

DISTRIBUIÇÃO UNIFORME DAS SOLUÇÕES PARETO-ÓTIMO NA IDENTIFICAÇÃO MULTIOBJETIVO DE SISTEMAS

SAMIR ANGELO MILANI MARTINS*, ERIVELTON GERALDO NEPOMUCENO*

*GCOM - Grupo de Controle e Modelagem
Departamento de Engenharia Elétrica
Universidade Federal de São João del-Rei
Praça Frei Orlando, 170 - Centro
36307-654 - São João del-Rei - MG - Brasil

Abstract— The multiobjective System Identification searches for solutions which considers, simultaneously, the minimization of the prediction error and take into account other auxiliary information. In this sense, it is common the use of a weighted sum of the objectives, producing different solutions for different weights. A still not solved problem consists in how to generate a weights vector which produces significative solutions for the pareto set composition. In this work, the relationship between the derivative of the function composed by the weighted-sum of the objectives related to the weights is explored. The obtained results allows to keep the same number of solutions, distributed uniformly in the pareto set.

Keywords— NARMAX, Multiobjective System Identification, Optimization.

Resumo— A identificação multiobjetivo procura obter soluções que leve em conta simultaneamente a minimização do erro de predição e alguma outra informação auxiliar. Uma das técnicas utilizadas é a soma ponderada, que produz diferentes soluções para diferentes valores do peso. Um problema encontrada é como gerar um vetor de pesos que produzam soluções significativas para a composição do conjunto Pareto-Ótimo. Neste trabalho, explora-se a relação entre a derivada da função composta pela soma ponderada dos funcionais individuais em relação aos pesos envolvidos. O resultado obtido permite manter o mesmo número de soluções, mas distribuí-las de modo mais uniforme no pareto.

Palavras-chave— NARMAX, Identificação Multiobjetivo de Sistemas, Otimização.

1 Introdução

A Identificação de Sistemas é uma das atividades mais antigas e relevantes da ciência. É a área do conhecimento que estuda maneiras de modelar e analisar sistemas na tentativa de encontrar algum padrão em observações (Aguirre, 2007; Billings, 2013). Define-se modelo como o conjunto de hipóteses sobre a estrutura ou comportamento de um sistema físico. Um modelo é uma abstração de um sistema real expresso por meio de equações.

Na Engenharia, essas técnicas podem ser empregadas na identificação de vários sistemas, como na identificação de aquecedores elétricos (Martins et al., 2009), conversores CC-CC (Yang et al., 2009), sistemas químicos, e em outras áreas, como sistemas biológicos, econômicos e outros ramos da ciência.

A literatura está repleta de técnicas de Identificação de Sistemas (Zhang and Billings, 2015; Ayala et al., 2015; Bassi et al., 2015; Aguirre, 2014; Martins et al., 2013; Nepomuceno et al., 2007; Johansen, 1996). São utilizadas várias representações matemáticas e computacionais, entre elas, podem ser citadas Redes Neurais, Lógica Fuzzy, Modelos NARMAX (*Nonlinear Autoregressive Moving Average model with exogenous input*) polinomiais, racionais e Modelo Baseado em Indivíduos. A representação NARMAX polinomial possui como pontos positivos o fato de permitir, com relativa facilidade, a incorporação de informações que se tem *a priori* do sis-

tema no modelo. Tais informações (pontos fixos, curva e ganho estático, por exemplo) podem não estar contidas nos dados dinâmicos. Sendo assim, o acréscimo dessas informações pode acrescentar qualidade no modelo, no que se refere à representatividade do sistema.

A partir de trabalhos de (Johansen, 1996) e colaboradores, a área de identificação de sistemas começou a se preocupar com a possibilidade de utilizar informações auxiliares sobre o sistema e não apenas dados na identificação de sistemas. Dessa preocupação, surge uma nova técnica, chamada de Identificação Multiobjetivo de Sistemas (IMS), que faz uso de técnicas de otimização multiobjetivo aplicadas à identificação de sistemas.

Uma técnica escalar de otimização multiobjetivo é o método da soma ponderada. Nele, o problema original é transformado em um problema de otimização escalar, composto pela soma ponderada dos objetivos, sendo que as ponderações associadas a cada peso quantificam a importância de cada função objetivo naquele problema monoobjetivo. Nesse caso, a combinação de diferentes valores de pesos gera um conjunto diferente de soluções eficientes.

