Chinese Journal of Chemical Engineering 22 (2014) 799–804

Contents lists available at ScienceDirect

Chinese Journal of Chemical Engineering

journal homepage: www.elsevier.com/locate/CJCHE

Process Model

A Selective Moving Window Partial Least Squares Method and Its Application in Process Modeling ☆

Ouguan Xu [1,*], Yongfeng Fu [1], Hongye Su [2], Lijuan Li [3]

[1] Zhijiang College, Zhejiang University of Technology, Hangzhou 310024, China
[2] State Key Laboratory of Industrial Control Technology, Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou 310027, China
[3] College of Automation and Electrical Engineering, Nanjing University of Technology, Nanjing 210009, China

ARTICLE INFO

Article history:
Received 15 May 2013
Received in revised form 14 September 2013
Accepted 16 October 2013
Available online 20 June 2014

Keywords:
SMW-PLS
Hydro-isomerization process
Selective updating strategy
Soft sensor

ABSTRACT

A selective moving window partial least squares (SMW-PLS) soft sensor was proposed in this paper and applied to a hydro-isomerization process for on-line estimation of para-xylene (PX) content. Aiming at the high frequency of model updating in previous recursive PLS methods, a selective updating strategy was developed. The model adaptation is activated once the prediction error is larger than a preset threshold, or the model is kept unchanged. As a result, the frequency of model updating is reduced greatly, while the change of prediction accuracy is minor. The performance of the proposed model is better as compared with that of other PLS-based model. The compromise between prediction accuracy and real-time performance can be obtained by regulating the threshold. The guidelines to determine the model parameters are illustrated. In summary, the proposed SMW-PLS method can deal with the slow time-varying processes effectively.

1. Introduction

Partial least squares (PLS) regression has many excellent attributes such as simple model structure, stable and robust performance, and fewer training samples needed [1,2], so it is widely used in process modeling [3,4], process control [5,6], process monitoring and fault diagnosis [2,7,8]. However, the PLS model may be locally valid and its performance will be degraded due to high level noise and disturbance in samples or time-varying industrial process such as catalytic decaying, equipment aging, or process drifting [9–13]. Many adaptive PLS models [9–16] have been proposed to deal with the dynamic behavior of processes. The basic recursive PLS was first proposed by Helland *et al.* [9] and then modified by Qin [10]. A representation was given for the old data that maintain the information without increasing the dimensionality. The block-wise recursive PLS algorithm developed by Qin [11] was extended from the basic form with a moving window and a forgetting factor. The algorithm could adapt the model based on new data and the old PLS model, avoiding re-modeling the old data. A fast moving window algorithm [12] was derived to update the kernel PLS model. The proposed approach adapted the parameters of inferential model with the dissimilarity between the new and oldest data. The time-varying characteristics of processes could also be dealt with effectively by moving window approach [11]. However, the effect of discarded sample on the model could not be evaluated properly. In this case, a new recursive PLS model was developed by Liu [14] through updating the mean and variance of the new sample and old ones, so part of historical information of the abandoned sample remained. The effective model was expanded to an online dual updating method by Mu *et al.* [15], integrating the model updating and the output offset updating. Since the dual updating strategy takes the advantages of the two methods, it is more effective in adapting process changes. A similar dual updating scheme was proposed by Ahmed *et al.* [16] for the prediction of melt index of a continuous polymerization process. The recursive PLS models are updated repeatedly either in block-wise or sample-wise once any new sample(s) is available, inducing a heavy load on the model manager or computational machine. In order to improve the real-time performance of the model, the frequency of model updating should be properly controlled. Among the proposed adaptive PLS models, the frequency of model updating is reduced with the dual updating method proposed by Mu *et al.* [15], since the model updating is activated at regular intervals. Different from the strategy proposed by Mu *et al.* [15], the decision of which updating method to be performed is based on the prediction error [16]. A novel adaptive modeling method was proposed by Lee *et al.* [17]. Depending on the model performance assessment, partial or complete adaptation is utilized to remodel the PLS method. The adaptive modeling method shows better updating performance and lower updating frequency compared to the block-wise recursive PLS modeling technique.

