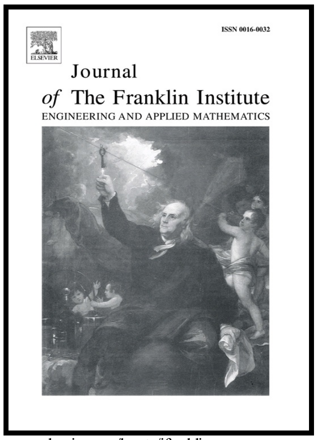
Author's Accepted Manuscript



www.elsevier.com/locate/jfranklin

Quality-related fault detection using linear and nonlinear principal component regression

Guang Wang, Hao Luo, Kaixiang Peng

PII: S0016-0032(16)30105-3
DOI: http://dx.doi.org/10.1016/j.jfranklin.2016.03.021
Reference: FI2572

To appear in: *Journal of the Franklin Institute*

Received date: 14 May 2015
Revised date: 20 January 2016
Accepted date: 27 March 2016

Cite this article as: Guang Wang, Hao Luo and Kaixiang Peng, Quality-relatec fault detection using linear and nonlinear principal component regression, *Journa of the Franklin Institute,* http://dx.doi.org/10.1016/j.jfranklin.2016.03.021

Quality-related fault detection using linear and nonlinear principal component regression

Guang Wang[1,*], Hao Luo[2], Kaixiang Peng[3]

Abstract

The issue of quality-related fault detection is a hot research topic in the process monitoring community in the recent five years. Several modifications based on partial least squares (PLS) have been proposed to solve the revelent problems for linear systems. For the systems with nonlinear characteristics, some modified algorithms based on kernel partial least squares (KPLS) have also been designed very recently. However, most of the existing methods suffer from the defect that their performances are not stable when the fault intensity increases. More importantly, there is no way yet to solve the linear and nonlinear problems in a uniform algorithm structure, which is very important for simplifying the design steps of fault detection systems. This paper aims to propose such approaches based on principal component regression (PCR) and kernel principal component regression (KPCR). Such that, relevant problems in linear and nonlinear systems can be solved in the same way. Two literature examples are used to test the performance of the proposed approaches.

Keywords: fault detection, quality-related, principal component regression, kernel principal component regression.

1. Introduction

Data-driven approaches have been receiving considerable attention in the field of process monitoring [1, 2, 3, 4] due to their easy implementation and less requirement for the underlying model. In recent years, the so-called quality-related fault detection has attracted wide attention in the process monitoring community. By modeling the relationship between quality and process variables, such methods classify the faults happened in the process space into two categories, namely those affect the final product quality and those do not affect. By reducing the alarm rates of the later, these methods can significantly reduce the unnecessary downtime of the plant, and then bring considerable economic benefits for practical applications. Moreover, quality-related fault detection can be used to guide the implementation of fault tolerant control (FTC) scheme more accurate and more targeted.

Principal component analysis (PCA) and PLS are the two most common used multivariate statistical analysis methods in process monitoring [5]. As PCA cannot establish the correlation between quality and process variables, it is naturally unable to be used for quality-related fault detection. The nature of PLS makes it without such problem, however, as revealed by Li et al. [6], the standard PLS performs an oblique decomposition on process variable space and the significant process variation related to the output might also be contained in the residual part. As a result, the utilized test statistics yet fault diagnostic results offered by PLS are problematic for quality-related fault detection. To overcome this shortcoming, Zhou et al. [7] first proposed a PLS-postprocessing-based approach, named total projection to latent structures (T-PLS), by further decomposing the score and load matrices of PLS, which finally divided the process space (spanned by the $\mathbf{X}$ matrix) into four subspaces with each part had a different correlation with the output space (spanned by the $\mathbf{Y}$ matrix). Such that, by designing appropriate statistics in these four subspaces faults with different correlations with $\mathbf{Y}$ can be classified. Soon later, Yin et al. [8] proposed a different approach, called modified partial least squares (M-PLS), which first estimated the regression coefficient matrix between $\mathbf{X}$ and $\mathbf{Y}$, and then projected $\mathbf{X}$ onto the null and the remaining subspaces of the coefficient matrix, respectively. Finally, the process space was decomposed into two orthogonal subspaces. Compared with T-PLS, M-PLS realizes an orthogonal decomposition on the process space, and it is more simple and effective than T-PLS for most cases. By taking advantages of the two methods, Qin et al. [9] developed a concurrent partial least squares (C-PLS) which was claimed more efficient. Considering the drawbacks of PLS-postprocessing-based approaches, Wang et al. [10] combined orthogonal signal correction (OSC) and M-PLS to develop an enhanced method. For nonlinear process monitoring, Peng et al. [11] and Zhang et al. [12] extended T-PLS and C-PLS into nonlinear versions, respectively. Recently, Zhang et al. [13] made a comprehensive and detailed comparison study for all these approaches.

[*]Bohai University, Jinzhou, 121013, China.
Email addresses: guang.wang@hit.edu.cn (Guang Wang), hao.luo@uni-due.de (Hao Luo), kaixiang@ustb.edu.cn (Kaixiang Peng)
[1]Research Institute of Intelligent Control and Systems, Harbin Institute of Technology, Harbin 150001, China.
[2]Institute for Automatic Control and Complex Systems, University of Duisburg-Essen, Duisburg, 47057, Germany.
[3]Key Laboratory for Advanced Control of Iron and Steel Process, School of Automation and Electrical Engineering, University of Science and Technology of Beijing, Beijing 100083, China.