

ESTIMAÇÃO DE MODELOS LPV USANDO MÁQUINA DE SUPORTE VETORIAL

Marcelo Mendes Lafetá Lima ¹; Rodrigo Alvite Romano ²

¹ Aluno de Iniciação Científica da Escola de Engenharia Mauá (EEM/CEUN-IMT);

² Professor da Escola de Engenharia Mauá (EEM/CEUN-IMT).

Resumo. *Os modelos lineares com parâmetros variantes no tempo (LPV), são eficientes para descrever não-linearidades e sistemas variantes no tempo. Neste trabalho é descrito um algoritmo de identificação de modelos LPV usando máquinas de suporte vetorial. Esta técnica é usada para aprender a dependência entre os coeficientes do modelo e um sinal variante no tempo que caracteriza o ponto de operação do modelo. O algoritmo utilizará o estimador de variáveis instrumentais, que permite a obtenção de estimativas consistentes para condições gerais de ruído nos dados experimentais. O algoritmo ofereceu bons resultados sendo capaz de reduzir a variância das estimações para uma situação genérica de ruído nos dados. Também foi possível reduzir o viés do estimador para situações em que a dinâmica do ruído é mais rápida do que a dinâmica da planta.*

Introdução

A necessidade por técnicas de identificação cada vez mais robustas tem crescido com o avanço da capacidade de processamento e a busca por representações mais fiéis dos fenômenos observados. Porém, parte majoritária dos fenômenos industriais são tanto não lineares quanto ruidosos, o que justifica o uso de algoritmos de alta complexidade para sua identificação. Modelos de identificação de sistemas LPV (*linear parameter-varying*) são comumente vistos como eficientes abordagens para identificação de não linearidades e de sistemas variantes no tempo. Os modelos LPV apresentam relações lineares como os LTI (*linear time-invariant*), porém seus parâmetros são funções de uma variável mensurável variante no tempo denominada variável de comutação (p_k).

Muitos trabalhos demonstram a eficiência da técnica de regressão *Support Vector Machine* (SVM) para identificação de sistemas a partir de modelos LPV. Como descrito em [1] o algoritmo de identificação SVM se torna polarizado para sistemas com ruído branco na saída. Em [1] este problema é amenizado com o uso de filtros bi-dimensionais para sistemas em que o ruído é branco na saída, eliminando a polarização da máquina de suporte vetorial. Porém a premissa de que o ruído nos dados é branco se distancia da realidade.

A técnica de variáveis instrumentais já é consagrada para identificação de casos em que a natureza do ruído não é especificada ([2], [3] e [4]). Em [2] a técnica de variáveis instrumentais é utilizada para o tratamento de modelos com ruído colorido na saída, mostrando resultados efetivos na eliminação da polarização do estimador. O uso da técnica de variáveis instrumentais implementada junto a técnica LS-SVM (*least-squares support vector machine*) [4], gera um estimador sem polarização, porém, como visto em [5], com variância nos resultados maior que a técnica LS-SVM. Por mais que o estimador fique com maior representatividade devido a eliminação da polarização, uma alta variância não é aceitável.

Com o objetivo de desenvolver um estimador não polarizado para ruído colorido na saída e com a variância dos resultados reduzida, neste trabalho desenvolve-se um modelo em espaços de estados aplicando a máquina de suporte vetorial, juntamente com a filtragem bi-dimensional elaborada em [1] e a técnica de variáveis instrumentais descrita em [2]. O algoritmo aqui desenvolvido é denominado IV-filter.

Material e Métodos

Problema proposto

Dado o sistema de tempo discreto representado em espaços de estados

$$x_{k+1} = (A + L(p_k))x_k + B(p_k)u_k \quad (1)$$

$$y_k = Cx_k \quad (2)$$

em que $x_k \in \mathbb{R}^{n_x}$ representa o vetor de estados, $u_k \in \mathbb{R}^{n_u}$ as entradas do sistema, $y_k \in \mathbb{R}$ a saída e $p_k: \mathbb{Z} \rightarrow \mathbb{R}^{n_p}$, chamada de variável de comutação, é conhecida e depende do ponto de operação. As matrizes A e C são constantes e escolhidas de forma arbitrária desde que o par (A, C) seja observável e os autovalores da matriz A sejam estáveis. As matrizes $L(p_k)$ e $B(p_k)$ são definidas através de funções de base f_r arbitrárias que assumem valores de acordo com p_k

$$L(p_k) = \sum_{r=1}^{n_f} L_r f_r(p_k),$$

$$B(p_k) = \sum_{r=1}^{n_f} B_r f_r(p_k).$$

Em que L_r e B_r são pesos dados as funções de base.

