RESEARCH ARTICLE | JULY 27 2018

Proposal of multivariate control chart using exponentially log-likelihood for detection of change in mean and variability

Yuma Ueno; Yasushi Nagata

AIP Conf. Proc. 1982, 020016 (2018)
https://doi.org/10.1063/1.5045422

Proposal of Multivariate Control Chart Using Exponentially Log-Likelihood for Detection of Change in Mean and Variability

Yuma Ueno [1, a)] and Yasushi Nagata [2, b)]

[1] Graduate Student, Waseda University (Japan).
3-4-1 Okubo, Shinjuku-ku Tokyo, 169-8555, Japan
[2] Professor, Waseda University (Japan).
3-4-1 Okubo, Shinjuku-ku Tokyo, 169-8555, Japan
[a)] uen-yum@toki.waseda.jp
[b)] ynagata@waseda.jp

Abstract. This study aims to detect various small changes in multivariate control charts. In previous studies, the MEWMA control chart was proposed as a detection of mean vector change, the MEWMC control chart was proposed as a detection of variance covariance matrix change, and the ELR control chart was proposed as a detection of the change of the mean vector and the variance covariance matrix. This study proposes two method using log-likelihood. The first method (MEWML control chart) uses the statistic obtained by directly weighting the log-likelihood. The second method (MEWMML control chart) uses obtained maximum likelihood estimate from log-likelihood using the maximum likelihood method. As a result of Monte Carlo simulations using the ARL evaluation index, the study shows that the MEWML control chart is useful for variance covariance matrix change, and the MEWMML control chart is the most useful for various patterns.

INTRODUCTION

In recent years, the complexity of data has been increasing. Thus, it is necessary to detect changes in plant processes and access logs, from multivariable, rather than single variable. Various change patterns—from large to small—exist, and useful methods for detecting them have already been proposed by many scholars. For example, Hotelling [1] proposed the Hotelling T^2 control chart using Mahalanobis distance. However, it can only detect large changes, not small ones. Therefore, Woodall and Ncube [2] proposed a multivariate cumulative sum control (MCUSUM) chart that accumulates the difference between the previous and the current terms as a method for detecting small changes. Furthermore, Lowry et al. [3] proposed a multivariate exponential weighted moving average (MEWMA) control chart using statistics weighted by exponential parameters. The MEWMA control chart is still a topic of research due to its usefulness in studying minor changes.

Additionally, Zou and Qiu [4] proposed a LEWMA control chart using Lasso, Jiang et al. [5] proposed a VS-MEWMA control chart using variable selection, and Nishimura et al. [6] proposed an AIC-MEWMA control chart that obtains the optimal combination of variables using AIC.

In the MEWMA control chart, generally, the population mean vector and the population covariance matrix are assumed to be known. However, Zamba and Hawkins [7,8] conducted numerical experiments assuming that they are unknown.

The MEWMA control chart is a method for detecting the change of the mean; it cannot detect the change of the variance covariance matrix. Therefore, Hawkins and Maboudou-Tchao [9] proposed a MEWMC control chart to detect the change of the variance covariance matrix.

Mathematical Methods and Computational Techniques in Science and Engineering II
AIP Conf. Proc. 1982, 020016-1–020016-6; https://doi.org/10.1063/1.5045422
Published by AIP Publishing. 978-0-7354-1698-7/$30.00

Zhang et al. [10] proposed an ELR control chart using log-likelihood. In it, the weighted mean and the variance covariance matrix are substituted into the log-likelihood to be the ELR statistic. However, the final statistic is log-likelihood, exponential weighted updating each of the mean vector and the variance covariance matrix.

