Quality Technology & Quantitative Management
Vol. 3, No. 4, pp. 415-436, 2006

QTQM

Multivariate Control Charts for Monitoring Covariance Matrix: A Review

Arthur B. Yeh[1], Dennis K. J. Lin[2] and Richard N. McGrath[3]

[1,3]Department of Applied Statistics and Operations Research, Bowling Green State University, Bowling Green, OH, USA
[2]Department of Supply Chain and Information Systems, The Pennsylvania State University, University Park, PA, USA
(*Received* March 2005, *accepted* October 2005)

Abstract: In this paper, we review multivariate control charts designed for monitoring changes in a covariance matrix that have been developed in the last 15 years. The focus is on control charts developed for multivariate normal processes, assuming that independent subgroups of observations or independent individual observations are sampled as process monitoring proceeds. Control charts developed between 1990 and 2005 are reviewed according to the types of the control chart: multivariate Shewhart chart, multivariate CUSUM chart and multivariate EWMA chart. In addition to these developments, a new multivariate EWMA control chart is proposed. We also discuss comparisons of chart performance that have been carried out in the literature, as well as the issue of diagnostics. Some potential future research ideas are also given.

Keywords: Conditional entropy, decomposition, generalized variance, likelihood ratio test, moving ranges, phase II process monitoring, projection pursuit.

1. Introduction

In the last two decades, there has been an increasing research interest in multivariate quality control, which is evidenced by the large number of papers published in statistical and quality journals. The recent development is certainly welcoming since in many industrial applications the quality of a product can be attributed to several correlated quality characteristics, all of which need to be controlled and monitored simultaneously. For example, in a process in wafer manufacturing called chemical mechanical planarization (CMP), the quality of a polished wafer depends on several correlated variables. Another indication of the growing popularity of these methods in industry is the availability of some of these quality control tools in widely-used commercial software packages such as Minitab. Numerous authors have pointed out that multivariate quality control, especially application of multivariate control charts, is an important area of research for the new century (see, for example, Woodall and Montgomery [35], Stoumbos *et al.* [27] and Woodall [34]). With newly developed advanced data acquisition techniques, computing power and commercial software, multivariate quality control will play a greater role in monitoring and improving manufacturing processes.

The majority of the research in the last 20 years focuses on developing multivariate control charts for monitoring shifts in process mean. Excellent reviews of these developments can be found in, for example, Wierda [33], Lowry and Montgomery [21], Mason *et al.* [25], and Montgomery [26]. Although these reviews also contain discussions on multivariate control charts for monitoring changes in a covariance matrix, the main coverage is limited to developments that have occurred prior to 1990, such as those found in Alt [1], Alt and Bedewi [2], Healy [14], and Alt and Smith [3]. Since 1990, numerous papers have been published which specifically discuss multivariate control charts for monitoring the covariance matrix. The main purpose of this paper is to give an updated review of these developments from 1990 to 2005. Our review focuses on control charts developed specifically for multivariate normal processes, assuming that independent subgroups of observations or independent individual observations are repeatedly sampled as process monitoring proceeds. In addition, the focus is also on control charts designed for Phase II process monitoring. For discussions of control charts designed for Phase I process control, see, for example, Alt [1], Wierda [33], Mason *et al.* [25], Sullivan and Woodall [29] and Vargas [32].