Em (Nepomuceno et al., 2007) foi desenvolvido um meio de codificar informações auxiliares de modo que as mesmas pudessem ser expressas em forma de erro de estimação. Tal abordagem permitiu a aplicação de técnicas escalares de otimização multiobjetivo (método da soma ponderada, por exemplo) na obtenção de parâmetros de

modelos NARMAX polinomiais.

Contudo, na prática, não se pode garantir facilmente uma boa diversidade de soluções do conjunto Pareto-ótimo, associada à distribuição das ponderações. Sendo assim, tem-se como foco o presente desenvolvimento que propõe uma metodologia para distribuição dos valores de ponderações, de modo a obter um conjunto de soluções Pareto-ótimas o mais uniformemente distribuídas possível.

2 Conceitos preliminares

2.1 Modelos NARMAX

Modelos NARMAX descrevem sistemas não-lineares por meio de equações de diferença, relacionando a saída atual em combinação das saídas e entradas passadas. O modelo NARMAX polinomial pode ser representado por:

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

em que $y(k)$ é a saída, $u(k)$ é a entrada exógena, e $e(k)$ é o sinal de ruído. $\Xi(k)$ representa o erro de predição. n_y , n_u , e n_e são as ordens da saída, da entrada exógena e do ruído modelado por processo de média móvel, respectivamente. Neste trabalho, F^ℓ é restrita a funções polinomiais não-lineares. $k = 1, \dots, N$, sendo $N \in \mathbb{Z}^+$ o número de passos discretos.

Podem ser destacadas grandes características deste tipo de representação matemática, tal como a baixa complexidade. Outra característica, explorada neste trabalho, é a relativa facilidade de se incorporar informações auxiliares no modelo, bem como a facilidade na obtenção de informações analíticas do sistema por meio do modelo.

2.2 Identificação Multiobjetivo de Sistemas

Pode-se compor uma função custo com a ponderação de todas as funções objetivo que devem ser minimizadas simultaneamente (Nepomuceno et al., 2007):

$$J_c(\hat{\theta}) = w_1 J_1 + w_2 J_2 + \dots + w_n J_n, \quad (2)$$

sendo $w_1 + w_2 + \dots + w_n = 1$, os pesos de cada função custo, $J_1, J_2, \dots, J_n$ são os valores das funções custos referentes a cada objetivo, e J_c o valor da função custo composta. Uma forma direta de se obter os parâmetros de um modelo NARMAX polinomial ($\hat{\theta}^*$) que minimize a função descrita na equação 2 é dada por (Nepomuceno et al., 2007):

$$\begin{aligned} \hat{\theta}^* &= \arg \min \sum_{i=1}^n w_i (v - G\hat{\theta})^T (v - G\hat{\theta}) \quad (3) \\ &= \left[\sum_{i=1}^n w_i G_i^T G_i \right]^{-1} \left[\sum_{i=1}^n w_i G_i^T v_i \right], \end{aligned}$$

sendo n o número de objetivos a serem levados em consideração. Variando os valores de $w_i, i = 1, 2, \dots, n$, obtém-se a curva Pareto-ótima (equação 3).

Uma vez já codificadas as informações auxiliares por meio de funções custo quadráticas, basta aplicar o método da soma ponderada, por exemplo, e obter o conjunto Pareto-ótimo. Contudo, a distribuição das soluções ao longo da fronteira Pareto não tem uma relação linear com a distribuição dos pesos associados à cada objetivo.

3 Metodologia

3.1 Análise do processo

Visando a identificação multiobjetivo de um modelo NARMAX polinomial, a função mono-objetivo, composta pela soma ponderada das funções custo estático e dinâmico, pode ser expressa por:

$$J_c = \omega_1 J_D + \omega_2 J_E, \quad (4)$$

sendo $\omega_1 + \omega_2 = 1$. Desse modo, pode-se reescrever a equação 4 da forma:

$$\begin{aligned} J_c &= \omega_1 J_D + (1 - \omega_1) J_E \\ J_c &= \omega_1 (J_D - J_E) + J_E. \end{aligned} \quad (5)$$

Conjecturou-se poder haver correlação entre o vetor dado pela distância euclidiana entre dois pontos consecutivos e a ponderação ω_1 . A correlação cruzada entre essas duas grandezas pode ser vista na Figura (1).

