

FORMULAÇÃO ITERATIVA BASEADA EM VARIÁVEIS INSTRUMENTAIS E BASES DE FUNÇÕES RACIONAIS PARA IDENTIFICAÇÃO DE SISTEMAS NO DOMÍNIO DO TEMPO

RICARDO SCHUMACHER*, GUSTAVO H. C. OLIVEIRA*

**Departamento de Engenharia Elétrica
Universidade Federal do Paraná (UFPR)
Curitiba, Paraná, Brasil*

Emails: schumacher.ric@gmail.com, gustavo@eletrica.ufpr.br

Abstract— System identification using linear models formed by rational basis functions (RBFs) plays an important role in system macromodeling. Nevertheless, estimating RBF models involves, in general, the minimization of an objective function that is nonlinear with respect to its variables. In order to address this issue, Sanathanan-Koerner/Steiglitz-McBride (SK/SM) iteration-based approaches approximate the initial nonlinear objective function by a sequence of linear objective functions, which are then easily minimized by the least-squares method. Within such context, this paper proposes a formulation using instrumental variables (IV) for system identification based on time-domain data. Unlike previously proposed approaches, this novel IV formulation ensures that the solution after the convergence of the iterations is a local optimum of the nonlinear objective function initially defined. When compared with its respective recently proposed standard formulation, the results contained in this paper show that the proposed IV formulation can present advantages in terms of model accuracy and convergence of the iterative algorithm. Such results consider models formed by two different RBFs: pole-residue type functions as well as Takenaka-Malmquist orthonormal functions.

Keywords— System Macromodeling, Models Formed by Rational Basis Functions, Sanathanan-Koerner/Steiglitz-McBride Iterations, Instrumental Variables.

Resumo— A identificação de sistemas usando modelos lineares formados por bases de funções racionais (BFRs) desempenha um papel fundamental na macromodelagem de sistemas. Todavia, a estimação de modelos BFRs envolve, a princípio, a minimização de uma função objetivo não-linear em relação às suas variáveis. Visando tratar desta questão, abordagens baseadas nas iterações de Sanathanan-Koerner/Steiglitz-McBride (SK/SM) aproximam a função objetivo não-linear inicial por uma sequência de funções objetivo lineares, que são então facilmente minimizadas pelo método dos mínimos quadrados. Neste contexto, este trabalho propõe uma formulação usando variáveis instrumentais (VI) que visa a identificação de sistemas com dados no domínio do tempo. Diferentemente de formulações propostas anteriormente, esta nova formulação VI garante que a solução encontrada após a convergência das iterações é um ótimo local da função objetivo não-linear inicialmente definida. Quando comparada com a sua respectiva formulação padrão proposta recentemente, os resultados contidos neste trabalho mostram que a formulação VI aqui proposta pode apresentar vantagens em termos de precisão dos modelos e convergência do processo iterativo. Tais resultados levam em conta modelos formados por duas diferentes BFRs: funções do tipo polo-resíduo e também as funções ortonormais de Takenaka-Malmquist.

Palavras-chave— Macromodelagem de Sistemas, Modelos Formados por Bases de Funções Racionais, Iterações de Sanathanan-Koerner/Steiglitz-McBride, Variáveis Instrumentais.

1 Introdução

Na macromodelagem de sistemas, o problema de se modelar um sistema (excessivamente) complexo é resolvido, dividindo-se e modelando-se de forma independente, sub-blocos e/ou dispositivos internos a este sistema. Neste contexto, a identificação de sistemas usando modelos lineares formados por bases de funções racionais (BFRs) tem desempenhado, especialmente na última década, um papel fundamental para a construção de algoritmos capazes de estimar modelos que atendem requisitos como passividade, estabilidade, causalidade e reciprocidade (Grivet-Talocia and Gustavsen, 2016).

Todavia, a estimação de modelos BFRs envolve, a princípio, a minimização de uma função objetivo não-linear em relação às suas variáveis. Visando tratar desta questão, abordagens baseadas nas iterações de Sanathanan-Koerner/Steiglitz-McBride (SK/SM) (Sanathanan and Koerner, 1963; Steiglitz and

McBride, 1965) aproximam a função objetivo não-linear inicial por uma sequência de funções objetivo lineares, que são então facilmente minimizadas pelo método dos mínimos quadrados.

A primeira técnica baseada nas iterações SK/SM para estimação de modelos BFRs, denominada de *Vector Fitting*, foi proposta no domínio da frequência por Gustavsen and Semlyen (1999) para modelos formados por funções do tipo polo-resíduo. Com o objetivo de melhorar a estabilidade numérica e a convergência do processo iterativo, Deschrijver et al. (2007) propuseram então a técnica *Orthonormal Vector Fitting* que, por sua vez, baseia-se na substituição das funções do tipo polo-resíduo pelas chamadas funções ortonormais de Takenaka-Malmquist.

Apesar do evidente sucesso na aplicação de abordagens SK/SM para estimação de modelos BFRs no domínio da frequência, abordagens similares para identificação com dados no domínio do tempo permanecem ainda pouco exploradas. Re-

centemente, Schumacher and Oliveira (2015) propuseram um método descrito em tempo discreto que permite estimar modelos formados tanto por funções do tipo polo-resíduo quanto pelas funções ortonormais de Takenaka-Malmquist. Conforme mostram os autores deste trabalho, tal método pode apresentar vantagens (em termos de precisão dos modelos e convergência do processo iterativo) em relação à técnica similar *Time-Domain Vector Fitting* proposta por Grivet-Talocia (2003), cuja formulação permite o uso apenas de funções do tipo polo-resíduo.

