

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



---

## Equivalência de Estruturas de Modelos Não Lineares

---

*Alípio Monteiro Barbosa*

Tese 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 Doutor em Engenharia Elétrica.

**Orientadores:** Dr. Luis Antonio Aguirre  
Dr. Ricardo Hiroshi Caldeira Takahashi

Belo Horizonte, 19 de agosto de 2015



---

# Dedicatória

---

*Aos meus pais, Dr. Alípio e Dra. Marta.*

*Ao meu afilhado, Lucas.*



---

# Agradecimentos

---

Agradeço a Deus!

Aos meus pais, eles são exemplos e fonte de inspiração, verdadeiros doutores. Aos meus irmãos, Érica e Felipe, por todo carinho.

Ao professor Luis Aguirre pela orientação, dedicação, paciência e por todo ensinamento. Obrigado, também, pelas várias conversas que transpunham os limites da engenharia.

Ao professor Ricardo Takahashi pelo apoio.

Ao professor Erivelton Nepomuceno, pela disponibilidade irrestrita, pela amizade e conselhos que vão além da ciência.

Aos amigos Eduardo Bento e João Paulo Vieira pelo companheirismo, ensinamentos e pela amizade.

Ao grande amigo Samir Milani. Obrigado pela ajuda sem “limites” nas mais diversas ocasiões. É um privilégio tê-lo como amigo.

Aos amigos da ELT-E, a melhor turma que já passou pelo CEFET-MG.

À Universidade de Sheffield (Department of Automatic Control and Systems Engineering - The University of Sheffield) pelo acolhimento e oportunidade de realizar o estágio de doutoramento. Ao professor Peter Fleming e ao grupo CODEM pela receptividade.

Aos colegas da Universidade Federal de São João del-Rei e aos colegas do Centro Federal de Educação Tecnológica de Minas Gerais.

À Universidade Federal de Minas Gerais e ao Programa de Pós-Graduação em Engenharia Elétrica pela oportunidade. Agradeço à todos os integrantes do CPH e MACSIN, funcionários, alunos e professores.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior CAPES, pelo suporte financeiro.



---

# Epígrafe

---

*É preciso amor para poder pulsar  
É preciso paz para poder sorrir  
É preciso chuva para florir.*

*Tocando em frente, Renato Teixeira e Almir Sater*

*Andá com fé eu vou  
Que a fé não costuma faiá...*

*Andar com fé, Gilberto Gil*

*“Give thanks and praise to the Lord  
and I will feel all right.”*

*One Love, Bob Marley*





---

---

# Sumário

---

<b>Resumo</b>	<b>xiii</b>
<b>Abstract</b>	<b>xv</b>
<b>Lista de Figuras</b>	<b>xix</b>
<b>Lista de Tabelas</b>	<b>xxi</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Objetivos e contribuições . . . . .	6
1.2 Estrutura da tese . . . . .	6
<b>2 Fundamentação Teórica</b>	<b>9</b>
2.1 Introdução . . . . .	9
2.2 Modelos polinomiais NARX e funções de custo . . . . .	10
2.2.1 Uso da curva estática . . . . .	12
2.3 Identificação multiobjetivo . . . . .	14
2.3.1 A solução Pareto-ótima . . . . .	14
2.4 Teste de hipótese . . . . .	17
2.4.1 Exemplo: teste de hipótese de Wilcoxon . . . . .	19
2.5 Conclusões do capítulo . . . . .	20
<b>3 Equivalência de estruturas de modelos</b>	<b>21</b>

3.1	Introdução . . . . .	21
3.2	Definição do problema . . . . .	23
3.3	Descrição da curva Pareto . . . . .	23
3.4	Seleção do conjunto $\mathcal{G}^*$ . . . . .	25
3.5	Verificação . . . . .	29
3.6	Conclusões do capítulo . . . . .	30
<b>4</b>	<b>Resultados numéricos</b>	<b>31</b>
4.1	Introdução . . . . .	31
4.2	Exemplos de simulação . . . . .	31
4.2.1	Exemplo 1 . . . . .	32
4.2.2	Exemplo 2 . . . . .	39
4.2.3	Influência do atraso puro de tempo . . . . .	46
4.3	Exemplo experimental . . . . .	50
4.4	Conclusões do capítulo . . . . .	60
<b>5</b>	<b>Incerteza intervalar</b>	<b>63</b>
5.1	Introdução . . . . .	63
5.2	$\mathcal{D}(P_j)$ e incerteza nos parâmetros . . . . .	64
5.2.1	Descrição da região de incerteza $\mathbf{D}$ . . . . .	65
5.2.2	Modelo com incerteza intervalar . . . . .	65
5.3	Exemplo . . . . .	68
5.4	Conclusões do capítulo . . . . .	74
<b>6</b>	<b>Conclusões e Perspectivas Futuras</b>	<b>75</b>
6.1	Introdução . . . . .	75
6.2	Discussões e contribuições . . . . .	75
6.3	Publicações . . . . .	77
6.4	Pesquisas futuras . . . . .	78

<b>Bibliografia</b>	<b>79</b>
<b>A Rotinas Computacionais</b>	<b>89</b>
A.1 teste-hipotese . . . . .	89
A.2 pareto . . . . .	90
A.3 build-r . . . . .	91
A.4 build-q . . . . .	92
A.5 build-pr . . . . .	93
A.6 simodeld . . . . .	93
<b>B Tabela Estatística</b>	<b>95</b>



---

# Resumo

---

A identificação de sistemas tem como finalidade obter modelos matemáticos que descrevam o comportamento de um sistema dinâmico a partir de medições. Um problema típico é selecionar uns poucos modelos entre muitas possibilidades. A construção de modelos não lineares, em particular a etapa de seleção de estruturas, apresenta desafios para os quais não há soluções conclusivas.

Tendo em vista que a quantidade e a qualidade de dados de identificação é limitada, nem sempre é possível encontrar a melhor estrutura do modelo. Também não parece necessário nem justificável procurar, em situações práticas, uma *única* estrutura de modelo. Este trabalho, portanto, propõe uma forma de selecionar, a partir de um conjunto de estruturas candidatas, aquelas que parecem mais consistentes com os dados. A solução proposta para este problema foi um método de discriminação de estruturas de modelos com base em um teste de hipótese não paramétrico e conjuntos Pareto referentes a cada estrutura candidata.

O resultado do método proposto é um subconjunto de estruturas de modelos que, do ponto de vista de curvas Pareto, não se distinguem entre si, dentro dos limites de confiança utilizados. Como subproduto, para cada estrutura representativa é possível obter uma região de incerteza  $\mathcal{D}(P)$  determinada no plano de Pareto. A região  $\mathcal{D}(P)$  é convertida em incerteza intervalar nos parâmetros que pode ser usada em métodos de controle robusto.

**Palavras-chave:** Identificação multiobjetivo de sistemas não lineares, seleção de estruturas, incerteza intervalar, identificação relevante para controle.



---

# Abstract

---

System identification aims to obtain mathematical models that describe the behavior of a dynamic system from measurements. A typical problem is to select a few models among many possibilities. The construction of nonlinear models, in particular the structure selection stage presents challenges for which there is no conclusive solutions.

In view of practical limitations it is not always possible to find *the* best model structure. It does not seem necessary or justifiable to seek, in practical situations, a *single best* model structure. This work proposes a way to select from a pool of candidate structures, a subset of model structures that is consistent with the data. The solution proposed to solve this problem is a procedure based on Pareto sets and hypothesis testing to discriminate model structures.

The result of the proposed method is a subset of model structures that is not distinguishable in terms of Pareto curves, for the used data, given a user-defined confidence level. As a byproduct, for each representative structure is possible to obtain an uncertainty region  $\mathcal{D}(P)$  determined on the Pareto plane. The region  $\mathcal{D}(P)$  is converted into parametric uncertainty that can be used in robust control methods.

**Keywords:** Nonlinear system identification, multiobjective, model structure selection, uncertainty, control relevant identification.





---

---

# Lista de Figuras

---

1.1	Esquema ilustrativo dos tipos de identificação . . . . .	2
1.2	Etapas da identificação de sistemas . . . . .	3
1.3	Esquema ilustrativo do problema de equivalência de estruturas .	5
2.1	Dados estáticos de uma válvula pneumática . . . . .	13
2.2	Representação de um conjunto Pareto-ótimo de um problema bi-objetivo . . . . .	16
3.1	Ilustração dos conjuntos Pareto e os respectivos modelos . . . .	24
3.2	Exemplo hipotético, força de conexão entre as estruturas de modelos . . . . .	28
4.1	Dados dinâmicos de identificação e validação, Exemplo 1 . . . .	32
4.2	Dados estáticos, Exemplo 1 . . . . .	33
4.3	Conjuntos Pareto, $P_0$ até $P_6$ . . . . .	34
4.4	Conjuntos Pareto, $P_0$ até $P_5$ . . . . .	35
4.5	População de conjuntos Pareto ótimos $\mathcal{D}(P_0)$ e curvas de Pareto	36
4.6	Relação entre populações de conjuntos Pareto ótimos, $\mathcal{D}(P_0)$ com $\mathcal{D}(P_1)$ até $\mathcal{D}(P_5)$ . . . . .	37
4.7	Força de conexão entre estruturas de modelos, Exemplo 1 . . . .	38
4.8	Força de conexão entre estruturas de modelos com o modelo líder, Exemplo 1 . . . . .	39

4.9	Resposta em malha aberta, Exemplo 1 . . . . .	40
4.10	Resposta em malha aberta, Exemplo 1, parâmetros estimados por MQE . . . . .	40
4.11	Resposta em malha fechada, Exemplo 1 . . . . .	41
4.12	Resposta em malha fechada, Exemplo 1, parâmetros estimados por MQE . . . . .	42
4.13	Dados dinâmicos de identificação e validação, Exemplo 2 . . . .	43
4.14	Dados estáticos, Exemplo 2 . . . . .	43
4.15	Conjuntos Pareto, $P_0$ até $P_3$ . . . . .	44
4.16	Relação entre populações de conjuntos Pareto ótimos, $\mathcal{D}(P_0)$ com $\mathcal{D}(P_1)$ até $\mathcal{D}(P_3)$ . . . . .	45
4.17	Resposta em malha aberta, Exemplo 2 . . . . .	47
4.18	Resposta em malha fechada, Exemplo 2 . . . . .	47
4.19	Dados dinâmicos de identificação e validação, Equação (4.3 . . .	48
4.20	Dados estáticos obtidos a partir da Equação (4.3) . . . . .	49
4.21	Relação entre populações de conjuntos Pareto ótimos, $\mathcal{D}(P_5)$ com $\mathcal{D}(P_1)$ até $\mathcal{D}(P_7)$ . . . . .	51
4.22	Conversor buck . . . . .	51
4.23	Dados de identificação – conversor Buck . . . . .	52
4.24	Dados de validação – conversor Buck . . . . .	53
4.25	Dados estáticos: conversor Buck . . . . .	54
4.26	Conjuntos Pareto, $P_1$ a $P_{13}$ . . . . .	55
4.27	Conjuntos Pareto, detalhe da Figura 4.27 . . . . .	55
4.28	Janela deslizante, representação hipotética . . . . .	56
4.29	Índice $I$ e força de conexão entre estruturas de modelos com o modelo líder, exemplo experimental . . . . .	57
4.30	Resposta em malha aberta, exemplo experimental, estimador bi-objetivo . . . . .	58

4.31	Resposta em malha aberta, exemplo experimental, estimador MQE . . . . .	59
4.32	Resposta em malha fechada, exemplo experimental, estimador bi-objetivo . . . . .	59
4.33	Resposta em malha fechada, exemplo experimental, estimador MQE . . . . .	60
5.1	Fluxograma ilustrativo da ligação entre identificação, incerteza e controle . . . . .	64
5.2	Ilustração da união das populações de conjuntos de Pareto, região de incerteza $\mathbf{D}$ . . . . .	66
5.3	Ilustração dos conjuntos Pareto que compõem a região $\mathbf{D}$ . . . .	67
5.4	Ilustração de $\mathcal{D}(P_0)$ a $\mathcal{D}(P_3)$ . . . . .	69
5.5	União da população de conjuntos de Pareto. Região $\mathbf{D}$ , exemplo ilustrado . . . . .	70
5.6	Região de incerteza $\mathbf{D}$ e região obtida a partir de $\mathcal{G}_{\mathbf{D}}(\hat{\theta}_{\mathbf{D}})$ . . . .	71
5.7	Resposta em malha aberta, estrutura $\mathcal{G}_{\mathbf{D}}$ . . . . .	73
5.8	Resposta em malha fechada, estrutura $\mathcal{G}_{\mathbf{D}}$ . . . . .	73



---

# Lista de Tabelas

---

2.1	Amostras A e B de tamanhos $N_A = 8$ e $N_B = 10$ . . . . .	19
2.2	Classificação das observações. As classificações destacadas com <i>asterisco</i> pertencem à amostra A. . . . .	20
4.1	Estruturas dos modelos $G_2$ , $G_3$ , $G_7$ e $G_{12}$ . . . . .	58
5.1	Valor médio dos parâmetros estimados e os desvios padrão, distribuição gaussiana. . . . .	71



# Introdução

---

*“Procedemos de forma ordenada porque pressupomos que haja ordem. Sem ordem não há problema a ser resolvido. Porque o problema é exatamente construir uma ordem ainda invisível de uma desordem visível e imediata.”*

Rubem Alves<sup>1</sup>

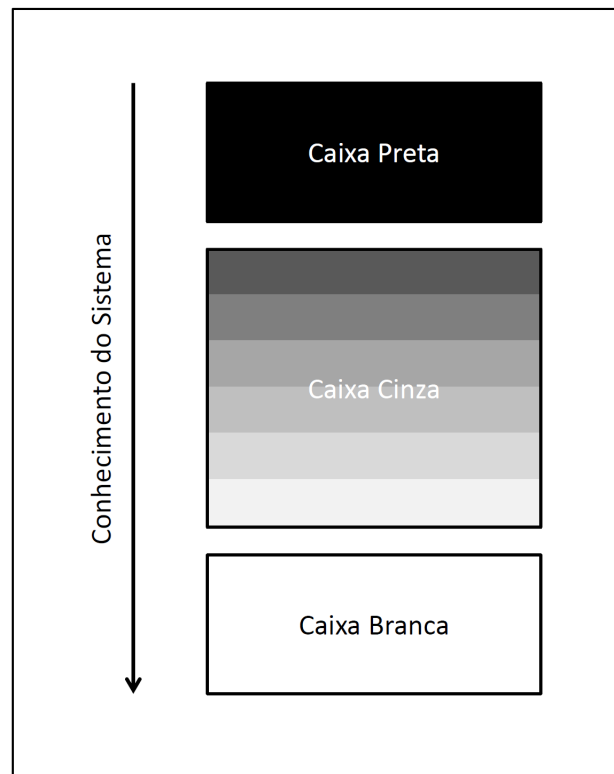
A identificação de sistemas tem como finalidade obter modelos matemáticos que descrevam o comportamento de um sistema dinâmico a partir de medições (Zhu, 2009; Ding et al., 2010; Isermann e Munchhof, 2011). Tais modelos matemáticos podem ser utilizados para analisar as relações de causa e efeito em dados observados, predição e projeto de controladores (Hong et al., 2008; Bazanella et al., 2010).

As técnicas de modelagem podem ser classificadas em três grupos denominados modelagem caixa branca, modelagem caixa preta e modelagem caixa cinza (Aguirre, 2007). Os modelos caixa branca são descritos pelas leis físicas do processo e, em geral, os parâmetros têm um significado físico (Garcia, 2005). Deve-se conhecer o sistema profundamente. Os modelos caixa preta são completamente desenvolvidos a partir dos dados dinâmicos de entrada e

---

<sup>1</sup>Alves (2005, p. 22).

saída. Em princípio nenhum conhecimento prévio do sistema é necessário. Por fim os modelos caixa cinza caracterizam-se por usar informação auxiliar (algum conhecimento prévio do sistema), que não se encontra no conjunto de dados dinâmicos usados na identificação. A modelagem caixa cinza pode ser vista como um procedimento classificado entre a identificação caixa branca e a identificação caixa preta (Corrêa e Aguirre, 2004). O tipo de conhecimento auxiliar utilizado, assim como o nível de conhecimento, varia de caso para caso. A Figura 1.1 mostra que a tonalidade do cinza varia de acordo com o nível e a quantidade de informação auxiliar usada no processo de modelagem (Sohlberg, 1998).



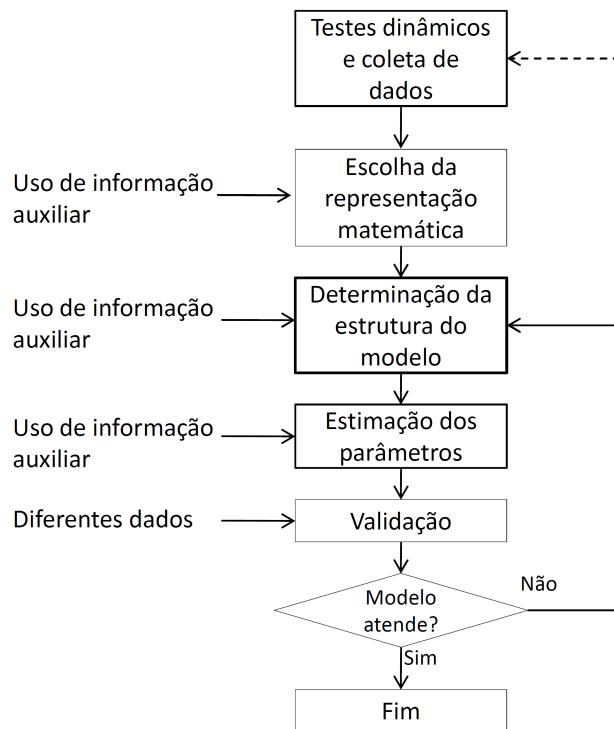
**Figura 1.1:** Esquema ilustrativo dos tipos de identificação.

Esquema ilustrativo dos tipos de identificação. Escala de cinza definindo os diferentes níveis de informação auxiliar.

Existem algumas etapas típicas no processo de construção de modelos matemáticos a partir de dados (Norton, 1986; Ljung, 1999): 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 1.2 é ilustrado o processo de identificação, onde foi adicionado o potencial uso de informação auxiliar.



**Figura 1.2:** Etapas da identificação de sistemas.

Destaque para o bloco de determinação da estrutura do modelo, um dos objetos de estudo do projeto.

O problema central da teoria de identificação de sistemas é justamente obter um modelo apropriado para representar o sistema de interesse. Ou seja, deseja-se selecionar um modelo particular entre uma classe de modelos. A construção de modelos, em particular a etapa de seleção de estruturas, apresenta desafios para os quais não há soluções conclusivas. Uma das primeiras abordagens eficazes para o caso de modelos polinomiais não lineares NARX (*Nonlinear Auto-Regressive 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 fo-

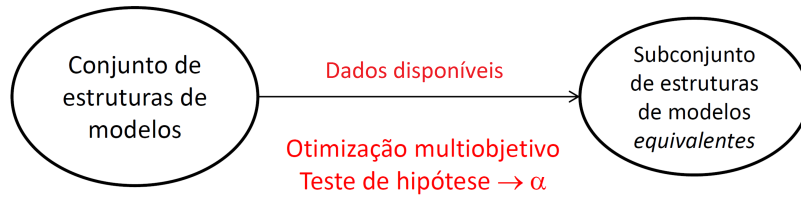
ram descritas na literatura para as quais soluções alternativas foram propostas (Aguirre e Billings, 1995; Aguirre et al., 1998; Mendes e Billings, 2001; Piroddi e Spinelli, 2003). Apesar do importante esforço feito, a determinação de estrutura de modelos não lineares ainda permanece sem uma solução definitiva. Em consequência, é possível encontrar diversos estudos recentes que abordam esse problema (Piroddi, 2008; Wei e Billings, 2008; Hong et al., 2008; Cantelmo e Piroddi, 2010; Aguirre et al., 2010; Barbosa et al., 2010; Baldacchino et al., 2012; Farina e Piroddi, 2012; Piroddi et al., 2012; Baldacchino et al., 2013; Falsone et al., 2015; Solares e Wei, 2015).

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 dos resíduos, 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 há evidências de que 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; Barbosa et al., 2011). No entanto, até agora, o uso de algoritmos multiobjetivo em problemas de seleção de estrutura manteve-se relativamente inexplorado.

Um problema abordado na identificação de sistemas é o de selecionar, a partir de um conjunto de dados e um conjunto de estruturas, as estruturas de modelos que parecem ser mais consistentes com os dados disponíveis. Neste tipo de problema, supõe-se que a fase de escolha dos regressores foi realizada e o que resta é a escolha de estruturas de modelos como um todo.

Esta tese investiga o referido problema usando um estimador bi-objetivo. Uma tentativa de usar a estimação multiobjetivo em problemas de seleção de estrutura foi proposta recentemente por Martins e colaboradores (Martins et al., 2013). O foco de tal trabalho consiste em determinar a estrutura nominal, para o caso simulado, ou a melhor estrutura para o caso experimental. Neste trabalho, no entanto, aborda-se uma questão um pouco diferente. Dado

um conjunto de estruturas candidatas e um conjunto de dados, parece inadequado supor que há uma única estrutura de modelo que seja compatível com os dados (Wit et al., 2012). Suponha que esteja disponível um conjunto de estruturas de modelos. A questão, ilustrada na Figura 1.3 resume-se em encontrar um subconjunto de estruturas que sejam indistinguíveis, em algum aspecto, para o conjunto de dados disponíveis.



**Figura 1.3:** Esquema ilustrativo do problema de equivalência de estruturas.

A questão consiste em selecionar um subconjunto de estruturas de modelos que seja igualmente consistente com os dados.

O problema investigado é análogo à perda de identificabilidade de uma estrutura. A identificabilidade de uma estrutura é uma propriedade que garante que os parâmetros possam ser unicamente determinados a partir de dados de entrada e saída do sistema (Bazanella et al., 2010; Bazanella e Rui, 2012). Ou seja, se houver perda de identificabilidade existirão infinitos valores para os parâmetros capazes de descrever os sinais de entrada e saída medidos do sistema. No problema abordado neste trabalho, o aumento da incerteza nos dados torna cada vez mais difícil discriminar entre “estruturas de modelos” semelhantes.

Esta tese apresenta um método para descartar, dentre um conjunto de estruturas de modelos candidatos, aqueles que são menos consistentes com os dados disponíveis. O método, formulado em termos de um teste de hipóteses não paramétrico, permite que o usuário escolha o nível de significância e este, por sua vez, facilita a interpretação dos resultados. O uso de testes semelhantes tem se mostrado útil em problemas de validação de modelos (Robinson e Froese, 2004; Halder e Bhattacharya, 2014).

As amostras utilizadas para compor o teste de hipótese representam as estruturas de modelos candidatas. Tais amostras são tomadas a partir das soluções de Pareto que resultam da estimação bi-objetivo. Como subproduto

do processo, o método produz uma região de incerteza no plano de Pareto que poderá ser utilizada para caracterizar a incerteza dos parâmetros. O conjunto obtido de estruturas de modelos, que são indistinguíveis a partir dos dados usados, contém informações que uma única estrutura de modelo não pode fornecer. Essa informação, pode auxiliar o procedimento de identificação para controle (Lovera, 2014; Potts et al., 2014), além de dar suporte à formulação de técnicas de controle robusto (Kothare et al., 1996; Ramos et al., 2008; Ottoni, 2013), ou seja, de controle que permita estabilizar a planta e garantir um desempenho mínimo em malha fechada para quaisquer modelos dentre aqueles previamente considerados.

## 1.1 Objetivos e contribuições

Em face do exposto na seção anterior, este trabalho tem por objetivo principal desenvolver uma metodologia de identificação de sistemas não lineares, sob enfoque na etapa de seleção de estruturas, que integre ferramentas de análise robusta, levando em consideração incertezas e aspectos multiobjetivos. Tal metodologia deve ser capaz de indicar a equivalência de estruturas de modelos não lineares.

Os objetivos desta tese podem ser assim resumidos:

1. contribuir com o procedimento de identificação multiobjetivo, em particular a etapa de seleção de estruturas;
2. propor um procedimento baseado na otimização bi-objetivo e teste de hipótese que, dado um conjunto de estruturas candidatas, selecione um subconjunto que seja consistente com os dados a partir de um nível de confiança definido pelo usuário;
3. extrair do método proposto uma região de incerteza e sugerir um procedimento para converter tal região em incerteza paramétrica.

## 1.2 Estrutura da tese

Esta tese está organizada em seis capítulos da seguinte forma:

- Capítulo 1: a motivação e os objetivos do trabalho são apresentados.
- Capítulo 2: traz uma revisão teórica e bibliográfica referente às principais ferramentas utilizadas ao longo deste trabalho. As propriedades de modelos NARX polinomiais e a definição matemática das funções de custo são apresentadas. O capítulo também demonstra a técnica de identificação multiobjetivo empregada, construção da curva de Pareto e os critérios de decisão utilizados. Por fim, uma seção aborda o teste estatístico aplicado.
- Capítulo 3: descreve o procedimento para discriminar estruturas de modelos não lineares.
- Capítulo 4: a metodologia proposta é aplicada em três exemplos simulados e em um exemplo experimental, verificando o desempenho e a robustez da técnica de discriminação de estruturas de modelos.
- Capítulo 5: descreve um procedimento para obter, a partir do método proposto, uma região de incerteza e converter tal região em incerteza paramétrica.
- Capítulo 6: uma discussão final sobre as contribuições do trabalho e perspectivas futuras é apresentada.

