Journal of Loss Prevention in the Process Industries 43 (2016) 212–224

Contents lists available at ScienceDirect

Journal of Loss Prevention in the Process Industries

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

Kernel PLS-based GLRT method for fault detection of chemical processes

Chiranjivi Botre [a], Majdi Mansouri [b], Mohamed Nounou [c,*], Hazem Nounou [b], M. Nazmul Karim [a]

[a] Chemical Engineering Department, Texas A&M University, College Station, TX 77843, USA
[b] Electrical and Computer Engineering Program, Texas A&M University at Qatar, Doha, Qatar
[c] Chemical Engineering Program, Texas A&M University at QATAR, Doha, Qatar

ARTICLE INFO

Article history:
Received 7 March 2016
Accepted 18 May 2016
Available online 20 May 2016

Keywords:
Kernel partial least square
Generalized likelihood ratio test
Continuously stirred tank reactor
Fault detection

ABSTRACT

Fault detection is essential for proper and safe operation of various chemical processes, and it has recently become even more important than ever before. In this paper, we extended our previous work (Mansouri et al. (2016)), which addresses the problem of fault detection of chemical systems using kernel principal component analysis (KPCA)-based generalized likelihood ratio test (GLRT), to widen its applicability for processes represented by input-output models. Specifically, hypothesis testing fault detection technique that are based on linear and nonlinear partial least squares (PLS) models are developed. For nonlinear PLS models, a kernel PLS (KPLS) modeling framework is utilized. KPLS has been widely used to model various nonlinear processes, such as distillation columns and reactors. Thus, in the current work, a KPLS-based GLRT fault detection method is developed, in which KPLS is used as a modeling framework and the KPLS model generated residuals are evaluated using a GLRT statistic. The fault detection performance of the developed KPLS-based GLRT method is illustrated through a simulated example representing a continuously stirred tank reactor (CSTR). The simulation results show that the KPLS-based GLRT method outperforms its linear PLS-based version, and that both of the aforementioned techniques provide clear advantages over the conventional linear and nonlinear PLS based statistics, i.e., T^2 and Q.

1. Introduction

Monitoring of the chemical processes is important for the safety of the plant and to ensure product quality is maintained. Process monitoring steps consists of detecting the fault in the system and taking corrective action against it (Hwang et al. (2010); Isermann (2006); Qin (2012); Venkatasubramanian et al. (2003a); Qingsong (2004); Venkatasubramanian et al. (2003b)).

Multivariate statistic methods are very effective for fault detection and diagnosis in chemical industry Hwang et al. (2010); Qin (2012). The partial least square (PLS) and principle component analysis (PCA) are two basic types of multivariate methods. PCA is among the most popular statistical methods used for modeling and faults detection problems (Yu (2011); Herve and Lynne (2010); Wang and Chen (2004); Diana and Tommasi (2002)), however, it provides a linear combinations of variables that demonstrate major trends in data set. In our previous work (Mansouri et al. (2016)), we have successfully applied kernel PCA (KPCA) based generalized likelihood ratio test (GLRT) for nonlinear fault detection of chemical system. However, KPCA is an input-space model and cannot take outcome measures into account and most chemical processes or many of them, such as distillation columns, are usually described by input-output models. PLS is an input output model and could be used to detect fault in both process and variables. It can also be used as a linear regression tool to predict the output variables form process variables. Hoteling T^2 and Q statistics are common statistical fault detection (FD) charts that are applied with PLS for process monitoring. However, the use of the conventional PLS (Lodhi and Yamanishi (2011)) through its two charts hoteling T^2 and Q could lead to missed detection and high false alarm rate.

However, chemical and refinery processes are complex and most of the variables are non-linear in nature and the fault

* Corresponding author.
E-mail address: mohamed.nounou@qatar.tamu.edu (M. Nounou).

http://dx.doi.org/10.1016/j.jlp.2016.05.023
0950-4230/

detection of this processes by linear PLS would lead to many missed diagnosis and non-reliable results. In literature, several nonlinear version of PLS are developed, a nonlinear iterative partial least square (NIPALS) algorithm is used to model PLS which was developed in Wold (1992). The authors in (Malthouse et al. (1997)) have presented a complicated artificial neural network to model nonlinear PLS. While, the authors in (Wold et al. (1984)) have proposed to use a quadratic function to relate scores in PLS algorithm. Thus, in this paper, we propose to use the kernel partial least square (KPLS) as a modeling framework. The KPLS is among the most well-known nonlinear statistical method (Rosipal and Trejo (2001)), it is the method for performing a nonlinear form of PLS. It is an input output model, which reduces the dimensionality of process variable and variables to extract scores and principle components and could be used to detect fault in both process and variables. The KPLS gives good general properties of nonlinear PLS by selecting appropriate kernel function (Rosipal and Trejo (2001)), radial basis function, polynomial function and sigmoid function are three common kernel function used. KPLS can also be used as a regression tool to predict the product variables from nonlinear process variables. KPLS approach works similar to PLS, it reduces the dimensionality of nonlinear process variables and variables by projecting into space with less dimensionality.

