Applying the ICA-Based Approach to Detect Faults in Processes

Mu-Chen Chen, National Chiao Tung University, Taiwan, R.O.C.
Chun-Chin Hsu, Chaoyang University of Technology, Taiwan, R.O.C.
Bharat Malhotra, Indian Institute of Technology, India
Manoj Kumar Tiwari, Indian Institute of Technology, India

The Asian Conference on Technology, Information & Society
Official Conference Proceedings 2014

Abstract
To save time, cost and labor, there are many studies that have been conducted about the detection of faults in industrial processes. Most of the previous studies used only Independent Component Analysis (ICA) or Principal Component Analysis (PCA) for detection, but they cannot form close enough boundaries to reject outliers. This paper proposes an ICA-based approach to detect outliers in a process by forming close boundaries. The basic idea of the proposed method is to apply ICA to convert original data into independent components, and then apply Durbin Watson (DW) criterion to select important independent components. Hereafter, Support Vector Data Description (SVDD) has been applied for outlier detection by forming much tighter boundaries. The efficiency of proposed ICA-based approach is investigated via a simulated multivariate process example. To demonstrate the identification capability of the proposed ICA-based approach, the traditional Hotelling's T^2 chart is constructed for the simulated data set.

Keywords: Fault detection, Independent Component Analysis, Principal Component Analysis, Hotelling's T^2 chart

Introduction

In Statistical Process Control (SPC), Hotelling's T^2 chart adopts the original variable to calculate Mahalanobis distance. A simple type of Multivariate Statistical Process Control (MSPC) is a multivariate chart which is an extended from univariate SPC methods like Hotelling's T^2 chart (Montgomery, 2001). This chart is unable to detect small changes taking place during the cycle. To remove this drawback, MEWMA and MCUSUM are developed to detect small process changes (Virojana *et al*., 2003). However, the fault detectability decreases as the number of variables increases. Principal Component Analysis (PCA) has been widely used for detecting faults because of its compatibility with many of the methods available in MSPC (Bakshi, 1998) and its ability to reduce attribute dimensions. However, PCA lacks in assuming state of variable, it assumes that latent variable follows Gaussian distribution, while in chemical processes variables hardly remains at stable state because of the uncontrolled disturbances.

Dynamic PCA (DPCA) has been proposed which uses an amplified matrix with time lagging variables (Ku *et al*., 1995). Serial correlation of data has been considered in this method. Along with other developed methods, DPCA has been widely used in many fields such as sensor fault detection (Luo *et al*., 1999 and Rato *et al*., 2013), multi-scale fault detection, multi-scale analysis, simultaneously monitoring, diagnosis in the wastewater treatment process (Yoo *et al*., 2002) and diagnosing an automatic controlled process (Tsung, 2000). Kernal PCA has been developed for nonlinear nonstationary process monitoring (Khediri *et al*., 2011).

Techniques based on Independent Component Analysis (ICA) have been developed (Kano *et al*., 2003 and Lee *et al*., 2003) recently. ICA is used for decomposition of data into linear combination of components independent from each other. Then, ICA monitoring on lagged variables, Dynamic ICA has also been used to detect faults. Chiang *et al*. (2004) has used Support Vector Machine (SVM) and Fisher discriminant analysis (FDA) for fault diagnosis. Guo *et al*. (2014) proposed envelope based dimension reduction for ICA in fault diagnosis.

ICA has been used extensively for fault detection since it gives more sophisticated results than PCA, and the ICA's power in fault detection has been ascertained. However, the improvement of ICA based fault detection approach is still in need because of the following limitations of ICA (Hyvärinen and Oja, 1997; Hyvärinen, 1999; Hsu *et al*., 2010): 1) the incapability of identifying important independent components (ICs), 2) the lack of understanding the influence of original variable on given IC, and 3) the loose boundary for enveloping all data points.

Understanding the abovementioned limitations of ICA, this paper aims at developing an efficient approach of fault detection and diagnosis for non-Gaussian processes. In the proposed approach, important ICs are selected with the help of Durbin Watson (DW) criterion (Dublin and Watson, 1950). Original variables are retraced to obtain a view that which original variable is influencing the identified important independent components. Further, Support Vector Data Description (SVDD)(Tax and Duin, 2004) has been applied for the purpose of tightening of the boundary surrounding the data points.