O Apêndice, apresentado após a lista de referências bibliográficas, traz as principais rotinas computacionais desenvolvidas durante a tese e a tabela estatística utilizada.



# 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 Alves<sup>1</sup>

## 2.1 Introdução

Este capítulo é uma exposição preliminar referente às principais ferramentas utilizadas ao longo deste trabalho. O objetivo é aproximar o leitor do conteúdo que serviu como base teórica.

Na Seção 2.2 a classe de modelos NARX polinomiais e a definição matemática das funções de custo são apresentadas. Na Seção 2.3 demonstra-se a técnica de identificação multiobjetivo, construção da curva de Pareto, e o critério de decisão utilizado. Por fim, uma breve exposição do teste de hipótese é fornecida na Seção 2.4.

---

<sup>1</sup>Alves (2005, p. 166).

## 2.2 Modelos polinomiais NARX e funções de custo

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

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

sendo  $n_y$  e  $n_u$  os máximos atrasos da saída e entrada, respectivamente. A constante  $d$  é o atraso puro de tempo e  $e(k)$  é ruído, neste caso do tipo branco. Além disso,

$$Z^N = \begin{bmatrix} u(1) & y(1) \\ u(2) & y(2) \\ \vdots & \vdots \\ u(N) & y(N) \end{bmatrix} \quad (2.2)$$

representa 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  $\ell \in \mathbb{N}$  e  $k = 1, \dots, N$ , sendo  $N \in \mathbb{Z}^+$  o número de passos discretos.

A Equação 2.1 pode ser expressa em forma de regressão linear, como

$$y(k) = \psi^T(k-1)\boldsymbol{\theta} + e(k), \quad (2.3)$$

em que  $\psi^T(k-1)$  é o vetor de regressores, que pode conter combinações lineares e não lineares dos termos de entrada e saída e incluir medições até o instante  $(k-1)$ , e  $\boldsymbol{\theta}$  é o vetor de parâmetros. Um modelo do tipo (2.3), tomado sobre uma massa de dados, gera restrições que podem ser apresentadas na forma de uma equação matricial, dada por (Aguirre, 2007):

$$\mathbf{y} = \Psi\hat{\boldsymbol{\theta}} + \boldsymbol{\xi} \quad (2.4)$$

em que o circunflexo indica valores estimados,  $\Psi$  é a matriz de regressão e  $\boldsymbol{\xi}$  é o vetor de resíduos, definido como a diferença entre os dados medidos  $y(k)$  e os dados de predição um passo à frente  $\psi^T(k-1)\hat{\boldsymbol{\theta}}$ , utilizando dados de estimação  $Z^N$ . O vetor de parâmetros  $\boldsymbol{\theta}$  pode ser estimado minimizando o produto interno do vetor de resíduos  $\|\boldsymbol{\xi}\|_2^2$  sobre  $Z^N$ , como detalhado abaixo.



**Definição 1.** O *problema de estimação de parâmetros mono-objetivo* consiste em resolver

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta} \in \mathbf{D}} J_{\text{LS}}(\hat{\boldsymbol{\theta}}, Z^N), \quad (2.5)$$

em que  $\mathbf{D}$  é um conjunto factível de soluções e a função de custo  $J_{\text{LS}}$  é dada por

$$J_{\text{LS}}(\hat{\boldsymbol{\theta}}, Z^N) = \sum_{k=1}^N \left( y(k) - f(\hat{\boldsymbol{\theta}}, Z^N) \right)^2, \quad (2.6)$$

em que  $f(\hat{\boldsymbol{\theta}}, Z^N)$  indica a predição de um passo à frente sobre os dados de estimação  $Z^N$ . No caso de um modelo que seja linear nos parâmetros, esta função de custo pode ser escrita como:

$$J_{\text{LS}}(\hat{\boldsymbol{\theta}}, Z^N) = (\mathbf{y} - \Psi \hat{\boldsymbol{\theta}})^T (\mathbf{y} - \Psi \hat{\boldsymbol{\theta}}). \quad (2.7)$$

□

A fim de desenvolver uma metodologia bi-objetivo, é necessário definir uma outra função de custo (Barbosa et al., 2011). Para o desenvolvimento apresentado as funções podem ser convexas ou não convexas (Chankong e Haimes, 1983). Uma alternativa que tem se mostrado muito útil é o erro em regime estacionário  $J_{\text{SF}}$ , como apresentado em Barroso et al. (2007) e dado por:

$$J_{\text{SF}}(\hat{\boldsymbol{\theta}}, \bar{Z}^M) = \sum_{k=1}^M \left( \bar{y}(k) - f(\hat{\boldsymbol{\theta}}, \bar{Z}^M) \right)^2. \quad (2.8)$$

em que  $\bar{y}(k)$  indica dados medidos em estado estacionário e  $f(\hat{\boldsymbol{\theta}}, \bar{Z}^M)$  os dados estimados, também em estado estacionário.

Para modelos lineares nos parâmetros  $J_{\text{SF}}$  pode ser escrito, convenientemente, como (Nepomuceno et al., 2007):

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

em que  $\bar{Z}^M$  são os dados em estado estacionário e  $\hat{\mathbf{y}} = Q S \hat{\boldsymbol{\theta}}$  indica a característica estática do sistema (Barroso et al., 2007),  $Q$  é a matriz de regressores

estáticos e  $S$  é uma matriz de constantes (zeros e uns) que mapeia o vetor de parâmetros nos coeficientes de agrupamento (Aguirre e Billings, 1995; Barroso et al., 2007, p. 323). A barra superior indica valores em estado estacionário, por exemplo,  $\bar{y} = y(k - i)$ ,  $i = 1, \dots, n_y$ , e  $\bar{u}$  é definido da mesma forma.

Em algumas situações, o estimador de mínimos quadrados é polarizado. Isso ocorre quando o ruído ou erro na equação de regressão não for branco (Aguirre, 2007), ou seja, quando o ruído for autocorrelacionado (colorido). Nesses casos pode-se estimar, além do modelo do processo, um modelo de ruído para evitar a polarização, como é feito no estimador estendido de mínimos quadrados (Ljung, 1999). Existem outros estimadores não tendenciosos que não requerem o ajuste de um modelo de ruído, como o estimador de variáveis instrumentais (Young, 1970) e o estimador bi-objetivo (Barroso et al., 2007; Aguirre et al., 2010). A ideia central do estimador bi-objetivo é utilizar as informações dinâmicas e estáticas para obter um conjunto de soluções Pareto-ótimo. E, por meio de um decisor, escolher um conjunto de parâmetros.

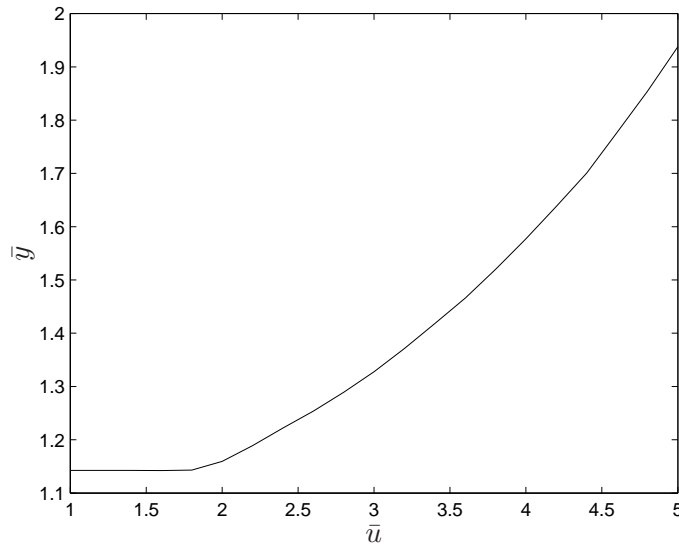
A próxima seção descreve uma justificativa para o uso da curva estática. Depois, na Seção 2.3, o estimador bi-objetivo e a solução Pareto-ótima são apresentados.

### 2.2.1 Uso da curva estática

O que se pretende é garantir que um modelo obtido a partir de dados dinâmicos tenha as características estáticas do sistema, mesmo quando os dados dinâmicos se encontram em uma faixa de operação mais estreita que os dados estáticos. Muitas vezes, é desejável encontrar modelos parcimoniosos com boas respostas estática e dinâmica (Jakubek et al., 2008).

Realizar testes para coleta de dados em estado estacionário, ou mesmo conhecer a curva estática de antemão é comumente mais fácil do que ter dados dinâmicos que excursionem o sistema em uma ampla faixa de operação (Abreu et al., 2012). A Figura 2.1 ilustra a curva estática de uma válvula pneumática obtida experimentalmente (Barbosa, 2010).

A estimação de modelos não lineares com essas características é muito difícil principalmente porque estática e dinâmica não são informações igualmente ponderadas em um único conjunto de dados. Neste ponto, as informações estática e dinâmica podem ser pensadas como sendo conflitantes (Barbosa,



**Figura 2.1:** Dados estáticos de uma válvula pneumática

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

2009; Aguirre et al., 2010).

Estruturas caixa preta flexíveis, tais como redes neurais artificiais, são capazes de aproximar de forma eficiente conjunto de dados de estimação. No entanto, um inconveniente é que uma vez que tais modelos são estimados, a informação estática não é facilmente disponível de forma analítica, ou seja, é mais difícil de extrair informação do sistema utilizando esse tipo de modelo. Deve ser observado que identificação caixa preta, mesmo quando correta, não garante necessariamente um bom desempenho em estado estacionário quando o modelo é não linear (Aguirre et al., 2000).

Quando os dados são de alguma forma conflitantes, é aconselhável usar abordagens multi-objetivo. Problemas de otimização restritos ou multi-objetivo como um meio para identificação caixa cinza foram considerados em (Corrêa et al., 2002; Aguirre et al., 2004; Aguirre e Furtado, 2007; Barroso et al., 2007; Aguirre et al., 2007).

## 2.3 Identificação multiobjetivo

Se em um problema de otimização qualquer houver mais de um objetivo que se deseje alcançar, e se os objetivos forem conflitantes entre si, haverá soluções que, comparadas com outras, 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 multi-objetivo consiste em obter o conjunto de soluções Pareto-ótimas e escolher a solução final por meio de um decisor.

As funções de custo  $J_{LS}(\hat{\theta}, Z^N)$  (2.7) e  $J_{SF}(\hat{\theta}, \bar{Z}^M)$  (2.9) são funcionais que dependem de  $\hat{\theta}$  e, portanto, é possível definir um problema bi-objetivo usando tais funcionais para a estimação de  $\hat{\theta}$ . Do mesmo modo é possível trabalhar simultaneamente com outras funções custo (Nepomuceno et al., 2007; Aguirre et al., 2010).

### 2.3.1 A solução Pareto-ótima

Em geral, não há uma única solução (modelo) que simultaneamente minimize as duas funções de custo  $J_{LS}$  e  $J_{SF}$ , em vez disso, diversas soluções (modelos) podem ser encontradas, no conjunto Pareto-ótimo  $\hat{\Theta}^*$  de soluções  $\hat{\theta}$  caracterizadas por (Chankong e Haimes, 1983):

$$\hat{\theta}^* \in \hat{\Theta}^* \Leftrightarrow \{\nexists \hat{\theta} \mid \mathbf{J}(\hat{\theta}) \leq \mathbf{J}(\hat{\theta}^*) \text{ e } \mathbf{J}(\hat{\theta}) \neq \mathbf{J}(\hat{\theta}^*)\}. \quad (2.10)$$

Na Equação 2.10 as relações de desigualdade são definidas como:

$$\mathbf{V} \leq \mathbf{Z} \Leftrightarrow v_i \leq z_i, \forall i \in 1 \cdots n \quad (2.11)$$

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

em que  $n$  é o número de objetivos,  $v_i$  e  $z_i$  são componentes de um vetor  $n$ -dimensional  $\mathbf{V}$  e  $\mathbf{Z}$ . O vetor de funções, neste trabalho, é definido como

$$\mathbf{J}(\hat{\theta}) = \begin{bmatrix} J_{LS}(\hat{\theta}, Z^N) & J_{SF}(\hat{\theta}, \bar{Z}^M) \end{bmatrix}^T; \quad (2.12)$$

portanto  $n = 2$ .

Os funcionais  $J_{\text{LS}}(\hat{\boldsymbol{\theta}}, Z^N)$  e  $J_{\text{SF}}(\hat{\boldsymbol{\theta}}, \bar{Z}^M)$  podem ser combinados em (2.12) para obter um novo funcional escalar. Uma forma de fazer isto consiste em escrever:

$$J_{\lambda}(\hat{\boldsymbol{\theta}}) = \lambda J_{\text{LS}}(\hat{\boldsymbol{\theta}}, Z^N) + (1 - \lambda) J_{\text{SF}}(\hat{\boldsymbol{\theta}}, \bar{Z}^M), \quad 0 \leq \lambda \leq 1. \quad (2.13)$$

**Definição 2.** O *problema de estimação bi-objetivo* consiste em resolver

$$\hat{\boldsymbol{\theta}}(\lambda) = \arg \min_{\hat{\boldsymbol{\theta}} \in \mathbf{D}'} J_{\lambda}(\hat{\boldsymbol{\theta}}), \quad (2.14)$$

para  $0 \leq \lambda \leq 1$ . O conjunto factível de soluções  $\mathbf{D}'$ , inclui  $\mathbf{D}$  como um caso particular (ver definição 1). □

No caso de modelos lineares nos parâmetros, as funções de custo podem ser expressas como dado em (2.7) e (2.9) e o conjunto bi-objetivo de soluções  $\hat{\boldsymbol{\theta}}(\lambda)$  pode ser encontrado resolvendo o seguinte problema de mínimos quadrados (Nepomuceno et al., 2007):

$$\hat{\boldsymbol{\theta}}(\lambda) = [\lambda \Psi^T \Psi + (1 - \lambda)(QS)^T QS]^{-1} \times [\lambda \Psi^T \mathbf{y} + (1 - \lambda)(QS)^T \bar{\mathbf{y}}]. \quad (2.15)$$

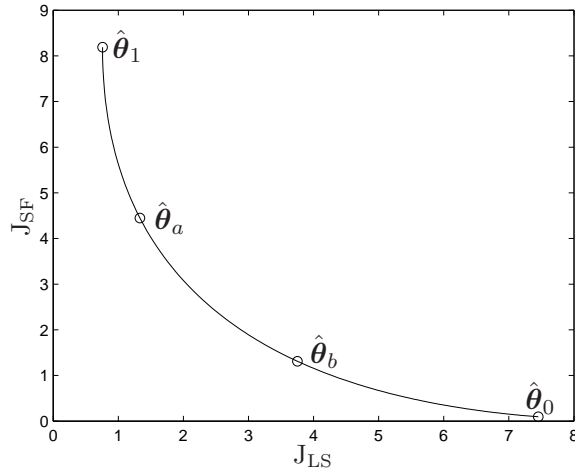
Para cada  $\lambda \in \mathbb{R}$ , encontra-se uma estimativa  $\hat{\boldsymbol{\theta}}(\lambda)$ , que será indicada por  $\hat{\boldsymbol{\theta}}_{\lambda}$ , no conjunto Pareto-ótimo,  $P$ . Na prática são necessários poucos valores de  $\lambda$  para construir a curva Pareto. As soluções mono-objetivo que minimizam  $J_{\text{LS}}(\hat{\boldsymbol{\theta}}_{\lambda})$  e  $J_{\text{SF}}(\hat{\boldsymbol{\theta}}_{\lambda})$  são encontradas, respectivamente, para  $\lambda = 1$  e  $\lambda = 0$ , como ilustrado na Figura 2.2.

Em alguns problemas é necessário seleccionar um único modelo da curva Pareto. Essa é a chamada fase de decisão. Uma maneira de realizar isso é usar o decisor de mínima correlação, conforme detalhado na sequência.

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

$$\tilde{\mathbf{y}}(k) = \hat{\boldsymbol{\psi}}^T(k-1) \hat{\boldsymbol{\theta}}, \quad (2.16)$$

em que  $\hat{\boldsymbol{\psi}}(k-1)$  é a matriz composta de regressores de simulação livre, ou seja, todos os regressores de saída são substituídos pelos valores correspondentes da



**Figura 2.2:** Representação de um conjunto Pareto-ótimo de um problema bi-objetivo.

A curva representada é um conjunto Pareto-ótimo hipotético de um problema bi-objetivo de otimização. As soluções  $\hat{\theta}_1$  e  $\hat{\theta}_0$  denotam os ótimos individuais das funções objetivo  $J_{LS}$  e  $J_{SF}$ , respectivamente.  $\hat{\theta}_a$  e  $\hat{\theta}_b$  são soluções pertencentes ao conjunto Pareto-ótimo.

saída simulada livremente  $\tilde{y}(k)$ . O erro de simulação livre  $\eta(k)$ , para o mesmo modelo candidato, é dado por

$$\eta(k) = y(k) - \tilde{y}(k). \quad (2.17)$$

**Definição 3.** O *decisor de mínima correlação* para um modelo com um vetor de parâmetros  $\hat{\theta}$ , é dado por (Barroso et al., 2007):

$$J_{\text{corr}}(\hat{\theta}) = \left| \frac{1}{N} \sum_{k=1}^N \eta(k) \tilde{y}(k) \right|, \quad (2.18)$$

em (2.18),  $N$  é o comprimento da série temporal. Um modelo sobre o Pareto pode ser escolhido usando  $J_{\text{corr}}$ . Por exemplo, escolhendo-se o modelo com o menor  $J_{\text{corr}}$  dentre um conjunto de candidatos.  $\square$

Dado um sistema com estrutura idêntica à do sistema real, se esse apresentar erro de simulação muito pequeno, ou for não correlacionado com a saída

do modelo, os parâmetros do modelo são próximos dos valores reais (Barroso, 2006).

## 2.4 Teste de hipótese

O teste de hipótese é um método de inferência estatística, e tem como objetivo extrair informações de uma ou mais amostras. A distribuição de probabilidades da estatística de teste paramétrico pressupõe uma forma particular das distribuições populacionais de onde as amostras foram recolhidas. Assim, se a distribuição é desconhecida, deve-se aplicar um teste *não paramétrico*, ou “*distribution-free test*”.

O teste de Wilcoxon é um teste não paramétrico para comparar as distribuições de duas amostras independentes (Gibbons e Chakraborti, 2011). Têm-se o *Wilcoxon Signed-Ranks* e o *Wilcoxon Rank-Sum*. O *Wilcoxon Rank-Sum*, descrito neste trabalho, é equivalente ao *Mann-Whitney U-test*, e deve ser utilizado quando os dados não estão emparelhados (Petrucelli et al., 1999). O único pressuposto exigido para a aplicação do teste de Wilcoxon é que as duas amostras sejam independentes e aleatórias, e que as variáveis em análise sejam numéricas ou ordinais.

Considere que  $A$  e  $B$  são amostras escolhidas aleatoriamente de duas populações, com  $N_A$  e  $N_B$  observações, respectivamente. Deseja-se saber se tais populações são estatisticamente equivalentes, ou seja, em que medida suas distribuições estão próximas. A hipótese nula e a hipótese alternativa são, portanto, definidas como:

$$H_0 : A \equiv B$$

$$H_1 : A \not\equiv B,$$

em que a equivalência ( $\equiv$ ) relaciona-se com a distribuição, como mencionado antes. A hipótese nula afirma que as duas amostras têm distribuições que não podem ser distinguidas com base nos dados disponíveis, ao passo que a hipótese alternativa afirma que as distribuições podem ser discriminadas.

Definidas as hipóteses, a partir dos conjuntos de dados, escolhem-se as

observações das duas amostras para formar uma única amostra de tamanho  $N_A + N_B$ . Tais dados são ordenados em ordem crescente, atribuindo o número 1 à observação de menor valor (o modelo com o menor  $J_\lambda(\hat{\theta})$ ) e o número  $N_A + N_B$  à observação de maior valor de  $J_\lambda(\hat{\theta})$ .

Calculam-se  $W_A$ , a soma dos números de ordem das observações da amostra  $A$  e  $W_B$ , soma dos números de ordem das observações da amostra  $B$ .

Computa-se o número de vezes que a amostra  $A$  ocorre à frente da amostra  $B$ , em um arranjo ordenado dos elementos das duas amostras. As quantidades  $U_A$  e  $U_B$  são definidas por:

$$\begin{aligned} U_A &= N_A N_B + \frac{N_B(N_B + 1)}{2} - W_B; \\ U_B &= N_A N_B + \frac{N_A(N_A + 1)}{2} - W_A. \end{aligned} \quad (2.19)$$

O teste estatístico  $U$  é dado por (Wilcoxon et al., 1972):

$$U = \min(U_A, U_B). \quad (2.20)$$

Por meio da estatística (2.20), calcula-se o valor- $p$  ( $p$ -Value). O valor- $p$  pode ser obtido por meio de tabelas (Petrucelli et al., 1999). Neste trabalho, os valores- $p$ , para o teste de Wilcoxon, foram obtidos usando a função **ranksum** (ver Apêndice A) do programa Matlab ©.

O valor- $p$  é a probabilidade de ocorrência de valores iguais ou superiores ao assumido pela estatística do teste, sob a hipótese de que  $H_0$  seja verdadeira. Portanto, um valor- $p$  pequeno significa que a probabilidade de obter um valor da estatística de teste como o observado é muito improvável. Quanto menor o valor- $p$ , mais evidência existe em favor de rejeitar a hipótese nula.

Finalmente, para tomada de decisão, define-se um parâmetro de ajuste  $\alpha$ , denominado nível de significância e estabelece-se a seguinte regra:

$$\begin{aligned} \text{Se valor-}p &\leq \alpha, \text{ rejeita-se } H_0, \text{ com um nível de significância } \alpha. \\ \text{Se valor-}p &> \alpha, \text{ não se rejeita } H_0. \end{aligned} \quad (2.21)$$



### 2.4.1 Exemplo: teste de hipótese de Wilcoxon

O exemplo a seguir tem o intuito de ilustrar uma aplicação do teste de hipótese. Considere duas amostras, A e B, dispostas na Tabela 2.1 (Walpole et al., 2012).

**Tabela 2.1:** Amostras A e B de tamanhos  $N_A = 8$  e  $N_B = 10$ .

Amostra A	2,1	4,0	6,3	5,4	4,8	3,7	6,1	3,3		
Amostra B	4,1	0,6	3,1	2,5	4,0	6,2	1,6	2,2	1,9	5,4

Deseja-se saber se tais populações são estatisticamente equivalentes dado um nível de significância  $\alpha = 0,05$ .<sup>1</sup>

Portanto, para a solução do problema, deve-se calcular  $W_A$  e  $W_B$ . As observações são ordenadas e enumeradas de acordo com a Tabela 2.2, e:

$$W_A = 4 + 8 + 9 + 10,5 + 13 + 14,5 + 16 + 18 = 93$$

$$W_B = 1 + 2 + 3 + 5 + 6 + 7 + 10,5 + 12 + 14,5 + 17 = 78$$

Computa-se  $U_A$  e  $U_B$ :

$$U_A = 8 \times 10 + \frac{10 \times (10 + 1)}{2} - 78 = 57;$$

$$U_B = 8 \times 10 + \frac{8 \times (8 + 1)}{2} - 93 = 23.$$

A estatística  $U$  (Equação 2.20) é dada por:

$$U = \min(U_A, U_B) = \min(57, 23) = 23. \quad (2.22)$$

De acordo com a Tabela A.17, disponível em Walpole et al. (2012, p. 761) e reproduzida no Apêndice B, a região crítica ocorre em  $U = 17$ . Portanto, para valores de  $U$  maiores que 17 não se rejeita a hipótese nula. Isto equivale, para o exemplo simulado, a um valor de probabilidade, valor- $p$ , igual a 0,0688 e, como o valor- $p$  foi maior que  $\alpha = 0,05$  (Equação 2.21), não se rejeita  $H_0$ . Portanto, não há evidências de que as distribuições sejam idênticas.

<sup>1</sup>Computacionalmente equivale a escrever  $[p, h] = \text{ranksum}(A, B, \alpha)$ , em que  $p$  indica o valor- $p$  e  $h$  se a hipótese nula foi rejeitada, conforme descrito no Apêndice A.

**Tabela 2.2:** Classificação das observações. As classificações destacadas com *asterisco* pertencem à amostra A.

Observações	Classificação
0,6	1
1,6	2
1,9	3
2,1	4*
2,2	5
2,5	6
3,1	7
3,3	8*
3,7	9*
4,9	10,5*
4,0	10,5
4,1	12
4,8	13*
5,4	14,5*
5,4	14,5
6,1	16*
6,2	17
6,3	18*

## 2.5 Conclusões do capítulo

O presente capítulo teve como objetivo apresentar ao leitor as técnicas de identificação de sistemas, em particular a estimação bi-objetivo. A solução Pareto-ótima foi apresentada, bem como o decisor a ser utilizado. O teste de hipótese, útil para discriminar modelos, foi também detalhado. Essas técnicas servirão como base para o método de equivalência de estruturas de modelos não lineares, apresentado no Capítulo 3.

# Equivalência de estruturas de modelos

---

*“Você paga um preço muito alto no mercado de ações por um consenso animado. A incerteza, na realidade, é amiga do comprador de valores a longo prazo.”*

Warren Buffett<sup>1</sup>

## 3.1 Introdução

Um modelo é uma versão simplificada de algum problema ou situação real. Não raro, mais de um modelo pode descrever um mesmo fenômeno, haja visto que cada pesquisador tem a liberdade de modelar o processo seguindo a metodologia que julgar mais adequada.