It is normally natural to add the log-likelihood; if the final decision statistic is log-likelihood, it is preferable to weight and update log-likelihood as it is. Consequently, we propose a multivariate exponentially weighted moving likelihood (MEWML) control chart that weights log-likelihood to be the statistic in order to detect more various changes. We also propose a multivariate exponentially weighted moving maximum likelihood (MEWMML) control chart that obtains the maximum likelihood estimate of the mean vector and the variance covariance matrix from the log-likelihood by maximum likelihood method, and substitutes it to the log-likelihood.

The rest of this paper is organized as follows. Section Ⅱ introduces previous studies. Section Ⅲ explains the proposed method. In section Ⅳ, simulation experiments of the proposed method and the existing method are performed, and the accuracy is compared. Finally, section Ⅴ concludes the study.

EXISTING METHOD

MEWMA control chart (Lowry et al.[3])

The MEWMA control chart proposed by Lowry et al. [3] weights the latest and historical accumulated data by an exponential parameter. This method has excellent detection power for small changes. At this time, we assume a multivariate normal distribution with the situation, as shown in Eq.(1)

$$\begin{cases} \boldsymbol{x}_t \sim N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \ (t = 1,2,\ldots,\tau) \\ \boldsymbol{x}_t \sim N_p(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t) \ (t = \tau+1, \tau+2, \ldots,) \end{cases} \tag{1}$$

We then perform exponential weighting on the mean vector based on $\lambda (0 \leq \lambda \leq 1)$:

$$\boldsymbol{w}_t = (1-\lambda)\boldsymbol{w}_{t-1} + \lambda(\boldsymbol{x}_t - \boldsymbol{\mu}_0), \ \boldsymbol{w}_0 = 0. \tag{2}$$

Next, the $MEWMA_t$ statistic is expressed in Eq.(3)

$$MEWMA_t = \boldsymbol{w}_t' \Sigma_{\boldsymbol{w}_t}^{-1} \boldsymbol{w}_t. \tag{3}$$

The MEWMA control chart detects the change in mean based on the statistic of Eq.(3). When the weighting among the variables are all equal, we obtain the variance-covariance matrix in Eq.(4)

$$\Sigma_{w_t} = \left[\frac{\lambda \left\{ 1 - (1-\lambda)^{2t} \right\}}{2-\lambda} \right] \Sigma_0. \tag{4}$$

In general, Eq.(4) can be approximated to Eq.(5)

$$\Sigma_{w_t} = \left[\frac{\lambda}{2-\lambda} \right] \Sigma_0, \tag{5}$$

as $i \to \infty$.

The equation above is the outline of the MEWMA control chart. In $\lambda = 1$, this control chart is equal to the Hotelling T^2 control chart. In general, the MEWMA control chart is said to be able to detect very small changes compared to Hotelling T^2 control chart [3].

MEWMC control chart (Hawkins and Maboudou-Tchao[9])

Hawkins and Maboudou-Tchao[9] proposed the MEWMC control chart to detect the variance covariance matrix change.

Multivariate normalization is performed in multivariate normal distribution according to Eq.(2.1). Specifically, using $\mathbf{A}$ which satisfies $\boldsymbol{A\Sigma_0 A'} = \boldsymbol{I}_p$, let $\boldsymbol{U}_t = \boldsymbol{A}(\boldsymbol{x}_t - \boldsymbol{\mu}_0)$. When $t < \tau$, we obtain $\boldsymbol{U}_t \sim N(\boldsymbol{0}, \boldsymbol{I}_p)$. Then, the statistic weighted by the exponential parameter λ is

$$\boldsymbol{S}_t = (1-\lambda)\boldsymbol{S}_{t-1} + \lambda \boldsymbol{U}_t \boldsymbol{U}_t', \tag{6}$$

where $\boldsymbol{S}_0 = \boldsymbol{I}_p$. The log-likelihood obtained by using $\boldsymbol{S}_t$ is expressed in Eq. (7)

$$MEWMC_t = tr(\boldsymbol{S}_t) - \log|\boldsymbol{S}_t| - p. \tag{7}$$

The MEWMC control chart detects the change in variance covariance matrix based on the statistic of Eq.(7).

