OSCAR WILFREDO RODRÍGUEZ RODRÍGUEZ

PRÉ-PROCESSAMENTO DE DADOS NA IDENTIFICAÇÃO DE
PROCESSOS INDUSTRIAIS

São Paulo

2015

OSCAR WILFREDO RODRÍGUEZ RODRÍGUEZ

PRÉ-PROCESSAMENTO DE DADOS NA IDENTIFICAÇÃO DE
PROCESSOS INDUSTRIAIS

Dissertação apresentada à Escola
Politécnica da Universidade de
São Paulo para obtenção do título
de Mestre em Ciências

Orientador:
Prof. Dr. Claudio Garcia

São Paulo

2015

OSCAR WILFREDO RODRÍGUEZ RODRÍGUEZ

PRÉ-PROCESSAMENTO DE DADOS NA IDENTIFICAÇÃO DE
PROCESSOS INDUSTRIAIS

Dissertação apresentada à Escola
Politécnica da Universidade de
São Paulo para obtenção do título
de Mestre em Ciências

Área de concentração:
Engenharia de Sistemas

Orientador:
Prof. Dr. Claudio Garcia

São Paulo

2015

Este trabalho é dedicado, com todo o meu apreço e carinho, aos meus pais Aniano e Marcela e aos meus irmãos Esther, Doris, Jesika e Carlos.

Agradecimentos

Ao professor Claudio Garcia, não apenas pela orientação e sábias palavras, mas principalmente pela amistosa convivência nos últimos anos.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo apoio financeiro.

Aos colegas de jornada Bruno Castro, Camilo Velasquez, Christiam Morales, Estela Mara de Oliveira, Fabian Nuñez, Jorge Alvarado, Roy Aguirre e Zoraida López pelos momentos de reflexão e descontração durante esta jornada.

Aos amigos Arianna Olivera, Ariagna Ramón, Ana Cárdenas, Carlos Chaves, Carmen Angulo, Chrisnael Saavedra, Michael Prieto, Robert Gavidea, Rosmeri Segura e Santos Tarrillo por me acompanhar e apoiar durante todo este tempo, tornando meu dia a dia mais alegre e descontraído.

Aos meus irmãos Esther, Doris Jesika e Carlos por todo apoio, amor e união, mesmo nos momentos mais difíceis. Por fim, gostaria de ressaltar a minha gratidão aos meus pais Aniano e Marcela, pois não sei como eu chegaria até aqui sem o exemplo de fibra e perseverança que vocês sempre demonstraram.

Resumo

Neste trabalho busca-se estudar as diferentes etapas de pré-processamento de dados na identificação de sistemas, que são: filtragem, normalização e amostragem. O objetivo principal é de acondicionar os dados empíricos medidos pelos instrumentos dos processos industriais, para que quando estes dados forem usados na identificação de sistemas, se possa obter modelos matemáticos que representem da forma mais próxima a dinâmica do processo real.

Vai-se também implementar as técnicas de pré-processamento de dados no software MatLab 2012b e vai-se fazer testes na Planta Piloto de Vazão instalada no Laboratório de Controle de Processos Industriais do Departamento de Engenharia de Telecomunicações e Controle da Escola Politécnica da USP; bem como em plantas simuladas de processos industriais, em que é conhecido *a priori* seu modelo matemático.

Ao final, vai-se analisar e comparar o desempenho das etapas de pré-processamento de dados e sua influência no índice de ajuste do modelo ao sistema real (*fit*), obtido mediante o método de validação cruzada. Os parâmetros do modelo são obtidos para predições infinitos passos à frente.

Palavras-chave: Filtragem. Normalização. Reamostragem. Pré-processamento de dados. Identificação de sistemas.

Abstract

This work aims to study the different stages of data pre-processing in system identification, as are: filtering, normalization and sampling. The main goal is to condition the empirical data measured by the instruments of industrial processes, so that when these data are used to identify systems, one can obtain mathematical models that represent more closely the dynamics of the real process.

It will also be implemented the techniques of preprocessing of data in MatLab 2012b and it will be performed tests in the Pilot Plant of Flow at the Laboratory of Industrial Process Control, Department of Telecommunications and Control Engineering from the Polytechnic School of USP; as well as with simulated plants of industrial processes where it is known a priori its mathematical model.

At the end, it is analyzed and compared the performance of the pre-processing of data and its influence on the index of adjustment of the model to the real system (*fit*), obtained by the cross validation method. The model parameters are obtained for infinite step-ahead prediction.

Keywords: Filtering. Standardization. Resampling. Preprocessing of data. Identification systems.

Lista de Figuras

2.1	Sinal de tempo contínuo $x(t)$	7
2.2	Representação de $x(t)$ como sinal de tempo discreto $x[n]$	8
2.3	Sinal par	8
2.4	Sinal ímpar	9
2.5	Sinal periódico, para $T = 2s$	10
2.6	Sinal aleatório	11
2.7	Esquema geral de um sistema em tempo contínuo	12
2.8	Esquema geral de um sistema em tempo discreto	13
2.9	Sistema invertível	14
2.10	Esquema geral de um sistema	17
2.11	Sistema de controle realimentado	18
2.12	Diagrama de blocos de um filtro não-recursivo FIR de ordem m . . .	20
3.1	Magnitude do filtro $H(e^{j\omega})$ para $h(n) = \{2, 3, 4, 3, 2\}$	26
3.2	Fase linear por partes do filtro $H(e^{j\omega})$ para $h(n) = \{2, 3, 4, 3, 2\}$. . .	26
3.3	Amplitude $A(e^{j\omega})$ do filtro $h(n) = \{2, 3, 4, 3, 2\}$	28
3.4	Fase linear $\theta(\omega)$ do filtro $h(n) = \{2, 3, 4, 3, 2\}$	28
3.5	Filtro passa-baixa ideal com frequência de corte ω_c	31
3.6	Resposta ao impulso de um filtro ideal para $\omega_c = 0, 6$ e comprimento $N = 41$	32
3.7	Especificações do filtro FIR	33

3.8	Processo de janelamento no domínio da frequência	36
3.9	Formas de várias funções de janelas	39
3.10	Filtro passa-baixas pelo método da janela no domínio da frequência .	39
3.11	Filtro passa-baixas pelo método da janela (dB)	40
3.12	Amostragem do filtro ideal com $N = 31$ e $f_c = 0,25\text{Hz}$	43
3.13	Resposta em frequência aproximada pelo método de amostragem em frequência com $N = 31$, $f_c = 0,25\text{Hz}$	44
3.14	Resposta em frequência aproximada pelo método de amostragem em frequência com $N = 30$ e $f_c = 0,25\text{Hz}$	44
3.15	Resposta em frequência do filtro de Chebyshev	49
3.16	Resposta em frequência do filtro de Chebyshev (dB)	50
3.17	Normalização de dados em faixas iguais	53
3.18	Variância de $\hat{a}_N$ plotado como intervalo do período de amostragem $T_s(\omega_0 = 1/\tau)$ (1) e T_s muito grande (2)	56
4.1	Diagrama P&ID da Planta Piloto de Vazão	60
4.2	Não-linearidade da Planta Piloto de Vazão	61
4.3	Dados coletados de entrada e saída	62
4.4	Espectro em potência (dB) do sinal $PV(t)$	63
4.5	Resposta em frequência do método da janela retangular	65
4.6	Resposta em frequência do método da janela de Bartlett	65
4.7	Resposta em frequência do método da janela de Hann	65
4.8	Resposta em frequência do método da janela de Hamming	66
4.9	Resposta em frequência do método da janela de Blackman	66
4.10	Resposta em frequência do método da janela de Kaiser	66
4.11	Resposta em frequência do método de amostragem em frequência. . .	67
4.12	Resposta em frequência do método de mínimo erro quadrático	67
4.13	Resposta em frequência do método de Chebyshev	67

4.14	Sinal filtrado pelo método da janela retangular	68
4.15	Sinal filtrado pelo método da janela de Bartlett	68
4.16	Sinal filtrado pelo método da janela de Hann	68
4.17	Sinal filtrado pelo método da janela de Hamming	69
4.18	Sinal filtrado pelo método da janela de Blackman	69
4.19	Sinal filtrado pelo método da janela de Kaiser	69
4.20	Sinal filtrado pelo método de amostragem em frequência	70
4.21	Sinal filtrado pelo método de mínimo erro quadrático	70
4.22	Sinal filtrado pelo método de Chebyshev	70
4.23	Função de covariância normalizada da planta de piloto de vazão . . .	75
4.24	Normalização de dados sem filtrar da Planta Piloto de Vazão (a) . . .	78
4.25	Normalização de dados sem filtrar da Planta Piloto de Vazão (b) . . .	79
4.26	Normalização de dados filtrados da Planta Piloto de Vazão	81
4.27	Simulação de uma planta virtual no Simulink	85
4.28	Resposta ao degrau da planta rápida e de baixo ganho	86
4.29	Dados coletados da planta rápida e de baixo ganho	87
4.30	Densidade espectral de potência do sinal de saída da planta rápida e de baixo ganho	88
4.31	Resposta ao degrau da planta rápida e de ganho alto	91
4.32	Dados coletados da planta rápida e de ganho alto	91
4.33	Densidade espectral de potência do sinal de saída da planta rápida e de baixo ganho	92
A.1	Etapas do procedimento do projeto de identificação	111
A.2	Estrutura do modelo ARX	122
A.3	Estrutura do modelo ARMAX	123
A.4	Estrutura do modelo ARARX	124

A.5	Estrutura do modelo ARARMAX	125
A.6	Estrutura do modelo de erro na saída	126
A.7	Estrutura do modelo Box-Jenkins	126

Lista de Tabelas

3.1	Resumo das características da função janela $w(n)$	40
4.1	Comprimento N do filtro passa-baixa projetado para $\Delta\omega = 0,0126$. .	64
4.2	<i>Fit</i> de validação cruzada com dados filtrados da Planta Piloto de Vazão para os filtros projetados	72
4.3	<i>Fit</i> de validação cruzada com dados sem filtrar Planta Piloto de Vazão para os filtros projetados	73
4.4	<i>Fit</i> de validação cruzada da Planta Piloto de Vazão para os métodos de reamostragem projetados com dados sem filtrar	76
4.5	<i>Fit</i> de validação cruzada da Planta Piloto de Vazão para os métodos de reamostragem projetados com dados filtrados pelo método de Chebyshev.	77
4.6	<i>Fit</i> de validação cruzada do modelo ARMAX para $n_a = n_b = n_c = a$ e $n_k = 0, a = 2, 3, 4$ com dados sem filtrar	80
4.7	textitFit de validação cruzada do modelo ARMAX para $n_a = n_b = n_c = a$ e $n_k = 0, a = 2, 3, 4$ com dados filtrados	82
4.8	<i>Fit</i> de validação cruzada da planta rápida e de baixo ganho para os filtros projetados	89
4.9	<i>Fit</i> de validação cruzada da planta rápida e de baixo ganho para os métodos de reamostragem projetados com dados filtrados	89
4.10	<i>Fit</i> de validação cruzada da planta rápida e de baixo ganho para os métodos de normalização	90
4.11	<i>Fit</i> de validação cruzada da planta rápida e de baixo ganho para os filtros projetados	93

4.12	<i>Fit</i> de validação cruzada da planta rápida e de ganho alto para os métodos de reamostragem projetados com dados filtrados	93
4.13	<i>Fit</i> de validação cruzada da planta rápida e de ganho alto para os métodos de normalização	94
A.1	Alguns casos especiais de modelos SISO	127

Simbologia e abreviações

Operadores e notação

$\hat{\theta}$: estimativa de θ ;

$\mathbb{N}$: conjunto de números naturais;

$\max(\cdot)$: valor máximo;

$\min(\cdot)$: valor mínimo;

∇ : operador gradiente;

$\|\cdot\|$: norma euclidiana;

$E[\cdot]$: operador de esperança matemática;

$H(e^{j\omega T}) \odot W(e^{j\omega T})$: convolução periódica das variáveis $H(e^{j\omega T})$ e $W(e^{j\omega T})$;

q : operador de deslocamento, $q^{-1}y(k) = y(k - 1)$;

s : operador matemático em tempo contínuo, $s = j\omega$;

$x(n) * h(n)$: produto de convolução finita dos sinais discretizados $x(n)$ e $h(n)$;

z : operador matemático em tempo discreto, $z = e^{j\omega T}$;

$\text{cov}(X, Y)$: covariância das variáveis X e Y ;

Símbolos gerais

$\Delta\omega_t$: faixa de transição entre a frequência de passagem e a frequência de rejeição.;

ω : frequência [rad/s];

$\omega(n)$: função de janelamento;

ω_c : frequência de corte;

ω_p : frequência de passagem;

ω_s : frequência de rejeição;

ω_{c_n} : frequência de corte normalizada;

σ : desvio padrão;

σ^2 : variância;

τ : constante de tempo;

$\varepsilon(t)$: erro de predição;

A_s : faixa de ondulações na faixa de rejeição [dB];

E : energia total do sinal;

E_2 : erro quadrático médio ou norma L_2 ;

E_∞ : norma L_∞ ;

f : frequência [Hz];

f_c : frequência de corte [Hz];

K, α, β : constantes de proporcionalidade;

N : comprimento do filtro;

P : potência média do sinal;

R_p : faixa de ondulações na faixa de passagem [dB];

$r_{y^*}(\tau)$: função de covariância linear do sinal $y^*(k)$;

$r_{y^{*2'}}(\tau)$: função de covariância não-linear do sinal $y^*(k)$;

T_s : período de amostragem;

t_s : tempo de acomodação;

T_{ch} : período de chaveamento;

u : média de um conjunto de valores;

Abreviações e acrônimos

fit: índice de ajuste do modelo ao sistema real;

BN: ruído binário;

dB: decibels;

DEP: densidade espectral de potência (*power spectral density*);

DTF: transformada discreta de Fourier (*discrete Fourier transform*);

FDP: função densidade de probabilidade (*function density probability*);

FI: indicador de vazão (*flow indicator*);

FIC: indicador e controlador de vazão (*flow indicator controller*);

FIR: resposta finita ao impulso (*finite impulse response*);

FIT: transmissor e indicador de vazão (*flow indicator transmitter*);

FV: válvula de controle de vazão (*flow valve*);

GBN: sinal binário generalizado (*generalized binary signal*);

IDTF: inversa da transformada discreta de Fourier (*inverse discrete Fourier transform*);

IIR: resposta infinita ao impulso (*infinite impulse response*);

ILD: diferenças de nível interneuronal (*interaural level differences*);

LIT: sistema linear invariante no tempo (*linear time-invariant*);

MIMO: sistema com múltiplas entradas e múltiplas saídas (*multiple input and multiple output*);

MPC: controle preditivo baseado em modelo (*model predictive control*);

P& ID: diagrama de tubulação e instrumentação (*piping and instrumentation diagram*);

PID: controlador proporcional-integral-derivativo (*proportional-integral-derivative*);

PRBS: sinal binário pseudo-aleatório (*pseudo-random binary sequence*);

PV: variável de processo (*process variable*);

RBS: sinal binário aleatório (*random binary signal*);

SD: *speed driver*;

SISO: sistema com uma única entrada e uma única saída (*single input and single output*);

Sumário

1	Introdução	1
1.1	Motivação	1
1.2	Objetivos	2
1.3	Justificativa	3
1.4	Contribuições propostas	3
1.5	Estrutura do texto	4
2	Sinais e sistemas	5
2.1	Sinais	6
2.1.1	Definição	6
2.1.2	Classificação	7
2.2	Sistemas	12
2.2.1	Definição	12
2.2.2	Propriedades	13
2.3	Conceitos de sistemas de controle	16
2.4	Revisão Bibliográfica	18
2.4.1	Filtragem	18
2.4.2	Normalização	20
2.4.3	Reamostragem	21

3	Pré-processamento de dados	22
3.1	Filtragem	22
3.1.1	Métodos de filtragem	23
3.1.2	Filtros FIR de fase linear	25
3.1.3	Passos para projetar um filtro digital	30
3.1.4	Projeto de filtros causais pelo método da janela	35
3.1.5	Projeto por amostragem em frequência	41
3.1.6	Projeto por mínimo erro quadrático	45
3.1.7	Projeto pelo método de Chebyshev	47
3.2	Normalização	50
3.2.1	Normalização por subtração da média	51
3.2.2	Normalização pelo valor máximo	51
3.2.3	Normalização pelo desvio padrão	51
3.2.4	Normalização padrão	52
3.2.5	Normalização pelo valor máximo sem média	52
3.2.6	Normalização em faixas iguais	53
3.3	Reamostragem	54
3.3.1	Método de autocovariância	54
3.3.2	Método pela constante do tempo	56
3.3.3	Método de décimo do tempo de acomodação	57
3.3.4	Método de décimo da constante de tempo	57
4	Resultados e discussões	58
4.1	Planta Piloto de Vazão	58
4.1.1	Descrição da Planta Piloto de Vazão	59
4.1.2	Coleta de dados	61
4.1.3	Avaliação de filtragem na identificação	62

4.1.4	Avaliação de reamostragem na identificação	74
4.1.5	Avaliação de normalização na identificação	78
4.2	Plantas virtuais	85
4.2.1	Planta rápida e de baixo ganho	86
4.2.2	Planta rápida e de ganho alto	90
5	Conclusões e sugestões para trabalhos futuros	96
5.1	Conclusões gerais	96
5.2	Sugestões para trabalhos futuros	97
	Referências Bibliográficas	99
A	Identificação de sistemas	104
A.1	Modelos Matemáticos	106
A.1.1	Tipos de modelo	106
A.1.2	Métodos de obtenção de modelos	108
A.2	Procedimento de identificação de modelos paramétricos	109
A.3	Sinal de entrada e saída	111
A.4	Modelo de sistemas LIT e tipos de estruturas	115
A.4.1	Modelo de sistemas lineares invariantes no tempo	115
A.4.2	Método dos Mínimos Quadrados	116
A.4.3	Erro de predição	119
A.4.4	Estruturas do modelo	121
A.5	Validação do modelo	128

Capítulo 1

Introdução

Este capítulo introdutório se inicia com uma discussão dos aspectos mais relevantes que motivaram a escolha do tema de pesquisa. A seguir, são descritos os objetivos a cumprir no presente trabalho (seção 1.2) e a justificativa para sua realização (seção 1.3). Na seção 1.4, se descrevem as principais contribuições do trabalho. Finalmente, se apresenta a estrutura da dissertação na seção 1.5.

1.1 Motivação

A indústria vai evoluindo constantemente e suas necessidades também vão mudando. Na década de 1950, o controle automático de processos substituiu o controle manual, em que o operador gerava o sinal de controle depois de observar um erro entre o valor desejado e a variável do processo; com o controle automático o operador passou só a realizar tarefas de supervisão e manutenção, e é um algoritmo de controle que recebe os sinais das variáveis do processo e gera e envia os sinais de controle para um atuador. Em quase todos os casos, o algoritmo de controle é o PID. No entanto, com o passar do tempo, a necessidade de melhorar as malhas de controle tem levado à busca não de um controle com parâmetros de sintonia estática, como é o algoritmo PID, mas de muitos algoritmos de controle adaptativos e/ou preditivos, pois muitos processos industriais são variantes no tempo, por causa disso, o processo vai mudando sua dinâmica e precisa-se também gerar um algoritmo que seja capaz

de obter um modelo que represente o comportamento da dinâmica do processo, seja *online* ou *offline*. A identificação de sistemas surge para satisfazer essa necessidade.

A identificação de sistemas é um campo de estudo que faz uso de métodos estatísticos para obter modelos matemáticos de sistemas dinâmicos, para isso, são requeridos valores medidos das entradas e das saídas. O modelo matemático obtido não tem nenhuma representação física, no entanto os dados obtidos pelo modelo identificado podem ser muito próximos aos valores das variáveis geradas pelo sistema com o mesmo sinal de entrada.

Desde a década de 1970, vem-se estudando, desenvolvendo e aplicando as técnicas de identificação de sistemas ao controle de processos industriais, obtendo-se bons resultados. Realizou-se uma pesquisa bibliográfica e encontrou-se que o estudo do pré-processamento de dados na identificação de sistemas é uma área ainda pouco estudada, sendo esse o objetivo do presente estudo, fazer a análise qualitativa e quantitativa das técnicas de filtragem, normalização e amostragem; e sua implicação e aplicação no rendimento do processo de identificação.

1.2 Objetivos

Estudar e desenvolver algoritmos das técnicas de filtragem, normalização e amostragem; e realizar a implementação no MatLab R2012b.

Estudar as considerações que tem-se que fazer para aplicar as técnicas de pré-processamento de dados de acordo com o tipo de variáveis industriais e suas faixas de valores.

Aplicar as técnicas de pré-processamento de dados em plantas simuladas que contam com variáveis com presença do ruído, com variáveis com faixas muito diferentes e com muitas taxas de amostragem e realizar análises quantitativas e qualitativas da melhoria do índice *fit* no processo da identificação de sistemas.

1.3 Justificativa

Na atualidade, as indústrias tanto químicas quanto petroquímicas estão usando controladores avançados, como o controlador preditivo baseado em modelo (MPC), que como seu nome indica, precisa de um modelo matemático que seja capaz de reproduzir o comportamento dinâmico do sistema, de forma que o controlador possa estimar ou prever o valor da saída do sistema, e conseqüentemente possa gerar os sinais de controle corrigindo os efeitos, tanto das perturbações quanto das trocas de carga do sistema, antes que estas gerem um erro entre a variável controlada e o valor de referência.

O cenário descrito implica procurar técnicas ou ferramentas que ajudem a melhorar os modelos matemáticos obtidos pelo método de identificação de sistemas. Está aí a justificativa e importância do presente estudo, pois dado que a identificação de sistemas é um método de natureza empírica, pretende-se pré-processar os dados empíricos das variáveis a serem usadas na identificação de sistemas. Finalmente avaliar, tanto quantitativamente o índice de ajuste do modelo ao sistema real (*fit*) na identificação, quanto qualitativamente os critérios que têm que se usar, dependendo do tipo de variável a identificar.

1.4 Contribuições propostas

A primeira contribuição originada deste trabalho é ter uma base experimental documentada da influência da etapa de pré-processamento de dados na identificação de processos industriais. Os métodos de identificação paramétrica foram amplamente estudados e há uma ampla documentação dos mesmos em diferentes artigos e livros, mas não da influência do pré-processamento de dados no ajuste do modelo ao sistema real.

Outra contribuição originada deste trabalho é a implementação dos algoritmos de filtragem, normalização e re-amostragem no software Matlab. Estes programas

podem servir como base para experimentos em novas aplicações industriais que precisem de identificação de sistemas.

Por fim, este trabalho pode ser usado como base de comparação entre métodos de filtragem não-recursivos (abordados ao longo do estudo) e métodos de filtragem recursivos. Os métodos de filtragem recursivos não foram alvo deste trabalho.

1.5 Estrutura do texto

O capítulo 1 apresenta uma breve introdução do conteúdo e a importância deste estudo, também definem-se os objetivos e sua justificativa.

O capítulo 2 trata das definições básicas de sinais e sistemas que são usadas durante o desenvolvimento do presente trabalho e que formam a base teórica. Depois disso, apresenta-se uma revisão bibliográfica dos estudos e/ou trabalhos que servem como apoio teórico do presente estudo.

O capítulo 3 trata do estudo dos métodos de pré-processamento, isto é, filtragem, normalização e reamostragem.

O capítulo 4 trata da coleta, pré-processamento de dados, obtenção de modelos e avaliação do desempenho do índice de ajuste do modelo ao sistema real (*fit*) na identificação de uma Planta Piloto de Vazão, instalada no Laboratório de Controle de Processos Industriais do Departamento de Engenharia de Telecomunicações e Controle da Escola Politécnica da USP. Também, trata a análise do desempenho do pré-processamento de dados em plantas virtuais e a análise dos resultados obtidos.

O capítulo 5 apresenta as principais conclusões do trabalho sumarizadas. Posteriormente, são citadas algumas propostas que visam dar continuidade ao tema explorado neste trabalho.

O apêndice A trata do estudo da metodologia de identificação de sistemas e algumas considerações, tanto na hora de escolher os sinais de excitação quanto no tratamento dos sinais coletados.

Capítulo 2

Sinais e sistemas

Ao longo da história para o desenvolvimento das sociedades, a humanidade procurou diferentes formas de se comunicar, tanto em seu dia a dia quanto com outras gerações. Para isso usou símbolos, pinturas, sons, obras de arte, entre outros.

Na atualidade, a comunicação é de vital importância na sociedade. Quando uma pessoa usa o celular, pesquisa na Internet, faz um pagamento em uma caixa eletrônica, faz um agendamento no hospital ou simplesmente fala, está se comunicando. A dita comunicação está baseada em duas partes importantes; o sinal, que envolve o meio físico; e o protocolo de comunicação, que é um conjunto de regras pré-definidas para interpretar os sinais.

Um sinal é definido como *uma função de uma ou mais variáveis, o qual veicula informações sobre a natureza de um fenômeno físico* [12]; ou também, como *algo que transmite informação* [28]; ou simplesmente como *um conjunto de dados ou informações*[19].

O alvo deste capítulo é o estudo dos fundamentos básicos dos sinais, que envolvem algumas definições importantes, seu processamento digital e a importância nos sistemas de controle industrial.

