

**UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS**

André Luiz Borges Barcellos

**Estudo de Técnicas de Identificação de Sistemas aplicado
a um Twin-Rotor**

São Carlos

2025

André Luiz Borges Barcellos

Estudo de Técnicas de Identificação de Sistemas aplicado a um Twin-Rotor

Monografia apresentada ao Curso de Engenharia Elétrica com Ênfase em Sistemas de Energia e Automação, da Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Engenheiro Eletricista.

Orientador: Prof. Dr. Marcos Rogério Fernandes

**São Carlos
2025**

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Ficha catalográfica elaborada pela Biblioteca Prof. Sérgio Rodrigues Fontes
e pelo Serviço de Comunicação e Marketing da EESC-USP,
com dados inseridos pelo(a) autor(a).

B242e

Barcellos, André Luiz Borges

Estudo de Técnicas de Identificação de Sistemas
aplicado a um Twin-Rotor / André Luiz Borges Barcellos ;
orientador Marcos Rogério Fernandes. -- São Carlos, 2025.

93 p.

Monografia - Graduação em Engenharia Elétrica com
ênfase em Sistemas de Energia e Automação -- Escola de
Engenharia de São Carlos da Universidade de São Paulo,
2025.

1. Identificação de sistemas. 2. Modelagem dinâmica.
3. MMQ. 4. LASSO. 5. Twin-rotor. I. Fernandes, Marcos
Rogério, orient. II. Título.

FOLHA DE APROVAÇÃO

Nome: André Luiz Borges Barcellos

Título: “Estudo de Técnicas de Identificação de Sistemas aplicado a um Twin-Rotor”

Trabalho de Conclusão de Curso defendido e aprovado
em 03 / 12 / 2025,

com NOTA 10,0 (dez, zero), pela Comissão
Julgadora:

Prof. Dr. Marcos Rogério Fernandes - Orientador - SEL/EESC/USP

**Prof. Dr. João Bosco Ribeiro do Val - Professor Titular
FEEC/UNICAMP**

**Prof. Dr. Melkzedekue Moraes de Alcantara Calabrese Moreira -
Professor Temporário SEL/EESC/USP**

**Coordenador da CoC-Engenharia Elétrica - EESC/USP:
Professor Associado José Carlos de Melo Vieira Júnior**

AGRADECIMENTOS

Agradeço primeiramente à Universidade de São Paulo (USP) e à Escola de Engenharia de São Carlos (EESC), por proporcionarem o ambiente e a infraestrutura propícios para meu desenvolvimento acadêmico. Em especial, estendo minha gratidão ao meu orientador, Prof. Dr. Marcos Rogério Fernandes, por sua inestimável disposição, apoio e orientação técnica que nortearam este trabalho e que serão fundamentais para minha carreira profissional.

Aproveito também para agradecer a todos os professores, servidores e colegas que acompanharam minha trajetória acadêmica e cujo conhecimento compartilhado foi essencial para minha formação. O caminho foi repleto de desafios e alegrias, e o suporte fornecido por todos, tanto nos momentos bons quanto nos difíceis, foi fundamental para a conclusão desta etapa.

Por fim, faltam-me palavras para expressar minha gratidão à minha família. Em particular, agradeço a meus pais por seu apoio e amor incondicionais, por todos os ensinamentos e por todo o esforço e carinho despendidos. Serei eternamente grato, visto que certamente não teria alcançado esta conquista sem todo o suporte que recebi.

*“If you wish to make an apple pie from scratch,
you must first invent the universe”*

Carl Sagan

RESUMO

BARCELLOS, A. L. B. **Estudo de Técnicas de Identificação de Sistemas aplicado a um Twin-Rotor**. 2025. 91 p. Monografia (Trabalho de Conclusão de Curso) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2025.

Os métodos de identificação de sistemas são essenciais para a engenharia moderna, dada a necessidade de modelos adequados para a implementação de técnicas de controle e para o monitoramento preditivo de processos industriais. Nesse contexto, é fundamental definir uma estrutura para otimizar o processo de modelagem, partindo da escolha de um sinal de excitação persistentemente excitante que assegure a excitação de um amplo espectro de frequências durante o experimento, até a seleção de representações robustas à presença de ruídos no sistema, utilizando estimadores, como o MMQ, para minimizar o erro de estimação. A metodologia escolhida deve também priorizar a obtenção de modelos com alta capacidade de generalização e menor dependência das características do sinal de entrada, uma vez que os processos de identificação do tipo caixa-preta tendem a incluir um elevado número de regressores pouco relevantes para o sistema de interesse. Para mitigar esse problema, é essencial aplicar técnicas como o LASSO, que reduzem a influência do *overfitting* ao selecionar regressores mais representativos do modelo. Embora não sejam o principal motivador do uso dessas técnicas, elas também contribuem para a eficiência computacional ao reduzir o custo associado ao processo de estimação. A eficácia desse procedimento foi demonstrada em experimentos com sistemas lineares e não lineares, nos quais as metodologias apresentadas foram comparadas aos processos tradicionais de identificação. Por fim, os conhecimentos obtidos foram aplicados a uma modelagem do tipo caixa cinza para um sistema de alta complexidade e forte acoplamento: o twin-rotor.

Palavras-chave: Identificação de Sistemas; Modelagem Dinâmica; MMQ; LASSO; Overfitting; Caixa Preta; Caixa Cinza; Twin-Rotor.

ABSTRACT

BARCELLOS, A. L. B. **Study of Systems Identification Techniques applied to a Twin-Rotor**. 2025. 91 p. Monograph (Conclusion Course Paper) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2025.

ABSTRACT

System identification methods are essential in modern engineering, given the need for satisfactory models to implement control techniques and to predictively monitor industrial processes. In this context, it is crucial to define a structured workflow to optimize the modeling process, starting from the choice of a persistently exciting input signal, which ensures the excitation of a broad frequency spectrum during the experiment, up to the selection of representations that are robust to eventual noise inserted into the system, making use of estimators such as Least Squares (LS) to minimize estimation error. The chosen methodology must also focus on obtaining models with high generalization capability and reduced dependence on the characteristics of the input signal, since black-box identification processes often involve a large number of regressors that are only weakly relevant to the system of interest. To address this, it is essential to apply techniques such as LASSO, which reduces overfitting by selecting the regressors that are more representative of the model. Although not the primary motivation for using such methods, they also improve efficiency by reducing the computational cost of estimation, thereby lowering processing demands and, consequently, the financial impact during modeling. The effectiveness of this process was demonstrated through experiments with linear and nonlinear systems, in which various methodologies were compared with traditional identification procedures. Finally, the knowledge obtained was applied to a grey-box modeling approach for a highly complex and coupled system: the twin-rotor.

Keywords: System Identification; Dynamic Modeling; Least Squares (LS); LASSO Regression; Overfitting Reduction; Black-Box Modeling; Grey-Box Modeling; Twin-Rotor System.

LISTA DE FIGURAS

Figura 1 – Comparação entre os processos relacionados ao método científico e à identificação de sistemas.	23
Figura 2 – Estudo analítico proposto por Chen.	24
Figura 3 – PDF da distribuição normal em função dos parâmetros $\mu_{\mathcal{N}}$ (média) e $\sigma_{\mathcal{N}}$ (desvio padrão).	33
Figura 4 – Validação Estatística de uma Sequência Aleatória Gerada no Matlab: Ruído Branco Gaussiano.	33
Figura 5 – Autocorrelação da sequência $a[k]$	34
Figura 6 – Distribuição normal multidimensional - visualização 3D. As funções $p(x)$ e $p(y)$ são distribuições gaussianas marginais.	35
Figura 7 – Escalas de tons cinzentos que ilustram diferentes abordagens “caixa-cinza”. Da região 1 à 3, a quantidade de informação sobre o sistema diminui.	38
Figura 8 – Sistema M do tipo SISO, apresentando entrada $u[k]$ e saída $y[k]$	41
Figura 9 – Exemplo de sinal do tipo <i>Chirp</i>	42
Figura 10 – Comparação entre senoides e variável aleatória, simulação de 1s e $F_s = 10^4 Hz$	43
Figura 11 – Comparação entre uma variável aleatória e um período do sinal PRBS, com simulação de 1s e $F_s = 10^4 Hz$. $N = 10^5$	44
Figura 12 – Comportamento da função de autocorrelação do sinal PRBS.	45
Figura 13 – Exemplo de implementação em nível de hardware, $n = 6$	45
Figura 14 – Comparação entre variável aleatória e um período do sinal PRMLS escolhido, simulação de 1s e $F_s = 10^4 Hz$. $N = 10^5$	47
Figura 15 – Comportamento da função de autocorrelação do sinal PRMLS escolhido.	47
Figura 16 – Sistema dado pela equação 3.8.	48
Figura 17 – Exemplo de uma função $J(\theta)$ de natureza quadrática.	51
Figura 18 – Exemplo de estimação a partir do MMQ - número máximo de 100 amostras.	54
Figura 19 – Exemplos de underfitting (à esquerda) e de overfitting (à direita).	57
Figura 20 – Comportamento da função objetivo $f(M, \theta)$ em vermelho, apresentando ponto de minimização em $\hat{\theta}$. Os termos de regularização impostos pela equação 3.50 são denotados pelas formas geométricas em azul. Da esquerda para a direita: Lasso ($q = 1$) e Ridge ($q = 2$).	58
Figura 21 – Diferentes formatos das restrições introduzidas em eq. (3.50).	59
Figura 22 – Exemplo de validação cruzada.	59

Figura 23 – Influência do parâmetro λ na estimação, com dados obtidos na seção 4.4, relativos à figura 36a. Uso da escala logarítmica no eixo x.	60
Figura 24 – Comparação entre o espectro de frequências do sinal PRBS de degrau unitário, simulação de $2.5s$ e $F_s = 10^3 Hz$	63
Figura 25 – Resposta ao sistema para diferentes distúrbios na entrada, simulação de $1s$ e $F_s = 10^4 Hz$	64
Figura 26 – Comparação do modelo obtido para $\rho = 10$ e saída ideal. Simulação de $1s$ e $F_s = 10^4 Hz$	64
Figura 27 – Circuito RC implementado no <i>Simulink</i>	65
Figura 28 – Comparação entre respostas do sistema RC ao sinal PRBS e degrau unitário, simulação de $2.5s$ e $F_s = 10^3 Hz$	66
Figura 29 – Comparação entre respostas em frequência do sistema RC ao sinal PRBS e degrau unitário, simulação de $2.5s$ e $F_s = 10^3 Hz$	66
Figura 30 – Comportamento da estimação do vetor de parâmetros em função de ρ , simulação de $1s$ e $F_s = 10^4 Hz$	67
Figura 31 – Dados do circuito RC obtidos em laboratório. Para estimação e validação: $5ms$, $F_s = 5 \cdot 10^5 Hz$; $25ms$, $F_s = 1 \cdot 10^5 Hz$	68
Figura 32 – Dados do circuito RC para o MMQ recursivo, com ampliação no instante em que a carga resistiva é aumentada. Simulação de $1s$, $F_s = 1 \cdot 10^4 Hz$	69
Figura 33 – Esquemático do circuito para aferição de temperatura.	70
Figura 34 – Circuito para aferição de temperatura, sensor NTC indicado.	70
Figura 35 – Entradas dos dados obtidos para o sistema de temperatura. Para estimação e validação: $10 min$, $F_s = 50 Hz$; $5 min$, $F_s = 50 Hz$	71
Figura 36 – Saídas estimadas para os dados obtidos do sistema de temperatura.	72
Figura 37 – Função de autocorrelação e espectro de frequências da entrada PRBS apresentada em 35a.	73
Figura 38 – Saídas estimadas para 100 regressores.	74
Figura 39 – Saída para 3 regressores, com o uso do método LASSO em um modelo de 100 parâmetros.	75
Figura 40 – Adição de um termo dependente de $y[k-1]^4 \cdot u[k-1]$ ao sistema obtido na seção 4.4.2	77
Figura 41 – Sistema de interesse.	79
Figura 42 – Conjunto de dados simulados, gerados com base no modelo de twin-rotor, para identificação.	84
Figura 43 – Resultado da identificação para saída y_1	85
Figura 44 – Resultado da identificação para saída y_2	85

LISTA DE TABELAS

Tabela 1 – Geração de um sinal PRBS	46
Tabela 2 – Classes de modelos notáveis obtidos a partir de 3.8.	48
Tabela 3 – Estimação dos parâmetros da reta $y = \theta_1 t + \theta_2$ em função do número de amostras	53
Tabela 4 – Dados de validação para o modelo da figura 26.	65
Tabela 5 – Componentes utilizados para a simulação.	65
Tabela 6 – Dados de validação para o modelo da figura 31.	68
Tabela 7 – Dados para o MMQ e o MMQR, $\lambda = 0.9$	69
Tabela 8 – Dados de estimação para o modelo da figura 36a	72
Tabela 9 – Dados de validação para o modelo da figura 36b	72
Tabela 10 – Dados de validação para o modelo da figura 36c	72
Tabela 11 – Dados de estimação para o modelo da figura 38a.	74
Tabela 12 – Dados de validação para o modelo da figura 38b.	74
Tabela 13 – Tempo necessário para estimar e validar o sistema, conforme a figura 34. Notebook modelo Lenovo Ideapad 82CGS00100.	75
Tabela 14 – Dados em função do modelo LASSO dado pela figura 39	76
Tabela 15 – Dados para o modelo não linear proposto.	77
Tabela 16 – Parâmetros dinâmicos referentes às equações 5.1 e 5.2	80
Tabela 17 – Parâmetros do modelo TRMS: significado e valores numéricos	80

LISTA DE ABREVIATURAS E SIGLAS

ARX	Autoregressive with eXogenous inputs.
DFT	Discrete Fourier Transform.
EDO	Equação Diferencial Ordinária.
FPE	Final Prediction Error.
FT	Função de Transferência.
i.i.d.	independente e identicamente distribuída.
LAC	Laboratório de Controle.
LASSO	Least Absolute Shrinkage and Selection Operator.
LQR	Linear Quadratic Regulator.
MLS	Maximum Length Sequence.
MMQ	Método dos Mínimos Quadrados.
MMQR	Método dos Mínimos Quadrados Recursivo.
MSE	Mean Squared Error.
NARX	Non-linear Autoregressive with eXogenous inputs.
PDF	Função de Densidade de Probabilidade.
PID	Proporcional, Integral e Derivativo.
PRBS	Pseudo Random Binary Sequence.
PRMLS	Pseudo Random Multi Level Sequence.
PSD	Power Spectral Density.
RMS	Root Mean Square.

LISTA DE SÍMBOLOS

N	Número de amostras.
λ	Parâmetro de penalização, fator de esquecimento.
$\mathbb{E}[a[k]]$	Esperança matemática de $a[k]$.
$\mathbb{S}_{aa}(e^{j\omega})$	PSD do sinal $a[k]$.
$\mathcal{F}[a[k]]$	Operador de Fourier.
\mathcal{N}	Função gaussiana, distribuição normal.
μ_a, \bar{a}	Média da variável $a[k]$.
ϕ_{ab}	Espectro cruzado entre as variáveis a e b .
σ_a^2	Variância da variável $a[k]$.
θ	Vetor de parâmetros.
\tilde{a}	Variável $a[k]$ normalizada.
ζ	Erro de estimação, vetor de resíduos.
$e[k]$	Erros de medição, distúrbios diversos presentes nos sinais aferidos.
k	Índice de tempo discreto.
r_{ab}	Correlação entre as variáveis a e b .
a'	Flutuação da variável $a[k]$.
$\ \cdot\ _\infty$	Norma infinita, valor máximo absoluto.

SUMÁRIO

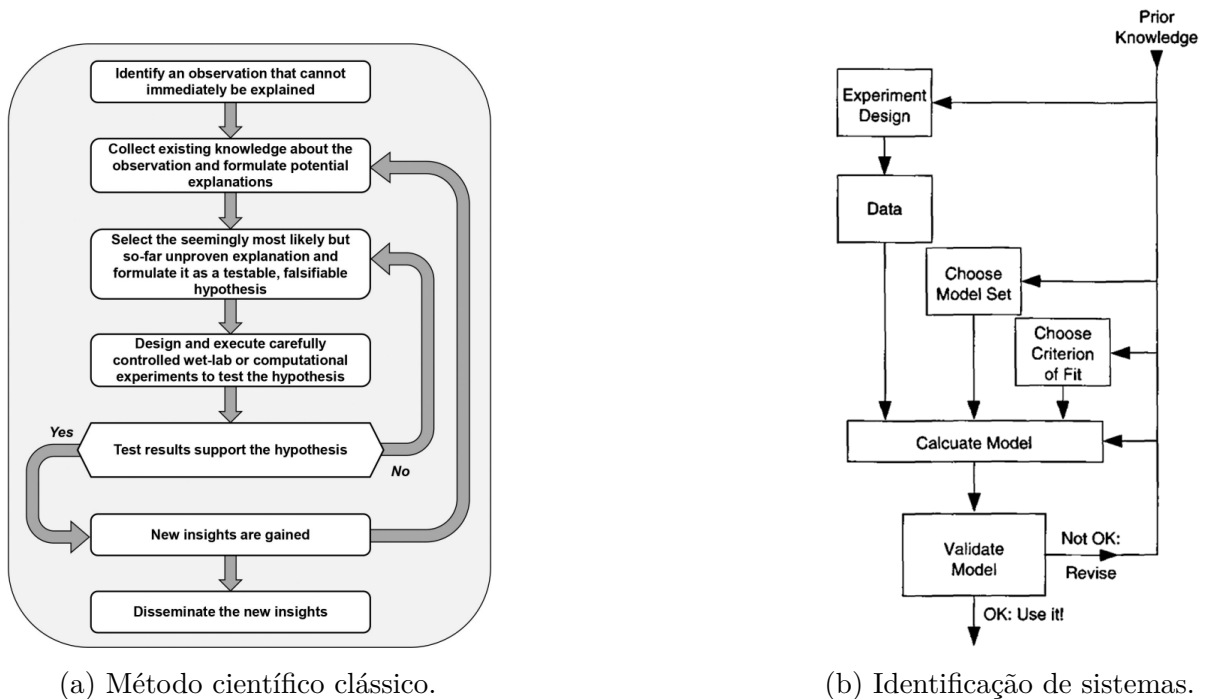
	Lista de Abreviaturas e Siglas	17
	Lista de Símbolos	19
1	INTRODUÇÃO	23
1.1	Objetivos	25
1.1.1	Objetivo geral	25
1.1.2	Objetivos específicos	25
1.2	Justificativa do trabalho	25
1.3	Procedimentos metodológicos	26
1.4	Estrutura do trabalho	26
2	REVISÃO BIBLIOGRÁFICA	29
2.1	Embasamento matemático e estatístico	29
2.1.1	Ergodicidade	29
2.1.2	Funções de correlação, autocorrelação e espectro	30
2.1.3	Variáveis aleatórias	31
2.1.3.1	Ruído branco gaussiano	32
2.2	Conhecimento prévio do sistema	35
2.2.1	Definições relacionadas à linearidade	35
2.2.1.1	Linearidade	36
2.2.1.2	Causalidade	37
2.2.1.3	Invariância no tempo	37
2.2.2	Classificação por nível de conhecimento	38
3	METODOLOGIA	41
3.1	Projeto do sinal de excitação	41
3.1.1	Sinal aleatório	42
3.1.2	Pseudo Random Binary Sequence (PRBS)	43
3.1.2.1	Geração de um PRBS	45
3.1.3	Pseudo Random Multi Level Sequence (PRMLS)	46
3.2	Métodos de representação	48
3.3	Teoria da estimação	50
3.3.1	O estimador de mínimos quadrados	50
3.3.2	O Método dos Mínimos Quadrados (MMQ) recursivo	54

3.3.3	Uso do Least Absolute Shrinkage and Selection Operator (LASSO) como ferramenta de seleção de regressores	56
3.3.3.1	Parâmetro de penalização λ	59
3.4	Validação dos modelos	61
4	EXPERIMENTOS	63
4.1	Sinal PRBS e degrau unitário	63
4.1.1	Comparação em um circuito RC: simulação computacional	65
4.2	Exemplo de identificação simples: circuito RC	68
4.3	Alteração paramétrica durante a identificação	68
4.4	Exemplo de identificação simples: dinâmica de temperatura	69
4.4.1	Influência do overfitting	73
4.4.2	Uso do LASSO como ferramenta de seleção de regressores	74
4.5	Identificação de sistemas não lineares	76
5	ESTUDO DE CASO: TWIN ROTOR	79
5.1	Discretização	81
5.2	Estimação de parâmetros	83
5.2.1	Identificação do modelo	84
6	CONCLUSÃO	87
6.1	Trabalhos futuros	87
	REFERÊNCIAS	89

1 INTRODUÇÃO

A capacidade de modelar com precisão a dinâmica de um sistema físico é crucial para diversas aplicações de engenharia, como o desenvolvimento de estratégias de controle, a otimização do desempenho e o monitoramento preditivo. Neste contexto, a identificação de sistemas é um procedimento fundamental que permite construir representações matemáticas de seus objetos de estudo a partir da análise de dados experimentais [Ljung 1999, Chen 1998]. Essa abordagem apresenta semelhanças conceituais com o método científico tradicional, partindo do princípio de que visa inferir os princípios construtivos de seu sistema de interesse. Tal comportamento é apresentado na figura 1. Ambas as práticas se fundamentam em observações e evidências empíricas para formular e validar um modelo ou hipótese que descreva de forma satisfatória os fenômenos observados, adotando um ciclo contínuo de validação e reformulação [Ljung 1999, Voit 2019].

Figura 1 – Comparação entre os processos relacionados ao método científico e à identificação de sistemas.



(a) Método científico clássico.

(b) Identificação de sistemas.

Fonte: [Voit 2019] (1a) e [Ljung 1999] (1b).

Ainda que similares, ambas as metodologias divergem fundamentalmente: o método científico tende a buscar a formulação que melhor represente seu objeto de estudo, enquanto a modelagem, no âmbito da identificação de sistemas, necessita apenas de uma representação que descreva os fenômenos de maneira satisfatória para a aplicação de interesse [Voit 2019, Chen 1998]. Essa diferença, ainda que sutil, é de suma importância: há inúmeros

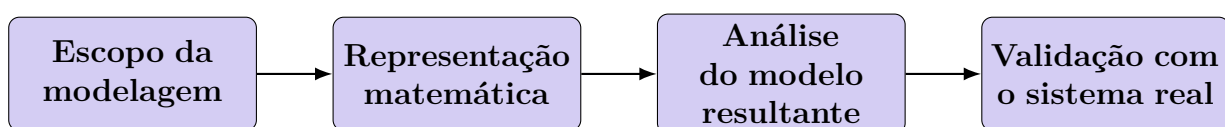
parâmetros e variáveis envolvidos na dinâmica de qualquer sistema, e até mesmo o mais simples dos resistores pode ser modelado tanto a partir da medição empírica de sua resistência quanto em função de suas diferentes curvas características [Ford e Hasbun 2019].

Neste contexto, a identificação de sistemas é tipicamente dividida em metodologias paramétricas e não paramétricas. As técnicas não paramétricas incluem a análise de resposta em frequência, utilizada para obter diagramas de Bode ou Nyquist, bem como o cálculo das funções de correlação e de coerência. Tais funções fornecem *insights* valiosos sobre a relação entre as variáveis de entrada e de saída, bem como sobre a influência do ruído nos dados observados. Mais recentemente, abordagens baseadas em redes neurais também ganharam destaque, aplicadas principalmente em sistemas com forte não-linearidade ou com poucas características conhecidas. Dessa maneira, embora este trabalho tenha como foco o uso de técnicas paramétricas, devido às suas capacidades de predição e simulação mais diretas, é essencial reconhecer a relevância e o uso das metodologias não paramétricas na prática da engenharia.

Assim, considerando o equilíbrio entre o detalhamento do modelo e a aplicação prática, não é factível obter uma representação excessivamente complexa, devido a limitações relacionadas a custos dos mais diversos tipos: financeiros, computacionais, temporais, etc. [Chen 1998]. Além disso, a abordagem escolhida deve considerar demais limitações, tais como os pontos onde podem ser realizadas as aferições, os desgastes relacionados ao processo de identificação, a inviabilidade de alteração do sinal de entrada e afins [Isermann e Münchhof 2010]. Dessa maneira, faz-se necessário delimitar as variáveis e os processos mais significativos, de modo a obter uma modelagem factível e com um custo associado satisfatório.

Assim sendo, um dos métodos disponíveis para encontrar tal representação paramétrica pode ser obtido a partir de um estudo analítico composto por quatro etapas distintas: escopo da modelagem, representação matemática, análise do modelo resultante e validação com o sistema real [Chen 1998], como ilustrado na figura 2. Outra abordagem apresentada em [Ljung 1999] adiciona uma etapa paralela à modelagem, envolvendo não apenas o objeto a ser estudado, mas também a definição de todo o experimento a ser realizado.

Figura 2 – Estudo analítico proposto por Chen.



Fonte: [Chen 1998].

Por fim, vale reforçar que tal procedimento, assim como a própria metodologia científica, é iterativo. Não é incomum, portanto, ser necessário ajustar os parâmetros ou obter mais dados sobre o objeto de estudo, a fim de eliminar discrepâncias ou aprimorar o modelo utilizado [Ljung 1999]. Caso necessário, deve-se considerar a possibilidade de efetuar um pré-processamento dos dados obtidos, tendo em vista eventuais ruídos e distúrbios inerentes ao sistema ou ao processo de medição [Juang 1994].

