ICISE 2019

پنجمین کنفرانس بین المللی
مهندسی صنایع وسیستم‌ها

5th International Conference on Industrial and Systems Engineering
• 4 & 5 Sep. 2019 • Ferdowsi University of Mashhad

۱۳ و ۱۴ شهریور ۱۳۹۸ • دانشگاه فردوسی مشهد

High dimensional process monitoring using Principle Component Analysis and T^2 chart

Zahra Jalilibal
Department of Industrial Engineering, Shahed University
Tehran, Iran
zjalili222@gmail.com

Seyed Meysam Mousavi
Department of Industrial Engineering, Shahed University
Tehran, Iran

Amirhossein Amiri
Department of Industrial Engineering, Shahed University
Tehran, Iran

Abstract— ***Statistical process monitoring is an essential need for industrial processes. Many of these processes apply principal component analysis to perform statistical process monitoring as its simplicity in computations. The PCA is used in this study to reduce dimension for monitoring high dimensional process which has complex computations. Fault detection charts that are commonly employed with the PCA method are the Hotelling T^2 statistic which are used for monitoring the process which is reduced by PCA. This study has two steps; first, the high dimensional process is reduced by applying PCA, and then, the reduced process is monitored.***

Keywords— Principal component analysis, High dimensional process monitoring, dimension reduction, Hotelling T^2.

I. INTRODUCTION

Statistical process monitoring is an essential need for industrial processes. Chemical and manufacturing plants require process control to ensure that their product meets the characteristics specified and desired by the client, for example, specific electrical properties for a semiconductor being manufactured [1].

Various clients may need products with different electrical properties, and this means that the process needs to be able to focus according to the situation which is given, with minimum cost. economic benefits are the second incentive for process control, for example, continuous monitoring during all stages of a process helps keep the process to be in-control by ensuring that faulty product is detected early and corrective actions are taken immediately to rectify the fault. This causes saving both time and money [2]. Safety is the third advantage of process control; as early detection of abnormal sensor readings can help prevent catastrophe. control charts are used widely, which are developed in statistical process monitoring field to identify when a system is deviating from normal behavior. Statistical process monitoring's purpose is to detect deviations from normal process behavior during two distinct phases of the process, which are called Phase I, and Phase II.

One of the most critical methods for dimension reduction is principal component analysis (PCA) which can perform monitoring on high-dimensional processes. Two fault charts are often used with the PCA model [3]: the Hotelling T statistic that measures the variations in the PCA model, and the Q statistic, that measures variations in the residual space produced by the PCA model. These charts have mainly been used to detect shifts in the mean.

Multivariate statistical analysis and statistical process monitoring methods have been carried out in an extended range of fields such as genomics, signal processing, and various industrial processes[4-7]. Principal component analysis (PCA) is one of the most commonly used multivariate techniques with various applications. PCA preserves as much variability as possible of the dataset by finding a new set of variables or principal components (PCs) that are linear combinations of those in the original dataset that successively maximize variance and are uncorrelated with each other. These new sets of principle components are obtained by computing an eigenvalue problem. PCA captures the variance in m variables of the original dataset in a reduced dimension by l remained PCs. Thus, l is often much smaller than m. In PCA, since the derived PCs are uncorrelated, the distance between data points is preserved, leading to a diagonal covariance matrix. However, to satisfy these geometric constraints, most PCs contain non-zero loadings in all the coordinates.

Many processes need multiple process variables to be monitored in a continuous way, and these processes can be complicated and high dimensional. Principal component analysis (PCA) is one of the most standard statistical methods to discover common factors in the general multivariate setting [8-10].

PCA is a popular dimensionality reduction technique used by many industries[11], first as its simplicity. PCA method shows a data matrix using a lower number of retained variables known as principal components that capture most of the process variations. This increases the computational efficiency and decreases the time required to detect and signal faults, making it particularly advantageous for fault detection. These advantages make it a popular choice for monitoring industrial process [12].

