Journal of Process Control 21 (2011) 462–474

Contents lists available at ScienceDirect

Journal of Process Control

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

Robust monitoring and fault reconstruction based on variational inference component analysis

Zhiqiang Ge*, Zhihuan Song*

State Key Laboratory of Industrial Control Technology, Institute of Industrial Process Control, Zhejiang University, Hangzhou 310027, Zhejiang, PR China

ARTICLE INFO

Article history:
Received 16 March 2010
Received in revised form
25 November 2010
Accepted 16 February 2011
Available online 15 March 2011

Keywords:
Robust process monitoring
Fault reconstruction
Bayesian PCA
Variational inference
Dimensionality selection

ABSTRACT

Probabilistic models such as probabilistic principal component analysis (PPCA) have recently caught much attention in the process monitoring area. An important issue of the PPCA method is how to determine the dimensionality of the latent variable space. In the present paper, one of the most popular Bayesian type chemometric methods, Bayesian PCA (BPCA) is introduced for process monitoring purpose, which is based on the recent developed variational inference algorithm. In this monitoring framework, the effectiveness of each extracted latent variable can be well reflected by a hyperparameter, upon which the dimensionality of the latent variable space can be automatically determined. Meanwhile, for practical consideration, the developed BPCA-based monitoring method is robust to missing data and can also give satisfactory performance under limited data samples. Another contribution of this paper is due to the proposal of a new fault reconstruction method under the BPCA model structure. Two case studies are provided to evaluate the performance of the proposed method.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

In the process monitoring area, data-based methods have recently become very popular, which is due to the following two reasons. Firstly, modern production processes are becoming very complex, as a result, traditional model-based monitoring methods become more and more complicated, costly, and sometimes impossible to use. Secondly, a huge number of process data have been collected by the widely used distributed control system (DCS), depending on which useful information and data features can be extracted and utilized for modeling and monitoring purposes. As a represented type of data-based method for process monitoring, statistical-based approaches such as principal component analysis (PCA) and partial least squares (PLS) have caught much attention since last decades [1–3]. To our best knowledge, those statistical-based approaches have been researched in different aspects, such as the nonlinear case, the dynamic case, the multiscale case, the batch process case, and the probabilistic case [4–13]. The topic of this paper is focused on the probabilistic case, which has recently caught attention in this area [14–16].

As a counterpart of PCA, the probabilistic principal component analysis (PPCA) has been employed for probabilistic monitoring purpose, which was originally proposed by Tipping and Bishop [17]. Different from PCA, PPCA not only presents the model in a probabilistic manner, but also can catch the noise information simultaneously. Under the PPCA-based monitoring framework, an important issue is the number selection of the retained principal components. One main concern of this paper is to determine this important number automatically when constructing the probabilistic model for monitoring. To this end, the Bayesian principal component analysis (BPCA) is introduced in the present paper, which was proposed by Bishop [18]. Later, the Bayesian PCA method was improved through the variational inference algorithm [19,20]. Compared to the original BPCA method, the variational inference based BPCA method is more computationally efficient, and can be optimized through a rigorous bound on the marginal log probability. Through the Bayesian treatment, the effective number of principal components can be determined automatically, which also has its probabilistic interpretation in the model structure.

In practice, the following two important issues always need to be considered in industrial processes: (1) missing values; (2) limited data samples. Although several research studies have been carried out to address those two problems, we intend to provide a robust modeling method in this new probabilistic monitoring framework. With the variational inference based Bayesian estima-

* Corresponding authors.
E-mail addresses: zqge@iipc.zju.edu.cn, gezhiqiang@gmail.com (Z. Ge), zhsong@iipc.zju.edu.cn (Z. Song).

0959-1524/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2011.02.005