1.1 Objetivos

1.1.1 Objetivo geral

O presente trabalho propõe-se a desenvolver e avaliar técnicas de identificação aplicáveis a sistemas dinâmicos, tanto lineares quanto não lineares.

1.1.2 Objetivos específicos

Em específico, são estabelecidas as seguintes metas para o desenvolvimento deste trabalho:

- Realizar um estudo aprofundado das principais técnicas de identificação de sistemas, bem como da teoria de estimação e métodos numéricos necessários;
- Implementar algoritmos de identificação de sistemas, tanto para modelos lineares quanto para modelos não-lineares, em ambiente de simulação computacional;
- Avaliar o desempenho comparativo entre diferentes técnicas de identificação de sistemas;
- Aplicar e validar os algoritmos de identificação em uma planta do tipo *twin-rotor* em um ambiente de simulação computacional;
- Documentar os resultados e as conclusões obtidos.

1.2 Justificativa do trabalho

A modelagem precisa e eficiente de sistemas dinâmicos é imperativa para a engenharia moderna, aplicando-se a contextos variados: desde o monitoramento de processos industriais até a utilização em sistemas de controle de redes de energia elétrica [Aguirre 2004]. Outra de suas aplicações notáveis está no projeto de sistemas de controle para diversas aplicações, como drones e até mesmo em plantas de caráter educacional, como o *twin-rotor*. A identificação de sistemas surge, portanto, como a ferramenta ideal para construir representações matemáticas a partir de dados empíricos, superando as limitações e os custos da modelagem puramente física. O desafio reside, contudo, em selecionar a

metodologia de estimação que ofereça o melhor equilíbrio entre a acurácia preditiva e a simplicidade computacional.

Neste sentido, a relevância do trabalho reside na análise comparativa e na aplicação de diversas técnicas de modelagem paramétrica de sistemas dinâmicos. Essa abordagem abrange desde a seleção adequada do sinal de excitação e da estrutura de representação até a escolha de estimadores, buscando sempre alternativas que apresentem características intrínsecas de rejeição ao ruído. Especificamente para os estimadores, o estudo utiliza o MMQ como referência para a implementação do LASSO; assim, é possível analisar quantitativamente o equilíbrio entre o custo computacional e a precisão do modelo obtido.

1.3 Procedimentos metodológicos

Inicialmente, foi realizada uma revisão bibliográfica da literatura existente, centrada nas técnicas usuais e recomendadas para a identificação correta de sistemas. Também foram considerados estudos de caso com foco em estratégias empregadas em plantas do tipo *twin-rotor*. Posteriormente, foram avaliadas as fundamentações teóricas e matemáticas necessárias à aplicação das metodologias estudadas, com enfoque em funções de correlação e autocorrelação.

Dentre as estratégias aplicadas, destacam-se a importância do uso de sinais persistentemente excitantes como entrada do sistema de interesse, bem como a aplicação do MMQ e a análise comparativa com o LASSO. Foco também na utilização do LASSO como ferramenta de seleção de regressores, devido ao seu grande potencial para mitigar problemas decorrentes do *overfitting*.

Desse modo, foi realizada a validação computacional dessas metodologias em sistemas lineares e não lineares, analisando a robustez e o nível de acurácia dos estimadores em função dos parâmetros e dos procedimentos empregados.

Por fim, os conhecimentos teóricos e práticos obtidos foram aplicados para identificar uma planta do tipo *twin-rotor*, o que demonstra a eficácia das metodologias estudadas em sistemas de alta complexidade.

1.4 Estrutura do trabalho

- Capítulo 1 - Introdução ao tema, bem como uma breve comparação entre o método científico e a identificação de sistemas. Proposta do trabalho, justificativas e objetivos;
- Capítulo 2 - Apresentação de conceitos necessários para um bom entendimento do trabalho;
- Capítulo 3 - Apresentação de técnicas de identificação de sistemas. Foco na escolha do sinal de entrada, do método de representação e dos estimadores utilizados;

- Capítulo 4 - Simulação e estimação de sistemas lineares e não lineares em ambiente computacional. Aplicação prática em sistemas de primeira ordem com características lineares e não lineares;
- Capítulo 5 - Identificação do sistema *twin-rotor*;
- Capítulo 6 - Conclusões finais do trabalho.

2 REVISÃO BIBLIOGRÁFICA

A identificação de sistemas dinâmicos apoia-se em diversas áreas, desde a teoria dos sinais até as ferramentas estatísticas e matemáticas. É igualmente importante, embora não necessário, possuir um certo conhecimento do objeto e do objetivo de estudo, conforme abordado no capítulo 1. Nesse âmbito, é interessante abordar estudos de caso notáveis na literatura: por exemplo, o *twin-rotor* é um sistema altamente não linear e acoplado, constituindo um excelente ambiente de testes para a aplicação de técnicas de identificação e controle.

Considerando os desafios apresentados anteriormente, dentre as técnicas usualmente utilizadas para obter uma representação do *twin-rotor* encontram-se: modelagem matemática padrão, identificação de sistemas do tipo caixa-preta, uso de algoritmos heurísticos e técnicas híbridas. Na mesma ordem anterior, os artigos [Rahideh e Shaheed 2008], [Ahmad e Chipperfield 2002], [Darus, Aldebrez e Tokhi 2004] e [Huu e Ismail 2016] são representativos de cada uma das metodologias apresentadas.

Por fim, o controle escolhido para o sistema beneficia-se diretamente da acurácia do modelo e também é objeto de estudo na literatura. Justifica-se, portanto, o uso de controladores Proporcional, Integral e Derivativo (PID) [Juang, Huang e Liu 2008] ou Fuzzy [Zeghlache *et al.* 2022], demonstrando eficácia na estabilização das dinâmicas diante de distúrbios e incertezas.

2.1 Embasamento matemático e estatístico

Inicialmente, faz-se importante definir conceitos relacionados à matemática e à estatística, imprescindíveis para o correto desenvolvimento do trabalho proposto. Demais formulações amplamente difundidas podem ser encontradas em [Grimmett e Stirzaker 2001], [Walpole *et al.* 2011] e [Hastie, Tibshirani e Wainwright 2015], ainda que as mais utilizadas sejam apresentadas ao longo do desenvolvimento do trabalho.

2.1.1 Ergodicidade

A definição de ergodicidade relaciona valores temporais a valores estatísticos, em especial ao comprovar a equivalência entre a esperança matemática (média estocástica) e a média temporal. Desse modo, para que um processo estocástico estacionário seja considerado ergódico, a média temporal de uma sequência $a[k]$ deve convergir para a sua esperança matemática em um número de amostras N que tende ao infinito [Birkhoff 1931]. Na prática, considerando N suficientemente grande, a média temporal é uma estimativa aceitável da média estocástica conforme a eq. (2.1):

$$\mathbb{E}[a[k]] = \frac{1}{N} \sum_{k=1}^N a[k]. \quad (2.1)$$

Do ponto de vista da identificação de sistemas, a ergodicidade é crucial por estabelecer a equivalência prática entre a média temporal de uma única e longa medição e a média estocástica de diversas medições da mesma dinâmica. Esse princípio é altamente desejável, pois permite obter estimativas estatisticamente consistentes dos parâmetros do sistema a partir de um único experimento, superando a inviabilidade de repetir medições sob as mesmas condições exatas. Contudo, essa propriedade não é universal, sendo crucial reconhecer que ela não se aplica a todos os sistemas físicos: notadamente, tal princípio é impraticável para dinâmicas altamente sensíveis às condições iniciais ou que exibem processos irreversíveis [Aguirre 2004].

2.1.2 Funções de correlação, autocorrelação e espectro

Dadas duas sequências finitas e ergódicas quaisquer $a[k]$ e $b[k]$ de comprimento N , o estimador da função de correlação cruzada $r_{ab}[\tau]$ para um deslocamento temporal τ é dado por eq. (2.2). Desse modo, r_{ab} representa o nível de similaridade entre a primeira sequência e uma versão defasada da segunda [Bakalis *et al.* 2023, Theodorsen 2016].

$$r_{ab}[\tau] = \frac{1}{N} \sum_{i=1}^N a[i] \cdot b[i + \tau]. \quad (2.2)$$

De maneira análoga, a função de autocorrelação mede o nível de similaridade entre uma sequência e uma versão deslocada de si mesma. Assim sendo, o estimador $r_{aa}[\tau]$ é obtido ao substituir $b[i + \tau]$ por $a[i + \tau]$ na eq. (2.2), resultando na formulação 2.3. Também é válido destacar que $r_{aa}[\tau]$ é uma função par para processos reais, conforme dinâmica representada na eq. (2.4). Considerando essa simetria, usualmente se representam apenas os valores de r_{aa} para $\tau \geq 0$.

$$r_{aa}[\tau] = \frac{1}{N} \sum_{i=1}^N a[i] \cdot a[i + \tau], \quad (2.3)$$

$$r_{aa}[\tau] = r_{aa}[-\tau]. \quad (2.4)$$

Outra abordagem interessante é averiguar que, pela função de autocorrelação comparar uma função com si mesma, ela também pode ser utilizada para obter *insights* sobre a taxa de variação do sinal de interesse em função do tempo. Isso ocorre visto que variações mais lentas implicam uma similaridade maior e, portanto, uma autocorrelação maior entre amostras subsequentes [Aguirre 2004].

Vale ressaltar que todo o desenvolvimento apresentado parte do princípio de que as sequências em estudo possuem caráter ergódico. Tal suposição permite o uso das variações em função do tempo discreto k , visto que as formulações usuais de correlação e autocorrelação dependem do operador \mathbb{E} (esperança matemática).

Por fim, a relação entre a correlação e o espectro é formalizada pelo Teorema de Wiener-Khinchin. Ao aplicar a transformada de Fourier ao estimador da função de correlação e de autocorrelação, obtemos as funções de espectro cruzado (ϕ_{ab}) e autoespectro (ϕ_{aa}):

$$\begin{cases} \phi_{ab}(e^{j\omega}) = \mathcal{F}[r_{ab}[k]], \\ \phi_{aa}(\omega) = \mathcal{F}[r_{aa}[k]]. \end{cases} \quad (2.5)$$

Assim como é possível estimar a velocidade de uma dinâmica a partir de sua autocorrelação, o autoespectro ϕ_{aa} indica como a energia ou potência do sinal analisado se distribui no domínio da frequência [Aguirre 2004]. Tal constatação também decorre do Teorema de Wiener-Khinchin, visto que ϕ_{aa} apresenta a mesma expressão que a Power Spectral Density (PSD).

2.1.3 Variáveis aleatórias

Por definição, uma variável aleatória é uma função $X : \Omega \rightarrow \mathbb{R}$ com propriedades dadas por $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}_a$ para todo $x \in \mathbb{R}$. Ou seja, variáveis aleatórias são, formalmente, funções mensuráveis (ou \mathcal{F}_a -mensuráveis) que associam cada resultado ω do espaço amostral Ω a um número real x . A condição apresentada garante que o evento $X(\omega) \leq x$ é um evento bem definido e, portanto, sua probabilidade pode ser calculada no espaço de probabilidade $(\Omega, \mathcal{F}_a, P)$ [Grimmett e Stirzaker 2001].

No contexto de processamento de sinais discretos, uma variável aleatória é frequentemente representada por uma sequência numérica aleatória ou por um processo estocástico discreto $\nu[k]$, em que k refere-se ao tempo discreto ou ao índice de uma dada amostra. Para analisar as propriedades de tais sequências, utilizam-se ferramentas como a adimensionalização, a flutuação e, crucialmente, a autocorrelação.

Primeiramente, considerando uma sequência $a[k]$ qualquer, sua adimensionalização, dada por $\tilde{a}[k]$, consiste em dividir $a[k]$ pelo seu valor máximo.

$$\tilde{a}[k] = \frac{a[k]}{\|a[k]\|_\infty}. \quad (2.6)$$

Além disso, a flutuação $a'[k]$ de uma sequência numérica $a[k]$ é a diferença entre $a[k]$ e seu valor médio $\bar{a}[k]$, representando o comportamento simétrico de $a[k]$ em relação ao eixo das abcissas.

$$a'[k] = a[k] - \bar{a}[k]. \quad (2.7)$$

No âmbito deste trabalho, uma sequência numérica $\nu[k]$ será considerada representativa de uma variável aleatória e, conseqüentemente, com o perfil de ruído branco, se sua flutuação $\nu'[k]$ apresentar uma função de autocorrelação $\tilde{r}_{\nu'\nu'}[\tau]$ com característica impulsiva. Dessa forma, $\tilde{r}_{\nu'\nu'}[\tau]$ é essencialmente não correlacionada com as defasagens $\tau \neq 0$, conforme apresentado na eq. (2.8).

$$\begin{cases} \tilde{r}_{\nu'\nu'}[0] = 1, \\ \tilde{r}_{\nu'\nu'}[\tau] = 0 \quad \forall \tau \neq 0. \end{cases} \quad (2.8)$$

Devido à natureza estatística inerente à função de autocorrelação obtida a partir de uma sequência de amostras com tamanho N , a condição ideal apresentada na eq. (2.8) pode ser relaxada: para que uma sequência finita seja considerada representativa de um ruído branco, basta que seus valores apresentem um índice de confiança de pelo menos 95%. Assim sendo, os valores de $\tilde{r}_{\nu'\nu'}[\tau]$ para $\tau \neq 0$ devem estar contidos no intervalo $\pm \frac{1.96}{\sqrt{N}}$ para que as características de interesse sejam garantidas, conforme a eq. (2.9) [Aguirre 2004].

Vale reforçar que tal definição, por meio de intervalos de confiança, é de caráter prático, sendo apenas indicativa da presença de aleatoriedade. A definição da autocorrelação para uma sequência aleatória com caráter ergódico é dada por eq. (2.8).

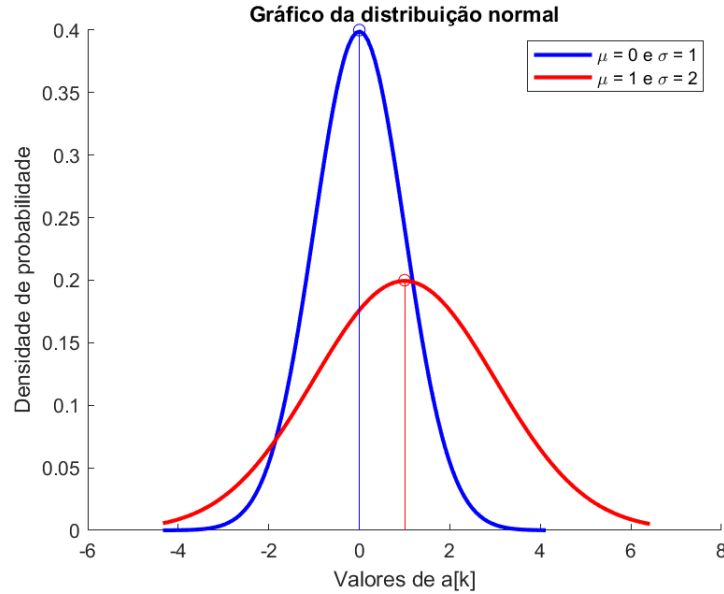
$$\begin{cases} \tilde{r}_{\nu'\nu'}[0] = 1, \\ -\frac{1.96}{\sqrt{N}} < \tilde{r}_{\nu'\nu'}[\tau] < \frac{1.96}{\sqrt{N}} \quad \forall \tau \neq 0. \end{cases} \quad (2.9)$$

2.1.3.1 Ruído branco gaussiano

Uma classe importante de ruído branco, amplamente utilizada em simulações computacionais devido às suas propriedades analíticas e à facilidade de geração, é o ruído branco gaussiano. O mesmo é obtido por uma sequência de variáveis aleatórias do tipo independente e identicamente distribuída (i.i.d.) que seguem a distribuição normal (ou gaussiana), conforme a Função de Densidade de Probabilidade (PDF) dada em eq. (2.10). A figura 3 ilustra o comportamento da distribuição normal em função dos parâmetros μ_N e σ_N .

$$\mathcal{N}(k, \mu_N, \sigma_N^2) = \frac{1}{\sqrt{2\pi\sigma_N^2}} \cdot \exp \left[-\frac{(k - \mu_N)^2}{2\sigma_N^2} \right]. \quad (2.10)$$

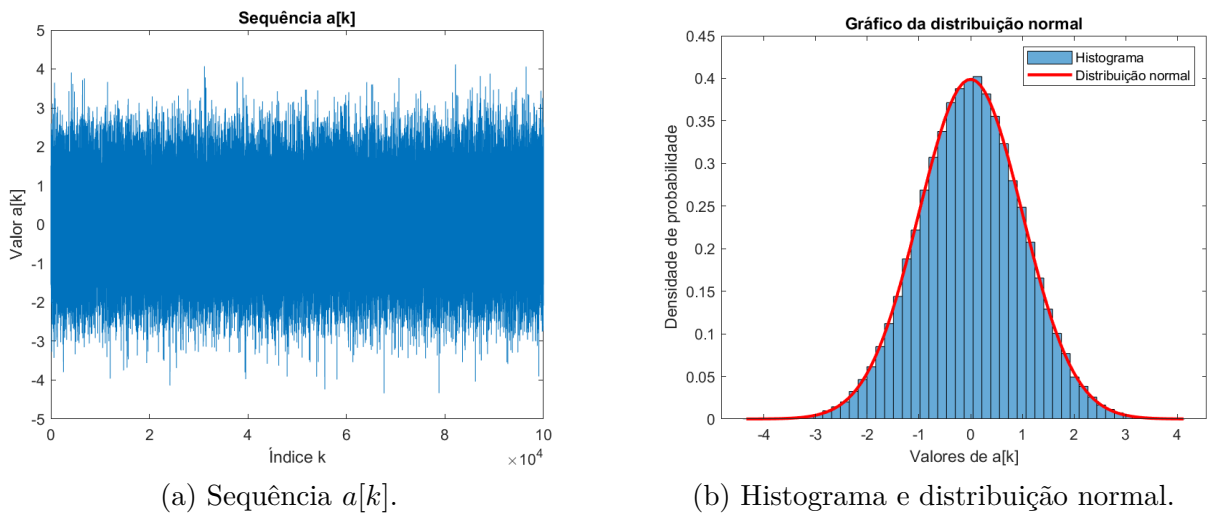
Figura 3 – PDF da distribuição normal em função dos parâmetros μ_N (média) e σ_N (desvio padrão).



Fonte: Autoria própria.

Gerando uma sequência $a[k]$ de tamanho 10^5 a partir do comando `randn()` no MATLAB e, posteriormente, agrupando os valores em 50 intervalos distintos, é possível calcular a função de densidade de probabilidade do sinal de interesse. Tal procedimento é realizado ao averiguar o percentual de amostras de $a[k]$ em cada um dos intervalos definidos, criando, assim, um histograma do processo estudado. Esse procedimento é apresentado na figura 4, na qual a distribuição amostral é comparada à PDF teórica da função gaussiana para $\mu_N = 0$ e $\sigma_N = 1$.

Figura 4 – Validação Estatística de uma Sequência Aleatória Gerada no Matlab: Ruído Branco Gaussiano.

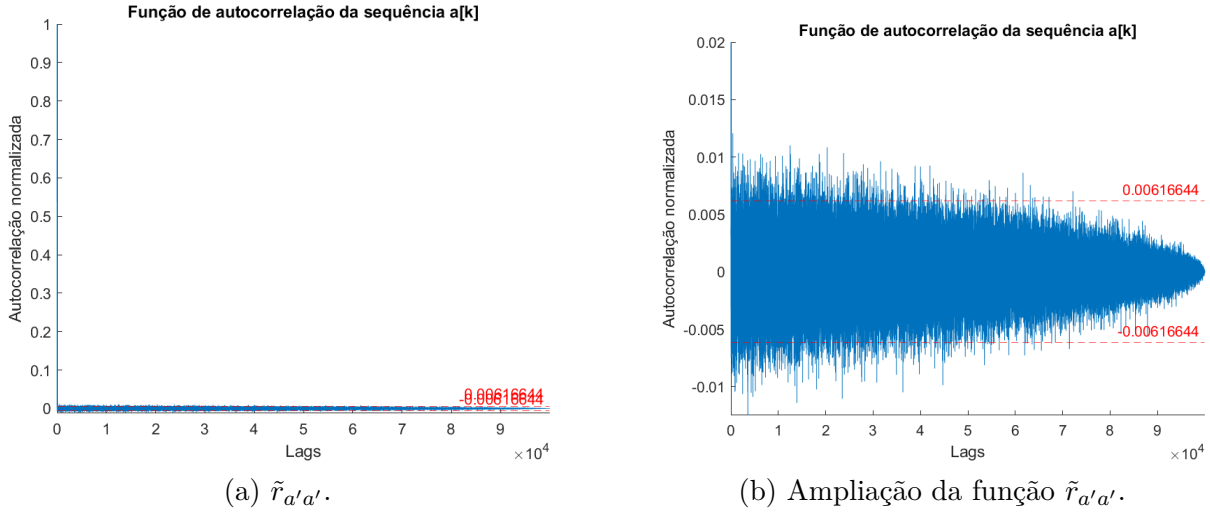


Fonte: Autoria própria.

O perfil impulsivo do sinal apresentado na figura 5, com valores próximos de zero

para $\tau \neq 0$, confirma que a sequência gerada no Matlab pode ser modelada como amostras de um ruído branco.

Figura 5 – Autocorrelação da sequência $a[k]$.



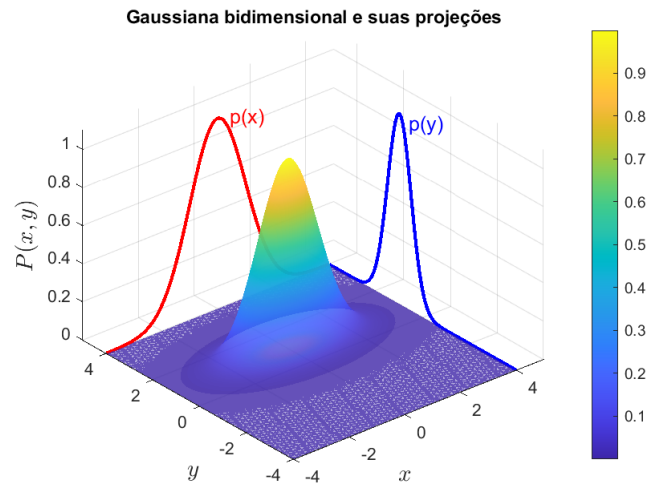
Fonte: Autoria própria.

Vale destacar, ainda, que os intervalos de confiança, denotados pelas linhas tracejadas em vermelho, são apenas indicativos de que a sequência de interesse possui caráter aleatório. Portanto, ainda que se observe uma certa extrapolação dos limites impostos em eq. (2.9) na figura 5b, é razoável considerar que o comando *randn()* do Matlab gera um ruído branco do tipo gaussiano.

Por fim, é necessário apresentar também o conceito de gaussianas multidimensionais. Essa abordagem é útil para modelar a distribuição conjunta de múltiplas variáveis aleatórias, em que a correlação entre elas desempenha um papel tão importante quanto suas médias e variâncias individuais. Considerando, então, que será estimada uma série de parâmetros, cada um relaciona-se a uma distribuição normal distinta, de modo que a distribuição resultante é completamente caracterizada pelo vetor de médias μ e pela matriz de covariância Σ . Essa matriz captura as variâncias e, crucialmente, as correlações entre os parâmetros, essenciais para modelar a distribuição conjunta dos parâmetros estimados neste trabalho, conforme ilustrado na figura 6 e na eq. (2.11).

$$\mathcal{N}(\mathbf{x}|\mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu)\right). \quad (2.11)$$

Figura 6 – Distribuição normal multidimensional - visualização 3D. As funções $p(x)$ e $p(y)$ são distribuições gaussianas marginais.



Fonte: Autoria própria.

2.2 Conhecimento prévio do sistema

Assim como abordado no capítulo 1, um certo nível de conhecimento prévio, tanto do objeto quanto do objetivo do estudo, é uma parte importante do processo de identificação, ainda que não seja essencial para uma modelagem satisfatória. Dessa forma, o quadro 1 agrega metodologias de classificação comumente utilizadas na literatura, que podem ser empregadas para definir o sistema de interesse.

Além disso, as seções 2.2.1 e 2.2.2 abordam, em maior detalhe, conceitos de suma importância para a seleção e validação de técnicas de identificação adequadas.

2.2.1 Definições relacionadas à linearidade

Embora seja comum agrupar sistemas em lineares e não lineares, existem inúmeras outras definições que devem ser atendidas, de modo a viabilizar a aplicação de diversas técnicas e simplificações, métodos esses comumente associados a sistemas lineares. Dessa forma, faz-se necessário definir não só o conceito de linearidade, mas também as demais propriedades de **sistemas lineares, causais e invariantes no tempo**.

Vale também destacar que sinais e sistemas são intrinsecamente conectados, embora as definições agregadas em 2.2.1 sejam utilizadas para categorizar o segundo.

