

Universidade Estadual de Campinas – UNICAMP
Faculdade de Engenharia Elétrica e de Computação – FEEC
Depto. de Eng. Computação e Automação Industrial – DCA

Desenvolvimento de Modelos Discretos de Volterra Usando Funções de Kautz

Autor: Alex da Rosa

Orientador: Dr. Wagner Caradori do Amaral

Co-orientador: Dr. Ricardo José Gabrielli Barreto Campello

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: **Engenharia de Computação.**

BANCA EXAMINADORA

Dr. Wagner Caradori do Amaral (FEEC/UNICAMP)

Dr. Paulo Augusto Valente Ferreira (FEEC/UNICAMP)

Dr. Luis Antonio Aguirre (UFMG)

Campinas-SP, 18 de fevereiro de 2005

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

R71d Rosa, Alex da
Desenvolvimento de Modelos Discretos de Volterra
Usando Funções de Kautz / Alex da Rosa – Campinas, SP:
[s.n.], 2005.

Orientadores: Wagner Caradori do Amaral, Ricardo
José Gabrielli Barreto Campello

Dissertação (Mestrado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

1. Sistemas Não-lineares. 2. Volterra, Series de.
3. Otimização Matemática. 4. Identificação de Sistemas.
I. Amaral, Wagner Caradori do. II. Campello, Ricardo José
Gabrielli Barreto. III. Universidade Estadual de Campinas.
Faculdade de Engenharia Elétrica e de Computação.
IV. Título.

Este trabalho é dedicado
aos meus pais, Ciro e
Vanda, e aos meus irmãos
Alan, Alessandra e Andréia.

Agradecimentos

Primeiramente, eu gostaria de agradecer ao meu orientador, professor Dr. Wagner Caradori do Amaral, por esses dois anos de supervisão. Especialmente pelo seu comprometimento em me orientar nesse trabalho de mestrado, bem como pelo tempo dispendido por ele ao ler os rascunhos dessa tese. Seus comentários críticos sobre meu trabalho tiveram enorme utilidade para o conteúdo e apresentação de meus argumentos e discussões.

Estendo meus agradecimentos ao professor Dr. Ricardo José Gabrielli Barreto Campello, pela co-orientação desse trabalho. Suas sugestões e comentários permitiram aprimorar a qualidade do trabalho.

Eu também gostaria de agradecer imensamente a todos os amigos e professores da UNICAMP, que deram sua contribuição para o desenvolvimento deste trabalho.

Sou muito agradecido aos professores que contribuíram para a minha vida acadêmica, desde alguns anos atrás, durante minha graduação na Universidade Federal de Goiás. Muitos agradecimentos aos colegas de Goiânia, que me ajudaram e me apoiaram.

Meus infinitos agradecimentos aos meus pais, Ciro e Vanda, aos meus irmãos Alan, Alessandra e Andréia, e à minha namorada Danielle. Mesmo a quase 900 quilômetros de distância, todos sempre me deram muita força para que eu atingisse meus objetivos.

À CAPES, pelo apoio financeiro durante o desenvolvimento deste trabalho.

A Deus.

Resumo

Este trabalho analisa a modelagem de sistemas não-lineares utilizando modelos de Wiener/Volterra com funções ortonormais de Kautz. Os modelos de Volterra são uma generalização do modelo resposta ao impulso para a descrição de sistemas não-lineares. Esses modelos necessitam de um número considerável de termos para a representação dos *kernels* de Volterra. Essa complexidade pode ser reduzida utilizando-se uma representação do tipo Wiener/Volterra, em que os *kernels* são desenvolvidos utilizando uma base de funções ortonormais. São discutidos aspectos da seleção dos parâmetros livres (pólos) que caracterizam essas funções, particularmente a seleção ótima dos pólos complexos das funções de Kautz. Este problema é resolvido minimizando-se o limitante superior do erro que surge a partir da aproximação truncada dos *kernels* de Volterra usando-se as funções de Kautz. Obtém-se a solução analítica para a escolha ótima de um dos parâmetros relacionados com o pólo de Kautz, sendo os resultados válidos para modelos Wiener/Volterra de qualquer ordem. Apresentam-se ainda resultados de simulações que ilustram a metodologia apresentada, bem como a modelagem de um sistema de levitação magnética.

Palavras-chave: Sistemas Não-Lineares, Séries de Volterra, Funções de Base Ortonormal, Otimização, Identificação de Processos.

Abstract

This work investigates the modelling of nonlinear systems using the Wiener/Volterra models with Kautz orthonormal functions. The Volterra models constitute a generalization of the impulse response model to describe nonlinear systems. Such models require a large number of terms for representing the Volterra kernels. However, this complexity can be reduced by using Wiener/Volterra models, in which the kernels are expanded using an orthonormal basis functions. Aspects about selection of the free parameters (poles) characterizing these functions are discussed, in particular the optimal selection of the complex poles of the Kautz functions. This problem is solved by minimizing the upper bound of the error arising from the truncated approximation of Volterra kernels using Kautz functions. An analytical solution for the optimal choice of one of the parameters related to the Kautz pole is thus obtained, with the results valid for any-order Wiener/Volterra models. Simulations that illustrate the methodology described above are presented. Also, the modelling of a magnetic levitation system is discussed.

Key-words: Nonlinear Systems, Volterra Series, Orthonormal Basis Functions, Optimization, System Identification.

Sumário

Lista de Figuras	v
Lista de Tabelas	vii
1 Introdução	1
1.1 Organização do Trabalho e Contribuições	2
2 Representação de Sistemas via Modelos de Volterra	4
2.1 Considerações Iniciais Sobre Sistemas	4
2.1.1 Classificação dos Sistemas	5
2.2 Modelos de Volterra	8
2.2.1 Séries de Volterra	8
2.3 O <i>Kernel</i> de Volterra de Segunda Ordem	12
2.3.1 Resposta ao Impulso Bidimensional	12
2.3.2 Exemplo 2.1	13
2.3.3 Exemplo 2.2	14
2.4 Simetrização do <i>kernel</i> de Volterra	15
2.5 Resumo	16
3 Bases de Funções Ortonormais	17
3.1 Introdução às Funções Ortogonais	17
3.2 Polinômios Ortogonais Discretos	18
3.3 Funções Ortonormais Generalizadas	19
3.4 Funções de Laguerre	21
3.4.1 Descrição das Funções de Laguerre em Espaço de Estados	22
3.5 Funções de Kautz	23
3.5.1 Descrição das Funções de Kautz em Espaço de Estados	24
3.6 Resumo	26
4 Aproximação de Modelos de Volterra Usando Funções de Kautz	27
4.1 Motivação Inicial	27
4.2 Modelos de Wiener/Volterra	28
4.3 Definição do Problema	30
4.3.1 Erro de Aproximação dos <i>Kernels</i> de Volterra	32
4.4 Desenvolvimento Matemático Preliminar	34
4.4.1 Proposições	38

4.5	Expansão Sub-Ótima dos <i>Kernels</i> de Volterra em Base de Kautz . . .	41
4.5.1	Algoritmo	45
4.6	Resumo e Contribuições	46
5	Simulações e Resultados	47
5.1	Exemplo Ilustrativo	47
5.1.1	Estudo da Simetrização do <i>Kernel</i> de Volterra	52
5.1.2	Comparação Entre Funções de Kautz e de Laguerre	53
5.2	Influência da Dinâmica do Sistema no Cálculo do Pólo de Kautz . . .	53
5.3	Estimação Simultânea dos <i>Kernels</i> de Volterra e do Pólo de Kautz . .	58
5.3.1	Modelo do Levitador Magnético	58
5.3.2	Resultados de Simulação	60
5.4	Resumo e Discussões	65
6	Conclusões e Comentários Finais	66
6.1	Contribuições	67
6.2	Perspectivas Futuras	67
A	Transformada Z Multidimensional	69
A.1	Definição de Transformada Z Multidimensional	69
A.1.1	Exemplo	69
B	Demonstrações	71
B.1	Demonstração da Equação (4.47)	71
B.2	Demonstração da Equação (4.49)	72
B.3	Demonstração da Equação (4.62)	76
	Referências Bibliográficas	81

Lista de Figuras

2.1	O mapeamento do operador \mathcal{H} em uma excitação $u \in U$ leva a uma resposta $y \in Y$	5
2.2	Modelo entrada/saída para sistemas físicos	6
2.3	Representação esquemática de um sistema caracterizado por uma série de Volterra	10
2.4	Sistema não-linear com <i>kernel</i> de segunda ordem	14
2.5	Representação em diagrama de blocos de um determinado sistema	14
3.1	Sistema a partir do qual são geradas as funções de Laguerre $\phi_n(k)$, $n = 1, 2, \dots$	22
3.2	Sistema a partir do qual se obtém as funções de Kautz $\psi_{2n}(k)$ e $\psi_{2n-1}(k)$, para $n = 1, 2, \dots$	24
5.1	<i>Kernel</i> da equação (5.1), para $\lambda_1 = 0,45$, $\lambda_2 = 0,7$, $\omega_1 = 100$ e $\omega_2 = 1$	48
5.2	Valor de $c_{\text{ótimo}}$, calculado pela equação (4.62) para cada valor de $b \in (-1, 1)$	50
5.3	Erro Quadrático Normalizado (EQN) entre o <i>kernel</i> real e sua aproximação com $M = 6$ funções de Kautz, calculado para cada valor de $b \in (-1, 1)$	50
5.4	<i>Kernel</i> aproximado $\tilde{h}_2(k_1, k_2)$ para $b = 0,593$ e $c = -0,25940$	51
5.5	Superfície de erro $h_2(k_1, k_2) - \tilde{h}_2(k_1, k_2)$ para $b = 0,593$ e $c = -0,25940$	51
5.6	Superfície de erro para a versão simétrica do <i>kernel</i> , com $b = 0,684$ e $c = -0,46949$	52
5.7	<i>Kernel</i> de segunda ordem da equação (5.8) para $\sigma = -0,2$ e $\omega = 2$	55
5.8	Pólo ótimo de Kautz quando se varia a frequência de oscilação ω do <i>kernel</i> , mantendo-se σ constante	56
5.9	Pólo ótimo de Kautz quando se varia o fator de decaimento σ do <i>kernel</i> , mantendo-se ω constante	56
5.10	Superfície de erro proveniente da aproximação do <i>kernel</i> da equação (5.8) com $\sigma = -0,2$ e $\omega = 2$, para o caso em que o pólo de Kautz utilizado é igual ao parâmetro ζ	57
5.11	Diagrama do levitador magnético	59
5.12	Acima: corrente elétrica na bobina 2 (sinal de entrada u). Abaixo: posição do disco 1 (sinal de saída y)	61

5.13	Evolução dos parâmetros de Kautz, de acordo com o número de iterações, para o <i>kernel</i> de primeira ordem	62
5.14	Evolução dos parâmetros de Kautz, de acordo com o número de iterações, para o <i>kernel</i> de segunda ordem	62
5.15	<i>Kernel</i> de primeira ordem estimado para o modelo (5.10), após 30 iterações do algoritmo	63
5.16	<i>Kernel</i> de segunda ordem estimado para o modelo (5.10), após 30 iterações do algoritmo	63
5.17	Saídas medida (em azul) e estimada (em vermelho), após 30 iterações do algoritmo	64
5.18	Erro de estimação: diferença entre os dados de validação de $y(k)$ e os dados da saída estimada pelo modelo	64
5.19	Evolução do Erro Quadrático Normalizado (EQN) entre os dados de validação da saída $y(k)$ e a sua estimativa	65

Lista de Tabelas

3.1	Coeficientes da equação a diferenças (3.3) para alguns polinômios ortogonais	18
5.1	Erro de aproximação do <i>kernel</i> da equação (5.1) de acordo com o número de funções de Kautz utilizadas na expansão	49
5.2	Comparação do erro de aproximação do <i>kernel</i> (5.1) de acordo com o número de funções utilizadas na expansão, para Laguerre e Kautz .	53
5.3	Valores finais dos parâmetros de Kautz, após 30 iterações do algoritmo	61

Capítulo 1

Introdução

Modelagem, análise e controle de sistemas dinâmicos têm sido de grande interesse de engenheiros por muito tempo [23]. Em poucos anos, o estudo dessas áreas cresceu em importância devido a várias razões. Antes dos computadores digitais, os cálculos usados nessas aplicações consumiam muito tempo, e os erros provenientes de simulações eram consideráveis. Hoje em dia, o uso de computadores se expandiu, permitindo estudar modelos mais detalhados e algoritmos complexos para o projeto e análise de sistemas [12]. Com o aumento do poder computacional, pôde-se impor especificações de desempenho mais rígidas nos projetos de engenharia, levando à necessidade de modelos mais detalhados do processo em estudo, especialmente no que diz respeito à predição de seu comportamento transitório. Nesse escopo, inserem-se os modelos de Volterra.

No presente trabalho, aborda-se o problema de identificação de sistemas utilizando modelos de Volterra. Esses modelos têm grande relevância na representação matemática de sistemas não-lineares, e representam uma generalização do modelo de resposta ao impulso, que relaciona a entrada e a saída de sistemas lineares [24], [26]. Os modelos de Volterra podem ser aplicados na previsão da saída de sistemas, para projetos de controle preditivo. Em [19], [21] e [22], estudou-se o controle preditivo de reatores químicos modelados por séries de Volterra usando base de funções ortonormais.

A utilização dos modelos de Volterra permite, entre outras coisas [14]:

- descrever um sistema não-linear como uma generalização do caso linear;
- obter uma relação explícita entre os sinais de entrada e saída;
- estudar sistemas excitados por sinais aleatórios;
- estender a sistemas não-lineares conceitos definidos somente para os lineares.

O número de termos necessários para a representação dos modelos de Volterra pode ser reduzido usando-se uma estrutura do tipo Wiener (modelos Wiener/Volterra). Essa representação possui uma dinâmica linear constituída por um conjunto de filtros de funções ortonormais, seguida de um mapeamento estático não-linear. A utilização de funções ortonormais permite uma redução na ordem da

representação do modelo, o que vem a simplificar os problemas de identificação e controle associados [6], [29].

Uma vez que as funções ortonormais são geralmente caracterizadas por parâmetros arbitrários, conforme será visto no Capítulo 3, na literatura são apresentados resultados para a seleção ótima dessas constantes. Em [15], resolveu-se o problema da escolha ótima do parâmetro que caracteriza as funções ortonormais de Laguerre, conhecido como pólo de Laguerre, para o caso linear. Outros trabalhos, [31] e [32], também obtiveram soluções analíticas para aquele pólo, no caso de sistemas lineares. Posteriormente, [6], [7], [8] e [9] apresentaram o problema da melhor escolha do pólo de Laguerre para sistemas não-lineares.

As funções de Laguerre são mais apropriadas para representar modelos de Volterra de sistemas super-amortecidos [36]. Na representação de sistemas com comportamento oscilatório (sub-amortecidos), é mais conveniente usar funções de Kautz. Em [33] e [34], enfocou-se o estudo de uma solução literal sub-ótima para o pólo de Kautz usado na representação de sistemas lineares. Outro trabalho [11] estudou a seleção sub-ótima do pólo de Kautz para a representação de sistemas não-lineares.

O objetivo principal do presente trabalho é obter a representação de sistemas físicos por meio de modelos Wiener/Volterra, usando-se funções ortonormais de Kautz. Nesse contexto, desenvolvem-se soluções analíticas globais ótimas para a parametrização das bases de Kautz utilizadas na expansão dos *kernels* de Volterra em séries de funções ortonormais. O interesse central do trabalho é apresentar resultados a respeito de “modelos ótimos” que simplifiquem suas implementações computacionais. Deseja-se também fazer uma comparação dessa representação quando se utiliza base de Kautz ou Laguerre, além de propor um algoritmo iterativo que permite estimar os *kernels* de um sistema a partir de dados de entrada e saída.

1.1 Organização do Trabalho e Contribuições

O presente trabalho está organizado da seguinte forma:

- **Capítulo 2 - Representação de Sistemas via Modelos de Volterra:** São discutidos os modelos discretos de Volterra e sua relevância na modelagem de sistemas físicos. Apresenta-se também alguns aspectos matemáticos dessa representação, além de um pequeno histórico de sua utilização. O capítulo apresenta as séries de Volterra como uma generalização do modelo de resposta ao impulso, bastante conhecido no estudo de sistemas lineares. Será abordado também o conceito de simetrização de *kernels* de Volterra, que se apresenta como uma ferramenta matemática para se obter uma representação de sistemas via modelos de Volterra com menor complexidade.
- **Capítulo 3 - Base de Funções Ortonormais:** Apresenta alguns conjuntos de funções ortonormais utilizados em identificação de sistemas. Entre elas, são apresentadas as funções ortonormais generalizadas. Uma ênfase maior é dada para as funções (discretas) de Laguerre e de Kautz, e essas últimas são o objeto maior de estudo do trabalho. Será visto que essas funções são caracterizadas

por parâmetros livres, o que motiva o estudo da escolha desses valores para uma melhor representação computacional dos modelos.

- **Capítulo 4 - Aproximação de Modelos de Volterra Usando Funções de Kautz:** Este capítulo aborda a representação de sistemas utilizando-se modelos Wiener/Volterra com funções ortonormais de Kautz. Tendo em vista que essas funções são parametrizadas por constantes arbitrárias (pólos de Kautz), enfoca-se o problema da escolha ótima desses parâmetros. A partir de uma proposição apresentada, uma equação analítica para essa escolha é obtida. Como as funções de Kautz são parametrizadas por pólos complexos, realiza-se um estudo que envolve a seleção sub-ótima desses pólos, para uma melhor representação de sistemas via modelos de Wiener/Volterra. Os resultados apresentados são válidos para modelos não-lineares de qualquer ordem, e representam uma generalização do caso linear estudado em [34]. A principal contribuição do capítulo, apresentada também em [11], é a generalização para o caso não-linear de resultados até então válidos somente para a representação de sistemas lineares.
- **Capítulo 5 - Simulações e Resultados:** Apresenta as simulações dos resultados obtidos no Capítulo 4. É ilustrada a seleção ótima dos parâmetros de Kautz, ressaltando-se a importância da escolha dos parâmetros livres que caracterizam as funções ortonormais utilizadas na modelagem de sistemas. Ilustra-se também a simetria de *kernels* de Volterra, comparando-se o caso simétrico com o caso assimétrico. Estuda-se ainda a influência da dinâmica do sistema na escolha do pólo de Kautz, i.e. como o seu comportamento dinâmico afeta o cálculo do melhor modelo Wiener/Volterra. É realizada também uma comparação, em termos do erro de aproximação, quando utiliza-se funções ortonormais de Laguerre e de Kautz. O objetivo é comparar a qualidade da aproximação de sistemas não-lineares em função da base ortonormal utilizada na sua representação matemática. Propõe-se ainda neste capítulo um algoritmo para estimação simultânea dos *kernels* do sistema e dos pólos de Kautz, aplicado à modelagem de um sistema de levitação magnética. Esse algoritmo representa uma generalização do algoritmo apresentado em [6] e [7], que utiliza as funções de Laguerre.
- **Capítulo 6 - Conclusões e Comentários Finais:** Apresenta os comentários finais sobre o trabalho, além de suas principais contribuições em identificação de sistemas não-lineares. Discute também as perspectivas para trabalhos futuros.

Capítulo 2

Representação de Sistemas via Modelos de Volterra

A modelagem de sistemas discutida neste capítulo é baseada na representação funcional em séries de Volterra. Embora seja uma ferramenta matemática, a sua aplicação à representação entrada/saída de sistemas pode ser discutida inicialmente sem a necessidade de um desenvolvimento matemático mais elaborado.

O presente capítulo tem o objetivo de apresentar a terminologia, a notação e as manipulações básicas relacionadas às representações de sistemas não-lineares por modelos de Volterra. Apresenta-se uma breve revisão sobre os tipos de sistemas (lineares e não-lineares, estáticos e dinâmicos, dentre outros), como forma de motivação inicial. Também será discutida brevemente a origem e o porquê do uso de séries de Volterra na representação de sistemas. Além disso, são apresentados alguns exemplos de como obter a representação de Volterra (*kernels*) a partir de um conhecimento prévio do sistema que se deseja modelar.

Discute-se ainda o conceito de simetrização, que é apresentado mais como uma ferramenta matemática que visa simplificar a representação de sistemas usando-se modelos de Volterra.

2.1 Considerações Iniciais Sobre Sistemas

A vida diária contém inúmeras situações de ação e reação. Em todos esses casos um conjunto de estímulos ativa a operação de algum processo, para produzir um resultado ou uma resposta [12]. A expressão “sistema” representa genericamente esse processo. Em suma, um sistema se refere a algum processo físico que gera um comportamento, em resposta a um estímulo de entrada [10], [17].

Um sistema pode ser definido no sentido matemático como sendo uma regra pela qual uma excitação u é mapeada em uma resposta y . Esta regra pode ser expressa [12], [14] simbolicamente da forma $y = \mathcal{H}[u]$. Em aplicações científicas, a resposta e a excitação são usualmente funções de uma variável independente, como posição ou tempo [26]. Se essas funções são dependentes somente do tempo, pode-se

escrever:

$$y(t) = \mathcal{H}[u(t)] \quad (2.1)$$

Na representação da equação (2.1), \mathcal{H} é denominado operador porque a função resposta $y(t)$ pode ser vista como sendo produzida a partir de uma operação aplicada à função de entrada $u(t)$. A afirmação de que $y(t)$ é a resposta de um sistema à excitação $u(t)$ significa que existe um operador \mathcal{H} (regra) que mapeia o sinal de entrada $u(t)$ em uma única saída $y(t)$.

Esse mapeamento pode ser visto como ilustrado na figura 2.1, onde o conjunto de todas as possíveis entradas é denotada por U , e o conjunto de todas as saídas possíveis é denotada por Y .

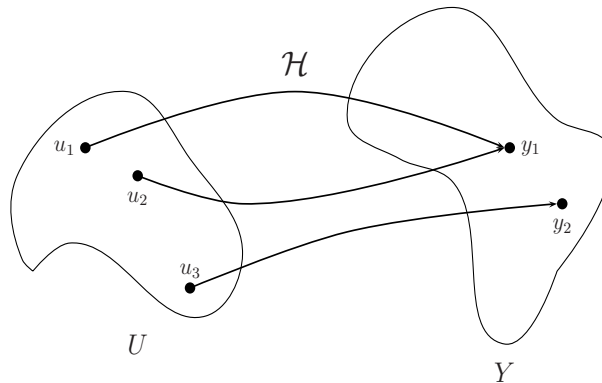


Figura 2.1: O mapeamento do operador \mathcal{H} em uma excitação $u \in U$ leva a uma resposta $y \in Y$

O mapeamento \mathcal{H} define relações matemáticas típicas, como equações diferenciais e equações matriciais, sempre relacionando entrada e saída. Na figura 2.1 acima, note que as entradas denotadas por u_1 e u_2 são ambas mapeadas na mesma saída y_1 . Entretanto, uma entrada não pode ser mapeada em mais do que uma saída.

2.1.1 Classificação dos Sistemas

Uma representação clássica de qualquer sistema é a conhecida como *black-box* (caixa-preta) [12], que pode ser visto como uma metodologia para modelagem de sistemas. Esta representação é essencial no que costuma ser chamado “abordagem de sistemas”. Nessa abordagem, cada elemento do sistema é visto como uma caixa-preta, e a análise é focada em como as conexões entre os elementos influenciam o comportamento total do sistema. Esse ponto de vista implica em um estudo menos detalhado da operação individual de cada elemento do sistema, para se ter um entendimento mais geral do mesmo.

O comportamento de um elemento caixa-preta é especificado por uma relação do tipo entrada/saída. Essa idéia pode ser apresentada como mostrado na figura 2.2.

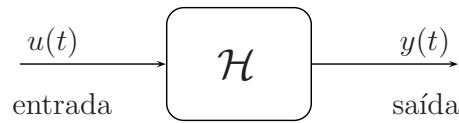


Figura 2.2: Modelo entrada/saída para sistemas físicos

Diagramas como o da figura 2.2 constituem um resumo conciso da relação entre causas e efeitos do sistema [12], [26], pois relacionam a entrada $u(t)$ com a saída $y(t)$ do processo.

Sistemas Estáticos e Dinâmicos

Em geral, o valor atual da saída de um elemento do sistema é o resultado do que aconteceu com esse elemento no passado, bem como o que o afeta no presente [12]. Define-se elemento dinâmico como sendo aquele cuja saída corrente depende das entradas e saídas passadas [17]. Analogamente, um elemento estático é um cuja saída em um dado instante de tempo depende somente da entrada naquele instante de tempo.

Os termos estático e dinâmico são usados para distinguir situações em que mudanças ocorrem (ou não ocorrem) nos elementos que estão sujeitos a esse efeito. A saída de um elemento estático pode mudar com o tempo somente se a entrada variar, e não mudará se a entrada permanecer constante. Entretanto se a entrada de um elemento dinâmico for constante ou removida, sua saída correspondente ainda pode mudar.

Modelos Discretos e Contínuos

Em alguns casos é conveniente visualizar a dinâmica do sistema em termos de uma variável contínua no tempo. Em outros casos, uma variável discreta para representação do tempo é utilizada [12]. Em engenharia, uma situação em que o uso de modelos discretos no tempo ocorre é quando o sistema contém um computador digital com o propósito de medição e controle. Um computador digital não é capaz de realizar medidas continuamente, mas somente amostras das variáveis medidas.

Modelos contínuos são aqueles para os quais o tempo toma valores no conjunto dos números reais. A medição do tempo é realizada de forma contínua, e não amostrada, como feita para modelos discretos.

Sistemas Variantes e Invariantes no Tempo

Alguns sistemas podem mudar sua estrutura com o tempo, outros não. Aqueles cuja estrutura permanece fixa com o passar do tempo são chamados invariantes no tempo, e aqueles cuja estrutura ou cujos parâmetros podem mudar são denominados variantes no tempo [30].

Formalmente falando, um sistema caracterizado pela relação $y(t) = \mathcal{H}[u(t)]$ é invariante no tempo se para qualquer T real, então $y(t - T) = \mathcal{H}[u(t - T)]$.

Sistemas Lineares e Não-Lineares

Um sistema, e conseqüentemente o operador \mathcal{H} que o define, é dito linear se o seguinte enunciado se verificar [14]:

se a resposta para uma excitação $u_1(t)$ for $y_1(t)$, e a resposta para uma excitação $u_2(t)$ for $y_2(t)$, então a saída para uma entrada $au_1(t) + bu_2(t)$ é $ay_1(t) + by_2(t)$, para quaisquer constantes a e b .

Por outro lado, um sistema é dito não-linear se o enunciado acima, conhecido como princípio da superposição, não valer para quaisquer sinais no tempo $u(t)$ e $y(t)$. Em outras palavras, um sistema linear é aquele para o qual a resposta a uma combinação linear de entradas é a mesma combinação linear das respostas a cada entrada individual.

No contexto deste trabalho, os sistemas lineares são representados pelo conhecido modelo de resposta ao impulso. Será visto posteriormente na seção 2.2.1 que uma extensão daquele modelo pode ser utilizada para caracterizar também sistemas não-lineares.

Sistemas Estocásticos e Determinísticos

Algumas vezes os valores das entradas e saídas de um sistema, ou os coeficientes de seu modelo, estão sujeitos a certos tipos de incerteza. Se essa incerteza não puder ser caracterizada por valores determinísticos, torna-se necessário utilizar um modelo estocástico. Nesse tipo de modelo, os coeficientes e as entradas são descritas em termos de variáveis aleatórias, que por sua vez são caracterizadas pelas respectivas distribuições de probabilidade [12].

Já nos sistemas determinísticos, não existem incertezas em relação aos parâmetros de seus modelos, e suas entradas e saídas podem ser descritas por funções determinísticas. Os coeficientes dos estímulos de entrada e dos modelos são completamente conhecidos.

Foram apresentados acima os principais tipos de modelos utilizados na representação de sistemas físicos. No presente trabalho, um enfoque será dado aos sistemas dinâmicos, discretos, invariantes no tempo e determinísticos. As equações que serão apresentadas no desenvolvimento do trabalho no Capítulo 4 são válidas para sistemas lineares e não-lineares.

O desenvolvimento do presente trabalho será apresentado considerando-se que os sistemas possam ser representados por modelos discretos no tempo. Isso se deve ao fato de que as simulações serão realizadas em computadores digitais, nos quais simulações digitais são mais facilmente implementadas. Entretanto os modelos contínuos equivalentes podem ser estudados e derivados a partir do que será apresentado, por simples analogia.

2.2 Modelos de Volterra

A representação de sistemas que será apresentada nesta seção foi introduzida na matemática pelo italiano Vito Volterra, no final do século XIX. Entretanto, o primeiro uso da série de Volterra para o estudo de sistemas não-lineares foi feito por Norbert Wiener [13], em meados de 1930. Volterra foi um pesquisador que se interessou em áreas como funções integrais, que levam seu nome. Além disso, ele se preocupou com o estudo de equações diferenciais não-lineares (Lotka/Volterra), propostas para descrever a dinâmica da populações em sistemas presa/predador.

