

UNIVERSIDADE FEDERAL DE ALAGOAS
Instituto de Computação

Sofia Lopes de Castro

**IDENTIFICAÇÃO DE SISTEMAS NÃO LINEARES
USANDO MODELOS NARX POLINOMIAIS**

Maceió
2025

Sofia Lopes de Castro

IDENTIFICAÇÃO DE SISTEMAS NÃO LINEARES USANDO MODELOS NARX POLINOMIAIS

Trabalho de Conclusão de Curso submetido
ao Instituto de Computação da Universidade
Federal de Alagoas como requisito parcial
para a obtenção do grau de Engenheira de
Computação.

Orientador: Prof. Dr. Ícaro Bezerra Queiroz de Araújo

Banca Avaliadora:

Glauber Rodrigues Leite	Prof. Dr., IC - UFAL
Andressa Martins Oliveira	Me.

Maceió
2025

**Catálogo na fonte
Universidade Federal de
Alagoas Biblioteca Central
Divisão de Tratamento Técnico**

Bibliotecária: Sâmela Rouse de Brito Silva – CRB-4 – 6023

C355i Castro, Sofia Lopes de.
Identificação de sistemas não lineares usando modelos NARX polinomiais /
Sofia Lopes de Castro. – 2025.
31 f. : il. color.

Orientador: Ícaro Bezerra Queiroz de Araújo.
Monografia (Trabalho de conclusão de curso em Engenharia de Computação)
– Universidade Federal de Alagoas. Instituto de Computação. Maceió, 2025.

Bibliografia: f. 29-31.

1. Sistemas dinâmicos não lineares. 2. Modelos NARX. 3. Identificação
de sistemas. I. Título.

CDU: 62-50

UNIVERSIDADE FEDERAL DE ALAGOAS
Instituto de Computação

IDENTIFICAÇÃO DE SISTEMAS NÃO LINEARES USANDO MODELOS NARX POLINOMIAIS

Trabalho de Conclusão de Curso submetido ao Instituto de Computação da Universidade Federal de Alagoas como requisito parcial para a obtenção do grau de Engenharia de Computação.

Aprovada em 28 de Abril de 2025:

Documento assinado digitalmente
 ICARO BEZERRA QUEIROZ DE ARAUJO
Data: 29/04/2025 08:24:43-0300
Verifique em <https://validar.iti.gov.br>

Ícaro Bezerra Queiroz de Araújo,
Prof. Dr., Orientador

Documento assinado digitalmente
 GLAUBER RODRIGUES LEITE
Data: 30/04/2025 12:55:07-0300
Verifique em <https://validar.iti.gov.br>

Glauber Rodrigues Leite,
Prof. Dr., IC - UFAL

Documento assinado digitalmente
 ANDRESSA MARTINS OLIVEIRA
Data: 30/04/2025 17:17:32-0300
Verifique em <https://validar.iti.gov.br>

Andressa Martins Oliveira,
Me.

Dedicatória

Aos meus pais.

Agradecimentos

À minha família, em especial ao meu pai João, *in memoriam*, que me inspirou em todos os passos da minha jornada e à minha mãe Ivete, minha primeira orientadora. Aos meus irmãos, Lucas e Júlia, por abrirem o caminho.

Aos meus professores, por todo o conhecimento transmitido durante a graduação e, em especial, ao meu orientador, Prof. Dr. Ícaro Bezerra Queiroz de Araújo, pelo direcionamento e imensa paciência. Aos professores que compuseram a banca avaliadora, Prof. Dr. Glauber Rodrigues Leite e Prof^a Me. Andressa Martins Oliveira.

Aos meus colegas da CGU, pelos ensinamentos profissionais. Aos colegas dos laboratórios LaCCAN e AI-SPARC, pelas trocas valiosas.

Aos meus amigos do IC, que tornavam a rotina mais leve, Álvaro, Rafael, Valério, João, Hugo, Kevin, Augusto, Pâmela e Roger. E aos meus amigos de antes, pelos momentos de apoio e também de distração. A Matheus, Anderson, Millena, Mariane, Rafaela, Lavínia, Rosana, Nayra e Luiza. E, em especial, ao meu namorado, Arthur, pelo carinho, ajuda e por acreditar em mim.

“Some say that knowledge is something that you never have”

Kate Bush, *Sat In Your Lap*

Resumo

Este trabalho apresenta a análise, implementação e comparação de dois métodos de identificação de sistemas baseados no modelo NARX polinomial: FROLS e SEMP, aplicados a três estudos de caso. O método FROLS utiliza uma abordagem de seleção progressiva, construindo um conjunto ortogonal de regressores e escolhendo, a cada etapa, aquele mais correlacionado com a saída do sistema. Por sua vez, o SEMP simula os modelos possíveis e adiciona apenas os regressores que contribuem significativamente para a redução do erro, descartando os excessivos. A comparação foi realizada a partir da identificação de três sistemas dinâmicos não lineares SISO, utilizando o RMSE como métrica de desempenho, com base na diferença entre a saída esperada e a saída estimada em simulação livre. Os resultados indicam que ambos os métodos são eficientes e robustos, com o SEMP apresentando desempenho superior em dois dos três casos analisados.

Palavras-chave: identificação de sistemas; sistemas não lineares; NARX; FROLS; SEMP.

Abstract

This work presents the analysis, implementation, and comparison of two system identification methods based on the polynomial NARX model: FROLS and SEMP, applied to three case studies. The FROLS method employs a forward selection approach, constructing an orthogonal set of regressors and selecting, at each step, the one most correlated with the system output. In turn, the SEMP method simulates possible model combinations and includes only the regressors that significantly contribute to error reduction, discarding redundant ones. The comparison was carried out through the identification of three nonlinear SISO dynamic systems, using RMSE as the performance metric, based on the difference between the expected output and the estimated output in free-run simulation. The results indicate that both methods are efficient and robust, with SEMP showing superior performance in two out of the three cases analyzed.

Keywords: system identification; nonlinear systems; NARX; FROLS; SEMP.

Lista de Figuras

1.1	Fluxograma do processo de identificação de sistemas.	15
3.1	Fotografia do sistema de ferro de solda.	30
3.2	Fotografia do sistema de tanques em cascata.	31
4.1	Gráfico de sinais de saída esperada e simulada do sistema simulado.	36
4.2	Gráfico de sinais de saída esperada e simulada do sistema de ferro de solda.	38
4.3	Gráfico de sinais de saída esperada e simulada do sistema de tanques em cascata.	39

Lista de Tabelas

4.1	Raiz Quadrada do Erro Quadrático Médio e quantidade de termos regressores do sistema simulado	37
4.2	Raiz Quadrada do Erro Quadrático Médio e quantidade de termos regressores do sistema de ferro de solda	38
4.3	Raiz Quadrada do Erro Quadrático Médio e quantidade de termos regressores do sistema de tanques em cascata	39

Lista de Abreviaturas

AR *Autoregressive*

ARMA *Autoregressive Moving Average*

ARMAX *Autoregressive Moving Average with eXogenous inputs*

ARX *Autoregressive with eXogenous inputs*

ERR *Error Reduction Ratio*

FROLS *Forward Regression Orthogonal Least Squares*

MIMO *Multiple Input Multiple Output*

MQ Mínimos Quadrados

MQO Mínimos Quadrados Ortogonais

MSE *Mean Squared Error*

MSSE *Mean Squared Simulation Error*

NAR *Nonlinear Autoregressive*

NARMA *Nonlinear Autoregressive Moving Average*

NARMAX *Nonlinear Autoregressive Moving Average with eXogenous inputs*

NARX *Nonlinear Autoregressive with eXogenous inputs*

RMSE *Root Mean Square Error*

SEMP *Simulation Error Minimisation with Pruning*

SISO *Single Input Single Output*

SRR *Simulation error Reduction Ratio*

Lista de Símbolos

$u(k)$ entrada do sistema.

$y(k)$ saída do sistema.

$e(k)$ ruído aditivo do sistema.

n_u atraso máximo de entrada.

n_y atraso máximo de saída.

ℓ grau de não linearidade dos modelos NARMAX.

Ψ matriz de regressores.

θ vetor de parâmetros.

ξ vetor de resíduo da saída.

Q matriz de regressores ortogonalizada.

A matriz triangular superior da ortogonalização de Gram-Schmidt.

g vetor de parâmetros na base ortogonal.

Sumário

1	Introdução	13
1.1	Motivação	13
1.2	Trabalhos relacionados	14
1.3	Objetivos	16
1.3.1	Objetivo Geral	16
1.3.2	Objetivos Específicos	16
1.4	Organização do Trabalho	17
2	Fundamentação Teórica	18
2.1	Representação de sistemas não lineares	18
2.2	Estimação de Parâmetros	20
2.2.1	Método de Mínimos Quadrados - MMQ	20
2.2.2	Mínimos Quadrados Ortogonais - MQO	21
2.3	Método <i>Forward Regression with Orthogonal Least Squares</i> - FROLS	24
2.4	Método <i>Simulation Error Minimization with Pruning</i> - SEMP	26
2.4.1	Simulation Error Reduction Ratio - SRR	26
2.4.2	Procedimento de poda	27
3	Metodologia	29
3.1	Descrição dos sistemas	29
3.1.1	Sistema Simulado	29
3.1.2	Sistema de Ferro de Solda	30
3.1.3	Sistema de Tanques em Cascata	31
3.2	Seleção de estrutura e estimação de parâmetros	32
3.3	Validação do modelo	35
4	Resultados	36
4.1	Sistema Simulado	36
4.2	Sistema de Ferro de Solda	37
4.3	Sistema de Tanques em Cascata	38
5	Conclusão	40
	Referências	41

1 Introdução

1.1 Motivação

Na indústria, a modelagem de sistemas desempenha um papel importante na otimização de processos e na maximização da eficiência operacional. Através da modelagem, é possível representar sistemas complexos com modelos análogos que reproduzem, de forma aproximada, as suas características e o seu comportamento, permitindo a análise e a previsão de seu desempenho sob diferentes condições. No entanto, modelar sistemas fenomenologicamente envolve identificar e considerar a sua estrutura, componentes e variáveis relevantes, além das leis físicas que regem os processos dentro do sistema, o que, para sistemas desconhecidos ou muito complexos, é desafiador e muitas vezes impraticável (Ljung, 1999). Nesses casos, utiliza-se o que é conhecido na engenharia de controle como identificação de sistemas.

Segundo Coelho e Coelho (2004), a identificação de sistemas é o processo de determinação de um modelo matemático, a partir de dados experimentais de entrada e saída, representando a dinâmica de um sistema em seus aspectos essenciais de forma adequada para um propósito particular. Na área de automação industrial, esses propósitos são predição, análise e projeto de sistemas de controle, supervisão, otimização e detecção de falha (Coelho e Coelho, 2004; Nelles, 2001).