Na identificação de um modelo não linear, a etapa de escolha de qual representação deve ser utilizada ainda é uma questão que parece não ter uma resposta definitiva. Há uma grande variedade de representações não lineares, tais como redes neurais artificiais (Haykin, 2001), funções de base radial (Casdagli, 1989), *wavelets* (Billings e Coca, 1999), séries de Volterra (Volterra, 1930; Doyle et al., 1995), modelos de blocos interconectados (Wiener, 1958; Aguirre et al., 2005), modelos polinomiais e racionais (Chen e Billings, 1989). Independente do tipo de representação usada, um problema comum a todas elas é a determinação da estrutura.

---

<sup>1</sup>Schroeders (2008, p. 486).

A seleção de estrutura representa uma etapa extremamente importante no processo de identificação de sistemas, o que justifica o número elevado de trabalhos recentes publicados na área (Piroddi e Leva, 2006; Pulecchi e Piroddi, 2007; Hong et al., 2008; Piroddi, 2008; Wei e Billings, 2008; Cantelmo e Piroddi, 2010; Aguirre et al., 2010; Barbosa et al., 2010; Barbosa, 2010; Aguirre et al., 2010; Baldacchino et al., 2012; Farina e Piroddi, 2012; Piroddi et al., 2012; Alves et al., 2012; Baldacchino et al., 2013; Falsone et al., 2015).

Um ponto comum aos trabalhos citados é a busca por um procedimento que encontre o modelo exato gerador dos dados. Em outras palavras, os autores almejam um método que seja capaz de incluir todos os regressores genuínos do modelo e excluir os regressores espúrios. Nos sistemas simulados, os autores mostraram que, em alguns casos, é possível resgatar as estruturas originais do modelo. Mais recentemente, Baldacchino e colaboradores propuseram um método de seleção de estrutura estocástica capaz de reproduzir o modelo original em situações em que os métodos de regressão falham (Baldacchino et al., 2013).

Contudo, nos casos em que os dados são medidos de um sistema, possivelmente físico, não existem regressores “genuínos”, nem “espúrios”. Portanto, a questão é se existe um procedimento de modelagem que deve indicar, a partir de um conjunto de dados, o modelo *verdadeiro*? Ou, baseado nos dados, existe apenas *um* modelo que melhor aproxime os dados (Wit et al., 2012)? Se a conjectura é realmente falsa, como isso se relaciona com o grande número de técnicas para a seleção de estruturas de modelos?

As perguntas fundamentais sobre a seleção do modelo não podem ser respondidas apenas prestando atenção aos detalhes matemáticos ou procedimentos envolvidos (Hájek, 2007). A seleção de modelos nos obriga a enfrentar questões fundamentais em estatística e olhar para o contexto no qual o modelo é usado. Do ponto de vista estatístico existem muitos aspectos da seleção de modelos a serem explorados (Tibshirani, 1996; Miranda, 2006; Henderson et al., 2010; Chakrabarti e Ghosh, 2011).

A abordagem de Longford (2005) destaca ainda que a seleção de modelo está intimamente ligada a problemas de decisão (Savage, 1951) e que o critério de escolha não deve resultar em um único modelo. Neste sentido, escolher um modelo é um trabalho de base estatística e a ideia de um modelo *verdadeiro*

pode ser desnecessária (Burnham e Anderson, 2002; Longford, 2005); o que se requer é uma teoria para tomada de decisão.

Neste contexto, descreve-se no presente Capítulo uma metodologia para discriminar estruturas de modelos não lineares a partir da técnica de identificação multiobjetivo, apresentada no Capítulo 2. Utilizando hipóteses estatísticas é possível escolher um subconjunto de estruturas de modelos, supostamente equivalente, dado um nível de significância. As estruturas escolhidas são posteriormente submetidas a testes para verificação da suposta equivalência.

## 3.2 Definição do problema

Sejam  $Z^N$  e  $\bar{Z}^M$  dados dinâmicos e estáticos de identificação, respectivamente, coletados de um sistema  $\mathcal{S}$ . Suponha disponível um conjunto de estruturas de modelos SISO (*Single-Input Single-Output*),  $\mathcal{G} = \{G_1, G_2, \dots, G_{N_m}\}$ . Não são necessárias outras hipóteses sobre essas estruturas. De fato, as estruturas de modelos em  $\mathcal{G}$  poderiam ser de diferentes classes. O objetivo é encontrar um subconjunto de estruturas de modelos,  $\mathcal{G}^* \subset \mathcal{G}$ , que são indistinguíveis, em algum aspecto, do ponto de vista dos dados  $Z^N$  e  $\bar{Z}^M$  disponíveis.

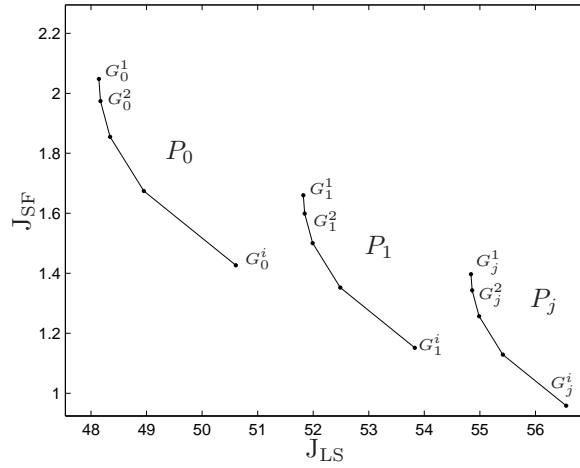
Na próxima seção é mostrado como produzir um conjunto de Pareto para cada estrutura de modelo. A curva Pareto incorpora, por construção, o desempenho do modelo sobre os conjuntos de dados  $Z^N$  e  $\bar{Z}^M$ .

## 3.3 Descrição da curva Pareto

Para o desenvolvimento a seguir, considere que os dados dinâmicos e estáticos são produzidos por um modelo  $G_0(\hat{\theta})$ , possivelmente não linear e SISO. Dada uma estrutura de modelo  $G_j$ , é possível determinar um vetor de parâmetros usando o estimador bi-objetivo (ver Definição 2 na Seção 2.3.1). O modelo estimado será indicado por  $G_j^i(\hat{\theta}_{ij})$ , ou simplesmente por  $G_j^i$ , em que  $i$  indica que o modelo é o  $i$ -ésimo no Pareto referente a  $G_j$  (ver Figura 3.1).

Dado um modelo  $G_j^i$  e os conjuntos de dados  $Z^N$  e  $\bar{Z}^M$ , as funções de custo  $J_{LS}$  e  $J_{SF}$  podem ser calculadas como detalhado na Seção 2.2. Portanto, o modelo  $G_j^i$  pode ser representado como um ponto no plano  $J_{LS} \times J_{SF}$ .

Se a  $j$ -ésima estrutura do modelo é utilizada e é permitido que o vetor de parâmetros assuma um conjunto de valores  $\{\hat{\theta}_{1j}, \hat{\theta}_{2j}, \dots, \hat{\theta}_{ij}\}$ , então, o conjunto de modelos  $G_j^i(\hat{\theta}_{ij})$ ,  $i = 1, 2, \dots$ , pode ser representado por um conjunto de pontos no plano  $J_{LS} \times J_{SF}$ , denominado *conjunto Pareto*, em que os conjuntos de parâmetros são resultados do procedimento bi-objetivo de estimação. Cada curva Pareto  $P_j$  (ver Figura 3.1) é composta por modelos  $\{G_j^1, G_j^2, \dots, G_j^i\} \in P_j$  de mesma estrutura e parâmetros diferentes. Cada estrutura de modelo dá origem a um conjunto Pareto diferente.



**Figura 3.1:** Ilustração dos conjuntos Pareto e os respectivos modelos.

Conjuntos Pareto  $\{P_0, P_1, \dots, P_j\}$  e seus respectivos modelos  $G_j^i$ , em que o índice  $j$  refere-se a cada conjunto Pareto (diferente estrutura) e o índice  $i$  refere-se aos modelos do mesmo conjunto Pareto (mesma estrutura).

Como a estrutura de todos os modelos em cada conjunto Pareto é a mesma, o que distingue um modelo do outro em  $\{G_j^1, G_j^2, \dots, G_j^i\}$  é o vetor de parâmetros. Portanto, o conjunto Pareto relacionado a  $G_j$  é na verdade uma família de vetores de parâmetros indicados por  $\hat{\theta}(\lambda, G_j, Z^N, \bar{Z}^M)$ ,  $0 \leq \lambda \leq 1$ . Para cada valor específico de  $\lambda$  o vetor de parâmetros correspondente será denotado simplesmente por  $\hat{\theta}_{ij}$  ou por  $\hat{\theta}_\lambda$ .

Portanto, um conjunto de Pareto no plano  $J_{LS} \times J_{SF}$  pode ser formalmente

descrito como:

$$P_j(G_j(\hat{\theta}_\lambda), Z^N, \bar{Z}^M): \{J_{\text{LS}}(\hat{\theta}_\lambda, Z^N) \times J_{\text{SF}}(\hat{\theta}_\lambda, \bar{Z}^M) \mid \hat{\theta}_\lambda \in \hat{\theta}(\lambda, G_j, Z^N, \bar{Z}^M)\}, \quad (3.1)$$

com  $0 \leq \lambda \leq 1$ . Cada estrutura de modelo  $G_j$  dá origem a um conjunto Pareto correspondente,  $P_j$ .

Tendo descrito formalmente um conjunto Pareto, o próximo passo é considerar o efeito da incerteza presente nos parâmetros estimados. A partir de (3.1), pode ser visto que a incerteza sobre o vetor de parâmetros refletirá diretamente sobre a incerteza do conjunto Pareto.

As amplitudes de tais incertezas são indiretamente utilizadas no teste de hipótese, descrito na Seção 2.4, quando determinam-se as amostras das duas populações. Portanto, para aplicar o teste de hipótese deve-se obter tais populações. Uma forma prática de fazer isso é:

$$\mathcal{D}(P_j) = \bigcup_{i=1}^{N_r} P_j(G_j(\hat{\theta}_\lambda), Z_i^N, \bar{Z}^M). \quad (3.2)$$

A população  $\mathcal{D}(P_j)$  relacionada ao conjunto Pareto  $P_j$  é a união de todos  $N_r$  conjuntos de Pareto produzidos pela *mesma* estrutura de modelo, mas com diferentes realizações de dados dinâmicos e diferentes parâmetros. Portanto, a dispersão de  $\mathcal{D}(P_j)$  indica a incerteza associada à estrutura  $G_j$ , aos parâmetros  $\hat{\theta}_\lambda$  e aos dados  $Z_i^N$ .

### 3.4 Seleção do conjunto $\mathcal{G}^*$

O ponto central da metodologia proposta pode ser apresentado da seguinte maneira. Considere dois modelos de diferentes estruturas  $G_j$  e  $G_k$  e o conjunto de dados  $Z^N$  e  $\bar{Z}^M$ . Seguindo o procedimento detalhado na Seção 3.3, calcula-se  $P_j$ ,  $\mathcal{D}(P_j)$ ,  $P_k$  e  $\mathcal{D}(P_k)$ . Se  $P_k$  está contido na população  $\mathcal{D}(P_j)$  ou se existe alguma sobreposição entre as populações  $\mathcal{D}(P_k)$  e  $\mathcal{D}(P_j)$ , seria razoável afirmar que as estruturas dos modelos  $G_j$  e  $G_k$  são distinguíveis para os conjuntos de dados disponíveis?

Em vista disso, as seguintes observações são apresentadas:

*Observação 1.* Na metodologia proposta, comparam-se conjuntos Pareto, que incorporam o desempenho de modelos considerando o erro de predição de um passo à frente sobre dados de validação ( $Z_v^{N'}$ ) e o erro de estado estacionário sobre os dados estáticos ( $\bar{Z}^M$ ). Os parâmetros são estimados usando o estimador bi-objetivo.

*Observação 2.* É plausível que dois modelos de estruturas indistinguíveis neste cenário, poderiam ser distinguidos utilizando um conjunto de dados dinâmicos mais informativo ou um estimador de parâmetros diferente. Essa observação é análoga à perda de identificabilidade de uma estrutura (Bazanella e Rui, 2012).

*Observação 3.* Deseja-se responder à pergunta: as distribuições de “ $\mathcal{D}(P_k)$  e  $\mathcal{D}(P_j)$  são distinguíveis? Ou seja, os modelos de estrutura  $G_j$  e  $G_k$  são distinguíveis em termos de suas populações de conjuntos Pareto ótimos?” A questão é avaliada por meio de um teste de hipótese em que a hipótese nula  $H_0$  é que a população e, portanto as estruturas do modelo, são *não* distinguíveis. Deseja-se saber se essa hipótese  $H_0$  pode ser rejeitada e com que confiança, com base nos dados disponíveis.

O procedimento final é descrito na sequência. Sejam  $Z_i^N$ ,  $i = 1, \dots, N_r$  realizações de dados dinâmicos e  $\bar{Z}^M$ , os dados estáticos. Suponha que há um conjunto inicial  $\mathcal{G}$  de  $N_m$  estruturas candidatas.

1. Utilizando (3.2), calcular as populações de conjunto Pareto  $\mathcal{D}(P_j)$ ,  $j = 1, \dots, N_m$ . O vetor de parâmetro  $\hat{\theta}_\lambda$  é estimado para cada realização de dados dinâmicos  $Z_i^N$ .
2. Escolher um par de estruturas de modelos,  $G_j$  e  $G_k$  para avaliação. Aleatoriamente selecionar um grupo de  $N_A$  conjuntos Pareto de  $\mathcal{D}(P_j)$  e de cada conjunto Pareto escolha o modelo baseado no decisor de mínima correlação (ver Definição 2.3.1). O grupo resultante com  $N_A$  modelos é denominado amostra A. Procedendo da mesma forma com  $\mathcal{D}(P_k)$ , obtém-se um outro grupo de  $N_B$  modelos denominado amostra B.
3. Escolher o valor de  $\alpha$  (por exemplo 1%) e aplicar o teste de hipótese (ver Seção 2.4).



- (a) Calcular  $U = \min(U_A, U_B)$ , ver (2.19);
  - (b) Usando  $U$ , calcular o valor- $p$ , ver após (2.19);
  - (c) Se o valor- $p \leq \alpha$ , então rejeita-se  $H_0$ .
4. Repetir passos 2 e 3  $N_h$  vezes (por exemplo  $N_h = 50$ ). As amostras A e B devem ser diferentes em cada vez.
  5. Calcular o percentual de vezes que  $H_0$  não foi rejeitada. Esse percentual, chamado de  $w_{jk}$ , é a força de ligação entre  $G_j$  e  $G_k$ .
  6. Repetir os passos de 2 a 5 para todas as estruturas de modelos no conjunto  $\mathcal{G}$  tomados aos pares.
  7. Para cada estrutura de modelo, calcular

$$I_j = \frac{1}{N_m - 1} \sum_{k=1, k \neq j}^{N_m - 1} w_{jk}, \quad j = 1, 2, \dots, N_m, \quad (3.3)$$

o modelo com maior índice  $I$  é definido como a estrutura de modelo *líder*.

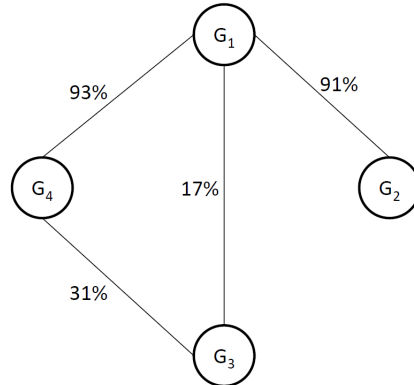
8. Definir um parâmetro de ajuste  $\beta$  (por exemplo 90%). Estruturas de modelos que conectam com a estrutura de modelo líder, com  $w_{jk} \geq \beta$  são consideradas  $\beta$ -equivalente com um nível de significância  $\alpha$  (ver Passo 3) e são designadas para o grupo  $\mathcal{G}^*$ .

A Figura 3.2 exemplifica numericamente o procedimento final, ilustrando a conexão entre as  $N_m = 4$  estruturas de modelos, após avaliação via simulação Monte Carlo. Neste caso  $\mathcal{G} = \{G_1, G_2, G_3, G_4\}$ .

Para cada estrutura de modelo, os índices (3.3) são  $I_1 = \frac{1}{3}(91 + 17 + 93) = 67,0\%$ ,  $I_2 = \frac{1}{3}(91 + 0 + 0) = 30,3\%$ ,  $I_3 = \frac{1}{3}(17 + 0 + 31) = 16,0\%$  e  $I_4 = \frac{1}{3}(93 + 0 + 31) = 41,3\%$ .

Como a estrutura de modelo  $G_1$  teve o maior índice, ele foi escolhido como a estrutura de modelo líder. Portanto, todas as outras estruturas de modelos conectadas com  $G_1$ , com  $w_{jk} \geq \beta$ , serão consideradas  $\beta$ -equivalentes.

Para um  $\beta$  de 90%, por exemplo,  $G_2$  e  $G_4$  são ditos  $\beta$ -equivalentes a  $G_1$ , mas não a  $G_3$ , com um nível de significância  $\alpha$  (ver Equação 2.21) e, portanto,  $\mathcal{G}^* = \{G_1, G_2, G_4\}$ . Claramente,  $\alpha$  e  $\beta$  são parâmetros definidos pelo usuário que afetam a interpretação dos resultados.



**Figura 3.2:** Exemplo hipotético, força de conexão entre as estruturas de modelos.

As conexões entre os pares de estruturas de modelos mostram o percentual de vezes que a hipótese nula ( $H_0$ ) não foi rejeitada.

### Características das constantes definidas pelo usuário

Um breve resumo dos principais parâmetros do procedimento é apresentado a seguir. Embora sejam parâmetros ajustados pelo usuário, valores padrão são sugeridos. No Capítulo 4 exemplos numéricos ilustrarão a influência de tais parâmetros.

*N - Tamanho do conjunto de dados dinâmico:* o comprimento dos dados é a quantidade de amostras que deverá ser coletada durante o experimento de identificação. Quando multiplicado pelo período de amostragem determina a duração total do experimento. A duração do teste em sistemas experimentais deverá ser minimizada com o objetivo de reduzir variações indesejadas no processo e o tempo de engenharia necessário para acompanhar o experimento. Para o caso do procedimento proposto, sugerem-se valores de amostras iguais ou superiores a 100, o que possibilitará a construção da população de conjuntos Pareto ótimo com no mínimo 50 indivíduos (ver Figura 4.28).

*M - Tamanho do conjunto de dados estático:* pode ser coletado durante o procedimento de identificação. No caso prático, aplica-se um entrada constante, espera-se o sistema entrar em regime permanente e coleta o sinal estático de saída. Esses dados podem ser obtidos a partir do registro de dados históricos

(Abreu et al., 2012). Sugerem-se valores de  $M$  iguais ou superiores a 10.

$N_r$  - *Número de realizações*: está diretamente relacionado com o tamanho da população de conjuntos Pareto ótimo. Para o caso prático, dependerá do tamanho do conjunto de dados dinâmicos. Podem ser obtidos por janela deslizante. Sugerem-se valores iguais ou superiores a 50. O tempo de processamento é pouco influenciado, permanecendo na ordem de segundos.

$N_h$  - *Número de repetições do teste de hipótese*: para o primeiro exemplo, o tempo necessário foi inferior a 20 segundos (para 200 execuções) e menos de 5 segundos para 50 execuções. O resultado foi o mesmo. Portanto, sugere-se 50 como um número padrão.

$N_A$  e  $N_B$  - *Número de amostras para o teste de hipótese*: Sugere-se  $N_A = N_B = 8$ . Nos testes realizados o aumento do número de amostras não alterou os resultados.

$\alpha$  - *Nível de significância*: está vinculado à probabilidade de rejeitar uma hipótese verdadeira. Um resultado que é *significante* ao nível de 1% é mais *significante* do que um resultado que é *significante* ao nível de 5%. No entanto, um teste ao nível de 1% é mais susceptível de não rejeitar a hipótese nula quando esta é falsa, do que um teste de 5%. Sugestão como valor padrão,  $\alpha = 1\%$ .

$\beta$  - *Taxa de equivalência entre modelos*: define os modelos que serão ditos equivalentes. Valores de  $\beta$  maiores reduzem o conjunto de modelos equivalentes. Ao mesmo tempo que para valores menores de  $\beta$  tem-se um conjunto mais populoso de modelos equivalentes, porém com uma *menor força* de equivalência. Sugere-se  $\beta = 90\%$ .

## 3.5 Verificação

O procedimento para encontrar  $\mathcal{G}^*$  foi proposto na seção anterior e com isso o objetivo definido na Seção 3.2 foi atingido. Os dois testes mencionados a seguir serão utilizados para verificar o procedimento e não fazem parte da

metodologia geral proposta. Nesta seção, o intuito não é validar os modelos, mas sim determinar algum critério para avaliar o resultado encontrado, de forma a verificar a equivalência no desempenho das estruturas de modelos  $\mathcal{G}$ . Em outras palavras, deseja-se verificar se os modelos são equivalentes entre si, e não se os modelos incorporam ou não as características de interesse do sistema.

A fim de fazer isso, os critérios de análise devem ser diferentes daqueles usados para a construção da curva Pareto ( $J_{LS}$  e  $J_{SF}$ ). É fundamental observar que os procedimentos desta seção *não* precisam ser realizados na prática.

Para cada Pareto (estrutura de modelo) escolhe-se o modelo baseado no decisor de mínima correlação (ver Definição 3), que já está disponível como um subproduto do procedimento. Além disso, uma vez que o método proposto refere-se à escolha de estruturas de modelos, há liberdade para escolher o estimador de parâmetros. Assim, o vetor de parâmetros será estimado utilizando o algoritmo de mínimos quadrados estendido (MQE) (Billings et al., 1989).

Portanto, para verificação, escolheram-se duas ferramentas de análise: i) resposta em malha aberta e ii) resposta em malha fechada. No caso da resposta em malha fechada, os ganhos foram os mesmos para todos os modelos. O sinal de entrada para a simulação da resposta em malha fechada é dado por:

$$u(k) = K_1 u(k-1) + K_2 [r(k) - y(k)], \quad (3.4)$$

que corresponde a uma lei de controle com ganhos  $K_1$  e  $K_2$ , em que  $r(k)$  é a referência e  $(r(k) - y(k))$  é o erro  $e(k)$ . Portanto,  $u(k) = K_1 u(k-1) + K_2 e(k)$ .

No domínio  $z$  pode-se escrever  $\frac{U(z)}{E(z)} = \frac{K_2 z}{z - K_1}$ . Logo, tem-se um integrador quando  $K_1 = 1$ . Quando  $K_1 = 0$  o controlador é puramente proporcional.

## 3.6 Conclusões do capítulo

Neste capítulo descreveu-se um procedimento para discriminar a estrutura de modelos não lineares a partir da técnica de identificação multiobjetivo e utilizando hipóteses estatísticas.

A metodologia apresentada será avaliada nos próximos capítulos. Exemplos numéricos mostrarão sua aplicação, bem como seu desempenho e robustez.

# Resultados numéricos

---

*“Logo que, numa inovação, nos mostram alguma coisa de antigo, ficamos sossegados.”*

Friedrich Nietzsche

## 4.1 Introdução

Este Capítulo tem como objetivo ilustrar a metodologia desenvolvida no Capítulo 3. Pretende-se verificar o desempenho e a robustez do critério de discriminação de estruturas de modelos, isto é, encontrar o conjunto  $\mathcal{G}^*$  e avaliar se os modelos com estruturas em  $\mathcal{G}^*$  são semelhantes em algum sentido. A segunda etapa será realizada utilizando os critérios mencionados na Seção 3.5. Para isso, serão estudados três exemplos simulados e um experimental.

## 4.2 Exemplos de simulação

Dentre os exemplos simulados, o primeiro exemplo (Seção 4.2.1), estudado por Bonin et al. (2010) e o segundo exemplo (Seção 4.2.2), têm o intuito de averiguar o desempenho do método proposto. Como a estrutura dos sistemas utilizados nesses dois exemplos é conhecida, o desempenho final do método é facilmente avaliado.

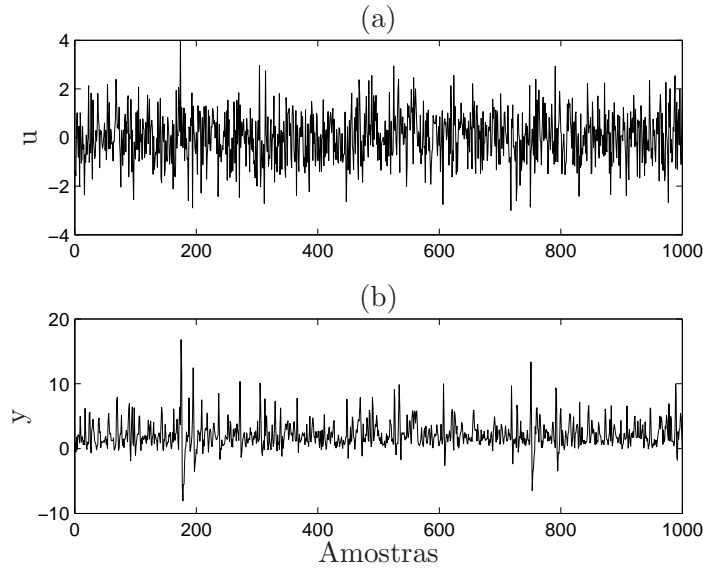
O terceiro exemplo é uma extensão da discussão, e tem como objetivo avaliar o efeito de incluir no conjunto inicial  $\mathcal{G}$  estruturas com atraso de tempo incorreto.