ELR control chart (Zhang et al.[10])

Zhang et al. [10] proposed the ELR control chart using the mean vector and the variance covariance matrix weighted by the exponential parameter λ. In this control chart, we first simultaneously weight the mean vector and the variance covariance matrix by the exponential parameter:

$$\boldsymbol{w}_t = (1-\lambda)\boldsymbol{w}_{t-1} + \lambda \boldsymbol{U}_t, \tag{8}$$

$$\boldsymbol{S}_t = (1-\lambda)\boldsymbol{S}_{t-1} + \lambda \boldsymbol{V}_t. \tag{9}$$

The multivariate normalized vector $\boldsymbol{U}_t$ is used for the MEWMC control chart. The variance covariance matrix $\boldsymbol{V}_t$, of each term in $\boldsymbol{S}_t$, is calculated using $\boldsymbol{w}_t$, weighting the mean vector,

$$\boldsymbol{V}_t = (\boldsymbol{U}_t - \boldsymbol{w}_t)(\boldsymbol{U}_t - \boldsymbol{w}_t)', \tag{10}$$

where $\boldsymbol{w}_0 = \boldsymbol{0}_p$, $\boldsymbol{S}_0 = \boldsymbol{I}_p$. Then, the final statistic is

$$ELR_t = \text{tr}(\boldsymbol{S}_t) - \log|\boldsymbol{S}_t| + \|\boldsymbol{w}_t\|^2, \tag{11}$$

where, $\|\cdot\|$ represents the L2 norm.

The ELR control chart detects the change in mean and variance covariance matrix based on the statistic of Eq. (11).

PROPOSED METHOD

MEWML control chart (proposed method 1)

MEWMA and MEWMC control charts can detect only specific changes, while the ELR control chart (Zhang et al. [10]), as per previous studies, is able to detect changes in the mean vector and variance covariance matrix.

However, because the log-likelihood, which is useful for measuring the degree of abnormality, must be the final statistic, it should be weighted.

Therefore, we propose a control chart for detection of small changes of various situations by weighting log-likelihood as it is.

We call this proposed method the MEWML control chart.

First, we define the negative log-likelihood doubled in Eq. (12)

$$l(x_t) = p\log(2\pi) + \log|\boldsymbol{\Sigma}_0| + \boldsymbol{U}_t'\boldsymbol{\Sigma}_0^{-1}\boldsymbol{U}_t, \tag{12}$$

where $\boldsymbol{U}_t$ is a multivariate normalized vector of $\boldsymbol{x}_t$, and, $\boldsymbol{\Sigma}_0 = \boldsymbol{I}_p$. The weighted Eq.(12) is defined in Eq.(13)

$$L(\boldsymbol{x}_t) = (1-\lambda)L(\boldsymbol{x}_{t-1}) + \lambda l(\boldsymbol{x}_t), \tag{13}$$

where $L(\boldsymbol{x}_0) = l(\boldsymbol{U}_0 = \boldsymbol{0}_p, \boldsymbol{\Sigma}_0 = \boldsymbol{I}_p) = p\log(2\pi)$.

Then, the final statistic of the MEWML is calculated by the likelihood ratio, defined in Eq.(14)

$$MEWML_t = L(\boldsymbol{x}_t) - L(\boldsymbol{x}_0). \tag{14}$$

The MEWML control chart detects the change in mean and variance covariance matrix based on the statistic of Eq.(14).

MEWMML chart control (proposed method 2)

The MEWML control chart, which is the proposed method 1, uses the log-likelihood as it is. However, we must originally substitute the maximum likelihood estimate into the log-likelihood.