Para realizar a identificação de um sistema arbitrário, a literatura dispõe de uma ampla gama de algoritmos capazes de gerar modelos de naturezas também diversas. Tradicionalmente, modelos lineares foram suficientes para descrever sistemas dinâmicos usados na indústria, tendo como benefícios sua simplicidade e facilidade de obtenção. No entanto, com o avanço tecnológico e a inserção de plantas complexas aos processos industriais, como manipuladores robóticos, e o aumento de sistemas mais restritos, como sistemas embarcados, contar com modelos lineares pode acarretar em desperdício e até mesmo impossibilitar o uso adequado, devido à perda de exatidão e à impossibilidade de reproduzir certos regimes dinâmicos (Nelles, 2001).

Uma resposta para o problema está no uso de modelos capazes de representar mais fielmente a não linearidade intrínseca de sistemas dinâmicos, reduzindo o nível de abstração. Os modelos que usam representações não lineares, como o modelo de blocos Hammerstein-Wiener, série de Volterra e os modelos baseados em equações a diferenças NARX (*Nonlinear Autoregressive Model with eXogenous inputs*) e NARMAX (*Nonlinear AutoRegressive Moving Average with eXogenous inputs*), por exemplo, são alternativas bastante utilizadas (Aguirre, 2015).

Os modelos baseados em NARX são divididos em dois tipos: racional e polinomial.

Enquanto o modelo racional destaca-se pela eficiência e exatidão na representação de sistemas não lineares, sua desvantagem reside na não linearidade dos parâmetros. Por outro lado, o modelo polinomial é linear nos parâmetros, o que o torna compatível com algoritmos de estimação de parâmetros baseados no método dos mínimos quadrados (Aguirre, 2015). Como os métodos de identificação de sistemas que foram utilizados neste trabalho são baseados no método de mínimos quadrados, trazem-se essas referências que serão posteriormente aprofundadas.

Dada a relevância do tema, neste trabalho é proposta a modelagem de três sistemas *benchmark* de tempo contínuo não lineares com base no modelo NARX polinomial, utilizando os métodos baseados na estimação de mínimos quadrados com seleção de estrutura, FROLS (*Forward Regression Orthogonal Least Squares*) e SEMP (*Simulation Error Minimization with Pruning*).

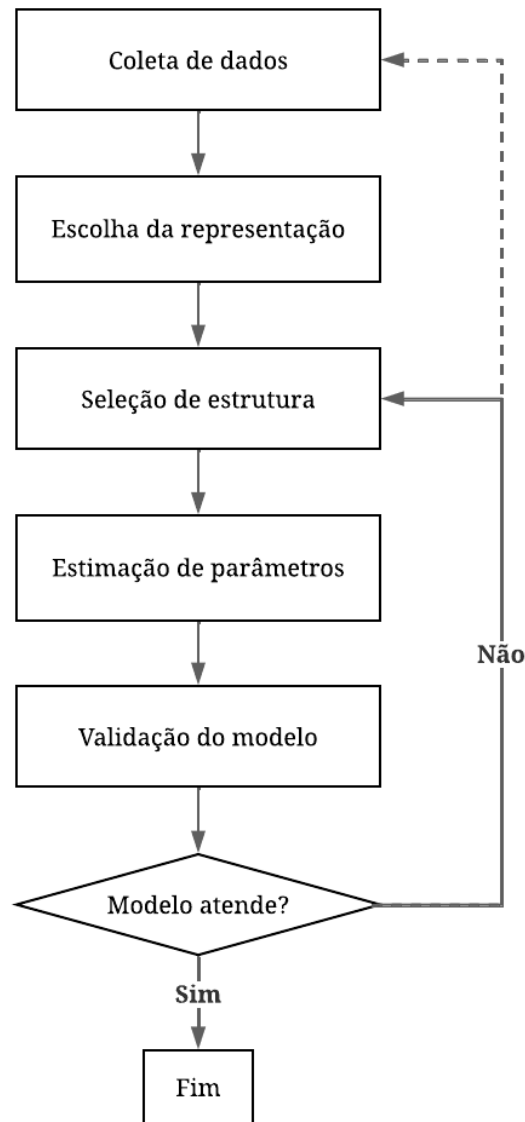
1.2 Trabalhos relacionados

O problema de modelagem e identificação é amplamente abordado na literatura científica, com trabalhos que exploram teoria e aplicação em várias áreas de conhecimento. Um dos que se destacam é Zadeh (1962), como um dos primeiros trabalhos a apresentar o estudo da teoria de sistemas, além de cunhar o termo “identificação de sistemas”, que Zadeh definiu como a determinação baseada na observação de dados de entrada e saída, de um sistema dentro de uma classe específica de sistemas, a qual o sistema em teste é equivalente (Zadeh, 1956; Ljung, 2008). Essa definição e terminologia estão fortemente associadas ao campo da teoria de sistemas, teoria de controle e automação; porém, técnicas de identificação de sistemas estão presentes nas mais diversas áreas, mesmo que sob outras nomenclaturas (Zadeh, 1956; Mehra, 1974).

Diversos trabalhos mostram a aplicação de identificação de sistemas em várias áreas, por exemplo, Gupta *et al.* (2018) e Shu *et al.* (2022) aplicam técnicas na área de controle, Low e Meghir (2017) e Mehra (1974) na área de econometria, Tsuchiya *et al.* (2017) para a área de biologia, Hirata *et al.* (2016) em medicina e Alaoui *et al.* (2012) na análise de performances artísticas.

Segundo Aguirre (2015), o processo de identificação de sistemas pode ser estruturado em cinco etapas fundamentais: testes dinâmicos e coleta de dados; escolha da representação matemática a ser usada; determinação da estrutura do modelo; estimação de parâmetros e validação do modelo. Nessa etapa, se o modelo for válido, ou seja, se atende aos requisitos, o processo pode ser encerrado. Caso contrário, deve-se continuar com a elaboração de um novo modelo, podendo-se voltar para a etapa de seleção de estrutura, ou até realizar uma nova coleta de dados, caso os dados disponíveis não sejam representativos do sistema, conforme ilustrado na Figura 1.1.

Figura 1.1 – Fluxograma do processo de identificação de sistemas.



Fonte: Adaptado de Aguirre (2015).

Dentre essas etapas, a estimação de parâmetros e a seleção da estrutura do modelo desempenham um papel importante, especialmente quando se trata de modelos não lineares, como os do tipo NARX. Nesse contexto, em Billings (2013) e Chen *et al.* (1988), é apresentado o método FROLS, que combina a estimação de parâmetros com a escolha da estrutura mais adequada do modelo, sendo particularmente útil na identificação de sistemas não lineares. De forma semelhante, o método SEMP, proposto por Piroddi e Spinelli (2003), foca em uma abordagem de simulação de modelos para a identificação de sistemas. O trabalho de Araújo *et al.* (2019) explora o uso dos métodos FROLS e SEMP para a identificação de modelos polinomiais NARX em um sistema não linear MIMO (*Multiple Input Multiple Output*), comparando-os por meio da métrica RMSE (*Root Mean Square Error*).

Ainda dentro do escopo de estimação de parâmetros e a seleção da estrutura, em Man-

gan *et al.* (2017), é proposto um método de detecção de modelo com identificação esparsa de dinâmicas não lineares para computação de critérios de informação AIC (*Akaike Information Criterion*) e BIC (*Bayesian Information Criterion*). Em Tavares *et al.* (2020), são identificados modelos NARX polinomiais de três sistemas utilizando ERR (Taxa de Redução de Erro), o critério AIC e o método de mínimos quadrados estendidos. Severino *et al.* (2016) exploram a estimação de parâmetros com métodos de otimização que envolvem algoritmo genético, otimização por enxame de partículas e algoritmo do morcego.

Nas últimas décadas, com o avanço da inteligência artificial, métodos de aprendizado de máquina têm ganhado espaço na identificação de sistemas. Alguns dos primeiros trabalhos são Narendra e Parthasarathy (1990) e Bittanti e Piroddi (1996), com o uso de redes neurais artificiais. O trabalho de Bijl *et al.* (2017) apresenta o algoritmo SONIG, que utiliza regressão de processos gaussianos esparsos para identificação de sistemas não lineares em tempo real, comparando o desempenho com modelos NARX. Em Bottegal *et al.* (2015), o problema de identificação é abordado com um método bayesiano baseado em *kernels* com métodos de cadeia de Markov Monte Carlo.

Trabalhos recentes envolvendo identificação de sistemas não lineares e seleção de estrutura para modelos polinomiais reforçam o interesse acadêmico no tema. Em Schär *et al.* (2024), por exemplo, é proposta uma nova representação F-NARX, que modela a resposta do sistema com base em características temporais dos sinais de entrada e saída. Já Gunawardena *et al.* (2024) apresenta a biblioteca de MATLAB NonSysId, que integra os métodos iOFR para regressão progressiva e PRESS para seleção de estrutura, criando um modelo robusto que não necessita de um *dataset* de validação separado.

1.3 Objetivos

1.3.1 Objetivo Geral

Comparar o desempenho dos métodos FROLS e SEMP para identificação de sistemas dinâmicos não lineares, considerando três diferentes cenários experimentais.

1.3.2 Objetivos Específicos

- Descrever os algoritmos FROLS e SEMP;
- Aplicar os algoritmos a um conjunto de dados de entrada e saída relativos a três sistemas não lineares de forma a identificar os modelos NARX que os descrevem;
- Avaliar e comparar a performance dos modelos identificados por ambos os métodos, com base na métrica RMSE.

1.4 Organização do Trabalho

Este trabalho encontra-se dividido nos seguintes capítulos: neste capítulo introdutório 1, foram apresentadas a motivação, alguns trabalhos relacionados à área com o intuito de destacar a relevância do tema abordado e os objetivos deste trabalho. No Capítulo 2 é apresentado um embasamento teórico sobre identificação de sistemas, representação de sistemas não lineares, estimação de parâmetros e a descrição dos métodos FROLS e SEMP. No Capítulo 3 é demonstrada a metodologia desenvolvida neste trabalho. No Capítulo 4 são expostos os resultados obtidos. Finalmente, no Capítulo 5 são apresentadas as discussões finais sobre o trabalho.

2 Fundamentação Teórica

A identificação de sistemas é um campo fundamental na engenharia e em diversas áreas das ciências aplicadas, cujo objetivo é construir modelos matemáticos que representem o comportamento de sistemas dinâmicos a partir de dados experimentais. Dentro desse contexto, a estimação de parâmetros e a seleção de estrutura desempenham um papel central.

Entre as abordagens modernas para esse problema, destaca-se o método FROLS, que permite a escolha incremental dos termos mais relevantes para o modelo. Além disso, técnicas como SEMP são empregadas para refinar a estrutura do modelo, reduzindo a redundância e melhorando sua capacidade preditiva.

