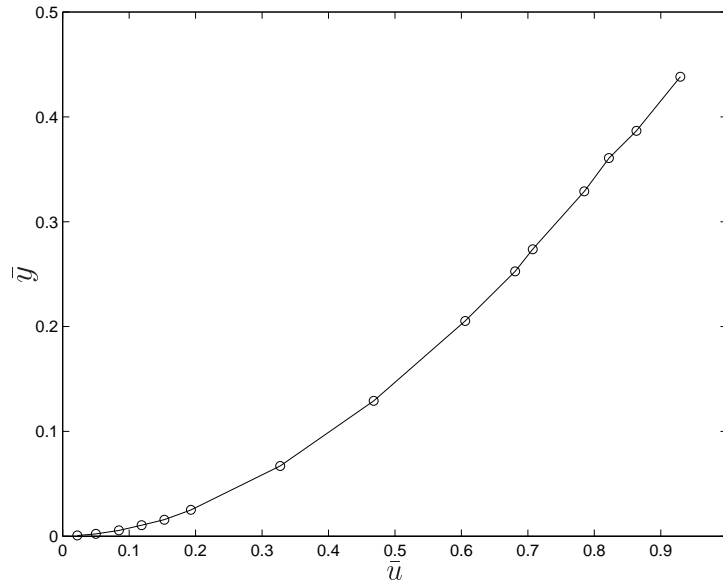


Figura 6.4: Dados estáticos - aquecedor elétrico

Curva estática do aquecedor elétrico, sendo: (\bar{u}) e (\bar{y}) em p.u. Os círculos determinam as medições realizadas. A linha foi traçada para facilitar a visualização do comportamento da curva estática.

6.2.4 Resultados

Utilizou-se a estrutura proposta em (Barroso, 2001, p. 65) acrescentando a parte MA (média móvel) $\Psi_{\xi}^T(k-1)\theta_{\xi}$, composta de cinco termos de ruído.

$$\begin{aligned}
 y(k) = & \theta_1 y(k-1) + \theta_2 u(k-2)u(k-1) + \theta_3 u(k-1)^2 \\
 & + \theta_4 y(k-2) + \theta_5 u(k-2)y(k-1) \\
 & + \theta_6 u(k-2)y(k-2) + \theta_7 u(k-2)^2 \\
 & + \Psi_{\xi}^T(k-1)\theta_{\xi}.
 \end{aligned} \tag{6.1}$$

A Figura 6.5 mostra o conjunto Pareto-ótimo obtido a partir da incorporação de conhecimento *a priori* da curva estática na estimação de parâmetros do modelo (6.1), de acordo com a metodologia proposta no Capítulo 4. As Figuras 6.6 e 6.7 mostram as simulações dinâmicas e estáticas dos modelos de (a-f), respectivamente. A Tabela 6.1 mostra os valores de RMSE estático e dinâmico para os modelos (a-f).

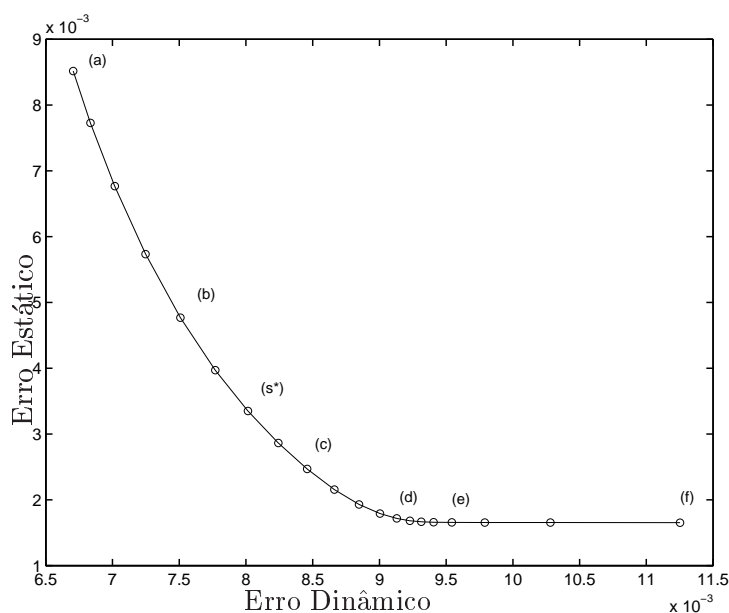


Figura 6.5: Pareto-ótimo - aquecedor elétrico

Conjunto Pareto-ótimo obtido a partir da incorporação de conhecimento *a priori* da curva estática na estimação de parâmetros do modelo (6.1) para o aquecedor elétrico. Cada (o) significa um modelo. No eixo das abcissas está o Erro Dinâmico que é o RMSE da predição livre, enquanto no eixo das ordenadas encontra-se o Erro Estático que é o RMSE da localização dos pontos fixos. As Figuras 6.6 e 6.7 mostram as simulações dinâmicas e estáticas, respectivamente, dos modelos de (a-f). O modelo (s*) é escolhido pelo critério de norma mínima dos objetivos normalizados e suas simulações estáticas e dinâmicas estão mostradas nas Figuras 6.8 e 6.9.

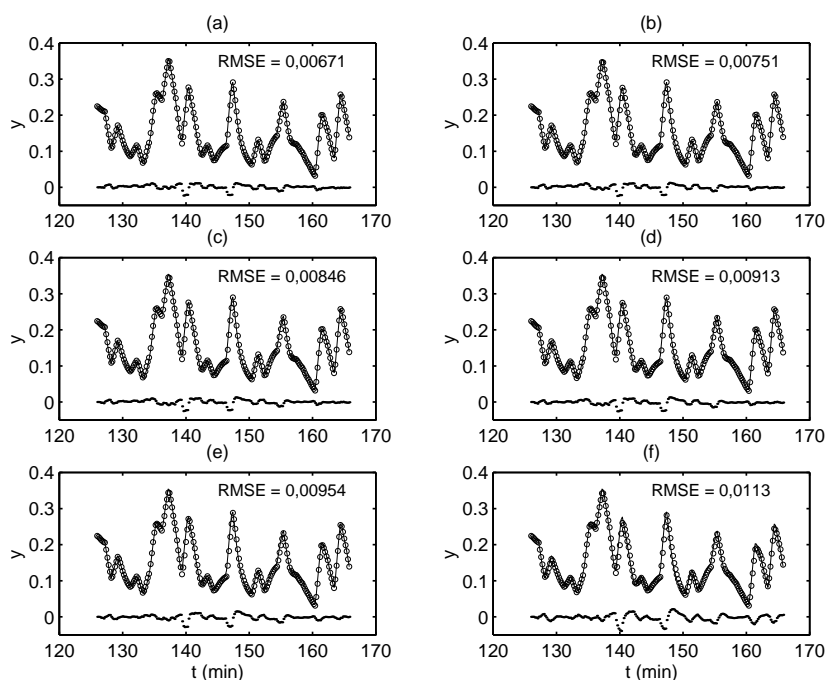


Figura 6.6: Validação dinâmica - aquecedor elétrico

Validação dinâmica dos modelos (a-f) para aquecedor elétrico retirados do conjunto Pareto-ótimo da Figura 6.5, sendo: (-) dados reais, (o) a resposta do modelo e (·) sinal de erro.

Tabela 6.1: RMSE da predição livre e da localização dos pontos fixos computados dos modelos (a-f) retirados da Figura 6.5 (aquecedor elétrico).

Modelo	(a)	(b)	(c)
Peso (w)	1,0000e-010	8,5616e-001	9,7931e-001
Erro dinâmico	0,00671	0,00750	0,00845
Erro estático	0,00852	0,00477	0,00247
Modelo	(d)	(e)	(f)
Peso (w)	9,9702e-001	9,9957e-001	9,9990e-001
Erro dinâmico	0,00913	0,00954	0,01125
Erro estático	0,00171	0,00166	0,00165

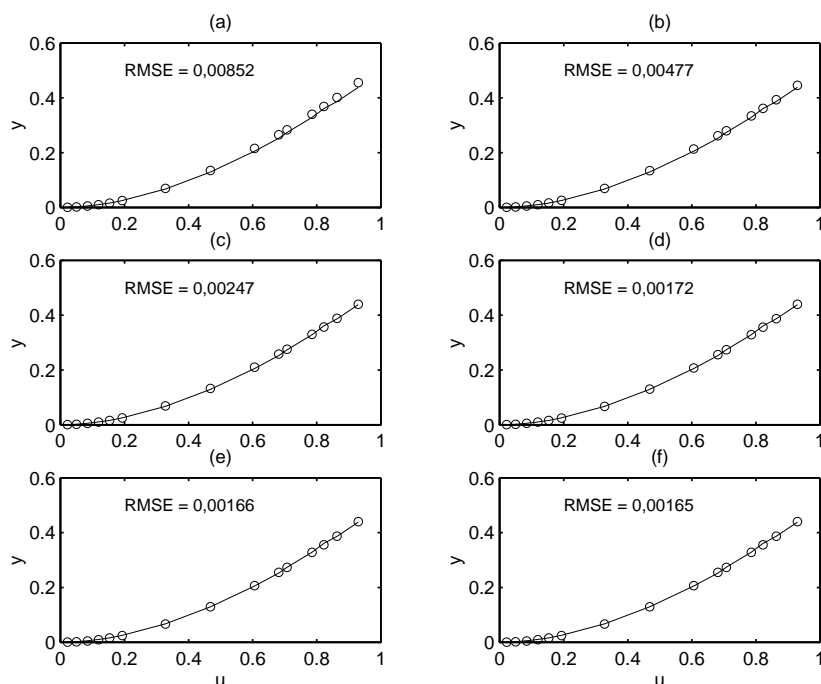


Figura 6.7: Validação estática - aquecedor estático

Validação estática dos modelos (a-f) para o aquecedor elétrico retirados do conjunto Pareto-ótimo da Figura 6.5, sendo: (-) dados reais e (o) a resposta do modelo.

A Figura 6.6 (a) apresenta um RMSE de 0,00671, utilizando $w = 1e - 10$. Nessas condições o efeito da utilização da informação da curva estática é desprezível, conseqüentemente, os resultados são praticamente idênticos aos obtidos por MQ. Para efeito comparativo, Barroso (2001) obteve um RMSE de 0,0094. A pequena melhora atingida deve-se ao fato de que o modelo (6.1) foi acrescido da média móvel (MA).

6.2.5 Problema de decisão

Parece uma atitude óbvia perguntar qual o melhor modelo dentro do conjunto *Pareto ótimo*. Em Engenharia, por exemplo, poder-se-ia estar interessado em projetar um controlador a partir de um melhor modelo. Entretanto, apesar de atraente, a escolha do melhor modelo a partir de um conjunto Pareto-ótimo é uma tarefa subjetiva. Chankong e Haimes (1983) apresentam algumas técnicas que abordam esse problema, chamado *problema de*

decisão, conforme visto na Seção 4.6.

Para esse exemplo foi empregado o critério de decisão que minimiza a *norma euclidiana dos objetivos normalizados*, conforme visto na Seção 4.6.2. Foi encontrado o modelo s^* , indicado na Figura 6.5. As Figuras 6.8 e 6.9 mostram a simulação dinâmica e estática do melhor modelo.

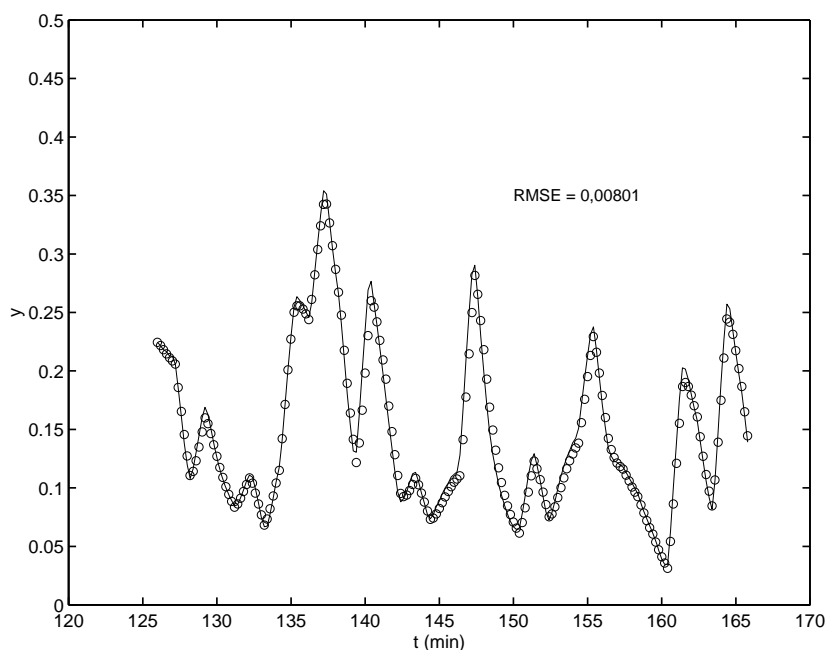


Figura 6.8: Validação dinâmica do melhor modelo - aquecedor elétrico

Validação dinâmica do melhor modelo para o aquecedor elétrico segundo critério da norma mínima dos objetivos normalizados. Sendo que (—) são os dados reais e (○) a resposta do modelo.

6.2.6 Discussão

A incorporação da informação da curva estática, conforme Seção 4.5.2, pode ser considerada uma generalização da incorporação da informação a respeito dos pontos fixos. O procedimento utilizado consiste em montar a matriz de regressores estáticos, Q , com os dados de entrada e saída da curva estática e utilizar o mapeamento linear $S\hat{\theta}$, Equação (4.24). A curva estática é descrita como a Equação (4.25), reescrita a seguir para conveniência do

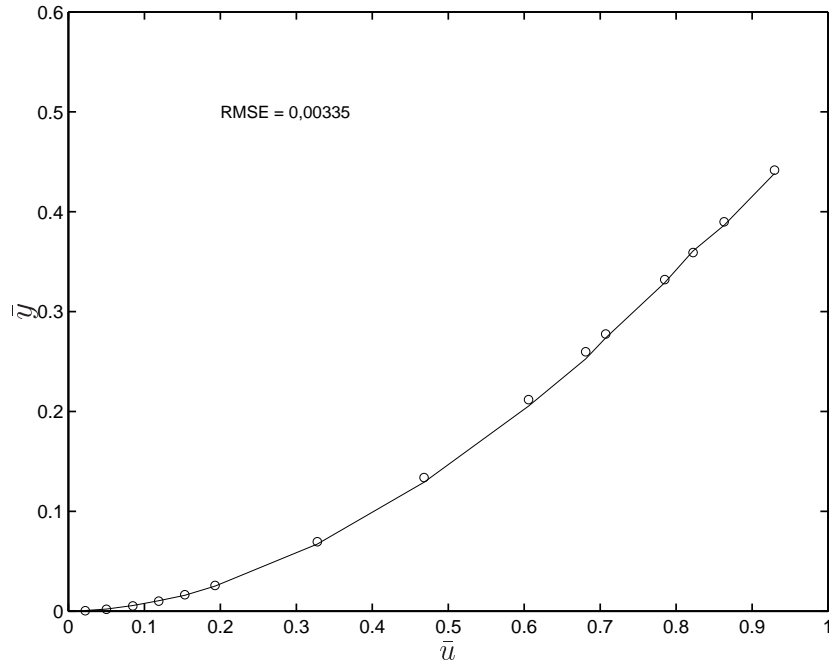


Figura 6.9: Validação estática do melhor modelo - aquecedor elétrico

Validação estática do melhor modelo para o aquecedor elétrico segundo critério da norma mínima. Sendo que (–) são os dados reais e (o) a resposta do modelo.

leitor.

$$\hat{\bar{y}} = QS\hat{\theta}.$$

Facilmente, encontra-se a função custo que relaciona a curva estática real e a estimada: $J_{CE}(\hat{\theta}) = (\bar{y} - \hat{\bar{y}})^T (\bar{y} - \hat{\bar{y}})$, Equação (4.26). Utilizando o Teorema 4.4.1, o vetor de parâmetros θ pode ser estimado de acordo com a Equação (4.27), reescrita a seguir:

$$\hat{\theta} = [w_1 \Psi^T \Psi + w_2 (QS)^T (QS)]^{-1} [w_1 \Psi^T y + w_2 (QS)^T \bar{y}].$$

Claramente, a Equação (4.26) é um problema convexo, tornando sua otimização trivial. Além disso, a estrutura matricial permite facilmente a implementação de modelos NARMAX (ver Pág. 66).