A formulação proposta por Schumacher and Oliveira (2015), entretanto, não garante que a solução encontrada após a convergência das iterações é um ótimo local ou global da função objetivo não-linear que define o problema de otimização inicial. Na verdade, tal propriedade para a solução encontrada não é garantida por nenhuma das tradicionais abordagens SK/SM presentes na literatura (Grivet-Talocia and Gustavsen, 2016).

Com base nestas considerações, o presente trabalho tem por objetivo reformular, usando-se variáveis instrumentais (VI), o método proposto em (Schumacher and Oliveira, 2015) a fim de garantir que a solução encontrada após a convergência das iterações é de fato um ótimo local da função objetivo não-linear inicial. Salienta-se que esta nova formulação VI não modifica a estrutura de modelo BFR utilizada, mas sim o critério usado para se estimar os seus parâmetros.

Este trabalho está organizado conforme segue. Na Seção 2, a formulação padrão proposta em (Schumacher and Oliveira, 2015) é revista dando-se enfoque na definição da função objetivo não-linear. Na realidade, em (Schumacher and Oliveira, 2015) uma notação simplificada é usada e a função objetivo não é explicitamente definida. Esta versão mais detalhada desempenha um papel fundamental no desenvolvimento da formulação VI, proposta na Seção 3. Na Seção 4, os resultados mostram que, diferentemente da formulação padrão, a formulação VI é capaz de identificar o real comportamento do sistema avaliado, cuja saída encontra-se corrompida por ruído colorido. Por fim, a Seção 5 ressalta as conclusões deste trabalho.

2 Formulação iterativa padrão

No domínio do tempo discreto, a relação entre a entrada escalar $u_0(k)$ e a saída escalar $y_0(k)$ de um sistema linear invariante no tempo pode ser descrita de acordo com a seguinte expressão:

$$y_0(k) = G_0(q)u_0(k) + v(k), \quad (1)$$

onde $v(k)$ representa o ruído aditivo presente na saída $y_0(k)$ e q denota o operador avanço, $q^j u_0(k) = u_0(k+j)$, $j \in \mathbb{Z}$.

O desafio de identificação tratado neste trabalho consiste em encontrar um modelo BFR

cujo comportamento dinâmico seja suficientemente próximo do comportamento dinâmico apresentado por $G_0(q)$. Tal modelo deve ter ainda uma estrutura do tipo

$$G(q, \mathbf{c}, \mathbf{a}) = c_0 + \sum_{i=1}^n c_i \Phi_i(q, \mathbf{a}), \quad (2)$$

onde c_0 e $\{c_i\}$ são coeficientes a serem estimados e $\mathbf{a}$ representa um conjunto de n polos que parametriza as funções racionais $\{\Phi_i(q, \mathbf{a})\}$. Tais polos também devem ser estimados. Destaca-se ainda que $\{\Phi_i(q, \mathbf{a})\}$ pode ser um conjunto formado tanto por funções do tipo polo-resíduo quanto pelas funções ortonormais de Takenaka-Malmquist (Schumacher and Oliveira, 2015).

O problema de otimização associado à estimação dos parâmetros da estrutura em (2) é, em princípio, não-linear em relação às suas variáveis, uma vez que os polos encontram-se presentes nos denominadores das funções racionais.

Neste contexto, abordagens baseadas nas iterações de SK/SM estimam $\mathbf{a}$ (iterativamente) por meio de uma estrutura de modelo genérica do tipo

$$\bar{G}(q, \theta, \bar{\mathbf{a}}) = \frac{B(q, \theta, \bar{\mathbf{a}})}{F(q, \theta, \bar{\mathbf{a}})} = \frac{B(q, \theta, \bar{\mathbf{a}})/\hat{F}(q, \bar{\mathbf{a}})}{\hat{F}(q, \theta, \bar{\mathbf{a}})/\hat{F}(q, \bar{\mathbf{a}})}, \quad (3)$$

onde $\bar{\mathbf{a}}$ é um conjunto pré-especificado de n polos,

$$\frac{B(q, \theta, \bar{\mathbf{a}})}{\hat{F}(q, \bar{\mathbf{a}})} = r_0 + \sum_{i=1}^n r_i \Phi_i(q, \bar{\mathbf{a}}), \quad (4)$$

$$\frac{F(q, \theta, \bar{\mathbf{a}})}{\hat{F}(q, \bar{\mathbf{a}})} = 1 + \sum_{i=1}^n d_i \Phi_i(q, \bar{\mathbf{a}}). \quad (5)$$

