ISBN 978-981-11-0008-6
Proceedings of 2016 6th International Workshop on Computer Science and Engineering (WCSE 2016)
Tokyo, Japan, 17-19 June, 2016, pp. 710-714
doi: 10.18178/wcse.2016.06.126

Dynamic Modeling on the MIMO System with Linear Programming Support Vector Regression and Its Application

Yanfang Yu+ and Li Jiang

Shanghai Institute of Technology, Shanghai 201418, China

Abstract. Few researches have been done for dynamic modeling on the multiple-input multiple-output system in actual industry so far. Therefore, this paper proposed a novel dynamic model to solve the multiple-input multiple-output problem. It respectively mapped the delayed output and the primal input into different nonlinear space, associated the kernel trick with the support vector regression technique to estimate the nonlinear dynamic problem. Simulation results of dynamic estimation of COOH content and degree of polymerization in the poly ethylene terephthalate production demonstrated that the present algorithm exactly approached the output of test set and reduced the error of test set respectively to the magnitude of 10^{-7} and 10^{-8}.

Keywords: linear programming support vector regression, dynamic modeling, kernel trick.

1. Introduction

Support vector machines are originally developed for pattern recognition by Vapnik [1], and later they are carried over to the case of support vector regression (SVR) [2] which conduces to the problem of function estimation and data modeling. Generally, SVR is readily used for establishing steady-state model [3, 4, 5, 6] of nonlinear system [7, 8]. However, actual industrial processes such as poly ethylene terephthalate (PET) production changes over time. Therefore, dynamic modelling on the industrial processes could constantly monitor the production output. So far, there were few studies on the dynamic model of SVR, and most of them were based on the single-input single-output (SISO) system [9, 10]. The ideas of these dynamic models were usually simple. That is, the original input variables and delayed feedback output together constituted the inputs of the dynamic model. But for multiple-input multiple-output (MIMO) or multi-input single-output (MISO) system, it is undesirable to carry on with the thought. Goethals [11] presented the MIMO dynamic model which was composed of the linear mapping of feedback output and the nonlinear mapping of the original input. The model had difficulty in many parameters selection and problem solving. Feng [12] also developed a dynamic model on the MIMO system. The algorithm regarded the output and each input attribute as the inputs of SVR, and had the same form as the SISO system. However, it became more complex due to multiple input attributes.

Based on the above consideration, this paper proposed a novel approach for dynamic modeling of the MIMO or MISO system. Respectively, it mapped the original input variables and the delayed feedback output into different nonlinear space, and then regarded their linear combination as the objective dynamic model. Besides, it also applied v-linear programming SVR (v-LPSVR) [13] to the primal problem, which restricted the target function to the linear function and simplified the problem solving. The experimental results of PET production indicated that the dynamic model of COOH content and degree of polymerization using the proposed method basically approximated the actual output of test set, and respectively restricted the mean squared error (MSE) of test sets to the magnitude of 10^{-7} and 10^{-8}.

+ Corresponding author. Tel.: + 86-02160873338; fax: + 86-02160873338.
E-mail address: yuyf@sit.edu.cn

2. Dynamic Modeling on the MIMO System

In a SISO system, the linear dynamic model of the autoregressive model with exogenous inputs (ARX)[9] form is as follows,

$$y_t = \sum_{i=1}^{n} a_i y_{t-i} + \sum_{j=0}^{m} b_j u_{t-j} \tag{1}$$

where u_t , y_t $\in$R, and respectively represent the original input and output variable. And m and n are respectively the delayed orders of the model. We can reconstruct the input variable $x_t = [y_{t-1}, \cdots, y_{t-n}, u_t, u_{t-1}, \cdots, u_{t-m}]$ of the SVR algorithm and then establish a dynamic model on the SISO system.

2.1. Dynamic model for the MIMO system

In the MIMO system, we propose to map the original input variables and the delayed feedback output into different nonlinear space. The dynamic model is written by

$$\begin{aligned}
& y_t = f_1(\vec{y}_t) + f_2(\vec{x}_t), \\
& f_1(\vec{y}_t) = w_1\varphi_1(\vec{y}_t) + b_1, \quad \vec{y}_t = [y_{t-1}, \cdots, y_{t-n}]^T, \\
& f_2(\vec{x}_t) = w_2\varphi_2(\vec{x}_t) + b_2, \quad b = b_1 + b_2, \\
& w_2\varphi_2(\vec{x}_t) = \textstyle\sum_{j=0}^{m} \omega_j \phi(u_{t-j}), \quad \vec{x}_t = [\phi(u_t), \phi(u_{t-1}), \cdots, \phi(u_{t-m})]^T,
\end{aligned} \tag{2}$$