To give a consistent treatment of the many multivariate control charts that will be discussed in subsequent sections, some definitions and notation are first introduced. Let $X=(X_1,X_2,\ldots,X_p)'$ denote the random variable that represents p correlated quality characteristics derived from a manufacturing process whose quality is to be monitored. When the process is in control, it is assumed that X follows a p-dimensional normal distribution, denoted by $N_p(\mu_0,\Sigma_0)$, where μ_0 is the in-control process mean and Σ_0 is the in-control process covariance matrix. On the other hand, when the process is out of control, it is assumed that X follows $N_p(\mu,\Sigma)$, where either $\mu\neq\mu_0$ or $\Sigma\neq\Sigma_0$ or both. If μ_0 and Σ_0 are not known, it is assumed that at the end of Phase I, k samples, each with size n, are available for estimating the parameters. From these k training samples, $\overline{\overline{X}}=\sum_{i=1}^{k}\overline{X}_i/k$ and $\overline{S}=\sum_{i=1}^{k}S_i/k$ can easily be computed which can be used to estimate μ_0 and Σ_0, respectively. Here $\overline{X}_i=\sum_{j=1}^{n}X_{ij}/n$ and $S_i=\sum_{j=1}^{n}(X_{ij}-\overline{X}_i)(X_{ij}-\overline{X}_i)'/(n-1)$ denote, respectively, the sample mean vector and sample covariance matrix of the ith training sample, $i=1,2,\ldots,k$. When the Phase II process monitoring begins, depending on the methodologies, either independent samples, each with size n, are taken or independent individual ($n=1$) observations are drawn. For the former, these n observations are denoted by $X_{t1},X_{t2},\ldots,X_{tn}$, $t=1,2,\ldots$, and the corresponding sample mean and sample covariance matrix are denoted by $\overline{X}_t=\sum_{j=1}^{n}X_{tj}/n$ and $S_t=\sum_{j=1}^{n}(X_{tj}-\overline{X}_t)(X_{tj}-\overline{X}_t)'/(n-1)$, respectively. As for the latter, the individual observations will simply be denoted by X_t.

The review will be done based on the types of chart being discussed. The types of chart include multivariate Shewhart, multivariate CUSUM and multivariate EWMA control charts. Specifically, Part I will review numerous multivariate Shewhart charts: Guerrero-Cusumano [9] in Section 2, Tang and Barnett [30, 31] in Section 3, Levinson *et al.* [19] in Section 4, Yeh and Lin [39] in Section 5, and Khoo and Quah [18] in Section 6. Part II will review two multivariate CUSUM charts: Hawkins [11, 13], Yeh *et al.* [38], Yeh *et al.* [37] and Huwang *et al.* [15] in Section 7, and Chan and Zhang [7] in Section 8. Part III will review a number of multivariate EWMA charts: Hawkins [11, 13], Yeh *et al.* [38], Yeh *et al.* [37] and Huwang *et al.* [15] in Section 9, Yeh *et al.* [40] in Section 10, Yeh *et al.* [38] in Section 11, Yeh *et al.* [37] in Section 12, and Huwang *et al.* [15] in Section 13. In Section 14, a new multivariate EWMA control chart is proposed. The new chart is essentially based on the EWMA of the determinants of the sample covariance matrices. Part IV will then discuss chart performance comparisons that have been carried out in the literature (Section 15), as well as the issue of diagnostics (Section 16). The performance is defined in terms of the average run length (ARL) of a control chart, where the run length is defined as the number of samples needed before an out-of-control signal is first detected on a control chart. Finally, in Section 17 a summary and discussion are given with suggestions of potential problems for future research.

PART I: MULTIVARIATE SHEWHART CONTROL CHARTS

2. A Control Chart Based on Conditional Entropy

The entropy of a random vector X may be regarded as a descriptive quantity which measures the extent to which the probability distribution is concentrated on a few points or dispersed over many points. Therefore, the entropy is a measure of dispersion, similar to standard deviation in the univariate case. For a p-dimensional multivariate random variable X, the entropy of X is defined as

$$H(x) = \int f(x)\, ln f(x) dx = E_f[-ln f(x)],$$

where $f(x)$ is the density function of X. If X follows $N_p(\mu_0, \Sigma_0)$, then the entropy is given by

$$H(x) = \frac{1}{2} p\, ln(2\pi e) + \frac{1}{2} ln |\Sigma_0|,$$

where $|A|$ denotes the determinant of a matrix A. Guerrero-Cusumano [9] suggested that the following alternative expression of $H(x)$ be considered:

$$H(x) = \frac{1}{2} p\, ln(2\pi e) + \frac{1}{2} 2\, ln |\Sigma_{d_0}^2| + \frac{1}{2} ln |P_0|$$

$$= \frac{1}{2} p\, ln(2\pi e) + \frac{1}{2} \Sigma_{i=1}^{p} ln(\sigma_{i0}^2) - T(x)$$

where $P_0 = \Sigma_{d_0}^{-1} \Sigma_0 \Sigma_{d_0}^{-1}$ is the correlation matrix, $\Sigma_{d_0} = diag(\sigma_{i0})$ with σ_{i0}, $i = 1, 2, ..., p$, being the in-control standard deviation for the ith component of X. The function $T(x)$ is called the mutual information of the random variable X. Estimating σ_{i0}^2 by the sample variance for ith component s_i^2, thus obtaining $\widehat{H}(x)$, and measuring the difference between sample and theoretical entropy, the author proposed the following statistic E for each of the samples taken when the monitoring begins:

$$E = \sqrt{\frac{n-1}{2p}} \Sigma_{i=1}^{p} ln(\frac{s_i^2}{\sigma_{i0}^2}). \tag{1}$$