A partir de um conjunto de amostras discretizadas no domínio do tempo $\{y_0(k), u_0(k)\}$, $k = 1, \dots, N$, os coeficientes

$$\theta = [r_0 \quad \dots \quad r_n \quad d_1 \quad \dots \quad d_n]^T$$

da estrutura em (3) são então estimados baseando-se na minimização do erro quadrático médio (EQM) entre a saída do sistema $y_0(k)$ e a saída da estrutura definida por $y(k, \theta, \bar{\mathbf{a}}) = \bar{G}(q, \theta, \bar{\mathbf{a}})u_0(k)$, ou seja, baseando-se na minimização da função objetivo não-linear:

$$\begin{aligned} V(\theta) &= \frac{1}{N} \sum_{k=1}^N (y_0(k) - \bar{G}(q, \theta, \bar{\mathbf{a}})u_0(k))^2, \\ &= \frac{1}{N} \sum_{k=1}^N \left(\frac{F(q, \theta, \bar{\mathbf{a}})}{F(q, \theta, \bar{\mathbf{a}})} y_0(k) - \frac{B(q, \theta, \bar{\mathbf{a}})}{F(q, \theta, \bar{\mathbf{a}})} u_0(k) \right)^2. \end{aligned} \quad (6)$$

O valor de θ que minimiza $V(\theta)$ será denotado por

$$\theta^* = \arg \min_{\theta} V(\theta). \quad (7)$$

Note-se que o problema de otimização que busca estimar θ^* é, inicialmente, não-linear em relação às variáveis em θ , uma vez que θ está presente no denominador de $V(\theta)$. Para resolver esta questão, o seguinte procedimento baseado nas iterações SK/SM é utilizado:

$$\hat{\theta}^{(t)} = \arg \min_{\theta} \frac{1}{N} \sum_{k=1}^N \left(\frac{F(q, \theta, \mathbf{a}^{(t)})}{\hat{F}(q, \mathbf{a}^{(t)})} y_0(k) - \frac{B(q, \theta, \mathbf{a}^{(t)})}{\hat{F}(q, \mathbf{a}^{(t)})} u_0(k) \right)^2, \quad (8)$$

ou, de forma equivalente,

$$\hat{\theta}^{(t)} = \arg \min_{\theta} \sum_{k=1}^N \left(\left(1 + \sum_{i=1}^n d_i \Phi_i(q, \mathbf{a}^{(t)}) \right) y_0(k) - \left(r_0 + \sum_{i=1}^n r_i \Phi_i(q, \mathbf{a}^{(t)}) \right) u_0(k) \right)^2. \quad (9)$$

Em (8) e (9), $\hat{\theta}^{(t)}$ é a estimativa para θ^* obtida na iteração t e $\mathbf{a}^{(t)}$ representa o conjunto pré-especificado de polos para o qual $\hat{\theta}^{(t)}$ é calculado.

Note-se que a Equação (9) descreve uma sequência iterativa de problemas de otimização que são agora lineares na variável θ . Tal equação pode ser ainda reescrita na forma compacta

$$\hat{\theta}^{(t)} = \arg \min_{\theta} \sum_{k=1}^N \left(y_0(k) - \mathbf{m}^T(k, \mathbf{a}^{(t)}) \theta \right)^2, \quad (10)$$

cujas solução analítica é então dada por:

$$\hat{\theta}^{(t)} = \left[\mathbf{M}^T(\mathbf{a}^{(t)}) \mathbf{M}(\mathbf{a}^{(t)}) \right]^{-1} \mathbf{M}^T(\mathbf{a}^{(t)}) \mathbf{y}, \quad (11)$$

onde

$$\begin{aligned} \mathbf{M}^T(\mathbf{a}^{(t)}) &= \begin{bmatrix} \mathbf{m}(1, \mathbf{a}^{(t)}) & \cdots & \mathbf{m}(N, \mathbf{a}^{(t)}) \end{bmatrix}, \\ \mathbf{y}^T &= \begin{bmatrix} y_0(1) & \cdots & y_0(N) \end{bmatrix}, \\ \mathbf{m}^T(k, \mathbf{a}^{(t)}) &= \begin{bmatrix} u_0(k) & \tilde{\mathbf{u}}^T(k) & -\tilde{\mathbf{y}}^T(k) \end{bmatrix}, \\ \tilde{\mathbf{u}}^T(k) &= \begin{bmatrix} \Phi_1(q, \mathbf{a}^{(t)}) & \cdots & \Phi_n(q, \mathbf{a}^{(t)}) \end{bmatrix} u_0(k), \\ \tilde{\mathbf{y}}^T(k) &= \begin{bmatrix} \Phi_1(q, \mathbf{a}^{(t)}) & \cdots & \Phi_n(q, \mathbf{a}^{(t)}) \end{bmatrix} y_0(k). \end{aligned}$$

A formulação iterativa padrão se resume assim em resolver a Equação (11) iterativamente.

Os termos $\{\mathbf{m}^T(k, \mathbf{a}^{(t)})\}$ presentes na matriz $\mathbf{M}(\mathbf{a}^{(t)})$ podem ser facilmente obtidos usando-se recursivamente as seguintes equações de estado:

$$\begin{cases} \tilde{\mathbf{u}}(k) &= \mathbf{A} \tilde{\mathbf{u}}(k-1) + \mathbf{B} u_0(k-1), \\ \tilde{\mathbf{y}}(k) &= \mathbf{A} \tilde{\mathbf{y}}(k-1) + \mathbf{B} y_0(k-1), \end{cases} \quad (12)$$

onde $\mathbf{A}(\mathbf{a}^{(t)}) \in \mathbb{R}^{(n \times n)}$ e $\mathbf{B}(\mathbf{a}^{(t)}) \in \mathbb{R}^{(n \times 1)}$ são matrizes definidas de tal forma que as funções de transferência da entrada (ou saída) para os estados são dadas pelas próprias funções racionais que compõem o modelo, ou seja,

$$(q\mathbf{I} - \mathbf{A})^{-1} \mathbf{B} = \begin{bmatrix} \Phi_1(q, \mathbf{a}^{(t)}) \\ \vdots \\ \Phi_n(q, \mathbf{a}^{(t)}) \end{bmatrix} \quad (13)$$