Visto n_x como a ordem de regressão do modelo apresentado, tem-se que A pode ser composta da seguinte maneira

$$A = \begin{bmatrix} 0 & \cdots & 0 & -\alpha_{n_x} \\ & & & -\alpha_{n_x-1} \\ & I_{n_x-1} & & \vdots \\ & & & -\alpha_1 \end{bmatrix}, \quad (3)$$

de forma a caracterizar o modelo de regressão, e garantindo que o par (A, C) seja observável torna-se conveniente para a escolha de parâmetros a matriz C na forma

$$C = [0 \quad \cdots \quad 0 \quad 1]. \quad (4)$$

Para

$$\mathcal{A} \triangleq \text{bloco diagonal } \{A^T, \dots, A^T\} \in \mathbb{R}^{2n_x n_f \times 2n_x n_f},$$

$$\mathcal{B} \triangleq \text{bloco diagonal } \{C^T, \dots, C^T\} \in \mathbb{R}^{2n_x n_f \times 2n_f},$$

$$F(p_k) \triangleq [f_1(p_k) \quad \cdots \quad f_{n_f}(p_k)]^T \in \mathbb{R}^{n_f},$$

e sendo \otimes o produto de Kronecker, é demonstrado em [1] que o modelo descrito em (1) e (2) pode apresentar um preditor na forma

$$\varphi_{k+1} = \mathcal{A} \varphi_k + \mathcal{B} \left(\begin{bmatrix} y_k \\ u_k \end{bmatrix} \otimes F(p_k) \right) \quad (5)$$

$$\hat{y}_k = \theta^T \varphi_k,$$

com $\theta = \text{vec} \left(\begin{bmatrix} L_1 & \cdots & L_{n_f} & B_1 & \cdots & B_{n_f} \end{bmatrix} \right) \in \mathbb{R}^{2n_x n_f}$, sendo $\text{vec}(\cdot)$ é o operador que coloca as colunas do argumento uma sobre a outra.

Em [1] é provado a partir da proposição acima que

$$\varphi_k = (qI_{2n_x n_f} - \mathcal{A})^{-1} \mathcal{B} \left(\begin{bmatrix} y_k \\ u_k \end{bmatrix} \otimes F(p_k) \right) = \frac{q^{n_x}}{\alpha(q)} (I_{2n_f} \otimes [q^{-n_x} \quad \cdots \quad q^{-1}]^T) \left(\begin{bmatrix} y_k \\ u_k \end{bmatrix} \otimes F(p_k) \right),$$

que reescrito assume a forma

$$\varphi_k = \frac{q^{n_x}}{\alpha(q)} [f_1(p_{k-n_x})y_{k-n_x} \quad \cdots \quad f_1(p_{k-1})y_{k-1} \quad \cdots \quad f_{n_f}(p_{k-n_x})y_{k-n_x} \quad \cdots \quad f_{n_f}(p_{k-1})y_{k-1} \\ f_1(p_{k-n_x})u_{k-n_x} \quad \cdots \quad f_1(p_{k-1})u_{k-1} \quad \cdots \quad f_{n_f}(p_{k-n_x})u_{k-n_x} \quad \cdots \quad f_{n_f}(p_{k-1})u_{k-1}]^T.$$

Assim

$$P\varphi_k = \frac{q^{n_x}}{\alpha(q)} \phi_k,$$

em que P é a matriz de permutação responsável por representar de forma compacta o modelo de regressão e sendo $z_k \triangleq [y_k \ u_k]^T$

$$\phi_k \triangleq [F^T(p_{k-n_x}) \otimes z_{k-n_x}^T \dots F^T(p_{k-1}) \otimes z_{k-1}^T]^T.$$

Como demonstrado em [1], é possível determinar o modelo de regressão proposto

$$\hat{Y} = \frac{q^{n_x}}{\alpha(q)} \Phi P \theta, \quad (6)$$

em que $\alpha(q)$ é o polinômio característico de A , $\hat{Y} \in \mathbb{R}^{N-n_x}$ é o vetor de saída estimado e $\Phi \in \mathbb{R}^{N-n_x \times 2n_x n_f}$ é um vetor de regressores dado por $\Phi = [\phi_{n_x+1} \dots \phi_N]^T$.

