

Otimização de Modelos BFO com Funções de Laguerre

Jeremias Barbosa Machado , Ricardo J. G. B. Campello , Wagner Caradori do Amaral(Orientador)

Departamento de Engenharia de Computação e Automação Industrial (DCA)
Faculdade de Engenharia Elétrica e de Computação (FEEC)
Universidade Estadual de Campinas (Unicamp)
Caixa Postal 6101, 13083-970 – Campinas, SP, Brasil

{jeremias,wagner}@dca.fee.unicamp.br, campello@icmc.usp.br

Abstract – This paper presents a new approach to optimization of linear dynamic systems models using orthonormal basis functions with Laguerre functions. This approach propose a method to select the poles of the proposed model. This selection is made by optimizing of the model by nonlinear algorithms of optimization, which requires to calculate the gradients of the orthonormal functions outputs with respect to its parameters. These gradients are calculated analytically and provide the exact direction of search in the optimization of functions parameters. In this context, this paper presents the approaches used for the identification and optimization dynamical system models via orthonormal basis functions. A simulation example illustrates the techniques of identification and modeling proposals.

Keywords – Dynamic linear Systems, Orthonormal Basis Functions, Identification, Modeling

1. Introdução

Modelos de sistemas dinâmicos que utilizam base de funções ortonormais - BFO apresentam vantagens em relação a outras abordagens de modelagem. Estes modelos possibilitam a incorporação de conhecimento prévio sobre a dinâmica do sistema [11, 15] o que reduz a ordem da representação do modelo, simplificando os problemas de identificação e controle associados [19, 10]. Outra característica da utilização de modelos BFO em sistemas dinâmicos é a ausência de realimentação da saída, isto é, não existe realimentação de erros de previsão, o que normalmente resulta em modelos mais precisos [14]. Devido à completude da base também é possível aumentar a sua capacidade de representação incrementando o número de funções ortonormais da base [10]. Os modelos BFO são parametrizados por polos e a escolha adequada destes leva a modelos com um número reduzido de termos. A escolha do(s) polo(s) pode ser feita através de informações a priori das características dinâmicas do sistema ou através de métodos de seleção do(s) polo(s). Em trabalhos como [18, 4, 3, 7] apresenta-se métodos analíticos para a seleção de polos para funções de Laguerre e Kautz, contudo, tais soluções são sub-ótimas pois em geral otimiza-se somente um subconjunto dos parâmetros da base. Alguns trabalhos como [1, 17] propõem procedimentos para seleção de polos de modelos BFO através de métodos numéricos. Em [9, 8] é proposta a otimização dos parâmetros de modelos com BFO (Kautz e Funções de Takenaka-Malmquist) através de métodos de otimização não-linear, empregando uma direção de busca dada pelo cálculo analítico dos gradientes das saídas das funções ortonormais.

Neste trabalho propõem-se um método para o cálculo analítico dos gradientes das funções de Laguerre utilizando a representação em espaço de estados. Tais gradientes, associados a métodos de otimização não-linear permitem a obtenção de modelos mais precisos e que não exijam um conhecimento a priori do sistema a ser modelado. A seguir, na Seção 2., faz-se uma revisão de modelagem com BFO's. Na Seção 3. descreve-se o

processo de otimização e o cálculo dos gradientes para os modelos BFO com funções de Laguerre é apresentado na Seção 3.. Na Seção 4. apresenta-se um exemplo que ilustra o desempenho das metodologias propostas. Finalmente, na Seção 5., apresentam-se as conclusões.

2. Modelos BFO

Modelos BFO podem ser genericamente representados em modelos de espaço de estados como [14, 5]:

$$\begin{aligned}\Psi(k+1) &= A_f \Psi(k) + B_f u(k) \\ \hat{y}(k) &= C_f^T \Psi(k)\end{aligned}\quad (1)$$

onde $\Psi(k) = [\psi_1(k) \dots \psi_n(k)]^T$ é o vetor de estados ortonormais (denominação usada por conveniência já que a ortonormalidade é uma propriedade da função e não dos estados) de ordem n , sendo estes estados coincidentes com as saídas dos filtros correspondentes às funções ortonormais de ordem equivalente, $u(k)$ a entrada do sistema e $\hat{y}(k)$ é a saída estimada. Este modelo pode ainda incluir uma parcela na parte estática da representação que modela o nível DC da saída do sistema sob estudo.