As características apresentadas são evidentes vantagens comparadas aos métodos propostos em (Corrêa, 2001; Barroso, 2001), em que é necessário

fazer um ajuste de curva polinomial para uma $f(\cdot)$ de tal forma a relacionar $\bar{y} = f(\bar{u})$ (Barroso, 2001, p. 46). Não se tem garantias de que tal ajuste seja convexo, tornando essa metodologia um pouco mais complexa.

O critério de norma mínima usado para encontrar a solução ótima, mostrou-se eficiente e prático. A normalização dos objetivos permite, até certo ponto, colocá-los numa mesma faixa, permitindo que se efetue o cômputo da norma euclidiana. Comparando o RMSE dinâmico e estático da solução ótima com o modelo caixa-preta, percebe-se que houve uma depreciação de 20% no índice dinâmico (de 0,0080 para 0,0067) e uma melhoria de 400% no índice estático (de 0,0034 para 0,0085). Isso confirma a eficiência da metodologia empregada nesse caso.

6.3 Conversor buck

6.3.1 Descrição do sistema

Como segundo exemplo para uso de informação da curva estática no processo de identificação será usado um conversor estático CC-CC buck. Esse conversor consiste em um sistema de regulação de tensão de carga controlado por um MOSFET IRF840. Atuando-se na porta (G) da chave semicondutora, obtém-se uma tensão contínua de saída, v_o a partir de uma fonte de tensão contínua, constante, v_d . Um esquema representativo do sistema pode ser visto na Figura 6.10.

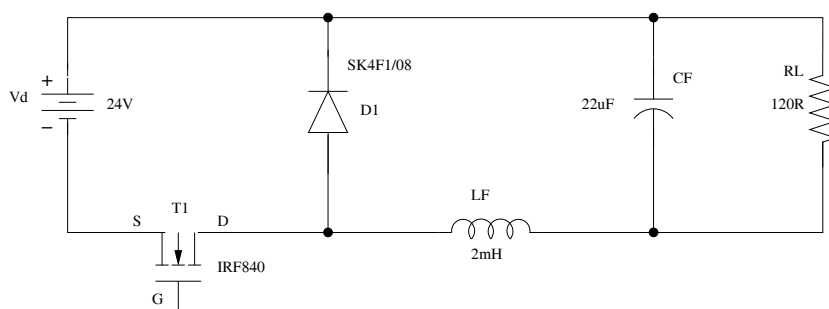


Figura 6.10: Conversor buck

Estrutura de um conversor estático CC-CC do tipo buck.

A razão cíclica é definida como a proporção de tempo em que a chave está ligada em relação ao período total de operação, T , ou, $D = T_{\text{ligado}}/T$.

Para variar essa razão, utilizaram-se técnicas de modulação por largura de pulso (PWM) a uma taxa de $1/T = 33\text{kHz}$, utilizando para tal um circuito integrado LM3524. A taxa de 33 kHz resultou em um modo de operação contínua, ou seja, a corrente através do indutor não se anulava.

Quando a razão cíclica está perto do valor unitário, a corrente através do indutor (LF) e a tensão na carga (RL), v_o , aumentam, pois a fonte v_d energiza a malha formada por ela, o capacitor (CF) e o indutor (LF). Quando $D \rightarrow 0$, a tensão v_o diminui com um regime dinâmico diferente. Esse fato caracteriza um regime dinâmico não-linear do sistema (Aguirre et al., 2000).

6.3.2 Teste dinâmico

Com a finalidade de excitar a dinâmica do conversor buck, utilizou-se um sinal do tipo PRBS.

Testes de resposta ao degrau foram realizados previamente com o intuito de levantar a constante de tempo predominante do sistema. A constante de tempo foi de aproximadamente 2 milissegundos, com um tempo de amostragem de $T_s = 10\mu\text{s}$. Partindo dessa informação foi construído o sinal PRBS.

Em sinais binários pseudo-aleatórios (PRBS), dois parâmetros são necessários para sua construção, o número de bits, b , e o tempo mínimo de chaveamento, T_b (Aguirre, 2000a).

O número mínimo de bits deve ser escolhido de forma a garantir que o período do sinal PRBS seja maior do que o tempo de acomodação do sistema, ou

$$2^b \times T_b > 5 \times 2 \text{ milissegundos}, \quad (6.2)$$

sendo o tempo de acomodação cinco vezes a constante de tempo dominante do sistema.

A partir da Equação (6.2), foi escolhido um número mínimo de 8 bits e $T_b = 670\mu\text{s}$ para gerar um sinal PRBS que fosse capaz de excitar a dinâmica do sistema.

Os sinais de identificação e validação foram superamostrados, com isso fez-se necessário decimar os sinais de um fator $\Delta = 12$, para que os sinais respeitassem a relação (2.4). Os sinais obtidos podem ser vistos nas Figuras 6.11 e 6.12.

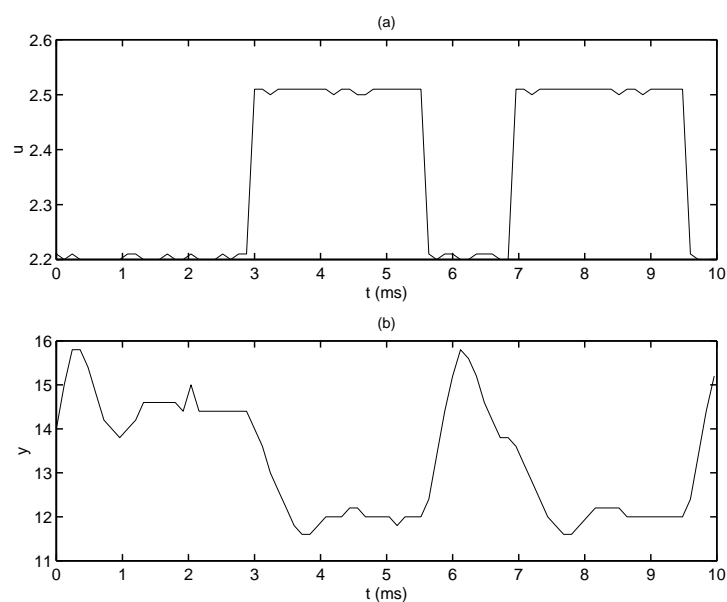


Figura 6.11: Dados de identificação - conversor buck

Dados de identificação do conversor buck (decimado de $\Delta = 12$ - primeira metade), sendo: (a) sinal de entrada do sistema, (b) sinal de saída do sistema.

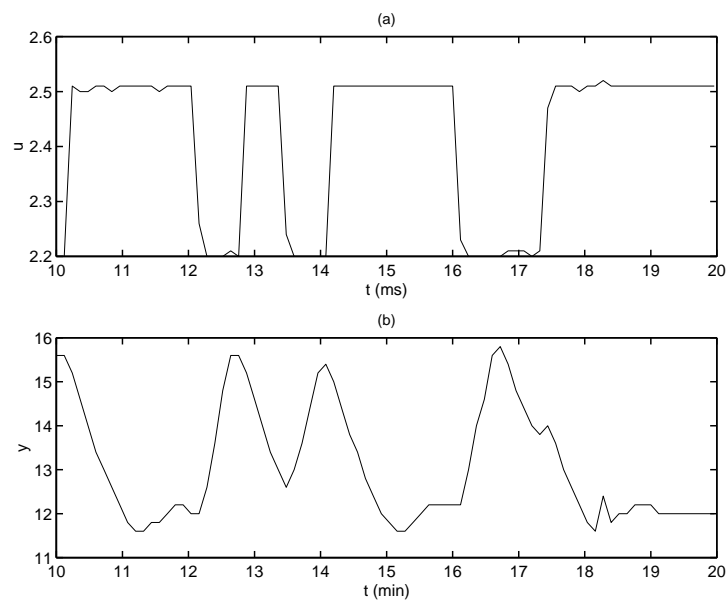


Figura 6.12: Dados de validação - conversor buck

Dados de validação do conversor buck (decimado de $\Delta = 12$ - segunda metade), sendo: (a) sinal de entrada do sistema, (b) sinal de saída do sistema.

6.3.3 Característica estática

A relação teórica entre a razão cíclica e a tensão em estado estacionário do conversor implementado é

$$\begin{aligned} V_o &= (1 - D)V_d \\ &= \left(1 - \frac{\bar{u} - 1}{3}\right) V_d \\ &= \frac{4V_d}{3} - \frac{V_d}{3}\bar{u}, \end{aligned} \quad (6.3)$$

sendo V_o a tensão na carga (RL), D é a razão cíclica, V_d a tensão constante de alimentação e \bar{u} é o valor em estado estacionário da entrada do modelo $u(k)$.

Os dados estáticos podem ser obtidos aplicando-se à Equação (6.3) valores de \bar{u} que abrangem todos os pontos de operação do sistema (1 a 4 V). A curva estática para o conversor buck pode ser vista na Figura 6.13. Observe que apenas na região delimitada pelas retas $(-\cdot-)$ possuem dados dinâmicos.

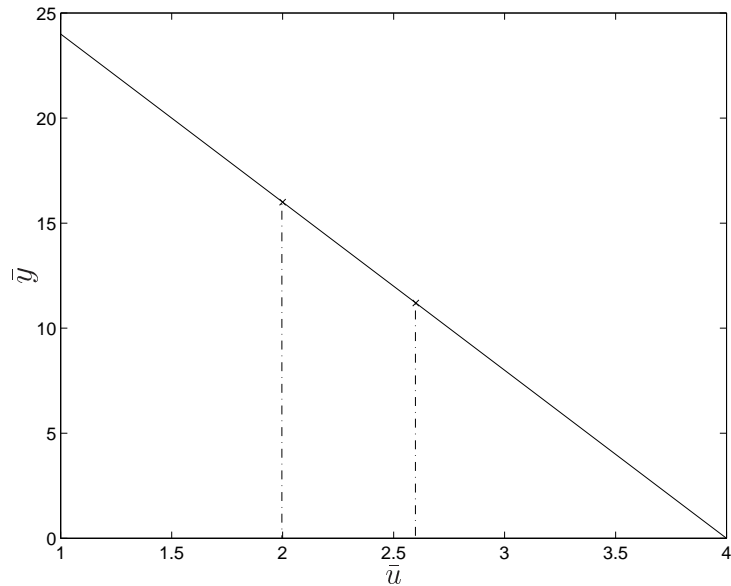


Figura 6.13: Dados estáticos - conversor buck

Curva estática do conversor buck, sendo os dados estáticos de entrada (\bar{u}) e de saída ($\bar{y} = V_o$) em volts. Os dados dinâmicos excursionam apenas a região delimitada pelas linhas $(-\cdot-)$.

Tabela 6.2: RMSE dinâmico e estático dos modelos (a-f) retirados da Figura 6.14 (conversor buck).

Modelo	(a)	(b)	(c)
Peso (w)	0,0500	0,3378	0,6782
Erro dinâmico	0,5916	0,6206	0,6486
Erro estático	0,9507	0,2120	0,1052
Modelo	(d)	(e)	(f)
Peso (w)	0,8436	0,9631	0,9990
Erro dinâmico	0,6669	0,6991	0,7214
Erro estático	0,0740	0,0296	0,0010

6.3.4 Resultados

A estrutura proposta em (Barroso, 2001, p. 86) foi utilizada para estimar os parâmetros utilizando a abordagem multiobjetivo descrita no Capítulo 4. No caso do conversor buck, foram acrescentados à estrutura a parte MA (média móvel) $\Psi_{\xi}^T(k-1)\theta_{\xi}$, composta de quatro termos de ruído

$$\begin{aligned}
 y(k) = & \theta_1 y(k-1) + \theta_2 y(k-2) + \theta_3 \\
 & + \theta_4 u(k-1) + \theta_5 u(k-1)^2 \\
 & + \theta_6 u(k-2)u(k-1) + \theta_7 u(k-2) \\
 & + \theta_8 u(k-2)^2 + \Psi_{\xi}^T(k-1)\theta_{\xi}.
 \end{aligned} \tag{6.4}$$

A Figura 6.14 mostra o conjunto Pareto-ótimo obtido a partir da incorporação de conhecimento *a priori* da curva estática na estimação de parâmetros do modelo (6.4). As Figuras 6.15 e 6.16 mostram as simulações dinâmicas e estáticas dos modelos de (a-f), respectivamente. O modelo s^* é uma possível escolha do melhor modelo e suas simulações estáticas e dinâmicas estão mostradas nas Figuras 6.17 e 6.18. A Tabela 6.2 mostra os valores de RMSE estático e dinâmico para os modelos (a-f).

6.3.5 Problema de decisão

No caso do conversor buck, o problema de decisão será abordado sob uma perspectiva subjetiva, sendo que o usuário final do modelo possui uma

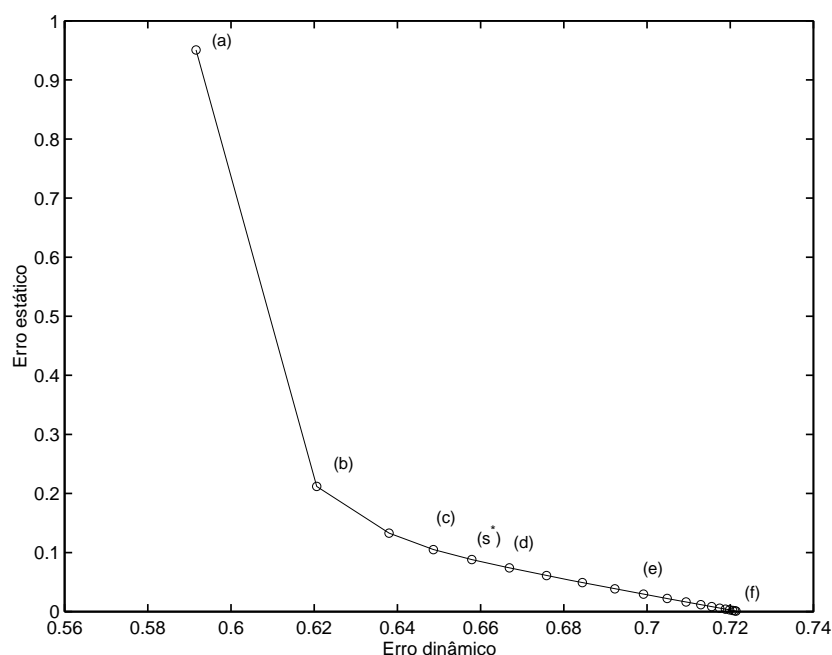


Figura 6.14: Pareto-ótimo - conversor buck

Conjunto Pareto-ótimo obtido a partir da incorporação de conhecimento *a priori* da curva estática na estimação de parâmetros do modelo (6.4) para o conversor buck. Cada (o) significa um modelo. No eixo das abcissas está o Erro Dinâmico que é o RMSE da predição livre, enquanto o eixo das ordenadas apresenta o Erro Estático que é o RMSE da localização dos pontos fixos. As Figuras 6.15 e 6.16 mostram as simulações dinâmicas e estáticas, respectivamente, dos modelos de (a-f). O modelo s^* foi escolhido por uma análise de preferência *a priori* e suas simulações estáticas e dinâmicas estão mostradas nas Figuras 6.17 e 6.18.

preferência *a priori*. Uma possível preferência poderia ser que o RMSE estático fosse inferior a 0,01. O modelo s^* representa exatamente essa exigência, como mostrado na Figura 6.14. As Figuras 6.17 e 6.18 mostram a resposta dinâmica e estática do modelo s^* , respectivamente.