Hence, the objective of this paper, is to address the problem of nonlinear fault detection so that the data are first modeled using the KPLS algorithm and then the faults are detected using generalized likelihood ratio test (GLRT). The KPLS is used to create the model and find nonlinear combinations of parameters which describe the major trends in a data set and GLRT is used to detect the faults and both are utilized to improve faults detection process. An alternative approach for fault detection is to use hypothesis testing based techniques, such the GLRT. The GLRT has been shown to provide good detection abilities for specified false alarm rates (Gustafsson (1996); Willsky et al. (1980); Dawdle et al. (1982)).

The fault detection performances of the KPLS-based GLRT is illustrated through a simulated continuously stirred tank reactor (CSTR) data. The results demonstrate the effectiveness of the KPLS-based GLRT method over the linear PLS-based GLRT and conventional KPLS methods for detection of single as well as multiple sensor faults and assessed using the false alarms and missed detection rates.

The rest of the paper is organized as the following. In Section 2, an introduction to PLS and KPLS methods is given, followed by descriptions of the two main detection indices, T^2 and Q, which are generally used with KPLS for fault detection. Then, the GLRT which is utilized in composite hypothesis testing is discussed in Section 3. After that, the KPLS-based GLRT method used for fault detection which integrates KPLS modeling and GLR statistical testing, is presented in Section 4. Next, in Section 5, the KPLS-based GLRT performance is studied through a simulated continuously stirred tank reactor data. At the end, the conclusions are presented in Section 6.

2. Partial least square and kernel partial least square methods description

First, in Section 2.1, we present the linear partial least square.

2.1. Partial least square (PLS) method

Let $X \in R^{N\times M}$ denotes an input data matrix having N observations and M variables, and $Y \in R^{N\times L}$ an output data matrix consists of L response variables. PLS is an input output model and can decompose both X and Y matrices and detect the fault in both X and Y variables, it is formally determined by two sets of linear equations: the inner model (the relations between the latent variables) and the outer model (the relations linking the latent variables and their associated observed variables) Geladi and Kowalski (1986a). The X and Y matrices are linked by score vectors (T and U) and principle components (P and U). The PLS model is given by Kourti and MacGregor (1995):

$$X = TP^T + E = \sum_{i=1}^{I} t_i p_i^t + E = \widehat{X} + E \quad (1)$$

$$X = UQ^T + F = \sum_{j=1}^{J} u_j q_j^t + F = \widehat{Y} + F, \quad (2)$$

where $\widehat{X}$ and $\widehat{Y}$ represent modeling matrices of X and Y successively, $E \in R^{N\times M}$ and $F \in R^{N\times L}$ are the residuals of X and Y respectively, $T = [t_1, t_2 \cdots t_I] \in \mathbb{R}^{M\times I}$ is the resulting input score matrix, $U = [u_1, u_2 \cdots u_J] \in \mathbb{R}^{N\times J}$ is the output score matrix, $P = [p_1^t, p_2^t \cdots p_I^t] \in \mathbb{R}^{M\times I}$ and $Q = [q_1^t, q_2^t \cdots q_J^t] \in R^{L\times J}$ represent the loading matrices, successively. The two matrices X and Y are generally pre-treated by centering and scaling to have mean zero and variance unity prior to PLS modeling.

The scores and the principle components are calculated from the NIPALS algorithm (presented in Algorithm 1). Each iteration calculates a column matrix of t, u, p, q and the residuals obtained from previous iteration is used as input to the next iteration, which makes sure that all scores and principle components are extracted from X and Y matrices. The NIPALS algorithm is presented in Algorithm 1.

The variability of data is extracted from X and Y matrices to get

Algorithm 1
NIPALS algorithm.

Input: $N \times M$ input data matrix X and $N \times L$ output data matrix Y
Output: Input score matrix $T = [t_1, t_2 \cdots t_I]$, Output score matrix $U = [u_1, u_2 \cdots u_J]$.

- Initialize output score as:

$$u = y_i \quad (3)$$

- Weights are given as:

$$w = u^T X \Big/ u^T u \quad (4)$$

- Normalize weight w:

$$w = w/||w|| \quad (5)$$

- Score vector of X:

$$t = Xw \Big/ w^T w \quad (6)$$

- Loading vector of X:

$$p = Xt^T \Big/ t^T t \quad (7)$$

- Loading vector of Y:

$$q = Yt^T \Big/ t^T t \quad (8)$$

- Score vector of Y:

$$u = Yq^T \Big/ q^T q \quad (9)$$

- Rescale weights, loading vectors and scores: $p = \frac{p}{norm(p)}$, $w = w \times norm(p)$, $t = t \times norm(p)$;
- Deflate matrices X and Y as:

$$X = X - tp^T \quad (10)$$

$$Y = Y - tq^T \quad (11)$$

Repeat step 2 to 8, to get next latent variables, the resulting input score matrices $T = [t_1, t_2 \cdots t_m] \in \mathbb{R}^{m\times N}$ and output score matrix $U = [u_1, u_2 \cdots u_m] \in \mathbb{R}^{m\times N}$.