Duas funções utilizadas com muita frequência na representação através de BFO's são as funções de Laguerre e de Kautz. A base de Laguerre é apropriada para representar sistemas com polos puramente reais ou com parte imaginária de valor reduzido. A base de Kautz é mais apropriada para representar sistemas com dinâmica oscilatória subamortecida por ser parametrizada por polos complexos conjugados [10, 5]. A base de funções ortonormais de Laguerre é caracterizada pela utilização de funções de transferência com apenas um polo real [10]:

$$\Phi_i(z) = \frac{\sqrt{1-p^2}}{z-p} \left(\frac{1-pz}{z-p} \right)^{i-1}, \quad i = 1, 2, \dots \quad (2)$$

onde $p = \{p : p \in \mathbb{R} \text{ e } |p| < 1\}$ é o polo que parametriza as funções ortonormais e i a ordem da função dada. As funções de Laguerre podem ser caracterizadas como modelos FIR com polos fora da origem, sendo o modelo FIR

com polos na origem um caso especial das BFO's [10, 5]. O modelo FIR normalmente requer um número elevado de termos para representar sistemas com dinâmica dominante lenta. Como pode-se notar através da equação (2) os modelos de ordem i podem ser escritos como modelos em cascata de ordem $i - 1$. A representação em cascata da função de Laguerre lhe confere a propriedade de recursividade já que a i -ésima função pode ser escrita em função da $(i-1)$ -ésima. Sendo assim, é possível representar a dinâmica de funções de Laguerre através de uma equação de estados, como apresentado na equação (1). A descrição detalhada do modelo em espaço de estados pode ser encontrada em [13]. Neste trabalho propõem-se uma metodologia para a busca do valor do polo p de Laguerre e dos parâmetros da expansão do modelo.

3. Ajuste dos parâmetros das BFO's

Os parâmetros a serem determinados na modelagem de sistemas dinâmicos lineares por base de funções ortonormais são os polos das funções ortonormais e os coeficientes da expansão da série de funções. Conforme discutido em [16, 2, 5, 13], uma vez selecionados os polos, os termos da expansão da série podem ser determinados através do já bem conhecido método de mínimos quadrados.

A seleção do(s) polo(s) da base de funções ortonormais que representará a dinâmica do sistema normalmente depende de um conhecimento prévio do sistema ou de métodos de otimização que geralmente envolvem algoritmos de busca que minimizem o erro entre a saída do sistema e a saída do modelo estimado como apresentado em [14, 10, 6]. Em [9, 6] é apresentada uma metodologia de otimização dos parâmetros de modelos de Kautz e GOBF com funções de Takenaka-Malmquist através de buscas exatas dadas pelos cálculos dos gradientes da saída do modelo estimado com relação aos parâmetros do modelo. Nesta seção otimiza-se, através de métodos de otimização não-linear [12] em especial empregando o método de Levenberg-Marquardt [14], os parâmetros de um modelo linear BFO com funções de Laguerre apresentando de maneira detalhada as equações para o cálculo dos gradientes necessário ao processo de otimização.

3.1. Estimação e Otimização dos Parâmetros dos modelos BFO's

A estratégia adotada consiste da otimização de ambos o polo p e os termos da expansão da série na matriz C minimizando uma função de custo com relação ao erro quadrático entre a saída do sistema e a saída do modelo:

$$\min_p J = \frac{1}{2} \sum_{k=0}^L (\hat{y}(k) - y(k))^2 \quad (3)$$

com $\hat{y}(k)$ sendo a saída estimada.

Para se resolver o problema de minimização é necessário calcular o gradiente da função de custo com relação aos parâmetros que se deseja otimizar. Para o cálculo dos gradientes através dos dados coletados do sistema será empregada a técnica de *back-propagation-through-time* [14]. Esta técnica decompõe a dinâmica do sistema sob estudo em uma sequência de representações estáticas, permitindo descrever as derivadas da saída do modelo somente em termos das condições iniciais e dos sinais de entrada, através de uma recursão de k passos no tempo.