Quadro 1 – Classificações comuns de sistemas dinâmicos [Aguirre 2004, Ljung 1999, Ogata 2010]

Critério de Classificação	Tipo	Descrição
Número de entradas e saídas	SISO	Uma entrada e uma saída (Single Input Single Output)
	MIMO	Múltiplas entradas e múltiplas saídas (Multiple Input Multiple Output)
Natureza da dinâmica	Determinístico	Saída totalmente previsível a partir da entrada
	Estocástico	Saída depende, total ou parcialmente, de características aleatórias
Realimentação	Malha aberta	Não há uso da saída no cálculo do sinal de controle
	Malha fechada	A saída é realimentada ao sistema de controle
Estabilidade	Estável	As saídas tendem a permanecer limitadas ao longo do tempo
	Instável	As saídas divergem ou crescem indefinidamente

2.2.1.1 Linearidade

Segundo [Ogata 2010], a linearidade de um sistema é definida pela combinação de duas propriedades: a superposição e a homogeneidade. Pelo princípio da superposição, a resposta a uma soma de sinais de entrada é a soma das respostas individuais a cada sinal (equação 2.12), enquanto a homogeneidade define que a resposta a um sinal de entrada escalonado por uma constante é a resposta individual escalonada pela mesma constante (2.13).

Matematicamente, se $y_i(t)$ é a saída correspondente ao sinal de entrada $u_i(t)$, as propriedades são expressas como:

$$\begin{cases} u_1(t) \Rightarrow y_1(t), \\ u_2(t) \Rightarrow y_2(t), \\ \therefore u_1(t) + u_2(t) \Rightarrow y_1(t) + y_2(t). \end{cases} \quad (2.12)$$

$$\begin{cases} u_1(t) \Rightarrow y_1(t), \\ \therefore \alpha \cdot u_1(t) \Rightarrow \alpha \cdot y_1(t). \end{cases} \quad (2.13)$$

A linearidade é satisfeita quando ambas as condições são combinadas em uma única relação:

$$\alpha_1 u_1(t) + \alpha_2 u_2(t) \Rightarrow \alpha_1 y_1(t) + \alpha_2 y_2(t). \quad (2.14)$$

Apesar de nem todos os sistemas serem inerentemente lineares, a linearização é uma metodologia poderosa. Ao assumir um modelo aproximado em torno de um ponto de operação, é possível modelar satisfatoriamente uma ampla gama de sistemas não lineares. Dessa forma, tal metodologia é uma técnica simples e suficientemente precisa para atender a muitos casos de interesse [Ogata 2010].

2.2.1.2 Causalidade

O princípio da causalidade estabelece que a resposta (saída) de um sistema no instante t_0 pode depender apenas dos sinais de entrada e das saídas em instantes anteriores ou no próprio instante t_0 . Em outras palavras, um sistema causal não pode ter uma resposta prévia a um evento futuro.

Formalmente, sistemas lineares e causais obedecem à relação dada em eq. (2.15), na qual $h[k]$ é a resposta ao impulso do referido sistema. Portanto, como a parcela de qualquer termo $u[i] \cdot h[k - i]$ para $i > t$ é nula, é possível substituir o limite superior da somatória de convoluções por k .

$$y[k] = \sum_{i=0}^{\infty} u[i] \cdot h[k - i] = \sum_{i=0}^k u[i] \cdot h[k - i]. \quad (2.15)$$

Uma implicação da causalidade na análise em frequência é que a função de transferência apresenta um número de polos maior ou igual ao de zeros; caso contrário, haveria dependência de amostras futuras da entrada, o que provocaria um ganho em alta frequência que tende ao infinito.

Também é importante notar que sistemas físicos reais, como o *twin-rotor*, são intrinsecamente causais: isso ocorre porque a inércia e outros fatores físicos fazem com que o sistema apresente um comportamento semelhante ao de um filtro passa-baixa, impedindo que a saída responda instantaneamente a mudanças bruscas na entrada. Embora sistemas teóricos ou computacionais (como no pós-processamento de dados) possam violar a causalidade, esta é um requisito fundamental para a modelagem de sistemas físicos usuais [Ogata 2010].

2.2.1.3 Invariância no tempo

De forma simplificada, um sistema invariante no tempo apresenta coeficientes (como α_1 e α_2 na formulação 2.14) independentes do tempo. Formalmente, se um sinal

de entrada $u_1(t)$ produz uma saída $y_1(t)$, então um atraso τ na entrada deve resultar no mesmo atraso na saída, de modo que os coeficientes associados a ambos, $u_1(t - \tau)$ e $y_1(t - \tau)$, permaneçam inalterados.

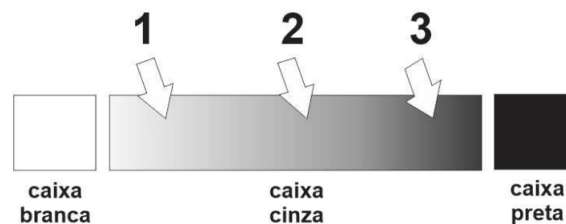
Ainda assim, a variação paramétrica é comum em sistemas físicos, podendo ser lenta (como o desgaste de componentes) ou rápida (como a ejeção de massa em um foguete). No entanto, para a maioria dos sistemas e em um horizonte de tempo determinado, considerar o sistema invariante no tempo é uma suposição aceitável [Ljung 1999].

Por outro lado, quando tal variação exerce uma grande influência no modelo utilizado, pode-se empregar técnicas de identificação em tempo real para atualizar dinamicamente os parâmetros variáveis [Ljung 1999]. Essas técnicas, frequentemente associadas ao controle adaptativo e à identificação em malha fechada, incluem o Filtro de Kalman e o Método dos Mínimos Quadrados Recursivo (MMQR).

2.2.2 Classificação por nível de conhecimento

Por fim, diversos métodos de representação exigem que as equações matemáticas definidoras do sistema de estudo sejam conhecidas. Faz-se necessário, portanto, categorizar os sistemas com base no nível de informação que é possível obter a partir deles, o que possibilitará, posteriormente, a escolha de uma representação adequada.

Figura 7 – Escalas de tons cinzentos que ilustram diferentes abordagens “caixa-cinza”. Da região 1 à 3, a quantidade de informação sobre o sistema diminui.



Fonte: [Fernandes, Oliveira e Souto 2017].

Primeiramente, métodos de identificação do tipo caixa branca pressupõem conhecimento total das dinâmicas do sistema de interesse, sendo muito utilizados em sistemas de menor complexidade. Dessa forma, embora seja possível aplicar técnicas de estimação paramétrica, também seria plausível estudar a planta de interesse de forma completamente analítica [Ljung 1999].

Por outro lado, métodos relacionados a um sistema de caixa-preta consideram que não há nenhum conhecimento prévio do objeto de estudo, sendo possível obter apenas amostras do sinal de entrada e do sinal de saída do mesmo. Assim, como abordado em [Chen 1998], pode ser necessária uma grande quantidade de amostras e um poder computacional

considerável. Ainda que seja possível encontrar uma modelagem satisfatória e robusta, tanto a perturbações quanto a alterações sistêmicas.

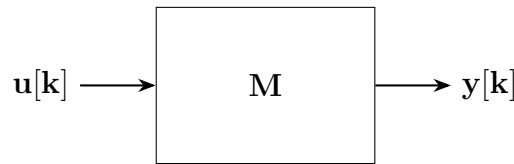
Por fim, métodos de caixa cinza consideram que as características do sistema de interesse estão parcialmente obscurecidas, seja por desconhecimento do funcionamento interno, seja pela dificuldade de mensurar alguns de seus parâmetros e dinâmicas. Dessa forma, é comum recorrer a modelagens físicas já conhecidas, ainda que parciais ou simplificadas [Aguirre 2004].

Vale também destacar que é possível utilizar diversas técnicas para agregar as informações previamente conhecidas às demais informações do sistema, seja por meio da correta definição dos regressores, do teorema de Bayes, da formulação matemática do filtro de Kalman ou até mesmo de redes neurais de diversas complexidades.

3 METODOLOGIA

Nesta seção, serão apresentados os métodos a partir dos quais serão efetuadas as medições e o processamento dos dados obtidos a partir do sistema de interesse, bem como os índices com base nos quais é possível validar o modelo obtido.

Figura 8 – Sistema M do tipo SISO, apresentando entrada $u[k]$ e saída $y[k]$.



Fonte: Autoria própria.

3.1 Projeto do sinal de excitação

Tendo em vista que é possível escolher a entrada do sistema de interesse, há uma ampla gama de sinais que podem ser utilizados para a identificação. Dentre os mais comuns, estão o uso da resposta ao degrau, da resposta impulsiva e da resposta em frequência do sistema por meio de entradas senoidais. Vale destacar, porém, que, embora amplamente utilizado, o método de obtenção da resposta em frequência de um sistema consiste na aquisição de um grande volume de dados.

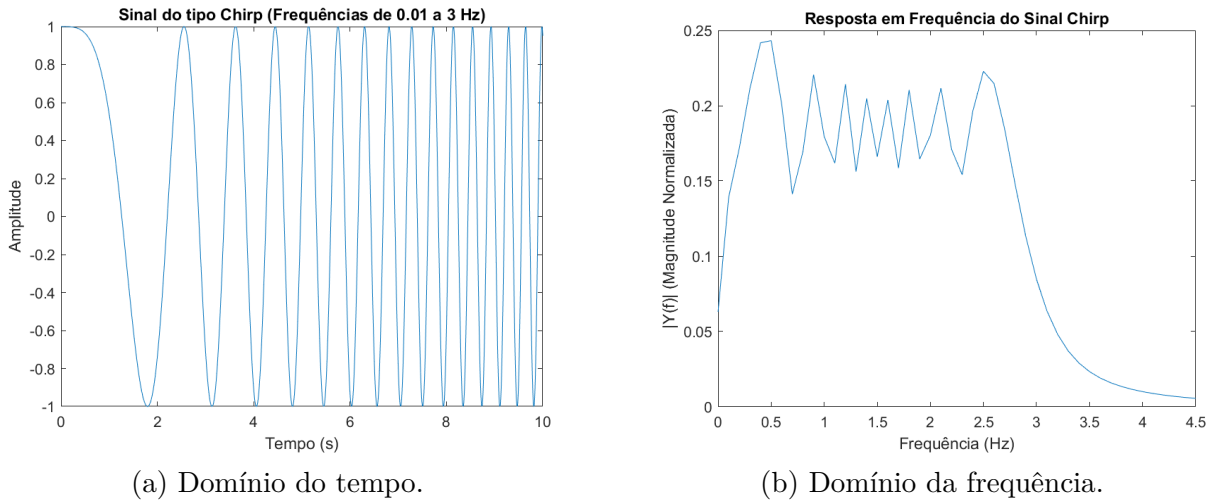
Tal afirmação parte do princípio de que a solução para excitar todo o espectro de frequências de forma satisfatória consiste, usualmente, na utilização de um conjunto de funções senoidais de frequências f_0 distintas, cada uma excitando uma frequência específica do sistema. Essa dinâmica é representada pela eq. (3.1).

$$\begin{cases} u_{\sin}(t) = \sin(2\pi f_0 t), \\ U_{\sin}(f) = \mathcal{F} \left[u_{\sin}(t) \right] = \frac{1}{2j} \left[\delta(f - f_0) - \delta(f + f_0) \right]. \end{cases} \quad (3.1)$$

Uma alternativa mais eficiente é utilizar sinais com um espectro de frequências mais rico, tais como sinais *chirp* (eq. (3.2) e figura 9) ou sinais de natureza aleatória. Em particular, para sinais aleatórios, tal afirmação baseia-se no fato de que um ruído branco contém informações em todo o espectro de frequências. Desse modo, a utilização de sinais de entrada bem projetados permite reduzir drasticamente a quantidade de dados necessários para obter uma resposta em frequência satisfatória para o sistema de interesse.

$$\begin{cases} u_{\text{chirp}}(t) = \sin\left(2\pi f_0 t + \pi \frac{f_1 - f_0}{T} t^2\right), & 0 \leq t \leq T, \\ U_{\text{chirp}}(f) \approx \begin{cases} \neq 0, & f \in \mathbb{F}, \\ 0, & f \notin \mathbb{F}, \end{cases} \\ \mathbb{F} = \{f \mid f_0 \leq |f| \leq f_1\}. \end{cases} \quad (3.2)$$

Figura 9 – Exemplo de sinal do tipo *Chirp*.



Fonte: Autoria própria.

3.1.1 Sinal aleatório

Assim como apresentado em 3.1, uma alternativa ao uso de sinais senoidais na identificação de sistemas é adquirir a resposta ao impulso. Entretanto, a excitação impulsiva ideal, $\delta(t)$, é impraticável em sistemas físicos, pois exigiria amplitude e largura de banda infinitas. Por esse motivo, algumas alternativas para uso em técnicas de identificação consistem em empregar excitações que aproximam a resposta ao impulso, tais como o ruído branco, Maximum Length Sequence (MLS), PRBS ou até mesmo o PRMLS [Farina 2000, Theodorsen 2016].

Para o ruído branco, um processo estocástico $\nu[k]$ é dito aleatório quando sua função de autocorrelação apresenta características semelhantes à de uma função de impulso. Esse resultado, como já definido em 2.1.2, implica que amostras do sinal são descorrelacionadas para todos os atrasos $\tau \neq 0$ [Bakalis *et al.* 2023, Theodorsen 2016].

$$r_{\nu\nu}[\tau] = \sigma_\nu^2 \delta[\tau]. \quad (3.3)$$

Considerando tal desenvolvimento, de acordo com as definições apresentadas em 2.1.2, a PSD é dada pela Discrete Fourier Transform (DFT) da autocorrelação.

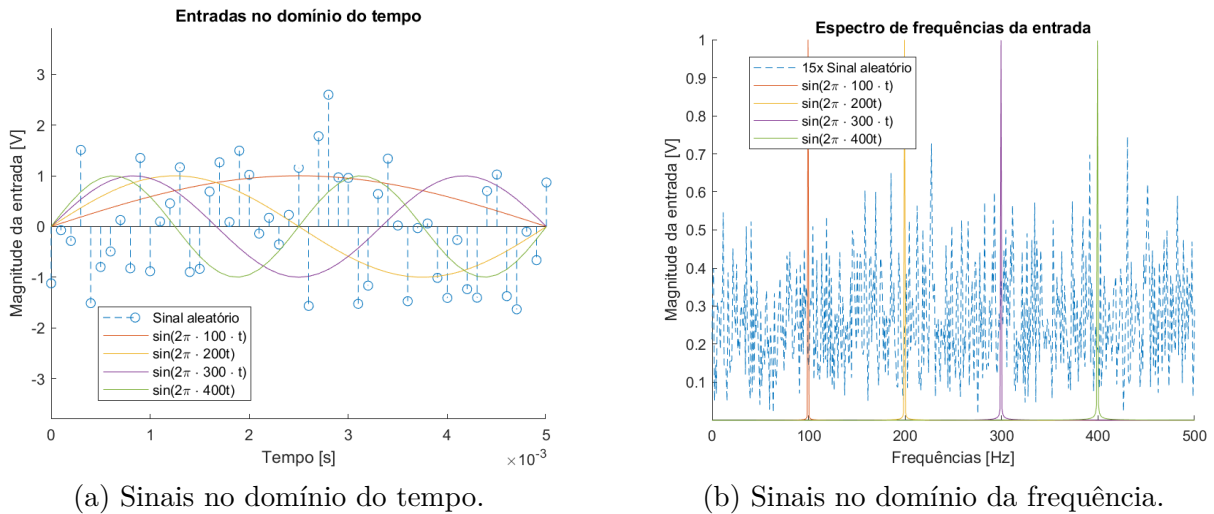
$$\phi_{vv}(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_{vv}[k] e^{-j\omega k}. \quad (3.4)$$

Nota-se que a definição introduzida pela equação 3.4 é análoga à apresentada na equação 2.5. Desse modo, substituindo 3.3 em 3.4 e resolvendo a DFT, é possível obter 3.5.

$$\phi_{vv}(e^{j\omega}) = \sum_{k=-\infty}^{\infty} \sigma^2 \delta[k] e^{-j\omega k} = \sigma^2. \quad (3.5)$$

Portanto, a PSD de um sinal com características aleatórias é constante para um número suficientemente grande de amostras N . Dessa forma, o ruído branco contém componentes em todas as frequências discretas, sendo, assim, um sinal persistentemente excitante [Aguirre 2004].

Figura 10 – Comparação entre senoides e variável aleatória, simulação de 1s e $F_s = 10^4 Hz$.



Fonte: Autoria própria.

3.1.2 PRBS

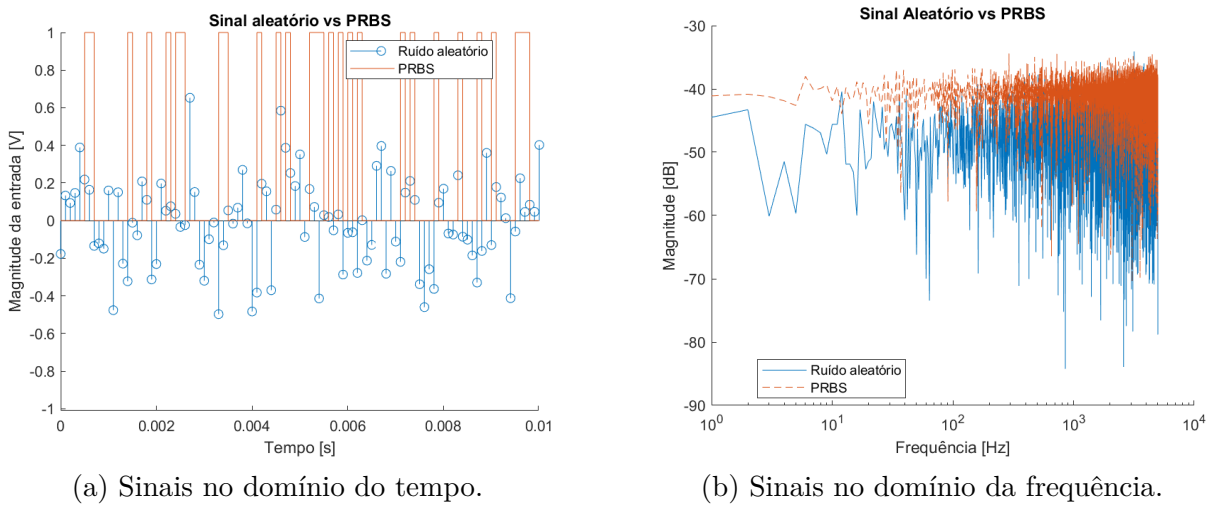
Ainda que sinais aleatórios excitem um espectro de frequências muito mais amplo do que o de um único senoide, restam várias questões associadas ao seu uso. Algumas delas residem no fato de que uma variação de nível em períodos de tempo muito curtos, uma de suas características, pode implicar um desgaste maior em atuadores mecânicos. Além disso, a inversão constante do sinal de entrada leva a diversos malfuncionamentos, avarias e impossibilidades, dependendo dos equipamentos de medição e dos atuadores intrínsecos ao sistema de estudo.

Além disso, a análise de um sistema com essas características na entrada pode mostrar-se ineficaz, visto que essas características variam a cada intervalo mínimo de aferição. Portanto, sistemas com dinâmicas muito mais lentas do que as do sinal aleatório terão pouco tempo para serem devidamente excitados e, conseqüentemente, necessitarão

de uma quantidade maior de dados para que seja realizada uma aquisição paramétrica satisfatória.

Uma das soluções possíveis para contornar esses problemas é o uso de um sinal PRBS, eventualmente adicionando um segurador à entrada do sinal, a fim de evitar desgastes desnecessários e auxiliar na identificação de dinâmicas mais lentas. Ao contrário de uma variável completamente aleatória, o PRBS possui valores binários; mas, como se vê na figura 11b, ainda excita satisfatoriamente todo o espectro de frequências.

Figura 11 – Comparação entre uma variável aleatória e um período do sinal PRBS, com simulação de 1s e $F_s = 10^4 \text{ Hz}$. $N = 10^5$.

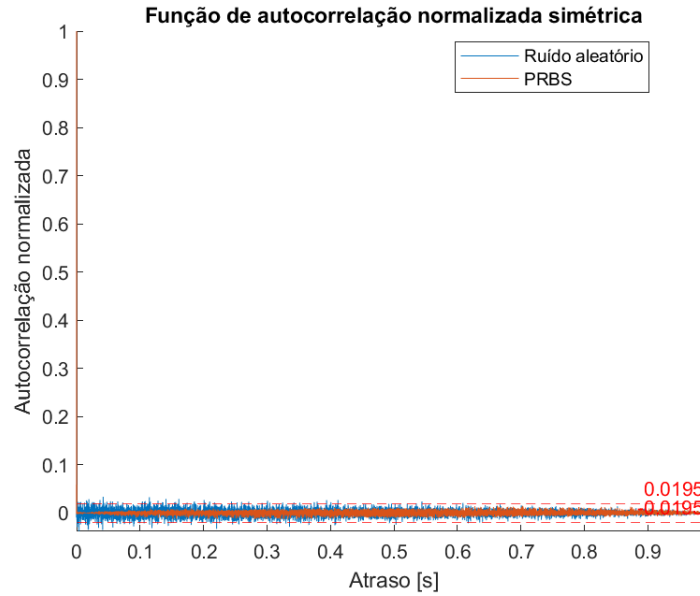


Fonte: Autoria própria.

Outra de suas particularidades é que o sinal PRBS é periódico, porém apresenta fortes características aleatórias ao longo de um de seus períodos. Dessa forma, ainda é possível aproveitar as características previamente abordadas em 3.1.

Por fim, vale destacar que, para fins comparativos, será sempre utilizada a função de autocorrelação normalizada simétrica para avaliar o comportamento aleatório dos sinais de interesse, conforme apresentado nas figuras 12 [Aguirre 2004].

Figura 12 – Comportamento da função de autocorrelação do sinal PRBS.



Fonte: Autoria própria

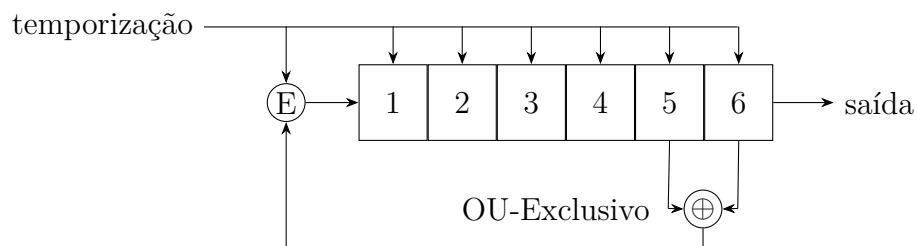
3.1.2.1 Geração de um PRBS

Duas das possíveis maneiras de gerar um sinal PRBS, uma focada na implementação em hardware e outra em uma representação matemática formal, são apresentadas a partir da equação 3.6 [Ljung 1999] e da imagem 13 [Aguirre 2004], em que $rem_2(\alpha)$ representa o resto da divisão de α por 2.

$$u(t) = rem_2 \left(\sum_{i=1}^n a_i \cdot u(t - i) \right). \quad (3.6)$$

Assim, em função do tamanho do período de aleatoriedade requisitado (N), é possível definir o número mínimo de bits, **n**, necessário para a implementação desejada. Assim, os valores **não nulos** de \mathbf{a}_k , bem como o esquema de ligação via hardware usando a porta ou-exclusivo (fig. 13), podem ser encontrados na tabela 1.

Figura 13 – Exemplo de implementação em nível de hardware, $n = 6$.



Fonte: [Aguirre 2004].

Tabela 1 – Geração de um sinal PRBS

n	$N = 2^n - 1$	\mathbf{a}_k	X-OR
2	3	1, 2	1, 2
3	7	2, 3	2, 3
4	15	1, 4	3, 4
5	31	2, 5	3, 5
6	63	1, 6	5, 6
7	127	3, 7	4, 7
8	255	1, 2, 7, 8	2, 3, 4, 8
9	511	4, 9	5, 9
10	1023	7, 10	7, 10
11	2047	9, 11	9, 11

Fonte: adaptado de [Ljung 1999] e [Aguirre 2004]

Vale destacar que é possível gerar um sinal PRBS de diversas maneiras e que os parâmetros da tabela 1 garantem o maior período possível para um número de bits dado. Assim sendo, as referências utilizadas não impedem que sejam efetuadas outras ligações ou arranjos que resultem em resultados semelhantes. Para o desenvolvimento deste trabalho, os valores \mathbf{a}_k serão sempre obtidos a partir de [Toker e Emara-Shabaik 2004].