Estimação de parâmetros

Com $Y \in \mathbb{R}^{N-n_x}$ sendo a saída conhecida do sistema, descreve-se $E \in \mathbb{R}^{N-n_x}$, como o erro de estimação

$$E = Y - \hat{Y} = Y - \frac{q^{n_x}}{\alpha(q)} \Phi P \theta; \quad (7)$$

logo, pode-se formular a função custo descrita em [2], $J \in \mathbb{R}^+$, dada por

$$J = \frac{1}{2} \theta^T \theta + \frac{\gamma}{2N^2} \|\Gamma^T E\|_2^2, \quad (8)$$

$$\Gamma = \frac{1}{\alpha(q)} \mathcal{S}, \quad (9)$$

em que os instrumentos são dados por $\mathcal{S} \in \mathbb{R}^{N-n_x \times 2n_x n_f}$, e $\Gamma \in \mathbb{R}^{N-n_x \times 2n_x n_f}$ é o vetor de instrumentos filtrados. O parâmetro $\gamma \in \mathbb{R}$ representa o valor que atribui peso à relação entre polarização do estimador e variância dos resultados. Sendo que a matriz de instrumentos é determinada na forma

$$\mathcal{S} = [\zeta_{n_x+1} \dots \zeta_N]^T$$

sendo

$$\zeta_k \triangleq \left[F^T(p_{k-n_x}) \otimes \begin{bmatrix} \hat{y}_{k-n_x} \\ u_{k-n_x} \end{bmatrix} \dots F^T(p_{k-1}) \otimes \begin{bmatrix} \hat{y}_{k-1} \\ u_{k-1} \end{bmatrix} \right]^T.$$

De [2] é possível determinar a seguinte condição de otimização para os parâmetros do modelo, sujeito a (7),

$$\hat{\theta} = \arg_{\theta} \min \frac{1}{2} \theta^T \theta + \frac{\gamma}{2N^2} \left\| \frac{1}{\alpha(q)} \mathcal{S}^T E \right\|_2^2. \quad (10)$$

Variáveis instrumentais

A técnica de variáveis instrumentais consiste em eliminar a polarização do estimador a partir da substituição das variáveis de regressão do modelo por instrumentos. Tais instrumentos são escolhidos visando minimizar a polarização \mathbf{b} do estimador. De modo que dado um modelo genérico descrito por $\mathbf{y} = \Phi \theta + \mathbf{e}$, estimado por $\hat{\theta} = A_b \mathbf{y}$, em que \mathbf{y} é a saída do modelo, Φ são os regressores do modelo, θ_b são os parâmetros do modelo, \mathbf{e} é o ruído presente nos dados e $\hat{\theta}_b$ é uma variável aleatória que corresponde aos parâmetros estimados. Tendo $E[\]$ como o operador esperança, a polarização, \mathbf{b} , pode ser caracterizada por

$$\begin{aligned} \mathbf{b} &= \lim_{N \rightarrow \infty} E[\hat{\theta}] - \theta \\ \mathbf{b} &= \lim_{N \rightarrow \infty} E[A_b(\Phi \theta + \mathbf{e})] - \theta \\ \mathbf{b} &= \lim_{N \rightarrow \infty} E[A_b \Phi \theta + A_b \mathbf{e}] - \theta \\ \mathbf{b} &= \lim_{N \rightarrow \infty} E[A_b \Phi - I] \theta - E[A_b \mathbf{e}]. \end{aligned} \quad (11)$$

De (11), é possível afirmar que \mathbf{b} somente é nulo para as situações em que $A_b \boldsymbol{\phi} = I$, a média $E[\mathbf{e}] = 0$ e A_b é descorrelacionada com o ruído \mathbf{e} , o que garante $E[A_b \mathbf{e}] = 0$.

A escolha dos instrumentos baseia-se na determinação da matriz A_b de forma a garantir a não correlação entre A_b e o ruído (\mathbf{e}). Considerando o estimador como o de mínimos quadrados, é possível determinar $A_b = [\mathcal{S}_b^T \boldsymbol{\phi}]^{-1} \mathcal{S}_b^T$, onde \mathcal{S}_b^T é o vetor de instrumentos,

$$\begin{aligned} \mathbf{b} &= \lim_{N \rightarrow \infty} [(\mathcal{S}_b^T \boldsymbol{\phi})^{-1} \mathcal{S}_b^T \mathbf{e}] = \lim_{N \rightarrow \infty} [(\mathcal{S}_b^T \boldsymbol{\phi})^{-1}] \lim_{N \rightarrow \infty} [\mathcal{S}_b^T \mathbf{e}], \\ \mathbf{b} &= \lim_{N \rightarrow \infty} [(\mathcal{S}_b^T \boldsymbol{\phi})] \mathbf{c}, \end{aligned} \quad (12)$$

em que \mathbf{c} é a matriz de correlação cruzada entre o vetor de instrumentos e o ruído. Vale constatar que o ruído descrito não é *a priori* branco e, de acordo com (12), a técnica de variáveis instrumentais garante a não polarização do estimador desde que a correlação cruzada entre os instrumentos e o ruído seja nula. Logo é possível validar a escolha de filtragem dos instrumentos em (9) que objetiva descorrelacionar o vetor de instrumentos \mathcal{S} do ruído presente na saída.