978-1-5090-1076-9/16/$31.00 ©2017 IEEE

In most applications, the original variables in a dataset have a physical meaning, and PCA is especially useful if the resulting PCs are composed of a small number of the original variables. For about a decade, improving the interpretability of PCs has been a topic of active research [13-17]. Rotation of PCs is a common practice wherein the rotated components are more comfortable to interpret without any loss of information. Once PCs are rotated it is not possible to preserve the property that the components be pairwise uncorrelated or the loadings are orthogonal [15]. Rotating one also has to choose the normalization to preserve either orthogonality or zero correlation. Besides, different normalization criteria can lead to different quantitative results. Moreover, in conventional PCA, the variance captured by each PC decreases monotonically. However, once the components are rotated; this property does not always hold true.

In discussing the monitoring of multivariate processes, Bisgaard [18] highlights the PCA, partial least squares, factor analysis, and canonical correlation analysis as applicable monitoring methods. These methods and their extensions have the property that they are capable of handling high-dimensional process data and time-dependence. All of them project the high-dimensional process onto a lower dimensional subspace and monitor the process behavior with respect to it. Woodall and Montgomery [19] provide a survey of multivariate process-monitoring techniques as well as motivations for their use. Other books and papers devote more attention to PCA process monitoring. Kourti [20] describes fundamental control charting procedures for latent variables, including PCA and PLS, but does not discuss many of the main methods for time-dependent data nor their extensions. Kruger and Xie [21] include a chapter covering the monitoring of high-dimensional, time-dependent processes but focus on one method only. Qin [4] provides a review of fault detection, identification and reconstruction methods for PCA process monitoring.

II. RESEARCH METHODOLOGY

A. Principle Component Analysis (PCA)

The PCA model is a dimensionality reduction technique that has been utilized by a variety of engineers and researchers in order carry out process monitoring. For a given process with m process variables, and n collected observations, the data matrix $\boldsymbol{X}$ can be represented using a linear sum of given variables [15]:

$$\boldsymbol{X = TP^T}, \qquad (1)$$

where, T and P, represent the principal component (PCs), and loading vectors, respectively. The loading vectors, which are orthogonal, can be obtained from the covariance matrix $(\boldsymbol{\Sigma})$ of $\boldsymbol{X}$ as shown [3]:

$$\boldsymbol{\Sigma = \frac{1}{n-1} X^T X = P \Lambda P^T \; with \; PP^T = P^T P = I_m}, \qquad (2)$$

where, $\boldsymbol{\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_m)}$ is a diagonal matrix and contains the eigenvectors provided by the (m) PCs, while $\boldsymbol{I_m}$ is the identity matrix. In order to improve computational efficiency, PCA utilizes dimensionality reduction, where a given data set can be represented using a lower number of principal components, given that there is a correlation between the variables. Some methods can be utilized to decide how many principal components to retain, e.g., scree plot, profile likelihood [22], cross-validation [23], and cumulative percent variance (CPV) [24]. CPV is popularly used as it is simple to compute and often provides reliable results. CPV is computed as [3]:

$$\boldsymbol{CPV(l) = \frac{\sum_{i=1}^{l} \lambda_i}{trace(\Sigma)} \times 100}, \qquad (3)$$

The number of principal components to retain is determined based on the minimum number of process variables it takes to meet a certain CPV, e.g., 95-99%. The original data matrix can now be represented utilizing only the (l) retained PCs and $(m - l)$ ignored principal components as [3]:

$$\boldsymbol{X = TP^T = \left[\hat{T}\tilde{T}\right]\left[\hat{P}\tilde{P}\right]^T}, \qquad (4)$$

Where, $\hat{T} \in R^{n\times 1}, \tilde{T} \in R^{n\times(m-1)}, \hat{P} \in R^{m\times 1}, \tilde{P} \in R^{m\times(m-1)}$. expansion of Eq. (4) gives [15]:

$$\boldsymbol{X = \hat{T}\hat{P}^T + \tilde{P}\tilde{T}^T = X\hat{P}\hat{P}^T + X\left(I_m - \hat{P}\hat{P}^T\right) = \hat{X} + E}, \qquad (5)$$

Where $\hat{x}$ represents the predicted data using only the retained (l) PCs, and E represents the model residuals, respectively.