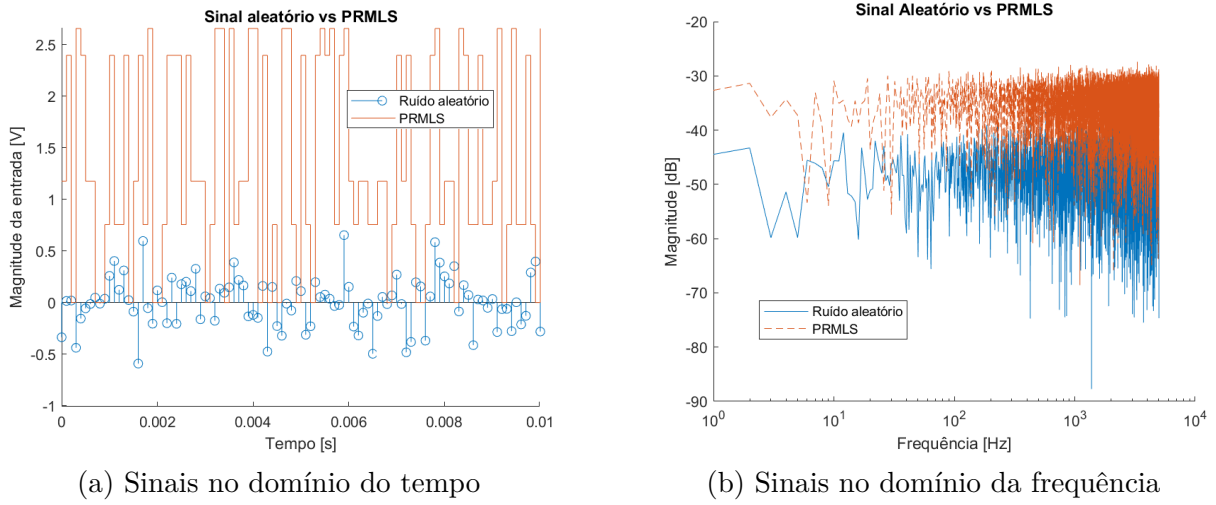
3.1.3 PRMLS

Análogo ao já abordado para o sinal PRBS, é importante mencionar sua versão multinível, o PRMLS. Extremamente útil para identificação de sistemas não lineares, o mesmo pode ser criado a partir da equação 3.7, análoga à 3.6. Para tal, foi implementada uma variação da metodologia introduzida por [Toker e Emara-Shabaik 2004] para obter os valores de \mathbf{a}_k .

$$u(t) = \sum_{i=1}^p a_i \cdot u(t - i) ; \{p \in \mathbb{N} / p \geq 2\}. \quad (3.7)$$

Dessa forma, serão calculados parâmetros aleatoriamente até obter-se um período máximo para o PRMLS, mapeando tais valores aleatórios para níveis ótimos estritamente positivos, a serem utilizados como valores do sinal de entrada. Um exemplo de PRMLS é dado a seguir:

Figura 14 – Comparação entre variável aleatória e um período do sinal PRMLS escolhido, simulação de 1s e $F_s = 10^4 \text{ Hz}$. $N = 10^5$

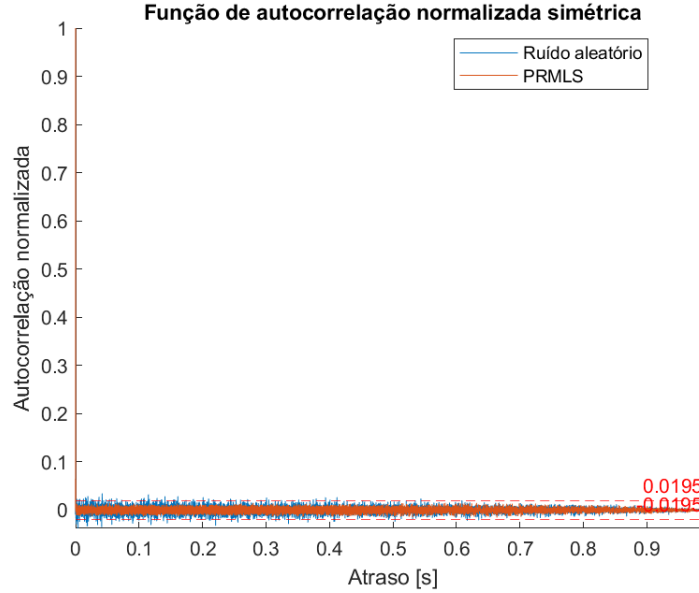


(a) Sinais no domínio do tempo

(b) Sinais no domínio da frequência

Fonte: Autoria própria.

Figura 15 – Comportamento da função de autocorrelação do sinal PRMLS escolhido.



Fonte: Autoria própria.

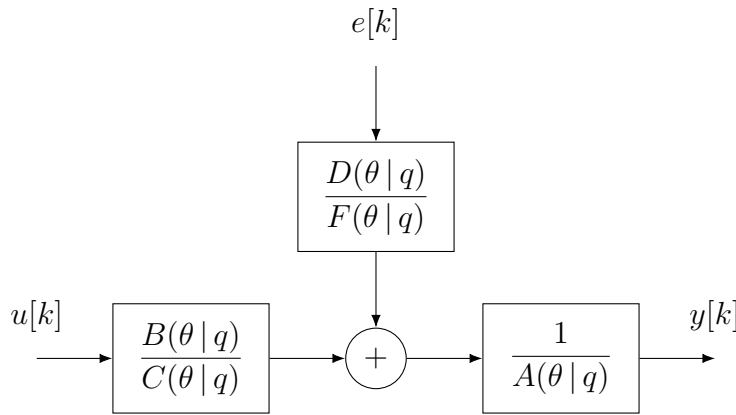
Vale também destacar que, a partir da definição de [Toker e Emara-Shabaik 2004], o PRMLS garante persistência de excitação para p maior que a ordem de não linearidade do sistema ($p > \mathbb{O}$). Por outro lado, isso também significa que o sinal PRBS clássico excita efetivamente somente sistemas sem nenhuma não linearidade ($p = 2 \therefore \mathbb{O} \leq 1$), ainda que apresente componentes em todas as frequências.

3.2 Métodos de representação

Embora representações clássicas, como o espaço de estados e as funções de transferência, sejam válidas, é comum recorrer a modelos polinomiais na identificação de sistemas em tempo discreto. A estrutura geral dessas classes de modelos é definida pela Equação 3.8 e ilustrada na figura 16.

$$A(\theta|q) y[k] = \frac{B(\theta|q)}{C(\theta|q)} u[k] + \frac{D(\theta|q)}{F(\theta|q)} e[k]. \quad (3.8)$$

Figura 16 – Sistema dado pela equação 3.8.



Fonte: Autoria própria.

Estes modelos polinomiais estão definidos no domínio do tempo discreto por meio do operador de atraso q^{-1} , que atua sobre as variáveis $u[k]$ ou $y[k]$ conforme a Equação 3.9. O operador q^{-1} no domínio do tempo discreto é o análogo ao termo z^{-1} na Transformada Z e permite escrever as equações de diferença como funções de transferência.

$$q^{-\alpha} u[k] = u[k - \alpha] : \alpha \in \mathbb{N}. \quad (3.9)$$

Tabela 2 – Classes de modelos notáveis obtidos a partir de 3.8.

Nome da estrutura	Funções utilizadas
FIR (<i>Finite Impulse Response</i>)	B
ARX	A, B
ARMAX	A, B, D
ARMA	A, D
ARARX	A, B, F
ARARMAX	A, B, D, F
OE (<i>Output Error</i>)	B, C
BJ (<i>Box-Jenkins</i>)	B, C, D, F

Fonte: [Ljung 1999].

Vale também mencionar que, por vezes, há a presença de um *delay* entre a dinâmica da saída e a dinâmica da entrada. Sua representação pode ser efetuada por meio da simples inserção de uma nova variável τ , de modo que as equações características do método Autoregressive with eXogenous inputs (ARX) sejam obtidas a partir de eq. (3.10). A eq. (3.11) apresenta um exemplo de estimação de um sistema de ordem 2.

$$\begin{cases} y[k] = \sum_{j=\tau+1}^{\alpha} a_{j-\tau} y[k-j] + \sum_{i=\tau}^{\beta} b_{i-\tau} u[k-i] + e[k], \\ y[k] = \frac{B(\theta|q|\tau)}{A(\theta|q)} u[k] + \frac{1}{A(\theta|q)} e[k]. \end{cases} \quad (3.10)$$

$$\begin{cases} y[k] = a_1 y[k-1] + a_2 y[k-2] + b_1 u[k-1] + b_2 u[k-2] + e[k] \\ y[k] = \frac{b_1 q^{-1} + b_2 q^{-2}}{1 - a_1 q^{-1} - a_2 q^{-2}} u[k] + \frac{1}{1 - a_1 q^{-1} - a_2 q^{-2}} e[k]. \end{cases} \quad (3.11)$$

Dessa forma, a popularidade dos modelos ARX na modelagem e identificação é explicada por sua formulação direta no domínio do tempo discreto e pela capacidade intrínseca de lidar com distúrbios e ruídos de medição.

Análogo ao já apresentado, o modelo Non-linear Autoregressive with eXogenous inputs (NARX) inclui regressores com características não lineares na equação 3.10. Um dos métodos para incluir tais dinâmicas consiste em adicionar termos polinomiais dependentes de u e y como regressores, incluindo produtos cruzados e potências elevadas dos sinais de entrada e de saída, que atuam como regressores não-lineares. Esses termos polinomiais, usualmente implementados para identificações do tipo caixa-preta, seguem a eq. (3.12):

$$u^i \cdot y^j : [i, j] \in \mathbb{N}. \quad (3.12)$$

Vale destacar que há um grande número de combinações possíveis a partir dessa aproximação polinomial, o que aumenta exponencialmente o custo computacional associado ao processo de estimação. Entretanto, um grande número de regressores também tende a introduzir o problema de overfitting, elevando a variância do modelo e diminuindo sua capacidade de generalização. Para contornar tal problemática é comum recorrer a métodos de seleção de regressores como o LASSO.

Além disso, dependendo do sistema modelado e da quantidade de informações previamente obtidas sobre o sistema, há uma grande separação entre a modelagem e o significado físico de cada um dos parâmetros encontrados. Dessa forma, o processo de validação para identificações do tipo caixa-preta torna-se mais complexo e abstrato, mesmo que ainda seja possível comprovar que o modelo encontrado é satisfatório.

3.3 Teoria da estimação

Tendo como base os dados referentes ao comportamento do sistema de interesse, faz-se necessário encontrar, a partir de eq. (3.8), os parâmetros de um modelo escolhido que aproximem o melhor possível o sistema real em sua estimativa. Esse problema tem origem histórica nas aplicações à astronomia e à geodésia do século XVIII, quando Legendre e Gauss formalizaram e aplicaram o MMQ para ajustar observações ruidosas [Legendre 1805, Gauss 1809].

Além disso, ao longo do século XX, o desenvolvimento teórico e computacional consolidou propriedades formais dos estimadores clássicos e abriu caminho para métodos que penalizam a complexidade do modelo, visando a um melhor desempenho preditivo em situações em que o sistema apresenta alta dimensionalidade [Greene 2018, Hastie, Tibshirani e Friedman 2009].

Com este objetivo, introduzido e analisado por Tibshirani, o método LASSO acrescenta um termo de regularização à formulação do MMQ. Essa penalização seleciona as variáveis mais significativas para a dinâmica do sistema, reduzindo a variância a custo de um aumento do viés [Tibshirani 1996, Hastie, Tibshirani e Friedman 2009]. Portanto, o LASSO pode ser utilizado como estratégia de seleção de regressores no âmbito da identificação de sistemas.

A partir da teoria da estimação, portanto, é possível encontrar algoritmos que selecionam e estimam os parâmetros de interesse, mantendo eficiência e robustez, minimizando possíveis efeitos decorrentes do ruído [Wasserman 2004].

3.3.1 O estimador de mínimos quadrados

A abordagem clássica do estimador por MMQ tem como objetivo minimizar a soma dos quadrados dos erros de estimação, ou seja, o custo quadrático $J(\theta)$, conforme a eq. (3.13). De sua formulação clássica, a saída estimada é obtida por $\hat{y} = M\theta$, com o objetivo de definir o vetor de parâmetros θ que minimiza $J(\theta)$ ($\min_{\theta} J(\theta) \in \mathbb{R}$). A matriz M é definida com base no modelo de representação a ser utilizado: como, durante o desenvolvimento do trabalho, será aplicado o método ARX, M é denominada matriz de regressores.

$$\min \left\{ \sum_i^N \|\hat{y}_i - y_i\|^2 \right\} = \min J(\theta). \quad (3.13)$$

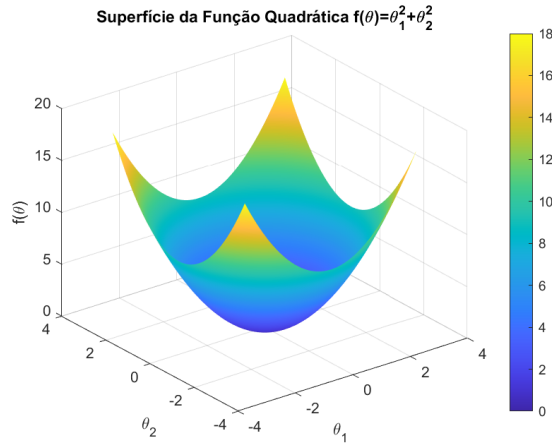
Dada uma função convexa qualquer $f(y)$, um método analítico comprovado para encontrar pontos de mínimo (ou máximo) é a solução por pontos estacionários, ou seja, pontos em que a derivada de $f(y)$ é zero. Dessa forma, assumindo que tal função é bem comportada e derivável até a segunda ordem, para a minimização, deve-se calcular θ a partir de eq. (3.14). As propriedades de derivação matricial são apresentadas em eq. (3.15).

$$\begin{cases} \frac{\partial}{\partial \theta} f(y) = 0, \\ \frac{\partial^2}{\partial \theta^2} f(y) \geq 0. \end{cases} \quad (3.14)$$

$$\begin{cases} \frac{\partial}{\partial \theta} \theta' A \theta = (A + A') \theta, \\ \frac{\partial}{\partial \theta} A' \theta = \theta. \end{cases} \quad (3.15)$$

Portanto, se a segunda derivada for estritamente positiva, a solução encontrada é única. Considerando, ainda, a natureza quadrática da função 3.13, ilustrada em fig. 17, é possível encontrar o valor de θ para o qual o erro associado à estimação seja mínimo. Vale também lembrar que, em ambiente computacional, certas operações matemáticas podem ser facilmente transformadas em suas versões matriciais, conforme proposto em eq. (3.17).

Figura 17 – Exemplo de uma função $J(\theta)$ de natureza quadrática.



Fonte: Autoria própria.

$$f(y, M, \theta) = \sum (\hat{y} - y)^2 = \sum \zeta^2, \quad (3.16)$$

$$\sum \zeta^2 = \zeta^T \zeta = \|\zeta\|^2. \quad (3.17)$$

Aplicando as definições na equação característica $\hat{y} = M\theta$:

$$\hat{y} = M\theta \quad \therefore \quad \zeta = M\theta - y, \quad (3.18)$$

$$\|\zeta\|^2 = (M\theta - y)^T (M\theta - y) = (\theta^T M^T - y^T) (M\theta - y), \quad (3.19)$$

$$\|\zeta\|^2 = \theta^T M^T M \theta - \theta^T M^T y - y^T M \theta + y^T y. \quad (3.20)$$

As dimensões das matrizes, $y \theta^T M^T$, são idênticas a $y^T M \theta$, visto que ambos os termos são escalares. Assim sendo, é possível aplicar a derivação de eq. (3.21), a fim de identificar um valor de θ que seja candidato à minimização de $J(\theta)$.

$$\| \zeta \|^2 = \theta^T M^T M \theta - 2\theta^T M^T y + y^T y, \quad (3.21)$$

$$\frac{\partial}{\partial \theta} f(y, M, \theta) = 0 \quad \therefore \quad 2M^T M \theta = 2M^T y, \quad (3.22)$$

$$\therefore \quad \boxed{\theta = (M^T M)^{-1} M^T y}. \quad (3.23)$$

Além disso, considerando as condições para minimização, faz-se necessário averiguar o comportamento da segunda derivada:

$$\frac{\partial}{\partial \theta} 2M^T M \theta - 2M^T y > 0 \quad \Longleftrightarrow \quad 2M^T M > 0. \quad (3.24)$$

Da própria construção do MMQ, é possível garantir que a minimização ocorre quando se utilizam entradas persistentemente excitantes, visto que levam a uma matriz M com posto completo e, portanto, a $M^T M$ ser definida positiva [Aguirre 2004].

Portanto, utilizando a definição de matriz de Penrose ($M^+ = (M^T M)^{-1} M^T$), é possível obter a definição do parâmetro θ que resulta em eq. (3.25). Pelas definições eq. (3.13) a eq. (3.24), fica claro que $M^T M$ definida positiva implica uma solução ótima e única para os parâmetros do modelo.

$$\boxed{\theta = M^+ \cdot y \Longleftrightarrow \sum (\hat{y} - y)^2 = \min \left\{ \sum (\hat{y} - y)^2 \right\}}. \quad (3.25)$$

De forma equivalente, a solução do MMQ pode ser expressa por meio de somatórias, como apresentado na equação 3.26 [Aguirre 2004]. Nesta notação, $\psi(i-1)$ representa o vetor de regressores (a i -ésima linha da matriz M) e N é o número de amostras:

$$\hat{\theta}_N = \left[\frac{1}{N} \sum_{i=1}^N \psi(i-1) \psi^T(i-1) \right]^{-1} \cdot \left[\frac{1}{N} \sum_{i=1}^N \psi(i-1) y(i) \right]. \quad (3.26)$$

É fundamental notar que o MMQ exige linearidade nos parâmetros (vetor θ), permitindo sua aplicação tanto em modelos de regressão lineares (como retas) quanto em modelos intrinsecamente não-lineares nos dados (como polinômios), desde que os parâmetros sejam lineares em relação ao modelo.

A fim de ilustrar o estimador proposto, foram identificados parâmetros da reta $y = 2t + 7$ ($\theta = [2, 7]$), considerando um pequeno ruído do tipo i.i.d. na saída do sistema. Dessa forma, a equação 3.27 descreve as quatro primeiras amostras do sistema ruidoso.

$$\begin{cases} 6.6552 = 2 \cdot 0 + 7 + e[0], \\ 8.9776 = 2 \cdot 1 + 7 + e[1], \\ 10.9123 = 2 \cdot 2 + 7 + e[2], \\ 13.4654 = 2 \cdot 3 + 7 + e[3]. \end{cases} \quad (3.27)$$

Considerando a formulação encontrada para o MMQ, é possível reescrever o sistema de equações eq. (3.27) a partir de eq. (3.28). Dessa forma, é possível aplicar a eq. (3.25) iterativamente à medida que aumenta o número de amostras obtidas, ou seja, utilizando um número crescente de linhas no processo de estimação. Portanto, alterando a ordem do sistema de equações aplicado a eq. (3.25), é possível averiguar a qualidade dos parâmetros estimados a partir da tabela 3.

$$\begin{bmatrix} 6.6552 \\ 8.9776 \\ 10.9123 \\ 13.4654 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} + e. \quad (3.28)$$

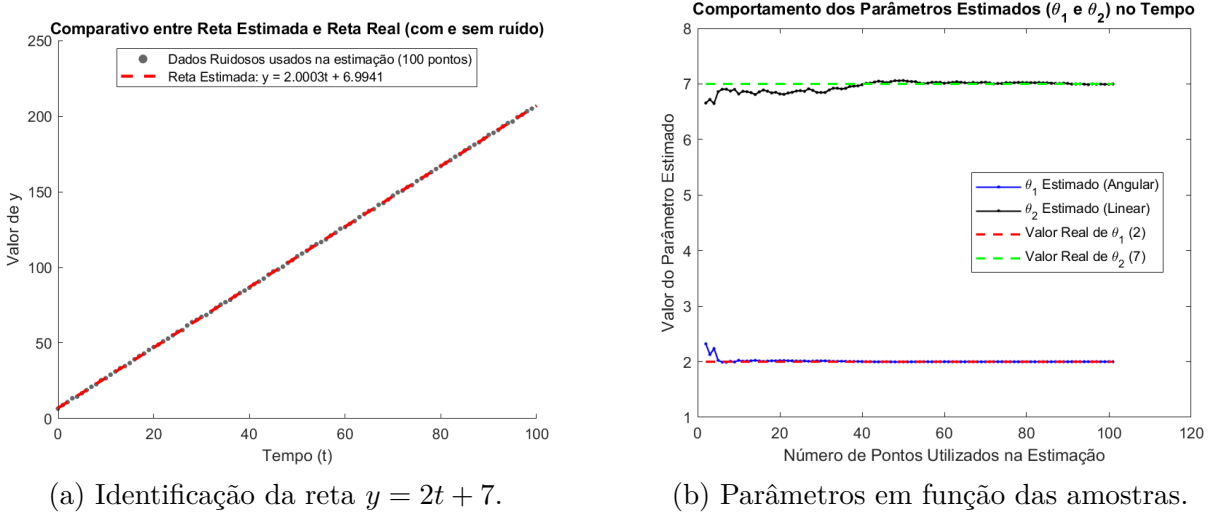
Tabela 3 – Estimação dos parâmetros da reta $y = \theta_1 t + \theta_2$ em função do número de amostras

Amostras	Coefficiente angular	Coefficiente linear
2	2.3223	6.6552
3	2.1285	6.7198
4	2.2365	6.6478

Fonte: Autoria própria.

A fig. 18 ilustra o resultado do mesmo processo de estimação. O gráfico 18a mostra a reta estimada a partir de 100 amostras de dados ruidosos, enquanto 18b apresenta a convergência dos parâmetros $\hat{\theta}_1$ (coeficiente angular) e $\hat{\theta}_2$ (coeficiente linear) em função do número de amostras utilizadas, destacando como o estimador se aproxima dos valores reais ($\theta_1 = 2$, $\theta_2 = 7$) à medida que mais dados são processados.

Figura 18 – Exemplo de estimação a partir do MMQ - número máximo de 100 amostras.



Fonte: Autoria própria.

3.3.2 O MMQ recursivo

O MMQR permite atualizar os parâmetros do modelo ($\hat{\theta}$) a cada nova aquisição de dados, eliminando a necessidade de recomputar a solução em batch (3.26) a cada passo. A derivação inicia-se reescrevendo a solução batch para um número de amostras c :

$$\hat{\theta}_c = \left[\sum_{i=1}^c \psi(i-1)\psi^T(i-1) \right]^{-1} \cdot \left[\sum_{i=1}^c \psi(i-1)y(i) \right]. \quad (3.29)$$

Desse modo, definindo as matrizes P_c e Q_c tal que $P_c = \left[\sum_{i=1}^c \psi(i-1)\psi^T(i-1) \right]^{-1}$ e $Q_c = \left[\sum_{i=1}^c \psi(i-1)y(i) \right]$, a equação 3.29 pode ser reescrita como 3.30.

$$\hat{\theta}_c = P_c \cdot Q_c. \quad (3.30)$$

Efetuando a separação do termo para $i = c$ tanto em P_c quanto em Q_c , é possível obter 3.31. Em seguida, devido à nomenclatura utilizada, é plausível reescrever 3.31 como 3.32.

$$\begin{cases} P_c^{-1} = \left[\sum_{i=1}^{c-1} \psi(i-1)\psi^T(i-1) \right] + \psi(c-1)\psi^T(c-1), \\ Q_c = \left[\sum_{i=1}^{c-1} \psi(i-1)y(i) \right] + \psi(c-1)y(c), \end{cases} \quad (3.31)$$

$$\begin{cases} P_c^{-1} = P_{c-1}^{-1} + \psi(c-1)\psi^T(c-1), \\ Q_c = Q_{c-1} + \psi(c-1)y(c). \end{cases} \quad (3.32)$$

Retomando a equação 3.30, é possível reescrevê-la para o instante $c - 1$, de modo a obter 3.33.

$$\begin{cases} Q_{c-1} = P_{c-1}^{-1} \hat{\theta}_{c-1}, \\ Q_c = P_{c-1}^{-1} \hat{\theta}_{c-1} + \psi(c-1)y(c). \end{cases} \quad (3.33)$$

Substituindo Q_c (de 3.33) em $\hat{\theta}_c = P_c Q_c$ (3.30), obtemos uma forma que ainda requer a inversão da matriz P_c^{-1} a cada passo eq. (3.34). Para evitar essa onerosa inversão matricial, aplicamos a Identidade da Inversão da Matriz (ou Identidade de Woodbury) 3.35 à relação $P_c^{-1} = P_{c-1}^{-1} + \psi(c-1)\psi^T(c-1)$ (3.32).

$$\hat{\theta}_c = P_c \left[P_{c-1}^{-1} \hat{\theta}_{c-1} + \psi(c-1)y(c) \right], \quad (3.34)$$

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}, \quad (3.35)$$

$$\begin{cases} A = P_{c-1}^{-1}, \\ U = \psi(c-1), \\ C = 1, \\ V = \psi^T(c-1). \end{cases} \quad (3.36)$$

Assim sendo, é possível inverter P_c^{-1} denotada na formulação 3.32 a partir de:

$$P_c = \left[P_{c-1}^{-1} + \psi(c-1)\psi^T(c-1) \right]^{-1}, \quad (3.37)$$

$$P_c = (A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}, \quad (3.38)$$

$$P_c = P_{c-1} - P_{c-1}\psi(c-1) \left[1 + \psi^T(c-1) \cdot P_{c-1} \cdot \psi(c-1) \right]^{-1} \psi^T(c-1)P_{c-1}, \quad (3.39)$$

$$\therefore P_c = P_{c-1} - \frac{P_{c-1}\psi(c-1)}{1 + \psi^T(c-1) \cdot P_{c-1} \cdot \psi(c-1)} \psi^T(c-1)P_{c-1}. \quad (3.40)$$

Definindo um ganho K_c de modo que a formulação 3.41 seja verdadeira, é possível reescrever 3.39 como 3.42.

$$\begin{cases} K_c = \frac{P_{c-1} \psi(c-1)}{1+a}, \\ a = \psi^T(c-1) P_{c-1} \psi(c-1), \end{cases} \quad (3.41)$$

$$P_c = P_{c-1} - K_c \psi^T(c-1)P_{c-1}. \quad (3.42)$$

Substituindo 3.42 em 3.34, obtém-se 3.44. É importante ressaltar que, a partir do uso de 3.30 e da primeira equação em 3.33 é possível simplificar os termos M_1 e M_3 .

$$\begin{cases} \hat{\theta}_c = M_1 + M_2 - M_3 - M_4, \\ M_1 = P_{c-1} P_{c-1}^{-1} \hat{\theta}_{c-1} = \hat{\theta}_{c-1}, \\ M_2 = P_{c-1} \psi(c-1) y(c), \\ M_3 = K_c \psi^T(c-1) P_{c-1} P_{c-1}^{-1} \hat{\theta}_{c-1} = K_c \psi^T(c-1) \hat{\theta}_{c-1}, \\ M_4 = K_c \psi^T(c-1) P_{c-1} \psi(c-1) y(c) = K_c \cdot a \cdot y(c), \end{cases} \quad (3.43)$$

$$\therefore \hat{\theta}_c = \hat{\theta}_{c-1} - K_c \psi^T(c-1) \hat{\theta}_{c-1} + [P_{c-1} \psi(c-1) - K_c \cdot a] y(c). \quad (3.44)$$

A partir da definição de K_c , verifica-se que $P_{c-1} \psi(c-1) = K_c(1+a)$. Substituindo na equação 3.44 e agrupando os termos dependentes de K_c , obtém-se 3.45.

$$\begin{cases} \hat{\theta}_c = \hat{\theta}_{c-1} - K_c \psi^T(c-1) \hat{\theta}_{c-1} + K_c y(c), \\ \hat{\theta}_c = \hat{\theta}_{c-1} + K_c [y(c) - \psi^T(c-1) \hat{\theta}_{c-1}]. \end{cases} \quad (3.45)$$

O MMQR básico pressupõe que todos os dados anteriores são igualmente relevantes. Em sistemas em que as características mudam lentamente, é comum introduzir o fator de esquecimento λ , com $0 < \lambda \leq 1$, para dar maior peso às amostras mais recentes e ‘esquecer’ as anteriores. O uso de λ resulta em equações formuladas em 3.46 [Goel, Bruce e Bernstein 2020].

$$\begin{cases} K_c = \frac{P_{c-1} \psi(c-1)}{\lambda + \psi^T(c-1) P_{c-1} \psi(c-1)}, \\ \hat{\theta}_c = \hat{\theta}_{c-1} + K_c [y(c) - \psi^T(c-1) \hat{\theta}_{c-1}], \\ P_c = \lambda^{-1} [P_{c-1} - K_c \psi^T(c-1) P_{c-1}]. \end{cases} \quad (3.46)$$

O sucesso do MMQR reside no fato de que, ao contrário do MMQ padrão, a única inversão necessária no algoritmo (3.46) é a do denominador do ganho K_c . Visto que $\lambda + \psi^T P_{c-1} \psi$ é um escalar, o alto custo computacional da inversão matricial é evitado, o que permite a execução eficiente em tempo real.