Neste capítulo, serão explorados esses conceitos e métodos, proporcionando uma base teórica para a formulação de modelos de sistemas.

2.1 Representação de sistemas não lineares

Apesar de aproximações lineares serem muitas vezes suficientes para descrever um sistema dinâmico – que é não linear por natureza –, há vários casos em que um modelo não linear é mais adequado. Os modelos NARX (*Nonlinear Autoregressive Model with eXogenous inputs*) e NARMAX (*Nonlinear AutoRegressive Moving Average with eXogenous inputs*) são modelos não lineares baseados em equações a diferenças que descrevem como um sistema evolui no tempo discreto em função de entradas e saídas atrasadas, sendo o NARMAX a versão do NARX que introduz termos de ruído $e(k)$, o que permite uma estimação de parâmetros mais precisa por acomodar os efeitos de ruído de medição, erros de modelagem e/ou perturbações não medidas (Billings, 2013). A Equação (2.1) mostra a representação geral do modelo NARMAX:

$$\begin{aligned} y(k) = F^\ell & \left[y(k-1), y(k-2), \dots, y(k-n_y), \right. \\ & u(k-d), u(k-d-1), \dots, u(k-d-n_u), \\ & \left. e(k-1), e(k-2), \dots, e(k-n_e) \right] + e(k) \end{aligned} \quad (2.1)$$

Onde $F^\ell(\cdot)$ é uma função não linear qualquer, $u(k)$, $y(k)$ e $e(k)$ são as entradas, saídas e ruído aditivo do sistema, respectivamente, ℓ é o grau de não linearidade, d é o atraso e n_u , n_y e n_e são os atrasos máximos de entrada, saída e ruído (Billings, 2013).

Vários modelos podem ser vistos como casos especiais do NARMAX. Se não forem incluídos os termos de ruído, e_{k-n_e} , têm-se os modelos NARX. Outros modelos podem

ser descritos como variantes. Se $\ell = 1$, ou seja, para uma função $F^\ell(\cdot)$ linear, trata-se de modelos ARMAX. Se $\ell = 1$ e não forem considerados os termos de entrada e de ruído, transforma-se em um modelo AR, ARX se entradas forem incluídas. Se em vez de incluir-se entradas, incluir-se termos de ruído, obtêm-se o ARMA. Se $\ell > 1$ e não houver termos de entrada, tem-se o NARMA. Se não houver termos de entrada ou de ruído, tem-se NAR.

Um dos modelos NARMAX mais comumente utilizados é a aproximação polinomial de grau ℓ (2.2)

$$\begin{aligned} y(k) = & \theta_0 + \sum_{i_1=1}^n \theta_{i_1} x_{i_1}(k) + \sum_{i_1=1}^n \sum_{i_2=1}^n \theta_{i_1 i_2} x_{i_1}(k) x_{i_2}(k) + \cdots \\ & + \sum_{i_1=1}^n \cdots \sum_{i_\ell=i_{\ell-1}}^n \theta_{i_1 i_2 \dots i_\ell} x_{i_1}(k) x_{i_2}(k) \cdots x_{i_\ell}(k) + e(k) \end{aligned} \quad (2.2)$$

com os parâmetros do modelo $\theta_{i_1 i_2 \dots i_m}$ constantes que devem ser estimadas para ajustar a estrutura escolhida aos dados de identificação e os máximos atrasos n_y , n_u , n_e (Billings, 2013).

Os modelos polinomiais são vantajosos em relação a outras representações, como a representação racional, pois, apesar de não lineares, eles podem ser lineares nos parâmetros (Aguirre *et al.*, 1998). Assim, tais modelos podem ser representados como (2.3), onde $p_m(k)$ com $m = 1, 2, \dots, M$ são os regressores formados pela combinação das variáveis atrasadas de entrada e saída do sistema e M é o número de regressores do modelo.

$$y(k) = \sum_{m=1}^M \theta_m p_m(k) + e(k) \quad (2.3)$$

Devido ao fato de esses modelos funcionarem analogamente a uma representação linear, é possível utilizar o método de mínimos quadrados (MMQ) ou algoritmos que se baseiam no MMQ para a escolha dos parâmetros $\theta_{m_1 m_2 \dots m_M}$ (Aguirre *et al.*, 1998).

Pode-se mostrar que, para um modelo NARMAX (2.2), o número total de termos candidatos do modelo é

$$M = \binom{n + \ell}{\ell} = \frac{(n + \ell)!}{n! \ell!} = \frac{(n + \ell)(n + \ell - 1) \cdots (n + 1)}{\ell!}. \quad (2.4)$$

Para ℓ e $n = n_u + n_y + n_e$ grandes, o número de termos candidatos do modelo no modelo completo inicial pode ser muito grande (Billings, 2013). Porém, alguns poucos termos já são suficientes para se obter um modelo adequado, escolhidos através do processo de seleção de estrutura, que consiste em escolher um subconjunto apropriado de um conjunto de regressores candidatos, conforme será detalhado posteriormente.

2.2 Estimação de Parâmetros

Métodos baseados em mínimos quadrados são amplamente utilizados na estimação de parâmetros de modelos devido à sua capacidade de fornecer estimativas ótimas no sentido da minimização do erro quadrático (Verhaegen e Verdult, 2007; Gauss e Stewart, 1995). Para a construção de modelos mais robustos, o método de Gram-Schmidt e sua versão modificada são frequentemente empregados para a ortogonalização das variáveis de regressão. Essa abordagem é particularmente relevante na representação de sistemas não lineares, onde a seleção de variáveis e a redução da complexidade do modelo são desafios críticos (Aguirre, 2015).

2.2.1 Método de Mínimos Quadrados - MMQ

Para estimar os parâmetros θ_m na Equação (2.3), o método dos mínimos quadrados (MMQ) assume o vetor de parâmetros estimados $\hat{\boldsymbol{\theta}} = [\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m]^\top$. Deste modo, o vetor de saídas estimadas é escrito

$$\hat{\mathbf{y}} = \boldsymbol{\Psi} \hat{\boldsymbol{\theta}} \quad (2.5)$$

em que

$$\boldsymbol{\Psi} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_M] = \begin{bmatrix} p_1(1) & p_2(1) & \cdots & p_M(1) \\ p_1(2) & p_2(2) & \cdots & p_M(2) \\ \vdots & \vdots & \ddots & \vdots \\ p_1(N) & p_2(N) & \cdots & p_M(N) \end{bmatrix}. \quad (2.6)$$

Quando a matriz $\boldsymbol{\Psi}$ é invertível, ou seja, não singular, o vetor de parâmetros $\hat{\boldsymbol{\theta}}$ pode ser determinado por:

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\Psi}^{-1} \hat{\mathbf{y}} \quad (2.7)$$

No entanto, quando $N > M$, o sistema de equações é dito sobredeterminado (Aguirre, 2015). Nesse caso, tem-se que $\boldsymbol{\Psi} \in \mathbb{R}^{N \times M}$, $\hat{\mathbf{y}} \in \mathbb{R}^{N \times 1}$ e $\hat{\boldsymbol{\theta}} \in \mathbb{R}^{M \times 1}$. Como $\boldsymbol{\Psi}$ não é uma matriz quadrada, ela não possui inversa. Para contornar essa limitação, multiplica-se a Equação (2.5) por $\boldsymbol{\Psi}^\top$, obtendo:

$$\boldsymbol{\Psi}^\top \hat{\mathbf{y}} = \boldsymbol{\Psi}^\top \boldsymbol{\Psi} \hat{\boldsymbol{\theta}} \quad (2.8)$$

Com isso, a matriz $\boldsymbol{\Psi}^\top \boldsymbol{\Psi}$ torna-se quadrada, permitindo a obtenção de $\hat{\boldsymbol{\theta}}$ por meio da equação

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{\Psi}^\top \boldsymbol{\Psi})^{-1} \boldsymbol{\Psi}^\top \mathbf{y}. \quad (2.9)$$

O termo $(\Psi^\top \Psi)^{-1} \Psi^\top$ é conhecido como matriz pseudo-inversa, que representa uma solução para a determinação do vetor de parâmetros quando há um número maior de observações do que de incógnitas (Aguirre, 2015; Coelho e Coelho, 2004).

Por fim, assume-se que o valor observado $\mathbf{y}$ pode ser expresso como:

$$\mathbf{y} = \hat{\mathbf{y}} + \xi \quad (2.10)$$

em que ξ representa o resíduo obtido pela diferença entre a saída observada e a saída estimada.

A estimativa dos parâmetros de um sistema pelo método dos mínimos quadrados foi apresentada. Esse método busca minimizar o resíduo quadrático, o que dá origem ao seu nome. A função de custo utilizada para essa minimização é definida como

$$J_{MQ} = \sum_{i=1}^N (\xi)^2 = \boldsymbol{\xi}^\top \boldsymbol{\xi} = \|\boldsymbol{\xi}\|^2. \quad (2.11)$$

A minimização dessa função leva à mesma solução obtida ao se utilizar a pseudo-inversa para estimar os parâmetros. Dessa forma, o método dos mínimos quadrados pode ser interpretado como a busca da melhor aproximação, no sentido da norma quadrática, entre a resposta do modelo e os valores observados.

A estimação por mínimos quadrados pode enfrentar dificuldades em diversas situações, principalmente quando há forte correlação entre os regressores, o que acontece na identificação de sistemas não lineares devido ao grande número de colunas de Ψ . Nesse caso, essa matriz pode ser mal condicionada, o que amplifica erros numéricos e resulta em coeficientes instáveis (Aguirre *et al.*, 1998). Uma abordagem eficaz para mitigar esses problemas é a ortogonalização dos regressores. Técnicas como a ortogonalização de Gram-Schmidt transformam a matriz Ψ em um conjunto ortonormal de variáveis, garantindo que suas colunas sejam linearmente independentes. Isso melhora a estabilidade numérica do MMQ e reduz a sensibilidade a ruídos nos dados.

2.2.2 Mínimos Quadrados Ortogonais - MQO

Nesta seção, são apresentados os métodos de Gram-Schmidt e Gram-Schmidt modificado para a ortogonalização dos regressores na estimação por mínimos quadrados ortogonais (MQO).