Máquinas de Suporte Vetorial

A fim de resolver (10), desenvolve-se a Lagrangiana, em que $\lambda \in \mathbb{R}^{N-n_x}$ é o vetor multiplicador de Lagrange,

$$\mathcal{L}(E, \theta, \lambda) = \frac{1}{2} \theta^T \theta + \frac{\gamma}{2N^2} \left(E^T \left[\frac{q^{n_x}}{\alpha(q)} \mathcal{S} \mathcal{S}^T \frac{q^{n_x}}{\alpha(q)} \right] E \right) - \lambda^T \left(E - Y + \frac{q^{n_x}}{\alpha(q)} \Phi P \theta \right), \quad (13)$$

e considerando as condições de Karush-Kuhn-Tucker (KKT)[1] para a Lagrangiana obtém-se

$$\frac{\partial \mathcal{L}}{\partial E} = \frac{\gamma}{N^2} \frac{q^{n_x}}{\alpha(q)} \mathcal{S} \mathcal{S}^T \frac{q^{n_x}}{\alpha(q)} E - \lambda = 0, \quad (14)$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = \theta - P^T \theta^T \frac{q^{n_x}}{\alpha(q)} \lambda = 0, \quad (15)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = E - Y + \frac{q^{n_x}}{\alpha(q)} \Phi P \theta = 0. \quad (16)$$

A partir de manipulações algébricas é possível determinar a seguinte relação

$$\left[\frac{\gamma}{N^2} \frac{q^{n_x}}{\alpha(q)} \mathcal{S} \mathcal{S}^T \frac{q^{n_x}}{\alpha(q)} \frac{q^{n_x}}{\alpha(q)} \Phi \Phi^T \frac{q^{n_x}}{\alpha(q)} + I_{N-n_x} \right] \lambda = \frac{\gamma}{N^2} \frac{q^{n_x}}{\alpha(q)} \mathcal{S} \mathcal{S}^T \frac{q^{n_x}}{\alpha(q)} Y. \quad (17)$$

Sendo $K_{\mathcal{S}\mathcal{S}} = \frac{q^{n_x}}{\alpha(q)} \mathcal{S} \mathcal{S}^T \frac{q^{n_x}}{\alpha(q)}$ e $K_{\Phi\Phi} = \frac{q^{n_x}}{\alpha(q)} \Phi \Phi^T \frac{q^{n_x}}{\alpha(q)}$, é possível determinar os multiplicadores de Lagrange por

$$\lambda = \left(\frac{1}{N^2} K_{\mathcal{S}\mathcal{S}} K_{\Phi\Phi} + \frac{I_{N-n_x}}{\gamma} \right)^{-1} \frac{1}{N^2} K_{\mathcal{S}\mathcal{S}} Y, \quad (18)$$

em que as matrizes $\Phi \Phi^T \in \mathbb{R}^{N-n_x \times N-n_x}$ e $\mathcal{S} \mathcal{S}^T \in \mathbb{R}^{N-n_x \times N-n_x}$, são desconhecidas. A fim de determinar algo que represente $\Phi \Phi^T$, o termo da linha i e coluna j é dado por:

$$\begin{aligned} [\Phi \Phi^T]_{i,j} &= \phi_{i+n_x}^T \phi_{j+n_x} \\ [\Phi \Phi^T]_{i,j} &= (F^T(p_i) \otimes z_i^T) (F(p_j) \otimes z_j) + (F^T(p_{i+n_x-1}) \otimes z_{i+n_x-1}^T) (F(p_{j+n_x-1}) \otimes z_{j+n_x-1}) \\ [\Phi \Phi^T]_{i,j} &= \sum_{m=0}^{n_x-1} (F^T(p_{i+m}) \otimes z_{i+m}^T) (F(p_{j+m}) \otimes z_{j+m}) = \sum_{m=0}^{n_x-1} F^T(p_{i+m}) F(p_{j+m}) z_{i+m}^T z_{j+m}. \end{aligned}$$

Assim, pelo denominado kernel-trick [6], o produto interno $F^T(p_{i+m}) F(p_{j+m})$ desconhecido pode ser caracterizado da seguinte forma

$$[\Phi \Phi^T]_{i,j} \triangleq \sum_{m=0}^{n_x-1} \Psi(p_{i+m}, p_{j+m}) z_{i+m}^T z_{j+m}. \quad (19)$$

O termo $\Psi(p_i, p_j)$ é uma função de kernel utilizada para representar o produto interno das funções $F^T(p_i)F(p_j)$. Com isso, fica claro a necessidade da filtragem bi-dimencional, ou seja, das linhas e colunas da matriz caracterizada pelo produto $\Phi\Phi^T$, visto que somente é possível representar o produto destas matrizes.

