

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

Programa de Pós-Graduação em Engenharia Elétrica

Universidade Federal de Minas Gerais

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

Fone: +55 3409-4866 - Fax: +55 3409-4850

alipiomonteiro@yahoo.com.br



Técnicas de Otimização Bi-Objetivo para a Determinação da Estrutura de Modelos NARX

Alípio Monteiro Barbosa

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. Luis Antonio Aguirre

Co-orientador: Dr. Erivelton Geraldo Nepomuceno

Belo Horizonte, 16 de dezembro de 2010

Dedicatória

“Tive toda sorte de privilégios. Tive a vantagem de viver numa casa em que as pessoas falavam sobre coisas interessantes, tive pais inteligentes e frequentei boas escolas. Duvido que pudesse ter sido criado por pais melhores que os meus. Isso foi de uma importância enorme. Meus pais não me deram dinheiro, e sinceramente eu não queria nenhum. Mas nasci no lugar e na hora certos. Tirei a sorte grande na Loteria do Ovário.”

Warren Buffett

Agradecimentos

Agradeço a Deus!

Aos meus pais, Dona Marta e Senhor Alípio. Eles são fera! Aos meus irmãos, Érica e Felipe, por todo carinho.

Aos meus amigos, grandes amigos!

Aos orientadores Luis Aguirre, Ricardo Takahashi e Erivelton Nepomuceno pela dedicação, paciência e ensinamentos.

Ao Professor Márcio Falcão pelo fundamental e constante suporte!

Ao CEFET - MG, Coordenação de Eletrônica, em especial ao Professor Rubens Santos.

Aos colegas do MACSIN, CPH e PPGEE.

Aos colegas do GCoM (UFSJ).

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, pelo suporte financeiro.

Epígrafe

“Há muitas razões para duvidar e uma só para crer.”

Carlos Drummond de Andrade

*“... Como um velho boiadeiro levando a boiada
Eu vou tocando os dias pela longa estrada eu vou
Estrada eu sou ...”*

Almir Sater e Renato Teixeira

Conteúdo

Resumo	xiii
Abstract	xv
Lista de Tabelas	xvii
Lista de Figuras	xxi
1 Introdução	1
1.1 Objetivo	3
1.2 Estrutura da dissertação	4
2 Fundamentação Teórica	5
2.1 Introdução	5
2.2 Modelos polinomiais NARX	5
2.2.1 Agrupamento de termos	6
2.2.2 Curva estática	7
2.3 Determinação de estrutura	9
2.3.1 Determinação de estrutura utilizando ERR	10
2.3.2 Critério de informação de Akaike	11
2.4 Identificação multiobjetivo	12
2.4.1 A solução Pareto-ótima	12
2.4.2 O decisor de mínima correlação	13

2.4.3	O decisor de melhor sincronismo	13
2.5	O método LASSO	16
2.6	Validação do modelo	16
2.7	Conclusões do capítulo	18
3	Procedimento Bi-Objetivo para Seleção de Estrutura	19
3.1	Introdução	19
3.2	Definição do problema	19
3.3	Procedimento bi-objetivo	21
3.3.1	Descrição e implementação da rotina I	22
3.3.2	Descrição e implementação da rotina II	24
3.4	Exemplo ilustrado	26
3.4.1	Seleção dos agrupamentos <i>verdadeiros</i>	28
3.4.2	Seleção dos regressores <i>verdadeiros</i>	30
3.5	Conclusões do capítulo	34
4	Resultados de Simulação	35
4.1	Introdução	35
4.2	Exemplos simulados	35
4.2.1	Exemplo 1	36
4.2.2	Exemplo 2	38
4.2.3	Exemplo 3	46
4.3	Conclusões do capítulo	57
5	Resultados Experimentais	59
5.1	Introdução	59
5.2	Conversor Buck	59
5.2.1	Descrição do sistema	60
5.2.2	Característica estática	62
5.2.3	Identificação de modelos	62

5.3	Válvula pneumática	67
5.3.1	Descrição do sistema	67
5.3.2	Descrição de testes e dados	70
5.3.3	Identificação de modelos	71
5.4	Conclusões do capítulo	77
6	Conclusão e Pesquisas Futuras	79
6.1	Introdução	79
6.2	Discussões finais	80
6.3	Pesquisas futuras	80
	Bibliografia	83
A	Rotinas Computacionais	91
A.1	build-r	91
A.2	build-q	92
A.3	build-pr	93
A.4	simodeld	94
A.5	valsinc	95
A.6	myccf	95
A.7	orthreg	96
A.8	ntermos	97
A.9	akaike	98
A.10	lasso	98

Resumo

A identificação de sistemas tem como finalidade obter um modelo matemático que descreva o comportamento de um sistema dinâmico a partir de medições.

Existem algumas etapas típicas no processo de construção de modelos matemáticos a partir de dados: i) execução de testes e coletas de dados, ii) escolha da classe de modelos, iii) escolha da estrutura do modelo, iv) estimação de parâmetros e v) validação de modelos. Neste trabalho investigou-se o problema da determinação de estruturas.

Desenvolveu-se um critério baseado em técnicas multiobjetivo para auxiliar na determinação de estrutura de modelos NARX polinomiais. Técnicas multiobjetivo, diferentemente da abordagem mono-objetivo, são capazes 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. Portanto, o objetivo consistiu em determinar a estrutura de um modelo NARX polinomial por meio de análises das curvas de Pareto. Curvas geradas pela variação combinacional de possíveis estruturas do modelo.

A metodologia proposta foi aplicada na identificação de três exemplos simulados e dois exemplos experimentais: um conversor buck e uma válvula pneumática. Os resultados simulados mostraram que é possível discriminar estruturas subparametrizadas e que, dado um modelo sobreparametrizado que contenha a estrutura nominal, é possível discriminar entre termos espúrios e genuínos. Os resultados experimentais indicaram que o procedimento bi-objetivo pode ser empregado como uma ferramenta auxiliar na determinação de estruturas de modelos polinomiais NARX.

Palavras-chave: Identificação de sistemas não-lineares, otimização bi-objetivo, seleção de estruturas de modelos.

Abstract

The system identification aims to obtain a mathematical model that describes the behavior of a dynamic system from measurements.

There are some typical steps in the process of building mathematical models from data: i) tests realization and data collection, ii) selection of the class of models, iii) structure detection, iv) parameters estimation and v) model validation. In this study we investigated the problem of structure detection.

In this work, one criterion based on multiobjective techniques was developed to help to determine the structure of polynomial NARX models. Multiobjective techniques, unlike the mono-objective approach, are able to find intermediate solutions that take into account the compromise between the goals, not just individual solutions. This set of solutions is called *Pareto-optimal*. Therefore, the aim was to determine the structure of a polynomial NARX model, using Pareto curves. The curves were generated by the combination of possible structures of the model.

The methodology proposed in this work was applied to identify three simulated examples and two experimental examples: a buck converter and a pneumatic valve.

Simulation results have shown that it is possible to distinguish between under and overparametrization scenarios and to classify spurious from genuine regressors in a model. Experimental results showed that the bi-objective method can assist in the selection of polynomial NARX model structures.

Keywords: Nonlinear system identification, biobjective optimization, model structure selection.

Lista de Tabelas

4.1	Regressores acrescentados e removidos de \mathcal{M}_1 (ver Equação 4.2).	41
4.2	Modelo final obtido pelo critério bi-objetivo a partir dos dados apresentados nas Figuras 4.15 e 4.16. Termos genuínos em negrito.	53
4.3	Modelo obtido pelo critério de informação de Akaike e ERR a partir dos dados apresentados na Figura 4.15. Conjunto inicial de 56 termos candidatos. Termos genuínos em negrito.	53
4.4	Modelo obtido pelo critério de informação de Akaike e ERR a partir dos dados apresentados na Figura 4.15. Conjunto inicial de 15 termos candidatos. Termos genuínos em negrito.	54
5.1	Índice RMSE do modelo (5.2) e número de parâmetros (n_p) comparado com outros modelos.	66

Lista de Figuras

1.1	Etapas da identificação de sistemas.	2
2.1	Representação geométrica das normas ℓ_1 , ℓ_2 e ℓ_∞	17
2.2	Representação gráfica do método LASSO e <i>ridge regression</i> . . .	17
3.1	Representação de um conjunto Pareto-ótimo de um problema bi- objetivo	20
3.2	Representação dos conjuntos Pareto-ótimos \mathcal{P}_1 , \mathcal{P}_2 e \mathcal{P}_3	21
3.3	Fluxograma do procedimento bi-objetivo para seleção de estru- tura, etapa de seleção dos agrupamentos	23
3.4	Métrica de desempenho: área	24
3.5	Fluxograma do procedimento bi-objetivo para seleção de estru- tura, etapa de seleção do modelo nominal	25
3.6	Dados dinâmicos de identificação e validação: Exemplo ilustrado	27
3.7	Dados estáticos: Exemplo ilustrado	27
3.8	Conjuntos Pareto dos modelos candidatos: seleção dos agrupa- mentos verdadeiros	30
3.9	Conjuntos Pareto dos modelos candidatos \mathcal{G}_1 , \mathcal{G}_6 e \mathcal{G}_8 a \mathcal{G}_{11} . . .	31
3.10	Curvas Pareto dos modelos candidatos \mathcal{M}_1 a \mathcal{M}_5	32
3.11	Curvas Pareto dos modelos candidatos \mathcal{M}_1 a \mathcal{M}_4	33
3.12	Erro máximo de sincronização	33
3.13	Custo de sincronização	34

4.1	Dados dinâmicos de identificação e validação: Exemplo 1	37
4.2	Dados estáticos: Exemplo 1	37
4.3	Conjuntos Pareto dos modelos candidatos: Exemplo 1	38
4.4	Função de autocorrelação do sinal de entrada – Exemplo 2	39
4.5	Dados dinâmicos de identificação e validação: Exemplo 2	40
4.6	Dados estáticos: Exemplo 2	40
4.7	Conjuntos Pareto dos modelos candidatos (ver Tabela 4.1)	42
4.8	Conjuntos Pareto dos modelos \mathcal{M}_1 a \mathcal{M}_{10} (ver Tabela 4.1)	43
4.9	Erro máximo de sincronização (ver Equação 2.31)	44
4.10	Custo da sincronização (ver Equação 2.30) para os mesmos mo- delos da Figura 4.9.	44
4.11	Curvas Pareto: região de incerteza.	45
4.12	Curvas Pareto: extensão do Exemplo 2.	47
4.13	Erro de sincronização: extensão do Exemplo 2.	47
4.14	Função de autocorrelação do sinal de entrada – Exemplo 3	48
4.15	Dados dinâmicos de identificação e validação: Exemplo 3	49
4.16	Dados estáticos: Exemplo 3	49
4.17	Conjunto Pareto dos modelos candidatos: seleção dos agrupa- mentos verdadeiros, Exemplo 3	51
4.18	Conjunto Pareto dos modelos candidatos: seleção dos regresso- res verdadeiros, Exemplo 3	52
4.19	Validação dinâmica do modelo 4.3	54
4.20	Validação estática do modelo 4.3	55
4.21	Resultado gráfico da aplicação do critério de informação de Akaike	56
5.1	Conversor buck	60
5.2	Dados de identificação – conversor buck	61
5.3	Dados de validação – conversor buck	61

5.4	Dados estáticos: conversor buck	63
5.5	Curvas Pareto – conversor buck	64
5.6	Validação dinâmica do modelo (5.2) – conversor buck	65
5.7	Validação estática do modelo (5.3) – conversor buck	66
5.8	Sistema de controle de nível	67
5.9	Diagrama funcional do sistema de controle de nível	68
5.10	Diagrama em blocos do sistema de aquisição e tratamento dos dados	69
5.11	Tela principal do sistema supervisorio	69
5.12	Entrada e saída da válvula	70
5.13	Dados estáticos da válvula	71
5.14	Validação dinâmica do modelo (5.4) da válvula pneumática . . .	73
5.15	Validação estática do modelo (5.4) da válvula pneumática . . .	73
5.16	Validação dinâmica do modelo (5.4) da válvula pneumática. Prioridade para J_{SF}	74
5.17	Validação estática do modelo (5.4) da válvula pneumática. Prioridade para J_{SF}	74
5.18	Curvas Pareto - válvula pneumática	76
5.19	Validação dinâmica do modelo (5.5) da válvula pneumática . . .	77
5.20	Validação estática do modelo (5.5) da válvula pneumática . . .	78

Introdução

“O estudante deve entender o problema. Mas não basta que ele o entenda. É necessário que ele deseje sua solução.”

George Pólya

A identificação de sistemas tem como finalidade obter um modelo matemático que descreva o comportamento de um sistema dinâmico a partir de medições.

Existem algumas etapas típicas no processo de construção de modelos matemáticos a partir de dados: i) execução de testes e coletas de dados, ii) escolha da classe de modelos, iii) escolha da estrutura do modelo, iv) estimação de parâmetros e v) validação de modelos. Na Figura 1.1 é ilustrado o processo de identificação.

A etapa de estimação de parâmetros é, possivelmente, a mais antiga, e remonta aos dias de Gauss e Legendre (Sorenson, 1970). Na primeira metade do século passado, o estimador de mínimos quadrados foi um importante algoritmo para estimar constantes desconhecidas em modelos que são lineares em tais parâmetros (Eykhoff, 1968). Infelizmente, esse estimador torna-se tendencioso quando os dados tem ruído correlacionado. Para minimizar essas dificuldades, uma outra família de estimadores foi desenvolvida e que passou a ser chamada de métodos de predição de erro, PEM (*prediction error methods*)

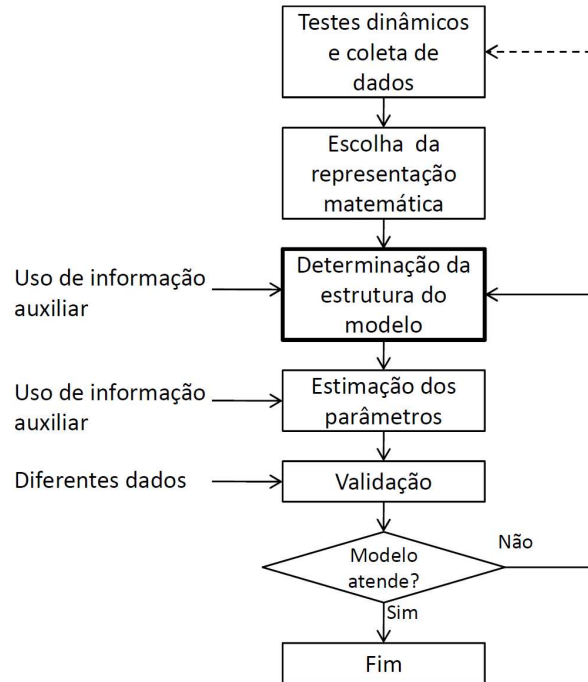


Figura 1.1: Etapas da identificação de sistemas.

(Ljung, 1999). Uma importante característica dessa família de estimadores é que, juntamente com o modelo do processo, é necessário estimar um modelo de ruído para evitar a polarização. Contudo, existem outros estimadores não tendenciosos que não requerem o ajuste de um modelo de ruído, como o estimador de variáveis instrumentais (Young, 1970). Apesar de haver problemas ainda sem solução adequada, tal como a estimação não tendenciosa de modelos com ruído nos dados de entrada (Wei e Billings, 2009), a área de estimação de parâmetros (de modelos lineares e pseudo-lineares nos parâmetros) pode ser considerada madura.

Por outro lado, a etapa de escolha da estrutura do modelo ainda apresenta grandes desafios para os quais não há soluções conclusivas. Uma das primeiras abordagens eficazes para o caso de modelos polinomiais NARX (*nonlinear autoregressive with exogenous inputs*) foi proposta por Billings e seus colaboradores (Billings et al., 1989). Esse procedimento, que se baseia na taxa de redução de erro (ERR *error reduction ratio*), tem falhas que foram descritas na literatura para as quais soluções alternativas foram propostas (Aguirre e

Billings, 1995; Aguirre et al., 1998; Mendes e Billings, 2001; Piroddi e Spinelli, 2003). Apesar do importante esforço feito, a determinação de estrutura de modelos não-lineares ainda permanece sem uma solução definitiva. Em consequência é possível encontrar diversos estudos recentes que abordam esse problema (Piroddi, 2008; Wei e Billings, 2008; Hong et al., 2008; Aguirre et al., 2010).

Um ingrediente relativamente inovador no contexto da identificação de sistemas surgiu na década de 90: era o uso de técnicas multiobjetivo na estimação de parâmetros de modelos (Johansen, 1996). Dentre as dificuldades práticas em tais procedimentos está a definição de objetivos adicionais ao da minimização da variância do erro de predição de um passo à frente, que é o objetivo clássico. Diversos estudos foram realizados ao longo de uma década investigando outras possibilidades (Nepomuceno et al., 2002; Barroso et al., 2006, 2007; Nepomuceno et al., 2007; Barbosa, 2009). Contudo, todos esses estudos concentraram-se no problema de estimação de parâmetros. O uso de técnicas multiobjetivo nesse problema foi de grande valia, pois o estimador bi-objetivo juntamente com um decisor adequado é não-tendencioso *sem a necessidade* do ajuste de modelos de ruído (Barroso et al., 2007).

Essa dissertação nasceu de uma motivação dupla. Por um lado, a eficácia das técnicas multiobjetivo (MO) na estimação de parâmetros e, por outro, o fato de ainda não haver soluções conclusivas para a determinação de estrutura de modelos serviram de motivação no sentido de adaptar as técnicas MO a tal problema.

1.1 Objetivo

O objetivo geral desse trabalho é desenvolver um critério baseado em técnicas multiobjetivo para auxiliar na determinação de estrutura de modelos NARX polinomiais.

Técnicas multiobjetivo, diferentemente da abordagem mono-objetivo, são capazes 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. Portanto, o objetivo específico consiste em determinar a estrutura de um modelo NARX polinomial por meio

de análises das curvas de Pareto. Curvas geradas pela variação combinacional de possíveis estruturas do modelo.

1.2 Estrutura da dissertação

Essa dissertação está organizada em seis capítulos da seguinte forma:

No Capítulo 1 a identificação de sistemas é introduzida, apresentando a motivação e justificativa para o assunto em questão. O objetivo da dissertação é definido e sua estrutura apresentada.

O Capítulo 2 traz uma revisão teórica e bibliográfica referente às principais ferramentas utilizadas ao longo deste trabalho. As propriedades dos modelos NARX e algumas técnicas existentes para seleção de estrutura e estimação de parâmetros são discutidas, e uma revisão envolvendo técnicas de identificação multiobjetivo é apresentada.

No Capítulo 3 é descrita a metodologia proposta para auxiliar na determinação da estrutura de modelos NARX polinomiais. Um exemplo simulado é estudado para ilustrar o método.

No Capítulo 4 a metodologia desenvolvida é aplicada em três exemplos simulados, verificando o desempenho e a robustez da técnica proposta.

No Capítulo 5 o método é avaliado na identificação de dois sistemas físicos: um conversor buck e uma válvula pneumática.

As considerações finais sobre o trabalho e as propostas para pesquisas futuras são apresentadas no Capítulo 6. Por fim, um apêndice traz as principais rotinas computacionais usadas na dissertação.

Fundamentação Teórica

“Enquanto o monjolo funcionar bem, ninguém pensará em substituí-lo. Guarde bem isso, porque o ajudará a entender as razões por que teorias são aceitas e mantidas.”

Rubem Alves¹

2.1 Introdução

Este capítulo é uma exposição preliminar referente às principais ferramentas utilizadas ao longo deste trabalho. Os objetivos são aproximar o leitor do conteúdo da literatura, que serviu como base teórica e elucidá-lo para algumas pressuposições.

As seções seguintes comentam brevemente algumas etapas da identificação de sistemas, as características e propriedades dos modelos NARX e algumas técnicas de seleção de estrutura e estimação de parâmetros. Na Seção 2.4 é feita uma revisão envolvendo técnicas de identificação multiobjetivo.

2.2 Modelos polinomiais NARX

Considere a classe de modelos definida como (Billings et al., 1989)

$$y(k) = f^\ell[y(k-1), \dots, y(k-n_y), u(k-d), u(k-n_u)], \quad (2.1)$$

¹(Alves, 2005, p. 166).

sendo n_y e n_u os máximos atrasos da saída e entrada, respectivamente. A constante d é o atraso puro de tempo. Além disso, $u(k)$ e $y(k)$ são respectivamente as séries temporais de entrada e saída de comprimento N . $f^\ell[\cdot]$ é alguma função polinomial de $u(k)$ e $y(k)$ com grau de não-linearidade $\ell \in \mathbb{N}$.

A Equação 2.1 pode ser expressa em forma preditiva, como

$$y(k) = \psi^T(k-1)\boldsymbol{\theta} + e(k). \quad (2.2)$$

Usando essa expressão como equação de regressão ao longo dos dados de estimação, pode-se escrever o conjunto de equações resultantes em forma matricial: $\mathbf{y} = \Psi\boldsymbol{\theta} + \mathbf{e}$, em que \mathbf{e} é o vetor do erro que aparece na equação de regressão. O vetor de parâmetros $\boldsymbol{\theta}$ pode ser estimado minimizando o produto interno do vetor de resíduos, ou seja, minimizando a função de custo

$$J_{LS}(\hat{\boldsymbol{\theta}}) = (\mathbf{y} - \Psi\hat{\boldsymbol{\theta}})^T(\mathbf{y} - \Psi\hat{\boldsymbol{\theta}}). \quad (2.3)$$

Os algoritmos de mínimos quadrados e os da família PEM minimizam o funcional J_{LS} em (2.3) sendo que eles diferem essencialmente pela maneira com que a predição de um passo à frente é construída.

A função de custo J_{LS} diminui à medida que a predição de um passo à frente do modelo se aproxima dos dados.

2.2.1 Agrupamento de termos

Analisando-se o modelo NARX (2.1) em estado estacionário para entradas contantes, tem-se

$$\begin{aligned} y(k-1) &= y(k-2) = \dots = y(k-n_y) \\ u(k-1) &= u(k-2) = \dots = u(k-n_u). \end{aligned} \quad (2.4)$$

Aplicando (2.4) em (2.1), após algumas manipulações (Aguirre e Billings, 1995), chega-se a

$$y(k) = \sum_{\substack{n_y, n_u \\ n_l, n_m}} c_{p, m-p}(n_l, \dots, n_m) \sum_{m=0}^l \sum_{p=0}^m y(k-1)^p u(k-1)^{m-p}. \quad (2.5)$$

O conjunto de termos da forma $y(k-i)^p u(k-j)^{m-p}$ é denominado agrupamentos de termos. Esses serão representados por $\Omega_{y^p u^{m-p}}$. A constante

$$\sum_{\substack{n_y, n_u \\ n_l, n_m}} c_{p, m-p}(n_l, \dots, n_m)$$

é o coeficiente do agrupamento de termos $y(k-1)^p u(k-1)^{m-p}$ e será representado por $\Sigma_{y^p u^{m-p}}$. Todos os termos pertencentes a um dado agrupamento de termos explicam o mesmo tipo de não-linearidade do modelo.

Os modelos NARX polinomiais são bastantes sensíveis à sobreparametrização de sua estrutura (Mendes e Billings, 1998). 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. De forma geral, o modelo polinomial, identificado a partir da análise dos agrupamentos de termos efetivos (agrupamentos genuínos) tem melhores chances de reproduzir as dinâmicas dos sistemas que se deseja estudar.

A importância dos agrupamentos e a análise dos seus coeficientes são tratadas com maior detalhe em (Aguirre, 2007). Pulecchi e Piroddi (2007) propuseram um método de seleção de agrupamentos para modelos polinomiais NARX, em que os conceitos de agrupamentos de termos e coeficientes de agrupamentos são utilizados na seleção de estrutura de modelos não-lineares, auxiliando na identificação de termos espúrios.

2.2.2 Curva estática

Por meio dos conceitos de agrupamento de termos e de coeficientes de agrupamento é possível derivar uma equação para o ganho estático (Aguirre, 2007; Nepomuceno, 2002; Corrêa e Aguirre, 2004).

O modelo NARX polinomial, visto na Seção 2.2, em estado estacionário pode ser escrito da seguinte forma (assumindo que sejam assintoticamente

estável em torno de (\bar{u}, \bar{y}) :

$$\hat{y} = \Sigma_0 + \Sigma_y \bar{y} + \Sigma_u \bar{u} + \sum_{m=1}^{l-1} \sum_{p=1}^{l-m} \Sigma_{y^p u^m} \bar{y}^p \bar{u}^m + \sum_{p=2}^l \Sigma_{y^p} \bar{y}^p + \sum_{m=2}^l \Sigma_{u^m} \bar{u}^m, \quad (2.6)$$

sendo que os termos de processo e seus respectivos parâmetros foram agrupados em: i) termo constante, Σ_0 ; ii) termos lineares em y , $\Sigma_y \bar{y}$; iii) termos lineares em u , $\Sigma_u \bar{u}$; iv) termos cruzados, $\sum_{m=1}^{l-1} \sum_{p=1}^{l-m} \Sigma_{y^p u^m} \bar{y}^p \bar{u}^m$; v) termos não-lineares em y , $\sum_{i=1}^l \Sigma_{y^i} \bar{y}^i$ e vi) termos não-lineares em u , $\sum_{i=1}^l \Sigma_{u^i} \bar{u}^i$.

A curva 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} . O ganho estático é dado pela seguinte relação (Corrêa e Aguirre, 2004):

$$\hat{K}(\bar{y}, \bar{u}) = \frac{\bar{y}}{\bar{u}} = \frac{\frac{\Sigma_0}{\bar{u}} + \Sigma_u + \sum_{m=2}^l \Sigma_{u^m} \bar{u}^{(m-1)}}{1 - \Sigma_y - \sum_{m=1}^{l-1} \sum_{p=1}^{l-m} \Sigma_{y^p u^m} \bar{y}^{(p-1)} \bar{u}^m - \sum_{p=2}^l \Sigma_{y^p} \bar{y}^{(p-1)}}. \quad (2.7)$$

Obtida a expressão para o ganho estático de um modelo, a expressão para a curva estática pode ser determinada por

$$\bar{y} = \frac{\Sigma_0 + \Sigma_u \bar{u} + \sum_{m=2}^l \Sigma_{u^m} \bar{u}^m}{1 - \Sigma_y - \sum_{m=1}^{l-1} \sum_{p=1}^{l-m} \Sigma_{y^p u^m} \bar{y}^{(p-1)} \bar{u}^m - \sum_{p=2}^l \Sigma_{y^p} \bar{y}^{(p-1)}}. \quad (2.8)$$

Determinada a curva estática, é possível definir uma função custo relacionada ao erro em regime estacionário. Similar a Equação 2.3, uma alternativa que tem se mostrado muito útil é (Barroso et al., 2007; Nepomuceno et al., 2007)

$$\begin{aligned} J_{\text{SF}}(\hat{\theta}) &= (\bar{y} - \hat{y})^T (\bar{y} - \hat{y}) \\ &= (\bar{y} - \mathbf{Q}\mathbf{S}\hat{\theta})^T (\bar{y} - \mathbf{Q}\mathbf{S}\hat{\theta}), \end{aligned} \quad (2.9)$$

em que a barra superior indica valores em estado estacionário e o circunflexo valores estimados.