where b, y_t $\in$R, and $u_t \in R^d$. Thus, the objective function is represented by

$$\begin{aligned}
\min \ & \frac{1}{2}\| w_1 \|^2 + \frac{1}{2}\| w_2 \|^2 + \frac{C_1}{\ell - r + 1}\sum_{t=r}^{\ell}(\xi_{1t} + \xi_{1t}^*) + \frac{C_2}{\ell - r + 1}\sum_{t=r}^{\ell}(\xi_{2t} + \xi_{2t}^*) \\
s.t. \ & w_1\varphi_1(\vec{y}_t) + b - y_t \le \xi_{1t} + \varepsilon_1, \\
& y_t - w_1\varphi_1(\vec{y}_t) - b \le \xi_{1t}^* + \varepsilon_1, \\
& w_2\varphi_2(\vec{x}_t) - (y_t - w_1\varphi_1(\vec{y}_t) - b) \le \xi_{2t} + \varepsilon_2, \\
& y_t - w_1\varphi_1(\vec{y}_t) - b - w_2\varphi_2(\vec{x}_t) \le \xi_{2t}^* + \varepsilon_2, \\
& \xi_{1t}, \xi_{1t}^*, \xi_{2t}, \xi_{2t}^*, t = r, \cdots, \ell, \\
& r = \max(m, n) + 1
\end{aligned} \tag{3}$$

The constant C_1, C_2 involve the trade-off between target function and deviation. ℓ denotes the size of training set, m, n is respectively the delayed input and output order, the slack variables ξ_1, ξ_2 with and without the asterisk are tolerated deviations from the tube width ε_1, ε_2. Then we construct the Lagrangian L of the objective problem, and give the conditions for optimality as follows,

$$\begin{aligned}
& \frac{\partial L}{\partial w_1} = 0 \to w_1 = \sum_{t=r}^{\ell}(\alpha_t^* - \alpha_t + \beta_t^* - \beta_t)\varphi_1(\vec{y}_t) \\
& \frac{\partial L}{\partial w_2} = 0 \to w_2 = \sum_{t=r}^{\ell}(\beta_t^* - \beta_t)\varphi_2(\vec{x}_t)
\end{aligned} \tag{4}$$

where the dual variables α, β with and without the asterisk are Lagrange multipliers.

2.2. Dynamic model based on ν-LPSVR

The ν-LPSVR approach views the insensitive tube width ε as a constrained variable, automatically determines the size of ε to fit the data, and effectively controls the number of support vectors and the fraction of errors. Then the above optimization problem is expressed by,

$$
\begin{aligned}
&\min \frac{1}{\ell-r+1}(\alpha_t^*+\alpha_t+\beta_t^*+\beta_t)+\frac{1}{\ell-r+1}(\beta_t^*+\beta_t)+\frac{C_1}{\ell-r+1}\sum\nolimits_{t=r}^{\ell}(\xi_{1t}+\xi_{1t}^*)+\frac{C_2}{\ell-r+1}\sum\nolimits_{t=r}^{\ell}(\xi_{2t}+\xi_{2t}^*)\\
&\quad +C_1 v_1\varepsilon_1+C_2 v_2\varepsilon_2\\
&s.t.\ \sum\nolimits_{t=r}^{\ell}(\alpha_t^*-\alpha_t+\beta_t^*-\beta_t)k_1(\vec{y}_t,\vec{y}_t)+b-y_t\le\xi_{1t}+\varepsilon_1,\\
&\quad y_t-\sum\nolimits_{t=r}^{\ell}(\alpha_t^*-\alpha_t+\beta_t^*-\beta_t)k_1(\vec{y}_t,\vec{y}_t)-b\le\xi_{1t}^*+\varepsilon_1,\\
&\quad \sum\nolimits_{t=r}^{\ell}(\beta_t^*-\beta_t)k_2(\vec{x}_t,\vec{x}_t)-(y_t-\sum\nolimits_{t=r}^{\ell}(\alpha_t^*-\alpha_t+\beta_t^*-\beta_t)k_1(\vec{y}_t,\vec{y}_t)-b)\le\xi_{2t}+\varepsilon_2,\\
&\quad y_t-\sum\nolimits_{t=r}^{\ell}(\alpha_t^*-\alpha_t+\beta_t^*-\beta_t)k_1(\vec{y}_t,\vec{y}_t)-b-\sum\nolimits_{t=r}^{\ell}(\beta_t^*-\beta_t)k_2(\vec{x}_t,\vec{x}_t)\le\xi_{2t}^*+\varepsilon_2,\\
&\quad \xi_{1t},\xi_{1t}^*,\xi_{2t},\xi_{2t}^*,t=r,\cdots,\ell,
\end{aligned}
\tag{5}
$$