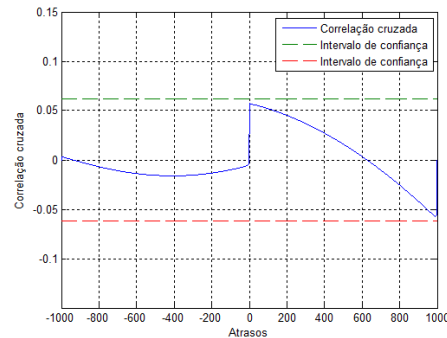


Figura 1: Correlação cruzada entre a distância de dois pontos consecutivos e o vetor de ponderações.

Para aplicar a metodologia aqui proposta, utilizou-se a massa de dados separada para identificação, totalizando 1000 modelos sendo o vetor de ponderações ω_1 formado por 1000 valores, igualmente espaçados entre zero e um. As linhas pontilhadas em vermelho e verde mostram os intervalos de confiança de 95%, ou seja, uma vez que a correlação cruzada entre a distância de duas soluções do conjunto Pareto-ótimo e o vetor ponderação ω_1 está compreendida nesse intervalo, pode-se garantir com uma confiança de 95% que não há correlação entre os sinais.

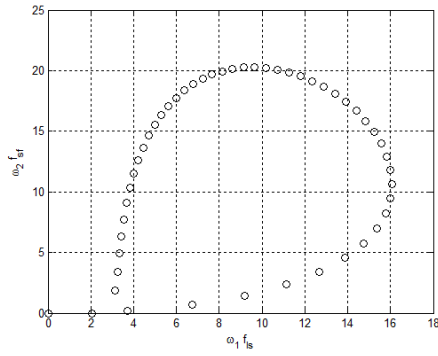


Figura 2: Parcelas da função mono-objetivo composta.

As Figuras (2) e (3) foram construídas visando encontrar algum padrão em observações, de modo a relacionar a variação do vetor peso com a diversidade no conjunto Pareto-ótimo.

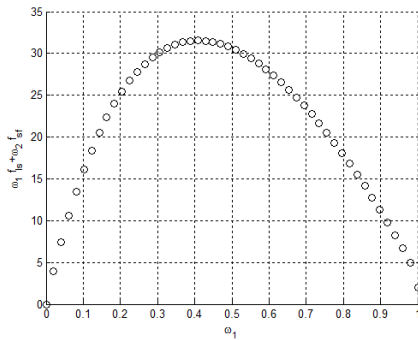


Figura 3: Função mono-objetivo composta.

Ao analisar a Figura (2), observa-se que a evolução do gráfico sugere uma maior quantidade de soluções presentes para valores intermediários de ω_1 . Para valores extremos de ω_1 ($\omega_1 \rightarrow 0$ ou $\omega_1 \rightarrow 1$), observa-se que há uma baixa densidade de soluções.

Ao analisar a Figura (3), evolução da função objetivo composta, tem-se que o módulo da derivada da função composta em função do peso ω_1 ($\frac{dJ_c}{d\omega_1}$) é elevado em seus extremos ($\omega_1 \rightarrow 0$ ou $\omega_1 \rightarrow 1$), onde há poucas soluções, e praticamente nula para valores intermediários. A derivada de J_c em função de ω_1 pode ser diretamente obtida a partir da equação (5) como sendo:

$$\frac{dJ_c}{d\omega_1} = J_D - J_E. \quad (6)$$

A partir dessas análises fica claro que o número de soluções em uma dada região Pareto-ótima está diretamente relacionado com o módulo da derivada da função composta, em função da ponderação. A Figura (4) apresenta o módulo da derivada exata da função composta J_c , calculada exatamente em cada ponto como a diferença entre os valores das funções individuais. Observa-se

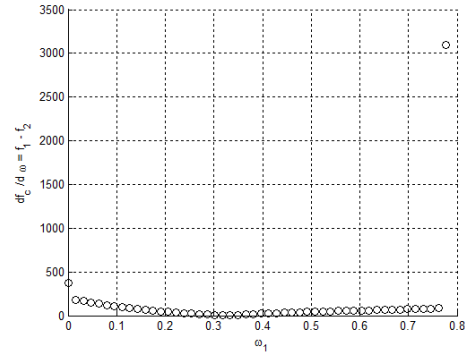


Figura 4: Derivada da função mono-objetivo composta.

um aumento significativo do módulo nos extremos ($\omega_1 \rightarrow 0$ ou $\omega_1 \rightarrow 1$), onde existe uma variação rápida de J_c .