Para ver o significado de $\hat{\mathbf{y}} = \mathbf{Q}\mathbf{S}\hat{\boldsymbol{\theta}}$, que indica a característica estática do modelo, considere novamente o seguinte. O comportamento em estado estacionário de um sistema assintoticamente estável pode ser obtido fixando-se a entrada a um valor constante \bar{u} para o qual a saída estabilizará em \bar{y} . Em geral \bar{u} e \bar{y} serão relacionados por $\bar{y} = f(\bar{y}, \bar{u})$, sendo que $f(\cdot)$ é a função estática. Para modelos NARX polinomiais, graças à linearidade nos parâmetros, é possível escrever:

$$\bar{y}_i = f(\bar{y}_i, \bar{u}_i) = \mathbf{q}_i^T \mathbf{S}\hat{\boldsymbol{\theta}}, \quad i = 1, 2, \dots, n_{\text{sf}} \quad (2.10)$$

sendo que n_{sf} é o número de pontos em estado estacionário (\bar{u}_i, \bar{y}_i) considerado e

$$\mathbf{q}_i^T = [1 \ \bar{y}_i \dots \bar{y}_i^\ell \ \bar{u}_i \dots \bar{u}_i^\ell \ F_{yu}] , \quad i = 1, 2, \dots, n_{\text{sf}} \quad (2.11)$$

sendo que F_{yu} representa todos os monômios não-lineares no modelo que envolvam $y(k)$ e $u(k)$, e

$$\mathbf{R} = \mathbf{S}\hat{\boldsymbol{\theta}} = [\Sigma_0 \ \Sigma_y \dots \Sigma_{y^\ell} \ \Sigma_u \dots \Sigma_{u^\ell} \ F_\Sigma]^T , \quad (2.12)$$

em que F_Σ indica os coeficientes de agrupamentos de todos os agrupamentos de termos em F_{yu} e $\mathbf{S}\hat{\boldsymbol{\theta}}$ é um vetor de constantes. Portanto $\hat{\mathbf{y}} = \mathbf{Q}\mathbf{S}\hat{\boldsymbol{\theta}}$, em que

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_{n_{\text{sf}}}^T \end{bmatrix} . \quad (2.13)$$

Os detalhes do desenvolvimento acima podem ser encontrados em Barroso et al. (2007) e na literatura ali referenciada.

2.3 Determinação de estrutura

A escolha da estrutura a ser utilizada para representar os regimes dinâmicos descritos pelos dados é um problema fundamental na identificação de sistemas. Embora vários critérios relacionados com a seleção de regressores em modelos

NARX polinomiais têm sido propostos (Zhang et al., 2007; Mendes e Billings, 2001; Wei e Billings, 2008, 2009; Maia e Mendes, 2008; Alves, 2009), a seleção de estrutura continua sendo um problema sem solução definitiva.

Na seleção de estruturas, um dos principais fatores é definir o número de termos candidatos. Em modelos NARX os aumentos no grau de não-linearidade ℓ e nos máximos atrasos n_y e n_u , provocam um aumento exponencial no número de termos candidatos. O número de termos de um modelo polinomial pode ser determinado pela seguinte expressão (Korenberg et al., 1988):

$$n_\theta = M + 1, \quad (2.14)$$

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

$$M = \sum_{i=1}^l n_i \quad (2.15)$$

$$n_i = n_{i-1} \frac{n_y + n_u + n_e + i - 1}{i}, n_0 = 1. \quad (2.16)$$

Várias técnicas podem ser usadas para fazer a seleção de estrutura. Nessa seção, são apresentados dois métodos: critério de informação de Akaike e taxa de redução de erro. Esse dois critérios, facilmente encontrados na literatura, servirá como referência para o método de seleção bi-objetivo proposto no Capítulo 3.

2.3.1 Determinação de estrutura utilizando ERR

A taxa de redução de erro (*error reduction ratio*) ou ERR (Billings et al., 1989; Aguirre, 2007) baseia-se na redução do erro de predição de um passo à frente. Esse critério associa a cada termo candidato um índice correspondente à contribuição na explicação da variância dos dados de saída, $\xi(k)$.

Seja:

$$\text{var} \{\xi(k)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \left[\mathbf{y}^T \mathbf{y} - \sum_{i=1}^N g_i^2 w_i^T w_i \right], \quad (2.17)$$

em que g_i indica os elementos do vetor de parâmetros \mathbf{g} e w_i indica os regres-

sores ortogonais e \mathbf{y} é o vetor contendo os dados de saída.

A cada termo acrescentado, a variância de $\xi(k)$ decresce de um fator igual a $\frac{1}{N}(g_1^2 w_i^T w_i)$, em que w_i indica o termo incluído e g_i o seu respectivo parâmetro. A redução do 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:

$$[ERR] = \frac{(g_i^2 w_i^T w_i)}{\mathbf{y}^T \mathbf{y}}. \quad (2.18)$$

O ERR é amplamente utilizado na seleção de estrutura de modelos. Escolhe-se o número de termos desejados e consideram-se aqueles que possuem os maiores valores de ERR.

2.3.2 Critério de informação de Akaike

Como descrito na seção anterior, uma maneira de ordenar hierarquicamente os termos candidatos é utilizando-se o critério ERR. Portanto, tendo ordenado os termos, ainda é necessário achar o ponto de parada, ou seja o número de termos que constituirá o modelo.

Há vários métodos para determinar o número de termos de um modelo. Destacam-se os critérios de informação, tais como o critério do erro final de predição (FPE) (Akaike, 1970), o critério de informação de Akaike (Akaike, 1974), o critério de informação de Bayes (Kashyap, 1977) entre outros. Nakamura et al. (2006) fizeram um estudo comparando alguns critérios de informação na modelagem de séries temporais.

Nesse trabalho, utilizou-se o critério de informação de Akaike com o intuito de comparar alguns resultados. De acordo com o critério de Akaike, o número de termos de um modelo deve minimizar a seguinte função:

$$AIC(n_\theta) = N \ln(\text{var} \{\xi(k)\}) + 2n_\theta, \quad (2.19)$$

sendo que N é o número de dados, $\text{var} \{\xi(k)\}$ é a variância do erro de modelagem (erro de predição de um passo à frente ou resíduos) e n_θ é o número de parâmetros do modelo.

O número de termos determinados a partir do AIC minimiza a variância dos resíduos de identificação partindo de um 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 pelo AIC pode ser visto como uma indicação do número “ótimo” de termos do modelo (Nepomuceno, 2002).

2.4 Identificação multiobjetivo

Se em um problema de otimização qualquer estiverem disponíveis mais de um objetivo que se deseja alcançar, e se esses objetivos forem conflitantes entre si, haverá soluções que, comparadas com outras soluções, serão melhores em algum ou alguns objetivos, mas piores em outro ou outros objetivos. Esse conjunto de soluções é denominado soluções eficientes ou soluções Pareto-ótimas. O problema multiobjetivo consiste em obter o conjunto de soluções Pareto-ótima e escolher a solução final por meio de um decisor.

As funções custo $J_{LS}(\hat{\theta})$ (2.3) e $J_{SF}(\hat{\theta})$ (2.9), definidas na Seção 2.2, são funcionais que dependem de $\hat{\theta}$ e, portanto, é possível definir um problema bi-objetivo usando tais funcionais para a estimação de $\hat{\theta}$.

2.4.1 A solução Pareto-ótima

Em geral, não há uma única solução (modelo) que simultaneamente minimize as duas funções de custo, em vez disso, diversas soluções (modelos) podem ser encontrados. Com o conjunto Pareto-ótimo $\hat{\Theta}^*$ de soluções $\hat{\theta}$ caracterizadas por (Chankong e Haimes, 1983):

$$\hat{\theta}^* \in \hat{\Theta}^* \Leftrightarrow \{\nexists \hat{\theta} : \mathbf{J}(\hat{\theta}) \leq \mathbf{J}(\hat{\theta}^*) \text{ e } \mathbf{J}(\hat{\theta}) \neq \mathbf{J}(\hat{\theta}^*)\}, \quad (2.20)$$

em que

$$\mathbf{J}(\hat{\theta}) = \begin{bmatrix} J_{LS}(\hat{\theta}) & J_{SF}(\hat{\theta}) \end{bmatrix}^T. \quad (2.21)$$

As soluções multiobjetivo $\hat{\theta}$ podem ser encontradas resolvendo o seguinte problema de mínimos quadrados para $0 \leq w \leq 1$ (Nepomuceno et al., 2007):

$$\hat{\theta}(w) = [w\Psi^T\Psi + (1-w)(QR)^TQR]^{-1} \times [w\Psi^T\mathbf{y} + (1-w)(QR)^T\bar{\mathbf{y}}] \quad (2.22)$$

Para cada w , encontra-se uma estimativa $\hat{\boldsymbol{\theta}}(w)$ no conjunto Pareto-ótimo. As soluções mono-objetivo que minimizam J_{SF} e J_{LS} são encontradas, respectivamente, para $w = 0$ e $w = 1$.

2.4.2 O decisor de mínima correlação

Para cada modelo candidato $\hat{\boldsymbol{\theta}}$, pertencente ao conjunto Pareto $\hat{\Theta}^*$, a série de dados de *simulação livre* é definida como:

$$\hat{y}(k) = \hat{\psi}^T(k-1)\hat{\boldsymbol{\theta}}, \quad (2.23)$$

em que $\hat{\psi}(k-1)$ é a matriz composta de regressores de simulação livre, ou seja, todos os regressores de saída são substituídos pelos valores correspondentes da saída simulada livremente $\hat{y}(k)$. O erro de simulação livre $\eta(k)$, para o mesmo modelo candidato, é dado por

$$\eta(k) = y(k) - \hat{y}(k). \quad (2.24)$$

Define-se o seguinte funcional:

$$J_{\text{corr}}(\hat{\boldsymbol{\theta}}) = \left| \frac{1}{N} \sum_{k=1}^N \eta(k)\hat{y}(k) \right|, \quad (2.25)$$

em que N é o comprimento da série temporal. Portanto, pode-se escolher *um* modelo sobre o Pareto tal que minimize

$$\min_{\hat{\boldsymbol{\theta}} \in \hat{\Theta}^*} J_{\text{corr}}(\hat{\boldsymbol{\theta}}). \quad (2.26)$$

A justificativa e a análise das características do decisor de mínima correlação definido em (2.26) podem ser encontradas em (Barroso, 2006; Barroso et al., 2007).

2.4.3 O decisor de melhor sincronismo

O decisor descrito na seção anterior é essencialmente estatístico. Em Aguirre et al. (2006) foi desenvolvido um procedimento baseado em sincronismo para escolher, dentre um conjunto de modelos, aquele que supostamente melhor

aproxima a dinâmica subjacente aos dados. No presente estudo esse procedimento também será utilizado como decisor.

Sejam dois osciladores, possivelmente não-lineares. Se tais osciladores tiverem dinâmicas muito próximas, em geral é possível sincronizar um deles (chamado escravo) ao outro (chamado mestre). Contudo, se a dinâmica desses osciladores forem distintas, será muito difícil sincronizar um ao outro. Em outras palavras, a capacidade de sincronizar dois osciladores pode servir como um teste para a hipótese deles terem dinâmicas próximas.

Na prática, apenas esse procedimento não é suficiente, pois em alguns casos é possível sincronizar dois osciladores mesmo que distintos. Contudo, em tais casos o esforço requerido para estabelecer o sincronismo é tipicamente muito maior. Sendo assim, é possível definir um procedimento pelo qual a proximidade da dinâmica de dois osciladores pode ser quantificada levando-se em conta (Aguirre et al., 2006): i) a qualidade do sincronismo atingível e ii) o esforço necessário para conseguir tal sincronismo.

No contexto de escolha de modelos o que se faz é considerar os dados como a trajetória produzida pelo oscilador mestre e verifica-se se o oscilador escravo (o modelo) sincroniza e com que custo. Para ver isso mais precisamente, seja $y(k)$ a saída e $u(k)$ a entrada. Assume-se que os dados são descritos pelo sistema

$$y(k) = g(\psi_{uy}(k-1)), \quad (2.27)$$

sendo $\psi_{uy}(k-1)$ o vetor de regressores que envolvem atrasos até o instante $k-1$ e g é uma função desconhecida que mapeia os regressores à saída futura $y(k)$. Assume-se também a existência de um modelo f que explica os dados como $y(k) = f(\hat{\psi}_{uy}(k-1)) + \xi(k)$, sendo que $f(\hat{\psi}_{uy}(k-1))$ é a predição de um passo à frente feita com a parcela *determinística* do modelo, e $\xi(k)$ é o respectivo erro. O vetor de variáveis independentes do modelo é indicado por $\hat{\psi}_{uy}(k-1)$, que *não* precisa ser o mesmo que $\psi_{uy}(k-1)$.

A fim de verificar se o modelo f sincronizará com os dados, implementa-se o seguinte esquema de acoplamento (Aguirre et al., 2006):

$$\hat{y}(k) = f(\hat{\psi}_{uy}(k-1)) - h(k-1), \quad (2.28)$$

em que $h(k) = c(y(k) - \hat{y}(k))$ e $c \in \mathbb{R}$ é uma constante. Em (2.28) além do

acoplamento, ocorreu uma substituição direta de \hat{y} (ver Equação 2.23) por y . Sendo assim, o erro de sincronismo é:

$$\begin{aligned}
 \eta(k) &= y(k) - \hat{y}(k) \\
 &= g(\psi_{uy}(k-1)) - [f(\hat{\psi}_{uy}(k-1)) - h(k-1)] \\
 &= c(y(k-1) - \hat{y}(k-1)) + [g(\psi_{uy}(k-1)) - f(\hat{\psi}_{uy}(k-1))] \\
 &= c\eta(k-1) + \xi(k-1).
 \end{aligned} \tag{2.29}$$

A expressão (2.29) é um modelo ARX em que a entrada exógena é o erro de modelagem $\xi(k-1)$. Assumindo que a parte autorregressiva é estável (condição que se verifica na prática para modelos minimamente razoáveis), logo se vê que à medida que o erro de modelagem ξ tende a zero, o erro de sincronismo $\eta(k)$ também irá a zero. Ou seja, o erro de sincronismo entre modelo e dados é uma medida do ajuste do modelo à dinâmica que produziu os dados.

O custo de sincronismo a ser utilizado neste trabalho é

$$J_{\text{rms}}(c) = \lim_{N \rightarrow \infty} \sqrt{\frac{1}{N} \sum_{k=1}^N \|h(k)\|^2}, \tag{2.30}$$

em que se considerou que $k = 1$ indica o instante a partir do qual o sincronismo é estabelecido (o transitório é desprezado). Como $h(k) = c(x(k) - \hat{y}(k))$ pode ser interpretado como a ação de controle requerida para manter o modelo sincronizado aos dados, é possível interpretar $J_{\text{rms}}(c)$ como a energia necessária para manter o modelo “perto” dos dados. A fim de sermos mais quantitativos na caracterização de “perto”, será usada a seguinte definição (Tôres e Aguirre, 2004). Dois sistemas pertencerão à classe ϵ de sincronismo se e somente se

$$\epsilon(c) = \frac{\max \|\eta(k)\|}{\max \|\eta(k)\| + 1}, \quad \forall k \geq 1. \tag{2.31}$$

em que $\eta(k)$ foi definido em (2.29) e a dependência entre a qualidade de sincronismo e o ganho proporcional c é explícita em (2.29) e está implícita em (2.31). Claramente $0 \leq \epsilon(c) \leq 1$ e $\epsilon(c) \approx 0$ implica $y(k) = \hat{y}(k)$, ou seja, sincronismo perfeito para aquele valor de c . Por outro lado, $\epsilon(c) \approx 1$ indica

comportamento assíncrono.

2.5 O método LASSO

O método LASSO (*Least absolute shrinkage and selection operator*), proposto por Tibshirani (1996), é obtido como a solução de um problema de estimação utilizando uma penalidade ℓ_1 (ilustrada na Figura 2.1) na minimização do problema:

$$J_{LASSO} = \|y - \hat{y}\| + \lambda \sum_{j=2}^q |\theta_j|, \quad (2.32)$$

em que $\|y - \hat{y}\|$ é a norma euclidiana dos resíduos, λ um parâmetro de regularização e θ_j o parâmetro estimado.

Uma vantagem do estimador LASSO é que ao invés de apenas forçar os parâmetros para zero, como no caso de técnicas como *ridge regression*, ele efetivamente permite que alguns parâmetros sejam identicamente iguais a zero, realizando simultaneamente um procedimento de redução (*shrinkage*) e seleção de modelos. A Figura 2.2 apresenta uma descrição visual do método, em que a solução pelo método LASSO (figura à esquerda), mostra o contorno elipsoidal tocando o vértice do quadrado (ℓ_1), correspondente ao coeficiente nulo. Na figura à direita, para a penalidade ℓ_2 , não há vértices e raramente o contorno coincide com a solução nula.

Contudo, o método apresenta dificuldades em determinar o parâmetro λ (2.32). O estimador LASSO pode ser polarizado, ou apresentar resultados “distorcidos” caso o valor de λ seja inadequado (Bonin et al., 2010).

Uma descrição detalhada do método LASSO pode ser encontrada no trabalho de Tibshirani (1996). Aplicações e limitações do método podem ser vistas em (Bonin et al., 2010; Laurini e Hotta, 2009; Miranda, 2006).

2.6 Validação do modelo

A validação de um modelo é a etapa final do processo de identificação. O objetivo é verificar se o modelo obtido incorpora as características de interesse

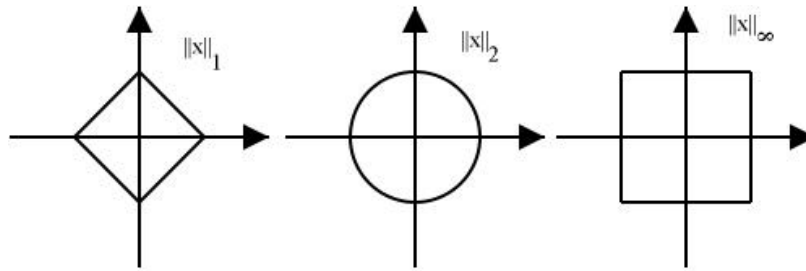


Figura 2.1: Representação geométrica das normas l_1 , l_2 e l_∞ .

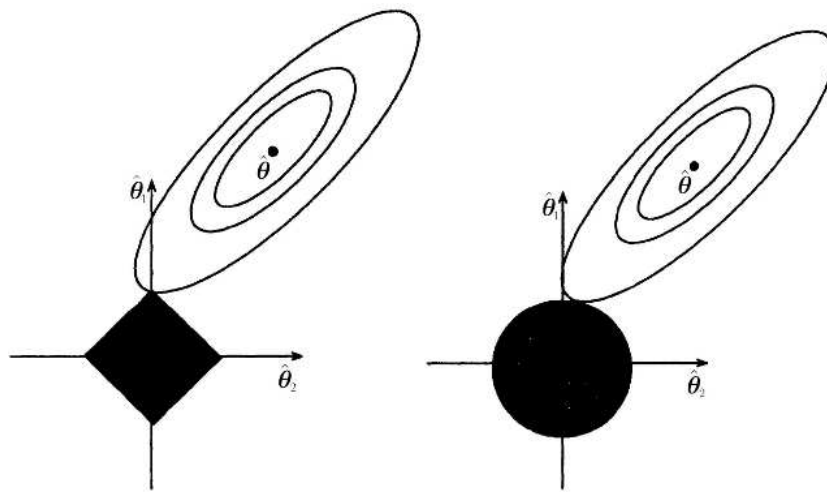


Figura 2.2: Representação gráfica do método LASSO e *ridge regression*.

Representação gráfica do método LASSO, figura à esquerda e *ridge regression*, figura à direita. O LASSO permite que alguns parâmetros sejam identicamente iguais a zero. Figura adaptada de Tibshirani (1996).

do sistema original, ou ainda, comparar o desempenho entre dois ou mais modelos.

Por essa razão, faz-se necessário a utilização de um critério para determinar se o modelo incorpora as características que lhe são exigidas. Neste trabalho foram considerados dois aspectos: i) simulação livre e ii) capacidade de representar o comportamento estático do sistema. Ambos aspectos quantificados pelo erro quadrático médio, RMSE (*root mean squared error*) que, para o caso dinâmico, normalizado, é escrito como:

$$\text{RMSE} = \frac{\sqrt{\sum_{k=1}^N [y(k) - \hat{y}(k)]^2}}{\sqrt{\sum_{k=1}^N [y(k) - \frac{1}{N} \sum_{k=1}^N y(k)]^2}}. \quad (2.33)$$

sendo que $\hat{y}(k)$ é a simulação livre do sinal e $\frac{1}{N} \sum_{k=1}^N y(k)$ é o valor médio do sinal medido $y(k)$. Para o caso estático, o mesmo índice, normalizado, é escrito como:

$$\text{RMSE} = \frac{\sqrt{\sum_{k=1}^N [\bar{y}(k) - \hat{y}(k)]^2}}{\sqrt{\sum_{k=1}^N [\bar{y}(k) - \frac{1}{N} \sum_{k=1}^N \bar{y}(k)]^2}}. \quad (2.34)$$

em que $\bar{y}(k)$ são os valores em estado estacionário medidos, $\frac{1}{N} \sum_{k=1}^N \bar{y}(k)$ é o valor médio dos dados estacionários medidos e $\hat{y}(k)$ são os dados obtidos a partir da simulação livre do sinal (ver Equação 2.8).

Quanto menor for o valor do RMSE, maior será a eficiência na recuperação da característica (dinâmica e/ou estática) do sistema.

O livro de Aguirre (2007) tem um capítulo dedicado ao estudo da validação de modelos, onde encontram-se sugestões e referências para o assunto.

2.7 Conclusões do capítulo

O presente capítulo teve como objetivo familiarizar o leitor com algumas técnicas de identificação de sistemas, em particular técnicas de identificação multiobjetivo. Essas técnicas servirão como base para o método de seleção de estrutura proposto e que será apresentado no Capítulo 3.

Procedimento Bi-Objetivo para Seleção de Estrutura

“Toda a arte é um problema de equilíbrio entre dois opostos.”

Cesare Pavese

3.1 Introdução

No Capítulo 2 foram apresentados os conceitos básicos e uma revisão de literatura no que concerne a identificação de sistemas, em especial alguns aspectos da identificação multiobjetivo e a etapa de seleção de estruturas.

Neste Capítulo¹ descreve-se a metodologia proposta para auxiliar na determinação da estrutura de modelos NARX polinomiais. Um exemplo numérico é apresentado para ilustrar o método.

3.2 Definição do problema

A Figura 3.1 representa o conjunto Pareto-ótimo, ou conjunto de soluções eficientes de um problema de otimização bi-objetivo. Isso significa que as soluções de Pareto são as melhores soluções entre as quais não existe um ordenamento, ou seja, não há como definir, a partir da avaliação dos funcionais

¹As principais idéias que deram origem ao procedimento proposto nesse capítulo estão em um artigo publicado no XVIII Congresso Brasileiro de Automática (Barbosa et al., 2010).

objetivo, que uma solução é melhor que a outra. Qualquer método de otimização que minimize f_1 e/ou f_2 , no melhor caso, encontrará soluções pertencentes àquele conjunto.

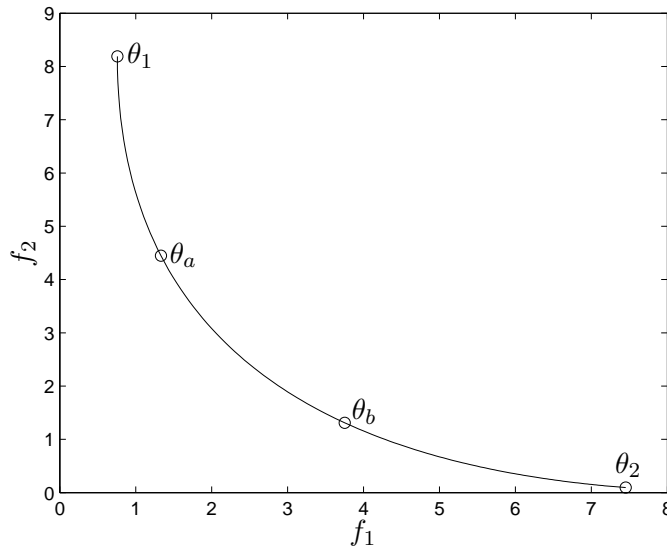


Figura 3.1: Representação de um conjunto Pareto-ótimo de um problema bi-objetivo.

A curva representada é um hipotético conjunto Pareto-ótimo de um problema bi-objetivo de otimização. As soluções θ_1 e θ_2 denotam os ótimos individuais das funções objetivos f_1 e f_2 . θ_a e θ_b são soluções pertencentes ao conjunto Pareto-ótimo.

Em identificação de sistemas, a etapa de estimação de parâmetros é um problema de otimização. Usualmente, minimiza-se o erro de predição de um passo à frente associado ao modelo dinâmico, ou seja, ajustando a estrutura matemática aos dados por meio dos parâmetros estimados. Contudo, mais de um tipo de informação no procedimento de estimação pode ser considerado. Isso remete a um problema de otimização vetorial no qual os objetivos tendem a ser conflitantes.

A partir de dois funcionais (erro estático e erro dinâmico), descritos no Capítulo 2, definiu-se o problema bi-objetivo de estimação. Com dois objetivos, tem-se um conjunto Pareto de possíveis soluções (ver Seção 2.4.1). Sendo assim, o problema consiste em estabelecer, a partir dos conjuntos Pareto-ótimos, um procedimento para discriminar de forma nítida o melhor modelo.

3.3 Procedimento bi-objetivo

Cada estrutura-candidata, ao ser submetida a um procedimento multiobjetivo de estimação de parâmetros, conduz a um conjunto Pareto-ótimo de parâmetros, conforme representado na Figura 3.1.

Quando são consideradas várias estruturas-candidatas, a estimação multiobjetivo dos parâmetros (sobre cada estrutura) conduz a vários conjuntos Pareto-ótimos de parâmetros. Um esboço desse conjunto de conjuntos Pareto-ótimos é representado na Figura 3.2. Na mesma figura, são também representados os vetores de parâmetros que se associam aos modelos de mínima correlação θ_j^1 , θ_j^2 e θ_j^3 , respectivamente provenientes dos conjuntos Pareto-ótimos de cada estrutura-candidata.

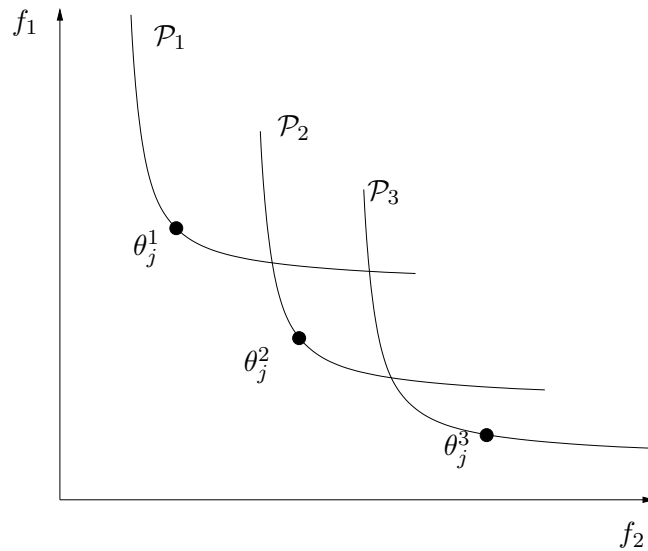


Figura 3.2: Representação dos conjuntos Pareto-ótimos \mathcal{P}_1 , \mathcal{P}_2 e \mathcal{P}_3 .