Therefore, a method that estimates the maximum likelihood estimate using the maximum likelihood method for the log-likelihood is proposed as method 2, called the MEWMML control chart. In it, we continuously update the log-likelihood weighted exponential parameter. Then, the log-likelihood at the *t* term is as shown in Eq. (15)

$$L(\boldsymbol{x}_t) = \lambda \sum_{k=1}^{t} (1-\lambda)^{t-k} l(\boldsymbol{x}_k). \tag{15}$$

Next, we obtain the mean vector $\boldsymbol{\mu}_t *$ and variance covariance matrix $\boldsymbol{\Sigma}_t *$ from Eq. (15) using the maximum likelihood method,

$$\boldsymbol{\mu}_t * = \frac{\sum_{k=0}^{t} (1-\lambda)^{t-k} \boldsymbol{U}_k}{\sum_{k=0}^{t} (1-\lambda)^{t-k}}, \tag{16}$$

$$\boldsymbol{\Sigma}_t * = \frac{\sum_{k=0}^{t} (1-\lambda)^{t-k} \boldsymbol{V}_k}{\sum_{k=0}^{t} (1-\lambda)^{t-k}}, \tag{17}$$

where, $\boldsymbol{V}_k = \boldsymbol{U}_k \boldsymbol{U}'_k$, $\boldsymbol{U}_0 = \boldsymbol{0}_p$, $\boldsymbol{\Sigma}_0 = \boldsymbol{I}_p$.

The maximum likelihood estimate estimated from maximum likelihood method is substituted into the log-likelihood to be the final statistic

$$MEWMML_t = tr(\boldsymbol{\Sigma}_t *) - \log\left|\boldsymbol{\Sigma}_t *\right| + \left\|\boldsymbol{\mu}_t *\right\|^2. \tag{18}$$

The MEWMML control chart detects the change in mean and variance covariance matrix based on the statistic of Eq.(18).

SIMULATION

We conduct a simulation to compare the performance of the existing and proposed methods the MEWMA, MEWMC, ELR, MEWML, and MEWMML control charts. We assume the same situation as Eq.(1) in section 2, and set τ=100 and state changes in the 101st term.

We set p=2, $\boldsymbol{\mu}_0 = \boldsymbol{0}_p$, $\boldsymbol{\Sigma}_0 = \boldsymbol{I}_p$.

Then, we also set like Eq.(19),(20),

$$\boldsymbol{\mu}_t = (\mu_1, \mu_2), \tag{19}$$

$$\Sigma_t = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}. \tag{20}$$

Here, we change $(\mu_1, \mu_2, \rho, \sigma_1, \sigma_2)$. Table 1 shows the setting of these parameters in the simulation.

We use ARL (average run length) as the evaluation index as its use to measure the performance of control charts is fully established. It is obtained by the average lengths of runs beyond the control limit.

The ARL in the normal state is called IC-ARL (in control ARL), while in the abnormal state, it is called OC-ARL (out of control ARL). The former should be higher, and the latter, lower. Since, IC-ARL = 200, 370 is generally used for analysis, we define IC-ARL=200. The IC-ARL is set to 200 by 20,000 simulations, allowing it to be the control limit line. The OC-ARL that measures performance under abnormal conditions is used as an evaluation index. Note that, the lower the OC-ARL, the better. The number of simulations performed is also 20,000.

Here, we consider the change of only the mean vector in the simulation pattern 1, the change of only the correlation in pattern 2, the change of only the variance in pattern 3, the change of the mean vector and the correlation in pattern 4, the change of the mean vector and the variance in pattern 5, the change of the correlation and the variance in pattern 6, and all parameters changes in pattern 7.

As shown in Table 1, the MEWMA control chart is useful when the mean vector changes, and the MEWMC control chart is useful when the correlation changes. The MEWML control chart, the proposed method 1, is useful when variance changes, but it cannot satisfactorily detect correlation change. Although the ELR and MEWMML control charts basically maintain good accuracy with any simulation pattern, the MEWMML control chart is better in many simulation patterns.

