Title: Multivariate Fault Detection and Classification using Interval Principal Component Analysis

Author: Nour Basha Mohamed Nounou Hazem Nounou

PII: S1877-7503(17)31245-0
DOI: https://doi.org/doi:10.1016/j.jocs.2018.04.017
Reference: JOCS 866

To appear in:

Received date: 11-11-2017
Revised date: 21-2-2018
Accepted date: 6-4-2018

Please cite this article as: Nour Basha, Mohamed Nounou, Hazem Nounou, Multivariate Fault Detection and Classification using Interval Principal Component Analysis, *<![CDATA[Journal of Computational Science]]>* (2018), https://doi.org/10.1016/j.jocs.2018.04.017

Multivariate Fault Detection and Classification using Interval Principal Component Analysis

Nour Basha[a], Mohamed Nounou[a], Hazem Nounou[b]

[a] *Chemical Engineering Department, Texas A&M University, Doha, Qatar*
[b] *Electrical and Computer Engineering Department, Texas A&M University, Doha, Qatar*

Abstract

Principal Component Analysis (PCA) is a linear data analysis tool that aims to reduce the dimensionality of a dataset, while retaining most of the variation found in it. It transforms the variables of a dataset into a new set, called the principal components, using linear combinations of the original variables. PCA is a powerful statistical technique used in research for fault detection, classification and feature extraction. Interval Principal Component Analysis (IPCA) is an extension to PCA designed to apply PCA to large datasets using interval data generated from single-valued samples. In this paper, three IPCA methods are compared: Centers IPCA, Midpoint-Radii IPCA, and Symbolic Covariance IPCA, and methods for fault detection and classification using interval data are described. Fault detection and classification applications are respectively carried out through two examples, one using synthetic and the other using real data, and the results are compared to those of the classical PCA.

The results show that IPCA methods have a higher detection rate than classical PCA, for the same false alarm rate. Moreover, IPCA methods are capable of differentiating the type of fault to a high degree of accuracy, unlike classical PCA. Interval centers were capable of detecting changes in mean, while interval radii were capable of detecting changes in variance. Furthermore, for data classification, the results show that MRIPCA had a higher classification precision than other IPCA methods and classical PCA.

Keywords: Principal Component Analysis, Interval Data, Midpoint-Radii, Symbolic Covariance, Fault Detection, Classification