Representação dos conjuntos Pareto-ótimos dos vetores de parâmetros que constituem as soluções multiobjetivo do problema de estimação de parâmetros bi-objetivo sobre três estruturas diferentes. As respectivas soluções de mínima auto-correlação θ_j^1 , θ_j^2 e θ_j^3 se encontram também representadas.

De posse de um conjunto de Paretos, correspondentes a diferentes estruturas, deve-se selecionar o melhor modelo. O procedimento de seleção de estrutura bi-objetivo aqui proposto é descrito em duas etapas, Rotina I e Rotina II, da seguinte maneira:

Procedimento de Seleção de Estrutura

Rotina I:

- 1: Considere uma coleção inicial de n estruturas-candidatas $\{\mathcal{G}_1, \dots, \mathcal{G}_n\}$;
- 2: Para $i = 1$ até n
 - 3: Determine, utilizando dados de identificação, o conjunto Pareto-ótimo \mathcal{P}_i dos parâmetros da i -ésima estrutura;
- 4: Identifique, por meio dos Conjuntos-Pareto, os agrupamentos potencialmente espúrios;
- 5: Eliminar os termos pertencentes aos agrupamentos potencialmente espúrios.

Rotina II:

- 1: Seja $\{\mathcal{M}_1, \dots, \mathcal{M}_m\}$ um conjunto previamente selecionado de estruturas-candidatas;
- 2: Para $i = 1$ até m
 - 3: Determine, utilizando dados de identificação, o conjunto Pareto-ótimo \mathcal{P}_i dos parâmetros da i -ésima estrutura;
 - 4: Determine o modelo de mínima autocorrelação dentre os modelos pertencentes a $\mathcal{P}_i : \theta_j^i$.
- 5: Seja $\Theta = \{\theta_j^1, \theta_j^2, \dots, \theta_j^m\}$, em que m indica o Pareto, e j pode assumir valores distintos entre 1 e \bar{n} . Sendo \bar{n} o número de modelos por Pareto. Escolher, utilizando dados de validação, o modelo θ^* que possuir o máximo sincronismo.

3.3.1 Descrição e implementação da rotina I

O algoritmo implementado na Rotina I tem o objetivo de discriminar os agrupamentos espúrios dos genuínos (ver Seção 2.2.1). Ou seja, o foco dessa etapa é eliminar os termos potencialmente espúrios pertencentes a agrupamentos potencialmente espúrios, reduzindo os termos candidatos.

O diagrama da Figura 3.3 ilustra os passos que compõem o algoritmo. Considere conhecidos os dados dinâmicos e estáticos do sistema e um conjunto de x termos candidatos, incluindo os termos corretos. A partir de uma estrutura sobreparametrizada de h agrupamentos e x termos, em sequência, elimina-se um dos agrupamentos e reestimam-se os parâmetros do modelo que tem os $h - 1$

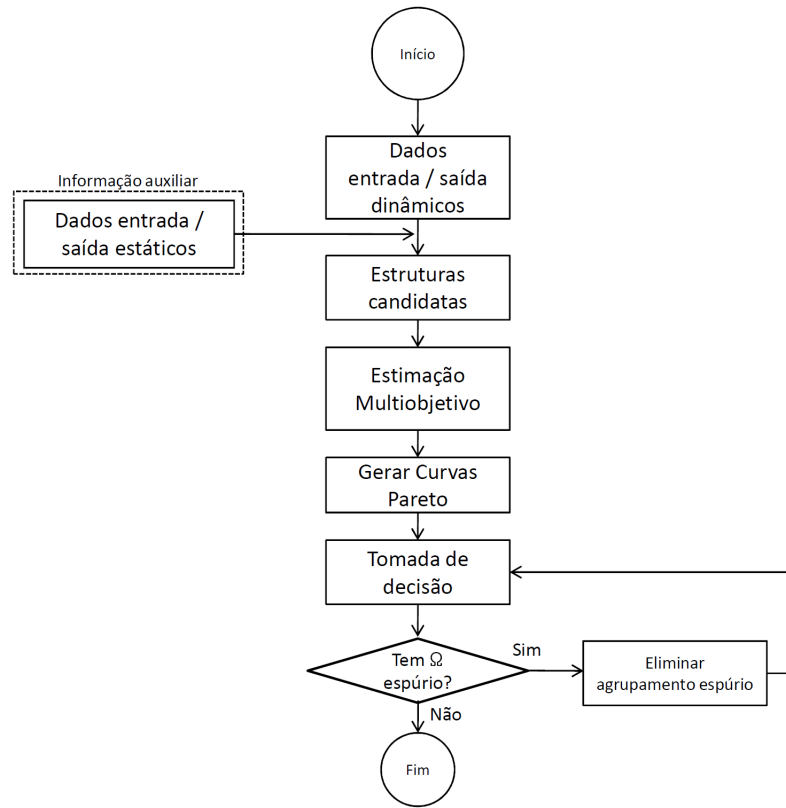


Figura 3.3: Fluxograma do procedimento bi-objetivo para seleção de estrutura, etapa de seleção dos agrupamentos.

agrupamentos restantes. Para os $h + 1$ modelos geram-se as Curvas-Pareto correspondente (ver Seção 2.4.1).

De posse das curvas Pareto, busca-se identificar os agrupamentos genuínos. Para isso, conjectura-se que a exclusão de agrupamentos genuínos afasta a curva Pareto da origem, ou seja, penaliza os custos estático e dinâmico. Nessa etapa aplica-se, como métrica para distinguir os agrupamentos genuínos, a análise da área formada pela curva Pareto, como descrito a seguir.

Ilustrada na Figura 3.4, o índice de desempenho é quantificado pela área definida do eixo à fronteira Pareto encontrada. Para problemas bi-objetivo de minimização, quanto menor a área melhor (Deb et al., 2002).

Por fim, elimina-se os termos pertencentes aos agrupamentos espúrios. De posse de um conjunto menor de termos candidatos, o próximo objetivo é determinar os regressores que são verdadeiros.

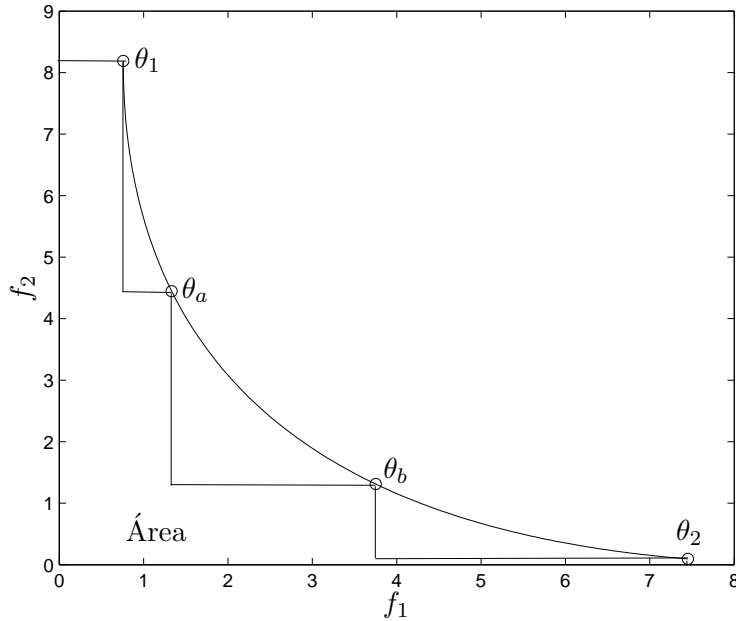


Figura 3.4: Métrica de desempenho: área.

Métrica aplicada para avaliação da curva Pareto. Em problemas de otimização com 3 objetivos tem-se o volume e no caso de quatro ou mais objetivos o “hipervolume”.

3.3.2 Descrição e implementação da rotina II

A Rotina II é uma continuação da rotina anterior. O diagrama da Figura 3.5 descreve os passos do algoritmo. O objetivo é “filtrar” o modelo. Ou seja, a partir de um conjunto de termos espúrios pertencentes a agrupamentos não-espúrios, deseja-se encontrar todos os termos genuínos.

Com os termos restantes, forma-se um conjunto m de estruturas-candidatas. Determina-se, por meio da estimação multiobjetivo, o conjunto Pareto-ótimo de todas as estruturas-candidatas. Em seguida, aplica-se o decisor de mínima autocorrelação e determina-se o modelo θ_i dentre os modelos pertencentes ao respectivo conjunto Pareto.

Os m diferentes modelos (estrutura + parâmetros) que se tornam disponíveis após a execução dos passos de determinação dos modelos de mínima correlação constituem um conjunto relativamente pequeno de candidatos a se-

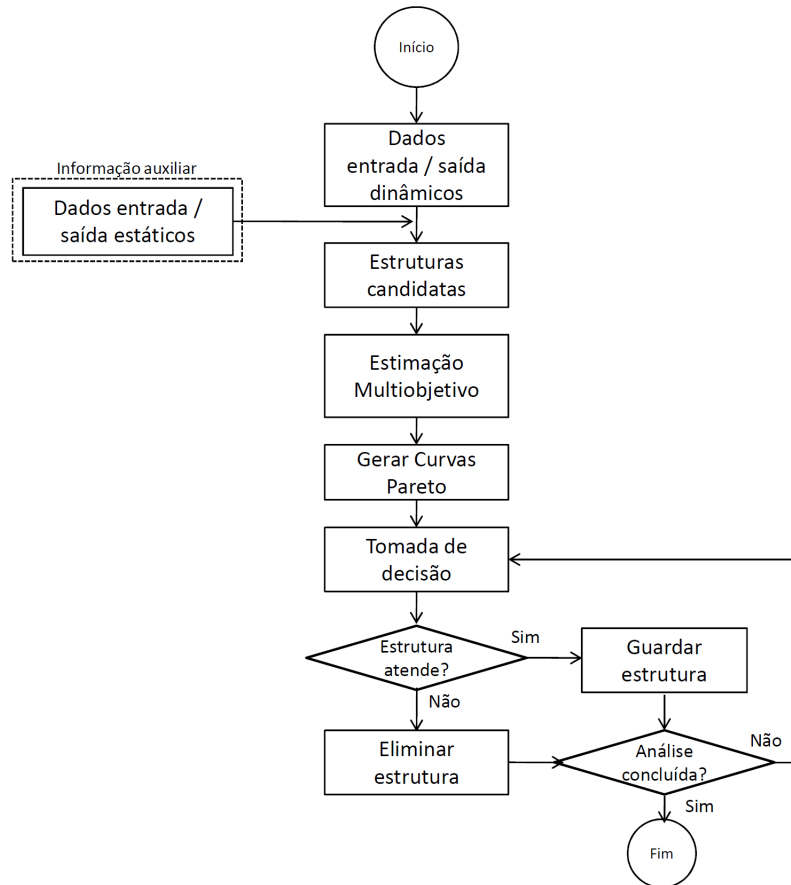


Figura 3.5: Fluxograma do procedimento bi-objetivo para seleção de estrutura, etapa de seleção do modelo nominal.

rem considerados. Se, dentre as estruturas inicialmente consideradas, houver um conjunto de termos estruturalmente adequado para descrever os dados em questão, este provavelmente resultará no aparecimento de um modelo completo (estrutura + parâmetros) adequado. Caso no conjunto inicial não tenha todo os termos corretos, espera-se que os termos que melhor representam o sistema sejam os escolhidos. Cabe então, a essa altura, realizar algum procedimento para discriminar, dentre as opções disponíveis, aquela que seria a mais adequada.

Deve-se notar, neste ponto, que a própria medida de correlação, definida na Seção 2.4.2, já exerce tal papel, pelo menos para realizar a discriminação entre os (possivelmente muitos) modelos definitivamente inadequados e os

(provavelmente poucos) modelos possivelmente adequados. Os primeiros apresentarão correlação significativamente maior que os segundos. No entanto, para a seleção, dentre vários modelos estruturalmente parecidos, daquele que poderia ser considerado mais adequado, é interessante adotar-se uma métrica que seja relativamente ortogonal em relação à informação já utilizada, na etapa anterior, da autocorrelação. Uma métrica capaz de carrear informação nova provavelmente será também capaz de produzir uma discriminação mais nítida do melhor modelo. Por esse motivo, a última etapa, de seleção do modelo, emprega a informação do “melhor sincronismo” e menor custo.

3.4 Exemplo ilustrado

O seguinte exemplo simulado, estudado por Bonin et al. (2010), tem o intuito de ilustrar o método proposto. No Capítulo 4, o mesmo exemplo será abordado levando em consideração diferentes aspectos.

Considere o modelo NARX, dado por

$$y(k) = 0,5y(k-1) + 0,8u(k-2) + u(k-1)^2 - 0,05y(k-2)^2 + 0,5 + e(k), \quad (3.1)$$

em que a entrada $u(k)$ foi gerada a partir de um sinal do tipo ruído branco com distribuição Gaussiana, média nula e variância unitária com ruído gerado da mesma forma, contudo, com variância igual a 0,05. A Figura 3.6 apresenta os sinais dinâmicos de entrada e saída do sistema. A característica estática do sistema é mostrada na Figura 3.7. O objetivo final é selecionar os 5 termos do modelo dado pela Equação 3.1 dentre um conjunto de 56 possíveis regressores.

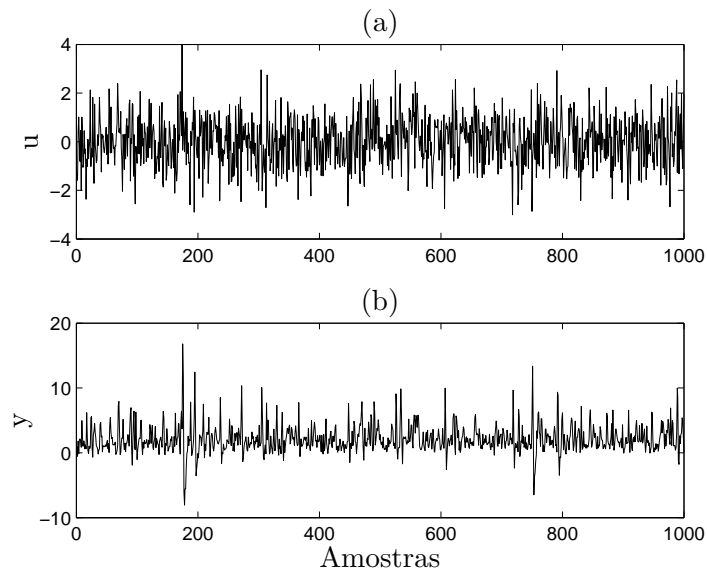


Figura 3.6: Dados dinâmicos de identificação e validação: Exemplo ilustrado.

Primeira metade dados de identificação e segunda metade de validação, sendo: (a) sinal de entrada do sistema, (b) sinal de saída do sistema, (dadospiroddi-ag@).

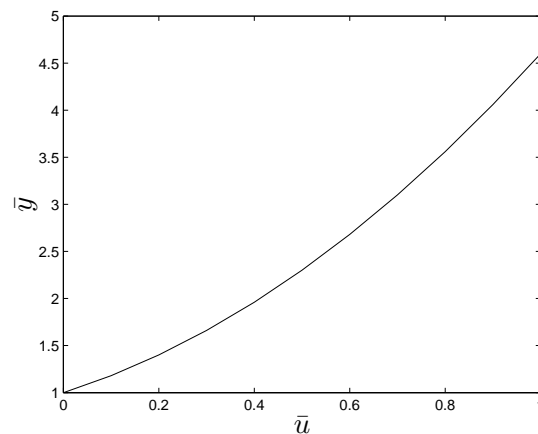


Figura 3.7: Dados estáticos: Exemplo ilustrado.

Característica estática do sistema representado pela Equação 3.1, (dadospiroddi-ag@).

3.4.1 Seleção dos agrupamentos *verdadeiros*

Conforme o algoritmo proposto, o primeiro passo consiste em definir, de forma geral, os termos que farão parte das estruturas-candidatas. Partindo da ideia dos pesquisadores Bonin et al. (2010), adotou-se a seguinte conduta. Considerou-se um modelo com grau de não-linearidade máximo $l = 3$ e máximos atrasos $n_u = 3$ e $n_y = 2$ e todas as possíveis combinações, resultando em um conjunto de 56 regressores candidatos distribuídos em 10 agrupamentos, listados a seguir.

$$\begin{aligned}
\Omega_0 &= [1]; \\
\Omega_y &= [y(k-1) \quad y(k-2)]; \\
\Omega_u &= [u(k-1) \quad u(k-2) \quad u(k-3)]; \\
\Omega_{y^2} &= [y(k-1)y(k-1) \quad y(k-1)y(k-2) \quad y(k-2)y(k-2)]; \\
\Omega_{yu} &= [y(k-1)u(k-1) \quad y(k-1)u(k-2) \quad y(k-1)u(k-3) \quad y(k-2)u(k-1) \\
&\quad y(k-2)u(k-2) \quad y(k-2)u(k-3)]; \\
\Omega_{u^2} &= [u(k-1)u(k-1) \quad u(k-1)u(k-2) \quad u(k-1)u(k-3) \quad u(k-2)u(k-2) \\
&\quad u(k-2)u(k-3) \quad u(k-3)u(k-3)]; \\
\Omega_{y^3} &= [y(k-1)y(k-1)y(k-1) \quad y(k-1)y(k-1)y(k-2) \quad y(k-1)y(k-2)y(k-2) \\
&\quad y(k-2)y(k-2)y(k-2)]; \\
\Omega_{y^2u} &= [y(k-1)y(k-1)u(k-1) \quad y(k-1)y(k-1)u(k-2) \quad y(k-1)y(k-1)u(k-3) \\
&\quad y(k-1)y(k-2)u(k-1) \quad y(k-1)y(k-2)u(k-2) \quad y(k-1)y(k-2)u(k-3) \\
&\quad y(k-2)y(k-2)u(k-1) \quad y(k-2)y(k-2)u(k-2) \quad y(k-2)y(k-2)u(k-3)]; \\
\Omega_{yu^2} &= [y(k-1)u(k-1)u(k-1) \quad y(k-1)u(k-1)u(k-2) \quad y(k-1)u(k-1)u(k-3) \\
&\quad y(k-1)u(k-2)u(k-2) \quad y(k-1)u(k-2)u(k-3) \quad y(k-1)u(k-3)u(k-3) \\
&\quad y(k-2)u(k-1)u(k-1) \quad y(k-2)u(k-1)u(k-2) \quad y(k-2)u(k-1)u(k-3) \\
&\quad y(k-2)u(k-2)u(k-2) \quad y(k-2)u(k-2)u(k-3) \quad y(k-2)u(k-3)u(k-3)]; \\
\Omega_{u^3} &= [u(k-1)u(k-1)u(k-1) \quad u(k-1)u(k-1)u(k-2) \quad u(k-1)u(k-1)u(k-3) \\
&\quad u(k-1)u(k-2)u(k-2) \quad u(k-1)u(k-2)u(k-3) \quad u(k-1)u(k-3)u(k-3) \\
&\quad u(k-2)u(k-2)u(k-2) \quad u(k-2)u(k-2)u(k-3) \quad u(k-2)u(k-3)u(k-3) \\
&\quad u(k-3)u(k-3)u(k-3)].
\end{aligned}$$

A estrutura inicial tem 56 termos e 10 agrupamentos. Deve-se organizar as estruturas-candidatas para executar o próximo teste: gerar as curvas Pareto

e verificar a possibilidade de eliminar os agrupamentos espúrios. Portanto, a partir da estrutura de 10 agrupamentos, em sequência, elimina-se um dos agrupamentos e reestimam-se os parâmetros do modelo que tem os 9 agrupamentos restantes. Dessa forma, obtém-se os seguintes 11 modelos:

$$\begin{aligned}
\mathcal{G}_1 &= \underline{\Omega}_0 + \underline{\Omega}_y + \underline{\Omega}_u + \underline{\Omega}_{y^2} + \Omega_{yu} + \underline{\Omega}_{u^2} + \Omega_{y^3} + \Omega_{y^2u} + \Omega_{yu^2} + \Omega_{u^3} \\
\mathcal{G}_2 &= \underline{\Omega}_y + \underline{\Omega}_u + \underline{\Omega}_{y^2} + \Omega_{yu} + \underline{\Omega}_{u^2} + \Omega_{y^3} + \Omega_{y^2u} + \Omega_{yu^2} + \Omega_{u^3} \\
\mathcal{G}_3 &= \underline{\Omega}_0 + \underline{\Omega}_u + \underline{\Omega}_{y^2} + \Omega_{yu} + \underline{\Omega}_{u^2} + \Omega_{y^3} + \Omega_{y^2u} + \Omega_{yu^2} + \Omega_{u^3} \\
\mathcal{G}_4 &= \underline{\Omega}_0 + \underline{\Omega}_y + \underline{\Omega}_{y^2} + \Omega_{yu} + \underline{\Omega}_{u^2} + \Omega_{y^3} + \Omega_{y^2u} + \Omega_{yu^2} + \Omega_{u^3} \\
\mathcal{G}_5 &= \underline{\Omega}_0 + \underline{\Omega}_y + \underline{\Omega}_u + \Omega_{yu} + \underline{\Omega}_{u^2} + \Omega_{y^3} + \Omega_{y^2u} + \Omega_{yu^2} + \Omega_{u^3} \\
\mathcal{G}_6 &= \underline{\Omega}_0 + \underline{\Omega}_y + \underline{\Omega}_u + \underline{\Omega}_{y^2} + \underline{\Omega}_{u^2} + \Omega_{y^3} + \Omega_{y^2u} + \Omega_{yu^2} + \Omega_{u^3} \\
\mathcal{G}_7 &= \underline{\Omega}_0 + \underline{\Omega}_y + \underline{\Omega}_u + \underline{\Omega}_{y^2} + \Omega_{yu} + \Omega_{y^3} + \Omega_{y^2u} + \Omega_{yu^2} + \Omega_{u^3} \\
\mathcal{G}_8 &= \underline{\Omega}_0 + \underline{\Omega}_y + \underline{\Omega}_u + \underline{\Omega}_{y^2} + \Omega_{yu} + \underline{\Omega}_{u^2} + \Omega_{y^2u} + \Omega_{yu^2} + \Omega_{u^3} \\
\mathcal{G}_9 &= \underline{\Omega}_0 + \underline{\Omega}_y + \underline{\Omega}_u + \underline{\Omega}_{y^2} + \Omega_{yu} + \underline{\Omega}_{u^2} + \Omega_{y^3} + \Omega_{yu^2} + \Omega_{u^3} \\
\mathcal{G}_{10} &= \underline{\Omega}_0 + \underline{\Omega}_y + \underline{\Omega}_u + \underline{\Omega}_{y^2} + \Omega_{yu} + \underline{\Omega}_{u^2} + \Omega_{y^3} + \Omega_{y^2u} + \Omega_{u^3} \\
\mathcal{G}_{11} &= \underline{\Omega}_0 + \underline{\Omega}_y + \underline{\Omega}_u + \underline{\Omega}_{y^2} + \Omega_{yu} + \underline{\Omega}_{u^2} + \Omega_{y^3} + \Omega_{y^2u} + \Omega_{yu^2},
\end{aligned}$$

em que os agrupamentos genuínos estão indicados em negrito e sublinhados.

Para cada uma das 11 estruturas (\mathcal{G}_1 a \mathcal{G}_{11}), realizou-se a estimação multiobjetivo (ver Seção 2.4) e traçaram-se as Curvas Pareto, como mostrado na Figura 3.8.

Os modelos \mathcal{G}_2 a \mathcal{G}_5 e \mathcal{G}_7 , modelos em que faltam algum agrupamento genuíno tiveram a curva Pareto visualmente mais afastadas da origem (isso pode ser quantificado pela área). Por outro lado, os modelos \mathcal{G}_1 , \mathcal{G}_6 e \mathcal{G}_8 a \mathcal{G}_{11} , melhor visualizado na Figura 3.9, permaneceram próximos um dos outros e mais próximos da origem em relação aos demais.

Portanto, conjectura-se que a falta de um agrupamento genuíno seja a causa do deslocamento da curva Pareto. Ou seja, agrupamentos que não influenciam (ou pouco influenciam) na área formada pela curva Pareto gerada, pode ser descartado. Seguindo essa premissa, é factível eliminar os termos pertencentes aos agrupamentos:

$$[\Omega_{yu} \quad \Omega_{y^3} \quad \Omega_{y^2u} \quad \Omega_{yu^2} \quad \Omega_{u^3}].$$

Isso implica excluir 41 regressores, o que equivale a uma redução de 73,2% do conjunto inicial de possíveis termos.

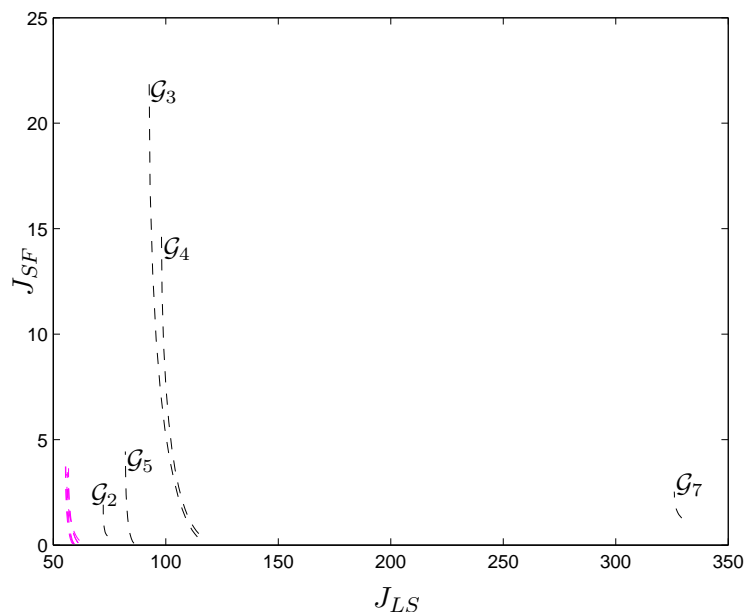


Figura 3.8: Conjuntos Pareto dos modelos candidatos: seleção dos agrupamentos verdadeiros.

As curvas Pareto tracejadas, na parte direita do gráfico correspondem a cada um dos modelos de 10 agrupamentos nos quais falta um agrupamento *verdadeiro*.

3.4.2 Seleção dos regressores *verdadeiros*

A etapa final consiste em executar a Rotina II. Selecionar, dentre os 15 termos restantes, os 5 genuínos. É preciso, antes, combinar os regressores a fim de formar as estruturas-candidatas. Um procedimento simples e intuitivo é organizar os termos como na descrição anterior. Ou seja, a partir da estrutura de 15 termos, em sequência, eliminar um dos termos e reestimar os parâmetros que tem os 14 termos restantes. Para esses modelos, gerar a curva Pareto correspondente e analisar a área, conforme métrica apresentada. Uma vez sendo as áreas próximas, aplica-se o decisor de *melhor sincronismo* (ver Seção 2.4.3).

Contudo, para o exemplo ilustrado, serão analisadas 5 estruturas candidatas apenas para mostrar o potencial das ferramentas. Escolheu-se as estruturas-candidatas, conforme descrito a seguir.