The mean vector is included in the calculation process of the variance-covariance matrix; it is impossible to calculate an optimum value if they are weighted, and update each other individually

On the other hand, in the MEWMML control chart, we update the log-likelihood itself to obtain the maximum likelihood estimate of the final log-likelihood. This way, we can obtain the optimum value of the mean vector and the variance covariance matrix.

Table.1 Simulation results of each method (ARL value)

$(\mu_1, \mu_2, \rho, \sigma_1, \sigma_2)$	*ELR*	*MEWMA*	*MEWMC*	*MEWML*	*MEWMML*
(0.0, 0.0, 0.0, 1.0, 1.0)	201.18	200.62	199.24	200.56	200.70
1.(0.3, 0.3, 0.0, 1.0, 1.0)	44.59	34.67	126.09	99.48	45.24
(0.6, 0.6, 0.0, 1.0, 1.0)	14.81	12.11	34.39	28.91	13.94
(0.9, 0.9, 0.0, 1.0, 1.0)	8.26	7.13	12.35	11.31	7.33
2.(0.0, 0.0, 0.3, 1.0, 1.0)	78.68	173.91	67.63	168.34	85.11
(0.0, 0.0, 0.6, 1.0, 1.0)	24.81	126.02	21.65	119.03	28.11
(0.0, 0.0, 0.9, 1.0, 1.0)	11.97	91.68	10.71	83.13	12.88
3.(0.0, 0.0, 0.0, 1.2, 1.2)	55.66	63.41	34.60	22.87	34.37
(0.0, 0.0, 0.0, 1.3, 1.3)	28.95	43.48	19.08	13.73	19.74
(0.0, 0.0, 0.0, 1.4, 1.4)	17.94	31.79	12.54	9.50	12.99
4.(0.3, 0.3, 0.3, 1.0, 1.0)	32.11	33.60	45.74	86.05	31.48
(0.6, 0.6, 0.6, 1.0, 1.0)	10.96	12.51	12.84	25.91	10.55
(0.6, 0.6, 0.3, 1.0, 1.0)	13.54	12.40	20.72	27.60	12.60
(0.3, 0.3, 0.6, 1.0, 1.0)	18.16	32.67	18.64	67.87	19.00
5.(0.3, 0.3, 0.0, 1.2, 1.2)	27.77	25.00	26.55	18.40	20.43
(0.6, 0.6, 0.0, 1.3, 1.3)	10.91	10.65	11.08	8.69	8.82
(0.6, 0.6, 0.0, 1.2, 1.2)	12.60	11.09	15.14	11.61	10.38
(0.3, 0.3, 0.0, 1.3, 1.3)	19.44	21.49	16.17	25.55	7.63
6.(0.0, 0.0, 0.3, 1.2, 1.2)	36.82	60.44	25.55	22.49	26.85
(0.0, 0.0, 0.9, 1.3, 1.3)	8.81	33.18	7.63	13.62	8.54
(0.0, 0.0, 0.3, 1.3, 1.3)	23.57	42.37	16.81	13.88	17.58
(0.0, 0.0, 0.9, 1.2, 1.2)	9.88	42.68	8.60	19.93	9.85
7.(0.3, 0.3, 0.3, 1.2, 1.2)	22.96	25.68	20.60	18.46	18.02
(0.6, 0.6, 0.6, 1.3, 1.3)	8.88	11.30	8.31	9.05	7.82
(0.3, 0.3, 0.6, 1.3, 1.3)	12.53	22.11	10.70	12.17	10.97
(0.3, 0.3, 0.3, 1.3, 1.3)	17.64	22.45	14.48	12.17	13.77
(0.3, 0.3, 0.6, 1.2, 1.2)	14.73	25.07	13.07	17.86	13.28
(0.6, 0.6, 0.3, 1.2, 1.2)	11.86	11.50	12.76	11.76	9.95
(0.6, 0.6, 0.6, 1.2, 1.2)	9.84	11.90	9.71	11.86	8.82
(0.6, 0.6, 0.3, 1.3, 1.3)	10.40	11.05	9.92	8.80	8.58

CONCLUSION

The MEWMA control chart for change of mean vector, the MEWMC control chart for change of variance covariance matrix, and the ELR control chart for change of either mean vector or variance covariance matrix have already been proposed.

