Journal of Fundamental Sciences Vol.7, No.2 (2011) 120-125.

ISSN 1823-626X

Journal of Fundamental Sciences

available online at http://jfs.ibnusina.utm.my

Measuring the Performance of an Eigenvalue Control Chart for Monitoring Multivariate Process Variability

Aishah Mohd Noor[1,2*] and Maman Abdurachman Djauhari[1]

[1]Department of Mathematics, Faculty of Science, Universiti Teknologi Malaysia, UTM Skudai, Johor, Malaysia
[2] Institute for Engineering Mathematics, Universiti Malaysia Perlis, UniMAP, Kangar, Perlis, Malaysia

Received 4 July 2011, Revised 25 August 2011, Accepted 10 September 2011, Available online 4 December 2011

ABSTRACT

In manufacturing process, it is very important to control and monitor the stability of a process such that a high quality product will be produced. The most common statistical tool used for monitoring the stability of a process is the control chart. In recent applications of control charting methods, there is a need to construct a control chart that is able to represent the behaviour of a multivariate process since in many manufacturing processes; quality of a product is determined by the joint-level of several quality characteristics. For this reason, in this paper, a new control chart is introduced for monitoring the stability of multivariate process in terms of the process variability. The proposed method is based on charting each of the eigenvalues of a covariance matrix. To show the efficiency of the proposed method, we conduct a simulation study and compare the performance of the proposed method with the existing method. A real example will be presented to illustrate the advantage of our proposed method.

|Statistical Process Control| Multivariate Dispersion| Covariance Matrix | Generalized Variance | Vector Variance |

http://dx.doi.org/10.11113/mjfas.v7n2.249

1. INTRODUCTION

Control chart is the most magnificent statistical and graphical tool in process control and monitoring. In multivariate data environment where the quality characteristic, p is greater than 1, $p > 1$ and these quality characteristics are correlated, the univariate control chart is no longer sufficient to represent shift in a process [1,2,3]. This is due to an increase of probability of false positive alarm, α . In statistical process control (SPC) methodology, the process monitoring via control charting method works like a series of hypothesis testing where α is refer to the probability of rejecting null hypothesis, H_0 while actually H_0 is true. In here, H_0 represent that a process has not shifted or the process remains in-control. In contrast, the alternative hypothesis, H_1 for which we are interested in investigating the occurrence of process shift. Thus, process monitoring is equivalent to conducting a repeated hypothesis testing on the process parameters of interest [1,3]. In multivariate process, there are two important population parameters to be monitored. One is the process target given by the mean vector $\boldsymbol{\mu}$ and the second is the process variability which characterized by the covariance matrix, $\Sigma_{p\times p}$.

*Corresponding author at: Department of Mathematics, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia
E-mail address: aishah.mn@unimap.edu.my (Aishah Mohd Noor)

In general, an efficient and effective control chart is a control chart that is able to detect a process shift quickly so that the production process will be stopped as soon as possible when out-of-control signal is occurred in order to take corrective actions [4]. In addition, a control chart with less sensitive in detecting out-of-control signal when actually the process is in-control will be desired [5]. When the process is in-control, α (also called as Type I error) should be small. If α is large, it would increase the cost of inspection since the product is actually 'good' but the process is stopped due to the false out-of-control signal. On the other hand, β or Type II error is the probability of failing to reject H_0 while actually the process has shifted. If β is large it would be crucial importance to take careful corrective actions since a large amount of 'bad' product are passed during the inspection process instead of making the true decision to quickly detect the process shift and stop the process.

In SPC there are two different phases, phase I and phase II stage. According to [1,6], phase I is used as a retrospective study where in this phase, the unknown process parameters, $\boldsymbol{\mu}$ and Σ are being estimated. Moreover, in this phase the appropriate control limits are calculated in order to establish an in-control state. Once we achieved in-control state, consequently for phase II study, our focus is in monitoring the future observations.