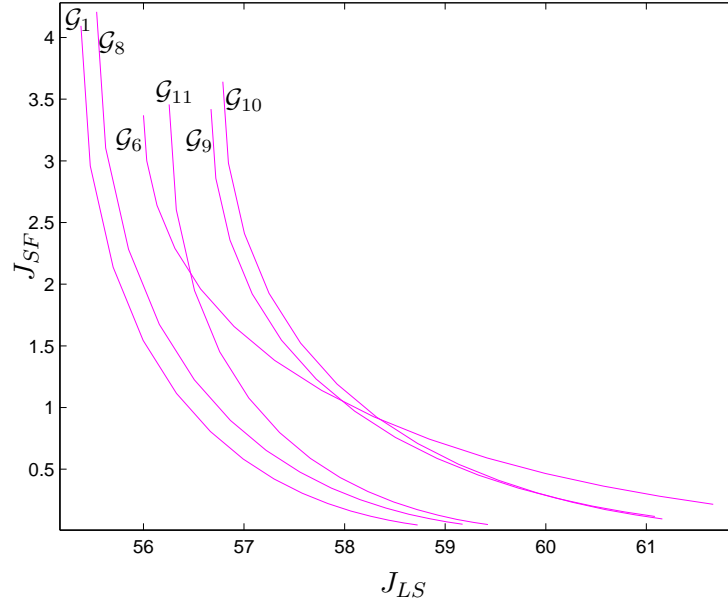


Figura 3.9: Conjuntos Pareto dos modelos candidatos \mathcal{G}_1 , \mathcal{G}_6 e \mathcal{G}_8 a \mathcal{G}_{11} .

As curvas Pareto, ampliadas, correspondem a cada um dos modelos de 10 agrupamentos nos quais estão presentes todos os 5 agrupamentos *verdadeiros*.

$$\begin{aligned}
 \mathcal{M}_1 &= 1 \quad y(k-1) \quad u(k-2) \quad y(k-2)^2 \quad u(k-1)^2; \\
 \mathcal{M}_2 &= 1 \quad y(k-1) \quad u(k-2) \quad y(k-2)^2 \quad u(k-1)^2 \quad y(k-2); \\
 \mathcal{M}_3 &= 1 \quad y(k-1) \quad u(k-2) \quad y(k-2)^2 \quad u(k-1)^2 \quad y(k-1)y(k-2); \\
 \mathcal{M}_4 &= 1 \quad y(k-1) \quad u(k-2) \quad y(k-2)^2 \quad u(k-1)^2 \quad u(k-1)u(k-2); \\
 \mathcal{M}_5 &= 1 \quad u(k-2) \quad y(k-2)^2 \quad u(k-1)^2 \quad y(k-2);
 \end{aligned}$$

em que \mathcal{M}_1 é a estrutura nominal, \mathcal{M}_2 a \mathcal{M}_4 são estruturas sobreparametrizadas com termos de agrupamentos genuínos e \mathcal{M}_5 uma estrutura com um termo espúrio, $y(k-2)$ do agrupamento genuíno Σ_y .

Para cada uma das cinco estruturas, fez-se o processo de estimação multi-objetivo e gerou-se a curva Pareto correspondente, como ilustra a Figura 3.10. Nessa figura, é importante ressaltar que os índices J_{SF} e J_{LS} foram calculados

utilizando-se *dados de estimação*. Pode-se constatar o desempenho inferior do modelo \mathcal{M}_5 , modelo no qual falta um regressor verdadeiro. A Figura 3.11 ilustra, de forma ampliada, os conjuntos Pareto correspondentes aos modelos sobreparametrizados mais o Pareto correspondente ao modelo nominal. Entre esses modelos, pela análise dos Pareto é difícil, a princípio, distingui-los. Dessa forma, aplica-se o decisor de *melhor sincronismo*.

As Figuras 3.12 e 3.13 mostram o erro e o custo de sincronismo, respectivamente. Analisando o erro de sincronismo, o modelo \mathcal{M}_5 é facilmente eliminado. Os modelos \mathcal{M}_2 , \mathcal{M}_3 e \mathcal{M}_4 são também eliminados na análise do erro e custo de sincronismo, apesar de possuírem valores bem próximos.

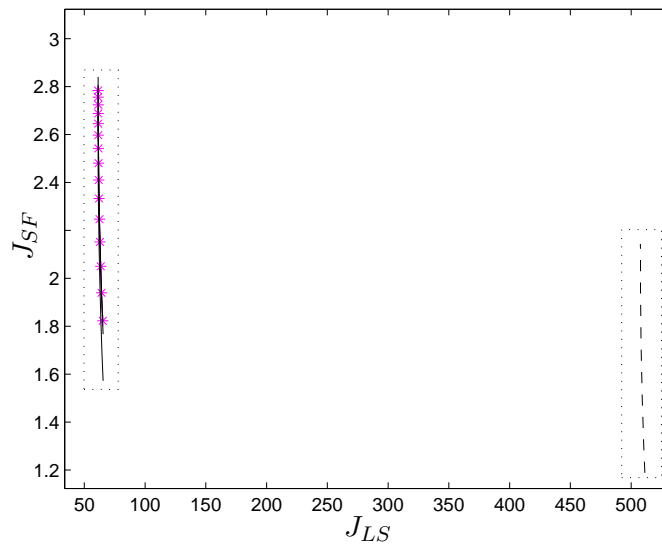


Figura 3.10: Curvas Pareto dos modelos candidatos \mathcal{M}_1 a \mathcal{M}_5 .

Curvas Pareto dos modelos candidatos. Destaque para dois grupos de curvas Pareto: um grupo afastado da origem e um grupo mais próximo (retângulo à esquerda, ampliado na Figura 3.11).

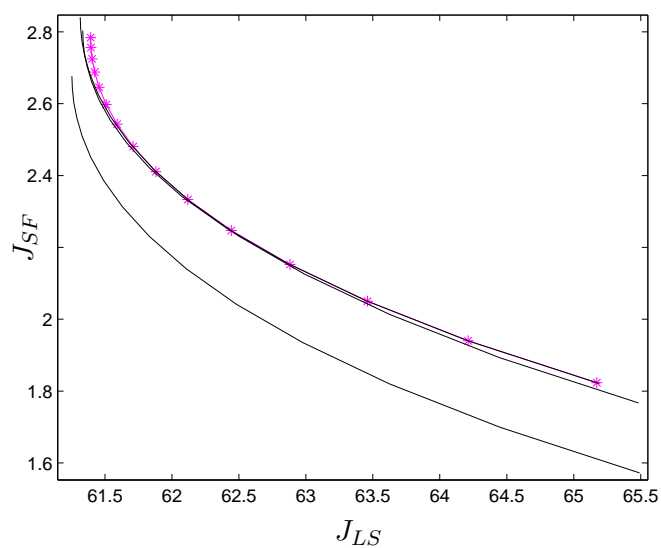


Figura 3.11: Curvas Pareto dos modelos candidatos \mathcal{M}_1 a \mathcal{M}_4 .

Curvas Pareto dos modelos candidatos \mathcal{M}_1 a \mathcal{M}_4 . ($-^*$) modelo nominal e ($-$) modelos sobreparametrizados.

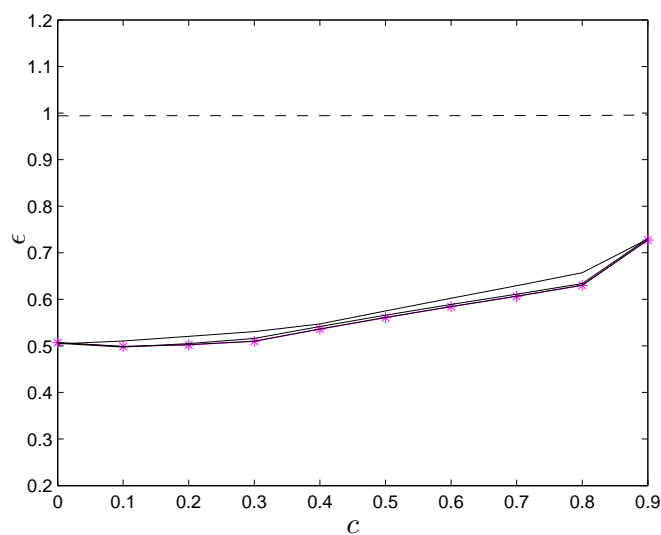


Figura 3.12: Erro máximo de sincronização.

Erro máximo de sincronização: ($-^*$) \mathcal{M}_1 , ($-$) \mathcal{M}_2 a \mathcal{M}_4 e ($--$) \mathcal{M}_5 .

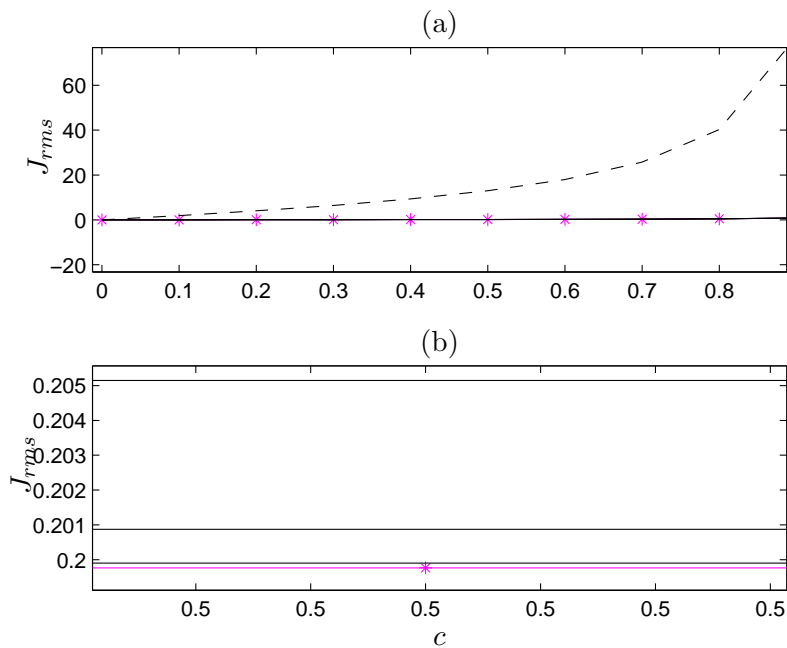


Figura 3.13: Custo de sincronização.

(a) Custo de sincronização para os mesmos modelos da Figura 3.12. (b) Figura em escala ampliada.

3.5 Conclusões do capítulo

Esse capítulo apresentou um procedimento baseado em técnicas multiobjetivo para auxílio na etapa de seleção de estruturas de modelos NARX polinomiais. Um exemplo ilustrado foi descrito com o objetivo de enfatizar e facilitar a compreensão do desenvolvimento sugerido.

Mostrou-se que, por meio dos conjuntos Pareto-ótimos formado pelos funcionais estático e dinâmico, juntamente com uma tomada de decisão adequada, é possível discriminar entre termos espúrios e genuínos. Vale ressaltar que a característica estática é relevante em uma série de aplicações e que são facilmente obtidas a partir de modelos NARX polinomiais.

A metodologia apresentada revelou-se poderosa e os próximos capítulos mostrarão sua aplicação, bem como seu desempenho e robustez em vários outros exemplos numéricos.

Resultados de Simulação

“Contra o positivismo que pára perante os fenômenos e diz: ‘Há apenas fatos’, eu digo: ‘Ao contrário, fatos é o que não há; há apenas interpretações’.”

Friedrich Nietzsche

4.1 Introdução

Este capítulo tem como objetivo aplicar a metodologia desenvolvida no Capítulo 3. Pretende-se verificar o desempenho e a robustez do critério de otimização bi-objetivo na determinação da estrutura de modelos NARX. Para isso, serão estudados três exemplos simulados.

4.2 Exemplos simulados

Dentre os exemplos simulados, o primeiro exemplo (Seção 4.2.1), estudado por Barroso et al. (2007), tem como objetivo mostrar um possível comportamento das curvas Pareto para modelos de estruturas diferentes. O segundo exemplo (Seção 4.2.2), sugerido por Piroddi e Spinelli (2003), tem o intuito de averiguar a aplicação das duas técnicas de tomada de decisão discutidas no Capítulo 2, bem como a caracterização de incertezas e particularidades. Na Seção 4.2.3 o terceiro exemplo é analisado. Proposto por Bonin et al. (2010), o objetivo desse terceiro exemplo é avaliar o procedimento quando o sinal de

entrada é persistentemente excitante, porém de menor ordem. Essa é um situação menos favorável, importante para avaliar a robustez do método.

Como a estrutura dos sistemas utilizados nesses três exemplos é conhecida, o desempenho final do método é facilmente avaliado. Ou seja, compara-se a estrutura obtida com a estrutura que gerou os dados. Além disso, avaliou-se e comparou-se o método proposto quanto à robustez e à complexidade computacional.

4.2.1 Exemplo 1

Barroso et al. (2007) determinaram, dentre uma grande gama de possíveis estruturas potenciais, por meio da taxa de redução de erro (ERR) e do critério de informação de Akaike, a estrutura do processo de polimerização, dada por:

$$y(k) = \begin{bmatrix} y(k-1) & y(k-2) & y(k-3) & y(k-4) & u(k-1) \\ u(k-1)^2 & u(k-4) & u(k-1)^4 & u(k-1)^3 & 1 \end{bmatrix}. \quad (4.1)$$

Os dados dinâmicos utilizados na identificação e na validação podem ser vistos na Figura 4.1. A Figura 4.2 mostra a curva estática do modelo. O modelo usado para produzir os dados¹ foi descrito por Ray (1972).

A partir da estrutura (Equação 4.1), adicionaram-se termos de agrupamentos espúrios e não-espúrios e retiraram-se termos, em princípio, importantes. Várias estruturas foram obtidas, classificadas em três grupos. O primeiro grupo é formado pela adição de termos espúrios pertencentes a agrupamentos espúrios, o segundo é formado pela adição de termos espúrios pertencentes a agrupamentos não-espúrios e o terceiro grupo formado pela retirada de termos não-espúrios.

Para cada nova estrutura, realizou-se a identificação multiobjetivo, aplicou-se o decisor de mínima correlação (Seção 2.4.2) e traçaram-se as curvas Pareto, como ilustra a Figura 4.3. Nessa figura, é importante ressaltar que os índices J_{SF} e J_{LS} foram calculados utilizando-se *dados de estimação*. O uso de dados de estimação para a construção dos conjuntos Paretos tem a vantagem de permitir discriminar mais claramente estruturas sobre e subparametrizadas.

¹Dados disponíveis em: <http://www.cpdee.ufmg.br/~MACSIN>.

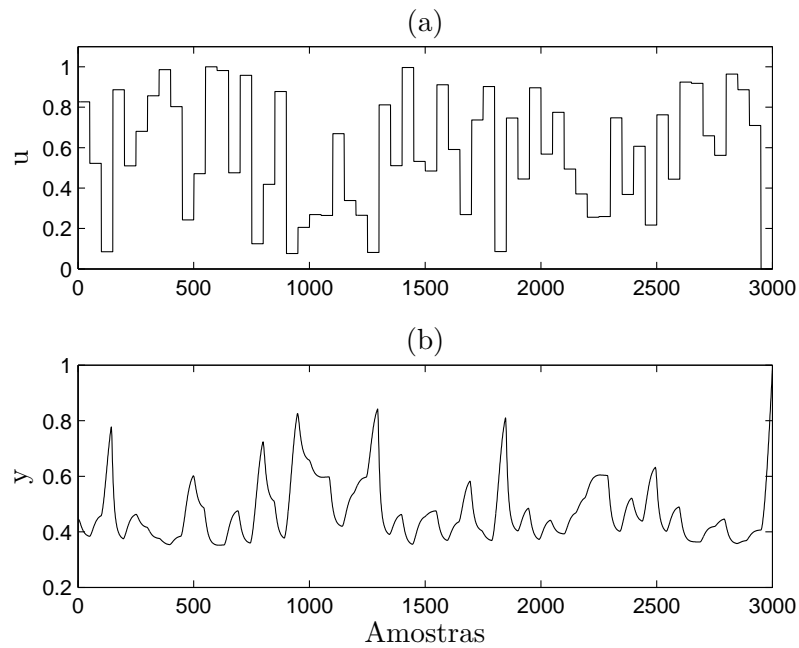


Figura 4.1: Dados dinâmicos de identificação e validação: Exemplo 1.

Dados dinâmicos de identificação (amostras 1:1500) e de validação (amostras 1501:3000), sendo: (a) sinal de entrada do sistema em p.u, (b) sinal de saída do sistema em p.u, (`dinamica@`).

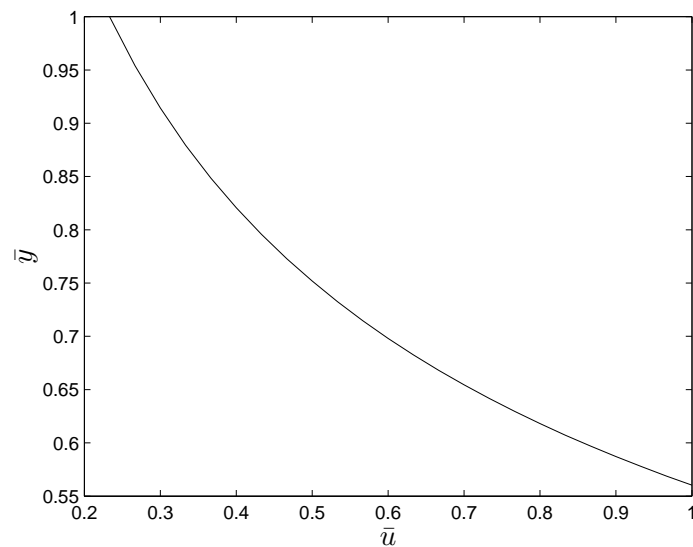


Figura 4.2: Dados estáticos: Exemplo 1.

Curva estática, sendo (\bar{u}) e (\bar{y}) em p.u, (`estatica@`).

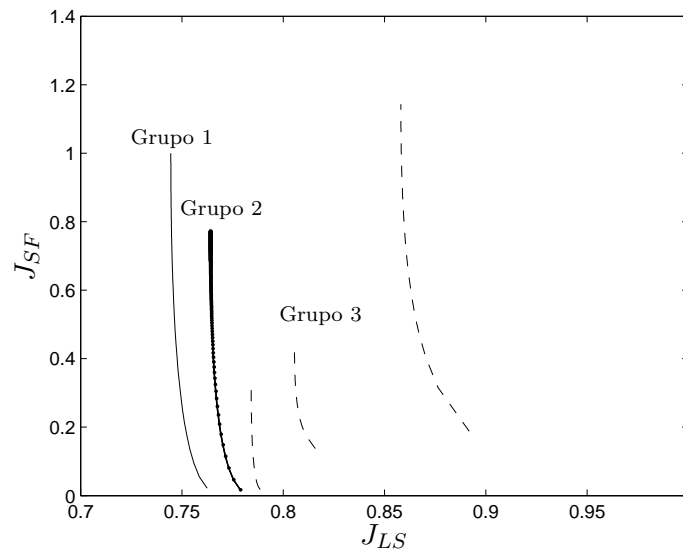


Figura 4.3: Conjuntos Pareto dos modelos candidatos: Exemplo 1.

Conjuntos Pareto dos modelos candidatos. Cada curva Pareto representa uma estrutura. Grupo 1: estruturas sobreparametrizadas com termos cruzados; Grupo 2: estruturas sobreparametrizadas e estrutura nominal, e Grupo 3: três estruturas subparametrizadas, indicadas por (—).

A inclusão de termos pertencentes a agrupamentos espúrios é pior que a inclusão de termos pertencentes a agrupamentos não-espúrios (Mendes e Billings, 2001), e as curvas Pareto de estruturas subparametrizadas ficam mais afastadas da origem. Infelizmente, há casos em que as curvas Pareto não aparecem tão claramente ordenadas e separadas. Contudo, ainda é possível distingui-las, como será ilustrado no próximo exemplo.

4.2.2 Exemplo 2

Considere o seguinte sistema representado pela Equação 4.2 (Piroddi e Spinelli, 2003):

$$y(k) = 0,5y(k-1) + 0,8u(k-2) + u(k-1)^2 - 0,05y(k-2)^2 + 0,5 + e(k), \quad (4.2)$$

em que a entrada $u(\cdot)$, cuja função de autocorrelação² (FAC) se encontra representada na Figura 4.4, foi gerada a partir de um sinal do tipo ruído branco com distribuição Gaussiana, média nula e variância unitária. O ruído $e(k)$ foi gerado da mesma forma, contudo, com variância igual a 0,05. A Figura 4.5 apresenta os sinais dinâmico de entrada e saída do sistema. Seguindo o método apresentado na Seção 2.2.2, gerou-se a curva estática do sistema, ilustrada na Figura 4.6.

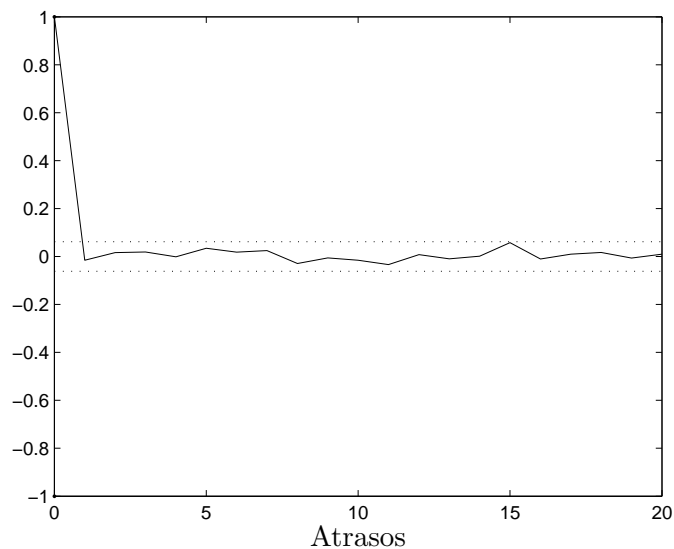


Figura 4.4: Função de autocorrelação do sinal de entrada – Exemplo 2.

A função de autocorrelação do sinal de entrada assemelha-se à de um ruído branco, que tem a característica impulsiva. As linhas pontilhadas indicam a região de confiança de 95%

²Função de autocorrelação normalizada entre -1 e 1. Nível de confiança de 95%, limites variando de $1,96/\sqrt{N}$ a $-1,96/\sqrt{N}$, em que N é o número de amostras do sinal.

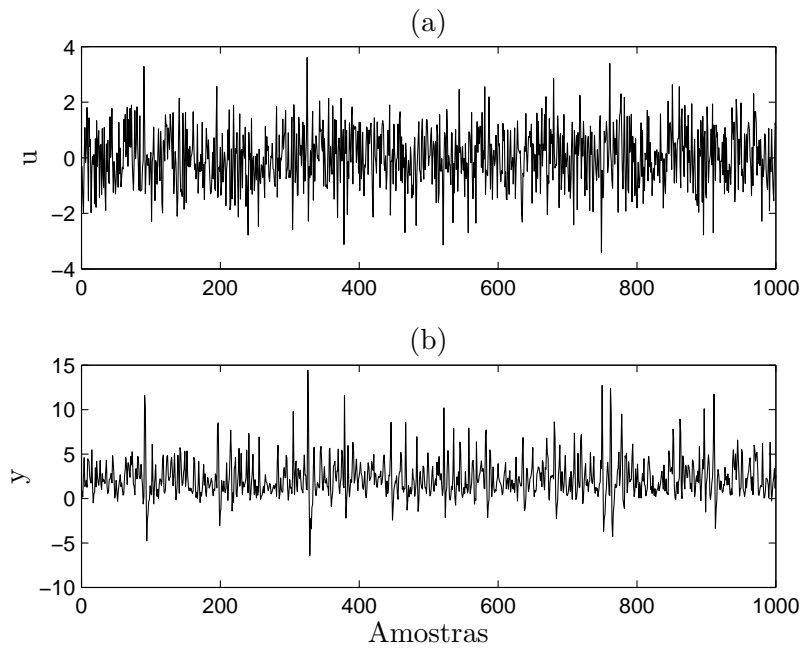


Figura 4.5: Dados dinâmicos de identificação e validação: Exemplo 2.

Dados dinâmicos de identificação (amostras 1:500) e de validação (amostras 501:1000), sendo: (a) sinal de entrada do sistema, (b) sinal de saída do sistema, (dadospiroddi@).

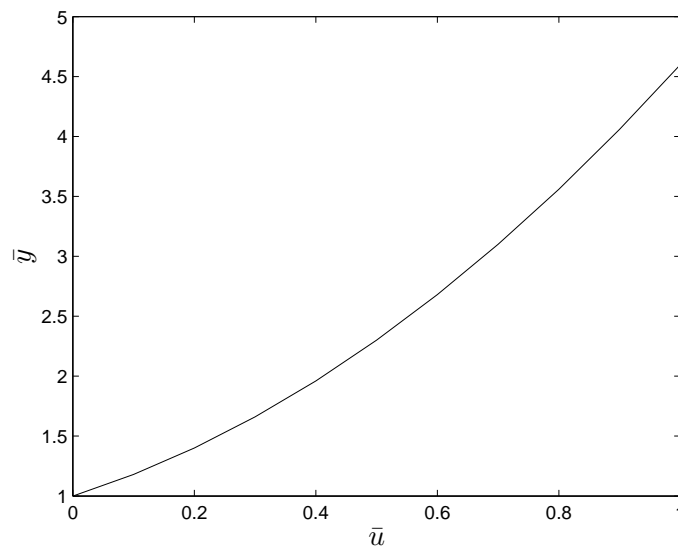


Figura 4.6: Dados estáticos: Exemplo 2.

Característica estática do sistema representado pela Equação 4.2, (dadospiroddi@).

A partir da estrutura que gerou os dados (\mathcal{M}_1), outras 13 diferentes estruturas foram obtidas acrescentando ou retirando termos de \mathcal{M}_1 . As estruturas foram classificadas em grupos, conforme Exemplo 1. A Tabela 4.1 resume os termos acrescentados e retirados tendo como base o modelo que gerou os dados.

Tabela 4.1: Regressores acrescentados e removidos de \mathcal{M}_1 (ver Equação 4.2).

Grupo	Ação	Regressor	Modelo
1	acrescentar	$y(k-1)^3$	\mathcal{M}_2
	acrescentar	$u(k-1)^3$	\mathcal{M}_3
	acrescentar	$y(k-1)u(k-1)$	\mathcal{M}_4
	acrescentar	$y(k-1)^2u(k-1)$	\mathcal{M}_5
	acrescentar	$y(k-1)u(k-1)^2$	\mathcal{M}_6
2	acrescentar	$y(k-2)$	\mathcal{M}_7
	acrescentar	$u(k-1)$	\mathcal{M}_8
	acrescentar	$u(k-2)^2$	\mathcal{M}_9
	acrescentar	$y(k-1)2$	\mathcal{M}_{10}
3	remover	$y(k-2)^2$	\mathcal{M}_{11}
	remover	$y(k-1)$	\mathcal{M}_{12}
	remover	$u(k-2)$	\mathcal{M}_{13}

Para cada um dos modelos, fez-se o processo de identificação multiobjetivo, aplicou-se o decisor de correlação e gerou-se a curva Pareto correspondente (Figura 4.7). Pode-se constatar o desempenho inferior dos modelos subparametrizados (\mathcal{M}_{11} a \mathcal{M}_{13}).

À semelhança do que foi feito no exemplo anterior, na Figura 4.7, os índices J_{SF} e J_{LS} foram calculados utilizando-se *dados de estimação*.