3.2 O Método Proposto

A partir das análises realizadas anteriormente, pôde-se concluir que a distribuição de soluções no conjunto Pareto-ótimo, bem como sua diversidade estão diretamente relacionados com a derivada da função custo composta J_D .

A princípio, como todo procedimento de identificação de sistemas, deve-se separar uma massa de dados para identificação (obtenção do modelo) e outra para validação.

Com a massa de dados de identificação, deve-se, por meio da técnica apresentada, estimar uma fronteira Pareto-ótima com N pontos. A partir daí, calcula-se $(N - 1)$ pontos intermediários de derivadas:

$$\frac{df a_i}{d\omega_1} = \frac{\frac{df a_{i+1}}{d\omega_1} + \frac{df a_{i-1}}{d\omega_1}}{2}, \quad (7)$$

sendo a um ponto na fronteira Pareto-ótimo. O ponto $\frac{df a_i}{d\omega_1}$ é uma estimativa média da derivada, utilizada para representar a região de pesos entre $\omega_{i-1} \leq \omega_1 \leq \omega_{i+1}$. Assim, tem-se uma estimativa da derivada média, um valor quantitativo representativo da i -ésima região de ponderações.

A Figura (5) ilustra tal procedimento. Em preto, encontram-se os valores de derivada calculados para a massa de dados de identificação, ao passo que em vermelho se encontram os valores médios estimados, responsáveis por representar uma faixa de valores específica ($\omega_{i-1} \leq \omega_1 \leq \omega_{i+1}$).

Como já discutido anteriormente, há uma relação inversamente proporcional entre o módulo da derivada e o número de soluções eficientes naquela região. Quanto maior o número de pontos da fronteira Pareto para os dados de identificação, menor será o intervalo ($\omega_{i-1} \leq \omega_1 \leq \omega_{i+1}$) e consequentemente mais precisa será a estimativa de $\frac{df a_i}{d\omega_1}$.

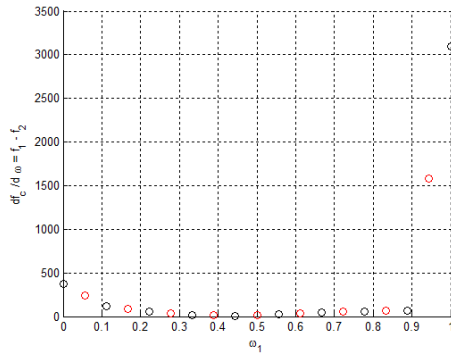


Figura 5: Estimativa de derivada na i-ésima região.

Sendo assim, é intuitivo considerar que em regiões com valores elevados de $\frac{df a_i}{d\omega_1}$ ($\omega_{i-1} \leq \omega_i \leq \omega_{i+1}$), deva haver um maior número de soluções associadas. Desse modo, define-se o número percentual de pontos da i-ésima região de pesos ω_i ($\omega_{i-1} \leq \omega_i \leq \omega_{i+1}$) como sendo:

$$NP(\%) = \text{round} \left(\frac{\frac{df a_i}{d\omega_1}}{\sum_{i=1}^{N-1} \frac{df a_i}{d\omega_1}} \right). \quad (8)$$

em que *round* é o operador que arredonda a fração entre parênteses para o número inteiro mais próximo.

Suponha que se deseje uma fronteira Pareto-ótima com 100 soluções. Considere que o i-ésimo ponto do vetor $NP(\%)$, correspondente à i-ésima região ($\omega_{i-1} \leq \omega_1 \leq \omega_{i+1}$) tenha valor 14. Isso sugere que 14% do número de pontos (14 pontos, nesse caso) deve ser distribuídos na região ($\omega_{i-1} \leq \omega_1 \leq \omega_{i+1}$), por meio de uma distribuição, aqui adotada linear, de ω_1 .

