Journal of Process Control 21 (2011) 1306–1317

Contents lists available at ScienceDirect

Journal of Process Control

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

Improved nonlinear PCA for process monitoring using support vector data description

Xueqin Liu[a], Kang Li[a,*], Marion McAfee[b], George W. Irwin[a]

[a] Intelligent Systems and Control Group, School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast BT9 5AH, UK
[b] Department of Mechanical and Electronic Engineering, Institute of Technology Sligo, Sligo, Ireland

ARTICLE INFO

Article history:
Received 3 February 2011
Received in revised form 29 May 2011
Accepted 12 July 2011
Available online 9 August 2011

Keywords:
Nonlinear principal component analysis
Support vector datadescription
Fast Recursive Algorithm
Principal curves
Radial basis function network

ABSTRACT

Nonlinear principal component analysis (PCA) based on neural networks has drawn significant attention as a monitoring tool for complex nonlinear processes, but there remains a difficulty with determining the optimal network topology. This paper exploits the advantages of the Fast Recursive Algorithm, where the number of nodes, the location of centres, and the weights between the hidden layer and the output layer can be identified simultaneously for the radial basis function (RBF) networks. The topology problem for the nonlinear PCA based on neural networks can thus be solved. Another problem with nonlinear PCA is that the derived nonlinear scores may not be statistically independent or follow a simple parametric distribution. This hinders its applications in process monitoring since the simplicity of applying predetermined probability distribution functions is lost. This paper proposes the use of a support vector data description and shows that transforming the nonlinear principal components into a feature space allows a simple statistical inference. Results from both simulated and industrial data confirm the efficacy of the proposed method for solving nonlinear principal component problems, compared with linear PCA and kernel PCA.

1. Introduction

The demand for efficient monitoring of product quality and guaranteeing safe operation in the process industry has motivated the development of statistically-based fault detection and diagnosis methods. Multivariate statistical methods such as principal component analysis (PCA) [1,2], partial least squares [3] and, more recently, independent component analysis [4,5] have been widely applied in the chemical industry for this purpose [6]. Of these PCA is the most popular because of its simplicity. However, conventional PCA is a linear projection method, so it may not be able to handle nonlinear system dynamics thus violating the underlying assumption that neglected minor components do not contain important information [7].

This has naturally motivated the development of nonlinear PCA approaches which can retain more information using fewer components. Examples include principal curves/surfaces [8], multi-layer auto-associative neural networks (ANNs) [9] including radial basis function (RBF) networks [7] and the kernel function approach [10,11].

Nonlinear extensions of PCA which rely on multi-layer ANNs were first developed by Kramer [9]. Unfortunately, ANNs are difficult to train and topological features like the number of nodes in each layer are difficult to determine [12,13]. Dong and McAvoy [8] and Liu et al. [14] incorporated principal curves into the neural network structure to overcome some of these deficiencies. Other nonlinear PCA methods include an input-training neural network [15,16] and the use of genetic programming [17]. For fault detection, the input-training network may fail to detect incipient faults since the nonlinear principal components (PCs) are computed from the recorded operating data to maximize its reconstruction. While the score variables of a linear PCA model asymptotically follow a Gaussian distribution, this can no longer be assumed for a nonlinear PCA model. In fact, such models produce score variables that may not be statistically independent or follow a simple parametric distribution [18].

Existing work on determining confidence limits for nonlinear PCs either utilize a kernel density estimation (KDE) [19], which is computationally expensive for determining confidence limit, or a local approach [18], to produce statistics that asymptotically follow a normal distribution.

An alternative nonlinear PCA extension is kernel PCA (KPCA) [10], which can compute PCs in a high-dimensional feature space using integral operators and nonlinear kernel functions. The core idea of KPCA is to first map the data space into a feature space using a nonlinear mapping and then to compute the PCs in this space. Further, this mapping function can be approximated using kernel functions in a similar fashion to an RBF network. KPCA does not

* Corresponding author.
E-mail address: k.li@qub.ac.uk (K. Li).

0959-1524/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2011.07.003