A Figura 4.8 ilustra, de forma ampliada, os conjuntos Paretos correspondentes aos modelos sobreparametrizados (\mathcal{M}_2 a \mathcal{M}_{10}) mais o Pareto correspondente ao modelo nominal, o que gerou os dados (\mathcal{M}_1). Um ponto importante a observar na Figura 4.8 é que alguns conjuntos Pareto se cruzam. Esse fenômeno também foi observado no exemplo do Capítulo 3. Se os conjuntos Pareto de dois modelos se cruzam significa que nenhuma das duas estruturas é “absolutamente” melhor do que a outra. Em situações como essa o uso de conjuntos Pareto é muito informativo. Vale a pena reforçar que no presente exemplo, assim como no exemplo do Capítulo 3 (ver Figura 3.8), esses cruzamentos só

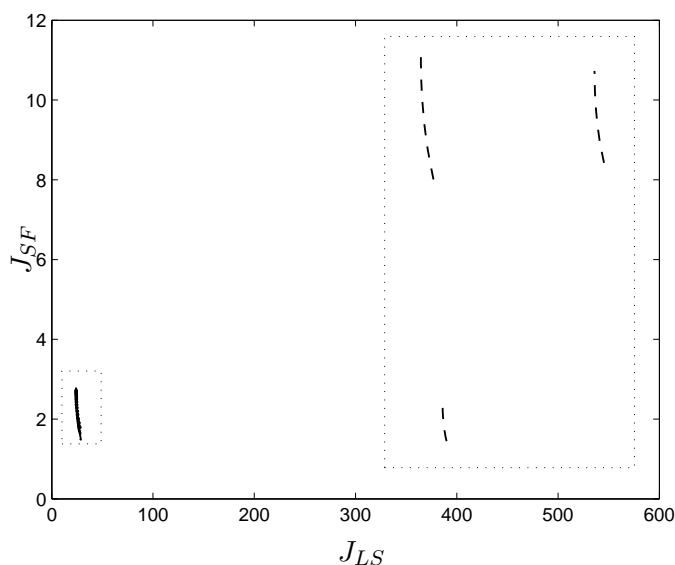


Figura 4.7: Conjuntos Pareto dos modelos candidatos (ver Tabela 4.1).

Conjuntos Pareto dos modelos candidatos. Destaque para os conjuntos Paretos do grupo 3 de modelos (retângulo maior à direita) e grupos 1 e 2 (retângulo menor à esquerda, ampliado na Figura 4.8).

se verificaram entre conjuntos Pareto de estruturas sobreparametrizadas.

Se uma “melhor” estrutura tiver que ser escolhida, então deve-se definir um critério. Aqui tanto o critério de mínima correlação (Seção 2.4.2) como o de sincronismo (Seção 2.4.3) poderiam ser utilizados. A fim de poder realizar uma simulação Monte Carlo, escolheu-se o critério de mínima correlação, pois é mais fácil de automatizar.

Portanto, foi produzido um conjunto de 500 séries de dados usando o mesmo modelo. Para cada uma dessas séries foram geradas as curvas Pareto para cada um dos 13 modelos, sendo que aplicou-se o decisor de mínima correlação para escolher um, dentre os modelos do Pareto. Portanto, para cada série tinham-se um conjunto de 13 modelos. Dentre esses 13, escolheu-se o “melhor”, novamente utilizando o critério de mínima correlação. Portanto, ao final, ao seguir esse procedimento para cada uma das séries, tinha-se 500 modelos escolhidos. A frequência com que cada modelo foi escolhido como sendo o “melhor” foi: \mathcal{M}_1 14,3%; \mathcal{M}_2 10,2%; \mathcal{M}_3 14,5%; \mathcal{M}_4 2,5%; \mathcal{M}_5 2,7%; \mathcal{M}_6 24,9%; \mathcal{M}_7 5,7%; \mathcal{M}_8

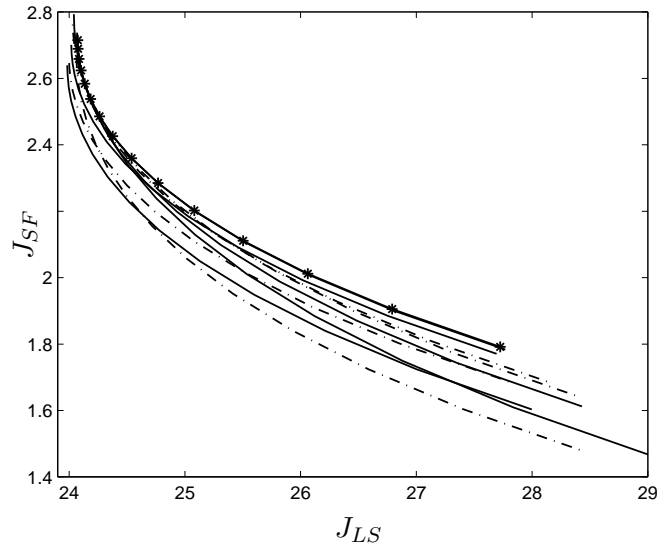


Figura 4.8: Conjuntos Pareto dos modelos \mathcal{M}_2 a \mathcal{M}_{10} (ver Tabela 4.1).

Conjuntos Pareto dos modelos \mathcal{M}_2 a \mathcal{M}_{10} . (—) correspondem às estruturas do grupo 1; (—·) correspondem às estruturas do grupo 2 e (-*) corresponde à estrutura nominal (4.2).

12,1%; \mathcal{M}_9 7,9%; \mathcal{M}_{10} 4,7%; \mathcal{M}_{11} , \mathcal{M}_{12} e \mathcal{M}_{13} 0%. Nesse estudo, chama a atenção o fato de o modelo \mathcal{M}_6 ter sido escolhido o melhor de acordo com o critério de mínima correlação 24,9% das vezes.

Uma outra forma de melhor distinguir esses modelos é aplicando-se o decisor de sincronismo, conforme discutido na Seção 2.4.3.

As Figuras 4.9 e 4.10 mostram o erro e o custo de sincronismo, respectivamente. Assim como o decisor de correlação, o decisor de sincronismo facilmente eliminou os modelos \mathcal{M}_{11} a \mathcal{M}_{13} . Analisando o erro de sincronismo é possível ainda eliminar os modelos \mathcal{M}_4 a \mathcal{M}_6 , que correspondem às três curvas (linha sólida) superiores do grupo inferior na Figura 4.9. Esse três modelos, do grupo 1, são modelos compostos por termos cruzados de agrupamentos espúrios. Os demais modelos concentram-se em uma região na qual tanto o erro quanto o custo de sincronismo são muito próximos e, portanto, parecem ser indiscrimináveis, ao menos pelo critério de sincronismo. É interessante notar que se o critério de sincronismo é utilizado, em vez do de mínima correlação, ao menos neste caso, o modelo \mathcal{M}_6 não seria escolhido como o melhor.

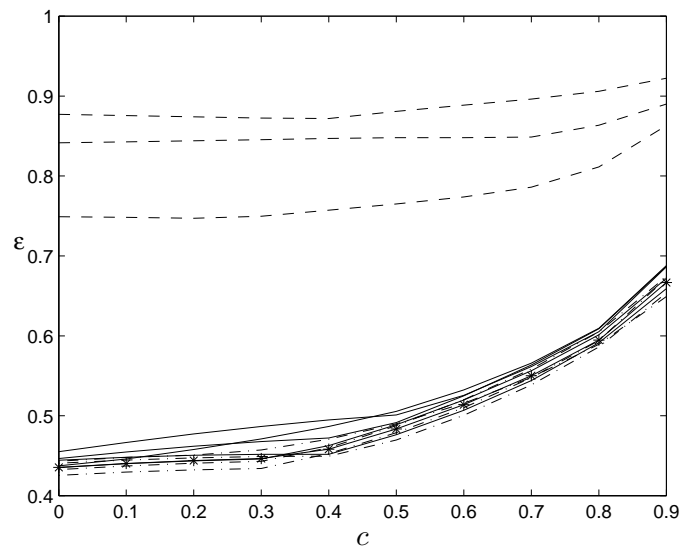


Figura 4.9: Erro máximo de sincronização (ver Equação 2.31).

Erro máximo de sincronização: (-*) \mathcal{M}_1 , (—) \mathcal{M}_2 a \mathcal{M}_6 , (-·) \mathcal{M}_7 a \mathcal{M}_{10} e (---) \mathcal{M}_{11} a \mathcal{M}_{13} .

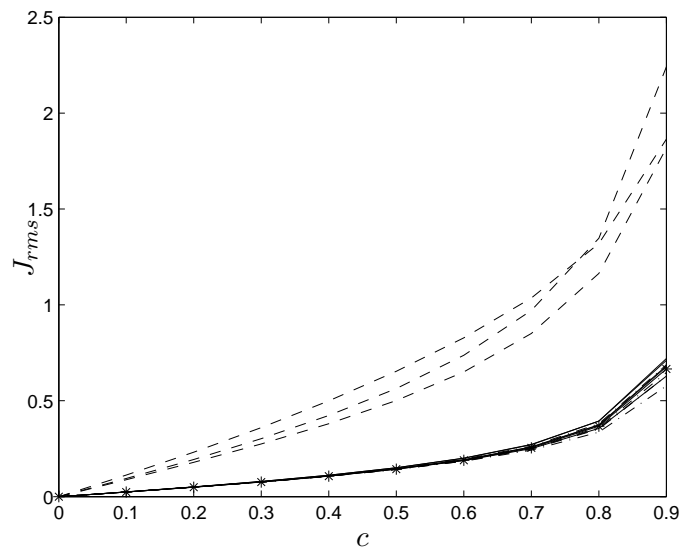


Figura 4.10: Custo da sincronização (ver Equação 2.30) para os mesmos modelos da Figura 4.9.

Um ponto central na discussão acima é perceber que não é necessário imaginar que em um determinado problema exista apenas uma estrutura “correta” que seja melhor para *qualquer* realização de dados e segundo *qualquer* critério. A Figura 4.11 justifica essa afirmativa. A partir da estrutura que gerou os dados, \mathcal{M}_1 , foi produzido um conjunto de 500 séries de dados, com realizações diferentes de ruído. Para cada uma das 500 séries gerou-se uma curva Pareto. Essas curvas delimitam uma *região de incerteza*, conforme ilustrado. As curvas Pareto referentes às estruturas que buscamos (\mathcal{M}_1 a \mathcal{M}_{10}) encontram-se todas dentro da *região de incerteza*³. Uma vez que todas as estruturas representam os dados, se tomarmos como referência o princípio da parcimônia⁴ a estrutura nominal seria a escolhida. Contudo, é possível chegar, por meio dos métodos estabelecidos, à estrutura verdadeira, como descrito a seguir.

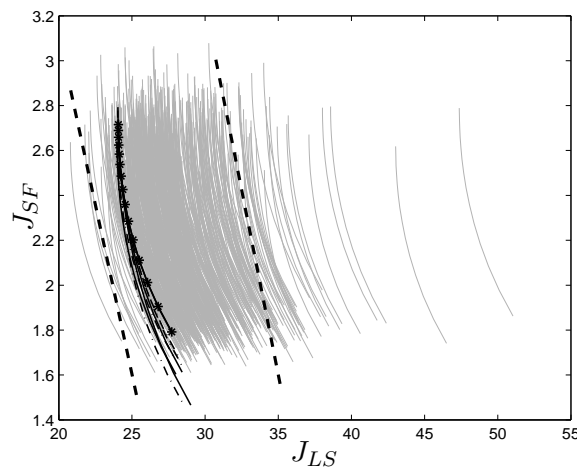


Figura 4.11: Curvas Pareto: região de incerteza.

Região de incerteza relacionada ao modelo nominal, modelo que gerou os dados (ver Equação 4.2). A média e o desvio padrão são: $\bar{J}_{LS} = 28,36$, $\sigma_{LS} = 3,18$ e $\bar{J}_{SF} = 2,37$, $\sigma_{SF} = 0,10$. Para traçar os limites no gráfico, considerou-se um intervalo de confiança de 95%. Os conjuntos Pareto dos modelos \mathcal{M}_1 a \mathcal{M}_{10} estão também representados.

³Carlos Fonseca e colaboradores (Carrano et al., 2007; Fonseca, 1995) utilizaram outras formas de analisar e traçar limites para curvas Pareto. Outros trabalhos podem ser encontrados em: <http://w3.ualg.pt/~cmfonsec/>.

⁴O princípio da parcimônia, também conhecido por navalha de Occam (“Occam’s razor”), diz que a explicação mais simples costuma ser a correta.

Extensão do Exemplo 2

A partir de um conjunto de dados como o gerado (Figuras 4.5), Piroddi e Spinelli (2003), usando o critério ERR, obtiveram uma estrutura com os seguintes 8 regressores: $y(k-1)$, $y(k-2)$, $u(k-1)^2$, $u(k-2)^2$, $u(k-1)$, $y(k-2)^2$, 1 e $u(k-2)$. Como os regressores 2, 4 e 5 são espúrios, esses autores desenvolveram um critério baseado no SRR (*simulation error reduction ratio*) para remover tais termos. Como apontado pelos autores, o procedimento proposto por eles é bastante intenso computacionalmente. A seguir, descreveremos dois procedimentos, um baseado em estimação multiobjetivo e outro baseado no sincronismo, ambos capazes de escolher, entre os 8 regressores, os 5 que são verdadeiros.

O procedimento é bastante simples. A partir da estrutura de 8 termos, em sequência, elimina-se um dos termos e reestimam-se os parâmetros do modelo que tem os 7 termos restantes. Para esse modelo gera-se a curva Pareto e/ou o gráfico de sincronismo. Quando o termo *retirado* é um termo do modelo nominal, a piora de desempenho do modelo de 7 termos restante é muito grande. Se, por outro lado, o termo removido for espúrio, o desempenho do modelo resultante permanece essencialmente inalterado ou melhora. Essa é a interpretação das Figuras 4.12 e 4.13. Note que para isso não é necessário saber qual é o modelo nominal, mas é necessário garantir que a estrutura nominal esteja incluída no modelo que é o “ponto de partida”.

4.2.3 Exemplo 3

Considere o seguinte sistema representado pela Equação 4.3 (Bonin et al., 2010). Esse exemplo é uma extensão do estudo de caso ilustrado na Seção 3.4.

$$\begin{aligned} h(k) &= 0,5h(k-1) + 0,8u(k-2) + u(k-1)^2 - 0,05h(k-2)^2 + 0,5 \\ y(k) &= h(k) + e(k), \end{aligned} \tag{4.3}$$

$$u(k) = 0,5u(k-1) + w(k), \tag{4.4}$$

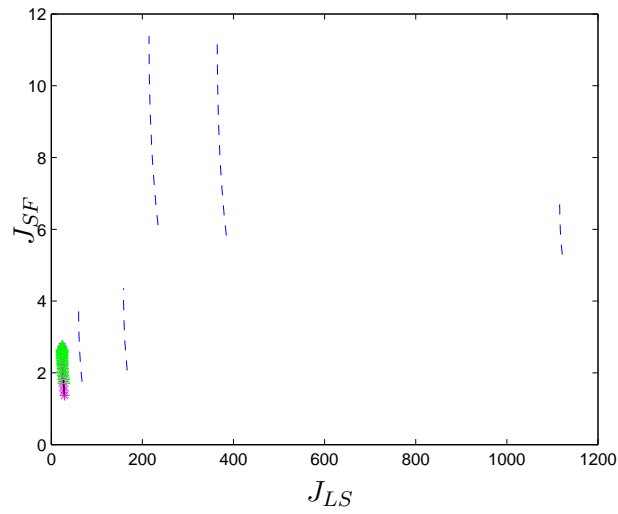


Figura 4.12: Curvas Pareto: extensão do Exemplo 2.

As curvas Pareto tracejadas, na parte direita do gráfico correspondem a cada um dos modelos de 7 termos nos quais faltam um regressor *verdadeiro*.

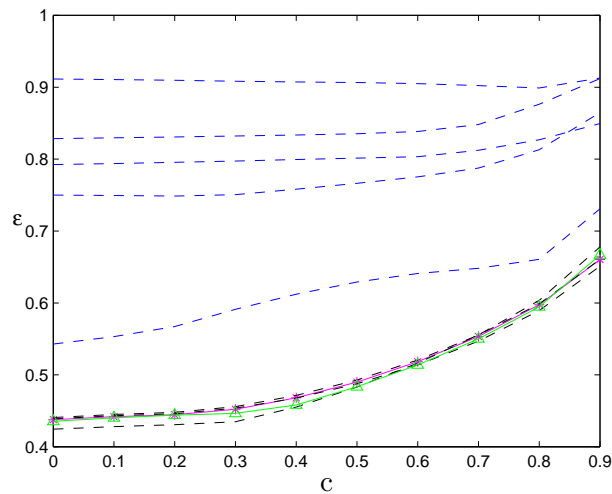


Figura 4.13: Erro de sincronização: extensão do Exemplo 2.

As cinco curvas superiores, correspondentes aos cinco modelos de 7 termos nos quais falta um dos 5 regressores verdadeiros, não sincronizam com a mesma qualidade.

em que $w(\cdot)$ é um ruído branco com distribuição gaussiana, média nula e variância unitária e $e(\cdot)$ definido da mesma forma, porém, com variância de 0,33.

O sistema, com ruído na saída, é excitado com uma entrada persistentemente excitante de ordem menor que infinito (Equação 4.4). Nesses casos, a seleção da estrutura correta tende a ser mais difícil, conforme relata Bonin et al. (2010). A Figura 4.14 mostra a função de autocorrelação do sinal de entrada aplicado. Os dados dinâmico e estático são apresentados nas Figuras 4.15 e 4.16, respectivamente.

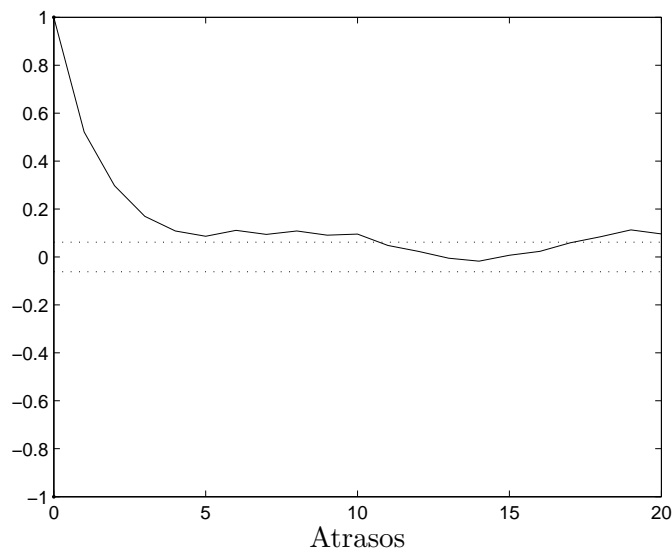


Figura 4.14: Função de autocorrelação do sinal de entrada – Exemplo 3.

A função de autocorrelação do sinal de entrada, relativamente lento, mostra-se diferente de um impulso (comparar com Figura 4.4). As linhas pontilhadas indicam a região de confiança de 95%.

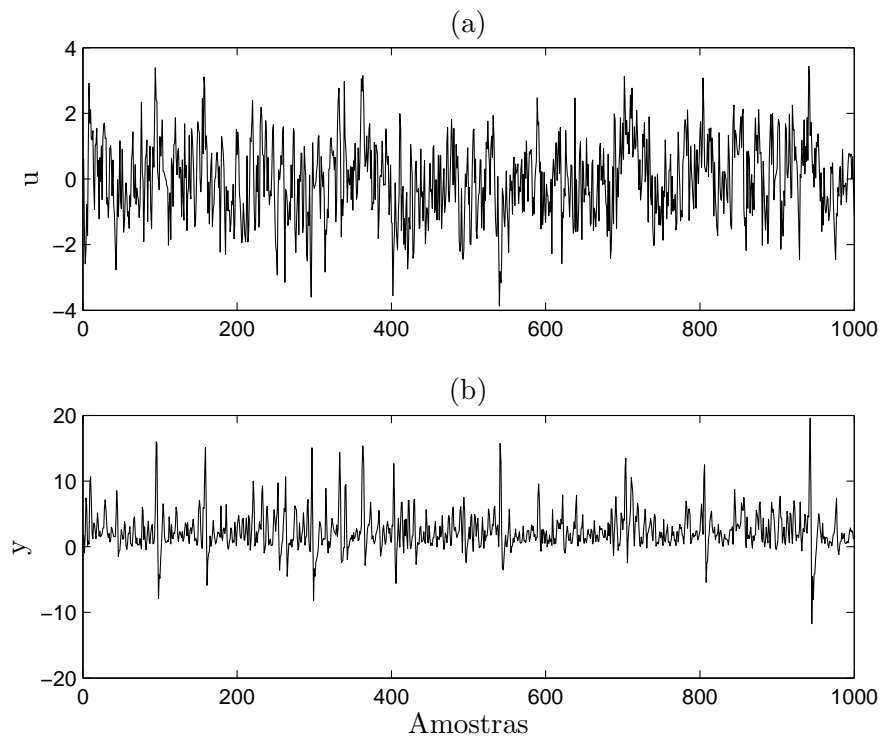


Figura 4.15: Dados dinâmicos de identificação e validação: Exemplo 3.

Dados dinâmicos de identificação (amostras 1:500) e de validação (amostras 501:1000), sendo: (a) sinal de entrada do sistema, (b) sinal de saída do sistema, (dadosel@).

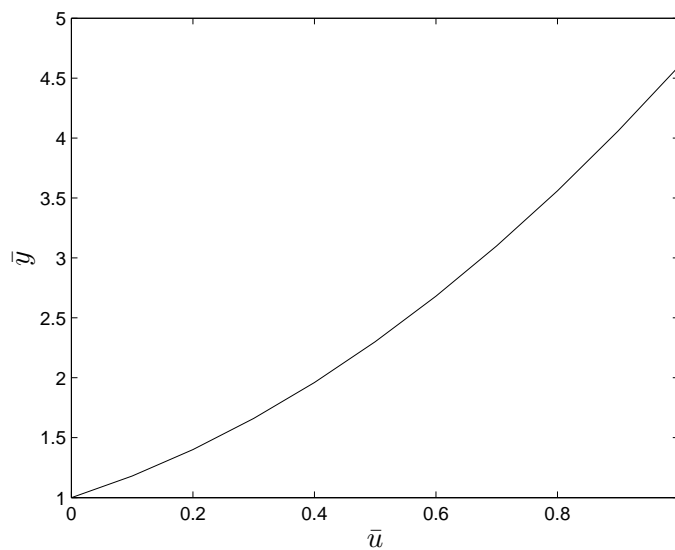


Figura 4.16: Dados estáticos: Exemplo 3.

Característica estática do sistema, (dadosel@).

Para compor a coleção inicial de estruturas candidatas, assim como em Bonin et al. (2010), utilizaram-se modelos NARX com grau de não-linearidade igual a 3, máximo atraso no sinal de entrada igual a 2 e máximo atraso no sinal de saída igual a 3. Conforme Equação 2.14 (ver Seção 2.3), obteve-se um total de 56 termos candidatos e 10 agrupamentos.

Bonin et al. (2010) aplicam o método do LASSO (ver Seção 2.5) para reduzir o número de termos candidatos. De 56 termos iniciais 29 foram eliminados, o que corresponde a uma redução de 51,8%. Contudo, para esse método é necessário escolher o parâmetro de regularização (λ).

A título de comparação, neste trabalho reduziu-se o conjunto de termos candidatos eliminando os falsos agrupamentos, por meio da análise do Pareto. Cinco agrupamentos foram eliminados, reduzindo o número de termos em 73,2% (de 56 regressores para 15). A Figura 4.17 ilustra as Curvas Pareto correspondentes às estruturas-candidatas.

De acordo com a metodologia proposta, após eliminar os agrupamentos espúrios deve-se selecionar os termos verdadeiros. Para isso é necessário combinar os regressores a fim de formar as estruturas-candidatas. Dentre todos os procedimentos estudados, o mais promissor tem sido a verificação termo a termo. Ou seja, a partir da estrutura de 15 termos, em sequência, elimina-se um dos termos e reestimam-se os parâmetros dos 14 termos restantes. Para cada modelo, gera-se a curva Pareto. Essa análise é condizente com os estudos de Bonin et al. (2010), Piroddi (2008) e Piroddi e Spinelli (2003).

Portanto, gerou-se 16 curvas Pareto, conforme representado na Figura 4.18. Cada curva Pareto corresponde a uma estrutura-candidata de 14 termos mais a estrutura composta pelos 15 termos, totalizando 16 curvas.

A Figura 4.18 mostra que os Paretos nos quais falta um regressor *verdadeiro* afastaram-se da origem. Ou seja, tomando como indicador a área (ver Seção 3.3.1), a piora das estruturas subparametrizadas é detectada.

Após identificar o 5 termos genuínos realizou-se o procedimento de estimação bi-objetivo (ver Seção 2.4) e aplicou-se o decisor de mínima correlação, chegando ao modelo final indicado na Tabela 4.2. As Figuras 4.19 e 4.20 mostram, respectivamente, a validação dinâmica e estática do modelo.

Com o objetivo de detalhar algumas discussões, aplicou-se o critério de informação de Akaike (ver Seção 2.3.2) e a taxa de redução de erro (ERR), de-

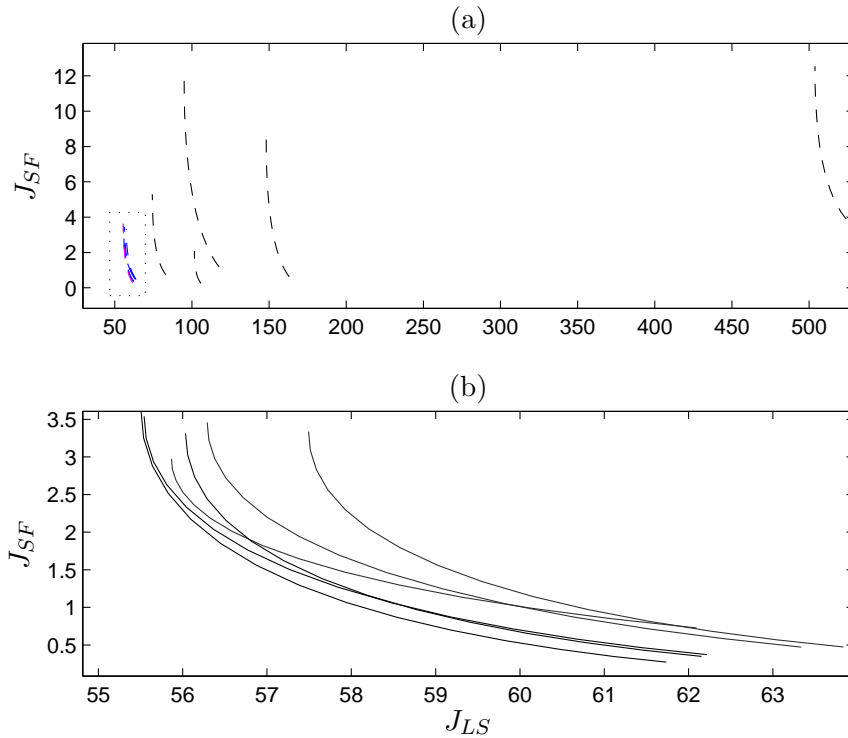


Figura 4.17: Conjunto Pareto dos modelos candidatos: seleção dos agrupamentos verdadeiros, Exemplo 3.