A seguir, serão apresentadas a série de Volterra e suas aplicações em identificação de sistemas não-lineares. Em seguida, serão abordados alguns aspectos matemáticos que mostram a utilização da série de Volterra como ferramenta para a modelagem de sistemas.

2.2.1 Séries de Volterra

O princípio da superposição enunciado na seção 2.1.1 se aplica a qualquer sistema linear. Nestes casos, o conhecimento de uma função temporal que caracterize o processo (por exemplo, a resposta ao impulso) é suficiente para determinar a saída para qualquer sinal de entrada. Na descrição de processos não-lineares, também é desejável se obter uma função temporal que o caracterize. O conhecimento dessa função pode oferecer as seguintes vantagens [14]:

- fornecer uma relação explícita entre os sinais de entrada e saída;
- permitir o estudo dos sistemas excitados por sinais aleatórios;
- facilitar o estudo de combinação de sistemas (em cascata, em paralelo, etc.).

A utilização de uma série de Volterra permite caracterizar um processo não-linear de forma a se ter uma idéia do seu significado físico [24]. Essa análise tem a grande vantagem conceitual de descrever um sistema não-linear como uma generalização direta do caso linear. Assim, a modelagem de um sistema por séries de Volterra considera os sistemas lineares como um subconjunto do caso não-linear. Descrever um sistema não-linear por meio de funções temporais que o caracterizem completamente é uma vantagem que pode ser utilizada na solução de problemas referentes a esses sistemas. Em outras palavras, as séries de Volterra tornam possível usar conceitos definidos somente para sistemas lineares, de forma a identificar também os processos não-lineares [14].

Uma série de Volterra descreve a saída de um sistema não-linear como a soma das respostas de operadores de primeira ordem, de segunda ordem, terceira ordem, e assim por diante. Cada operador é descrito, no domínio do tempo ou da frequência, por uma função multivariável chamada *kernel* de Volterra. Seja o caso de um sistema com uma única entrada e única saída (SISO – *Single Input Single Output*). Denotando o sinal de entrada por $u(k)$, matematicamente pode-se escrever a saída

$y(k)$ do processo como [14], [26]:

$$\begin{aligned}
 y(k) = & \sum_{\tau_1=-\infty}^{+\infty} h_1(\tau_1)u(k - \tau_1) + \sum_{\tau_1=-\infty}^{+\infty} \sum_{\tau_2=-\infty}^{+\infty} h_2(\tau_1, \tau_2)u(k - \tau_1)u(k - \tau_2) + \\
 & + \cdots + \sum_{\tau_1=-\infty}^{+\infty} \cdots \sum_{\tau_n=-\infty}^{+\infty} h_n(\tau_1, \dots, \tau_n)u(k - \tau_1) \cdots u(k - \tau_n)
 \end{aligned} \tag{2.2}$$

Ou ainda:

$$y(k) = y_{\text{linear}} + y_{\text{quadrático}} + y_{\text{cúbico}} + \cdots \tag{2.3}$$

$$y(k) = \mathcal{H}_1[u(k)] + \mathcal{H}_2[u(k)] + \mathcal{H}_3[u(k)] + \cdots \tag{2.4}$$

onde:

$$\mathcal{H}_n[u(k)] = \sum_{\tau_1=-\infty}^{+\infty} \cdots \sum_{\tau_n=-\infty}^{+\infty} h_n(\tau_1, \dots, \tau_n)u(k - \tau_1) \cdots u(k - \tau_n) \tag{2.5}$$

A função $h_n(\tau_1, \dots, \tau_n)$ é denominada *kernel* de Volterra de ordem n . Se o processo em questão for causal (i.e., se a saída em um certo instante de tempo depender da entrada unicamente em instantes anteriores), a equação (2.5) é considerada somente para $\tau_l \geq 0$ (com $l = 1, 2, \dots, n$). A forma funcional do tipo mostrado em (2.4) e (2.5) indica que a saída $y(k)$ pode ser vista como uma função dos valores passados de $u(k)$.

A equação (2.2) é conhecida como Série de Volterra e depende, além da entrada $u(k)$, das funções $h_n(\tau_1, \tau_2, \dots, \tau_n)$. É também uma generalização direta do modelo linear de convolução, que descreve a saída de sistemas lineares em termos da entrada e da resposta ao impulso:

$$y(k) = \sum_{\tau=-\infty}^{+\infty} h(\tau)u(k - \tau) \tag{2.6}$$

em que $h(\tau)$ é a resposta ao impulso do sistema linear. A equação (2.6) caracteriza os modelos lineares com dinâmica invariante no tempo [13], pois relaciona o sinal de entrada $u(k)$ com a saída $y(k)$. A dinâmica do sistema fica completamente determinada pela resposta ao impulso $h(k)$.

Desde que o termo $\mathcal{H}_n[u(k)]$ em (2.5) contém contribuições de grau n para a entrada $u(k)$, nota-se que os modelos de Volterra representam uma generalização não-linear da equação de convolução (2.6). Neste contexto, n representa o grau de não-linearidade do sistema. Pela equação (2.2), os *kernels* h_n são generalizações de dimensão n ($n = 1, 2, \dots, \infty$) da função de resposta ao impulso de um sistema linear.

Assim, a resposta ao impulso $h(\tau)$ do sistema linear é generalizada, e estendida para os *kernels*

$$h_1(\tau_1) \quad h_2(\tau_1, \tau_2) \quad h_3(\tau_1, \tau_2, \tau_3) \quad \dots$$

A função $h_n(\tau_1, \dots, \tau_n)$ pode ser interpretada como a resposta ao impulso de dimensão n , e a função \mathcal{H}_n em (2.5) é chamada funcional de ordem n . Será discutido posteriormente que em muitos casos, a série pode ser truncada depois de alguns termos. Isso pode ser feito quando se garante a convergência da série na equação (2.2).

Tratada de uma forma esquemática, a série de Volterra pode ser ilustrada como na figura 2.3, em que a saída $y(k)$ é vista como uma soma das contribuições de cada funcional \mathcal{H}_j (com $j = 1, 2, \dots, n$) aplicado à entrada $u(k)$.

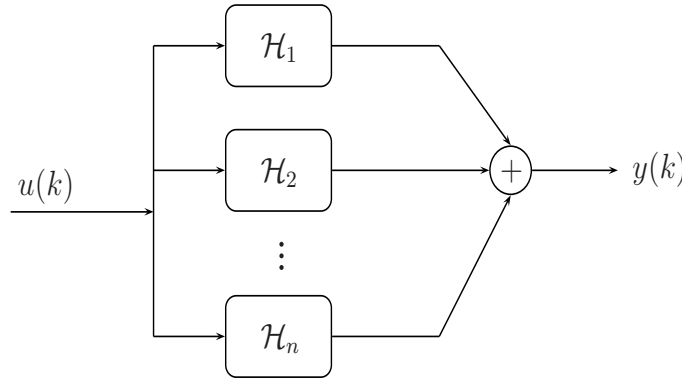


Figura 2.3: Representação esquemática de um sistema caracterizado por uma série de Volterra

A série de Volterra (2.4) pode ser interpretada como uma série de potências com memória. Para entender esse ponto de vista, considere que a entrada $u(k)$ seja alterada por um fator multiplicador a , tal que a nova entrada seja $au(k)$. A nova saída calculada pelas equações (2.2) e (2.4) fica:

$$\begin{aligned} y(k) &= \sum_{n=1}^{\infty} \mathcal{H}_n[au(k)] \\ &= \sum_{n=1}^{\infty} a^n \mathcal{H}_n[u(k)] \end{aligned} \quad (2.7)$$

que é uma série de potências em a . A equação (2.7) é uma série com memória [26] porque a representação funcional de (2.5) é uma generalização da convolução (2.6).

Para sistemas causais, a representação do modelo (2.2) é do tipo:

$$y(k) = \mathcal{O}(\{u(\tau)\}_{\tau=-\infty}^{\tau=k}) \quad (2.8)$$

em que \mathcal{O} é um operador (não-linear, no caso mais geral) que representa o sistema modelado pela equação (2.2). Em termos matemáticos, \mathcal{O} pode ser um funcional,

ou seja, é uma operação sobre uma função onde o resultado é um número.

Como a série de Volterra é uma soma infinita, devem haver condições de convergência associadas, que devem ser satisfeitas para garantir que a representação de um sistema via série de Volterra tenha significado físico. Usualmente, essas condições exigem que a entrada $u(k)$ seja limitada. Em alguns casos, é possível investigar a convergência da série de Volterra usando uma série de potências [26]. Entretanto, raramente a série é obtida em uma forma fechada (expressão analítica). Na maioria dos casos, somente uma série truncada pode ser usada, e que forneça uma estimativa satisfatória da representação quando são desprezados os termos de ordem superior. A série funcional em (2.4) é única e converge dentro de um raio de convergência $r > 0$ que satisfaça $|u(k)| < r$. Se a série de Taylor que representa o operador \mathcal{O} da equação (2.8) tiver convergência garantida, é possível truncá-la a partir de um certo número de termos. Detalhes formais sobre a convergência da série de Volterra não serão discutidos nesse trabalho, e podem ser encontrados em [24].

A partir do último termo considerado na aproximação da série, as parcelas de maior ordem podem ser desprezadas. Essa simplificação reduz o esforço computacional requerido em simulações, e também é capaz de representar com boa precisão grande parte dos operadores \mathcal{O} não-lineares. Ainda, qualquer operador contínuo, invariante no tempo e estável pode ser aproximado por uma série de Volterra [3].

A representação causal da aproximação com η termos da série em (2.2) é escrita como:

$$y(k) = \sum_{n=1}^{\eta} \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_n=0}^{\infty} h_n(\tau_1, \dots, \tau_n) \prod_{j=1}^n u(k - \tau_j) \quad (2.9)$$

Na modelagem de sistemas reais e estáveis, os *kernels* de Volterra apresentam três propriedades importantes [14]:

- $h_n(\tau_1, \tau_2, \dots, \tau_n)$ podem ser construídos de forma a serem simétricos, se já não o forem
- $\lim_{\tau_l \rightarrow \infty} h_n(\tau_1, \tau_2, \dots, \tau_n) = 0 \quad l = 1, 2, \dots, n$
- $h_n(\tau_1, \tau_2, \dots, \tau_n) = 0 \quad \forall \tau_i < 0$

O primeiro item acima é verificado fazendo-se uma troca cíclica dos argumentos do *kernel* assimétrico, e será estudado na seção 2.4. Já a segunda propriedade surge apenas na modelagem de sistemas em que a resposta a uma entrada desaparece com o tempo, até não contribuir significativamente para a saída do sistema (sistemas *fading memory*). E o terceiro item vem do fato de se admitir que o sistema seja causal e com condições iniciais nulas, isto é, a saída de um sistema físico realizável depende meramente do passado da entrada.

Considere que determinado sistema seja causal, e que seja possível modelá-lo por meio de uma série de Volterra. Admitindo-se que o sinal de entrada $u(k)$ seja nulo

para $k < 0$, a resposta à entrada $u(k)$ pode ser expressa como:

$$\begin{aligned}
 y(k) = & \sum_{\tau_1=0}^k h_1(\tau_1)u(k - \tau_1) + \sum_{\tau_1=0}^k \sum_{\tau_2=0}^k h_2(\tau_1, \tau_2)u(k - \tau_1)u(k - \tau_2) + \\
 & + \cdots + \sum_{\tau_1=0}^k \cdots \sum_{\tau_n=0}^k h_n(\tau_1, \dots, \tau_n)u(k - \tau_1) \cdots u(k - \tau_n)
 \end{aligned} \tag{2.10}$$

Os modelos considerados neste trabalho são SISO, tal como o modelo discreto em (2.10). Não serão estudados os modelos MIMO (*Multiple Input - Multiple Output*), que podem ser obtidos por simples analogia a partir do que foi apresentado anteriormente [14].

2.3 O *Kernel* de Volterra de Segunda Ordem

Conforme discutido na seção 2.2.1, a equação (2.6) define um operador linear, onde a resposta a uma combinação linear de sinais de entrada é a mesma combinação linear da resposta a cada sinal individualmente. Esse é o clássico modelo de resposta ao impulso, onde a função $h_1(\tau_1)$ pode ser vista também como o *kernel* de Volterra de primeira ordem.

Na presente seção, essa abordagem é estendida para o estudo mais aprofundado do *kernel* de segunda ordem de Volterra $h_2(\tau_1, \tau_2)$. Para esse caso, a resposta a uma combinação linear de sinais é uma operação bilinear aplicada ao sinal de entrada, conforme apresenta a seção 2.3.1.

Voltando ao modelo dado em (2.10), a segunda parcela da série é dada por:

$$y_2(k) = \sum_{\tau_1=0}^k \sum_{\tau_2=0}^k h_2(\tau_1, \tau_2)u(k - \tau_1)u(k - \tau_2) \tag{2.11}$$

que representa a contribuição do *kernel* de segunda ordem $h_2(k_1, k_2)$ para o sinal de saída $y(k)$. Em notação de operadores, pode-se escrever ainda $y_2(k) = \mathcal{H}_2[u(k)]$.

2.3.1 Resposta ao Impulso Bidimensional

O *kernel* de Volterra de segunda ordem $h_2(\tau_1, \tau_2)$ caracteriza o operador de Volterra de segunda ordem $\mathcal{H}_2[\cdot]$, isto é, a resposta deste a qualquer sinal de entrada pode ser determinada, uma vez que o *kernel* seja conhecido. Surge então uma questão sobre a interpretação física do *kernel*. Para um entendimento maior a esse respeito, considere o caso em que se deseja determinar a resposta ao impulso para um sistema com modelo de Volterra de grau 2. Fazendo $u(k) = \delta(k)$, onde $\delta(k)$ representa a

função impulso unitário, a equação (2.11) é reescrita como:

$$\begin{aligned}
 y_2(k) &= \mathcal{H}_2[\delta(k)] \\
 &= \sum_{\tau_1=0}^k \sum_{\tau_2=0}^k h_2(\tau_1, \tau_2) \delta(k - \tau_1) \delta(k - \tau_2) \\
 &= h_2(k, k)
 \end{aligned} \tag{2.12}$$

Observe que a resposta ao impulso obtida acima não caracteriza genericamente um sistema de segunda ordem. Isto é, ela é determinada pelo *kernel* de segunda ordem somente ao longo da linha $\tau_1 = \tau_2$.

Entretanto, o *kernel* $h_2(\tau_1, \tau_2)$ pode ser interpretado como uma resposta ao impulso bidimensional, conforme apresentado a seguir. Inicialmente, considere $y(k)$ como sendo a resposta de um operador de Volterra de segunda ordem a uma entrada $u(k)$ composta da soma $u_1(k) + u_2(k)$. A resposta $y(k)$ pode ser expressa da forma:

$$\begin{aligned}
 y(k) &= \mathcal{H}_2[u(k)] \\
 &= \mathcal{H}_2[u_1(k) + u_2(k)] \\
 &= \mathcal{H}_2[u_1(k)] + 2\mathcal{H}_2\{u_1(k), u_2(k)\} + \mathcal{H}_2[u_2(k)]
 \end{aligned} \tag{2.13}$$

em que $\mathcal{H}_2[\cdot]$ é a representação funcional do operador de Volterra de segunda ordem, e também:

$$\mathcal{H}_2\{u_1(k), u_2(k)\} = \sum_{\tau_1=0}^k \sum_{\tau_2=0}^k h_2(\tau_1, \tau_2) u_1(k - \tau_1) u_2(k - \tau_2) \tag{2.14}$$

é denominado operador bilinear de Volterra [26]. Esse termo cruzado $\mathcal{H}_2\{u_1(k), u_2(k)\}$ pode ser isolado em (2.13) para se obter o operador bilinear de segunda ordem.

Os exemplos que serão apresentados a seguir ilustram a utilização de modelos de Volterra como ferramenta matemática na modelagem de sistemas dinâmicos.

2.3.2 Exemplo 2.1

A tarefa de identificar um sistema utilizando série de Volterra consiste em estimar os *kernels* $h_n(\tau_1, \dots, \tau_n)$ do modelo. Entretanto esse problema é complexo, já que os *kernels* são funções n -dimensionais não-parametrizadas [24]. O exemplo a seguir ilustra como estimar o *kernel* de segunda ordem para um determinado sistema.

Considere a conexão não-linear de dois sub-sistemas, mostrados na figura 2.4. Os dois sub-sistemas dessa figura são lineares, caracterizados por suas respectivas respostas ao impulso, $h_a(k)$ e $h_b(k)$. Eles são descritos pela clássica equação de convolução:

$$y_i(k) = \sum_{\tau=0}^k h_i(\tau) u(k - \tau) \quad i = a, b \tag{2.15}$$

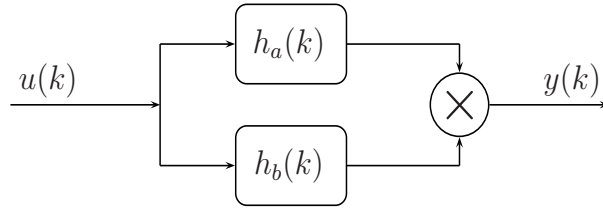


Figura 2.4: Sistema não-linear com *kernel* de segunda ordem

O sistema completo é descrito por:

$$\begin{aligned}
 y(k) &= y_a(k)y_b(k) \\
 &= \sum_{\tau_1=0}^k h_a(\tau_1)u(k-\tau_1) \sum_{\tau_2=0}^k h_b(\tau_2)u(k-\tau_2) \\
 &= \sum_{\tau_1=0}^k \sum_{\tau_2=0}^k h_a(\tau_1)u(k-\tau_1)h_b(\tau_2)u(k-\tau_2) \quad (2.16)
 \end{aligned}$$

Observando-se a equação acima, nota-se claramente que o *kernel* de segunda ordem do sistema é:

$$h_2(\tau_1, \tau_2) = h_a(\tau_1)h_b(\tau_2) \quad (2.17)$$

Portanto, o sistema da figura 2.4 pode ser representado por meio de um modelo de Volterra de segunda ordem, conforme visto acima. O *kernel* de Volterra (de segunda ordem) para esse sistema é o produto das respostas ao impulso individuais de cada um dos dois sub-sistemas que compõem o sistema completo.

2.3.3 Exemplo 2.2

Seja um sistema não-linear, representado pelo diagrama de blocos ilustrado na figura 2.5.

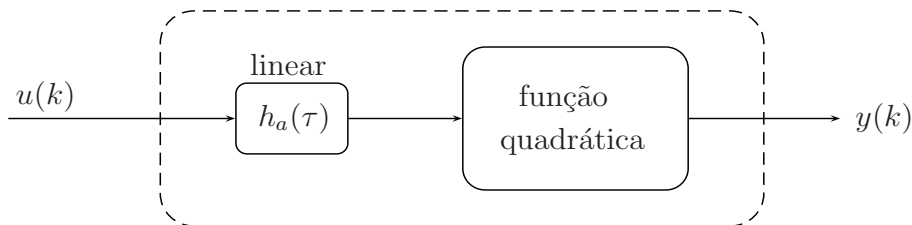


Figura 2.5: Representação em diagrama de blocos de um determinado sistema

Este sistema é composto por um primeiro estágio linear, com resposta ao impulso $h_a(\tau)$. O segundo estágio consiste de um dispositivo com característica quadrática,

conhecido como *squarer*. O sinal na sua saída é igual ao quadrado do sinal de entrada para todo instante de tempo, i.e., se a entrada do squarer for $z(k)$, sua saída é $z^2(k)$. Portanto:

$$\begin{aligned}
 y(k) &= \left[\sum_{\tau=0}^k h_a(\tau)u(k-\tau) \right]^2 \\
 &= \sum_{\tau_1=0}^k h_a(\tau_1)u(k-\tau_1) \sum_{\tau_2=0}^k h_a(\tau_2)u(k-\tau_2) \\
 &= \sum_{\tau_1=0}^k \sum_{\tau_2=0}^k h_a(\tau_1)h_a(\tau_2)u(k-\tau_1)u(k-\tau_2) \quad (2.18)
 \end{aligned}$$

A equação (2.18) é do tipo apresentado em (2.11), em que $h_2(\tau_1, \tau_2) = h_a(\tau_1)h_a(\tau_2)$ é o *kernel* de Volterra de segunda ordem para o sistema mostrado na figura 2.5. Deve ser notado que modelos como o da equação (2.18) são descritos por uma operação de convolução bidimensional. Observa-se que modelos que envolvem *kernels* de dimensão superior possuem elevado custo computacional.

Conforme discutido na seção 2.2.1, os *kernels* são simétricos, ou podem ser feitos simétricos. Como obter a versão simétrica de um *kernel* assimétrico é o objetivo da seção seguinte.

2.4 Simetrização do *kernel* de Volterra

Na análise dos *kernels* de Volterra pode haver a necessidade de se realizar operações que envolvam reordenamento dos argumentos τ 's em $h_n(\tau_1, \dots, \tau_n)$. Qualquer *kernel* assimétrico pode ser simetrizado, de maneira que a análise pode ser realizada considerando-se somente sua versão simétrica, sem qualquer perda de generalidade [26]. O objetivo principal da simetrização é que considerar determinado *kernel* simétrico requer a obtenção de um número menor de parâmetros. Isso vem a simplificar os cálculos computacionais necessários.

A regra geral para simetrizar um *kernel* assimétrico é dada em [24], e pode ser escrita como:

$$h_n^{\text{sim}}(\tau_1, \dots, \tau_n) = \frac{1}{n!} \sum_{\pi(\cdot)} h_n^{\text{assim}}(\tau_{\pi(1)}, \dots, \tau_{\pi(n)}) \quad (2.19)$$

onde $\pi(\cdot)$ denota todas as permutações dos inteiros $1, 2, \dots, n$.

A partir dessa definição, um *kernel* de segunda ordem é simétrico se $h_2(\tau_1, \tau_2) = h_2(\tau_2, \tau_1)$. Note que o *kernel* do exemplo 2.2 é simétrico, já que:

$$h_2(\tau_1, \tau_2) = h_a(\tau_1)h_a(\tau_2) = h_a(\tau_2)h_a(\tau_1) = h_2(\tau_2, \tau_1)$$

Como exemplo, suponha que o *kernel* de terceira ordem de um determinado sistema seja dado por:

$$h_3(\tau_1, \tau_2, \tau_3) = g(\tau_1)g(\tau_2)g(\tau_3)f(\tau_1 + \tau_2) \quad (2.20)$$

Aplicando-se a definição dada em (2.19) à equação (2.20), tem-se que:

$$\begin{aligned}
 h_3^{\text{sim}}(\tau_1, \tau_2, \tau_3) &= \frac{1}{6} \left[2g(\tau_2)g(\tau_1)g(\tau_3)f(\tau_2 + \tau_1) + 2g(\tau_3)g(\tau_2)g(\tau_1)f(\tau_2 + \tau_3) + \right. \\
 &\quad \left. + 2g(\tau_1)g(\tau_3)g(\tau_2)f(\tau_1 + \tau_3) \right] \\
 &= \frac{1}{3} g(\tau_2)g(\tau_1)g(\tau_3) \left[f(\tau_2 + \tau_1) + f(\tau_2 + \tau_3) + f(\tau_1 + \tau_3) \right] \quad (2.21)
 \end{aligned}$$

Embora a versão simétrica do *kernel* usualmente seja caracterizada por expressões matemáticas mais complexas que uma versão assimétrica [24], a simetriação pode ser um procedimento importante. Isso se justifica pelo fato de que, em muitos casos, as propriedades dos sistemas podem estar relacionadas de forma mais simples com as propriedades do *kernel* simétrico do que com as propriedades de sua versão assimétrica. Além disso, a determinação de apenas uma das partes simétricas é suficiente para a caracterização completa do *kernel*.

2.5 Resumo

Foram apresentados neste capítulo alguns conceitos importantes no estudo de sistemas físicos, dentre eles o conceito de modelos de Volterra. Neste aspecto, foi ilustrada a utilização daqueles modelos na modelagem de sistemas não-lineares. Suas principais vantagens residem no fato de que aqueles modelos descrevem um sistema não-linear como uma generalização do caso linear, o que permite estender a sistemas não-lineares conceitos definidos somente para os lineares.

De acordo com as discussões apresentadas, principalmente por meio dos exemplos ilustrativos, pôde-se ter uma idéia da dificuldade em estimar os *kernels* de um sistema, especialmente porque os *kernels* são funções n -dimensionais não-parametrizadas. O presente capítulo enfocou também a idéia de simetriação de *kernels* de Volterra, que constitui uma ferramenta matemática capaz de fornecer vantagens na representação computacional de sistemas não-lineares.

Capítulo 3

Bases de Funções Ortonormais

O uso de séries ortonormais de funções na representação de sinais foi estudado primeiramente por Lee e Wiener [13], [34] em meados da década de 1930. Esse tema é um importante tópico de pesquisa nas áreas de automação, controle e processamento de sinais, e vem se constituindo um objeto de estudo amplo nos últimos anos [34]. Em vista das dificuldades encontradas no problema de aproximação de funções, uma abordagem usando desenvolvimento ortogonal se apresenta como uma importante alternativa. Nessa abordagem procura-se representar os *kernels* de um determinado sistema em termos de uma base ortonormal para o espaço de interesse.

O matemático Norbert Wiener desenvolveu uma representação de processos não-lineares fazendo uso das funções ortogonais de Laguerre. As propriedades de ortonormalidade dessas funções resultam em um modelo geral mais simples [26]. Nesse capítulo, é apresentada uma introdução às funções ortonormais discretas. São discutidos os principais conjuntos de funções ortonormais utilizados na representação de sinais. Posteriormente, dá-se maior atenção às funções de Kautz (tema central do presente trabalho) e sua relevância na aproximação de sistemas invariantes no tempo.

3.1 Introdução às Funções Ortogonais

Uma razão para o uso de base de funções ortogonais é que o problema de aproximação correspondente é de solução simples [36]. A idéia de utilizar funções ortonormais em identificação e controle é clássica. Essas funções serão de relevância na abordagem do problema de identificação estudado no presente trabalho. Esta seção discutirá as funções ortogonais discretas e apresentará os principais conjuntos ortonormais utilizados na representação de sinais.

Seja $k = 1, 2, \dots$ uma seqüência crescente de números inteiros. O conjunto de funções $\{p_m(k)\}$ é denominado ortonormal [16] no intervalo I se:

$$\sum_{k \in I} p_m(k)p_n(k) = \begin{cases} 0 & \text{se } m \neq n \\ 1 & \text{se } m = n \end{cases} \quad (3.1)$$

Uma definição equivalente a (3.1) para funções definidas no domínio complexo z é

a seguinte:

$$\frac{1}{2\pi j} \oint_C P_m(z) P_n(z^{-1}) \frac{dz}{z} = \begin{cases} 0 & \text{para } m \neq n \\ 1 & \text{para } m = n \end{cases} \quad (3.2)$$

onde $P_m(z) = \mathcal{Z}[p_m(k)]$ e a região C é o círculo unitário.

A seguir, são apresentados os principais conjuntos de funções ortonormais utilizados em representação de sinais e identificação de sistemas.

3.2 Polinômios Ortogonais Discretos

Algumas funções ortogonais amplamente utilizadas em aproximação de funções são os polinômios de Krawtchouk, Charlier e Meixner [4]. Existem ainda os polinômios discretos de Chebyshev, que não estão relacionados com equações a diferenças, e não serão tratados aqui. O intervalo de ortogonalidade I em (3.1) é $\{0, 1, \dots, L\}$ para os polinômios de Krawtchouk, e $\{0, 1, \dots, \infty\}$ para Charlier e Meixner.

Esses polinômios são geralmente denotados por $p_n(k)$, e estão relacionados com a seguinte equação a diferenças de segunda ordem:

$$f(k+1)p_n(k+1) + [\sigma_n - f(k+1) - g(k)]p_n(k) + g(k)p_n(k-1) = 0 \quad (3.3)$$

onde os coeficientes σ_n e as funções $f(k)$ e $g(k)$ são dados na tabela 3.1.

Tabela 3.1: Coeficientes da equação a diferenças (3.3) para alguns polinômios ortogonais

	Krawtchouk	Charlier	Meixner
Peso $w(k)$	$\binom{L}{k} \rho^k (1-\rho)^{L-k}$	$\frac{e^{-a} a^k}{k!}$	$\frac{b c^k}{k!}$
Parâmetros	ρ	a	b e c
$f(k)$	$\rho(L+1-k)$	a	$c(k+b-1)$
$g(k)$	$(1-\rho)k$	k	k
σ_n	n	n	$n(1-c)$
κ_n^2	$\binom{L}{n} \rho^n (1-\rho)^n$	1	$\frac{n! b}{c^n (1-c)^b}$

Nessa tabela, os parâmetros que caracterizam cada um dos polinômios ortogonais definidos anteriormente são restritos aos intervalos seguintes: ρ e c estão no intervalo $(-1, 1)$ e os parâmetros a e b são números reais positivos.