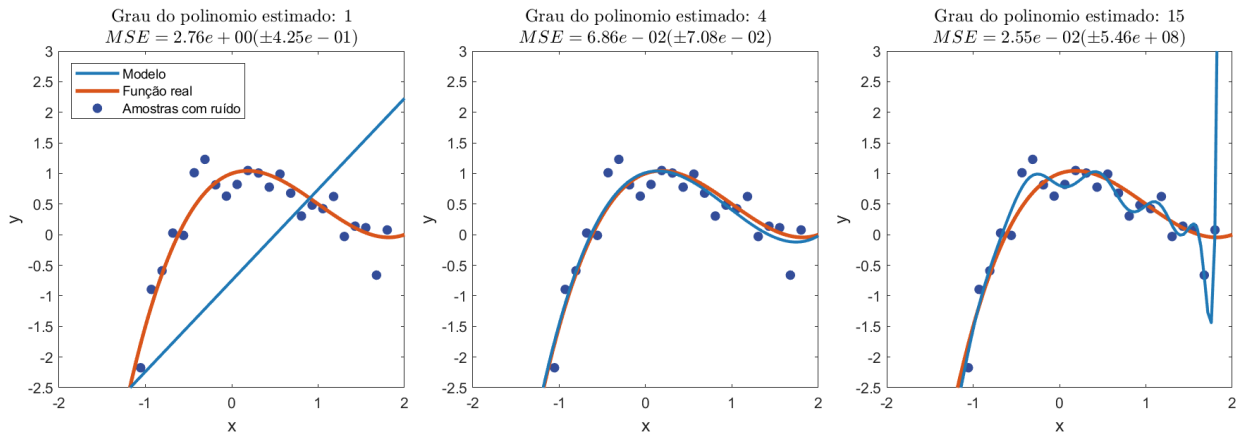
3.3.3 Uso do LASSO como ferramenta de seleção de regressores

A partir da metodologia apresentada em 3.3.1 e 3.2, é possível obter uma aproximação paramétrica ótima a partir do MMQ. Deve-se considerar, porém, que o custo computacional associado à estimação pode se tornar desnecessariamente elevado e que a possível inserção de regressores não relevantes pode provocar o efeito de *overfitting*.

Isso ocorre quando foi definida uma quantidade elevada de regressores, que pode ser superior ao número necessário para caracterizar o sistema de interesse. Isso tende a ocorrer por desconhecimento ou até mesmo por uma interpretação equivocada do objeto de estudo.

Nesses casos, com base nas limitações previamente abordadas no capítulo 1, faz-se necessário limitar o número de regressores. Além disso, outro problema oriundo do uso de uma quantidade superior à necessária é o *overfitting*, o que pode levar a um modelo enviesado ou demasiadamente contaminado por ruídos. Tal dinâmica é representada pela figura 20

Figura 19 – Exemplos de underfitting (à esquerda) e de overfitting (à direita).



Fonte: Autoria própria.

Para tais casos, é útil utilizar técnicas como a regressão LASSO, que identifica e atenua a influência de regressores irrelevantes para o sistema de interesse. Assim sendo, espera-se que a utilização de métodos como este acarrete uma diminuição da precisão do modelo nos dados obtidos, em troca de maior eficácia na descrição do sistema como um todo. Em termos de *machine learning*, o desempenho nos dados de treinamento é inferior, mas tende a ser igual ou superior ao dos dados de validação.

Assim, definem-se duas funções distintas: $f(M, \theta)$, que minimiza o problema original, e $g(M, \theta)$, um termo de regularização que busca generalizar melhor os dados obtidos e eliminar o *overfitting*. Vale destacar que o parâmetro λ , abordado separadamente, nada mais é do que um peso adicionado à função objetivo, indicando o quão importante é essa correção em relação ao desempenho da função original $f(M, \theta)$.

$$\min \left\{ f(M, \theta) + \lambda g(M, \theta) \right\}. \quad (3.47)$$

Dessa forma, as regressões do tipo Ridge e LASSO são dadas pelas equações 3.48 e 3.49. É importante destacar que a regressão de Ridge é mais utilizada para reduzir o viés

que os dados de treinamento podem introduzir no modelo, enquanto o método LASSO é extremamente útil para eliminar regressores e, portanto, dinâmicas, que possuem pouca ou nenhuma influência na saída do sistema. Em específico, para o LASSO, aplica-se a regularização baseada na norma L1, que tende a zerar os parâmetros de menor relevância à medida que aumenta λ .

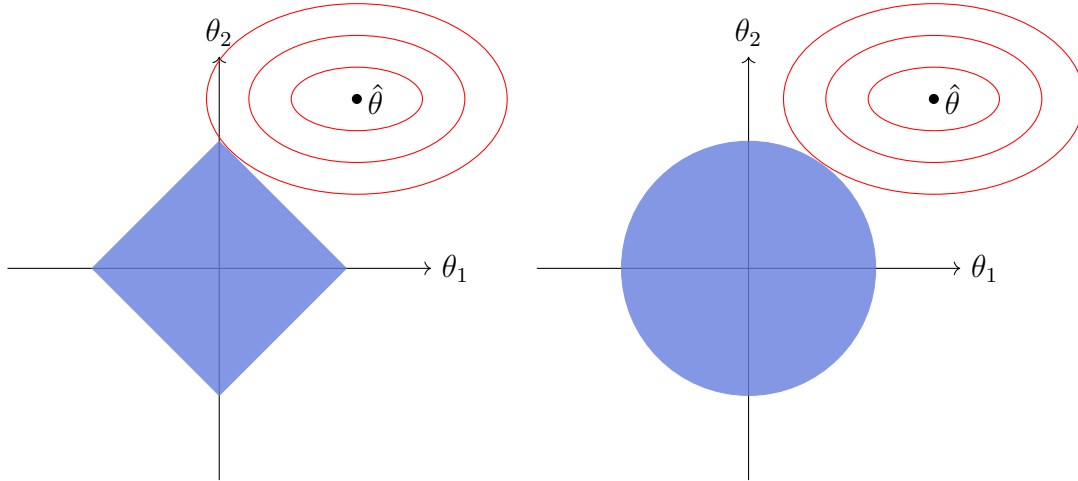
$$\min \left\{ \frac{1}{n} \sum (\hat{y}_i - y_i)^2 + \lambda \sum (\theta_i)^2 \right\}, \quad (3.48)$$

$$\min \left\{ \frac{1}{n} \sum (\hat{y}_i - y_i)^2 + \lambda \sum |\theta_i| \right\}. \quad (3.49)$$

Vale também destacar que é possível substituir a função $g(M, \theta)$ a partir de eq. (3.50), desde que a função objetivo original $f(M, \theta)$ respeite a nova restrição imposta. Desse modo, a fig. 20 representa tal dinâmica para Ridge e LASSO, tratando $\theta = [\theta_1, \theta_2]$, e $\hat{\theta}$ o ponto em que $f(M, \theta)$ é mínimo.

$$\sum |\theta_i|^q < t^q. \quad (3.50)$$

Figura 20 – Comportamento da função objetivo $f(M, \theta)$ em vermelho, apresentando ponto de minimização em $\hat{\theta}$. Os termos de regularização impostos pela equação 3.50 são denotados pelas formas geométricas em azul. Da esquerda para a direita: Lasso ($q = 1$) e Ridge ($q = 2$).

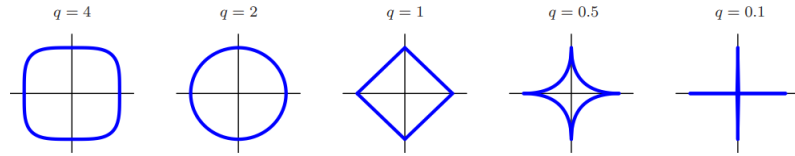


Fonte: [Hastie, Tibshirani e Wainwright 2015].

Além das regularizações L1 (LASSO) e L2 (Ridge), existe uma família de restrições similares que resulta em diversos métodos de otimização distintos, todos os quais podem ser representados pela estrutura dada na equação 3.47, apenas alterando a natureza das funções $f(M, \theta)$ e $g(M, \theta)$. As restrições correspondentes a essas regularizações são apresentadas na fig. 21, notando-se que para L1 ($q = 1$) e L2 ($q = 2$) o formato é o mesmo do apresentado em fig. 20.

$$g(M, \theta) = \sum |\theta_i|^q. \quad (3.51)$$

Figura 21 – Diferentes formatos das restrições introduzidas em eq. (3.50).



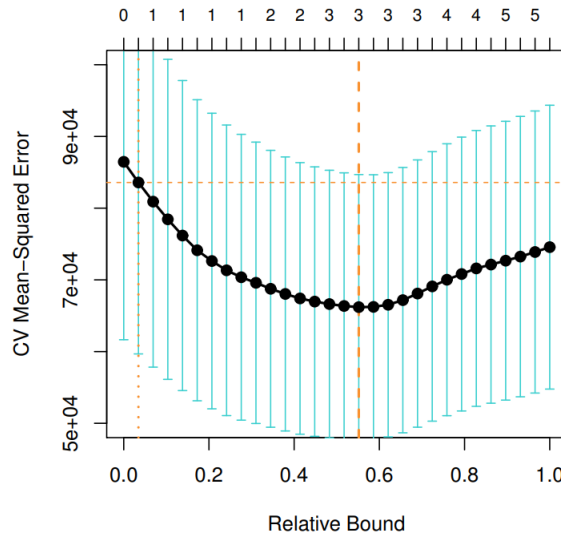
Fonte: adaptado de [Hastie, Tibshirani e Wainwright 2015].

3.3.3.1 Parâmetro de penalização λ

É importante também destacar a influência do parâmetro λ na modelagem do problema. Inicialmente, é possível afirmar que, para $\lambda = 0$, a equação de interesse nada mais é do que o método dos mínimos quadrados já abordado, acrescido de um fator $\frac{1}{n}$. A influência de valores elevados de λ , porém, pode não ser tão intuitiva.

Considerando que o problema de interesse é classificado como uma otimização por minimização, para pesos maiores em $g(M, \theta)$, os próprios parâmetros terão de ser fortemente penalizados, o que indica uma menor variação de y em função dos regressores. Para o método LASSO, tal comportamento é ainda mais evidente, visto que os regressores de menor influência podem ser simplesmente eliminados da equação do sistema.

Figura 22 – Exemplo de validação cruzada.



Fonte: [Hastie, Tibshirani e Wainwright 2015].

Desse modo, uma das maneiras de verificar um valor ideal de λ é realizar a validação cruzada dos dados, como ilustrado na fig. 22. Isso nada mais é do que excluir uma das linhas de M e sua correspondente em y , verificando, para uma ampla gama de valores de

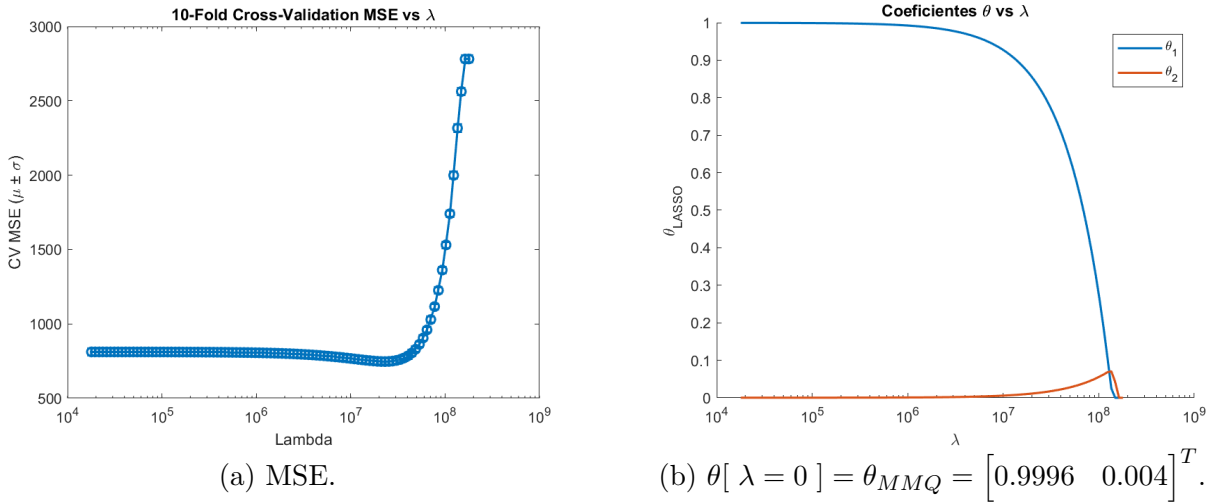
λ , a regressão LASSO (ou Ridge) correspondente. Depois, apenas para a linha excluída, calcule o erro quadrático médio da saída.

Ao efetuar o procedimento em todas as linhas, é possível agrupar os erros de todas as amostras em vetores, um por λ , e confirmar suas médias e variâncias. Assim, é possível verificar qual λ leva ao menor Mean Squared Error (MSE), conforme apresentado na fig. 22.

É importante destacar que tal processo acarreta um grande custo computacional, visto que requer o cálculo do erro associado à retirada de N amostras. Dessa forma, será utilizada uma abordagem distinta, na qual os dados amostrados serão separados em um número determinado de *folds*. Dessa forma, é necessário efetuar N_{folds} operações, o que representa um custo bem inferior ao associado a N loops de *cross-validation*.

Por fim, a fig. 23 ilustra tanto o processo de validação cruzada (23a) quanto a tendência do LASSO a zerar os coeficientes à medida que o parâmetro λ aumenta (23b).

Figura 23 – Influência do parâmetro λ na estimação, com dados obtidos na seção 4.4, relativos à figura 36a. Uso da escala logarítmica no eixo x.



Fonte: Autoria própria.

Dessa forma, mesmo que o custo computacional associado à implementação do LASSO seja significativo, a depender da necessidade e do nível de invariância temporal do sistema, tal etapa pode ser aplicada a partir de um pré-processamento sobre os dados obtidos experimentalmente. Portanto, considerando uma aplicação em tempo real, é possível se utilizar do MMQ, ou até mesmo de sua versão recursiva, para identificar apenas os parâmetros mais significativos encontrados. Assim sendo, o uso de regularizações como o LASSO tende a diminuir não só o efeito do *overfitting*, mas também o custo computacional associado a utilização de toda a lógica de identificação.

3.4 Validação dos modelos

Também é de suma importância verificar se o modelo obtido condiz com o sistema de interesse. Para tal, no escopo do trabalho, serão utilizados três parâmetros principais: o MSE, o fit dos dados e o Final Prediction Error (FPE). Os demais indicadores também serão definidos, embora não sejam utilizados diretamente na validação dos dados obtidos.

Primeiramente, o indicador R^2 baseia-se na diferença entre o valor estimado (\hat{y}) e o valor médio (\bar{y}) do sinal de interesse, bem como no erro de estimação. Assim sendo, valores próximos ao unitário indicam uma boa modelagem e, para valores superiores a 0.7, a representação é considerada confiável [Gupta, Stead e Ganti 2024].

$$R^2 = \frac{\sum(\hat{y} - \bar{y})^2}{\sum(y - \hat{y})^2 + (\hat{y} - \bar{y})^2}. \quad (3.52)$$

Como o R^2 , também é comum utilizar o parâmetro Fit para avaliar a adequação da modelagem em relação aos dados utilizados para obtê-la.

$$Fit = 100 \cdot \left[1 - \frac{\|y - \hat{y}\|}{\|y - \bar{y}\|} \right]. \quad (3.53)$$

Além disso, é possível utilizar a própria norma de diferença de vetores (eq. (3.54)) para identificar a similaridade entre eles. De uma perspectiva geométrica, tal resultado nada mais é do que o comprimento do vetor de diferenças, ou seja, a distância entre o vetor estimado e o vetor com informações aferidas diretamente no sistema.

$$\|v\| = \sqrt{\sum_{i=0}^n v_i^2}. \quad (3.54)$$

Assim, utilizando a definição de norma (eq. (3.54)), o MSE calcula um valor médio associado à diferença entre os parâmetros de interesse, usualmente expressa por $y - \hat{y}$ ou $\theta - \hat{\theta}$.

$$MSE(v[k]) = \frac{\|v[k]\|^2}{N}; \quad \{k \in \mathbb{N}/1 < k < K\}. \quad (3.55)$$

O FPE também é uma métrica interessante de obter, pois quantifica a capacidade do sistema de evitar o overfitting.

$$FPE = \sigma_\zeta^2 \cdot \frac{N + d}{N - d}. \quad (3.56)$$

Por fim, dadas as funções de correlação cruzada e autocorrelação dos sinais de entrada e de saída de um sistema qualquer, é possível obter o espectro de coerência a

partir de 3.57. Assim, como abordado em [Aguirre 2004], é possível considerar confiáveis os valores obtidos para frequências em que $\hat{\gamma}_{uy}(\omega) \geq 0.6$.

$$\hat{\gamma}_{uy}(\omega) = \sqrt{\frac{|\hat{\phi}(e^{j\omega})|^2}{|\hat{\phi}_u(\omega)| |\hat{\phi}_y(\omega)|}} . \quad (3.57)$$

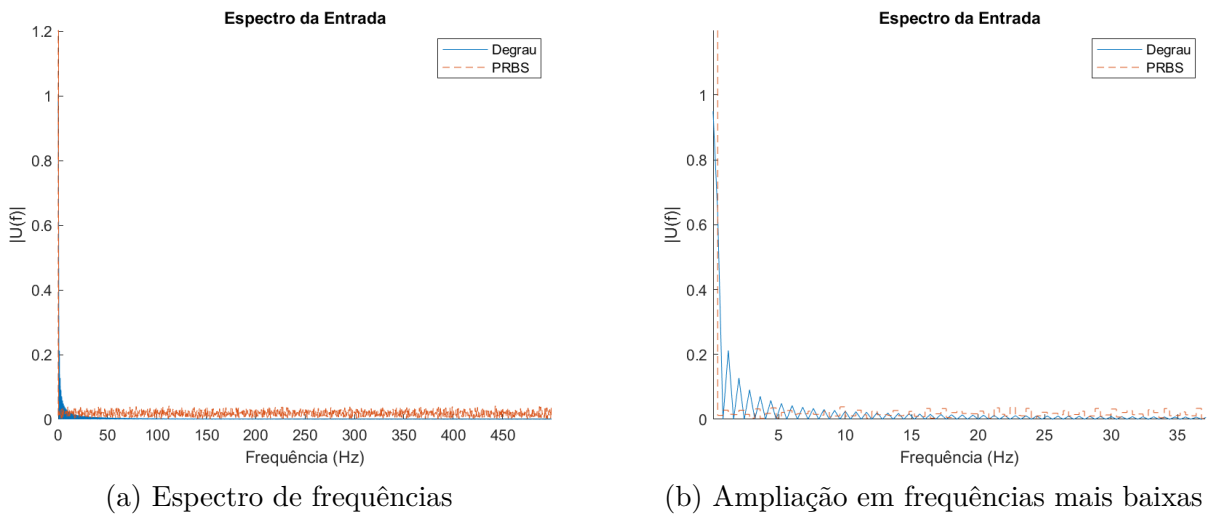
4 EXPERIMENTOS

A fim de ilustrar e validar os métodos definidos no capítulo 3, foram primeiramente realizados testes em ambientes computacionais e laboratoriais para sistemas de menor complexidade, tanto lineares quanto não lineares. Tal procedimento foi realizado com o intuito de aumentar o nível de familiaridade com processos de identificação antes de aplicar as mesmas metodologias a sistemas mais complexos, tais como o *twin-rotor*.

4.1 Sinal PRBS e degrau unitário

Inicialmente, é interessante comparar o espectro de frequências, bem como a qualidade da identificação do sistema de interesse, em função do sinal de entrada escolhido. Com tal objetivo em mente, serão avaliados os resultados em função da entrada PRBS proposta e da resposta ao degrau, método amplamente utilizado na identificação. Também serão utilizados todos os procedimentos relevantes a sistemas lineares, apresentados no capítulo 3.

Figura 24 – Comparação entre o espectro de frequências do sinal PRBS de degrau unitário, simulação de $2.5s$ e $F_s = 10^3 Hz$.



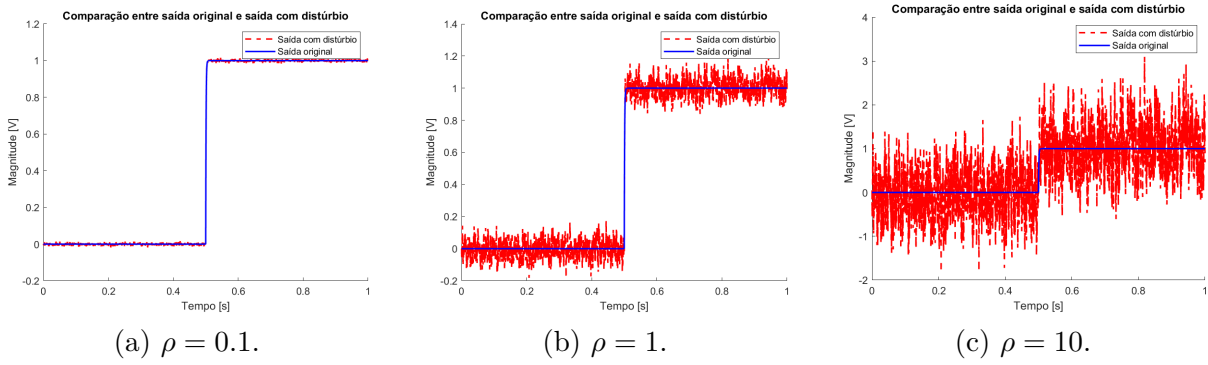
Fonte: Autoria própria.

Comparando fig. 24a e fig. 24b, é possível afirmar que uma variável com caráter aleatório possui componentes em uma faixa de frequências muito superior à de sinais convencionais, como o degrau unitário. Portanto, ainda que seja possível obter representações precisas a partir da identificação por um degrau unitário, fica claro que o uso de sinais de banda larga, como o PRBS, é preferível. Tal afirmação se mostra ainda mais precisa para sistemas de alta complexidade, como já abordado previamente.

Além disso, é interessante definir uma variável ρ que contém a informação sobre o quão significativa é a magnitude do ruído em relação ao sinal de entrada de interesse. Como exemplo, foi obtida a resposta de um sistema RC para diferentes valores de ρ , apresentada na figura 25.

$$\rho = \frac{\|e[k]\|_{\infty}}{\|u[k]\|_{\infty}}. \quad (4.1)$$

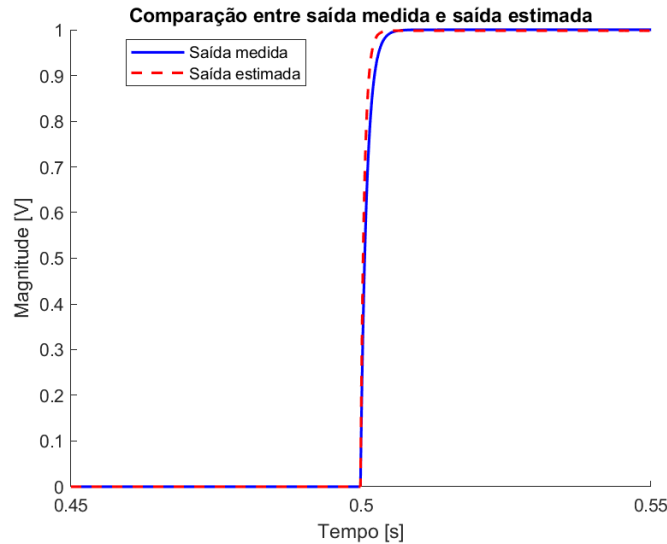
Figura 25 – Resposta ao sistema para diferentes distúrbios na entrada, simulação de 1s e $F_s = 10^4 Hz$.



