PCA-based fault isolation and prognosis with application to pump

Gary R. Halligan · S. Jagannathan

Received: 15 October 2009 / Accepted: 29 November 2010 / Published online: 15 December 2010
© Springer-Verlag London Limited 2010

Abstract In this paper, the use of Linear and Kernel PCA for fault isolation and prognosis is explored since PCA is normally utilized for detection and isolation. Vector projection and statistical analysis were utilized to isolate and predict faults in the PCA domain. Linear PCA was applied to data collected from experiments on a one half horsepower centrifugal water pump both for normal and faulty operation consisting of the four fault scenarios: impeller failure, seal failure, inlet pressure sensor failure, and a filter clog. Upon close observation of the behavior of the principal component scores, it was determined that the linear PCA does not adequately isolate and predict the failures. Therefore, Kernel PCA, utilizing a Gaussian kernel, was applied to the same data sets. Analysis of the behavior shows that the principal component scores gained from the Kernel PCA performed better than linear PCA.

Keywords Principal component analysis · Fault isolation · Fault prognosis · Rotating machinery

1 Introduction

Principal component analysis (PCA) is a multivariate statistical analysis technique which is very commonly used for dimension reduction, fault detection, and isolation. PCA renders this by constructing a model from the eigenvalues and eigenvectors of the covariance matrix of the data to be analyzed. Dimension reduction is achieved by eliminating the eigenvectors, or PCs, associated with sufficiently small eigenvalues. The fault detection and diagnosis comes from observing the relationships between the variables and monitoring any changes in these relationships due to the occurrence of faults in the system.

There are many types of PCA including: linear, or conventional, Moving PCA, Adaptive PCA, Exponentially Weighted PCA [1], and kernel PCA. As with any analysis technique, each of the techniques mentioned above has advantages and disadvantages many of which are discussed at length in [1]. While conventional PCA is well suited for application to a linear system, its greatest weaknesses are assumptions that the data belong to a Gaussian distribution and relationships among the variables are linear [2] in nature. As a result of these assumptions, conventional PCA is ill-suited for processes with multiple operation modes. An additional weakness of conventional PCA is its inability to recognize outliers [3]. In short, Moving PCA, Adaptive PCA, and Exponentially Weighted PCA are all less sensitive to outliers than linear PCA and are better suited for applications with multiple operating conditions provided the additional computational complexity can be accommodated [1]. In [4], the authors suggest that faults based on the behavior of the PCs can be isolated by identifying the out of control PCs. The isolation mentioned in [4] requires observing the changes in the variables feeding the out of control PCs and combining that information with the expertise of the operator to relate the observed behavior to a given fault.

On the other hand, Wang et al. [5] aimed at detecting the failure of a rolling bearing on an electric motor. The
measurements taken on the test apparatus were vibration of the bearing housing and acoustic signal to monitor the noise generated by the apparatus. In addition to these readings, eight additional non-dimensional statistical features were extracted. These statistical features consisted of Shape, Kurtosis, Crest, Skewness, Impulse, Second-order moment, Clearance, and K Factor. Both linear (conventional) and kernel PCA, utilizing a Gaussian kernel, were applied to the data, and the results were compared. Additionally, both variations of PCA were applied to just the vibration and acoustic signals, and again to the vibration and acoustic signals including the eight statistical features. It was found that due to the nonlinear nature of the machinery in the experiment, the kernel PCA performed better than the linear PCA both with and without the statistical features. It was also found that the analysis which included the statistical features performed better than without the statistical features.

In Wang et al. [6], the aim was again to detect the failure of rolling bearings on an electric motor. Again, the acoustic and vibration signals were recorded and the same eight statistical features were calculated. However, this time, a genetic algorithm was included in the analysis. Both linear and kernel PCA were applied to the data with and without the genetic algorithm. It was found that incorporating the genetic algorithm with PCA improved the performance of both versions of PCA, under the same conditions, kernel PCA performed better than linear PCA.

Lachouri et al. [7] used a Daubechies-1 wavelets for analysis and processing of vibration signals to detect the failure of rolling bearings. The use of the wavelet was intended to decompose the signals into approximations. Multiscale PCA was then applied to each of the resulting matrices. The authors in [7] concluded that the results obtained were satisfactory and the method ensured a good and accurate diagnosis.

To the best knowledge of the authors, there has not been any work done in extending PCA into the area of fault prognosis. In this paper, the proposed methodology uses only the PC scores and basic regression algorithms to isolate and predict faults. Once the PCA components and associated thresholds are identified offline, the approach can be implemented in real time to detect abnormal data and determine which known fault is the most likely to be occurring. Once the type of fault has been identified, the time to failure can then be estimated online using least squares regression from a combination of appropriate PCA components.

In addition past literature [1–7] indicated that PCA was typically applied to processes to aid in the detection and rarely applied to individual machines. In the few cases of the application of PCA to individual pieces of machinery [5–7], PCA was used in conjunction with other approaches such as wavelets, or genetic algorithms, or by simply introducing statistical features such as kurtosis and skewness. For this paper, PCA was applied to data from a single piece of rotating machinery, a centrifugal water pump, and isolation and prognosis are verified using experimental data.

Therefore the main contributions of this paper include the extension of the PCA-based schemes for fault isolation and prognosis as well as the application of the PCA-based detection, isolation and prognosis scheme onto the rotating machinery which in the case is a centrifugal pump.

The remainder of the paper is organized as follows: Section 2 will cover the methodology of both linear and kernel PCA and their extension to isolation and prognosis while Section 3 will discuss a case study involving a centrifugal water pump. Conclusions are discussed in Section 4, and finally, the references are included in Section 5.

2 PCA methodology

2.1 Linear PCA

The first step in PCA is the standardization of the data. Standardization simply means that the data are scaled so that it has a mean of zero and a variance of one. The standardization is defined as

\[ X = (X_0 - \alpha U)S^{-1} \]  

(1)

where \( X_0 \) is a \( M \times N \) matrix of raw data, \( \alpha \) is an \( M \times 1 \) matrix with all entries equal to one, \( U \) is a \( 1 \times N \) matrix with each column entry being the mean of the corresponding column in \( X_0 \) and \( S \) is an \( N \times N \) diagonal matrix containing the standard deviations of the columns in \( X_0 \). Since \( S \) is a diagonal matrix, \( S^{-1} \) is a diagonal matrix with the inverse of the standard deviations. PCA calculates principal components, or PCs, that are thought to contain the essence of the data. These PCs describe varying amounts of variance observed in the data. The dimension reduction feature mentioned earlier comes into play when the user wishes to account for only a percentage of the variance. The first step in the calculation of the PCs is the calculation the covariance matrix, \( C \), of the data matrix \( X \) which is given by

\[ C = (M - 1)^{-1}X^TX \]  

(2)

The PCs mentioned above are the eigenvectors associated with the eigenvalues of the covariance matrix. The eigenvalues are equal to the variance explained by each of the PCs [8]. This means that the relative magnitudes of the eigenvalues are a measure of the relative importance of the corresponding PCs. In other words, the PCs associated with larger eigenvalues are more important than the PCs.