Note-se que a dependência de $\mathbf{A}(\mathbf{a}^{(t)})$, $\mathbf{B}(\mathbf{a}^{(t)})$, $\tilde{\mathbf{u}}(k, \mathbf{a}^{(t)})$ e $\tilde{\mathbf{y}}(k, \mathbf{a}^{(t)})$ nos polos em $\mathbf{a}^{(t)}$ foi omitida nas equações anteriores para facilitar a notação. Note-se também a partir da Equação (13) que $\mathbf{A}(\mathbf{a}^{(t)})$ e $\mathbf{B}(\mathbf{a}^{(t)})$ definem a BFR utilizada no modelo, ou seja, o conjunto $\{\Phi_i(q, \mathbf{a}^{(t)})\}$. Schumacher and Oliveira (2015) apresentam uma explicação detalhada sobre como essas duas matrizes devem ser construídas para que o conjunto $\{\Phi_i(q, \mathbf{a}^{(t)})\}$ seja dado por funções do tipo polo-resíduo, ou ainda, pelas funções ortonormais de Takenaka-Malmquist.

A definição do conjunto de polos $\mathbf{a}^{(t)}$, por sua vez, tem por objetivo tornar a sequência de problemas lineares de otimização em (8) próxima ao problema inicialmente não-linear definido em (7). Para tanto, define-se $\mathbf{a}^{(t)}$ de forma a se satisfazer:

$$\hat{F}(q, \mathbf{a}^{(t)}) = F(q, \hat{\theta}^{(t-1)}, \mathbf{a}^{(t-1)}). \quad (14)$$

A condição acima é satisfeita fazendo-se com que as raízes dos dois polinômios sejam iguais. Em outras palavras, os polos em $\mathbf{a}^{(t)}$ devem ser iguais às raízes de $F(q, \hat{\theta}^{(t-1)}, \mathbf{a}^{(t-1)})$, ou seja,

$$\mathbf{a}^{(t)} = \text{autovalores} \left(\mathbf{A}(\mathbf{a}^{(t-1)}) - \mathbf{B}(\mathbf{a}^{(t-1)}) \mathbf{d}(\hat{\theta}^{(t-1)}) \right), \quad (15)$$

onde¹ $\mathbf{d}(\hat{\theta}^{(t-1)}) = [d_1^{(t-1)} \cdots d_n^{(t-1)}]$. O lado direito da Equação (15) representa um meio para se calcular as raízes de $F(q, \hat{\theta}^{(t-1)}, \mathbf{a}^{(t-1)})$ a partir da realização em espaço de estados para a expressão $F(q, \hat{\theta}^{(t-1)}, \mathbf{a}^{(t-1)}) / \hat{F}(q, \mathbf{a}^{(t-1)})$, obtida quando $\mathbf{A}(\mathbf{a}^{(t-1)})$ e $\mathbf{B}(\mathbf{a}^{(t-1)})$ são usadas na definição da sua equação de estado. Para mais detalhes, ver (Grivet-Talocia and Gustavsen, 2016).

Para finalizar esta seção, discute-se agora a convergência desta formulação padrão.

Assumindo-se que $\hat{\theta}^{(t)}$ e $\mathbf{a}^{(t)}$ convergem para $\hat{\theta}$ e $\hat{\mathbf{a}}$ na iteração $t = c$, então $\hat{\theta}^{(c)} = \hat{\theta}^{(c+1)} = \hat{\theta}$ e $\mathbf{a}^{(c)} = \mathbf{a}^{(c+1)} = \hat{\mathbf{a}}$. Portanto, pela Equação (14),

$$\begin{aligned} \hat{F}(q, \mathbf{a}^{(c+1)}) &= F(q, \hat{\theta}^{(c)}, \mathbf{a}^{(c)}) = \\ &= F(q, \hat{\theta}^{(c+1)}, \mathbf{a}^{(c+1)}) = F(q, \hat{\theta}, \hat{\mathbf{a}}) = \hat{F}(q, \hat{\mathbf{a}}). \end{aligned} \quad (16)$$

Tendo como base estas equivalências, é fundamental observar a partir das Equações (3), (4) e (5), que o modelo obtido após a convergência das iterações tem a forma desejada em (2),

$$\bar{G}(q, \hat{\theta}, \hat{\mathbf{a}}) = \hat{r}_0 + \sum_{i=1}^n \hat{r}_i \Phi_i(q, \hat{\mathbf{a}}), \quad (17)$$

uma vez que

$$1 + \sum_{i=1}^n \hat{d}_i \Phi_i(q, \hat{\mathbf{a}}) = \frac{F(q, \hat{\theta}, \hat{\mathbf{a}})}{\hat{F}(q, \hat{\mathbf{a}})} = 1.$$

Por fim, ressalta-se que a formulação padrão apresentada nesta seção não permite garantir que $\hat{\theta}$ (valor para o qual $\hat{\theta}^{(t)}$ converge) é um mínimo local da função objetivo original $V(\theta)$ definida em (6), ou ainda, que $\hat{\theta} = \theta^*$ (ver Equação (7)).