6.3.6 Discussão

O caso do conversor buck é mais difícil do que o caso do aquecedor elétrico pois os dados dinâmicos excursionam apenas em uma faixa de operação

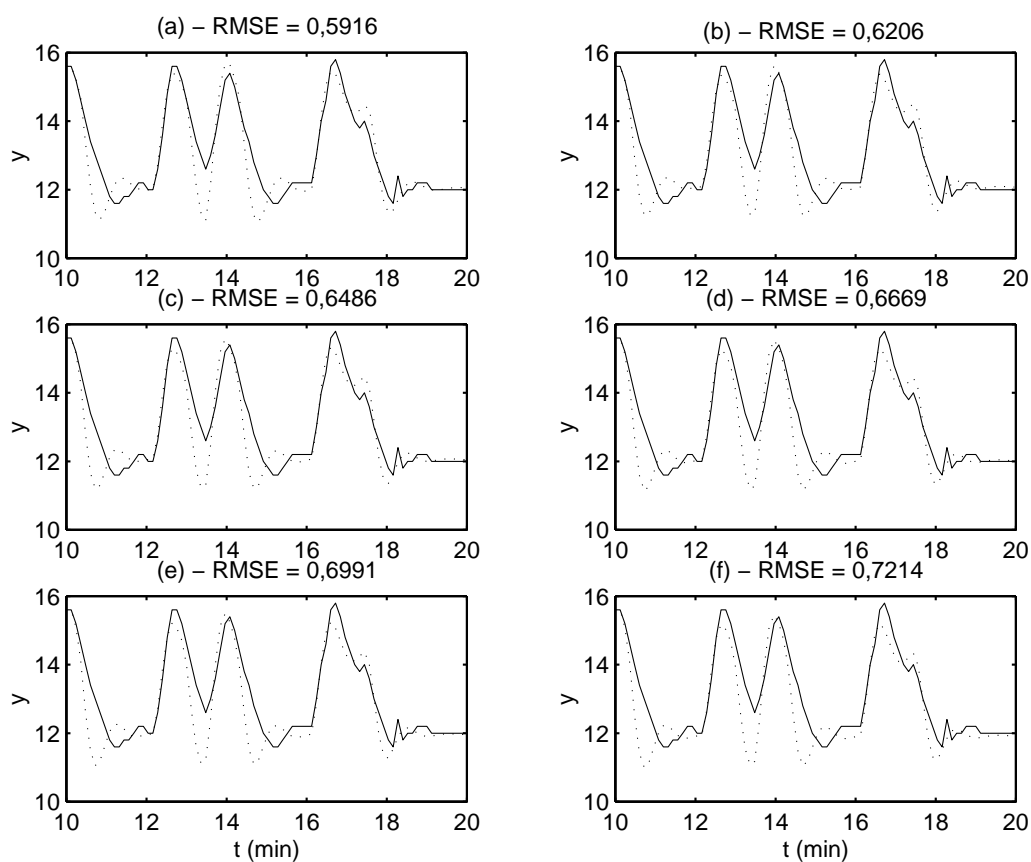


Figura 6.15: Validação dinâmica - conversor buck

Caso conversor buck - Validação dinâmica dos modelos (a-f) retirados do conjunto Pareto-ótimo da Figura 6.14, sendo: (-) dados reais e (·) a resposta do modelo.

relativamente estreita, conforme ilustrado na Figura 6.13.

A escolha da solução ótima, partindo de um caso hipotético de preferência *a priori*, mostrou a importância da abordagem multiobjetivo, pois permitiu ao usuário escolher o modelo que mais se adequasse às suas exigências, com o menor prejuízo para os outros objetivos.

6.4 Conclusões do capítulo

O presente capítulo mostrou duas aplicações de identificação multiobjetivo de sistemas não-lineares, utilizando a curva estática como informação

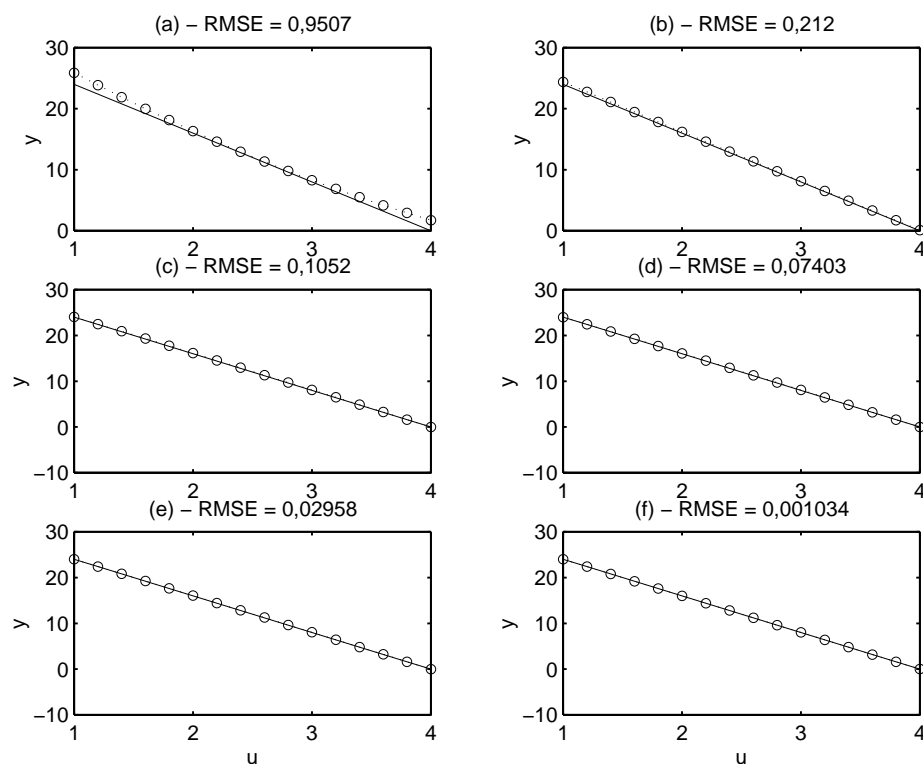
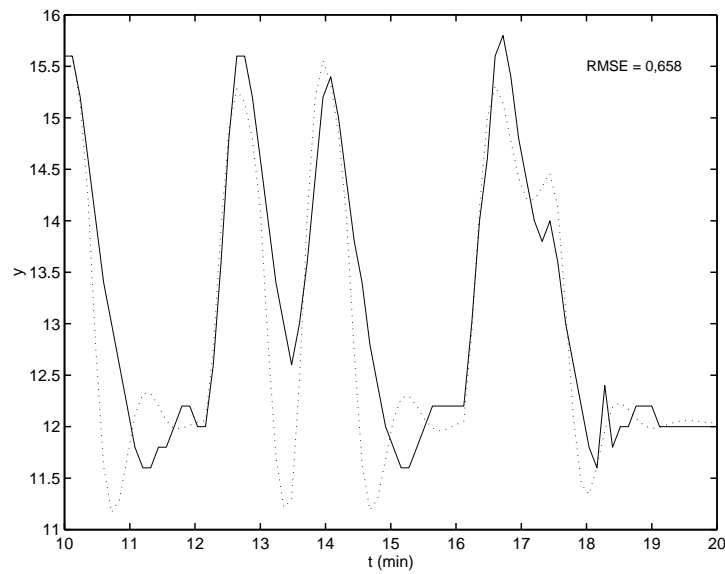


Figura 6.16: Validação estática - conversor buck

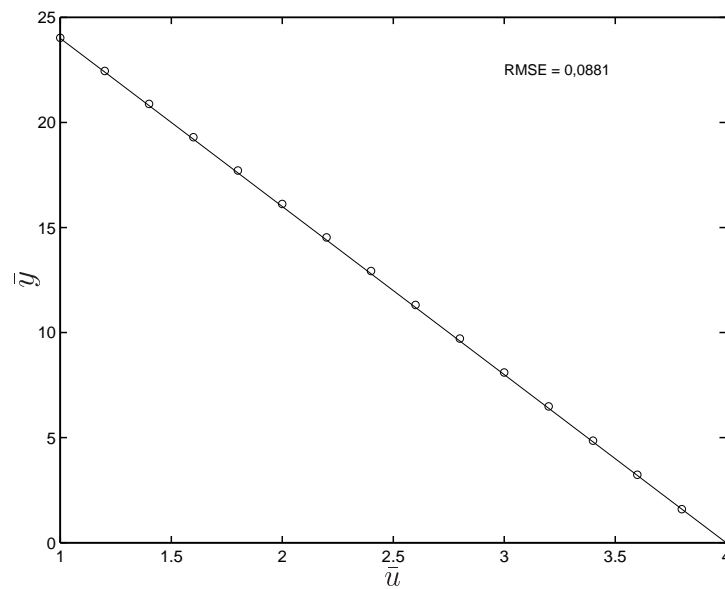
Caso conversor buck - Validação estática dos modelos (a-f) retirados do conjunto Pareto-ótimo da Figura 6.14, sendo: (-) dados reais e (o) a resposta do modelo.

auxiliar. Os dois sistemas estudados foram: um aquecedor elétrico (Cassini, 1999) e um conversor buck (Aguirre et al., 2000). No caso do aquecedor elétrico a informação auxiliar foi obtida experimentalmente, enquanto para o conversor, a curva estática foi obtida através do conhecimento da física do processo. Conforme discutido no Capítulo 2, a identificação no caso do conversor buck trata-se de um procedimento *caixa-cinza*, enquanto no caso do aquecedor, trata-se de um procedimento *caixa-preta*, tendo em vista a diferença de origem da informação auxiliar.

A abordagem empregada mostrou-se eficiente por várias razões. Primeiramente, a abordagem multiobjetivo propiciou resultados tão bons quanto aqueles mostrados por (Barroso, 2001), apresentando duas vantagens: i) permite o acesso a soluções intermediárias, que em certos casos podem ser mais

Figura 6.17: Validação dinâmica do modelo s^* - conversor buck

Caso conversor buck - Validação dinâmica do modelo s^* sendo: (-) os dados reais e (·) a resposta do modelo.

Figura 6.18: Validação estática do modelo s^* - conversor buck

Caso conversor buck - Validação estática do modelo s^* sendo: (-) os dados reais e (o) a resposta do modelo.

apropriadas do que as soluções individuais de cada função objetivo; ii) não necessita de procedimentos intermediários para ajustar a curva estática, tornando a metodologia mais simples.

Além disso, para casos como o do conversor buck em que os dados disponíveis não excursionam em toda a faixa de operação do sistema, a metodologia mostra-se eficiente para encontrar um modelo mais global.

Identificação Multiobjetivo de um TCSC

“Não existe nenhum caminho lógico que nos conduza às grandes leis do universo. Elas só podem ser atingidas por meio de intuições baseadas em algo semelhante a um amor intelectual pelos objetos da experiência.”

Albert Einstein

7.1 Introdução¹

O objetivo deste capítulo é utilizar a informação da curva estática para identificação multiobjetivo de um TCSC (*Thyristor Controlled Series Capacitor* ou capacitor série controlado a tiristor). A entrada e saída são o ângulo de disparo do tiristor e a reatância do TCSC, respectivamente. A curva estática desse dispositivo apresenta uma região de ressonância. A agregação de tal informação é um desafio devido: (i) o comportamento ressonante raramente aparece na operação normal do dispositivo (tornando muito difícil a obtenção de bons modelos a partir somente da série temporal de dados); (ii) a utilização de modelos do TCSC para análise/controlado não será adequada se tal comportamento ressonante não for assimilado pelo modelo.

Utilizou-se a base conceitual descrita no Capítulo 4 para a agregação de informação acerca da resposta estática. Em particular, neste capítulo é feita

¹Os principais resultados apresentados nesse capítulo estão em um artigo submetido ao XIV Congresso Brasileiro de Automática (Nepomuceno et al., 2002a).

a agregação de informação acerca da resposta estática *ressonante* presente na operação de um TCSC, sendo que os dados de entrada e saída são o ângulo de disparo e a reatância do TCSC, respectivamente. O TCSC é modelado no Simulink/Matlab (Neto et al., 2000).

Diferentemente dos dois exemplos apresentados no Capítulo 6, o aquecedor elétrico e o conversor buck, o TCSC não possuía um modelo caixa-preta previamente identificado. O presente Capítulo irá mostrar todas as etapas da identificação, e não somente a incorporação da informação acerca da curva estática, constituindo-se assim em um estudo de caso abrangente. Assim, pelo fato de se constituir em um exemplo completo de identificação e possuir uma particularidade na curva estática (uma região de ressonância) foi dedicado um Capítulo ao estudo de caso do TCSC.

O restante do Capítulo está organizado da seguinte forma. A Seção 2 descreve em linhas gerais o TCSC. A metodologia para simulação e obtenção dos dados de entrada e saída são apresentadas na Seção 3. Os resultados e uma breve discussão estão na Seção 4, e por fim, a Seção 5 traz as considerações finais.

7.2 O TCSC

7.2.1 Relevância do TCSC

No fim da década de 80 e início da década de 90, um conceito geral em termos da operação e planejamento de sistemas de potência foi proposto e amplamente discutido em inúmeros congressos e publicações: FACTS (*Flexible AC Transmission System* ou sistema flexível de transmissão AC), incorporando um grande número de possibilidades para a melhor utilização dos sistemas de energia existentes (Hingorani, 1994). A tecnologia FACTS envolve um controle rápido de vários parâmetros inter-relacionados do sistema de potência (tensão, impedância, ângulo de fase, corrente, potência ativa e reativa) usando dispositivos eletrônicos de alta potência e alta velocidade. Entre os dispositivos FACTS propostos, possivelmente o TCSC apresentou os melhores resultados em termos de desempenho e flexibilidade (Gama, 1999). O TCSC foi usado em muitas aplicações: para controlar o fluxo de potência, para amortecer oscilações do sistema, para suprimir a ressonância subsíncrona, para limitar correntes de falta, entre outras. Conseqüentemente, a modelagem desse dispositivo passou a ser de grande importância, tanto para

análise quanto para o projeto de controladores (Neto et al., 2000).

Os programas convencionais de modelagem de sistemas de potência, tais como EMTP e ATP e, recentemente, programas mais flexíveis e mais amigáveis tais como Matlab, já foram usados para modelar o TCSC (Paserba et al., 1995; Gama e Tenorio, 2000). É de grande interesse que um modelo, baseado em dados de entrada e saída do dispositivo, permita a sua conexão direta em algum ponto da rede, ou em um sistema maior dentro de um programa da simulação. Isso é justificável para casos em que os sistemas são muito complexos e não é possível obter uma descrição por meio de modelagem caixa-branca.

7.2.2 Descrição do TCSC

Basicamente, o TCSC é composto de um capacitor em paralelo a um TCR (par de tiristores anti-paralelos em série com um reator). O capacitor e o TCR são dispostos em série com a linha de transmissão. A Figura 7.1 mostra o circuito básico de um TCSC. O capacitor pode ser protegido por um *metal oxide varistor* (MOV), ou varistor de óxido metálico. Para simplificação, o MOV não será considerado. A Figura 7.2 ilustra a variação da impedância no TCSC. Os tiristores (T1 e T2) transformam o reator L em uma reatância variável, que, estando em paralelo com o capacitor C, resulta numa impedância (Z) variável.