Gráfico (a) em tamanho real, destaque para os conjuntos Pareto dos modelos com todos os agrupamentos *verdadeiros* (retângulo à esquerda) e (b) em escala ampliada, conjuntos Pareto nos quais falta um agrupamento *verdadeiro*.

finida na Seção 2.3.1, para selecionar os regressores. Inicialmente trabalhou-se com o conjunto inicial de 56 termos candidatos e, em seguida, com o conjunto pré-selecionado de 15 termos. O critério de informação de Akaike indicou a utilização de 11 termos para um conjunto inicial de 56 termos candidatos e 7 termos quando o conjunto inicial continha 15 regressores, conforme mostra a Figura 4.21. Os termos foram ordenados por meio do critério ERR e os parâmetros correspondentes estimados utilizando o algoritmo de mínimos quadrados estendido (MQE). As Tabelas 4.3 e 4.4 resumem os principais valores encontrados.

O critério ERR corretamente selecionou os 5 termos genuínos. O AIC não

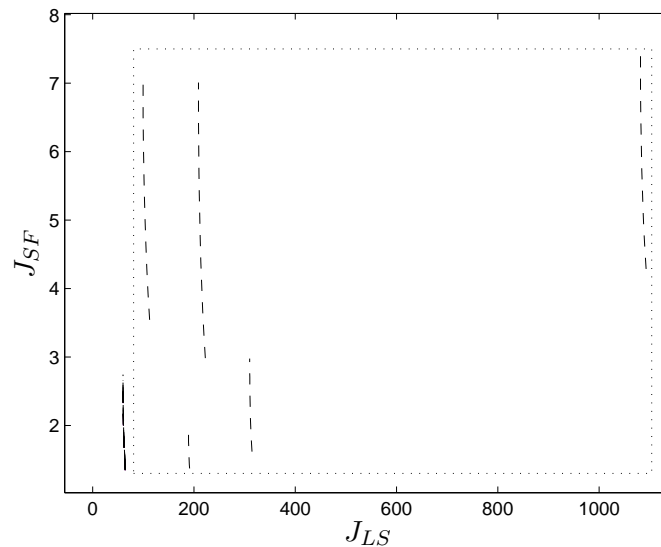


Figura 4.18: Conjunto Pareto dos modelos candidatos: seleção dos regressores verdadeiros, Exemplo 3.

Curvas Pareto dos modelos candidatos. Destaque para os conjuntos Pareto dos modelos nos quais falta um regressor *verdadeiro*.

soube cortar o modelo no ponto correto, mas isso pode ser feito facilmente de outras formas. Além disso, o valor do ERR bem como o valor dos parâmetros dos regressores falsos parecem ser estatisticamente não significativos. Pode-se afirmar que, tanto no critério bi-objetivo quanto nos critérios AIC e ERR, foi possível obter os 5 termos genuínos do sistema.

Complexidade computacional

O custo computacional de um algoritmo é o número de operações com vírgula flutuante requeridas para a sua execução. De um modo geral, o conhecimento do número exato de operações efetuadas por um certo algoritmo não é essencial. Basta, em vez disso, determinar sua ordem de grandeza em função de um parâmetro associado à dimensão do problema (Quarteroni e Fausto, 2006).

A complexidade computacional do procedimento proposto é quantificada levando em consideração o número aproximado de parâmetros estimados. Essa

Tabela 4.2: Modelo final obtido pelo critério bi-objetivo a partir dos dados apresentados nas Figuras 4.15 e 4.16. Termos genuínos em negrito.

Regressor	Ação	Parâmetro estimado	Parâmetro real
1	acrescentado	0,5531	0,5
$y(k-1)$	acrescentado	0,5131	0,5
$y(k-2)$	removido	–	–
$u(k-1)$	removido	–	–
$u(k-2)$	acrescentado	0,8415	0,8
$u(k-3)$	removido	–	–
$y(k-1)^2$	removido	–	–
$y(k-1)y(k-2)$	removido	–	–
$y(k-2)^2$	acrescentado	-0,0506	-0,05
$u(k-1)^2$	acrescentado	0,9889	1
$u(k-1)u(k-2)$	removido	–	–
$u(k-1)u(k-3)$	removido	–	–
$u(k-2)^2$	removido	–	–
$u(k-2)u(k-3)$	removido	–	–
$u(k-3)^2$	removido	–	–

Tabela 4.3: Modelo obtido pelo critério de informação de Akaike e ERR a partir dos dados apresentados na Figura 4.15. Conjunto inicial de 56 termos candidatos. Termos genuínos em negrito.

Regressor	ERR	Parâmetro estimado	Parâmetro real
$u(k-1)^2$	$7,2796 \times 10^{-1}$	1,0236	1
$y(k-1)$	$7,5403 \times 10^{-2}$	0,5172	0,5
$y(k-2)^2$	$1,2259 \times 10^{-1}$	-0,0489	-0,05
$u(k-2)$	$5,8649 \times 10^{-2}$	0,7638	0,8
1	$8,4226 \times 10^{-3}$	0,4957	0,5
$u(k-2)^3$	$8,4194 \times 10^{-5}$	0,0110	0
$y(k-2)u(k-1)^2$	$9,6740 \times 10^{-5}$	-0,0109	0
$y(k-2)u(k-1)u(k-2)$	$4,3808 \times 10^{-5}$	0,0068	0
$u(k-2)u(k-3)$	$7,0585 \times 10^{-5}$	-0,0530	0
$y(k-1)^2$	$3,9710 \times 10^{-5}$	-0,0037	0
$y(k-1)u(k-2)u(k-3)$	$7,9171 \times 10^{-5}$	0,0074	0

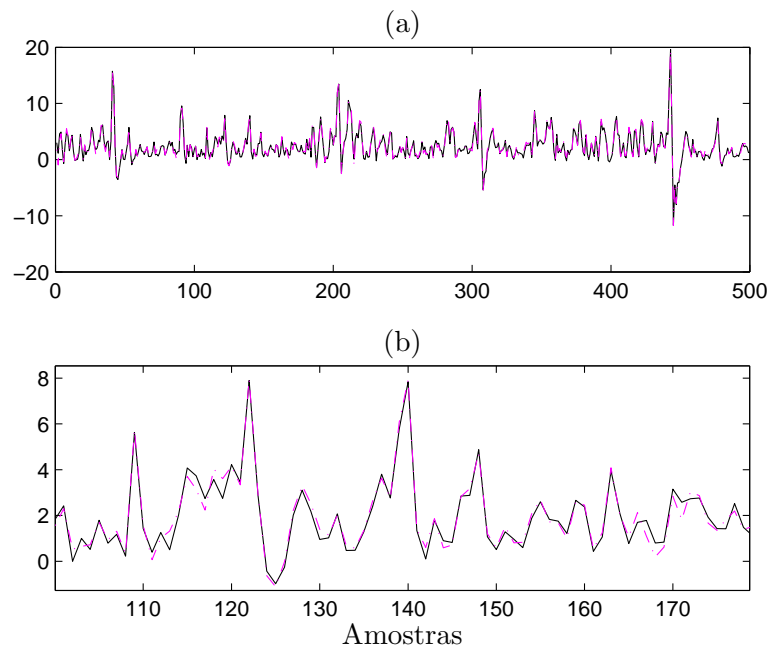


Figura 4.19: Validação dinâmica do modelo 4.3.

(a) Validação dinâmica de modelo. (—) dados reais e (---) simulação livre do modelo (4.3) para uma janela de dados subsequente à janela dos dados usados para identificação, RMSE = 0,1434. (b) Figura em escala ampliada.

Tabela 4.4: Modelo obtido pelo critério de informação de Akaike e ERR a partir dos dados apresentados na Figura 4.15. Conjunto inicial de 15 termos candidatos. Termos genuínos em negrito.

Regressor	ERR	Parâmetro estimado	Parâmetro real
$u(k-1)^2$	$7,2796 \times 10^{-1}$	0,9941	1
$y(k-1)$	$7,5402 \times 10^{-2}$	0,5221	0,5
$y(k-2)^2$	$1,2257 \times 10^{-1}$	-0,0481	-0,05
$u(k-2)$	$5,8648 \times 10^{-2}$	0,8178	0,8
1	$8,4226 \times 10^{-3}$	0,5122	0,5
$y(k-1)^2$	$5,7324 \times 10^{-5}$	-0,0187	0
$y(k-1)^2$	$6,5973 \times 10^{-5}$	-0,0015	0

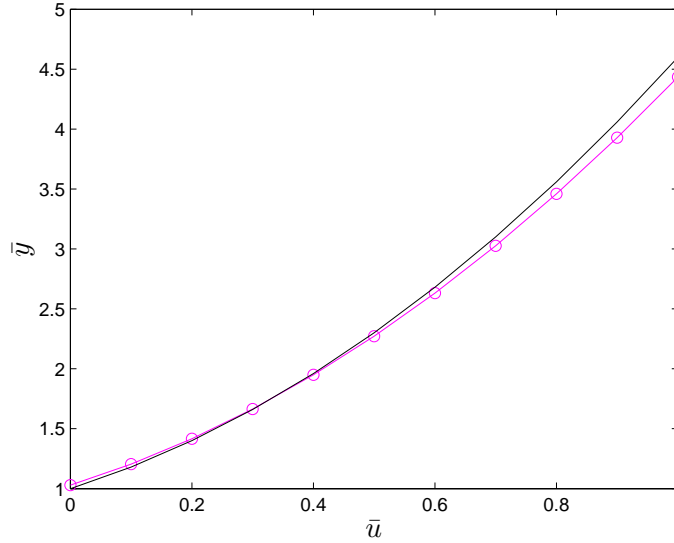


Figura 4.20: Validação estática do modelo 4.3.

Característica estática (—) dados reais, e (---) dados simulados, obtidos a partir do modelo (4.3). RMSE = 0,0668.

métrica permite uma comparação preliminar com outras metodologias presentes na literatura⁵.

A Equação 4.5 é uma aproximação do número de parâmetros estimados por meio do método bi-objetivo para seleção de estrutura.

$$Q\hat{\theta}_{OM} = (n_{pp} \times n_{t_1} \times n_{p_1}) + (n_{pp} \times n_{t_2} \times n_{p_2}), \quad (4.5)$$

sendo que a primeira parcela da soma representa a primeira etapa de seleção (Rotina I) e a segunda parcela a segunda etapa (Rotina II). Sendo n_{pp} o número de pontos considerados para formar a curva Pareto, n_{t_1} e n_{t_2} o número de termos candidatos na primeira e segunda etapa, respectivamente. n_{p_1} e n_{p_2} o número de curvas Pareto analisadas em ambas etapas.

O método proposto recentemente por Bonin et al. (2010), nomeado de SEM-LASSO, não quantifica claramente a complexidade computacional. A Equação 4.6 é uma tentativa de mensurar o custo do método e poder compará-

⁵Essa comparação só pode ser feita se houver uma semelhança entre os dois procedimentos avaliados.

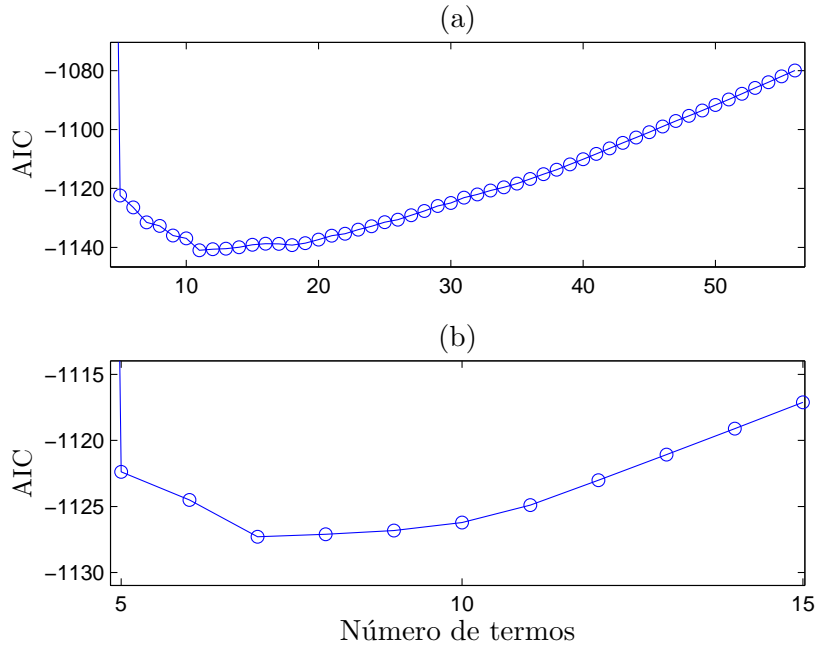


Figura 4.21: Resultado gráfico da aplicação do critério de informação de Akaike.

Resultado gráfico da aplicação do critério de informação de Akaike para o conjunto de dados apresentados na Figura 4.15. (a) conjunto inicial de 56 termos candidatos e (b) conjunto inicial de 15 termos candidatos. Figura em escala ampliada.

lo com o estudo realizado neste trabalho.

$$Q\hat{\theta}_{SM} = (\lambda \times n_{t_1}) + (n_{t_2} \times n_m), \quad (4.6)$$

em que λ é um parâmetro de ponderação, n_{t_1} e n_{t_2} são os números de termos candidatos na primeira etapa (LASSO) e segunda etapa (SEM) do processo e n_m o número de modelos avaliados.

Em ambos os procedimentos o custo computacional do método cresce exponencialmente com o número de termos candidatos. Para o presente exemplo, tem-se aproximadamente:

$$Q\hat{\theta}_{OM} = (15 \times 56 \times 11) + (15 \times 14 \times 16) = 12.600;$$

$$Q\hat{\theta}_{SM} = (500 \times 56) + (27 \times 168) = 32.536.$$

O tempo computacional gasto⁶ pelo método proposto, para resolver o problema de seleção de estrutura, foi de aproximadamente 24,3 segundos e para o SEM-LASSO de aproximadamente 250 segundos⁷. À medida que o grau de não-linearidade e máximos atrasos aumentam, o número de parâmetros candidatos aumentam drasticamente em ambos os métodos. Contudo, o método proposto mostrou-se menos custoso do ponto de vista computacional. Por outro lado, para utilizar o método SEM-LASSO não é necessário um segundo funcional, ou seja, não precisa dos dados estáticos.

4.3 Conclusões do capítulo

Esse capítulo discutiu e avaliou as características do critério bi-objetivo para a seleção da estrutura de modelos NARX polinomiais. A análise gráfica das curvas Pareto foi ilustrada, mostrando sua importância, principalmente quando há um cruzamento entre as curvas.

A metodologia apresentada no Capítulo 3 foi aplicada a três exemplos simulados. No primeiro exemplo ficou evidente o comportamento das curvas Pareto a partir de dois funcionais: erro estático e erro dinâmico. Conjectura-se que curvas Pareto construídas a partir de outros funcionais, conflitantes entre si, resultariam em gráficos similares.

O segundo exemplo mostrou que por meio do critério bi-objetivo é possível discriminar estruturas subparametrizadas e que, dado um modelo sobreparametrizado que contenha a estrutura nominal, é possível utilizando técnicas bi-objetivo e de sincronismo discriminar entre termos espúrios e genuínos. O mesmo exemplo mostrou ainda a possibilidade da existência de uma *região de incerteza*, ou seja, uma região na qual estruturas sobreparametrizadas, a princípio, seriam equivalentes entre si.

No terceiro caso, estudou-se um sistema com ruído na saída e com sinal

⁶Simulação realizada em um computador portátil, processador de 2GHz e memória RAM de 3Gb.

⁷O cálculo do tempo computacional realizado pelo algoritmo SEM-LASSO é uma aproximação tomando como base o número de iterações necessárias.

de entrada relativamente lento. O procedimento de identificação bi-objetivo mostrou-se eficiente na escolha dos termos genuínos, mesmo em condições não ideais de identificação. Contudo, ficou evidente a necessidade de ter, no meio do conjunto inicial de termos candidatos, os regressores verdadeiros. A solução para esse caso é trabalhar com um conjunto relativamente grande de termos candidatos, uma vez que o método prevê uma filtragem dos agrupamentos espúrios na etapa inicial. Entretanto, essa opção aumenta o tempo computacional do método e deve ser avaliada com cuidado.

Os resultados simulados mostram que o procedimento bi-objetivo parece ser uma alternativa adequada. No Capítulo 5 dois exemplos experimentais são analisados.

Resultados Experimentais

“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 ¹

5.1 Introdução

Esse capítulo descreve a aplicação do procedimento bi-objetivo para seleção de estruturas em dois sistemas experimentais: um conversor CC–CC do tipo buck e uma válvula pneumática. O objetivo desses dois exemplos é aplicar e analisar o desempenho do método em sistemas nos quais a estrutura correta é desconhecida.

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

5.2 Conversor Buck

O presente exemplo descreve a seleção da estrutura de um conversor estático CC–CC do tipo buck, como mostrado na Figura 5.1. O mesmo sistema foi modelado em outros trabalhos (Aguirre et al., 2000; Corrêa et al., 2002; Barroso

¹(Pepper, 1959, p. 32)

e Nepomuceno, 2004; Corrêa e Aguirre, 2004; Martins et al., 2009) e com isso é possível fazer uma comparação mais abrangente.

Esse exemplo é muito importante pois inclui uma série de dificuldades que tornam o problema de identificação ainda mais complexo (Barroso, 2006). Como os dados dinâmicos são limitados a uma faixa estreita de operação, a dinâmica do sistema não é completamente excitada e, com isso, técnicas convencionas de detecção de estrutura podem ser ineficientes.

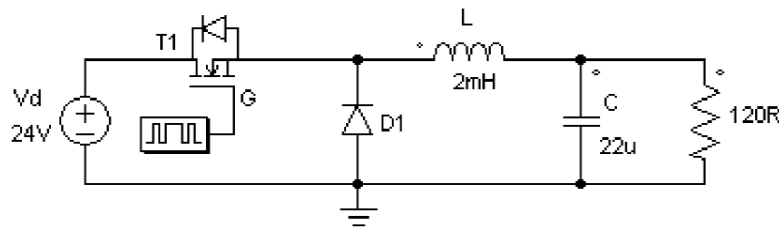


Figura 5.1: Conversor buck.

Estrutura de um conversor CC-CC do tipo buck.

5.2.1 Descrição do sistema

Os dados dinâmicos de identificação (Figura 5.2) e validação (Figura 5.3) utilizados são os mesmos considerados em Aguirre et al. (2000). Atuando-se na porta (G) do MOSFET (IRF840), obtém-se uma tensão contínua de saída, V_o a partir de uma fonte de tensão contínua, V_d . Durante todo o teste a tensão da fonte é mantida constante em 24V.

O ciclo de trabalho é definido pela proporção do tempo em que a chave está ligada em relação ao período total, $D = T_{on}/T$. Para variar D , utilizaram-se técnicas de modulação por largura de pulso, PWM (*pulse width modulation*) a uma taxa de 33kHz, utilizando para tal um circuito integrado LM3524. A taxa de 33KHz resultou em um modo de operação contínuo, ou seja, a corrente através do indutor era sempre maior que zero.

Quando o ciclo de trabalho tende à unidade, a corrente através do indutor e V_o aumentam, pois a fonte V_d energiza a malha formada por ela, o capacitor e o indutor. Quando D tende a zero, V_o diminui com um regime dinâmico diferente. Esse fato caracteriza um regime dinâmico não-linear do sistema.

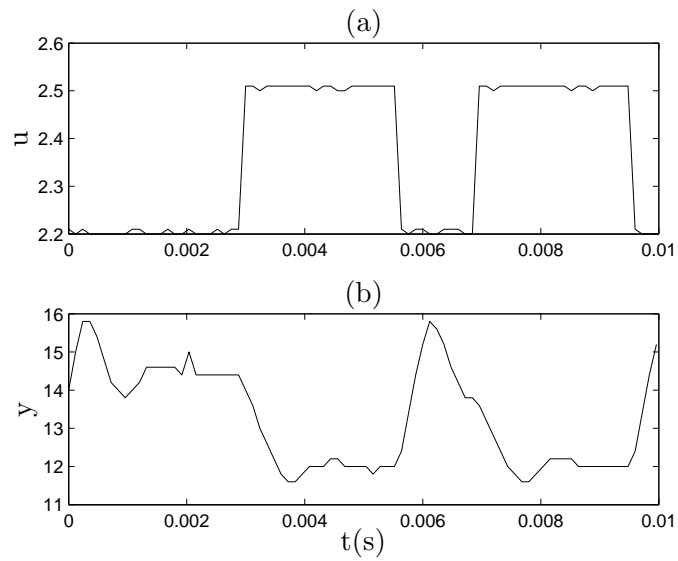


Figura 5.2: Dados de identificação – conversor buck.

Dados de identificação do conversor buck, sendo: (a) sinal de entrada do sistema, (b) sinal de saída do sistema, (buck@).

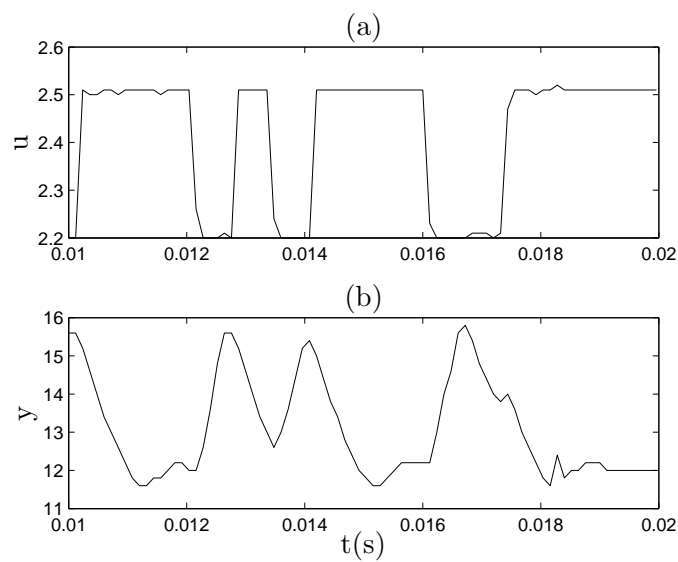


Figura 5.3: Dados de validação – conversor buck.

Dados de validação do conversor buck, sendo: (a) sinal de entrada do sistema, (b) sinal de saída do sistema, (buck@).

5.2.2 Característica estática

Em estado estacionário o conversor implementado possui a seguinte relação teórica entre o ciclo de trabalho e a tensão de entrada:

$$\begin{aligned}
 V_o &= DV_d \\
 &= (1 - D')V_d = \left(1 - \frac{\bar{u} - 1}{3}\right) V_d \\
 &= \frac{4V_d}{3} - \frac{V_d}{3}\bar{u},
 \end{aligned} \tag{5.1}$$

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

Os dados estáticos podem ser obtidos aplicando-se à Equação 5.1 valores de \bar{u} que abranjam todos os pontos de operação do sistema (1 a 4V). A curva estática para o conversor buck pode ser vista na Figura 5.4. Os dados dinâmicos são limitados a uma estreita faixa em comparação aos possíveis pontos de operação, como destacado na característica estática.

5.2.3 Identificação de modelos

A curva estática do sistema (Figura 5.4) obtida de forma analítica pode ser escrita de forma geral como um mapeamento não-linear do tipo $f(x) = ax + b$. Esse tipo de relação estática pode ser conseguida em modelos NARX polinomial por meio de uma estrutura com grau de não-linearidade unitário acrescido do termo constante.

Para determinar o conjunto inicial de termos candidatos considerou-se o máximo atraso no sinal de entrada e no sinal de saída igual a 7. Conforme Equação 2.14 (ver Seção 2.3), obteve-se um total de 15 termos candidatos, distribuídos em 3 agrupamentos, listados a seguir.

$$\begin{aligned}
 \Omega_0 &= [1]; \\
 \Omega_y &= [y(k-1) \ y(k-2) \ y(k-3) \ y(k-4) \ y(k-5) \ y(k-6) \ y(k-7)]; \\
 \Omega_u &= [u(k-1) \ u(k-2) \ u(k-3) \ u(k-4) \ u(k-5) \ u(k-6) \ u(k-7)].
 \end{aligned}$$

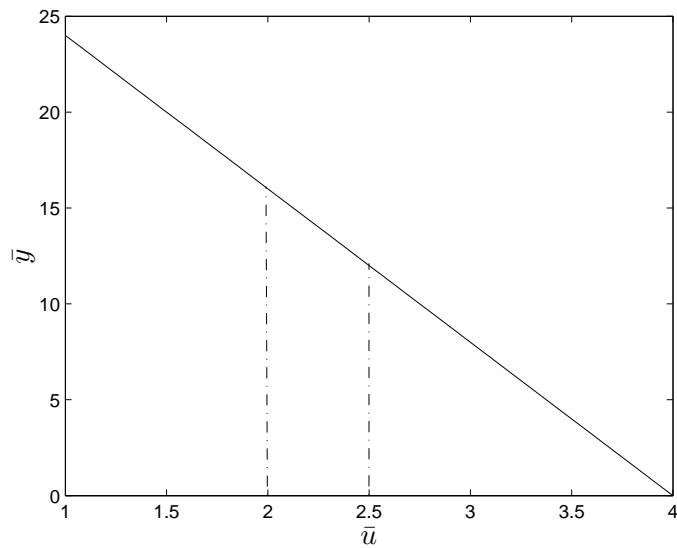


Figura 5.4: Dados estáticos: conversor buck

Curva estática do conversor buck, sendo os dados estáticos de entrada (\bar{u}) e de saída (\bar{y}) em volts. Os dados dinâmicos excursionam apenas a região delimitada pelas linhas pontilhadas, (`buckest@`).

De acordo com a metodologia proposta deve-se eliminar os agrupamentos espúrios. Contudo, como o conjunto inicial de regressores é formado apenas por 3 agrupamentos, não foi necessário executar a primeira etapa do algoritmo (ver Seção 3.3.1). Portanto, busca-se selecionar os regressores verdadeiros. Para isso é necessário combinar os regressores a fim de formar as estruturas-candidatas.

A partir da estrutura de 15 termos, em sequência, elimina-se um dos termos e reestimam-se os parâmetros dos 14 termos restantes. Para cada modelo, gera-se a curva Pareto. Cada curva Pareto corresponde a uma estrutura-candidata de 14 termos mais a estrutura composta pelos 15 termos, totalizando 16 curvas. As curvas traçadas são ilustradas na Figura 5.5.

A Figura 5.5a mostra as 16 curvas. Três grupos são facilmente identificáveis. O grupo (b), ampliado na Figura 5.5b, corresponde a uma curva Pareto, no qual falta o termo *constante*, que é importante. O grupo (c), ampliado na

Figura 5.5c, representa duas curvas Pareto nas quais faltam os termos $y(k-1)$ e $u(k-1)$. E o grupo (d), ampliado na Figura 5.5d, corresponde aos agrupamentos nos quais os termos retirados não influenciaram tanto o desempenho do modelo. A menos por duas curvas, que apresentaram um erro estático maior, conforme destacado. Esses termos são $y(k-7)$ e $u(k-7)$.