Método de Gram-Schmidt

O objetivo do método de Gram-Schmidt é fatorar a matriz de regressores Ψ na forma:

$$\Psi = Q\mathbf{A} \quad (2.12)$$

onde $\mathbf{A}$ é uma matriz triangular superior com diagonal unitária, dada por:

$$\mathbf{A} = \begin{bmatrix} 1 & \alpha_{1,2} & \alpha_{1,3} & \cdots & \alpha_{1,M} \\ 0 & 1 & \alpha_{2,3} & \cdots & \alpha_{2,M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & \alpha_{M-1,M} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.13)$$

e $\mathbf{Q}$ é uma matriz de dimensão $N \times M$ cujas colunas são ortogonais, ou seja:

$$\mathbf{Q} = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \cdots \quad \mathbf{q}_M] \quad (2.14)$$

satisfazendo a relação:

$$\mathbf{Q}^\top \mathbf{Q} = \mathbf{D} \quad (2.15)$$

onde $\mathbf{D}$ é uma matriz diagonal definida positiva. A matriz $\mathbf{A}$

O método Gram-Schmidt ortogonaliza as colunas de $\mathbf{\Psi}$ iterativamente. Na i -ésima iteração, a coluna $\mathbf{p}_i$ deve ser tornada ortogonal em relação às $i - 1$ colunas previamente calculadas. Esse procedimento é definido pelas equações:

$$\mathbf{q}_1 = \mathbf{p}_1 \quad (2.16)$$

$$\alpha_{j,i} = \frac{\mathbf{q}_j^\top \mathbf{p}_i}{\mathbf{q}_j^\top \mathbf{q}_j}, \quad 1 \leq j < i \quad (2.17)$$

$$\mathbf{q}_i = \mathbf{p}_i - \sum_{j=1}^{i-1} \alpha_{j,i} \mathbf{q}_j. \quad (2.18)$$

A partir da equação normal:

$$\mathbf{\Psi}^\top \mathbf{\Psi} \hat{\boldsymbol{\theta}} = \mathbf{\Psi}^\top \mathbf{y} \quad (2.19)$$

e utilizando a decomposição $\mathbf{\Psi} = \mathbf{Q}\mathbf{A}$, obtém-se:

$$\mathbf{A}^\top \mathbf{Q}^\top \mathbf{Q} \mathbf{A} \hat{\boldsymbol{\theta}} = \mathbf{A}^\top \mathbf{Q}^\top \mathbf{y} \quad (2.20)$$

$$\mathbf{A}^\top \mathbf{D} \mathbf{A} \hat{\boldsymbol{\theta}} = \mathbf{A}^\top \mathbf{Q}^\top \mathbf{y} \quad (2.21)$$

$$\mathbf{D} \mathbf{A} \hat{\boldsymbol{\theta}} = \mathbf{Q}^\top \mathbf{y} \quad (2.22)$$

$$\mathbf{A}\hat{\boldsymbol{\theta}} = \mathbf{D}^{-1}\mathbf{Q}^\top \mathbf{y}. \quad (2.23)$$

Os parâmetros do modelo no espaço ortogonal são fornecidos por $\hat{\mathbf{g}} = \mathbf{A}\hat{\boldsymbol{\theta}}$. Substituindo também $\mathbf{D} = \mathbf{Q}^\top \mathbf{Q}$:

$$\hat{\mathbf{g}} = (\mathbf{Q}^\top \mathbf{Q})^{-1} \mathbf{Q}^\top \mathbf{y}. \quad (2.24)$$

Os parâmetros no modelo ortogonalizado são, então, obtidos como:

$$\hat{g}_i = \frac{\mathbf{y}^\top \mathbf{q}_i}{\mathbf{q}_i^\top \mathbf{q}_i}, \quad i = 1, 2, \dots, M. \quad (2.25)$$

Finalmente, os parâmetros do modelo original são recuperados por:

$$\hat{\boldsymbol{\theta}} = \mathbf{A}^{-1} \hat{\mathbf{g}}. \quad (2.26)$$

Método de Gram-Schmidt Modificado

O método de Gram-Schmidt modificado (MGS) busca reduzir a sensibilidade numérica e o mau condicionamento frequentemente observados no método clássico.

Na $i - 1$ iteração, indicada entre parênteses, a coluna $\mathbf{p}_i$ é utilizada para gerar a coluna ortogonal $\mathbf{q}_i$ conforme:

$$\mathbf{q}_i = \mathbf{p}_i^{(i-1)} \quad (2.27)$$

$$\alpha_{i,j} = \frac{\mathbf{q}_i^\top \mathbf{p}_j^{(i-1)}}{\mathbf{q}_i^\top \mathbf{q}_i}, \quad j = i + 1, \dots, M \quad (2.28)$$

$$\mathbf{p}_j^{(i)} = \mathbf{p}_j^{(i-1)} - \alpha_{i,j} \mathbf{q}_i, \quad j = i + 1, \dots, M. \quad (2.29)$$

Após a ortogonalização, os parâmetros do modelo ortogonalizado são estimados como:

$$\hat{g}_i = \frac{\mathbf{q}_i^\top \mathbf{y}^{(i-1)}}{\mathbf{q}_i^\top \mathbf{q}_i} \quad (2.30)$$

com a atualização:

$$\mathbf{y}^{(i)} = \mathbf{y}^{(i-1)} - \hat{g}_i \mathbf{q}_i. \quad (2.31)$$

Embora o MGS seja mais estável numericamente, sua aplicação em modelos com grande número de regressores pode ser inviabilizada pelo crescimento da complexidade computacional. Como alternativa, estratégias de detecção de estrutura podem ser adotadas para reduzir a dimensionalidade do problema.

2.3 Método *Forward Regression with Orthogonal Least Squares* - FROLS

O algoritmo FROLS (do inglês, *Forward Regression with Orthogonal Least Squares*), também conhecido como algoritmo de regressão ortogonal avançada (OFR, do inglês, *Orthogonal Forward Regression*), seleciona e classifica esses termos significativos e se tornou um dos algoritmos padrão mais utilizados para a detecção da estrutura do modelo não linear (Chen *et al.*, 1988; Billings, 2013).

Baseado na estrutura do método de mínimos quadrados ortogonais (MQO), o método FROLS propõe a inserção iterativa dos termos do modelo. A cada iteração do algoritmo, busca-se o regressor que minimiza a Taxa de Redução de Erro (ERR, do inglês *Error Reduction Ratio*), uma taxa que indica a melhoria obtida na representação do sistema através da inclusão do termo no modelo. De maneira similar ao MQO, uma técnica de ortogonalização da matriz de regressores análoga ao método de Gram-Schmidt modificado é utilizada.

Inicialmente, considera-se o modelo linear nos parâmetros (2.3) com $n_e = 0$, ou seja, um modelo NARX. Seja $\mathbf{y} = [y(1), \dots, y(N)]^\top$ um vetor de saídas medidas em N instantes de tempo, e $\mathbf{p}_m = [p_m(1), p_m(2), \dots, p_m(N)]^\top$ um vetor formado pelo m -ésimo termo candidato do modelo, onde $m = 1, 2, \dots, M$. Seja $\Psi = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_M\}$ um dicionário composto pelas M bases candidatas. Com muita frequência, o conjunto Ψ de dimensão finita será redundante. O problema de seleção de termos do modelo é equivalente a encontrar um subconjunto de dimensão total $\Psi_{M_0} = \{\psi_1, \psi_2, \dots, \psi_{M_0}\} = \{\mathbf{p}_{i_1}, \mathbf{p}_{i_2}, \dots, \mathbf{p}_{i_{M_0}}\}$ de M_0 ($M_0 \leq M$) bases, a partir do dicionário Ψ , onde $\psi_m = \mathbf{p}_{i_m}$, $i_k \in \{1, 2, \dots, M\}$ e $k = 1, 2, \dots, M_0$, de modo que $\mathbf{y}$ possa ser satisfatoriamente aproximado usando uma combinação linear de $\psi_1, \psi_2, \dots, \psi_{M_0}$ como mostrado abaixo:

$$\mathbf{y} = \theta_1 \psi_1 + \dots + \theta_{M_0} \psi_{M_0} + \mathbf{e} \quad (2.32)$$

ou, em forma compacta matricial,

$$\mathbf{y} = \Psi_{M_0} \boldsymbol{\theta} + \mathbf{e} \quad (2.33)$$

onde a matriz $\Psi_{M_0} = [\psi_1, \dots, \psi_{M_0}]$ é assumida como tendo posto completo de colunas, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_{M_0}]^\top$ é um vetor de parâmetros, e $\mathbf{e}$ é o erro de aproximação.

A primeira etapa começa com o modelo completo inicial (2.3) e o dicionário completo inicial $\Psi = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_M\}$. Para $m = 1, 2, \dots, M$, seja $\mathbf{q}_m = \mathbf{p}_m$ e $\sigma = \mathbf{y}^\top \mathbf{y}$, calcula-se:

$$g_m^{(1)} = \frac{\mathbf{y}^\top \mathbf{q}_m}{\mathbf{q}_m^\top \mathbf{q}_m} \quad (2.34)$$

$$\text{ERR}^{(1)}[m] = \frac{\left(g_m^{(1)}\right)^2 (\mathbf{q}_m^\top \mathbf{q}_m)}{\sigma}. \quad (2.35)$$

Seja

$$\ell_1 = \arg \max_{1 \leq m \leq M} \left\{ \text{ERR}^{(1)}[m] \right\}, \quad (2.36)$$

o que implica

$$\text{ERR}[\ell_1] = \max \left\{ \text{ERR}^{(1)}[m] : 1 \leq m \leq M \right\}. \quad (2.37)$$

A primeira base significativa pode então ser selecionada como $\boldsymbol{\psi}_1 = \mathbf{p}_{\ell_1}$ e o primeiro vetor ortogonal associado pode ser escolhido como $\mathbf{q}_1 = \mathbf{p}_{\ell_1}$.

Assuma que um subconjunto $\boldsymbol{\Psi}_{s-1}$, consistindo em $(s-1)$ bases significativas $\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \dots, \boldsymbol{\psi}_{s-1}$, foi determinado na etapa $(s-1)$, e que as $(s-1)$ bases selecionadas foram transformadas em um novo grupo de bases ortogonalizadas $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{s-1}$. Ao passo de número s , seja $m \neq \ell_1, m \neq \ell_2, \dots, m \neq \ell_{s-1}$. Para $m = 1, 2, \dots, M$, calcule

$$\mathbf{q}_m^{(s)} = \mathbf{p}_m - \sum_{r=1}^{s-1} \frac{\mathbf{p}_m^\top \mathbf{q}_r}{\mathbf{q}_r^\top \mathbf{q}_r} \mathbf{q}_r, \mathbf{p}_m \in \boldsymbol{\Psi} - \boldsymbol{\Psi}_{s-1} \quad (2.38)$$

$$g_m^{(s)} = \frac{\mathbf{y}^\top \mathbf{q}_m^{(s)}}{(\mathbf{q}_m^{(s)})^\top \mathbf{q}_m^{(s)}} \quad (2.39)$$

$$\text{ERR}^{(s)}[m] = \frac{(g_m^{(s)})^2 \left[(\mathbf{q}_m^{(s)})^\top \mathbf{q}_m^{(s)} \right]}{\sigma} \quad (2.40)$$

Seja

$$\ell_s = \arg \max_{1 \leq m \leq M} \left\{ \text{ERR}^{(s)}[m] \right\} \quad (2.41)$$

a s -ésima base mais significativa pode então ser escolhida como $\boldsymbol{\psi}_s = \mathbf{p}_{\ell_s}$ e a s -ésima base ortogonal pode ser feita $\mathbf{q}_s = \mathbf{q}_{\ell_s}^{(s)}$.