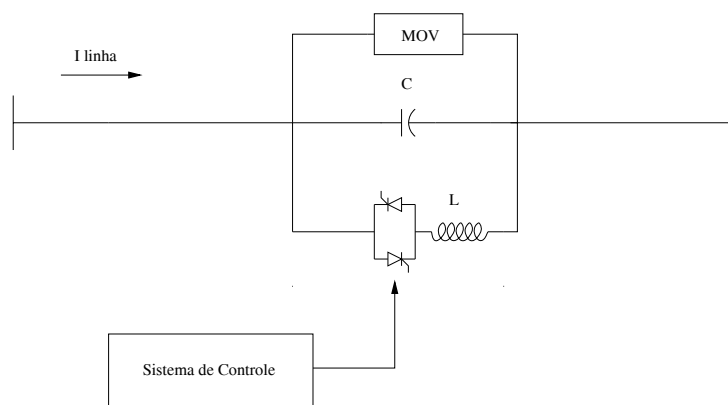


Figura 7.1: Diagrama *on-line* do TCSC

Diagrama *on-line* do TCSC, mostrando o circuito básico. O capacitor pode ser protegido por um *metal oxide varistor* (MOV), ou varistor de óxido metálico.

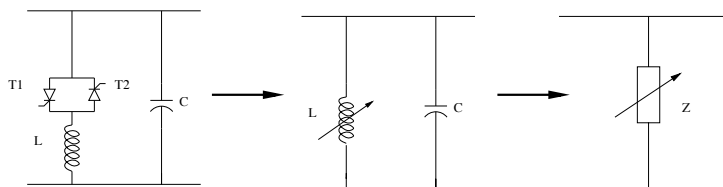


Figura 7.2: Variação de impedância em um TCSC

Esquema ilustrativo da variação de impedância em um TCSC. Os tiristores (T1 e T2) transformam o reator L em uma reatância variável, que, estando em paralelo com o capacitor C, resulta numa impedância (Z) variável.

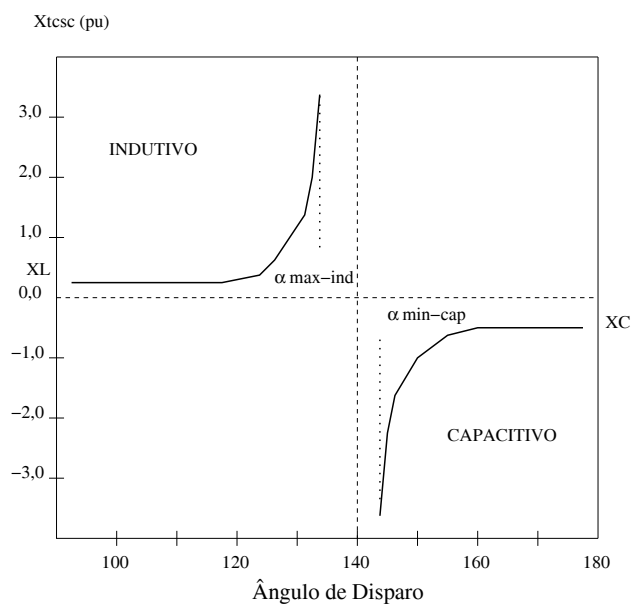


Figura 7.3: Impedância de um TCSC para a frequência fundamental

Característica da impedância de um TCSC hipotético para a frequência fundamental. Pode-se ver a característica não-linear do dispositivo e sua região ressonante.

O dispositivo pode ser operado em três modalidades diferentes: (i) *contornado*: os tiristores estão operando em capacidade máxima; (ii) *bloqueado*: os tiristores são obstruídos; e (iii) *vernier*: a condução dos tiristores é controlada por um sinal na porta e conseqüentemente o TCSC passa a ter uma

reatância controlada nas regiões indutivas e capacitivas. Esse último caso é o que será empregado. Os ângulos de disparo do tiristor (α) podem variar de 90° até um valor indutivo máximo ($90^\circ + \alpha_{\text{ind}}$) na escala de operação e de 180° até um valor capacitivo mínimo ($180^\circ - \alpha_{\text{cap}}$). O valor máximo de reatância indutiva e o valor mínimo de reatância capacitivo devem ser ajustados com alguma margem de segurança no projeto do dispositivo para impedir a ressonância em paralelo entre o capacitor e o TCR na frequência fundamental.

A Figura 7.3 mostra a impedância característica em estado permanente do TCSC. Pode-se ver na Figura 7.3 a característica não-linear do dispositivo e sua região ressonante. O conhecimento dessa característica de ressonância desempenha um papel importante no estabelecimento dos limites de operação do dispositivo. As avaliações do equipamento devem também ser feitas de tal forma que esses limites sejam respeitados.

A Figura 7.4 mostra os limites para reatâncias capacitivas e indutivas em função da corrente de linha. Os ângulos de disparo fornecem os limites para uma reatância capacitiva máxima (curva A) e para uma reatância indutiva máxima (curva E), ambos com correntes em níveis baixos. Em valores de correntes mais elevados o limite da reatância capacitiva é inversamente proporcional à corrente de linha (curva C) e é dada pela avaliação máxima da tensão do capacitor. O limite da reatância indutiva é também inversamente proporcional à corrente de linha para correntes mais elevadas (curva F) e é dada pelo efeito de aquecimento provocado pelos harmônicos da corrente fundamental. Para correntes muito elevadas de linha, o limite capacitivo mínimo é conseguido quando os tiristores estão obstruídos (curva B) e limite indutivo mínimo quando os tiristores estão na condução cheia (D). As correntes dos tiristores podem também limitar a corrente de linha (curva G).

7.3 Metodologia

A Figura 7.5 mostra o sistema-teste modelado no Simulink/Matlab (Neto et al., 2000). O sistema é basicamente composto por um gerador, por uma linha da transmissão com um TCSC e por uma carga resistiva. Os parâmetros usados foram:

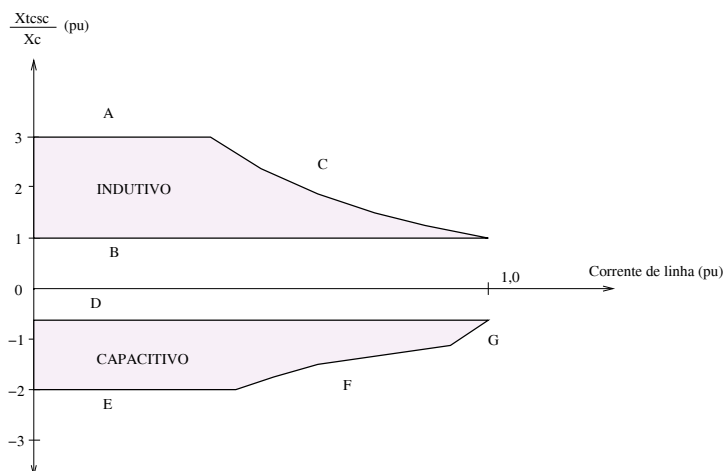


Figura 7.4: Relação típica entre impedância do TCSC e corrente de linha

Relação típica entre impedância do TCSC e corrente de linha. Estão também ilustrados os limites para reatâncias capacitivas e indutivas em função da corrente de linha.

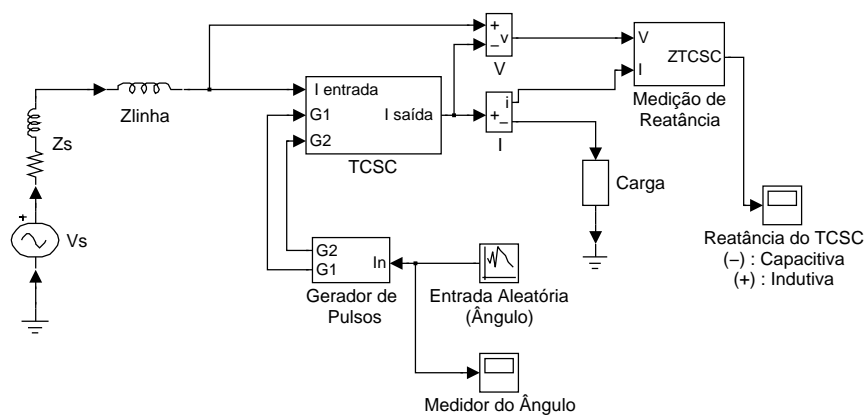


Figura 7.5: Sistema de potência com um TCSC modelado no Simulink

Sistema de potência com um TCSC modelado no Simulink. A entrada do sistema é o ângulo de disparo e a saída, a reatância do TCSC.

- Fonte da tensão: 65 kV, 60 Hz;
- Impedância em série: $4,5 \Omega$ e 132 mH;
- Capacitância do TCSC: $212 \mu\text{F}$;

- Indutância da linha de transmissão: 265 mH;
- Indutância do TCSC: 15 mH;
- Carga resistiva: 100 Ω .

O sistema foi excitado por uma entrada aleatória, agindo como o ângulo de disparo dos tiristores. Cada amostra de entrada foi mantida constante por 0,25 s a fim de garantir tempo suficiente para que o sistema respondesse adequadamente. O passo de integração usado na simulação foi $T_i = 30 \mu\text{s}$. A reatância do sistema foi computada usando procedimento padrão e foi usada para compor os dados de entrada e saída. Esses dados foram decimados por um fator de decimação $\Delta = 250$ para gerar os dados de trabalho. A Figura 7.6 mostra uma série temporal simulada com entrada e saída. A Figura 7.7 mostra a janela dos dados, entre 6 s a 8 s, em que há um valor ressonante. Essa série temporal foi quebrada em duas séries menores. Os dados de 1 s a 4 s foram usados para o processo de identificação, enquanto os dados entre 4 s e 6,5 s foram usados para validar os modelos.

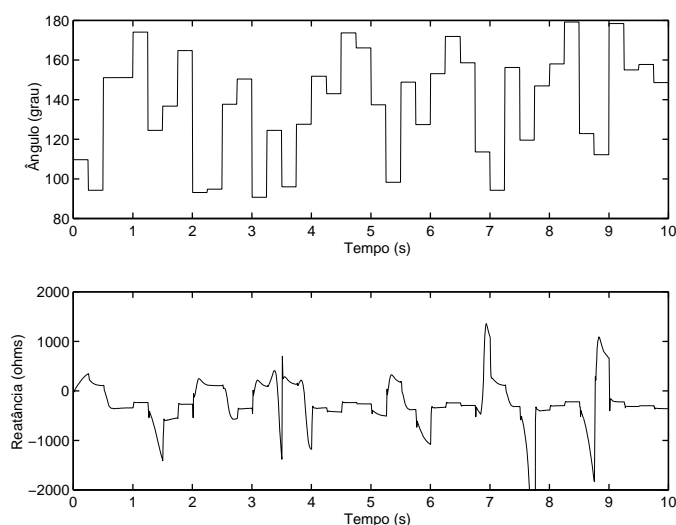


Figura 7.6: Simulação do sistema-teste com um TCSC

Simulação do sistema-teste com um TCSC, entrada (ângulo de disparo do tiristor) aleatória, tendo como saída a reatância do TCSC. O passo de integração foi $30 \mu\text{s}$ e os dados foram decimados por um fator de 250, gerando os dados de identificação e validação. Os dados entre 1 s e 4 s foram usados para o procedimento de identificação, enquanto os dados de validação situam-se entre 4 s e 6,5 s.

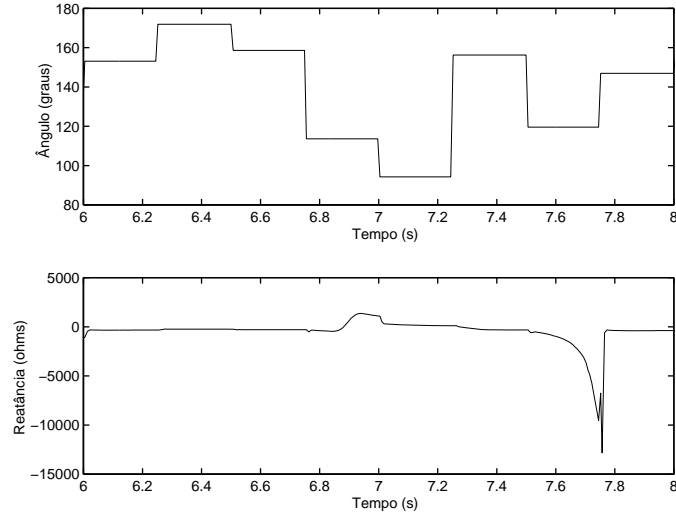


Figura 7.7: Simulação do sistema-teste com um TCSC - dados de 6 s a 8 s

Simulação do sistema-teste com um TCSC. Janela de dados da Figura 7.6 correspondente ao intervalo de 6 s a 8 s.

A informação a respeito da resposta de estática ressonante não está fortemente presente nas séries temporais simuladas, e pode ser obtida de duas maneiras. Primeiramente, é possível calcular uma aproximação para a relação analítica entre o ângulo de disparo e a reatância equivalente através das equações do circuito. A Figura 7.8 mostra uma simplificação de um sistema contendo um TCSC. A corrente $i(t)$ é dada por (Sen, 1989):

$$i(t) = \frac{\sqrt{2}V_0}{Z} \left[\text{sen}(wt - \phi) - \text{sen}(\alpha - \phi)e^{-\frac{R}{X_L}(t - \frac{\alpha}{w})} \right] \quad (7.1)$$

em que:

- $Z = \sqrt{R^2 + X_L^2}$: impedância total da carga RL ;
- α : ângulo de disparo dos tiristores;
- $\phi = \text{arctg}\left(\frac{X_L}{R}\right)$: ângulo da carga RL ;
- $e(t) = \sqrt{2}V_0\text{sen}(wt)$: tensão da fonte.

Considerando $L \gg R$ e assim $\phi = \pi/2$, a Equação 7.1 torna-se:

$$i(t) = \frac{\sqrt{2}V_0}{X_L}(\cos(\alpha) - \cos(\omega t)). \quad (7.2)$$

Nessas condições, a componente fundamental de $i(t)$ é

$$I_1 = \frac{\sqrt{2}V_0}{\pi X_L}[(2\Lambda - \text{sen}2\Lambda)\cos(\omega t)], \quad (7.3)$$

em que $\Lambda = \pi - \alpha$. E por inspeção

$$X_{\text{Leq}} = \frac{\pi X_L}{2\Lambda - \text{sen}2\Lambda}. \quad (7.4)$$

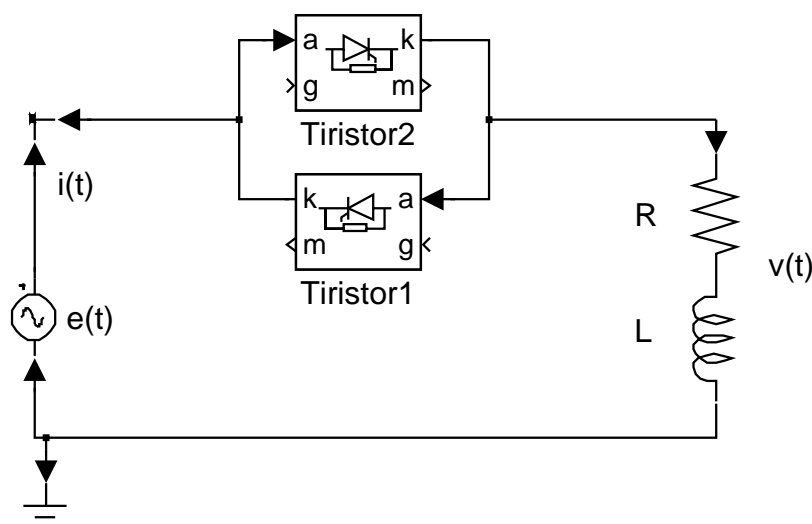


Figura 7.8: Simplificação do sistema com um TCSC

Simplificação do sistema com um TCSC para descrição matemática da relação entre ângulo de disparo e reatância equivalente.