3.3 Estudo de caso

O sistema estudado foi um conversor CC-CC Buck com dados estáticos e dinâmicos exatamente iguais àqueles descritos em (Nepomuceno et al., 2007). Foram considerados dados de identificação estáticos aqueles obtidos quando ($0 \leq \bar{u} \leq 4$) e de validação àqueles obtidos quando ($4 \leq \bar{u} \leq 8$).

3.4 Comentários Acerca do Método Proposto

A aplicação da técnica se dá em problemas de otimização multiobjetivo aplicados à determinação de parâmetros em modelos NARMAX polinomiais. Para o caso, a composição de uma função custo pela combinação linear das funções objetivos individuais permite a existência, por construção, de $\frac{dJ_c}{d\omega_1}$, independente do número de objetivos considerados.

Para um problema com um maior número de objetivos, acredita-se ser direta a expansão da técnica. Assim como no problema bi-objetivo, há

uma menor densidade de soluções em regiões que apresentam uma maior variação da função custo composta, em módulo (módulo de $\frac{dJ_c}{d\omega_1}$), entende-se que haverá uma menor densidade de soluções naquelas regiões em que a norma do vetor composto pelas derivadas $\frac{dJ_c}{d\omega_i}$ ($i = 1, \dots, N-1$, sendo N o número de objetivos) tiver seu valor elevado.

4 Resultados e Discussões

A Figura (6) apresenta a correlação cruzada entre o módulo da derivada de J_c , em função da distância entre dois pontos subsequentes. A linha contínua (em azul) indica o valor da correlação entre os sinais e as linhas tracejadas (verde e vermelho) os intervalos de confiança de 95%.

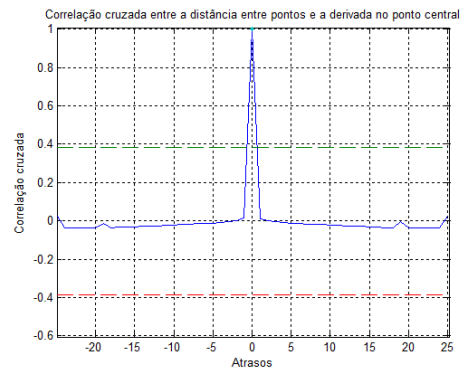


Figura 6: Correlação cruzada entre a derivada de J_c em função da distância entre dois pontos subsequentes.

Existe uma correlação clara (de valor unitário) entre a variação da função custo composta em função da variação do peso em função da distância entre dois pontos subsequentes. Dessa forma, ratifica-se a teoria proposta nesse trabalho em que se considera haver uma relação direta entre a distância de pontos na fronteira Pareto-ótimo com a derivada da função objetivo composta naquela região.

O valor da derivada e a consequente distribuição de pontos proposta pela equação (8) foram aplicadas à massa de dados de identificação. Uma vez obtida, por meio de $NP(\%)$, a distribuição percentual de pontos, utilizou-se os dados de validação para obter um conjunto de 100 soluções utilizando a abordagem aqui proposta. Foram obtidos com propósitos comparativos 100 soluções considerando 100 valores de ω_1 , linearmente espaçados entre ($0 \leq \omega_1 \leq 1$). Os resultados podem ser vistos nas Figuras (7) e (8). Os resultados obtidos considerando os dados de identificação se encontram nas Figuras (9) e (10).

Como pode ser visto, em ambas as massas de dados (identificação e validação), há uma melhor povoação do conjunto Pareto-ótimo em regiões extremas, quando uma das funções objetivo tendem a ser minimizadas. Isso é de se esperar

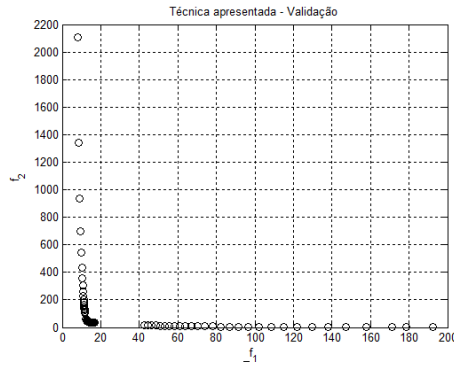


Figura 7: Conjunto de 100 soluções Pareto-ótimas obtido por meio da técnica proposta. Dados de validação.

