CUSUM Chart to Monitor Dispersion Matrix for Multivariate Normal Process

Duk-Joon Chang[1] · Yong-Man Kwon[2] · Yeon-Woong Hong[3]

Abstract

Cumulative sum(CUSUM) control charts for monitoring dispersion matrix under multivariate normal process are proposed. Performances of the proposed CUSUM charts are measured in terms of average run length(ARL) by simulation. Numerical results show that small reference values of the proposed CUSUM chart is more efficient for small shifts in the production process.

Keywords : dispersion matrix, ARL, sample statistics

1. Introduction

In many industries, there exist multiple quality variables to define the quality of output and the quality is often characterized by joint levels of quality variables rather than a single quality characteristic. And shifts in the components of dispersion matrix for the related quality variables are often important.

The CUSUM chart directly incorporates all of the information in the sequence of sample values by plotting the cusum of the deviation of the sample values from the target value. The standard Shewhart chart, although simple to understand and apply, uses only the informations in the current sample and is thus relatively inefficient in detecting small shifts of the process. CUSUM chart is a good alternative to the Shewhart chart when the detection of small shifts in a process are important.

The multivariate procedures to quality control were first introduced by Hotelling (1947) and became popular in recent years. Woodall and Ncube(1985) considered a single multivariate CUSUM procedure for monitoring the means of multivariate normal process. They used p two-sided univariate CUSUM charts simultaneously, and evaluated the performance of the collection of the scheme. Croiser(1988) and Pignatiello and Runger(1990) considered new multivariate CUSUM control schemes

1) Professor, Department of Statistics, Changwon National University, Changwon, 641-773, Korea
E-mail : djchang@sarim.changwon.ac.kr

2) Associate Professor, Department of Computer Science and Statistics, Chosun University, Gwangju, 501-759, Korea

3) Associate Professor, Division of Internet Industrial Engineering, Dongyang University, Yongju, 750-012, Korea

that accumulate past sample information for each parameter and then form a univariate CUSUM statistic from the multivariate data for monitoring the mean vector of a multivariate normal process.

In this paper, we propose multivariate CUSUM procedures for monitoring the dispersion matrix Σ of correlated quality characteristics under multivariate normal process.

2. Multivariate Sample Statistics

Assume that the quality vector $\underline{X}' = (X_1, X_2, \cdots, X_p)$ are jointly distributed as p-variate normal distribution $N_p(\underline{\mu}, \Sigma)$. We take a sequence of independent random vectors $\underline{X}_1, \underline{X}_2, \underline{X}_3, \cdots$, where $\underline{X}_i = (\underline{X}'_{i1}, \cdots, \underline{X}'_{in})'$ is a sample of observations at the sampling time i and $\underline{X}_{ij} = (X_{ij1}, \cdots, X_{ijp})'$.

Let $\underline{\theta}_0 = (\underline{\mu}_0, \Sigma_0)$ be the known target process parameters for $\underline{\theta} = (\underline{\mu}, \Sigma)$ of p quality variables, where $\underline{\mu}$ is mean vector and Σ is dispersion matrix of $\underline{X}'$. For simplicity in this paper, we assume that $\underline{\mu}_0 = \underline{0}'$, all diagonal elements of Σ_0 are 1 and off-diagonal elements of Σ_0 are 0.3.

For the multivariate procedure for monitoring Σ, we can consider a multivariate version of $T_i = (n-1)S_i^2/\sigma_0^2$ in univariate procedure as

$$V_i = \sum_{j=1}^{n}(\underline{X}_{ij} - \underline{\overline{X}}_i)'\Sigma_0^{-1}(\underline{X}_{ij} - \underline{\overline{X}}_i) = tr(A_i\Sigma_0^{-1}) \tag{2.1}$$

where the statistic $A_i = \sum_{j=1}^{n}(\underline{X}_{ij} - \underline{\overline{X}}_i)(\underline{X}_{ij} - \underline{\overline{X}}_i)'$. Hotelling(1947) proposed that the statistic V_i can be used to monitor the process dispersion matrix Σ of p quality variables.