The statistic E is distributed asymptotically as a univariate standard normal distribution, denoted by $N(0,1)$. The upper control limit (UCL) and the lower control limit (LCL) of the conditional entropy chart be calculated by

$$UCL = gp\left[G'(\frac{n-1}{2}) - ln(\frac{n-1}{2})\right] + z_{\alpha/2} k \sqrt{pG''(\frac{n-1}{2}) + \frac{2}{n-1} trace(P_0 - I)^2}, \tag{2}$$

$$LCL = gp\left[G'(\frac{n-1}{2}) - ln(\frac{n-1}{2})\right] - z_{\alpha/2} k \sqrt{pG''(\frac{n-1}{2}) + \frac{2}{n-1} trace(P_0 - I)^2}, \tag{3}$$

where $g = (2(n-1)/p)^{1/2}$, $G'(\cdot)$ and $G''(\cdot)$ are, respectively, the first and second derivative of the natural logarithm of the gamma function, $trace(A)$ is the trace of a matrix A, and z_α is the $1-\alpha$ quantile of $N(0,1)$. Note that in the conditional entropy chart, Σ_0 is assumed known and that $n > p$ to ensure that sample covariance matrix has full rank.

3. A Control Chart Based on the Decomposition of S_t

Tang and Barnett [30, 31] proposed a multivariate Shewhart chart that is based on decomposing S_t into a sum of a series of independent χ^2 statistics. Assuming that Σ_0 is known, define $S_j(\Sigma_j)$ and $S_{*k}(\Sigma_{*k})$ respectively as the sample (population) covariance matrix of the first j variables and of the last k variables. The sample covariance matrix can be partitioned into (for simplicity of the discussion, we will drop the subscript t in this section):

$$S = \begin{pmatrix} S_{j-1} & S_{(j-1)\times(p-j+1)} \\ S'_{(j-1)\times(p-j+1)} & S_{*p-j+1} \end{pmatrix},$$

where $S'_{(j-1)\times(p-j+1)} = (S_{j,j-1}, S_{j+1,j-1}, \ldots, S_{p,j-1})$ and $S_{k,j}$ represents the row vector of sample covariances between the k th variable and each of the first j variables. Note that Σ_0 can similarly be partitioned by replacing sample statistics with the corresponding population parameters. Further define the conditional sample variance of the j th variable given the first $j-1$ variables as

$$s^2_{j\cdot 1,2,\ldots,j-1} = s_j^2 - S'_{j,j-1} S^{-1}_{j-1} S_{j,j-1}$$

$$(\sigma^2_{j\cdot 1,2,\ldots,j-1} = \sigma_j^2 - \Sigma'_{j,j-1} \Sigma^{-1}_{j-1} \Sigma_{j,j-1}).$$

In addition, the conditional sample covariance matrix of the last $p-j+1$ variables given the first $j-1$ variables can be expressed as

$$S_{j,j+1\ldots,p\cdot 1,2,\ldots,j-1} = S_{*p-j+1} - S'_{(j-1)\times(p-j+1)} S^{-1}_{j-1} S_{(j-1)\times(p-j+1)}$$

$$(\Sigma_{j,j+1\ldots,p\cdot 1,2,\ldots,j-1} = \Sigma_{*p-j+1} - \Sigma'_{(j-1)\times(p-j+1)} \Sigma^{-1}_{j-1} \Sigma_{(j-1)\times(p-j+1)}).$$

Also let d_j (θ_j), $j = 2, 3, \ldots, p$, denote the vector of sample (population) regression coefficients when each of the last $p-j+1$ variables is regressed on the $(j-1)$ th variable while the first $j-2$ variables are held fixed. The d_j can be expressed as

$$d_j = \frac{\left\{S_{(j-1)\times(p-j+1)} - S'_{j-1,j-2} S^{-1}_{j-2} (S'_{j,j-2} S'_{j+1,j-2} S'_{p,j-2})'\right\}'}{s^2_{j-1} - S'_{j-1,j-2} S^{-1}_{j-2} S'_{j-1,j-2}},$$

and likewise θ_j can similarly be expressed by replacing sample statistics with population parameters. Note that d_2 (θ_2) should be interpreted as the vector of unconditional sample (population) regression coefficients when each of the last $p-1$ variables is regressed on the first variable.