B. Hotelling's T^2 and SPE Statistics

When the PCA model is used for monitoring industrial processes, Hotelling's $T^2$of the retained PCs and the squared prediction error (SPE or Q) statistics of the PCA model are often used for fault detection [25]. T^2 index is the squared Mahalanobis distance of the retained PCs, designed to measure the variability of the mean and covariance within these PCs. Q statistic is the measure of the lack of fit for the PCA model. Due to the signal averaging aspect of the PCA, the combinations of variables obtained are often the robust indicators of process conditions compared to individual variables [26]. The procedure of using the conventional PCA to perform process monitoring is introduced in this subsection. When a new sample vector *x* (after scaling with the sample mean and variance obtained from the normal data, i.e., data from a "good" process) becomes available, it is projected with the help of the PCA model to PC space (PCS) and the residual space (RS). Hotelling's T^2 is calculated by the Mahalanobis distance of a score vector, *t*, in PCS, and Q is the Euclidean distance of a residual vector in the RS, which are given by [27]:

$$\boldsymbol{T_i^2 = t_i \Lambda t_i^T = x_i P_l \Lambda^{-1} P_l^T x_i^T} \qquad (6)$$

$$\boldsymbol{Q_i = e_i^T e_i} \tag{7}$$

Where $\boldsymbol{\Lambda = diag(\lambda_1, \lambda_2, \dots, \lambda_m)}$, $\boldsymbol{e_i}$ is the error at ith observation vector, and $I \in R^{m\times m}$ denotes the identity matrix. The status of the process can then be monitored by these two scalars $T^2$and Q.

The Hotelling's T^2measures the variation in latent variables sub-space [28]. The approximated control limits of Hotelling's T^2 statistic, with an acceptable false alarm rate, α, can be determined from the normal operating data in several ways by applying the probability distribution assumptions [4, 27]. The upper control limits (UCL) is calculated as follows:

$$\boldsymbol{T_\alpha^2 = \frac{(n-1)l}{(n-l)} F_\alpha(l, n-l)} \tag{8}$$

$\boldsymbol{F_\alpha(l, n-l)}$ is the upper α ، percentile of the F-distribution with freedoms degree of $\boldsymbol{l}$ and $\boldsymbol{(n-1)}$.

The process is considered faulty if either Hotelling's T^2 statistic for a new observation exceeds the control limit which is computed in Eq. (8). Once either Hotelling's T^2 statistic detects a fault, a contribution plot is often used to identify the root cause. Contribution plots provide information about which variables contribute to the distance between the points in a Hotelling's T^2 chart and the sample mean of the data. The high contribution of a variable implies it could be the main source of the faulty signal. The process is considered faulty if either Hotelling's T^2 statistics for a new observation exceeds the corresponding control limit.

III. Simulation Study

The data are generated from a normal distribution with mean vector [1,2,1,2,3,4,5,2,6,1,4,7] and orthogonal covariance matrix 2. To compute the in-control and out-of-control ARL, we used 100000 simulation runs. The upper control limit (UCL) for the proposed and T^2 chart is set by equation (8) to achieve a desired in-control ARL$_0$.

Simulated data is used in MINITAB 18, and PCA is applied on the dataset. Results are depicted in Table 1 which are the computations of eigenvalue, eigenvectors, the proportion of each variable variance regards to total variance and cumulative percent variance. Also, PC scores are calculated in Table 4, and coefficient of each variable are computed in Table 5 which are used in computing loadings.