In this paper, we introduced a new control charting method for monitoring multivariate process variability for phase II stage. This paper is organized as follows; in section

2, we recalled one of the existing control chart based on vector variance (VV) statistic. The proposed control charting method is discussed in section 3. In section 4, a simulation experiment is presented to investigate the performance of proposed chart compared with VV chart based on chart's average run length (ARL). To show the advantage of our proposed method, a real industrial example is illustrated in section 5. Finally, a conclusion based on simulation results is discussed.

2. VECTOR VARIANCE CHART

Assume that a historical data set (HDS) is available from a process during phase I stage. Let $\mathbf{X}_{ij} = \left(X_{ijk}\right)^T$ be the p x 1 observations vector which represent the jth observation measure on kth quality characteristic of corresponding ith subgroup of size n for j = 1,2,...,n, k = 1,2,…,p and i = 1,2,…,m . Suppose $\mathbf{X}_{ij} = \left(X_{ijk}\right)^T$ follows a p-variate normal distribution, $N_p(\boldsymbol{\mu}, \Sigma)$ and the observations vectors are independent and identically distributed (i.i.d). In our study, the subgroup m is of each size n > 1. In multivariate setting, process variability is characterized by a p x p positive definite covariance matrix Σ [7]. When the process is in-control state, the HDS which consist of $\mathbf{X}_{ij}$'s are used to estimate the in-control process parameters denoted by $\boldsymbol{\mu}_0$ and Σ_0 respectively such that $\mathbf{X}_{ij}'s \sim N_p(\boldsymbol{\mu}_0, \Sigma_0)$.

In this study, we focused on the occurrence of shift in a process covariance matrix in phase II stage for which $\mathbf{Y}_{ij}'s \sim N_p(\boldsymbol{\mu}_0, \Sigma_f)$. Let the ith subgroup sample covariance matrix is denoted by S_i and is estimated from HDS, hence S_i is given by:

$$\mathrm{S}_i = \frac{1}{n-1}\sum_{j=1}^{n}\left(\mathbf{X}_{ij} - \overline{\mathbf{X}}_i\right)\left(\mathbf{X}_{ij} - \overline{\mathbf{X}}_i\right)^T$$

where $\overline{\mathbf{X}}_i = \left(\overline{X}_{i1}, \overline{X}_{i2}, ..., \overline{X}_{ip}\right)^T$ and

$$\overline{\mathbf{X}}_i = \left(\frac{1}{n}\sum_{j=1}^{n} X_{ij1}, \frac{1}{n}\sum_{j=1}^{n} X_{ij2}, ..., \frac{1}{n}\sum_{j=1}^{n} X_{ijp},\right)^T \text{ for } i = 1,2,\ldots,m.$$

When m subgroups are obtained and the sample mean vector $\overline{\mathbf{X}}_i$ is calculated for each subgroup, thus the average of all sample mean vectors is denoted by $\overline{\overline{\mathbf{X}}}$ while the average of all sample covariance matrices is given by $\overline{\mathrm{S}}$. These statistics $\overline{\overline{\mathbf{X}}}$ and $\overline{\mathrm{S}}$ could be written in the form of:

$$\overline{\overline{\mathbf{X}}} = \left(\frac{1}{m}\sum_{i=1}^{m}\overline{\mathbf{X}}_{i1}, \frac{1}{m}\sum_{i=1}^{m}\overline{\mathbf{X}}_{i2}, ..., \frac{1}{m}\sum_{i=1}^{m}\overline{\mathbf{X}}_{ip},\right)^T \text{ and } \quad \overline{\mathrm{S}} = \frac{1}{m}\sum_{i=1}^{m} S_i$$

respectively.