Fonte: Autoria própria.

Vale ressaltar que, para fins de identificação, ainda é possível obter um modelo satisfatório mesmo em condições nas quais ρ indica a presença de um ruído significativo nos dados. A imagem 26 foi obtida por meio da aplicação do MMQ a um modelo ARX nos dados de 25c, utilizando a mesma entrada sem distúrbios para fins comparativos. Para a simulação, foi utilizado o circuito de primeira ordem, apresentado na fig. 27 e na tabela 5.

Figura 26 – Comparação do modelo obtido para $\rho = 10$ e saída ideal. Simulação de 1s e $F_s = 10^4 Hz$.



Fonte: Autoria própria.

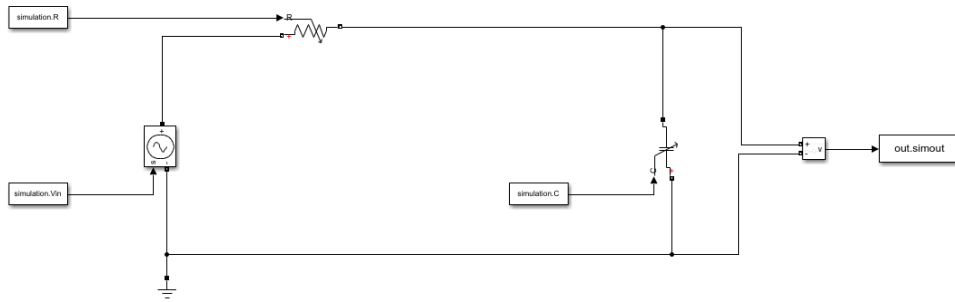
Tabela 4 – Dados de validação para o modelo da figura 26.

Validação da saída y	Valor obtido
MSE	$5.1418 \cdot 10^{-29}$
FIT	100%
FPE	$2.2213 \cdot 10^{-16}$

Fonte: Autoria própria.

4.1.1 Comparação em um circuito RC: simulação computacional

A partir dos exemplos apresentados nas figuras 25 e 26, é possível abordar em mais detalhes o comportamento do circuito RC de interesse, bem como as demais características e resultados relevantes à identificação de sistemas.

Figura 27 – Circuito RC implementado no *Simulink*.

Fonte: Autoria própria.

Tabela 5 – Componentes utilizados para a simulação.

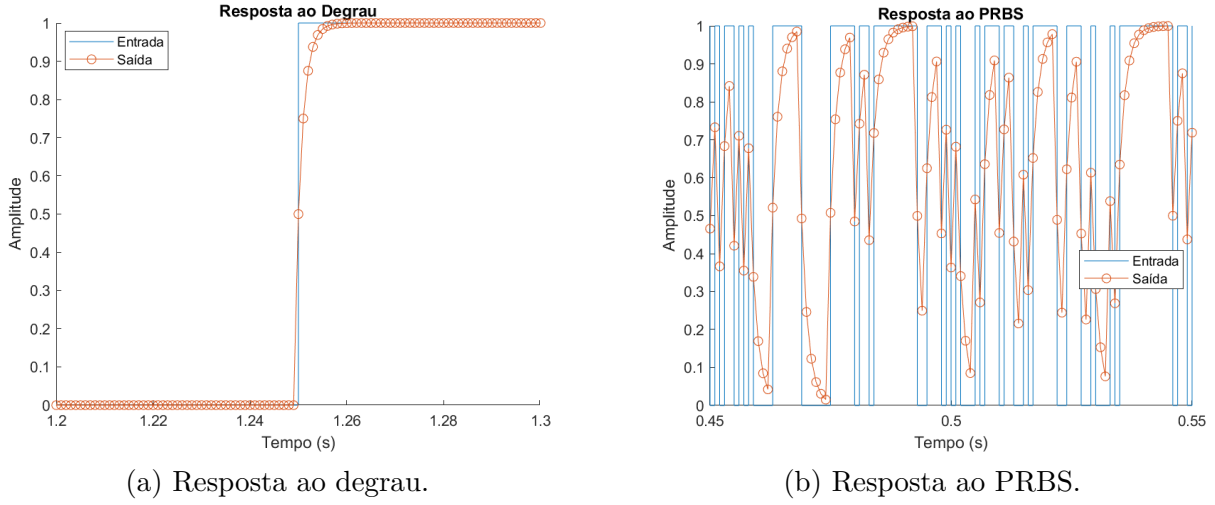
Componente	Valor associado
R	$1 \text{ } k\Omega$
C	$1 \text{ } \mu F$

Fonte: Autoria própria.

$$H_{RC} = \frac{1}{RCs + 1} = \frac{1}{10^{-3}s + 1} = \frac{10^3}{s + 10^3}. \quad (4.2)$$

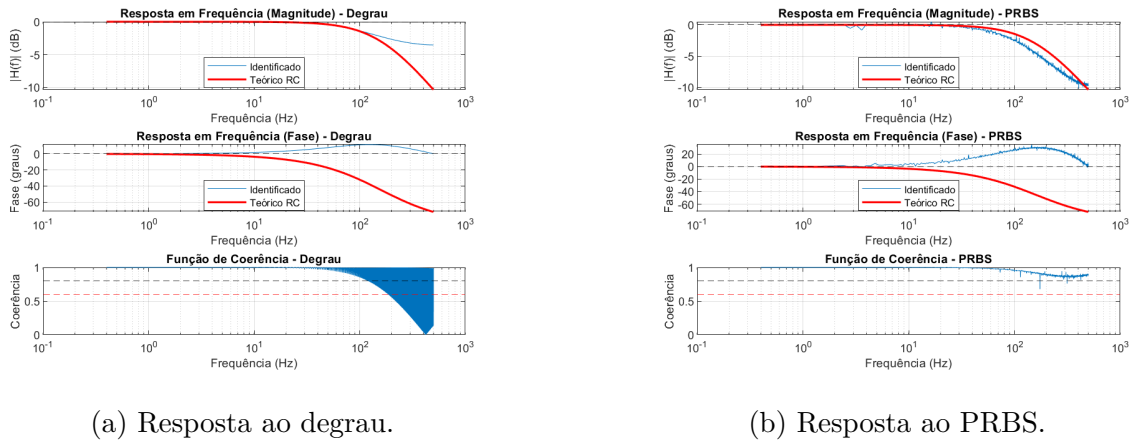
Desse modo, foi obtida a resposta em frequência em função de ambas as entradas degrau e PRBS. Além disso, considerando que se encontra em ambiente computacional, também é possível averiguar a qualidade da estimação diretamente a partir do vetor de parâmetros θ

Figura 28 – Comparação entre respostas do sistema RC ao sinal PRBS e degrau unitário, simulação de $2.5s$ e $F_s = 10^3 Hz$.



Fonte: Autoria própria.

Figura 29 – Comparação entre respostas em frequência do sistema RC ao sinal PRBS e degrau unitário, simulação de $2.5s$ e $F_s = 10^3 Hz$.



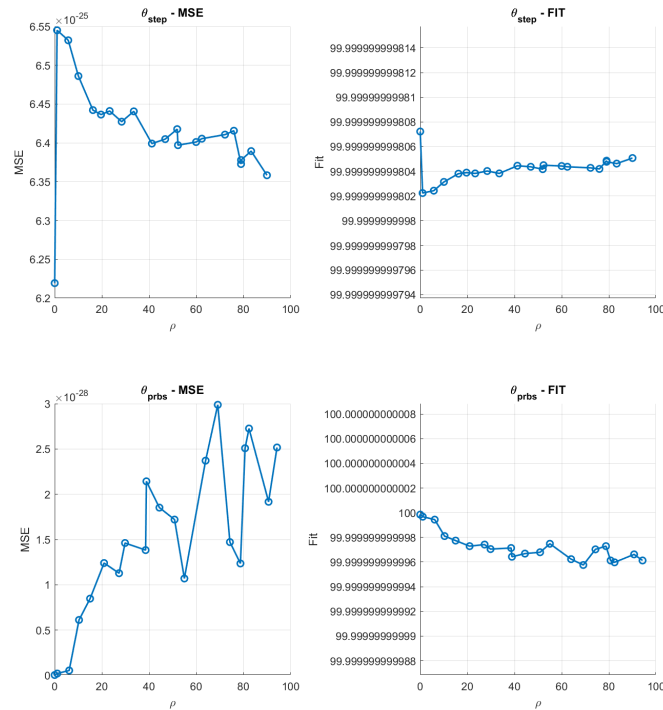
Fonte: Autoria própria.

Avaliando, primeiramente, a resposta em frequência apresentada na figura 29, é possível afirmar que o resultado obtido para a entrada PRBS representa, de maneira mais fiel, o comportamento esperado do circuito, característica que se manifesta com maior intensidade em frequências mais altas. Dessa forma, também é possível afirmar que os resultados obtidos na fase do sistema não foram satisfatórios para nenhum dos sinais de entrada escolhidos.

Além disso, é possível afirmar que a função de coerência é representativa da qualidade do modelo na frequência de interesse, em especial da magnitude da resposta em frequência. Assim sendo, ainda que formalmente represente o quanto a entrada explica o comportamento observado na saída do sistema, a função de coerência também é indicativa da qualidade da estimativa em cada frequência.

A figura 30 apresenta o comportamento dos parâmetros MSE e FIT em função do aumento do nível de ruído no sistema (aumento de ρ), tanto para uma entrada PRBS quanto para uma entrada do tipo degrau unitário. É notável que, considerando a escala utilizada, há pouca variação na qualidade da estimação com o aumento de ρ . Ainda assim, devem ser considerados diversos aspectos que também afetam esta influência do ruído no processo de identificação, como não linearidades, a presença de acoplamento entre entradas e saídas, a ordem do sistema ou até mesmo o número de medições obtidas do mesmo. Ainda assim, é possível constatar que a aplicação do MMQ a modelos do tipo ARX apresenta características intrínsecas de rejeição ao ruído.

Figura 30 – Comportamento da estimação do vetor de parâmetros em função de ρ , simulação de 1s e $F_s = 10^4 Hz$.



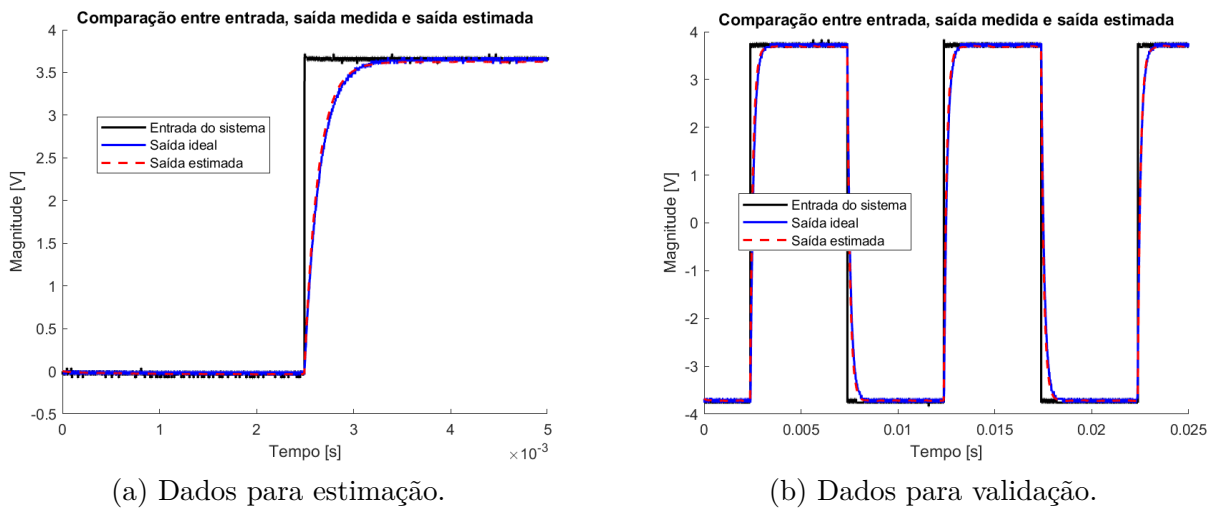
Fonte: Autoria própria.

É importante observar que o modelo identificado a partir da entrada PRBS apresenta uma modelagem mais próxima dos dados do sistema real diante de distúrbios, como mostrado em 30. Além disso, vale destacar que se espera uma certa variabilidade nos dados obtidos, considerando as características estocásticas do distúrbio utilizado. Foi também necessário aumentar a taxa de amostragem para $F_s = 10^4 Hz$ para a obtenção da figura, considerando que o parâmetro FIT_{step} iniciava em 15% mesmo na modelagem ideal.

4.2 Exemplo de identificação simples: circuito RC

Com base no que foi apresentado para o circuito computacional, também foram realizadas a aferição e validação de um circuito de primeira ordem utilizando-se de dados reais obtidos em laboratório, resultando na figura 31. Percebe-se que os resultados para os dados de validação, assim como para o caso simulado, são tão satisfatórios quanto os obtidos para a estimação. Os parâmetros foram obtidos a partir da resposta ao degrau, posteriormente ajustados em função da nova frequência de amostragem e comparados com um sinal de onda quadrada.

Figura 31 – Dados do circuito RC obtidos em laboratório. Para estimação e validação: $5ms$, $F_s = 5 \cdot 10^5 Hz$; $25ms$, $F_s = 1 \cdot 10^5 Hz$.



Fonte: Autoria própria.

Tabela 6 – Dados de validação para o modelo da figura 31.

Validação da saída y	Valor para estimação	Valor para validação
MSE	$78 \cdot 10^{-4}$	$173 \cdot 10^{-4}$
FIT	97.55%	96.34%
FPE	$78 \cdot 10^{-4}$	$173 \cdot 10^{-4}$

Fonte: Autoria própria.

Desse modo, embora não seja preferível para a modelagem, é possível obter uma representação satisfatória do sistema de interesse a partir de uma entrada degrau. Vale destacar apenas que tal afirmação é aplicável a circuitos de menor complexidade e com poucas influências não lineares, assim como o utilizado nesta seção.

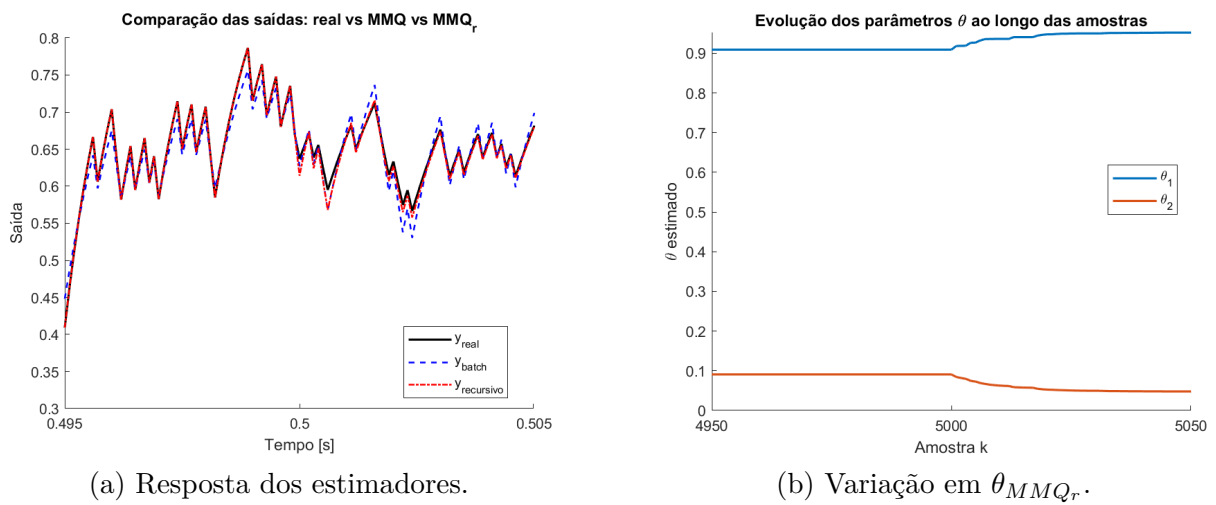
4.3 Alteração paramétrica durante a identificação

Considerando a possibilidade de o sistema de interesse depender de alterações paramétricas ao longo do tempo, é interessante avaliar se o método utilizado consegue

identificar tais variações durante o processo de estimação. Essa capacidade é extremamente interessante, por exemplo, para a aplicação em sistemas com controle adaptativo, nos quais já se espera uma variação durante o processo de identificação e controle. Para tal, foi aplicado o MMQR descrito em seção 3.3.2 a uma alteração da carga resistiva do circuito denotado em fig. 27, visto que ele já permite essa variação por meio da alteração de variáveis no *workspace* do Matlab.

$$R[k] \in \{0.5s \leq k \leq 1s\} = 2 \cdot R[k] \in \{0s \leq k < 0.5s\} \quad (4.3)$$

Figura 32 – Dados do circuito RC para o MMQ recursivo, com ampliação no instante em que a carga resistiva é aumentada. Simulação de $1s$, $F_s = 1 \cdot 10^4 Hz$



Fonte: Autoria própria.

Tabela 7 – Dados para o MMQ e o MMQR, $\lambda = 0.9$

Validação da saída y	Valor para MMQ	Valor para MMQR
MSE	$4.7156 \cdot 10^{-4}$	$3.1261 \cdot 10^{-7}$
FIT	76.93%	99.41%
FPE	$4.7175 \cdot 10^{-4}$	$3.1274 \cdot 10^{-7}$

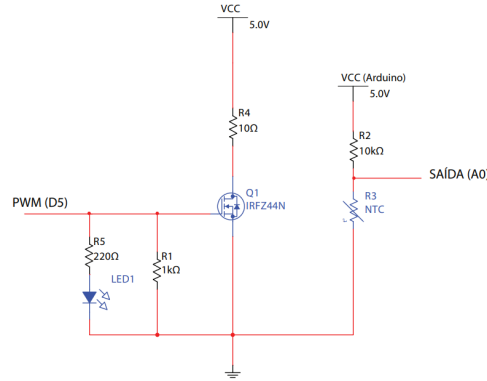
Fonte: Autoria própria.

Desse modo, a partir da figura 32a e da tabela 7, é possível afirmar que o estimador recursivo apresenta um desempenho muito superior ao do estimador MMQ em sistemas variantes no tempo.

4.4 Exemplo de identificação simples: dinâmica de temperatura

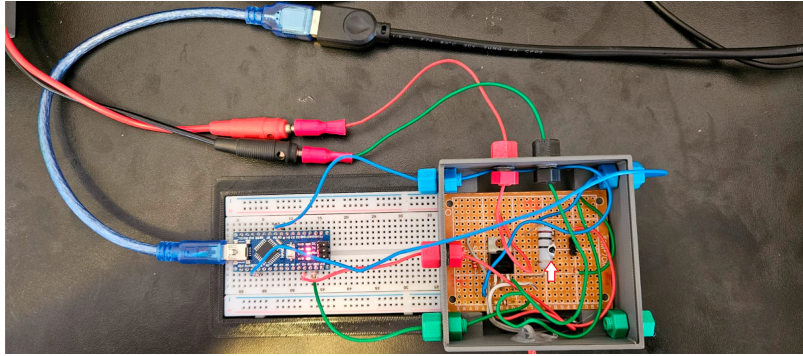
Análoga à realizada na seção 4.1.1, a aferição de temperatura de um circuito resistivo foi realizada, conforme apresentado em fig. 34. Os dados obtidos em função da entrada PRBS foram utilizados para estimar um modelo, tanto pelo método ARX quanto por meio de uma função de transferência, utilizando o comando *tfest*.

Figura 33 – Esquemático do circuito para aferição de temperatura.



Fonte: [Fernandes 2025].

Figura 34 – Circuito para aferição de temperatura, sensor NTC indicado.



Fonte: [Fernandes 2025].

Considerando que o sistema em estudo envolve dinâmicas de temperatura, a temperatura foi aproximada por uma Função de Transferência (FT) de primeira ordem, conforme apresentada na equação 4.4. Desse modo, é possível desenvolver eq. (4.4) para obter a Equação Diferencial Ordinária (EDO) 4.6.

$$G(s) = \frac{k}{Ts + 1}, \quad (4.4)$$

$$Y(s) \cdot (Ts + 1) = k \cdot U(s), \quad (4.5)$$

$$T \cdot \frac{d}{dt}y(t) + y(t) = k \cdot u(t). \quad (4.6)$$

Considerando uma discretização da formulação apresentada em 4.6, dada pela aproximação $\frac{d}{dt}y(t) = \frac{y[k] - y[k-1]}{T_s}$, é possível obter eq. (4.9) a partir de eq. (4.7).

$$T \cdot \frac{y[k] - y[k-1]}{T_s} + y[k] = k \cdot u[k], \quad (4.7)$$

$$\frac{T + T_s}{T_s} \cdot y[k] = \frac{T}{T_s} y[k - 1] + k \cdot u[k], \quad (4.8)$$

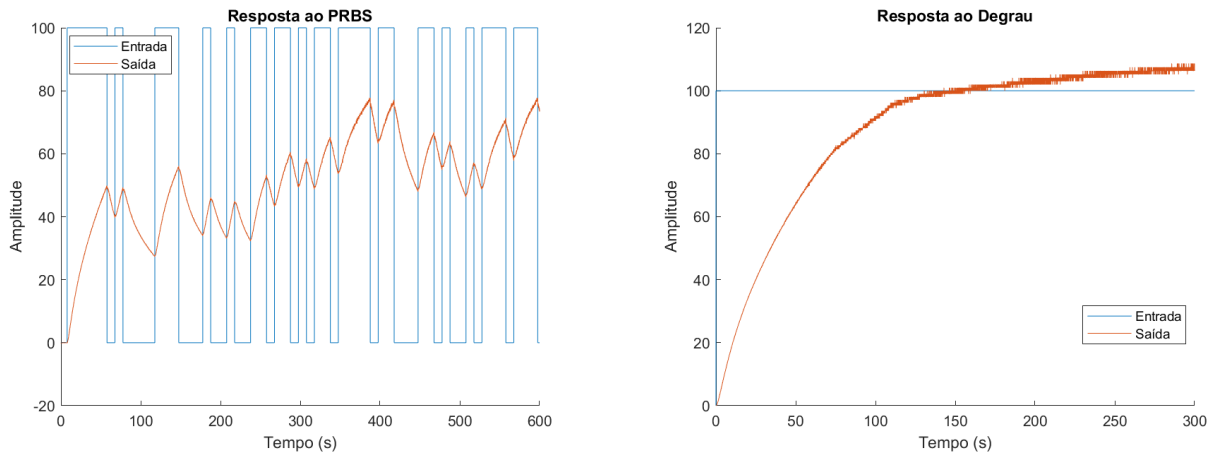
$$\cdot y[k] = \frac{T}{T + T_s} y[k - 1] + \frac{kT_s}{T + T_s} \cdot u[k]. \quad (4.9)$$

Por fim, é possível reescrever eq. (4.9) de modo a obter eq. (4.11), tal que se encontrem $\theta_1 = \frac{T}{T+T_s}$ e $\theta_2 = \frac{kT_s}{T+T_s}$ que resultem na formulação do MMQ para o modelo de interesse:

$$\cdot y[k] = \theta_1 \cdot y[k - 1] + \theta_2 \cdot u[k], \quad (4.10)$$

$$\cdot y[k] = \begin{bmatrix} y[k - 1] & u[k] \end{bmatrix} \cdot \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}. \quad (4.11)$$

Figura 35 – Entradas dos dados obtidos para o sistema de temperatura. Para estimação e validação: 10 min, $F_s = 50Hz$; 5 min, $F_s = 50Hz$.



(a) Dados para estimação

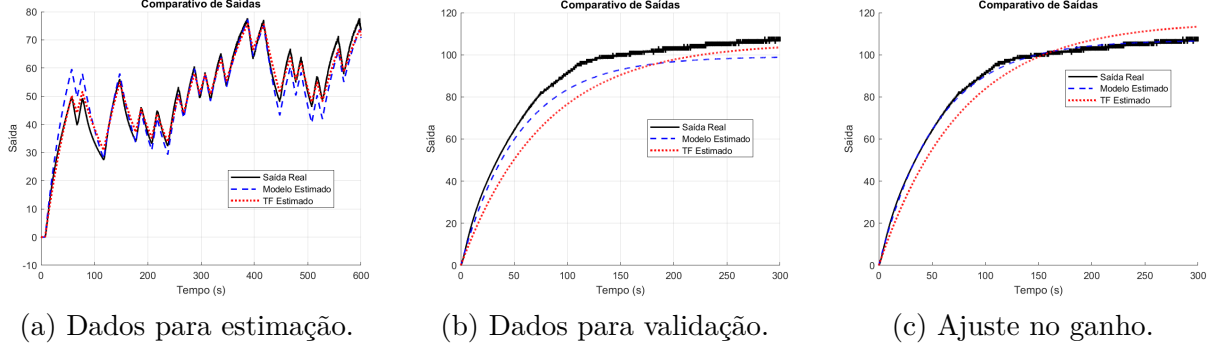
(b) Dados para validação

Fonte: Autoria própria.