2.1 Sinais

2.1.1 Definição

Sinal é um termo que deriva do latim *signālis*, trata-se de uma variável que pode ser transmitida por qualquer meio e contém informação relativa à natureza ou ao comportamento de um fenômeno físico.

Ao longo do tempo e do avanço tecnológico, as aplicações de sinais foram mudando. A seguir se descrevem alguns exemplos:

- No transporte. Quando um semáforo muda de cor de vermelho ao verde, o meio pelo qual a onda luminosa é transmitido é o meio ambiente e a informação que proporciona para o motorista é a permissão para seguir a diante;
- Na instrumentação industrial. O medidor de vazão eletromagnético consta de uma bobina montada à volta da tubulação, que ao passar uma intensidade de corrente constante através dela, gera um campo magnético constante. Quando a vazão passa através deste campo magnético gerado por dita bobina, gera-se uma corrente induzida que é medida por dois eletrodos instalados em cada lado da tubulação. Dado que o campo magnético é constante, a intensidade da corrente induzida é proporcional à vazão do líquido na tubulação. O meio pelo qual a informação da medida de vazão é transmitida é elétrico e a informação que ela proporciona é a medida de vazão.
- Nos sistemas de controle. Em uma malha fechada de controle SISO, o controlador recebe o valor medido da variável a ser controlada, compara com um valor de referência predefinido e, através de um algoritmo de controle, gera um sinal que é enviado a um atuador que modifica o processo. O meio de transmissão é elétrico (mA) e a informação envolvida nele é tanto a entrada do controlador, que é o valor medido da variável controlada, quanto sua saída, que é a amplitude de ação do atuador.

2.1.2 Classificação

Nesta subseção vai-se considerar o estudo dos sinais unidimensionais em função do tempo, isto é, para um instante de tempo existe um único valor do sinal.

A seguir se apresentam os métodos de classificação do sinal, baseado em diferentes recursos [12].

Sinais de tempo contínuo e tempo discreto

Esta classificação está baseada na forma como o sinal está definido em função do tempo. Um sinal $x(t)$ é dito ser de tempo contínuo se está definido para todo tempo t . Na Figura 2.1 se apresenta um exemplo de um sinal de tempo contínuo.

Figura 2.1: Sinal de tempo contínuo $x(t)$.

Um sinal de tempo discreto está definido somente para instantes isolados de tempo. O sinal de tempo discreto usualmente é derivado de um sinal de tempo contínuo amostrado a uma taxa uniforme. Seja T o período de amostragem e n um número inteiro, pode-se representar o tempo t como $t = nT$ e o sinal em tempo discreto (amostrado) está definido por $x(t) = x(nT)$. Por conveniência, na notação vai-se considerar: $x[n] = x(nt)$, para $n = 0, \pm 1, \pm 2, \dots$. Na Figura 2.2 se apresenta o sinal $x(t)$ e o sinal amostrado $x[n]$ para um período de amostragem de $T = 1s$.

Figura 2.2: Representação de $x(t)$ como sinal de tempo discreto $x[n]$.

Sinais pares e ímpares

Define-se um sinal de tempo contínuo como um sinal par, se ele satisfizer a condição

$$x(-t) = x(t), \quad (2.1)$$

isto é, os sinais pares são simétricos em relação ao eixo das ordenadas. A Figura 2.3 mostra o exemplo de um sinal par.

Figura 2.3: Sinal par.

Define-se um sinal de tempo contínuo como um sinal ímpar, se ele satisfizer a condição

$$x(-t) = -x(t) \quad (2.2)$$

isto é, os sinais pares são assimétricos em relação ao eixo das ordenadas. A Figura 2.4 mostra o exemplo de um sinal ímpar.

Figura 2.4: Sinal ímpar.

Sinais periódicos e não-periódicos

Diz-se que um sinal $x(t)$ é periódico se satisfaz a seguinte condição:

$$x(t) = x(t + T), \quad \forall t \in \mathbb{R}, \quad T \in \mathbb{R}^+, \quad (2.3)$$

onde T é o período fundamental do sinal. O período fundamental T define a duração de um ciclo completo de $x(t)$.

Define-se a frequência fundamental do sinal f como o recíproco do período fundamental T . A frequência fundamental f é medida em Hertz (Hz) ou ciclos por segundo e está representada na equação (2.4).

$$f = \frac{1}{T} \quad (2.4)$$

A frequência angular fundamental do sinal medido em radianos por segundo (rad/s) é definida na equação (2.5).

$$\omega = \frac{2\pi}{T} \quad (2.5)$$

Na Figura 2.5 se mostra um sinal $x(t)$ com período fundamental igual a $T = 2s$, isto é, $x(t) = x(t + 2)$.

Figura 2.5: Sinal periódico, para $T = 2s$.

Diz-se que um sinal $x(t)$ é assimétrico ou não-periódico se não existe nenhum valor de T que satisfaça a equação (2.3).

Sinais determinísticos e sinais aleatórios

Um sinal determinístico é um sinal sobre o qual não existe nenhuma incerteza com respeito ao seu valor em qualquer instante de tempo. Daí que os sinais determinísticos podem ser modelados como funções do tempo completamente especificadas [12]. Os sinais dos processos industriais são considerados determinísticos, daí que o comportamento dinâmico do sistema pode ser representado mediante um modelo matemático com um certo grau de incerteza. A Figura 2.1 representa um sinal determinístico que está definido pela expressão matemática:

$$0,5 \cdot \sin(0,2\pi t) + 0,5 \cdot \sin(0,3\pi t).$$

Um sinal aleatório é aquele sobre o qual há incerteza antes de sua ocorrência real [12]. O sinal aleatório não pode ser definido por uma expressão matemática, mas pode ser representado por suas características estocásticas como a média, a variância e a autocorrelação. Na Figura 2.6 se mostra uma realização de um sinal aleatório.

Figura 2.6: Sinal aleatório.

Sinais de energia e sinais de potência

Diz-se que um sinal é de energia se e somente se a energia total do sinal satisfizer a condição

$$0 < E < \infty$$

onde E é a energia total do sinal e está definida tanto no tempo contínuo $x(t)$, pela equação (2.6); quanto no tempo discreto $x[n]$, pela equação (2.7).

$$E = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} x^2(t) dt \quad (2.6)$$

$$E = \sum_{n=-\infty}^{\infty} x^2[n] \quad (2.7)$$

Diz-se que um sinal é de potência se e somente se a potência média do sinal satisfizer a condição

$$0 < P < \infty$$

onde P é a potência média do sinal e está definida tanto no tempo contínuo $x(t)$, pela equação (2.8); quanto no tempo discreto $x[n]$, pela equação (2.9).

$$P = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt \quad (2.8)$$

$$P = \lim_{N \rightarrow \infty} \frac{1}{2N} \sum_{n=-N}^N x^2[n] \quad (2.9)$$

As classificações de energia e potência média de sinais são mutuamente exclusivas. Em especial, um sinal de energia tem potência média zero, enquanto que um sinal de potência tem energia infinita. Normalmente, os sinais periódicos e sinais aleatórios são vistos como sinais de potência, enquanto os sinais determinísticos ou não-periódicos são sinais de energia [12].

2.2 Sistemas

2.2.1 Definição

Um sistema muitas vezes é visto como uma coleção de componentes que recebem uma excitação forçada, chamada entrada, $x(t)$, e produzem uma resposta $y(t)$ chamada de saída [6]. O sistema não precisa necessariamente ser físico, também pode ser conceitual ou abstrato.

Figura 2.7: Esquema geral de um sistema em tempo contínuo.

Na Figura 2.7 apresenta-se o esquema geral de um sistema. O sistema conta com uma ou mais entradas que interagem com o sistema e modifica sua saída ou conjunto de saídas.

No presente trabalho, um sistema é um modelo matemático que relaciona seus sinais de entrada (excitações) com seus sinais de saída (respostas).

2.2.2 Propriedades

Sistemas em tempo contínuo e em tempo discreto

Diz-se que um sistema é em tempo contínuo quando seus sinais de entrada e de saída estão no tempo contínuo, isto é $x(t) \rightarrow y(t)$ (vide Figura 2.7). Se os sinais de entrada e de saída estão no tempo discreto $x[n] \rightarrow y[n]$, então o sistema é dito de tempo discreto (vide Figura 2.8).

Figura 2.8: Esquema geral de um sistema em tempo discreto.

Em geral, o modelo do sistema em tempo contínuo é representado por equações diferenciais e o modelo do sistema em tempo discreto por equações de diferenças.

Memória

Diz-se que um sistema é sem memória, instantâneo, ou estático se seu sinal de saída em qualquer instante depende somente do sinal de entrada nesse mesmo instante, isto é, não depende de nenhum valor passado ou futuro do sinal de entrada. Caso contrário, se diz que o sistema possui memória ou é dinâmico.

Na equação (2.10) apresenta-se o modelo de um sistema sem memória, onde K é um ganho de proporcionalidade.

$$y(t) = K \cdot x(t) \quad (2.10)$$

Quando um sistema não é instantâneo se diz que é um sistema dinâmico, isto porque não depende só dos sinais de entrada nesse instante mas também dos valores dos sinais de entrada passados. Um exemplo típico é o modelo de um capacitor C que tem a corrente $x(t)$ como entrada e a tensão $y(t)$ como saída, apresentada na equação (2.11).

$$y(t) = \frac{1}{C} \int_0^t x(\tau) d\tau \quad (2.11)$$

Invertibilidade

Diz-se que um sistema é invertível se sinais de entrada diferentes geram diferentes sinais de saída. Se um sistema é invertível, então existe um sistema inverso que, quando é excitado com o sinal de saída, produz o sinal de entrada original. Na equação (2.12) se apresenta um exemplo de um sistema invertível no tempo contínuo e no tempo discreto.

$$\begin{aligned} S_1 &: x(t) \rightarrow y(t) = 3 \cdot x(t) \\ S_2 &: x[n] \rightarrow y[n] = \sum_{k=-\infty}^n x[k] \end{aligned} \quad (2.12)$$

Diz-se que um sistema é não-invertível se existirem dois ou mais sinais de entrada diferentes que geram um mesmo sinal de saída. Na equação (2.13) se apresenta um exemplo de um sistema não-invertível no tempo contínuo e no tempo discreto.

$$\begin{aligned} S_3 &: x(t) \rightarrow y(t) = x^2(t) \\ S_4 &: x[n] \rightarrow y[n] = x[n] - x[n-1] \end{aligned} \quad (2.13)$$

Se o sistema S for invertível então existe um outro sistema, chamado de sistema inverso S^{-1} , que ligado à saída do sistema S produz como sua saída a entrada de S (vide Figura 2.9).

Figura 2.9: Sistema invertível.

Causalidade

Diz-se que um sistema é causal se sua saída em qualquer instante do tempo depende somente dos valores presente e passado da entrada. O termo causali-

dade denota a relação causa-efeito. Estes sistemas também são chamados de não-antecipatórios, pois o sistema não antecipa, nem depende dos valores futuros da entrada.

Se a saída do sistema depende de valores futuros, o sistema é chamado não-causal ou antecipativo. Na equação (2.14) se apresenta um exemplo de um sistema não-causal.

$$\begin{aligned} S_1 &: x(t) \rightarrow y(t) = x(t+2) \\ S_2 &: x[n] \rightarrow y[n] = x[-n] \end{aligned} \quad (2.14)$$

Estabilidade

Diz-se que um sistema é estável se para sinais de entrada delimitados, geram-se sinais de saída delimitados. Isto é, seja k_1 e k_2 duas constantes reais e finitas, se considera que se a entrada está definida por

$$|x| \leq k_1,$$

então a saída está definida por

$$|y| \leq k_2.$$

No caso contrário se diz que o sistema é instável.

Invariância

Um sistema é invariante no tempo se sua conduta e suas características são fixas no tempo. Também pode-se dizer que um sistema é invariante no tempo se um deslocamento no tempo (atraso ou avanço) no sinal de entrada resulta em um deslocamento igual no sinal de saída. Na equação (2.15) se apresenta um sistema invariante no tempo tanto contínuo quando discreto para qualquer valor real de t_0 ou inteiro de n_0 .

$$\begin{aligned} x(t) \rightarrow y(t) &: x(t-t_0) \rightarrow y(t-t_0) \\ x[n] \rightarrow y[n] &: x[n-n_0] \rightarrow y[n-n_0] \end{aligned} \quad (2.15)$$

Um sistema é dito variante no tempo quando não obedece a equação (2.15).

Linearidade

Diz-se que um sistema é linear em tempo contínuo ou discreto, caso se aplique a propriedade da superposição. Isto é, se a entrada consiste na soma ponderada de vários sinais, então a saída é a superposição (soma ponderada) das respostas do sistema a cada um desses sinais. Então o sistema é linear se cumpre a propriedade de aditividade e homogeneidade apresentadas nas equações (2.16) e (2.17), respectivamente.

$$x_1(t) + x_2(t) \rightarrow y_1(t) + y_2(t) \quad (2.16)$$

$$x_1(t) \rightarrow \alpha \cdot y_1(t) \quad (2.17)$$

Pode-se combinar as equações (2.16) e (2.17) na equação (2.18) para tempo contínuo e discreto.

$$\begin{aligned} \alpha \cdot x_1(t) + \beta \cdot x_2(t) &\rightarrow \alpha \cdot y_1(t) + \beta \cdot y_2(t) \\ \alpha \cdot x_1[n] + \beta \cdot x_2[n] &\rightarrow \alpha \cdot y_1[n] + \beta \cdot y_2[n] \end{aligned} \quad (2.18)$$

2.3 Conceitos de sistemas de controle

Ao longo do desenvolvimento deste estudo se faz referência a diferentes conceitos da teoria de controle de sistemas [8, 10, 18, 27, 29], visto que a aplicação deste estudo procura obter melhores modelos matemáticos para aumentar a eficiência de controle nos processos industriais. Por causa disto, definem-se alguns conceitos básicos a seguir.

Planta

Define-se como uma parte ou um conjunto de partes que trabalham conjuntamente para realizar uma operação determinada; ou também um objeto real a ser controlado. Ao longo deste estudo, uma planta é o objetivo que se deseja identificar para melhorar o seu controle.

Processo

Define-se como uma operação caracterizada por uma série de trocas graduais, contínuas, sequenciais e que têm um determinado resultado ou valor final. O resultado ou valor final é o que se deseja controlar.

Sistema de controle

Define-se como um conjunto de componentes que ao interagir entre si buscam um determinado objetivo.

Figura 2.10: Esquema geral de um sistema.

Na Figura 2.10 apresenta-se o esquema geral de um sistema. O sistema conta com uma ou mais entradas que ativam a dinâmica do sistema e modificam a saída ou conjunto de saídas do sistema; no controle de processos as entradas do sistema (processo) são usadas para modificar as saídas para um valor de referência. O sistema conta também com uma ou mais entradas de perturbação, elas são de natureza aleatória e não desejadas, pois levam as saídas do sistema para fora dos valores de referência.

Perturbação

Define-se como todo sinal não desejado que modifica o valor da saída fora do valor de referência. Pode ser de dois tipos: interna, quando a perturbação é gerada dentro do sistema, ou externa, quando a perturbação é gerada fora do sistema; neste último caso pode-se representar como uma entrada do sistema.

Sistema de controle realimentado

Refere-se ao sistema que procura minimizar ou eliminar a diferença que existe entre o sinal de saída (variável controlada) e uma entrada de referência, em um instante de tempo determinado. O controlador usa dita diferença como um ponto de partida para gerar um sinal de controle que modifique o sistema. Na Figura 2.11 apresenta-se o diagrama de blocos de um sistema de controle realimentado.

2.4 Revisão Bibliográfica

2.4.1 Filtragem

Nesta subseção realiza-se uma revisão dos principais trabalhos publicados, que são a sustentação teórica e conceitual acerca do tema da filtragem. Em [31, 34, 41, 43, 50] classifica-se o projeto dos filtros digitais em: filtros IIR (*Infinite Impulse Response*) e filtros FIR (*Finite Impulse Response*).

- A saída do filtro IIR está definida por:

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{n=0}^M b_n z^{-n}}{\sum_{n=0}^N a_n z^{-n}} = \frac{b_0 + b_1 z^{-1} + \dots + b_M z^{-M}}{1 + a_1 z^{-1} + \dots + a_N z^{-N}}; \quad (2.19)$$

onde $a_0 = 1$, b_n e a_n são os coeficientes do filtro. O comprimento do filtro é chamado de N se $a_N \neq 0$. O filtro definido na equação (2.19) pode ser representado pela equação diferencial que está definida por:

$$y(n) = -\sum_{k=1}^N a(k)y(n-k) + \sum_{k=0}^M b(k)x(n-k) \quad (2.20)$$

sendo sua forma equivalente:

$$\sum_{k=0}^N a(k)y(n-k) = \sum_{k=0}^M b(k)x(n-k); \quad a(0) = 1$$

Da equação (2.19), pode-se ver que ao aplicar um filtro tipo IIR se adiciona pólos e zeros ao sistema que se pretende identificar e, portanto, modificam sua dinâmica. Por isto, ao longo deste estudo não vai-se trabalhar com filtros IIR.

- A saída do filtro FIR está definida por:

$$y(n) = \sum_{m=0}^{N-1} h(m)x(n-m) \quad (2.21)$$

onde $h(m)$ representa os coeficientes da resposta ao impulso, $x(n)$ é a sequência de entradas (conjunto de dados a ser filtrado) e $y(n)$ é a sequência de saídas (conjunto de dados filtrados).

A função de transferência do filtro FIR é:

$$H(z) = \sum_{k=0}^{N-1} h(k)z^{-k} \quad (2.22)$$

Da equação (2.22), pode-se ver que ao aplicar um filtro tipo FIR se adiciona somente zeros ao sistema que se pretende identificar e portanto não modifica sua dinâmica.

Um dos principais trabalhos com respeito aos filtros FIR é [36], nele sintetizam-se algumas técnicas dos filtros não-recursivos como: *Convolução direta*, que é usada quando está completamente definido a resposta ao impulso do filtro, sendo a saída filtrada definida na equação (2.21) e sua representação gráfica se mostra na Figura 2.12.

Como se pode ver na Figura 2.12, a entrada $x(n)$ passa por uma sequência de m blocos de retardo que a atrasam; depois de cada bloco o sinal é multiplicado pelo correspondente coeficiente da resposta ao impulso. A soma total desses produtos é a saída do sinal filtrado $y(n)$.

Figura 2.12: Diagrama de blocos de um filtro não-recursivo FIR de ordem m .

Em [4, 24, 31] desenvolveram-se tanto os conceitos básicos da teoria da filtragem, quanto os diferentes métodos de projetos de filtros não-recursivos. Definem-se também os quatro tipos em que se classificam os filtro FIR de fase linear, dependendo de sua simetria e da paridade ou imparidade do comprimento N .

Em [45, 35, 48] apresentam-se o desenvolvimento teórico da filtragem pelo método de frequência de amostragem. Nestes trabalhos busca-se determinar o resposta ao impulso finito de um filtro FIR $h(n)$, baseado no método de aproximação ao filtro ideal, para isso o filtro ideal é amostrado em N partes. Em [25], apresenta-se a forma para se implementar o dito filtro na linguagem de programação FORTRAN.

2.4.2 Normalização

Nesta subseção vai-se fazer uma revisão dos principais trabalhos publicados a respeito da normalização de dados.

Em [47] apresenta-se um resumo dos principais modelos de normalização. O autor define um banco de ensaios que compara sete modelos de normalização para nove funções de sensibilidade. Os dados experimentais foram coletado mediante sensores que gravam, dos neurônios do cérebro, as variações das diferenças de nível interneuronal (*Interaural Level Differences*, ILD) em referência à posição azimutal de uma fonte sonora. Os principais modelos de normalização apresentados são normalização por correção da média, normalização pelo valor máximo de cada variável, normalização pelo desvio padrão e normalização padrão.

Em [14] apresenta-se o modelo de normalização por subtração da média, onde a média μ é tirada do conjunto de dados de cada variável, por causa disso as faixas

dos dados normalizados ficam em torno do zero, mas suas amplitudes ainda são não padronizadas.

Em [42] apresenta-se o modelo de normalização pelo desvio padrão, onde a média μ é tirada do conjunto de dados de cada variável e o resíduo é dividido pelo desvio padrão, por causa disso as faixas dos dados normalizados ficam em torno do zero e sua amplitude é reduzida proporcionalmente à dispersão dos dados.

2.4.3 Reamostragem

Os sistemas industriais e em geral, a maioria dos sistemas reais, são processos que se encontram no domínio do tempo contínuo. Para o controle, supervisão e armazenamento de suas variáveis, e para a identificação de seus modelos, precisa-se que as ditas variáveis sejam amostradas. Define-se o intervalo de amostragem T_s como o período entre duas amostras.

Em [1], o autor indica que a forma de escolher o tempo de amostragem é fazendo uso do Teorema de Shannon, que diz que a frequência de amostragem tem que ser “de 2 a 10 vezes a maior frequência de interesse”. Mas o problema é que nem sempre a frequência de interesse é conhecida *a priori*. Então o autor recomenda fazer uma amostragem do sinal com um período muito pequeno $y^*(k)$ e avaliar os efeitos causados pela super-amostragem nas funções de autocovariância linear e não-linear. A quantificação desses efeitos e o processo de escolha do tempo de amostragem recomendado vão ser apresentados na subseção 3.3.

Em [22] o autor faz uma avaliação da forma de minimizar a perda de informação devido à escolha do tempo de amostragem. Para isso, faz uma avaliação da variância do parâmetro do modelo quando se determina o preditor por duas formas: a primeira quando se faz uso do modelo de erro de saída para $T_s = 1/\tau$, a segunda quando se faz uso de um integrador como um simples filtro de pré-amostragem e se usa o modelo de erro de saída para T_s suficientemente grande. Os resultados obtidos são que a escolha ótima do intervalo de amostragem encontra-se em torno da constante de tempo τ do sistema e na prática recomenda-se usar como tempo de amostragem a décima parte da constante de tempo dominante do sistema.

Capítulo 3

Pré-processamento de dados

Nas últimas décadas, a identificação de sistemas usando estruturas de modelos lineares foi intensamente explorada, tanto que atualmente existem ferramentas padronizadas e bem estabelecidas para tratar tal problema. Inclusive em aplicações de processos não-lineares, pode-se aplicar aproximações lineares em torno de um ponto de operação com resultados satisfatórios.

No entanto, não há muitos estudos práticos da influência do pré-processamento de dados na identificação de sistemas.

Ao longo do presente capítulo vai-se realizar o estudo das técnicas de pré-processamento de dados, compostas por: filtragem, normalização e reamostragem. Técnicas que são utilizadas no capítulo 4 na avaliação da sua influência no índice *fit* de validação de modelos.

3.1 Filtragem

Os filtros digitais têm duas grandes classificações: os filtros *não-recursivos*, cuja principal característica é que têm uma resposta de duração finita ao impulso (FIR); e os *filtros recursivos* que têm uma resposta infinita ao impulso (IIR).

Os filtros FIR são estáveis, pois só adicionam zeros ao sistema, mas o comprimento do filtro pode ser muito maior e portanto requer muitas operações computacionais.

Os filtros IIR podem tornar-se instáveis e/ou alterar a dinâmica do sistema, pois não só adicionam zeros mas também polos; no entanto o comprimento do filtro é muito menor em comparação com os filtros FIR e não requerem muitas operações computacionais na sua implementação.

Considera-se o comprimento do filtro como o número de coeficientes da resposta ao pulso unitário de um filtro digital FIR ou IIR.

3.1.1 Métodos de filtragem

Ao projetar um filtro, deve-se escolher entre dois métodos: aproximação ou realização. O problema do *método de aproximação* está na escolha de parâmetros ou coeficientes de uma função de transferência $H(z)$, definida na equação (3.1), cuja resposta em frequência se aproxime de uma resposta ideal (Figura 3.5) ou às vezes a uma resposta em frequência desejada, de acordo com uma aplicação específica. Esta aproximação normalmente é feita usando a resposta em frequência. O problema do *método de realização* reside na escolha de uma estrutura que permita implementar a função de transferência $H(z)$. Esta estrutura pode ter a forma de um diagrama de circuitos se o filtro é para ser construído de componentes; ou pode ser um programa a ser utilizado em um computador de uso geral ou de um microprocessador para processamento do sinal.

Ao longo desta seção, vai-se estudar o projeto de filtros digitais mediante o método de aproximação, que basicamente consta das seguintes etapas:

1. Escolher uma resposta ideal ou desejada de um filtro no domínio da frequência.
2. Escolher um tipo e o comprimento do filtro a projetar.
3. Escolher um método ou algoritmo que permita estimar os parâmetros do filtro.
4. Escolher um critério que permita medir a qualidade da aproximação entre o filtro desejado e o filtro projetado.

A longo do estudo, vai-se assumir que o filtro digital é causal e que pode ser representado pela função de transferência (3.1).

$$H(z) = \frac{b_0 + b_1 z^{-1} + \cdots + b_{N-1} z^{-(N-1)}}{1 + a_1 z^{-1} + \cdots + a_M z^{-M}} \quad (3.1)$$

onde $N - 1$ é o grau do polinômio do numerador; M é o grau do polinômio do denominador; $b_0, b_1, b_2, \dots, b_{N-1}$ e $a_1, a_2, \dots, a_M$ são os parâmetros a serem estimados.