For the purpose of on-line application, more attentions have been paid to the real-time performance of the model. In this paper, a selective modeling strategy is proposed for the moving window PLS to decrease the model adaptation frequency by a preset threshold. The model

☆ Supported by the National Natural Science Foundation of China (61203133, 61203072), and the Open Project Program of the State Key Laboratory of Industrial Control Technology (ICT1214).

* Corresponding author.
E-mail address: ogxu@zjut.edu.cn (O. Xu).

http://dx.doi.org/10.1016/j.cjche.2014.05.012
1004-9541/

will be updated if its prediction accuracy exceeds a confidence limit, otherwise, the model remains unchanged. The selective moving window PLS (SMW-PLS) method is applied to an industrial example to predict para-xylene (PX) content in the hydro-isomerization process of C_8-aromatics.

2. SMW-PLS Method

2.1. PLS method

Given a pair of input and output data sets $\boldsymbol{X}$ and $\boldsymbol{Y}$ that have been standardized, where $\boldsymbol{X} \in \Re^{n \times m}$ and $\boldsymbol{Y} \in \Re^{n \times l}$. The linear relationship of the two matrices can be expressed as

$$\boldsymbol{Y}=\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{e} \tag{1}$$

where $\boldsymbol{\beta}$ is the coefficient matrix and $\boldsymbol{e}$ is the noise vector.

The PLS regression builds a linear model by decomposing matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ into bilinear terms,

$$\boldsymbol{X} = \boldsymbol{t}_1\boldsymbol{p}_1^{\mathrm{T}} + \boldsymbol{E}_1 \tag{2}$$

$$\boldsymbol{Y}=\boldsymbol{u}_1\boldsymbol{q}_1^{\mathrm{T}}+\boldsymbol{F}_1 \tag{3}$$

where $\boldsymbol{E}_1$ and $\boldsymbol{F}_1$ are the residual matrices, $\boldsymbol{p}_1$ and $\boldsymbol{q}_1$ are the loading vectors, respectively, and $\boldsymbol{t}_1$ and $\boldsymbol{u}_1$ are the latent score vectors of the first PLS factors determined by

$$\boldsymbol{t}_1=\boldsymbol{X}\boldsymbol{w}_1, \boldsymbol{u}_1=\boldsymbol{Y}\boldsymbol{c}_1 \tag{4}$$

where $\boldsymbol{w}_1$ and $\boldsymbol{c}_1$ are the normalized eigenvectors of corresponding dominant eigenvalue of matrices $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}$ and $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$, respectively.

The relationship between $\boldsymbol{t}_1$ and $\boldsymbol{u}_1$ is

$$\boldsymbol{u_1}=b_1\boldsymbol{t}_1+\boldsymbol{r}_1 \tag{5}$$

where b_1 is the regression coefficient and $\boldsymbol{r}_1$ is the residual vector.

After the calculation of the first factor, if $\boldsymbol{t}_1$ and $\boldsymbol{u}_1$ do not contain enough information, the second factor is calculated by decomposing the residuals $\boldsymbol{E}_1$ and $\boldsymbol{F}_1$ with the same procedure for the first factor. The procedure is repeated until the model accuracy is satisfied. Finally, the value $\hat{\boldsymbol{\beta}}$ can be estimated by

$$\hat{\beta}=\overline{\boldsymbol{W}}\boldsymbol{B}\boldsymbol{Q}^{\mathrm{T}} \tag{6}$$

where $\overline{\boldsymbol{W}} = [\overline{\boldsymbol{w}}_1, \overline{\boldsymbol{w}}_2, \ldots, \overline{\boldsymbol{w}}_a]$, $\overline{\boldsymbol{w}}_1 = \boldsymbol{w}_1$, $\overline{\boldsymbol{w}}_h = \boldsymbol{w}_h \prod_{i=1}^{h-1}\left(\boldsymbol{I}_m - \mathbf{w}_i\boldsymbol{p}_i^{\mathrm{T}}\right)$, $h = 2, \cdots, a$, and a is the number of latent variables. $\boldsymbol{T} = [\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_a]$ and $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_a]$ are the vectors of score matrices of input sets $\boldsymbol{X}$ and output sets $\boldsymbol{Y}$. Define $\boldsymbol{B} = \mathrm{diag}[b_1, b_2, \ldots, b_a]$ and $\boldsymbol{Q} = [\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_a]$. More information about the nonlinear iterative partial least squares method can be referred to [9].