Para fins de validação, também foi efetuado um ajuste manual no ganho encontrado para ambas as metodologias de estimação utilizadas, visando maximizar a semelhança entre os sinais de validação e os obtidos em laboratório. Tal ajuste mostrou-se necessário, visto que, embora o MMQ e o estimador de FT utilizados tenham obtido resultados satisfatórios tanto na estimação quanto na dinâmica do sistema como um todo, não conseguiram reproduzir fielmente o ganho na validação. Desse modo, a partir da figura 36c e da tabela 10, é visível que o uso do método ARX se comprova como uma alternativa mais fidedigna à dinâmica do sistema de interesse. Vale também mencionar que, como abordado em [Ford e Hasbun 2019], existem diversas dinâmicas não lineares associadas ao comportamento

dissipativo de um resistor, que não foram consideradas em nenhum dos modelos (ARX e FT) utilizados.

Figura 36 – Saídas estimadas para os dados obtidos do sistema de temperatura.



Fonte: Autoria própria.

Tabela 8 – Dados de estimação para o modelo da figura 36a

Validação da saída y	Valor para ARX	Valor para FT
MSE	14.2718	5.2588
FIT	75.21%	84.95%
FPE	14.2747	5.2595

Fonte: Autoria própria.

Tabela 9 – Dados de validação para o modelo da figura 36b

Validação da saída y	Valor para ARX	Valor para FT
MSE	34.4915	98.5323
FIT	76.99%	61.11%
FPE	34.5053	95.5586

Fonte: Autoria própria.

Tabela 10 – Dados de validação para o modelo da figura 36c

Validação da saída y	Valor para ARX	Valor para FT
MSE	2.2571	38.7603
FIT	94.11%	75.61%
FPE	2.2580	38.7707

Fonte: Autoria própria.

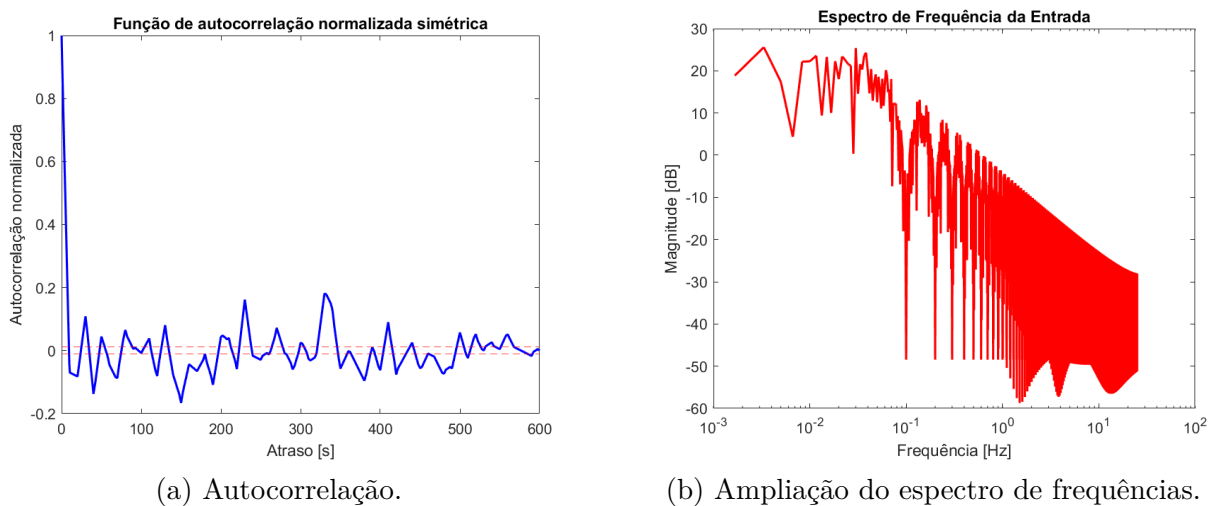
$$\hat{y} = 1.07 \cdot \begin{bmatrix} y[k-1] & u[k] \end{bmatrix} \begin{bmatrix} 0.9996 \\ 0.0004 \end{bmatrix}. \quad (4.12)$$

Por fim, é importante destacar que o gráfico da função de autocorrelação do sinal PRBS utilizado não atende aos limites impostos em 2.9. Para fins de comprovação do nível

de excitação do sistema, basta adquirir o espectro de frequências da entrada. Portanto, a partir da comparação entre 37b e 24, é possível confirmar que o resultado obtido ainda é satisfatório no domínio da frequência, embora ocorra uma grande atenuação em dinâmicas mais velozes.

Tal efeito decorre da inserção de um segurador na entrada PRBS usual, evitando que dinâmicas excessivamente rápidas prejudiquem o funcionamento e a vida útil do dispositivo de chaveamento utilizado. Outro fator importante a se considerar é que componentes de frequência mais elevada possuem pouca ou nenhuma influência na dinâmica de interesse, tendo em vista a escala de tempo utilizada e a característica de os sistemas físicos atuarem como filtros passa-baixa.

Figura 37 – Função de autocorrelação e espectro de frequências da entrada PRBS apresentada em 35a.

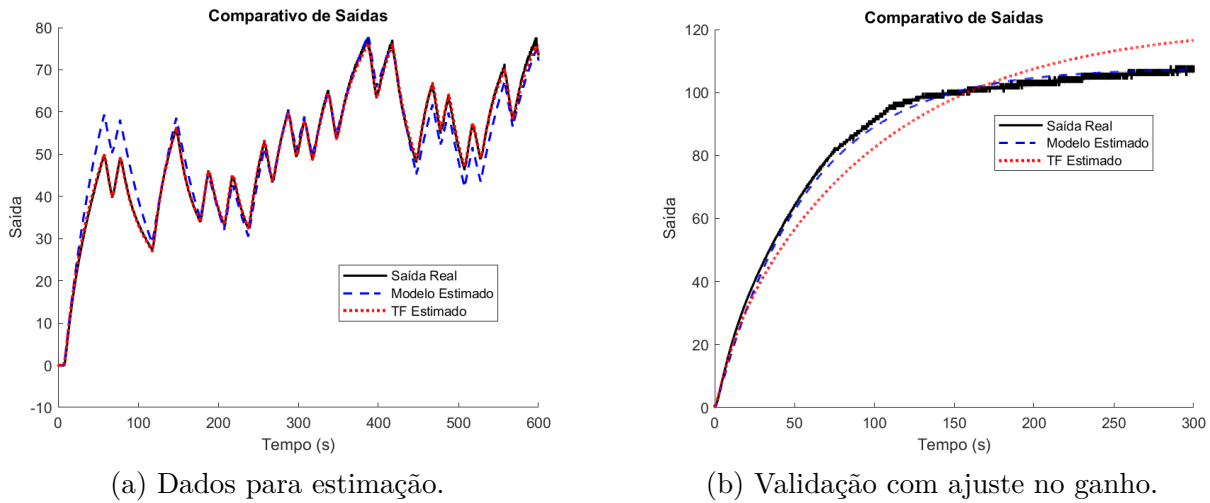


Fonte: Autoria própria.

4.4.1 Influência do overfitting

Com base nos dados obtidos para o circuito real em laboratório, com o intuito de avaliar a influência do *overfitting* sobre os dados de validação do modelo, foi novamente efetuada a estimação do circuito de primeira ordem, agora com 50 regressores na entrada e 50 na saída. Como visto na comparação entre as figuras 36 e 38, embora o modelo represente satisfatoriamente os dados de estimação, a simples inserção de um número maior de regressores pode prejudicar o sistema.

Figura 38 – Saídas estimadas para 100 regressores.



Fonte: Autoria própria.

Tabela 11 – Dados de estimação para o modelo da figura 38a.

Validação da saída y	Valor para ARX	Valor para FT
MSE	13.9081	0.3069
FIT	75.53%	96.37%
FPE	14.0011	0.3090

Fonte: Autoria própria.

Tabela 12 – Dados de validação para o modelo da figura 38b.

Validação da saída y	Valor para ARX	Valor para FT
MSE	2.9527	44.8955
FIT	93.27%	73.75%
FPE	2.9923	45.4980

Fonte: Autoria própria.

É importante mencionar que o modelo estimado não apresentou melhora significativa com o aumento da ordem do sistema, mesmo com a utilização de uma grande quantidade de regressores. O mesmo ocorre para os dados de validação, em que houve apenas uma pequena queda nos indicadores utilizados.

4.4.2 Uso do LASSO como ferramenta de seleção de regressores

Levando em consideração os resultados obtidos, é possível afirmar que o *overfitting* leva a um aumento do viés do modelo estimado, visto que há um melhor desempenho no treinamento, a custa de uma perda de generalização. Assim, ainda que o impacto encontrado em 4.4.1 não tenha sido considerável, tal tendência pode se mostrar significativa para a qualidade do modelo em casos nos quais a quantidade de regressores seja ainda mais elevada.

Além disso, o tempo necessário para calcular os parâmetros, bem como o custo computacional associado ao processo de estimação, aumentam com o número de operações associadas a uma ordem superior. Desse modo, como abordado em 3.3.3, é possível utilizar o algoritmo LASSO para identificar os parâmetros mais significativos do sistema, evitando problemas de desempenho ou um custo computacional muito elevado.

Tabela 13 – Tempo necessário para estimar e validar o sistema, conforme a figura 34. Notebook modelo Lenovo Ideapad 82CGS00100.

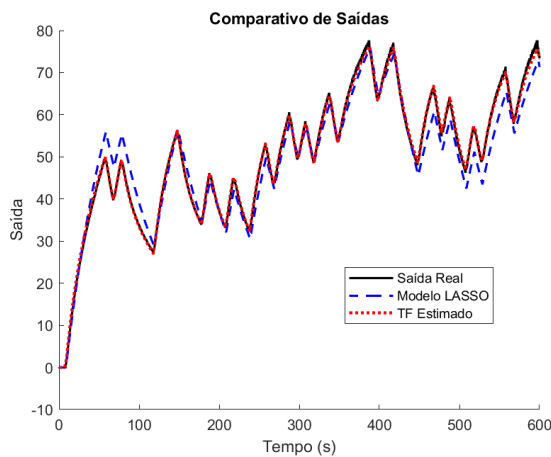
2 regressores	100 regressores
2.6133 [s]	14.8803 [s]

Fonte: Autoria própria.

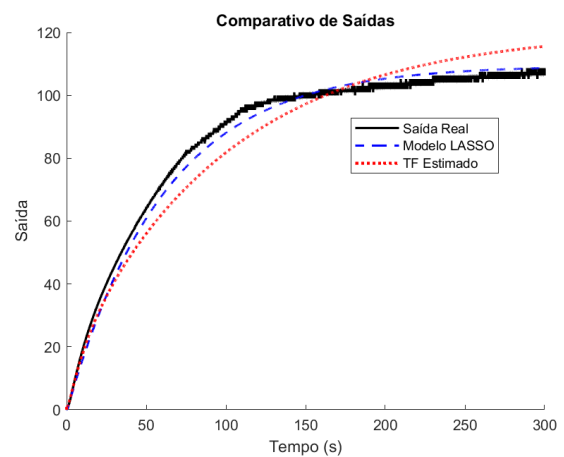
Dessa forma, os valores de θ_{LASSO} foram analisados iterativamente, a fim de reduzir gradativamente o número de regressores e recuperar apenas os mais significativos para a modelagem do sistema. A figura 39 apresenta a saída do método LASSO, em que os regressores mais significativos foram $y[k-1]$, $y[k-2]$ e $u[k-1]$.

Vale destacar que, devido ao elevado custo computacional de modelar um sistema de ordem 10 por meio do método LASSO, foram utilizados 5 regressores na entrada e 5 na saída. Além disso, por conhecimento do sistema de interesse, é, de fato, impossível que uma representação de ordem tão elevada denote de maneira verossímil a dinâmica desejada.

Figura 39 – Saída para 3 regressores, com o uso do método LASSO em um modelo de 100 parâmetros.



(a) Dados para estimação.



(b) Validação com ajuste no ganho.

Fonte: Autoria própria.

Tabela 14 – Dados em função do modelo LASSO dado pela figura 39

Validação da saída y	Valor para estimação	Valor para validação
MSE	14.3428	2.1815
FIT	75.15%	94.21%
FPE	14.3447	2.1821

Fonte: Autoria própria.

$$\hat{y} = 1.095 \cdot \begin{bmatrix} \hat{y}[k-1] & \hat{y}[k-2] & u[k-1] \end{bmatrix} \begin{bmatrix} 0.5196 \\ 0.4799 \\ 0.0005 \end{bmatrix}. \quad (4.13)$$

Analisando os dados obtidos, é possível afirmar que, para casos em que não se conhece o sistema de interesse, o método LASSO oferece uma boa estimativa para a identificação dos regressores mais significativos. Além disso, é perceptível uma qualidade inferior dos dados de estimação em troca de uma melhora na validação do modelo, seja em relação à representação dada pela equação 4.12 ou à ilustrada na figura 38. Dessa forma, é possível estimar um modelo a partir do MMQ usual, ou até mesmo sua versão recursiva, com base nos dados encontrados a partir do LASSO.

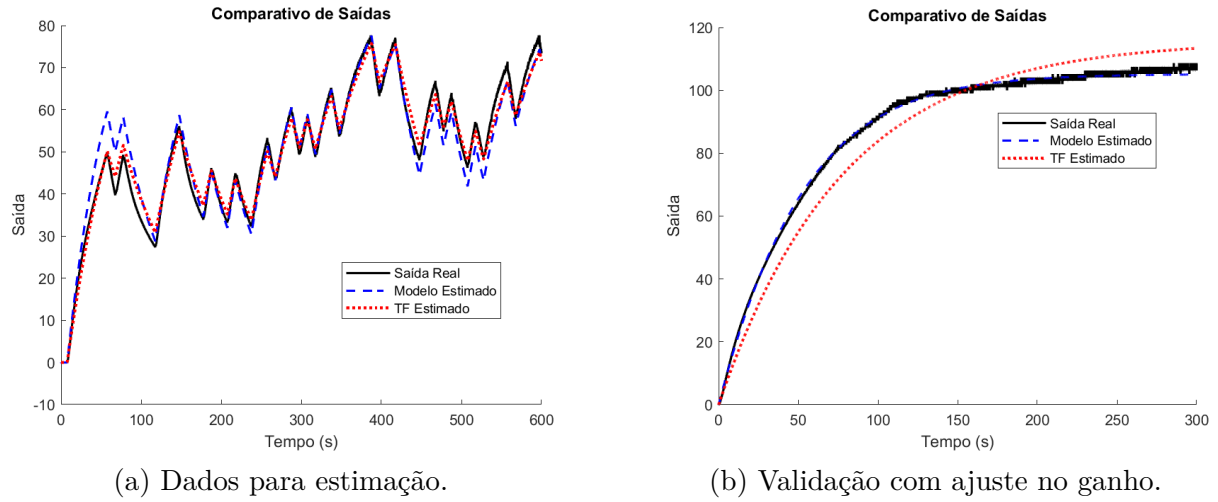
4.5 Identificação de sistemas não lineares

Com base no desenvolvimento das seções 4.1.1 e 4.4, é possível aplicar todas as técnicas já validadas e comprovadas a fim de obter uma modelagem satisfatória do *twin-rotor*. Serão também introduzidas as classes de modelos não lineares dadas pelo NARX, embora as métricas para aferir o desempenho dessas sejam idênticas às apresentadas em 3.4.

Levando em consideração as diversas não linearidades relacionadas ao processo de aquecimento e resfriamento de um resistor, conforme abordado em [Ford e Hasbun 2019], foram adicionados termos polinomiais não lineares à equação característica 4.13, resultando na modelagem indicada na figura 40. Assim sendo, as características fundamentais do sistema são definidas pelas componentes lineares, enquanto os demais resíduos podem ser aproximados por termos polinomiais ou até mesmo por redes neurais [Aguirre 2004].

Neste caso, foi introduzido um termo dependente da relação entre a entrada e a saída do sistema, além de um termo de quarta ordem na saída e de um atraso unitário.

Figura 40 – Adição de um termo dependente de $y[k-1]^4 \cdot u[k-1]$ ao sistema obtido na seção 4.4.2



Fonte: Autoria própria.

Tabela 15 – Dados para o modelo não linear proposto.

Validação da saída y	Valor para estimação	Valor para validação
MSE	14.8011	1.1119
FIT	74.75%	95.87%
FPE	14.8031	1.1122

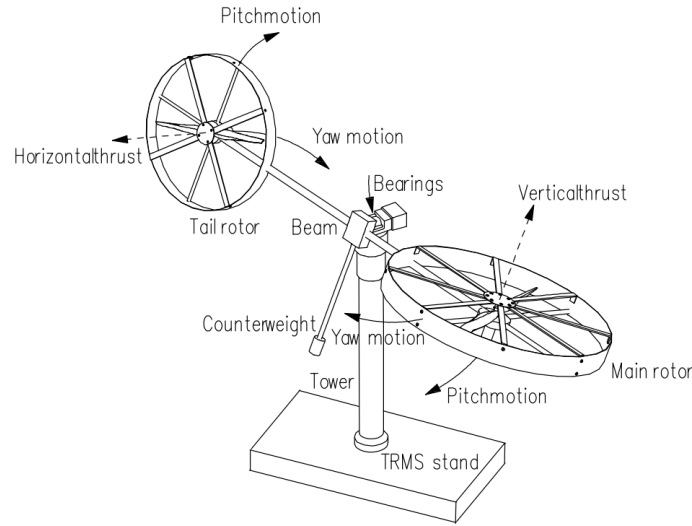
Fonte: Autoria própria.

Com base na tabela 15, é possível afirmar que a inclusão de apenas um termo não linear aumentou significativamente a precisão do modelo, especialmente nos dados de validação. Embora a dependência de um termo de quarta ordem tenha sido deduzida a partir da relação entre a saída aferida e a temperatura do sistema [Aguirre 2004], é perfeitamente plausível aplicar um modelo NARX do tipo caixa-preta e, posteriormente, delimitar os parâmetros mais significativos por meio do LASSO.

5 ESTUDO DE CASO: TWIN ROTOR

Por fim, aplicando os métodos apresentados ao longo do desenvolvimento do trabalho, foi possível identificar o tipo caixa cinza do modelo computacional do *twin-rotor* presente no Laboratório de Controle (LAC). Para tal, foi realizada a simulação das equações fornecidas pelo próprio fabricante da planta, apresentadas em eq. (5.1) a eq. (5.4). As tabelas tabela 16 e tabela 17 apresentam as dinâmicas e os parâmetros sugeridos para a modelagem, enquanto fig. 41 ilustra o comportamento do *twin-rotor*.

Figura 41 – Sistema de interesse.



Fonte: retirado de [Huu e Ismail 2016].

Vale também destacar que, como abordado em seção 3.3.1, desde que haja linearidade entre os regressores e os parâmetros, ainda é possível utilizar todo o desenvolvimento do MMQ e do LASSO. Por fim, também será utilizada a entrada PRMLS, conforme desenvolvida em seção 3.1.3.

$$\ddot{\psi} = \frac{M_1 - M_{FG} - M_{B1} - M_G}{I_1}, \quad (5.1)$$

$$\ddot{\varphi} = \frac{M_2 - M_{B2} - M_R}{I_2}, \quad (5.2)$$

$$\dot{\tau}_i = -\frac{1}{T_i}\tau_i + \frac{k_i}{T_i}u_i, \quad (5.3)$$

$$\dot{M}_R = -\frac{1}{T_p}M_R + \frac{k_c \cdot T_0}{T_p}\dot{\tau}_1 + \frac{k_c}{T_p}\tau_1. \quad (5.4)$$

As variáveis M_{B1} e M_{B2} , bem como os demais estados intermediários, são descritas em mais detalhes na tabela 16. Também foram incluídos termos para simplificar as formulações 5.1 a 5.4, sendo eles $i \in \{1, 2\}$ e $\kappa \in \{\kappa_1, \kappa_2\}$, tais que $\kappa_1 = \psi$ e $\kappa_2 = \varphi$.

Tabela 16 – Parâmetros dinâmicos referentes às equações 5.1 e 5.2

Parâmetro	Formulação	Significado físico
M_i	$a_i \cdot \tau_i^2 + b_i \cdot \tau_i$	Característica estática
τ_i	$\frac{k_i}{T_{is+1}} \cdot u_i$	Dinâmica dos motores
M_{FG}	$M_g \cdot \sin(\psi)$	<i>Momentum</i> da gravidade
M_{Bi}	$B_{1i} \cdot \dot{\kappa}_i + B_{2i} \cdot \text{sign}(\dot{\kappa}_i)$	<i>Momentum</i> de fricção
M_G	$K_{gy} \cdot M_1 \cdot \varphi \cdot \cos(\psi)$	<i>Momentum</i> do giroscópio
M_R	$k_c \frac{T_0+1}{T_p+1} \tau_1$	<i>Momentum</i> de acoplamento

Fonte: Autoria própria.

Tabela 17 – Parâmetros do modelo TRMS: significado e valores numéricos

Parâmetro	Descrição	Valor	Unidade
I_1	Momento de inércia do rotor vertical	$6.8 \cdot 10^{-2}$	$\text{kg} \cdot \text{m}^2$
I_2	Momento de inércia do rotor horizontal	$2 \cdot 10^{-2}$	$\text{kg} \cdot \text{m}^2$
a_1	Parâmetro de característica estática	$1.35 \cdot 10^{-2}$	–
b_1	Parâmetro de característica estática	$9.24 \cdot 10^{-2}$	–
a_2	Parâmetro de característica estática	$2 \cdot 10^{-2}$	–
b_2	Parâmetro de característica estática	$9 \cdot 10^{-2}$	–
M_g	<i>Momentum</i> da gravidade	0.32	$\text{N} \cdot \text{m}$
B_{11}	Parâmetro do <i>momentum</i> de fricção	$6 \cdot 10^{-3}$	$\text{N} \cdot \text{m} \cdot \text{s}/\text{rad}$
B_{21}	Parâmetro do <i>momentum</i> de fricção	$1 \cdot 10^{-3}$	$\text{N} \cdot \text{m} \cdot \text{s}^2/\text{rad}$
B_{12}	Parâmetro do <i>momentum</i> de fricção	$1 \cdot 10^{-1}$	$\text{N} \cdot \text{m} \cdot \text{s}/\text{rad}$
B_{22}	Parâmetro do <i>momentum</i> de fricção	$1 \cdot 10^{-2}$	$\text{N} \cdot \text{m} \cdot \text{s}^2/\text{rad}$
K_{gy}	Parâmetro do <i>momentum</i> de giroscópico	0.05	s/rad
k_1	Ganho do motor de movimento vertical	1.1	–
k_2	Ganho do motor de movimento horizontal	0.8	–
k_c	Ganho do <i>momentum</i> de acoplamento	-0.2	–
T_1	Constante de tempo do motor vertical	1.1	–
T_2	Constante de tempo do motor horizontal	1	–
T_0	Constante de tempo do zero de acoplamento	3.5	–
T_p	Constante de tempo do polo de acoplamento	2	–
u_1	Tensão de controle aplicada ao motor vertical	–	V
u_2	Tensão de controle aplicada ao motor horizontal	–	V
ψ	Ângulo do movimento vertical (pitch)	–	rad
φ	Ângulo do movimento horizontal (yaw)	–	rad

Fonte: [Feedback Instruments Ltd. 2013].