A função de transferência discreta de um filtro FIR pode ser obtida a partir da equação (3.1) e expressa como:

$$H(z) = \sum_{n=0}^{N-1} h(n) z^{-n} \quad (3.2)$$

onde $h(n) = b_n$ para $n = 0, 1, \dots, N - 1$.

A função de transferência $H(z)$ vai ser empregada para projetar filtros digitais tipo FIR no domínio da frequência, isto é, para $z = e^{j\omega T}$ onde T é o período de amostragem. Ao longo do estudo vai-se considerar, por simplicidade e padronização, o período de amostragem igual a $T = 1$ s. Então a equação (3.2) resulta em:

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} h(n) e^{-j\omega n} \quad (3.3)$$

onde N é o comprimento do filtro; $h(n)$ são os parâmetros do filtro FIR para $n = 0, 1, \dots, N - 1$ que se deseja estimar e ω é a frequência em radianos por segundo definida por:

$$\omega = 2\pi f$$

onde f é a frequência em Hertz ou ciclos por segundo.

Em termos gerais, o problema do método de aproximação para filtros tipo FIR reside na estimativa dos parâmetros $h(n)$ da equação (3.3), tal que se minimize a diferença entre a resposta ideal ou desejada $D(e^{j\omega})$ e a resposta real $H(e^{j\omega})$ no domínio da frequência para o intervalo de $-\pi$ até π .

Devido ao método de implementação, o filtro FIR é chamado também de filtro não-recursivo ou filtro de convolução, às vezes também é intitulado filtro de média móvel.

O comprimento do filtro FIR por definição é finito, portanto sua saída pode ser escrita como o somatório da convolução finita [2, 15, 17, 28, 31, 39, 38, 40, 50]:

$$y(n) = x(n) * h(m) = \sum_{m=0}^{N-1} h(m)x(n-m) \quad (3.4)$$

onde $y(n)$ é a saída da convolução entre o sinal de entrada $x(n)$ e a resposta ao impulso unitário $h(m)$ de comprimento N .

Com uma mudança dos índices das variáveis, (3.4) resulta em:

$$y(n) = \sum_{m=n}^{n-N+1} h(n-m)x(m) \quad (3.5)$$

3.1.2 Filtros FIR de fase linear

A seguir, estudam-se duas características importantes dos filtros FIR: a fase linear e a simetria.

Fase linear

O filtro FIR definido na equação (3.3) possui valores complexos, portanto pode ser representado como:

$$H(e^{j\omega}) = R(e^{j\omega}) + jI(e^{j\omega}) \quad (3.6)$$

onde o módulo $M(e^{j\omega})$ e a fase $\phi(e^{j\omega})$ estão definidos por:

$$M(e^{j\omega}) = |H(e^{j\omega})| = \sqrt{R^2 + I^2} \quad (3.7)$$

$$\phi(e^{j\omega}) = \angle H(e^{j\omega}) = \arctan\left(\frac{I}{R}\right) \quad (3.8)$$

Portanto,

$$H(e^{j\omega}) = M(e^{j\omega})e^{j\phi(e^{j\omega})} \quad (3.9)$$

Na equação (3.9) surge um problema: $M(e^{j\omega})$ não é uma função analítica e $\phi(e^{j\omega})$ é linear por partes. Este problema se apresenta nas Figuras 3.1 e 3.2. Para resolver este problema, se introduz uma função de amplitude $A(e^{j\omega}) \in R$ que pode assumir valores positivos e negativos.

Figura 3.1: Magnitude do filtro $H(e^{j\omega})$ para $h(n) = \{2, 3, 4, 3, 2\}$.

Figura 3.2: Fase linear por partes do filtro $H(e^{j\omega})$ para $h(n) = \{2, 3, 4, 3, 2\}$.

A função de resposta em frequência impulsiva de comprimento N apresentada na equação (3.3), vai ser deslocada, assim resultando:

$$H(e^{j\omega}) = e^{-j\omega M} \sum_{n=0}^{N-1} h(n) e^{j\omega(M-n)} \quad (3.10)$$

ou

$$H(e^{j\omega}) = e^{-j\omega M} \left[h_0 e^{j\omega M} + h_1 e^{j\omega(M-1)} + \dots + h_{N-1} e^{j\omega(M-N+1)} \right] \quad (3.11)$$

onde

$$M = \frac{N-1}{2} \quad (3.12)$$

ou

$$M = N - M - 1$$

Aplicando a fórmula de Euler na equação (3.11), resulta:

$$H(e^{j\omega}) = e^{-j\omega M} \left[(h_0 + h_{N-1}) \cos(\omega M) + j(h_0 - h_{N-1}) \sin(\omega M) \right. \\ \left. + (h_1 + h_{N-2}) \cos(\omega(M-1)) \right. \\ \left. + j(h_1 - h_{N-2}) \sin(\omega(M-1)) + \dots \right] \quad (3.13)$$

A equação (3.11) pode ser descrita como:

$$H(e^{j\omega}) = A(e^{j\omega})e^{(\alpha+\beta\omega)} \quad (3.14)$$

onde α pode conter valores de $\alpha = 0$ ou $\alpha = \pi/2$ e $A(e^{j\omega})$ é a amplitude do filtro $H(e^{j\omega})$.

- Quando $\alpha = 0$, diz-se que o filtro apresenta simetria par e se expressa como:

$$h(n) = h(N - n - 1) \quad (3.15)$$

daí a equação (3.14) resulta em:

$$H(e^{j\omega}) = A(e^{j\omega})e^{-j\omega M} \quad (3.16)$$

e

$$A(e^{j\omega}) = \sum_{n=0}^{M-1} 2h(n) \cos(\omega(M - n)) + h(M) \quad (3.17)$$

Realizando-se uma mudança de variáveis na equação (3.17), ela pode ser reescrita como:

$$A(e^{j\omega}) = \sum_{n=1}^M 2h(M - n) \cos(\omega n) + h(M) \quad (3.18)$$

Para N par $A(e^{j\omega})$ em (3.17) resulta:

$$A(e^{j\omega}) = \sum_{n=0}^{N/2-1} 2h(n) \cos(\omega(M - n)) \quad (3.19)$$

A equação (3.19) pode ser reescrita como:

$$A(e^{j\omega}) = \sum_{n=1}^{N/2} 2h(N/2 - n) \cos(\omega(n - 1/2)) + h(M) \quad (3.20)$$

- Quando $\alpha = \pi/2$, diz-se que o filtro apresenta simetria ímpar e se expressa como:

$$h(n) = -h(N - n - 1) \quad (3.21)$$

A equação (3.13) resulta:

$$H(e^{j\omega}) = e^{-j\omega M} \left[\begin{aligned} &(h_0 - h_{N-1}) \cos(\omega M) + j(h_0 + h_{N-1}) \sin(\omega M) \\ &+ (h_1 - h_{N-2}) \cos(\omega(M - 1)) \\ &+ j(h_1 + h_{N-2}) \sin(\omega(M - 1)) + \dots \end{aligned} \right] \quad (3.22)$$

A amplitude para N par está definida por:

$$A(e^{j\omega}) = \sum_{n=0}^{M-1} 2h(n) \sin(\omega(M-1)) \quad (3.23)$$

$$A(e^{j\omega}) = \sum_{n=0}^{N/2-1} 2h(n) \sin(\omega(M-n)) \quad (3.24)$$

O filtro FIR com coeficientes reais é chamado PAR, quando a sequência conjugada dos coeficientes é simétrica; e é chamado ÍMPAR, quando a sequência conjugada dos coeficientes é anti-simétrica.

A amplitude $A(e^{j\omega})$ do filtro FIR é uma função analítica e a fase do filtro é totalmente linear, como se pode ver nas Figuras 3.3 e 3.4.

Figura 3.3: Amplitude $A(e^{j\omega})$ do filtro $h(n) = \{2, 3, 4, 3, 2\}$.

Figura 3.4: Fase linear $\theta(\omega)$ do filtro $h(n) = \{2, 3, 4, 3, 2\}$.

Periodicidade

A periodicidade pode ser avaliada partindo-se da equação (3.3) e acrescentando-se a fase em 2π , isto resulta:

$$\begin{aligned}
 H(e^{j(\omega+2\pi)}) &= \sum_{n=0}^{N-1} h(n) e^{-j(\omega n + 2\pi n)} \\
 &= \sum_{n=0}^{N-1} h(n) e^{-j\omega n} \cdot e^{-j2\pi n} \\
 &= H(e^{j\omega})
 \end{aligned} \tag{3.25}$$

onde ω é a frequência angular (radianos/segundo) e f está em Hertz (ciclos/segundo).

$$\omega = 2\pi f$$

Portanto, da equação (3.25), fica demonstrado que a variável ω é periódica, com período de 2π .

Tipos de filtros de fase linear

Os filtros FIR de fase linear classificam-se em 4 tipos [28, 31, 49] dependendo da simetria de sua resposta impulsiva e da paridade ou imparidade do comprimento N .

Das equações (3.18), (3.20), (3.23) e (3.24), resulta:

- Tipo 1: $h(n)$ é simétrica e N é ímpar

$$H(e^{j\omega}) = \left[\sum_{n=1}^M 2h(M-n) \cos(\omega n) + h(M) \right] e^{-jM\omega} \tag{3.26}$$

- Tipo 2: $h(n)$ é simétrica e N é par

$$H(e^{j\omega}) = \left[\sum_{n=1}^{N/2} 2h\left(\frac{N}{2} - n\right) \cos\left(\omega\left(n - \frac{1}{2}\right)\right) \right] e^{-jM\omega} \tag{3.27}$$

- Tipo 3: $h(n)$ é anti-simétrica e N é ímpar

$$H(e^{j\omega}) = j \left[\sum_{n=1}^M 2h(M-n) \sin(\omega n) + h(M) \right] e^{-jM\omega} \tag{3.28}$$

- *Tipo 4: $h(n)$ é anti-simétrica e N é par*

$$H(e^{j\omega}) = j \left[\sum_{n=1}^{N/2} 2h\left(\frac{N}{2} - n\right) \sin\left(\omega\left(n - \frac{1}{2}\right)\right) \right] e^{-jM\omega} \quad (3.29)$$

onde

$$M = \frac{N-1}{2}$$

e sendo a condição de simetria e anti-simetria $h[n] = h[N-n-1]$, $h[n] = -h[N-n-1]$, respectivamente.

Pode-se observar que $H(e^{j\omega})$ nos quatro tipos tem como fator comum $e^{-jM\omega}$, o qual na transformada de Laplace é igual a e^{-Ms} , isto é, no tempo contínuo $H(e^{j\omega})$ tem um atraso de M períodos de amostragem. Caso se deseje eliminar a defasagem gerada pela filtragem dos sinais, tem-se que defasar os sinais filtrados em $-M$ períodos de amostragem, sendo isso só possível em uma filtragem *offline*.

3.1.3 Passos para projetar um filtro digital

O procedimento geral para projetar filtros digitais consiste em alguns passos, que são abordados na presente subseção:

- Partindo-se de uma resposta em frequência ideal ou desejada $D(e^{j\omega})$, pretende-se encontrar uma resposta em frequência $H(e^{j\omega})$ semelhante a $D(e^{j\omega})$.
- A realização do filtro é feita procurando-se otimizar alguma medição de rendimento do mesmo, tais como: minimizar o comprimento N do filtro; minimizar a largura da faixa de transição ou reduzir o erro na faixa de passagem e/ou na faixa de rejeição.
- Definir as especificações de projeto do filtro, com elas, vai-se obter ou estimar os parâmetros do filtro.

Filtro ideal passa-baixa

Dado que a resposta em frequência de um filtro é sempre periódica com respeito à variável ω , com um período de 2π , a especificação do filtro só precisa ser feita

por um período no intervalo de $[-\pi, \pi]$. Dado que o objetivo do presente estudo está orientado ao projeto de filtros passa-baixa, por sua aplicação nos processos industriais, e sabendo que os filtros de tipos 1 e 2 são simétricos, então a especificação da resposta pode ser avaliada somente no intervalo positivo $[0, \pi]$.

Na Figura 3.5 é apresentado o filtro ideal passa-baixa, com fase zero. Este é definido por sua resposta em frequência como:

$$D(e^{j\omega}) = \begin{cases} 1 & , \quad |\omega| < \omega_c \\ 0 & , \quad \omega_c < |\omega| < \pi \end{cases} \quad (3.30)$$

onde ω_c é a frequência de corte.

Figura 3.5: Filtro passa-baixa ideal com frequência de corte ω_c .

A resposta impulsiva do filtro ideal de duração infinita pode ser calculada a partir da transformada inversa de Fourier no tempo discreto (IDFT) mostrada em (3.31).

$$\begin{aligned} h_D(n) &= \mathcal{F}^{-1}[D(e^{j\omega})] = \frac{1}{2\pi} \int_{-\pi}^{\pi} D(e^{j\omega}) \cdot e^{j\omega n} d\omega \\ h_D(n) &= \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} 1 \cdot e^{j\omega n} d\omega \\ h_D(n) &= \frac{\sin(\omega_c n)}{\pi n} \\ h_D(n) &= \omega_c \frac{\sin(\omega_c n)}{\pi n \omega_c} \\ h_D(n) &= \frac{\omega_c}{\pi} \frac{\sin\left(\frac{\omega_c n}{\pi} \pi\right)}{\left[\pi \frac{\omega_c n}{\pi}\right]} \\ h_D(n) &= \frac{\omega_c}{\pi} \text{sinc}\left(\frac{\omega_c n}{\pi}\right) \end{aligned} \quad (3.31)$$

onde a função *sinc* é definida por:

$$\text{sinc}(x) = \frac{\sin(x\pi)}{\pi x}$$

A equação (3.31) pode ser reescrita como:

$$h_D(n) = \begin{cases} \frac{\omega_c}{\pi} & , \quad n = 0 \\ \frac{\omega_c}{\pi} \text{sinc}\left(\frac{\omega_c n}{\pi}\right) & , \quad n \neq 0 \end{cases} \quad (3.32)$$

De forma geral, o centro do filtro que naturalmente é na origem, pode ser deslocado de α , tal que o extremo esquerdo da resposta ao impulso unitário desejado $h_D(n)$ coincida com a origem. Isto é, com $\alpha = (N - 1)/2$, $h_D(n)$ é representado na equação (3.33) e ilustrado na Figura 3.6.

$$h_D(n) = \begin{cases} \frac{\omega_c}{\pi} & , \quad n = 0 \\ \frac{\omega_c}{\pi} \text{sinc}\left(\frac{\omega_c(n-\alpha)}{\pi}\right) & , \quad n \neq 0 \end{cases} \quad (3.33)$$

Figura 3.6: Resposta ao impulso de um filtro ideal para $\omega_c = 0,6$ e comprimento $N = 41$.

Na Figura 3.6 pode-se observar que os extremos do filtro passa-baixa ideal não convergem a zero, pois o filtro foi truncado para um comprimento $N = 41$. Para representar corretamente o filtro ideal passa-baixa, o comprimento do filtro deveria ser infinito, por isso, numericamente não é possível ser implementado.

Projeto das especificações do filtro

Dado que o filtro ideal passa-baixa (filtro desejado) é não realizável, é necessário definir algumas especificações para o filtro $H(e^{j\omega})$ em relação ao filtro desejado.

Estas especificações podem ser classificadas em duas formas [33]: especificações absolutas e especificações relativas; e se mostram na Figura 3.7.

Figura 3.7: Especificações do filtro FIR: (a) absoluto, (b) relativo [33].

(a) *Especificações absolutas*: conforme indicado na Figura 3.7(a), possuem as seguintes características:

- A faixa $[\omega_p, \omega_s]$ é chamada de *faixa de transição* e expressa como $\Delta\omega_t = \omega_s - \omega_p$. Não existe nenhuma restrição da resposta em frequência nesta faixa.
- A faixa $[0, \omega_p]$ é chamada de *faixa de passagem*. O termo δ_p é a tolerância a ondulações na faixa de passagem;
- A faixa $[\omega_s, \pi]$ é chamada de *faixa de rejeição*. O termo δ_s é a tolerância a ondulações na faixa de rejeição; e

(b) *Especificações relativas*: conforme indicado na Figura 3.7(b), possuem as seguintes características:

- R_p é a faixa de ondulações toleradas na faixa de passagem expressa em dB ; e
- A_s é a faixa de ondulações toleradas na faixa de rejeição expressa em dB .

Os parâmetros apresentados nas especificações absolutas e relativas estão relacionadas por:

$$R_p = -20 \log_{10} \frac{1 - \delta_p}{1 + \delta_p} > 0 (\approx 0) \quad (3.34)$$

e

$$A_s = -20 \log_{10} \frac{\delta_s}{1 + \delta_p} > 0 (\gg 1) \quad (3.35)$$

O objetivo ideal durante o processo de projeto de um filtro FIR é que os parâmetros δ_p , δ_s , $\Delta\omega_t$ e o comprimento N do filtro sejam os menores possíveis. Contudo, na prática, no projeto do filtro se minimiza um parâmetro e os outros ficam fixos.

Geralmente, o filtro pode ser expresso por suas duas componentes (vide na equação (3.14)), a amplitude e a fase. Dado que o objetivo do presente estudo é projetar filtros passa-baixas, estes filtros FIR devem ser de tipo 1 ou 2, equações (3.26) ou (3.27), respectivamente. Para filtros tipos 3 ou 4, quando $\omega = 0$ na equação (3.22), resulta $A(e^{j\omega}) = 0$, que não é consistente para um filtro passa-baixas.

A fase para um filtro com simetria par (filtros tipo 1 e 2) é:

$$\angle H(e^{j\omega}) = -M\omega$$

sendo que o sinal negativo representa um atraso de M períodos de intervalo de amostragem.

Medição do erro

A medição do erro é muito importante ao estimar os coeficientes do filtro $H(e^{j\omega})$ em função da aproximação do filtro desejado $D(e^{j\omega})$. O erro é definido por:

$$E(e^{j\omega}) = W(e^{j\omega}) [D(e^{j\omega}) - H(e^{j\omega})] \quad (3.36)$$

onde $W(e^{j\omega})$ é uma função real de ponderação de erro. Vai-se considerar $E(e^{j\omega})$ como uma medição escalar do erro, conhecida como *norma do erro*.

A seguir se definem diferentes formas de calcular a norma do erro, que são usualmente usadas no projeto de filtros:

- Erro quadrático médio ou norma L_2 :

$$E_2 = \left[\frac{1}{2\pi} \int_B |E(e^{j\omega})|^2 d\omega \right]^{1/2} \quad (3.37)$$

- Chebishev ou norma L_∞ :

$$E_\infty = \max_{\omega \in B} |E(e^{j\omega})| \quad (3.38)$$

onde B é a faixa de frequências de interesse sobre a qual se deseja minimizar a norma do erro. A faixa de frequências B está definida no intervalo: $B \subset [-\pi, \pi]$.

3.1.4 Projeto de filtros causais pelo método da janela

O janelamento é uma método de projetar filtros digitais mediante o truncamento, em ambos os lados, de um filtro ideal $h_D(n)$ mediante uma função (chamada de *janela*). Assim, pode-se obter filtros FIR causais e de fase linear. O filtro projetado $h(n)$ é definido na equação (3.39).

$$h(n) = h_D(n)w(n) \quad (3.39)$$

onde $h(n)$ é o filtro projetado e $w(n)$ é a função janela.

No domínio da frequência, a resposta do filtro causal $H(e^{j\omega})$ é definida pela convolução periódica de $H_D(e^{j\omega})$ com a janela $W(e^{j\omega})$, definida na equação (3.40) e explicada graficamente na Figura 3.8.

$$\begin{aligned} H(e^{j\omega}) &= H_D(e^{j\omega}) \odot W(e^{j\omega}) \\ H(e^{j\omega}) &= \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} W(e^{j(\omega-\lambda)}) d\lambda \end{aligned} \quad (3.40)$$

Da Figura 3.8, pode-se observar que:

- Seja N o comprimento do filtro ideal, então a largura do lóbulo principal da sua resposta em frequência é $2\pi/N$;

Figura 3.8: Processo de janelamento no domínio da frequência [34].

- A resposta em frequência da convolução periódica (3.40) dá uma versão distorcida da resposta em frequência do filtro ideal;
- A largura da faixa de transição de $H(e^{j\omega})$ é proporcional à largura do lóbulo principal de $W(e^{j\omega})$, portanto proporcional a $1/N$; e
- Os lóbulos laterais de $W(e^{j\omega})$ geram ondulações similares nas bandas de passagem e de rejeição.

Ao projetar um filtro FIR pelo método de janelamento, surgem dois problemas:

- (a) *Problema da largura da banda de transição*, dado que a largura da banda de transição é proporcional a $1/N$, então quando $N \rightarrow \infty$, a resposta em frequência de $H(e^{j\omega})$ será igual à resposta em frequência do filtro ideal desejado $H_D(e^{j\omega})$. Mas computacionalmente é inviável que o comprimento do filtro tenda ao infinito; e
- (b) *Ondulações na banda de passagem e rejeição* de $H(e^{j\omega})$, as quais são geradas pelas ondulações do lóbulo secundário $W(e^{j\omega})$ e podem ser justificadas pelo fato que dada uma descontinuidade abrupta na frequência de corte ω_c , esta leva a uma distribuição de energia por todo o espectro, por causa do *aliasing*. No estudo das séries de Fourier se determinou que, quando uma função com descontinuidade é aproximada mediante séries de Fourier, aparecem descontinuidades na proximidade da descontinuidade. Quando o número de termos aumenta ($N \rightarrow \infty$), o nível de oscilação vai diminuindo até zero, exceto na descontinuidade, onde aparece uma oscilação aproximada de 8,9% da amplitude da descontinuidade. Este fenômeno é conhecido como *efeito Gibbs*.

O problema de ondulações na banda de passagem pode ser corrigido, em certa medida, se os lóbulos secundários de $W(e^{j\omega})$ tiverem uma tendência a diminuir. Contudo, isso implica em um aumento da largura do lóbulo principal, que leva a aumentar a banda de transição em $H(e^{j\omega})$. Portanto, pode-se reduzir a energia dos lóbulos secundários mediante um custo na largura da banda de transição.

A seguir, apresentam-se as principais funções de janelamento considerando que o filtro ideal $h_D(n)$ é centrado em $n = 0$:

- A. *Janela retangular*: é a função mais simples, mas também a que oferece o pior desempenho da atenuação na faixa de rejeição.

$$w(n) = \begin{cases} 1 & , \quad 0 \leq n \leq N-1 \\ 0 & , \quad \text{outro caso} \end{cases} \quad (3.41)$$

B. *Bartlett*:

$$w(n) = \begin{cases} \frac{2n}{N-1} & , \quad 0 \leq n \leq \frac{N-1}{2} \\ 2 - \frac{2n}{N-1} & , \quad \frac{N-1}{2} < n \leq N-1 \\ 0 & , \quad \text{outro caso} \end{cases} \quad (3.42)$$

C. *Hann*:

$$w(n) = \begin{cases} 0,5 \left[1 - \cos \left(\frac{2\pi n}{N-1} \right) \right] & , \quad 0 \leq n \leq N-1 \\ 0 & , \quad \text{outro caso} \end{cases} \quad (3.43)$$

D. *Hamming*:

$$w(n) = \begin{cases} 0,54 - 0,46 \cos \left(\frac{2\pi n}{N-1} \right) & , \quad 0 \leq n \leq N-1 \\ 0 & , \quad \text{outro caso} \end{cases} \quad (3.44)$$

E. *Blackman*:

$$w(n) = \begin{cases} 0,42 - 0,5 \cos \left(\frac{2\pi n}{N-1} \right) + 0,08 \cos \left(\frac{4\pi n}{N-1} \right) & , \quad 0 \leq n \leq N-1 \\ 0 & , \quad \text{outro caso} \end{cases} \quad (3.45)$$

F. *Kaiser*:

$$w(n) = \frac{I_0 \left[\beta \sqrt{1 - \left(1 - \frac{2n}{N-1} \right)^2} \right]}{I_0[\beta]} , \quad 0 \leq n \leq N-1 \quad (3.46)$$

onde β controla a mínima atenuação na faixa de rejeição e $I_0[\cdot]$ é a função de Bessel modificada de grau zero, dada por:

$$I_0(x) = 1 + \sum_{k=0}^{\infty} \left[\frac{(x/2)^k}{k!} \right]^2 \quad (3.47)$$

onde o comprimento do filtro está definido pela equação (3.48).

$$N \simeq \frac{A_s - 7,95}{2,285\Delta\omega} \quad (3.48)$$

Na Figura 3.9 apresentam-se as formas de vários tipos de janelas, para $N = 60$. Nas Figuras 3.10 e 3.11 apresenta-se o projeto de um filtro passa-baixas com frequência de corte $\omega_c = \pi/3$ para janelas de comprimento $N = 60$.

Figura 3.9: Formas de várias funções de janelas.

Figura 3.10: Filtro passa-baixas pelo método da janela no domínio da frequência, para $\omega_c = \pi/3$ e $N = 60$.

Figura 3.11: Filtro passa-baixas pelo método da janela (dB), com $\omega_c = \pi/3$ e $N = 60$.

Na Tabela 3.1 [31] apresenta-se um resumo das funções janela e suas características com respeito à largura da faixa de transição $\Delta\omega$, ao comprimento do filtro N e à mínima faixa de atenuação.