O segundo procedimento consiste em varrer o ângulo de disparo de 90° a 180° , utilizando o modelo que foi empregado para gerar as séries temporais simuladas, armazenando os valores em regime permanente. A Figura 7.9 mostra a reatância do TCSC em função da variação do ângulo de disparo. O segundo método foi escolhido nesse trabalho para gerar os dados de iden-

tificação auxiliares. A razão dessa escolha é a tentativa de diminuir os erros inerentes à linearização de (7.4).

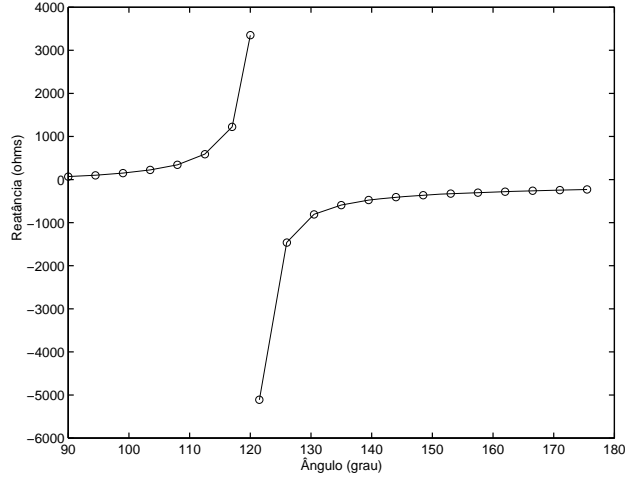


Figura 7.9: Reatância em regime permanente do TCSC

Reatância em regime permanente do TCSC. Os valores negativos referem-se a reatâncias capacitivas.

A identificação do TCSC pode ser vista como um problema multiobjetivo de otimização, em que as duas funções objetivo são: o erro quadrático de predição e o erro quadrático da diferença entre a resposta em regime permanente da reatância do modelo e do sistema original (7.9). A resposta de estado permanente do sistema é uma função linear nos parâmetros no modelo NARMAX, e pode facilmente ser expressa como $A\hat{\theta} = \hat{\mathbf{b}}$, em que A é um operador linear, $\hat{\theta}$ é o vetor os parâmetros e $\hat{\mathbf{b}}$ a informação a respeito da resposta estática ressonante. Assim é possível construir as seguintes funções de custo:

$$J_{\text{MQ}}(\hat{\theta}) = (\mathbf{y} - \Psi\hat{\theta})^T (\mathbf{y} - \Psi\hat{\theta}) \quad (7.5)$$

e

$$J_{\text{CE}}(\hat{\theta}) = (\mathbf{b} - A\hat{\theta})^T (\mathbf{b} - A\hat{\theta}). \quad (7.6)$$

Utilizando o Teorema 4.4.1 chega-se à seguinte Equação para estimação dos parâmetros do modelo NARMAX:

$$\hat{\theta} = [w\Psi^T\Psi + (1-w)A^T A]^{-1}[w\Psi^T\mathbf{y} + (1-w)A^T\mathbf{b}]. \quad (7.7)$$

7.4 Resultados e discussão

O melhor modelo identificado sem o uso de informação auxiliar é mostrado na Equação (7.8).

$$\begin{aligned}
 y(k) = & +0,16099 \times 10^{+0} u(k-1)y(k-1) \\
 & -0,24229 \times 10^{-3} u(k-1)u(k-1)u(k-1) \\
 & -0,88739 \times 10^{+1} y(k-1) \\
 & -0,65486 \times 10^{-3} u(k-1)u(k-1)y(k-1) \\
 & +0,37955 \times 10^{+0} u(k-1) \\
 & +0,26710 \times 10^{-1} u(k-1)u(k-1) \\
 & +\Psi_{\xi}^T(k-1)\hat{\theta}_{\xi} + \xi(k).
 \end{aligned} \tag{7.8}$$

A parte MA $\Psi_{\xi}^T(k-1)\hat{\theta}_{\xi}$ foi estimada para reduzir a polarização, mas não é usada para simulação.

Um outro modelo, incorporando a informação auxiliar, como descrita na Equação (7.7), é apresentado na Equação (7.9). Este modelo foi encontrado com $w = 0,0954$, e com a mesma estrutura do modelo (7.8). O valor apropriado de w foi encontrado pela inspeção de uma amostra de 50 elementos do conjunto Pareto-ótimo que foi encontrado variando w de 0 a 1. Essa abordagem de inspeção, constitui-se em um problema de *decisão* do tipo *apresentação de preferências a posteriori*. Neste caso o critério de preferência a *posteriori* foi o adequado ajuste da resposta estática na região de operação do TCSC.

$$\begin{aligned}
 y(k) = & +0,16931 \times 10^{+0} u(k-1)y(k-1) \\
 & -0,21315 \times 10^{-3} u(k-1)u(k-1)u(k-1) \\
 & -0,93511 \times 10^{+1} y(k-1) \\
 & -0,69119 \times 10^{-3} u(k-1)u(k-1)y(k-1) \\
 & +0,90720 \times 10^{+0} u(k-1) \\
 & +0,17864 \times 10^{-1} u(k-1)u(k-1) \\
 & +\Psi_{\xi}^T(k-1)\hat{\theta}_{\xi} + \xi(k).
 \end{aligned} \tag{7.9}$$

A Figura 7.10 (alto) compara a saída do modelo (7.8) com os dados originais. Nesse caso o RMSE é 0,2945. A Figura 7.10 (baixo) mostra a saída dinâmica do modelo (7.9). O RMSE para essa resposta dinâmica é 0,3374.

A Figura 7.11 mostra as respostas estáticas dos modelos, considerando os valores apresentados na Figura 7.3. A faixa do ângulo de disparo (α) do tiristor foi escolhida de 90° a 115° (faixa indutiva), e de 135° a 180° (faixa

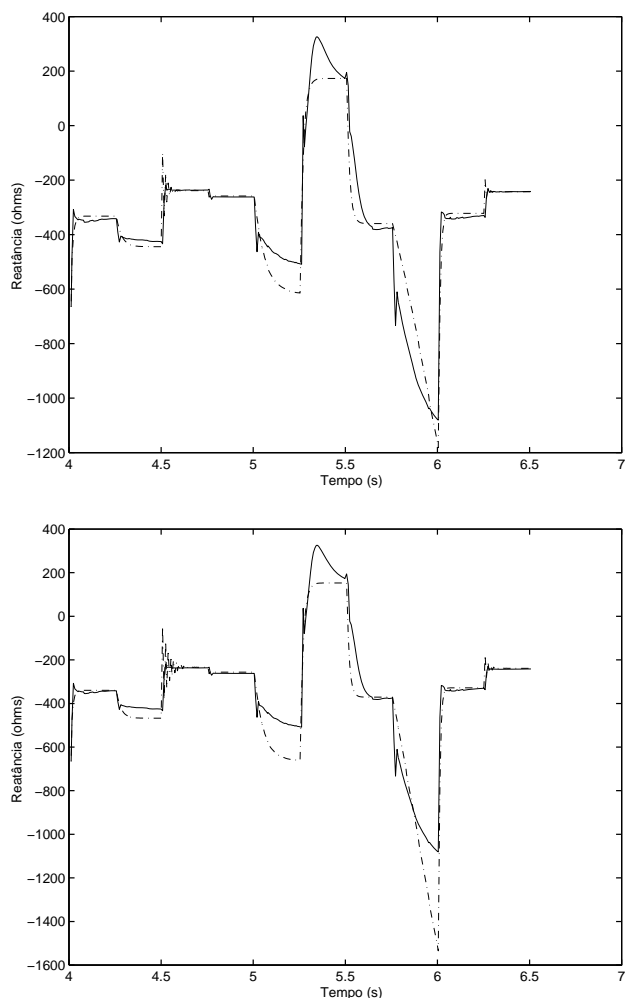


Figura 7.10: Validação dinâmica para o TCSC

Testes de validação dinâmica para o TCSC. Acima: modelo (7.8), estimado por MQ convencionais. (—) dados originais, (— · —) saída do modelo. Abaixo: modelo (7.9), o qual teve os seus parâmetros estimados usando informação auxiliar a respeito da resposta estática ressonante. (—) dados originais, (— · —) saída do modelo. Em ambos os casos, os dados originais correspondem a uma janela de dados entre 4 s e 7,5 s da Figura 7.6.

capacitiva). Os valores de 115° a 135° são irrelevantes, pois estão fora da faixa de operação permitida para o dispositivo, e não são empregados aqui para finalidades da comparação. A Figura 7.11, no alto, mostra a resposta estática do modelo mono-objetivo convencional (7.8), que conduz a um RMSE de

1,7636. A mesma Figura, em baixo, mostra a resposta estática do modelo multiobjetivo (7.9). O RMSE para a resposta estática, nesse caso, é 0,4031.

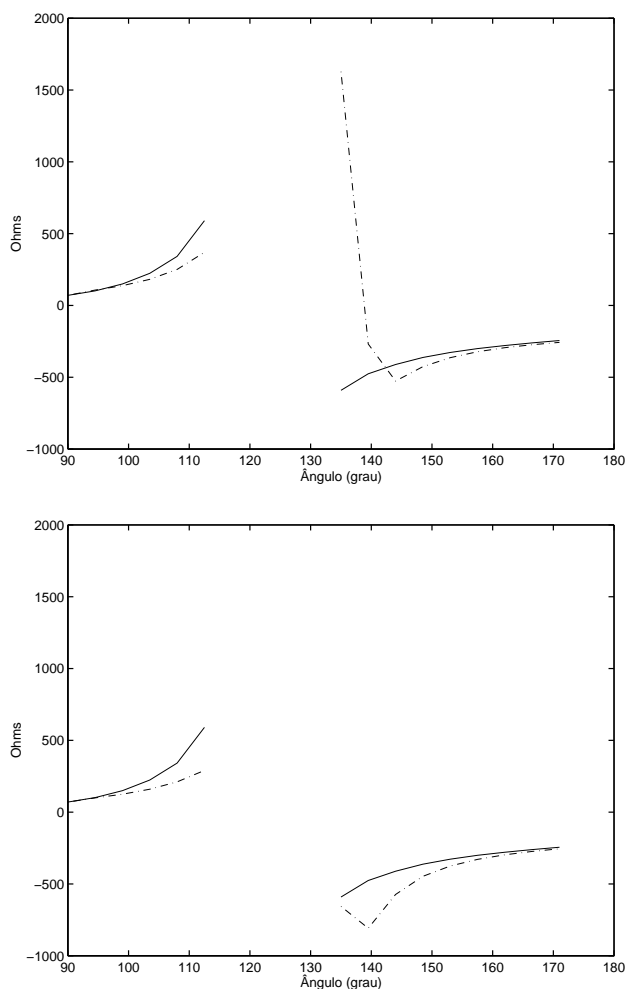


Figura 7.11: Validação estática para o TCSC

Reatância em regime permanente para dois modelos. Os valores negativos representam reatância capacitiva. Acima: método convencional de MQ. (—) dados originais, (---) modelo (7.8). Abaixo: modelo usando abordagem multiobjetivo. (—) dados originais, (---) modelo (7.9).

A resposta dinâmica de ambos os modelos é na prática equivalente e seu RMSE pode ser considerado razoável para diversas aplicações, tais como projeto de controle. O RMSE para a resposta dinâmica do modelo (7.9) aumentou cerca de 30%, mas o RMSE para a resposta estática diminuiu 400%,

quando comparado com o modelo convencional (7.8). A Figura 7.11 permite também a comparação de uma outra característica de modelo importante: o modelo multiobjetivo retém, de certo modo, a simetria entre a região de operação indutiva e a capacitiva, enquanto o modelo mono-objetivo convencional não foi capaz de assimilar essa propriedade. Conseqüentemente, parece razoável indicar que a metodologia de agregação multiobjetivo de informação é uma abordagem adequada para aumentar a qualidade do modelo, no que diz respeito à resposta estática.

7.5 Conclusões do capítulo

A metodologia multiobjetivo proposta no Capítulo 4 foi aplicada neste capítulo para modelar um TCSC usando, como informação auxiliar, a resposta estática. O TCSC é um sistema peculiar, porque sua resposta estática apresenta uma *região ressonante*, cuja caracterização não pode ser feita apenas a partir da série temporal disponível. Os resultados mostram uma considerável melhoria da resposta estática na região de operação usando a informação auxiliar, com uma pequena diminuição na qualidade do ajuste dinâmico.

Conclusão e Pesquisas Futuras

“Eu sustento que a única finalidade da ciência está em aliviar a miséria humana.”

Bertolt Brecht

8.1 Discussões finais

Um dos mais citados artigos a respeito de identificação de sistemas não-lineares, de autoria do pesquisador inglês Stephen A. Billings, (Billings, 1980) relata:

“Identificação de sistemas não-lineares é um problema muito difícil e não há uma técnica que possa ser recomendada para todos os casos. Todos os algoritmos considerados apresentam vantagens e desvantagens, sendo que cada um deve ser julgado de acordo com o problema em investigação. Isso inevitavelmente deixa nas mãos do especialista delicadas decisões e considerações, e futuras pesquisas serão requisitadas para aprimorar algoritmos de identificação e seleção de estrutura, reduzir a quantidade de amostras necessárias e simplificar as técnicas de medição.”

Incontáveis trabalhos surgiram nas décadas de 80 e 90, atentando para os problemas enumerados por Billings e vários outros. A identificação de sistemas não-lineares tornou-se incontestavelmente relevante para os mais diversos campos da ciência e tecnologia. Várias representações matemáticas

foram desenvolvidas, tais como: *série de Volterra* (Bedrosia e Rice, 1971; Billings, 1980; Bui et al., 2001), *modelos de Hammerstein e de Wiener* (Wigren, 1993; Pearson e Pottmann, 2000; Coelho, 2002), *modelos polinomiais contínuos* (Gouesbet e Letellier, 1994), *funções de base radial* (Haykin, 1994; Braga et al., 2000), *redes neurais artificiais*, (Haykin, 1994; Braga et al., 2000), *wavelets* (Daubechies, 1988; Cao et al., 1995; Combastel et al., 2002), e *modelos NARMAX polinomiais* (Leontaritis e Billings, 1985b) e *modelos NARMAX racionais* (Chen e Billings, 1989).

Nas últimas duas décadas os mais diferentes sistemas foram identificados: sistemas biológicos, químicos, físicos, percorrendo desde batimentos cardíacos (Goldberger, 1991; Schiff et al., 1994), passando por sistemas de comunicação ótica (Kennedy, 2000) até o fluxo de deposição de poluentes em uma estação de esgotos (Bechmann et al., 1999), para citar apenas alguns. Consideráveis avanços foram presenciados, mas algumas questões ainda estão abertas. Qual é a melhor representação para uma particular aplicação? Como selecionar uma estrutura de uma específica representação? Qual é o melhor procedimento para validar um modelo?

Uma outra importante questão, freqüentemente presente em situações práticas e que passou a ser estudada com mais atenção no início da década de 90, é como melhor identificar um sistema que apresentasse um conjunto de dados finitos, corrompidos por ruído e normalmente de caráter local, mas que possuísse alguma informação auxiliar? A presente dissertação teve como objetivo contribuir para responder essa pergunta, criando uma metodologia multiobjetivo para incorporação de informação auxiliar na identificação de sistemas não-lineares.

8.2 Resumo das contribuições

As principais contribuições presentes nessa dissertação são:

- Elucidação dos conceitos de identificação caixa-cinza, informação *a priori*, informação auxiliar e identificação multiobjetivo;
- Elaboração de uma formulação multiobjetivo para incorporação de informação auxiliar;
- Elaboração de uma formulação computacional através do Teorema 4.4.1, que foi demonstrado nesse trabalho;

- Descrição da incorporação de informação acerca de pontos fixos e curva estática;
- Emprego do decisor de equilíbrio da polarização e variância para escolha da solução ótima;
- Elaboração de um novo decisor baseado na norma mínima dos objetivos normalizados.

Nas seções seguintes, essas contribuições são brevemente detalhadas.