### 4.2.1 Exemplo 1

Considere o seguinte sistema (Bonin et al., 2010):

$$y(k) = 0,5 + 0,5y(k-1) + 0,8u(k-2) + u(k-1)^2 - 0,05y(k-2)^2 + e(k), \quad (4.1)$$

em que a entrada  $u(\cdot)$  é um sinal do tipo ruído branco com distribuição Gaussiana, média nula e variância unitária. O ruído  $e(k)$  é branco, com distribuição Gaussiana e com variância igual a 0,1. A Equação (4.1) foi utilizada para produzir os dados dinâmicos,  $Z^N$ , com  $N = 500$  e estáticos,  $\bar{Z}^M$ , com  $M = 11$ . A Figura 4.1 apresenta os sinais dinâmicos de entrada e saída do sistema (4.1). A característica estática desse sistema é mostrada na Figura 4.2.

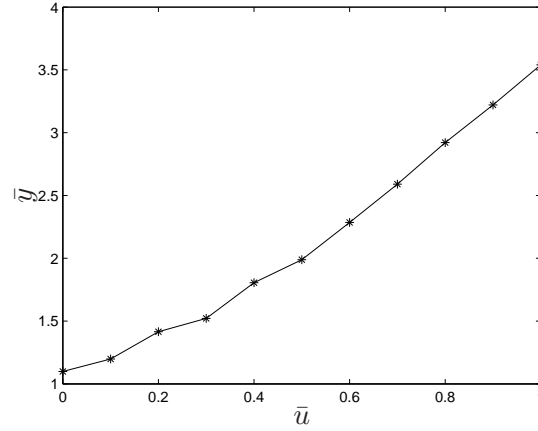


**Figura 4.1:** Dados dinâmicos de identificação e validação, Exemplo 1.

Primeira metade dados de identificação e segunda metade de validação, sendo: (a) sinal de entrada, (b) sinal de saída (da-dospiroddiNV01@).

Bonin e colegas (Bonin et al., 2010) avaliaram várias estruturas de modelos e sete foram previamente selecionadas pelo algoritmo de busca dos autores.

Questões referentes à seleção de estrutura propriamente dita não são abordadas neste trabalho. Aqui presume-se existir um conjunto de estruturas can-



**Figura 4.2:** Dados estáticos, Exemplo 1.

Característica estática do sistema representado pela Equação 4.1, com  $M = 11$  número de dados em estado estacionário (dadospiroddiNV01@).

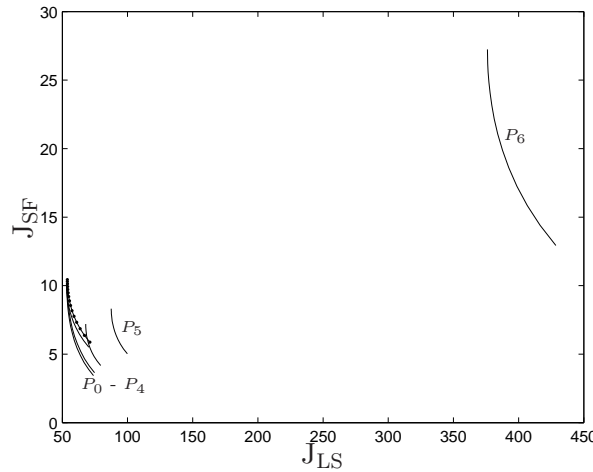
didatas oriundas de um procedimento de identificação de sistemas. Portanto, a objetivo é verificar a equivalência entre as estruturas candidatas e não a qualidade e o desempenho. As sete estruturas avaliadas são:

$$\begin{aligned}
 G_0 &: [1 \ y(k-1) \ u(k-2) \ u(k-1)^2 \ y(k-2)^2]; \\
 G_1 &: [1 \ y(k-1) \ u(k-2) \ u(k-1)^2 \ y(k-2)^2 \ \underline{y(k-2)} \ \underline{u(k-2)^2}]; \\
 G_2 &: [1 \ y(k-1) \ u(k-2) \ u(k-1)^2 \ y(k-2)^2 \ \underline{y(k-2)} \ \underline{u(k-1)} \ \underline{u(k-2)^2}]; \\
 G_3 &: [1 \ y(k-1) \ u(k-2) \ u(k-1)^2 \ y(k-2)^2 \ \underline{u(k-3)u(k-1)y(k-1)} \\
 &\quad \underline{u(k-1)u(k-3)^2}]; \\
 G_4 &: [1 \ y(k-1) \ u(k-2) \ u(k-1)^2 \ \underline{y(k-2)^3}]; \\
 G_5 &: [y(k-1) \ u(k-2) \ u(k-1)^2 \ \underline{u(k-1)u(k-3)} \ \underline{y(k-2)^3}]; \\
 G_6 &: [1 \ y(k-1) \ u(k-1)^2 \ \underline{y(k-2)^3} \ y(k-2)^2].
 \end{aligned}$$

Em situações práticas o modelo nominal  $G_0$  não é conhecido, no entanto, neste exemplo, ele será incluído na análise para fins de ilustração. Nas estruturas dos modelos acima, os termos sublinhados são regressores espúrios. Tais

regressores não são descartados após aplicar o teste *t de student* com um nível de significância de 95%.

A Figura 4.3 mostra os conjuntos Pareto dos modelos  $G_0$  até  $G_6$ . O conjunto Pareto correspondente ao modelo  $G_6$  está visivelmente mais afastado da origem, o que indica que a estrutura correspondente pode ser eliminada com segurança. Os conjuntos Pareto restantes são apresentados na Figura 4.4. Um ponto importante a observar na Figura 4.4 é que *alguns conjuntos Pareto se cruzam*, o que significa que nenhuma das duas estruturas é “absolutamente” melhor do que a outra no que diz respeito às funções de custo escolhidas. Em situações como essa o uso de conjuntos Pareto é muito informativo.



**Figura 4.3:** Conjuntos Pareto,  $P_0$  até  $P_6$ .

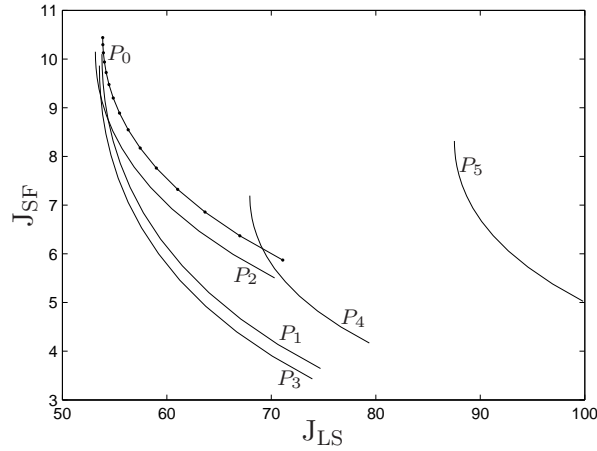
Representação dos conjuntos Pareto referente às estruturas  $G_0$  até  $G_6$ .

A seguir, o procedimento descrito no Capítulo 3 será aplicado. Como as estruturas dos modelos  $G_0$  até  $G_5$  demonstram ter desempenho semelhante para os dados disponíveis, estes serão tomados como o conjunto inicial.

Neste exemplo, as populações de conjuntos Pareto ótimos  $\mathcal{D}(P_j)$  foram obtidas por simulação de Monte Carlo. Um procedimento diferente será apresentado na Seção 4.3.

A partir de  $G_0$  (4.1), foi produzido um conjunto de 1000 séries de dados dinâmicos,  $Z_i^N$ ,  $i = 1 \dots 1000$  ( $N = 500$ ), e estáticos,  $\bar{Z}^M$ , com  $M = 11$ , usando diferentes realizações de ruído com  $\sigma^2 = 0,1$ .





**Figura 4.4:** Conjuntos Pareto,  $P_0$  até  $P_5$ .

*Asterisco* corresponde à estrutura nominal,  $G_0$ , e linha contínua às estruturas  $G_1$  até  $G_5$ .

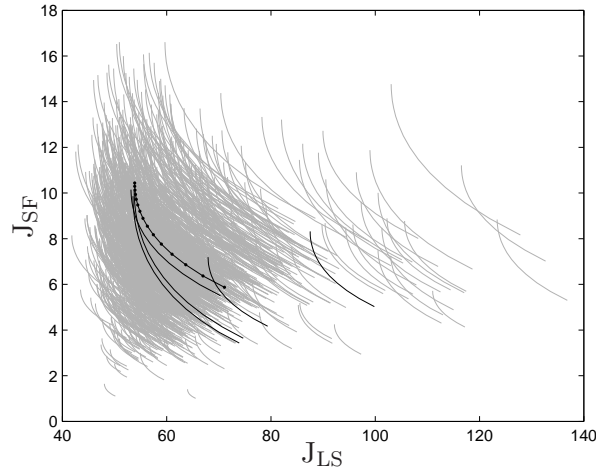
### Seleção do conjunto $\mathcal{G}^*$

A seguir passaremos pelos passos descritos na Seção 3.4.

*Passo 1.*  $\mathcal{D}(P_j)$ ,  $j = 0, \dots, 5$  é calculado usando a Equação (3.2). Para cada série, estimaram-se os parâmetros da estrutura nominal,  $G_0$ , utilizando-se o estimador bi-objetivo (ver Definição 2). As populações de conjuntos Pareto ótimos resultantes são apresentadas na Figura 4.5, onde a população  $\mathcal{D}(P_0)$  (Equação 3.2) aparece como a região sombreada. O mesmo procedimento é feito para as outras estruturas de modelos restantes. As populações  $\mathcal{D}(P_j)$ ,  $j = 0, \dots, 5$  são mostradas na Figura 4.6.

Claramente  $P_i \in \mathcal{D}(P_0)$ ,  $i = 0, 1, \dots, 4$ , e é possível que  $P_5 \in \mathcal{D}(P_0)$  (Figura 4.5). Observa-se claramente uma considerável sobreposição de  $\mathcal{D}(P_0)$  com  $\mathcal{D}(P_1)$ ,  $\mathcal{D}(P_2)$  e  $\mathcal{D}(P_3)$ , e alguma sobreposição de  $\mathcal{D}(P_0)$  com  $\mathcal{D}(P_4)$  e  $\mathcal{D}(P_5)$  (Figura 4.6). Para quantificar a equivalência aplicou-se o teste de hipótese (passos 2 e 3). Ressalta-se que as amostras têm distribuições não Gaussianas.

*Passo 2 até Passo 6.* Escolheu-se para todos os pares de estruturas de modelos  $N_A = N_B = 8$ ,  $N_h = 50$  com um nível de significância  $\alpha = 1\%$ . Depois de aplicar o teste de hipótese é possível computar o percentual de não rejeição, ou seja, percentual de vezes que a hipótese nula ( $H_0$ ) não foi rejeitada, ou  $A \equiv B$  (ver Seção 2.4). A força de ligação entre todos os pares de estruturas



**Figura 4.5:** População de conjuntos Pareto ótimos  $\mathcal{D}(P_0)$  e curvas de Pareto.

Os Paretos em cinza constituem  $\mathcal{D}(P_0)$ , *asterisco* corresponde ao Pareto  $P_0$  e linha contínua aos Paretos  $P_1$  até  $P_5$ .

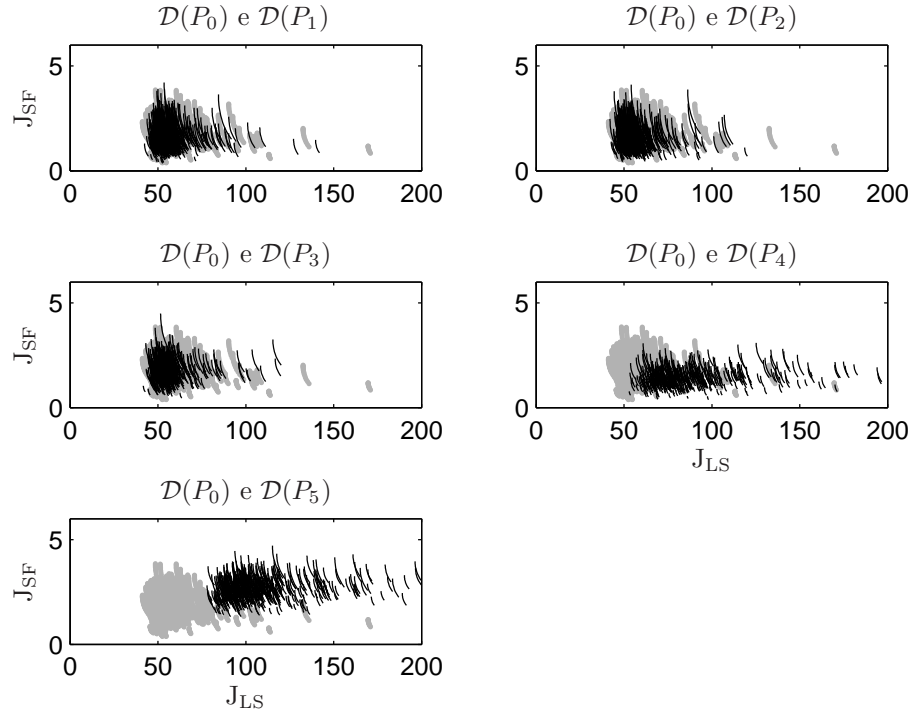
de modelos é ilustrada na Figura 4.7. A não rejeição entre o modelo nominal,  $G_0$ , e os demais foram:  $w_{0,1} = 93\%$ ,  $w_{0,2} = 92,2\%$ ,  $w_{0,3} = 91\%$ ,  $w_{0,4} = 16,5\%$  e  $w_{0,5} = 10\%$ .

*Passo 7.* Calculou-se o índice  $I$  (Equação 3.3) e obteve-se o seguinte resultado:  $I_0 = 60,7\%$ ,  $I_1 = 59,7\%$ ,  $I_2 = 60,3\%$ ,  $I_3 = 60,4\%$ ,  $I_4 = 28,5\%$  e  $I_5 = 24,6\%$ . Portanto, a estrutura de modelo  $G_0$  foi escolhida como a estrutura líder. Neste exemplo, a escolha confirma o fato de os dados terem sido gerados usando  $G_0$ . A força de conexão da estrutura líder com as demais estruturas de modelo é ilustrada na Figura 4.8.

*Passo 8.* Escolheu-se como parâmetro de corte  $\beta = 90\%$  e, portanto,  $\mathcal{G}^* = \{G_0, G_1, G_2, G_3\}$  é o subconjunto de estruturas que são  $\beta$ -equivalentes com um nível de significância  $\alpha = 1\%$  e com  $\beta = 90\%$ .

Deve-se notar que para qualquer valor de  $\beta$  na faixa de  $16,5\% < \beta < 91\%$  ter-se-ia o mesmo subconjunto  $\mathcal{G}^*$ .

Para esse exemplo  $N_m = 6$ ,  $N_r = 1000$  e  $N_h = 50$ . O tempo de simulação para realizar o teste, em um computador pessoal, foi de 4,9 segundos. Para  $N_h = 200$ , o resultado foi exatamente o mesmo com um tempo de 18,7 segundos.



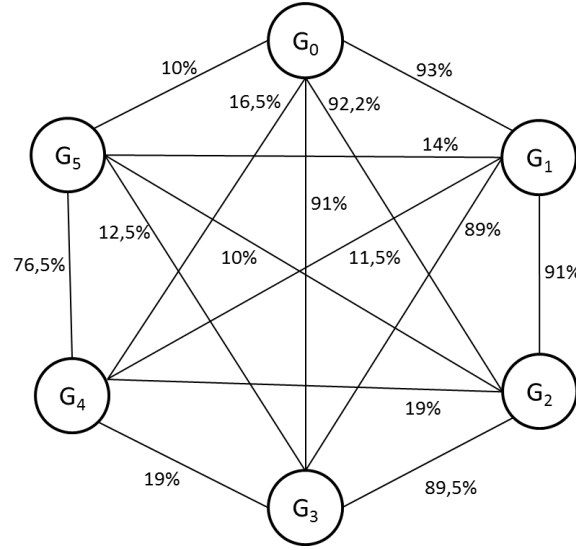
**Figura 4.6:** Relação entre populações de conjuntos Pareto ótimos,  $\mathcal{D}(P_0)$  com  $\mathcal{D}(P_1)$  até  $\mathcal{D}(P_5)$ .

Comparação das dimensões das populações de conjuntos Pareto ótimos.  $\mathcal{D}(P_0)$  (cinza) com  $\mathcal{D}(P_1)$  até  $\mathcal{D}(P_5)$  (preto).

### Verificação

Nesta seção comparou-se o desempenho dos seis modelos por meio dos dois testes descritos na Seção 3.5. Avaliaram-se as estruturas de modelos  $G_0$  até  $G_5$ . Os parâmetros foram obtidos utilizando o estimador bi-objetivo, seguido do decisor de mínima correlação (ver Definição 3, Seção 2.3.1). Também estimaram-se os parâmetros das estruturas por meio do estimador de mínimos quadrados estendido (MQE), nesse caso utilizam-se apenas os dados dinâmicos.

As respostas em malha aberta e em malha fechada são apresentadas nas Figuras 4.9 até 4.12. Para a resposta em malha fechada escolheu-se  $K_1 = 0$  (lei de controle puramente proporcional) e  $K_2 = 0,2$ . O sinal de referência foi alterado, de 0 para 0,5 na iteração  $k = 50$ , como mostrado nas Figuras 4.11 e 4.12. As condições iniciais são  $u(i) = 0$ ,  $i = -2, -1, 0$  e  $y(i) = 0,5$ ,  $i = -1, 0$ .



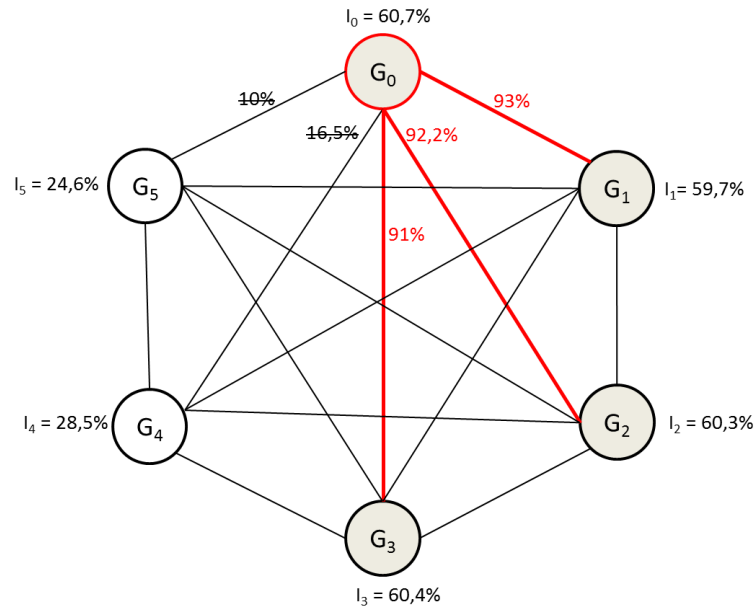
**Figura 4.7:** Força de conexão entre estruturas de modelos, Exemplo 1.

Percentual de vezes que  $H_0$  não foi rejeitada entre todos os pares de estruturas de modelos.

O desempenho do controlador não foi levado em consideração. Nesta seção, o objetivo é verificar a similaridade dos comportamentos dos modelos. Entretanto, o erro em estado estacionário, apresentado na resposta em malha fechada, é justificado uma vez que o controlador é puramente proporcional.

Todos os testes mostram que o modelo com a estrutura  $G_5$  apresentou um comportamento mais distante de  $G_0$ . O modelo com estrutura  $G_4$  é visivelmente mais próximo dos modelos com estrutura em  $\mathcal{G}^*$  que o modelo com estrutura  $G_5$ , embora alguma diferença possa ser vista na Figura 4.9 (menor sobressinal). Quando os parâmetros foram estimados usando MQE, e nesse caso utilizam-se apenas os dados dinâmicos, o comportamento de  $G_4$  e  $G_5$ , tanto na resposta em malha aberta (Figura 4.10) quanto na resposta em malha fechada (Figura 4.12), apresentaram diferenças ainda maiores.

Uma vez que o desempenho dos modelos com estrutura em  $\mathcal{G}^*$  não são idênticos, todos os elementos de  $\mathcal{G}^*$  devem ser utilizados no projeto de um controlador, ao menos para determinar limites do comportamento dinâmico do sistema controlado. Outra utilização é no projeto do tipo *minmax* que destina-se a encontrar o melhor desempenho no pior caso. Esses procedimentos seriam mais significativos do que a aplicação simples do modelo nominal no projeto



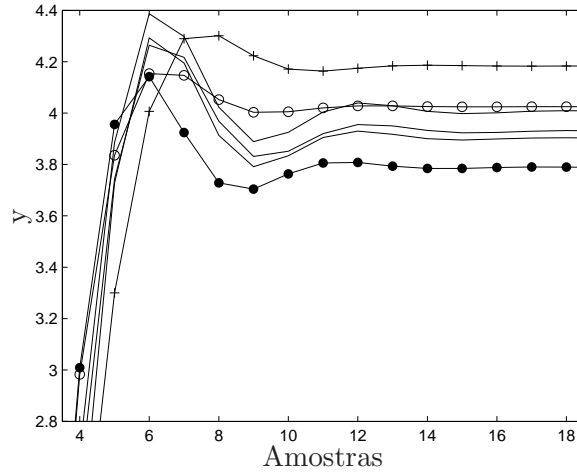
**Figura 4.8:** Força de conexão entre estruturas de modelos com o modelo líder, Exemplo 1.

Índices  $I$  das estruturas de modelos avaliadas. Destaque para a estrutura de maior índice e o percentual de não rejeição entre o modelo líder e os demais.

do controlador.

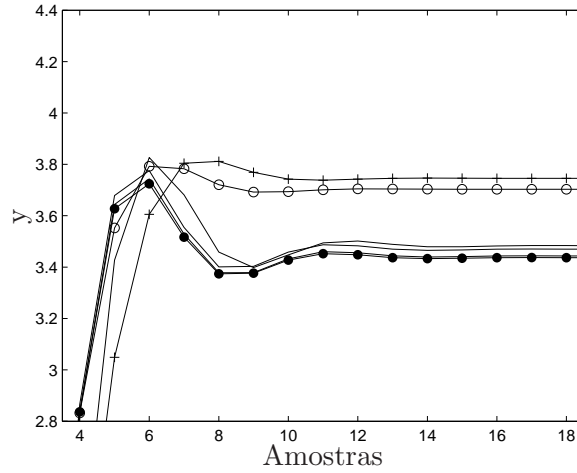
### 4.2.2 Exemplo 2

O segundo exemplo é um processo com atraso puro de tempo e maior número de termos. Os parâmetros têm ordens de grandezas diferentes. O mesmo exemplo foi usado em um teste de seleção de estruturas pelo próprio autor e, para o conjunto de dados simulados, não foi possível, com a técnica de otimização bi-Objetivo para a determinação de estrutura utilizada (Barbosa et al., 2010), reproduzir o modelo original.



**Figura 4.9:** Resposta em malha aberta, Exemplo 1.

(—●) modelo com estrutura nominal,  $G_0$ , (—) modelos  $G_1$  até  $G_3$ , (—○) modelo  $G_4$  e (—+) modelo  $G_5$ .

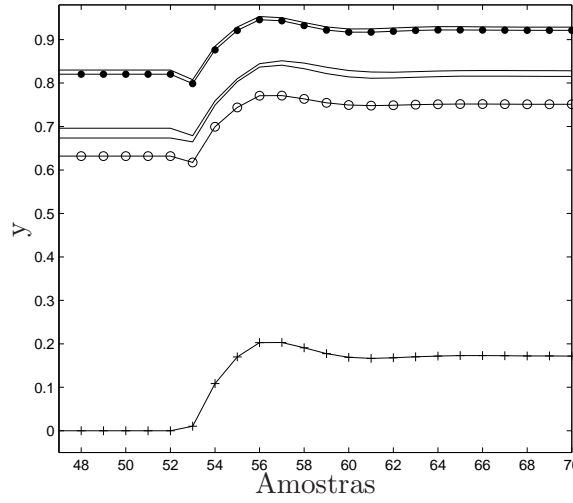


**Figura 4.10:** Resposta em malha aberta, Exemplo 1, parâmetros estimados por MQE.

(—●) modelo com estrutura nominal,  $G_0$ , (—) modelos  $G_1$  até  $G_3$ , (—○) modelo  $G_4$  e (—+) modelo  $G_5$ . MQE com modelo de ruído  $e(k) = c_1\nu(k-1) + c_2\nu(k-2) + \nu(k)$  e dez iterações.

Considere o seguinte sistema:

$$\begin{aligned}
 y(k) &= 1,4945y(k-1) - 6,4596 \times 10^{-1}y(k-2) + 3,4655 \times 10^{-1}u(k-4) \\
 &\quad - 2,0921 \times 10^{-1}u(k-4)^2 + 3,5457 \times 10^{-2}u(k-5)u(k-4) - 2,2714 \times 10^{-2} \\
 &\quad - 1,7591 \times 10^{-2}u(k-5)y(k-3) + 9,8811 \times 10^{-3}y(k-3)^2 + e(k), \\
 u(k) &= 0,9u(k-1) + w(k),
 \end{aligned} \tag{4.2}$$



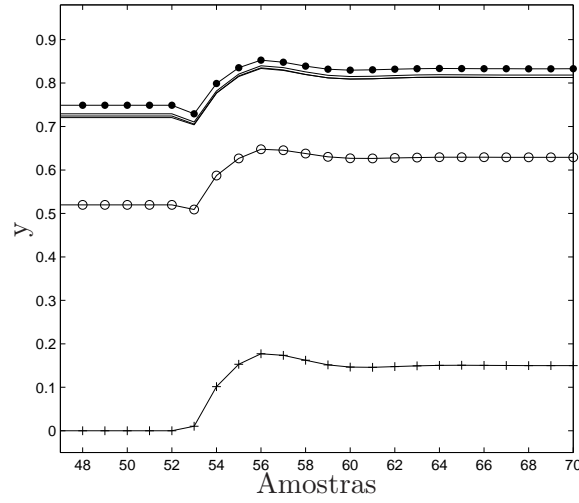
**Figura 4.11:** Resposta em malha fechada, Exemplo 1.