5.1 Discretização

Como o presente trabalho tem como objetivo efetuar a identificação paramétrica no domínio discreto, faz-se necessário discretizar as equações definidoras do sistema de interesse. Tal processo também é de suma importância para definir corretamente os regressores na modelagem do tipo caixa cinza, evitando o *overfitting*. Aplica-se, portanto, o mesmo princípio utilizado em eq. (4.4) a eq. (4.11), a saber: $\frac{d}{dt}x \approx \frac{x[k] - x[k-1]}{T_s}$, com T_s como tempo de amostragem. Dessa forma, pode-se tomar a seguinte equivalência

$$\dot{x} = f(x) \Rightarrow \frac{x[k] - x[k-1]}{T_s} = f(x[k-1]).$$

A partir das equações diretamente relacionadas às entradas do *twin-rotor*, é interessante iniciar o processo de discretização retomando as equações 5.3 e 5.4. Primeiramente, para τ_i , obtêm-se:

$$\begin{aligned} \dot{\tau}_i &\approx \frac{\tau_i[k] - \tau_i[k-1]}{T_s} = -\frac{1}{T_i}\tau_i[k-1] + \frac{k_i}{T_i}u_i[k-1], \\ \therefore \tau_i[k] &= (1 - \frac{T_s}{T_i})\tau_i[k-1] + \frac{T_s k_i}{T_i}u_i[k-1], \end{aligned} \quad (5.5)$$

Os termos M_i para $i \in \{1, 2\}$ podem ser reescrito a partir de eq. (5.6) e M_{FG} de eq. (5.7). As saídas ψ e φ foram renomeadas para y_1 e y_2 , para manter a consistência com a nomenclatura utilizada no restante do trabalho.

$$M_i[k] = a_i \cdot \tau_i^2[k] + b_i \cdot \tau_i[k], \quad (5.6)$$

$$M_{FG}[k] = M_g \cdot \sin(y_1[k]). \quad (5.7)$$

Tomando agora o termo M_{Bi} e já efetuando a troca de nomenclatura das saídas, é possível obter eq. (5.9). Além disso, a partir da definição dada em eq. (5.8), é possível omitir a dependência de $\frac{1}{T_s}$.

$$\text{sign}(x) = \begin{cases} 1, & \forall x > 0, \\ 0, & \text{se } x = 0, \\ -1, & \forall x < 0. \end{cases} \quad (5.8)$$

$$M_{Bi}[k] = \frac{B_{1i}}{T_s}y_i[k] - \frac{B_{1i}}{T_s}y_i[k-1] + B_{2i}\text{sign}(y_i[k] - y_i[k-1]). \quad (5.9)$$

Da mesma maneira, o parâmetro M_G é dado por

$$M_G[k] = K_{gy} \cdot M_1[k] \cdot y_2[k] \cdot \cos(y_1[k]). \quad (5.10)$$

Por fim, para o parâmetro M_R :

$$\dot{M}_R \approx \frac{M_R[k] - M_R[k-1]}{T_s} = -\frac{1}{T_p} M_R[k-1] + \frac{k_c \cdot T_0}{T_p} \cdot \frac{\tau_1[k-1] - \tau_1[k-2]}{T_s} + \frac{k_c}{T_p} \tau_1[k-1],$$

$$\therefore M_R[k] = (1 - \frac{T_s}{T_p}) M_R[k-1] + \frac{k_c \cdot T_0}{T_p} \cdot (\tau_1[k-1] - \tau_1[k-2]) + \frac{T_s k_c}{T_p} \tau_1[k-1].$$

Desse modo, é possível definir o conjunto de equações descrito pela formulação eq. (5.11). Para simplificação, foi alterada a notação para incluir os parâmetros auxiliares α e β .

$$\begin{cases} \tau_i[k] = \alpha_i \tau_i[k-1] + \beta_i u_i[k-1], \\ M_i[k] = a_i \cdot \tau_i^2[k] + b_i \cdot \tau_i[k], \\ M_{FG}[k] = M_g \cdot \sin(y_1[k]), \\ M_{Bi}[k] = \frac{B_{1i}}{T_s} y_i[k] - \frac{B_{1i}}{T_s} y_i[k-1] + B_{2i} \text{sign}(y_i[k] - y_i[k-1]), \\ M_G[k] = K_{gy} \cdot M_1[k] \cdot y_2[k] \cdot \cos(y_1[k]), \\ M_R[k] = \alpha_3 \cdot M_R[k-1] + \beta_3 \cdot \tau_1[k-1] - \beta_4 \cdot \tau_1[k-2]. \end{cases} \quad (5.11)$$

Por fim, aplicando o mesmo às equações das saídas do sistema (eq. (5.1) e eq. (5.2)), obtêm-se eq. (5.12) e eq. (5.13).

$$\ddot{\psi} \approx \frac{y_1[k] - 2y_1[k-1] - y_1[k-2]}{T_s^2} = \frac{M_1[k-1] - M_{FG}[k-1] - M_{B1}[k-1] - M_G[k-1]}{I_1}, \quad (5.12)$$

$$\ddot{\phi} \approx \frac{y_2[k] - 2y_2[k-1] - y_2[k-2]}{T_s^2} = \frac{M_2[k-1] - M_{B2}[k-1] - M_R[k-1]}{I_2}. \quad (5.13)$$

Portanto, o modelo não-linear completo do twin-rotor discretizado é dado por

$$y_1[k] = 2y_1[k-1] - y_1[k-2] + T_s^2 \frac{M_1[k-1] - M_{FG}[k-1] - M_{B1}[k-1] - M_G[k-1]}{I_1},$$

$$\begin{aligned}
y_2[k] &= 2y_2[k-1] + y_2[k-2] + T_s^2 \frac{M_2[k-1] - M_{B2}[k-1] - M_R[k-1]}{I_2}, \\
M_1[k-1] &= a_1 \cdot \tau_1^2[k-1] + b_1 \cdot \tau_1[k-1], \\
M_2[k-1] &= a_2 \cdot \tau_2^2[k-1] + b_2 \cdot \tau_2[k-1], \\
\tau_1[k-1] &= \alpha_1 \tau_1[k-2] + \beta_1 u_1[k-1], \\
\tau_2[k-1] &= \alpha_2 \tau_2[k-2] + \beta_2 u_2[k-1], \\
M_{FG}[k-1] &= M_g \cdot \sin(y_1[k-1]), \\
M_{B1}[k-1] &= \frac{B_{11}}{T_s} y_1[k-1] - \frac{B_{11}}{T_s} y_1[k-2] + B_{21} \text{sign}(y_1[k-1] - y_1[k-2]), \\
M_{B2}[k-1] &= \frac{B_{12}}{T_s} y_2[k-1] - \frac{B_{12}}{T_s} y_2[k-2] + B_{22} \text{sign}(y_2[k-1] - y_2[k-2]), \\
M_G[k-1] &= K_{gy} \cdot M_1[k-1] \cdot y_2[k-1] \cdot \cos(y_1[k-1]), \\
M_R[k-1] &= \alpha_3 \cdot M_R[k-2] + \beta_3 \cdot \tau_1[k-1] - \beta_4 \cdot \tau_1[k-2]
\end{aligned}$$

5.2 Estimação de parâmetros

Considerando todo o desenvolvimento realizado até o momento, as mesmas técnicas do capítulo 4 serão aplicadas à identificação de um modelo computacional do *twin-rotor*. A fim de obter uma modelagem mais realista, serão inseridos ruídos na saída do sistema.

Por simplicidade, suponha

$$M_1[k] = \gamma_1 u_1[k], \quad (5.14)$$

$$M_2[k] = \gamma_2 u_2[k], \quad (5.15)$$

$$M_R[k] = \beta_3 \gamma_1 u_1[k] - \beta_4 \gamma_1 u_1[k-1] \quad (5.16)$$

Dessa forma, tem-se

$$\begin{aligned}
y_1[k] &= 2y_1[k-1] + y_1[k-2] + \frac{T_s^2 \gamma_1}{I_1} u_1[k-1] - \frac{T_s^2}{I_1} M_g \sin(y_1[k-1]) - \frac{T_s^2}{I_1} \frac{B_{11}}{T_s} y_1[k-1] + \\
&\quad \frac{T_s^2}{I_1} \frac{B_{11}}{T_s} y_1[k-2] - \frac{T_s^2}{I_1} B_{21} \text{sign}(y_1[k-1] - y_1[k-2]) \\
&\quad - \frac{T_s^2}{I_1} K_{gy} \gamma_1 u_1[k-1] y_2[k-1] \cos(y_1[k-1]), \quad (5.17)
\end{aligned}$$

e

$$\begin{aligned}
y_2[k] &= 2y_2[k-1] + y_2[k-2] + \frac{T_s^2}{I_2} \gamma_2 u_2[k-1] - \frac{T_s^2}{I_2} \frac{B_{12}}{T_s} y_2[k-1] + \\
&\quad \frac{T_s^2}{I_2} \frac{B_{12}}{T_s} y_2[k-2] - \frac{T_s^2}{I_2} B_{22} \text{sign}(y_2[k-1] - y_2[k-2]) \\
&\quad - \frac{T_s^2}{I_2} \beta_3 \gamma_1 u_1[k-1] - \frac{T_s^2}{I_2} \beta_4 \gamma_1 u_1[k-2]. \quad (5.18)
\end{aligned}$$

Assim, fica evidente o acoplamento existente entre as saídas $y_1[k]$, $y_2[k]$ e os sinais de entrada $u_1[k-1]$, $u_2[k-1]$.

Note ainda que

$$y_1[k] = \theta_1 y_1[k-1] + \theta_2 y_1[k-2] + \theta_3 u_1[k-1] + \theta_4 \sin(y_1[k-1]) + \theta_5 \text{sign}(y_1[k-1] - y_1[k-2]) + \theta_6 u_1[k-1] y_2[k-1] \cos(y_1[k-1]), \quad (5.19)$$

e

$$y_2[k] = \xi_1 y_2[k-1] + \xi_2 y_2[k-2] + \xi_3 u_2[k-1] + \xi_4 \text{sign}(y_2[k-1] - y_2[k-2]) + \xi_5 u_1[k-1] + \xi_6 u_1[k-2]. \quad (5.20)$$

Logo, conclui-se, através de uma abordagem caixa-cinza, que a saída y_1 é dada pela combinação linear dos regressores:

$$\{y_1[k-1], y_1[k-2], u_1[k-1], \sin(y_1[k-1]), \text{sign}(y_1[k-1] - y_1[k-2]), u_1[k-1] y_2[k-1] \cos(y_1[k-1])\}.$$

Já a saída y_2 é dada pela combinação linear dos regressores

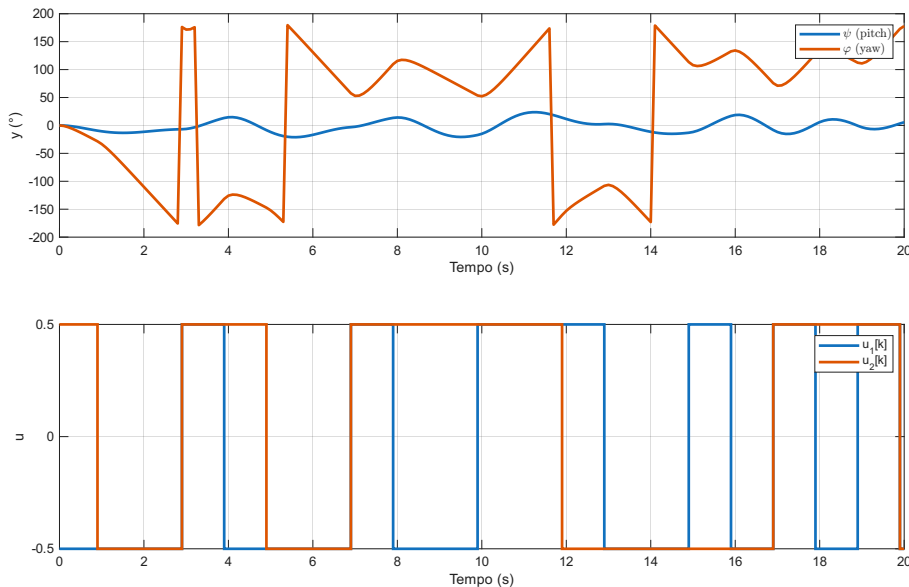
$$\{y_2[k-1], y_2[k-2], u_2[k-1], \text{sign}(y_2[k-1] - y_2[k-2]), u_1[k-1], u_1[k-2]\}.$$

Consequentemente, a estimação dos parâmetros do modelo pode ser realizada usando o MMQ, dado um conjunto de entrada-saída $\mathcal{D} = \{y_1[k], y_2[k], u_1[k], u_2[k]\}_{k=0}^N$.

5.2.1 Identificação do modelo

Usando as equações diferenciais do sistema *twin-rotor*, foi realizada a simulação do sistema com uma entrada PRBS em cada canal u_1 e u_2 . Os dados simulados resultantes são apresentados na figura 42.

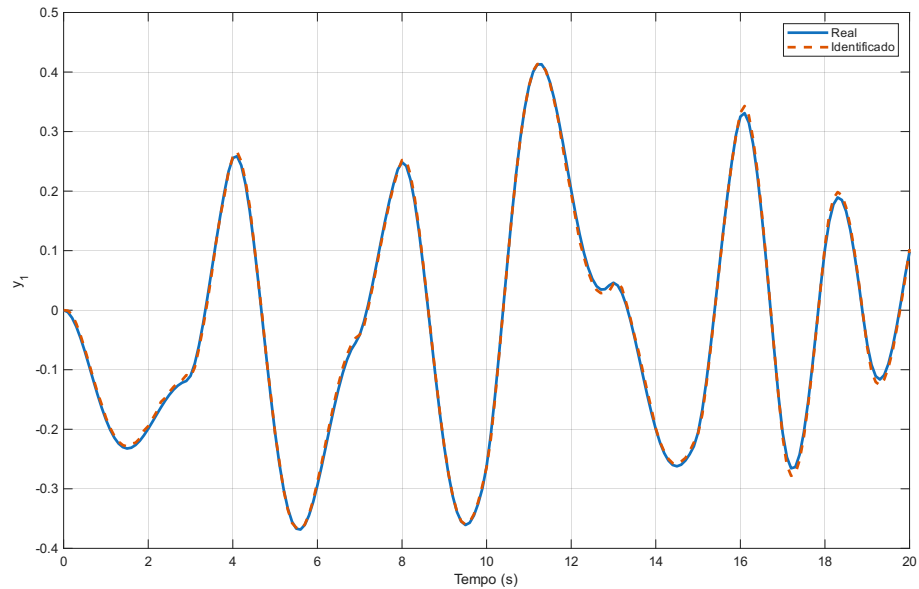
Figura 42 – Conjunto de dados simulados, gerados com base no modelo de twin-rotor, para identificação.



Fonte: autoria própria.

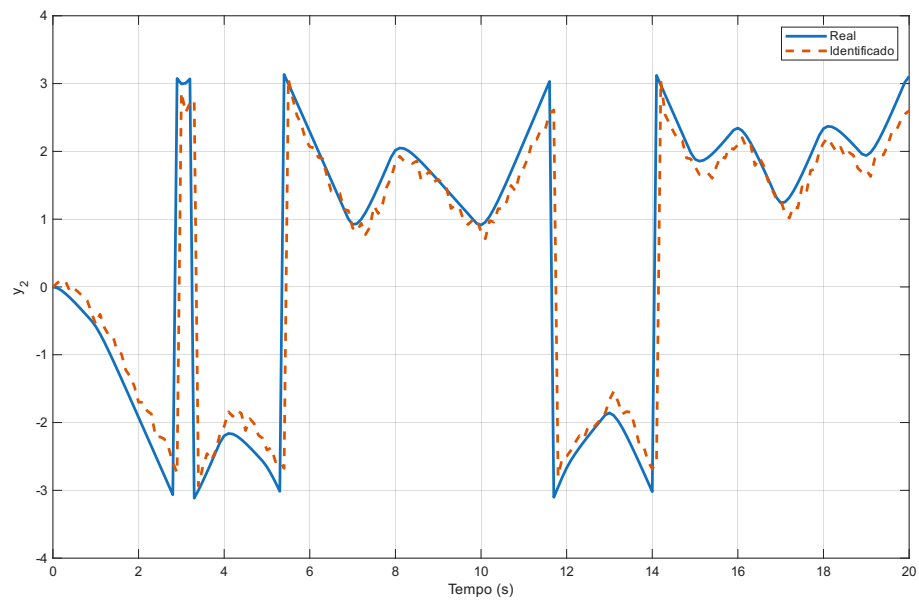
Usando a família de regressões definida pela abordagem caixa-cinza descrita anteriormente na seção 5.2, aplicou-se o MMQ para estimar os parâmetros do modelo. A figura 43 apresenta o resultado da identificação para a saída $y_1[k]$. Já a figura 44 apresenta o resultado da identificação para a saída $y_2[k]$.

Figura 43 – Resultado da identificação para saída y_1 .



Fonte: autoria própria.

Figura 44 – Resultado da identificação para saída y_2 .



Fonte: autoria própria.

6 CONCLUSÃO

Este trabalho foi desenvolvido com o intuito de identificar técnicas adequadas para a estimação paramétrica de sistemas lineares e não lineares. Os objetivos propostos foram alcançados por meio do desenvolvimento, da ilustração e da validação de diversas metodologias de modelagem de sistemas dinâmicos, culminando na aplicação bem-sucedida de um modelo dinâmico do *twin-rotor*.

A análise comparativa entre as técnicas demonstrou a importância da escolha estratégica do sinal de excitação, com o PRBS e o PRMLS mostrando-se mais eficazes do que o degrau unitário na excitação do espectro de frequências das plantas analisadas. Em relação à estimação, o MMQ mostrou-se eficaz em todos os casos em que foi aplicado.

Em especial, o uso do LASSO também se mostrou uma ferramenta valiosa para lidar com o *trade-off* entre a precisão do modelo e sua capacidade de generalização. Ao penalizar e eliminar regressores de baixa relevância, o LASSO não só se comprovou eficaz na redução do custo computacional do processo de estimação, como também em evitar o *overfitting*, resultando em modelos com maior capacidade de generalização.

No estudo de caso do *twin-rotor*, a abordagem de caixa cinza conseguiu integrar efetivamente o conhecimento das equações físicas à estimação de parâmetros no domínio discreto, permitindo obter modelos que descreveram satisfatoriamente as dinâmicas acopladas do sistema. Portanto, a aplicação das técnicas estudadas em um sistema de alta complexidade confirma a validade do desenvolvimento teórico e prático proposto.

6.1 Trabalhos futuros

Por fim, considerando a quantidade de técnicas que podem ser aplicadas para modelar satisfatoriamente o sistema de interesse, há diversos aspectos que merecem atenção especial e que podem levar ao desenvolvimento de inúmeros trabalhos futuros. Primeiramente, considerando a dinâmica de temperatura abordada na seção 4.4, é possível também determinar a validade de identificar modelos distintos para o aquecimento e o resfriamento, visto que ambos envolvem parâmetros, ou até mesmo equacionamentos, com características distintas. Além disso, é válido avaliar o desempenho do LASSO em comparação com as demais alternativas de regularização apresentadas em seção 4.4.2, tendo em vista que todas têm a capacidade de contornar os efeitos do *overfitting* e o aumento do custo computacional. Ainda é possível avaliar o desempenho das diferentes estruturas apresentadas na tabela 2, em particular, para estimações caixa-preta não lineares envolvendo o sistema de temperatura e o *twin-rotor*.

Além disso, no experimento realizado com o *twin-rotor*, pode-se estender a aborda-

gem para dados coletados diretamente de uma planta física real, explorar variações na família de regressores e realizar uma identificação paramétrica baseada no LASSO. Assim, é possível buscar regressores que melhorem a precisão do modelo em uma planta twin-rotor real, com base nas ferramentas apresentadas neste trabalho.

Do ponto de vista do controle, também é possível desenvolver técnicas de identificação em malha fechada, a fim de modelar adequadamente sistemas que necessitam intrinsecamente de *feedback* para sua estabilidade. Por fim, também é possível aplicar as técnicas de modelagem abordadas ao longo do trabalho para projetar um controlador para a planta *twin-rotor* presente no LAC, tais como controladores do tipo *sliding mode* ou até mesmo Linear Quadratic Regulator (LQR).

REFERÊNCIAS

- AGUIRRE, L. A. **Introdução à Identificação de Sistemas: Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais**. 2. ed. Belo Horizonte, Brasil: Editora UFMG, 2004. ISBN 978-85-7041-400-7.
- AHMAD, A. J. T. M. O.; CHIPPERFIELD, S. M. Dynamic modelling and open-loop control of a twin rotor multi-input multi-output system. **Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering**, v. 216, sep 2002.
- BAKALIS, E. *et al.* Daughter coloured noises: The legacy of their mother. **Entropy**, MDPI, v. 26, n. 9, sep 2023. ISSN 1099-4300. Disponível em: <https://doi.org/10.3390/e25091371>.
- BIRKHOFF, G. D. Proof of the ergodic theorem. **Proceedings of the National Academy of Sciences of the United States of America**, v. 17, n. 12, p. 656–660, 1931.
- CHEN, C.-T. **Linear System Theory and Design**. [*S.l.: s.n.*]: Oxford University Press, 1998.
- DARUS, I. Z. M.; ALDEBREZ, F. M.; TOKHI, M. O. Parametric modelling of a twin rotor system using genetic algorithms. *In: Control, Communications and Signal Processing, 2004. First International Symposium on*. [*S.l.: s.n.*], 2004. p. 115–118.
- FARINA, A. **Impulse Response Measurement Techniques and their Applicability in the Real World**. 2000. AES Convention Paper 108. Disponível em: <http://pcfarina.eng.unipr.it/Public/Papers/226-AES108.pdf>.
- Feedback Instruments Ltd. **Twin Rotor MIMO System: Control Experiments**. Ed. 02. Park Road, Crowborough, East Sussex, TN6 2QX, United Kingdom, 2013. Manual 33-949S, Part No. 1160-33949S. For use with MATLAB. Disponível em: <https://www.feedback-instruments.com>.
- FERNANDES, M. R. **Experimento 2 - Controle de Temperatura - parte 1**. 2025. Roteiro de Laboratório, SEL0359 - Laboratório de Controle Digital, 2º Semestre de 2025. Última atualização: 5 de setembro de 2025.
- FERNANDES, M. R.; OLIVEIRA, M. O. F. d.; SOUTO, R. F. Construção de um protótipo de helicóptero de baixo custo para estudos em identificação de sistemas. *In: Anais do XIII Simpósio Brasileiro de Automação Inteligente (SBAI)*. Porto Alegre, RS: Sociedade Brasileira de Automática, 2017. p. 1177–1183. Disponível em: <https://www.researchgate.net/publication/324150302>.
- FORD, K. S.; HASBUN, J. E. Modeling the temperature behavior of an rlc circuit. **Georgia Journal of Science**, v. 77, n. 2, 2019. Disponível em: <https://digitalcommons.gaacademy.org/gjs/vol77/iss2/13>.
- GAUSS, C. F. **Theoria motus corporum coelestium in sectionibus conicis solem ambientium**. Hamburg: Sumtibus F. Perthes et I. H. Besser, 1809. 246 p. Disponível em: https://archive.org/details/bub_gb_ORUOAAAAQAAJ.

GOEL, A.; BRUCE, A. L.; BERNSTEIN, D. S. Recursive least squares with variable-direction forgetting: Compensating for the loss of persistency. **arXiv preprint arXiv:2003.03523**, mar 2020. Disponível em: <https://arxiv.org/abs/2003.03523>.

GREENE, W. H. **Econometric Analysis**. 8. ed. [*S.l.: s.n.*]: Pearson, 2018. ISBN 0134461363.

GRIMMETT, G. R.; STIRZAKER, D. R. **Probability and Random Processes**. Third. Oxford and New York: Oxford University Press, 2001.

GUPTA, A.; STEAD, T. S.; GANTI, L. Determining a meaningful r-squared value in clinical medicine. **Academic Medicine & Surgery**, University Medical Press, oct 27 2024.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. 2. ed. Springer, 2009. ISBN 9780387848570. Disponível em: <https://link.springer.com/book/10.1007/978-0-387-84858-7>.

HASTIE, T.; TIBSHIRANI, R.; WAINWRIGHT, M. **Statistical Learning with Sparsity: The Lasso and Generalizations**. [*S.l.: s.n.*]: CRC Press, 2015. v. 143. (Monographs on Statistics and Applied Probability, v. 143).

HUU, T. D.; ISMAIL, I. B. Modelling of twin rotor mimo system. *In: 2016 2nd IEEE International Symposium on Robotics and Manufacturing Automation (ROMA)*. [*S.l.: s.n.*], 2016. p. 1–6.

ISERMANN, R.; MÜNCHHOF, M. **Identification of Dynamic Systems: An Introduction with Applications**. [*S.l.: s.n.*]: Springer, 2010.

JUANG, J.-G.; HUANG, M.-T.; LIU, W.-K. Pid control using presearched genetic algorithms for a mimo system. **IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)**, v. 38, n. 5, p. 716–727, 2008.

JUANG, J.-N. **Applied System Identification**. [*S.l.: s.n.*]: Prentice Hall, 1994.

LEGENDRE, A.-M. **Nouvelles méthodes pour la détermination des orbites des comètes**. Paris: F. Didot, 1805. 80 p. Disponível em: <https://archive.org/details/nouvellesmethode00legegoog>.

LJUNG, L. **System Identification: Theory for the User**. [*S.l.: s.n.*]: Prentice Hall, 1999.

OGATA, K. **Modern Control Engineering**. 5. ed. [*S.l.: s.n.*]: Prentice Hall, 2010.

RAHIDEH, A.; SHAHEED, M. Mathematical dynamic modelling of a twin-rotor multiple input-multiple output system. **Control Engineering Practice**, v. 16, n. 3, p. 241–259, 2008.

THEODORSEN, A. **Gaussian processes on compact metric spaces: spectral characterisation and rates of approximation**. 2016. Disponível em: <https://arxiv.org/abs/1609.01607>.

TIBSHIRANI, R. Regression shrinkage and selection via the lasso. **Journal of the Royal Statistical Society. Series B (Methodological)**, Wiley for the Royal Statistical Society, v. 58, n. 1, p. 267–288, 1996. Disponível em: <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>.

TOKER, O.; EMARA-SHABAIK, H. E. Construction of optimal multi-level pseudo-random sequences. **IEEE Transactions on Automatic Control**, Institute of Electrical and Electronics Engineers (IEEE), v. 40, n. 5, p. 891–896, may 2004. ISSN 0018-9286. Disponível em: <http://dx.doi.org/10.1109/TAC.2004.828311>.

VOIT, E. O. Perspective: Dimensions of the scientific method. **PLoS Computational Biology**, Public Library of Science (PLOS), v. 15, n. 9, p. e1007279, sep 2019. ISSN 1553-7358, 1553-734X. Disponível em: <https://doi.org/10.1371/journal.pcbi.1007279>.

WALPOLE, R. E. *et al.* **Probability and Statistics for Engineers & Scientists**. 9. ed. [*S.l.: s.n.*]: Pearson Education, 2011. ISBN 978-0-321-62911-1.

WASSERMAN, L. **All of Statistics: A Concise Course in Statistical Inference**. Springer, 2004. ISBN 0387402721. Disponível em: <https://link.springer.com/book/10.1007/978-0-387-21736-9>.

ZEGHLACHE, S. *et al.* Twin rotor mimo system experimental validation of robust adaptive fuzzy control against wind effects. **IEEE Systems Journal**, v. 16, n. 1, p. 409–419, 2022.