Here the kernel function k_1 is the popular and powerful kernel function, Gaussian radial basis function kernel.

$$
k_1(\vec{y}_i,\vec{y}_j)=\exp(-\|\vec{y}_i-\vec{y}_j\|^2/2\sigma_1^2) \tag{6}
$$

And the kernel function k_2 is of the form,

$$
\begin{aligned}
k_2(\vec{x}_i,\ \vec{x}_j)&=\vec{x}_i^{\mathrm{T}}\cdot\vec{x}_j=[\phi(u_i),\phi(u_{i-1}),\cdots,\phi(u_{i-m})]\cdot[\phi(u_j),\phi(u_{j-1}),\cdots,\phi(u_{j-m})]^T\\
&=k(u_i,u_j)+\cdots+k(u_{i-m},u_{j-m})\\
&=\exp(-||u_i-u_j||^2/2\sigma_2^2)+\cdots+\exp(-||u_{i-m}-u_{j-m}||^2/2\sigma_2^2)
\end{aligned}
\tag{7}
$$

Then the linear programming problem is represented in the matrix form,

$$
\begin{aligned}
&\text{minimize } c^{\mathrm{T}}\cdot S\\
&\text{subject to } A\cdot S\le \mathrm{b}'\\
&S\ge[\vec{0},\vec{0},\vec{0},\vec{0},\vec{0},\vec{0},\vec{0},\vec{0},0,0,-\infty]\in R^{8(l-r+1)+3}
\end{aligned}
\tag{8}
$$

where

$$
\begin{aligned}
&c=[\frac{1}{l-r+1}\vec{1},\frac{1}{l-r+1}\vec{1},\frac{2}{l-r+1}\vec{1},\frac{2}{l-r+1}\vec{1},\frac{C_1}{l-r+1}\vec{1},\frac{C_1}{l-r+1}\vec{1},\frac{C_2}{l-r+1}\vec{1},\frac{C_2}{l-r+1}\vec{1},C_1v_1,C_2v_2,0]\in R^{8(l-r+1)+3}\\
&S=[\vec{\alpha},\vec{\alpha}^*,\vec{\beta},\vec{\beta}^*,\vec{\xi}_1,\vec{\xi}_1^{\,*},\vec{\xi}_2,\vec{\xi}_2^{\,*},\varepsilon_1,\varepsilon_2,b']\in R^{8(l-r+1)+3}\\
&K_1=\begin{bmatrix}k_1(\vec{y}_r,\vec{y}_r), & \cdots & ,k_1(\vec{y}_r,\vec{y}_l)\\ \vdots & \vdots & \vdots\\ k_1(\vec{y}_l,\vec{y}_r), & \cdots & ,k_1(\vec{y}_l,\vec{y}_l)\end{bmatrix}\quad K_2=\begin{bmatrix}k_1(\vec{y}_r,\vec{y}_r)+\vec{x}_r\cdot\vec{x}_r, & \cdots & ,k_1(\vec{y}_r,\vec{y}_l)+\vec{x}_r\cdot\vec{x}_l\\ \vdots & \vdots & \vdots\\ k_1(\vec{y}_l,\vec{y}_r)+\vec{x}_l\cdot\vec{x}_r, & \cdots & ,k_1(\vec{y}_l,\vec{y}_l)+\vec{x}_l\cdot\vec{x}_l\end{bmatrix}\\
&A=\begin{bmatrix}-K_1 & K_1 & -K_1 & K_1 & -I & \vec{\vec{0}} & \vec{\vec{0}} & \vec{\vec{0}} & -\vec{1} & \vec{0} & \vec{1}\\ K_1 & -K_1 & K_1 & -K_1 & \vec{\vec{0}} & -I & \vec{\vec{0}} & \vec{\vec{0}} & -1 & \vec{0} & -\vec{1}\\ -K_1 & K_1 & -K_2 & K_2 & 0 & \vec{\vec{0}} & -I & \vec{\vec{0}} & \vec{0} & -\vec{1} & \vec{1}\\ K_1 & -K_1 & K_2 & -K_2 & \vec{\vec{0}} & \vec{\vec{0}} & \vec{\vec{0}} & -I & \vec{0} & -\vec{1} & -\vec{1}\end{bmatrix}\in R^{4(l-r+1)\times[8(l-r+1)+3]}\\
&b'=[y,-y,y,-y]\in R^{4(l-r+1)}
\end{aligned}
\tag{9}
$$