Os polinômios $p_n(k)$ da equação (3.3) são ortogonais utilizando uma função de ponderação $w(k)$, também mostrada na tabela 3.1. Nesse contexto, a ortonormalidade das funções $p_n(k)$ é caracterizada por:

$$\sum_{k \in I} w(k) p_n(k) p_m(k) = \begin{cases} 0 & \text{se } m \neq n \\ \kappa_n^2 & \text{se } m = n \end{cases} \quad (3.4)$$

A equação (3.3) é geradora dos polinômios de Krawtchouk, Charlier e Meixner. Agora, serão definidas suas respectivas funções.

As funções de Krawtchouk são definidas como:

$$p_n(k) = \binom{L}{n}^{-1/2} \binom{L}{k}^{1/2} p^{(k-n)/2} (\sqrt{1-p})^{L-k-n} \mathcal{K}_n(k) \quad (3.5)$$

As funções de Charlier são definidas como:

$$p_n(k) = a^{n/2} \frac{e^{-a/2} a^{k/2}}{\sqrt{k!}} \mathcal{C}_n(k) \quad (3.6)$$

As funções de Meixner são definidas como:

$$p_n(k) = (-1)^n \frac{c^{(n+k)/2} (1-c)^{b/2} \sqrt{b}}{\sqrt{n!k!b}} \mathcal{M}_n(k) \quad (3.7)$$

onde $\mathcal{K}_n(k)$, $\mathcal{C}_n(k)$ e $\mathcal{M}_n(k)$ são os polinômios de Krawtchouk, Charlier e Meixner respectivamente [4].

As funções apresentadas acima não são geralmente utilizadas na aproximação ortogonal de sistemas dinâmicos. Isso se deve ao fato de que a caracterização dessas funções não está relacionada diretamente com a dinâmica do sistema. Essa condição pode ser obtida utilizando as funções ortogonais generalizadas (em particular as funções de Laguerre e de Kautz), apresentadas a seguir.

3.3 Funções Ortonormais Generalizadas

Suponha que a saída de um filtro digital linear seja a seqüência $\{y(k)\}$, formada a partir de uma seqüência de entrada $\{u(k)\}$. Se o filtro for do tipo recursivo [5], a saída é gerada pela relação:

$$y(k) = \sum_{n=0}^R a_n u(k-n) + \sum_{n=1}^S b_n y(k-n) \quad (3.8)$$

que usa as saídas passadas bem como as entradas passadas e atuais para produzir a saída corrente. O valor de S pode ser qualquer inteiro positivo, e R pode ser zero ou qualquer inteiro positivo.

A partir da equação (3.8), pode-se inferir o seguinte:

$$\sum_{k=0}^{\infty} y(k) z^k = \sum_{n=0}^R a_n \sum_{k=0}^{\infty} u(k-n) z^k + \sum_{n=1}^S b_n \sum_{k=0}^{\infty} y(k-n) z^k \quad (3.9)$$

Usando a seguinte propriedade da transformada Z

$$\begin{aligned}\sum_{k=0}^{\infty} y(k-n)z^k &= z^n Y(z) \\ \sum_{k=0}^{\infty} u(k-n)z^k &= z^n U(z)\end{aligned}$$

pode-se reescrever a equação (3.9) como:

$$Y(z) = \sum_{n=0}^R a_n z^n U(z) + \sum_{n=1}^S b_n z^n Y(z) \quad (3.10)$$

ou ainda:

$$Y(z) = \frac{U(z) \sum_{n=0}^R a_n z^n}{1 - \sum_{n=1}^S b_n z^n} \quad (3.11)$$

Quando o sinal de entrada é $u(k) = \delta(k)$, onde $\delta(k)$ representa a função impulso unitário, tem-se que $U(z) = \mathcal{Z}[\delta(k)] = 1$. A função de transferência $\Gamma(z) = Y(z)/U(z)$ do filtro é então dada por:

$$\Gamma(z) = \frac{\sum_{n=0}^R a_n z^n}{1 - \sum_{n=1}^S b_n z^n} \quad (3.12)$$

Uma outra maneira de se escrever $\Gamma(z)$ é a forma fatorada

$$\Gamma(z) = \frac{a_0 \prod_{n=1}^R (1 - \alpha_n z)}{\prod_{n=1}^S (1 - \beta_n z)} \quad (3.13)$$

onde os coeficientes α_n e β_n devem ser selecionados de tal forma que satisfaçam a condição de ortogonalidade dada na equação (3.2).

Quando tem-se β_n como sendo um pólo complexo, a equação anterior resulta

em um conjunto mais fácil de ser implementado computacionalmente. É possível provar [5] então que as funções $\Gamma(z)$ ficam da forma:

$$\Gamma(z) = \frac{C(z - a_1) \prod_{n=1}^{j-1} (\beta_n - z)(\bar{\beta}_n - z)}{\prod_{n=1}^j (1 - \beta_n z)(1 - \bar{\beta}_n z)} \quad (3.14)$$

onde C é uma constante que depende de α_n e β_n , e $\bar{\beta}_n$ representa o complexo conjugado de β_n . As funções em (3.14) são denominadas funções ortonormais generalizadas. Essa série de funções constitui, sob certas condições, uma base completa em $[0, \infty)$. Essas condições envolvem [25] basicamente a exigência $|\beta_n| < 1$. Alguns casos particulares dessas funções são as funções de Laguerre e de Kautz, objeto de estudo das seções 3.4 e 3.5.

3.4 Funções de Laguerre

A equação (3.14) para $\beta_n = c$, onde c é um número real, representa o caso especial de Laguerre. As funções de Laguerre são definidas em termos de sua transformada Z como segue [36]:

$$\Phi_n(z) = \sqrt{1 - c^2} \frac{z}{z - c} \left(\frac{1 - cz}{z - c} \right)^{n-1} \quad n = 1, 2, \dots \quad (3.15)$$

onde o parâmetro c presente na definição (3.15) é conhecido como pólo de Laguerre. Esse parâmetro é arbitrário, exceto pelo fato de que deve estar dentro do círculo unitário, ou seja, deve-se ter $|c| < 1$.

Uma vantagem das funções de Laguerre é que suas transformadas são funções racionais com uma estrutura repetitiva bastante simples [27]. Desde que cada função de Laguerre $\phi_n(k) = \mathcal{Z}^{-1}[\Phi_n(z)]$ é admitida ser nula para $k < 0$, torna-se possível construir um sistema causal que seja linear e invariante no tempo, com resposta ao impulso igual a $\phi_n(k)$. Um diagrama de blocos que sintetiza essa idéia é obtido diretamente a partir de (3.15), e é mostrado na figura 3.1.

O caso $c = 0$ representa o modelo FIR – *Finite Impulse Response* [28]. Alocando o parâmetro c próximo aos pólos dominantes do sistema a ser aproximado, o número de termos necessários para representá-lo adequadamente pode ser reduzido consideravelmente. Apesar disso, as funções de Laguerre possuem as seguintes desvantagens:

- sistemas com pólos dominantes situados dispersamente no plano complexo não são muito bem descritos usando-se um número reduzido de funções;
- a presença de pólos que produzam oscilações pode causar problemas computacionais termos da velocidade de decaimento dos *kernels*, já que eles sempre ocorrem aos pares de números complexos conjugados.

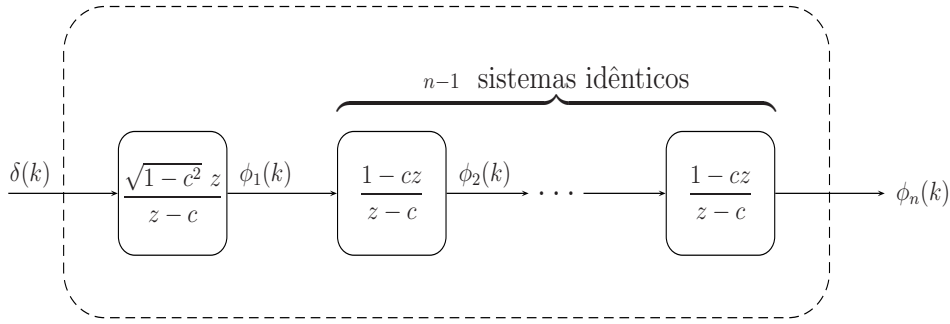


Figura 3.1: Sistema a partir do qual são geradas as funções de Laguerre $\phi_n(k)$, $n = 1, 2, \dots$

Uma alternativa para superar essas dificuldades é utilizar os pólos β como sendo números complexos. Esse é o caso das funções de Kautz, que serão apresentadas na seção 3.5.

3.4.1 Descrição das Funções de Laguerre em Espaço de Estados

Esta seção tem por objetivo apresentar um modelo matemático em espaço de estados para as funções de Laguerre. Uma breve descrição deste esse modelo é dada a seguir.

Qualquer conjunto de funções ortonormais pode ser representado em espaço de estados como segue [6]:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}u(k+1) \quad (3.16)$$

em que $u(k)$ é o sinal de entrada e $\mathbf{x}(k) = [x_1(k) \ x_2(k) \ \dots \ x_n(k)]'$ representa o vetor de estados ortonormais.

A equação a diferenças (3.16) é a representação do modelo em espaço de estados de um sistema linear invariante no tempo, com entrada $u(k)$ e estados $x_j(k)$ ($j = 1, 2, \dots, n$). No caso em que o sinal de entrada do referido sistema é a função impulso unitário, ou seja $u(k) = \delta(k)$, o estado $\mathbf{x}(k)$ representa o vetor contendo as funções ortonormais $\phi(k)$. Resolvendo-se então para $\phi(k)$, é direto mostrar que a solução de (3.16) é dada por:

$$\phi(k) = \mathbf{A}^k \mathbf{B} \quad k \geq 0 \quad (3.17)$$

Para as funções de Laguerre em particular, definidas na equação (3.15), as matrizes \mathbf{A} e \mathbf{B} são dadas por [20]:

$$\mathbf{A} = \begin{bmatrix} c & 0 & 0 & \dots & 0 \\ 1-c^2 & c & 0 & \dots & 0 \\ (-c)(1-c^2) & 1-c^2 & c & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (-c)^{n-2}(1-c^2) & (-c)^{n-3}(1-c^2) & \dots & \dots & c \end{bmatrix} \quad (3.18)$$

$$\mathbf{B} = \sqrt{1-c^2} [1 \quad -c \quad (-c)^2 \quad \dots \quad (-c)^{n-1}]' \quad (3.19)$$

A solução (3.17) admite como condição inicial $\phi(0) = \mathbf{B}$. Esse fato também pode ser verificado usando-se a equação (3.15), que permite concluir que para cada função de Laguerre $\phi_n(k)$ no domínio do tempo, tem-se:

$$\phi_n(0) = \lim_{z \rightarrow \infty} \Phi_n(z) = \sqrt{1-c^2} (-c)^{n-1} \quad (3.20)$$

Conseqüentemente, o vetor de estados $\phi(0)$ pode ser escrito como:

$$\phi(0) = \sqrt{1-c^2} [(-c)^0 \quad (-c)^1 \quad (-c)^2 \quad \dots \quad (-c)^{n-1}]' = \mathbf{B}$$

Portanto, as n primeiras funções de Laguerre podem ser obtidas pela equação (3.17).

3.5 Funções de Kautz

Um outro caso particular da equação (3.14) ocorre quando $\beta_n = \beta$, $\forall n$ ($\beta \in \mathbb{C}$). Para esse caso, tem-se as funções discretas de Kautz $\psi_l(k)$, definidas através de suas transformadas Z como [36]:

$$\begin{aligned} \Psi_{2n}(z) &= \frac{\sqrt{(1-c^2)(1-b^2)} z}{z^2 + b(c-1)z - c} \left(\frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c} \right)^{n-1} \\ \Psi_{2n-1}(z) &= \frac{\sqrt{1-c^2} z(z-b)}{z^2 + b(c-1)z - c} \left(\frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c} \right)^{n-1} \end{aligned} \quad (3.21)$$

em que $n = 1, 2, \dots$ e as constantes b e c são parâmetros reais que caracterizam a função $\Psi_l(z)$, sendo $b = (\beta + \bar{\beta})/(1 + \beta\bar{\beta})$ e $c = -\beta\bar{\beta}$. Os parâmetros estão restritos a $|b| < 1$ e $|c| < 1$, e β é denominado o pólo de Kautz. A definição acima permite deduzir que:

$$\Psi_{2n-1}(z) = \frac{z-b}{\sqrt{1-b^2}} \Psi_{2n}(z) \quad (3.22)$$

e essa relação será utilizada no Capítulo 4.

A partir da definição das funções de Kautz em (3.21), é possível construir um diagrama de blocos equivalente, como mostrado na figura 3.2.

O diagrama ilustrado na figura 3.2 consiste de uma conexão de um sistema causal com função

$$A(z) = \frac{\sqrt{(1-c^2)(1-b^2)} z}{z^2 + b(c-1)z - c}$$

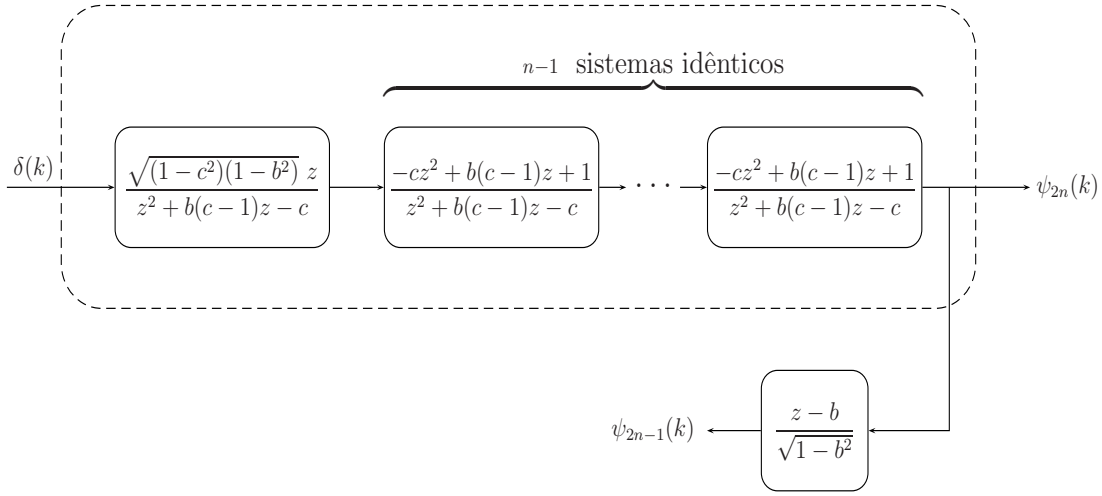


Figura 3.2: Sistema a partir do qual se obtém as funções de Kautz $\psi_{2n}(k)$ e $\psi_{2n-1}(k)$, para $n = 1, 2, \dots$

em cascata com $n - 1$ sistemas causais, cada um com função de transferência

$$B(z) = \frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c}$$

As funções discretas no tempo são escritas como $\psi_l(k)$, ou como $\psi_l(k, b, c)$ sempre que se desejar exibir os parâmetros. O conjunto $\{\psi_n\}$ é completo no espaço de Lebesgue $\ell^2[0, \infty)$ para $|\beta| < 1$, de forma que qualquer sinal de energia finita pode ser aproximado com uma precisão pré-estabelecida, fazendo-se o truncamento de uma expansão infinita [34].

3.5.1 Descrição das Funções de Kautz em Espaço de Estados

As funções de Kautz definidas pelas expressões em (3.21) também possuem uma representação em espaço de estados, na forma geral (3.16). Para as funções de Kautz, as matrizes \mathbf{A} e \mathbf{B} são dadas por [20]:

$$\mathbf{A} = T_N A_N T_N^{-1} \quad (3.23)$$

$$\mathbf{B} = T_N B_N \quad (3.24)$$

onde:

$$T_N = \begin{bmatrix} T_1 & & & \\ & T_1 & & \\ & & \ddots & \\ & & & T_1 \end{bmatrix} \quad (3.25)$$

com T_N contendo N blocos, formados por:

$$T_1 = \begin{bmatrix} -m_1 n_1 & n_1 \\ n_2 & -m_2 n_2 \end{bmatrix} \quad (3.26)$$

e

$$A_N = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ A_2 & A_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ A_n & A_{n-1} & \cdots & A_1 \end{bmatrix} \quad (3.27)$$

$$\mathbf{B} = [B_1 \ B_2 \ \cdots \ B_n]' \quad (3.28)$$

com

$$A_1 = \begin{bmatrix} 0 & 1 \\ r^2 & 2\mathbf{Re}(\beta) \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 \\ 1 - r^4 & 2(r^2 - 1)\mathbf{Re}(\beta) \end{bmatrix}$$

$$A_i = \begin{bmatrix} 0 & 0 \\ 0 & r^2 \end{bmatrix}^{i-2} A_2 \quad i \geq 2 \quad (3.29)$$

$$B_i^T = \begin{bmatrix} 0 & 0 \\ 0 & r^2 \end{bmatrix}^{i-1} \begin{bmatrix} 0 \\ \sqrt{1 - r^2} \end{bmatrix} \quad (3.30)$$

e os seguintes parâmetros:

$$r = \sqrt{\mathbf{Re}(\beta)^2 + \mathbf{Im}(\beta)^2} \quad (3.31)$$

$$m_1 = \mathbf{Re}(\beta) \quad (3.32)$$

$$m_2 = \mathbf{Re}(\beta) \left[\frac{1 - \mathbf{Re}(\beta)^2 - \mathbf{Im}(\beta)^2}{1 - \mathbf{Re}(\beta)^2 + \mathbf{Im}(\beta)^2} \right] \quad (3.33)$$

$$n_1^2 = \frac{[1 + \mathbf{Re}(\beta)^2 + \mathbf{Im}(\beta)^2]^2 - 4\mathbf{Re}(\beta)^2}{[1 + m_1^2][1 + \mathbf{Re}(\beta)^2 + \mathbf{Im}(\beta)^2]^2 - 4m_1\mathbf{Re}(\beta)} \quad (3.34)$$

$$n_2^2 = \frac{[1 + \mathbf{Re}(\beta)^2 + \mathbf{Im}(\beta)^2]^2 - 4\mathbf{Re}(\beta)^2}{[1 + m_2^2][1 + \mathbf{Re}(\beta)^2 + \mathbf{Im}(\beta)^2]^2 - 4m_2\mathbf{Re}(\beta)} \quad (3.35)$$

onde $\mathbf{Re}(\beta)$ e $\mathbf{Im}(\beta)$ representam a parte real e imaginária do pólo β de Kautz, respectivamente.

3.6 Resumo

Abordou-se neste capítulo algumas das principais funções ortonormais utilizadas em identificação de sistemas. Viu-se que alguns dos aspectos mais importantes quando se escolhe a base de funções a ser utilizada são:

- É vantajoso usar o menor número possível de funções, a fim de garantir um modelo de menor ordem;
- O erro de aproximação deve tender a zero quando o número de parâmetros tende a infinito.

Nesse contexto, foram apresentadas as funções de Laguerre e de Kautz, que representam um caso particular das funções ortonormais generalizadas. Conforme foi discutido, essas funções são representadas por parâmetros que permitem caracterizá-las como sistemas físicos causais. É importante salientar, no entanto, que a sua utilização somente é válida para sistemas com memória finita, que serão estudados no Capítulo 4.

Capítulo 4

Aproximação de Modelos de Volterra Usando Funções de Kautz

Este capítulo tem por finalidade discutir o problema de expansão ótima de modelos de Volterra. Apresenta-se uma solução sub-ótima para o desenvolvimento dos *kernels* de Volterra usando-se as funções de Kautz, por meio da seleção de parâmetros livres dessas funções. Conforme será discutido, a escolha ótima simultânea desses parâmetros ainda não é conhecida, e assim propõe-se um método sub-ótimo para esta escolha.

Primeiramente, define-se o problema de expansão de *kernels* de Volterra utilizando funções de Kautz. Em seguida, apresenta-se uma expressão analítica para a melhor escolha de um dos parâmetros das funções de Kautz quando o outro parâmetro é considerado fixo (constante). Essa solução baseia-se na generalização de um resultado apresentado em [34], e é obtido minimizando-se um limitante superior do erro de aproximação dos *kernels* de Volterra.

4.1 Motivação Inicial

Conforme foi discutido no Capítulo 2, os modelos de Volterra possuem as vantagens de permitir descrever um sistema não-linear como uma generalização do caso linear, além de estender a sistemas não-lineares conceitos definidos somente para os lineares. A principal desvantagem desses modelos, entretanto, é que eles necessitem de um número considerável de termos para a representação dos *kernels* de Volterra [1]. Para superar essa complexidade, uma saída é a utilização de modelos Wiener/Volterra, em que os *kernels* são desenvolvidos utilizando-se uma base de funções ortonormais. Esses modelos permitem que um número menor de termos sejam utilizados em suas representações, quando comparado com modelos de Volterra. E usando-se o fato de que as funções ortonormais são geralmente caracterizadas por parâmetros arbitrários (ver seções 3.3, 3.4 e 3.5), muitos trabalhos têm se dedicado a resolver o problema da seleção ótima dessas constantes.

Um dos primeiros trabalhos que se preocupou com esse problema foi [18], que obteve uma solução analítica para a melhor escolha do pólo de Laguerre para uma

classe particular de sistemas lineares. Posteriormente [15] resolveu-se esse problema para o caso linear em geral, ao minimizar uma função de custo quadrática. Trabalhos posteriores, [31] e [32], também obtiveram soluções analíticas para o pólo de Laguerre, abordando o problema de uma forma diferente das anteriores. Mais recentemente, [6], [7], [8] e [9] apresentaram o problema da melhor escolha do pólo de Laguerre para modelos de Volterra de qualquer ordem. Em todos esses trabalhos, uma solução analítica que dependesse unicamente do sistema foi obtida.

Em [28], [33] e [34], um estudo equivalente foi realizado com as funções ortonormais de Kautz. Entretanto o primeiro desses trabalhos se dedicou somente à obtenção de condições sobre sistemas lineares para a obtenção do pólo ótimo de Kautz. Os outros dois trabalhos enfocaram o estudo de uma solução literal sub-ótima¹ para o pólo de Kautz usado na representação de sistemas lineares. Foram obtidos resultados para os casos discreto e contínuo, respectivamente.

A idéia original do presente trabalho era a obtenção de uma solução conjunta ótima para ambos os parâmetros que caracterizam as funções de Kautz. Entretanto essa foi uma das dificuldades encontradas, e até o momento tal resultado não foi obtido. O problema maior enfrentado, conforme já observado em [34], é o fato de que as funções de Kautz, apresentadas na seção 3.5, não possuem uma equação a diferenças que as caracterize. Isso dificulta a abordagem analítica que foi adotada para o caso de Laguerre em [6], [7], [8], [9], [15] e [32].

Assim, obteve-se a solução analítica ótima de um dos parâmetros de Kautz, em função do outro. Esse problema foi resolvido em [34] para o caso linear, e aqui estende-se esses resultados para modelos de Volterra de qualquer ordem, conforme também apresentado em [11].

4.2 Modelos de Wiener/Volterra

O objetivo desta seção é apresentar um desenvolvimento matemático que permite a utilização de funções ortonormais na expansão dos *kernels* de Volterra. O modelo de Volterra da equação (2.10) será reconsiderado aqui para o caso em que o sinal de entrada possui uma determinada característica. Isso torna possível representar os *kernels* de Volterra por meio de uma expansão em base de funções ortonormais.

A justificativa para a representação de sistemas via séries de Volterra é baseada em uma abordagem bastante intuitiva, e é apresentada com mais detalhes em [24]. É natural observar a saída $y(k)$ de um sistema não-linear em um instante de tempo discreto k como dependente (de uma forma não-linear, no caso mais geral possível) de todos os valores da entrada $u(k)$ em instantes de tempo anteriores a k . Isto é, $y(k)$ depende de $u(k - T)$ para todo $T \geq 0$. É conveniente observar k como o instante de tempo atual, e supor que a saída atual depende de todos os valores do passado da entrada.

Esse ponto de vista leva à seguinte idéia: se para todo $T = 0, 1, 2, \dots$ a entrada $u(k)$ puder ser caracterizada por $u_1(k), u_2(k), \dots$, então a saída $y(k)$ pode ser

¹Essa solução sub-ótima se deve ao fato de que o problema da escolha ótima usando uma abordagem analítica está em estudo e ainda não foi obtida.

representada como uma função f dessas quantidades:

$$y(k) = f(u_1(k), u_2(k), \dots) \quad (4.1)$$

O primeiro passo para prosseguir nessa linha de pensamento é encontrar uma caracterização para o passado do sinal de entrada. Suponha agora que se fixe o instante k , considerando que o sinal $u(k-T)$, $T \in \{0, 1, 2, \dots\}$, seja quadraticamente somável:

$$\sum_{T=0}^{\infty} u^2(k-T) < \infty \quad (4.2)$$

Além disso, considere um conjunto de funções $\{\phi_1(T), \phi_2(T), \dots\}$ que constituem uma base ortonormal em $[0, \infty)$, ou seja:

$$\sum_{T=0}^{\infty} \phi_m(T)\phi_n(T) = \begin{cases} 0 & \text{para } m \neq n \\ 1 & \text{para } m = n \end{cases} \quad (4.3)$$

Então, o valor do sinal de entrada em qualquer instante de tempo anterior a k pode ser escrito da forma:

$$u(k-T) = \sum_{i=1}^{\infty} u_i(k)\phi_i(T) \quad (4.4)$$

onde, devido à ortogonalidade do conjunto $\{\phi_n\}$, tem-se que:

$$u_i(k) = \sum_{T=0}^{\infty} u(k-T)\phi_i(T) \quad (4.5)$$

Embora o instante de tempo k seja considerado fixo, o desenvolvimento acima leva a uma caracterização dos instantes passados de $u(k)$ em termos de $u_1(k), u_2(k), \dots$ sem levar k em consideração. Com essa caracterização em mente, pode-se expandir a função $f(u_1(k), u_2(k), \dots)$ em uma série de potências [24], tal que a saída em qualquer instante de tempo k é dada por:

$$y(k) = \sum_{i=1}^{\infty} a_i u_i(k) + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} a_{i_1, i_2} u_{i_1}(k) u_{i_2}(k) + \dots \quad (4.6)$$

Para se ter um maior entendimento sobre a representação Wiener/Volterra, basta substituir (4.5) em (4.6) para obter:

$$\begin{aligned} y(k) &= \sum_{i=1}^{\infty} \sum_{T=0}^{\infty} a_i u(k-T)\phi_i(T) + \\ &+ \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \sum_{T_1=0}^{\infty} \sum_{T_2=0}^{\infty} a_{i_1, i_2} u(k-T_1)\phi_{i_1}(T_1) u(k-T_2)\phi_{i_2}(T_2) + \dots \end{aligned} \quad (4.7)$$

Comparando a equação (4.7) com (2.10), pode-se então escrever o seguinte:

$$\begin{aligned}
 h_1(\tau_1) &= \sum_{i=1}^{\infty} a_i \phi_i(\tau_1) \\
 h_2(\tau_1, \tau_2) &= \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} a_{i_1, i_2} \phi_{i_1}(\tau_1) \phi_{i_2}(\tau_2) \\
 &\vdots \\
 h_n(\tau_1, \dots, \tau_n) &= \sum_{i_1=1}^{\infty} \cdots \sum_{i_n=1}^{\infty} a_{i_1, \dots, i_n} \prod_{j=1}^n \phi_{i_j}(\tau_j)
 \end{aligned} \tag{4.8}$$

Este desenvolvimento mostra como pode se utilizar uma base de funções ortonormais no desenvolvimento dos *kernels* de Volterra. A forma geral dessa expansão é dada pela equação (4.8), onde cada *kernel* h_n é escrito como uma combinação dos elementos da base ortonormal $\{\phi_l\}$.

A vantagem prática da representação de modelos de Volterra em base de funções ortonormais (modelos Wiener/Volterra) está situada em sua eficiência. Frequentemente é possível aproximar somas infinitas como a série de Volterra, usando-se expansão em funções ortonormais. Usualmente, um truncamento nas somas infinitas da equação (4.7) requer um número menor de termos em relação ao truncamento em modelos de Volterra, para se atingir a mesma precisão. E esse fato torna-se um atrativo para a modelagem de sistemas.

4.3 Definição do Problema

Conforme discutido na seção 4.1, muitos trabalhos têm dado atenção à obtenção de uma expressão analítica para o pólo ótimo de Laguerre que leve à melhor representação de sistemas por modelos Wiener/Volterra. Entretanto esse problema ainda está em aberto com relação às funções de Kautz, já que uma expressão analítica para os parâmetros b e c dessas funções não foi obtida ainda.