Para fundamentar esta proposição, note-se a partir das equivalências em (16) que a Equação (8) pode ser reescrita na iteração $t = c + 1$ como sendo:

$$\hat{\theta} = \arg \min_{\theta} \frac{1}{N} \sum_{k=1}^N \left(\frac{F(q, \theta, \hat{\mathbf{a}})}{F(q, \hat{\theta}, \hat{\mathbf{a}})} y_0(k) - \frac{B(q, \theta, \hat{\mathbf{a}})}{F(q, \hat{\theta}, \hat{\mathbf{a}})} u_0(k) \right)^2.$$

Como pode ser percebido, a função objetivo minimizada nesta expressão não coincide necessariamente com $V(\theta)$. Na realidade, a equação acima

¹Caso não exista nenhum conhecimento prévio aproximado acerca dos polos de $G_0(q)$ na primeira iteração, pode-se fazer $\mathbf{a}^{(0)} = \mathbf{0}$.

mostra que o m nimo global θ^*   obtido somente se o denominador da fun  o objetivo for parametrizado pelo pr prio θ^* .

3 Formula  o iterativa usando vari veis instrumentais (formula  o VI)

O objetivo desta se  o  , usando-se vari veis instrumentais, reformular o m todo da se  o anterior a fim de garantir que a solu  o encontrada ap s a converg ncia das itera  es   de fato um  timo local da fun  o objetivo inicial $V(\theta)$. Para tanto, se um determinado $\hat{\theta}_{OL}$   um  timo local de $V(\theta)$, ent o ele necessariamente satisfaz $\frac{\partial V(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}_{OL}} = \mathbf{0}$, ou ainda, pela regra da cadeia:

$$\sum_{k=1}^N \mathbf{p}(k, \hat{\theta}_{OL}, \bar{\mathbf{a}}) \left(\frac{F(q, \hat{\theta}_{OL}, \bar{\mathbf{a}})}{F(q, \hat{\theta}_{OL}, \bar{\mathbf{a}})} y_0(k) - \frac{B(q, \hat{\theta}_{OL}, \bar{\mathbf{a}})}{F(q, \hat{\theta}_{OL}, \bar{\mathbf{a}})} u_0(k) \right) = \mathbf{0}, \quad (18)$$

sendo $\mathbf{p}(k, \theta, \bar{\mathbf{a}})$ o vetor de vari veis instrumentais definido por

$$\mathbf{p}(k, \theta, \bar{\mathbf{a}}) = \frac{\partial}{\partial \theta} y(k, \theta, \bar{\mathbf{a}}) = \frac{\partial}{\partial \theta} \bar{G}(q, \theta, \bar{\mathbf{a}}) u_0(k). \quad (19)$$

Ao t rmino desta se  o, mostra-se com detalhes como os termos $\{\mathbf{p}(k, \theta, \bar{\mathbf{a}})\}$ podem ser calculados.

  poss vel observar pela Equa  o (18) que o problema de se encontrar $\hat{\theta}_{OL}$   tamb m n o-linear em rela  o  s suas vari veis. Buscando resolver esse problema de maneira similar ao realizado na se  o anterior (Equa  o (8)), prop e-se aqui o seguinte processo iterativo de regress o por vari veis instrumentais:

$$\sum_{k=1}^N \mathbf{p}(k, \hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)}) \left(\frac{F(q, \hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)})}{\hat{F}(q, \mathbf{a}^{(t)})} y_0(k) - \frac{B(q, \hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)})}{\hat{F}(q, \mathbf{a}^{(t)})} u_0(k) \right) = \mathbf{0}. \quad (20)$$

Esta express o implica que, a cada itera  o, $\hat{\theta}_{VI}^{(t)}$ deve satisfazer (baseando-se na not  o j  introduzida na se  o anterior):

$$\sum_{k=1}^N \mathbf{p}(k, \hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)}) \left(y_0(k) - \mathbf{m}^T(k, \mathbf{a}^{(t)}) \hat{\theta}_{VI}^{(t)} \right) = \mathbf{0},$$

ou ainda, na forma matricial equivalente,

$$\mathbf{P}^T(\hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)}) [\mathbf{y} - \mathbf{M}(\mathbf{a}^{(t)}) \hat{\theta}_{VI}^{(t)}] = \mathbf{0}, \quad (21)$$

sendo

$$\mathbf{P}(\hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)}) = \begin{bmatrix} \mathbf{p}^T(1, \hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)}) \\ \vdots \\ \mathbf{p}^T(N, \hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)}) \end{bmatrix}. \quad (22)$$

Isolando-se ent o o vetor de par metros $\hat{\theta}_{VI}^{(t)}$ em (21), tem-se a seguinte solu  o anal tica para a formula  o usando vari veis instrumentais:

$$\hat{\theta}_{VI}^{(t)} = [\mathbf{P}^T(\hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)}) \mathbf{M}(\mathbf{a}^{(t)})]^{-1} \mathbf{P}^T(\hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)}) \mathbf{y}. \quad (23)$$

Comparando-se esta solu  o com a solu  o padr o em (11),   poss vel observar que elas diferem entre si uma vez que a formula  o usando vari veis instrumentais requer tamb m a constru  o da matriz $\mathbf{P}(\hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)})$. Al m disso, diferentemente do caso padr o, esta nova abordagem permite garantir que um  timo local da fun  o objetivo n o-linear inicial $V(\theta)$   obtido quando as itera  es em (23) convergem. Tal proposi  o   confirmada a seguir.