By applying PCA method on the multivariate data, which is simulated, it is concluded from scree plot that omits five variables data and consider first 7 variables as the slope of figure from component number 7 to 8 has a considerable change, so that the dimension is reduced. The red line in Figure 2 demonstrates the frontier of components which should be separated into two groups, one of the groups remain in the process monitoring, and the other one omit form computations.

TABLE 1:PCA computations

Variable	Eigenvalue	Proportion	Cumulative
X1	2.17	0.181	0.181
X2	1.856	0.155	0.336
X3	1.63	0.136	0.471
X4	1.384	0.115	0.587
X5	1.148	0.096	0.683
X6	0.972	0.081	0.764
X7	0.823	0.069	0.832
X8	0.641	0.053	0.886
X9	0.523	0.044	0.929
X10	0.932	0.033	0.962
X11	0.264	0.022	0.984
X12	0.0192	0.016	1.000

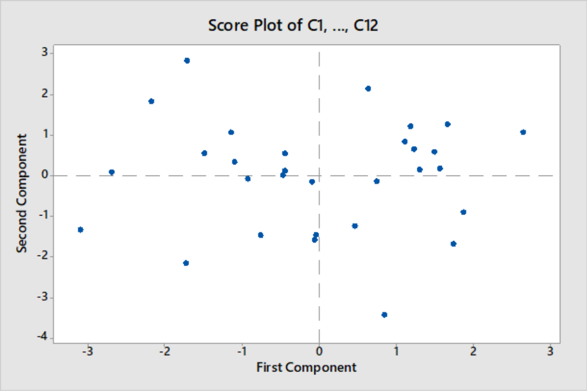

Fig. 1: Score plot of components

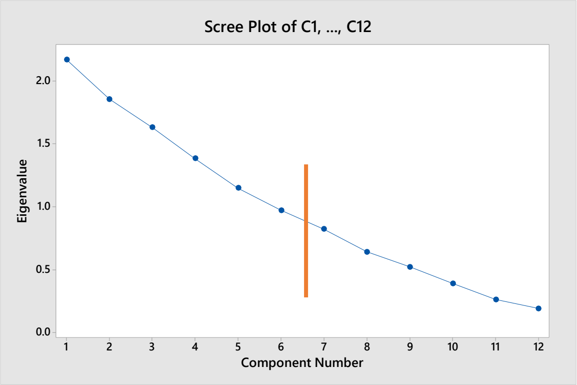

Fig. 2: Scree plot of variables

Table 6 presents the data left from the first step, (dimension reduction by PCA) which will be used, in the next step for monitoring the reduced process data set. 7 variables are left from dimension reduction and each variable has 30 observations. The average run length (ARL) is the average number of samples that need to be collected before an out-of-control alarm is declared. Theoretically, one would expect to

see a high ARL$_0$, when the process is in control, and a low ARL$_1$, when the process is out-of-control. In this study, to the ARL for Hotelling's T and Q statistic when using PCA are compared.
The control limit is computed according to Eq. (8):

$$T_{\alpha}^{2}=\frac{(n-1)l}{(n-l)}F_{\alpha}(l,n-l)=T_{0.05}^{2}=\frac{(12-1)\times 7}{(12-7)}F_{0.05}(7,12-7)$$
$$=15.4\times 3.97=61.138.$$

This quantity is substituted as UCL in the MATLAB code for monitoring step by T^2 chart.

A monitoring procedure can be implemented efficiently using a two-step algorithm. First, a sufficiently large number of observations from a good process are collected to obtain the PCA model. As the new observations become available, they are projected onto the loading matrix to obtain the corresponding PC scores. The monitoring statistics PCA T^2 is then calculated.
The complete monitoring procedure is summarized as follows:

1) Sufficient data is acquired when a process is operated under normal operating conditions. Each column (variable) of the data matrix is standardized, i.e., scaled to zero mean and unit variance.