The Proposed Fault Detection and Diagnosis Approach

Firstly, the basics of ICA, DW and SVDD are presented. Independent component analysis (ICA) is an approach which supposes statistical independence of non-Gaussian source signals for dividing a multivariate signal into smaller components. Readers are referred to (Lee *et al*., 2003) for the further details of ICA. ICA can be of many types such as Jade (Rutledge and Bouveresse, 2013), FastICA (Hyvärinen, 1999), etc. In this paper, we use FastICA.

Hyvärinen (1999) invented FastICA which is an efficient algorithm for ICA. FastICA has many advantages over normal ICA model which are discussed later in this paper. Centering and whitening of the input vector data x has to be done before applying FastICA algorithm (Hyvärinen, 1999). The FastICA is finds maximum of the non-Gaussianity (Hyvärinen, 1999; Hyvärinen and Oja, 1997) and it is based on fixed point iteration scheme. By approximative Newton iteration FastICA can be derived.

Recently, Durbin Watson (DW) criterion has been widely used in many models, and recently found its application in ICA (Ammari *et al*., 2011). Initially this methodology was proposed for measuring signal/noise ratio (Dublin and Watson, 1950). In absence of any king of noise, the DW value tends toward 0, and if signal consist of only noise, it will incline to 2. This criterion has been used for validation of the multivariate models (Rutledge *et al*., 2002; Gourvénec *et al*., 2002 and Gomez-Carracedo *et al*., 2007). Readers are referred to Dublin and Watson (1950) for the further details of DW.

In Support Vector Data Description (SVDD), it is assumed that the data is enclosed by a hyper sphere with minimum volume (Tax and Duin, 2004). We try to minimize probability of accepting outliers by minimizing the volume of the enclosed space. SVDD is able to make tighter boundaries around the data points. Readers are referred to (Tax and Duin, 2004) for the further details.

The work of SVDD is to map the target data nonlinearly into a higher dimensional feature space and construct a separating hyperplane with maximum margin there. It is probable to find the dividing hyperplane without explicitly carrying out the map into the feature space by using kernel function. The dot product $x_i \cdot x_j$ can also be avoided by implementing kernel function. Kernel function is any kind of a function that follows Mercer's Theorem (Sch¨olkopf *et al*., 1999). The most often used kernels are polynomial kernel and radial basis function (RBF) (Sun and Tsung, 2003). In this paper, RBF has been used as the kernel.

After introducing the basics of ICA, DW and SVDD, we describe the architecture of the proposed fault detection and diagnosis approach for non-Gaussian processes. The algorithmic procedure for ICA-DW-SVDD consists of two phases, offline training and online monitoring. After data preprocessing which includes centering and whitening of data the much advanced algorithm FastICA is applied first to obtain ICs. A large number of ICs causes involvement of noise in the data, therefore the DW algorithm is applied to get important ICs. They can be selected by observing the DW color plot. For training purpose selected important ICs are used to obtain SVDD parameters which are further used while monitoring of testing data. SVDD algorithm is used enclosing data points in much tighter boundary and for detecting outliers.

After detection of outliers, original variables are retraced in order to get a clear view of the factors affecting faults, and this is done with the help of proposed retracing algorithm.

A Simulation Example

In this paper, the proposed approach is applied to monitor a simple multivariate process, and this simulation work is similar to (Lee *et al*., 2004 and Ku *et al*., 1995) in which there are five monitored variables in a dynamic process as follows:

$$Z(k) = \begin{bmatrix} .118 & -.191 & .287 \\ .847 & .264 & .943 \\ -.333 & .514 & -.217 \end{bmatrix} \times Z(k-1) + \begin{bmatrix} 1 & 2 \\ 3 & -4 \\ -2 & 1 \end{bmatrix} \times U(k-1)$$

$$Y(k) = V(k) + Z(k)$$

where Y is the output with three variables (y_1, y_2, y_3). Y is the normal distributed random vector with zero mean and variance of 0.1. U is the input with

$$U(k) = \begin{bmatrix} .811 & -.226 \\ .477 & .415 \end{bmatrix} \times U(k-1) + \begin{bmatrix} .193 & .689 \\ -.320 & -.749 \end{bmatrix} \times W(k-1)$$

where W is a random vector following uniform distribution over interval (-2, 2). Input is U and output is Y, for process monitoring total five variables (y_1, y_2, y_3, y_4, y_5) are used.