Tabela 3.1: Resumo das características da função janela $w(n)$.

Nome da janela	Largura na banda de transição $\Delta\omega$		Atenuação na faixa de rejeição
	Aproximação	Valor exato	
Retangular	$\frac{4\pi}{N}$	$\frac{1,8\pi}{N}$	21 dB
Bartlett	$\frac{8\pi}{N}$	$\frac{6,1\pi}{N}$	25 dB
Hann	$\frac{8\pi}{N}$	$\frac{6,2\pi}{N}$	44 dB
Hamming	$\frac{8\pi}{N}$	$\frac{6,6\pi}{N}$	53 dB
Blackman	$\frac{12\pi}{N}$	$\frac{11\pi}{N}$	74 dB

3.1.5 Projeto por amostragem em frequência

É o projeto mais simples e fácil de implementar, mas também não possui muito controle sobre o erro de aproximação nas frequências que não foram amostradas.

O método consiste em determinar os coeficientes do filtro $H(e^{j\omega})$, dada a amostragem da resposta em frequência de um filtro desejado $D(e^{j\omega})$. Ao longo deste estudo, o filtro desejado vai ser o filtro ideal.

Se as amostras de frequência são de igual espaçamento, a transformada discreta de Fourier (DFT) pode ser usada. A DFT de uma resposta ao impulso dá a amostragem da resposta em frequência e a transformada inversa de Fourier (IDFT) das amostras de uma resposta em frequência desejada dá a resposta ao impulso.

Seja a transformada discreta de Fourier da resposta ao impulso definida por:

$$C(k) = \sum_{n=0}^{N-1} h(n)e^{-j2\pi nk/N} \quad (3.49)$$

onde $k = 0, 1, \dots, N-1$.

Por comparação entre (3.3) e (3.49), se obtém:

$$C(k) = H(e^{j\omega}) \Big|_{\omega=2\pi k/N} = H\left(\frac{2\pi k}{N}\right) \quad (3.50)$$

Para N amostras da resposta em frequência de igual comprimento de (3.50), os coeficientes do filtro FIR de comprimento N são dados pela IDFT e definidos por:

$$h(n) = \frac{1}{N} \sum_{k=0}^{N-1} C(k) \cdot e^{j2\pi nk/N} \quad (3.51)$$

Quando $H(e^{j\omega})$ é de fase linear, a equação (3.51) é simplificada pelas equações (3.26) e (3.27) para os tipos 1 e 2 de um filtro FIR de fase linear.

A frequência discretizada em N amostras de igual comprimento está definida por:

$$\omega = \frac{2\pi k}{N}, \quad k = 0, 1, \dots, N-1 \quad (3.52)$$

1. *Filtro tipo 1:*

Substituindo-se (3.52) em (3.17), resulta:

$$A(k) = \sum_{n=0}^{M-1} 2h(n) \cdot \cos\left(\frac{2\pi(M-n)k}{N}\right) + h(M) \quad (3.53)$$

A resposta ao impulso $h(n)$ é calculada mediante a IDTF de (3.53), isto resulta em:

$$\begin{aligned} h(n) &= \frac{1}{N} \sum_{k=0}^{N-1} e^{-j2\pi Mk/N} A(k) e^{j2\pi nk/N} \\ h(n) &= \frac{1}{N} \sum_{k=0}^{N-1} A(k) e^{-j2\pi(n-M)k/N} \end{aligned} \quad (3.54)$$

Dado que $A(k) = A(n-k)$, (3.54) resulta em:

$$h(n) = \frac{1}{N} \left[A(0) + \sum_{k=1}^M 2A(k) \cdot \cos\left(\frac{2\pi(n-M)k}{N}\right) \right] \quad (3.55)$$

2. *Filtro tipo 2:*

Substituindo-se (3.52) em (3.19), resulta em:

$$A(k) = \sum_{n=0}^{N/2-1} 2h(n) \cdot \cos\left(\frac{2\pi(M-n)k}{N}\right) \quad (3.56)$$

A resposta ao impulso $h(n)$ é calculada mediante a IDTF de (3.56), isto resulta em:

$$\begin{aligned} h(n) &= \frac{1}{N} \sum_{k=0}^{N-1} A(k) e^{-j2\pi Mk/N} \cdot A(k) e^{j2\pi nk/N} \\ h(n) &= \frac{1}{N} \sum_{k=0}^{N-1} A(k) e^{-j2\pi(n-M)k/N} \end{aligned} \quad (3.57)$$

Dado que $A(k) = A(n-k)$, (3.57) resulta em:

$$h(n) = \frac{1}{N} \left[A(0) + \sum_{k=1}^{N/2-1} 2A(k) \cdot \cos\left(\frac{2\pi(n-M)k}{N}\right) \right] \quad (3.58)$$

Note que, dada a simetria par (3.15), só é necessário calcular M amostras.

Então, a resposta ao impulso unitário, que é o vetor dos coeficientes do filtro FIR $H(e^{j\omega})$, que se aproxima de um filtro desejado (filtro ideal passa-baixa) está determinada pelas equações (3.55) e (3.58).

A Figura 3.12 mostra a amostragem do filtro ideal desejado $H_D(e^{j\omega})$ e as Figuras 3.13 e 3.14 mostram os filtros aproximados pelo método de amostragem em frequência (tipos 1 e 2).

Figura 3.12: Amostragem do filtro ideal com $N = 31$ e $f_c = 0,25\text{Hz}$.

Das Figuras 3.13 e 3.14 se observa que:

- O erro de aproximação nas frequências amostradas é zero.
- O erro de aproximação nas outras frequências vai depender de quão abrupta é a mudança da resposta em frequência, o pior caso é um degrau.
- O erro é maior perto das fronteiras das bandas e menor dentro das bandas.

Figura 3.13: Resposta em frequência aproximada pelo método de amostragem em frequência com $N = 31$, $f_c = 0,25\text{Hz}$.

Figura 3.14: Resposta em frequência aproximada pelo método de amostragem em frequência com $N = 30$ e $f_c = 0,25\text{Hz}$.

3.1.6 Projeto por mínimo erro quadrático

O método do mínimo erro quadrático (LSE) busca determinar os pesos da resposta impulsiva de um filtro digital de fase linear, baseado na minimização do quadrado do erro, que foi definido na equação (3.37). O erro de aproximação é a diferença entre a resposta em frequência desejada $D(e^{j\omega})$ e a resposta em frequência que se pretende projetar $H(e^{j\omega})$. Assim, a discretização do erro de aproximação (3.36) é definida por:

$$E = \sum_{k=0}^{L-1} |H(\omega_k) - D(\omega_k)|^2 \quad (3.59)$$

onde $k = 0, 1, 2, \dots, L - 1$.

A resposta em frequência é amostrada em L períodos de igual comprimento, sendo:

$$\omega_k = \frac{2\pi k}{L} \quad (3.60)$$

Restringindo-se o problema para filtros lineares, o erro de aproximação pode ser calculado em função da amplitude que se pretende projetar $A(e^{j\omega})$ e a amplitude desejada $A_d(e^{j\omega})$. Redefinindo-se (3.59), resulta em:

$$E = \sum_{k=0}^{L-1} \left| A\left(\frac{2\pi k}{L}\right) - A_d\left(\frac{2\pi k}{L}\right) \right|^2 \quad (3.61)$$

Para um filtro de tipo 1 e comprimento do filtro ímpar, a amplitude de $H(e^{j\omega})$ (vide (3.18)) com sua frequência amostrada por (3.60), resulta em A_k :

$$A_k = \sum_{n=0}^{M-1} 2h(n) \cos\left(\frac{2\pi(M-n)k}{L}\right) \quad (3.62)$$

onde $k = 1, 2, \dots, L - 1$.

A equação (3.62) pode ser escrita em sua forma matricial como:

$$A_d = \mathbf{A}h$$

onde:

$$\begin{aligned}
 h &= [h_0, h_1, \dots, h_{N-1}]^T \\
 A_d &= [A_d(\omega_0), A_d(\omega_1), \dots, A_d(\omega_{L-1})]^T \\
 \mathbf{A} &= \begin{bmatrix} 1 & \cos(\frac{2\pi(M-0)}{L}) & \dots & \cos(\frac{2\pi(M-0)(L-1)}{L}) \\ 1 & \cos(\frac{2\pi(M-1)}{L}) & \dots & \cos(\frac{2\pi(M-1)(L-1)}{L}) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cos(\frac{2\pi(M-N+1)}{L}) & \dots & \cos(\frac{2\pi(M-N+1)(L-1)}{L}) \end{bmatrix}^T
 \end{aligned}$$

onde h é um vetor coluna com N elementos, A_d é o vetor coluna da amplitude da resposta em frequência desejada, o qual tem L elementos e $\mathbf{A}$ é a matriz de cossenos de dimensão $N \times L$.

O problema do método de minimização do erro quadrático (3.61) pode ser definido da seguinte forma:

$$\min_h \boldsymbol{\epsilon}^T \boldsymbol{\epsilon} = \min_h |\mathbf{A}h - A_d|^2 \quad (3.63)$$

$$\begin{aligned}
 \boldsymbol{\epsilon}^T \boldsymbol{\epsilon} &= (h^T \mathbf{A}^T - A_d^T)(\mathbf{A}h - A_d) \\
 &= h^T \mathbf{A}^T \mathbf{A}h - h^T \mathbf{A}^T A_d - A_d^T \mathbf{A}h + A_d^T A_d \\
 &= h^T \mathbf{A}^T \mathbf{A}h - 2h^T \mathbf{A}^T A_d + A_d^T A_d
 \end{aligned}$$

Dado que se deseja minimizar $\boldsymbol{\epsilon}^T \boldsymbol{\epsilon}$, determina-se seu gradiente e iguala-se a zero. Com isso, obtém-se um mínimo global:

$$\nabla_h \{\boldsymbol{\epsilon}^T \boldsymbol{\epsilon}\} = 2\mathbf{A}^T \mathbf{A}h^* - 2\mathbf{A}^T A_d = 0 \quad (3.64)$$

Isolando h^* da equação (3.64), se obtém a resposta ao pulso que minimiza o erro (3.61). O termo h^* é o vetor dos coeficientes do filtro projetado e é definido por:

$$h^* = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T A_d \quad (3.65)$$

Portanto, h^* é o mínimo global da função de erro $|\mathbf{A}h - A_d|^2$, sendo esta a resposta ao impulso da frequência ótima h^* , tal que $\mathbf{A}h^*$ tem a amplitude da resposta em frequência o mais próxima ao comportamento de um filtro ideal passa-baixa.

Para um filtro do tipo 2 (vide equação (3.19)) o procedimento é o mesmo.

3.1.7 Projeto pelo método de Chebyshev

O método de Chebyshev, chamado também “método de projeto de filtros ótimos de ondulações constantes”, procura projetar filtros que permitam distribuir uniformemente o erro, assim, o filtro possui um comprimento menor de coeficiente N .

O projeto do filtro é proposto como um critério de projeto ótimo, no sentido que o erro de aproximação entre a resposta de frequência ideal e a projetada, seja distribuído uniformemente nas faixas de passagem e rejeição, minimizando o erro máximo em cada uma delas.

O problema está definido por:

$$\min_P \{ \|E(\omega)\|_\infty \} = \min_P \left\{ \max_{\omega \in F} \{ |E(\omega)| \} \right\}, \quad (3.66)$$

onde F é o conjunto de faixas de frequências prescritas.

Parks e McClellan [30] resolveram o problema da equação (3.66) aplicando o teorema da alternância (Teorema 3.1).

Teorema 3.1 (Teorema da alternância)

Se $P(\omega)$ é uma combinação linear de $(L+1)$ funções cosseno, isto é,

$$P(\omega) = \sum_{n=0}^L p(n) \cos(\omega n), \quad (3.67)$$

a condição necessária e suficiente para $P(\omega)$ que seja a aproximação de Chebyshev para uma função $D(\omega)$ contínua em F , um subconjunto compacto de $[0, \pi]$, é que a função-erro $E(\omega)$ apresente, no mínimo, $(L+2)$ frequências com extremos em F . Em outras palavras, devem existir pelo menos $(L+2)$ pontos ω_k em F , onde $\omega_0 < \omega_1 < \dots < \omega_{L+1}$, tais que

$$E(\omega_k) = -E(\omega_{k+1}), \quad k = 0, 1, \dots, L \quad (3.68)$$

e

$$|E(\omega_k)| = \max_{\omega \in F} \{ |E(\omega)| \}, \quad k = 0, 1, \dots, L+1. \quad (3.69)$$

Em [7], descreve-se brevemente o algoritmo de trocas de Remez, que dá solução ao problema de aproximação de Chebyshev, mediante a procura das frequências dos extremos de $E(\omega)$, descrita nos seguintes passos:

- (i) Atribua uma estimativa inicial às frequências dos extremos $\omega_0, \omega_1, \dots, \omega_{L+1}$ selecionando $(L+2)$ frequências igualmente espaçadas nas faixas especificadas para o filtro desejado.
- (ii) Encontre $P(\omega_k)$ e δ tais que

$$W_q(\omega_k)(D_q(\omega_k) - P(\omega_k)) = (-1)^k \delta, \quad \text{para } k = 0, 1, \dots, L+1. \quad (3.70)$$

A equação (3.70) pode ser escrita em forma matricial e admite uma solução analítica (vide [37]). Uma abordagem mais eficiente calcula δ por:

$$\delta = \frac{a_0 D_q(\omega_0) + a_1 D_q(\omega_1) + \dots + a_{L+1} D_q(\omega_{L+1})}{\frac{a_0}{W_q(\omega_0)} + \frac{a_1}{W_q(\omega_1)} + \dots + \frac{(-1)^{L+1} a_{L+1}}{W_q(\omega_{L+1})}}, \quad (3.71)$$

onde

$$a_k = \prod_{i=0, i \neq k}^{L+1} \frac{1}{\cos(\omega_k) - \cos(\omega_i)} \quad (3.72)$$

- (iii) Use o interpolador de Lagrange na forma baricêntrica para $P(\omega)$, isto é,

$$P(\omega) = \begin{cases} c_k & , \quad \omega = \omega_k \in \{\omega_0, \omega_1, \dots, \omega_L\} \\ \frac{\sum_{k=0}^L \frac{\beta_k}{\cos(\omega) - \cos(\omega_k)} c_k}{\sum_{k=0}^L \frac{\beta_k}{\cos(\omega) - \cos(\omega_k)}} & , \quad \omega \neq \omega_0, \omega_1, \dots, \omega_L \end{cases} \quad (3.73)$$

onde

$$c_k = D_q(\omega_k) - (-1)^k \frac{\delta}{W_q(\omega_k)} \quad (3.74)$$

$$\beta_k = \prod_{i=0, i \neq k}^L \frac{1}{\cos(\omega_k) - \cos(\omega_i)} = a_k (\cos(\omega_k) - \cos(\omega_{L+1})) \quad (3.75)$$

- (iv) Avaliar $|E(\omega)|$ em um conjunto denso de frequências. Se $|E(\omega)| \leq |\delta|$ para todas as frequências do conjunto, a solução ótima foi resolvida; vá para o próximo passo. Se $|E(\omega)| > |\delta|$ para algumas frequências, um novo conjunto

de candidatos a extremos tem que ser escolhido para picos de $|E(\omega)|$. Dessa maneira, força-se δ a crescer e a convergir para seu limite superior. Se houver mais que $(L + 2)$ picos em $E(\omega)$, mantenha as localizações dos $(L + 2)$ picos de $|E(\omega)|$ com maiores valores, certificando-se que as extremidades das faixas sejam sempre mantidas e retorne ao passo (II).

- (i) Como $P(\omega)$ é uma soma de $(L + 1)$ cossenos com frequências variando de zero a L , então é também uma soma de $(2L + 1)$ exponenciais complexas com frequências variando de $-L$ a L . Então, $p(n)$ pode ser recuperado amostrando-se $P(\omega)$ em $2L + 1$ frequências igualmente espaçadas $\omega = 2\pi n/(2L + 1)$, para $n = 0, 1, \dots, 2L$ e calculando-se sua IDTF.

O algoritmo descrito foi implementado em [25].

Nas Figuras 3.15 e 3.16 apresenta-se um filtro projetado pelo método de Chebyshev com as seguintes especificações: frequência de corte de $f_c = 25$ Hz; faixa de transição de $\Delta f = 5$ Hz; ondulações na faixa de passagem e rejeição de $\delta_1 = 0,1$ e $\delta_2 = 0,01$ respectivamente; e comprimento do filtro de $N = 126$.

Figura 3.15: Resposta em frequência do filtro de Chebyshev.

Figura 3.16: Resposta em frequência do filtro de Chebyshev (dB).

3.2 Normalização

A normalização de dados é uma técnica muito usada em diferentes áreas, principalmente na área de Engenharia Biomédica, onde se precisa processar dados como: ritmos cardíacos, temperatura do corpo, pressão sanguínea, entre outros; os quais apresentam faixas muito variadas uns dos outros. Isso, quando se precisa processar os dados ou procurar uma correlação entre duas ou mais variáveis, gera problemas numéricos ou computacionais, geralmente pelo fato do arredondamento.

Em [47] apresentam-se sete modelos de normalização de dados: normalização por subtração da média, normalização pelo valor máximo de cada variável, normalização pelo desvio padrão, normalização padrão, normalização por um único valor máximo, normalização logarítmica e normalização de probabilidade de massa total. Destes sete modelos, só os primeiros quatro vão ser apresentados neste estudo, por sua aplicação em identificação de sistemas.

Além dos modelos anteriores, apresentam-se mais dois modelos: normalização pelo valor máximo sem média e normalização em faixas iguais.

3.2.1 Normalização por subtração da média

Neste tipo de normalização, procura-se tirar a média μ do conjunto de dados de cada variável. A média da nova faixa de dados normalizados é zero e define-se na equação (3.76) o vetor de dados normalizados.

$$V_n(i) = X_n(i) - \mu \quad (3.76)$$

onde V_n é o vetor de dados normalizados, X_n é o vetor de dados originais e μ é a média de X_n , definida por:

$$\mu = \frac{\sum_{i=1}^N X_n(i)}{N} \quad (3.77)$$

Nesta técnica, as faixas dos dados normalizados dos sinais continuam tendo amplitudes discrepantes, porém ficam em torno de zero.

3.2.2 Normalização pelo valor máximo

Neste tipo de normalização, procura-se que os dados normalizados fiquem em uma faixa que tem como valor máximo a unidade. Isso acontece porque o conjunto de dados de cada variável é dividido pelo valor máximo dessa mesma variável. Dessa forma, pode-se ter um conjunto de variáveis e cada um deles com uma faixa similar. A equação (3.78) define este tipo de normalização como:

$$V_n(i) = \frac{X_n(i)}{\max(X_n(i))} \quad (3.78)$$

3.2.3 Normalização pelo desvio padrão

Neste tipo de normalização, a faixa de dados normalizados é proporcional à dispersão de dados, que está definida pelo desvio padrão σ . Os dados normalizados definem-se como:

$$V_n(i) = \frac{X_n(i)}{\sigma} \quad (3.79)$$

onde

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_n(i) - \mu)^2} \quad (3.80)$$

Nesta técnica, o nível de compressão de dados tem relação com o nível de dispersão para cada variável de forma independente e cada uma delas com um valor diferente, por causa disto as faixas dos conjuntos de dados normalizados ainda estão misturadas.

3.2.4 Normalização padrão

Neste tipo de normalização, tira-se a média μ (equação (3.77)) da faixa de dados reais e divide-se pelo desvio padrão σ (equação (3.80)). Dessa forma, os dados normalizados não só são proporcionais à dispersão dos dados de cada variável de forma independente, mas também encontram-se em torno de zero. A normalização está definida pela seguinte equação:

$$V_n(i) = \frac{X_n(i) - \mu}{\sigma} \quad (3.81)$$

Nesta técnica, o nível de compressão dos dados tem relação com o nível de dispersão para cada variável de forma independente. Por causa disto, as faixas dos conjuntos de dados normalizados ainda estão misturadas, mas encontram-se em torno do zero.

3.2.5 Normalização pelo valor máximo sem média

Neste tipo de normalização, procura-se que os dados normalizados estejam em uma faixa que tem como valor máximo a unidade, mas com as faixas dos conjuntos de dados em torno do zero. A equação (3.82) define este tipo de normalização como:

$$V_n(i) = \frac{X_n(i) - \mu}{\max(|X_n(i) - \mu|)} \quad (3.82)$$

Neste tipo de normalização, as faixas dos conjuntos de dados normalizados encontram-se limitados por:

- +1 quando a diferença entre o valor máximo e a média é maior que a diferença entre o valor mínimo e a média; e
- -1 quando a diferença entre o valor máximo e a média é menor que a diferença entre o valor mínimo e a média.

3.2.6 Normalização em faixas iguais

Neste tipo de normalização, procura-se que os dados normalizados estejam em uma mesma faixa, que pode ser ± 1 , $\pm 0,5$ ou $\pm Z$, correspondente ao valor máximo e mínimo, como se pode ver na Figura 3.17.

Figura 3.17: Normalização de dados em faixas iguais.

Na Figura 3.17 pode-se ver a equação de uma linha reta, que interpola a faixa do conjunto de dados limitados pelo valor máximo D_{max} e o valor mínimo D_{min} , para uma nova faixa de dados normalizados proporcionalmente entre os valores de $+Z$ e $-Z$. A equação da linha reta está definida por:

$$Y = A + BX \quad (3.83)$$

onde

$$A = -Z \left[\frac{D_{max} + D_{min}}{D_{max} - D_{min}} \right]$$

$$B = \frac{2Z}{D_{max} - D_{min}}$$

Na equação (3.83) os parâmetros A e B permitem ajustar a faixa do conjunto de dados não-normalizados X proporcionalmente para um conjunto de dados normalizados Y . Neste tipo de normalização os dados normalizados não necessariamente estão distribuídos em torno do zero, isto porque a média não é nula para a maioria dos casos.

3.3 Reamostragem

3.3.1 Método de autocovariância

Nesta seção, apresenta-se um critério para determinar o período de reamostragem, quando a constante de tempo dominante τ não é conhecida. Para isto o critério faz uma análise da autocovariância [1].

Seja $y(t)$ o sinal de trabalho do sistema real no domínio do tempo e $y^*(k)$ o registro superamostrado, isto é, em que o intervalo de amostragem é muito pequeno. O primeiro objetivo é determinar um $\Delta \in \mathbb{N}$ tal que: $y(k) = y^*(\Delta k)$. Para fazer isso tem que se verificar o grau de correlação (redundância) entre observações adjacentes do sinal $y^*(k)$.

Calculam-se as funções de covariância definidas na equação (3.84), para quantificar os efeitos gerados pela sobreamostragem do sinal $y^*(k)$ em uma função linear e uma não-linear:

$$r_{y^*}(\tau) = E[(y^*(k) - \overline{y^*(k)})(y^*(k - \tau) - \overline{y^*(k)})]$$

$$r_{y^{*2}}(\tau) = E[(y^{*2}(k) - \overline{y^{*2}(k)})(y^{*2}(k - \tau) - \overline{y^{*2}(k)})], \quad (3.84)$$

onde $E[\cdot]$ faz referência à esperança matemática. Considera-se o sinal $y^*(k)$ ergódico, portanto substitui-se a esperança matemática pela média temporal.

O objetivo de se usar $r_{y^{*2'}}(\tau)$, além de $r_{y^*}(\tau)$, é poder detectar algumas correlações não-lineares que porventura estejam nos dados. A escolha da taxa de decimação Δ pode ser feita da seguinte maneira: dado o sinal superamostrado $y^*(k)$, determinam-se as funções na equação (3.84) e seus primeiros mínimos, τ_{y^*} e $\tau_{y^{*2'}}$ respectivamente. O menor desses mínimos passará a ser o valor de trabalho, isto é:

$$\tau_m^* = \min[\tau_{y^*}, \tau_{y^{*2'}}] \quad (3.85)$$

onde τ_m^* é medido em número de atrasos. A escolha de decimação Δ deve ser tal que as funções de autocovariância do sinal decimado $y(k) = y^*(\Delta k)$ satisfaçam:

$$10 \leq \tau_m \leq 20 \quad (3.86)$$

onde τ_m é definido para o sinal decimado $y(t)$ de maneira análoga a τ_m^* para o sinal original $y^*(\Delta \cdot t)$, de modo que $\tau_m = \tau_m^*/\Delta$. Os limites inferior e superior da equação (3.86) podem ser relaxados para 5 e 25 respectivamente, estes dados foram obtidos empiricamente [1].

Nota-se que as funções de autocovariância de $y(t)$ e $y^*(t)$ são idênticas, com exceção do fator de escala Δ , então por semelhança, pode-se inferir que:

$$10 \cdot \Delta \leq \tau_m^* \leq 20 \cdot \Delta \quad (3.87)$$

Em resumo, τ_m^* é obtido das funções de autocovariância mediante as equações (3.84) e (3.85), logo ao ser substituído na equação (3.87) pode-se obter Δ . Finalmente, o período de reamostragem T_s é igual a Δ vezes o período de superamostragem do sinal.