Assumindo que a realoca  o de polos ocorra de acordo com a Equa  o (14) e que, al m disso, $\hat{\theta}_{VI}^{(t)}$ e $\mathbf{a}^{(t)}$ convergem para $\hat{\theta}_{VI}$ e $\hat{\mathbf{a}}$ na itera  o $t = c$, ent o $\hat{\theta}_{VI}^{(c)} = \hat{\theta}_{VI}^{(c+1)} = \hat{\theta}_{VI}$, $\mathbf{a}^{(c)} = \mathbf{a}^{(c+1)} = \hat{\mathbf{a}}$. Portanto, pela pr pria Equa  o (14),

$$\hat{F}(q, \mathbf{a}^{(c+1)}) = F(q, \hat{\theta}_{VI}^{(c)}, \mathbf{a}^{(c)}) = F(q, \hat{\theta}_{VI}^{(c+1)}, \mathbf{a}^{(c+1)}) = F(q, \hat{\theta}_{VI}, \hat{\mathbf{a}}) = \hat{F}(q, \hat{\mathbf{a}}). \quad (24)$$

Dessa forma, na itera  o $t = c + 1$ a Equa  o (20) pode ser reescrita como sendo

$$\sum_{k=1}^N \mathbf{p}(k, \hat{\theta}_{VI}, \hat{\mathbf{a}}) \left(\frac{F(q, \hat{\theta}_{VI}, \hat{\mathbf{a}})}{F(q, \hat{\theta}_{VI}, \hat{\mathbf{a}})} y_0(k) - \frac{B(q, \hat{\theta}_{VI}, \hat{\mathbf{a}})}{F(q, \hat{\theta}_{VI}, \hat{\mathbf{a}})} u_0(k) \right) = \mathbf{0}. \quad (25)$$

Esta express o coincide com a condi  o de otimalidade local imposta em (18) e, portanto, $\hat{\theta}_{VI}$   de fato um  timo local de $V(\theta)$.

Por fim, salienta-se que   necess ria uma estimativa inicial $\hat{\theta}_{VI}^{(t-1)}$ para se dar in cio   formula  o VI (ver Equa  o (23)). Esta estimativa inicial pode ser obtida utilizando-se, durante um determinado n mero de itera  es, a formula  o padr o em (11) e, em seguida, fazendo-se $\hat{\theta}_{VI}^{(t-1)} = \hat{\theta}^{(t-1)}$.

3.1 C lculo aproximado do vetor de instrumentos $\mathbf{p}(k, \hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)})$

Nesta se  o, mostra-se como os termos $\{\mathbf{p}(k, \hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)})\}$ podem ser aproximados de modo a simplificar a montagem da matriz $\mathbf{P}(\hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)})$. De acordo com as defini  es realizadas em (3), (4), (5) e (19), tem-se que

$$\mathbf{p}(k, \theta, \bar{\mathbf{a}}) = \frac{\partial}{\partial \theta} \left(\frac{r_0 + \sum_{i=1}^n r_i \Phi_i(q, \bar{\mathbf{a}})}{1 + \sum_{i=1}^n d_i \Phi_i(q, \bar{\mathbf{a}})} \right) u_0(k). \quad (26)$$

Dessa forma, existem tr s tipos poss veis de derivada de primeira ordem:

$$\frac{\partial}{\partial r_0} \bar{G}(q, \theta, \bar{\mathbf{a}}) = \frac{1}{1 + \sum_{i=1}^n d_i \Phi_i(q, \bar{\mathbf{a}})}, \quad (27)$$

$$\frac{\partial}{\partial r_j} \bar{G}(q, \theta, \bar{\mathbf{a}}) = \frac{1}{1 + \sum_{i=1}^n d_i \Phi_i(q, \bar{\mathbf{a}})} \Phi_j(q, \bar{\mathbf{a}}), \quad (28)$$

$$\frac{\partial}{\partial d_j} \bar{G}(q, \theta, \bar{\mathbf{a}}) = \frac{-\left(r_0 + \sum_{i=1}^n r_i \Phi_i(q, \bar{\mathbf{a}})\right)}{\left(1 + \sum_{i=1}^n d_i \Phi_i(q, \bar{\mathbf{a}})\right)^2} \Phi_j(q, \bar{\mathbf{a}}), \quad (29)$$

com $j = 1, \dots, n$. Com base nestas express es, pode-se notar que o c culo da matriz $\mathbf{P}(\hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)})$   bastante custoso computacionalmente, uma vez que tais equa  es devem ser calculadas para $k = 1, \dots, N$ e ainda, a cada itera  o.