Neste trabalho foram utilizadas as funções de base radial gaussianas devido sua capacidade de representação, com $\sigma \in \mathbb{R}_+^*$

na representação de produto interno das funções por

$$\Psi(p_i, p_j) = \exp\left(-\frac{\|p_i - p_j\|_2^2}{\sigma^2}\right). \quad (20)$$

O mesmo pode ser aplicado para o produto interno do vetor de instrumentos $\mathcal{S}\mathcal{S}^T$. O par de parâmetros (γ, σ) é conhecido como hyper-parâmetros, os quais são determinados de forma livre.

Escolha dos filtros

A partir de (5) é possível perceber que a matriz A é responsável pela dinâmica do preditor, sendo assim seus autovalores podem ser capazes de filtrar o ruído dos dados. Com isto, os filtros determinados por $\alpha(q)$ foram escolhidos a partir da arquitetura dos filtros de *Butterworth*, como descrito em [1], com

$$s_m = \omega_c e^{j\frac{2m+n_x-1}{2n_x}}, \quad (21)$$

sendo ω_c a frequência de corte do filtro. Para a filtragem bi-dimencional é necessário escolher os autovalores da matriz A (3) que fazem a sintonia do filtro responsável por mitigar o efeito do ruído de alta frequência presente nos dados.

A estrutura de $\alpha(q)$ como os polos do filtro de *Butterworth* foi utilizada devido ao fato de que ω_c é uma grandeza física intuitiva capaz de trazer informações como a natureza do ruído presente no sistema e características interpretativas do modelo.

Para os valores candidatos de frequência $\Omega = \{\omega_1, \dots, \omega_{n_\omega}\}$, foi utilizado um processo de otimização sem o uso de derivadas para estimar a frequência de corte do filtro. Para cada frequência foi determinado uma função custo $J(\omega_c)$ que indica a representatividade do modelo, para $\omega_c \in \Omega$. A partir da técnica de baricentro [7]

$$\omega_c^* = \frac{\sum_{v=1}^{n_\omega} \omega_v e^{-\mu J(\omega_v)}}{\sum_{v=1}^{n_\omega} e^{-\mu J(\omega_v)}}, \quad (22)$$

sendo que $e^{-\mu J(\omega_v)}$ é responsável por dar peso maior aos termos de melhor performance. O termo $\mu \in \mathbb{R}_+^*$ é usado como peso para controlar os valores da função custo. Para valores altos de μ , mais o baricentro (ω_c^*) tende ao valor de Ω que minimiza $J(\omega_c)$ e promove a melhor representatividade.

Resultados e Discussão

É proposta a seguinte planta de resolução relacionada ao *Aström system* [8],

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} a_{11}(p_k) & 1 \\ a_{21}(p_k) & 0 \end{bmatrix} x_k + \begin{bmatrix} b_1(p_k) \\ b_2(p_k) \end{bmatrix} u_k \\ y_k &= [1 \ 0] x_k + v_k \end{aligned}$$

sendo

$$\begin{aligned} a_{11}(p_k) &= 0,35 \operatorname{sinc}(\pi^2 p_k) + 1,4 \\ a_{21}(p_k) &= 5p_k^2 - 0,8 \\ b_1(p_k) &= \begin{cases} 1,5 & , \text{ para } p_k > 0,125 \\ 1 + 4p_k & , \text{ para } |p_k| \leq 0,125 \\ 0,5 & , \text{ para } p_k < -0,125 \end{cases} \\ b_2(p_k) &= \begin{cases} 0 & , \text{ para } p_k > 0,125 \\ 0,5 - 4p_k & , \text{ para } |p_k| \leq 0,125 \\ 1 & , \text{ para } p_k < -0,125. \end{cases} \end{aligned}$$

Foram criadas três condições de espectro de ruído: ruído branco determinado por $v_o(k) = e(k)$, um ruído de banda estreita representado em (23), e um de banda larga em (24),

$$v_1(k) = \frac{1}{1-0,6 z^{-1}} e(k), \quad (23)$$

$$v_2(k) = \frac{1}{1-0,9 z^{-1}} e(k). \quad (24)$$

Produziram-se ensaios de Monte Carlo com 100 rodadas para cada condição de ruído em cada valor de SNR (0, 5, 10, 15 e 20 dB). Foi utilizado um sistema com $N = 1000$ amostras, a entrada $u(k)$ é um ruído branco binário com média nula e $p(k) = [u(k) \ y(k)]$, o que caracteriza um modelo *quasi*-LPV. Analisou-se a representatividade de quatro diferentes algoritmos para o sistema descrito: o modelo LS-SVM [1] onde os polos do filtro se encontram na origem do plano Z (não há filtragem); o modelo 2D-filter proposto em [1]; o método IV-SVM, com $\alpha(q) = 1$ em (9), referente a abordagem descrita em [4]; e o modelo IV-filter onde os instrumentos são sujeitos ao mecanismo de filtragem.