As bases mais significativas subsequentes podem ser selecionadas da mesma maneira, passo a passo. Em cada etapa, a base com a maior capacidade de representar o vetor de saída $\mathbf{y}$ é selecionada. O procedimento de seleção pode ser encerrado quando o valor de ESR, definido na Equação 2.42, for menor que um limite pré-definido ρ :

$$\text{ESR} = 1 - \sum_{s=1}^{M_0} \text{ERR}(s) \leq \rho. \quad (2.42)$$

O modelo final é a combinação linear dos M_0 termos significativos selecionados dos M termos candidatos $\{\mathbf{p}_i\}_{i=1}^M$:

$$y(k) = \sum_{i=1}^{M_0} g_i \mathbf{q}_i(k) + e(k) \quad (2.43)$$

que é equivalente a

$$y(k) = \sum_{m=1}^{M_0} \beta_{\ell_m} \mathbf{p}_{\ell_m}(k) + e(k) \quad (2.44)$$

onde o vetor de parâmetros $\boldsymbol{\beta} = [\beta_{\ell_1}, \beta_{\ell_2}, \dots, \beta_{\ell_{M_0}}]^\top$ é calculado a partir da equação triangular $\mathbf{A}\boldsymbol{\beta} = \mathbf{g}$, com $\mathbf{g} = [g_1, g_2, \dots, g_{M_0}]^\top$ e

$$\mathbf{A} = \begin{bmatrix} 1 & \alpha_{1,2} & \alpha_{1,3} & \cdots & \alpha_{1,M_0} \\ 0 & 1 & \alpha_{2,3} & \cdots & \alpha_{2,M_0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & \alpha_{M_0-1,M_0} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.45)$$

As entradas α_{ij} tq. $1 \leq i < j \leq M_0$ na matriz $\mathbf{A}$ são dadas no algoritmo FROLS apresentado no Capítulo 3. Definindo $\theta_m = \beta_{i_m}$ e $\boldsymbol{\psi}_m = \mathbf{p}_{i_m}$, com $m = 1, 2, \dots, M_0$ e $i_m \in \{1, 2, \dots, M\}$, a Equação (2.44) se tornará a Equação (2.32), estando estimados, então, os regressores e os seus respectivos parâmetros.

2.4 Método *Simulation Error Minimization with Pruning* - SEMP

O método SEMP (do inglês, *Simulation Error Minimization with Pruning*) é um algoritmo para a seleção de modelos, particularmente eficaz para a seleção de modelos NARX polinomiais. Ele foi proposto em Piroddi e Spinelli (2003) como uma melhoria para a classe de métodos *forward-regression orthogonal estimator*, ao substituir o ERR pela métrica SRR (*Simulation Error Reduction Ratio*) e criar um sistema de poda de regressores excessivos, que potencializa a seleção de estrutura e torna o modelo mais enxuto. O SEMP opera minimizando o erro de simulação e adicionando um regressor por vez ao modelo. Durante esse processo, ele também realiza uma regularização que equilibra a precisão do modelo com seu tamanho na estimação de parâmetros.

2.4.1 Simulation Error Reduction Ratio - SRR

A Taxa de Redução do Erro de Simulação (do inglês, *Simulation Error Reduction Ratio*, SRR) foi pensada como um critério para atender a uma série de exigências que buscam aperfeiçoar a seleção correta de modelos. As propriedades propostas são:

1. O termo com o maior valor do critério a cada iteração deve ser um dos regressores efetivamente presentes no sistema;
2. Regressores que não são necessários para explicar a dinâmica do sistema devem apresentar um valor significativamente mais baixo do critério em comparação com os regressores corretos;
3. Os termos correspondentes aos regressores corretos devem ter valores comparáveis do critério.

Um critério que atende melhor a essas exigências em comparação com o ERR é o SRR, definida como a redução do Erro Médio Quadrático de Simulação (do inglês, *Mean Squared Simulation Error*, MSSE) obtida pela inclusão de um regressor no modelo, normalizada pela variância da saída:

$$[SRR]_j = \frac{MSSE(\mathcal{M}_i) - MSSE(\mathcal{M}_{i+1})}{\frac{1}{N} \sum_{t=1}^N y^2(t)} \quad (2.46)$$

onde $\mathcal{M}_i$ é o modelo obtido até a i -ésima iteração e $\mathcal{M}_{i+1}$ é o modelo candidato na iteração subsequente, com a inclusão do j -ésimo termo.

2.4.2 Procedimento de poda

Durante o processo de identificação de um sistema, a relevância dos regressores incluídos pode mudar; regressores inicialmente avaliados como essenciais podem, posteriormente, ser compensados por outros regressores mais significativos. Portanto, para melhorar o processo de seleção de regressores, aqueles que não contribuem mais tão significativamente para a qualidade do modelo devem ser eliminados. O procedimento sistemático de eliminação proposto é chamado de “poda”.

Em princípio, o regressor menos significativo em um modelo é aquele cuja eliminação aumenta minimamente o índice de desempenho do modelo. Na abordagem proposta, a seleção e eliminação de regressores são realizadas iterativamente durante a execução do algoritmo de identificação. A iteração começa com a inclusão de um novo regressor no modelo atual. Então, o seguinte procedimento iterativo de poda é aplicado:

1. Para cada regressor selecionado, avalia-se o submodelo resultante após sua exclusão, estimam-se novamente os parâmetros e calcula-se o respectivo índice de desempenho;
2. O regressor menos significativo é identificado como aquele associado ao submodelo com o melhor valor do índice de desempenho;
3. Se o mesmo submodelo for também melhor do que o modelo obtido na iteração anterior, o regressor é efetivamente eliminado;

4. Se a eliminação de um regressor ocorreu, uma nova verificação de regressores elimináveis é feita, caso contrário, a iteração termina.

Ao final de uma iteração completa do algoritmo de identificação, ou um novo regressor é adicionado ao modelo atual ou substitui um ou mais termos, desde que essa troca melhore a precisão do modelo. A lógica por trás do algoritmo é que, dados dois modelos com o mesmo índice de desempenho, o menor é o preferido. Nota-se que o mecanismo de poda pode, ocasionalmente, eliminar um regressor correto, mas este será normalmente reintroduzido em uma etapa posterior.

Em síntese, ao comparar os métodos FROLS e SEMP, descritos nesse capítulo, observa-se uma distinção fundamental na forma como avaliam a relevância dos regressores. O FROLS estima essa importância com base em critérios analíticos aplicados ao conjunto de regressores candidatos, selecionando os mais representativos para compor o modelo. Por outro lado, o SEMP adota uma abordagem baseada em simulação, avaliando o impacto de cada regressor a partir do desempenho do modelo já construído e progressivamente ajustado através de comparações constantes entre saída real e simulada. Sendo assim, entre os dois métodos, o SEMP tende a ter maior custo computacional, enquanto o FROLS é mais leve, porém menos preciso sob certas condições (Araújo *et al.*, 2019).

3 Metodologia

Como visto anteriormente, de maneira resumida, os passos para o processo de identificação de um sistema são:

- execução de experimentos para coleta de dados de entrada e saída
- escolha da representação matemática a ser utilizada
- seleção de estrutura do modelo
- estimação de parâmetros
- validação do modelo

Deste modo, a metodologia utilizada neste trabalho segue a mesma estrutura. Este capítulo tem início com a exposição dos sistemas abordados, discutindo a coleta dos dados, suas características físicas e aplicações. Em seguida, são expostos os processos de identificação dos modelos NARX com os algoritmos FROLS e SEMP, que combinam as etapas de seleção de estrutura e estimação de parâmetros. Por fim, é discutida a etapa de validação dos modelos.

3.1 Descrição dos sistemas

Para o teste dos algoritmos implementados, foram analisados como estudo de caso três sistemas SISO, conforme descritos a seguir.

Os valores de hiperparâmetros dos modelos – atraso máximo de entrada n_u , atraso máximo de saída n_y , grau de não linearidade ℓ – e o limite ρ para critério de parada foram definidos de forma empírica após uma série de experimentos. Diversas combinações desses parâmetros foram testadas, sendo os modelos gerados avaliados com base na métrica RMSE e na quantidade de termos selecionados, a fim de priorizar modelos precisos e simples.

3.1.1 Sistema Simulado

O primeiro sistema a ser considerado é um sistema numérico proposto em Billings (2013) descrito pela equação a diferenças 3.1:

$$y(k) = -0,605y(k-1) - 0,163y^2(k-2) + 0,588u(k-1) - 0,240u(k-2) + \xi(k). \quad (3.1)$$

onde $\xi(k)$ é uma sequência de ruído branco gaussiano com média zero e desvio padrão $\sigma_\xi = 0,1$. Esse sistema foi escolhido por já ser conhecido na literatura como compatível com os algoritmos utilizados neste trabalho, tendo apresentado bons resultados tanto em Billings (2013), quanto em outros trabalhos como Araújo (2019).

Como entrada do sistema para excitar o sistema, foi considerado um sinal senoidal descrito em 3.2. Já o sinal usado na validação é descrito pela Equação 3.3. A partir disso, foram coletadas $N = 1000$ amostras de entrada e saída.

$$u_{treino}(t) = 0,5\text{sen}(0,7t) + 0,25\text{sen}(1,4t) + 0,5\text{sen}(0,35t) \quad (3.2)$$

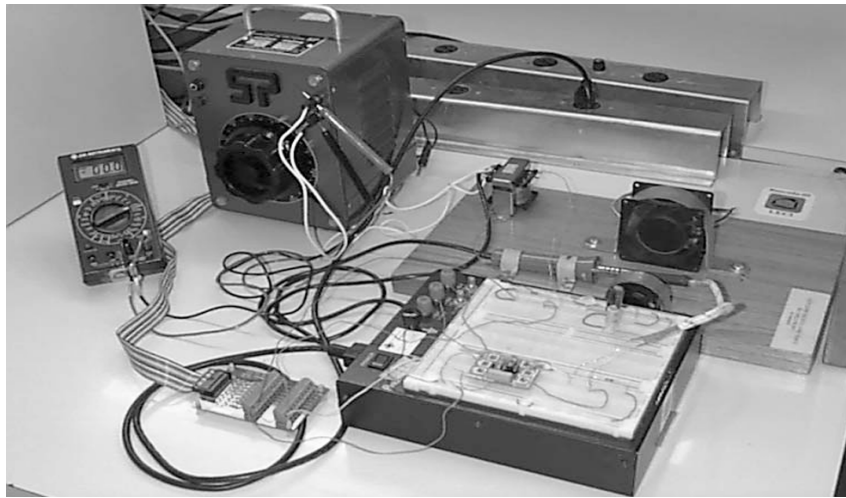
$$u_{teste}(t) = 0,25\text{sen}(0,7t) + 0,1\text{sen}(1,4t) + 0,5\text{sen}(0,35t) + 0,2\text{sen}(t) + 0,4\text{sen}(0,5t) \quad (3.3)$$

Para a construção dos modelos NARX, foram considerados os valores de grau de não linearidade $\ell = 3$ e os atrasos máximos de entrada e saída $n_u = 2$ e $n_y = 2$. No algoritmo FROLS, o valor do critério de parada foi $\rho = 0,5$, no SEMP, $\rho = 0,0001$.