3.3.2 Método pela constante do tempo

Em [22], o autor faz uma avaliação da influência do período de amostragem com respeito à variância do parâmetro estimado $\hat{a}_N$ de um sistema SISO de primeiro ordem para duas condições, tanto para $T_s = 1/\tau$ quanto para T_s muito grande. A relação entre a variância e o período de amostragem se mostra na Figura 3.18.

Figura 3.18: Variância de $\hat{a}_N$ plotado como intervalo do período de amostragem $T_s(\omega_0 = 1/\tau)$ (1) e T_s muito grande (2).

Da Figura 3.18 conclui-se que:

1. A seleção ótima do período de amostragem está em torno da constante do tempo, isto é:

$$T_s \approx \tau \quad (3.88)$$

2. Uma seleção muito pequena do período de amostragem é melhor que uma seleção muito grande.

3.3.3 Método de décimo do tempo de acomodação

Em [44], os autores recomendam como uma regra prática para o cálculo do período de amostragem T_s , considerar 10% do tempo de acomodação ou de estabilização da resposta ao degrau, isto é:

$$T_s = \frac{t_s}{10} \quad (3.89)$$

Deve-se ter em conta, que é muito pior usar um período de amostragem muito grande do que muito pequeno [22].

3.3.4 Método de décimo da constante de tempo

Em [11], o autor recomenda como uma regra prática para o cálculo do período de amostragem T_s , considerar 10% da constante do tempo, isto é:

$$T_s = \frac{\tau}{10} \quad (3.90)$$

Capítulo 4

Resultados e discussões

Neste capítulo pretende-se avaliar o desempenho da metodologia de pré-processamento de dados (capítulo 3) na identificação de sistemas, apresentada no capítulo anterior. Primeiramente, vai-se avaliar o desempenho da filtragem, normalização e reamostragem dos dados coletados da Planta Piloto de Vazão com o intuito de analisar a influência no índice de ajuste do modelo ao sistema real. Logo, avalia-se o desempenho do pré-processamento de dados em um ambiente totalmente simulado com o intuito de analisar outras condições de sinais que não foram avaliadas na identificação da Planta Piloto de Vazão. Além disso, questões relacionadas com a análise qualitativa também são abordadas.

4.1 Planta Piloto de Vazão

Nesta seção vai-se descrever, testar e identificar a Planta Piloto de Vazão, instalada no Laboratório de Controle de Processos Industriais do Departamento de Engenharia de Telecomunicações e Controle da Escola Politécnica de USP.

O objetivo do teste é a coleta de dados para seu posterior pré-processamento, visando realizar uma análise da influência do pré-processamento na identificação de sistemas. A coleta de dados, o pré-processamento de sinais e a identificação de sistemas são feitos usando o Matlab 2012b.

4.1.1 Descrição da Planta Piloto de Vazão

O objetivo da Planta Piloto de Vazão é permitir o estudo e testes de: identificação, controle e supervisão; além disto também permite o estudo e testes do comportamento do posicionador da válvula de controle. Mas neste estudo vai-se fazer uso da planta piloto para realizar testes de identificação de sistemas, mediante o uso do cartão de aquisição de dados da *National Instruments* como *interface* para gerar sinais de controle e coletar dados do medidor de vazão.

A Planta Piloto de Vazão por sua natureza é ruidosa. Isto pelo fato que o fluxo de água ao passar pela tubulação, que apresenta rugosidade em sua parte interior, gera atrito. Também, por que na tubulação apresentam-se reduções, cotovelos, equipamento de controle e medição que geram perdas de carga localizadas. Mas principalmente, pela dinâmica da bomba ao gerar a propulsão da água.

De forma geral, a Planta Piloto de Vazão está composta por: uma bomba centrífuga, um *driver* de velocidade que regula a rotação da bomba, uma válvula de controle que regula a vazão na linha e finalmente o medidor de vazão (medição indireta por diferencial de pressão na placa de orifício) que apresenta ruído na medição (vide na Figura 4.1). O objetivo principal da planta piloto é regular a vazão da água na tubulação.

A Figura 4.1 mostra o diagrama de tubulação e instrumentação (P&ID) da Planta Piloto de Vazão, que é composto por:

- Um tanque de armazenamento, que abastece água ao sistema e permite seu refluxo;
- um *driver* de velocidade SD (*Speed Driver*) que regula a rotação da bomba, gerando a pressão na linha;
- uma válvula pneumática de controle FV que permite modular a passagem da água no sistema. O sinal de excitação do sistema será inserido através dela;

Figura 4.1: Diagrama P&ID da Planta Piloto de Vazão.

- um medidor local de vazão FI (*Flow Indicator*);
- um medidor de vazão com indicação local FIT (*Flow Indicator Transmitter*);
e
- o controlador de vazão FIC (*Flow Indicator Controller*) que recebe o sinal do medidor de vazão e gera o sinal de controle para o atuador da válvula de controle. O controle da planta de vazão pode trabalhar em modo manual ou automático.

4.1.2 Coleta de dados

Durante a coleta de dados, um ponto a ter em conta é que a válvula de controle FV é do tipo ar para fechar, isto é, quando o sinal de controle é $u(t) = 0\%$ a válvula está totalmente aberta e quando o sinal de controle é $u(t) = 100\%$ a válvula de controle está totalmente fechada.

Assim também, o sinal de controle $u(t)$ e a variável do processo $PV(t)$ (sinal de vazão) foram normalizados. Utilizou-se um escalonamento proporcional de 0% para o valor mínimo e de 100% para o valor máximo de cada sinal.

Para determinar a não-linearidade do processo, gerou-se sinais de controle $u(t) = 0\%, 5\%, 10\%, \dots, 100\%$ e coletaram-se os valores em estado estável da saída do sistema. A Figura 4.2 ilustra a característica não linear do processo.

Figura 4.2: Não-linearidade da Planta Piloto de Vazão.

No experimento vai-se considerar, de forma arbitrária, que o ponto de operação é de 50% da variável de processo $PV(t)$. Da Figura 4.2 pode-se ver, que a região linear em torno do ponto de operação está na faixa de 40% – 80% da $PV(t)$ e na faixa de 15% – 35% do sinal de controle $u(t)$. O sinal gerado para excitar o sistema dentro da região linear e os valores de saída do sistema são vistos na Figura 4.3.

Figura 4.3: Dados coletados de entrada e saída. O sinal de entrada é um sinal PRBS.

Na Figura 4.3 apresentam-se os dados coletados da planta de vazão, que foi excitada com um sinal PRBS na válvula de controle com um tempo de amostragem T_s de 0,01 segundos (sobre-amostrado) e com uma duração do experimento de 363 segundos. O sinal de saída do processo é ruidoso, isto pode ser devido ao ruído gerado no instrumento de medição, ao ruído gerado pelas interferências eletromagnéticas ou ao ruído presente no processo de vazão.

4.1.3 Avaliação de filtragem na identificação

Quando pretende-se realizar a filtragem do sinal, a primeira questão a resolver é: qual é a frequência de corte do filtro que atenuará o ruído?

Para resolver isso, tem que se avaliar o sinal no domínio da frequência, especificamente a potência espectral do sinal. Nela pode-se distinguir as componentes de frequência do sinal do processo (normalmente baixas frequências) e as componentes do ruído. Na Figura 4.4 apresenta-se o espectro de potência da medição de vazão da Planta Piloto.

Figura 4.4: Espectro em potência (dB) do sinal $PV(t)$.

Na Figura 4.4, se observa que a componente principal do processo apresenta uma frequência de $f_0 = 0,4$ Hz. Para determinar a frequência de corte do filtro a ser projetado se fizeram testes com frequências de: $f_c = 0,5; 0,6; \dots 1,5$ Hz. Como critério de escolha se considerou a menor frequência que não distorce o sinal. Portanto, neste caso de estudo, considera-se a frequência de corte do filtro passa-baixa igual a $f_c = 1$ Hz e sua frequência de corte normalizada está definida por:

$$f_{c_n} = \frac{f_c}{F_s/2} = \frac{1}{100/2} = 0,02$$

onde F_s é a frequência de amostragem do sinal. Finalmente, a frequência de corte normalizada é:

$$\begin{aligned} \omega_{c_n} &= 2\pi f_{c_n} \\ \omega_{c_n} &= 2\pi 0,02 \\ \omega_{c_n} &= 0,0628 \end{aligned} \tag{4.1}$$

Considerou-se uma faixa de transição do filtro igual a 20% da frequência de corte ω_{c_n} , sendo este o ponto médio da faixa de transição. Portanto, $\omega_p = 0,05655$, $\omega_s = 0,0628$ e $\Delta\omega = 0,0126$.

Definiram-se as bandas de oscilação permitidas na faixa de passagem e de rejeição em $\delta_1 = 0,01$ e $\delta_2 = 0,01$. Mediante as equações (3.34) e (3.35), determinaram-se

suas correspondentes oscilações permitidas em dB para $R_p = 0,1737$ e $A_s = 40,09$.

Finalmente, o comprimento do filtro N foi definido para cada um dos métodos da janela. Para isso, considerou-se o valor de $\Delta\omega$ igual ao “valor exato” da Tabela 3.1. O comprimento pelo método da janela de Kaiser, amostragem de frequência, mínimo erro quadrático e Chebyshev foram calculados mediante a equação (3.48). Na Tabela 4.1 apresenta-se o comprimento do filtro para cada um dos métodos.

Tabela 4.1: Comprimento N do filtro passa-baixa projetado para $\Delta\omega = 0,0126$.

Nome da janela	Comprimento do Filtro
Retangular	451
Bartlett	1526
Hann	1551
Hamming	1651
Blackman	2751
Kaiser	1121
Amostragem em frequência	1121
Mínimo erro quadrático	1121
Chebyshev	1121

A resposta em frequência dos filtros projetados mediante: o método da janela (3.41) a (3.46), o método da amostragem em frequências (3.55) e (3.58), o método de mínimo erro quadrático (3.65) e o método de Chebyshev (3.66) são ilustrados nas Figuras 4.5 a 4.13; respectivamente.

Figura 4.5: Resposta em frequência do método da janela retangular.

Figura 4.6: Resposta em frequência do método da janela de Bartlett.

Figura 4.7: Resposta em frequência do método da janela de Hann.

Figura 4.8: Resposta em frequência do método da janela de Hamming.

Figura 4.9: Resposta em frequência do método da janela de Blackman.

Figura 4.10: Resposta em frequência do método da janela de Kaiser.

Figura 4.11: Resposta em frequência do método de amostragem em frequência.

Figura 4.12: Resposta em frequência do método de mínimo erro quadrático.

Figura 4.13: Resposta em frequência do método de Chebyshev.

O sinal de saída $PV(t)$ filtrado para os diferentes métodos de filtragem, em uma janela de tempo de 100s, se mostra nas Figuras 4.14 a 4.22.

Figura 4.14: Sinal filtrado pelo método da janela retangular.

Figura 4.15: Sinal filtrado pelo método da janela de Bartlett.

Figura 4.16: Sinal filtrado pelo método da janela de Hann.

Figura 4.17: Sinal filtrado pelo método da janela de Hamming.

Figura 4.18: Sinal filtrado pelo método da janela de Blackman.

Figura 4.19: Sinal filtrado pelo método da janela de Kaiser.

Figura 4.20: Sinal filtrado pelo método de amostragem em frequência.

Figura 4.21: Sinal filtrado pelo método de mínimo erro quadrático.

Figura 4.22: Sinal filtrado pelo método de Chebyshev.

Dividiram-se os dados sobre-amostrados e filtrados em duas partes, a primeira metade para a etapa de identificação e a outra metade para a etapa de validação. Na etapa de avaliação, consideram-se dois casos: validação com dados filtrados e validação com dados sem filtrar (dados originais). Além disso, os dados usados durante o processo de identificação foram normalizados pelo método de normalização por subtração da média.

Escolheu-se a estrutura ARMAX por ser um dos modelos mais usados nos processos industriais. No experimento não foi gerada nenhuma perturbação. O processo de vazão caracteriza-se por ter baixa ordem, por isso, testou-se para ordens $n_a = n_b = n_c = a$, onde $a = 2, 3, 4$. Ademais, o parâmetro n_k foi fixado em zero ($n_k = 0$), dado que no processo de vazão considera-se que não há tempo morto.

Filtragem e identificação de sistemas com dados de validação filtrados

Na validação cruzada foram testadas previsões de 1, 10, 100 e infinitos passos à frente para os dados sem filtrar e para cada um dos filtros projetados. Esta etapa forneceu os índices de ajuste dos modelos (*fit*), que são vistos na Tabela 4.2.

Na Tabela 4.2, observa-se que para previsões 1 passo à frente, os modelos têm um alto índice *fit*. Porém, para previsões mais à frente, o *fit* é muito menor. Também observa-se que o modelo para grau 2 não representa bem a dinâmica do sistema identificado.

Na Tabela 4.2, para o modelo de grau 3, nota-se que não há uma vantagem de se usar dados filtrados na identificação do processo. Contudo, para o modelo de grau 4, observa-se que apresenta um melhor desempenho com um máximo de 3,37% de melhora no índice de ajuste do modelo.

Tabela 4.2: *Fit* de validação cruzada com dados filtrados da Planta Piloto de Vazão para os filtros projetados.

	Nome da janela	Passos à frente			
		1	10	100	inf
Modelo de grau 2	Sem filtro	90,80	90,76	90,36	89,86
	Retangular	99,98	99,77	92,09	86,08
	Bartlett	100,00	99,86	92,41	80,71
	Hann	100,00	99,86	91,78	76,83
	Hamming	100,00	99,86	92,49	81,12
	Blackman	100,00	99,86	91,54	75,84
	Kaiser	100,00	99,86	92,58	81,57
	Amostragem em Frequência	99,99	99,80	92,49	84,01
	Mínimo Erro Quadrático	100,00	99,82	92,30	82,50
	Chebyshev	99,92	99,73	92,15	86,42
Modelo de grau 3	Sem filtro	90,84	90,81	90,51	90,05
	Retangular	99,98	99,85	95,55	94,58
	Bartlett	100,00	99,98	93,95	85,14
	Hann	100,00	99,99	92,54	79,12
	Hamming	100,00	99,98	93,67	83,66
	Blackman	100,00	99,99	92,42	77,64
	Kaiser	100,00	99,97	93,86	85,03
	Amostragem em Frequência	99,99	99,87	94,78	89,50
	Mínimo Erro Quadrático	100,00	99,91	93,66	85,32
	Chebyshev	99,93	99,83	94,86	90,79
Modelo de grau 4	Sem filtro	90,81	90,75	90,42	89,95
	Retangular	99,98	99,85	95,56	92,58
	Bartlett	100,00	99,99	96,10	91,63
	Hann	100,00	100,00	95,08	85,40
	Hamming	100,00	99,99	95,98	90,27
	Blackman	100,00	100,00	95,26	84,45
	Kaiser	100,00	99,99	96,09	91,01
	Amostragem em Frequência	99,99	99,89	96,54	92,51
	Mínimo Erro Quadrático	100,00	99,93	95,45	92,48
	Chebyshev	99,93	99,86	95,88	93,32

Da Tabela 4.2 e considerando a validação para predição com infinitos passos à frente (por ser este o caso mais rigoroso), nota-se que o melhor modelo obtido é com os dados filtrados pelo método de Chebyshev ($fit=92,7576\%$) para um modelo ARMAX de grau 4. O modelo resultante é apresentado na equação (A.32), sendo o mínimo erro quadrático $MSE = 4,777 \cdot 10^{-05}$ e onde os seus parâmetros estimados são:

$$A(q) = 1 - 3,966q^{-1} + 5,903q^{-2} - 3,906q^{-3} + 0,9698q^{-4};$$

$$B(q) = 6,587 \cdot 10^{-05} - 0,0002156q^{-1} + 0,0002343q^{-2} - 8,475 \cdot 10^{-05}q^{-3};$$

$$C(q) = 1 - 3,578q^{-1} + 4,808q^{-2} - 2,878q^{-3} + 0,6481q^{-4}.$$

Filtragem e identificação de sistemas com dados de validação sem filtrar

Nesta sub-subseção dividiu-se os dados sobre-amostrados em duas partes: a primeira metade para a etapa de identificação, nesta etapa os dados foram filtrados; a outra metade para validação cruzada com dados sem filtrar.

Tabela 4.3: *Fit* de validação cruzada com dados sem filtrar da Planta Piloto de Vazão para os filtros projetados.

Nome do filtro	Modelo ARMAX		
	Grau 2	Grau 3	Grau 4
Sem filtro	89,86	90,05	89,95
Retangular	83,52	89,49	89,49
Bartlett	78,76	82,70	87,57
Hann	75,24	77,35	82,91
Hamming	79,16	81,42	86,80
Blackman	74,32	75,99	82,10
Kaiser	79,67	82,50	87,39
Amostragem em Frequência	81,67	86,26	89,49
Mínimo Erro Quadrático	80,43	82,84	88,34
Chebyshev	85,46	89,44	90,65

Na etapa de identificação, escolheu-se a estrutura ARMAX e os dados foram

normalizados pelo método de subtração da média. Na etapa de validação, testou-se o modelo obtido só para infinitos passos à frente por ser o mais rigoroso. Os índices de ajuste do modelo ao sistema (*fit*) são mostrados na Tabela 4.3.

Da Tabela 4.3, pode-se observar que não há uma grande vantagem no *fit* de identificação quando se validam os modelos obtidos com dados sem filtrar. Isto, devido a que os dados sem filtrar apresentam uma grande dispersão com respeito aos dados gerados pelos modelo obtido.

4.1.4 Avaliação de reamostragem na identificação

Para realizar a reamostragem dos dados sobre-amostrados, que foram coletados empiricamente da Planta Piloto de Vazão, deve ter-se em conta: o tempo de sobre-amostragem $\tau_m = 0,01s$, o tempo de acomodação $t_s = 12s$ e a constante de tempo $\tau = 3s$.

No caso de reamostragem pelo método de autocovariância (subseção 3.3.1), se deve ter em conta o primeiro mínimo da função de covariância. Este mínimo é mostrado na Figura 4.23, onde $\tau_m^* = 869$. O fator de escala Δ , apresentado na equação (3.87), está dentro do intervalo:

$$43,45 \leq \Delta \leq 86,90$$

Portanto, vai-se considerar o fator de escala no ponto médio do intervalo, isto é, $\Delta = 65$.

Figura 4.23: Função de covariância normalizada da planta de piloto de vazão.

Testaram-se os métodos de reamostragem e a sua influência na identificação para o modelo ARMAX de segundo, terceiro e quarto grau ($n_a = n_b = n_c = a$, para $a = 2, 3, 4$). Considerou-se o modelo sem retardo de tempo ($n_k = 0$), por ser uma característica típica das plantas de vazão.

Os testes foram feitos tanto para dados sem filtrar quanto para dados filtrados pelo método de Chebyshev (vide subseção 4.1.3). Esta etapa forneceu os índices de ajuste dos modelos (*fit*) obtidos a partir da validação cruzada, vistos nas Tabelas 4.4 e 4.5.

Tabela 4.4: *Fit* de validação cruzada da Planta Piloto de Vazão para os métodos de reamostragem projetados com dados sem filtrar.

		Passos à frente			
		1	10	100	inf
Modelo de grau 2	Dados não-amostrados	90,80	90,76	90,36	89,86
	Método de autocovariância	88,30	86,95	90,14	87,11
	Método de décimo da constante de tempo	90,51	90,00	90,83	89,90
	Método de décimo do tempo de acomodação	86,10	82,51	82,42	82,42
	Método pela constante do tempo	76,03	75,22	73,83	73,83
Modelo de grau 3	Dados não-amostrados	90,84	90,81	90,51	90,05
	Método de autocovariância	88,44	87,12	90,13	87,35
	Método de décimo da constante de tempo	90,64	90,20	91,03	90,16
	Método de décimo do tempo de acomodação	86,31	82,60	82,56	82,56
	Método pela constante do tempo	75,74	76,74	74,38	74,38
Modelo de grau 4	Dados não-amostrados	90,81	90,75	90,42	89,95
	Método de autocovariância	88,36	87,11	90,18	87,36
	Método de décimo da constante de tempo	90,61	90,24	91,02	90,14
	Método de décimo do tempo de acomodação	86,34	84,15	91,32	82,92
	Método pela constante do tempo	75,18	76,52	74,34	74,34

Tabela 4.5: *Fit* de validação cruzada da Planta Piloto de Vazão para os métodos de reamostragem projetados com dados filtrados pelo método de Chebyshev.

		Passos à frente			
		1	10	100	inf
Modelo de grau 2	Dados não-amostrados	99,92	99,73	92,15	86,42
	Método de autocovariância	94,89	90,66	92,46	90,68
	Método de décimo da constante de tempo	98,31	91,71	92,62	91,72
	Método de décimo do tempo de acomodação	90,98	84,91	84,88	84,88
	Método pela constante do tempo	75,84	74,82	73,66	73,66
Modelo de grau 3	Dados não-amostrados	99,93	99,83	94,86	90,79
	Método de autocovariância	95,56	91,05	92,43	91,17
	Método de décimo da constante de tempo	98,97	94,34	94,67	94,20
	Método de décimo do tempo de acomodação	91,06	85,59	92,04	85,11
	Método pela constante do tempo	75,53	76,03	74,10	74,10
Modelo de grau 4	Dados não-amostrados	99,93	99,86	95,88	93,32
	Método de autocovariância	95,54	91,07	92,49	91,21
	Método de décimo da constante de tempo	99,07	94,53	94,91	94,29
	Método de décimo do tempo de acomodação	90,88	85,30	92,93	85,28
	Método pela constante do tempo	76,00	75,51	73,88	73,88

Das Tabelas 4.4 e 4.5, pode-se observar:

- *Avaliação de reamostragem com dados sem filtrar*, da Tabela 4.4, observa-se que não há uma vantagem significativa no índice de ajuste do modelo (*fit*) quando os dados foram reamostrados pelos métodos de autocovariância e de décimo da constante de tempo, comparados a quando não foram reamostrados. Aliás, quando os dados foram reamostrados pelos métodos de décimo do tempo de acomodação e da constante de tempo, o índice de ajuste do modelo apresenta um valor muito menor.

- *Avaliação de reamostragem com dados filtrados pelo método de Chebyshev*, da Tabela 4.5, observa-se que melhora o desempenho do índice de ajuste do modelo (*fit*) para infinitos passos à frente, comparado a quando os dados não foram reamostrados. Para modelo de grau 3, no método de décimo da constante de tempo melhora em 3,41%. Para modelo de grau 4, no método de décimo da constante de tempo melhora em 0,97%. Contudo, nos outros métodos não há nenhuma melhoria, ainda mais, o desempenho piora.

4.1.5 Avaliação de normalização na identificação

Na coleta dos dados empíricos da Planta Piloto de Vazão, tanto o sinal de entrada ($4 - 20$ mA) quanto o sinal de saída ($0 - 10$ l/s), foram escalonados em faixas de $0 - 100\%$.

Os dados coletados sem filtrar foram normalizados pelos métodos apresentados na seção 3.2 e mostrados graficamente nas Figuras 4.24 e 4.25.

Figura 4.24: Normalização de dados sem filtrar da Planta Piloto de Vazão (a).

Figura 4.25: Normalização de dados sem filtrar da Planta Piloto de Vazão (b).

Realizou-se a identificação de sistemas usando os dados normalizados e sem filtrar. Obteve-se os índices de ajuste do modelo ao sistema real (*fit*) mediante a validação cruzada para predições de 1, 10, 100 e infinitos passos à frente. Na Tabela 4.6 se apresenta um resumo do índice *fit* quando as ordens do modelo ARMAX são $n_a = n_b = n_c = a$ e $n_k = 0$ para $a = 2, 3, 4$.

Tabela 4.6: *Fit* de validação cruzada do modelo ARMAX para $n_a = n_b = n_c = a$ e $n_k = 0, a = 2, 3, 4$ com dados sem filtrar.

	Nome da janela	Passos à frente			
		1	10	100	inf
Modelo de grau 2	Dados sem normalizar	85,97	13,33	-1315,83	-31,05
	Normalização por subtração da média	90,51	90,00	90,83	89,90
	Normalização pelo valor máximo	85,97	19,01	-1209,98	-30,07
	Normalização pelo desvio padrão	85,97	15,81	-1157,44	-42,21
	Normalização padrão	90,42	89,68	90,61	89,68
	Normalização pelo valor máximo sem média	90,50	89,97	90,88	89,96
	Normalização em faixas iguais [-0,5 ; 0,5]	90,43	89,62	90,19	89,16
	Normalização pelo escalonamento [0 ; 1]	86,05	19,11	-532,05	-329,55
Modelo de grau 3	Dados sem normalizar	90,42	87,83	90,05	89,11
	Normalização por subtração da média	90,64	90,20	91,03	90,16
	Normalização pelo valor máximo	90,41	89,61	89,96	88,99
	Normalização pelo desvio padrão	90,32	89,37	89,95	89,27
	Normalização padrão	90,54	89,85	90,69	89,81
	Normalização pelo valor máximo sem média	90,62	90,16	91,01	90,16
	Normalização em faixas iguais [-0,5 ; 0,5]	90,56	89,82	90,35	89,44
	Normalização pelo escalonamento [0 ; 1]	86,04	20,34	-494,81	-386,85
Modelo de grau 4	Dados sem normalizar	90,54	89,83	90,07	89,21
	Normalização por subtração da média	90,61	90,24	91,02	90,14
	Normalização pelo valor máximo	90,52	89,76	84,96	89,09
	Normalização pelo desvio padrão	90,42	89,50	89,77	89,32
	Normalização padrão	90,50	89,90	90,68	89,80
	Normalização pelo valor máximo sem média	90,59	90,20	91,00	90,14
	Normalização em faixas iguais [-0,5 ; 0,5]	90,53	89,86	90,34	89,42
	Normalização pelo escalonamento [0 ; 1]	90,54	89,83	90,52	89,17

Os dados coletados e filtrados pelo método de Chebyshev (vide subseção 4.1.3) e reamostrados pelo método do décimo da constante de tempo, foram normalizados pelos métodos apresentados na seção 3.2 e mostrados graficamente na Figura 4.26.