2.2. Moving window PLS method

Industrial processes often experience time-varying changes, such as catalyst decaying, sensor and process drifting, as well as degradation of efficiency. Several adaptive algorithms [9,11,15,18] have been proposed to update the model based on new process data that reflect process changes. Xu and Chen [18] have shown that the simple PLS model presents the best real-time performance among the PLS-based models, but its accuracy and tracing performance are the worst. In order to improve the prediction accuracy and tracing performance of the model, the moving window approach is proposed for the simple PLS method. The model will be updated once a new sample is available while the oldest sample is discarded for the training data. As a result, the number of training samples is kept constant.

2.3. Strategy of selective updating for the moving window PLS method

The moving window PLS model needs updating repeatedly if any new sample is available. This high updating frequency is time-consuming. Hence, an approach to reduce the updating frequency is greatly required.

A strategy of selective updating for the model is proposed. The idea is as follows. If the model predicted value ($\hat{y}_k$) is consistent with the actual measurement value (y_k) of the process, and the prediction error is less than a preset threshold, that means the current model reflects the behavior of process exactly and does not need updating. The new sample (a pair of $[\boldsymbol{x}_k, y_k]$) is added to the training sample set and the oldest one is omitted. Subsequently, the sampling data set is renewed and the number of training samples remains unchanged. When the error between $\hat{y}_k$ and y_k exceeds the preset threshold, the available pair of $[\boldsymbol{x}_k, y_k]$ will be incorporated into the training sample set, while the oldest one is abandoned. Then the model updating will be activated with the latest sampling data set. A relative prediction error bound which is a small positive number, is introduced as a threshold for measurement of complexity conditions. A new sample will be introduced to sampling set and the model is updated once Eq. (7) is satisfied.

$$RE_k = \left|\frac{y_k - \hat{y}_k}{\hat{y}_k}\right| > \delta \tag{7}$$

where RE is the absolutely relative error, and δ is a predefined positive threshold, which can be set flexibly according to the demands and will be discussed in the next section.

2.4. Procedure of the on-line algorithm

With the above moving window PLS model and selective updating strategy, the procedure of on-line modeling of a continuous process can be designed as follows.

Step 1. Select the length of training samples, N, and an appropriate threshold, δ, given a pair of training data sets $\boldsymbol{X}_N$ and $\boldsymbol{Y}_N$, where $\boldsymbol{X}_N \in \Re^{N \times m}$ and $\boldsymbol{Y}_N \in \Re^{N \times l}$.

Step 2. Calculate the regression coefficient $\hat{\boldsymbol{\beta}}$ by Eq. (6).

Step 3. Predict the model output with $\hat{\boldsymbol{\beta}}$ and the newly available measurement $\boldsymbol{X}$ by

$$\hat{\boldsymbol{Y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}. \tag{8}$$

Step 4. If a new sample(s) is available, add the new sample(s) into the training data set and discard the oldest one(s) to compose a new sampling data set, go to Step 5; or return to Step 2.

Step 5. If Eq. (7) is satisfied, go to Step 2; or return to Step 3.

3. Modeling of a Hydro-isomerization Process

3.1. Hydro-isomerization process of C_8-aromatics

In the present work, an industrial hydro-isomerization process of C_8-aromatics is studied, which is one of the important parts of PX joint process in a refinery. Its schematic diagram is shown in Fig. 1. The raffinate (C_8-aromatics with lean PX) from Eluxyl adsorption unit, together with the C_8-naphtha and C_8-paraffin (C_8(N + P)) fractions from recycle column T302 and hydrogen from recycle compressor C301, exchange heat in heat exchanger E301 with the effluent from reactor R301. Then the mixture is heated to its required temperature