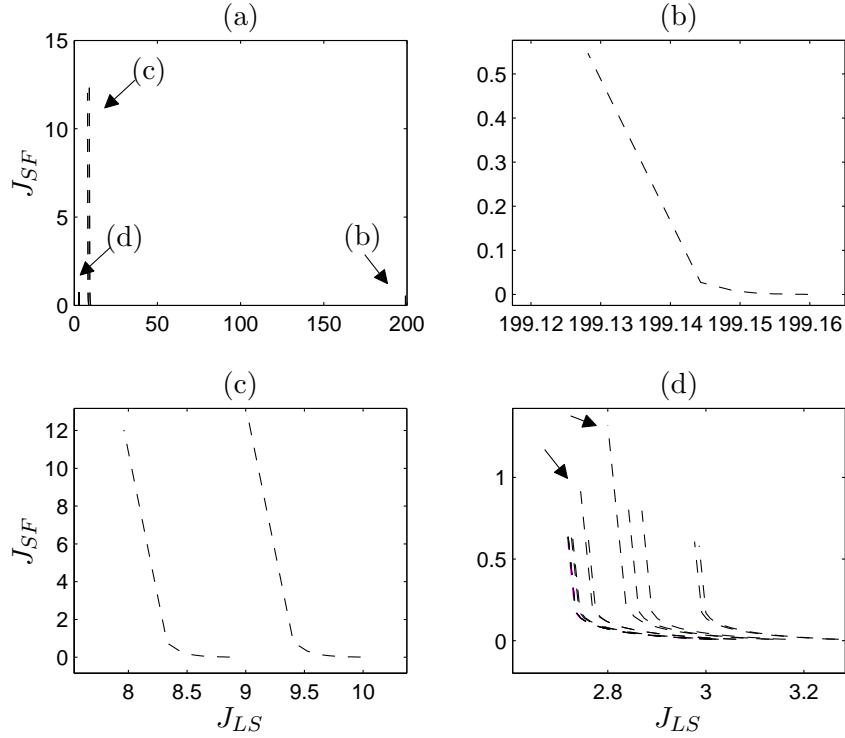


Figura 5.5: Curvas Pareto – conversor buck.

Curvas Pareto para identificação do conversor buck. (a) Conjuntos Pareto para cada uma das 16 estruturas avaliadas, (b-d) curvas Pareto ampliadas conforme indicado em (a).

Após identificar os termos $[y(k-1) \ u(k-1) \ y(k-7) \ u(k-7) \ 1]$, realizou-se o procedimento de estimação bi-objetivo (ver Seção 2.4) chegando ao modelo

$$y(k) = 0,6307y(k-1) - 0,0654y(k-7) - 3,7745u(k-1) + 0,2951u(k-7) + 13,9331. \quad (5.2)$$

A simulação livre do modelo (5.2) para os dados de validação é mostrada na Figura 5.6. Pode-se perceber que há alguns desajustes nas extremidades de algumas oscilações.

Seguindo-se o procedimento descrito na Seção 2.2.2, pode-se fazer a análise do modelo (5.2) em estado estacionário. Procedendo assim, chega-se à seguinte equação:

$$\bar{y} = \frac{13,9331 - 3,4794\bar{u}}{1 - 0,5652}, \quad (5.3)$$

que é mostrada na Figura 5.7 juntamente com a característica estática analítica do conversor. Como pode ser visto, o ajuste é bem aproximado.

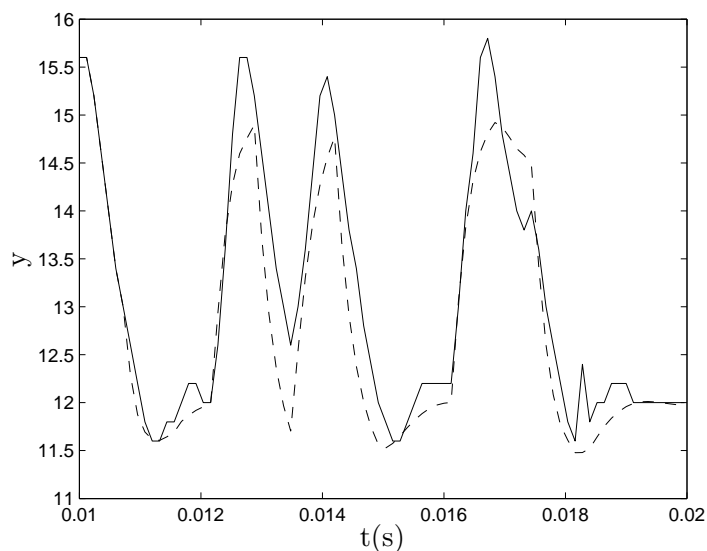


Figura 5.6: Validação dinâmica do modelo (5.2) – conversor buck.

Validação dinâmica do modelo do conversor buck sendo: (—) os dados reais e (---) simulação livre do modelo.

Como esse exemplo foi modelado em outros trabalhos é possível compará-lo. Em todos os trabalhos a métrica utilizada foi o índice RMSE (ver Seção 2.6). A Tabela 5.1 apresenta uma comparação dos resultados.

O modelo (5.2) tem um RMSE dinâmico próximo em relação ao melhor modelo, com uma expressiva melhora estática e um número reduzido de pa-

râmetros. Ou seja, a partir de um modelo simples (grau de não-linearidade unitário) foi possível aproximar a curva estática sem deteriorar muito a característica dinâmica.

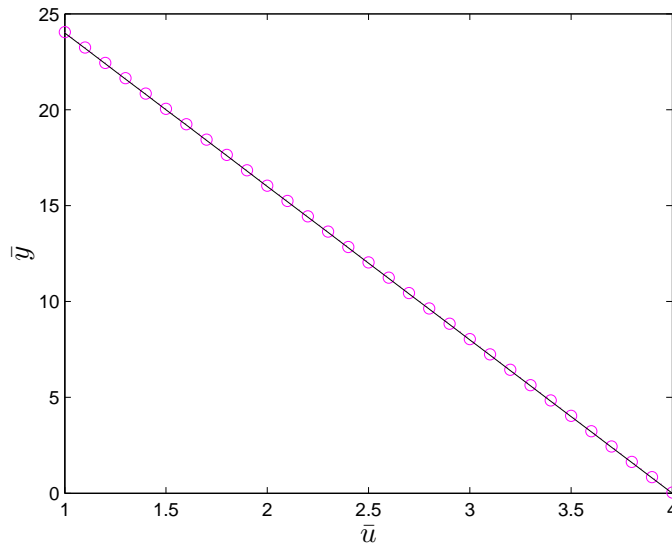


Figura 5.7: Validação estática do modelo (5.2) – conversor buck.

Validação estática do modelo do conversor buck sendo: (—) os dados reais e (o) a resposta do modelo.

Tabela 5.1: Índice RMSE do modelo (5.2) e número de parâmetros (n_p) comparado com outros modelos.

Modelo	RMSE Dinâmica	RMSE Estática	n_p
modelo (5.2)	0,3706	0,0056	5
Barroso (2006)	0,3691	0,0122	6
Barroso (2001)	0,6750	0,0028	7
Aguirre et al. (2000)	0,4906	0,9750	8
Corrêa (2001)	0,4851	0,2746	8

5.3 Válvula pneumática

O objetivo no presente exemplo é a identificação de modelos para uma válvula pneumática. Além de dinamicamente válidos, deseja-se que os modelos sejam também estaticamente válidos.

Diferente do exemplo anterior, a válvula estudada não possuía um modelo previamente identificado. Portanto, o estudo mostrará todas as etapas de identificação, e não só o procedimento de seleção de estrutura bi-objetivo.

5.3.1 Descrição do sistema



Figura 5.8: Sistema de controle de nível.

Sistema de controle de nível *Level Process Station 3503MO* (Lab-Volt, 2006). Destaque para a válvula pneumática, retângulo tracejado.

Ilustrado na Figura 5.8, o processo geral consiste em um sistema de bombeamento de água e controle de nível. Desenvolvido pela empresa canadense, Lab-Volt[®], a planta é composta por um reservatório que, por meio de uma bomba elétrica, permite o envio de água para um tanque cilíndrico de diâmetro uniforme. A vazão é manipulada por uma válvula de atuação pneumática. A Figura 5.9 mostra o diagrama funcional da planta, com destaque para a válvula pneumática. Detalhes de ligação e componentes da planta são descritos em Lab-Volt (2006).

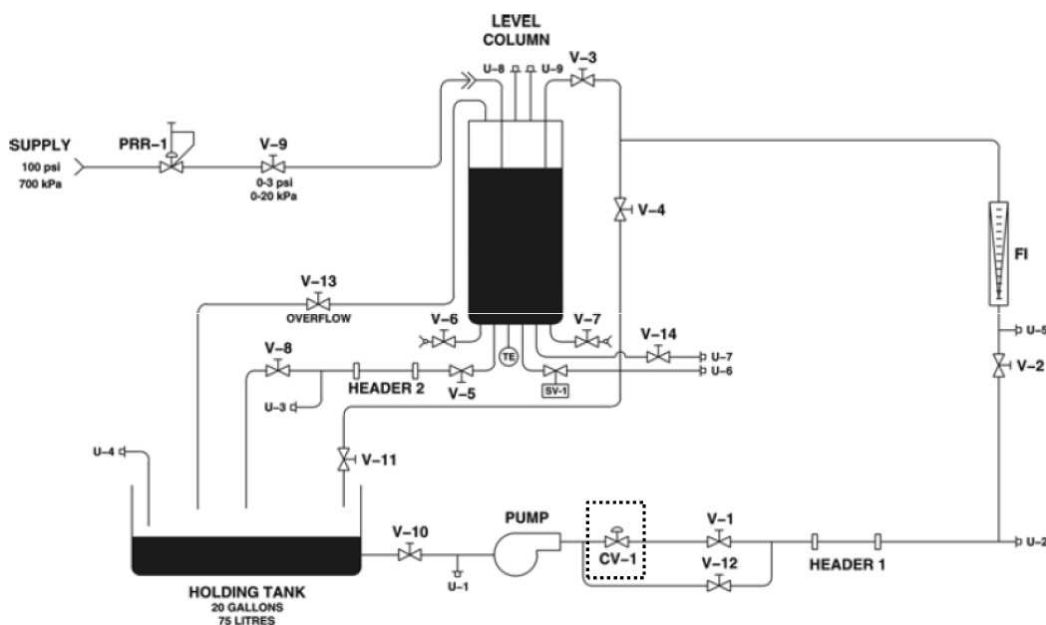


Figura 5.9: Diagrama funcional do sistema de controle de nível.

Diagrama funcional do sistema de controle de nível (Lab-Volt, 2006). Destaque para a válvula pneumática, retângulo tracejado.

A válvula pneumática foi o sistema em estudo, em que o sinal de entrada $u(k)$ é a pressão² que comanda sua abertura e fechamento. A saída $y(k)$ é a posição da válvula, e é medida usando-se um sensor resistivo que registra o deslocamento da válvula.

²A entrada é uma tensão, que após ser convertida em corrente, é convertida em pressão. Note que as características dos conversores V/I e I/P estão embutidos nos dados (dinâmicos e estáticos).

O procedimento de atuação e aquisição é ilustrado na Figura 5.10. Gera-se computacionalmente um sinal, e, por meio de uma placa de aquisição/atuação aplica-o em um transdutor de corrente-pressão. A entrada, em pressão, atua na válvula (processo). O sinal de deslocamento da válvula, após ser convertido em tensão, é armazenado. O sistema de gerenciamento dos sinais de entrada e saída, ilustrado na Figura 5.11, foi desenvolvido no Labview[®].

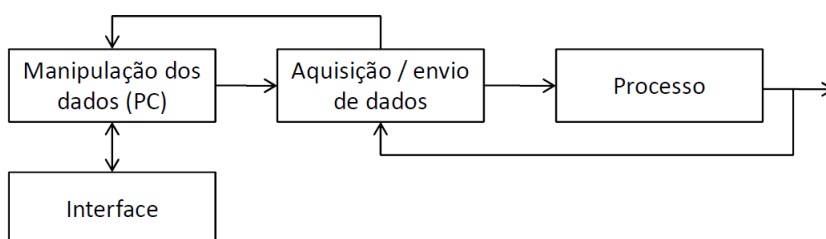


Figura 5.10: Diagrama em blocos do sistema de aquisição e tratamento dos dados.

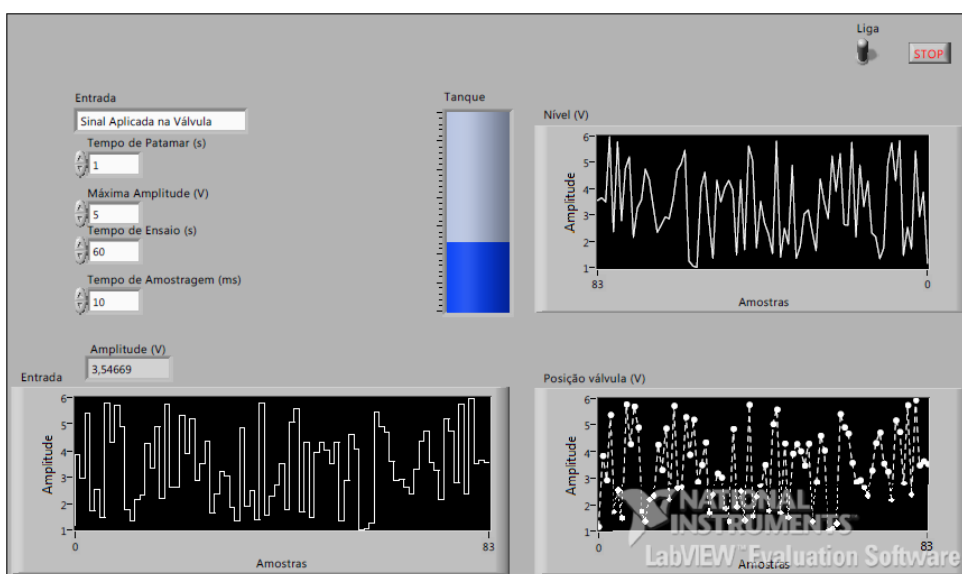


Figura 5.11: Tela principal do sistema supervisorio.

Sistema supervisorio onde é verificado a entrada aplicada e monitorado os sinais de saída (posição da válvula e nível do tanque). Apesar de constantemente monitorado, os dados do nível do tanque não foram utilizados nesse trabalho.

5.3.2 Descrição de testes e dados

Dois tipos diferentes de testes foram efetuados no sistema: testes dinâmicos e testes estáticos. Nos testes dinâmicos, projetou-se um sinal aleatório e aplicou-se tal sinal na entrada. Durante o teste, cada valor aleatório do sinal de entrada era mantido constante por 1,2 segundos, compatível com a menor constante de tempo de interesse. A taxa de amostragem escolhida foi de 0,006 segundos, para a coleta de dados durante os testes. Cada teste dinâmico durou aproximadamente 80 segundos. Os sinais de identificação e validação foram superamostrados, com isso fez-se necessário decimar os sinais de um fator $\Delta = 95$, para gerar os dados de trabalho (ver (Aguirre, 2007, p. 463)). A Figura 5.12 mostra um série temporal com os sinais de entrada e saída.

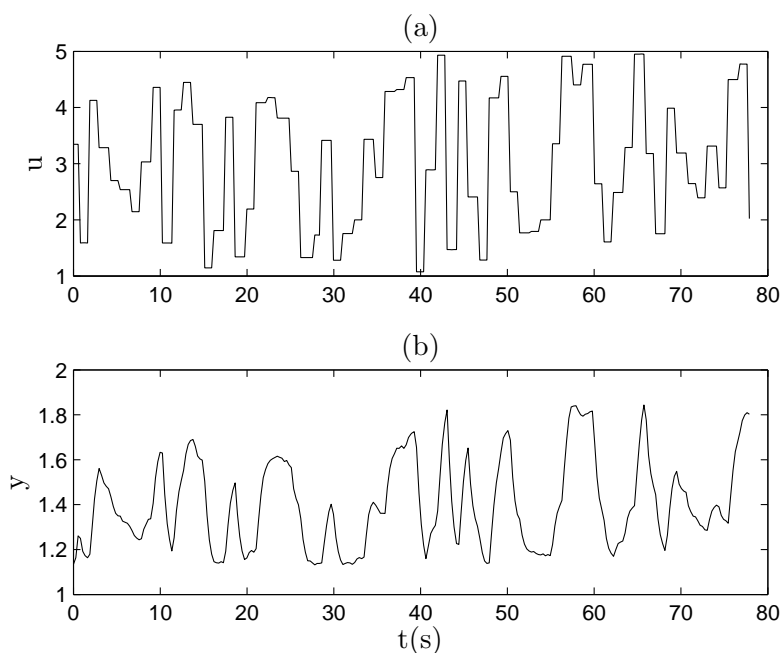


Figura 5.12: Entrada e saída da válvula.

Entrada e saída da válvula. (a) dados de entrada em volts.
e (b) dados de saída em volts, (dinamica15@).

Nos testes estáticos, para cada valor do sinal de entrada esperava-se a estabilização da posição da válvula. Esse conjunto de dois valores estacionários podem ser representados como um ponto no gráfico. Tal gráfico representa a

característica estática, ou seja, a característica do sistema em estado estacionário. A Figura 5.13 mostra a relação estática medida da válvula.

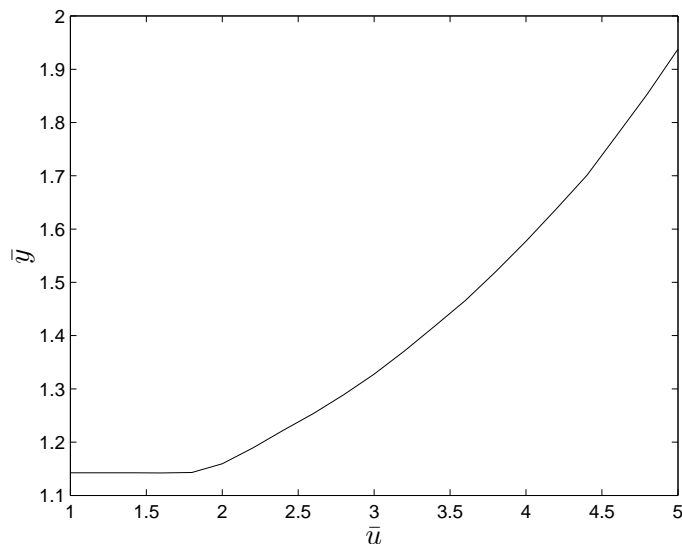


Figura 5.13: Dados estáticos da válvula

Curva estática da válvula, sendo: \bar{u} e \bar{y} dados estáticos de entrada e saída em volts, (`estatica2@`).

5.3.3 Identificação de modelos

Antes de aplicar o procedimento bi-objetivo, verificou-se o desempenho dos critérios de informação de Akaike e ERR. Trabalhou-se com um conjunto inicial de 56 termos candidatos (grau de não-linearidade igual a 3, máximo atraso no sinal de entrada igual a 2 e máximo atraso no sinal de saída igual a 3). O AIC sugeriu a utilização de 12 termos, que ordenados segundo o ERR são: $y(k-1)$, $u(k-1)$, 1 , $y(k-1)u(k-1)^2$, $y(k-2)^2u(k-3)$, $y(k-2)u(k-1)^2$, $u(k-3)^2$, $y(k-1)^2y(k-2)$, $y(k-1)u(k-3)^2$, $y(k-2)u(k-1)$, $u(k-1)^2$, $u(k-3)$.

O modelo final, estimado utilizando o estimador bi-objetivo é dado por:

$$\begin{aligned}
y(k) = & -0,3026y(k-1) + 0,0174u(k-1) + 1,1031 + 0,0180y(k-1)u(k-1)^2 \\
& - 0,0190y(k-2)u(k-1)^2 + 0,0409u(k-3)^2 - 0,0202y(k-1)u(k-3)^2 \\
& + 0,0014y(k-2)u(k-1) + 0,0099u(k-1)^2 - 0,0090u(k-3) \\
& - 0,0363y(k-2)^2u(k-3) + 0,2519y(k-1)^2y(k-2). \tag{5.4}
\end{aligned}$$

A Figura 5.14 mostra a simulação livre do modelo (5.4). O desempenho estático é ilustrado na Figura 5.15. Uma análise quantitativa é dada pelo índice RMSE, apresentado na legenda da figura correspondente.

A fim de verificar a possibilidade de melhorar a característica estática, forçou-se o estimador bi-objetivo a dar prioridade máxima à minimização do funcional estático (J_{SF}). As Figuras 5.16 e 5.17 mostram, respectivamente, os desempenhos estático e dinâmico do modelo. Por meio do índice RMSE foi possível verificar uma melhora tanto do erro estático quanto do erro dinâmico.

Com o intuito de fomentar uma comparação e avaliar possíveis melhorias no modelo, aplicou-se o critério bi-objetivo para seleção de estrutura, como detalhado a seguir.

Aplicação do critério bi-objetivo

Como descrito no procedimento bi-objetivo (Capítulo 3), deve-se determinar um conjunto inicial de termos candidatos. De acordo com a característica estática observada, considerou-se um modelo de grau de não-linearidade $l = 2$. Os máximos atrasos escolhidos foram $n_y = n_u = 3$. Conforme Equação 2.14, obteve-se 28 termos candidatos. A esse conjunto de termos, acrescentou-se ainda 10 termos pertencentes ao agrupamento Ω_{u^3} , resultando no conjunto inicial de 38 termos candidatos, distribuídos em 7 agrupamentos, listados a seguir.

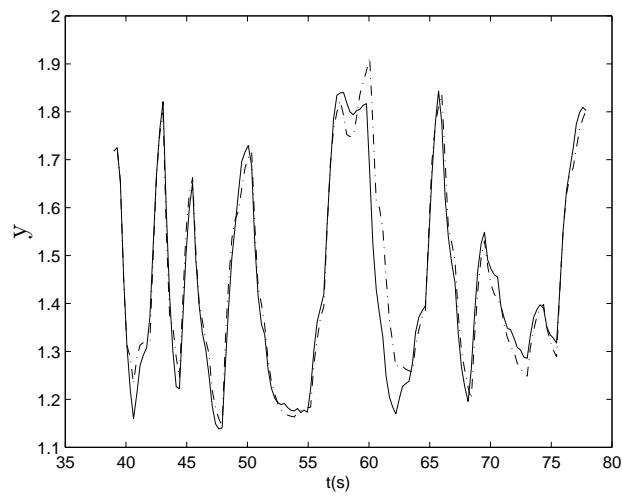


Figura 5.14: Validação dinâmica do modelo (5.4) da válvula pneumática.

Validação dinâmica, sendo: (—) dados medidos e (---) simulação livre do modelo (5.4). RMSE = 0,2796.

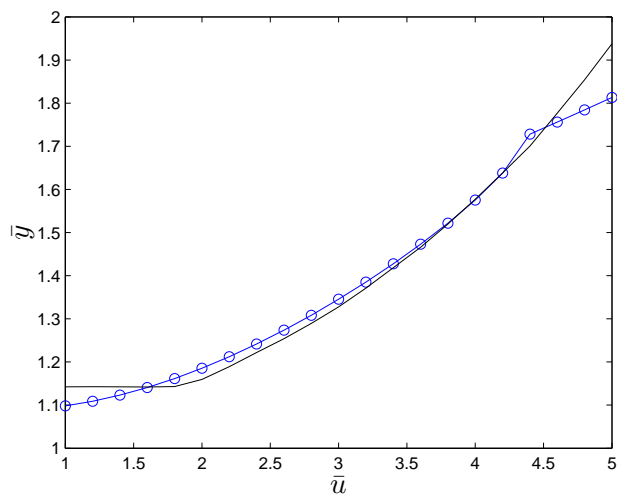


Figura 5.15: Validação estática do modelo (5.4) da válvula pneumática.

Característica estática medida (—), e (o) simulada, obtida a partir do modelo (5.4). RMSE = 0,1245.

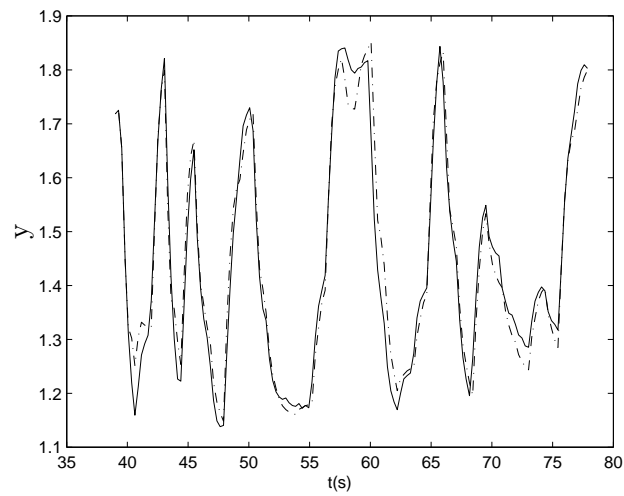


Figura 5.16: Validação dinâmica do modelo (5.4) da válvula pneumática. Prioridade para J_{SF} .

Validação dinâmica, sendo: (—) dados medidos e (---) simulação livre do modelo (5.4). RMSE = 0,2054.

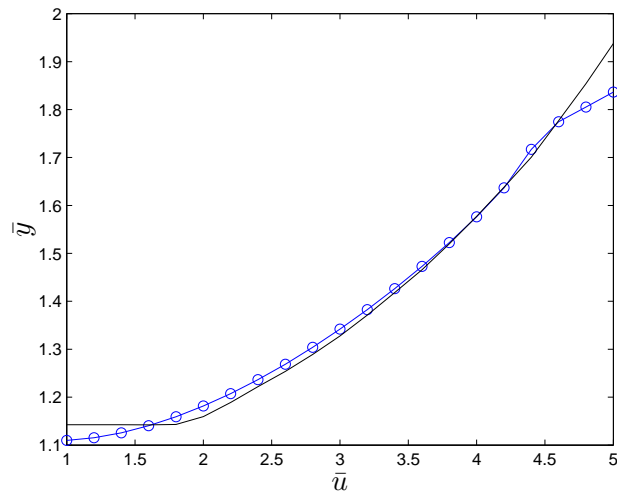


Figura 5.17: Validação estática do modelo (5.4) da válvula pneumática. Prioridade para J_{SF} .

Característica estática medida (—), e (o) simulada, obtida a partir do modelo (5.4). RMSE = 0,0967.

$$\begin{aligned}
\Omega_0 &= [1]; \\
\Omega_y &= [y(k-1) \quad y(k-2) \quad y(k-3)]; \\
\Omega_u &= [u(k-1) \quad u(k-2) \quad u(k-3)]; \\
\Omega_{y^2} &= [y(k-1)y(k-1) \quad y(k-1)y(k-2) \quad y(k-2)y(k-2) \quad y(k-1)y(k-3) \\
&\quad y(k-2)y(k-3) \quad y(k-3)y(k-3)]; \\
\Omega_{u^2} &= [u(k-1)u(k-1) \quad u(k-1)u(k-2) \quad u(k-1)u(k-3) \quad u(k-2)u(k-2) \\
&\quad u(k-2)u(k-3) \quad u(k-3)u(k-3)]; \\
\Omega_{yu} &= [y(k-1)u(k-1) \quad y(k-1)u(k-2) \quad y(k-1)u(k-3) \quad y(k-2)u(k-1) \\
&\quad y(k-2)u(k-2) \quad y(k-2)u(k-3) \quad y(k-3)u(k-1) \quad y(k-3)u(k-2) \\
&\quad y(k-3)u(k-3)]; \\
\Omega_{u^3} &= [u(k-1)u(k-1)u(k-1) \quad u(k-1)u(k-1)u(k-2) \quad u(k-1)u(k-1)u(k-3) \\
&\quad u(k-1)u(k-2)u(k-2) \quad u(k-1)u(k-2)u(k-3) \quad u(k-1)u(k-3)u(k-3) \\
&\quad u(k-2)u(k-2)u(k-2) \quad u(k-2)u(k-2)u(k-3) \quad u(k-2)u(k-3)u(k-3) \\
&\quad u(k-3)u(k-3)u(k-3)].
\end{aligned}$$

Organizou-se as estruturas-candidatas para executar a próxima etapa: gerar as curvas Pareto e verificar a possibilidade de eliminar os agrupamentos espúrios. A partir da estrutura de 7 agrupamentos, em sequência, eliminou-se um dos agrupamentos, realizou-se o procedimento de estimação bi-objetivo e traçou-se a curva Pareto.