A planta utilizada é artificial, logo o objetivo não foi de encontrar o modelo que melhor represente os parâmetros a_{11} , a_{21} , b_1 e b_2 , mas sim reproduzir o comportamento da saída sem ruído (\mathbf{y}) para uma dada entrada (\mathbf{u}). Para avaliar a capacidade assertiva do modelo utilizou-se o critério de *best fit rate* (BFR), pois retorna a representatividade do modelo a partir dos valores estimados,

$$\text{BFR}(\%) = 100 \cdot \max \left(1 - \frac{\|Y - \hat{Y}\|_2}{\|Y - \bar{Y}\|_2}, 0 \right), \quad (25)$$

definindo como Y o vetor de saídas real da planta, \hat{Y} o vetor de valores estimados pelo modelo e \bar{Y} a média comum do vetor de saídas real da planta, e de modo que seu complementar

$$J(\omega_c) = \min \left(\frac{\|Y - \hat{Y}\|_2}{\|Y - \bar{Y}\|_2}, 1 \right) \quad (26)$$

é utilizado como função custo para o cálculo da frequência de corte a partir do baricentro (22) para o conjunto de frequências candidatas $\Omega = \{0.05, \dots, 1.0\}\pi$ (rad/s). Para a determinação dos hiperparâmetros γ e σ , foi utilizada uma busca em grade com validação cruzada a fim do par ótimo (γ^* , σ^*) que minimiza a função custo (26), para cada um dos modelos abordados. Levando em conta as assertivas descritas foi possível determinar a tabela com os seguintes resultados para cada natureza e intensidade de ruído.

Tabela 1 - Referente aos valores de BFR e desvio padrão para ruído branco, na forma: BFR (Desvio padrão).

Método	0 dB	5 dB	10 dB	15 dB	20 dB
LS-SVM	31.91 (2.51)	39.94 (2.04)	48.05 (2.05)	62.51 (2.06)	79.73 (1.51)
2D-filter	71.44 (3.33)	83.91 (2.27)	91.79 (1.31)	95.34 (0.69)	97.18 (0.37)
IV-SVM	30.10 (20.7)	57.82 (19.9)	79.22 (5.39)	88.43 (1.87)	93.00 (1.15)
IV-filter	79.25 (3.91)	86.58 (1.78)	91.16 (1.97)	86.40 (2.16)	94.48 (8.70)

Tabela 2 - Referente aos valores de BFR e desvio padrão para ruído de banda larga, na forma: BFR (Desvio padrão).

Método	0 dB	5 dB	10 dB	15 dB	20 dB
LS-SVM	43.20 (2.75)	51.05 (2.58)	63.93 (2.55)	80.77 (1.93)	91.58 (1.04)
2D-filter	66.99 (4.56)	80.62 (2.96)	90.06 (1.43)	95.01 (0.75)	97.12 (0.41)
IV-SVM	60.83 (15.6)	78.84 (4.35)	87.07 (1.76)	93.05 (1.07)	95.60 (0.63)
IV-filter	77.40 (3.82)	84.37 (2.31)	88.55 (1.85)	93.56 (1.30)	95.19 (0.96)

Tabela 3 - Referente aos valores de BFR e desvio padrão para ruído de banda estreita, na forma: BFR (Desvio padrão).

Método	0 dB	5 dB	10 dB	15 dB	20 dB
LS-SVM	43.86 (2.97)	65.11 (2.59)	82.53 (1.92)	92.35 (1.07)	96.53 (0.48)
2D-filter	50.95 (4.84)	74.56 (3.18)	87.87 (1.91)	94.73 (0.84)	96.73 (0.42)
IV-SVM	74.57 (6.32)	85.67 (2.49)	92.25 (1.24)	95.59 (0.66)	97.29 (0.35)
IV-filter	76.77 (5.35)	85.93 (2.43)	91.96 (1.35)	95.50 (0.84)	97.04 (0.54)

A partir das Tabelas 1, 2 e 3, é possível inferir que o método IV-filter apresenta maior acurácia (BFR) para situações com baixo SNR quando comparado com o método 2D-filter. Fica claro que para valores de SNR maiores que 10 dB a abordagem 2D-filter é vantajosa, enquanto que para valores menores que 5 dB o modelo IV-filter é mais propício, devido à relação entre tempo de processamento computacional e representatividade (BFR).