A total of 1,000 observations are sampled for each simulation. The first 500 observations are used as a training data set and the remainders are used for on-line process monitoring. At observation 500, a step change of w_1 by 3 is introduced. This means that the first 500 training observations are not contaminated by outliers, and the remaining 500 data are faulty. The FastICA algorithm is applied to the generated data, and 5 ICs are obtained as a result of this. Results of DW criterion represents that 4 ICs are important, while 1 IC contains mainly noise. The 4 important ICs have been divided into target and outlier data sets. First 500 data are selected for testing, and the rest for outlier detection. In the training data, no outlier is selected. And results are obtained for number of normal data in test data as shown in Table 1. Actually, there is no normal data in test data.

Table 1: The results of simulation example.

Value of sigma	Detection rate (%)
3	99
4	100
5	100
6	98
7	96
8	96
9	94
10	92

There, for sigma equal to 4 and 5, error is equal to 0, i.e. every abnormal data has been verified as outlier data. We choose the value of sigma equals to 4 for further calculation. The hyper spherical boundary rejects 500 outliers, and makes tighter boundary enclosing only 500 normal data. Figure 5 illustrates the graph of distances of generated data from the centroid of the data points. Width parameter of RBF kernel function's is found out using sensitivity analysis and using sigma value equal to 4 in SVDD, it itself sets threshold according to the number of abnormal data we consider for training. It has been seen in Figure 1 that threshold has been set around 0.71 by SVDD to enclose all normal data and hence rejects all other points above threshold. Those points above the threshold should be considered as faulty.

Conclusion

In this paper, we proposed an approach based on ICA, DW and SVDD for the purpose of fault detection and diagnosis in non-Gaussian processes. The fault detection problem is converted to a one-class data description problem. The proposed monitoring method uses FastICA to get ICs. The DW algorithm is used to get important ICs, which further used in SVDD algorithm for enclosing data points in much tighter boundary and for detecting outliers. The model based on SVDD expresses fault-free data distribution using a hyper sphere with a close-fitting boundary. The proposed approach is validated on a simulation example.

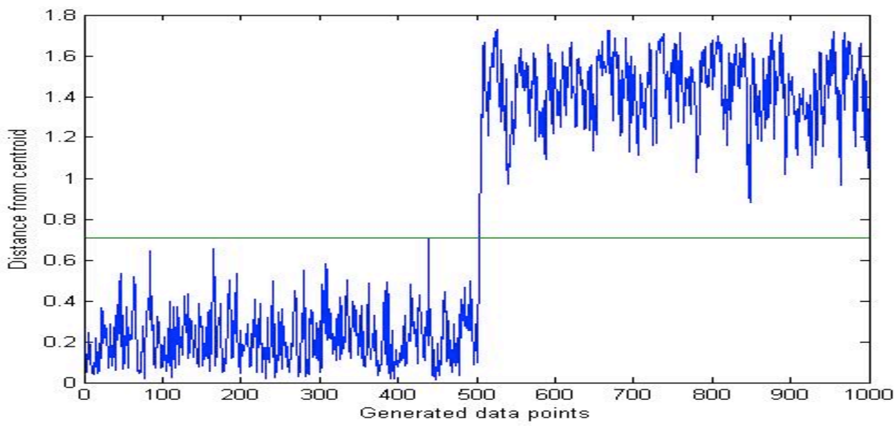

Figure 1: The distances of generated data from the centroid of data points.

Acknowledgements

This work is partially supported by National Science Council, Taiwan, R.O.C under grant NSC 100-2410-H-009-013-MY3.

References

Ammari, F., Cordella, C.B.Y., Boughanmi, N. and Rutledge D.N., 2011. Independent Component Analysis applied to 3D-front face fluorescence of spectra of edible oils to study the antioxidant effect of Nigella sativa L. extract on the thermal stability of heated oils, *Chemometrics and Intelligent Laboratory Systems*.

Bakshi, B.R., 1998. Multiscale PCA with application to multivariate statistical process monitoring, *American Institute of Chemical Engineering Journal* 44 (7), 1596–1610.

Chiang, L. H., Kotanchek, M. E. and Kordon A. K., 2004. Fault diagnosis based on Fisher discriminant analysis and support vector machines, *Computers & Chemical Engineering*, 28(8), 1389-1401.

Hsu, C.C., Chen, M.C and Chen, L.S., 2010. A novel process monitoring approach with dynamic independent component analysis, *Control Engineering Practice* 18(3), 242–253.

Durbin, J. and Watson, G.S., 1950. Testing for serial correlation in least squares regression*, Biometrika*, 37, 409–428.