Conforme procedimento apresentado no Capítulo 3 e ilustrado na Seção 3.4, apenas o agrupamento Ω_{yu} foi eliminado. Isso implica reduzir o conjunto inicial de possíveis termos em 23,7%, ou seja, a “melhor” estrutura deve ser escolhida entre 29 termos.

Conforme exemplo anterior, os 29 termos candidatos foram combinados gerando 30 estruturas-candidatas e, conseqüentemente, 30 curvas Pareto, ilustradas na Figura 5.18. As curvas destacadas correspondem às estruturas que apresentaram maior área. Estruturas nas quais faltam os termos: 1, $y(k-1)$, $y(k-3)$, $u(k-3)$, $u(k-3)^2$, $u(k-1)^3$ e $u(k-3)^3$.

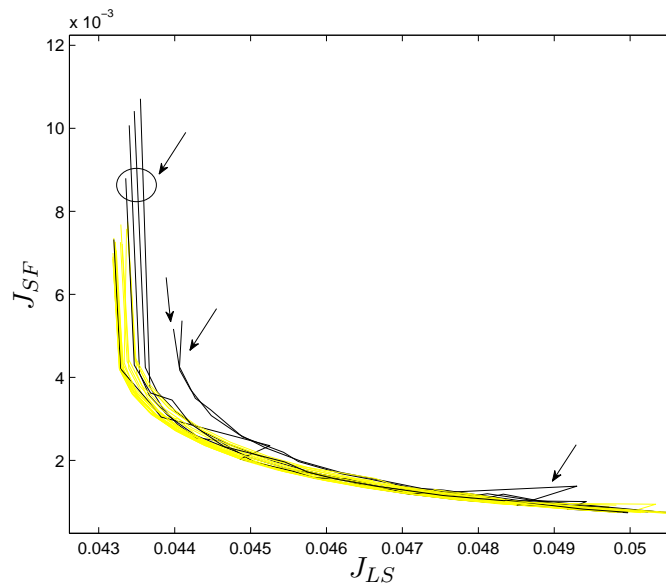


Figura 5.18: Curvas Pareto - válvula pneumática.

Curvas Pareto para identificação da válvula pneumática. Destaque para as curvas nas quais falta um regressor importante.

Identificados os “melhores” termos, realizou-se o procedimento de estimação bi-objetivo (ver Seção 2.4) e obteve-se o seguinte modelo:

$$y(k) = 0,9564 + 0,1324y(k-1) + 0,1482y(k-3) - 0,2019u(k-3) + 0,0823u(k-3)^2 + 0,0035u(k-1)^3 - 0,0086u(k-3)^3. \quad (5.5)$$

Seguindo-se o procedimento descrito na Seção 2.2.2, a função que descreve a característica estática do modelo (5.5) é

$$\bar{y} = \frac{0,9564 - 0,2019\bar{u} + 0,0823\bar{u}^2 - 0,0051\bar{u}^3}{0,7194}. \quad (5.6)$$

A simulação livre do modelo é mostrada na Figura 5.19. A Equação 5.6 é mostrada na Figura 5.20 junto com a característica estática medida. Essas duas

figuras mostram que o desempenho do modelo, selecionado pelo critério bi-objetivo, foi melhor do que o desempenho do modelo selecionado pelos critérios ERR e AIC. Ou seja, com 7 termos o modelo (5.5) ajustou a curva estática e dinâmica com erros menores. Essa análise é quantificada pelo índice RMSE, apresentado na legenda da figura correspondente.

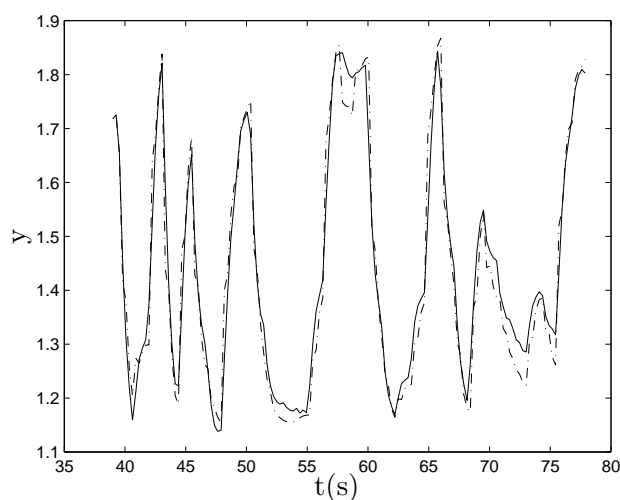


Figura 5.19: Validação dinâmica do modelo (5.5) da válvula pneumática.

Validação dinâmica, sendo: (—) dados medidos e (---) simulação livre do modelo (5.5). RMSE = 0,2238.

5.4 Conclusões do capítulo

A metodologia proposta no Capítulo 3 foi aplicada nesse capítulo para modelar dois processos experimentais: o conversor buck e a válvula pneumática. O conversor buck foi largamente estudado e possibilitou comparar os resultados encontrados. A abordagem empregada para o conversor propiciou resultados dinâmicos tão bons quanto os encontrados na literatura, apresentando a vantagem de um melhor ajuste da curva estática.

A identificação da válvula pneumática foi realizada no contexto desse trabalho. Preparou-se o sistema de aquisição, a interface de interação e realizou-se os testes para coleta dos dados. A partir dos dados coletados, aplicou-se o pro-

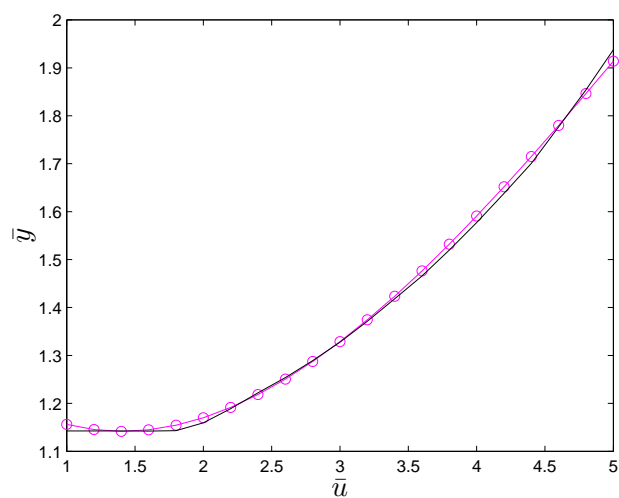


Figura 5.20: Validação estática do modelo (5.5) da válvula pneumática.

Característica estática medida (—), e (o) equação (5.6) obtida a partir do modelo (5.5). RMSE = 0,0386.

cedimento para identificação de modelos que representassem as características dinâmicas e estáticas da válvula.

De forma geral, os resultados obtidos indicam que o procedimento pode ser empregado como uma ferramenta auxiliar para seleção de estruturas de modelos NARX polinomiais.

Conclusão e Pesquisas Futuras

“A ideia de um método que contenha princípios científicos inalteráveis e absolutamente obrigatórios que rejam os assuntos científicos se defronta com dificuldades ao ser confrontada com os resultados da investigação histórica. Descobrimos que não existe uma única regra, por mais plausível que pareça, por mais alicerçada sobre a epistemologia, que não seja desrespeitada numa ou noutra ocasião. É evidente que tais transgressões não ocorrem acidentalmente (...) mas são antes necessárias ao progresso.”

Paul Feyerabend¹

6.1 Introdução

Este capítulo tem como objetivo apresentar as principais contribuições deste trabalho. Em seguida, faz-se um balanço dessas contribuições. Por fim, as perspectivas de pesquisas futuras são apresentadas, advindas do encontro de questões cuja investigação não foi conduzida neste trabalho.

¹(Feyerabend, 1970, p. 15).

6.2 Discussões finais

Neste trabalho investigou-se o problema da determinação de estruturas em modelos polinomiais NARX. Após apresentar uma revisão teórica e bibliográfica sobre o tema deste trabalho, foram mostrados um capítulo de desenvolvimento do método proposto e dois capítulos de resultados simulados e experimentais.

Os resultados simulados mostram que por meio do critério bi-objetivo é possível discriminar estruturas subparametrizadas e que, dado um modelo sobreparametrizado que contenha a estrutura nominal, é possível utilizando técnicas multiobjetivo e de sincronismo discriminar entre termos espúrios e genuínos. Foi possível constatar ainda a robustez do método mesmo em condições não ideais de identificação.

No Capítulo 5 dois exemplos experimentais foram realizados. A identificação do conversor buck mostrou que o critério proposto pode alcançar bons resultados. Foi obtido um modelo válido em toda faixa de operação, tanto de sua dinâmica quanto do seu regime permanente, ou seja, modelos que conseguiram localizar sua curva estática sem degradar sua representação dinâmica.

O segundo exemplo experimental identificou modelos para uma válvula pneumática. Esse último estudo de caso possibilitou executar todos os passos de um procedimento de identificação, além de permitir a avaliação da técnica desenvolvida, que mostrou-se como uma ferramenta auxiliar na determinação de estrutura de modelos NARX polinomiais.

6.3 Pesquisas futuras

Ao longo do trabalho realizado, algumas possibilidades de investigação foram levantadas, mas não investigadas. Entre elas, pode-se destacar:

- i) Verificar se o método proposto é válido para o caso de não haver, por exemplo, dados estáticos. É possível executar o mesmo algoritmo se o segundo objetivo for outro?
- ii) Como investigar situações com mais de dois objetivos disponíveis? É possível trabalhar com todos os funcionais?

-
- iii) Os exemplos simulados mostraram nos três casos que é possível obter o modelo correto quando há, dentre o conjunto de termos candidatos os termos verdadeiros. Se no conjunto de candidatos inicial não houver os termos genuínos, como ocorrerá a seleção de estrutura?
 - iv) No segundo exemplo do Capítulo 4, verificou-se a possibilidade da existência de uma região de incerteza, ou seja, uma região na qual estruturas sobreparametrizadas são equivalentes entre si. Se existe uma região de incerteza, acredita-se que seja possível utilizar esse conhecimento para síntese de controladores robustos. Um ponto que pode ser investigado.
 - v) Incorporar o estimador LASSO na Rotina II (ver Seção 3.3.2) do método bi-objetivo.
 - vi) Determinar uma forma sistemática de obter o conjunto inicial de termos candidatos e delimitar as situações em que o estimador bi-objetivo pode não responder adequadamente.

Referências Bibliográficas

- Aguirre, L. A. (2007). *Introdução à Identificação de Sistemas: técnicas lineares e não-lineares aplicadas a sistemas reais*. Editora da UFMG, Belo Horizonte, 3 edição.
- Aguirre, L. A., Barbosa, B. H. G., e Braga, A. P. (2010). Prediction and simulation errors in parameter estimation for nonlinear systems. *Mechanical Systems and Signal Processing*, 24(8):2855–2867.
- 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. (1995). Improved structure selection for nonlinear models based on term clustering. *International Journal of Control*, 62(3):569–587.
- Aguirre, L. A., Donoso-Garcia, P. F., e Santos-Filho, R. (2000). Use of a priori information in the identification of global nonlinear models-a case study using a buck converter. *IEEE J CASI*, 47(7):1081–1085.
- Aguirre, L. A., Furtado, E. C., e Tôrres, L. A. B. (2006). Evaluation of dynamical models: Dissipative synchronization and other techniques. *Physical Review E*, 74(019612):1–16.
- Aguirre, L. A., Rodrigues, G. G., e Jácome, C. R. F. (1998). Identificação de sistemas não-lineares utilizando modelos narmax polinomiais - uma revisão e novos resultados. *SBA Controle & Automação*, 9(2):90–106.

- Akaike, H. (1970). Statistical predictor identification. In *Annals of the Institute of Statistical Mathematics*, volume 22, páginas 203–217.
- Akaike, H. (1974). A new look at the statistical model identification. In *IEEE Transactions on Automatic Control*, volume 19, páginas 716–723.
- Alves, M. A. (2009). *Utilização de Auto-Consistência como Ferramenta Auxiliar na Seleção de Estrutura de Modelos NARX Polinomiais*. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- Alves, R. (2005). *Filosofia da Ciência – Introdução ao Jogo e suas Regras*. Editora Loyola, São Paulo, 9 edição.
- Barbosa, A. M., Takahashi, R. H. C., e Aguirre, L. A. (2010). Técnicas de otimização bi-objetivo aplicadas a problemas de determinação de estrutura em modelos polinomiais NARX. In *XVIII Congresso Brasileiro de Automática*, páginas 3189–3196, Bonito – MS.
- Barbosa, B. H. G. (2009). *Computação Evolucionária e Máquinas de Comitê na Identificação de Sistemas Não-Lineares*. Tese de Doutorado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- 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.
- Barroso, M. F. S. (2006). *Otimização Bi-Objetivo Aplicada à Estimação de Parâmetros e Modelos Não-Lineares: Caracterização e Tomada de Decisão*. Tese de Doutorado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- Barroso, M. F. S. e Nepomuceno, E. G. (2004). Multiobjective identification of a buck dc-dc converter. In *XV Congresso Brasileiro de Automática*, páginas 1–6, Gramado – Rs.

- Barroso, M. F. S., Takahashi, R. H. C., e Aguirre, L. A. (2006). Identificação bi-objetivo de um sistema de bombeamento mais turbina hidráulica. In *XVI Congresso Brasileiro de Automática*, páginas 1762–1767, Salvador – BA – Brasil.
- Barroso, M. F. S., Takahashi, R. H. C., e Aguirre, L. A. (2007). Multi-objective parameter estimation via minimal correlation criterion. *Journal of Process Control*, 17(4):321–332.
- 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.
- Bonin, M., Seghezza, V., e Piroddi, L. (2010). NARX model selection based on simulation error minimization and LASSO. *IET Control Theory & Applications*, 4(7):1157–1168.
- Carrano, E. G., Takahashi, R. H. C., Fonseca, C. M. M., e Neto, O. M. (2007). Bi-objective combined facility location and network design. In *4th International Conference on Evolutionary Multi-Criterion Optimization*, páginas 486–500, Matsushima – Japan.
- Chankong, V. e Haimes, Y. Y. (1983). *Multiobjective decision making: theory and methodology*. North-Holland, New York.
- 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.
- Corrêa, M. V. e Aguirre, L. A. (2004). Identificação não-linear caixa-cinza: uma revisão e novos resultados. *SBA Controle & Automação*, 15(2):109–126.
- Corrêa, M. V., Aguirre, L. A., e Saldanha, R. R. (2002). Using steady-state prior knowledge to constrain parameter estimates in nonlinear system identification. *IEEE Transactions on circuits and systems I*, 49(9):1376–1381.

- Deb, K., Pratap, A., Agarwal, S., e Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197.
- Eykhoff, P. (1968). Process parameter and state estimation. *Automatica*, 4(4):205.
- Feyerabend, P. K. (1970). *Contra el Método*. Barcelona, Ariel.
- Fonseca, C. M. M. (1995). *Multiobjective Genetic Algorithms with Application to Control Engineering Problems*. Tese de Doutorado, Department of Automatic Control and Systems Engineering, The University of Sheffield.
- Hong, X., Mitchell, R. J., Chen, S., Harris, C. J., Li, K., e Irwin, G. W. (2008). Model selection approaches for non-linear system identification: a review. *Int. J. Systems Sci.*, 39(10):925–946.
- Johansen, T. A. (1996). Identification of non-linear systems using empirical data and prior knowledge - an optimization approach. *Automatica*, 32:337.
- Kashyap, R. L. (1977). A bayesian comparison of different classes of dynamical models using empirical data. In *IEEE Transactions on Automatic Control*, volume 22, páginas 715–727.
- Korenberg, M., Billings, S. A., Liu, Y. P., e McIlroy, P. J. (1988). Orthogonal parameter estimation algorithm for non-linear stochastic systems. *International Journal of Control*, 48(1):193–210.
- Lab-Volt (2006). *Mobile Instrumentation and Process Control Training System, Series 3500*. <http://www.labvolt.com>.
- Laurini, M. P. e Hotta, L. K. (2009). Modelos de fatores latentes generalizados para curvas de juros em múltiplos mercados. Relatório técnico, Insper – Instituto de Ensino e Pesquisa, <http://www.insper.edu.br/en/node/366>.
- Ljung, L. (1999). *System Identification: Theory for the User*. Prentice-Hall, New Jersey, 2 edição.

- Maia, R. D. e Mendes, E. M. A. M. (2008). Identificação de estrutura e parâmetros de modelos de sistemas dinâmicos não-lineares utilizando programação genética. In *XVII Congresso Brasileiro de Automática*, páginas 1–6.
- Martins, S. A. M., Braga, D. C. S., Nepomuceno, E. G., Gomes, T. V., e Reis, M. L. F. (2009). Investigation of static curve information for multiobjective system identification. *Journal of Computational Interdisciplinary Sciences*, 1:149–157.
- Mendes, E. M. A. M. e Billings, S. A. (1998). On overparametrization of nonlinear discret systems. *International Journal of Bifurcation and Chaos*, 8(3):535–556.
- Mendes, E. M. A. M. e Billings, S. A. (2001). An alternative solution to the model structure selection problem. *IEEE Transactions on Systems, Man and Cybernetics, Part A*, 31(6):597–608.
- Miranda, C. G. (2006). *O método LASSO para modelos de COX e sua comparação com propostas tradicionais de seleção de variáveis*. Dissertação de Mestrado, Departamento de Estatística, Programa de Pós-Graduação, Universidade Federal de Minas Gerais.
- Nakamura, T., Judd, K., Mees, A. I., e Small, M. (2006). A comparative study of information criteria for model selection. *International Journal of Bifurcation and Chaos*, 16(8):2153–2175.
- Nepomuceno, E. G. (2002). *Identificação multiobjetivo de sistemas não-lineares*. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- Nepomuceno, E. G., Takahashi, R. H. C., e Aguirre, L. A. (2007). Multiobjective parameter estimation: Affine information and least-squares formulation. *International Journal of Control*, 80(6):863–871.
- Nepomuceno, E. G., Takahashi, R. H. C., Aguirre, L. A., Neto, O. M., e Mendes, E. M. A. M. (2002). Multiobjective nonlinear system identification: a case study with thyristor controlled series capacitor. In *XIV Congresso Brasileiro de Automática*, páginas 2156–2161, Natal – RN – Brasil.

- Pepper, K. (1959). *The Logic of Scientific Discovery*. Harper & Row, Nova York.
- Piroddi, L. (2008). Simulation error minimization methods for narx model identification. *International Journal of Modelling, Identification and Control*, 3(4):392–403.
- Piroddi, L. e Spinelli, W. (2003). An identification algorithm for polynomial NARX models based on simulation error minimization. *International Journal of Control*, 76(17):1767–1781.
- Pulecchi, T. e Piroddi, L. (2007). A cluster selection approach to polynomial NARX identification. In *Proceedings of the 2007 American Control Conference*, páginas 852–857.
- Quarteroni, A. e Fausto, S. (2006). *Cálculo Científico com MATLAB e Octave*. Springer.
- Ray, W. H. (1972). On the mathematical modeling of polymerization reactors,. *Polymer Reviews*, 8(1):1–56.
- Sorenson, H. W. (1970). Least-squares estimation: from Gauss to Kalman. *IEEE Spectrum*, 7:63–68.
- Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society, Series B*, 58(1):267–288.
- Tôrres, L. A. B. e Aguirre, L. A. (2004). Information transmission by controlling nonlinear oscillators. *Physica D*, 196(3-4):387–406.
- Wei, H. L. e Billings, S. (2008). Model structure selection using an integrated forward orthogonal search algorithm assisted by squared correlation and mutual information. *International Journal of Modelling, Identification and Control*, 3(4):341–356.
- Wei, H. L. e Billings, S. (2009). Improved parameter estimates for non-linear dynamical models using a bootstrap method. *International Journal of Modelling, Identification and Control*, 82(11):1991–2001.

-
- Young, P. C. (1970). An instrumental variable method for real-time identification of a noisy process. *Automatica*, 6(2):271–287.
- Zhang, L. F., Zhu, Q. M., e Longden, A. (2007). A set of novel correlation tests for nonlinear system variables. *International Journal of Systems Science*, 38(1):47–60.

Rotinas Computacionais

“Parece que Isaac Newton estava certo quando disse que pôde ver além só por estar apoiado em ombros de gigantes.”

Lutz Bornmann¹

Esse apêndice tem como objetivo apresentar as principais rotinas computacionais empregadas na dissertação. Todas rotinas foram desenvolvidas no *MATLABTM* (Matrix Laboratory, MathWorks, 2007), dada a forma simplificada de apresentação e manipulação de vetores e matrizes e pelo histórico do grupo de pesquisas MACSIN, que possui a maioria de suas rotinas implementadas nessa linguagem.

A.1 build-r

Sintaxe

$[R] = \text{build-r}(\text{modelo})$

¹Lutz Bornmann, pesquisador de bibliometria da Sociedade Max Planck em Munique, Alemanha, autor de artigo que mostra que os artigos mais propensos a ter maior impacto citam outros *papers* de alto impacto (Nature, 14/10/10).

Descrição

Retorna o mapeamento linear R utilizado na estimação multiobjetivo dos parâmetros ou na obtenção dos coeficientes de agrupamentos.

A codificação descrita a seguir é utilizada nessa e nas demais funções que utilizam a estrutura do modelo como informação.

- constante $\Rightarrow 0$;
- $y(k - i) \Rightarrow 100i$;
- $u(k - i) \Rightarrow 200i$;
- $e(k - i) \Rightarrow 300i$;

As linhas dessa matriz representam as combinações dos termos acima, como exemplificado a seguir:

$$modelo = \begin{bmatrix} 1001 & 1001 & 1001 \\ 2001 & 1001 & 0 \\ 0 & 0 & 0 \\ 2001 & 2001 & 2002 \end{bmatrix} = \begin{bmatrix} y(k-1)^3 \\ u(k-1)y(k-1) \\ constante \\ u(k-1)^2u(k-2) \end{bmatrix}.$$

Entrada

modelo : conjunto de termos

Saída

R : mapeamento linear dos coeficientes de agrupamento

João Paulo Vieira - 25/02/10 - UFSJ/São João del-Rei

Alípio Barbosa - 27/03/10 - UFMG/Belo Horizonte

A.2 build-q

Sintaxe

$$[Q, \text{agrup}] = \text{build-q}(R, \text{modelo}, ye, ue)$$

Descrição

Retorna a matriz de regressores estáticos utilizada na estimação Multiobjetivo dos parâmetros.

Entrada

R : matriz dos coeficientes de agrupamento (ver A.1 *build-r*)
modelo : conjunto de termos
 ue : dados estático de entrada
 ye : dados estático de saída

Saída

Q : matriz de regressores estáticos

João Paulo Vieira - 25/02/10 - UFSJ/São João del-Rei

Alípio Barbosa - 27/03/10 - UFMG/Belo Horizonte

A.3 build-pr

Sintaxe

$$[\Psi] = \text{build-pr}(\text{modelo}, u, y)$$

Descrição

Retorna a matriz de regressores Ψ .

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.4 simodeld

Sintaxe

$[y] = \text{simodeld}(\text{modelo}, x_0, u, y_0, e)$

Descrição

Retorna a predição infinitos passos a frente.

Entrada

modelo : modelo identificado
 x_0 : parâmetros
 u : dados de entrada
 y_0 : condições iniciais
 e : ruído

Saída

y : saída predita infinitos passos à frente

Eduardo Mendes - 03/09/94

ACSE - Sheffield

A.5 valsinc

Sintaxe

$$[cus, cla] = \text{valsinc}(u, y, e, gn, \text{arquivo})$$

Descrição

Realiza a validação de um modelo NARMAX identificado utilizando sincronismo com a série de dados de validação.

Entrada

u : dados de entrada
 y : dados de saída
 e : ruído (caso usado na simulação)
 gn : vetor de ganhos: [ganho inicial: incremento: ganho final]
arquivo : arquivo contendo os modelos a serem avaliados

Saída

cus : custo de sincronismo
 cla : classe de sincronismo

Edgar Furtado - 09/09/04

Alípio Barbosa - 12/03/10

UFMG - Belo Horizonte

A.6 myccf

Sintaxe

$$[r] = \text{myccf}(c, \text{lag}, f_1, f_2, cor)$$

Descrição

Retorna a correlação entre dois sinais.

Entrada

- c : matriz ($N \times 2$) que contém os sinais a serem analisados
- lag : máximo atraso de correlação entre os sinais
- f_1 : se $f_1 = 1$ a função de correlação é calculada de $-lag/2$ a $lag/2$.
Se $f_1 = 0$ a função de correlação é calculada de 0 a lag
- f_2 : se $f_2 = 1$ mostra resultado gráfico. Se $f_2 = 0$
não mostra resultado gráfico, apenas valores em r .
- cor : ruído

Saída

- r : função de correlação normalizada

Luis Aguirre - Sheffield - maio/91

Alípio Barbosa - dezembro/09 - UFMG/Belo Horizonte

A.7 orthreg

Sintaxe

$$[modelo, x, e, v] = orthreg(modelo, u, y, [n_p \ n_e], N)$$

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 selecionados
 x : matriz contendo: [(parâmetros) (ERR) (desvio)]
 e : resíduo
 v : variância

Eduardo Mendes - 11/08/94

ACSE - Sheffield

A.8 ntermos

Sintaxe

$[nt] = \text{ntermos}(l, ny, nu, ne)$

Descrição

Determina o número de termos de um modelo polinomial.

Entrada

l : grau de não-linearidade
 ny : máximo atraso em y
 nu : máximo atraso em u
 ne : máximo atraso de ruído

Saída

nt : número de termos de processo e de ruído no modelo

Alípio Barbosa - 04/09/10

UFMG - Belo Horizonte

A.9 akaike

Sintaxe

$[f, mint] = \text{akaike}(\text{modelo}, u, y, f_1, f_2)$

Descrição

Mostra através do critério de informação de Akaike o número aproximado de termos de processo a ser considerado na seleção de estrutura.

Entrada

modelo : conjunto de 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 : sequência contendo os dados do critério de Akaike

$mint$: número de termos a ser considerado

Giovani Rodrigues e Eduardo Mendes - 03/09/94

Márcio Barroso - 10/04/01

A.10 lasso

Sintaxe

$$[w, it, jlasso] = \text{lasso}(psi, y, t)$$

Descrição

Resolve um problema de estimação utilizando uma penalidade l_1 (Tibshirani, 1996).

Entrada

psi : grau de não-linearidade
 y : dados de entrada
 t : vetor de penalidade

Saída

w : parâmetros estimados
 it : número de iteração
 $jlasso$: custo

Adaptado por Alípio Barbosa - 04/04/10
UFMG - Belo Horizonte