8.2.1 Elucidação de conceitos

Alguns conceitos relativos à Identificação de Sistemas estão confusos na literatura. Assim, os conceitos de identificação caixa-cinza, informação *a priori*, informação auxiliar foram elucidados e foi proposto o conceito de identificação multiobjetivo:

- *Informação auxiliar*: denota qualquer informação utilizada na identificação adicionalmente ao critério dos MQ;
- *Informação a priori*: é a informação auxiliar sabida anteriormente ao início do procedimento de identificação;
- *Identificação caixa-cinza*: é a identificação de sistemas que utiliza informação auxiliar de natureza estrutural do sistema em conjunto com dados de entrada e saídas medidos;
- *Identificação multiobjetivo*: é a identificação que utiliza um ou mais tipos de informação auxiliar de qualquer procedência, associada ao critério de minimização do erro quadrático de predição.

8.2.2 Formulação multiobjetivo

A formulação multiobjetivo consiste em tratar cada tipo de informação auxiliar e o critério convencional de MQ como um objetivo e gerar um conjunto de soluções chamado *conjunto Pareto-ótimo* ou *conjunto de soluções eficientes*. A solução é *eficiente* se e somente se não existir nenhuma outra solução que minimize um dos objetivos sem degradar ao menos um outro objetivo.

8.2.3 Formulação computacional

Para determinar o conjunto Pareto-ótimo foi empregado o *problema de ponderação*, P_w . Para os casos em que os objetivos puderem ser expressos como somatório de erros quadráticos foi demonstrado, nesse trabalho, o Teorema 4.4.1 que permite calcular o vetor de parâmetros θ , ao minimizar o vetor de objetivos $\mathbf{J}(\hat{\theta}) = J_i(\hat{\theta}) = (\mathbf{v}_i - G_i \hat{\theta})^T (\mathbf{v}_i - G_i \hat{\theta})$:

$$\hat{\theta}^* = \left[\sum_{i=1}^n w_i G_i^T G_i \right]^{-1} \left[\sum_{i=1}^n w_i G_i^T \mathbf{v}_i \right].$$

8.2.4 Informação auxiliar: pontos fixos e curva estática

Utilizando as propriedades apresentadas dos modelos NARMAX na Seção 2.4 foram elaborados mecanismos que permitem a incorporação da informação da localização dos pontos fixos e da curva estática. A metodologia proposta permite usar a formulação computacional do Teorema 4.4.1, um avanço, se comparado com outros trabalhos (Corrêa, 2001; Barroso, 2001) que necessitam de uma etapa intermediária de ajuste de curva.

Foram escolhidos dois sistemas caóticos para a utilização de informação auxiliar acerca dos pontos fixos: o circuito de Chua e o mapa senoidal com não-linearidades cúbicas. Foram duas as razões para a escolha de sistemas caóticos: grande relevância desse estudo em áreas como biologia, meteorologia, matemática, engenharia, entre outros, e por se tratarem de sistemas mais complexos que poderiam por à prova a metodologia proposta.

No estudo de caso do circuito de Chua foram investigadas duas situações: na primeira os dados de identificação foram gerados por um modelo NARMAX previamente identificado; na segunda a esses mesmos dados de identificação foi adicionado ruído. Os resultados mostraram o esperado: para casos sem ruído a identificação caixa-preta é praticamente equivalente à identificação multiobjetivo. Para casos com ruído, a identificação multiobjetivo obteve melhores resultados. Além disso, mesmo se a informação auxiliar sobre pontos fixos for obtida através dos dados, o algoritmo de estimação de pontos fixos (Aguirre e Souza, 1998) é relativamente mais robusto. A solução ótima foi escolhida através de duas técnicas: análise de preferência a posteriori de caráter qualitativo, baseada no espaço de estados, nos valores estimados do maior expoente de Lyapunov e nos diagramas de bifurcação; e por um decisor que procura equilibrar a polarização e variância. Para esse último caso,

a solução ótima equilibrada perfeitamente a polarização e variância, mas não assimilou adequadamente as propriedades dinâmicas do sistema original.

No estudo de caso do mapa senoidal, foi verificado que a imposição dos pontos fixos melhora a reconstrução do mapa de primeiro retorno. Além disso, o parâmetro de ponderação, w , serviu como parâmetro de bifurcação, levando a acreditar que a identificação multiobjetivo pode constituir-se em um mecanismo de perturbação sistemático de modelos dinâmicos, cujo resultado imediato é a possibilidade de promover pequenas alterações em modelos inadequados, mas que estejam na vizinhança de soluções adequadas, assim melhorando o desempenho de tal modelo.

A metodologia mostrou-se eficiente em ambos os estudos de casos para incorporação da informação de pontos fixos.

Quanto à incorporação da curva estática, foram escolhidos dois sistemas reais: um aquecedor elétrico e um conversor estático CC-CC do tipo buck. A informação auxiliar foi obtida empiricamente para o aquecedor elétrico e através da física do processo para o conversor buck. Em ambos os casos, a metodologia mostrou-se eficiente em encontrar soluções intermediárias que podem ser, em certas situações, mais apropriadas do que soluções individuais de cada função objetivo. Para o caso do conversor buck, a metodologia foi capaz de encontrar um modelo mais global.

Como último estudo de caso, a identificação multiobjetivo foi aplicada para a modelagem do TCSC. Nesse caso, todas as etapas da identificação foram realizadas. O modelo caixa-preta obtido por técnicas convencionais foi aprimorado consideravelmente com a utilização da informação acerca da curva estática. Esse estudo foi de particular interesse, porque a curva estática apresenta uma região de ressonância, tornando a identificação consideravelmente difícil. A solução pertencente ao Pareto-ótimo que melhor aproximou o comportamento dinâmico e comportamento estático na região de operação do dispositivo foi escolhida.

8.2.5 Problema de decisão

Para o problema de decisão foi empregado o critério de equilíbrio entre variância e polarização proposto em (Teixeira, 2001). Verificou-se para a identificação do circuito de Chua que esse método, apesar de encontrar uma solução de equilíbrio, não foi capaz de assimilar as propriedades dinâmicas do sistema.

Como método alternativo, foi desenvolvido nesse trabalho um método denominado de norma mínima. Tal critério, discutido na Seção 4.6.2. normaliza os valores de cada objetivo em uma escala de 0 a 1, e então, escolhe a solução que apresentar a menor norma euclidiana. Esse método é mais simples do que o anterior e de certa forma permite que os objetivos, por estarem normalizados, sejam comparados.

8.3 Conclusões

A utilização de informação auxiliar através de uma abordagem multiobjetivo mostrou-se realizável e apresentou bons resultados nos casos estudados. A determinação de soluções pertencentes ao *conjunto de Pareto* permite o estudo, nesse conjunto, da maneira como variam propriedades relevantes dos modelos e, com base nessa análise, a escolha dos modelos identificados mais adequados às necessidades específicas de cada situação. A possibilidade dessa análise, em síntese, é a principal vantagem da metodologia multiobjetivo.

8.4 Pesquisas futuras

A partir do que foi apresentado e discutido nessa dissertação, podem-se enumerar algumas propostas de pesquisas futuras. Algumas eram pertencentes *a priori* ao escopo do trabalho, mas não foram realizadas devido ao tempo escasso e outras surgiram ao longo da investigação. A esperança é que tais propostas possam ser um dia conduzidas, já que constituem, em princípio, áreas promissoras para a área de Identificação Multiobjetivo de Sistemas. As propostas serão apresentadas na forma de perguntas a serem respondidas:

1. De acordo com a representação matemática e com o sistema em investigação como escolher a informação auxiliar mais relevante?
2. Como investigar situações que três ou mais tipos de informação auxiliar estejam disponíveis;
3. Como elaborar novos decisores e criar mecanismos de comparação entre os resultados obtidos por decisores diferentes? Seria possível utilizar correlação de ruídos, como um decisor?

-
4. Como utilizar Identificação Multiobjetivo para a determinação de estruturas de uma forma sistemática? Haveria formas de elaborar funcionais para a seleção de estruturas em conjunto com o ERR?
 5. Como sistematizar o processo de perturbação de modelos dinâmicos utilizando o parâmetro w ?
 6. Haveria vantagens em usar outras técnicas de solução para problemas multiobjetivo, tais como, P_χ e P_ϵ ?

Bibliografia

- Aguirre, L. (1994). Term clustering and the order selection of linear continuous systems. *Journal of the Franklin Institute-Engineering and Applied Mathematics*, 331B(4):403–415.
- Aguirre, L. (1997). On the structure of nonlinear polynomial models: higher order correlation functions, spectra, and term clusters. *IEEE Transactions on Circuits and Systems Part I: Fundamental Theory and Applications*, 44(5):450–453.
- Aguirre, L. e Souza, A. (1998). An algorithm for estimating fixed points of Dynamical Systems from time series. *International Journal of Bifurcation and Chaos*, 8(11):2203–2213.
- Aguirre, L. A. (1995). A nonlinear correlation function for selecting the delay time in dynamical reconstructions. *Physics Letters A*, 203(2,3):88–94.
- Aguirre, L. A. (2000a). *Introdução à Identificação de Sistemas*. Editora da UFMG.
- Aguirre, L. A. (2000b). A nonlinear dynamical approach to system identification. *IEEE Circuits & Systems Newsletter Society*, 11(2):10–23, 47.
- Aguirre, L. A. e Billings, S. A. (1994). Validating identified nonlinear models with chaotic dynamics. *International Journal of Bifurcation and Chaos*, 4(1):109–125.
- Aguirre, L. A. e Billings, S. A. (1995a). Dynamical effects of overparametrization in nonlinear models. *Physica D*, 80(1–2):26–40.

- Aguirre, L. A. e Billings, S. A. (1995b). Improved structure selection for nonlinear models based on term clustering. *International Journal of Control*, 62(3):569–587.
- Aguirre, L. A. e Corrêa, M. V. (2002). Using prior knowledge to constrain parameter estimation in nonlinear system identification. *IEEE Transactions on Circuits and Systems Part I: Fundamental Theory and Applications*. (accepted).
- Aguirre, L. A., Donoso-Garcia, P. F., e Santos-Filho, R. (2000). Use of *a priori* information in the identification of global nonlinear model – a case study using a buck converter. *IEEE Transactions on Circuits and Systems Part I: Fundamental Theory and Applications*, 47(7):1081–1085.
- Aguirre, L. A., Freitas, U., Letellier, C., e Maquet, J. (2001). Structure selection techniques applied to continuous-time nonlinear models. *Physica D*, 158(1–4):1–18.
- Aguirre, L. A. e Jácome, C. R. (1998). Cluster analysis of NARMAX models for signal-dependent systems. *IEE Proc. - Control Theory and Applications*, 145(4):409–414.
- Aguirre, L. A., Letellier, C., e Maquet, J. (2002). Induced one-parameter bifurcations in identified nonlinear models. *International Journal of Bifurcation and Chaos*, 12(1):135–145.
- Aguirre, L. A. e Mendes, E. M. A. M. (1996). Global nonlinear polynomial models: structure, term cluster and fixed points. *International Journal of Bifurcation and Chaos*, 6(2):279–294.
- Aguirre, L. A., Rodrigues, G., e Mendes, E. (1997). Nonlinear identification and cluster analysis of chaotic attractor from a real implementation of chua’s circuit. *International Journal of Bifurcation and Chaos*, 7(6):1411–1423.
- Aguirre, L. A. e Tôrres, L. A. B. (1997). Fixed point stability of Chua’s circuit: A case study with a real circuit. *Journal of Circuits Systems and Computers*, 7(2):111–115.

- Alessandri, A. e Parisini, T. (1997). Nonlinear modelling of complex large-scale plants using neural networks and stochastic approximation. *IEEE Transactions on Systems, and Cybernetics - Part A*, 27(6):750–757.
- Amaral, G. F. V. (2001). *Uso de Redes Neurais e Conhecimento a priori na Identificação de Sistemas Dinâmicos Não-Lineares*. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Artuso, R., Casati, G., e Lombardi, R. (1993). Periodic orbit theory of anomalous diffusion. *Physical Review Letters*, 71(1):62–64.
- Aström, K. J. e Wittenmark, B. (1990). *Computer-controlled systems - theory and design*. Prentice–Hall International, New York, U.S.A.
- Bai, E. W. e Sastry, S. S. (1986). Parameter identification using prior information. *International Journal of Control*, 44(2):455–473.
- Barany, E. (2001). Identification in the presence of symmetry: Oscillator networks. *IEEE Transaction on Automatic Control*, 46(3):476–481.
- Barroso, M. F. S. (2001). *Métodos de Otimização Mono-Objetivo Aplicados à Identificação Caixa-Cinza de Sistemas Não-Lineares*. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Bechmann, H., Nielsen, M. K., Madsen, H., e Niels, K. P. (1999). Grey-box modelling of pollutant loads from a sewer system. *Urban Water*, 1(1):71–78.
- Bedrosian, E. e Rice, S. O. (1971). Output properties of Volterra systems (nonlinear systems with memory) driven by harmonic and Gaussian inputs. *Proceedings of the IEEE*, 59(12):1688–1707.
- Billings, S. A. (1980). Identification of nonlinear systems — a survey. *IEE Proc. D Control Theory and Applications*, 127(6):272–285.
- Billings, S. A. e Chen, S. (1989). Identification of non-linear rational systems using a prediction error estimation algorithm. *International Journal of Systems Science*, 20(3):467–494.

- Billings, S. A., Chen, S., e Korenberg, M. J. (1989). Identification of MIMO nonlinear systems using a forward-regression orthogonal estimator. *International Journal of Control*, 49(6):2157–2189.
- Billings, S. A. e Coca, D. (1999a). Discrete wavelet models for identification and qualitative analysis of chaotic systems. *International Journal of Bifurcation and Chaos*, 9(7):2157–2189.
- Billings, S. A. e Coca, D. (1999b). Discrete wavelet models for identification and qualitative analysis of chaotic systems. *International Journal of Bifurcation and Chaos*, 9(7):1263–1284.
- Billings, S. A. e Voon, W. S. F. (1984). Least squares parameter estimation algorithms for non-linear systems. *International Journal of Systems Science*, 15(6):601–615.
- Billings, S. D. e Voon, W. S. F. (1983). Structure detection and model validity tests in the identification of nonlinear systems. *IEE Proc. D Control Theory and Applications*, 130(4):193–199.
- Billings, S. D. e Voon, W. S. F. (1986). Correlation model validity tests for nonlinear models. *International Journal of Control*, 44(1):235–244.
- Bohlin, T. (1991). *Interactive System Identification: Prospects and Pitfalls*. Springer-Verlag.
- Bohlin, T. (1994). A case study of grey box identification. *Automatica*, 30(2):307–318.
- Bohlin, T. e Graebe, S. T. (1995). Issues in nonlinear stochastic grey box identification. *International Journal of Adaptive Control and Signal Processing*, 9(6):465.
- Braake, H., Roubos, J., e Babuska, R. (1999). *Fuzzy Logic Control*, páginas 205–226. World Scientific, London, UK.
- Braga, A. P., Carvalho, A. C. P. L. F., e Ludemir, T. B. (2000). *Redes Neurais Artificiais: teoria e aplicações*. Livraria Técnica Científica - LTC.
- Broomhead, D. S. e Lowe, D. (1988). Multivariable functional interpolation and adaptive networks. *Complex Systems*, 2:321–355.