In this paper, we further propose new MEWML and MEWMML control charts to detect change of either mean vector or variance covariance matrix, and their compare accuracy.

Both of the proposed methods weight log-likelihood exponentially. The MEWML control chart uses the weighted log-likelihood as the final statistic. On the other hand, the MEWMML control chart estimates the mean vector and the variance covariance matrix of the weighted log-likelihood by the maximum likelihood method, and then, substitutes them to the log-likelihood. Finally, it is used as the final statistic.

As a result of simulation using the ARL evaluation index, we find that the MEWML control chart is useful when the variance changes, while the MEWMML control chart is comprehensively the most useful method. The MEWMML control chart, the proposed method, was superior in most simulation situations compared to the existing study with the ELR control chart.

It is possible to detect small changes in various situations promptly by weighting the log-likelihood itself, calculating the maximum likelihood estimate, assigning it to the log-likelihood, and then, using it as the statistic.

To conclude, in actual situations, when a pattern of abnormality cannot be predicted or when various pattern changes are expected, the MEWMML control chart should be used. When a change in mean vector is expected, the MEWMA control chart should be used. When a change in the correlation of variance covariance matrix is expected, the MEWMC control chart should be used. When a change in the variance of variance covariance matrix is expected, the MEWML control chart should be used.

ACKNOWLEDGMENTS

Thanks to the guidance and cooperation of many people, we are able to reach the completion of this paper.

REFERENCES

[1] Hotelling, H. (1947): "Multivariate Quality Control-Illustrated by the Air Testing of Bombsights", in *Techniques of Statistical Analysis*, eds. Eisenhart, C., Hastay, M.W. and Wallis, W.A., New York: McGraw-Hill, pp.111-184.

[2] Woodall, W. H. and Ncube, M. M. (1985): "Multivariate CUSUM Quality-Control Procedures", *Technometrics*, 27, pp.285-292.

[3] Lowry, C.A., Woodall, W.H., Champ, C.W. and Rigdon, S.E. (1992): "A multivariate exponentially weighted moving average control chart", *Technometrics*, 34 (1), pp.46–53.

[4] Zou, C. and Qiu, P. (2009): "Multivariate Statistical Process Control Using LASSO", *Journal of the American Statistical Association*, 104 (488), pp.1586-1596.

[5] Jiang, W., Wang, K. and Tsung, F. (2012): "A variable-selection-based multivariate EWMA chart for process monitoring and diagnosis", *Journal of Quality Technology*, 44 (3), pp.209–230.

[6] Nishimura, K., Matsuura, S. and Suzuki, H. (2015): "Multivariate EWMA control chart based on a variable selection using AIC for multivariate statistical process monitoring", *Statistics and Probability Letters*, 104, pp.7-13.

[7] Zamba, K.D. and Hawkins D.M. (2006): "A Multivariate Change-Point Model for Statistical Process Control", Technometrics, 48 (4), pp539-548.

[8] Zamba, K.D. and Hawkins D.M. (2009): "A Multivariate Change-Point Model for Change in Mean Vector and/or Covariance Structure", *Journal of Quality Technology*, 41 (3), pp.285-303.

[9] Hawkins, D.M. and Maboudou-Tchao, E. M. (2008): "Multivariate Exponentially Weighted Moving Covariance Matrix", *Technometrics*, 50, pp.155-166.

[10] Zhang, J., Li, Z. and Wang, Z. (2010): "A multivariate control chart for simultaneously monitoring process mean and variability", *Computational Statistics and Data Analysis*, 54, pp.2244-2252.