Another sample statistic for Σ can be considered by using the likelihood ratio test(LRT) statistic for testing $H_0: \Sigma = \Sigma_0$ vs $H_1: \Sigma \neq \Sigma_0$ where $\underline{\mu}_0$ is known. The regions above the upper control limit(UCL) corresponds to the rejection region. For the ith sample, likelihood ratio λ_i can be expressed as

$$\lambda_i = n^{-\frac{np}{2}} \cdot |A_i\Sigma_0^{-1}|^{\frac{n}{2}} \cdot \exp\left[-\frac{1}{2}tr(\Sigma_0^{-1}A_i) + \frac{1}{2}np\right].$$

Let TV_i be $-2\ln\lambda_i$. Then

$$TV_i = tr(A_i\Sigma_0^{-1}) - n\ln|A_i| + n\ln|\Sigma_0| + np\ln n - np, \tag{2.2}$$

and, the statistic TV_i can be used as the sample statistic for Σ. If the sample statistic V_i or TV_i plots above the UCL, the process is deemed out-of-control state and assignable causes are sought.

When the process is in-control, the sample statistic V_i has a chi-squared distribution with $(n-1)p$ degrees of freedom. But, the components of Σ shifts from Σ_0 then it is difficult to obtain the exact distribution of V_i. And, it is difficult to obtain the exact distribution of TV_i when the process is in-control or out-of-control states. Thus, to evaluate the performances of the charts based on the sample statistics V_i or TV_i, it is necessary to carry out computer simulations.

3. Multivariate Shewhart Chart

A Shewhart chart has a good ability to detect large changes in monitored parameter quickly and is easy to implement the process. However, the Shewhart chart is slow to detect small or moderate changes of the parameters.

Since the control limits for a multivariate Shewhart chart based on the sample statistic V_i in (2.1) would be set as $\{0, \chi^2_{1-\alpha}[(n-1)p]\}$, a Shewhart chart based on V_i signals whenever

$$V_i \geq \chi^2_{1-\alpha}[(n-1)p]. \tag{3.1}$$

And the control limits for a multivariate Shewhart chart based on the sample statistic TV_i in (2.2) would be set by using percentage point of TV_i, a Shewhart chart based on TV_i signals whenever

$$TV_i \geq h_{TV(S)} \tag{3.2}$$

where $h_{TV(S)}$ can be obtained to satisfy a specified in-control ARL by simulation.

4. Multivariate CUSUM Chart

The CUSUM chart was originally proposed by Page(1954). This chart is maintained by taking samples and plotting a cumulative sum of differences between sample statistic and the target value in time order on the chart. The CUSUM chart is efficient when the detection of small shifts in a production process is important.

A multivariate CUSUM chart based on the statistic V_i in (2.1) is given by

$$Y_{V,i} = \max\{Y_{V,i-1}, 0\} + (V_i - k_V) \tag{4.1}$$

where $Y_{V,0} = \omega_V I_{(\omega_V \geq 0)}$ and the reference value $k_V \geq 0$. This chart for dispersion matrix signals whenever $Y_{V,i} \geq h_V$.

And for the CUSUM chart based on TV_i can also be constructed as

$$Y_{TV,i} = \max\{Y_{TV,i-1}, 0\} + (TV_i - k_{TV}), \tag{4.2}$$

where $Y_{TV,i} = \omega_{TV} \cdot I_{(\omega_{TV} \geq 0)}$ and $k_{TV} \geq 0$. This multivariate CUSUM chart signals whenever $Y_{TV,i} \geq h_{TV(C)}$.

When the process is on-target, we can obtain decision interval h_V by Markov chain or integral equation approach to satisfy a specified in-control ARL. And when the process parameters in Σ have changed, the performances and properties of the chart in (4.1) can be evaluated by simulation.

Since it is difficult to obtain the exact performances of multivariate CUSUM scheme based on TV_i in (4.2), the percentage point and properties of this chart can be evaluated by simulation under the process parameters of the production process are on-target or changed.