Os *kernels* de Volterra $h_\eta(\tau_1, \dots, \tau_\eta)$ podem ser aproximados usando-se uma expansão em base de funções ortonormais, desde que eles sejam estáveis e que sejam absolutamente somáveis [6]:

$$\sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_\eta=0}^{\infty} |h_\eta(\tau_1, \dots, \tau_\eta)| < \infty \tag{4.9}$$

Suponha que um dado sistema causal e invariante no tempo possa ser modelado via séries de Volterra, ou seja, que a sua resposta $y(k)$ a uma entrada $u(k)$ seja dada pela equação (2.10). Considere também que cada *kernel* possua as características citadas na seção 2.2.1: $h_\eta(\tau_1, \dots, \tau_l, \dots, \tau_\eta) = 0$ para $\tau_l < 0$ ($l = 1, 2, \dots, \eta$); e também que $\lim_{\tau_l \rightarrow \infty} h_\eta(\tau_1, \dots, \tau_l, \dots, \tau_\eta) = 0$. Essas duas condições são uma consequência da equação (4.9), que garante que o sistema seja do tipo *fading memory*.

Deseja-se então aproximar o *kernel* $h_\eta(k_1, \dots, k_\eta)$ por uma expansão em série ortonormal. Matematicamente, um *kernel* de Volterra desenvolvido em base de funções de Kautz $\{\psi_n\}$ é escrito da forma:

$$h_\eta(k_1, \dots, k_\eta) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_\eta} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \quad (4.10)$$

em que $\psi_l(k)$ são as funções de Kautz no domínio do tempo, i.e., são a transformada Z inversa de $\Psi_l(z)$. No domínio da frequência complexa z , a equação (4.10) é escrita como:

$$H_\eta(z_1, \dots, z_\eta) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_\eta} \prod_{j=1}^{\eta} \Psi_{i_j}(z_j) \quad (4.11)$$

onde os coeficientes $\alpha_{(\cdot)}$ são obtidos a partir de:

$$\alpha_{i_1, \dots, i_\eta} = \sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} h_\eta(k_1, \dots, k_\eta) \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \quad (4.12)$$

devido à ortogonalidade do conjunto $\{\psi_l\}$.

Para sistemas *fading memory*, que são estudados neste trabalho, a equação (4.10) é utilizada limitando-se o número de termos do somatório. Como a obtenção de seus modelos é realizada por computadores digitais, o que se faz na prática é limitar a quantidade de funções usadas a um número M que seja o suficiente para garantir uma boa aproximação. Portanto, tem-se:

$$\tilde{h}_\eta(k_1, \dots, k_\eta) = \sum_{i_1=1}^M \cdots \sum_{i_\eta=1}^M \alpha_{i_1, \dots, i_\eta} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \quad (4.13)$$

O modelo de Volterra (2.10), resultante da expansão dos *kernels* de Volterra usando-se um número M finito de funções ortonormais $\{\psi_n\}$, pode então ser expresso como:

$$\begin{aligned} y(k) &= \sum_{\tau_1=0}^k \sum_{i_1=1}^M \alpha_{i_1} \psi_{i_1}(k_1) u(k - \tau_1) + \\ &+ \sum_{\tau_1=0}^k \sum_{\tau_2=0}^k \sum_{i_1=1}^M \sum_{i_2=1}^M \alpha_{i_1, i_2} \psi_{i_1}(k_1) \psi_{i_2}(k_2) u(k - \tau_1) u(k - \tau_2) + \\ &+ \cdots + \sum_{\tau_\eta=0}^k \cdots \sum_{\tau_\eta=0}^k \sum_{i_1=1}^M \cdots \sum_{i_\eta=1}^M \alpha_{i_1, \dots, i_\eta} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) u(k - \tau_j) \end{aligned} \quad (4.14)$$

4.3.1 Erro de Aproximação dos *Kernels* de Volterra

A presente seção destina-se a obter uma expressão para o erro de aproximação de um *kernel* de Volterra usando uma base de funções ortonormais. A expressão para o erro que será apresentada ao fim da seção é válida para qualquer base ortonormal que seja utilizada na aproximação. Entretanto dar-se-á ênfase à base de funções de Kautz durante o desenvolvimento matemático, que é o tema de maior relevância deste trabalho.

De acordo com o problema definido anteriormente, deseja-se escolher b e c (com $|b| < 1$ e $|c| < 1$) no sentido de minimizar o erro da aproximação do *kernel* de Volterra em base de Kautz. Considerando a seguinte norma:

$$\|h_\eta\|^2 = \sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} h_\eta^2(k_1, \dots, k_\eta) \quad (4.15)$$

define-se a agora o erro quadrático proveniente da aproximação do *kernel* h_η como sendo:

$$\text{Erro} \triangleq \|h_\eta - \tilde{h}_\eta\|^2 \quad (4.16)$$

Usando as equações (4.10) e (4.13), pode-se escrever:

$$\begin{aligned} h_\eta(k_1, \dots, k_\eta) - \tilde{h}_\eta(k_1, \dots, k_\eta) &= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_\eta} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) + \\ &- \sum_{i_1=1}^M \cdots \sum_{i_\eta=1}^M \alpha_{i_1, \dots, i_\eta} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \\ &= \sum_{i_1=M+1}^{\infty} \cdots \sum_{i_\eta=M+1}^{\infty} \alpha_{i_1, \dots, i_\eta} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \quad (4.17) \end{aligned}$$

Voltando a (4.16) e usando a equação (4.17), pode-se escrever então:

$$\begin{aligned}
\text{Erro} &= \\
&= \sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} \left(\sum_{i_1=M+1}^{\infty} \cdots \sum_{i_\eta=M+1}^{\infty} \alpha_{i_1, \dots, i_\eta} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \right)^2 \\
&= \sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} \left[\left(\sum_{i_1=M+1}^{\infty} \cdots \sum_{i_\eta=M+1}^{\infty} \alpha_{i_1, \dots, i_\eta} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \right) \left(\sum_{i'_1=M+1}^{\infty} \cdots \sum_{i'_\eta=M+1}^{\infty} \alpha_{i'_1, \dots, i'_\eta} \prod_{j=1}^{\eta} \psi_{i'_j}(k_{j'}) \right) \right] \\
&= \sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} \left[\sum_{i_1=M+1}^{\infty} \cdots \sum_{i_\eta=M+1}^{\infty} \sum_{i'_1=M+1}^{\infty} \cdots \sum_{i'_\eta=M+1}^{\infty} \alpha_{i_1, \dots, i_\eta} \alpha_{i'_1, \dots, i'_\eta} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \prod_{j'=1}^{\eta} \psi_{i'_j}(k_{j'}) \right] \\
&= \sum_{i_1=M+1}^{\infty} \cdots \sum_{i_\eta=M+1}^{\infty} \sum_{i'_1=M+1}^{\infty} \cdots \sum_{i'_\eta=M+1}^{\infty} (\alpha_{i_1, \dots, i_\eta} \alpha_{i'_1, \dots, i'_\eta}) \left[\sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \prod_{j'=1}^{\eta} \psi_{i'_j}(k_{j'}) \right]
\end{aligned}$$

E usando-se o fato de que o conjunto $\{\psi_l\}$ é ortonormal, a equação anterior resulta em:

$$\text{Erro} = \sum_{i_1=M+1}^{\infty} \cdots \sum_{i_\eta=M+1}^{\infty} \alpha_{i_1, \dots, i_\eta}^2 \quad (4.18)$$

Também tem-se que:

$$\begin{aligned}
&\sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} h_\eta^2(k_1, \dots, k_\eta) = \\
&= \sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} \left(\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_\eta} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \right)^2 \\
&= \sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} \left[\left(\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_\eta} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \right) \left(\sum_{i'_1=1}^{\infty} \cdots \sum_{i'_\eta=1}^{\infty} \alpha_{i'_1, \dots, i'_\eta} \prod_{j=1}^{\eta} \psi_{i'_j}(k_{j'}) \right) \right] \\
&= \sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} \left[\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \sum_{i'_1=1}^{\infty} \cdots \sum_{i'_\eta=1}^{\infty} \alpha_{i_1, \dots, i_\eta} \alpha_{i'_1, \dots, i'_\eta} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \prod_{j'=1}^{\eta} \psi_{i'_j}(k_{j'}) \right] \\
&= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \sum_{i'_1=1}^{\infty} \cdots \sum_{i'_\eta=1}^{\infty} (\alpha_{i_1, \dots, i_\eta} \alpha_{i'_1, \dots, i'_\eta}) \left[\sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} \prod_{j=1}^{\eta} \psi_{i_j}(k_j) \prod_{j'=1}^{\eta} \psi_{i'_j}(k_{j'}) \right] \\
&= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_\eta}^2 \quad (4.19)
\end{aligned}$$

que é uma consequência da ortonormalidade do conjunto $\{\psi_l\}$. Usa-se agora as equações (4.18) e (4.19) para definir o Erro Quadrático Normalizado (EQN) como a seguir:

$$\begin{aligned}
 \text{EQN} &\triangleq \frac{\text{Erro}}{\|h_\eta\|^2} \\
 &= \frac{\sum_{i_1=M+1}^{\infty} \cdots \sum_{i_\eta=M+1}^{\infty} \alpha_{i_1, \dots, i_\eta}^2}{\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_\eta}^2} \\
 &= 1 - \frac{\sum_{i_1=1}^M \cdots \sum_{i_\eta=1}^M \alpha_{i_1, \dots, i_\eta}^2}{\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_\eta}^2} \tag{4.20}
 \end{aligned}$$

A equação (4.20) expressa que o erro da aproximação de um *kernel* de Volterra h_η depende exclusivamente dos coeficientes $\alpha_{(\cdot)}$ do seu desenvolvimento em base ortonormal. A expressão do erro de aproximação também é independente da base de funções ψ_l utilizada na expansão em (4.10).

Nota-se também que a equação (4.20) confirma o que foi discutido na seção 3.1, ou seja, o erro de aproximação será reduzido quando se aumenta o número M de funções usadas na aproximação. Deve ser observado que $\text{EQN} \rightarrow 0$ quando $M \rightarrow \infty$.

4.4 Desenvolvimento Matemático Preliminar

O desenvolvimento seguinte é baseado no fato de que as funções de Laguerre e de Kautz se relacionam no domínio z por meio de uma função que é dependente de apenas um dos parâmetros de Kautz. Isso permite fazer quase todo o desenvolvimento até se chegar à solução desejada para a base de Kautz, utilizando-se a base de funções de Laguerre. Dessa forma o problema de Kautz é transformado em um problema de Laguerre, que já foi estudado por [6], [7], [8], [9] e [15], e possui solução conhecida.

Primeiramente, definindo as funções $P(z)$ e $Q(z)$ como a seguir:

$$P(z) = \frac{z(z-b)}{1-bz} \tag{4.21}$$

$$Q(z) = \frac{b-z}{1-bz} \tag{4.22}$$

tem-se que:

$$\begin{aligned}
P(Q(z)) &= \frac{Q(z)(Q(z) - b)}{1 - bQ(z)} \\
&= \frac{\left(\frac{b-z}{1-bz}\right) \left(\frac{b-z}{1-bz} - b\right)}{1 - b\left(\frac{b-z}{1-bz}\right)} \\
&= \frac{b-z}{1-bz} \cdot \frac{-z + zb^2}{1-b^2} \\
&= \frac{-z(b-z)}{1-bz} = P(z)
\end{aligned} \tag{4.23}$$

Ou seja, as funções $P(z)$ e $Q(z)$, como definidas acima, satisfazem a relação $P(z) = P(Q(z))$. Partindo então da equação (3.21) e da definição de $Q(z)$, pode-se verificar que:

$$\Psi_{2n-1}(Q(z)) = \Psi_{2n-1}(z) \tag{4.24}$$

A partir da equação (3.15), que define as funções de Laguerre, pode-se inferir também o seguinte:

$$\begin{aligned}
\Phi_n(P(z)) &= \sqrt{1-c^2} \frac{P(z)}{P(z)-c} \left(\frac{1-cP(z)}{P(z)-c} \right)^{n-1} \\
&= \sqrt{1-c^2} \cdot \frac{\frac{z(z-b)}{1-bz}}{\frac{z(z-b)}{1-bz} - c} \cdot \left(\frac{1 - \frac{cz(z-b)}{1-bz}}{\frac{z(z-b)}{1-bz} - c} \right)^{n-1} \\
&= \sqrt{1-c^2} \frac{z(z-b)}{z^2 + b(c-1)z - c} \left(\frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c} \right)^{n-1} \\
&= \Psi_{2n-1}(z)
\end{aligned} \tag{4.25}$$

Portanto, as funções de transferência de Kautz e de Laguerre estão relacionadas por $\Phi_n(P(z)) = \Psi_{2n-1}(z)$.

Sejam $\alpha_{i_1, \dots, i_\eta}$ os coeficientes da expansão do *kernel* $h_\eta(k_1, \dots, k_\eta)$ em funções de Kautz, como em (4.12). Sendo $\Phi_n(z)$ as funções de Laguerre, conforme a equação (3.15), defina então as seguintes funções:

$$G_{\text{par}}(z_1, \dots, z_\eta) \triangleq \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1, \dots, 2i_\eta} \prod_{j=1}^{\eta} \Phi_{i_j}(z_j) \tag{4.26}$$

$$G_{\text{impar}}(z_1, \dots, z_\eta) \triangleq \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1-1, \dots, 2i_\eta-1} \prod_{j=1}^{\eta} \Phi_{i_j}(z_j) \tag{4.27}$$

que podem ser escritas no domínio do tempo como:

$$g_{\text{par}}(k_1, \dots, k_\eta) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1, \dots, 2i_\eta} \prod_{j=1}^{\eta} \phi_{i_j}(k_j) \quad (4.28)$$

$$g_{\text{impar}}(k_1, \dots, k_\eta) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1-1, \dots, 2i_\eta-1} \prod_{j=1}^{\eta} \phi_{i_j}(k_j) \quad (4.29)$$

Agora, defina também:

$$\begin{aligned} H_{\text{par}}(z_1, \dots, z_\eta) &\stackrel{\Delta}{=} G_{\text{par}}(P(z_1), \dots, P(z_\eta)) \\ &= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1, \dots, 2i_\eta} \prod_{j=1}^{\eta} \Phi_{i_j}(P(z_j)) \\ &= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1, \dots, 2i_\eta} \prod_{j=1}^{\eta} \Psi_{2i_j-1}(z_j) \\ &= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1, \dots, 2i_\eta} \prod_{j=1}^{\eta} \frac{z_j - b}{\sqrt{1 - b^2}} \Psi_{2i_j}(z_j) \\ &= \left(\prod_{j=1}^{\eta} \frac{z_j - b}{\sqrt{1 - b^2}} \right) \cdot \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1, \dots, 2i_\eta} \prod_{j=1}^{\eta} \Psi_{2i_j}(z_j) \quad (4.30) \end{aligned}$$

$$\begin{aligned} H_{\text{impar}}(z_1, \dots, z_\eta) &\stackrel{\Delta}{=} G_{\text{impar}}(P(z_1), \dots, P(z_\eta)) \\ &= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1-1, \dots, 2i_\eta-1} \prod_{j=1}^{\eta} \Phi_{i_j}(P(z_j)) \\ &= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1-1, \dots, 2i_\eta-1} \prod_{j=1}^{\eta} \Psi_{2i_j-1}(z_j) \quad (4.31) \end{aligned}$$

Usando a equação (4.11), segue que o *kernel* $H_\eta(z_1, \dots, z_\eta)$ é escrito como:

$$\begin{aligned} H_\eta(z_1, \dots, z_\eta) &= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_\eta} \prod_{j=1}^{\eta} \Psi_{i_j}(z_j) \\ &= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \left[\alpha_{2i_1-1, \dots, 2i_\eta-1} \prod_{j=1}^{\eta} \Psi_{2i_j-1}(z_j) + \alpha_{2i_1, \dots, 2i_\eta} \prod_{j=1}^{\eta} \Psi_{2i_j}(z_j) \right] \\ &= H_{\text{impar}}(z_1, \dots, z_\eta) + \left(\prod_{j=1}^{\eta} \frac{\sqrt{1 - b^2}}{z_j - b} \right) H_{\text{par}}(z_1, \dots, z_\eta) \quad (4.32) \end{aligned}$$

Sendo a função complexa $Q(z)$ como definida anteriormente, pode-se escrever de acordo com (4.32) que:

$$\begin{aligned}
H_\eta(Q(z_1), \dots, Q(z_\eta)) &= H_{\text{impar}}(Q(z_1), \dots, Q(z_\eta)) + \\
&\quad + \left(\prod_{j=1}^{\eta} \frac{\sqrt{1-b^2}}{Q(z_j) - b} \right) H_{\text{par}}(Q(z_1), \dots, Q(z_\eta)) \\
&= G_{\text{impar}}(P(Q(z_1)), \dots, P(Q(z_\eta))) + \\
&\quad + \left(\prod_{j=1}^{\eta} \frac{\sqrt{1-b^2}}{Q(z_j) - b} \right) G_{\text{par}}(P(Q(z_1)), \dots, P(Q(z_\eta))) \\
&= G_{\text{impar}}(P(z_1), \dots, P(z_\eta)) + \\
&\quad + \left(\prod_{j=1}^{\eta} \frac{\sqrt{1-b^2}}{Q(z_j) - b} \right) G_{\text{par}}(P(z_1), \dots, P(z_\eta)) \\
&= H_{\text{impar}}(z_1, \dots, z_\eta) + \left(\prod_{j=1}^{\eta} \frac{\sqrt{1-b^2}}{Q(z_j) - b} \right) H_{\text{par}}(z_1, \dots, z_\eta)
\end{aligned} \tag{4.33}$$

As equações (4.32) e (4.33) formam um sistema linear de equações para $H_{\text{par}}(z_1, \dots, z_\eta)$ e $H_{\text{impar}}(z_1, \dots, z_\eta)$. Não é difícil deduzir que:

$$\begin{aligned}
H_{\text{par}}(z_1, \dots, z_\eta) &= G_{\text{par}}(P(z_1), \dots, P(z_\eta)) \\
&= \frac{H_\eta(z_1, \dots, z_\eta) - H_\eta(Q(z_1), \dots, Q(z_\eta))}{\left(\prod_{j=1}^{\eta} \frac{\sqrt{1-b^2}}{z_j - b} \right) - \left(\prod_{j=1}^{\eta} \frac{\sqrt{1-b^2}}{Q(z_j) - b} \right)}
\end{aligned} \tag{4.34}$$

$$\begin{aligned}
H_{\text{impar}}(z_1, \dots, z_\eta) &= G_{\text{impar}}(P(z_1), \dots, P(z_\eta)) = \\
&= \frac{\left(\prod_{j=1}^{\eta} \frac{1}{z_j - b} \right) H_\eta(Q(z_1), \dots, Q(z_\eta)) - \left(\prod_{j=1}^{\eta} \frac{1}{Q(z_j) - b} \right) H_\eta(z_1, \dots, z_\eta)}{\left(\prod_{j=1}^{\eta} \frac{1}{z_j - b} \right) - \left(\prod_{j=1}^{\eta} \frac{1}{Q(z_j) - b} \right)}
\end{aligned} \tag{4.35}$$

Então para cada $j \in \{1, 2, \dots, \eta\}$, a mudança de variáveis $P(z_j) \triangleq z'_j$ pode ser utilizada para a obtenção de $G_{\text{par}}(z_1, \dots, z_\eta)$ e $G_{\text{impar}}(z_1, \dots, z_\eta)$ por meio

das equações (4.34) e (4.35), com² $z_j = P^{-1}(z'_j)$. O conhecimento dos *kernels* $H_\eta(z_1, \dots, z_\eta)$ do sistema permite que essas funções sejam obtidas.

Entretanto os cálculos matemáticos necessários a essa obtenção podem ser bastante complicados, por causa da mudança de variáveis necessária. Outra dificuldade da aplicação do resultado derivado acima é o fato de que ele requer o conhecimento dos *kernels* do sistema de uma forma fechada, ou seja, uma expressão matemática que descreva a dinâmica de cada kernel. Para contornar esses problemas, propõe-se um método para o cálculo de G_{par} e G_{impar} no domínio do tempo, conforme a proposição enunciada a seguir. A utilização do resultado dessa proposição pode ser útil quando o *kernel* não estiver disponível na forma de uma expressão matemática fechada.

4.4.1 Proposições

O cálculo das funções $G_{\text{par}}(z_1, \dots, z_\eta)$ e $G_{\text{impar}}(z_1, \dots, z_\eta)$ usando-se as equações (4.34) e (4.35) pode se tornar uma tarefa tediosa e complicada. Como será visto na seção 4.5, essas funções são necessárias no cálculo do parâmetro sub-ótimo c das funções de Kautz. Uma saída é o cálculo de G_{par} e G_{impar} no domínio do tempo, como apresentado na proposição 4.4.1.

Proposição 4.4.1. *Seja $h_\eta(\tau_1, \dots, \tau_\eta)$ o kernel de Volterra de ordem η . As funções $G_{\text{par}}(z_1, \dots, z_\eta)$ e $G_{\text{impar}}(z_1, \dots, z_\eta)$, definidas em (4.26) e (4.27) respectivamente, são expressas no domínio do tempo por:*

$$g_{\text{par}}(k_1, \dots, k_\eta) = \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_\eta=0}^{\infty} h_\eta(\tau_1, \dots, \tau_\eta) \prod_{j=1}^{\eta} \hat{\psi}_{2(k_j+1)}(\tau_j) \quad (4.36)$$

$$g_{\text{impar}}(k_1, \dots, k_\eta) = \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_\eta=0}^{\infty} h_\eta(\tau_1, \dots, \tau_\eta) \prod_{j=1}^{\eta} \hat{\psi}_{2(k_j+1)-1}(\tau_j) \quad (4.37)$$

em que $\hat{\psi}_{(\cdot)}(\tau)$ denotam as funções de Kautz com $c = 0$.

Prova:

Escrevendo a equação (4.26) no domínio do tempo e usando (4.12), tem-se:

$$\begin{aligned} g_{\text{par}}(k_1, \dots, k_\eta) &= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1, \dots, 2i_\eta} \prod_{j=1}^{\eta} \phi_{i_j}(k_j) \\ &= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \left(\sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_\eta=0}^{\infty} h_\eta(\tau_1, \dots, \tau_\eta) \prod_{j=1}^{\eta} \psi_{2i_j}(\tau_j) \right) \prod_{j=1}^{\eta} \phi_{i_j}(k_j) \\ &= \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_\eta=0}^{\infty} h_\eta(\tau_1, \dots, \tau_\eta) \left(\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \prod_{j=1}^{\eta} \psi_{2i_j}(\tau_j) \phi_{i_j}(k_j) \right) \end{aligned} \quad (4.38)$$

²Nesse contexto, a notação $P^{-1}(z)$ refere-se à função inversa da função $P(z)$.

Agora, considere o termo entre parênteses da equação anterior. Sua transformada Z em relação a τ_j pode ser escrita da seguinte forma:

$$\begin{aligned}
& \mathcal{Z} \left[\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \prod_{j=1}^{\eta} \psi_{2i_j}(\tau_j) \phi_{i_j}(k_j) \right] = \\
& = \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \left\{ \prod_{j=1}^{\eta} \left[\frac{\sqrt{(1-c^2)(1-b^2)} z_j}{z_j^2 + b(c-1)z_j - c} \left(\frac{-cz_j^2 + b(c-1)z_j + 1}{z_j^2 + b(c-1)z_j - c} \right)^{i_j-1} \phi_{i_j}(k_j) \right] \right\} \\
& = \left[\prod_{j=1}^{\eta} \frac{\sqrt{(1-c^2)(1-b^2)} z_j}{z_j^2 + b(c-1)z_j - c} \right] \cdot \left[\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \prod_{j=1}^{\eta} \phi_{i_j}(k_j) \left(\frac{-cz_j^2 + b(c-1)z_j + 1}{z_j^2 + b(c-1)z_j - c} \right)^{i_j-1} \right] \\
& \tag{4.39}
\end{aligned}$$

Então fazendo-se a mudança de variáveis

$$w_j \triangleq \frac{z_j^2 + b(c-1)z_j - c}{-cz_j^2 + b(c-1)z_j + 1} \tag{4.40}$$

e usando as funções de transferência $\Phi_l(z)$ em (3.15), pode-se inferir ainda sobre a equação (4.39) que:

$$\begin{aligned}
\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \prod_{j=1}^{\eta} \phi_{i_j}(k_j) w_j^{1-i_j} & = \prod_{j=1}^{\eta} \left(w_j \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \phi_{i_j}(k_j) w_j^{-i_j} \right) \\
& = \prod_{j=1}^{\eta} \left[\frac{w_j \sqrt{1-c^2}}{w_j + c} \left(\frac{1 + cw_j}{w_j + c} \right)^{k_j} \right] \\
& \tag{4.41}
\end{aligned}$$

onde foi utilizada a seguinte relação válida para as funções de Laguerre $\phi_l(k)$:

$$\begin{aligned}
\sum_{l=1}^{\infty} \phi_l(k) w^{-l} &= \sum_{l=1}^{\infty} \mathcal{Z}^{-1} [\Phi_l(z)] w^{-l} \\
&= \sqrt{1-c^2} \mathcal{Z}^{-1} \left\{ \sum_{l=1}^{\infty} \frac{z}{z-c} \left(\frac{1-cz}{z-c} \right)^{l-1} w^{-l} \right\} \\
&= \sqrt{1-c^2} \mathcal{Z}^{-1} \left\{ \frac{z}{z-c} \left[\frac{w^{-1}}{1-w^{-1} \left(\frac{1-cz}{z-c} \right)} \right] \right\} \\
&= \sqrt{1-c^2} \mathcal{Z}^{-1} \left\{ \frac{z}{z(w+c) - (1+wc)} \right\} \\
&= \frac{\sqrt{1-c^2}}{w+c} \mathcal{Z}^{-1} \left\{ \frac{z}{z - \left(\frac{1+wc}{w+c} \right)} \right\} \\
&= \frac{\sqrt{1-c^2}}{w+c} \left(\frac{1+cw}{w+c} \right)^k \tag{4.42}
\end{aligned}$$

Utilizando o resultado (4.41) e retornando à variável z , a equação (4.39) fica, após algumas simplificações algébricas:

$$\begin{aligned}
\mathcal{Z} \left[\sum_{i_1=1}^{\infty} \cdots \sum_{i_{\eta}=1}^{\infty} \prod_{j=1}^{\eta} \psi_{2i_j}(\tau_j) \phi_{i_j}(k_j) \right] &= \prod_{j=1}^{\eta} \left[\frac{\sqrt{1-b^2}}{z_j-b} \left(\frac{1-bz_j}{z_j-b} \right)^{k_j} z_j^{-k_j} \right] \\
&= \prod_{j=1}^{\eta} \mathcal{Z} \left\{ [\psi_{2(k_j+1)}(\tau_j)]_{c=0} \right\} \\
&= \mathcal{Z} \left\{ \prod_{j=1}^{\eta} [\psi_{2(k_j+1)}(\tau_j)]_{c=0} \right\} \tag{4.43}
\end{aligned}$$

Logo, tem-se que:

$$\sum_{i_1=1}^{\infty} \cdots \sum_{i_{\eta}=1}^{\infty} \prod_{j=1}^{\eta} \psi_{2i_j}(\tau_j) \phi_{i_j}(k_j) = \prod_{j=1}^{\eta} \hat{\psi}_{2(k_j+1)}(\tau_j) \tag{4.44}$$

onde $\hat{\psi}_n(\tau_j) \triangleq [\psi_n(\tau_j)]_{c=0}$. Substituindo então a equação (4.44) em (4.38), resulta

que:

$$\begin{aligned} g_{\text{par}}(k_1, \dots, k_\eta) &= \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_\eta=0}^{\infty} h_\eta(\tau_1, \dots, \tau_\eta) \left(\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \prod_{j=1}^{\eta} \psi_{2i_j}(\tau_j) \phi_{i_j}(k_j) \right) \\ &= \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_\eta=0}^{\infty} h_\eta(\tau_1, \dots, \tau_\eta) \prod_{j=1}^{\eta} \hat{\psi}_{2(k_j+1)}(\tau_j) \end{aligned}$$

■

A demonstração para a função $g_{\text{impar}}(k_1, \dots, k_\eta)$ é análoga.

Proposição 4.4.2. *Considere as funções $\hat{\psi}_{2n}(k)$, definidas na proposição 4.4.1, que são obtidas das funções de Kautz $\psi_{2n}(k)$ fazendo-se $c = 0$. Aquelas funções estão relacionadas com as funções de Laguerre $\phi_n(k)$, por meio da relação $\hat{\psi}_{2n}(k) = \phi_n(k - n)$.*

Prova:

Fazendo $c = 0$ em (3.21), segue que:

$$\begin{aligned} \hat{\Psi}_{2n}(z) &= \frac{\sqrt{1-b^2} z}{z^2 - bz} \left(\frac{-bz + 1}{z^2 - bz} \right)^{n-1} \\ &= z^{-n} \sqrt{1-b^2} \frac{z}{z-b} \left(\frac{1-bz}{z-b} \right)^{n-1} \\ &= z^{-n} \Phi_n(z) \end{aligned} \tag{4.45}$$

onde redefiniu-se o parâmetro c em (3.15) como b . Para condições iniciais nulas, vale a seguinte propriedade da transformada Z unilateral:

$$z^{-1}X(z) = \mathcal{Z}[x(k-1)] \tag{4.46}$$

Fazendo-se uso n vezes da propriedade acima, tem-se que a equação (4.45) pode ser escrita no domínio do tempo como:

$$\hat{\psi}_{2n}(k) = \phi_n(k - n)$$

■

4.5 Expansão Sub-Ótima dos *Kernels* de Volterra em Base de Kautz

O problema da escolha ótima do pólo das funções discretas de Laguerre $\phi_n(k)$ utilizando uma abordagem analítica foi apresentado em [6], [7], [8], [9], [15] e [29]. As

funções de Laguerre possuem uma propriedade que torna esse tipo de solução mais simples que no caso de Kautz.