- Brown, R., Rulkov, N., e Tufflaro, N. (1994). Nonlinear prediction of chaotic time series. *Physical Review*, 50(6):4488–4508.
- Bui, F. M., Li, J. C., Bott, K., e Mintchev, M. P. (2001). Volterra series modelling and compensation of non-linear distortions caused by susceptibility difference artefacts related to the presence of ferromagnetic implants in magnetic resonance imaging. *Medical Engineering and Physics*, 23(3):207–215.
- Cao, L. C., Hong, Y., Fang, H., e He, G. (1995). Predicting chaotic time series with wavelet network. *Physica D*, 85:225–238.
- Casdagli, M., Eubank, S., Farmer, J. D., e Gibson, J. (1991). State space reconstruction in the presence of noise. *Physica D*, 51:52–98.
- Cassini, C. C. S. (1999). *Estimação Recursiva de Características Estáticas Não Lineares Utilizando Modelos Polinomiais NARMAX*. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Chankong, V. e Haimes, Y. Y. (1983). *Multiobjective decision making: theory and methodology*. North-Holland (Elsevier), New York.
- Chen, C.-T. (1999). *Linear System Theory and Design*. Oxford University Press, 3 edition.
- Chen, S., Billings, S., e Luo, W. (1989). Orthogonal least-squares methods and their applications to non-linear system identification. *International Journal of Control*, 50(5):1873–1896.
- Chen, S. e Billings, S. A. (1989). Representations of non-linear systems: the NARMAX model. *International Journal of Control*, 49(3):1013–1032.
- Chua, L. O. e Hasler, M. G. E. (1993). Special issue on chaos in nonlinear electronic circuits. *IEEE Transactions on Circuits and Systems Part I: Fundamental Theory and Applications*, 40:10–11.
- Codreanu, S. (2002). Desynchronization and chaotification of nonlinear dynamical systems. *Chaos, Solitons & Fractals*, 13(4):839–843.

- Coelho, M. C. S. (2002). *Modelos de Hammerstein e Wiener: conexões com modelos NARX e sua aplicação em identificação de sistemas não-lineares*. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Coelho, M. C. S., Corrêa, M. V., e Aguirre, L. A. (2001). Obtenção de modelos Hammerstein e modelos Wiener. *XXIV CNMAC - Congresso Nacional de Matemática Aplicada e Computacional, Belo Horizonte, Minas Gerais, Brasil*, 1:133.
- Combastel, C., Lesecq, S., Petropol, S., e Gentil, S. (2002). Model-based and wavelet approaches to induction on-line fault detection. *Control Engineering Practice*, 10(5):595–606.
- Corrêa, M. V. (2001). *Identificação caixa-cinza de sistemas não-lineares utilizando representações NARMAX racionais e polinomiais*. Tese de Doutorado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Cubillos, F. A., Alvarez, P. I., Pinto, J. C., e Lima, E. L. (1996). Hybrid-neural modeling for particulate solid drying processes. *Power Technology*, 87:153–160.
- Cubillos, F. A. e Lima, E. L. (1997). Identification and optimizing control of a rougher flotation circuit using an adaptable hybrid-neural model. *Minerals Engineering*, 10(7):707–721.
- Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41:909–996.
- Draper, N. R. e Smith, H. (1998). *Applied Regression Analysis*. Wiley series in probability and statistics, 3rd edition.
- Duran, M. A. e White, B. S. (1994). Bayesian estimation applied to effective heat transfer coefficients in a packed bed. *Chemical Engineering Science*, 50(3):495–510.
- Eskinat, E., Johnson, S. H., e Luyben, W. L. (1993). Use of auxiliary information in system identification. *Industrial & Engineering Chemistry Research*, 32:1981–1992.

- Eykhoff, P. (1981). Trends and Progress in System Identification. In *IFAC series*. Pergamon, Oxford.
- Fiedler-Ferrara, N. e Prado, C. P. C. (1994). *Caos - Uma Introdução*. Editora Edgard Blücher.
- Floriani, E., Dudok de Wit, T., e Le Gal, P. (2000). Nonlinear interactions in a rotating disk flow: From a Volterra model to the Ginzburg-Landau equation. *Chaos*, 10(4):834–847.
- Fonseca, C. M., Mendes, E. M. A. M., Fleming, P. J., e Billings, S. A. (1993). Nonlinear model term selection with genetic algorithms. In *IEEE/IEEE Workshop on Natural Algorithms in Signal Processing*, páginas 27/1–27/8, Chelmsford, Essex.
- Forsell, U. e Lindskog, P. (1997). Combining semi-physical and neural network modeling: an example of its usefulness. In *Proceedings of the 11th IFAC Symposium on System Identification*, volume 4, páginas 795–798, Kitakyushu, Fukuoka, Japan.
- Freitas, U. S. (2001). *Uso de Técnicas de Detecção de Estrutura na Identificação de Modelos Dinâmicos Contínuos Não-Lineares Polinomiais*. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Gama, C. (1999). Brazilian north-south interconnection control application and operating experience with a TCSC. In *Proceedings of IEEE/Power Engineering Society - 1999 Summer Meeting*, páginas 1103–1108, Edmonton, Canada.
- Gama, C. e Tenorio, A. R. M. (2000). Improvements for power systems performance: Modeling, analysis and benefits of TCSCs. In *Conf. Proc. IEEE PES Winter Meeting*, Singapore.
- Garcia, C. (1997). *Modelagem e simulação de processos industriais e de sistemas eletromecânicos*. EDUSP, São Paulo.
- Georgiou, I. T. e Scharz, I. B. (1999). Dynamics of large scale couples structural/mechanical systems: A singular perturbation/proper orthogonal decomposition approach. *SIAM Journal on Applied Mathematics*, 59(4):1178–1207.

- Gibson, J. F., Farmer, J. D., Casdagli, M., e Eubank, S. (1992). An analytic approach to practical state space reconstruction. *Physica D*, 57(1–2):1–20.
- Gleick, J. (1987). *Chaos, making a new science*. Abacus, London.
- Goldberger, A. L. (1991). Getting to the heart of chaos: nonlinear dynamics in clinical cardiology. In *In: Proceedings of the First Experimental Chaos Conference*, Singapore: World Scientific.
- Gouesbet, G. e Letellier, C. (1994). Global vector-field reconstruction by using a multivariate polynomial l_2 approximation on nets. *Physical Review E*, 49(6):4955–4972.
- Grayling, A. C. (1996). *Epistemology*. Blackwell Publishers Ltd. Tradução de Paulo Ghiraldelli. (<http://www.filosofia.pro.br/textos/a-epistemologia.htm>).
- Greblicki, W. (1996). Nonlinearity estimation in Hammerstein systems based on ordered observations. *IEEE Transactions on Signal Processing*, 44(5):1224–1233.
- Haber, R. (1985). Nonlinearity tests for dynamic processes. In *IFAC identification and system parameter estimation*, páginas 409–413, York, UK.
- Haykin, S. (1994). *Neural Networks - a comprehensive foundation*. London: McMillan.
- Haykin, S. e Van Veen, B. (2001). *Sinais e Sistemas*. Porto Alegre: Bookman. Tradução: Santos, José Carlos Barbosa dos.
- Herbert, J. A. e Tulleken, A. F. (1993). Grey-box modeling and identification using physical knowledge and Bayesian techniques. *Automatica*, 29(2):285–308.
- Hingorani, N. G. (1994). *FACTS Technology and Opportunities*. Proceeding of Colloquium on Flexible AC Transmission Systems (FACTS): the Key to Increased Utilisation of Power Systems.
- Holanda, A. B. (1988). *Dicionário Aurélio Básico da Língua Portuguesa*. Rio de Janeiro: Nova Fronteira.

- Hotelling, H. (1933). Analysis of a complex of statistical variable into principal components. *Journal of Educational Psychology*, 24:10–21.
- Irving, A. D. e Dewson, T. (1997). Determining mixed linear-nonlinear coupled differential equations from multivariate discrete time series sequences. *Physica D*, 102:15–36.
- Jaynes, E. T. (1996). *Probability Theory: The Logic of Science*. St. Louis, USA, Washington University (Fragmented Edition).
- Jácome, C. R. F. (1996). *Uso de Conhecimento Prévio na Identificação de Modelos Polinomiais NARMAX*. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Johansen, T. (1996). Identification of non-linear systems using empirical data and prior knowledge - an optimization approach. *Automatica*, 32(3):337–356.
- Johansen, T. (1997). Constrained and regularized system identification. In *Proceedings of the 11th IFAC Symposium on System Identification*, Kitakyushu, Fukuoda, Japan.
- Johansen, T. A. (2000). Multi-objective identification of FIR models. In *Proceedings of 12th IFAC Symposium on System Identification 2000*, Santa Barbara, USA.
- Jolliffe, I. T. (1986). *Principal Component Analysis*. New York: Springer.
- Jorgensen, S. B. e Hangos, K. M. (1995). Grey box modelling for control: Qualitative models as a unifying framework. *International Journal of Adaptive Control and Signal Processing*, 9(6):547–562.
- Kadtke, J. B., Brush, J., e Holzfuss, J. (1993). Global dynamical equations and Lyapunov exponents from noisy chaotic time series. *International Journal of Bifurcation and Chaos*, 3(3):607–616.
- Kárný, M., Halousková, A., e Nedoma, P. (1995a). Recursive approximation by ARX model: A tool for grey box modelling. *International Journal of Adaptive Control and Signal Processing*, 9(6):525–546.

- Kárný, M., Nedoma, P., Böhm, J., e Halousková, A. (1995b). Approximate ARX model estimation for jacketing adaptive systems. In Banyasz Cs, editor, *In: Adaptive Systems in Control and Signal Processing. Preprints*, páginas 129–134, Budapest. IFAC.
- Kennedy, M. P. (1992). Robust OP Amp realization of Chua's circuit. *Frequenz*, 46(3–4):66–80.
- Kennedy, M. P. (2000). Chaotic modulation for robust digital communications over multipath channels. *International Journal of Bifurcation and Chaos*, 10(4):695–718.
- Korenberg, M., Billings, S., Liu, Y., e Mcilroy, P. (1988). Orthogonal parameter estimation algorithm for non-linear stochastic systems. *International Journal of Control*, 48(1):193–210.
- Kugiumtzis, D. (1996). State space reconstruction parameters in the analysis of chaotic time series – the role of time series length. *Physica D*, 95:13–28.
- Lainscsek, C. S. M., Schürerer, F., e Kadtke, J. (1998). A general form for global dynamical data models for three-dimensional systems. *International Journal of Bifurcation and Chaos*, 8(5):899–914.
- Leontaritis, I. J. e Billings, S. A. (1985a). Input-output parametric models for non-linear systems - part i: deterministic non-linear systems. *International Journal of Control*, 41(2):303–328.
- Leontaritis, I. J. e Billings, S. A. (1985b). Input-output parametric models for non-linear systems - part ii: stochastic non-linear systems. *International Journal of Control*, 41(2):329–344.
- Letellier, C., Maquet, J., Le Sceller, L., e Aguirre, L. A. (1998). On the non-equivalence of observables in phase-space reconstructions from recorded time series. *Journal of Physics A*, 31:7913–7927.
- Lindskog, P. (1996). *Methods, Algorithms and Tools for System Identification Based on Prior Knowledge*. Tese de Doutorado, Department of Electrical Engineering, Linköping University, Sweden.
- Lindskog, P. e Ljung, L. (1995). A comparison between semi-physical and black-box neural net modeling: a case study. In Bulsari, A. B. e Kallio, S.,

- editors, *Proceedings of the International Conference: Engineering Applications of Artificial Neural Networks, Helsinki, Finland*, páginas 235–238.
- Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall, London.
- Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of Atmospheric Sciences*, 20:283–293.
- Lorito, F. (1998). Identification of a grey-box model of nonlinear current transformer for simulations purposes. *Control Engineering Practice*, 6:1331–1339.
- Martins, C. R. (1992). *Século XIX: o nascimento da ciência contemporânea*, capítulo Plank e o Nascimento da Mecânica Quântica. Centro de Lógica, Epistemologia e História da Ciência, Unicamp.
- MathWorks, T. (1997). *Wavelet Toolbox User's Guide*.
- Mendes, E. M. A. M. e Billings, S. A. (1998). On overparametrization of nonlinear discrete systems. *International Journal of Bifurcation and Chaos*, 8(3):535–556.
- Murakami, K. e Seborg, D. E. (2000). Constrained parameter estimation with application to blending operations. *Journal of Process Control*, 10:195–202.
- Nepomuceno, E. G., Takahashi, R. H. C., Aguirre, L. A., Neto, O. M., e Mendes, E. M. A. M. (2002a). Multiobjective nonlinear system identification: a case study with thyristor controlled series capacitor (TCSC). *Anais do XIV Congresso Brasileiro de Automática*. (Aceito para publicação).
- Nepomuceno, E. G., Takahashi, R. H. C., Amaral, G. F. V., e Aguirre, L. A. (2002b). Nonlinear identification using prior knowledge of fixed points: a multiobjective approach. *International Journal of Bifurcation and Chaos*. (accepted).
- Neto, O. M., Jesus, E. G., Nepomuceno, E. G., e Mendes, E. M. A. M. (2000). Controle de um TCSC utilizando representação em espaço de estados e controlador integral. *Anais do XIII Congresso Brasileiro de Automática*, páginas 1085–1089.

- Omlin, M. e Reichert, P. (1999). A comparison of techniques for the estimation of model prediction uncertainty. *Ecological Modelling*, 115:45–59.
- Packard, N. H., Crutchfield, J. P., Farmer, J. D., e Shaw, R. S. (1980). Geometry from a time series. *Physical Review Letters*, 45(9):712–716.
- Paserba, J. J., Miller, N. W., Larsen, E. V., e Piwko, R. J. (1995). A thyristor controlled series compensation model for power system stability analysis. *IEEE Transactions on Power Delivery*, 10(3).
- Pearson, R. K. e Pottmann, M. (2000). Gray-box identification of block-oriented nonlinear models. *Journal of Process Control*, 10(301–315).
- Petrack, M. H. e Wigdorowitz, B. (1997). A priori nonlinear model structure selection for system identification. *Control Engineering Practice*, 5(8):1053–1062.
- Peyton-Jones, J. e Billings, S. (1989). Recursive algorithm for computing the frequency response of a class of non-linear difference equation models. *International Journal of Control*, 50(5):1925–1940.
- Poincaré, H. (1957). *Les methodes nouvelles de la mecanique celeste*. New York: Dover.
- Popper, K. R. (1974). *A Lógica da Pesquisa Científica (trad.)*. (Logic der Forschung, 5a. ed. 1973; 1a. ed. 1934). Ed. Cultrix.
- Pottmann, M. e Pearson, R. (1998). Block-oriented NARMAX models with output multiplicities. *AIChE Journal*, 44(1):131–140.
- Rosenstein, M. T., Collins, J. J., e De Luca, C. J. (1994). Reconstruction expansion as a geometry-based framework for choosing proper delay times. *Physica D*, 73:82–98.
- Sauer, T., Yorke, J. A., e Casdagli, M. (1991). Embedology. *Journal of Statistical Physics*, 65(3/4):579–616.
- Schiff, S. J., Jerger, K., Duong, D. H., Chang, T., e Spano, M. L. Ditto, W. L. (1994). Controlling chaos in the brain. *Nature*, 370(6491):615–620.