Analisando-se individualmente os termos que compõem a equação (3), o gradiente da função de custo com relação aos parâmetros em C é dado por:

$$\begin{aligned} \nabla_C J &= \sum_{k=0}^L (y(k) - \hat{y}(k)) \nabla_C \hat{y}(k) \\ &= \sum_{k=0}^L (y(k) - \hat{y}(k)) \Psi(k) \end{aligned}$$

Já o gradiente de J com relação ao parâmetro p é definido por:

$$\begin{aligned} \frac{\partial J}{\partial p} &= \sum_{k=0}^L (y(k) - \hat{y}(k)) \sum_{i=1}^n g_i \frac{\partial \psi_i(k)}{\partial p} \\ &= \sum_{k=0}^L (y(k) - \hat{y}(k)) C^T \frac{\partial \Psi_i(k)}{\partial p} \end{aligned}$$

Neste trabalho utilizar-se-á o cálculo analítico destes gradientes da função de custo com relação ao(s) parâmetro(s) p e C , fornecendo assim o valor exato do gradiente para a otimização do(s) polo(s) e dos parâmetros da expansão das funções ortonormais através de métodos de otimização não-linear. A seguir é apresentada a solução para o cálculo dos gradientes citados para os modelos BFO com funções de Laguerre.

3.1.1. Gradiente para funções de Laguerre

Para o cálculo do gradiente utilizar-se-á a representação em espaço de estados do modelo de Laguerre como descrito na eq (1). Contudo, uma modificação é necessária já que se deseja conhecer as saídas dos filtros, ou seja, os estados nos instantes k para derivá-las com relação ao polo. Sendo assim, o modelo em espaço de estados para as funções de Laguerre será escrito como:

$$\Psi(k+1) = A_f \Psi(k) + B_f u(k) \quad (4)$$

onde A_f e B_f são as mesmas definidas na seção 2. para as funções de Laguerre. Desenvolvendo-se o modelo (4) é possível verificar que a solução para $\Psi(k)$ será dada por:

$$\Psi(k) = A_f^k \Psi(0) + \sum_{i=0}^{k-1} A_f^i B_f u(k-1-i) \quad (5)$$

O gradiente de $\Psi(k)$ com relação ao polo p pode então ser calculado através da derivada da equação (5) com relação ao polo, como apresentado na equação (6):

$$\frac{\partial \Psi(k)}{\partial p} = \frac{\partial A_f^k}{\partial p} \Psi(0) + \sum_{i=0}^{k-1} \left(\frac{\partial A_f^i}{\partial p} B_f + A_f^i \frac{\partial B_f}{\partial p} \right) u(k-1-i) \quad (6)$$

sendo $\frac{\partial A_f^k}{\partial p}$ dado por:

$$\frac{\partial A_f^k}{\partial p} = \sum_{j=1}^k A_f^{j-1} \frac{\partial A_f}{\partial p} A_f^{k-j} \quad (7)$$

Das equações (6) e (7) torna-se necessário calcular as derivadas de A_f e B_f com relação ao polo p . O valor destas derivadas é dado por:

$$\frac{\partial A_f}{\partial p} = \begin{bmatrix} dap_{11} & 0 & 0 & \cdots & 0 \\ dap_{21} & dap_{22} & 0 & \cdots & 0 \\ dap_{31} & dap_{32} & dap_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ dap_{n1} & dap_{n2} & dap_{n3} & \cdots & dap_{nn} \end{bmatrix}$$

onde i é igual à linha e j é igual à coluna do elemento dap_{ij} , sendo este dado por:

- No casos em que $j > i$:
 $\cdot dap_{ij} = 0$;
- Com $i = j$:
 $\cdot dap_{ij} = 1$;
- Com $j < i$:
 $\cdot dap_{ij} = (-1)(i-j-1)(-p)^{(i-j-2)} * \dots$
 $\dots * (1-p^2) + (-p)^{(i-j-1)}(-2p)$, tal que $i \leq n$
 $(i-j-1) \geq 0$;

Já para a matriz B_f , tem-se que a sua derivada com relação ao polo p é dada por:

$$\frac{\partial B_f}{\partial p} = [dbp_{11} \quad dbp_{21} \quad dbp_{31} \quad \cdots \quad dbp_{n1}]^T$$

com seus elementos sendo:

$$dbp_{i1} = \frac{-p}{\sqrt{1-p^2}} (-p)^{(i-1)} + \frac{1-p^2}{-p} (i-1)(-1)(-p)^{(i-2)}$$

onde i é a linha do elemento dbp .