We first provide an overview of the most widely used generalized variance (GV) chart. GV chart is proposed by [1]. This chart is based on plotting the ith future sample GV, $\left|S_f\right|$ developed by [8]. The control limits of this chart are constructed using the first and second moment of $\left|S_f\right|$. The limitation of GV lies on its property of determinant of a matrix. If GV = 0, that is in this case there exist a linear combination of at least one variable to another variables or there is a variable with zero variance [9]. Due to the limitation of GV, [9] proposed a new statistical measure for monitoring multivariate process variability. If sample GV is defined by $\left|S_f\right|$, VV is calculated from the sum of all diagonal elements of S_f^2, $Tr\left(S_f^2\right)$. For geometric interpretation of VV, [10] discussed the details. Unlike GV chart, VV chart is more sensitive to small shift.

Based on the asymptotic distribution of sample VV and for m subgroups of each size n available in phase I stage, $\overline{S}$ is calculated from phase I and with probability of false alarm $\alpha = 0.0027$, the UCL and LCL for VV chart is given by (see [11] for details explanation):

$$UCL = \hat{\theta} + 3\hat{\eta}\frac{\sqrt{n}}{n-1}$$

$$LCL = \hat{\theta} - 3\hat{\eta}\frac{\sqrt{n}}{n-1}$$

where

$$\hat{\theta} = \frac{n+1}{n-1}Tr\left(\overline{S}^2\right)$$

and

$$\hat{\eta}^2 = 8Tr\left(\overline{S}^4\right)$$

When LCL is negative, we set LCL to zero. Thus, the process is said to be out-of-control when there is a point plots outside the control limits above.

It can be shown that both GV and VV statistics are function of real and independent eigenvalues. If GV is the product of all eigenvalues, VV is the sum of squares of all eigenvalues. From the development of GV chart by [1] in 1985 and the new measure of process variability based on VV in 2008 by [9], and later on, [11] proposed to simultaneously used GV and VV in order to describe a better understanding of changes in covariance structure. [1,12,13] discussed the limitation of GV statistic as a measure of multivariate process variability. They showed that two sample covariance matrices may give the same GV but the correlation structures of both matrices are different. Similarly, two sample covariance matrices may give the same VV but different GV. Thus, based on these limitations, we proposed to monitor each eigenvalues of a covariance matrix in order to further investigate the behaviour of shift in covariance structure.

3. PROPOSED CHART

Let $\boldsymbol{\lambda}^{\bar{S}} = \left[\lambda_1^{\bar{S}}, \lambda_2^{\bar{S}}, ..., \lambda_p^{\bar{S}}\right]^T$ be the vector of eigenvalues of $\bar{S}$. Suppose $\hat{\lambda}_{ij}$ for j = 1,2, …,p and i = 1,2, …,m is the jth eigenvalue of ith future covariance matrix Σ_f and $\hat{\lambda}_j$ for all j are real and independent of each other [14]. The probabilistic distribution of $\hat{\lambda}_j$ will be used in order to determine the control limits for the proposed chart such that $1 - P(LCL_{\hat{\lambda}_j} < \hat{\lambda}_j < UCL_{\hat{\lambda}_j}) = \sqrt[p]{\alpha}$ where $\alpha = 0.0027$. We employed the conventional univariate Shewhart control limits in order to compute the control limits. Recall that Σ_f is a positive definite matrix and $\hat{\lambda}_j$ of Σ_f are assumed distinct (see theorem 8.3.3 [14]), the asymptotic distribution of $\hat{\lambda}_j$ is said to be :

$$\hat{\lambda}_j \sim N\left(\lambda_j^{\Sigma_0}, \frac{2}{n-1}\left(\lambda_j^{\Sigma_0}\right)^2\right)$$

for j = 1, 2,.., p and $\lambda_j^{\Sigma_0}$ is the jth eigenvalue of Σ_0. It should be noted that in our study Σ_0 is estimated by $\bar{\mathrm{S}}$. Thus, the associated control limits with $\alpha = 0.0027$ for an individual eigenvalue chart are given by:

$$\text{UCL} = \lambda_j^{\Sigma_0} + L\left(\sqrt{\frac{2}{n-1}\left(\lambda_j^{\Sigma_0}\right)^2}\right)$$

$$\text{LCL} = \lambda_j^{\Sigma_0} - L\left(\sqrt{\frac{2}{n-1}\left(\lambda_j^{\Sigma_0}\right)^2}\right)$$

for j = 1, 2, …, p , n is the sample size of ith future subgroup and $L = \Phi(Z_v)$ where $v = \sqrt[p]{\alpha}/2$ and $\Phi(.)$ is the cumulative distribution function of the standard normal distribution. According to the control limits above, a process is said to be out-of-control when at least one of $\hat{\lambda}_j$ gives an out of control signal.

4. PERFORMANCE ANALYSIS

When constructing a control chart, the aspect of how quick is the chart signal when actual change is occurred will be crucial importance. In addition, this chart should be less sensitive to the probability of false positive alarm α and false negative alarm β. As discussed in section 1, process monitoring is equivalent to conducting a repeated hypothesis testing. The average run length (ARL) is used to measure the performance of proposed chart. When the process is in-control state, the in-control ARL is denoted by ARL_0 where $ARL_0 = \frac{1}{\alpha}$. On the other hand, when the process is in out-of-control state or there is an occurrence of shift in process covariance matrix or changes in covariance structure, the out-of-control ARL is denoted by ARL_1 and is given by $ARL_1 = \frac{1}{1-\beta}$. Thus, from the above definition, the power of a control chart is depend on how small is it's ARL_1 and how large is it's ARL_0. A large ARL_0 will be required so that the probability of false positive alarm is nearly small, that is the process is declare out-of-control when in fact that the process has not shifted. In contrast, a small ARL_1 is desired since we want to detect the occurrence of shift in the process covariance matrix as soon as the process has shifted.

In order to show the performance of our proposed method, a simulation study is conducted using MATLAB R2010b to compare the performance of VV chart and the proposed chart based on ARL_1. In this paper we presented ARL_1 for only the largest eigenvalue (LEV) chart. For this experiment, we generated random data from in-control process $N_p(\mathbf{0}, I_p)$ where m = 100 of each size $n = 30$ with p = 2 and p = 3 in order to calculate the control limits. The run length is then computed based on the number of subgroups required before the first out-of-control signal is occurred. In this experiment, ARL_1 is calculated from the average of 1000 simulated run lengths. In order to provide a meaningful comparison, we set the control limits for all charts so that $ARL_0 \approx 250$.

It should be noted that the simulated ARL_0 will be different than that theoretical value $ARL_0 \approx 370$ for $\alpha = 0.0027$ when parameters are estimated and there are limited number of observations available for setting the control limits. We consider three types of shift for Σ_f in order to measure the performance of proposed chart compared with VV chart. The types of shift considered in our experiment are similar to the one presented in [9].

Suppose in-control covariance matrix:

$$\Sigma_0 = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} = I_p$$

The types of shift in Σ_f in order to calculate ARL_1 are:

1. The shift in σ_1^2 where $\Sigma_f = \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$
2. The shift in σ_1^2 and σ_2^2 where $\Sigma_f = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$
3. The covariance shift where $\Sigma_f = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$

The results of our experiments for above types of shift for Σ_f are presented in Table 1.0, Table 2.0 and Table 3.0 respectively. For the first type of shift, we learned that ARL_1 for LEV chart is smaller than ARL_1 for VV chart. However, in Table 2.0, when variance shifts in all variables, ARL_1 for VV chart is smaller than ARL_1 for LEV chart. Interestingly, in Table 3.0, ARL_1 for LEV chart outperform than that ARL_1 for VV chart. The experiment results for p = 3 are not reported in here and are equivalent to the results when p = 2.