Para os casos de ruído branco (Tabela 1) e ruído de banda larga (Tabela 2) a variância nos resultados diminui do modelo IV-SVM para o modelo IV-filter. Isso se dá devido a filtragem dos instrumentos em (9). Portanto para situações de SNR menores que 5 dB, é possível afirmar que o aumento da variância nos resultados descrito em [5] pôde ser amenizado.

O mesmo não é evidenciado para ruído de banda estreita devido a estrutura do filtro passa-baixas descrito em [1]. A estrutura de filtragem passa-baixas pode eliminar espectros pertencentes à planta na tentativa de eliminar o ruído, sendo necessário o desenvolvimento de outra estrutura de filtragem mais adequada para o problema. O desenvolvimento de um filtro capaz de viabilizar a filtragem de ruídos com dinâmica parecida à da planta acontece em paralelo com este trabalho.

Para casos de banda estreita, a filtragem dos instrumentos não é uma escolha favorável devido a diferença não significativa entre os resultados do modelo 2D-filter e IV-filter. Porém, é possível observar na Tabela 3 que a utilização da técnica de variáveis instrumentais pode ser uma abordagem propícia. Considerando tanto a acurácia quanto a variância dos resultados a técnica de variáveis instrumentais se torna uma interessante abordagem para sistemas em que o ruído é banda estreita.

Afim de ressaltar a vantagem da utilização da abordagem IV-filter, nas Figuras 1, 2 e 3 abaixo, foram montados os histogramas para as situações das Tabelas 1, 2 e 3, respectivamente, em que o valor de SNR vale 0 dB.

Figura 1 - Histograma dos resultados do modelo para ruído branco com SNR = 0 dB.

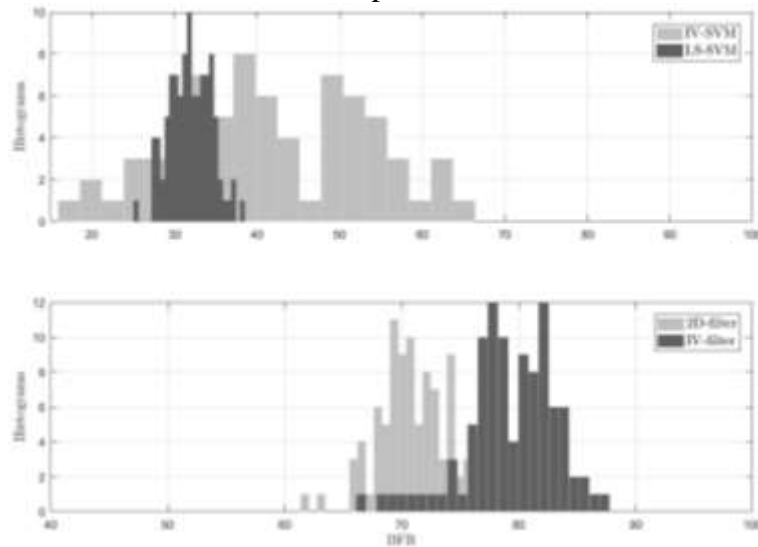


Figura 2 - Histograma dos resultados do modelo para ruído de banda larga com SNR = 0 dB.

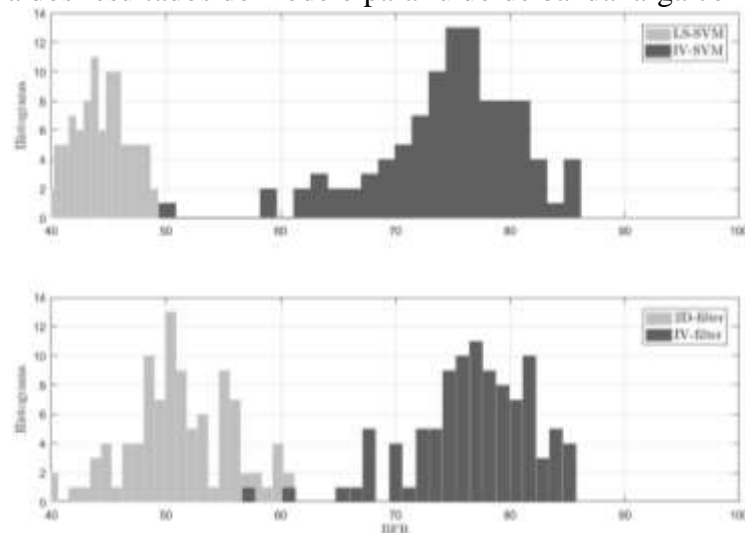
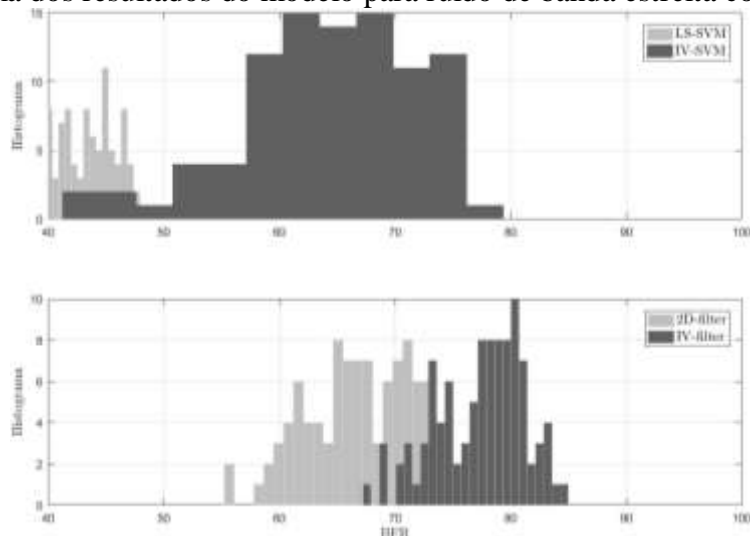