Figura 4.26: Normalização de dados filtrados da Planta Piloto de Vazão.

Realizou-se a identificação de sistemas usando os dados normalizados e filtrados. Obteve-se os índices de ajuste do modelo ao sistema real (*fit*) mediante a validação cruzada para predições de 1, 10, 100 e infinitos passos à frente mostrados na Tabela 4.7.

Tabela 4.7: *Fit* de validação cruzada do modelo ARMAX para $n_a = n_b = n_c = a$ e $n_k = 0$, $a = 2, 3, 4$ com dados filtrados.

	Nome da janela	Passos à frente			
		1	10	100	inf
Modelo de grau 2	Dados sem normalizar	98,26	24,23	-357,12	-31,62
	Normalização por subtração da média	98,31	91,71	92,62	91,72
	Normalização pelo valor máximo	98,26	30,72	-336,85	-28,26
	Normalização pelo desvio padrão	98,26	29,51	-797,20	-44,68
	Normalização padrão	98,31	91,21	91,69	90,69
	Normalização pelo valor máximo sem média	98,31	91,68	92,63	91,74
	Normalização em faixas iguais [-0,5 ; 0,5]	98,31	91,26	91,74	90,62
	Normalização pelo escalonamento [0 ; 1]	98,27	33,83	-176,94	-301,80
Modelo de grau 3	Dados sem normalizar	98,46	76,34	76,75	73,47
	Normalização por subtração da média	98,97	94,34	94,67	94,20
	Normalização pelo valor máximo	98,46	70,37	75,73	72,16
	Normalização pelo desvio padrão	98,46	66,79	80,09	77,80
	Normalização padrão	98,95	93,58	93,79	93,18
	Normalização pelo valor máximo sem média	98,97	94,21	94,57	94,20
	Normalização em faixas iguais [-0,5 ; 0,5]	98,94	93,26	93,01	92,51
	Normalização pelo escalonamento [0 ; 1]	98,46	78,10	74,93	71,06
Modelo de grau 4	Dados sem normalizar	98,95	93,88	94,19	93,52
	Normalização por subtração da média	99,07	94,53	94,91	94,29
	Normalização pelo valor máximo	98,94	92,77	78,59	93,32
	Normalização pelo desvio padrão	98,92	93,03	92,11	92,93
	Normalização padrão	99,05	93,91	94,22	93,49
	Normalização pelo valor máximo sem média	99,07	94,41	94,73	94,29
	Normalização em faixas iguais [-0,5 ; 0,5]	99,05	93,73	93,29	92,58
	Normalização pelo escalonamento [0 ; 1]	98,94	93,66	94,20	93,33

Das Tabelas 4.6 e 4.7, pode-se observar:

- *Avaliação de reamostragem com dados sem filtrar*, da Tabela 4.6.

Observam-se para o modelo de grau 2, valores negativos no índice de ajuste do modelo ao sistema real (*fit*) quando se usam métodos que não subtraem a média do vetor de dados. O modelo de grau 2 não é um bom modelo para representar o sistema.

Observam-se para os modelos de grau 3 e 4, que não há vantagem no desempenho do índice de ajuste do modelo ao sistema real (*fit*) entre os diferentes métodos de normalização, mesmo com os dados sem normalizar.

- *Avaliação de reamostragem com dados filtrados pelo método de Chebyshev*, da Tabela 4.7.

Observa-se para o modelo de grau 2, que existem valores negativos no índice de ajuste do modelo ao sistema real (*fit*) quando se usam métodos que não subtraem a média do vetor de dados. O modelos de grau 2 não é um bom modelo para representar o sistema.

Observa-se para o modelo de grau 3, que os métodos de normalização por subtração da média e normalização pelo valor máximo sem média, apresentam um mesmo e alto desempenho no índice *fit* com um valor de 94,20% para infinitos passos à frente.

Observa-se para o modelo de grau 4, que os métodos de normalização por subtração da média e normalização pelo valor máximo sem média apresentam um mesmo e alto desempenho no índice *fit*, com um valor de 94,29% para infinitos passos à frente.

Discussões

De forma geral, se observa que os melhores resultados foram obtidos quando se realizou a identificação com dados filtrados e normalizados. Os métodos de normalização por subtração da média e normalização pelo valor máximo sem média geram o melhor desempenho no índice de ajuste do modelo ao sistema real e os valores do *fit* obtidos são os mesmos.

Da Tabela 4.5, nota-se que o melhor modelo obtido é com os dados filtrados pelo método de Chebyshev e normalizados pelos métodos que realizam a subtração da média (*fit* = 94,29%) para um modelo ARMAX de grau 4. O modelo resultante é apresentado na equação (A.32), sendo o mínimo erro quadrático $MSE = 0,007357$ e onde os seus parâmetros estimados são:

$$\begin{aligned} A(q) &= 1 - 1,628q^{-1} + 1,075q^{-2} - 0,5907q^{-3} + 0,1923q^{-4}; \\ B(q) &= 0,001198 - 0,01652q^{-1} - 0,03035q^{-2} - 0,01872q^{-3}; \\ C(q) &= 1 + 0,8733q^{-1} - 0,1455q^{-2} - 0,4783q^{-3} + 0,01504q^{-4}. \end{aligned}$$

Pode-se observar que a filtragem de dados melhora o índice de ajuste do modelo ao sistema real até em 3,37% (vide Tabela 4.2). Contudo, o filtro de Chebyshev projetado possui um comprimento $N = 1121$ (vide Tabela 3.1), isto implica um alto número de operações numéricas pelo processador durante a filtragem do sinal.

Os métodos de reamostragem pelo método da autocovariância e pelo método de décimo da constante de tempo, segundo as Tabelas 4.4 e 4.5, geram um melhor desempenho do índice *fit*. Além disso, a vantagem desses métodos de reamostragem não só é que permitem obter um modelo mais representativo do sistema real, mas também permitem obtê-lo com um menor número de operações matemáticas, por conter um menor número de dados na hora de realizar a identificação do sistema.

Os métodos de normalização aplicados à identificação da Planta Piloto de Vazão não apresentam uma melhora significativa no índice de ajuste do modelo ao sistema, isto pode ser devido a que as faixas das variáveis de entrada e saída estão entre 0% e 100%. Contudo, os métodos que subtraem a média (normalização por subtração da média, normalização padrão e normalização pelo valor máximo sem média) mostraram melhor rendimento do índice *fit*.

4.2 Plantas virtuais

Nos processos industriais as principais variáveis de controle são: *nível*, que caracteriza-se por ser um sistema lento (depende da sua capacitância), com presença de ruído e ganho relativamente baixo; *pressão*, que caracteriza-se por ser um sistema rápido, com presença de ruído e ganho tanto alto quanto baixo; *temperatura*, que caracteriza-se por ser um sistema lento, sem ruído e de ganho tanto baixo quanto alto; e *vazão*, que caracteriza-se por ser um sistema rápido, com presença de ruído e ganho tanto alto quanto baixo.

Para avaliar a rapidez da dinâmica de um sistema, se utiliza a constante de tempo (τ). Diz-se que o sistema é rápido se a constante de tempo é menor que um minuto, isto é, $\tau \leq 60\text{s}$. Diz-se que o sistema é lento se a constante de tempo é maior que um minuto, isto é, $\tau > 60\text{s}$.

A maioria dos processos industriais podem ser representados por um modelo de grau dois. Portanto, nesta seção vai-se considerar um sistema linear invariante no tempo (LIT) de grau dois para quatro diferentes tipos de plantas virtuais, que podem representar as características dinâmicas das variáveis de nível, pressão, temperatura e vazão.

Figura 4.27: Simulação de uma planta virtual no Simulink.

Na Figura 4.27 apresenta-se o diagrama de blocos de uma planta simulada no software Simulink. O sinal de entrada do sistema é um sinal PRBS que excita a planta virtual. A função de transferência da planta virtual está composta de um vetor numerador **num(s)** que tem relação com os zeros e um vetor denominador **den(s)** que tem relação com os polos do sistema. O coletor de dados armazena os dados dos sinais de entrada e saída da planta virtual.

A planta virtual da Figura 4.27 vai ser avaliada para quatro diferentes tipos de dinâmicas do sistema: rápida e de baixo ganho; rápida e de alto ganho; lenta e de baixo ganho; e lenta e de alto ganho. Também, para cada um deles vai-se avaliar as técnicas de pré-processamento de dados quando a planta apresenta ruído branco com variância de 2%, 5% e 10% da amplitude de ΔPV .

4.2.1 Planta rápida e de baixo ganho

Vai-se considerar um planta de segundo grau com constante de tempo de $\tau = 3s$ e um tempo de acomodação de $t_s = 9,2s$. Os polos do sistema estão localizados em $p_1 = -0,5$ e $p_2 = -1$. O ganho da planta é $K = 2$.

Figura 4.28: Resposta ao degrau da planta rápida e de baixo ganho.

A Figura 4.28 apresenta a resposta dinâmica da planta virtual. O tempo de simulação e coleta de dados foi 500s com um tempo de super-amostragem de 0,01s. Contudo, na Figura 4.29 se apresentam só os primeiros 100s.

Avaliou-se a densidade espectral de potência (DEP) do sinal de saída da planta (vide Figura 4.30) e se observa que a componente principal do processo apresenta uma frequência de $f_0 = 0,4\text{Hz}$. Portanto, vai-se considerar $f_c = 1\text{Hz}$ como frequência de corte do filtro passa-baixa e sua frequência de corte normalizada está definida

Figura 4.29: Dados coletados da planta rápida e de baixo ganho para (a) ruído com 2% ΔPV , (b) ruído com 5% ΔPV e (c) ruído com 10% ΔPV .

por:

$$f_{c_n} = \frac{f_c}{F_s/2} = \frac{1}{100/2} = 0,02$$

onde F_s é a frequência de amostragem do sinal. Finalmente, a frequência de corte normalizada é:

$$\begin{aligned}\omega_{c_n} &= 2\pi f_{c_n} \\ \omega_{c_n} &= 0,0628\end{aligned}\tag{4.2}$$

Considerou-se uma faixa de transição do filtro igual a 20% da frequência de corte ω_{c_n} , sendo este o ponto médio da faixa de transição. Portanto, $\omega_p = 0,05655$, $\omega_s = 0,0628$ e $\Delta\omega = 0,0126$.

Definiram-se as bandas de oscilação permitidas na faixa de passagem e de rejeição em $\delta_1 = 0,01$ e $\delta_2 = 0,01$. Mediante as equações (3.34) e (3.35), determinaram-se suas correspondentes oscilações permitidas em dB para $R_p = 0,1737$ e $A_s = 40,09$.

Finalmente, o comprimento do filtro N foi definido para cada um dos métodos da janela. Para isso, considerou-se o valor de $\Delta\omega$ igual ao “valor exato” da Tabela

3.1. O comprimento pelo método da janela de Kaiser, amostragem de frequência, mínimo erro quadrático e Chebyshev foram calculados mediante a equação (3.48).

Figura 4.30: Densidade espectral de potência do sinal de saída da planta rápida e de baixo ganho.

O sinal de saída da planta virtual foi filtrado pelos métodos da janela retangular; janela de Bartlett; janela de Hann; janela de Hamming; janela de Blackman; janela de Kaiser; amostragem em frequência, mínimo erro quadrático; e Chebyshev.

Os dados filtrados foram divididos em duas partes iguais: a primeira parte foi usada para o processo de estimação dos parâmetros do modelo ARMAX identificado; a segunda parte foi usada tanto para a validação cruzada com previsões de infinitos passos à frente, quanto para a obtenção do índice de ajuste do modelo ao sistema. Esta etapa forneceu os índices (*fit*) que são vistos na Tabela 4.8.

Tabela 4.8: *Fit* de validação cruzada da planta rápida e de baixo ganho para os filtros projetados.

Nome da janela	Amplitude do ruído		
	2% ΔPV	5% ΔPV	10% ΔPV
Sem filtro	93,61	91,99	87,52
Retangular	92,87	93,07	93,32
Bartlett	91,91	93,18	93,30
Hann	92,37	93,13	93,26
Hamming	92,50	93,12	93,26
Blackman	92,63	93,11	93,27
Kaiser	92,27	93,14	93,27
Amostragem em frequência	93,14	93,24	93,33
Mínimo erro quadrático	93,04	93,09	93,29
Chebyshev	92,58	93,00	93,25

Na Tabela 4.8 se observa que quando o ruído não é muito severo (2% ΔPV) não há uma vantagem entre filtrar ou não os dados. Contudo, quando o ruído é mais severo, o desempenho do índice *fit* para dados sem filtrar cai, mas o desempenho para dados filtrados se mantém.

O melhor desempenho do *fit* para um ruído muito severo é de 93,33% quando os dados foram filtrados pelo método de amostragem em frequência.

Tabela 4.9: *Fit* de validação cruzada da planta rápida e de baixo ganho para os métodos de reamostragem projetados com dados filtrados.

Nome da janela	Sinal com ruído			Dados filtrados
	2% ΔPV	5% ΔPV	10% ΔPV	
Dados não-amostrados	93,61	91,99	87,52	91,45
Método de autocovariância	86,27	85,00	82,09	86,42
Método de décimo da constante de tempo	93,43	91,95	87,33	93,27
Método de décimo do tempo de acomodação	76,62	75,72	74,36	75,67
Método pela constante do tempo	55,25	55,37	54,85	55,20

Na Tabela 4.9 pode-se observar que só a reamostragem pelo método de décimo da constante de tempo apresenta um bom desempenho do índice de ajuste ao sistema (*fit*). Os outros métodos de reamostragem apresentam um pobre desempenho.

Os dados filtrados e reamostrados pelo método de décimo da constante de tempo foram normalizados. Estes dados foram usados na identificação de sistemas e seu índice de ajuste ao sistema real é fornecido na Tabela 4.10.

Tabela 4.10: *Fit* de validação cruzada da planta rápida e de baixo ganho para os métodos de normalização.

Nome da janela	Amplitude do ruído		
	$2\%\Delta PV$	$5\%\Delta PV$	$10\%\Delta PV$
Dados sem normalizar	93,18	91,80	87,22
Normalização por subtração da média	93,43	91,95	87,33
Normalização pelo valor máximo	93,18	91,80	87,22
Normalização pelo desvio padrão	93,18	91,80	87,22
Normalização padrão	93,43	91,95	87,33
Normalização pelo valor máximo sem média	93,43	91,95	87,33
Normalização em faixas iguais [-0,5 ; 0,5]	93,29	91,89	87,35
Normalização pelo escalonamento [0 ; 1]	93,29	91,74	86,79

Na Tabela 4.10 pode-se observar que não existe uma vantagem entre usar algum dos métodos de normalização. Isto devido a que tanto a variável de entrada quanto a variável de saída do sistema estão quase na mesma faixa de valores.

4.2.2 Planta rápida e de ganho alto

Vai-se considerar um planta de segundo grau com constante de tempo de $\tau = 3s$ e um tempo de acomodação de $t_s = 9,2s$. Os polos do sistema estão localizados em $p_1 = -0,5$ e $p_2 = -1$. O ganho da planta é $K = 50$.

Figura 4.31: Resposta ao degrau da planta rápida e de ganho alto.

A Figura 4.31 apresenta a resposta dinâmica da planta virtual. O tempo de simulação e coleta de dados foi 500s com um tempo de super-amostragem de 0,01s. Contudo, na Figura 4.32 se apresentam só os primeiros 100s.

Figura 4.32: Dados coletados da planta rápida e de ganho alto para (a) ruído com 2% ΔPV , (b) ruído com 5% ΔPV e (c) ruído com 10% ΔPV .

Dado que os polos do sistema são os mesmos que na subseção 4.2.2, então a dinâmica da planta virtual e a análise da frequência de corte normalizada também é a mesma. Portanto, a frequência de corte do filtro passa-baixa é $f_c = 1$ Hz. Na Figura 4.33 apresenta-se a densidade espectral de potência.

Figura 4.33: Densidade espectral de potência do sinal de saída da planta rápida e de baixo ganho.

O sinal de saída da planta virtual foi filtrado pelos métodos da janela retangular; janela de Bartlett; janela de Hann; janela de Hamming; janela de Blackman; janela de Kaiser; amostragem em frequência, mínimo erro quadrático; e Chebyshev.

Os dados filtrados foram divididos em duas partes iguais: a primeira parte foi usada para o processo de estimação dos parâmetros do modelo ARMAX identificado; a segunda parte foi usada tanto para a validação cruzada com predições de infinitos passos à frente, quanto para a obtenção do índice de ajuste do modelo ao sistema. Esta etapa forneceu os índices (*fit*) que são vistos na Tabela 4.11.

Tabela 4.11: *Fit* de validação cruzada da planta rápida e de ganho alto para os filtros projetados.

Nome da janela	Amplitude do ruído		
	2% ΔPV	5% ΔPV	10% ΔPV
Sem filtro	92,64	91,07	86,28
Retangular	92,35	92,35	92,23
Bartlett	92,25	89,41	92,25
Hann	92,30	92,36	92,22
Hamming	92,31	92,36	92,22
Blackman	92,32	92,36	92,22
Kaiser	92,29	92,36	92,22
Amostragem em frequência	92,42	92,43	92,30
Mínimo erro quadrático	92,38	92,37	92,26
Chebyshev	92,26	92,28	92,38

Na Tabela 4.11 se observa que quando o ruído não é muito severo (2% ΔPV) não há uma vantagem entre filtrar ou não os dados. Contudo, quando o ruído é mais severo, o desempenho do índice *fit* para dados sem filtrar cai, mas o desempenho para dados filtrados se mantém.

O melhor desempenho do *fit* para um ruído muito severo é de 92,38% quando os dados foram filtrados pelo método de Chebyshev.

Tabela 4.12: *Fit* de validação cruzada da planta rápida e de ganho alto para os métodos de reamostragem projetados com dados filtrados.

Nome da janela	Sinal com ruído			Dados filtrados
	2% ΔPV	5% ΔPV	10% ΔPV	
Dados não-amostrados	97,87	94,30	87,95	91,31
Método de autocovariância	86,68	85,31	83,02	86,93
Método de décimo da constante de tempo	92,84	91,32	86,47	92,43
Método de décimo do tempo de acomodação	72,73	71,66	70,53	72,37
Método pela constante do tempo	52,31	52,31	50,50	51,44

Na Tabela 4.12 pode-se observar que só a reamostragem pelo método de décimo da constante de tempo apresenta um bom desempenho do índice de ajuste ao sistema (*fit*). Os outros métodos de reamostragem apresentam um pobre desempenho.

Os dados filtrados e reamostrados pelo método de décimo da constante de tempo foram normalizados. Estes dados foram usados na identificação de sistemas e seu índice de ajuste ao sistema real é fornecido na Tabela 4.13.

Tabela 4.13: *Fit* de validação cruzada da planta rápida e de ganho alto para os métodos de normalização.

Nome da janela	Passos à frente		
	$2\%\Delta PV$	$5\%\Delta PV$	$10\%\Delta PV$
Dados sem normalizar	92,88	91,34	86,48
Normalização por subtração da média	92,87	91,32	86,47
Normalização pelo valor máximo	92,88	91,34	86,48
Normalização pelo desvio padrão	92,88	91,34	86,48
Normalização padrão	92,84	91,32	86,47
Normalização pelo valor máximo sem média	92,84	91,32	86,47
Normalização em faixas iguais [-0,5 ; 0,5]	92,21	91,16	86,46
Normalização pelo escalonamento [0 ; 1]	92,88	91,33	86,22

Na Tabela 4.13 pode-se observar que não existe uma vantagem entre usar algum dos métodos de normalização.

Discussões

Os casos para quando a planta virtual é lenta ($\tau \gg 1$ min) com ganho baixo e alto não vão ser abordados, pois os sistemas com dinâmica lenta apresentam menos complexidade, tanto no processo de pré-processamento de dados quanto na identificação de sistemas.

Quando os sistemas apresentam uma dinâmica lenta, a componente de sua frequência natural é mais próxima ao zero. Portanto, a separação das componentes de frequências do sistema com respeito às componentes de altas frequências (como do ruído) é mais diferenciada.

Pode-se perceber que quando o sistema é uma planta rápida e com um baixo ganho, as técnicas de normalização e reamostragem não geram uma melhora significativa do índice de ajuste do modelo (*fit*).

Pode-se perceber que quanto mais severo é o ruído, aumenta a vantagem de se usar as técnicas de filtragem.

Capítulo 5

Conclusões e sugestões para trabalhos futuros

Os aspectos mais relevantes referentes às análises realizadas foram apontados à medida que os resultados eram apresentados. Portanto, neste capítulo pretende-se apenas enfatizar as principais conclusões. Em seguida, também são sugeridas algumas possibilidades para dar continuidade ao tema abordado neste trabalho.

5.1 Conclusões gerais

O projeto de filtros digitais tipos FIR pode ser aplicado na filtragem de sinais como uma etapa prévia à identificação de sistemas, dado que tanto o acondicionamento de dados quanto a estimação dos parâmetros do modelo são realizados em modo *off-line*. De fato, o filtro FIR de fase linear acrescenta um atraso de $M = \frac{N-1}{2}$ tempo de amostragem. Dependendo do comprimento N e consequentemente do retardo de tempo do filtro, é que o filtro projetado pode ser ou não aplicado no controle de processos.

Na hora de projetar um filtro digital, deve-se procurar guardar uma relação entre o custo computacional e o benefício no índice de ajuste ao modelo. A diferença no índice *fit* entre os diferentes métodos de filtragem não é significativa, mas o comprimento do filtro projetado às vezes é muito significativo.

A normalização de dados não apresenta uma vantagem significativa para faixas semelhantes de valores das variáveis de entrada e saída. Contudo, quando o ruído é severo, o desempenho do índice de ajuste do modelo ao sistema, melhora. Os métodos de normalização que subtraem a média (normalização por subtração da média, normalização padrão e pelo valor máximo sem média) apresentam os mesmos resultados, mas também, proporcionam o melhor desempenho do índice de ajuste do modelo ao sistema (*fit*).

A técnica de reamostragem mediante o método de autocovariância e o método de décimo da constante de tempo, proporcionam os melhores resultados. Além disso, ao precisar de menos amostras durante a estimativa dos parâmetros do modelo, o modelo é mais genérico e também mais eficiente quanto ao número de operações matemáticas requeridas.

5.2 Sugestões para trabalhos futuros

A seguir são sugeridas algumas perspectivas de extensão deste trabalho:

- Em muitos casos os filtros projetados tipo FIR, apresentam um comprimento N muito grande, que deriva em um excessivo custo computacional. Nos filtros tipo IIR o comprimento do filtro é muito menor e mais fácil de se implementar mas acrescenta sua dinâmica ao sinal filtrado. Se os sinais de saída e de entrada são filtrados, mesmo que o sinal de entrada (PRBS ou GBN) que excita o sistema não apresente ruído, a dinâmica da planta a ser identificada não varia. Portanto, se sugere realizar uma análise de pré-processamento de dados mediante filtros digitais tipo IIR;
- O conceito de pré-processamento de dados pode ser aplicado não só para identificação de processos lineares invariantes no tempo (LIT), mas também para identificação de sistemas não-lineares.

-
- Por fim, uma extensão natural deste trabalho seria a aplicação dos programas desenvolvidos a dados de outros processos reais.

Referências Bibliográficas

- [1] Aguirre, L. A. *Introdução à identificação de sistemas-técnicas lineares e não-lineares aplicadas a sistemas reais*. UFMG, Belo Horizonte, 3ª ed., 2007.
- [2] Bozic, S. M.; Chance, R.J. *Digital filters and signal processing in electronic engineering: theory, applications, architecture, code*. Elsevier, Chichester, 1998.
- [3] Chen, B.; Zhu, Y.; Hu, J.; Principe, J. C. *System parameter identification: information criteria and algorithms*. Newnes, 2013.
- [4] Chen, W-K. *Passive, active, and digital filters*. CRC Press, Chicago, 3ª ed., 2009. ISBN 978-1-4200-5885-7.
- [5] Close, C. M.; Frederick, D. K.; Newell, J. C. *Modeling and analysis of dynamic systems*. John Wiley & Sons, New York, 3rd ed., 2001.
- [6] Corinthios, M. *Signals, systems, transforms, and digital signal processing with MATLAB*. CRC Press, 2009.
- [7] Diniz, P. S.; Silva, E. A.; Netto, S. L. *Processamento digital de sinais: projeto e análise de sistemas*. Bookman Editora, 2nd ed., 2014.
- [8] Dorf, R.; Bishop, R. *Modern control systems*. Prentice-Hall, Upper Saddle River, 11th ed., 2008. ISBN 978-013-206710-2.
- [9] Godfrey, K.R.; Barker, H.A.; Tucker, A.J. Comparison of perturbation signals for linear system identification in the frequency domain. In *Control Theory and Applications, IEE Proceedings*, volume 146, p. 535–548, 1999.
- [10] Goodwin, G.; Graebe, S.; Salgado, M. *Control system design*. Prentice-Hall, Valparaiso, Chile, 2001.