Here, $\vec{1}$ and $\vec{0}$ are all (ℓ–r+1) ×1 column vectors, and K_1, K_2, $\vec{\vec{0}}$ and I are (ℓ–r+1) × (ℓ–r+1) matrices. Then one searches for the solution S via the simplex algorithm or the interior-point method.

3. Simulation Results

PET is a commodity product for many different applications. Many continuous processes are well established for the industrial production of PET. Here, we regarded reactor temperature, pressure and level as the input variables and COOH content and degree of polymerization as the output, and then developed a dynamic model of COOH content and degree of polymerization based on above technique. The dataset comprised 900 continuous instances with 5 percent of Gaussian noise. The first three-quarters of samples were training set, the remaining were test set. Some samples were illustrated as Fig. 1 and 2.

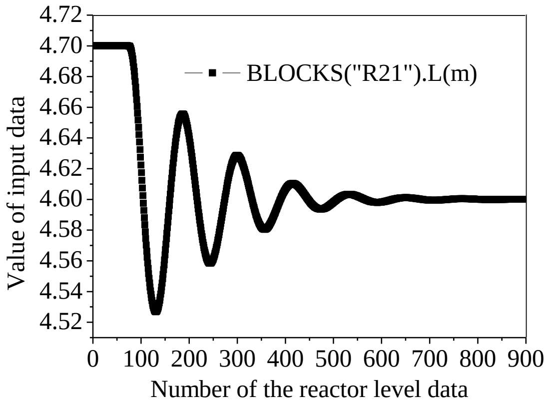

Fig. 1: Reactor level data.

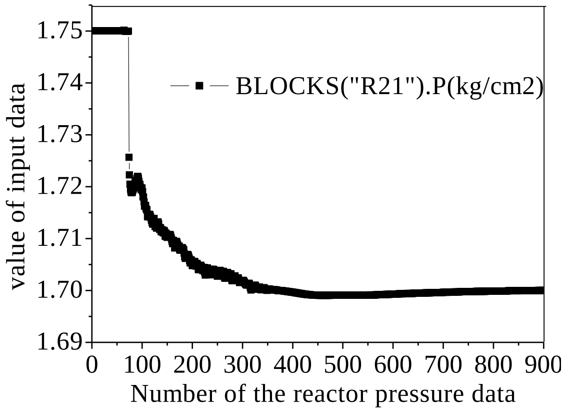

Fig. 2: Reactor pressure data.

Besides, we evaluated the model performance by MSE of all test sets. In the experiment, let m = 0 and n = 1. The six model parameters were chosen by the cultural algorithms [14] as shown in Table 1.

Table 1: Parameter settings and performance of the dataset

Model	C_1, C_2	v_1, v_2	σ_1, σ_2	Number of support vectors	MSE of test set
COOH content	6655, 0.05	0.56, 0.247	0.0001, 1.34	130	4.3E-7
Degree of polymerization	9661, 84782	0.71, 0.25	83.3, 0.0025	157	9.4E-8

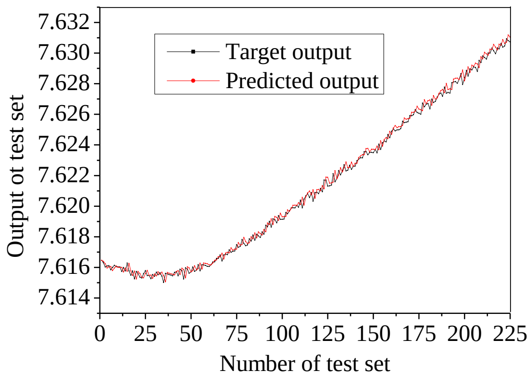

Fig. 3: Prediction of degree of polymerization.