Com isso, o método proposto nessa seção é baseado, em sua maior parte, nas funções de Laguerre, que são parametrizadas por somente uma variável livre. Conforme discutido no início da seção 4.4, no presente trabalho resolve-se o problema de Kautz baseado no problema de Laguerre. Uma expressão analítica para a melhor escolha do parâmetro c (para um dado b) será apresentada posteriormente. Propõe-se estender os resultados de [34] (obtidos para o caso linear) para o *kernel* de ordem η , conforme discutido a seguir.

Partindo da função de transferência que define as funções de Laguerre, dada por (3.15), não é difícil deduzir que, para $j = 1, 2, \dots, \eta$, a seguinte relação é válida (ver demonstração no Apêndice B):

$$-c(k+1)\phi_{i_j}(k+1) + [(1+c^2)k+1]\phi_{i_j}(k) - ck\phi_{i_j}(k-1) = i_j(1-c^2)\phi_{i_j}(k) \quad (4.47)$$

Nesse contexto, a constante real c representa um dos parâmetros das funções de Kautz. E como $|c| < 1$, pode-se tomar uma base de Laguerre nesse parâmetro.

Considere uma função não-nula $x(k_1, \dots, k_\eta)$, que é nula para todo $k_j < 0$ ($j = 1, 2, \dots, \eta$). Suponha também que x seja absolutamente somável em $[0, \infty)$, isto é:

$$\sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} |x(k_1, \dots, k_\eta)| < \infty$$

Seja o problema de se expandir a função x em base de funções de Laguerre. De acordo com a equação (4.12), os coeficientes $\vartheta_{i_1, \dots, i_\eta}$ dessa expansão são dados por:

$$\vartheta_{i_1, \dots, i_\eta} = \sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} x(k_1, \dots, k_\eta) \prod_{j=1}^{\eta} \phi_{i_j}(k_j) \quad (4.48)$$

A partir das equações (4.47) e (4.48), pode-se demonstrar³ que:

$$\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (i_1 + \cdots + i_\eta) \vartheta_{i_1, \dots, i_\eta}^2 = \frac{-2c\mu_1(x) + (1+c^2)\mu_2(x) + \eta\mu_3(x)}{1-c^2} \quad (4.49)$$

onde os momentos $\mu_1(x)$, $\mu_2(x)$ e $\mu_3(x)$ dependem unicamente de $x(k_1, \dots, k_\eta)$, e são dados por:

$$\mu_1(x) = \sum_{l=1}^{\eta} \left(\sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x(k_1, \dots, k_l, \dots, k_\eta) x(k_1, \dots, k_l - 1, \dots, k_\eta) \right) \quad (4.50)$$

³Uma demonstração detalhada dessa igualdade é apresentada no Apêndice B.

$$\mu_2(x) = \sum_{l=1}^{\eta} \left(\sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_{\eta}=0}^{\infty} k_l x^2(k_1, \dots, k_l, \dots, k_{\eta}) \right) \quad (4.51)$$

$$\mu_3(x) = \sum_{k_1=0}^{\infty} \cdots \sum_{k_{\eta}=0}^{\infty} x^2(k_1, \dots, k_{\eta}) \quad (4.52)$$

Considere agora a expansão das funções g_{par} e g_{impar} em base de Laguerre. Denominando os coeficientes dessas expansões de $\gamma_{i_1, \dots, i_{\eta}}$ e $\rho_{i_1, \dots, i_{\eta}}$, respectivamente, tem-se que:

$$g_{\text{par}}(k_1, \dots, k_{\eta}) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_{\eta}=1}^{\infty} \gamma_{i_1, \dots, i_{\eta}} \prod_{j=1}^{\eta} \phi_{i_j}(k_j) \quad (4.53)$$

$$g_{\text{impar}}(k_1, \dots, k_{\eta}) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_{\eta}=1}^{\infty} \rho_{i_1, \dots, i_{\eta}} \prod_{j=1}^{\eta} \phi_{i_j}(k_j) \quad (4.54)$$

Comparando-se as equações (4.28) e (4.29) com (4.53) e (4.54), verifica-se que os coeficientes $\alpha_{(\cdot)}$ da expansão do *kernel* $h_{\eta}(k_1, \dots, k_{\eta})$ em funções de Kautz estão relacionados com os coeficientes da expansão de g_{par} e g_{impar} em base de Laguerre, da seguinte forma:

$$\begin{aligned} \gamma_{i_1, \dots, i_{\eta}} &= \alpha_{2i_1, \dots, 2i_{\eta}} & \forall i_1, \dots, i_{\eta} \\ \rho_{i_1, \dots, i_{\eta}} &= \alpha_{2i_1-1, \dots, 2i_{\eta}-1} & \forall i_1, \dots, i_{\eta} \end{aligned}$$

Seja a seguinte desigualdade [9]:

$$\sum_{i_1=1}^{\infty} \cdots \sum_{i_{\eta}=1}^{\infty} (i_1 + \cdots + i_{\eta}) \alpha_{i_1, \dots, i_{\eta}}^2 \geq \eta(M+1) \sum_{i_1=M+1}^{\infty} \cdots \sum_{i_{\eta}=M+1}^{\infty} \alpha_{i_1, \dots, i_{\eta}}^2 \quad (4.55)$$

onde M representa o número de termos da base de Kautz que serão utilizados na aproximação do *kernel* h_{η} . Note que M tem justamente esse significado na equação

(4.13). A partir de (4.55), pode-se inferir o seguinte:

$$\begin{aligned}
\sum_{i_1=M+1}^{\infty} \cdots \sum_{i_\eta=M+1}^{\infty} \alpha_{i_1, \dots, i_\eta}^2 &\leq \frac{1}{\eta(M+1)} \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (i_1 + \cdots + i_\eta) \alpha_{i_1, \dots, i_\eta}^2 \\
&= \frac{1}{\eta(M+1)} \sum_{i_1 \text{ par}}^{\infty} \cdots \sum_{i_\eta \text{ par}}^{\infty} (i_1 + \cdots + i_\eta) \alpha_{i_1, \dots, i_\eta}^2 + \\
&+ \frac{1}{\eta(M+1)} \sum_{i_1 \text{ impar}}^{\infty} \cdots \sum_{i_\eta \text{ impar}}^{\infty} (i_1 + \cdots + i_\eta) \alpha_{i_1, \dots, i_\eta}^2 \\
&= \frac{1}{\eta(M+1)} \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (2i_1 + \cdots + 2i_\eta) \alpha_{2i_1, \dots, 2i_\eta}^2 + \\
&+ \frac{1}{\eta(M+1)} \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} [(2i_1 - 1) + \cdots + (2i_\eta - 1)] \alpha_{2i_1-1, \dots, 2i_\eta-1}^2 \\
&\leq \frac{2}{\eta(M+1)} \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (i_1 + \cdots + i_\eta) \gamma_{i_1, \dots, i_\eta}^2 + \\
&+ \frac{2}{\eta(M+1)} \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (i_1 + \cdots + i_\eta) \rho_{i_1, \dots, i_\eta}^2
\end{aligned} \tag{4.56}$$

Pode-se reconhecer na equação acima a expressão para o erro de aproximação do *kernel* de Volterra $h_\eta(k_1, \dots, k_\eta)$, conforme a equação (4.18). Usando então a igualdade (4.49) para reescrever a equação (4.56), tem-se que:

$$\begin{aligned}
\text{Erro} &\leq 2 \cdot \left[\frac{-2c\mu_1(g_{\text{par}}) + (1+c^2)\mu_2(g_{\text{par}}) + \eta\mu_3(g_{\text{par}})}{\eta(M+1)(1-c^2)} \right] + \\
&+ 2 \cdot \left[\frac{-2c\mu_1(g_{\text{impar}}) + (1+c^2)\mu_2(g_{\text{impar}}) + \eta\mu_3(g_{\text{impar}})}{\eta(M+1)(1-c^2)} \right] \\
&= \frac{2(m_2c^2 - 2m_1c + m_3)}{\eta(M+1)(1-c^2)}
\end{aligned} \tag{4.57}$$

onde as variáveis m_p ($p = 1, 2, 3$) são dadas pelas equações a seguir:

$$m_1 = \mu_1(g_{\text{par}}) + \mu_1(g_{\text{impar}}) \tag{4.58}$$

$$m_2 = \mu_2(g_{\text{par}}) + \mu_2(g_{\text{impar}}) \tag{4.59}$$

$$m_3 = \mu_2(g_{\text{par}}) + \mu_2(g_{\text{impar}}) + \eta\mu_3(g_{\text{par}}) + \eta\mu_3(g_{\text{impar}}) \tag{4.60}$$

4.5 Expansão Sub-Ótima dos *Kernels* de Volterra em Base de Kautz 45

Nesse contexto, $g_{\text{par}}(k_1, \dots, k_\eta)$ e $g_{\text{impar}}(k_1, \dots, k_\eta)$ são dadas por (4.36) e (4.37), respectivamente. Pode-se então resolver o seguinte problema de otimização:

$$\min_{|c| < 1} L(c) = \frac{2(m_2 c^2 - 2m_1 c + m_3)}{\eta(M+1)(1-c^2)} \quad (4.61)$$

para derivar uma expressão analítica para o parâmetro c de Kautz. A solução para o problema (4.61) encontra-se no Apêndice B, sendo dada por:

$$c_{\text{otimo}} = \begin{cases} \xi - \sqrt{\xi^2 - 1} & \text{se } \xi > 1 \\ \xi + \sqrt{\xi^2 - 1} & \text{se } \xi < -1 \end{cases} \quad (4.62)$$

onde $\xi = (m_2 + m_3)/(2m_1)$. Para o caso $\xi = 1$, ver detalhes explicativos na seção B.3 do Apêndice B.

A equação (4.62) é uma expressão analítica para a melhor escolha do coeficiente c das funções de Kautz segundo o critério (4.61). De acordo com o método proposto acima, essa equação é utilizada quando se admite conhecido o parâmetro b , geralmente escolhido *a priori* para o cálculo de g_{par} e g_{impar} usando-se as equações (4.36) e (4.37). Esse desenvolvimento pode ser visto como uma solução sub-ótima para o problema de se determinar os parâmetros b e c que minimizam um limitante superior do erro expresso em (4.16). Essa solução é sub-ótima devido ao fato de que ambos os parâmetros b e c não são otimizados simultaneamente.

4.5.1 Algoritmo

O procedimento apresentado anteriormente pode ser entendido mais facilmente quando sintetizado em um algoritmo. Com essa finalidade, apresenta-se a seguir a seqüência de passos necessária à solução do problema proposto neste trabalho. Para cada ordem do *kernel*, executar os passos:

1. Escolher um valor inicial para parâmetro b de Kautz ($|b| < 1$). Dependendo desta escolha, melhores resultados para a aproximação do *kernel* podem ser obtidos.
2. Tendo em mãos os *kernels* de Volterra de um determinado sistema, calcular as funções g_{par} e g_{impar} usando as equações (4.36) e (4.37), respectivamente.
3. Calcular os momentos $\mu_1(x)$, $\mu_2(x)$ e $\mu_3(x)$ usando-se as equações (4.50), (4.51) e (4.52), respectivamente, para $x = g_{\text{par}}$ e $x = g_{\text{impar}}$.
4. Calcular as variáveis m_1 , m_2 e m_3 usando as equações (4.58), (4.59) e (4.60) respectivamente.
5. Utilizar a equação (4.62) para determinar o parâmetro ótimo c relativo ao valor de b escolhido no passo 1.

Se o modelo obtido pelo algoritmo acima não oferecer bons resultados, deve-se escolher um outro valor para o parâmetro b no passo 1.

4.6 Resumo e Contribuições

Este capítulo apresentou resultados matemáticos referentes ao problema de aproximação de modelos de Volterra usando funções ortonormais de Kautz. Foram obtidas equações necessárias à solução desse problema, e uma expressão simplificada para o erro de aproximação ortonormal do *kernel* de Volterra foi obtida.

Discutiu-se também o problema da melhor seleção de um dos parâmetros de Kautz, do ponto de vista do erro de aproximação de modelos Wiener/Volterra. Otimizar simultaneamente os dois parâmetros que caracterizam o pólo de Kautz, em termos de uma abordagem analítica, é um problema que ainda está em aberto. Dessa forma, optou-se por abordar o mesmo problema minimizando-se o limitante superior do erro de aproximação. A contribuição principal é a obtenção de uma expressão analítica para a escolha sub-ótima do pólo de Kautz, para modelos Wiener/Volterra de qualquer ordem. Com esse resultado, é possível determinar (por meio de uma equação literal) um dos parâmetros de Kautz, e essa solução pode ser interpretada como uma equação analítica que fornece o pólo sub-ótimo de Kautz.

Outro assunto abordado foi a generalização de um teorema até então válido somente para modelos Wiener/Volterra de primeira ordem (que representam os sistemas lineares). Neste capítulo, esse teorema foi generalizado para modelos de qualquer ordem, o que permite que seus resultados possam ser aplicados no estudo de sistemas não-lineares.

Capítulo 5

Simulações e Resultados

No presente capítulo, serão apresentadas algumas simulações para ilustrar os resultados matemáticos deduzidos no capítulo anterior. Também será discutida a simetria dos *kernels* de Volterra, assunto abordado no Capítulo 2. Esse procedimento pode ser útil na caracterização das propriedades do sistema com base nas propriedades dos *kernels* que o representam. Serão comparados os resultados obtidos nas simulações implementadas, usando-se as versões assimétrica e simétrica do *kernel*.

Faz-se também uma comparação entre a aproximação de sistemas não-lineares usando séries ortonormais de Kautz e de Laguerre. É apresentado um estudo comparativo entre a utilização dessas duas bases, de acordo com suas propriedades matemáticas. Apresentam-se resultados numéricos que mostram a importância da escolha da base de funções utilizadas na representação de modelos Wiener/Volterra.

Discute-se ainda a influência do comportamento dinâmico do sistema sobre o cálculo do pólo de Kautz que melhor o representa. Sistemas com oscilações (sub-amortecidos) são melhor aproximados quando a base de Kautz é utilizada [34], enquanto que sistemas super-amortecidos podem ser bem implementados usando-se a base de funções de Laguerre. Propõe-se ainda um algoritmo iterativo para estimação simultânea dos *kernels* de Volterra e dos pólos sub-ótimos de Kautz. Esse algoritmo é utilizado na modelagem de um sistema de levitação magnética.

5.1 Exemplo Ilustrativo

Os resultados matemáticos derivados no Capítulo 4 são importantes na aproximação ortonormal de sinais. Foi apresentada uma expressão analítica para a escolha sub-ótima do pólo de Kautz. A presente seção dedica-se a ilustrar esses resultados, por meio de simulações.

Suponha que para um determinado sistema, o *kernel* de Volterra de segunda ordem seja dado por:

$$h_2(k_1, k_2) = (k_1 - 2k_2) \exp(-\lambda_1 k_1 - \lambda_2 k_2) \cos(\omega_1 k_1 + \omega_2 k_2) \quad (5.1)$$

para $k_1 \geq 0$ e $k_2 \geq 0$. Considera-se também que $h_2(k_1, k_2) = 0$ para valores negativos de k_1 ou k_2 (sistema causal).

O *kernel* $h_\eta(k_1, \dots, k_\eta)$ é separável se for possível expressá-lo como o produto de η funções dependentes de um único atributo, ou matematicamente se:

$$h_\eta(k_1, \dots, k_\eta) = \prod_{j=1}^{\eta} f_j(k_j) \quad (5.2)$$

onde as funções f_j são dependentes individualmente apenas do atributo k_j . Nota-se que o *kernel* (5.1) é não-separável.

Para o *kernel* da equação (5.1), λ_i e ω_i ($i = 1, 2$) são constantes reais¹ que determinam a dinâmica do sistema. As constantes λ_i podem ser vistas como o fator de decaimento do *kernel*, i.e. dão a idéia da velocidade com que o *kernel* vai a zero à medida que $k_i \rightarrow \infty$. E as constantes ω_i podem ser interpretadas como as frequências com que o *kernel* oscila em cada direção axial.

Conforme discutido na seção 3.1, o uso de funções de Kautz permite obter melhores aproximações de sistemas com comportamento oscilatório, em relação às funções de Laguerre. O *kernel* dado pela equação (5.1) é oscilatório amortecido, no sentido de ser caracterizado por uma função cosseno modulada por uma exponencial decrescente. A figura 5.1 ilustra o gráfico de $h_2(k_1, k_2)$ para $\lambda_1 = 0,45$, $\lambda_2 = 0,7$, $\omega_1 = 100$ e $\omega_2 = 1$.

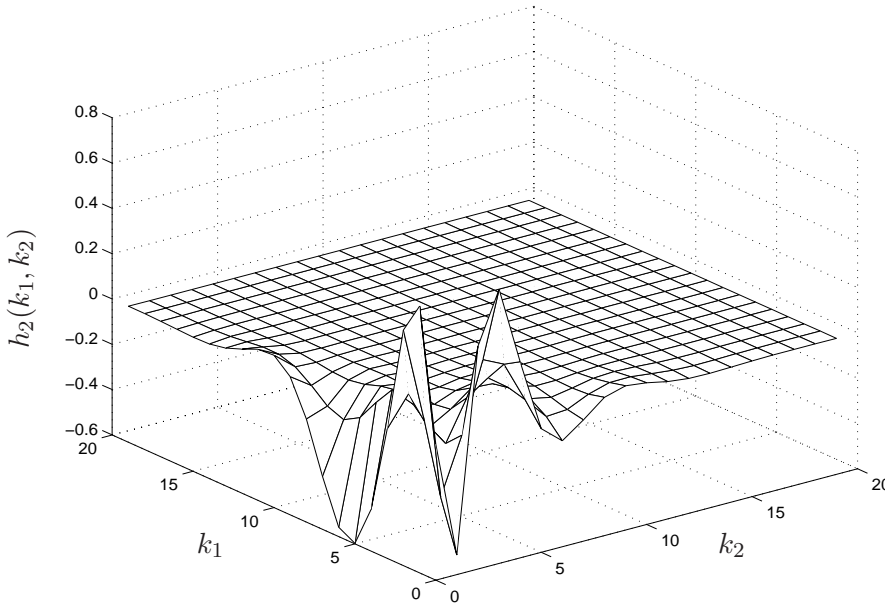


Figura 5.1: *Kernel* da equação (5.1), para $\lambda_1 = 0,45$, $\lambda_2 = 0,7$, $\omega_1 = 100$ e $\omega_2 = 1$

As simulações cujos resultados são apresentados a seguir foram baseadas no método proposto na seção 4.5: para um dado valor de b , determina-se o valor ótimo de c . O termo “ótimo” é utilizado no sentido de minimizar o limitante superior do erro quadrático entre o *kernel* e a sua aproximação usando uma expansão em base

¹Note que deve-se ter $\lambda_i > 0$ para que o *kernel* da equação (5.1) seja estável.

de Kautz, conforme o desenvolvimento do Capítulo 4. Uma primeira simulação foi realizada para $b = 0, 4$; utilizando as equações (4.58), (4.59) e (4.60), obtém-se:

- $m_1 = -1, 33352$
- $m_2 = 1, 99252$
- $m_3 = 4, 68624$

O parâmetro ótimo c calculado via (4.62) é $c_{\text{ótimo}} = -0, 20833$. A equação (4.13) é então utilizada para se determinar a aproximação do *kernel*. Para $M = 2, 4, 6$, o Erro Quadrático Normalizado (EQN) associado a essa aproximação pode ser calculado pela equação (4.20), sendo mostrado na tabela 5.1.

Tabela 5.1: Erro de aproximação do *kernel* da equação (5.1) de acordo com o número de funções de Kautz utilizadas na expansão

Número de funções (M)	EQN = $\frac{\ h_2 - \tilde{h}_2\ ^2}{\ h_2\ ^2}$
2	0,73525
4	0,28299
6	0,05877

Posteriormente, variou-se o parâmetro b em todo o intervalo $(-1, 1)$. Para cada valor de b nesse intervalo, calculou-se o valor de $c_{\text{ótimo}}$ usando a equação (4.62). A figura 5.2 mostra o gráfico de $c_{\text{ótimo}}$ calculado para cada valor de $b \in (-1, 1)$.

Variando-se o parâmetro b em todo o intervalo de valores que ele pode assumir, cada par $(b; c_{\text{ótimo}})$ resulta em um erro de aproximação. O par que melhor aproxima o sistema é aquele em que o erro assume o menor valor. Com isso, obtém-se o par $(b; c)$ que melhor aproxima o *kernel* da equação (5.1). O erro proveniente dessa aproximação, calculado para cada par $(b; c_{\text{ótimo}})$, é mostrado na figura 5.3.

A figura 5.3 permite concluir que o valor de b para o qual o erro quadrático é mínimo é igual a $b = 0, 593$. Para esse valor de b , o valor de $c_{\text{ótimo}}$ calculado foi de $c_{\text{ótimo}} = -0, 25940$. A aproximação de (5.1) com $M = 6$ funções de Kautz $\psi_l(k)$ fica:

$$\tilde{h}_2(k_1, k_2) = \sum_{i_1=1}^6 \sum_{i_2=1}^6 \alpha_{i_1, i_2} \psi_{i_1}(k_1) \psi_{i_2}(k_2) \quad (5.3)$$

cujo gráfico está ilustrado na figura 5.4.

O erro proveniente dessa aproximação é a diferença entre os gráficos das figuras 5.1 e 5.4, ilustrado na figura 5.5. Esse gráfico representa a superfície de erro entre o *kernel* real, dado por (5.1), e sua aproximação dada pela equação (5.3).

A figura 5.5 foi obtida para o par de parâmetros $(b; c) = (0, 593; -0, 25940)$, que forneceu nas simulações o menor erro quadrático normalizado dessa aproximação,

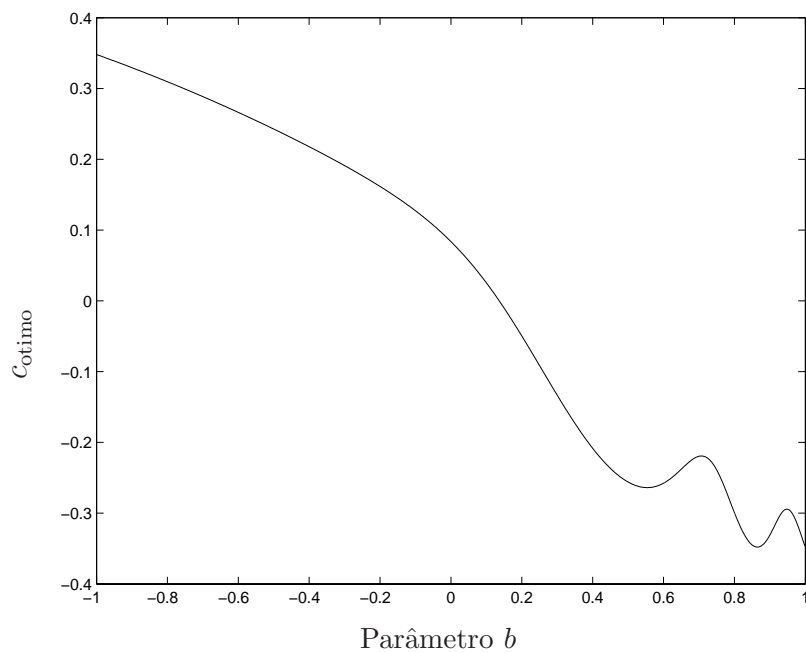


Figura 5.2: Valor de c_{otimo} , calculado pela equação (4.62) para cada valor de $b \in (-1, 1)$

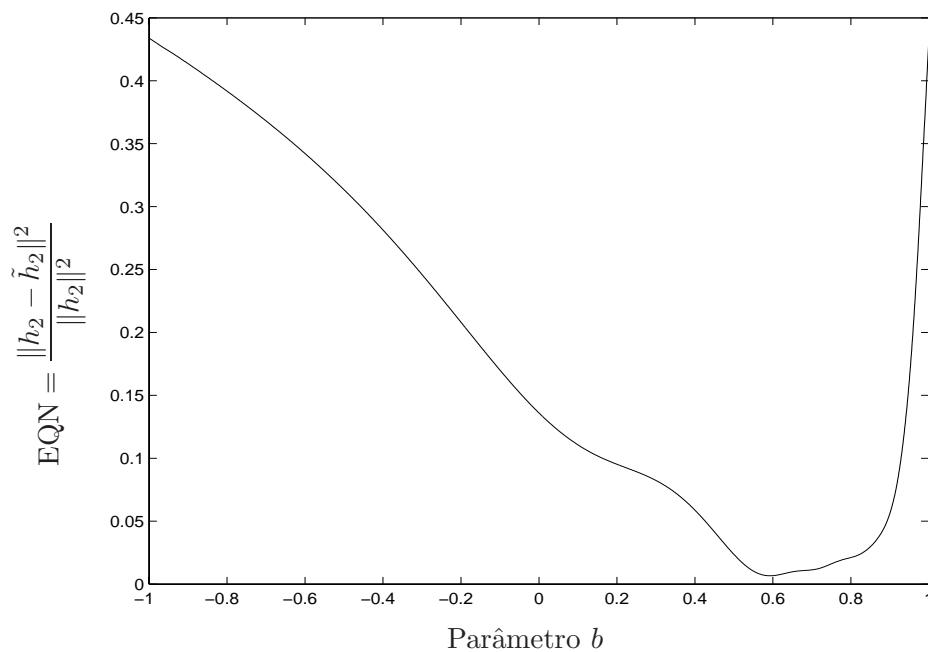


Figura 5.3: Erro Quadrático Normalizado (EQN) entre o *kernel* real e sua aproximação com $M = 6$ funções de Kautz, calculado para cada valor de $b \in (-1, 1)$

$\text{EQN} = 6,621 \cdot 10^{-3}$. Então, o pólo de Kautz pode ser calculado usando-se as relações

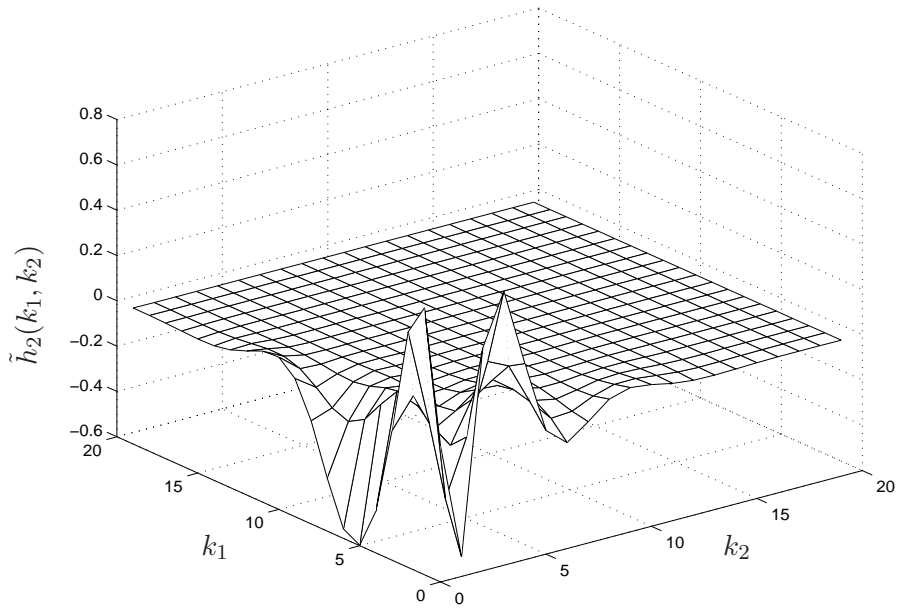


Figura 5.4: *Kernel* aproximado $\tilde{h}_2(k_1, k_2)$ para $b = 0,593$ e $c = -0,25940$

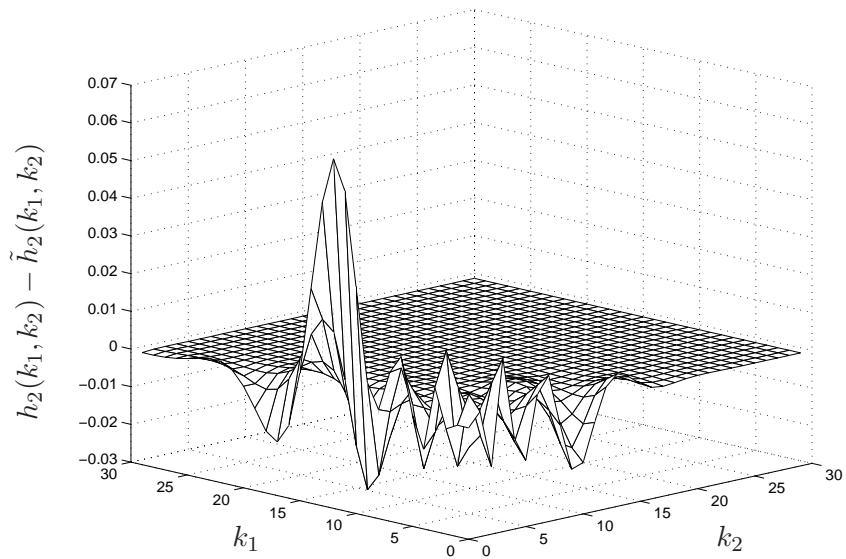


Figura 5.5: Superfície de erro $h_2(k_1, k_2) - \tilde{h}_2(k_1, k_2)$ para $b = 0,593$ e $c = -0,25940$

entre os parâmetros b e c e o pólo β :

$$b = \frac{\beta + \bar{\beta}}{1 + \beta\bar{\beta}} \quad c = -\beta\bar{\beta}$$

obtendo-se o par de pólos de Kautz $\beta = 0,37341 \pm i0,34636$, que melhor representam o *kernel* de segunda ordem dado por (5.1).