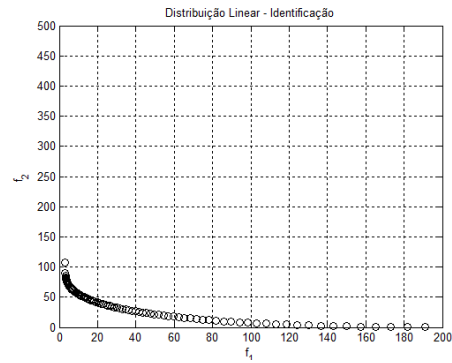


Figura 10: Conjunto de 100 soluções Pareto-ótimas obtido por variação linear de ω_1 . Dados de identificação.

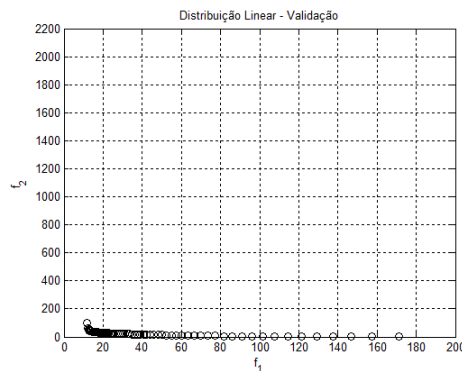


Figura 8: Conjunto de 100 soluções Pareto-ótimas obtido por variação linear de ω_1 . Dados de validação.

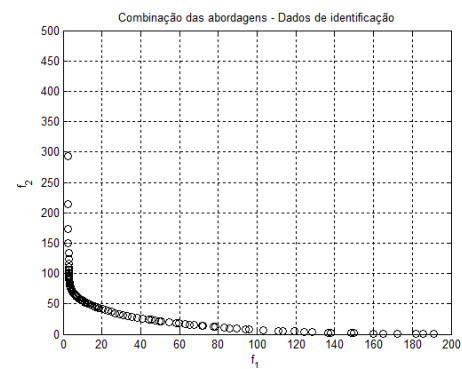


Figura 11: Combinação da distribuição linear de ω_1 e da técnica proposta. Dados de identificação.

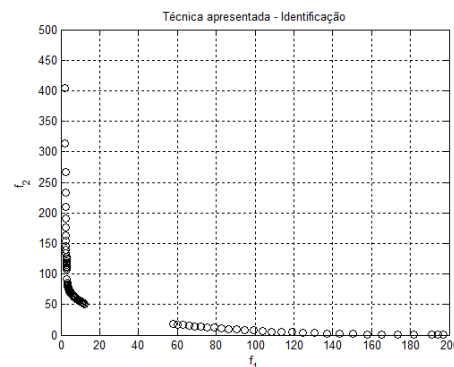


Figura 9: Conjunto de 100 soluções Pareto-ótimas obtido por meio da técnica proposta. Dados de identificação.

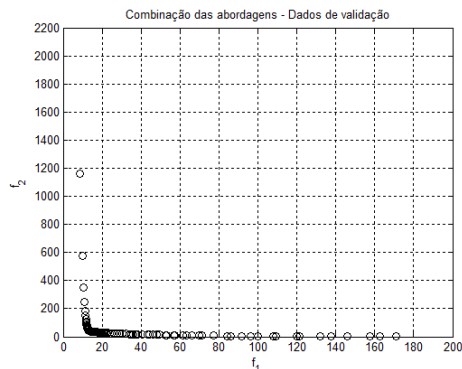


Figura 12: Combinação da distribuição linear de ω_1 e da técnica proposta. Dados de validação.

uma vez que, de acordo com a Figura (4), há uma grande variação, em módulo, da função objetivo composta (J_c) em função do peso ω_1 nessa região. Dessa forma, o método proposto busca por um maior número de soluções nessas regiões onde a derivada tem, em módulo, valores elevados.

Porém, como o número de soluções é limitado, há uma perda de diversidade em regiões centrais do conjunto Pareto-ótimo, quando comparada a técnica com a distribuição linear do peso ω_1 .

Uma vez que a técnica apresentada é eficiente em encontrar soluções nos extremos da função objetivo composta (soluções ótimas individuais) e que a distribuição linear é eficiente em encontrar soluções intermediárias, pode-se propor, como nas Figuras (11) e (12), uma combinação das duas técnicas de modo a obter um conjunto Pareto-ótimo bem representativo. Dessa forma, pode-se propor que 50% do número de pontos sejam distribuídos via distribuição linear e os outros 50% via técnica aqui apresentada. Assim, pode-se obter um conjunto representativo, bem distribuído, da fronteira

Pareto-ótimo do problema de otimização multiobjetivo.