Neste trabalho, a aproxima  o

$$\frac{F(q, \theta, \bar{\mathbf{a}})}{\hat{F}(q, \bar{\mathbf{a}})} = 1 + \sum_{i=1}^n d_i \Phi_i(q, \bar{\mathbf{a}}) \approx 1 \quad (30)$$

  aplicada nas Equa  es (27), (28) e (29). Com essa aproxima  o, o problema de se calcular $\mathbf{P}(\hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)})$ pode ser substancialmente simplificado, j  que o vetor de instrumentos $\mathbf{p}(k, \hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)})$ pode ser agora aproximado por (ver Equa  o (13) para dedu  o):

$$\mathbf{p}(k, \hat{\theta}_{VI}^{(t-1)}, \mathbf{a}^{(t)}) \approx \begin{bmatrix} u_0(k) \\ \bar{\mathbf{u}}(k, \mathbf{a}^{(t)}) \\ -\left(r_0^{(t-1)} + \sum_{i=1}^n r_i^{(t-1)} \Phi_i(q, \mathbf{a}^{(t)})\right) \bar{\mathbf{u}}(k, \mathbf{a}^{(t)}) \end{bmatrix}.$$

  essencial ressaltar que esta aproxima  o n o altera as conclus es feitas em rela  o   converg ncia da formula  o VI uma vez que, pela Equa  o (24), a aproxima  o imposta em (30) torna-se uma igualdade quando o algoritmo converge.

4 Resultados

O estudo de caso apresentado nesta se  o tem por objetivo identificar o sistema de terceira ordem definido por

$$G_0(q) = \frac{0,1389 + 0,4166q^{-1} + 0,4166q^{-2} + 0,1389q^{-3}}{1 + 0,4218q^{-1} + 0,2499q^{-2} - 0,5609q^{-3}}. \quad (31)$$

Para se gerar o conjunto de dados de estima  o, considera-se que a entrada $u_0(k)$   uma seq ncia aleat ria (com 1000 amostras) distribu da uniformemente no intervalo $[-\varsigma, \varsigma]$, sendo $\varsigma = 2,5$. A sa da $y_0(k)$, por sua vez,   corrompida por um ru do $v(k)$ colorido. Especificamente, considera-se $v(k) = H_0(q)e(k)$, sendo $H_0(q)$ o filtro

$$H_0(q) = \frac{0,1080 + 0,2160q^{-1} + 0,1080q^{-2}}{1 - 0,8799q^{-1} + 0,3119q^{-2}},$$

e $e(k)$ uma seq ncia aleat ria obtida a partir de uma distribu  o Gaussiana com m dia 0 e vari ncia $0,6^2$.

No processo de estima  o dos modelos, s o utilizadas 25 itera  es e modelos de terceira ordem² ($n = 3$). Sup e-se que n o existe nenhum conhecimento pr vio aproximado acerca dos polos de $G_0(q)$ na primeira itera  o e, portanto, faz-se $\mathbf{a}^{(0)} = \mathbf{0}$. Al m disso, duas BFRs diferentes s o consideradas para as formula  es VI (proposta neste trabalho) e padr o (proposta por Schumacher and Oliveira (2015)): as fun  es do tipo polo-res duo e as fun  es ortonormais de Takenaka-Malmquist. Dessa forma, no total, quatro t cnicas s o aqui comparadas:

- formula  o **padr o** usando fun  es do tipo polo-res duo (PADR O-FPR);
- formula  o **padr o** usando fun  es ortonormais de Takenaka-Malmquist (PADR O-FOTM);
- formula  o **VI** usando fun  es do tipo polo-res duo (VI-FPR);
- formula  o **VI** usando fun  es ortonormais de Takenaka-Malmquist (VI-FOTM).

Salienta-se ainda que, durante as duas primeiras itera  es, as t cnicas baseadas na formula  o VI utilizam a solu  o da formula  o padr o para gerar as suas respectivas estimativas iniciais $\hat{\theta}_{VI}^{(t-1)}$.

Para fins de compara  o entre as t cnicas, utiliza-se aqui o  ndice REQM (raiz quadrada do erro quadr tico m dio). Tais valores de REQMs foram calculados, a cada itera  o, atribuindo-se   estrutura desejada em (2) a estimativa atual $\mathbf{a}^{(t)}$ para o conjunto de polos $\mathbf{a}$ e, em seguida, estimando-se os seus coeficientes. Tal estimativa pode ser facilmente obtida a partir da solu  o (11) considerando-se $\hat{\theta}^{(t)} = \hat{\mathbf{c}}^{(t)} = [c_0^{(t)} \dots c_n^{(t)}]^T$ e fazendo-se $\mathbf{m}^T(k, \mathbf{a}^{(t)}) = [u_0(k) \ \bar{\mathbf{u}}^T(k)]$ na constru  o da matriz $\mathbf{M}(\mathbf{a}^{(t)})$. Pela Equa  o (17) pode-se observar que, quando as estimativas convergem, $c_0^{(t)}$ e $\{c_i^{(t)}\}$ tendem   $\hat{r}_0$ e $\{\hat{r}_i\}$, respectivamente.

A Figura 1 mostra o  ndice REQM obtido para cada t cnica durante o processo iterativo.   poss vel observar a partir desta figura que as t cnicas baseadas na formula  o VI apresentam, ap s a converg ncia, modelos com um menor REQM. Um menor  ndice REQM indica que as sa das dos modelos VI-FPR e VI-FOTM encontram-se mais pr ximas da sa da com ru do colorido $y_0(k)$. Tamb m   poss vel observar na Figura 1 que as t cnicas propostas neste trabalho convergem com um menor n mero de itera  es.