O pico ilustrado no erro de aproximação do *kernel*, na figura 5.5, pode ser proveniente do fato de ele ser assimétrico, ou seja, $h_2(k_1, k_2) \neq h_2(k_2, k_1)$. A amplitude do pico pode também ser explicada pela norma utilizada para o cálculo do erro de aproximação do *kernel*. Na seção seguinte, o mesmo sistema é simulado considerando-se sua versão simétrica.

5.1.1 Estudo da Simetrização do *Kernel* de Volterra

Para avaliar a simetrização de *kernels* de Volterra, conceito apresentado na seção 2.4, suponha agora que o *kernel* da equação (5.1) seja simetrizado. Usando a equação (2.19), a versão simétrica desse *kernel* é dada por:

$$h_2^{\text{sim}}(k_1, k_2) = \frac{h_2(k_1, k_2) + h_2(k_2, k_1)}{2} \quad (5.4)$$

onde $h_2(k_1, k_2)$ é tal como na equação (5.1).

Fazendo-se uma análise semelhante àquela ilustrada nas figuras 5.2 e 5.3, o par de parâmetros de Kautz que melhor aproxima o *kernel* simetrizado da equação (5.4) é $(b; c) = (0,684; -0,46949)$. Para esses valores, obtém-se $\text{EQN} = 7,4553 \cdot 10^{-3}$, e a figura 5.6 ilustra a superfície de erro de aproximação desse *kernel*, quando utilizadas $M = 6$ funções de Kautz.

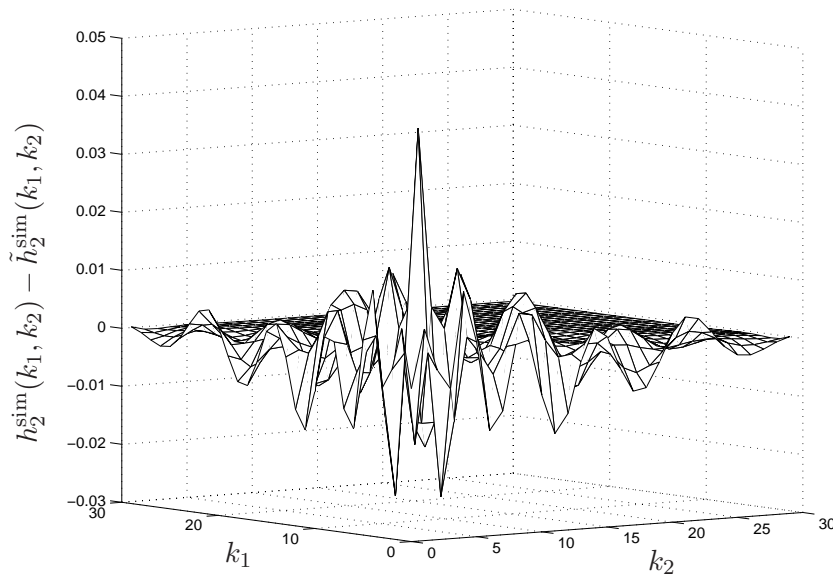


Figura 5.6: Superfície de erro para a versão simétrica do *kernel*, com $b = 0,684$ e $c = -0,46949$

Uma observação na figura 5.6 permite concluir que o pico proveniente do erro de aproximação do *kernel* $h_2^{\text{sim}}(k_1, k_2)$ é menor do que aquele apresentado na figura 5.5, que representa a versão original do *kernel* (sem simetrização). Isso se deve ao fato de que a “energia” do pico da figura 5.5 foi distribuída nas duas direções, após a simetrização do *kernel*.

5.1.2 Comparação Entre Funções de Kautz e de Laguerre

Discutiu-se na seção 3.1 que as funções de Kautz permitem melhores aproximações que as funções de Laguerre, na representação de sistemas com dinâmica oscilatória. Isso se deve ao fato de que as funções de Kautz são parametrizadas por pólos complexos, que caracterizam sistemas com oscilações. Já as funções de Laguerre são parametrizadas por um pólo real.

Para se comparar a qualidade da representação de sistemas com oscilação, usando modelos Wiener/Volterra, o *kernel* da equação (5.1) foi também estudado para o caso de Laguerre. O pólo ótimo de Laguerre pode ser calculado usando-se o método apresentado em [9], por exemplo, sendo igual a 0,40393. A tabela 5.2 apresenta o Erro Quadrático Normalizado (EQN) proveniente da aproximação desse sistema, comparando o caso em que é utilizado o pólo ótimo de Laguerre com o caso de Kautz.

Tabela 5.2: Comparação do erro de aproximação do *kernel* (5.1) de acordo com o número de funções utilizadas na expansão, para Laguerre e Kautz

Número de funções (M)	EQN (Kautz)	EQN (Laguerre)
2	0,82514	0,92533
4	0,02425	0,10157
6	0,00662	0,01431

Pode-se notar na tabela 5.2 que o erro de aproximação do *kernel* quando se utiliza funções de Laguerre é consideravelmente maior, comparado com o caso de Kautz. Com $M = 8$ funções de Laguerre, obtém-se $\text{EQN} = 2,0312 \cdot 10^{-3}$, valor quase obtido com somente 6 funções de Kautz. Como era de se esperar, sistemas com dinâmica oscilatória como o da equação (5.1) são melhor representados por modelos Wiener/Volterra quando a base utilizada é a das funções de Kautz, em comparação com base de Laguerre.

O estudo realizado acima é importante na escolha da base ortonormal a ser utilizada, o que depende fortemente da característica dinâmica do sistema. Na seção a seguir, serão estudados aspectos que ilustram o tipo de pólo que melhor representa determinado sistema, de acordo com a sua dinâmica.

5.2 Influência da Dinâmica do Sistema no Cálculo do Pólo de Kautz

Conforme apresentado na seção 3.4, as funções de Laguerre são parametrizadas pelo pólo real c . O problema da escolha ótima desse pólo foi estudado, em termos de soluções analíticas, por [6], [7], [8], [9] e [15]. Nestes trabalhos, concluiu-se que sistemas com dinâmica dominante de primeira ordem modelados por base ortonormal de Laguerre admitem como pólo ótimo o pólo que dita aquela dinâmica. Seria natural então imaginar que o cálculo dos pólos que melhor aproximam um sistema

representado via modelos Wiener/Volterra siga o próprio pólo que caracteriza sua dinâmica.

Nesta seção estuda-se o cálculo do pólo de Kautz para sistemas com dinâmica de segunda ordem, ou seja, regida por um par de pólos complexos conjugados. O objetivo é verificar o comportamento do pólo ótimo de Kautz em função da dinâmica (pólos) do sistema. Considera-se o caso em que os *kernels* são simétricos, por questão de simplicidade.

Sistemas com dinâmica de segunda ordem possuem *kernels* de Volterra expressos no domínio complexo z por:

$$H_\eta(z_1, \dots, z_\eta) = \prod_{l=1}^{\eta} \left(C_l \frac{z_l}{z_l - \zeta} + \bar{C}_l \frac{z_l}{z_l - \bar{\zeta}} \right) \quad (5.5)$$

onde C_l são constantes complexas, ζ e $\bar{\zeta}$ são os pólos complexos conjugados que caracterizam a dinâmica do *kernel*. Equivalentemente, no domínio do tempo pode-se escrever (5.5) como:

$$h_\eta(k_1, \dots, k_\eta) = \prod_{l=1}^{\eta} (C_l \zeta^{k_l} + \bar{C}_l \bar{\zeta}^{k_l}) \quad k_l \geq 0 \quad (5.6)$$

Seja então o problema do cálculo do pólo ótimo de Kautz para um sistema com dinâmica conforme discutido acima. Usando a idéia da equação (5.6) para $\eta = 2$, a escolha de $C_l = \bar{C}_l = 1/2$ (com $l = 1$ e 2), permite escrever o *kernel* de Volterra de segunda ordem desse sistema como:

$$h_2(k_1, k_2) = \left(\frac{\zeta^{k_1} + \bar{\zeta}^{k_1}}{2} \right) \cdot \left(\frac{\zeta^{k_2} + \bar{\zeta}^{k_2}}{2} \right) \quad (5.7)$$

Realizando a mudança de variáveis $\zeta = \exp(\rho)$ com $\rho = \sigma + i\omega$, o *kernel* (5.7) é expresso como:

$$\begin{aligned} h_2(k_1, k_2) &= \left(\frac{e^{\rho k_1} + e^{\bar{\rho} k_1}}{2} \right) \cdot \left(\frac{e^{\rho k_2} + e^{\bar{\rho} k_2}}{2} \right) \\ &= \left(\frac{e^{(\sigma+i\omega)k_1} + e^{(\sigma-i\omega)k_1}}{2} \right) \cdot \left(\frac{e^{(\sigma+i\omega)k_2} + e^{(\sigma-i\omega)k_2}}{2} \right) \\ &= e^{\sigma(k_1+k_2)} \cos(\omega k_1) \cos(\omega k_2) \end{aligned} \quad (5.8)$$

onde deve-se ter $\sigma = \mathbf{Re}(\rho) < 0$ para que o sistema seja assintoticamente estável.

A relação $\zeta = e^{\sigma+i\omega} = e^\sigma (\cos \omega + i \sin \omega)$ permite concluir que $|\zeta| = \exp(\sigma)$, e portanto $\sigma < 0$ implica $|\zeta| < 1$, o que garante equivalentemente a estabilidade do sistema. Além disso, $\omega = \mathbf{Im}(\rho)$ pode ser visto como a frequência com que o *kernel*

oscila em cada direção axial, e a relação $\zeta = e^{\sigma}(\cos \omega + i \sin \omega)$ permite concluir também que ω é a defasagem associada ao pólo ζ .

A figura 5.7 ilustra o *kernel* da equação (5.8) para $\sigma = -0,2$ e $\omega = 2$. Neste caso, tem-se $\zeta = \exp(-0,2 + 2i)$.

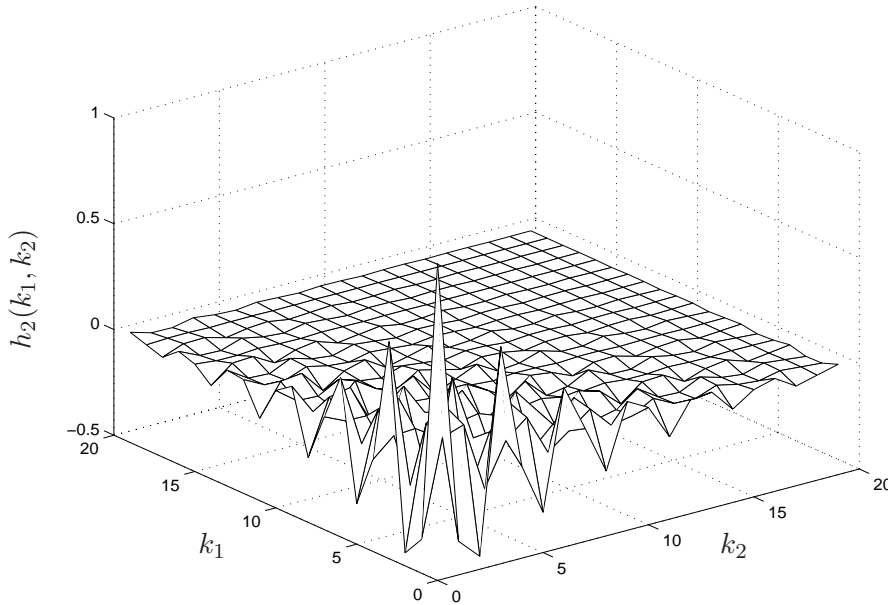


Figura 5.7: *Kernel* de segunda ordem da equação (5.8) para $\sigma = -0,2$ e $\omega = 2$

Para estudar a influência da dinâmica do sistema no cálculo do pólo ótimo de Kautz, variou-se os valores de σ (fator de decaimento) e ω (frequência de oscilação) dentro da região de estabilidade do *kernel* da equação (5.7). As figuras 5.8 e 5.9 ilustram² os resultados obtidos dessa simulação.

Conforme comentou-se anteriormente, $\sigma < 0$ garante a estabilidade do sistema. Quando $\sigma \rightarrow 0$ (puramente oscilatório), o sistema atinge a região-limite de estabilidade. Esse fato pode ser verificado na figura 5.8, onde os pólos são calculados para alguns valores de ω mantendo-se σ constante. O sistema em questão admite pólo ótimo de Kautz próximo do limite do círculo unitário ($|\zeta| \rightarrow 1$) à medida que o fator de decaimento σ do *kernel* se aproxima de 0^- . E como $|\zeta| = \exp(\sigma)$, o pólo de Kautz situa-se próximo da origem ($|\zeta| \rightarrow 0$) quando σ se torna mais negativo.

Foi discutido também que ω representa a oscilação do pólo ζ . Isso pode ser observado na figura 5.9, em que o pólo ótimo de Kautz é calculado para alguns valores de σ , mantendo-se ω constante. O valor de ω é igual ao ângulo (em radianos) que cada pólo possui com relação ao eixo real do plano complexo. Em particular, $\omega \rightarrow 0$ indica que o sistema não possui oscilações, de acordo com a equação (5.8). E a relação $\zeta = e^{\sigma}(\cos \omega + i \sin \omega)$ fornece um pólo puramente real, igual a $\zeta = e^{\sigma}$. Isso é intuitivo, pois o *kernel* seria caracterizado unicamente por uma exponencial

²Como os pólos ilustrados aparecem aos pares conjugados, por questão de conveniência, mostra-se somente um dos pólos.

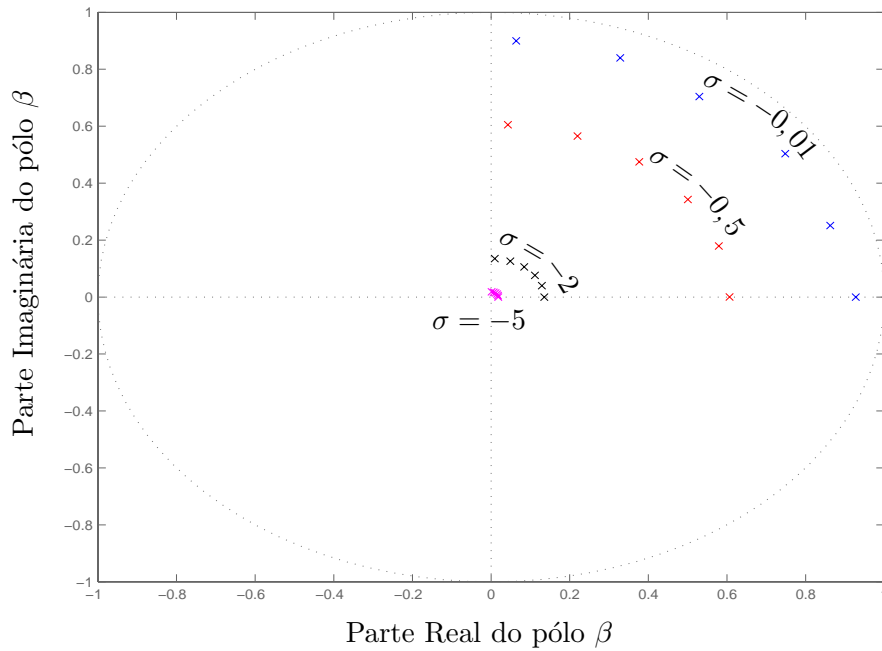


Figura 5.8: Pólo ótimo de Kautz quando se varia a frequência de oscilação ω do *kernel*, mantendo-se σ constante

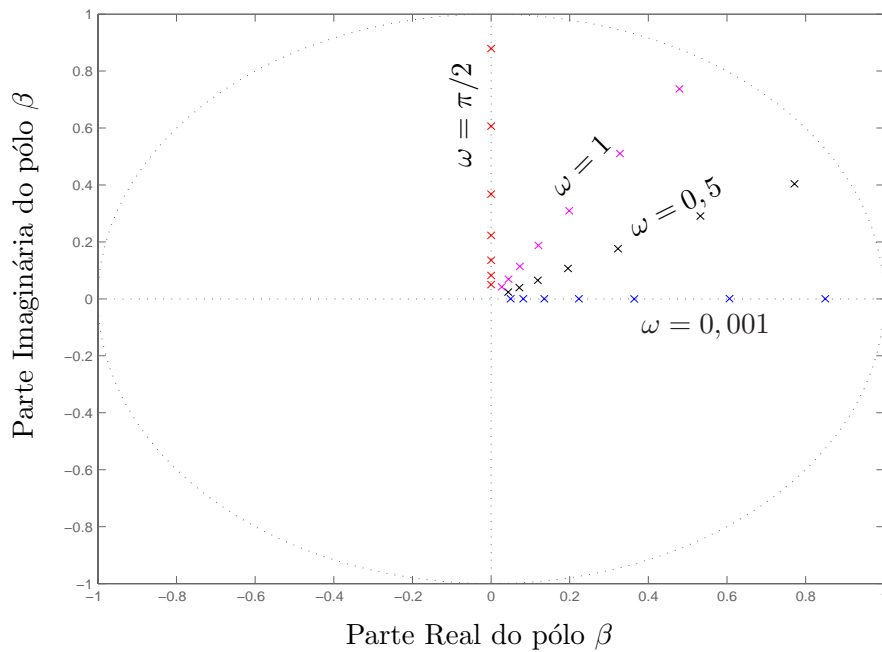


Figura 5.9: Pólo ótimo de Kautz quando se varia o fator de decaimento σ do *kernel*, mantendo-se ω constante

decrecente, o que significa decaimento puro sem oscilações. Com $\omega = \pi/2$ tem-se $\zeta = ie^\sigma$, que é um pólo puramente complexo. Isso também pode ser observado na

figura 5.9, onde os pólos de Kautz situam-se sobre o eixo imaginário.

Considere agora o caso em que se deseja comparar um sistema que possua dinâmica de segunda ordem (pólos complexos conjugados) com a sua aproximação em base de Kautz. Suponha que um determinado sistema físico possua *kernel* de segunda ordem do tipo da equação (5.8), para $\zeta = \exp(-0,2 + 2i) = -0,3407122128 + i0,7444697670$. Utilizando o método proposto na seção 4.5 ao se escolher o valor de $b = (\zeta + \bar{\zeta})/(1 + \zeta\bar{\zeta}) = -0,40796$, o respectivo valor de c calculado foi igual a $c_{\text{ótimo}} = -0,67032$. Com esse par $(b; c_{\text{ótimo}}) = (-0,40796; -0,67032)$, obteve-se como pólo calculado $\beta = -0,3407122127 + i0,7444697667$, muito próximo ao valor de ζ .

A qualidade da aproximação do sistema quando o pólo de Kautz utilizado no modelo Wiener/Volterra é igual ao pólo dominante do sistema pode ser observada na figura 5.10. Ilustra-se o erro de aproximação entre o *kernel* da equação (5.8) para $\sigma = -0,2$ e $\omega = 2$, e sua expansão com $M = 6$ funções de Kautz parametrizadas no pólo ótimo β .

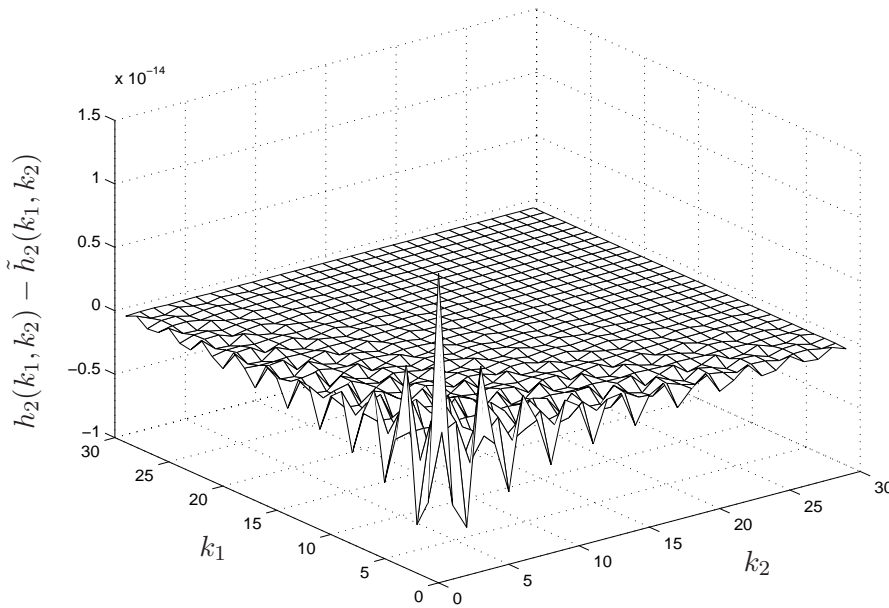


Figura 5.10: Superfície de erro proveniente da aproximação do *kernel* da equação (5.8) com $\sigma = -0,2$ e $\omega = 2$, para o caso em que o pólo de Kautz utilizado é igual ao parâmetro ζ

A superfície do erro de aproximação é muito próxima de zero para quaisquer valores de k_1 e k_2 , o que verifica a qualidade da aproximação. É possível mostrar que o mesmo ocorre para quaisquer valores de σ e ω dentro da região de estabilidade do sistema. Esse resultado mostra empiricamente que o pólo de Kautz β que melhor representa um *kernel* do tipo (5.5) é igual ao pólo ζ que rege a dinâmica daquele *kernel*.

5.3 Estimação Simultânea dos *Kernels* de Volterra e do Pólo de Kautz

A utilização do resultado matemático obtido na seção 4.5 e a aplicação da proposição 4.4.1 requer o conhecimento *a priori* dos *kernels* do sistema. Quando essa informação não está disponível, eles devem ser estimados a partir de um conjunto de dados entrada/saída de um determinado sistema. Propõe-se a seguir um algoritmo que estima os *kernels* do processo, ao mesmo tempo em que calcula o pólo sub-ótimo de Kautz de acordo com o método apresentado no Capítulo 4. Esse processo continua iterativamente, até a convergência dos parâmetros. Esse algoritmo representa uma generalização do algoritmo apresentado em [6], [7] e [9], que utiliza as funções de Laguerre. Neste trabalho não são estudadas as condições de convergência deste algoritmo.

Algoritmo

1. Selecionar a ordem do modelo de Volterra a ser utilizado, como na equação (2.10). Modelos de ordem superior a 2 possuem elevado custo computacional.
2. Escolher um truncamento dos *kernels* a serem estimados, ou seja, o valor de k_1, k_2, \dots, k_η acima dos quais os *kernels* de Volterra são considerados nulos. Normalmente, esses valores podem adotados ao se observar o tempo de resposta do sistema.
3. Escolher o número M de funções de Kautz que serão utilizados na aproximação do *kernel* $h_\eta(k_1, \dots, k_\eta)$.
4. Selecionar um pólo de Kautz inicial, para cada *kernel*.
5. A partir dos dados de entrada e saída disponíveis, estimar via mínimos quadrados os coeficientes $\alpha_{(\cdot)}$ da equação (4.14).
6. Calcular os *kernels* de Volterra utilizando a equação (4.13).
7. Obter o novo pólo de Kautz a partir dos *kernels* estimados no passo 6, conforme o procedimento ilustrado nas figuras 5.2 e 5.3, e retornar ao passo 5.

Para ilustrar o desempenho desse algoritmo, foram utilizados dados medidos em laboratório de um sistema de levitação magnética, descrito a seguir.

5.3.1 Modelo do Levitador Magnético

O levitador magnético com dois ímãs permanentes e duas bobinas se encontra no Laboratório de Controle e Servomecanismos da FEEC/UNICAMP, sendo esquematicamente descrito de acordo com a figura 5.11.

O sistema é composto por duas bobinas que, ao serem percorridas por correntes elétricas, criam campos magnéticos que interagem com os campos dos respectivos

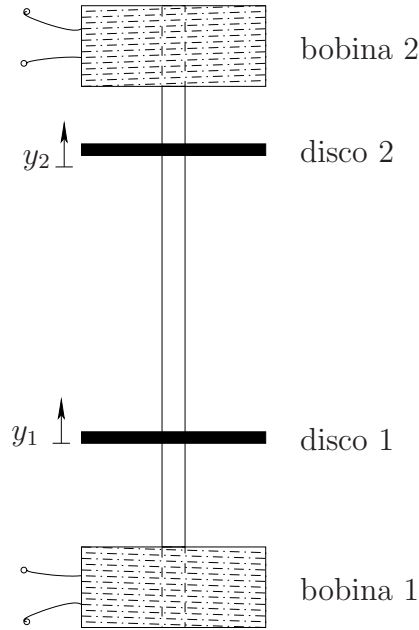


Figura 5.11: Diagrama do levitador magnético

discos magnéticos permanentes. Assim, surgem forças mecânicas sobre esses magnetos, que irão movimentá-los de acordo com a corrente elétrica que atravessa cada bobina. O sentido da corrente em cada bobina e a polaridade de cada ímã são tais que eles criam uma força mecânica resultante sobre cada disco. Essas forças são do tipo:

- Repulsiva para o magneto inferior acionado pela bobina inferior;
- Atrativa para o magneto superior acionado pela bobina superior.

Considere então o movimento do disco magnético inferior. Sendo m_1 a massa do disco, c_2 o coeficiente de atrito e g a aceleração da gravidade, a equação que descreve o movimento do disco é:

$$m_1 \ddot{y}_1 + c_1 \dot{y}_1 + F_{m12} = F_{u11} - F_{u21} - m_1 g \quad (5.9)$$

onde F_{u11} é a força magnética entre a bobina 1 e o disco 1, F_{u21} é a força magnética entre a bobina 2 e o disco 1, e F_{m12} é a força magnética entre os discos. Sendo ainda que:

$$F_{m12} = \frac{c}{(y_c + y_2 - y_1 + d)^4}$$

$$F_{u11} = \frac{i_1}{a(k_s y_1 + b)^4}$$

$$F_{u21} = \frac{i_2}{a(y_c + k_s y_1 + b)^4}$$

com:

- m_1 : massa do disco magnético 1;
- c_1 : coeficiente de atrito viscoso do disco 1 com o ar;
- y_1 : altura do disco 1;
- y_2 : altura do disco 2;
- i_1 : corrente elétrica na bobina 1;
- i_2 : corrente elétrica na bobina 2;
- y_c : distância entre as bobinas 1 e 2;
- a, b, c, d : constantes reais.

A seção 5.3.2 apresenta resultados de simulação do algoritmo acima, realizada com dados entrada/saída do levitador magnético.

5.3.2 Resultados de Simulação

Os dados experimentais foram obtidos mantendo-se constante a corrente elétrica na bobina 1. O sistema é instável em malha aberta, e a aplicação de uma corrente constante em uma das bobinas faz com que o sistema se estabilize. Para o estudo deste sistema, não considerou-se a presença de ruído no sinal de entrada $u(k)$ e no sinal de saída $y(k)$.

Aplicando-se degraus de amplitudes variáveis de corrente na bobina 2 (entrada u), mediu-se então a posição do disco 1 (saída y). A figura 5.12 mostra os dados de entrada e saída do levitador magnético que são utilizados para a fase de estimação do modelo. Enfatiza-se que esses dados foram normalizados no intervalo $[-1, 1]$, para evitar problemas numéricos na implementação computacional.