( $\bullet$ ) modelo com estrutura nominal,  $G_0$ , (—) modelos  $G_1$  até  $G_3$ , ( $\circ$ ) modelo  $G_4$  e ( $+$ ) modelo  $G_5$ .

em que  $w(\cdot)$  é ruído branco com distribuição uniforme, média nula e variância 0,08. O sinal  $e(k)$  é ruído branco com distribuição gaussiana, média nula e variância igual a 0,16.

A Equação (4.2) foi utilizada para produzir os dados dinâmicos,  $Z^N$ , com  $N = 500$  e estáticos,  $\bar{Z}^M$ , com  $M = 11$ . A Figura 4.13 apresenta os sinais dinâmicos de entrada e saída do sistema. A característica estática do sistema é mostrada na Figura 4.14.

Quatro estruturas de modelos foram avaliadas. A estrutura de modelo com estrutura nominal ( $G_0$ , Equação 4.2),  $G_1$ ,  $G_2$  e  $G_3$ . As estruturas foram obtidas utilizando taxa de redução de erro (Billings et al., 1989) e são apresentadas a seguir. Assim como no exemplo anterior, o modelo com estrutura nominal  $G_0$  é conhecido e será incluído para efeitos de comparação.



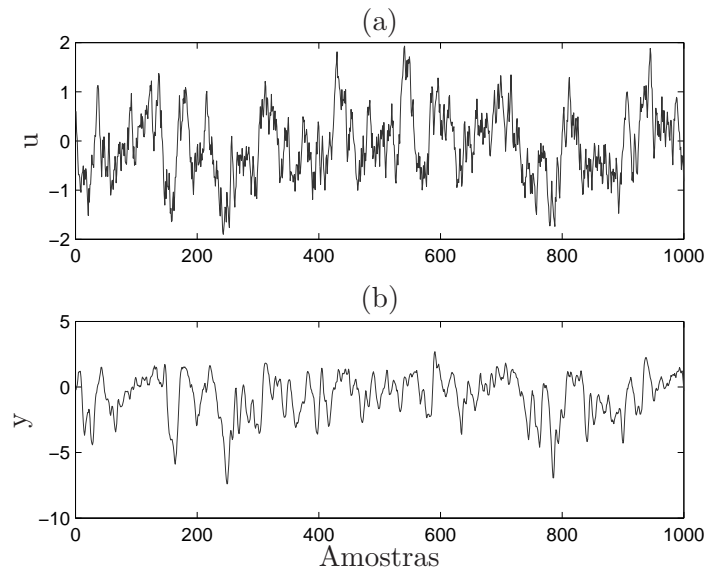
**Figura 4.12:** Resposta em malha fechada, Exemplo 1, parâmetros estimados por MQE.

(—●) modelo com estrutura nominal,  $G_0$ , (—) modelos  $G_1$  até  $G_3$ , (—○) modelo  $G_4$  e (—+) modelo  $G_5$ . MQE com modelo de ruído  $e(k) = c_1\nu(k-1) + c_2\nu(k-2) + \nu(k)$  e dez iterações.

$$\begin{aligned}
 G_0 &: [1 \quad y(k-1) \quad y(k-2) \quad u(k-4) \quad u(k-4)^2 \\
 &\quad u(k-5)u(k-4) \quad u(k-5)y(k-3) \quad y(k-3)^2] \\
 G_1 &: [y(k-1) \quad y(k-2) \quad u(k-4) \quad u(k-4)^2 \quad \underline{u(k-5)} \\
 &\quad \underline{u(k-1)u(k-2)} \quad \underline{u(k-1)u(k-6)} \quad \underline{u(k-3)^2} \\
 &\quad \underline{u(k-1)u(k-3)y(k-1)} \quad \underline{u(k-4)u(k-6)y(k-1)}] \\
 G_2 &: [y(k-1) \quad y(k-2) \quad u(k-4) \quad u(k-4)^2 \\
 &\quad \underline{u(k-4)y(k-1)} \quad \underline{u(k-5)u(k-4)}] \\
 G_3 &: [y(k-1) \quad y(k-2) \quad u(k-4) \quad u(k-4)^2 \\
 &\quad \underline{u(k-4)y(k-1)} \quad \underline{u(k-5)u(k-4)} \quad \underline{u(k-6)y(k-1)} \\
 &\quad \underline{u(k-4)^3} \quad \underline{u(k-5)^2u(k-4)} \quad \underline{u(k-5)}]
 \end{aligned}$$

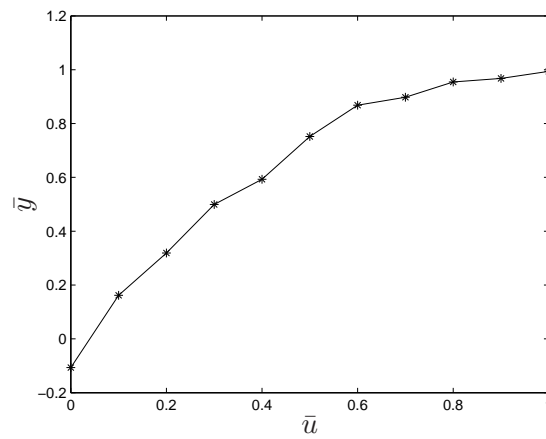
A Figura 4.15 mostra as curvas Pareto dos modelos  $G_0$  até  $G_3$ . Diferente





**Figura 4.13:** Dados dinâmicos de identificação e validação, Exemplo 2.

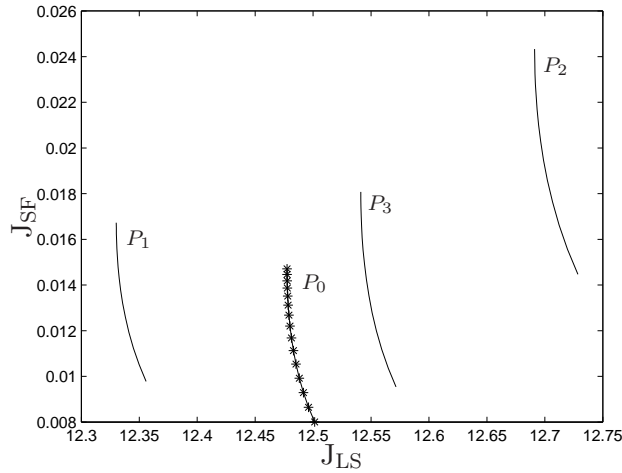
Primeira metade dados de identificação e segunda metade de validação, sendo: (a) sinal de entrada do sistema, (b) sinal de saída do sistema, (dadosaguirre@).



**Figura 4.14:** Dados estáticos, Exemplo 2.

Característica estática do sistema representado pela Equação 4.2, com  $M = 11$  número de dados em estado estacionário (dadosaguirre@).

do exemplo anterior as curvas não se cruzam. Entretanto, a curva referente ao modelo com estrutura nominal não é dominante e encontra-se entre as demais curvas.



**Figura 4.15:** Conjuntos Pareto,  $P_0$  até  $P_3$ .

Representação dos conjuntos Pareto referente aos modelos  $G_0$  até  $G_3$ .

A seguir, o procedimento descrito no Capítulo 3 será aplicado. O conjunto inicial é formado por  $G_0$  até  $G_3$  que apresentaram curvas Pareto próximas.

Assim como no exemplo anterior, para calcular a população de conjuntos Pareto ótimos  $\mathcal{D}$ , no Passo 1, utilizou-se simulação Monte Carlo.

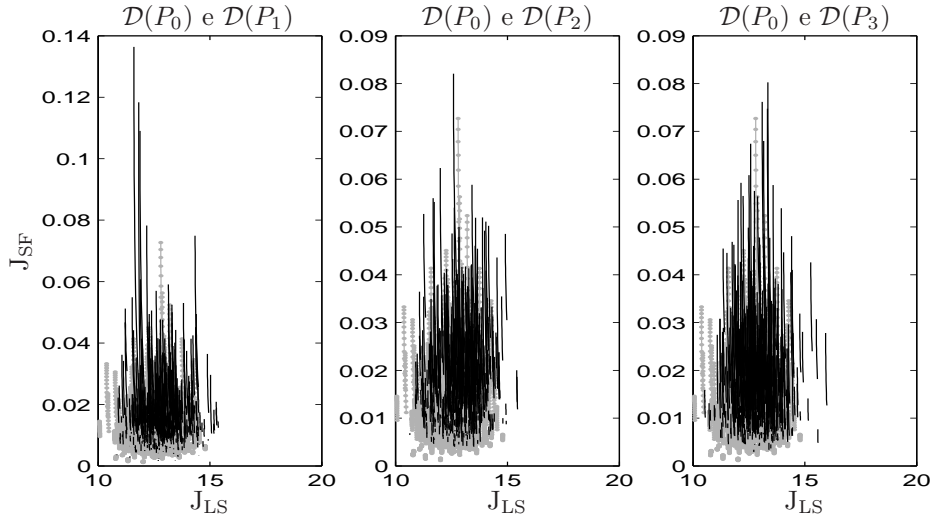
### Seleção do conjunto $\mathcal{G}^*$

A partir de  $G_0$ , foi produzida um conjunto de 500 séries de dados dinâmicos,  $Z^N$ , e estáticos,  $\bar{Z}^M$ , com  $N = 500$  e  $M = 11$ , respectivamente, usando diferentes realizações de ruído com  $\sigma^2 = 0,16$ .

*Passo 1.*  $\mathcal{D}(P_j)$ ,  $j = 0, \dots, 3$  é calculado usando a Equação (3.2). Para cada série, estimaram-se os parâmetros da estrutura nominal,  $G_0$ , utilizando-se o estimador bi-objetivo (ver Definição 2). As populações de conjuntos Pareto ótimos resultantes são apresentadas na Figura 4.16, onde a população  $\mathcal{D}(P_0)$  (Equação 3.2) aparece como a região sombreada. O mesmo procedimento é feito para as outras estruturas de modelos restantes. As populações de con-

juntos Pareto ótimos  $\mathcal{D}(P_j)$ ,  $j = 0, \dots, 3$  são mostradas na Figura 4.16.

Claramente  $P_i \in \mathcal{D}(P_0)$ ,  $i = 0, \dots, 3$ . Entretanto, para confirmar isto, utilizou-se o teste de hipótese (passos 2 e 3). Ressalta-se que as amostras não têm distribuições Gaussianas.



**Figura 4.16:** Relação entre populações de conjuntos Pareto ótimos,  $\mathcal{D}(P_0)$  com  $\mathcal{D}(P_1)$  até  $\mathcal{D}(P_3)$ .

Comparação das dimensões das populações de conjuntos Pareto ótimos.  $\mathcal{D}(P_0)$  (cinza) com  $\mathcal{D}(P_1)$  até  $\mathcal{D}(P_3)$  (preto). Exemplo 2.

*Passo 2 até Passo 6.* Assim como no exemplo anterior, escolheu-se para todos os pares de estruturas de modelos  $N_A = N_B = 8$ ,  $N_h = 50$  com um nível de significância  $\alpha = 1\%$ . Aplicou-se o teste de hipótese e a não rejeição entre o modelo com estrutura nominal,  $G_0$ , e os demais foram:  $w_{0,1} = 79,5\%$ ,  $w_{0,2} = 76,5\%$  e  $w_{0,3} = 65,55\%$ .

*Passo 7.* Calculou-se o índice  $I$  (Equação 3.3) e obteve-se o seguinte resultado:  $I_0 = 73,8\%$ ,  $I_1 = 72,5\%$ ,  $I_2 = 72,0\%$  e  $I_3 = 63,0\%$ . Portanto, a estrutura de modelo  $G_0$  foi escolhida como a estrutura de modelo líder, confirmando o fato de os dados terem sido obtidos a partir de  $G_0$ .

*Passo 8.* Para um parâmetro de corte  $\beta = 90\%$ , tem-se  $\mathcal{G}^* = \{\}$ , ou seja, o subconjunto de estruturas que são  $\beta$ -equivalentes é vazio. Cabe ao usuário definir se é pertinente ajustar o parâmetro  $\beta$ . Para um parâmetro de corte  $\beta = 65\%$ , tem-se  $\mathcal{G}^* = \{G_0, G_1, G_2, G_3\}$  como o subconjunto de estruturas que

são  $\beta$ -equivalentes com um nível de significância  $\alpha = 1\%$  e com  $\beta = 65\%$ .

Deve-se notar que valores de  $\beta$  na faixa de  $65,5\% < \beta < 76,5\%$  resultariam em um subconjunto  $\mathcal{G}^*$  diferente, sem  $G_3$ , que teve uma conexão com  $G_0$  de  $65,5\%$ .

### Verificação

Para efeitos de verificação, comparou-se o desempenho dos quatro modelos por meio dos dois testes descritos na Seção 3.5. Avaliaram-se as estruturas de modelo  $G_0$  até  $G_3$ . Os parâmetros foram obtidos utilizando o estimador bi-objetivo, seguido do decisor de mínima correlação (ver Definição 3, Seção 2.3.1).

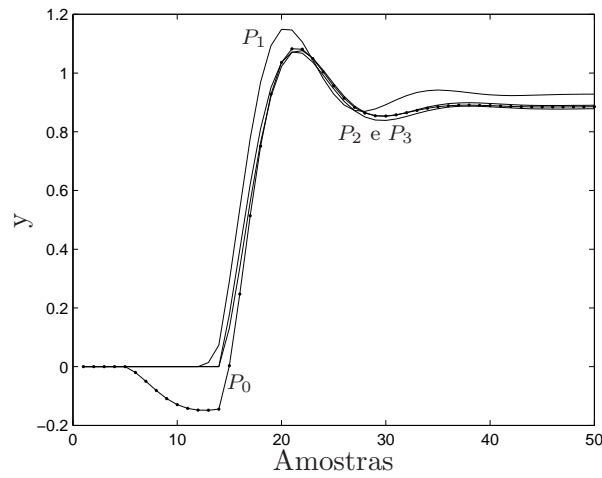
As respostas em malha aberta e em malha fechada são apresentadas nas Figuras 4.17 e 4.18. Para a resposta em malha fechada escolheu-se  $K_1 = 1$  (lei de controle com integração) e  $K_2 = 0,08$ . O sinal de referência foi ajustado, de 0 para 0,5 na iteração  $k = 250$ . As condições iniciais foram nulas em  $u$  e 0,5 em  $y$ .

Os dois testes mostraram comportamentos próximos dos modelos avaliados. O modelo  $G_2$  merece um destaque. Embora  $G_2$  tenha apenas 6 termos, 5 genuínos e 1 espúrio, seu desempenho em malha aberta foi muito similar ao do modelo com estrutura nominal.

### 4.2.3 Influência do atraso puro de tempo

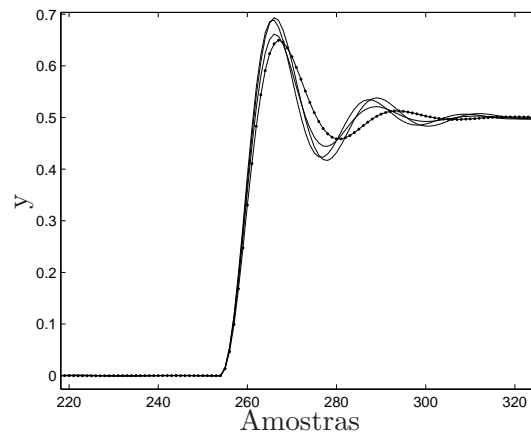
No exemplo anterior foi utilizado um modelo com atraso puro de tempo ( $d = 4$ ) e parâmetros de diferentes ordens de grandeza. Com um parâmetro de corte  $\beta = 90\%$ , a metodologia aplicada apontou um  $\mathcal{G}^* = \{\}$ . Para um parâmetro de corte  $\beta = 65\%$ , obteve-se  $\mathcal{G}^* = \{G_0, G_1, G_2, G_3\}$  como o subconjunto de estruturas que são  $\beta$ -equivalentes com um nível de significância  $\alpha = 1\%$ . Tal resultado aponta como sendo  $\beta$ -equivalentes estruturas com atraso puro de tempo diferentes, por exemplo  $G_0$  com  $d = 4$  e  $G_1$  com  $d = 1$ .

A estimação do atraso puro de tempo é tema de pesquisa de vários pesquisadores (Drakunov et al., 2006; Waschburger e Galvão, 2013; Zheng et al., 2013). No procedimento proposto considerou-se conhecido o conjunto inicial  $\mathcal{G}$  de  $N_m$  estruturas candidatas. Tal conjunto é obtido previamente de um



**Figura 4.17:** Resposta em malha aberta, Exemplo 2.

(-·-) modelo com estrutura nominal,  $G_0$ , (—) modelos  $G_1$  até  $G_3$ .



**Figura 4.18:** Resposta em malha fechada, Exemplo 2.

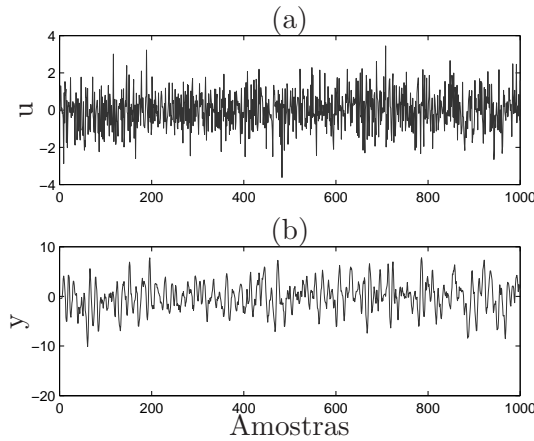
(-·-) modelo com estrutura nominal,  $G_0$ , (—) modelos  $G_1$  até  $G_3$ .

processo de identificação de sistemas, e pode conter elementos com atrasos de tempo diferentes. Portanto, o objetivo do exemplo é discutir o efeito de incluir no conjunto inicial  $\mathcal{G}$  estruturas com atrasos de tempo distintos.

Considere o seguinte sistema:

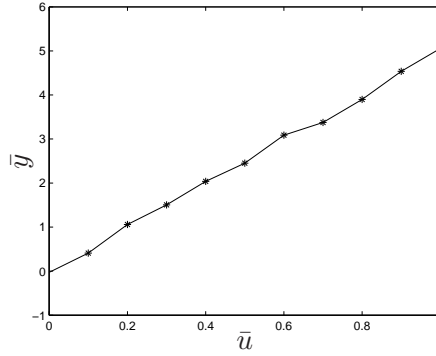
$$y(k) = 1,5y(k-1) - 0,7y(k-2) + 1,0u(k-5) + e(k), \quad (4.3)$$

em que a entrada  $u(\cdot)$  é um sinal do tipo ruído branco com distribuição Gaussiana, média nula e variância unitária. O ruído  $e(k)$  é branco, com distribuição Gaussiana e com variância igual a 0,2. A Equação (4.3) foi utilizada para produzir os dados dinâmicos,  $Z^N$ , com  $N = 500$  e estáticos,  $\bar{Z}^M$ , com  $M = 11$ . A Figura 4.19 apresenta os sinais dinâmicos de entrada e saída do sistema (4.3). A característica estática desse sistema é mostrada na Figura 4.20.



**Figura 4.19:** Dados dinâmicos de identificação e validação, Equação (4.3).

Primeira metade dados de identificação e segunda metade de validação, sendo: (a) sinal de entrada, (b) sinal de saída (dadosatr@).



**Figura 4.20:** Dados estáticos obtidos a partir da Equação (4.3).

Característica estática do sistema representado pela Equação 4.3, com  $M = 11$  número de dados em estado estacionário (`dadosatr@`).

Com o intuito de verificar exclusivamente o atraso puro de tempo, outras seis estruturas foram especificadas para compor o conjunto inicial  $\mathcal{G}$  de estruturas candidatas. Das sete estruturas candidatas,  $G_5$  é a estrutura nominal (Equação 4.3), que gerou os dados, as demais estruturas têm os mesmos agrupamentos, mesmo número de termos, diferenciando apenas do atraso puro de tempo, conforme mostrado a seguir.

$$\begin{aligned}
 G_1 &: [y(k-1) \ y(k-2) \ \underline{u(k-1)}]; \\
 G_2 &: [y(k-1) \ y(k-2) \ \underline{u(k-2)}]; \\
 G_3 &: [y(k-1) \ y(k-2) \ \underline{u(k-3)}]; \\
 G_4 &: [y(k-1) \ y(k-2) \ \underline{u(k-4)}]; \\
 G_5 &: [y(k-1) \ y(k-2) \ \underline{u(k-5)}]; \\
 G_6 &: [y(k-1) \ y(k-2) \ \underline{u(k-6)}]; \\
 G_7 &: [y(k-1) \ y(k-2) \ \underline{u(k-7)}].
 \end{aligned}$$

Nas estruturas dos modelos acima, os termos sublinhados são regressores espúrios, com atraso puro de tempo  $1 \leq d \leq 7$ . As estruturas dos modelos  $G_1$  até  $G_7$  foram tomadas como o conjunto inicial e obtiveram-se as populações

de conjuntos Pareto ótimos  $\mathcal{D}(P_j)$  para cada estrutura, via simulação Monte Carlo.

As populações de conjuntos Pareto ótimos  $\mathcal{D}(P_j)$ ,  $j = 1, \dots, 7$  foram obtidas por meio da Equação 3.2 e são ilustradas na Figura 4.21, em que a região sombreada é a população  $\mathcal{D}(P_5)$ .

Por meio da Figura 4.21 é possível observar claramente que  $\mathcal{D}(P_5) \cap \mathcal{D}(P_j) = \{\}$ ,  $j = 1, \dots, 7$  e  $j \neq 5$ . Nenhuma sobreposição de  $\mathcal{D}(P_5)$  com as demais regiões é constatada, e fica claro que o atraso de tempo incorreto resulta no afastamento das populações de conjuntos Pareto da origem. Por outro lado, do ponto de vista do custo estático, nota-se que  $J_{SF}$  é insensível ao atraso puro de tempo.

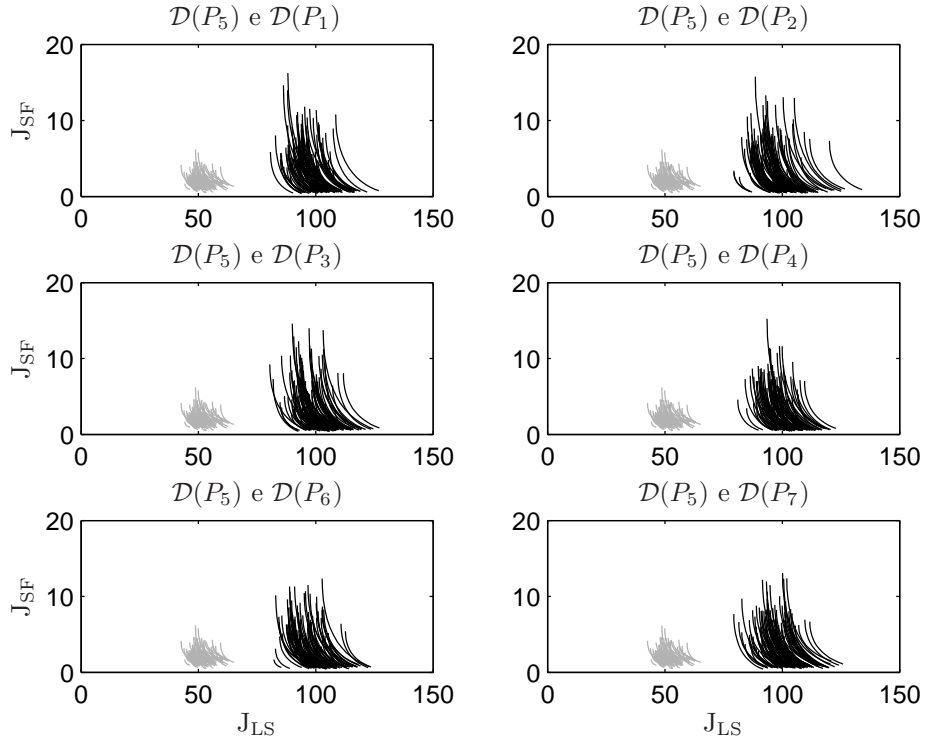
Entretanto, como pode-se observar (Figura 4.21), as populações  $\mathcal{D}(P_j)$ ,  $j = 1, \dots, 7$  e  $j \neq 5$ , encontram-se agrupadas em uma mesma região do gráfico (com a média de  $J_{LS} \simeq 100$  e a média de  $J_{SF} \simeq 10$ ). Tal efeito implica uma forte ligação ( $w_{jk}$ ) entre as estruturas de modelos  $G_j$ ,  $j = 1, \dots, 7$  e  $j \neq 5$ . Consequentemente o modelo nominal não é escolhido como modelo *líder* e ainda apresenta uma fraca conexão com o modelo *líder*. Tal efeito resultou em  $\mathcal{G}^* = \{G_1, G_2, G_3, G_4, G_6, G_7, \}$ , para um  $\beta = 90\%$ .

Os resultados deste exemplo mostraram que, em um conjunto de estruturas candidatas igualmente incorretas, o procedimento indica como equivalentes tais estruturas. O método não faz distinção, de forma direta, à seleção de estrutura propriamente. O resultado do procedimento é um conjunto de estruturas equivalentes e questões como qualidade e desempenho não são levadas em conta.