As each sample of n observations is drawn, one calculates

$$T = \Sigma_{j=1}^{2p-1} Z_j^2, \tag{4}$$

where

$$Z_1 = \Phi^{-1}\left\{\chi^2_{n-1}\left[\frac{(n-1)s_1^2}{\sigma_1^2}\right]\right\},$$

$$Z_j = \Phi^{-1}\left\{\chi^2_{n-j}\left[\frac{(n-1)s^2_{j\cdot 1,2,\ldots,j-1}}{\sigma^2_{j\cdot 1,2,\ldots,j-1}}\right]\right\}, \text{ for } j = 2, 3, \ldots, p,$$

$$Z_{p+1} = \Phi^{-1}\left\{\chi^2_{p-1}\left[(n-1)s_1^2(d_2-\theta_2)'\Sigma^{-1}_{2,3,\ldots,p\cdot 1}(d_2-\theta_2)\right]\right\},$$

and

$$Z_{p+j-1} = \Phi^{-1}\left\{\chi^2_{p-j+1}\left[(n-1)s^2_{j-1\cdot 1,2,\ldots,j-2}(d_j-\theta_j)'\Sigma^{-1}_{j,j+1,\ldots,p\cdot 1,2,\ldots,j-1}(d_j-\theta_j)\right]\right\}, \text{ for } j = 3, 4, \ldots, p.$$

Note that $\Phi^{-1}\{\cdot\}$ is the inverse of the distribution function of $N(0,1)$ and $\chi^2_\nu[x] = P(\chi^2_\nu \le x)$ is the distribution function of a χ^2 distribution with ν degrees of freedom.

When the process is in control, Z_j's are independent and identically distributed (i.i.d.) as $N(0,1)$, and therefore T is distributed as χ^2_{2p-1}. Thus the control chart can be established by plotting T's against sampling sequence and an out-of-control signal is detected as soon as T exceeds UCL which can be determined from χ^2_{2p-1}. Note that the decomposition is not unique since it depends on how the p variables are arranged. It was suggested that the p variables be arranged in decreasing order of importance from 1 to p. Furthermore, μ_0 and Σ_0 are assumed known and it is required that $n > p$.

The authors also discussed possible extensions of T statistics to cases when (i) Σ_0 is unknown and can be estimated by $\bar{S}$ and (ii) $1 < n < p$ when S may not be of full rank. In the case when Σ_0 is unknown, one essentially replaces population parameters in Σ_0 by their counterparts in $\bar{S}$. In the case when $1 < n < p$, one first transforms S to a matrix of reduced dimension W by $W = ASA$, where A is a full rank $(n-1)\times p$ matrix of constants. The same methods of decomposing and combining independent statistics can be applied to W.

4. A Control Chart Based on Testing $H_0 : \Sigma = \Sigma_0$

By treating the problem as testing $H_0 : \Sigma = \Sigma_0$ v.s. $H_a : \Sigma \ne \Sigma_0$ based on two independent samples, one being the training samples and the other being the repeatedly drawn independent samples when monitoring begins, Levinson *et al.* [19] proposed the following statistic, for $t \ge 1$,

$$mM_t = m\left[(k+1)(n-1)ln|S_p| - k(n-1)ln|\bar{S}| - (n-1)ln|S_t|\right], \quad (5)$$

where
$$m = 1 - \left[\frac{1}{k(n-1)} + \frac{1}{(n-1)} - \frac{1}{(k+1)(n-1)}\right] \times \left[\frac{2p^2+3p-1}{6(p+1)}\right]$$

and

$$S_p = \frac{k(n-1)\bar{S} + (n-1)S_t}{(k+1)(n-1)}$$

When the process is in control (i.e., $\Sigma = \Sigma_0$), the distribution of mM_t follows

$\chi^2_{p(p+1)/2}$. Therefore, the UCL and LCL can be determined based on $\chi^2_{p(p+1)/2}$. Note that in this chart, Σ_0 need not be known, since it is estimated by $\bar{S}$. Furthermore, it is assumed that $n > p$ to ensure that S_t has full rank.