5. Concluding Remarks

In order to evaluate the performances and compare the properties of the proposed charts, the charts should have the same ARL when the process is in-control and some kinds of standards for comparison are necessary.

In our computation, the numerical results were obtained when the ARL of the in-control state was approximately equal to 370.4 and the sample size for each variable was five for $p=2\sim4$.

Since the performance of the charts depends on the components of Σ, it is not possible to investigate all of the different ways in which Σ could change. Hence, we consider the following typical types of shifts for comparison in the process parameters.

(1) V_i : σ_{10} of Σ_0 is increased to $[1+(4i-3)/10]$.

(2) C_i : ρ_{120} and ρ_{210} of Σ_0 are changed to $[0.3+(2i-1)/10]$

(3) (V_i, C_i) for $i=1,2,3$.

(4) S_i : Σ_0 is changed to $c_i\Sigma_0$ where $c_i=[1+(3i-2)/10]^2$.

After the design parameters of the proposed charts have been determined, the ARL values of all the proposed types of shifts for Shewhart and CUSUM charts were obtained by simulation with 10,000 iterations.

[Table 1] ARL performances for dispersion matrix Σ ($p=2$)

types of shifts	Shewhart		CUSUM					
	V_i	TV_i	based on V_i			based on TV_i		
no shift	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4
V_1	151.6	313.1	70.4	76.6	85.9	256.9	273.6	282.7
V_2	12.0	24.6	10.1	8.5	7.8	16.1	14.8	15.2
V_3	3.8	5.1	5.0	4.1	3.7	5.5	4.6	4.3
C_1	465.7	337.0	805.2	695.2	629.3	292.2	307.1	314.2
C_2	466.8	183.4	3470	1780	1223	58.2	71.4	87.6
C_3	437.6	52.9	6637	2279	1263	12.0	10.9	11.4
(V_1, C_1)	187.1	297.3	107.2	116.8	127.5	228.4	247.1	258.9
(V_2, C_2)	16.0	22.6	13.2	11.3	10.7	14.2	13.0	13.2
(V_3, C_3)	5.0	4.3	6.4	5.3	4.8	4.5	3.8	3.5
S_1	79.9	271.8	32.6	32.0	34.5	194.6	212.4	224.7
S_2	6.7	20.9	6.7	5.5	5.0	12.8	11.6	11.6
S_4	1.5	2.0	2.5	2.0	1.8	2.6	2.2	2.1
			$k_V=8.5$	$k_V=9.0$	$k_V=9.5$	$k_{TV}=4$	$k_{TV}=4.5$	$k_{TV}=5$

[Table 2] ARL performances for dispersion matrix Σ ($p=3$)

types of shifts	Shewhart		CUSUM					
	V_i	TV_i	based on V_i			based on TV_i		
no shift	370.4	370.4	370.4	370.4	370.4	370.3	370.3	370.4
V_1	177.7	340.6	86.1	91.5	99.5	302.2	311.9	318.0
V_2	15.8	53.9	12.6	10.7	9.8	24.5	25.2	26.4
V_3	4.5	8.8	6.2	5.0	4.5	7.2	6.6	6.3
C_1	419.7	354.4	623.3	564.2	530.6	321.6	329.1	334.2
C_2	403.8	244.7	1663	1121	880.8	103.5	119.9	136.4
C_3	309.0	101.3	3019	1502	996.9	18.5	19.1	20.9
(V_1, C_1)	203.8	328.2	116.3	123.9	133.6	274.5	288.6	296.8
(V_2, C_2)	19.2	46.3	15.6	13.3	12.4	20.3	20.4	21.3
(V_3, C_3)	5.6	6.9	7.5	6.2	5.5	5.7	5.2	4.9
S_1	65.3	299.6	26.6	24.5	24.7	213.8	229.0	242.0
S_2	4.5	27.2	5.8	4.7	4.1	12.3	11.9	11.9
S_4	1.2	1.8	2.2	1.8	1.6	2.4	2.1	2.0
			$k_V=12.5$	$k_V=13.0$	$k_V=13.5$	$k_{TV}=9$	$k_{TV}=9.5$	$k_{TV}=10$

The performances for all the proposed types of shifts are given in [Table 1] through [Table 3]. From the numerical results, we found the following properties. When a shift in variance components in Σ has occured, control chart based on the statistic V_i is efficient. And a shift in correlation coefficients in Σ have occured, control procedure based on the TV_i will be recommended. When shifts for both variances and correlation coefficients have occurred, the CUSUM procedure based on TV_i will be recommended.