-
- [11] Gustavsson, I. Survey of applications of identification in chemical and physical processes. *Automatica*, 11(1):3–24, 1975.
 - [12] Haykin, S.; Van Veen, B. *Signals and systems*. John Wiley & Sons, 2003.
 - [13] Isermann, R.; Münchhof, M. *Identification of dynamic systems: an introduction with applications*. Springer-Verlag Berlin Heidelberg, 2011.
 - [14] Jackson, E. *A user's guide to principal components*. John Wiley and Sons, 1991. ISBN 978-0-4716-2267-3.
 - [15] Jackson, L. *Digital filtering and signal processing with MATLAB exercises*. Kluwer Academic Publishers, Rhode Island, USA, 3rd ed., 1996.
 - [16] Keesman, K. J. *System identification: an introduction*. Springer, New York, 2011.
 - [17] King, R. A.; Ahmadi, M.; Gorgui-Naguib, R; Kwabwe, A.A.; Azimi-Sadjadi, M. *Digital filtering in one and two dimensions: design and applications*. Perseus Publishing, 1989.
 - [18] Kuo, B.; Golnaraghi, F. *Automatic control systems*. John Wiley and Sons New York, 9th ed., 2010. ISBN 978-0470-04896-2.
 - [19] Lathi, B. P.; Green, R. *Essentials of digital signal processing*. Cambridge University Press, New York, USA, 2014.
 - [20] Leondes, C.T. *System Identification and Adaptive Control: Advances in theory and applications*. Academic Press, 1987.
 - [21] Ljung, L.; Söderström, T. *Theory and practice of recursive identification*. MIT Press,, Cambridge, MA, 1983.
 - [22] Ljung, Lennart. *System identification: theory for the user*. Prentice Hall, New Jersey, 2nd ed., 1999.
 - [23] Ljung, Lennart. Convergence analysis of parametric identification methods. *Automatic Control, IEEE Transactions on*, 23(5):770–783, 1978.

- [24] Madisetti, V. *Digital signal processing fundamentals*. CRC press, Georgia, USA, 2nd ed., 2010. ISBN 978-1-4200-4606-9.
- [25] McClellan, J. H.; Parks, T. W.; Rabiner, L. A computer program for designing optimum fir linear phase digital filters. *IEEE Transactions on Audio and Electroacoustics*, 21(6):506–526, 1973.
- [26] Nikiforuk, P.N.; Gupta, M.M. A bibliography on the properties, generation and control system applications of shift-register sequences. *International Journal of Control*, 9(2):217–234, 1969.
- [27] Ogata, K.; Yang, Y. *Modern control engineering*. Prentice-Hall, New Jersey, USA, 5th ed., 2010. ISBN 978-013-615673-4.
- [28] Oppenheim, A. V.; Schafer, R. W.; Buck, J. R.; others, . *Discrete-time signal processing*. Prentice-hall, Upper Saddle River, USA, 3rd ed., 2009.
- [29] Paraskevopoulos, P.N. *Modern Control Engineering*. CRC Press, 2002. ISBN 0-8247-8981-4.
- [30] Parks, T; McClellan, J. Chebyshev approximation for nonrecursive digital filters with linear phase. *Circuit Theory, IEEE Transactions on*, 19(2):189–194, 1972.
- [31] Parks, T. W.; Burrus, C. S. *Digital filter design*. Wiley-Interscience, USA, 1987.
- [32] Pintelon, R.; Schoukens, J. *System identification: a frequency domain approach*. John Wiley & Sons, 2012.
- [33] Proakis, J.; Ingle, V. *Digital Signal Processing Using Matlab*. Cengage Learning, 3rd ed., 2012.
- [34] Proakis, J. G.; Manolakis, D. *Digital signal processing: principles, algorithms, and applications*. Prentice Hall, 4th ed., 2006.
- [35] Rabiner, L.; Schafer, R. W. Recursive and nonrecursive realizations of digital filters designed by frequency sampling techniques. *Audio and Electroacoustics, IEEE Transactions*, 19(3):200–207, 1971.

- [36] Rabiner, L. R.; Gold, B.; McGonegal, C. A. An approach to the approximation problem for nonrecursive digital filters. *IEEE Transactions on Audio and Electroacoustics*, AU-18(2):83–106, Junho 1970.
- [37] Rabiner, Lawrence; McClellan, James H; Parks, Thomas W. Fir digital filter design techniques using weighted chebyshev approximation. *Proceedings of the IEEE*, 63(4):595–610, 1975.
- [38] Rodríguez, O.W.R.; Garcia, C. Projeto e implementação de filtros digitais não-recursivos para identificação de sistemas em processos industriais. *In: Anais do XX International Congress of Electrical, Electronic, Systems Engineering, INTERCON*, I:42–46, 2013.
- [39] Rodríguez, O.W.R.; Garcia, C. Filtragem e identificação de sistemas em processos industriais. *In: Anais do XVI Latin American Congress of Automatic Control, CLCA*, 2014.
- [40] Rorabaugh, C. B. *Digital filter designer's handbook: featuring C routines*. McGraw-Hill Professional, 1993.
- [41] Schlichthärle, D. *Digital filters: basics and design*. Springer, 2nd ed., 2011.
- [42] Sharma, S. *Applied multivariate techniques*. John Wiley, 1995. ISBN 978-0-4713-1064-8.
- [43] Shenoi, B. A. *Introduction to digital signal processing and filter design*. John Wiley & Sons, 2005.
- [44] Söderström, T.; Stoica, P. G. *System identification*. Prentice Hall, 1989.
- [45] Stubberud, P. A. A computationally efficient technique for designing frequency sampling filters. *IEEE Transactions on Circuits and Systems*, 44:45–50, Janeiro 1997.
- [46] Tulleken, H. J. A. F. Generalized binary noise test-signal concept for improved identification-experiment design. *Automatica*, 26(1):37–49, 1990.
- [47] Uragun, B.; Rajan, R. Developing an appropriate data normalization method. *In: 10th International Conference on Machine Learning and Applications*, 2: 195 – 199, Novembro 2011.

-
- [48] Wang, J.; Ren, L.; Yang, P.; Gao, N. Frequency sampling based fir digital filter design. *In: Computer-Aided Industrial Design and Conceptual Design, CAIDCD '06. 7th International Conference on*, p. 1 – 4, Novembro 2006.
- [49] Williams, D.; Madisetti, V. *Digital signal processing handbook*. CRC Press, Inc., 1997.
- [50] Winder, S. *Analog and digital filter design*. Newnes, 2002.
- [51] Zadeh, L. From circuit theory to system theory. *Proceedings of the IRE*, 50(5): 856–865, 1962.
- [52] Zhu, Y aStoicaP. . *Identification of multivariable industrial processes: for simulation, diagnosis and control*. Springer-Verlag New York, Inc., 1993.
- [53] Zhu, Y. *Multivariable system identification for process control*. Elsevier, 2001.

Apêndice A

Identificação de sistemas

A identificação de sistemas procura relacionar os sinais de entrada e saída de sistemas dinâmicos. Em [51], se define a identificação de sistemas como “a determinação, com base na observação de entrada e de saída, de um sistema dentro de uma classe específica de sistemas para o qual o sistema em teste, é equivalente”. Segundo [23], se define que “o procedimento de identificação é baseado em três entidades: os dados, o conjunto de modelos e o critério de avaliação. Identificação, então, é selecionar o modelo de um conjunto de modelos que descreve os melhores dados, de acordo com o critério”. Segundo [16], a identificação de sistemas é considerada como “um modelo aproximado para aplicações específicas sobre a base de dados observados e o conhecimento a priori de um sistema”. Segundo [44], “a identificação de sistemas é o campo do modelamento de sistemas dinâmicos a partir de dados experimentais”.

Portanto, a identificação de sistemas tem como objetivo: determinar um modelo matemático de um sistema dinâmico de interesse (seja um sistema linear ou não-linear), mediante o processamento de dados empíricos coletados (variáveis medidas das entradas e saídas) e uma estrutura de modelo do sistema (assume-se que a estrutura seja conhecida).

O objetivo de obter modelos matemáticos de sistemas dinâmicos varia dependendo de seu uso nas diferentes linhas de pesquisa ou de aplicações prática. Por exemplo, na área de ciências econômicas é usado para prever o comportamento das ações de uma empresa na bolsa de valores e avaliar o risco do investimento;

na aeronáutica é usado para gerar simuladores de treinamento de aviões e avaliar a capacidade e instinto do piloto em diferentes cenários; nos processos industriais são usados tanto para a sintonia de controladores clássicos, quanto para prever o comportamento do processo e gerar o sinal de controle nos controladores avançados, como o controlador preditivo baseado em modelo (MPC).

A identificação de sistemas pode ser classificada de diversos modos:

- Dependendo do tipo de modelo obtido:
 - *Métodos não-paramétricos*, que permitem obter modelos representados em curvas ou Tabelas. Alguns destes métodos são: análise da resposta transitória, análise da resposta em frequência, análise da correlação espectral, entre outros.
 - *Métodos paramétricos*, que permitem obter o modelo em função de um vetor de parâmetros. Os métodos mais usados no cálculo dos parâmetros são: o método dos mínimos quadrados e o método das variáveis instrumentais.
- Dependendo do momento da coleta de dados:
 - *Identificação off-line (em batelada)*, quando o cálculo dos parâmetros do modelo é obtido só depois de haver coletado os dados das entradas e das saídas do sistema durante o momento que o processo foi excitado nas suas entradas. Este método de identificação normalmente é usado quando é um sistema invariante no tempo ou quando as mudanças dos parâmetros são mínimas em um longo tempo.
 - *Identificação on-line (recursiva)*, quando o modelo dinâmico do processo é continuamente gerado, isto é, os parâmetros do modelo vão se atualizando continuamente, conforme se vai coletando os novos dados das entradas e das saídas. Normalmente, este tipo de identificação é requerido para sintonizar controladores adaptativos *on-line*.

A.1 Modelos Matemáticos

No controle de processos (sistemas dinâmicos) é fundamental conhecer o comportamento dinâmico do processo quando é submetido a diferentes condições. Por exemplo, supondo que o operador de uma planta industrial conhece empiricamente o comportamento da saída do processo para diferentes sinais de entrada, então ele poderia mexer nos sinais de entrada para levar o sinal de saída do processo para um valor de referência proposto. Mas, em muitos casos este tipo de experiência é muito complexo e às vezes impossível.

De forma geral, o modelo pode ser definido como uma ferramenta que permite prever o comportamento de um sistema, sem a necessidade de experimentar sobre ele. Nos processos industriais o modelo pode ser separado em duas partes: o modelo da planta e o modelo das perturbações.

A.1.1 Tipos de modelo

Um sistema físico pode ser modelado de diferentes formas. Uma classificação baseada no grau do formalismo matemático que possui o modelo [21], apresenta-se a seguir:

1. *Modelos mentais ou intuitivos*, o conhecimento do comportamento do sistema é resumido de forma não-analítica na mente de uma pessoa, isto é, carece de formalismo matemático. Por exemplo, para dirigir uma bicicleta se requer um conhecimento intuitivo da força que tem que se aplicar no pedal, em relação ao aumento da velocidade, e da força nos freios, em relação à diminuição da velocidade.
2. *Modelos não-paramétricos ou gráficos*, as propriedades do sistema encontram-se resumidas em um gráfico ou uma Tabela. Por exemplo: um sistema linear pode ser descrito mediante sua resposta ao impulso ou ao degrau ou pelos diagramas de Bode ou de Nyquist.

3. *Modelos paramétricos ou matemáticos*, em aplicações de maior complexidade, pode ser necessário utilizar modelos que descrevam as relações entre as variáveis do sistema mediante expressões matemáticas, como podem ser as equações diferenciais (para sistemas contínuos) ou de diferenças (para sistemas discretos). Em função do tipo de sistema e da representação matemática utilizada, os sistemas podem se classificar em [5, 44]:

- *Determinístico ou estocástico*, um modelo é determinístico quando a relação entre as entradas e as saídas pode ser expressa de forma exata mediante uma equação. Em contraste, um modelo estocástico contém termos aleatórios que tornam impossível um cálculo exato da saída. Estes últimos, definem-se mediante conceitos probabilísticos ou estatísticos.
- *Dinâmico ou estático*, um modelo é estático quando a saída depende unicamente da entrada no instante de tempo atual. Este tipo de modelo não possui “memória”, daí que o efeito da entrada é instantâneo. Em um modelo dinâmico, as saídas evoluem com o tempo depois da aplicação de uma determinada entrada. Neste último tipo de modelo, para conhecer o valor atual da saída é necessário conhecer o tempo transcorrido desde a aplicação da entrada.
- *Contínuo ou discreto*, os modelos de sistemas contínuos trabalham com sinais contínuos e são representadas mediante equações diferenciais. Os modelos de sistemas discretos trabalham com sinais que foram amostrados, sendo descritos mediante equações de diferenças.

Portanto, todo modelo matemático relaciona as entradas e as saídas do sistema mediante equações. Daí que os modelos matemáticos normalmente sejam chamados de modelos paramétricos, dado que podem ser definidos por uma estrutura e um número finito de parâmetros.

A.1.2 Métodos de obtenção de modelos

Para o estudo científico do comportamento dos sistemas de um fenômeno físico ou de um processo industrial, normalmente são escolhidos os modelos matemáticos ou paramétricos. Em termos gerais, existem duas formas de obter o modelo matemático de um sistema [44]:

- *Modelagem teórica*, refere-se à obtenção de modelos partindo de leis físicas, econômicas, etc.

Por exemplo: para obter o modelo do nível de um tanque de água, realiza-se um balanço de massa, em que a vazão mássica total de entrada é igual à massa total de saída mais a massa acumulada no tanque. Para obter o modelo de um trocador de calor, realiza-se um balanço de energia. No primeiro exemplo, considera-se que a influência dos efeitos da temperatura são desprezíveis, no segundo exemplo, dependendo do caso de interesse, pode-se considerar os efeitos da acumulação de partículas sólidas na tubulação desprezíveis.

Portanto, quando se realiza a modelagem teórica, dependendo do caso de interesse do estudo, pode-se desprezar alguns efeitos do sistema que não são significativos na aplicação. Daí que nenhum modelo matemático pode representar exatamente as características de um sistema real.

- *Identificação de sistemas*, refere-se à obtenção de modelos dinâmicos a partir de dados experimentais. Isto inclui tanto o conjunto de dados para a identificação dos parâmetros do modelo, quanto para o conjunto de dados de validação do modelo.

O modelo obtido representa só as características dos efeitos do sistema que aconteceram durante a coleta de dados. Isto é, se durante a coleta de dados não ocorre numa perturbação, esta não vai ser representada no modelo obtido. Além disso, os parâmetros do modelo são usados só para obter uma boa representação do comportamento do sistema e não possuem nenhuma interpretação física.

Na prática, o ideal é fazer uma mistura dos dois métodos. Mediante o uso da identificação de sistemas, obtêm-se modelos bastante exatos, mas quando mais se conhece acerca das leis físicas que regulam o processo, o processo de identificação (na hora de escolher uma estrutura do modelo e o número de parâmetros) é mais facilitado.

A.2 Procedimento de identificação de modelos paramétricos

Ao longo da dissertação vai-se considerar a obtenção de modelos paramétricos de sistemas lineares invariantes no tempo, como o foco do estudo. O modelo paramétrico vai ser representado por um vetor de parâmetros, que é assumido como constante.

Segundo [16, 23, 44, 51], a obtenção do modelo paramétrico mediante a identificação *off-line* possui três componentes principais: dados observados (de entradas e de saídas), conjunto de modelos candidatos e critério de avaliação para a seleção. Estes três componentes regulam diretamente o desempenho da identificação, incluindo a exatidão, o índice de convergência e a complexidade computacional do algoritmo de identificação [22].

Em termos gerais, o processo de identificação compreende os passos a seguir:

1. *Coleta de dados de entrada e saída*, as variáveis de entrada e saída do sistema são amostradas com um período de amostragem definido e coletadas durante um tempo determinado. O objetivo nesta etapa é que o conjunto de dados contenha a maior informação possível da dinâmica do sistema.
2. *Pré-processamento de dados*, os dados coletados geralmente estão acompanhados de ruído, mas também podem estar super-amostrados. Portanto, os dados antes de serem usados são processados para tirar estas imperfeições.

3. *Escolha da estrutura do modelo*, nesta etapa escolhe-se dentre um conjunto de estruturas já prontas, a qual é a mais próxima ao comportamento do sistema a identificar. Precisa-se para isso de conhecimento prévio do sistema. Quando os parâmetros de ajuste do modelo têm uma correspondência e interpretação física denomina-se *caixa cinza*; quando os parâmetros de ajuste do modelo não têm nenhuma relação física é denominada como *caixa preta*.
4. *Estimação dos parâmetros do modelo*, mediante o uso dos dados coletados e o método de mínimos quadrados, se procede a estimar os valores dos parâmetros do modelo.
5. *Validação do modelo*, a avaliação da qualidade do modelo está tipicamente baseada no desempenho dos modelos, quando eles tentam reproduzir os dados medidos. Isto é, a aproximação do critério da medição da diferença (ou semelhança) entre o modelo e o sistema real, e permite determinar quão boa é a estimativa do sistema [3]. Se o modelo for suficientemente bom, então parar o processo; mas se não for, o processo de identificação deverá se repetir e ir para uma etapa anterior.

As causas que podem fazer com que o modelo seja deficiente [22] são:

- o procedimento numérico falhou em encontrar o melhor modelo de acordo com o critério empregado;
- o critério não foi bem escolhido;
- o conjunto de modelos não era apropriado; e
- o conjunto de dados não era suficientemente informativo para prover uma orientação efetiva na seleção dos modelos.

A Figura A.1 ilustra as etapas do procedimento de identificação [11].

Figura A.1: Etapas do procedimento do projeto de identificação.

A.3 Sinal de entrada e saída

Em um projeto de identificação de sistemas, a primeira etapa é a escolha dos sinais de entrada e saída.

Nas aplicações industriais, em que o objetivo da identificação do modelo está relacionada ao controle de processos, os sinais de saída são as variáveis que se deseja controlar e os sinais de entrada são as variáveis manipuladas da planta. Neste caso, considera-se a dinâmica do atuador (válvula de controle ou *driver* de velocidade) e do medidor como parte de processo.

No caso de sistemas MIMO de n variáveis de entrada e m variáveis de saída, para determinar a relação que existe entre a dinâmica de uma variável de saída com respeito a uma ou mais variáveis de entrada, pode-se realizar uma análise de correlação de cada saída por cada entrada do sistema. Além disso, no caso de sistemas lineares invariantes no tempo (LIT) e devido ao princípio de superposição, o modelo do sistema MIMO é equivalente a obter m modelos MISO.

No caso de sistemas SISO, o sinal de entrada no processo de identificação é determinante nas estimativas dos parâmetros do modelo. Contudo, nem todos os processos permitem ser excitados por sinais de entrada padrão. O problema pode ser dividido em três sub-problemas [11]:

- *A forma*, pode ser degrau, senoidal ou onda quadrada.
- *A amplitude*, deve ser tal que a saída do sistema esteja dentro da faixa linear na qual pretende-se identificar (o modelo obtido é um modelo linear invariante no tempo de um processo não-linear, válido para uma região linear escolhida para um ponto de operação).
- *A característica da frequência*, deve promover alterações nas componentes de frequências do processo, isto é, obter a maior informação possível do sistema.

Os sinais de entrada padrão mais frequentemente usados são:

- (a) Degrâu, é um sinal que comuta os valores em um instante de tempo determinado como se mostra na equação (A.1). Este deve ser tal que maximize a mudança de valores na saída, dentro da região linear que corresponde ao ponto de operação que se deseja trabalhar. Fornece o tempo de subida, sobressinal, ganho estático, constante de tempo dominante e o tempo morto. Para gerar o sinal, precisa-se definir a amplitude e o valor inicial da entrada. Geralmente é usado na identificação não-paramétrica.

$$u(t) = \begin{cases} u_f & , \quad t < t_0 \\ u_i & , \quad t \geq t_0 \end{cases} \quad (\text{A.1})$$

onde: u_i é o sinal de entrada inicial, u_f é o sinal de entrada final e t_0 é o instante de tempo em que acontece o chaveamento.

- (b) Sinal binário aleatório (RBS), é um sinal periódico de onda quadrada que comuta somente entre dois valores possíveis. O tempo mínimo de permanência (T_{RBS}) é o mínimo de intervalos de amostragem depois do qual se permite comutar o valor [44]. Em termos práticos, o tempo T_{RBS} está definido pelas seguintes relações [1]:

- Com respeito à menor constante do tempo (τ_{min}) do sistema:

$$\frac{\tau_{min}}{10} \leq T_{RBS} \leq \frac{\tau_{min}}{3} \quad ,$$

neste caso, na identificação de sistemas lineares costuma-se escolher T_{RBS} mais próximo do limite inferior; e

- Com respeito ao intervalo de amostragem (T)

$$3 \cdot T \leq T_{RBS} \leq 5 \cdot T$$

- (c) Sinal pseudo-aleatório binário (PRBS), é um sinal periódico de onda quadrada que comuta entre dois valores possíveis (razão pela qual recebe a denominação de “binário”). O sinal é de natureza determinística e, dado que sua função de correlação se assemelha à função de correlação do ruído branco, é dito “pseudo-aleatório”. Um estudo mais detalhado pode-se encontrar em [9, 26]. A escolha dos níveis do sinal é normalmente limitada, tal que maximize a mudança de valores do sinal de saída sem comprometer seu funcionamento adequado, isto é, levar o sistema a operar fora da faixa linear [1].

- (d) Ruído binário generalizado (GBN), é um sinal que faz uma mudança entre dois valores possíveis. Tem uma configuração similar ao “ruído binário” (BN), com a diferença que a probabilidade de chaveamento p não está fixado em “0,5”, mas se encontra em uma faixa $0 < p < 1$. O sinal GBN possui média zero e é de natureza não-determinística. Foi proposto em [46] e sua regra de chaveamento está definida pela equação (A.2)

$$\begin{aligned} P[u(t) &= -u(t-1)] = p \\ P[u(t) &\neq u(t-1)] = 1 - p \end{aligned} \quad (\text{A.2})$$

onde $u(t)$ é o sinal no instante atual e $u(t-1)$ é o sinal no instante anterior. O sinal GBN $u(t)$ assume dois valores $-a$ e a .

Seja R um valor sorteado de forma aleatória e $u(t) = a$. Se $R < p$, então a variável troca de sinal $u(t+1) = -u(t) = -a$. Se $R \geq p$ então a variável não troca de sinal $u(t+1) = u(t) = a$.

Define-se o tempo de chaveamento T_{ch} como o intervalo decorrido entre dois chaveamentos consecutivos, estando sua unidade definida em unidades de amostragem. O tempo médio de chaveamento é definido pela equação (A.3)

$$E(T_{ch}) = \frac{T_{min}}{p} \quad (\text{A.3})$$

onde T_{min} é o tempo mínimo de chaveamento, na prática normalmente é assumido na faixa entre 3 a 5 intervalos de amostragem.

O tempo médio de chaveamento também está relacionado com o tempo de acomodação ao degrau t_s , conforme se mostra na equação (A.4)

$$E(T_{ch}) = \frac{T_s}{n} \quad (\text{A.4})$$

onde n é um parâmetro arbitrário escolhido para ajustar a frequência do sinal gerado.

A.4 Modelo de sistemas LIT e tipos de estruturas

Um modelo de um sistema é uma descrição de suas propriedades, adequado para certo propósito. O modelo não precisa ser uma descrição verdadeira e exata do sistema [22].

A.4.1 Modelo de sistemas lineares invariantes no tempo

Um modelo linear invariante no tempo pode ser especificado por sua resposta ao impulso $\{g(k)\}_1^\infty$, do espectro $\phi_v(\omega) = \sigma \cdot |H(e^{j\omega})|^2$ da perturbação aditiva e, possivelmente, da função densidade de probabilidade (FDP) da perturbação $e(t)$ [22]. Portanto, um modelo completo seria dado por:

$$y(t) = G(q)u(t) + H(q)e(t) \quad f_e(x) = FDP \text{ de } e \quad (\text{A.5})$$

onde:

$$G(q) = \sum_{k=1}^{\infty} g(k)q^{-k}$$

$$H(q) = 1 + \sum_{k=1}^{\infty} h(k)q^{-k}$$

Definir o modelo enumerando as sequências infinitas $\{g(k)\}$, $\{h(k)\}$ junto com a função $f_e(x)$, na maioria dos casos, é praticamente inviável. Portanto, na prática, é melhor trabalhar com estruturas que permitam a especificação de G e H em termos de um número de valores finitos. Tal é assim, que as funções de transferência racionais são típicos exemplos disto. Inclusive, na maioria das vezes, a função de densidade de probabilidade (PDF) não é especificada como uma função, mas descrita em função de umas poucas características numéricas, como os momentos de primeira e segunda ordem [22].

$$\begin{cases} E[e(t)] &= \int x f_e(x) dx &= 0 \\ E[e^2(t)] &= \int x^2 f_e(x) dx &= \sigma^2 \end{cases} \quad (\text{A.6})$$

Também é muito comum assumir que $e(t)$ tenha uma distribuição gaussiana, onde a PDF é especificada completamente pela equação (A.6).