Deseja-se então encontrar um modelo que relaciona os dados de entrada (u) e saída (y) da figura 5.12 como uma representação do tipo (2.10) para $\eta = 2$, isto é:

$$y(k) = \sum_{\tau_1=0}^k h_1(\tau_1)u(k - \tau_1) + \sum_{\tau_1=0}^k \sum_{\tau_2=0}^k h_2(\tau_1, \tau_2)u(k - \tau_1)u(k - \tau_2) \quad (5.10)$$

onde os *kernels* h_1 e h_2 devem ser expandidos usando-se as funções de base ortonormal de Kautz $\psi_l(k)$, ou seja:

$$h_1(k_1) = \sum_{i_1=1}^M \alpha_{i_1} \psi_{i_1}(k_1) \quad (5.11)$$

$$h_2(k_1, k_2) = \sum_{i_1=1}^M \sum_{i_2=1}^M \alpha_{i_1, i_2} \psi_{i_1}(k_1) \psi_{i_2}(k_2) \quad (5.12)$$

Por questão de simplicidade, admite-se nesta simulação que o *kernel* de segunda ordem é simétrico. Para a estimação dos *kernels* de Volterra a partir dos coeficientes

5.3 Estimação Simultânea dos *Kernels* de Volterra e do Pólo de Kautz

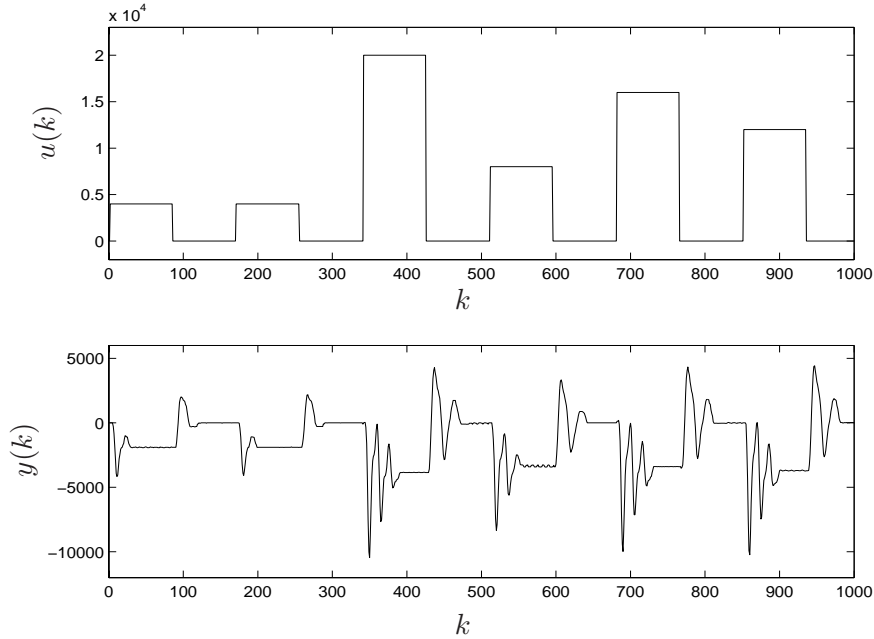


Figura 5.12: Acima: corrente elétrica na bobina 2 (sinal de entrada u). Abaixo: posição do disco 1 (sinal de saída y)

$\alpha_{(\cdot)}$, foram utilizadas $M = 6$ funções de Kautz.

Escolheu-se o par inicial de pólos de Kautz $\beta = 0,5 \pm 0,5i$ para ambos os *kernels*. Com esses valores, tem-se $b = 0,66666$ e $c = -0,5$. A evolução dos valores dos parâmetros b e c , calculados para os dois kernels, está ilustrada nas figuras 5.13 e 5.14.

Os valores finais dos parâmetros e pólos de Kautz estão mostrados na tabela 5.3, para os *kernels* de primeira e segunda ordem.

Tabela 5.3: Valores finais dos parâmetros de Kautz, após 30 iterações do algoritmo

Ordem do <i>kernel</i> (η)	b	c_{otimo}	β
1	0,90357	-0,68024	$0,75911 \pm 0,32248i$
2	0,92277	-0,69416	$0,78166 \pm 0,28839i$

Utilizando os valores dos pólos de Kautz mostrados na tabela 5.3, pode-se estimar os *kernels* $h_1(k_1)$ e $h_2(k_1, k_2)$ de acordo com as equações (5.11) e (5.12), respectivamente. As figuras 5.15 e 5.16 ilustram os *kernels* de primeira a segunda ordem estimados pelo algoritmo.

Depois de 30 iterações do algoritmo, os parâmetros de Kautz “convergiram” no sentido de que a diferença entre seus valores correntes e os respectivos valores finais se tornou desprezível. Agora, a equação (5.10) é utilizada para estimar a saída do sistema, usando-se um outro conjunto de dados para a validação do modelo. Obtém-

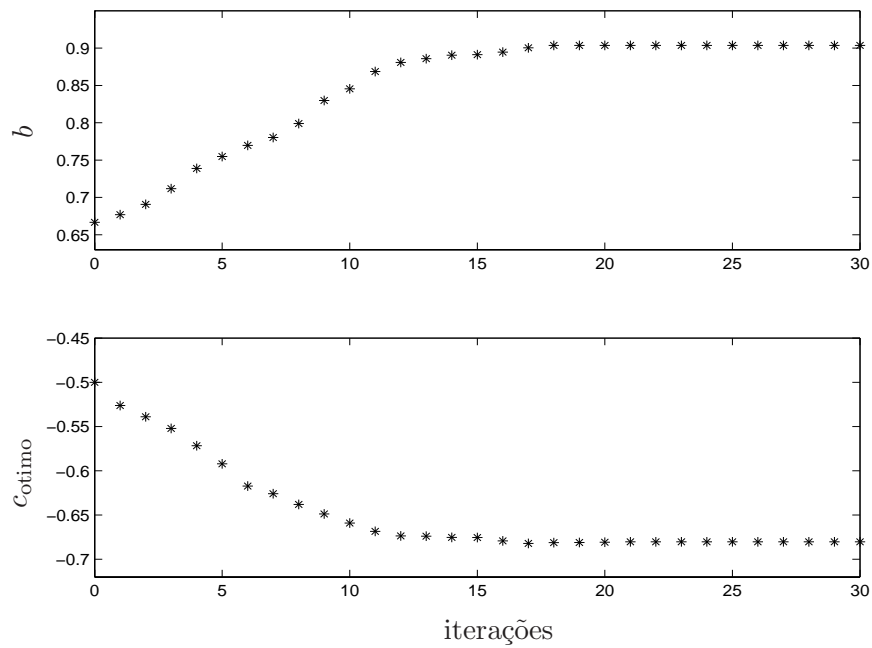


Figura 5.13: Evolução dos parâmetros de Kautz, de acordo com o número de iterações, para o *kernel* de primeira ordem

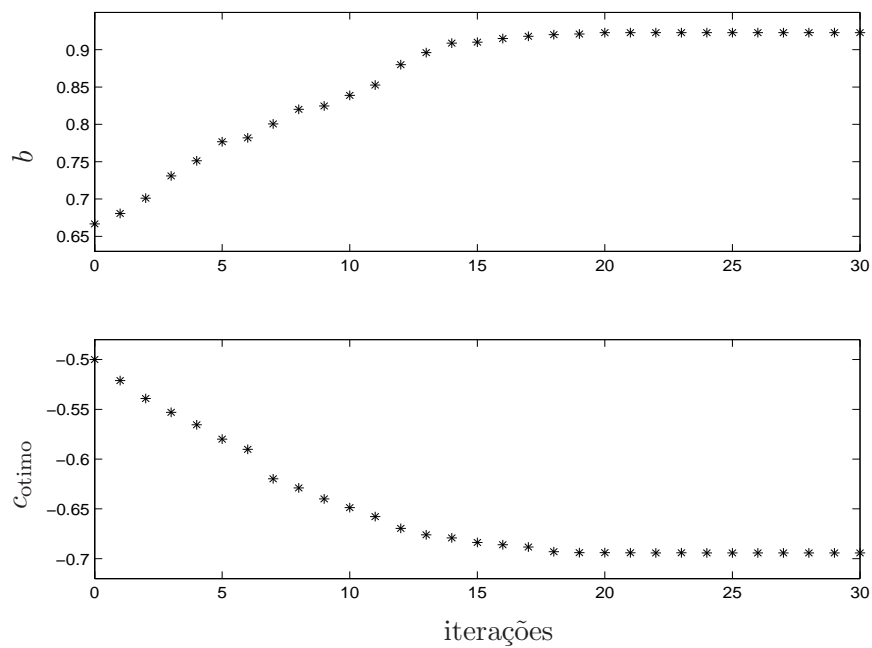


Figura 5.14: Evolução dos parâmetros de Kautz, de acordo com o número de iterações, para o *kernel* de segunda ordem

se então a previsão de $y(k)$, mostrada na figura 5.17 juntamente com o seu valor real (dados de validação). A figura 5.18 ilustra o erro proveniente dessa estimação.

5.3 Estimação Simultânea dos *Kernels* de Volterra e do Pólo de Kaut

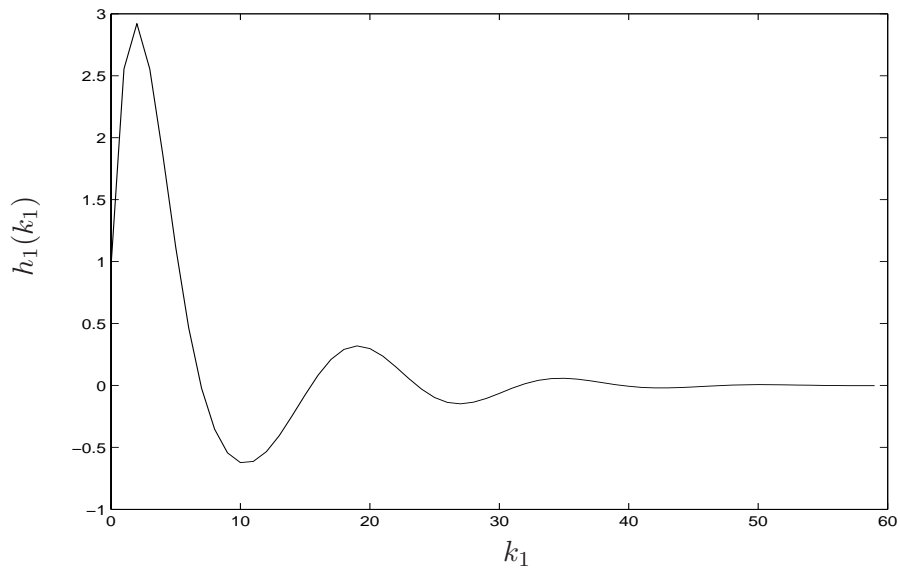


Figura 5.15: *Kernel* de primeira ordem estimado para o modelo (5.10), após 30 iterações do algoritmo

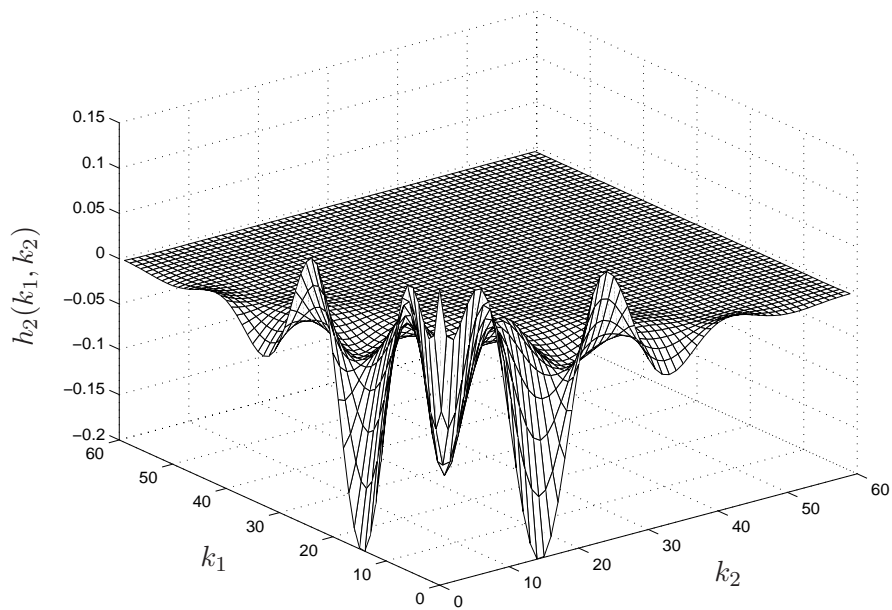


Figura 5.16: *Kernel* de segunda ordem estimado para o modelo (5.10), após 30 iterações do algoritmo

A qualidade da estimação pode ainda ser observada na figura 5.19, que mostra a evolução do Erro Quadrático Normalizado (EQN) entre a saída medida $y(k)$ e sua estimativa calculada pelo algoritmo. O valor final do erro quadrático, após 30 iterações do algoritmo, é $\text{EQN} = 5,07 \cdot 10^{-3}$.

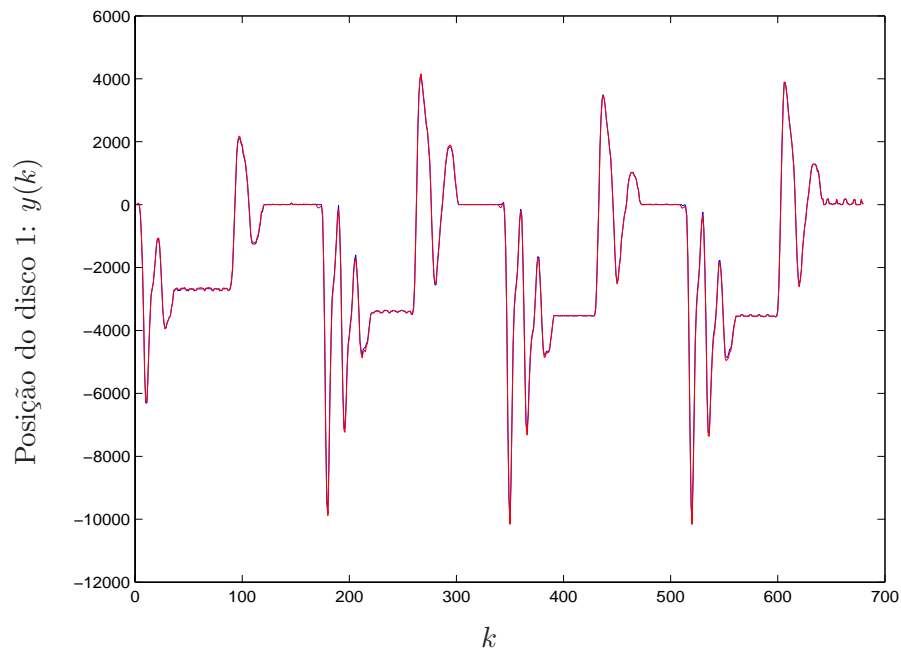


Figura 5.17: Saídas medida (em azul) e estimada (em vermelho), após 30 iterações do algoritmo

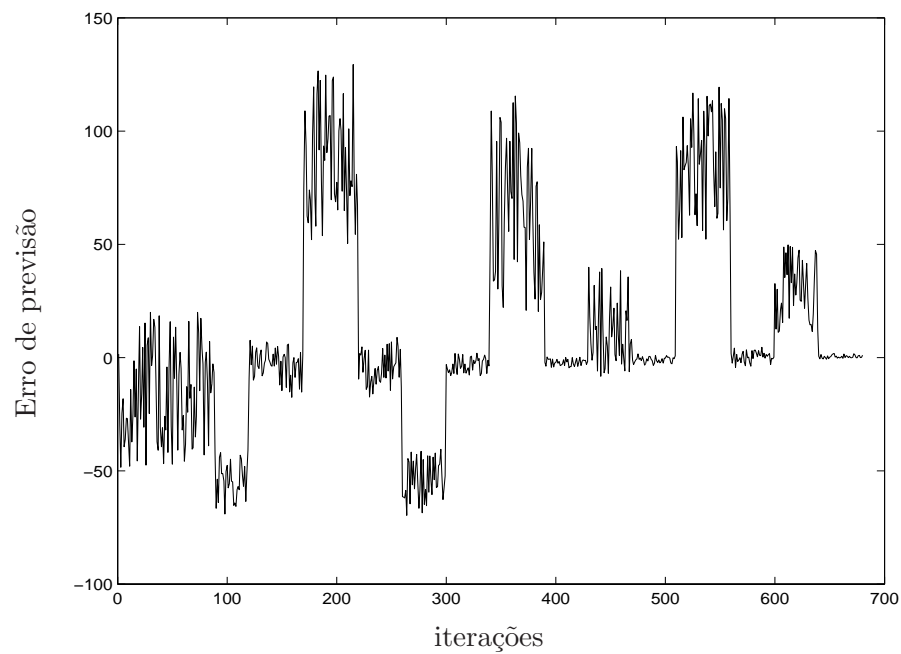


Figura 5.18: Erro de estimação: diferença entre os dados de validação de $y(k)$ e os dados da saída estimada pelo modelo

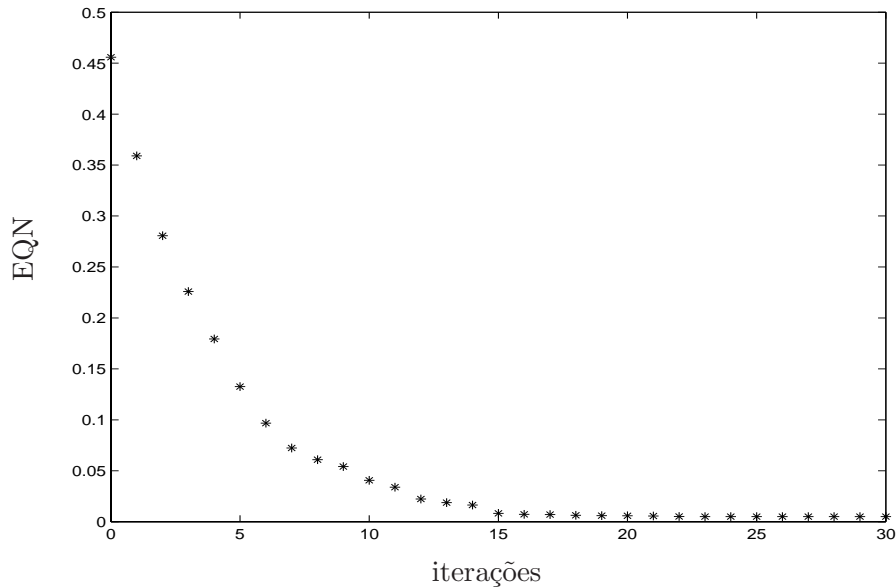


Figura 5.19: Evolução do Erro Quadrático Normalizado (EQN) entre os dados de validação da saída $y(k)$ e a sua estimativa

5.4 Resumo e Discussões

Este capítulo apresentou exemplos ilustrativos dos resultados matemáticos obtidos no Capítulo 4. Foi ilustrado inicialmente o exemplo de um sistema com *kernel* de Volterra de segunda ordem não-separável e assimétrico. Para esse exemplo, obteve-se o pólo sub-ótimo de Kautz calculado pelo método proposto no capítulo anterior.

Discutiu-se também a questão da simetrização de *kernels* e a sua importância na modelagem computacional de sistemas físicos. Os resultados das simulações (assimétrica e simétrica) puderam ser comparados qualitativamente, por meio de aproximações usando-se a base de Kautz. Simulações computacionais mostraram ainda que sistemas com dinâmica oscilatória são melhor representados usando-se uma base de funções de Kautz, em comparação com Laguerre, para um mesmo número de parâmetros utilizados no modelo.

Outro assunto estudado foi a influência do comportamento dinâmico do sistema no cálculo do pólo de Kautz. Para *kernels* com dinâmica de segunda ordem, pôde-se comprovar empiricamente que o pólo ótimo de Kautz é o próprio pólo que caracteriza essa dinâmica (pólos complexos conjugados). Esse fato foi verificado ao se variar o fator de decaimento e a frequência de oscilação desses *kernels*, calculando-se seu respectivo pólo ótimo.

Um algoritmo para estimação simultânea dos *kernels* de Volterra e dos pólos de Kautz também foi sintetizado. Ele foi aplicado à estimação dos *kernels* de um sistema de levitação magnética, a partir da medição de dados de entrada e saída. Sua implementação permitiu obter um modelo Wiener/Volterra de segunda ordem que representa adequadamente a dinâmica daquele sistema.

Capítulo 6

Conclusões e Comentários Finais

Nas últimas décadas, o estudo de modelagem e controle de sistemas dinâmicos tem tido grande interesse na engenharia. O estudo dessas áreas tem crescido em importância, o que levou à necessidade de modelos mais detalhados para descrever a dinâmica dos processos físicos. Tendo isso em mente, este trabalho abordou a representação de sistemas não-lineares utilizando-se modelos Wiener/Volterra.

Os modelos de Volterra representam uma generalização dos modelos lineares de resposta ao impulso para a descrição de sistemas dinâmicos não-lineares. Contudo, esses modelos necessitam de um número considerável de termos para a representação dos *kernels* de Volterra. Esse número de termos pode ser reduzido utilizando-se uma representação do tipo Wiener/Volterra, em que os *kernels* são desenvolvidos utilizando uma base de funções ortonormais. Esses modelos são capazes de aproximar arbitrariamente bem os sistemas *fading memory* (memória finita). Por sua vez, os modelos Wiener/Volterra podem ter sua complexidade reduzida realizando-se uma seleção ótima da base de funções ortonormais, o que permite uma redução no grau de sua representação. Esse aspecto leva ao cálculo de pólos ótimos das funções ortonormais, assunto abordado aqui.

O presente trabalho estudou o problema da determinação ótima dos parâmetros livres que caracterizam as funções ortonormais utilizadas na expansão dos *kernels* de Volterra. Enfatizou-se a seleção ótima de um dos parâmetros das funções de Kautz, que são parametrizadas por um par de pólos complexos conjugados. O principal resultado nesse contexto foi a obtenção de uma solução analítica global para a expansão sub-ótima dos *kernels* de Volterra usando funções ortonormais de Kautz. Essa solução minimiza o limite superior do erro de aproximação associado ao desenvolvimento truncado dos *kernels* em um número finito de funções, permitindo reduzir a complexidade de representação do modelo resultante.

Outro ponto abordado foi a comparação de resultados quando utiliza-se funções ortonormais de Kautz ou de Laguerre na representação dos modelos Wiener/Volterra. Resultados de simulações ilustraram que sistemas com comportamento oscilatório são melhor representados utilizando-se as funções de Kautz, comparando-se com a base de Laguerre, para um mesmo número de funções. Essa comparação foi realizada avaliando-se o erro quadrático obtido na expansão de *kernels* de Volterra, ao se utilizar essas duas bases de funções ortonormais.

Estudou-se também a influência do comportamento dinâmico do sistema no cálculo do pólo ótimo de Kautz. Resultados de simulação mostraram que o modelo Wiener/Volterra que melhor representa sistemas com dinâmica de segunda ordem (par de pólos complexos conjugados) é aquele em que o pólo de Kautz selecionado é o pólo característico do sistema.

Propôs-se ainda um algoritmo iterativo para estimação dos *kernels* de Volterra a partir de um conjunto de dados entrada/saída. Este algoritmo permite a obtenção simultânea dos *kernels* e do pólo de Kautz relacionado a esses *kernels*. Para ilustrar a aplicação desse algoritmo, foi apresentado o processo de identificação de um sistema de levitação magnética do Laboratório de Controle e Servomecanismos da FEEC/UNICAMP. Simulações computacionais ilustraram que o algoritmo permitiu obter um modelo que representa adequadamente as dinâmicas do sistema.

6.1 Contribuições

A principal contribuição do presente trabalho foi a obtenção de uma expressão analítica para um dos parâmetros das funções de Kautz, em função do outro parâmetro. Essa solução pode ser interpretada como a escolha sub-ótima do pólo de Kautz, obtida a partir da minimização de um limitante superior do erro de aproximação dos *kernels* de Volterra. Outra contribuição foi a dedução de um teorema, até então válido para modelos Wiener/Volterra de primeira ordem, estendido para modelos de ordem superior. Uma vez que o procedimento apresentado neste trabalho para a obtenção do pólo sub-ótimo de Kautz requer o cálculo de algumas funções auxiliares, aquele teorema permite que tais funções sejam obtidas.

Uma outra contribuição foi a apresentação de um modelo de Volterra para um sistema de levitação magnética, usando-se funções de Kautz. Um modelo equivalente, usando-se funções de Laguerre, já foi apresentado em [6], [7] e [9]. Aqui, esse modelo já existente foi estendido com a utilização da base de funções ortonormais de Kautz.

6.2 Perspectivas Futuras

As perspectivas para trabalhos futuros estão baseadas no estudo das funções ortonormais generalizadas [5], cuja utilização em identificação de sistemas dinâmicos tem sido objeto de estudo nas últimas décadas [35]. A descrição matemática daquelas funções é obtida pelo estudo de filtros digitais recursivos, conforme discutiu-se no presente trabalho. No aspecto para pesquisas futuras, inclui-se o cálculo (por meio de abordagens analíticas) de bases ótimas de funções ortonormais generalizadas utilizadas na representação de modelos Wiener/Volterra.

Uma segunda perspectiva está relacionada com a otimização simultânea, em termos de uma abordagem analítica, dos parâmetros de Kautz. Como esse é um problema ainda em aberto, ele pode vir a ser objeto de estudo futuro. Ao invés da otimização de um limitante superior do erro de aproximação dos *kernels* de Volterra usando-se funções de Kautz, um futuro trabalho seria a otimização do próprio erro

de aproximação dos *kernels*. Ainda neste aspecto, uma perspectiva seria o estudo de outras normas para o cálculo desse erro.

Um outro estudo futuro envolve o cálculo de pólos de Kautz diferentes para cada direção do *kernel*, o que poderia fornecer melhores resultados em termos de representação dos modelos de Volterra. Ou ainda, a obtenção do pólo ótimo de Kautz usando-se uma metodologia análoga à que foi apresentada neste trabalho: otimizar um dos parâmetros de Kautz, quando o outro parâmetro é considerado fixo.

Outra perspectiva é o estudo de condições de convergência do algoritmo de estimação simultânea dos *kernels* de Volterra e do pólo de Kautz para o desenvolvimento daqueles. Conforme foi comentado, esse algoritmo possui apenas um caráter intuitivo, permitindo assim que condições formais de convergência possam ser estudadas.

Apêndice A

Transformada Z Multidimensional

Em alguns pontos deste trabalho, houve a necessidade de se utilizar a transformada Z . O presente apêndice tem o objetivo de apresentar esse operador e exemplificar a sua utilização no estudo de sinais discretos no tempo.

A.1 Definição de Transformada Z Multidimensional

Seja a seguinte definição [13]:

Definição A.1.1. *Considere uma seqüência discreta de n variáveis, $x(k_1, \dots, k_n)$. Suponha que ela seja nula se algum dos inteiros k_j ($j = 1, 2, \dots, n$) for negativo. Defina-se a sua transformada Z multidimensional como sendo o operador:*

$$\mathcal{Z}[x(k_1, \dots, k_n)] = X(z_1, \dots, z_n) \triangleq \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} x(k_1, \dots, k_n) z_1^{-k_1} \cdots z_n^{-k_n} \quad (\text{A.1})$$

A definição acima é uma série de potências não-positivas das variáveis complexas z_j , e pode ser vista como uma generalização da transformada Z monovariável.

A.1.1 Exemplo

Como exemplo de aplicação da definição acima, suponha que se deseje calcular a transformada Z da seguinte seqüência de 2 variáveis:

$$x(k_1, k_2) = k_1 - k_1 \lambda^{-k_2} \quad (\text{A.2})$$

onde λ é uma constante real não-nula. Aplicando-se a definição (A.1) para o caso em que $n = 2$, tem-se que:

$$\begin{aligned}
X(z_1, z_2) &= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} x(k_1, k_2) z_1^{-k_1} z_2^{-k_2} \\
&= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} (k_1 - k_1 \lambda^{-k_2}) z_1^{-k_1} z_2^{-k_2} \\
&= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} k_1 z_1^{-k_1} z_2^{-k_2} - \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} k_1 \lambda^{-k_2} z_1^{-k_1} z_2^{-k_2} \\
&= \left(\sum_{k_1=0}^{\infty} k_1 z_1^{-k_1} \right) \cdot \left(\sum_{k_2=0}^{\infty} z_2^{-k_2} \right) - \left(\sum_{k_1=0}^{\infty} k_1 z_1^{-k_1} \right) \cdot \left(\sum_{k_2=0}^{\infty} (\lambda z_2)^{-k_2} \right)
\end{aligned} \tag{A.3}$$

e, lembrando de resultados bem conhecidos para a transformada Z monovariável, a equação (A.3) resulta que:

$$\begin{aligned}
X(z_1, z_2) &= \frac{z_1}{(z_1 - 1)^2} \cdot \frac{z_2}{z_2 - 1} - \frac{z_1}{(z_1 - 1)^2} \cdot \frac{z_2}{z_2 - \lambda^{-1}} \\
&= \frac{(1 - \lambda^{-1})z_1 z_2}{(z_1 - 1)^2 (z_2 - 1) (z_2 - \lambda^{-1})}
\end{aligned} \tag{A.4}$$

Apêndice B

Demonstrações

Esse apêndice destina-se a apresentar as demonstrações de proposições e de algumas equações. Tais demonstrações foram omitidas no corpo principal do trabalho, por conveniência.