Os modelos BFRs estimados em $t = 25$ (ou seja, obtidos ap s a converg ncia das itera  es) s o agora validados representando-se os mesmos pela raz o entre dois polin mios em q . O modelo

²Note que o sistema a ser identificado tamb m   de terceira ordem. Diferentes valores para n tamb m poderiam ter sido utilizados.

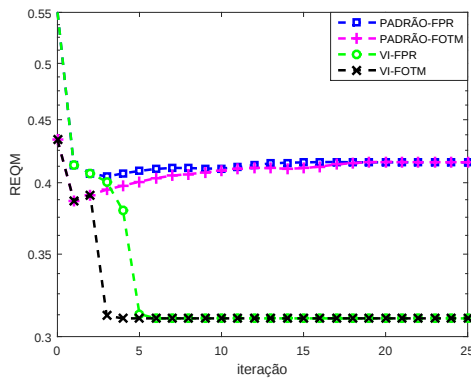


Figura 1: REQm no processo de estimação dos modelos.

equivalente obtido para as técnicas VI-FPR e VI-FOTM foi:

$$\bar{G}(q, \hat{\theta}, \hat{a}) = \frac{0,1311 + 0,4139q^{-1} + 0,4161q^{-2} + 0,1453q^{-3}}{1 + 0,4172q^{-1} + 0,2452q^{-2} - 0,5669q^{-3}}. \quad (32)$$

Já para as técnicas PADRÃO-FPR e PADRÃO-FOTM, o modelo equivalente obtido foi:

$$\bar{G}(q, \hat{\theta}, \hat{a}) = \frac{0,1387 - 0,0358q^{-1} - 0,3325q^{-2} + 0,2382q^{-3}}{1 - 2,6247q^{-1} + 2,2813q^{-2} - 0,6485q^{-3}}. \quad (33)$$

Comparando-se os modelos em (32) e (33) com o sistema em (31), é possível perceber que os modelos obtidos a partir da formulação VI apresentam coeficientes que estão mais próximos dos coeficientes de $G_0(q)$. Dessa forma, pode-se concluir que a formulação VI proposta neste trabalho foi capaz de estimar o real comportamento dinâmico do sistema sob investigação. Isso também pode ser visto através da Figura 2. Esta figura compara parte da saída $G_0(q)u_0(k)$ (ou seja, da saída do sistema sem o efeito do ruído $v(k)$) com as saídas dos modelos em (32) e (33).

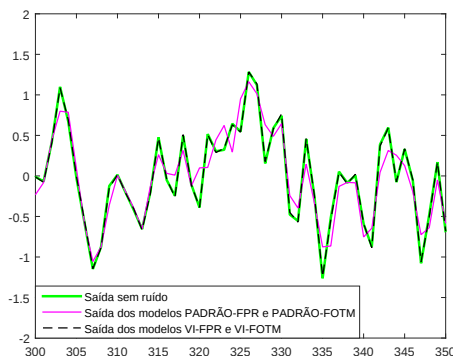


Figura 2: Saídas sem ruído e dos modelos obtidos após a convergência.

5 Conclusões

Neste trabalho, foi proposta uma formulação VI para identificação de sistemas no domínio do

tempo. Tal formulação baseia-se em modelos formados por BFRs e nas iterações de SK/SM. Diferentemente de formulações propostas anteriormente, a formulação VI proposta neste trabalho garante que a solução encontrada após a convergência das iterações é um ótimo local da função objetivo não-linear $V(\theta)$. Os resultados obtidos mostram que a formulação VI pode apresentar vantagens em termos de precisão dos modelos (menor (R)EQM) e convergência do processo iterativo (a formulação VI converge com um menor número de iterações). Em trabalhos futuros, pretende-se avaliar quantitativamente a influência do nível de ruído na qualidade das estimativas.

Agradecimentos

Os autores deste trabalho agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro.

Referências

- Deschrijver, D., Haegeman, B. and Dhaene, T. (2007). Orthonormal vector fitting: A robust macromodeling tool for rational approximation of frequency domain responses, *IEEE Transactions on Advanced Packaging* **30**(2): 216 – 225.
- Grivet-Talocia, S. (2003). Package macromodeling via time-domain vector fitting, *Microwave and Wireless Components Letters, IEEE* **13**(11): 472 – 474.
- Grivet-Talocia, S. and Gustavsen, B. (2016). *Passive Macromodeling*, John Wiley and Sons.
- Gustavsen, B. and Semlyen, A. (1999). Rational approximation of frequency domain responses by vector fitting, *IEEE Transactions on Power Delivery* **14**(3): 1052 – 1061.
- Sanathanan, C. K. and Koerner, J. (1963). Transfer function synthesis as a ratio of two complex polynomials, *IEEE Transactions on Automatic Control* **8**(1): 56–58.
- Schumacher, R. and Oliveira, G. H. C. (2015). Uma nova abordagem *Vector Fitting* para identificação de sistemas com dados no domínio do tempo, *XII Simpósio Brasileiro de Automação Inteligente (SBAI 2015)*, pp. 283–288.
- Steiglitz, K. and McBride, L. E. (1965). A technique for the identification of linear systems, *IEEE Transactions on Automatic Control* **10**(4): 461–464.