Table 1 ARL_1 (VV) and ARL_1 (LEV) for shift in first variable

k	VV	LEV
1.1	141.49	134.72
1.2	57.62	47.09
1.3	34.31	27.01
1.4	15.29	14.74
1.6	6.58	5.733
1.8	3.51	3.23
2	2.20	1.84

Table 2 ARL_1 (VV) and ARL_1 (LEV) for shift in both variables

k	VV	LEV
1.1	53.97	101.42
1.2	15.86	22.11
1.3	8.50	13.17
1.4	4.38	6.29
1.6	1.92	2.32
1.8	1.39	1.61
2	1.14	1.29

Table 3 ARL_1 (VV) and ARL_1 (LEV) for covariance shift

ρ	VV	LEV
0.1	210.38	172.42
0.2	117.81	86.15
0.3	63.91	31.56
0.4	37.27	14.96
0.5	22.70	12.10
0.6	13.78	7.45
0.7	9.77	4.83
0.8	7.51	3.33
0.9	4.82	2.65

5. AN EXAMPLE

An example of proposed method applied to a real industrial problem in order to monitor a flange production is presented in this section. By using HDS for m = 20, n = 5 and p = 3 for which the characteristics are thickness at the nozzle, thickness at the wall and thickness at the base, $\overline{S}$ is given by:

$$\overline{S} = \begin{bmatrix} 0.5643 & 0.1122 & 0.0467 \\ 0.1122 & 0.3020 & 0.0503 \\ 0.0467 & 0.0503 & 0.2675 \end{bmatrix}$$

It should be noted that $\overline{S}$ is the average of all covariance matrices issued from HDS. The associated vector of eigenvalues of $\overline{S}$ is given by:

$$\boldsymbol{\lambda} = \left(0.6165 \quad 0.2892 \quad 0.2280\right)^T$$

For phase II stage, 10 new subgroups of equal size are collected in order to monitor the shift in covariance structure. In Table 4.0, we presented the values of VV and the largest eigenvalue (LEV) of the future sample covariance matrices S_f for f = 1,2, …,10. To construct the control limits for our proposed chart, we first determine the value of $L = \Phi(Z_v)$ as discussed in section 3. Hence, for m = 10, n = 5 and p = 3, the upper and lower control limits of proposed chart are UCL = 1.2610 and LCL = 0. Next, for VV chart, we obtained the values of $Tr\left(\overline{S}^2\right)$ and $Tr\left(\overline{S}^4\right)$ where the values are given by 0.5157 and 0.1542 respectively. Here, $\omega_0 = 0.7735$ and $\eta^2 = 1.2332$. Consequently, the control limits for VV chart are UCL = 2.6359 and LCL = 0. By using the control limits above for both control charts and results in Table 4.0, we then plotted the corresponding control charts given by Figure 1.0 and Figure 2.0 respectively.

Table 4 The values of VV and the largest eigenvalue for each new subgroup

m	VV	LEV
1	0.2376	0.48609
2	0.0783	0.25272
3	0.1963	0.42979
4	0.6397	0.77145
5	0.3701	0.58618
6	0.0407	0.19852
7	0.0438	0.20861
8	0.1856	0.41425
9	1.8894	1.35460
10	0.1770	0.41204