- Schwartzman, S. (2001). *Um espaço para a ciência: a formação da comunidade científica no Brasil*. Ministério da Ciência e Tecnologia, Centro de Estudos Estratégicos, Brasília.
- Seborg, D. E., Edgar, T. F., e Mellichamp, D. A. (1989). *A Process Dynamics and Control*. New York: John Wiley.
- Sen, P. C. (1989). *Principles of Electrical Machines and Power Electronics*. John Wiley & Sons, Ontario.
- Shannon, C. E. (1949). Inductorless chua's circuit. *Proceeding of the Institute of Radio Engineers*, 37(1):10–21.
- Simutis, R., Oliveira, R., Manikowski, Azevedo, S. F. d., e Lübbert, A. (1997). How to increase the performance of models for process optimization and control. *Journal of Biotechnology*, 59:73–89.
- Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Delyon, B., Glorennec, P., Hjalmarsson, H., e Juditsky, A. (1995). Nonlinear black-box modeling in system identification: A unified overview. *Automatica*, 31:1691.
- Suykens, J. A. K. e Vandewalle, J. (1995). Learning a simple recurrent neural state space model to behave like Chua's double scroll. *IEEE Transactions on Circuits and Systems Part I: Fundamental Theory and Applications*, 42(8):499–502.
- Takahashi, R. H. C. (1998). *Controle Singular de Sistemas Incertos*. Tese de Doutorado, Faculdade de Engenharia Elétrica e Computação, Universidade de Campinas.
- Takahashi, R. H. C. (2001). *Notas de Aula da Disciplina Otimização Multiobjetivo*. Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais, Belo Horizonte.
- Takens, F. (1981). Detecting strange attractors in turbulence. In Rand, D. A. e Young, L. S., editors, *Dynamical systems and turbulence, Lecture Notes in Mathematics, vol. 898*, páginas 366–381. Springer Verlag, Berlin.
- Tan, K. C., Li, Y., Gawthrop, P. J., e Glidle, A. (1997). Evolutionary grey-box modelling for nonlinear systems. In *In Proceedings of the 2nd International Conference on Genetic Algorithms in Engineering Systems: Innovations and Applications*.

- Teixeira, R. A. (2001). *Treinamento de redes neurais artificiais através de otimização multi-objetivo: uma nova abordagem para o equilíbrio entre a polarização e a variância*. Tese de Doutorado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Thompson, M. L. e Kramer, M. A. (1994). Modeling chemical processes using prior knowledge and neural networks. *AIChE Journal*, 40(8):1328–1340.
- Tikhonov, A. N. e Arsenin, V. Y. (1977). *Solutions of Ill-posed Problems*. Winston, Washington DC.
- Tôrres, L. e Aguirre, L. (2000). Inductorless chua's circuit. *Electronic Letters*, 36(23):1915–1916.
- Tôrres, L. A. B. (2001). *Controle e Sincronismo de Osciladores Caóticos*. Tese de Doutorado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Tseng, C. H. (2002). Bandpass sampling criteria for nonlinear systems. *IEEE Transactions on Signal Processing*, 50(3):568–577.
- Tulleken, H. J. A. F. (1993). Grey-box modelling and identification using physical knowledge and Bayesian techniques. *Automatica*, 29(2):285–308.
- Weibert, K. e Main, J. (2001). Deterministic diffusion in one-dimensional maps—calculation of diffusion constants by harmonic inversion of periodic orbit sums. *Physics Letters A*, 292(1-2):120–124.
- Weyer, E., Gábor, S., e Katalin, H. (2000). Grey box fault detection of heat exchangers. *Control Engineering Practice*, 8:121–131.
- Wigren, T. (1993). Recursive prediction error identification using the nonlinear Wiener model. *Automatica*, 17(1):1011–1025.
- Williams, H. S. (1904). *A History of Science: in Five Volumes. Volume II*. Harper & Brothers. (Disponível em <http://etext.lib.virginia.edu/toc/modeng/public/Wil2Sci.html>).

Rotinas Computacionais

“É sempre oportuno salientar que, olhada a educação como investimento e aceito que, por isso, a nação realize nela grande inversão de dinheiros públicos, indo ao encontro do povo para educá-lo, em vez de esperar que cada qual se eduque ou deixe de educar, consoante seus caprichos, e tornando barato o ensino e todo elemento necessário a esse objetivo, muito grande se faz a responsabilidade dos que, dentro desse sistema, funcionam como professores ou como alunos.”

José Reis¹

Esse Apêndice tem como objetivo apresentar as principais rotinas computacionais empregadas na dissertação. Todas rotinas foram desenvolvidas no Matlab, dada a facilidade de programação desse programa e pelo histórico do grupo de pesquisas MACSIN, que possui a maioria de suas rotinas implementadas nessa linguagem. As rotinas se destinam a modelos NARMAX polinomiais SISO (*Single-Input and Single Output*, ou uma entrada e uma saída).

A.1 genterms

Sintaxe

$[\text{modelo}, n_t] = \text{genterms}(\ell, n_y, n_u, n_e)$

¹Um dos fundadores da Sociedade Brasileira para o Progresso da Ciência (SBPC), presidente do CNPq e um dos principais divulgadores da Ciência que o Brasil já teve. Faleceu no dia 16 de maio de 2002. Minhas homenagens!

Descrição

Utilizada para gerar o conjunto dos termos candidatos.

Entrada

ℓ : grau de não-linearidade
 n_y : máximo atraso do regressor de saída
 n_u : máximo atraso do regressor de entrada
 n_e : máximo atraso do regressor de ruído

Saída

n_t : número de termos candidatos
modelo : matriz contendo os termos de processo e ruído

Eduardo Mendes - 11/08/94 - Carlos Fonseca
ACSE - Sheffield

A.2 mcand

Sintaxe

[modelo]=mcand(modelo, p , m)

Descrição

Permite retirar do conjunto de termos candidatos aqueles que fazem parte dos agrupamentos de termos considerados espúrios.

Entrada

modelo : conjunto de termos candidatos
 p : agrupamento da saída
 n_u : agrupamento da entrada

Saída

modelo : novo conjunto de termos candidatos

Giovani Rodrigues

Eduardo Mendes - 11/08/95

A.3 akaike

Sintaxe

$[f, min_{n_t}] = \text{akaike}(\text{modelo}, u, y, f_1, f_2)$

Descrição

Calcula o número aproximado de termos que um modelo necessita baseado no critério de informação de Akaike.

Entrada

modelo : conjunto dos termos candidatos

u : dados de entrada

y : dados de saída

f_1 : (1) mostra resultado gráfico, (-1) não mostra resultado gráfico

f_2 : (1) retorna o número da figura

Saída

f : seqüência contendo os dados do critério de Akaike

min_{n_t} : número de termos a ser considerado

Giovani Rodrigues

Eduardo Mendes - 11/08/95

A.4 myhouse

Sintaxe

$[err] = \text{myhouse}(\Psi, n_p)$

Descrição

Calcula o valor de ERR de cada termo candidato.

Entrada

Ψ : matriz de regressores
 n_p : número de termos de processo

Saída

err : vetor do ERR para cada termo candidato

Grupo MACSIN

A.5 orthreg

Sintaxe

$[\text{modelo}, x, e, v] = \text{orthreg}(\text{modelo}, u, y, [n_p n_e], N)$

Descrição

Seleciona a estrutura e estima os parâmetros via MQE.

Entrada

modelo : conjunto de termos candidatos
 u : dados de entrada
 y : dados de saída
 n_p : número de termos de processo
 n_e : número de termos de ruído
 N : número de iteração de ruído

Saída

modelo : conjunto de termos selecionado
 x : matriz contendo: [(parâmetros) (err) (desvio)]
 e : resíduo
 v : variância

Eduardo Mendes - 11/08/94
ACSE - Sheffield

A.6 simodeld

Sintaxe

$[y]=\text{simodeld}(\text{modelo},x0,u,y0,e)$

Descrição

Retorna a predição infinitos passos à frente.

Entrada

modelo : modelo identificado
 $x0$: parâmetros
 u : dados de saída
 $y0$: condições iniciais
 e : ruído

Saída

y : saída predita infinitos passos à frente

Eduardo Mendes - 03/09/94
ACSE - Sheffield

A.7 build_pr

Sintaxe

$[\Psi]=\text{build_pr}(\text{modelo},u,y)$

Descrição

Retorna a matriz de regressores.

Entrada

modelo : conjunto de termos
 u : dados de entrada
 y : dados de saída

Saída

Ψ : matriz de regressores

Eduardo Mendes - 11/08/94
ACSE - Sheffield

A.8 build_no

Sintaxe

$[\Psi^*]=\text{build_no}(\text{modelo},u,y,e)$

Descrição

Retorna a matriz de regressores estendida.

Entrada

modelo : conjunto de termos
 u : dados de entrada
 y : dados de saída
 e : dados de ruído

Saída

Ψ^* : matriz de regressores estendida

Eduardo Mendes - 11/08/94

ACSE - Sheffield

A.9 fpoints

Sintaxe

$[pf]=\text{fpoints}(\text{modelo},x0)$

Descrição

Retorna os pontos fixos de um modelo NAR.

Entrada

modelo : conjunto de termos

$x0$: parâmetros

Saída

pf : matriz de regressores estendida

Grupo MACSIN

A.10 fixedp

Sintaxe

$[\text{sigma},\text{FP},\text{fp},\text{fpstd}]=\text{fixedp}(y, L,\text{delta},\text{see})$

Descrição

Estima pontos fixos usando uma estrutura não-linear de treze termos e série temporal.

Entrada

y : série temporal
 L : janela de dados
delta : passo entre as janelas
see : (1) mostra gráficos

Saída

sigma : matriz dos coeficientes de agrupamento para cada janela
FP : pontos fixos estimados para cada janela
fp : estimação dos pontos fixos
fpstd : desvio padrão da estimação

Álvaro Souza
MACSIN - UFMG

A.11 **matrizs**

Sintaxe

[S]=matrizs(modelo)

Descrição

Retorna o mapeamento linear do espaço de parâmetros para o espaço de coeficientes de agrupamento ($\Sigma = S\theta$).

Entrada

modelo : conjunto de termos

Saída

S : mapeamento linear

Erivelton Geraldo Nepomuceno - 19/06/2002
MACSIN - UFMG

A.12 lyapunov

Sintaxe

$[ml]=\text{lyapunov}(\text{modelo}, x_0, y_0, n, T_s)$

Descrição

Retorna o maior expoente de Lyapunov (ML) de um modelo NARMA, derivado de uma série temporal.

Entrada

modelo : conjunto de termos (apenas regressores de saída)
 x_0 : parâmetros
 y_0 : condições iniciais
 n : número de amostras usado nos cálculos
 T_s : tempo de amostragem

Saída

ml : maior expoente de Lyapunov

Giovani Rodrigues
Erivelton Geraldo Nepomuceno - 19/06/2002
MACSIN - UFMG

A.13 mapa1

Sintaxe

$[x, z]=\text{mapa1}(y, \text{lag}, \text{sec}, \text{sim})$

Descrição

Calcula e plota o mapa de primeiro retorno.

Entrada

y : série temporal
 lag : atraso usado no espaço de fases
 sec : valor da seção
 sim : (1) série simétrica, (0) série não-simétrica

Saída

x : dados do mapa de primeiro retorno
 z : dados simétricos do mapa de primeiro retorno

Erivelton Geraldo Nepomuceno - 19/06/2002
MACSIN - UFMG

A.14 im

Sintaxe

$[\text{modelo}, X0, w, r_d, r_i] = \text{im}(\text{modelo}, u, y, n_p, n_e, u_v, y_v, pf, Q, S, ue, ye, f_1)$

Descrição

Identificação multiobjetivo usando pontos fixos ou curva estática como informação auxiliar.

Entrada

m	: conjunto de termos
u	: dados de entrada
y	: dados de saída
n_p	: número de termos de processo
n_e	: número de termos de ruído
u_v	: dados de entrada (validação)
y_v	: dados de saída (validação)
pf	: pontos fixos do sistema original (informação auxiliar)
ue	: dados de entrada estática (informação auxiliar)
ye	: dados de saída estática (informação auxiliar)
Q	: matriz de regressores estáticos
S	: mapeamento linear
f_1	: (1) uso de pontos fixos, (2) uso da curva estática

Saída

modelo	: conjunto de termos
X_0	: matriz de parâmetros do conjunto Pareto-ótimo
w	: pesos usados no problema P_w
r_d	: erro quadrático para cada modelo do Pareto-ótimo (validação dinâmica)
r_i	: erro quadrático para cada modelo do Pareto-ótimo (informação auxiliar)

Erivelton Geraldo Nepomuceno - 19/06/2002
MACSIN - UFMG

A.15 decisor

Sintaxe

$[p^*]=\text{decisor}(\text{modelo}, X_0, r_d, r_i, f_1)$

Descrição

Escolha da solução ótima pertencente ao Pareto-ótimo.

Entrada

- modelo : conjunto de termos
 X_0 : matriz de parâmetros do Pareto-ótimo
 r_d : erro quadrático para cada modelo do
Pareto-ótimo (validação dinâmica)
 r_i : erro quadrático para cada modelo do
Pareto-ótimo (informação auxiliar)
 f_1 : (1) decisor: equilíbrio entre variância e polarização
(2) decisor: norma mínima dos objetivos normalizados

Saída

- p^* : solução ótima

Erivelton Geraldo Nepomuceno - 19/06/2002

MACSIN - UFMG

Índice

- autocorrelação
 - linear, 13
 - não-linear, 13
- conhecimento *a priori*, 3
- critério de Akaike, 30
- decimação, 13, 127
- Engenharia, 3
- epistemologia, 5
 - justificativa epistemológica, 7
- ERR, 29, 44
- esperança matemática, 13
- estudos de caso
 - aquecedor elétrico, 102
 - circuito de Chua, 76
 - conversor buck, 111
 - mapa senoidal, 93
 - TCSC, 121
- Identificação de Sistemas, 1
 - caixa-branca, 40
 - caixa-cinza, 37, 41
 - caixa-preta, 11, 40
 - classificação, 40
 - comparação, 43
 - etapas, 11, 29
 - determinação de estruturas, 28
 - escolha da representação, 14
 - estimação de parâmetros, 31
 - testes dinâmicos, 12
 - validação, 34
- identificação multiobjetivo, 55
 - estado da arte, 37
 - motivação, 3
 - Pareto-ótimo, 58
 - Teorema, 62
- informação auxiliar, 4
 - comparação com *a priori*, 37
 - curva estática, 66
 - definição, 38
 - pontos fixos, 65
- mínimos quadrados estendidos, 32
- mínimos quadrados restritos, 49
- matriz
 - de regressores, 31
 - de regressores estáticos, 66
 - informação, de, 32
 - matriz estendida regressores, 34
 - pseudo inversa, 32
 - S, mapeamento linear, 67
- metodologia mono-objetivo, 4
- metodologia multiobjetivo, 4
- modelos, 2
 - autônomos, 2
 - contínuos, 2

- dinâmicos, 2
 - discretos, 2
 - estáticos, 2
 - lineares, 2
 - não-autônomos, 2
 - não-lineares, 2
- não-linearidade
 - detecção de, 14
 - diagramas de bifurcação, 35
 - expoentes de Lyapunov, 35
 - mapa primeiro retorno, 35
- NARMAX, 3
 - propriedades, 21
 - agrupamento de termos, 22
 - coef. de agrupamento, 23
 - curva estática, 66
 - pontos fixos, 24
 - pontos fixos, simetria, 26
 - sobreparametrização, 23
- norma euclidiana, 17, 65
- Pareto-ótimo, 58, 60, 61, 66, 67, 69, 75, 82, 89, 96, 106, 115, 137
- problema da decisão, 67
 - necessidade, 68
 - norma mínima, 70, 109, 137
 - polarização e variância, 69, 89
 - preferências *a posteriori*, 68
 - preferências *a priori*, 68
 - preferências *progressivas*, 68
- representação matemática
 - Hammerstein e Wiener, 15
 - modelos polin. contínuos, 16
 - NARMAX, 19
 - RBF, 17
 - redes neurais, 3, 18
 - série de Volterra, 15
 - wavelets, 18
- RMSE, 35
- rotinas
 - akaike, 159
 - build_no, 162
 - build_pr, 161
 - decisor, 167
 - fixedp, 163
 - fpoints, 163
 - genterms, 157
 - im, 166
 - lyapunov, 165
 - mapa1, 165
 - matrizs, 164
 - mcand, 158
 - myhouse, 159
 - orthreg, 160
 - simodeld, 161
- sinais de excitação, 12
 - PRBS, 12
- sistema, 1
- tempo de amostragem, 13
 - Shannnon, teorema de, 13
- Transformada
 - Fourier, de, 19
 - Laplace, de, 15
 - wavelet, de, 18