2) PCA is applied to the data matrix, and the loadings $P\in R^{m\times m}$and the eigenvalue matrices $\boldsymbol{\Lambda}=\boldsymbol{diag}(\boldsymbol{\lambda_1},\boldsymbol{\lambda_2},\dots,\boldsymbol{\lambda_m})$ are obtained.

3) The threshold of cumulative percent variation, η is specified. The first l PCs that capture more than η percent variability are retained.

4) Loadings are obtained for l retained PCs.

5) The control limits are computed according to Eqs. (8).

TABLE 2: ARL1 (detecting shift in the mean)

Fault no.	Type	Shifts	ARL(PCA+T^2)
1	**Mean-3**	**-3**	**1.0490**
2	**Mean-2.5**	**-2.5**	**1.2710**
3	**Mean-2**	**-2**	**2.0880**
4	**Mean-1.5**	**-1.5**	**5.0500**
5	**Mean-1**	**-1**	**18.3550**
6	**Mean-0.5**	**-0.5**	**86.6430**
7	**Mean**	**0**	**196.1900**
8	**Mean+0.5**	**0.5**	**94.5780**
9	**Mean+1**	**1**	**19.6500**
10	**Mean+1.5**	**1.5**	**5.2090**
11	**Mean+2**	**2**	**2.0790**
12	**Mean+2.5**	**2.5**	**1.2700**
13	**Mean+3**	**3**	**1.0380**

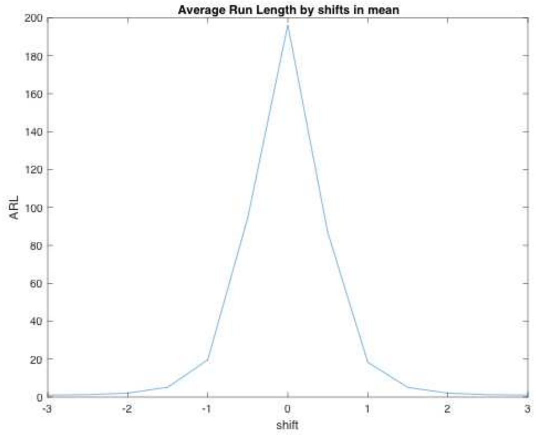

Fig 3:ARL1 (detecting shift in the mean)

IV. CONCLUSION

PCA is often used to transform the variables to a reduced dimension. Traditionally, the Hotelling's T^2 statistics are used for process fault detection with PCA.
By no means is this study, we propose a hybrid method for monitoring multivariate high dimensional process and future work should most certainly be focused on handling autocorrelation and non-stationarity simultaneously in a multivariate high dimensional process. Nowadays, works have focused on examining the performance of methods intended for only one type of dynamic data, but combinations of the two remain unexplored.

REFERENCES

1. Qin, S.J., Cherry G., Good R. ., *Semiconductor manufacturing process control and monitoring: A fab-wide framework.* Journal of Process Control, 2006. **16**(3): p. 179-191.
2. Shewhart, W. and W. Deming, *Statistical method from the viewpoint of quality control. Pennsylvania, PA*. 1939, Lancaster press, Inc.
3. Sheriff, M.Z., Botre, C. Mansouri, M., *Process Monitoring Using Data-Based Fault Detection Techniques: Comparative Studies*, in *Fault Diagnosis and Detection*. 2017, InTech.
4. Joe Qin, S., *Statistical process monitoring: basics and beyond.* Journal of Chemometrics: A Journal of the Chemometrics Society, 2003. **17**(8-9): p. 480-502.
5. Gibbons, M.R., *Multivariate tests of financial models: A new approach.* Journal of financial economics, 1982. **10**(1): p. 3-27.
6. Castells, F., Laguna, P. Sornmo, L. Bellman,A. Principal component analysis in ECG signal processing. EURASIP Journal on Advances in Signal Processing, 2007. (1): p. 074580.
7. Zou, H., T. Hastie, and R. Tibshirani, *Sparse principal component analysis.* Journal of computational and graphical statistics, 2006. **15**(2): p. 265-286.
8. Jackson, J.E., *Principal components and factor analysis: part I—principal components.* Journal of Quality Technology, 1980. **12**(4): p. 201-213.
9. Jackson, J.E., *Principal components and factor analysis: part II—additional topics related to principal components.* Journal of Quality Technology, 1981. **13**(1): p. 46-58.
10. Johnson, R.A. and D. Wichern, *Applied multivariate statistical analysis. Prentice Hall, Englewood Cliffs, NJ.* Applied multivariate statistical analysis. Prentice-Hall, Englewood Cliffs, NJ., 1992: p12-17.