3.1.2 Sistema de Ferro de Solda

O sistema a seguir será utilizado para demonstrar a aplicação da metodologia em um sistema real. Ele é composto por um ferro de solda e um termopar responsável por medir a diferença de potencial elétrico resultante da variação de temperatura. Esse sistema foi inicialmente descrito e identificado em Cassini (1999) e é mostrado na Figura 3.1.

Figura 3.1 – Fotografia do sistema de ferro de solda.



Fonte: Cassini (1999).

A entrada do sistema é a tensão elétrica em um divisor de tensão, coletada por uma placa de aquisição de dados (PCL 711s). Com a variação de tensão, tem-se a variação de temperatura no ferro de solda, que é medida com um termopar. A saída do sistema

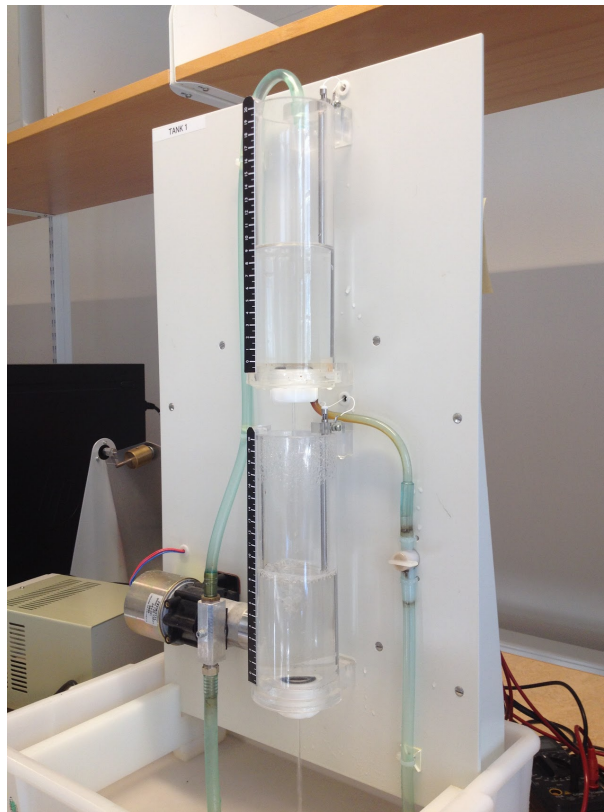
corresponde, então, à tensão de saída do circuito amplificador, conectado ao termopar. A taxa de amostragem utilizada foi de 12 segundos, com $N = 1254$ amostras.

Como hiperparâmetros para a modelagem do sistema, foram utilizados $n_u = 2$ de entrada, $n_y = 2$ de saída e grau de não linearidade $\ell = 2$. Para o critério de parada, foi considerado $\rho = 0,0001$ para o método FROLS e $\rho = 0,00001$ para o método SEMP.

3.1.3 Sistema de Tanques em Cascata

O sistema de tanques em cascata é um sistema de controle de nível de fluxo que consiste em dois tanques alimentados por uma bomba (Figura 3.2). Esse sistema *benchmark* foi descrito em Schoukens e Noël (2017).

Figura 3.2 – Fotografia do sistema de tanques em cascata.



Fonte: Schoukens *et al.* (2016).

O sinal de entrada controla a bomba de água que leva a água do reservatório até o tanque de água superior. A água do tanque superior passa por uma pequena abertura para o tanque inferior e, por fim, vai para o reservatório por uma pequena abertura no tanque inferior. A variável de entrada é a tensão da bomba e a saída é o nível de água no tanque inferior.

Esse sistema apresenta não linearidades devido ao fenômeno de transbordamento, que acontece quando a amplitude do sinal de entrada é muito grande, o que faz com que o tanque de cima transborde e parte da água flua para o tanque de baixo e parte para

o reservatório (Schoukens *et al.*, 2016). Este *dataset* já veio dividido em treinamento e validação. Os sinais de entrada de treino e validação são sinais multisenso com $N = 1024$ amostras e o período de amostragem é igual a 4 segundos.

Depois de uma série de testes, os hiperparâmetros escolhidos para a modelagem desse sistema foram $n_u = 3$ e $n_y = 5$ para o atraso máximo de entrada dos métodos FROLS e SEMP, respectivamente; $n_y = 1$ para o atraso de entrada e grau de não linearidade $\ell = 3$. Para o critério de parada, foi considerado $\rho = 0,00001$ para o método FROLS e $\rho = 0,000001$ para o método SEMP.

3.2 Seleção de estrutura e estimação de parâmetros

Para o processo de identificação dos sistemas descritos anteriormente, foi utilizada a linguagem de programação *Python*. Primeiramente, cada conjunto de dados foi dividido em subconjuntos de treinamento e validação, adotando-se, em geral, a proporção de 75% para treinamento e 25% para validação. A única exceção foi o sistema de tanques, cujos dados já estavam previamente divididos na proporção de 50% e 50%.

Em seguida, a matriz de regressores candidatos Ψ foi construída com base nos dados de treinamento e nos hiperparâmetros n_u , n_y e ℓ , conforme:

$$\Psi = \begin{bmatrix} y(0) & \cdots & y(1-n_y) & u(0) & \cdots & u(1-n_u) & y(0)^2 & \cdots & y(0)u(0) & \cdots & u(1-n_u)^\ell \\ y(1) & \cdots & y(2-n_y) & u(1) & \cdots & u(2-n_u) & y(1)^2 & \cdots & y(1)u(1) & \cdots & u(2-n_u)^\ell \\ y(2) & \cdots & y(3-n_y) & u(2) & \cdots & u(3-n_u) & y(2)^2 & \cdots & y(2)u(2) & \cdots & u(3-n_u)^\ell \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ y(N-1) & \cdots & y(N-n_y) & u(N-1) & \cdots & u(N-n_u) & y(N-1)^2 & \cdots & y(N-1)u(N-1) & \cdots & u(N-n_u)^\ell \end{bmatrix}$$

A partir dessa matriz, foram aplicados separadamente os algoritmos FROLS e SEMP — responsáveis pela seleção da estrutura do modelo e pela estimação de seus parâmetros — com o objetivo de comparar seus desempenhos.

O Algoritmo 1 apresenta o primeiro passo do algoritmo FROLS, que envolve a inicialização e seleção do primeiro termo significativo, conforme apresentado na seção 2.3. Os demais passos, incluindo a seleção iterativa dos termos significativos e o cálculo dos parâmetros do modelo, estão descritos no Algoritmo 2 (Billings, 2013).

Algoritmo 1 Inicialização FROLS

- 1: **Passo 1:** Inicialize o dicionário completo $\Psi = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_M\}$ e $\sigma = \mathbf{y}^\top \mathbf{y}$
 - 2: **for** $m = 1, 2, \dots, M$ **do**
 - 3: Defina $\mathbf{q}_m = \mathbf{p}_m$
 - 4: Calcule $g_m^{(1)} = \frac{\mathbf{y}^\top \mathbf{q}_m}{\mathbf{q}_m^\top \mathbf{q}_m}$
 - 5: Calcule $\text{ERR}^{(1)}[m] = \frac{(g_m^{(1)})^2 (\mathbf{q}_m^\top \mathbf{q}_m)}{\sigma}$
 - 6: **end for**
 - 7: Encontre $\ell_1 = \arg \max_{1 \leq m \leq M} \left\{ \text{ERR}^{(1)}[m] \right\}$
 - 8: Defina $a_{11} = 1$
 - 9: Defina $\mathbf{q}_1 = \mathbf{p}_{\ell_1}$
 - 10: Defina $g_1 = g_{\ell_1}^{(1)}$
 - 11: Defina $\text{err}[1] = \text{ERR}^{(1)}[\ell_1]$
-

Algoritmo 2 FROLS para $s \geq 2$

- 1: **Passo s** ($s \geq 2$): Sejam $m \neq \ell_1, m \neq \ell_2, \dots, m \neq \ell_{s-1}$
 - 2: **for** $s \geq 2$ **do**
 - 3: **for** $m = 1, 2, \dots, M$ **do**
 - 4: Calcule $\mathbf{q}_m^{(s)} = \mathbf{p}_m - \sum_{r=1}^{s-1} \frac{\mathbf{p}_m^\top \mathbf{q}_r}{\mathbf{q}_r^\top \mathbf{q}_r} \mathbf{q}_r$
 - 5: Calcule $g_m^{(s)} = \frac{\mathbf{y}^\top \mathbf{q}_m^{(s)}}{(\mathbf{q}_m^{(s)})^\top \mathbf{q}_m^{(s)}}$
 - 6: Calcule $\text{ERR}^{(s)}[m] = (g_m^{(s)})^2 \frac{[(\mathbf{q}_m^{(s)})^\top \mathbf{q}_m^{(s)}]}{\sigma}$
 - 7: **end for**
 - 8: Encontre $\ell_s = \arg \max_{1 \leq m \leq M} \left\{ \text{ERR}^{(s)}[m] \right\}$
 - 9: Defina $\mathbf{q}_s = \mathbf{q}_{\ell_s}^{(s)}$
 - 10: Defina $g_s = g_{\ell_s}^{(s)}$
 - 11: **for** $r = 1, 2, \dots, s-1$ **do**
 - 12: Calcule $a_{r,s} = \frac{\mathbf{q}_r^\top \mathbf{p}_{\ell_s}}{\mathbf{q}_r^\top \mathbf{q}_r}$
 - 13: **end for**
 - 14: Defina $a_{ss} = 1$
 - 15: Defina $\text{err}[s] = \text{ERR}^{(s)}[\ell_s]$
 - 16: A busca é finalizada no passo M_0 quando o ESR for menor que um limiar predefinido:
 - 17: $ESR = 1 - \sum_{s=1}^{M_0} \text{err}(s) \leq \rho$
 - 18: $\boldsymbol{\theta} = \mathbf{g} \mathbf{A}^{-1}$
 - 19: **end for**
 - 20: **return** $\ell, \boldsymbol{\theta}$
-

O método SEMP é resumido no Algoritmo 3, a seguir.