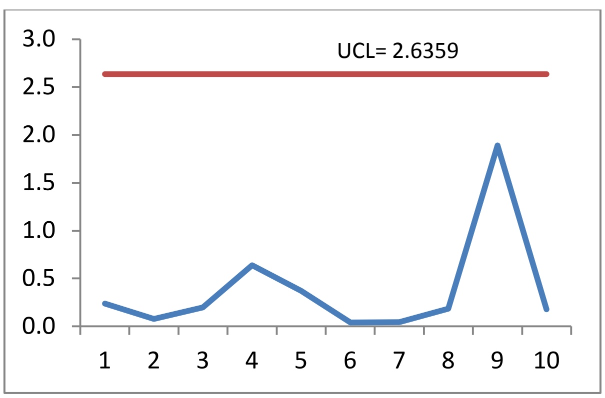

Fig. 1 VV chart

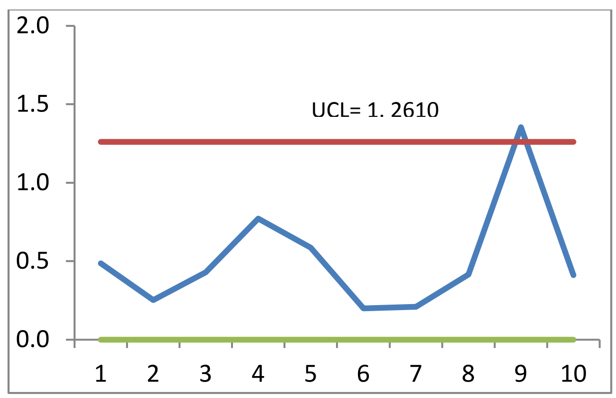

Fig. 2 $\hat{\lambda}_1$ chart

In Figure 2.0, there is an out-of-control signal occurred at the 9th subgroup. However, the signal is not able to detect by using VV chart. This example illustrates the effectiveness of proposed control chart compared with the VV chart for detecting the changes in process variability. This finding is in line with our simulation experiment for which ARL_1 for LEV chart outperform than ARL_1 for VV chart when the quality characteristics are correlated for which this is our main reason in choosing multivariate control chart.

6. CONCLUSION

In this paper, we presented a multivariate Shewhart control chart for monitoring multivariate process variability. The proposed method is promising since the statistic is independent of each other. The simulation experiments indicate that LEV chart and VV chart could describe a different behaviour of covariance structure. We learn that the covariance shift of a covariance matrix could be best explained by our proposed chart.

ACKNOWLEDGEMENTS

We sincerely acknowledge financial support from Ministry of Higher Education (MOHE) under the FRGS vote number 4F013 and Research University grant (RUG) vote number Q.J13000.7126.02H18. We also would like to thank the Government of Malaysia and University Technology Malaysia (UTM) in supporting research projects. The first author would like to express her gratitude to University Malaysia Perlis (UniMAP) for the opportunity to conduct her PhD research at UTM.

REFERENCES

[1] F. B. Alt, ESS, John Wiley & Sons, New York, 5(1985), 110-119.
[2] F. B. Alt and N. D. Smith, Handbook of Statis., Elsevier Science Publisher, North Holland, 7(1988), 333-351.
[3] D. C. Montgomery, Introduction to Statistical Quality Control, John Wiley & Sons, New York, 2009.
[4] F. Aparisi, J. Jabaloyes and A. Carrion, Comm. Statist. Theory Methods, 28(1999), 2671–2686.
[5] W. H. Woodall, J. Qual. Techno., 32(2000), 341-350.
[6] N. J. A. Vargas and C. J. Lagos, Qual. Eng., 19(2007), 191-196.
[7] A. Grigoryan and D. He, Int. J. Prod. Res., 43(2005), 715-730.
[8] S. S. Wilks, Biometrika, 24(1932), 471-494.
[9] M. A. Djauhari, M. Mashuri and D. E. Herwindiati, Comm. Statist. Theory Methods, 37(2008), 1742-1754.
[10] M. A. Djauhari, MATEMATIKA, 27(2011), 51-57.
[11] M. A. Djauhari, IEEE Xplore, (2011), 620-625.
[12] R. A. Johnson and D. W. Wichern, Applied Multivariate Statistical Analysis, Prentice Hall, New Jersey, 2007.
[13] M. A. Djauhari and M. Ismail, S. Afr. J. Ind.Eng., 21(2010), 207-215.
[14] K. V. Mardia, J. T. Kent and J. M. Bibby, Multivariate Analysis, *Academic Press, New York*, 1979.