### 4.3 Exemplo experimental

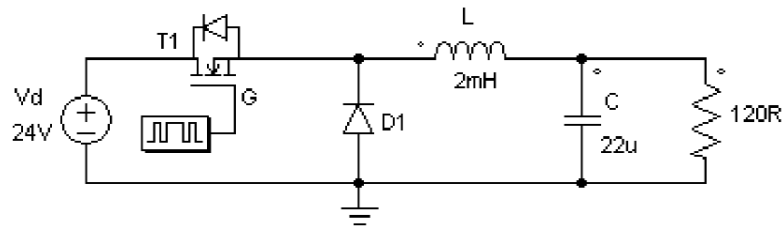
Nesta seção um exemplo experimental é apresentado. O procedimento foi usado para obter modelos para um conversor estático CC-CC do tipo Buck. O objetivo é analisar o desempenho do método em sistemas nos quais a estrutura correta é desconhecida. O sistema estudado é descrito brevemente, apresentando os testes estáticos e dinâmicos utilizados para obtenção dos dados. Considere o conversor conforme mostrado na Figura 4.22 (Aguirre et al., 2000; Corrêa et al., 2002; Barroso e Nepomuceno, 2004; Corrêa e Aguirre, 2004; Martins et al., 2009).





**Figura 4.21:** Relação entre populações de conjuntos Pareto ótimos,  $\mathcal{D}(P_5)$  com  $\mathcal{D}(P_1)$  até  $\mathcal{D}(P_7)$ .

Comparação das dimensões das populações de conjuntos Pareto ótimos.  $\mathcal{D}(P_5)$  (cinza) com  $\mathcal{D}(P_1)$  até  $\mathcal{D}(P_7)$  (preto).



**Figura 4.22:** Conversor Buck.

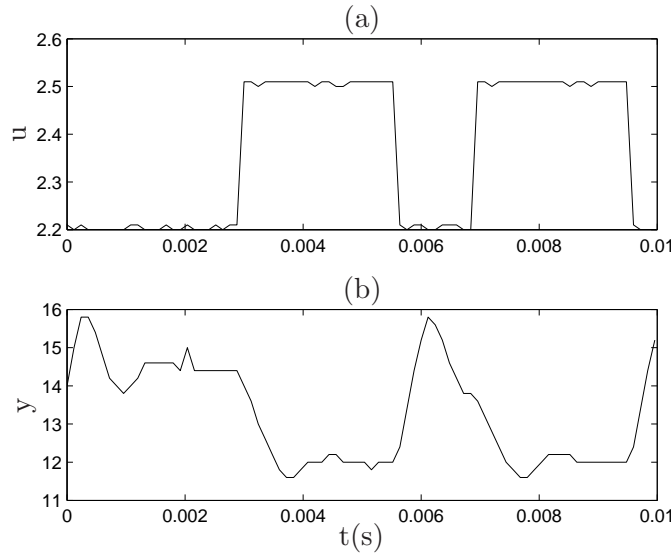
Estrutura de um conversor CC-CC do tipo Buck.

Os dados dinâmicos  $Z^N$  de identificação (Figura 4.23) e validação (Figura 4.24) utilizados são os mesmos considerados em Aguirre et al. (2000), com  $N = 2000$  e decimados por um fator de 12.

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.

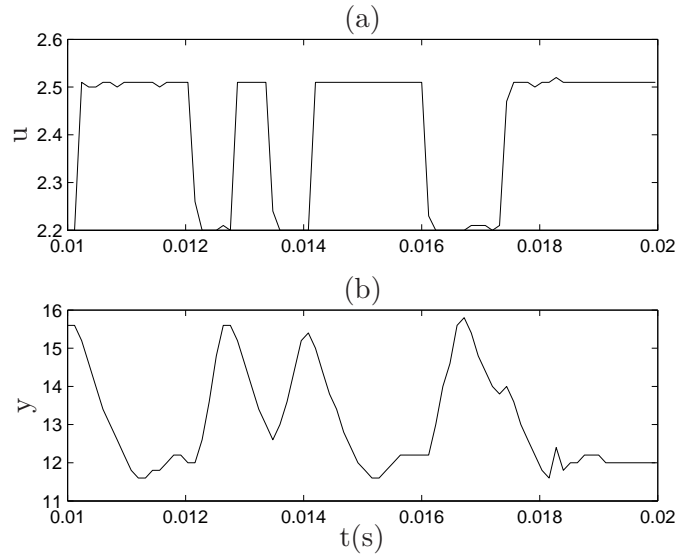
Quando o ciclo de trabalho tende à unidade, a corrente através do indutor e  $V_o$  aumentam, pois a fonte  $V_d$  energiza a malha formada por ela, o capacitor e o indutor. Quando  $D$  tende a zero,  $V_o$  diminui com um regime dinâmico diferente. Esse fato caracteriza um regime dinâmico não linear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



**Figura 4.23:** 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@**).

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



**Figura 4.24:** Dados de validação – conversor Buck.

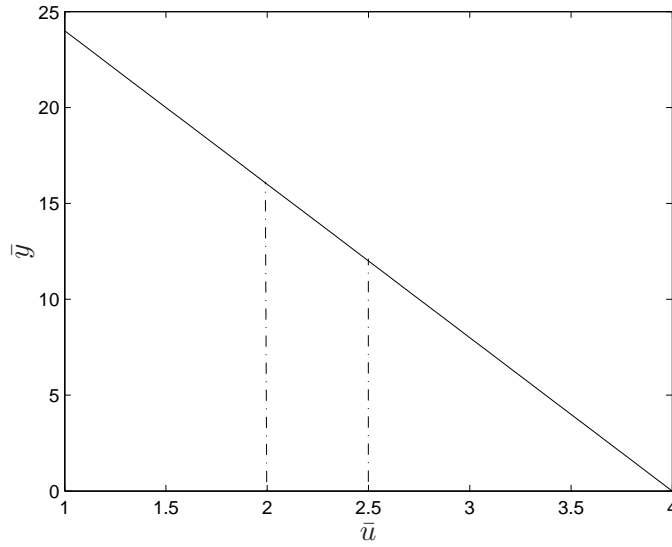
Dados de validação do conversor Buck, sendo: (a) sinal de entrada do sistema, (b) sinal de saída do sistema, (buck@).

$$\begin{aligned}
 V_o &= DV_d \\
 &= (1 - D')V_d = \left(1 - \frac{\bar{u} - 1}{3}\right) V_d \\
 &= \frac{4V_d}{3} - \frac{V_d}{3}\bar{u},
 \end{aligned} \tag{4.4}$$

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 4.4 valores de  $\bar{u}$  que abranjam todos os pontos de operação do sistema (1 a 4V). No exemplo  $\bar{Z}^M$  com  $M = 11$  foi obtido a partir de (4.4). A curva estática para o conversor Buck pode ser vista na Figura 4.25. Os dados dinâmicos são limitados a uma estreita faixa (12 V a 16 V aproximadamente) em comparação aos possíveis pontos de operação, como destacado na característica estática.

Baseado na identificação mono-objetivo (estimador de mínimos quadrados), treze potenciais modelos foram selecionados. Tais modelos foram obtidos uti-

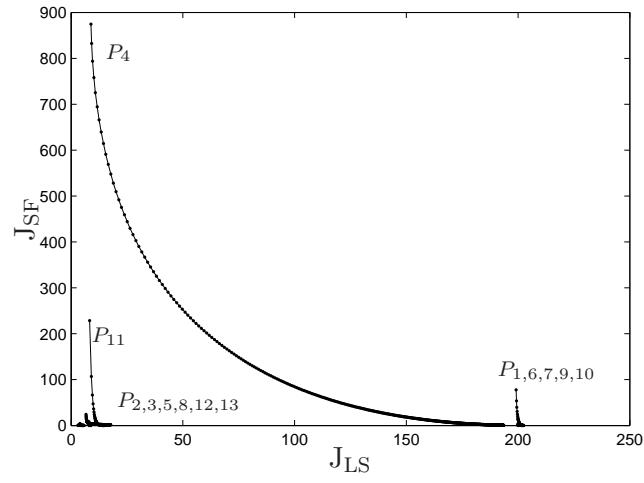


**Figura 4.25:** Dados estáticos: conversor Buck.

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

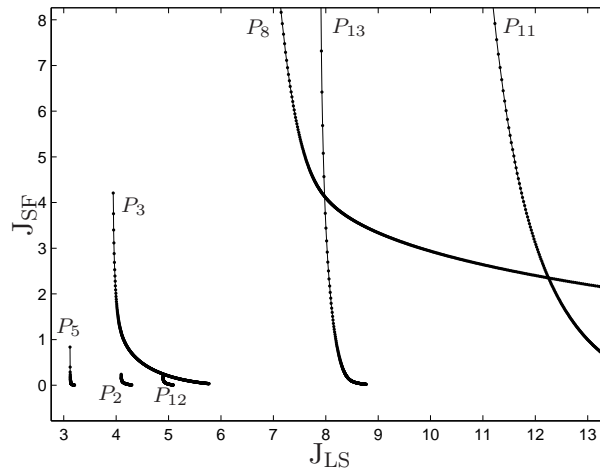
lizando taxa de redução de erro e, em alguns casos, coincidem com os modelos selecionados em outros trabalhos publicados (Martins et al., 2009; Barroso e Nepomuceno, 2004; Corrêa e Aguirre, 2004; Corrêa et al., 2002; Aguirre et al., 2000). As estruturas correspondentes aos modelos  $\mathcal{G} = \{G_1 \dots G_{13}\}$  ( $N = 13$ ) e suas respectivas curvas Pareto estão representadas na Figura 4.26. A Figura 4.27 mostra com detalhes as curvas Pareto mais próximas da origem. É possível observar a interseção entre várias curvas Pareto.

Devido ao fato deste exemplo ser experimental, utilizou-se um único conjunto de dados, dividido em janelas, conforme ilustrado na Figura 4.28. Os dados originais com 168 amostras foi dividido em 84 janelas com 84 amostras sobrepostas. Assim,  $Z_i^N$ ,  $i = 1 \dots 84$  ( $N = 84$ ). Para cada estrutura de modelo, determinaram-se 84 conjuntos de Pareto. Estes compõem a correspondente população de conjuntos Pareto  $\mathcal{D}(P)$ .



**Figura 4.26:** Conjuntos Pareto,  $P_1$  a  $P_{13}$ .

Representação dos conjuntos Pareto referentes aos modelos  $P_1$  a  $P_{13}$ .



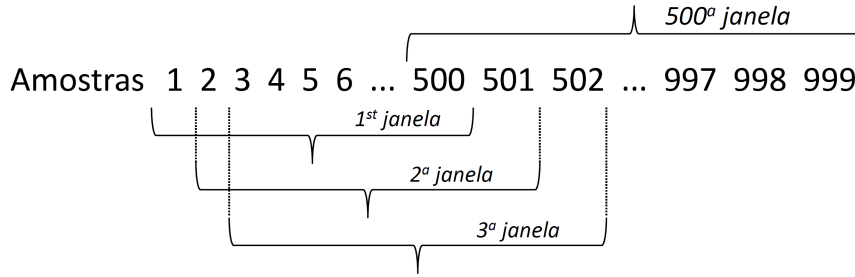
**Figura 4.27:** Conjuntos Pareto, detalhe da Figura 4.27.

Representação dos conjuntos Pareto mais próximos da origem.

### Seleção do conjunto $\mathcal{G}^*$

*Passo 1.*  $\mathcal{D}(P_j)$ ,  $j = 1, 2, \dots, 13$ , é calculado usando a Equação (3.2).

*Passo 2 até Passo 6.* Assim como nos exemplos anteriores, escolheu-se para



**Figura 4.28:** Janela deslizante, representação hipotética.

Representação de uma massa de dados dividida em 500 janelas contendo 500 amostras cada.

todos os pares de estruturas de modelos  $N_A = N_B = 8$ ,  $N_h = 50$  com um nível de significância  $\alpha = 1\%$ . Depois de aplicar o teste de hipótese computou-se o percentual de não rejeição  $w_{jk}$ .

*Passo 7.* Calculou-se o índice  $I$  (Equação 3.3) e obteve-se o seguinte resultado:  $I_1 = 11,7\%$ ,  $I_2 = 44,7\%$ ,  $I_3 = 44,4\%$ ,  $I_4 = 15,4\%$ ,  $I_5 = 21,8\%$ ,  $I_6 = 10,1\%$ ,  $I_7 = 14,0\%$ ,  $I_8 = 33,9\%$ ,  $I_9 = 15,5\%$ ,  $I_{10} = 23,5\%$ ,  $I_{11} = 37,2\%$ ,  $I_{12} = 48,6\%$  e  $I_{13} = 35,4\%$ . Portanto, escolheu-se a estrutura de modelo  $G_{12}$  como a estrutura de modelo líder. Os índices e a força de conexão da estrutura líder com as demais estruturas de modelo é ilustrada na Figura 4.29.

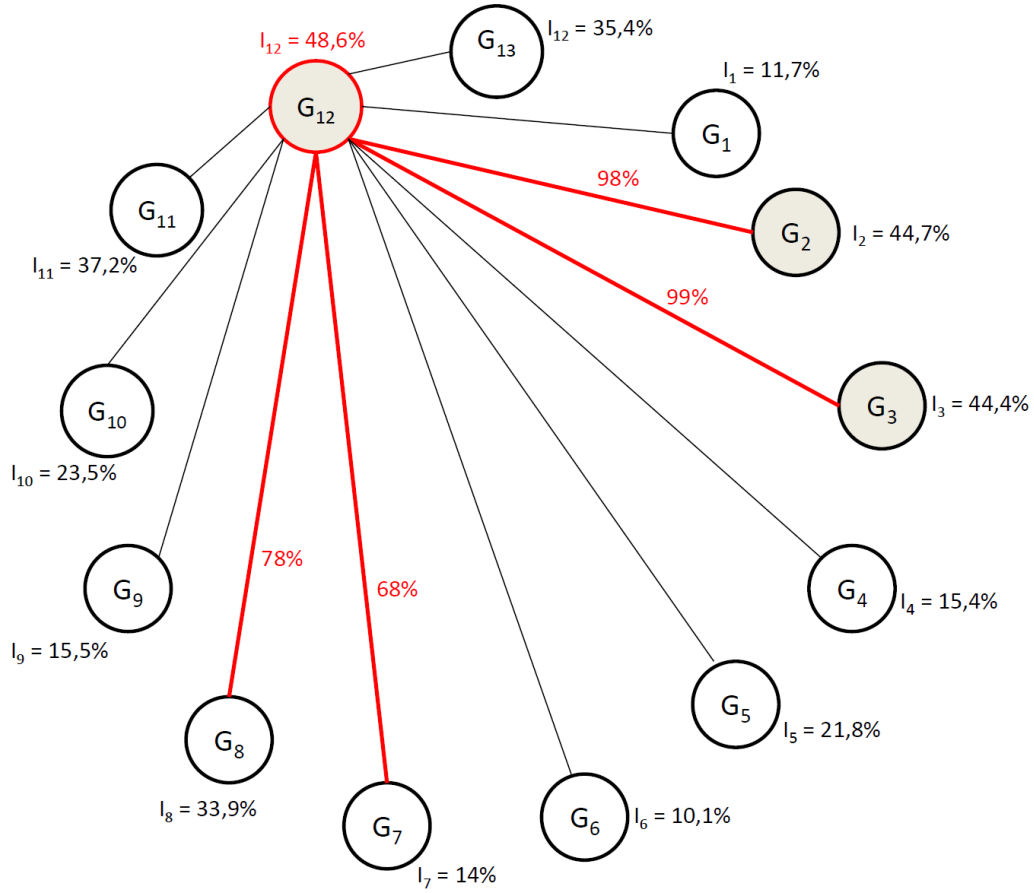
*Passo 8.* Como no primeiro exemplo, escolheu-se  $\beta = 90\%$  e, o resultado final foi  $\mathcal{G}^* = \{G_{12}, G_2, G_3\}$ .

Neste exemplo  $w_{12,3} = 99\%$ ,  $w_{12,2} = 98\%$ ,  $w_{12,8} = 78\%$ ,  $w_{12,7} = 68\%$ . Os outros pesos foram menores. Para este caso qualquer valor  $\beta$  na faixa de  $78\% < \beta < 98\%$  resultaria no mesmo subconjunto final  $\mathcal{G}^*$ .

## Verificação

Assim como nos exemplos anteriores, utilizou-se o estimador bi-objetivo com o decisor de mínima correlação e o MQE para estimar os vetores de parâmetros de cada estrutura de modelo. Desta forma, foi possível comparar o desempenho dos modelos.

As respostas em malha aberta e em malha fechada são apresentadas nas Figuras 4.30 até 4.33. Os modelos com estruturas  $G_2$  e  $G_3$  estão agrupados



**Figura 4.29:** Índice  $I$  e força de conexão entre estruturas de modelos com o modelo líder, Exemplo experimental.

Percentual de vezes que  $H_0$  não foi rejeitada entre estruturas de modelos com o modelo líder.

com  $G_{12}$  e, de fato, essas três estruturas de modelos compõem  $\mathcal{G}^*$ . Entre os modelos que não estão contidos em  $\mathcal{G}^*$ , o desempenho de  $G_7$  foi um dos mais próximos de  $G_{12}$  e, foi incluído nos demais gráficos para fins de comparação.

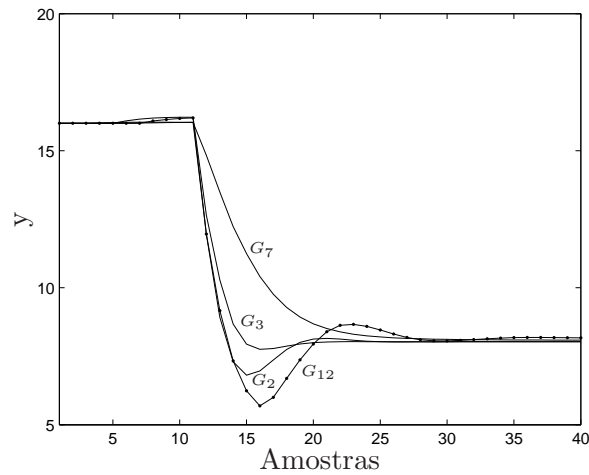
Para a resposta em malha fechada escolheu-se  $K_1 = 1$  (lei de controle com integração) e  $K_2 = -0,1$ , que é negativo devido ao ganho negativo do conversor Buck. O sinal de referência foi alterado de 8 V para 16 V na iteração  $k = 50$ , como mostrado nas Figuras 4.32 e 4.33. As condições iniciais são  $u(i) = 3$ ,  $i = -6, \dots, 0$  e  $y(i) = 8$ ,  $i = -6, \dots, 0$ .

Os resultados mostram que, de fato, o desempenho de modelos com estruturas  $G_2$ ,  $G_3$  e  $G_{12}$  é semelhante. É interessante observar que as estruturas

$G_2$  e  $G_3$  são afins, o que é teoricamente mais adequado para as estruturas de modelos do conversor Buck (Aguirre et al., 2000). As estruturas dos modelos  $G_2$ ,  $G_3$ ,  $G_7$  e  $G_{12}$  são apresentadas na Tabela 4.1.

**Tabela 4.1:** Estruturas dos modelos  $G_2$ ,  $G_3$ ,  $G_7$  e  $G_{12}$ .

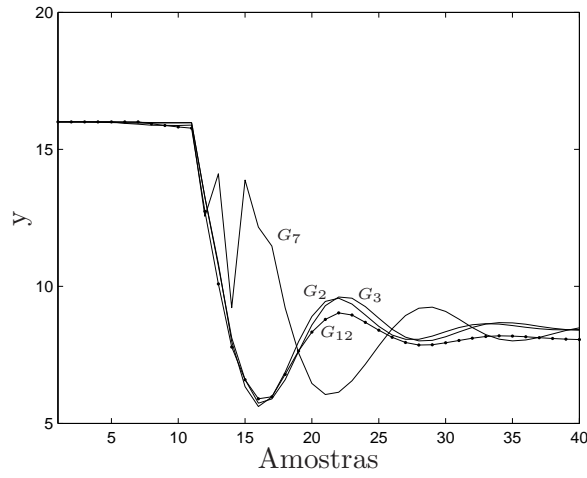
Modelo	Regressores
$G_{12}$	$1 \ y(k-1) \ u(k-1) \ u(k-5)y(k-1)$ $u(k-7)y(k-1) \ u(k-1)y(k-5) \ u(k-4)y(k-7)$
$G_2$	$1 \ y(k-1) \ y(k-2) \ y(k-3) \ u(k-1) \ u(k-3)$
$G_3$	$1 \ y(k-1) \ y(k-2) \ y(k-3) \ u(k-1) \ u(k-5)$
$G_7$	$y(k-1) \ y(k-2) \ u(k-5)y(k-1) \ u(k-1)y(k-1)$ $y(k-3)y(k-1) \ u(k-5)y(k-2) \ u(k-4)y(k-1) \ u(k-1)y(k-2)$



**Figura 4.30:** Resposta em malha aberta, exemplo experimental, estimador bi-objetivo.

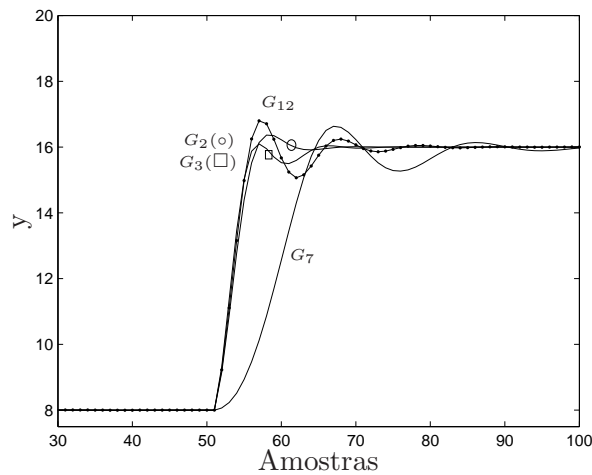
Estruturas de modelos  $G_2$ ,  $G_3$ ,  $G_7$  e  $G_{12}$ .





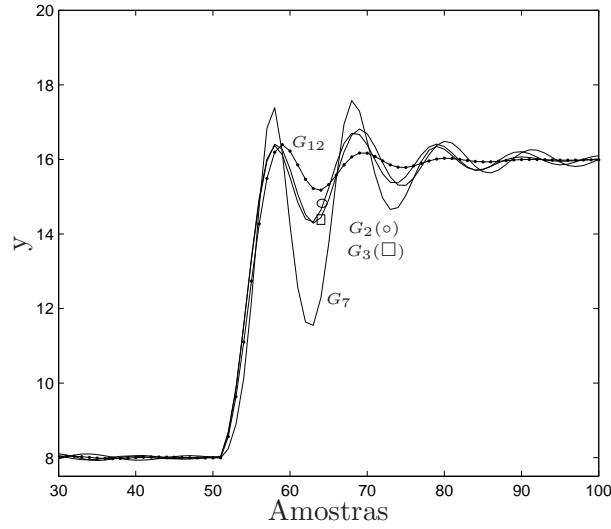
**Figura 4.31:** Resposta em malha aberta, exemplo experimental, estimador MQE.

Estruturas de modelos  $G_2$ ,  $G_3$ ,  $G_7$  e  $G_{12}$ . MQE com modelo de ruído  $e(k) = c_1\nu(k-1) + c_2\nu(k-2) + \nu(k)$  e dez iterações.



**Figura 4.32:** Resposta em malha fechada, exemplo experimental, estimador bi-objetivo.

$K_1 = 1$  e  $K_2 = -0,1$ . Estruturas de modelos  $G_2$  (círculo),  $G_3$  (quadrado),  $G_7$  e  $G_{12}$ .



**Figura 4.33:** Resposta em malha fechada, exemplo experimental, estimador MQE.

$K_1 = 1$  e  $K_2 = -0,1$ . Estruturas de modelos  $G_2$  (círculo),  $G_3$  (quadrado),  $G_7$  e  $G_{12}$ . MQE com modelo de ruído  $e(k) = c_1\nu(k-1) + c_2\nu(k-2) + \nu(k)$  e dez iterações.

## 4.4 Conclusões do capítulo

Neste capítulo discutiu-se e avaliou-se as características do critério de discriminação de estruturas de modelos.

A metodologia apresentada no Capítulo 3 foi aplicada a três exemplos simulados e um experimental. No primeiro exemplo, um exemplo padrão da literatura específica, o desempenho do método foi averiguado. No segundo exemplo trabalhou-se com um modelo com parâmetros de diferentes ordens de grandeza e atraso puro de tempo, e novamente o método foi avaliado, dado que o modelo que gerou os dados é conhecido e obtido por meio de simulação. O terceiro exemplo, desenvolvido pelo próprio autor, teve como objetivo averiguar o método quando têm-se estruturas com atrasos de tempo distintos.

O exemplo experimental é uma aplicação para o caso do conversor *Buck*, que já foi amplamente estudado na literatura. Nesse caso não há modelo nominal e um conjunto de modelos foi obtido e validado, para representar o sistema.

---

No próximo capítulo a metodologia proposta é utilizada como subproduto para descrever um outro procedimento para obter, a partir do método proposto, uma região de incerteza. Tal região representa a incerteza paramétrica e estrutural do sistema.



# Incerteza intervalar

---

*“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 Feyerabend<sup>1</sup>

## 5.1 Introdução

Este capítulo, que é prospectivo e preliminar, tem como objetivo descrever uma metodologia para converter a população de conjuntos de Pareto,  $\mathcal{D}(P_j)$ , em incerteza intervalar nos parâmetros. A transmutação de um conjunto de estruturas de modelos  $\mathcal{G}^*$  para um modelo incerto, é ilustrado na Figura 5.1, destacado pelos retângulos tracejados.