Algoritmo 3 SEMP

```

1: Inicialize  $V_y = \text{variância de } \mathbf{y}$ 
2: Defina  $\mathbf{P} = []$ 
3: Defina  $\mathbf{Q} = \Psi$ 
4: for  $i = 1, 2, \dots, M$  do
5:   for  $j = 1, 2, \dots, \text{cols}(\mathbf{Q})$  do
6:     Concatene  $\mathbf{P}_j = [\mathbf{P} \mid \mathbf{q}_j]$ 
7:     Calcule  $\boldsymbol{\theta}_j = (\mathbf{P}_j^\top \mathbf{P}_j)^{-1} (\mathbf{P}_j^\top \mathbf{y})$ 
8:     Calcule  $\hat{\mathbf{y}}_j = \mathbf{P}_j \boldsymbol{\theta}_j$ 
9:     Calcule  $J_j = \text{MSE}(\mathbf{y}, \hat{\mathbf{y}}_j)$ 
10:    Calcule  $\text{SRR}[j] = \frac{J_{\text{ant}} - J_j}{V_y}$ 
11:  end for
12:  Encontre  $\ell = \underset{1 \leq j \leq \text{size}(\text{SRR})}{\arg \max} \{ \text{SRR}[j] \}$ 
13:  if  $J_\ell < J_{\text{ant}}$  and  $|J_{\text{ant}} - J_\ell| > \rho$  then
14:    Adicione  $\mathbf{Q}_\ell$  a  $\mathbf{P}$ 
15:    Remova  $\mathbf{Q}_\ell$  de  $\mathbf{Q}$ 
16:    Atualize  $\boldsymbol{\theta} = \boldsymbol{\theta}_\ell$ ,  $J_{\text{ant}} = J_\ell$ 
17:  else
18:    break
19:  end if
20:  if  $\text{cols}(\mathbf{P}) > 1$  then
21:    for  $k = 1, 2, \dots, \text{cols}(\mathbf{P})$  do
22:       $\mathbf{R} = \mathbf{P}$  sem a  $k$ -ésima coluna
23:       $\boldsymbol{\theta}_k = (\mathbf{R}^\top \mathbf{R})^{-1} \mathbf{R}^\top \mathbf{y}$ 
24:       $\hat{\mathbf{y}}_k = \mathbf{R} \boldsymbol{\theta}_k$ 
25:       $J_k = \text{MSE}(\mathbf{y}, \hat{\mathbf{y}}_k)$ 
26:      if  $J_k < J_{\text{ant}}$  then
27:        Remova  $\mathbf{P}_k$  de  $\mathbf{P}$ 
28:        Remova  $\ell_k$  de  $\ell$ 
29:        Atualize  $\boldsymbol{\theta} = \boldsymbol{\theta}_k$ ,  $J_{\text{ant}} = J_k$ 
30:      break
31:    end if
32:  end for
33: end if
34: end for
35: return  $\ell, \boldsymbol{\theta}$ 

```

Como é possível observar nos pseudocódigos acima, ambos os algoritmos foram imple-

mentados para retornar as mesmas variáveis: ℓ , que representa a ordem em que os M_0 regressores da matriz de candidatos Ψ devem ser incluídos no modelo final e θ , que são os parâmetros calculados para cada regressor. A implementação dos dois algoritmos pode ser encontrada no repositório do GitHub¹.

3.3 Validação do modelo

A validação é a etapa final na identificação de sistemas, e o processo todo pode ser realizado repetidamente até que se obtenha um modelo matemático satisfatório segundo os critérios de validação. Como nenhum modelo representa perfeitamente a realidade, considera-se adequado aquele que satisfaz os requisitos do uso pretendido, como precisão, robustez ou capacidade de predição (Aguirre, 2015).

Nessa fase, avalia-se se o modelo é capaz de reproduzir adequadamente a dinâmica do sistema real e de prever sua resposta com dados distintos dos de treinamento, evitando a polarização. Idealmente, o modelo também deve ser parcimonioso, ou seja, representar o sistema com o menor número possível de termos, garantindo confiabilidade e adequação ao uso pretendido (Cassini, 1999).

Analisando em simulação livre as saídas estimadas utilizando os dados de validação, foi considerada a métrica Raiz Quadrada do Erro Quadrático Médio (RMSE, *Root Mean Squared Error*), conforme a Equação 3.4, como parâmetro de avaliação e comparação dos modelos obtidos pelos métodos FROLS e SEMP.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{k=1}^N (y(k) - \hat{y}(k))^2} \quad (3.4)$$

¹<https://github.com/soficaastro/is>

4 Resultados

Nesta seção, são apresentados os resultados encontrados com a implementação dos algoritmos FROLS e SEMP para identificação dos sistemas apresentados anteriormente. Os resultados serão analisados graficamente e comparados utilizando a métrica RMSE.

4.1 Sistema Simulado

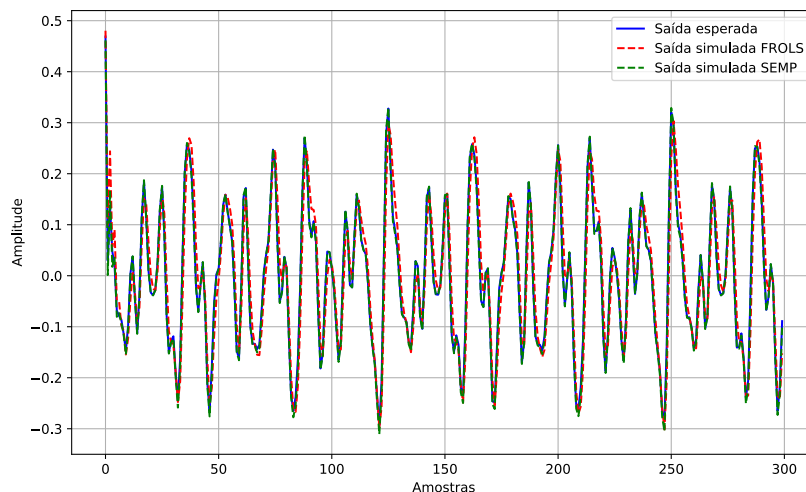
O primeiro caso de estudo foi um sistema simulado conforme apresentado em Billings (2013) na Equação 3.1. Para este sistema, os modelos NARX obtidos pelos métodos FROLS e SEMP são, respectivamente, os seguintes:

$$y(k) = 0,4287u(k-1) - 0,6771y(k-1) \quad (4.1)$$

$$y(k) = 0,6001u(k-1) - 0,6129y(k-1) - 0,2427u(k-2) - 0,0300u(k-2)^2 \quad (4.2)$$

O resultado da identificação do sistema é apresentado comparativamente na Figura 4.1, em que a curva azul representa a saída esperada do sistema e as curvas vermelha e verde tracejadas representam as saídas estimadas pelos modelos identificados pelos algoritmos FROLS e SEMP, respectivamente, em simulação livre.

Figura 4.1 – Gráfico de sinais de saída esperada e simulada do sistema simulado.



Fonte: Elaborado pela autora.

Pode-se observar que as saídas estimadas acompanharam de forma muito próxima a saída real. Os resultados satisfatórios eram esperados, uma vez que se trata de um sistema simulado, cuja dinâmica é bem compreendida, e que foi retirado de trabalhos anteriores nos quais os métodos de identificação demonstraram bom desempenho.

A Tabela 4.1 apresenta a comparação do desempenho entre os dois algoritmos, evidenciando um erro RMSE menor para o método SEMP em relação ao FROLS na tarefa de estimativa para este sistema. Considerando que a escala de variação da saída é relativamente pequena, essa diferença de desempenho se mostra particularmente relevante.

Tabela 4.1 – Raiz Quadrada do Erro Quadrático Médio e quantidade de termos regressores do sistema simulado

Método	RMSE	M_0
FROLS	0,0416	2
SEMP	0,0059	4

Fonte: Elaborado pela autora.

4.2 Sistema de Ferro de Solda

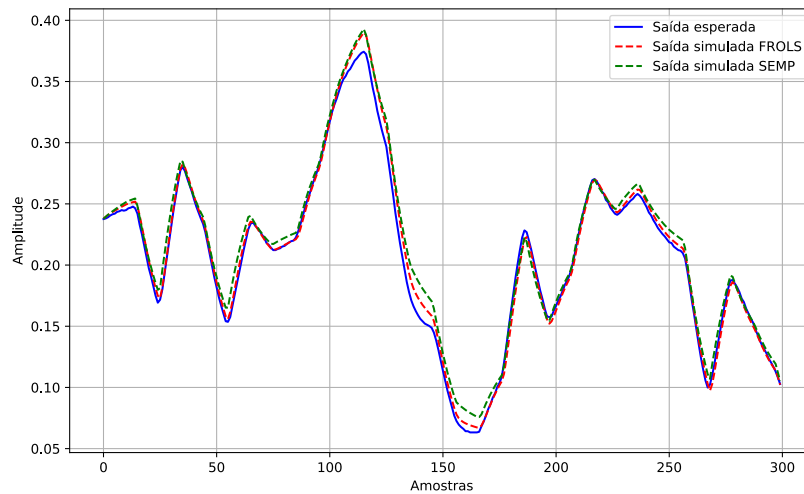
Este estudo de caso utiliza um sistema real apresentado em Cassini (1999). Os modelos NARX encontrados pelos algoritmos FROLS e SEMP são apresentados nas equações 4.3 e 4.4 respectivamente:

$$y(k) = 1,3760y(k-1) + 0,0219u(k-1)u(k-2) - 0,4167y(k-2) \quad (4.3)$$

$$y(k) = 0,9356y(k-1) + 0,0355u(k-1)u(k-2) \quad (4.4)$$

É possível comparar os resultados na Figura 4.2, em que a curva azul representa a saída esperada do sistema, enquanto as curvas vermelha e verde tracejadas correspondem às saídas produzidas pelos modelos identificados pelos algoritmos FROLS e SEMP, respectivamente, em simulação livre.

Figura 4.2 – Gráfico de sinais de saída esperada e simulada do sistema de ferro de solda.



Fonte: Elaborado pela autora.

Os sinais de saída gerados pelos modelos obtidos apresentam uma boa correspondência com a saída esperada do sistema. Isso indica que ambas as técnicas utilizadas foram capazes de capturar a dinâmica essencial do sistema de ferro de solda, fornecendo modelos adequados para a sua representação.

A Tabela 4.2 compara o desempenho dos dois algoritmos, revelando uma performance superior do método FROLS em relação ao SEMP, com um erro RMSE de 0,0059 frente a 0,0092. Novamente, por se tratar de um sistema com baixa variação na amplitude, essa diferença de desempenho se mostra particularmente relevante, indicando uma maior precisão do modelo identificado pelo FROLS neste caso específico.