Desta maneira calcula-se analiticamente o gradiente das saídas das funções de Laguerre com relação ao parâmetro p utilizando-se a técnica de retropropagação através do tempo, como apresentada na seção 3.1.

4. Resultados Experimentais

Esta seção exemplifica a aplicação do modelo e a implementação do processo de otimização propostos neste trabalho. Para ilustrar a aplicação de modelos de Laguerre,

considere o sistema linear estável de segunda ordem com polo real de multiplicidade 2.

$$F(z) = \frac{0,08037}{z^2 - 1,433z + 0,5134}, \quad \text{com } t_s = 0,1 \text{ seg} \quad (8)$$

cujos polos estão em $z = 0,7165$.

Este sistema tem como resposta ao degrau uma saída super-amortecida, o que, neste caso justifica a abordagem de Laguerre resultando em modelos com número reduzido de funções. Partindo-se do pressuposto que em sistemas reais não se conhece a ordem do sistema sob análise, os testes para se obter o modelo foram iniciados com uma função de Laguerre. O sinal de entrada $u(k)$ é aleatório com distribuição uniforme entre -1 e 1. A tabela 1 apresenta os resultados comparativos entre os modelos otimizados.

Analisando-se os resultados apresentados na tabela 1 verifica-se que os modelos com somente um polo não foram capazes de modelar o sistema sob estudo. Os modelos com duas funções modelaram perfeitamente com o polo das funções convergindo para um polo próximo ao polo exato do sistema. Já com três funções o polo ficou um pouco mais distante e o modelo com uma precisão semelhante. A figura 1 mostra o resultado do sistema estimado antes e depois do processo de otimização com duas funções de Laguerre.

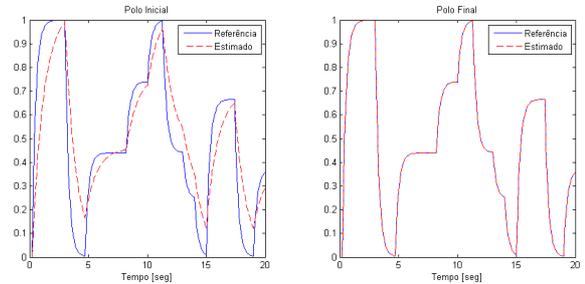


Figura 1. Comparação entre os polos inicial e final

5. Conclusões

Este artigo apresentou uma nova proposta para a otimização de modelos de sistemas dinâmicos através de BFO com funções de Laguerre na forma de espaço de estados. Foram apresentadas expressões analíticas para os gradientes das funções de Laguerre com relação aos parâmetros do modelo. Utilizando-se estas expressões é possível se obter direções de busca exatas que são utilizadas na otimização dos modelos. A metodologia descrita apresenta vantagens sobre outros modelos pois o cálculo dos gradientes é feito em batelada a partir de sinais de entrada e saída do sistema sob análise e o sistema é identificado sem conhecimento a priori do mesmo. Resultados de simulação ilustram a eficácia da metodologia proposta permitindo-se obter modelos BFO com uma excelente

Tabela 1. Resultados dos modelos de Laguerre para função de 2ª ordem.

| N. de Funções | Iter. | Polo Inicial | Polo Final | Avaliação Função de Custo |
|---------------|-------|--------------|------------|---------------------------|
| 1 | 8 | 0,1000 | 0,8664 | 0,5808 |
| 1 | 3 | 0,9000 | 0,7360 | 0,5808 |
| 2 | 7 | 0,1000 | 0,7160 | $3,1111.10^{-15}$ |
| 2 | 5 | 0,5000 | 0,7160 | $9,3224.10^{-18}$ |
| 3 | 10 | 0,5000 | 0,7124 | $6,8997.10^{-8}$ |

aproximação com polos convergindo para os polos reais do sistema modelado. Extensões deste trabalho estão sob investigação e estão sendo aplicadas na modelagem de sistemas não-lineares e a outras funções ortonormais.