O modelo completo (A.5) em termos de um número finito de valores ou coeficientes, é conhecido como modelo paramétrico e envolve métodos de estimação e predição conhecidos como “métodos de identificação paramétricos”.

Muitas vezes, não é possível determinar, *a priori*, estes coeficientes a partir do conhecimento das leis físicas que governam o comportamento do sistema. É por isso que a determinação do valor numérico de todos ou alguns desses coeficientes deve ser deixada para os procedimentos de estimação. Isto é, os coeficientes em questão entram no modelo como parâmetros a serem determinados, os quais são denotados pelo vetor θ , que contém todos os parâmetros a serem estimados. Então, agora a descrição do modelo é a seguinte [22]:

$$y(t) = G(q, \theta) \cdot u(t) + H(q, \theta) \cdot e(t) \quad (\text{A.7})$$

$$f_e(x, \theta), \quad \text{a FDP de } e(t)$$

$$\{e(t)\} = \text{ruído branco}$$

Vale notar que este não é um modelo, mas sim um conjunto de modelos, e cabe ao procedimento de estimação selecionar o membro do conjunto que aparente ser o mais adequado para o propósito em questão [22].

A.4.2 Método dos Mínimos Quadrados

O método dos mínimos quadrados é uma técnica estatística, que procura determinar os parâmetros de ajuste que a matriz de regressores requer, para determinar uma estimativa da saída do sistema. Em [3, 13, 20, 22, 21, 32] se formaliza o método.

Seja um processo descrito pela equação de diferenças a seguir:

$$y(t) + a_1y(t-1) + \dots + a_ny(t-n) = b_1u(t-1) + \dots + b_mu(t-m) + e(t) \quad (\text{A.8})$$

onde $y(t-1), y(t-2), \dots, y(t-n)$ e $u(t-1), u(t-2), \dots, u(t-m)$ são os dados empíricos da saída $y(t)$ e da entrada $u(t)$ respectivamente; e $e(t)$ representa todas as perturbações não medidas do processo.

Reescrevendo-se a equação (A.8) em função da saída atual do sistema, tem-se:

$$y(t) = -a_1y(t-1) - \dots - a_ny(t-n) + b_1u(t-1) + \dots + b_mu(t-m) + e(t) \quad (\text{A.9})$$

Pode-se reescrever a equação (A.9) em sua forma matricial como:

$$y(t) = \varphi(t)\theta^T + e(t) \quad (\text{A.10})$$

onde:

$$\theta = [a_1 \ \dots \ a_n \ \ b_1 \ \dots \ b_m]$$

$$\varphi(t) = [-y(t-1) \ \dots \ -y(t-n) \ \ u(t-1) \ \dots \ u(t-m)];$$

e o valor estimado de $y(t)$, dado os parâmetros estimados $\hat{\theta}$, é definido pela equação (A.11):

$$\hat{y}(t|\hat{\theta}) = \varphi(t)\hat{\theta}^T \quad (\text{A.11})$$

A partir das equações (A.10) e (A.11), resulta:

$$y(t) = \hat{y}(t|\hat{\theta}) + \varepsilon(t) \quad (\text{A.12})$$

onde o termo $\varepsilon(t)$ representa o erro da predição do modelo. Este erro não está correlacionado à saída do sistema.

Uma medida do desvio dos dados reais com respeito aos dados estimados, é a variância, que está definida por:

$$V_N = \frac{1}{N} \sum_{t=1}^N [\varepsilon(t)]^2 \quad (\text{A.13})$$

onde V_N é a variância do desvio entre os dados reais e os dados estimados de $y(t)$ e N é o comprimento dos dados coletados.

Da equação (A.13), se deduz que V_N é sempre positiva; e que quando o desvio do dado real $y(t)$ com respeito ao dado estimado $\hat{y}(t|\hat{\theta})$ aumenta, a variância também aumenta. Daí que o objetivo deste método é estimar o vetor de parâmetros $\hat{\theta}$ que minimiza a variância. Este objetivo pode-se definir como:

$$\min_{\theta} \mathbf{V}_N(\hat{\theta}, \varphi) \quad (\text{A.14})$$

Substituindo-se (A.11) em (A.14):

$$\mathbf{V}_N = \frac{1}{N} \sum_{t=1}^N [y(t) - \hat{y}(t|\hat{\theta})]^2 \quad (\text{A.15})$$

Dado que se deseja minimizar V_N , determina-se seu gradiente e iguala-se a zero. Com isso, obtêm-se um mínimo global:

$$\nabla_h \{\mathbf{V}_N\} = \frac{d}{d\theta} \mathbf{V}_N(\hat{\theta}, \varphi) = \frac{d}{d\theta} \left[\frac{1}{N} \sum_{t=1}^N [y(t) - \varphi^T(t)\hat{\theta}]^2 \right] = 0 \quad (\text{A.16})$$

$$\frac{2}{N} \sum_{t=1}^N \varphi(t) [y(t) - \varphi^T(t)\hat{\theta}] = 0$$

onde $\hat{\theta}$ é tal que:

$$\sum_{t=1}^N \varphi(t)y(t) = \sum_{t=1}^N \varphi(t)\varphi^T(t)\hat{\theta}$$

ou

$$\hat{\theta} = \left[\sum_{t=1}^N \varphi(t)\varphi^T(t) \right]^{-1} \left[\sum_{t=1}^N \varphi(t)y(t) \right] \quad (\text{A.17})$$

então $\hat{\theta}$ é a solução do problema da equação (A.14), isto é, o vetor de parâmetros estimados $\hat{\theta}$, tal que minimiza o desvio entre os dados reais e estimados de $y(t)$.

A.4.3 Erro de predição

A seguir se apresenta um processo com perturbações em sua forma completa, que foi introduzido na equação (A.7):

$$y(t) = G(q) \cdot u(t) + H(q) \cdot e(t) \quad (\text{A.18})$$

$$f_e(x, \theta), \quad \text{a FDP de } e(t)$$

onde $e(t)$ é uma perturbação do processo. O termo $e(t)$ responde pelo efeito de todas as perturbações não medidas e ruídos.

O modelo do processo descrito na equação (A.18) está definido pela equação de diferenças a seguir:

$$y(t) = \hat{G}(q) \cdot u(t) + \hat{H}(q) \cdot \varepsilon(t) \quad (\text{A.19})$$

onde $\varepsilon(t)$ é o erro de modelagem. O termo $\varepsilon(t)$ responde pelos efeitos de todas as perturbações não medidas, ruídos, erros na coleta de dados, erros numéricos na estimação dos parâmetros do modelo e por quaisquer outros erros não identificados. Portanto, $\varepsilon(t)$ descreve o desvio nos dados entre o processo real e o sistema estimado.

No modelo do processo descrito na equação (A.19), note que quando $\hat{G} = 0$, então $\hat{H} = I$, onde I é a matriz identidade. Isto é, quando se estima $\hat{y}(t)$ sem considerar o efeito das entradas no processo, a saída estimada do processo representa só os efeitos dos erros de predição.

O preditor um passo à frente pode ser escrito como [22]:

$$\begin{aligned}
 y(t) &= \hat{G}(q)u(t) + \hat{H}(q)\varepsilon(t) \\
 &= \hat{G}(q)u(t) + \hat{H}(q)\varepsilon(t) - \varepsilon(t) + \varepsilon(t) \\
 &= \hat{G}(q)u(t) + (\hat{H}(q) - I)\varepsilon(t) + \varepsilon(t) \\
 &= \hat{G}(q)u(t) + [\hat{H}(q) - I]\hat{H}^{-1}(q)(y(t) - \hat{G}(q)u(t)) + \varepsilon(t) \\
 &= \hat{G}(q)u(t) + [I - \hat{H}^{-1}(q)][y(t) - \hat{G}(q)u(t)] + \varepsilon(t) \\
 &= \hat{G}(q)u(t) + y(t) - \hat{H}^{-1}(q)y(t) - \hat{G}(q)u(t) + \hat{H}^{-1}(q)\hat{G}(q)u(t) + \varepsilon(t) \\
 &= [I - \hat{H}^{-1}(q)]y(t) + \hat{H}^{-1}(q)\hat{G}(q)u(t) + \varepsilon(t)
 \end{aligned} \tag{A.20}$$

Portanto, a saída do processo pode ser definida por:

$$y(t) = [I - \hat{H}^{-1}(q)]y(t) + \hat{H}^{-1}(q)\hat{G}(q)u(t) + \varepsilon(t) \tag{A.21}$$

Caso se despreze o erro $\varepsilon(t)$ na equação (A.21), resulta:

$$\hat{y}(t) = [I - \hat{H}^{-1}(q)]y(t) + \hat{H}^{-1}(q)\hat{G}(q)u(t) \tag{A.22}$$

A equação (A.22) representa o preditor da saída do processo (equação (A.18)) um passo à frente.

Sabendo que:

$$\hat{G}(q) = \frac{\hat{B}(q)}{\hat{A}(q)} \quad e \quad \hat{H}(q) = \frac{1}{\hat{A}(q)} \tag{A.23}$$

Substituindo-se (A.23) em (A.22), resulta:

$$\hat{y}(t) = [1 - \hat{A}(q)] \cdot y(t) + \hat{B}(q) \cdot u(t) \tag{A.24}$$

Os termos $\hat{A}(q)$ e $\hat{B}(q)$ são os parâmetros estimados do modelo, estes são calculados mediante o método dos mínimos quadrados, como definido na equação (A.17).

Dado que a saída do modelo no instante t pode ser predita usando as entradas $t - 1, t - 2, \dots$; então a equação (A.24) pode-ser reescrita como:

$$\hat{y}(t/t - 1) = [1 - \hat{A}(q)] \cdot y(t) + \hat{B}(q) \cdot u(t) \tag{A.25}$$

onde o erro de predição um passo à frente é definido como:

$$\varepsilon(t) = y(t) - \hat{y}(t) \tag{A.26}$$

A.4.4 Estruturas do modelo

Nesta subseção apresentam-se as famílias de modelos que resultam de reescrever a equação (A.18), analisada de diferentes formas. O modo mais imediato de parametrizar G e H é representá-los como funções racionais. Estas estruturas de modelo são conhecidas como modelos caixa-preta [1, 22].

Estrutura ARX

A relação mais simples de entrada/saída de um sistema pode ser escrita como um modelo representado em equações de diferenças lineares, conforme visto na equação (A.8). Assim:

$$y(t) + \hat{a}_1 y(t-1) + \dots + \hat{a}_{n_a} y(t-n_a) = \hat{b}_1 u(t-1) + \dots + \hat{b}_{n_b} u(t-n_b) + \varepsilon(t) \quad (\text{A.27})$$

onde $\varepsilon(t)$ é o modelo do erro da equação [22].

Neste caso, os parâmetros a ajustar são:

$$\hat{\theta} = [\hat{a}_1 \dots \hat{a}_{n_a} \quad \hat{b}_1 \dots \hat{b}_{n_b}]^T \quad (\text{A.28})$$

Definem-se os polinômios $\hat{A}(q)$ e $\hat{B}(q)$ como:

$$\begin{aligned} \hat{A}(q) &= 1 + \hat{a}_1 \cdot q^{-1} + \hat{a}_2 \cdot q^{-2} \dots \hat{a}_{n_a} \cdot q^{-n_a} \\ \hat{B}(q) &= \hat{b}_1 \cdot q^{-1} + \hat{b}_2 \cdot q^{-2} \dots \hat{b}_{n_b} \cdot q^{-n_b} \end{aligned} \quad (\text{A.29})$$

Então, as funções $\hat{G}(q)$ e $\hat{H}(q)$ estão definidas em função dos parâmetros $\hat{A}(q)$ e $\hat{B}(q)$ como se mostra a seguir:

$$y(t) = \frac{\hat{B}(q)}{\hat{A}(q)} u(t) + \frac{1}{\hat{A}(q)} \varepsilon(t) \quad (\text{A.30})$$

A equação (A.30) representa a estrutura do modelo tipo ARX, onde AR refere-se à estrutura do modelo de perturbação $1/\hat{A}(q)$; e X à entrada extra $u(t)$, chamada de variável exógena em econometria [1, 22].

O diagrama de blocos da equação (A.30) pode ser visto na Figura A.2.

Figura A.2: Estrutura do modelo ARX.

Na Figura A.2 pode-se observar que o efeito do modelo da perturbação é adicionado só depois da resposta dinâmica do modelo do processo, que do ponto de vista físico é inconsistente.

O preditor um passo à frente para a estrutura ARX se mostra na equação (A.31). Assim:

$$\hat{y}(t/t-1) = [1 - \hat{A}(q)] \cdot y(t) + \hat{B}(q) \cdot u(t) \quad (\text{A.31})$$

Estrutura ARMAX

A principal desvantagem de usar a estrutura de modelo tipo ARX, vista na equação (A.27), é a falta de liberdade na descrição do termo de perturbação. Contudo, é possível acrescentar maior flexibilidade ao modelo, adicionando um termo conhecido como média móvel (*moving average*) do ruído branco.

A representação matemática do modelo ARMAX é descrita na equação (A.32).

$$A(q) \cdot y(t) = B(q) \cdot u(t) + C(q) \cdot \varepsilon(t) \quad (\text{A.32})$$

O modelo ARMAX descrito na equação (A.32) pode ser descrito em equação de diferenças lineares, como:

$$\begin{aligned} y(t) + \hat{a}_1 y(t-1) + \dots + \hat{a}_{n_a} y(t-n_a) &= \hat{b}_1 u(t-1) + \dots + \hat{b}_{n_b} u(t-n_b) + \\ &\quad \varepsilon(t) + \hat{c}_1 \varepsilon(t-1) + \dots + \hat{c}_{n_c} \varepsilon(t-n_c) \end{aligned} \quad (\text{A.33})$$

Neste caso, os parâmetros a ajustar são:

$$\hat{\theta} = [\hat{a}_1 \dots \hat{a}_{n_a} \quad \hat{b}_1 \dots \hat{b}_{n_b} \quad \hat{c}_1 \dots \hat{c}_{n_c}]^T \quad (\text{A.34})$$

Definem-se os polinômios $\hat{A}(q)$, $\hat{B}(q)$ e $\hat{C}(q)$ como:

$$\begin{aligned}\hat{A}(q) &= 1 + \hat{a}_1 \cdot q^{-1} + \hat{a}_2 \cdot q^{-2} \cdots \hat{a}_{n_a} \cdot q^{-n_a} \\ \hat{B}(q) &= \hat{b}_1 \cdot q^{-1} + \hat{b}_2 \cdot q^{-2} \cdots \hat{b}_{n_b} \cdot q^{-n_b} \\ \hat{C}(q) &= \hat{c}_1 \cdot q^{-1} + \hat{c}_2 \cdot q^{-2} \cdots \hat{c}_{n_c} \cdot q^{-n_c}\end{aligned}\tag{A.35}$$

Então, as funções $\hat{G}(q)$ e $\hat{H}(q)$ estão definidas em função dos parâmetros $\hat{A}(q)$, $\hat{B}(q)$ e $\hat{C}(q)$ como se mostra a seguir:

$$y(t) = \frac{\hat{B}(q)}{\hat{A}(q)}u(t) + \frac{\hat{C}(q)}{\hat{A}(q)}\varepsilon(t)\tag{A.36}$$

onde

$$\hat{G}(q) = \frac{\hat{B}(q)}{\hat{A}(q)} \quad e \quad \hat{H}(q) = \frac{\hat{C}(q)}{\hat{A}(q)}\tag{A.37}$$

A equação (A.36) representa a estrutura do modelo tipo ARMAX; onde AR refere-se à estrutura do modelo de perturbação, MA refere-se à média móvel e X à entrada extra $u(t)$, chamada de variável exógena em econometria.

O diagrama de blocos da equação (A.36) pode ser visto na Figura A.3.

Figura A.3: Estrutura do modelo ARMAX.

Na prática, a estrutura do modelo tipo ARMAX virou uma ferramenta padrão em controle, tanto na descrição quanto para o projeto do controle.

O preditor um passo à frente para a estrutura ARMAX é obtido substituindo-se (A.37) em (A.22). Resulta:

$$\hat{C}(q)\hat{y}(t/t-1) = [\hat{C}(q) - \hat{A}(q)]y(t) + \hat{B}(q)u(t)\tag{A.38}$$

Observação: uma versão da estrutura do modelo ARMAX com uma integração (I) forçada é conhecida com o nome de **ARIMAX**. Este modelo é frequentemente usado para descrever sistemas com perturbações lentas. É obtido substituindo-se $\delta y = y(t) - y(t-1)$ no lugar de $y(t)$ na equação (A.36) ou, equivalentemente, $\hat{A}(q)$ por $(1 - q^{-1}) \cdot \hat{A}(q)$.

Estrutura ARARX

Ao invés de modelar o termo do ruído da equação (A.27) com média móvel, como foi feito em (A.33), pode-se descrever esta perturbação como uma auto-regressão, daí o nome da estrutura do modelo [22]:

$$\hat{A}(q)y(t) = \hat{B}(q)u(t) + \frac{1}{\hat{D}(q)}\varepsilon(t) \quad (\text{A.39})$$

onde

$$\hat{D}(q) = 1 + \hat{d}_1 \cdot q^{-1} + \hat{d}_2 \cdot q^{-2} \dots \hat{d}_{n_d} \cdot q^{-n_d}$$

O diagrama de blocos da estrutura do modelo ARARX se mostra na Figura A.4.

Figura A.4: Estrutura do modelo ARARX.

Estrutura ARARMAX

Uma forma mais geral, poder-se-ia usar uma descrição ARMA para o modelo de perturbação, obtendo-se a estrutura ARARMAX:

$$\hat{A}(q)y(t) = \hat{B}(q)u(t) + \frac{\hat{C}(q)}{\hat{D}(q)}\varepsilon(t) \quad (\text{A.40})$$

Sendo o diagrama de blocos mostrado na Figura A.5.

Figura A.5: Estrutura do modelo ARARMAX.

Estrutura erro na saída

Até agora, as estruturas vistas possuem o polinômio $\hat{A}(q)$ como fator comum em seus denominadores. Do ponto de vista físico, seria mais natural parametrizar estas funções de transferência de forma independente.

Suponha então que a relação entre a entrada e uma saída não perturbada $w(t)$ possa ser representada como uma equação de diferenças linear e que as perturbações consistam em ruído branco na medição [22]:

$$w(t) + \hat{f}_1 w(t-1) + \dots + \hat{f}_{n_f} w(t-n_f) = \hat{b}_1 u(t-1) + \dots + \hat{b}_{n_b} u(t-n_b) \quad (\text{A.41})$$

$$y(t) = w(t) + \varepsilon_{OE}(t) \quad (\text{A.42})$$

O modelo pode ser escrito como:

$$y(t) = \frac{\hat{B}(q)}{\hat{F}(q)}u(t) + \varepsilon_{OE}(t) \quad (\text{A.43})$$

O diagrama do blocos da estrutura do modelo de erro na saída apresenta-se na Figura A.6.

Figura A.6: Estrutura do modelo de erro na saída.

Estrutura Box-Jenkins

Uma evolução natural da estrutura de erro na saída (A.43) é também modelar as propriedades do erro na saída. Descrevendo-se isto como um modelo ARMA [22], resulta:

$$y(t) = \frac{\hat{B}(q)}{\hat{F}(q)}u(t) + \frac{\hat{D}(q)}{\hat{D}(q)}\varepsilon(t) \quad (\text{A.44})$$

Sendo o diagrama de blocos mostrado na Figura A.7.

Figura A.7: Estrutura do modelo Box-Jenkins.

Estrutura geral de modelo linear

A seguir, apresenta-se a equação (A.45), que é a estrutura geral de modelo linear:

$$A(q) \cdot y(t) = \frac{B(q)}{F(q)}u(t) + \frac{C(q)}{D(q)}\varepsilon(t) \quad (\text{A.45})$$

onde q^{-1} é um operador de atraso, isto é, $y(t)q^{-1} = y(t-1)$, $\varepsilon(t)$ é erro na medição e $A(q)$, $B(q)$, $C(q)$, $D(q)$ e $F(q)$ são os polinômios definidos a seguir [1, 22]:

$$\begin{aligned}
 A(q) &= 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a} \\
 B(q) &= b_1 q^{-1} + \dots + b_{n_b} q^{-n_b} \\
 C(q) &= 1 + c_1 q^{-1} + \dots + c_{n_c} q^{-n_c} \\
 D(q) &= 1 + d_1 q^{-1} + \dots + d_{n_d} q^{-n_d} \\
 F(q) &= 1 + f_1 q^{-1} + \dots + f_{n_f} q^{-n_f}
 \end{aligned} \tag{A.46}$$

A escolha da estrutura do modelo consiste na seleção do grau dos polinômios $A(q)$, $B(q)$, $C(q)$, $D(q)$ e $F(q)$, isto é, n_a , n_b , n_c , n_d e n_f .

Esta seleção pode ser até de 32 estruturas diferentes, porém, na Tabela A.1 apresentam-se as estruturas mais usadas.

Tabela A.1: Alguns casos especiais de modelos SISO [22].

Polinômio da equação (A.45)	Nome da estrutura do modelo
B	FIR
A, B	ARX
A, B, C	ARMAX
A, C	ARMAX
A, B, D	ARARX
A, B, C, D	ARARMAX
B, F	OE (erro na saída)
B, C, D, F	BJ (Box-Jenkins)

O preditor um passo à frente pode ser determinado a partir de (A.22):

$$\hat{y}(t/t-1, \hat{\theta}) = \frac{\hat{D}(q) \cdot \hat{B}(q)}{\hat{C}(q) \cdot \hat{F}(q)} u(t) + \left[1 - \frac{\hat{D}(q) \cdot \hat{A}(q)}{\hat{C}(q)} \right] y(t) \tag{A.47}$$

A.5 Validação do modelo

Depois de passar pelas etapas anteriores de identificação de sistemas e, tendo em mãos um modelo obtido a partir de dados experimentais, a questão é: o modelo obtido é bom? Quão bom é? E caso se tenha obtido não só um modelo, mas uma família de modelos, a pergunta é: qual dos modelos é melhor? Isto no caso de se ter modelos com estruturas diferentes ou com diferentes graus dos polinômios.

A resposta às questões anteriores é relativa, e isso só vai depender do objetivo da aplicação em questão. Por exemplo, pode-se desejar estimar ou prever o comportamento de uma planta industrial vários passos à frente, para gerar um sinal de controle tal que minimize o erro na saída, como é o caso do controle MPC. Em outros casos, quando o modelo é obtido para projetar a sintonia de controladores, a predição vários passos à frente não é muito importante.

Portanto, o modelo provavelmente é representativo do sistema em apenas alguns aspectos, ele será considerado válido se incorporar aquelas características do sistema que são fundamentais para a aplicação em questão [1].

Simulação

Um dos métodos mais usados na validação de sistemas é a comparação entre os dados medidos e os dados simulados (obtidos do modelo identificado). O método é muito simples, daí que é muito comum, mas precisa ter algumas considerações.

Na coleta de dados, precisa-se obter dados para identificação e dados para validação. Os dados de identificação são usados para estimar os parâmetros da estrutura do modelo obtido. Os dados de validação são usados para compará-los com os dados simulados do modelo. Ambos os tipos de dados devem ser coletados quando o sistema está operando em condições semelhantes.

A importância de não usar os mesmos dados quando se estima o modelo e quando se valida é que se perde o princípio de generalização, isto é, a capacidade que

tem o modelo para reproduzir ou prever um outro conjunto de dados observados do mesmo sistema. Portanto, refere-se à capacidade de generalização do modelo [1].

Os processos industriais, em geral, são sistemas não-lineares variantes no tempo, mas podem ser considerados lineares invariantes no tempo em uma região em torno do ponto de operação, portanto o modelo reproduzirá o comportamento dinâmico do sistema na mesma região.

Índices de validação de modelos

1. *Índice de avaliação da ordem da estrutura do modelo (RV)*, a seleção da ordem do modelo é importante, pois uma estrutura sobre-ajustada pode levar a computações desnecessariamente complicadas para encontrar as estimativas dos parâmetros do modelo, contudo, uma estrutura de modelo sub-ajustada pode ser muito imprecisa [53].

Para determinar a ordem da estrutura do modelo se usa a função de custo do erro na saída. O objetivo é que este erro seja menor que um valor determinado. Para isso, utiliza-se o índice a seguir [52]:

$$RV = \frac{\sum_{t=1}^N [\varepsilon_{OE}(t)]^2}{\sum_{t=1}^N [y(t)]^2} \quad (\text{A.48})$$

Sugere-se que um valor adequado de RV esteja entre 10% e 20% [52].

2. *Índice de ajuste do modelo (fit)*: o *fit* é o termo que mede o grau de ajuste do modelo estimado com respeito aos dados de validação e é definido em [1]:

$$fit(\%) = 100 \cdot \left(1 - \frac{\|\hat{y}(t) - y(t)\|}{\|y(t) - \bar{y}\|} \right) \quad (\text{A.49})$$