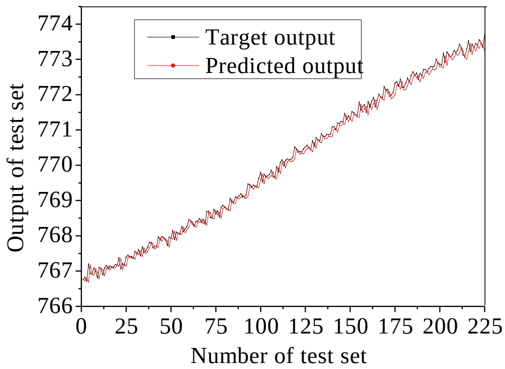

Fig. 4: Prediction of COOH content.

Fig. 3 illustrated that the degree of polymerization test set varied with time, the solid denoted the target output, and the point line was the forecast output of test set. Obviously, the two lines basically coincided. This depicted the proposed method obtained minimum generalization error. From Fig. 3, the two curves were not smooth due to data noise. Fig. 4 showed the change of COOH content over time. It had the similar conclusion with the model of the degree of polymerization. That further indicated the presented method was feasible to construct dynamic model of the MIMO system.

4. Conclusions

It is undesirable to build steady-state model of the practical industrial process varying with time. In order to follow the dynamic trend of the MIMO system by SVR, we respectively made nonlinear transformation for the input and the feedback output, and ingeniously associated the Gaussian kernel with the problem optimization. The PET simulation results were encouraging. This demonstrated the proposed approach achieved excellent generalization performance. It is more important that it provides a good way available for dynamic modeling on the MIMO system. The further work is determining the model order or changing original structure form to explore the nonlinear dynamic system.

5. Acknowledgements

This work was supported by Talent introduction of Shanghai Institute of Technology of China (Project No.2011-38).

6. References

[1] V. Vapnik. *Statistical Learning Theory*. John Wiley & Son, New York, 1998.

[2] V. Vapnik, S. Golowich, A. Smola. Support vector method for function approximation, regression estimation, and signal processing. *Advances in neural information processing systems 9*, 1996, pp: 281-287.

[3] A. Gretton, A. Doucet, R. Herbrich, et al. Support vector regression for black-box system identification. *Proc. of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing*, 2001, pp. 341-344.

[4] Z. Huang, B. Wang, Y. Li, et al. Modeling of modern automotive petrol engine performance using support vector machines. *Journal of Zhejiang University Science A*, 2005, 6A (1): 1-8.

[5] H. Zhou, J. P. Zhao., L. G. Zheng, et al. Modeling NO_x emissions from coal-fired utility boilers using support vector regression with ant colony optimization. *Engineering Application of Artificial Intelligence*, 2012, 25: 147-158.

[6] X. Wang, Z. Zou, X. Hou, et al. System identification modeling of ship manoeuvring motion based on ε-support vector regression. *Journal of Hydrodynamics*, 2015, 27(4):502-512.

[7] H. Rong, G. Zhang, C. Zhang. Application of support vector machines to nonlinear system identification. *Proc. Autonomous Decentralized systems*, 2005, pp: 501-507.

[8] H. Zhang, Z. Han, R. Feng, et al. Support vector machine-based nonlinear system modeling and control. *Journal of systems engineering and electronics*, 2003, 14(3): 53-58.

[9] D. A. Finan, H. Zisser, L. Jovanovic, et al. Identification of linear dynamic models for type 1 diabetes: a simulation study. *Proc. of 2006 IFAC International Symposium on Advanced Control of Chemical processes*, 2006, pp: 503-508.

[10] M. Ye, X. Wang. Modeling of nonlinear dynamic system using nu-support vector machines. *Proc. of the Fifth World Congress on Intelligent Control and Automation*, 2004, 1: 272-275.

[11] I. Goethals, K. Pelckmans, J.A.K. Suykens, et al. Identification of MIMO Hammerstein models using least squares support vector machines. *Automatica*, 2005, 41(7): 1263-1272.

[12] K. Feng, J. Lu, J. Chen. Nonliear model predictive control based on support vector machine and genetic algorithm. *Chinese Journal of Chemical Engineering*, 2015, 23:2048-2062.

[13] A. Smola, B. Schölkopf, G. Rätsch. Linear programs for automatic accuracy control in regression. *Proc. of the Ninth International Conference on Artificial Neural Networks*, 1999, 2: 575 - 580.

[14] R. G. Reynolds. An introduction to cultural algorithms. *Proc. of the Third Annual Conference on Evolution Programming*, 1994, pp: 131-136.