11. George, J.P., Z. Chen, and P. Shaw, *Fault detection of drinking water treatment process using PCA and hotelling's T2 chart.* World Academy of Science, Engineering and Technology, 2009. **50**: p. 970-975.

12. Chiang, L.H., E.L. Russell, and R.D. Braatz, *Fault detection and diagnosis in industrial systems*. 2000: Springer Science & Business Media.

13. Jolliffe, I.T., *Introduction*. 2002: Springer.

14. Jolliffe, I.T., *Rotation of principal components: some comments.* Journal of Climatology, 1987. **7**(5): p. 507-510.

15. Jolliffe, I.T., *Rotation of principal components: choice of normalization constraints.* Journal of Applied Statistics, 1995. **22**(1): p. 29-35.

16. Jolliffe, I.T., *Rotation of ill-defined principal components.* Applied Statistics, 1989: p. 139-147.

17. Richman, M.B., *Rotation of principal components.* Journal of climatology, 1986. **6**(3): p. 293-335.

18. Bisgaard, S., *The future of quality technology: From a manufacturing to a knowledge economy & from defects to innovations.* Quality Engineering, 2012. **24**(1): p. 30-36.

19. Woodall, W.H. and D.C. Montgomery, *Some current directions in the theory and application of statistical process monitoring.* Journal of Quality Technology, 2014. **46**(1): p. 78-94.

20. Kourti, T., *Application of latent variable methods to process control and multivariate statistical process control in industry.* International Journal of adaptive control and signal processing, 2005. **19**(4): p. 213-246.

21. Kruger, U. and L. Xie, *Statistical Monitoring of Complex Multivatiate Processes: With Applications in Industrial Process Control*. 2012: John Wiley & Sons.

22. Zhu, M. and A. Ghodsi, *Automatic dimensionality selection from the scree plot via the use of profile likelihood.* Computational Statistics & Data Analysis, 2006. **51**(2): p. 918-930.

23. Diana, G. and C. Tommasi, *Cross-validation methods in principal component analysis: a comparison.* Statistical Methods and Applications, 2002. **11**(1): p. 71-82.

24. Mansouri, M., et al., *Statistical fault detection of chemical process-comparative studies.* J. Chem. Eng. Process Technol, 2016. **7**(1): p. 1-10.

25. Chiang, L.H., E.L. Russell, and R. Braatz, *Fault detection and diagnosis in industrial systems*. 2001, IOP Publishing.

26. Wise, B.M. and N.B. Gallagher, *The process chemometrics approach to process monitoring and fault detection.* Journal of Process Control, 1996. **6**(6): p. 329-348.

27. Kourti, T. and J.F. MacGregor, *Process analysis, monitoring and diagnosis, using multivariate projection methods.* Chemometrics and intelligent laboratory systems, 1995. **28**(1): p. 3-21.

28. Tong, C.-d., X.-f. Yan, and Y.-x. Ma, *Statistical process monitoring based on improved principal component analysis and its application to chemical processes.* Journal of Zhejiang University SCIENCE A, 2013. **14**(7): p. 520-534.