Figura 3 - Histograma dos resultados do modelo para ruído de banda estreita com SNR = 0 dB.



A partir das figuras, fica evidente a redução de variância existente sobre a técnica de variáveis instrumentais quando a filtragem dos instrumentos é utilizada. Também fica clara a ineficiência da filtragem para os casos em que o ruído é de banda estreita. Logo, a filtragem mostrou-se interessante para situações em que a dinâmica do ruído é mais rápida que a da planta. Já a técnica de variáveis

instrumentais é eficiente para situações de valores baixos de SNR (alta influência do ruído nos dados), independente da condição de ruído presente nos dados.

Conclusão

Este trabalho apresenta um algoritmo de identificação de sistemas que utiliza a técnica de variáveis instrumentais em que os instrumentos são filtrados pela abordagem descrita em [1]. A técnica busca ser capaz de eliminar a polarização do modelo proposto em [1] em ambientes ruidosos nos quais a natureza do ruído não é necessariamente branca. Foi possível estimar um modelo no espaço de estados, robusto a ruído de natureza genérica na saída. Aplicado em uma planta de simulação, o modelo IV-filter, proposto por este trabalho, possibilitou melhoras tanto de representatividade quanto a redução da variância dos resultados para plantas com ruído colorido na saída. Foram comparados os resultados dos modelos LS-SVM e 2D-filter descritos em [1], o modelo IV-SVM [4] e o modelo IV-filter. A análise dos resultados evidencia que para ruído branco e de banda larga a técnica IV-filter se demonstra eficaz, enquanto para ruído de banda estreita, a utilização do modelo IV-SVM é mais atrativa. Este algoritmo foi implementado para a solução de problemas de *flutter* na área de aerodinâmica para fins literários revelando bons resultados. É almejado o desenvolvimento de um *paper* para um congresso internacional de automática, para o qual será mais detalhado o desenvolvimento computacional dos algoritmos. Podem existir desenvolvimentos seguintes a este trabalho que consiste no aumento da liberdade dos filtros, tornando-os variantes no tempo, assim como o aumento da complexidade da estrutura de filtragem.

Referências Bibliográficas

- [1] Romano, R.A.; dos Santos, P.L.; Pait, F.; Perdicoulis, T.A.; Ramos, J. (2016). Machine Learning Barycenter Approach to Identifying LPV State-Space Models. In *American Control Conference (ACC)*, 6351-6356.
- [2] Laurain, V.; Tóth, R.; Piga, D.; Zheng, W.X. (2015). An instrumental least squares support vector machine for nonlinear system identification. In *Automatica*, **54**, 340-347.
- [3] Laurain, V. (2012). Nonparametric Identification of LPV Models under General Noise Conditions: An LS-SVM Based Approach. In *IFAC Proceedings Volumes*, **45(16)**, 1761-1766.
- [4] Rizvi, S.Z.; Velni, J.M.; Tóth, R.; Meskin, N. (2015). An IV-SVM-based Approach for Identification of State-Space LPV Models under Generic Noise Conditions. In *Decision and Control (CDC), IEEE*, **54**, 7380-7385.
- [5] Laurain, V.; Gilson, M.; Tóth, R.; Garnier, H. (2010). Refined instrumental variable methods for identification of LPV Box-Jenkins models. In *Automatica*, **46**, 959-967.
- [6] Vapnik, V. (1998). *Statistical Learning Theory*. New York: Wiley-Interscience.
- [7] Pait, F.; Colón, D. (2014). A barycenter method for direct optimization. In *SIAM Conference on Optimization*.
- [8] Ljung, L. (1999). *System Identification: theory for the user*, 2^a Ed. Upper Saddle River.