B.1 Demonstração da Equação (4.47)

As funções discretas de Laguerre foram definidas na seção 3.5, pela equação (3.15). Elas têm como função de transferência:

$$\begin{aligned}\Phi_n(z) &= \sqrt{1-c^2} \frac{z}{z-c} \left(\frac{1-cz}{z-c} \right)^{n-1} \\ &= \sqrt{1-c^2} \frac{z(1-cz)^{n-1}}{(z-c)^n}\end{aligned}\tag{B.1}$$

Derivando a equação (B.1) em relação à variável complexa z , tem-se:

$$\begin{aligned}\frac{\partial \Phi_n(z)}{\partial z} &= \sqrt{1-c^2} \cdot \\ &\cdot \left[\frac{(1-cz)^{n-1}(z-c)^n + z(-c)(n-1)(1-cz)^{n-2}(z-c)^n}{(z-c)^{2n}} + \right. \\ &\quad \left. - \frac{nz(z-c)^{n-1}(1-cz)^{n-1}}{(z-c)^{2n}} \right] \\ &= \sqrt{1-c^2} \frac{z(1-cz)^{n-1}}{(z-c)^n} \left[z^{-1} - \frac{c(n-1)}{1-cz} - \frac{n}{z-c} \right] \\ &= \Phi_n(z) \left[\frac{1-cz^{-1} - n(1-c^2)}{(1-cz)(z-c)} \right]\end{aligned}\tag{B.2}$$

A equação anterior é reescrita da seguinte forma:

$$[(1 - cz)(z - c)] \cdot \frac{\partial \Phi_n(z)}{\partial z} = [1 - cz^{-1} - n(1 - c^2)] \cdot \Phi_n(z) \quad (\text{B.3})$$

Seendo $\mathcal{Z}[x_k] = \sum_{k=0}^{+\infty} x_k z^{-k} = X(z)$ a transformada Z unilateral de uma seqüência x_k discreta no tempo ($k \in \mathbb{N}$), com condições iniciais nulas, valem as seguintes propriedades:

$$\frac{dX(z)}{dz} = -z^{-1} \mathcal{Z}[kx_k] \quad (\text{B.4})$$

$$z^{-1} X(z) = \mathcal{Z}[x_{k-1}] \quad (\text{B.5})$$

Usando as relações acima, e após algumas manipulações elementares, a equação (B.3) fica:

$$[1 - cz^{-1} - cz + c^2] \cdot \mathcal{Z}[k\phi_n(k)] = [cz^{-1} - 1 + n(1 - c^2)] \cdot \mathcal{Z}[\phi_n(k)] \quad (\text{B.6})$$

E no domínio do tempo pode-se escrever:

$$\begin{aligned} (1 + c^2)k\phi_n(k) - c(k-1)\phi_n(k-1) - c(k+1)\phi_n(k+1) &= \\ &= c\phi_n(k-1) - \phi_n(k) + n(1 - c^2)\phi_n(k) \end{aligned} \quad (\text{B.7})$$

ou ainda:

$$-c(k+1)\phi_n(k+1) + [(1 + c^2)k + 1]\phi_n(k) - ck\phi_n(k-1) = n(1 - c^2)\phi_n(k)$$

■

B.2 Demonstração da Equação (4.49)

As funções discretas de Laguerre $\phi_{i_j}(k)$ satisfazem a equação a diferenças dada em (4.47):

$$-c(k+1)\phi_{i_j}(k+1) + [(1 + c^2)k + 1]\phi_{i_j}(k) - ck\phi_{i_j}(k-1) = i_j(1 - c^2)\phi_{i_j}(k) \quad (\text{B.8})$$

A partir dela, pode-se escrever para qualquer $l \in \{1, 2, \dots, \eta\}$ o seguinte:

$$\begin{aligned}
& -c(k_l + 1)\phi_{i_l}(k_l + 1) \prod_{\substack{j=1 \\ j \neq l}}^{\eta} \phi_{i_j}(k_j) + [(1 + c^2)k_l + 1]\phi_{i_l}(k_l) \prod_{\substack{j=1 \\ j \neq l}}^{\eta} \phi_{i_j}(k_j) + \\
& -ck_l\phi_{i_l}(k_l - 1) \prod_{\substack{j=1 \\ j \neq l}}^{\eta} \phi_{i_j}(k_j) = i_l(1 - c^2)\phi_{i_l}(k_l) \prod_{\substack{j=1 \\ j \neq l}}^{\eta} \phi_{i_j}(k_j)
\end{aligned} \tag{B.9}$$

Sendo a função $x(k_1, \dots, k_l, \dots, k_\eta)$ conforme definida na seção 4.5, defina agora a variável:

$$\begin{aligned}
\wp_l & \triangleq \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} x(k_1, \dots, k_l, \dots, k_\eta) \cdot \\
& \cdot \left\{ -c(k_l + 1)x(k_1, \dots, k_l + 1, \dots, k_\eta) + [(1 + c^2)k_l + 1]x(k_1, \dots, k_l, \dots, k_\eta) + \right. \\
& \quad \left. - ck_lx(k_1, \dots, k_l - 1, \dots, k_\eta) \right\}
\end{aligned} \tag{B.10}$$

e usando (4.10) para $\phi_{i_l}(k)$, a equação anterior pode ser reescrita como:

$$\begin{aligned}
\wp_l & = \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} x(k_1, \dots, k_l, \dots, k_\eta) \cdot \\
& \cdot \left\{ -c(k_l + 1) \sum_{i_1=1}^{\infty} \cdots \sum_{i_l=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_l, \dots, i_\eta} \phi_{i_l}(k_l + 1) \prod_{\substack{j=1 \\ j \neq l}}^{\eta} \phi_{i_j}(k_j) + \right. \\
& \quad + [(1 + c^2)k_l + 1] \sum_{i_1=1}^{\infty} \cdots \sum_{i_l=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_l, \dots, i_\eta} \phi_{i_l}(k_l) \prod_{\substack{j=1 \\ j \neq l}}^{\eta} \phi_{i_j}(k_j) + \\
& \quad \left. - ck_l \sum_{i_1=1}^{\infty} \cdots \sum_{i_l=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_l, \dots, i_\eta} \phi_{i_l}(k_l - 1) \prod_{\substack{j=1 \\ j \neq l}}^{\eta} \phi_{i_j}(k_j) \right\}
\end{aligned} \tag{B.11}$$

Ainda manipulando a variável φ_l , pode-se inferir que:

$$\begin{aligned} \varphi_l = & \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} x(k_1, \dots, k_l, \dots, k_\eta) \cdot \left(\sum_{i_1=1}^{\infty} \cdots \sum_{i_l=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_l, \dots, i_\eta} \cdot \right. \\ & \cdot \left. \left\{ -c(k_l + 1)\phi_{i_l}(k_l + 1) \prod_{\substack{j=1 \\ j \neq l}}^{\eta} \phi_{i_j}(k_j) + [(1 + c^2)k_l + 1]\phi_{i_l}(k_l) \prod_{\substack{j=1 \\ j \neq l}}^{\eta} \phi_{i_j}(k_j) + \right. \right. \\ & \left. \left. - ck_l\phi_{i_l}(k_l - 1) \prod_{\substack{j=1 \\ j \neq l}}^{\eta} \phi_{i_j}(k_j) \right\} \right) \end{aligned} \quad (\text{B.12})$$

Agora, usando a relação em (B.9), a equação anterior fica:

$$\begin{aligned} \varphi_l = & \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} x(k_1, \dots, k_l, \dots, k_\eta) \cdot \\ & \cdot \left(\sum_{i_1=1}^{\infty} \cdots \sum_{i_l=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_l, \dots, i_\eta} i_l (1 - c^2) \phi_{i_l}(k_l) \prod_{\substack{j=1 \\ j \neq l}}^{\eta} \phi_{i_j}(k_j) \right) \\ = & (1 - c^2) \sum_{i_1=1}^{\infty} \cdots \sum_{i_l=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} i_l \alpha_{i_1, \dots, i_l, \dots, i_\eta} \cdot \\ & \cdot \underbrace{\left(\sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} x(k_1, \dots, k_l, \dots, k_\eta) \prod_{j=1}^{\eta} \phi_{i_j}(k_j) \right)}_{\alpha_{i_1, \dots, i_l, \dots, i_\eta}} \\ = & (1 - c^2) \sum_{i_1=1}^{\infty} \cdots \sum_{i_l=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} i_l \alpha_{i_1, \dots, i_l, \dots, i_\eta}^2 \end{aligned} \quad (\text{B.13})$$

Portanto das equações (B.10) e (B.13) tem-se para qualquer l ($l = 1, 2, \dots, \eta$)

que:

$$\begin{aligned}
& -2c \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x(k_1, \dots, k_l - 1, \dots, k_\eta) x(k_1, \dots, k_l, \dots, k_\eta) + \\
& + (1 + c^2) \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x^2(k_1, \dots, k_l, \dots, k_\eta) + \\
& + \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} x^2(k_1, \dots, k_l, \dots, k_\eta) = \tag{B.14} \\
& = (1 - c^2) \sum_{i_1=1}^{\infty} \cdots \sum_{i_l=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} i_l \alpha_{i_1, \dots, i_l, \dots, i_\eta}^2
\end{aligned}$$

Onde fez-se uso da seguinte igualdade:

$$\begin{aligned}
& \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} (k_l + 1) x(k_1, \dots, k_l + 1, \dots, k_\eta) x(k_1, \dots, k_l, \dots, k_\eta) = \\
& = \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x(k_1, \dots, k_l - 1, \dots, k_\eta) x(k_1, \dots, k_l, \dots, k_\eta) \tag{B.15}
\end{aligned}$$

já que a função $x(k_1, \dots, k_l, \dots, k_\eta)$ é, por hipótese, nula para qualquer $k_l < 0$.

Voltando então à equação (B.14), pode-se escrever ainda:

$$\begin{aligned}
& -2c \sum_{l=1}^{\eta} \left(\sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x(k_1, \dots, k_l - 1, \dots, k_\eta) x(k_1, \dots, k_l, \dots, k_\eta) \right) + \\
& + (1 + c^2) \sum_{l=1}^{\eta} \left(\sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x^2(k_1, \dots, k_l, \dots, k_\eta) \right) + \\
& + \sum_{l=1}^{\eta} \left(\sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} x^2(k_1, \dots, k_l, \dots, k_\eta) \right) = \tag{B.16} \\
& = (1 - c^2) \sum_{l=1}^{\eta} \left(\sum_{i_1=1}^{\infty} \cdots \sum_{i_l=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} i_l \alpha_{i_1, \dots, i_l, \dots, i_\eta}^2 \right)
\end{aligned}$$

Usando as expressões para os momentos definidos na seção 4.5, a equação anterior

pode ser escrita como:

$$-2c\mu_1(x) + (1 + c^2)\mu_2(x) + \eta\mu_3(x) = (1 - c^2) \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (i_1 + \cdots + i_\eta) \alpha_{i_1, \dots, i_\eta}^2 \quad (\text{B.17})$$

E finalmente chega-se a:

$$\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (i_1 + \cdots + i_\eta) \alpha_{i_1, \dots, i_\eta}^2 = \frac{-2c\mu_1(x) + (1 + c^2)\mu_2(x) + \eta\mu_3(x)}{1 - c^2}$$

■

B.3 Demonstração da Equação (4.62)

Na seção 4.5, definiu-se o seguinte problema de otimização, dado pela equação (4.61):

$$\min_{|c| < 1} L(c) = \frac{2(m_2c^2 - 2m_1c + m_3)}{\eta(M + 1)(1 - c^2)} \quad (\text{B.18})$$

Para resolver esse problema, primeiramente calcula-se a derivada da função de custo $L(c)$ em relação à variável c . Tem-se que:

$$\frac{\partial L(c)}{\partial c} = \frac{2[(2m_2c - 2m_1)(1 - c^2) - (-2c)(m_2c^2 - 2m_1c + m_3)]}{\eta(M + 1)(1 - c^2)^2} \quad (\text{B.19})$$

Então igualando a equação (B.19) a zero, chega-se à seguinte equação algébrica do segundo grau em c :

$$m_1c^2 - (m_2 + m_3)c + m_1 = 0$$

Fazendo $\xi = (m_2 + m_3)/(2m_1)$, a equação anterior possui as soluções:

$$c_{1,2} = \xi \pm \sqrt{\xi^2 - 1} \quad (\text{B.20})$$

Agora, note que:

$$\sum_{l=1}^{\eta} \left(\sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l \left[x(k_1, \dots, k_l, \dots, k_\eta) \pm x(k_1, \dots, k_l - 1, \dots, k_\eta) \right]^2 \right) \geq 0 \quad (\text{B.21})$$

Na seção 4.3, foi suposto que o *kernel* $h_\eta(k_1, \dots, k_\eta)$ é absolutamente somável em $[0, \infty)$. Essa consideração é feita para que seja possível expandi-lo em uma base

de funções ortonormais, como na expansão (4.10). A condição de ser absolutamente somável implica que o *kernel* é estável, e seu valor pode ser desprezado para um determinado $T > k_l$ ($l = 1, 2, \dots, \eta$). Se os coeficientes do *kernel* $h_\eta(k_1, \dots, k_\eta)$ são desprezados para $T > k_l$, o mesmo poderá ocorrer com as funções $g_{\text{par}}(k_1, \dots, k_\eta)$ e $g_{\text{impar}}(k_1, \dots, k_\eta)$, respectivamente nas equações (4.36) e (4.37). Conseqüentemente, as funções g_{par} e g_{impar} serão estáveis (absolutamente somáveis). Isso exclui a condição de igualdade na equação (B.21), que somente seria obtida se g_{par} e g_{impar} (e conseqüentemente h_η) fossem instáveis ou nulos.

Considerando então a desigualdade (B.21) como sendo estrita, pode-se inferir o seguinte:

$$\begin{aligned} & \sum_{l=1}^{\eta} \left(\sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x^2(k_1, \dots, k_l, \dots, k_\eta) + \right. \\ & \pm 2 \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x(k_1, \dots, k_l, \dots, k_\eta) x(k_1, \dots, k_l - 1, \dots, k_\eta) + \\ & \left. + \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x^2(k_1, \dots, k_l - 1, \dots, k_\eta) \right) > 0 \end{aligned} \quad (\text{B.22})$$

E a partir de (B.22), pode-se escrever também:

$$\begin{aligned} & \sum_{l=1}^{\eta} \left(\sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x^2(k_1, \dots, k_l, \dots, k_\eta) + \right. \\ & \pm 2 \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x(k_1, \dots, k_l, \dots, k_\eta) x(k_1, \dots, k_l - 1, \dots, k_\eta) + \\ & \left. + \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x^2(k_1, \dots, k_l, \dots, k_\eta) + \right. \\ & \left. - \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=1}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} (k_l - 1) x^2(k_1, \dots, k_l - 1, \dots, k_\eta) + \right. \\ & \left. + \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=1}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_l x^2(k_1, \dots, k_l - 1, \dots, k_\eta) \right) > 0 \end{aligned} \quad (\text{B.23})$$

ou ainda:

$$\begin{aligned} \sum_{l=1}^{\eta} \left(2 \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_{\eta}=0}^{\infty} k_l x^2(k_1, \dots, k_l, \dots, k_{\eta}) + \right. \\ \left. \pm 2 \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_{\eta}=0}^{\infty} k_l x(k_1, \dots, k_l, \dots, k_{\eta}) x(k_1, \dots, k_l - 1, \dots, k_{\eta}) + \right. \\ \left. + \sum_{k_1=0}^{\infty} \cdots \sum_{k_l=0}^{\infty} \cdots \sum_{k_{\eta}=0}^{\infty} x^2(k_1, \dots, k_l, \dots, k_{\eta}) \right) > 0 \end{aligned} \quad (\text{B.24})$$

Usando as expressões para os momentos definidos na seção 4.5, a equação acima pode ser escrita como segue:

$$2\mu_2(x) \pm 2\mu_1(x) + \eta\mu_3(x) > 0$$

Fazendo primeiramente $x = g_{\text{par}}$ e depois $x = g_{\text{impar}}$, a equação anterior resulta nas duas desigualdades seguintes:

$$2\mu_2(g_{\text{par}}) \pm 2\mu_1(g_{\text{par}}) + \eta\mu_3(g_{\text{par}}) > 0 \quad (\text{B.25})$$

$$2\mu_2(g_{\text{impar}}) \pm 2\mu_1(g_{\text{impar}}) + \eta\mu_3(g_{\text{impar}}) > 0 \quad (\text{B.26})$$

Somando as equações (B.25) e (B.26), tem-se que:

$$2\mu_2(g_{\text{par}}) + 2\mu_2(g_{\text{impar}}) + \eta\mu_3(g_{\text{par}}) + \eta\mu_3(g_{\text{impar}}) > \pm [2\mu_1(g_{\text{par}}) + 2\mu_1(g_{\text{impar}})]$$

que permite concluir o seguinte:

$$\left| \frac{2\mu_2(g_{\text{par}}) + 2\mu_2(g_{\text{impar}}) + \eta\mu_3(g_{\text{par}})\eta\mu_3(g_{\text{impar}})}{2\mu_1(g_{\text{par}}) + 2\mu_1(g_{\text{impar}})} \right| > 1 \quad (\text{B.27})$$

Usando as relações em (4.58), (4.59) e (4.60), chega-se finalmente a:

$$\left| \frac{m_2 + m_3}{2m_1} \right| = |\xi| > 1 \quad (\text{B.28})$$

Para $\xi > 1$, tem-se:

$$\begin{aligned} \xi &> 1 \\ -2\xi &< -2 \\ \xi^2 + 1 - 2\xi &< \xi^2 + 1 - 2 \\ (\xi - 1)^2 &< \xi^2 - 1 \\ |\xi - 1| &< \sqrt{\xi^2 - 1} \end{aligned} \quad (\text{B.29})$$

Como $\xi > 1$ implica que $|\xi - 1| = \xi - 1$, pode-se escrever ainda:

$$\begin{aligned} |\xi - 1| &< \sqrt{\xi^2 - 1} \\ \xi - 1 &< \sqrt{\xi^2 - 1} \\ \xi - \sqrt{\xi^2 - 1} &< 1 \\ c_1 &< 1 \end{aligned} \tag{B.30}$$

Para $\xi < -1$, tem-se:

$$\begin{aligned} \xi &< -1 \\ 2\xi &< -2 \\ \xi^2 + 1 + 2\xi &< \xi^2 + 1 - 2 \\ (\xi + 1)^2 &< \xi^2 - 1 \\ |\xi + 1| &< \sqrt{\xi^2 - 1} \end{aligned} \tag{B.31}$$

Como $\xi < -1$ implica que $|\xi + 1| = -(\xi + 1)$, pode-se escrever ainda:

$$\begin{aligned} |\xi + 1| &< \sqrt{\xi^2 - 1} \\ -\xi - 1 &< \sqrt{\xi^2 - 1} \\ \xi + \sqrt{\xi^2 - 1} &> -1 \\ c_2 &> -1 \end{aligned} \tag{B.32}$$

Além disso, pode-se escrever também $\xi^2 > \xi^2 - 1$, que implica em $|\xi| > \sqrt{\xi^2 - 1}$. Então, tem-se que:

se $\xi > 1$:

$$\begin{aligned} |\xi| &> \sqrt{\xi^2 - 1} \\ \xi &> \sqrt{\xi^2 - 1} \\ \xi - \sqrt{\xi^2 - 1} &> 0 \\ c_1 &> 0 \end{aligned} \tag{B.33}$$

se $\xi < -1$:

$$\begin{aligned} |\xi| &> \sqrt{\xi^2 - 1} \\ -\xi &> \sqrt{\xi^2 - 1} \\ \xi + \sqrt{\xi^2 - 1} &< 0 \\ c_2 &< 0 \end{aligned} \tag{B.34}$$

Finalmente, agrupando (B.30) com (B.33), e (B.32) com (B.34), conclui-se que $|c_{1,2}| < 1$. Ou seja, ambas as equações dadas na equação (B.20) podem ser soluções do problema (4.61). Portanto:

$$c_{\text{otimo}} = \begin{cases} \xi - \sqrt{\xi^2 - 1} & \text{se } \xi > 1 \\ \xi + \sqrt{\xi^2 - 1} & \text{se } \xi < -1 \end{cases}$$

■

Para verificar se a equação (4.62) é realmente solução do problema (4.61), faz-se necessário o uso do seguinte lema [2]:

Lema B.3.1. *Seja uma função $f : \mathcal{X} \rightarrow \mathbb{R}$ definida como:*

$$f(x) = \frac{\varpi(x)}{v(x)} \tag{B.35}$$

onde \mathcal{X} é um conjunto convexo aberto, e:

- ϖ é uma função convexa, diferenciável e não-negativa para todo $x \in \mathcal{X}$;
- v é uma função côncava, diferenciável e positiva para todo $x \in \mathcal{X}$.

Tem-se que f é uma função pseudo-convexa em \mathcal{X} .

A partir das equações (4.51) e (4.59), pode-se concluir que $m_2 > 0$; conseqüentemente, a função $\varpi(c) \triangleq 2(m_2c^2 - 2m_1c + m_3)$ é convexa. Além disso, $\varpi(c)$ é não-negativa e diferenciável para todo $c \in (-1, 1)$. Também, a função $v(c) \triangleq 1 - c^2$ é côncava, diferenciável e positiva para qualquer $c \in (-1, 1)$. Aplicando-se o resultado do lema anterior, conclui-se que a função $L(c)$ em (4.61) é pseudo-convexa para $|c| < 1$. Com isso, a solução da equação $\partial L(c)/\partial c = 0$ é um mínimo global de $L(c)$. Conclui-se então que (4.62) é realmente solução do problema (4.61).

Referências Bibliográficas

- [1] Aguirre, L. A. **Introdução à Identificação de Sistemas – Técnicas Lineares e não-Lineares Aplicadas a Sistemas Reais**. Editora UFMG, 2004.
- [2] Bazaraa, M. S., Sherali, H. D., Shetty, C. M. **Nonlinear Programming: Theory and Algorithms**. John Wiley & Sons, 1993.
- [3] Boyd, S., Chua, L. O. **Fading Memory and the Problem of Approximation Nonlinear Operators with Volterra Series**. *IEEE Transactions on Circuits and Systems*, **32**(11):1150-1161, novembro/1985.
- [4] Brinker, A. C., Belt, H. J. W. **Optimal Free Parameters in Orthonormal Approximations**. *IEEE Transactions on Signal Processing*, **46**(8):2081-2087, agosto/1998.
- [5] Broome, P. W. **Discrete Orthonormal Sequences**. *Journal of Association Computing Machinery*, **12**(2):151-168, abril/1965.
- [6] Campello, R. J. G. B. **Arquiteturas e Metodologias para Modelagem e Controle de Sistemas Complexos Utilizando Ferramentas Clássicas e Modernas**. Tese de Doutorado, DCA/FEEC/UNICAMP. Campinas-SP, Brasil, 2002.
- [7] Campello, R. J. G. B., Amaral, W. C. **Desenvolvimento Ótimo de Modelos de Volterra em Funções Ortonormais de Laguerre**. Em: *XIV Congresso Brasileiro de Automática*. Natal-RN, Brasil, 2002.
- [8] Campello, R. J. G. B., Favier, G., Amaral, W. C. **Optimal Expansions of Discrete-Time Volterra Models Using Laguerre Functions**. Em: *Proc. 13th IFAC Symposium on System Identification*. Rotterdam, The Netherlands, 2003.
- [9] Campello, R. J. G. B., Favier, G., Amaral, W. C. **Optimal Expansions of Discrete-Time Volterra Models Using Laguerre Functions**. *Automatica*, **40**(5):815-822, maio/2004.
- [10] Chen, C. T. **Linear System Theory and Design**. Oxford University Press, 1998.

- [11] Da Rosa, A., Amaral, W. C., Campello, R. J. G. B. **Choice of Free Parameters in Expansions of Discrete-Time Volterra Models Using Kautz Functions**. Em: *16th IFAC World Congress*. Praha, Czech Republic, 2005 (aceito para publicação).
- [12] DeCarlo, R. A. **Linear Systems – A State Variable Approach with Numerical Implementation**. Prentice Hall, 1989.
- [13] Doyle III, F. J., Pearson, R. K. e Ogunnaike, B.A. **Identification and Control Using Volterra Models**. Springer, 2002.
- [14] Eykhoff, P. **System Identification: Parameter and State Estimation**. John Wiley & Sons, 1974.
- [15] Fu, Y. e Dumont, G. A. **An Optimum Time Scale for Discrete Laguerre Network**. *IEEE Transactions on Automatic Control*, **38**(6):934-938, junho/1993.
- [16] Gotlieb, M. J. **Concerning Some Polynomials Orthogonal on a Finite or Enumerable Set of Points**. *American Journal of Mathematics*, **60**(2):453-458, abril/1938.
- [17] Luenberger, D. G. **Introduction to Dynamic Systems: Theory, Models and Applications**. John Wiley & Sons, 1979.
- [18] Masnadi-Shirazi, M. A., Ahmed, N. **Optimum Laguerre Networks for a Class of Discrete-Time Systems**. *IEEE Transactions on Signal Processing*, **39**(9):2104-2108, setembro/1991.
- [19] Oliveira, G. H. C., Lopez, J. F., Favier, G., Amaral, W. C. **Review of Predictive Control Methods Using Laguerre Functions**. Em: *IEEE/IMACS/CESA*. Lille, France, 1996.
- [20] Oliveira, G. H. C. **Controle Preditivo para Processos com Incertezas Estruturadas Baseado em Série de Funções Ortonormais**. Tese de Doutorado, DCA/FEEC/UNICAMP. Campinas-SP, Brasil, 1997.
- [21] Oliveira, G. H. C., Amaral, W. C. **Identificação e Controle Preditivo de Processos Não-lineares Utilizando Séries de Volterra e Bases de Funções Ortonormais**. Em: *XIII Congresso Brasileiro de Automática*. Florianópolis-SC, Brasil, 2000.
- [22] Oliveira, G. H. C., Amaral, W. C. **CRHPC Using Volterra Models and Orthonormal Basis Functions: An Application to CSTR Plants**. Em: *IEEE Conference on Control Applications*. Istanbul, Turquia, 2003.
- [23] Palm III, William J. **Modeling, Analysis and Control of Dynamic Systems**. John Wiley & Sons, 1983.

- [24] Rugh, W. J. **Nonlinear System Theory – The Volterra/Wiener Approach**. The Johns Hopkins University Press, 1991.
- [25] Sarroukh, B. E., Eijndhoven, S. J. L., Brinker, A. C. **An Iterative Solution for the Optimal Poles in a Kautz Series**. Em: *International Conference on Acoustics, Speech and Signal Processing – ICASSP’2001*. Salt Lake City, Utah/USA, 2001.
- [26] Schetzen, M. **The Volterra and Wiener Theories of Nonlinear Systems**. Robert Krieger Publishing Company, 1989.
- [27] Silva, T. A. M. O. **Optimality Conditions for Laguerre Networks**. *IEEE Transactions on Signal Processing*, **42**(9):2528-2530, setembro/1994.
- [28] Silva, T. A. M. O. **Optimality Conditions for Truncated Kautz Networks with Two Periodically Repeating Complex Conjugate Poles**. *IEEE Transactions on Automatic Control*, **40**(2):342-346, fevereiro/1995.
- [29] Silva, T. A. M. O. **On the Determination of the Optimal Pole Position of Laguerre Filters**. *IEEE Transactions on Signal Processing*, **43**(9):2079-2087, setembro/1995.
- [30] Söderström, T., Stoica, P. M. **System Identification**. Prentice Hall, 1989.
- [31] Tanguy, N., Morvan, R., Vilbé, P., Calvez, L. C. **Optimum Choice of Free Parameter in Orthonormal Approximations**. *IEEE Transactions on Automatic Control*, **40**(10):1811-1813, outubro/1995.
- [32] Tanguy, N., Morvan, R., Vilbé, P., Calvez, L. C. **Online Optimization of the Time Scale in Adaptive Laguerre-Based Filters**. *IEEE Transactions on Signal Processing*, **48**(4):1184-1187, abril/2000.
- [33] Tanguy, N., Morvan, R., Vilbé, P., Calvez, L. C. **Pertinent Parameters for Kautz Approximation**. *Electronic Letters*, **36**(8):769-771, abril/2000.
- [34] Tanguy, N., Morvan, R., Vilbé, P., Calvez, L. C. **Pertinent Choice of Parameters for Discrete Kautz Approximation**. *IEEE Transactions on Automatic Control*, **47**(5):783-787, maio/2002.
- [35] Van den Hof, P. M. J., Heuberger, P. S. C., Bokor, J. **System Identification with Generalized Orthonormal Basis Functions**. *Automatica*, **31**(12):1821-1834, dezembro/1995.
- [36] Wahlberg, B. **System Identification Using Kautz Models**. *IEEE Transactions on Automatic Control*, **39**(6):1276-1282, junho/1994.