Referências

- [1] A.C. Den Brinker B.E. Sarroukh, S.J.L. Van Eijndhoven. An iterative solution for the optimal poles in a kautz series. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 6:3949–3952, 2001.
- [2] R. J. G. B. Campello. *Arquiteturas e Metodologias para Modelagem e Controle de Sistemas Complexos Utilizando Ferramentas Clássicas e Modernas*. Tese de Doutorado, DCA/FEEC/UNICAMP, 2002.
- [3] R. J. G. B. Campello, W. C. Amaral, and G. Favier. A note on the optimal expansion of Volterra models using Laguerre functions. *Automatica*, 42:689–693, 2006.
- [4] R. J. G. B. Campello, G. Favier, and W. C. Amaral. Optimal expansions of discrete-time Volterra models using Laguerre functions. *Automatica*, 40:815–822, 2004.
- [5] R. J. G. B. Campello, G. H. C. Oliveira, and W. C. Amaral. Identificação e Controle de Processos via Desenvolvimentos em Séries Ortonormais: Partes A (Identificação) e B (Controle). *Controle & Automação*, 18(3):298–332, 2007.
- [6] A. da Rosa. *Identificação de Sistemas Não-Lineares Usando Modelos de Volterra Baseados em Funções Ortonormais de Kautz e Generalizadas*. Tese de Doutorado, DCA/FEEC/UNICAMP, 2009.
- [7] A. da Rosa, R. J. G. B. Campello, and W. C. Amaral. Choice of free parameters in expansions of discrete-time Volterra models using Kautz functions. *Automatica*, 43(6), 2007.
- [8] A. da Rosa, R. J. G. B. Campello, and W. C. Amaral. Exact Search Directions for Optimization of Linear and Nonlinear Models Based on Generalized Orthonormal Functions. *IEEE Transactions on Automatic Control*, 54(12):2757–2772, 2009.
- [9] A. da Rosa, R.J.G.B. Campello, and W.C. Amaral. Cálculo de direções de busca exatas para otimização de filtros de laguerre e de kautz. In *XVII Congresso Brasileiro de Automática*, volume 1, 2008.
- [10] P. S. C. Heuberger, P. M. J. Van den Hof, and B. Wahlberg. *Modelling and Identification with Rational Orthogonal Basis Functions*. Springer, 2005.
- [11] Van den Hof P. M. J. e Bosgra O. H. Heuberger, P. S. C. A generalized orthonormal basis for linear dynamical systems. *IEEE Transactions on Automatic Control*, 40(3):451–465, 1999.
- [12] David G. Luenberger. *Linear and Nonlinear Programming*. Springer, 2nd edition, 2003.
- [13] Jeremias Barbosa Machado. Modelagem e controle preditivo utilizando multimodelos. Master's thesis, Universidade Estadual de Campinas, Campinas, Brasil, Fevereiro 2007.
- [14] O. Nelles. *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Springer-Verlag, 2001.
- [15] F. Ninness, B. e Gustafsson. A unifying construction of orthonormal bases for system identification. *IEEE Transactions on Automatic Control*, 42(4):515–521, 1997.
- [16] G. H. C. Oliveira. *Controle Preditivo para Processos com incertezas estruturadas baseado em séries de funções ortonormais*. PhD thesis, Universidade Estadual de Campinas, Campinas, 1997.
- [17] Shah S.L. Partwardhan, S.C. From data to diagnosis and control using generalized orthonormal basis filters, Part I: Development of state observers. *Journal of Process Control*, 15(7):819–835, 2005.
- [18] N. Tanguy, R. Morvan, P. Vilbé, and L. C. Calvez. Pertinent choice of parameters for discrete Kautz approximation. *IEEE Trans. on Automatic Control*, 47:783–787, 2002.
- [19] Heuberger P. S. C. e Bokor J. Van den Hof, P. M. J. System identification with generalized orthonormal basis functions. *Automatica*, 31(12):1821–1834, 1995.