Tabela 4.2 – Raiz Quadrada do Erro Quadrático Médio e quantidade de termos regressores do sistema de ferro de solda

Método	RMSE	M_0
FROLS	0,0059	3
SEMP	0,0092	2

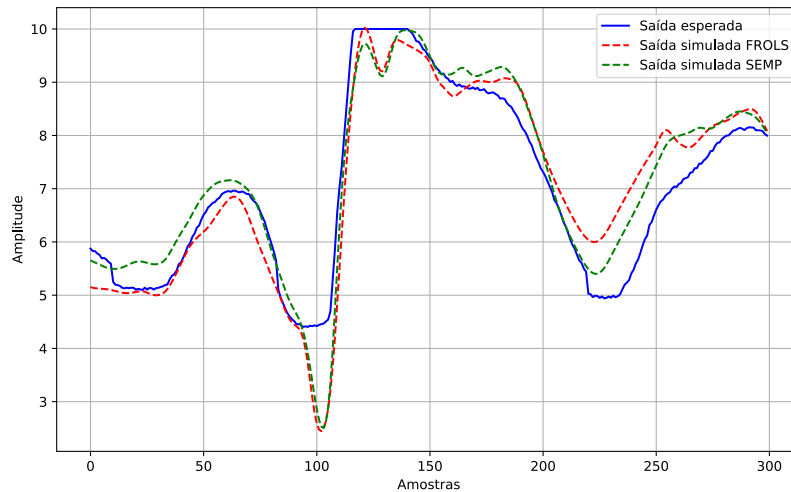
Fonte: Elaborado pela autora.

4.3 Sistema de Tanques em Cascata

O último sistema analisado neste trabalho foi o sistema de tanques em cascata, utilizando os dados coletados por Schoukens e Noël (2017). O resultado gráfico é mostrado na Figura 4.3, onde a curva azul representa a saída esperada do sistema, enquanto as curvas

vermelha e verde tracejadas representam, respectivamente, as saídas em simulação livre dos modelos obtidos pelos algoritmos FROLS e SEMP.

Figura 4.3 – Gráfico de sinais de saída esperada e simulada do sistema de tanques em cascata.



Fonte: Elaborado pela autora.

Observa-se que, neste caso de estudo, as saídas estimadas pelos modelos não apresentaram uma aproximação tão boa em relação à saída real do sistema, quando comparadas aos demais casos analisados. Esse resultado indica uma performance inferior dos modelos identificados.

Conforme apresentado na Tabela 4.3, os modelos encontrados apresentaram uma estrutura consideravelmente mais complexa em comparação aos casos anteriores, com um número elevado de termos selecionados (M_0). Além disso, os valores de RMSE obtidos foram significativamente mais elevados em relação aos demais sistemas avaliados. Embora o erro do SEMP tenha sido inferior, ambos os modelos identificados apresentaram limitação em reproduzir com precisão a dinâmica do sistema.

Tabela 4.3 – Raiz Quadrada do Erro Quadrático Médio e quantidade de termos regressores do sistema de tanques em cascata

Método	RMSE	M_0
FROLS	0,7799	34
SEMP	0,5568	43

Fonte: Elaborado pela autora.

5 Conclusão

A identificação de sistemas é essencial para a análise e controle de sistemas dinâmicos não lineares cuja representação matemática não é previamente conhecida. O foco principal deste trabalho foi a análise e comparação dos algoritmos FROLS e SEMP para a identificação de sistemas baseados em modelos polinomiais NARX, utilizando-se de três diferentes cenários e níveis de complexidade.

A partir dos resultados obtidos, nota-se que os algoritmos FROLS e SEMP apresentam bom desempenho geral na identificação dos sistemas. As limitações desses algoritmos foram evidenciadas para sistemas mais complexos, como o caso de estudo do sistema de tanques em cascata. Comparativamente, o método SEMP apresentou melhor desempenho na maioria dos casos analisados; porém, é importante notar que a escolha do método mais adequado pode depender do contexto específico do problema, sendo o SEMP um método mais custoso computacionalmente e que tende a produzir modelos mais complexos.

Este trabalho reforça a importância de estratégias eficientes de seleção de estrutura em modelos não lineares e abre espaço para investigações futuras com métodos híbridos ou baseados em aprendizado de máquina, que possam aprimorar a precisão da identificação em sistemas mais complexos e altamente não lineares. Além disso, propõe-se a exploração dos algoritmos FROLS e SEMP no contexto de sistemas MIMO, os quais, devido à sua natureza multivariada, tendem a gerar um grande número de regressores.

Referências

- Aguirre, L. A. (2015). *Introdução à Identificação de Sistemas - Técnicas Lineares e Não Lineares: Teoria e Aplicação*. Editora UFMG, Belo Horizonte, Brasil, 4. ed.
- Aguirre, L. A., Rodrigues, G. G., e Jácome, C. R. F. (1998). Identificação de Sistemas Não-Lineares Utilizando Modelos NARMAX Polinomiais – Uma Revisão e Novos Resultados. *SBA Controle & Automação*, 9(2):90–106.
- Alaoui, S. F., Caramiaux, B., Serrano, M., e Bevilacqua, F. (2012). Movement qualities as interaction modality. Em *Proceedings of the Designing Interactive Systems Conference on - DIS '12*.
- Araújo, I. B. Q. d. (2019). *Maximização da Correntropia por Simulação com Poda aplicado a detecção de estruturas e estimação de parâmetros de modelos NARX*. Tese de doutorado, Universidade Federal do Rio Grande do Norte, Natal, Brasil.
- Araújo, I. B. Q. d., Cavalcante, G., Lúcio, Y., e Araújo, F. (2019). Nonlinear system identification of a MIMO quadruple tanks system using NARX model. *Przegląd Elektrotechniczny*.
- Bijl, H., Schön, T., van Wingerden, J.-W., e Verhaegen, M. (2017). System identification through online sparse Gaussian process regression with input noise. *IFAC Journal of Systems and Control*, 2.
- Billings, S. A. (2013). *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*. John Wiley & Sons, United Kingdom.
- Bittanti, S. e Piroddi, L. (1996). Nonlinear identification and control of a heat exchanger: A neural network approach. *Journal of the Franklin Institute*, 334(1):135–153.
- Bottegal, G., Pillonetto, G., e Hjalmarsson, H. (2015). Bayesian kernel-based system identification with quantized output data. *IFAC-PapersOnLine*, 48(28):455–460.
- Cassini, C. C. d. S. (1999). *Estimação Recursiva de Características Estáticas Não Lineares Utilizando Modelos Polinomiais NARMAX*. Dissertação de mestrado, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Chen, S., Billings, S. A., e Luo, W. (1988). Orthogonal Least Squares Methods and Their Application to Nonlinear System Identification. Relatório técnico.
- Coelho, A. A. R. e Coelho, L. d. S. (2004). *Identificação de sistemas dinâmicos lineares*. Editora da UFSC, Florianópolis, Brasil, 1. ed.

- Gauss, C. F. e Stewart, G. W. (1995). *Theory of the Combination of Observations Least Subject to Errors, Part One, Part Two, Supplement*. Society for Industrial and Applied Mathematics.
- Gunawardena, R., Lang, Z.-Q., e He, F. (2024). NonSysId: A nonlinear system identification package with improved model term selection for NARMAX models.
- Gupta, S., Gupta, R., e Padhee, S. (2018). Parametric system identification and robust-controller design for liquid–liquid heatexchanger system. *IET Control Theory Applications*, 12(10):1474–1482.
- Hirata, Y., Morino, K., Suzuki, T., Guo, Q., Fukuhara, H., e Aihara, K. (2016). System identification and parameter estimation in mathematical medicine: examples demonstrated for prostate cancer. *Higher Education Press and Springer-Verlag Berlin Heidelberg*, 4(1):13–19.
- Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall, New Jersey, USA, 2. ed.
- Ljung, L. (2008). Perspectives on system identification. *IFAC Proceedings Volumes*, 41(2):7172–7184. 17th IFAC World Congress.
- Low, H. e Meghir, C. (2017). The Use of Structural Models in Econometrics. *Journal of Economic Perspectives*, 31(2):33–58.
- Mangan, N. M., Kutz, J. N., Brunton, S. L., e Proctor, J. L. (2017). Model selection for dynamical systems via sparse regression and information criteria. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 473(2204):20170009.
- Mehra, R. H. (1974). Identification in Control and Econometrics: Similarities and Differences. *Annals of Economic and Social Measurement*, 3(1).
- Narendra, K. e Parthasarathy, K. (1990). Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks*, 1(1):4–27.
- Nelles, O. (2001). *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*. Springer, Germany, 1. ed.
- Piroddi, L. e Spinelli, W. (2003). An identification algorithm for polynomial NARX models based on simulation error minimization. *International Journal of Control*, 76(17):1767–1781.
- Schoukens, M., Mattsson, P., Wigren, T., e Noël, J. P. (2016). Cascaded tanks benchmark combining soft and hard nonlinearities. Em *Workshop on Nonlinear System Identification Benchmarks 2016*.

- Schoukens, M. e Noël, J. P. (2017). Three Benchmarks Addressing Open Challenges in Nonlinear System Identification. Em *Proceedings of the 20th World Congress of the International Federation of Automatic Control (IFAC)*, pages 448–453, Toulouse, France.
- Schär, S., Marelli, S., e Sudret, B. (2024). Feature-centric nonlinear autoregressive models.
- Severino, A. G. V., Araújo, B. Q. d., Linhares, L., e Araújo, F. M. U. d. (2016). Metaheurísticas para Estimação de Parâmetros na Identificação de Sistema Não Lineares Utilizando Modelos NARMAX Polinomiais. Em *Congresso Nacional de Engenharia Mecânica*.
- Shu, Y., Liu, Y., Xu, Z., Zhao, X., e Chen, M. (2022). Optimization of Hydraulic Fine Blanking Press Control System Based on System Identification. *Modeling, Simulation and Control of Flexible Manufacturing Systems*.
- Tavares, L. A., Abreu, P. E. O. G. B., e Aguirre, L. A. (2020). Identification of NARX Models for Compensation Design. *ArXiv*.
- Tsuchiya, T., Fujii, M., Matsuda, N., Kunida, K., Uda, S., Kubota, H., e Kuroda, K. K. S. (2017). System identification of signaling dependent gene expression with different time-scale data. *PLOS Computational Biology*, 13(12).
- Verhaegen, M. e Verdult, V. (2007). *Filtering and System Identification: A Least Square Approach*. Cambridge University Press, New York, USA, 1. ed.
- Zadeh, L. A. (1956). On the Identification Problem. *IRE Transactions on Circuit Theory*, 3(4):277–281.
- Zadeh, L. A. (1962). From Circuit Theory to System Theory. *Proceedings of the IRE*, 50(5):856–865.