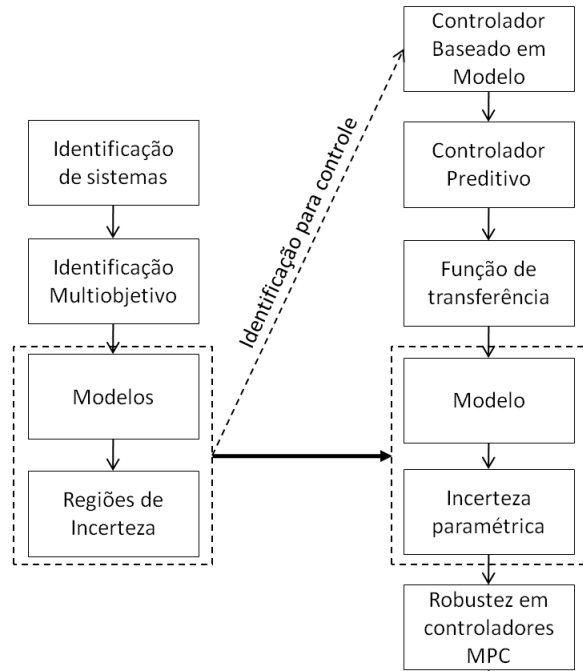
A identificação de modelos é usualmente a tarefa mais significativa e demandada no trabalho de implementação e manutenção de sistemas de controle que usam controle preditivo baseado em modelos. Segundo Rivera et al. (2003),

---

<sup>1</sup>(Feyerabend, 1970, p. 15).

os esforços gastos na modelagem podem consumir até 75% do custo de um projeto de controle.

O modelo com incerteza paramétrica apresentado a seguir pode ser utilizado no projeto de controladores. As próximas seções descrevem a região de incerteza obtida a partir da população de conjuntos de Pareto e o procedimento para obter o modelo com incerteza paramétrica.



**Figura 5.1:** Fluxograma ilustrativo da ligação entre identificação, incerteza e controle.

Destaque para o conjunto de modelos ( $\mathcal{G}^*$ ) e incerteza e a transição para um modelo com incerteza paramétrica aplicada em um controlador preditivo baseado em modelo.

## 5.2 $\mathcal{D}(P_j)$ e incerteza nos parâmetros

No Capítulo 3 foi descrito um procedimento para obter um conjunto  $\mathcal{G}^*$  de modelos  $\beta$ -equivalentes. O conjunto é formado por modelos de estruturas diferentes que não são distinguíveis para os conjuntos de dados disponíveis.

Para cada modelo  $G_j$  do conjunto  $\mathcal{G}^*$ , tem-se um Pareto  $P_j$  associado. Como a estrutura de todos os modelos em cada conjunto Pareto é a mesma, o

que distingue um modelo do outro é o vetor de parâmetros.

Portanto a incerteza do vetor de parâmetros refletirá a dispersão dos conjuntos Pareto. A população de conjuntos de Pareto é dada por (ver Equação 3.2):

$$\mathcal{D}(P_j) = \bigcup_{i=1}^{N_r} P_j(G_j(\hat{\theta}_\lambda), Z_i^N, \bar{Z}^M),$$

em que a população  $\mathcal{D}(P_j)$  relacionada ao conjunto Pareto  $P_j$  é a união de todos  $N_r$  conjuntos de Pareto produzidos pela *mesma* estrutura de modelo, mas com diferentes realizações de dados dinâmicos. Portanto, a dispersão de  $\mathcal{D}(P_j)$  indica a incerteza associada à estrutura  $G_j$ , aos parâmetros  $\hat{\theta}_\lambda$  e aos dados  $Z_i^N$ .

### 5.2.1 Descrição da região de incerteza $\mathbf{D}$

Uma vez que cada Pareto é composto por modelos de mesma estrutura e parâmetros diferentes e cada estrutura de modelo dá origem a um conjunto Pareto diferente, a região de incerteza é formada pela união de todos os conjuntos de populações de Pareto, assim definida

$$\mathbf{D} = \mathcal{D}(P_1) \cup \mathcal{D}(P_2) \cup \dots \cup \mathcal{D}(P_j), \quad (5.1)$$

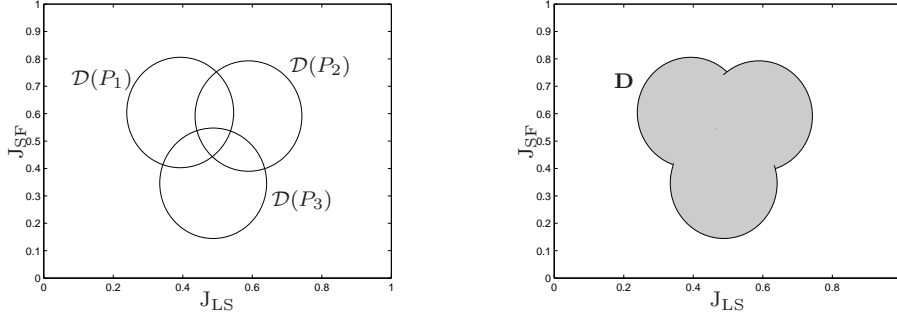
em que  $j$  indica o  $j$ -ésimo elemento do conjunto de estruturas de modelos  $\beta$ -equivalentes ( $\mathcal{G}^*$ ), definido na Seção 3.4.

A Figura 5.2 ilustra a união das populações de conjuntos de Pareto, tal união resulta na região de incerteza  $\mathbf{D}$ , representada como pontos no plano  $J_{LS} \times J_{SF}$ . Ou seja, a região  $\mathbf{D}$  é o resultado da união das populações de conjuntos de Pareto. Tais populações são obtidas por meio de estimação multiobjetivo de parâmetros.

Do conjunto  $\mathbf{D}$  extraem-se informações para obter os intervalos dos parâmetros do modelo incerto, conforme detalhado na próxima seção.

### 5.2.2 Modelo com incerteza intervalar

O procedimento para obtenção do modelo com incerteza intervalar é descrito na sequência. Suponha que há um conjunto inicial  $\mathcal{G}^*$  composto por  $j$



(a) População de conjuntos de Pareto  $\mathcal{D}(P_j)$ ,  $j = 1, 2, 3$ .

(b) União da população de conjuntos de Pareto  $\mathcal{D}(P_j)$ ,  $j = 1, 2, 3$ .

**Figura 5.2:** Ilustração da união das populações de conjuntos de Pareto, região de incerteza  $\mathbf{D}$ .

modelos  $\beta$ -equivalentes. Sejam  $Z_i^N$ ,  $i = 1, \dots, N_r$  realizações de dados dinâmicos e  $\bar{Z}^M$ , os dados estáticos.

1. Escolher uma estrutura do conjunto  $\mathcal{G}^*$ . Sugere-se escolher a estrutura do modelo *líder*. A estrutura escolhida é denotada por  $\mathcal{G}_{\mathbf{D}}$ .  
Em princípio todos os  $j$  modelos do conjunto  $\mathcal{G}^*$  são equivalentes e a escolha de qual utilizar é facultativa. Pode-se, por exemplo, escolher o modelo *líder*, como sugerido, ou escolher um modelo baseado no princípio da parcimônia.
2. Estimar, por meio da técnica de Mínimos Quadrados Estendidos (MQE), os parâmetros  $\hat{\theta}_{\mathbf{D}}$  para a estrutura escolhida  $\mathcal{G}_{\mathbf{D}}$ .
3. Simular o modelo  $\mathcal{G}_{\mathbf{D}}(\hat{\theta}_{\mathbf{D}})$  nos dados de validação e calcular  $J_{LS}$  e  $J_{SF}$ . Colocar os valores no plano  $J_{LS} \times J_{SF}$ , correspondente à região  $\mathbf{D}$ .
4. Tomar um outro conjunto de parâmetros, a partir de uma distribuição Gaussiana com centro em  $\hat{\theta}_{\mathbf{D}}$  e desvio padrão  $\sigma_{\hat{\theta}_{\mathbf{D}}}$ .
5. Repetir os passos 3 e 4  $N$  vezes (por exemplo  $N = 100$ ). Ajustar  $\sigma_{\hat{\theta}_{\mathbf{D}}}$  se necessário.
6. O algoritmo termina quando os valores de  $J_{LS}$  e  $J_{SF}$  calculados a partir de  $\mathcal{G}_{\mathbf{D}}(\hat{\theta}_{\mathbf{D}})$  se sobrepõem à região  $\mathbf{D}$ , definida em (5.1).



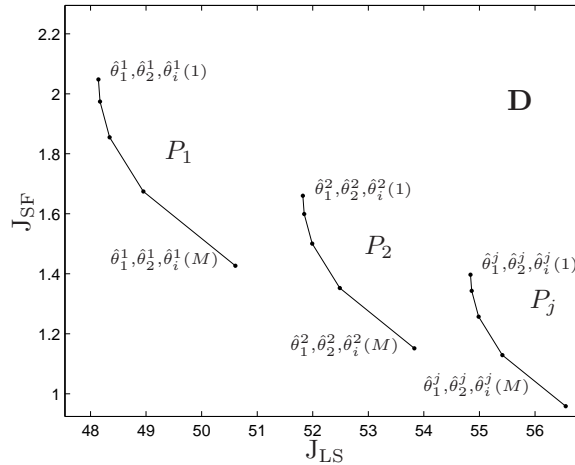
Ao final do algoritmo é obtida uma população de conjuntos de Pareto referentes à estrutura  $\mathcal{G}_{\mathbf{D}}$ , que ocupa uma região no plano  $J_{\text{LS}} \times J_{\text{SF}}$  parecida à ocupada por  $\mathbf{D}$ .

A população de conjuntos de Pareto representa a região de incerteza. Como a estrutura de todos os modelos em cada conjunto Pareto é a mesma, o que distingue um modelo do outro é o vetor de parâmetros.

A fim de traduzir a região de incerteza em valores paramétricos obtêm-se os parâmetros do modelo relacionado a cada ponto das curvas Pareto, de tal forma que:

$$\begin{aligned} \{\hat{\theta}_1^1, \hat{\theta}_2^1 \dots \hat{\theta}_i^1\} &\in P_1 \\ \{\hat{\theta}_1^2, \hat{\theta}_2^2 \dots \hat{\theta}_i^2\} &\in P_2 \\ \{\hat{\theta}_1^j, \hat{\theta}_2^j \dots \hat{\theta}_i^j\} &\in P_j \end{aligned} \quad (5.2)$$

em que  $j$  é o número da população de conjuntos de Pareto referentes à estrutura  $\mathcal{G}_{\mathbf{D}}$  e  $i$  são os parâmetros do modelo  $\mathcal{G}_{\mathbf{D}}(\hat{\theta}_{\mathbf{D}})$  (dimensão de  $\hat{\theta}_{\mathbf{D}}$ ). A Figura 5.3 ilustra a nomenclatura utilizada.



**Figura 5.3:** Ilustração dos conjuntos Pareto que compõem a região  $\mathbf{D}$ .

Conjuntos Pareto  $\{P_1, P_2, \dots, P_j\}$  e seus respectivos modelos  $\mathcal{G}_j^i$ , em que o índice  $j$  refere-se a cada conjunto Pareto da região  $\mathbf{D}$  e o índice  $i$  são os parâmetros dos modelos.  $M$  é o número de pontos da curva Pareto.

Os limites mínimo e máximo são definidos em cada parâmetro. Por exemplo,  $\underline{\hat{\theta}}_1$  é o mínimo de  $\{\hat{\theta}_1^1, \hat{\theta}_1^2 \cdots \hat{\theta}_1^j\}$ , o máximo  $\overline{\hat{\theta}}_1$  é definido de forma análoga. Finalmente, o politopo de incertezas,  $\mathcal{C}$ , é dado por:

$$\mathcal{C} = \left\{ \underline{\hat{\theta}}_n^k \leq \hat{\theta}_{\mathbf{D}} \leq \overline{\hat{\theta}}_n^k, \quad \text{com } k = 1 \cdots j \quad \text{e } n = 1 \cdots i. \right\}. \quad (5.3)$$

O modelo  $\mathcal{G}_{\mathbf{D}}(\hat{\theta}_{\mathbf{D}})$  definido no politopo  $\mathcal{C}$ , incorpora incerteza intervalar nos parâmetros e pode ser utilizado no projeto de controladores robustos baseados em modelo, conforme ilustrado no fluxograma da Figura 5.1.

As metodologias de projeto de controladores baseados em modelo carecem de um cuidado relacionado à robustez. É necessário controladores que preservem a estabilidade e desempenho, apesar das imprecisões ou incertezas dos modelos.

Um dos principais desafios dos controladores robustos é garantir a viabilidade de soluções e consequentemente garantir a estabilidade (Rodrigues e Odloak, 2000; Ramos et al., 2008). Não levar em conta as incertezas pode resultar em um problema de otimização inviável.

### 5.3 Exemplo

O seguinte exemplo é uma extensão daquele abordado na Seção 4.2.1 e tem o intuito de ilustrar a obtenção da incerteza intervalar.

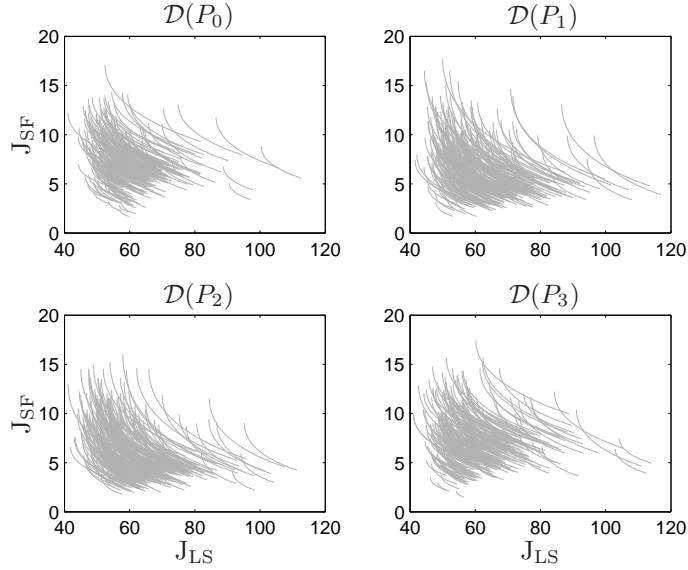
Seja conhecido o conjunto  $\mathcal{G}^* = \{G_0, G_1, G_2, G_3\}$  de estruturas que são  $\beta$ -equivalentes com um nível de significância  $\alpha = 1\%$  e com  $\beta = 90\%$ . Os termos das estruturas são:

$$\begin{aligned} G_0 &: [1 \ y(k-1) \ u(k-2) \ u(k-1)^2 \ y(k-2)^2]; \\ G_1 &: [1 \ y(k-1) \ u(k-2) \ u(k-1)^2 \ y(k-2)^2 \ y(k-2) \ u(k-2)^2]; \\ G_2 &: [1 \ y(k-1) \ u(k-2) \ u(k-1)^2 \ y(k-2)^2 \ y(k-2) \ u(k-1) \ u(k-2)^2]; \\ G_3 &: [1 \ y(k-1) \ u(k-2) \ u(k-1)^2 \ y(k-2)^2 \ u(k-3)u(k-1)y(k-1) \\ &\quad u(k-1)u(k-3)^2]. \end{aligned}$$

Utilizando a Equação (3.2), reapresentada na Seção 5.2, têm-se as populações de conjunto Pareto  $\mathcal{D}(P_j)$ ,  $j = 0, \dots, 3$  referente a cada estrutura de  $\mathcal{G}^*$ ,

conforme ilustrado na Figura 5.4.

A região de incerteza é formada a partir da união de todos os conjuntos de populações de Pareto no plano  $J_{LS}$  e  $J_{SF}$ . Apresentada na Figura 5.5, a região é dada por  $\mathbf{D} = \mathcal{D}(P_0) \cup \mathcal{D}(P_1) \cup \mathcal{D}(P_2) \cup \mathcal{D}(P_3)$ .



**Figura 5.4:** Ilustração de  $\mathcal{D}(P_0)$  a  $\mathcal{D}(P_3)$ .

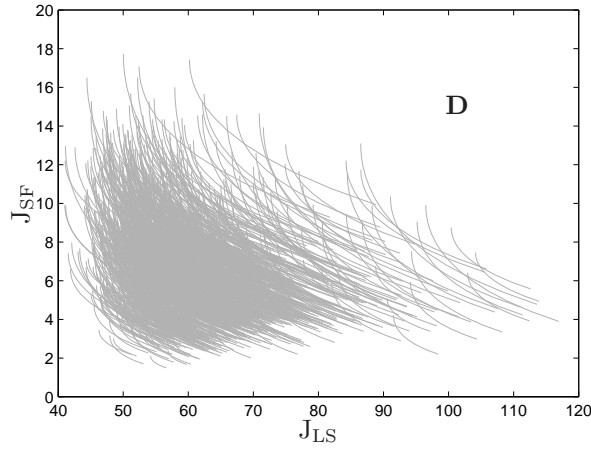
População de conjuntos de Pareto referente aos modelos do grupo  $\mathcal{G}^*$ .

Conhecidos  $\mathcal{G}^*$ ,  $\mathcal{D}(P_j)$  e  $\mathbf{D}$  e as realizações de dados dinâmicos  $(Z_i^N)$  e estáticos  $(\bar{Z}^M)$ , a próxima etapa é obter o modelo com incerteza intervalar nos parâmetros. A seguir passaremos pelos passos descritos na Seção 5.2.2.

*Passo 1.* Escolher a estrutura do modelo *líder* como  $\mathcal{G}_{\mathbf{D}}$ .

$$\mathcal{G}_{\mathbf{D}} : [1 \ y(k-1) \ u(k-2) \ u(k-1)^2 \ y(k-2)^2].$$

*Passo 2.* Para a estrutura  $\mathcal{G}_{\mathbf{D}}$  escolhida, estimar, por Mínimos Quadrados Estendidos, os parâmetros  $\hat{\theta}_{\mathbf{D}}$ . O modelo  $\mathcal{G}_{\mathbf{D}}(\hat{\theta}_{\mathbf{D}})$ , estimado por MQE com modelo de ruído  $e(k) = c_1\nu(k-1) + c_2\nu(k-2) + \nu(k)$  e dez iterações é dado por:



**Figura 5.5:** União da população de conjuntos de Pareto. Região **D**, exemplo ilustrado.

Região **D** no plano  $J_{LS} \times J_{SF}$ , obtida a partir de  $\mathcal{D}(P_0) \cup \mathcal{D}(P_1) \cup \mathcal{D}(P_2) \cup \mathcal{D}(P_3)$ , conforme Equação 5.1.

$$\begin{aligned} y(k) = & 0,5508 + 0,4910y(k-1) + 0,8180u(k-2) \\ & + 0,9836u(k-1)^2 - 0,0502y(k-2)^2. \end{aligned} \quad (5.4)$$

*Passo 3.* Simulou-se o modelo  $\mathcal{G}_{\mathbf{D}}(\hat{\theta}_{\mathbf{D}})$ , descrito pela Equação 5.4, nos dados de validação e calculou-se  $J_{LS}$  e  $J_{SF}$ . Os valores obtidos são dispostos no plano  $J_{LS} \times J_{SF}$ , correspondente à região **D**.

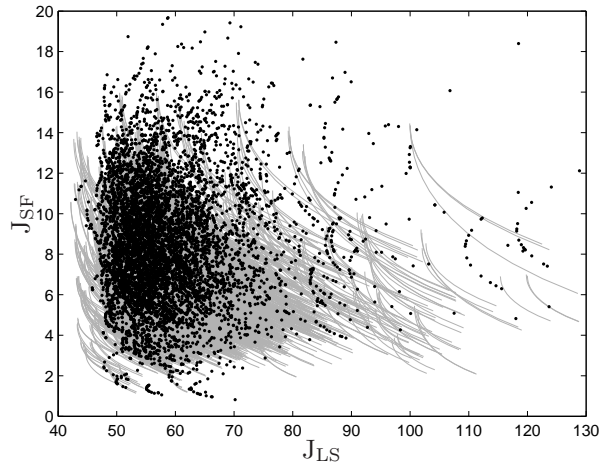
*Passo 4 e Passo 5.* Define-se outro conjunto de parâmetros a partir de uma distribuição Gaussiana com centro em  $\hat{\theta}_{\mathbf{D}}$  e desvio padrão  $\sigma_{\hat{\theta}_{\mathbf{D}}}$ . Simula-se novamente o modelo (5.4), calcula-se  $J_{LS}$  e  $J_{SF}$  e representa os valores no plano (*passo 3*).

*Passo 6.* Neste exemplo executou-se o procedimento (*passos 3 e 4*) 100 vezes, número de repetições suficiente para que os valores de  $J_{LS}$  e  $J_{SF}$  calculados a partir de  $\mathcal{G}_{\mathbf{D}}(\hat{\theta}_{\mathbf{D}})$  se sobrepõem à região **D**, conforme ilustrado na Figura 5.6. Os valores médio dos parâmetros e os respectivos desvios padrão, são apresentados na Tabela 5.1.

A população de conjuntos de Pareto referentes à estrutura  $\mathcal{G}_{\mathbf{D}}$  representa a região de incerteza e a estrutura de todos os modelos em cada conjunto Pareto

**Tabela 5.1:** Valor médio dos parâmetros estimados e os desvios padrão, distribuição gaussiana.

Regressores	Parâmetros ( $\hat{\theta}_{\mathbf{D}}$ )	Desvio padrão
1	0,5508	0,0265
$y(k-1)$	0,4910	0,0084
$u(k-2)$	0,8180	0,0160
$u(k-1)^2$	0,9836	0,0108
$y(k-2)^2$	-0,0502	0,0011



**Figura 5.6:** Região de incerteza  $\mathbf{D}$  e região obtida a partir de  $\mathcal{G}_{\mathbf{D}}(\hat{\theta}_{\mathbf{D}})$ .

Ilustração da sobreposição da região obtida a partir de  $\mathcal{G}_{\mathbf{D}}(\hat{\theta}_{\mathbf{D}})$ , pontos pretos, sobre a região de incerteza  $\mathbf{D}$ , linhas em cinza.

é a mesma. O que distingue um modelo do outro é o vetor de parâmetros. Os valores mínimos e máximos de cada parâmetro da população de Pareto definem o politopo de incertezas (5.2), dado por:

$$\mathcal{C} = \left\{ \begin{array}{l} 0,4616 \leq \hat{\theta}_1 \leq 0,6500 \\ 0,4628 \leq \hat{\theta}_2 \leq 0,5224 \\ 0,7642 \leq \hat{\theta}_3 \leq 0,8779 \\ 0,9474 \leq \hat{\theta}_4 \leq 1,0239 \\ -0,0540 \leq \hat{\theta}_5 \leq -0,0460 \end{array} \right. . \quad (5.5)$$

Os parâmetros  $\hat{\theta}_1, \dots, \hat{\theta}_5$  variam no politopo de incerteza. Portanto,  $\mathcal{G}_{\mathbf{D}}(\hat{\theta}_{\mathbf{D}})$ , em que  $\hat{\theta}_{\mathbf{D}}$  pode assumir qualquer valor em  $\mathcal{C}$  (5.5), representa as estruturas  $\{G_0, G_1, G_2, G_3\}$ , estruturas do conjunto  $\mathcal{G}^*$  de modelos  $\beta$ -equivalentes.

## Verificação

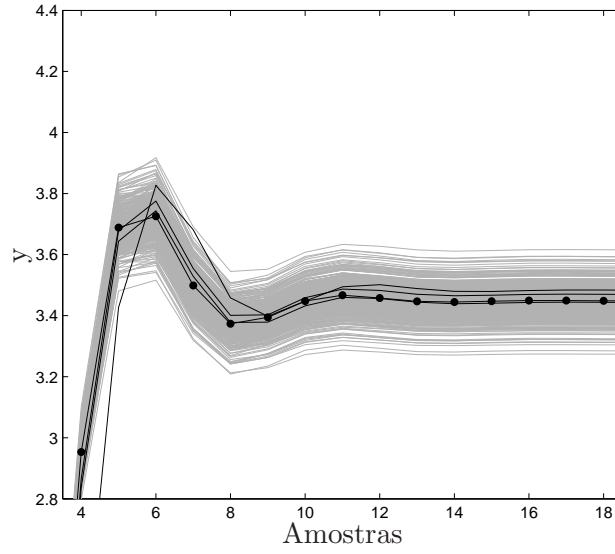
O desempenho do modelo obtido foi verificado. Avaliou-se a estrutura  $\mathcal{G}_{\mathbf{D}}$  para diferentes parâmetros do politopo de incertezas  $\mathcal{C}$  (5.5). Para verificação, escolheram-se duas ferramentas de análise: i) resposta em malha aberta e ii) resposta em malha fechada, conforme descrito na Seção 3.5.

As respostas em malha aberta e em malha fechada são apresentadas nas Figuras 5.7 e 5.8. Para a resposta em malha fechada escolheu-se  $K_1 = 0$  (lei de controle puramente proporcional) e  $K_2 = 0,2$ . O sinal de referência foi alterado, de 0 para 0,5 na iteração  $k = 50$ , como mostrado na Figura 5.8. As condições iniciais são  $u(i) = 0$ ,  $i = -2, -1, 0$  e  $y(i) = 0,5$ ,  $i = -1, 0$ .

O erro em estado estacionário, apresentado na resposta em malha fechada, é justificado uma vez que o controlador é puramente proporcional. Entretanto, o principal intuito é verificar se o modelo com estrutura  $\mathcal{G}_{\mathbf{D}}$  inclui os demais modelos do conjunto  $\mathcal{G}^*$  e não avaliar o desempenho do controlador propriamente dito.