Gomez-Carracedo, M.P., Andrade, J.M., Rutledge, D.N. and Faber, N.M., 2007. Selecting the optimal number of partial least squares components for the calibration of attenuated total reflectance mid-infrared spectra of undersigned kerosene samples, *Analytica Chimica Acta*, 585, 253-265.

Gourvénec, S., Massart, D.L. and Rutledge, D.N., 2002. Determination of the number of components during mixture analysis using the Durbin–Watson criterion in the orthogonal projection approach and in the simple-to-used interactive self modeling mixture analysis approach, *Chemometrics and Intelligent Laboratory Systems*, 61, 51-61.

Guo, Y., Na, J., Li, B. and Fung, R.F., 2014. Envelope extraction based dimension reduction for independent component analysis in fault diagnosis of rolling element bearing, *Journal of Sound and Vibration*, 333(13), 2983-2994.

Hyvärinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis, *IEEE Transactions on Neural Networks*, *10* (3), 626-634.

Hyvärinen, A. and Oja, E., 1997. A fast fixed-point algorithm for independent component analysis, *Neural Computation*, 9(7), 1483-1492.

Kano, M., Tanaka, S., Hasebe, S., Hashimoto, I. and Ohno, H., 2003. Monitoring independent components for fault detection, *American Institute of Chemical Engineering Journal*, 49 (4), 969-976.

Khediri, I.B., Limam, M. and Weihs, C., 2011. *Variable window adaptive Kernel Principal Component Analysis for nonlinear nonstationary process monitoring*, Computer and Industrial Engineering, 61(3), 437-446.

Ku, W., Storer, R.H. and Georgakis, C., 1995. Disturbance detection and isolation by dynamic principal component analysis, *Chemometrics and Intelligent Laboratory Systems*, 30, 179-196.

Lee, J.M., Qin, S.J. and Lee, I.B., 2007. Fault detection of non-linear process using kernel independent component analysis, *The Canadian Journal of Chemical Engineering*, 85, 526-536.

Lee, J.M., Yoo, C.K. and Lee, I.B., 2003a. Statistical process monitoring with independent component analysis, *Journal of Process Control*, 14 (5), 467-485.

Lee, J.M., Yoo, C.K. and Lee, I.B., 2003b. Statistical process monitoring with multivariate exponentially weighted moving average and independent component analysis, *Journal of Chemical Engineering of Japan*, 36 (5), 563-577.

Lee, J.M., Yoo, C.K. and Lee, I.B., 2004b. Statistical monitoring of dynamic processes based on dynamic independent component analysis, *Chemical Engineering Science*, 59, 2995-3006.

Lee, J.M., Yoo,C.K. and Lee,I.B. 2004a. Statistical process monitoring with independent component analysis, *Journal of Process Control*, 14, 467-485.

Luo, R., Misra, M. and Himmelblau, D.M., 1999. *Sensor fault detection via multiscale analysis and dynamic PCA*, *Industrial and Engineering Chemistry Research*, 38, 1489-1495.

Montgomery, D. C., 2001, *Introduction to Statistical Quality Control*, 4th ed., New York: Wiley.

Rato, T.J. and Reis, M.S., 2013. Fault detection in the Tennessee Eastman benchmark process using dynamic principal components analysis based on decorrelated residuals (DPCA-DR), *Chemomatrics and Intelligent Laboratory Systems*, 125,101-108.

Rutledge, D.N. and Barros, A.S., 2002. Durbin–Watson statistic as a morphological estimator of information content, *Analytica Chimica Acta*, 454, 277-295.

Rutledge, D.N. and Bouveresse, D.J.R., 2013. *Independent Component Analysis with Jade Algorithm*, 50, 22-32.

Sch¨olkopf, B., Burges, C.J.C. and Smola, A.J. 1999. *Advances in Kernel Methods.* The MIT Press.

Sun, R. and Tsung, F., 2003. A kernel-distance-based multivariate control chart using support vector methods, *International Journal of Production Research*, 41(13), 2975-2989.

Tax, D.M.J and Duin, R.P.W., 2004. Support vector data description, *Machine Learning*, 54, 45-66.

Tsung, F., 2000. Statistical monitoring and diagnosis of automatic controlled processes using dynamic PCA, *International Journal of Production Research*, 38 (3), 625-637.

Yoo, C.K., Choi, S.W. and Lee, I.-B., 2002. Dynamic monitoring method for multiscale fault detection and diagnosis in MSPC, *Industrial and Engineering Chemistry Research*, 41, 4303-4317.

Contact email: ittchen@mail.nctu.edu.tw