5 Considerações Finais

O presente trabalho apresentou uma nova técnica de distribuição dos pesos ω_i em um problema de otimização multiobjetivo solucionado pelo método escalar da soma ponderada, visando obter um conjunto de soluções eficientes com maior representatividade do conjunto Pareto-ótimo original. Como problema estudado, utilizou-se a codificação originalmente apresentada em (Nepomuceno et al., 2007) para estimação de parâmetros de modelos NARMAX polinomiais via abordagens multiobjetivo via método escalar da soma ponderada.

A técnica se baseia na relação entre o módulo da derivada da função composta pela soma ponderada dos funcionais individuais e em sua relação com a distribuição de soluções ao longo da fronteira Pareto-ótimo. Pôde-se, por meio de tal técnica, obter um conjunto de soluções em regiões não povoadas quando utilizada a distribuição linear dos pesos associados aos objetivos (particularmente em regiões onde a ponderação tende ou à unidade ou à zero).

A técnica, aplicada nesse trabalho em um problema de otimização bi-objetivo, pode ser estendida à um problema multiobjetivo, substituindo o módulo da derivada pela norma do vetor composto pelas derivadas da função composta em função das ponderações ω_i . Outro ponto é que, para N objetivos, é necessário o cálculo de $N - 1$ derivadas, uma vez que um dos valores de peso pode ser descrito em função dos demais.

Por fim, tem-se que a combinação entre a técnica apresentada e a distribuição linear dos pesos (50% das soluções obtidas via distribuição linear e 50% via técnica apresentada) fornece uma maior diversidade de soluções e uma boa representatividade da fronteira Pareto-ótima do problema de otimização multiobjetivo.

Agradecimentos

À Fapemig, Capes e CNPq pelo apoio.

Referências

- Aguirre, L. A. (2007). *Introdução à Identificação de Sistemas: Técnicas lineares e não-lineares aplicadas a sistemas reais*, Editora da UFMG. 3ª edição.
- Aguirre, L. A. (2014). Identification of smooth nonlinear dynamical systems with non-smooth steady-state features, *Automatica* **50**(4): 1160 – 1166.
- Ayala, H. V. H., Habineza, D., Rakotondrabe, M., Klein, C. E. and Coelho, L. S. (2015). Non-linear Black-box System Identification through Neural Networks of a Hysteretic Piezoelectric Robotic Micromanipulator, *17th IFAC Symposium on System Identification*.
- Bassi, F. C., Martins, S. A. M. and Nepomuceno, E. G. (2015). Influência do Máximo Atraso na Determinação do Tamanho de Modelos via Critério de Informação de Akaike Multi-objetivo, *Anais da Conferência Brasileira de Dinâmica, Controle e Aplicações*.
- Billings, S. A. (2013). *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*, West Sussex: John Wiley & Sons.
- Johansen, T. A. (1996). Identification of nonlinear systems using empirical data and prior knowledge - an optimization approach, *Automatica* **32**(3): 337–356.
- Martins, S. A. M., Braga, D. C. S., Nepomuceno, E. G., Gomes, T. V. and Reis, M. L. F. (2009). Investigation of the static curve information for multiobjective system identification, *Journal of Computational Interdisciplinary Sciences* **1**(2): 149–157.
- Martins, S. A. M., Nepomuceno, E. G. and Barroso, M. F. S. (2013). Improved structure detection for polynomial NARX models using a multiobjective error reduction ratio, *Journal of Control, Automation and Electrical Systems* **24**: 764–772.
- Nepomuceno, E. G., Takahashi, R. H. C. and Aguirre, L. A. (2007). Multiobjective parameter estimation for non-linear systems: affine information and least-squares formulation, *International Journal of Control* **80**(6): 863–871.
- Yang, X., Zhang, H. and Ma, X. (2009). Modeling and stability analysis of cascade buck converters with n power stages, *Mathematics and Computers in Simulation* **80**(3): 533 – 546.
- Zhang, B. and Billings, S. A. (2015). Identification of continuous-time nonlinear systems: The nonlinear difference equation with moving average noise (NDEMA) framework, *Mechanical Systems and Signal Processing* **60**: 810–835.