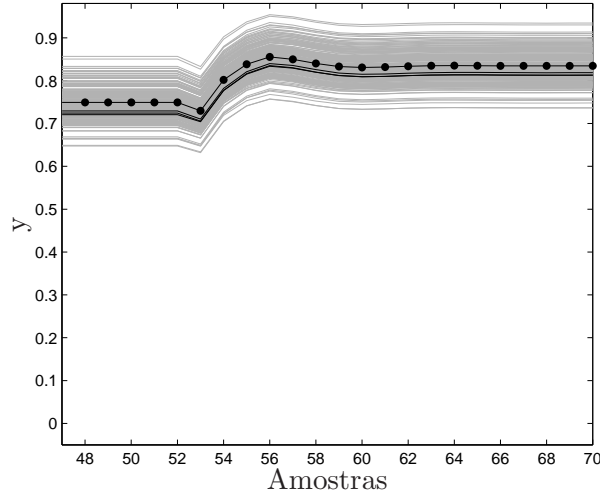
Para as respostas em malha aberta e em malha fechada, os modelos com estrutura  $G_1$ ,  $G_2$  e  $G_3$  encontram-se visivelmente inclusos na faixa de variação do modelo com estrutura  $\mathcal{G}_{\mathbf{D}}$ , mesmo sendo de estruturas diferentes. Na resposta em malha aberta (Figura 5.7),  $G_3$  apresentou uma maior variação, principalmente durante o regime transitório. De forma geral, pode-se dizer que o modelo  $\mathcal{G}_{\mathbf{D}}$ , definido no politopo de incerteza, inclui os modelos  $G_1$ ,  $G_2$  e  $G_3$ , ao menos nos dois testes realizados.

Além disto, a faixa de resposta pode ser utilizada para determinar limites do comportamento dinâmico do sistema controlado, ou em aplicações que destinam-se a encontrar o melhor desempenho no pior caso.



**Figura 5.7:** Resposta em malha aberta, estrutura  $\mathcal{G}_D$ .

(—●) modelo  $\mathcal{G}_D$  com parâmetros definidos por MQE e em cinza modelos com os parâmetros definidos no politopo de incerteza, (—) modelos  $G_1$  até  $G_3$ .



**Figura 5.8:** Resposta em malha fechada, estrutura  $\mathcal{G}_D$ .

(—●) modelo  $\mathcal{G}_D$  com parâmetros definidos por MQE e em cinza modelos com os parâmetros definidos no politopo de incerteza, (—) modelos  $G_1$  até  $G_3$ .

## 5.4 Conclusões do capítulo

Neste capítulo descreveu-se um procedimento para converter a população de conjuntos de Pareto, definida no Capítulo 3, em incerteza intervalar nos parâmetros. O resultado foi um modelo no qual os parâmetros são definidos em um politopo de incerteza e que contém características dos modelos de  $\mathcal{G}^*$ . Um exemplo ilustrado foi descrito com o objetivo de enfatizar e facilitar a compreensão do desenvolvimento sugerido.

O procedimento pode ser aplicável na obtenção de modelos com incerteza intervalar para o projeto de controladores baseados em modelo, tema para pesquisas futuras, apontado no próximo capítulo.



# Conclusões e Perspectivas Futuras

---

*“Quando um cientista enuncia uma lei ou uma teoria, ele está contando como se processa a ordem, está oferecendo um modelo da ordem. Agora ele poderá prever como a natureza vai se comportar no futuro. É isto que significa testar uma teoria: ver se, no futuro, ela se comporta da forma como o modelo previu.”*

Rubem Alves<sup>1</sup>

## 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.

## 6.2 Discussões e contribuições

Quando a quantidade de dados é limitada e possui incertezas, torna-se cada vez mais difícil distinguir entre um conjunto de estruturas de modelos

---

<sup>1</sup>Alves (2005, p. 29).

semelhantes. Em outras palavras, não parece necessário nem justificável procurar, em situações práticas, uma única estrutura de modelo. Assumindo isso, o problema abordado neste trabalho foi a forma de selecionar, a partir de um conjunto de estruturas de modelo candidatas aquelas que parecem mais consistentes com os dados.

A solução proposta para este problema foi um método de discriminação de estruturas de modelos com base em um teste de hipótese e conjuntos Pareto referentes a cada estrutura candidata. O resultado é um conjunto de estruturas de “modelos semelhantes” escolhidos a partir de um conjunto inicial de estruturas candidatas. O procedimento foi testado em dados medidos e simulados. Os modelos obtidos foram comparados e o desempenho das estruturas selecionadas foi semelhante em ambos os testes de malha aberta e malha fechada.

As principais características do procedimento são:

- i) o uso de otimização bi-objetivo para construir o Pareto permite utilizar um conjunto adicional de dados. Neste trabalho foram utilizados dados de estado estacionário;
- ii) o uso do teste estatístico de hipóteses permite que o usuário ajuste o grau de confiança na escolha de estruturas de modelos;
- iii) o resultado é um subconjunto de estruturas de modelos que, do ponto de vista de curvas Pareto, não se distinguem entre si, dentro dos limites de confiança utilizados;
- iv) os conjuntos de curvas Pareto também podem ser encontrados em problemas de otimização não-convexa, o método proposto é igualmente aplicável nesses casos;
- v) para cada estrutura de modelo obtém-se uma população de conjuntos de Pareto  $\mathcal{D}(P)$ . A mesma população foi utilizada para delinear uma região de incerteza  $\mathbf{D}$ ;
- vi) existem implicações em escolher o conjunto de estruturas candidatas. Por exemplo, pode-se ter um conjunto inicial de estruturas candidatas igualmente ruins. Neste caso, o procedimento apontará a equivalência entre modelos ruins. Em outras palavras, o procedimento não indica se

um modelo é bom ou ruim, mas sim se eles são equivalentes ou não. Isto ficou evidente no terceiro exemplo do quarto capítulo, Seção 4.2.3. Portanto, um cuidadoso procedimento de validação ainda é fundamental.

- vii) não é esperado que o procedimento leve em conta questões como qualidade e desempenho do conjunto de estruturas candidatas. Por exemplo, questões como atraso puro de tempo devem ser estimadas por algum outro método, e o conjunto  $\mathcal{G}$  deveria ter estruturas com essas informações.

Um procedimento para converter a população de conjuntos de Pareto, em incerteza intervalar nos parâmetros também foi proposto. O resultado foi um modelo no qual os parâmetros são definidos em um politopo de incerteza e que contém características dos modelos de  $\mathcal{G}^*$ . Uma vez que  $\mathcal{G}_D$  é obtido a partir de um conjunto  $\mathcal{G}^*$  de modelos  $\beta$ -equivalentes, de diferentes estruturas, pode-se dizer que a incerteza é definida tanto nos parâmetros quanto na estrutura.

## 6.3 Publicações

1. Barbosa, A. M., Takahashi, R. H. C. e Aguirre, L. A. Equivalence of Nonlinear Model Structures Based on Pareto Uncertainty. IET Control Theory & Applications, 2015. // doi: 10.1049/iet-cta.2015.0408.
2. Ottoni, A. G. S., Barbosa, A. M., Takahashi, R. H. C. e Aguirre, L. A. Controle Preditivo Baseado em Modelo Linear com Incerteza Intervalar: Formulação e Análise de Estabilidade. In: XI Simpósio Brasileiro de Automação Inteligente, 2013, Fortaleza. IX Simpósio Brasileiro de Automação Inteligente, 2013.
3. Martins, S. A. M., Barbosa, A. M. e Nepomuceno, E. G. Determinação do Número de Termos de Modelos Polinomiais: O Critério de Informação de Akaike Multiobjetivo. In: XI Simpósio Brasileiro de Automação Inteligente, 2013, Fortaleza. IX Simpósio Brasileiro de Automação Inteligente, 2013.

## 6.4 Pesquisas futuras

A partir do que foi apresentado e discutido nesta tese, pode-se enumerar algumas propostas de pesquisas futuras. A esperança é que tais propostas possam ser um dia conduzidas, já que constituem, em princípio, área promissoras para a área de identificação multiobjetivo, e identificação para controle. As propostas são enumeradas a seguir.

1. Avaliar o modelo com incerteza intervalar em controladores baseados em modelo.
2. Desenvolver uma métrica capaz de atestar que todos os modelos do conjunto  $\mathcal{G}^*$  são igualmente *relevantes* para o projeto de controle.
3. Estudar o projeto ótimo de experimentos de identificação para controle robusto.
4. Estabelecer possíveis relações entre a qualidade da região de incerteza  $\mathbf{D}$  e os objetivos de controle robusto.
5. Verificar o procedimento de equivalência de modelos para representações matemáticas diferentes e funções objetivos diferentes, tanto convexa quanto não convexa.

---

## Referências Bibliográficas

---

- Abreu, L. F., Barbosa, B. H. G., Teixeira, A. F., e Aguirre, L. A. (2012). Uso de informação auxiliar no treinamento de redes neurais: um procedimento bi-objetivo. In *XIX Congresso Brasileiro de Automática*, pp. 1719–1726, Campina Grande.
- 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., Alves, G. B., e Corrêa, M. V. (2007). Steady-state performance constraints for dynamical models based on RBF networks. *Engineering Applications of Artificial Intelligence*, 20:924–935.
- 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., Barroso, M. F. S., Saldanha, R. R., e Mendes, E. M. A. M. (2004). Imposing steady-state performance on identified nonlinear polynomial models by means of constrained parameter estimation. *Proc. IEE Part D: Control Theory and Applications*, 151(2):174–179.
- 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., Coelho, M. C. S., e Corrêa, M. V. (2005). On the interpretation and practice of dynamical differences between Hammerstein and Wiener models. *IEE Proc. Control Theory and Appl.*, 152(4):349–356.

- 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 Transactions on Circuits and Systems I*, 47(7):1081–1085.
- Aguirre, L. A. e Furtado, E. C. (2007). Building dynamical models from data and prior knowledge: the case of the first period-doubling bifurcation. *Physical Review E*, 76:046219.
- 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.
- Alves, M. A., Corrêa, M. V., e Aguirre, L. A. (2012). Use of self-consistency in the structure selection of NARX polynomial models. *International Journal of Modelling, Identification and Control*, 15(1):1–12.
- Alves, R. (2005). *Filosofia da Ciência – Introdução ao Jogo e suas Regras*. Editora Loyola, São Paulo, 9 edição.
- Baldacchino, T., Anderson, S. R., e Kadiramanathan, V. (2012). Structure detection and parameter estimation for NARX models in a unified EM framework. *Automatica*, 48:857–865.
- Baldacchino, T., Anderson, S. R., e Kadiramanathan, V. (2013). Computational system identification for bayesian NARMAX modelling. *Automatica*, 49:2641–2651.
- Barbosa, A. M. (2010). *Técnicas de otimização bi-objetivo para a determinação da estrutura de modelos NARX*. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- 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*, pp. 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.
- Barbosa, B. H. G., Aguirre, L. A., Martinez, C. B., e Braga, A. P. (2011). Black and gray-box identification of a hydraulic pumping system. *IEEE Transactions on Control Systems Technology*, 19(2):398–406.
- 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*, pp. 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*, pp. 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.
- Bazanella, A. S., Gevers, M., e Miskovic, L. (2010). Closed-loop identification of MIMO systems: a new look at identifiability and experiment design. *European Journal of Control*, 16(3):228–239.
- Bazanella, A. S. e Rui, R. (2012). Identification of arbitrary model parameterizations: The uses of RITT’s algorithm. In *XIX Congresso Brasileiro de Automática*.
- Billings, S. A., Chen, S., e Korenberg, M. J. (1989). Identification of MIMO nonlinear systems using a forward-regression orthogonal estimator. *International Journal of Control*, 49(6):2157–2189.
- Billings, S. A. e Coca, D. (1999). Discrete wavelet models for identification and qualitative analysis of chaotic systems. *International Journal of Bifurcation and Chaos*, 9(7):1263–1284.

- 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.
- Burnham, K. P. e Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer-Verlag, 2 edição.
- Cantelmo, C. e Piroddi, L. (2010). Adaptive model selection for polynomial narx models. *IET Control Theory & Applications*, 4(12):2693–2706.
- Casdagli, M. (1989). Nonlinear prediction of chaotic time series. *Physica D*, 35:335–356.
- Chakrabarti, A. e Ghosh, J. K. (2011). AIC, BIC, and recent advances in model selection. In *Handbook of the philosophy of science*. Bandyopadhyay, P. S. and Forster M. R.
- Chankong, V. e Haimes, Y. Y. (1983). *Multiobjective decision making: theory and methodology*. North-Holland, New York, Elsevier.
- Chen, S. e Billings, S. A. (1989). Representation of non-linear systems: the NARMAX model. *International Journal of Control*, 49(3):1013–1032.
- 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.
- Ding, F., Liu, P. X., e Liu, G. (2010). Multiinnovation least-squares identification for system modeling. *IEEE J SMCB*, 40(3):767–778.
- Doyle, F. J., Ogunnaike, B. A., e Pearson, R. K. (1995). Nonlinear model-based control using second-order Volterra models. *Automatica*, 31(5):697–714.



- Drakunov, S. V., Perruquetti, W., Richard, J. P., e Belkoura, L. (2006). Delay identification in time-delay systems using variable structure observers. *Annual Reviews in Control*, 30:143–158.
- Falsone, A., Piroddi, L., e Prandini, M. (2015). A randomized algorithm for nonlinear model structure selection. *Automatica*, 60:227–238.
- Farina, M. e Piroddi, L. (2012). Identification of polynomial input/output recursive models with simulation error minimisation methods. *International Journal of Systems Science*, 43(2):111–123.
- Feyerabend, P. K. (1970). *Contra el Método*. Barcelona, Ariel.
- Garcia, C. (2005). *Modelagem e simulação de processos industriais e de sistemas eletromecânicos*. USP - EDUSP, São Paulo, 2 edição.
- Gibbons, J. D. e Chakraborti, S. (2011). *Nonparametric Statistical Inference*. Chapman & Hall/CRC Press, 5th edição.
- Halder, A. e Bhattacharya, R. (2014). Probabilistic model validation for uncertain nonlinear systems. *Automatica*, 50(8):2038–2050.
- Haykin, S. (2001). *Redes neurais: princípios e prática*. Bookman.
- Henderson, L., N. D., Goodman, J. B., Tenenbaum, e Woodward, J. F. (2010). The structure and dynamics of scientific theories: a hierarchical bayesian perspective. *Philosophy of Science*, 77:172–200.
- Hájek, A. (2007). The reference class problem is your problem too. *Synthese*, 156:563–585.
- 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.
- Isermann, R. e Munchhof, M. (2011). *Identification of Dynamic Systems: An Introduction with Applications*. Springer.
- Jakubek, S., Hametner, C., e Keuth, N. (2008). Total least squares in fuzzy system identification: An application to an industrial engine. *Eng. Appl. Artif. Intell.*, 21(8):1277–1288.

- Johansen, T. A. (1996). Identification of non-linear systems using empirical data and prior knowledge - an optimization approach. *Automatica*, 32:337.
- Kothare, M. V., Balakrishnan, V., e Morari, M. (1996). Robust constrained model predictive control using linear matrix inequalities. *Automatica*, 32(10):1361–1379.
- Ljung, L. (1999). *System Identification: Theory for the User*. Prentice-Hall, New Jersey, 2 edição.
- Longford, N. T. (2005). Model selection and efficiency. Is “Which model...?” the right question? *Journal of the Royal Statistical Society: Series A*, 168(3):469–472.
- Lovera, M. (2014). *Control-oriented Modelling and Identification: Theory and practice*. IET, The Institution of Engineering and Technology.
- 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.
- Martins, S. A. M., Nepomuceno, E. G., e Barroso, M. F. S. (2013). Improved structure detection for polynomial NARX models using a multiobjective error reduction ratio. *Journal of Control, Automation and Electrical Systems*, 24:764–772.
- 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.
- 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*, pp. 2156–2161, Natal – RN – Brasil.
- Norton, J. P. (1986). *An introduction to identification*. London: Academic Press.
- Ottoni, A. G. S. (2013). Controle preditivo baseado em modelo com incertezas intervalares. Texto de Qualificação.
- Petrucelli, J. D., Nandram, B., e Chen, M. (1999). *Applied Statistic for Engineers and Scientists*. Prentice Hall, Inc.
- Piroddi, L. (2008). Simulation error minimization methods for NARX model identification. *International Journal of Modelling, Identification and Control*, 3(4):392–403.
- Piroddi, L., Farina, M., e Lovera, M. (2012). Black box model identification of nonlinear input-output models: a Wiener–Hammerstein benchmark. *Control Engineering Practice*, 20:1109–1118.
- Piroddi, L. e Leva, A. (2006). Model structure selection based on polygonal curve approximation techniques. In *Proc. 45th IEEE Conference on Decision and Control*, pp. 805–810.
- 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.
- Potts, A. S., Romano, R. A., e Garcia, C. (2014). Improving performance and stability of MPC relevant identification methods. *Control Engineering Practice*, 22:20–33.
- Pulecchi, T. e Piroddi, L. (2007). A cluster selection approach to polynomial NARX identification. In *Proceedings of the 2007 American Control Conference*, pp. 852–857.

- Ramos, C., Martínez, M., Sanchis, J., e Herrero, J. M. (2008). Robust and stable predictive control with bounded uncertainties. *Journal of Mathematical Analysis and Applications*, 342:1003–1014.
- Rivera, D. E., Lee, H., Braun, M. W., e Mittelmann, H. D. (2003). “Plant friendly” system identification: a challenge for the process industries. In *13th IFAC Symposium on System Identification*, pp. 917–922.
- Robinson, A. P. e Froese, R. E. (2004). Model validation using equivalence tests. *Ecological Modelling*, 176:349–358.
- Rodrigues, M. A. e Odloak, D. (2000). Output feedback MPC with guaranteed robust stability. *Journal of Process Control*, 10:557–572.
- Savage, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical Association*, 46:55–67.
- Schroeders, A. (2008). *A Bola de Neve - Warren Buffett e o negócio da vida*. Sextante.
- Sohlberg, B. (1998). *Supervision and Control for Industrial Processes - Using Grey Box Models, Predictive Control and Fault Detection Methods*. Springer-Verlag.
- Solares, J. R. A. e Wei, H. L. (2015). Nonlinear model structure detection and parameter estimation using a novel bagging method based on distance correlation metric. *Nonlinear Dynamics*, 82:201–215.
- Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society, Series B*, 58(1):267–288.
- Volterra, V. (1930). *Theory of functions*. Blackie and Sons.
- Walpole, R. E., Myers, R. H., Myers, S. L., e Ye, K. E. (2012). *Probability and Statistics for Engineers and Scientists*. Pearson.
- Waschburger, R. e Galvão, R. K. H. (2013). Time delay estimation in discrete-time state-space models. *SignalProcessing*, 93:904–912.

- 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.
- Wiener, N. (1958). *Nonlinear problems in random theory*. In Wiley, J. and Sons, New York.
- Wilcoxon, F., Katti, S. K., e Wilcox, R. A. (1972). Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test. *Institute of Mathematical Statistics, Selected Tables in Mathematical Statistics, American Mathematical Society*, 1:171–259.
- Wit, E., Heuvel, E. V. D., e Romeijn, J. (2012). All models are wrong...: an introduction to model uncertainty. *Statistica Neerlandica*, 66(3):217–236.
- Young, P. C. (1970). An instrumental variable method for real-time identification of a noisy process. *Automatica*, 6(2):271–287.
- Zheng, G., Barbot, J. P., e Boutat, D. (2013). Identification of the delay parameter for nonlinear time-delay systems with unknown inputs. *Automatica*, 49(6):1755–1760.
- Zhu, Y. (2009). System identification for process control: recent experience and outlook. *International Journal of Modelling, Identification and Control*, 6(2):89–103.



## Rotinas Computacionais

---

*“- E como os fregueses o chamam?  
- Os doutor me chamam assim, ó: “Ô, negão!” Eu acho até que é carinhoso.  
- O senhor chama eles de doutor?  
- Pra mim todo mundo é doutor. Pisou no aeroporto é doutor. É ó, doutor, como vai, doutor, é pra já, doutor....”*

Entrevista com Adail José da Silva<sup>1</sup>

Esse apêndice tem como objetivo apresentar as principais rotinas computacionais desenvolvidas no trabalho e que servirão de base para os estudos futuros. Todas rotinas foram desenvolvidas no *MATLAB*<sup>TM</sup> (Matrix Laboratory, MathWorks, 2011b), por meio de *script* e do pacote *SIMULINK*<sup>®</sup>.

### A.1 teste-hipotese

*Sintaxe*

$[p,h] = \text{ranksum}(amostra1, amostra2, alpha)$

---

<sup>1</sup>Entrevista com Adail José da Silva, carregador de malas do Aeroporto Salgado Filho, em Porto Alegre, para a coluna semanal de reportagem do jornal Zero Hora, 1999.

*Descrição*

Retorna o resultado do teste de hipótese (*Wilcoxon*) entre as duas amostras (modelos) avaliados.

*Entrada*

*amostra1* : amostra do modelo 1  
*amostra2* : amostra do modelo 2  
*alpha* : nível de significância

*Saída*

*p* : o valor da probabilidade, valor-p (*p-Value*)  
*h* : indica se a hipótese nula foi rejeitada para o valor de *alpha*

*MATLAB*<sup>TM</sup> (Matrix Laboratory, MathWorks, 2011b)

Adaptado por Alípio Barbosa - 15/06/14 - UFMG/Belo Horizonte

## A.2 pareto

*Sintaxe*

$[jls, jsf] = \text{pareto}(\text{modelo}, u, y, ue, ye, nmod, \text{flag})$

*Descrição*

Gera o gráfico do conjunto Pareto Ótimo e retorna o conjunto de dados de custo dinâmico e estático que compõem a curva Pareto.

*Entrada*



*modelo* : conjunto de termos (ver item A.3).  
*u* : dados dinâmicos de entrada  
*y* : dados dinâmicos de saída  
*ue* : dados estáticos de entrada  
*ye* : dados estáticos de saída  
*nmod* : número de pontos da curva Pareto  
*flag* : habilita a saída gráfica se  $\text{flag} = 1$ , desabilita se  $\text{flag} = 0$

#### Saída

*jls* : vetor com os dados de custo dinâmico  
*jsl* : vetor com os dados de custo estático

Alípio Barbosa - 15/06/15 - UFMG/Belo Horizonte

## A.3 build-r

#### Sintaxe

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

#### 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.4 build-q

*Sintaxe*

`[Q, agrup] = build-q(R, 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.3 *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.5 build-pr

*Sintaxe*

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

*Descrição*

Retorna a matriz de regressores  $\Psi$ .

*Entrada*

$\text{modelo}$  : conjunto de termos

$u$  : dados de entrada

$y$  : dados de saída

*Saída*

$\Psi$  : matriz de regressores

Eduardo Mendes - 11/08/94

ACSE - Sheffield

## A.6 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

## Tabela Estatística

---

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

Lutz Bornmann<sup>1</sup>

Esse apêndice tem como objetivo facilitar o acesso do leitor à tabela estatística utilizada para ilustrar a aplicação do teste de hipótese. A tabela apresentada é baseada no teste de Wilcoxon (*Wilcoxon Rank Sum Test*) e está também disponível em Walpole et al. (2012).

---

<sup>1</sup>Lutz 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).

Table A.17 (continued) Critical Values for the Wilcoxon Rank-Sum Test

One-Tailed Test at $\alpha = 0.025$ or Two-Tailed Test at $\alpha = 0.05$																			
$n_1$	$n_2$																		
	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
1																			
2					0	0	0	0	1	1	1	1	1	2	2	2	2		
3		0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8		
4	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	13		
5		2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20		
6			5	6	8	10	11	13	14	16	17	19	21	22	24	25	27		
7				8	10	12	14	16	18	20	22	24	26	28	30	32	34		
8					13	15	17	19	22	24	26	29	31	34	36	38	41		
9						17	20	23	26	28	31	34	37	39	42	45	48		
10							23	26	29	33	36	39	42	45	48	52	55		
11								30	33	37	40	44	47	51	55	58	62		
12									37	41	45	49	53	57	61	65	69		
13										45	50	54	59	63	67	72	76		
14											55	59	64	67	74	78	83		
15												64	70	75	80	85	90		
16													75	81	86	92	98		
17														87	93	99	105		
18															99	106	112		
19																113	119		
20																	127		

One-Tailed Test at  $\alpha = 0.05$  or Two-Tailed Test at  $\alpha = 0.1$ 

$n_1$	$n_2$																		
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1																		0	0
2			0	0	0	1	1	1	1	2	2	3	3	3	3	4	4	4	
3	0	0	1	2	2	3	4	4	5	5	6	7	7	8	9	9	10	11	
4		1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18	
5			4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25	
6				7	8	10	12	14	16	17	19	21	23	25	26	28	30	32	
7					11	13	15	17	19	21	24	26	28	30	33	35	37	39	
8						15	18	20	23	26	28	31	33	36	39	41	44	47	
9							21	24	27	30	33	36	39	42	45	48	51	54	
10								27	31	34	37	41	44	48	51	55	58	62	
11									34	38	42	46	50	54	57	61	65	69	
12										42	47	51	55	60	64	68	72	77	
13											51	56	61	65	70	75	80	84	
14												61	66	71	77	82	87	92	
15													72	77	83	88	94	100	
16														83	89	95	101	107	
17															96	102	109	115	
18																109	116	123	
19																	123	130	
20																		138	

Tabela estatística. Teste de hipótese. *Wilcoxon Rank Sum Test*. Retirado de Walpole et al. (2012, p. 761).