When small or moderate shifts in the process have occurred, CUSUM charts with small reference values will be recommended. Therefore, we recommend CUSUM chart based on TV_i to monitor process dispersion Σ in the multivariate normal process when small or moderate shifts have occured.

[Table 3] ARL performances for dispersion matrix Σ ($p=4$)

types of shifts	Shewhart		CUSUM					
	V_i	TV_i	based on V_i			based on TV_i		
no shift	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4
V_1	199.9	357.3	97.9	103.6	111.9	316.3	322.8	325.8
V_2	19.3	118.7	14.9	12.6	11.6	43.5	41.8	42.8
V_3	5.3	19.3	7.2	5.9	5.2	13.8	12.0	11.0
C_1	397.6	359.7	540.7	506.5	488.8	331.8	334.7	339.9
C_2	374.3	291.6	1107	859.3	721.2	130.8	140.9	153.6
C_3	291.8	166.4	1795	1106	815.8	33.8	32.1	32.8
(V_1, C_1)	219.1	349.6	123.6	130.5	139.9	290.7	298.0	305.9
(V_2, C_2)	22.5	100.7	17.7	15.1	14.0	35.7	33.6	33.9
(V_3, C_3)	6.2	14.2	8.5	7.0	6.2	10.7	9.2	8.4
S_1	55.3	326.6	23.2	20.7	20.0	215.1	226.6	237.2
S_2	3.4	47.9	5.2	4.2	3.7	17.9	15.7	14.7
S_3	1.4	5.7	2.9	2.4	2.1	6.3	5.3	4.8
S_4	1.1	2.0	2.0	1.7	1.5	3.5	2.9	2.6
			$k_V=16.5$	$k_V=17$	$k_V=17.5$	$k_{TV}=16$	$k_{TV}=16.5$	$k_V=17$

References

[1] Alt, F. B. (1982). Multivariate Quality Control in *The Encyclopedia of Statistical Sciences*, eds. S. Kotz and Johnson, John Wiley, New York.

[2] Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed., Wiley, New York.

[3] Chang, T. C. and Gan, F. F. (1995). A Cumulative Sum Control Chart for Monitoring

Process Variance, *Journal of Quality Technology,* Vol. 27, 109-119.
[4] Croiser, R. B. (1988). Multivariate Generalizations of Cumulative Sum Quality-Control Schemes, *Technometrics*, Vol. 30, 291-303.
[5] Crowder, S. V. and Hamilton, M. D.(1992). An EWMA for Monitoring a Process Standard Deviation, *Journal of Quality Techonology,* Vol. 24, 12-21.
[6] Hotelling, H. (1947). Multivariate Quality Control, *Techniques of Statistical Analysis,* McGraw-Hill, New York, 111-184.
[7] Page, E. S. (1990). Continuous Inspection Schemes, *Biometrika,* Vol. 41, 110-114.
[8] Pignatiello, J. J., Jr. and Runger, G. C. (1990). Comparisons of Multivariate CUSUM Charts, *Journal of Quality Technology,* Vol. 22, 173-186.
[9] Reynolds, M. R., Jr. and Amin, R. W. and Arnold, J. C. (1990). CUSUM Charts with Variable Sampling Intervals, *Technometrics,* Vol. 32, 371-384.
[10] Whang, W. and Spedding, T. A. (2000). A Synthetic Control Chart for Detecting Small Shifts in the Process Mean, *Journal of Quality Technology*, Vol. 32, 32-38.
[11] Woodall, W. H. and Ncube, M. M. (1985). Multivariate CUSUM Quality Control Procedure, *Technometrics*, Vol. 27, 285-292.