5. A Control Chart Based on Probability Integral Transformation

In an effort to develop a single multivariate control chart to simultaneously monitor changes in the process mean and covariance matrix, Yeh and Lin [39] proposed using the probability integral transformation to transform different statistics into the same random variable. Thus, different statistics can be combined and plotted on a single control chart. The part that deals with covariance matrix can be written as, for $t \geq 1$,

$$v_t = P\left(\prod_{i=1}^{p} F_{n-1-i,k(n-1)-k+1-i} \leq (\prod_{i=1}^{p} \frac{k(n-1)-k+1-i}{n-1-i}) \times \frac{|(n-1)S_t|}{|k(n-1)\bar{S}|} \right). \tag{6}$$

Here $F_{n-1-i,k(n-1)-k+1-i}$ denotes an F distribution with $n-1-i$ and $k(n-1)-k+1-i$ degrees of freedom, and the p F-distributions in the product are independent. For each sample of size n, the v_t is the probability that the random variable $\prod_{i=1}^{p} F_{n-1-i,k(n-1)-k+1-i}$ is less than or equal to the observed

$$(\prod_{i=1}^{p} \frac{k(n-1)-k+1-i}{n-1-i}) \times \frac{|(n-1)S_t|}{|k(n-1)\bar{S}|}.$$

When the process is in-control, v_t's are a sequence of i.i.d. $U(0,1)$ random variables. Therefore, the control limits can be set up based on $U(0,1)$. For example, for comparable 3σ limits, UCL and LCL can be set to equal to .99865 and .00135, respectively.

6. A Control Chart Based on Individual Observations

Assuming that Σ_0 is known and applying an idea from univariate moving range charts, Khoo and Quah (2003) proposed the following statistic, for $t \geq 1$,

$$M_{t+1} = \frac{1}{2}(X_{t+1} - X_t)'\Sigma_0^{-1}(X_{t+1} - X_t). \tag{7}$$

When the process is in-control, the distribution of M_t follows a χ^2_p so that the UCL and LCL can readily be obtained from χ^2_p. Note that this chart is specifically designed for $n = 1$. However, the M_t's are not independent.

PART II: MULTIVARIATE CUSUM CONTROL CHARTS

7. Multiple CUSUM Charts Based on Regression Adjusted Variables

Hawkins [11, 13] proposed a multivariate control chart for monitoring the process mean based on regression adjusted variables. In the discussion he also mentioned, though did not explain in detail, that the same idea coupled with his earlier work in Hawkins [10] can be extended to constructing multivariate control charts for monitoring process variability. This idea is expanded and discussed in more detail in a number of recent works by Yeh *et al.* [38], Yeh *et al.* [37] and Huwang *et al.* [15].

Note that in the context of regression adjusted variables, μ_0 and Σ_0 are assumed to be known. One first computes, for $t \geq 1$, the following

$$Z_t = [diag(\Sigma_0^{-1})]^{-1/2} \Sigma_0^{-1}(X_t - \mu_0),$$

where $Z_t = (Z_{t1}, Z_{t2}, ..., Z_{tp})'$. When the process is in-control, Z_t is distributed as $N_p(0, I_p)$, where I_p is a $p \times p$ identity matrix. In order to detect changes in the variance of the jth component Z_{tj}, $j = 1,2,...,p$, one further defines the following statistics

$$W_{tj} = \frac{|Z_{tj}|^{1/2} - .822}{.349}. \tag{8}$$

When the process is in control, W_{tj} is approximately distributed as $N(0,1)$. On the other hand, if the distribution of Z_{tj} changes from $N(0,1)$ to $N(0,\sigma^2)$ ($\sigma \neq 1$), the distribution of W_{tj} changes approximately to $N(2.355(\sigma^{1/2} - 1), \sigma)$. Therefore, one can construct the usual univariate CUSUM chart to monitor mean shifts in W_{tj} (and thus variance changes in Z_{tj}). Consequently, p such CUSUM charts can be combined in very much the same way as was suggested in Woodall and Ncube [36] to obtain the so-called multiple CUSUM control charts.

Specifically, one calculates, for $t \geq 1$ and $j = 1,2,...,p$,

$$S_{tj}^{+} = \max(0, S_{(t-1)j}^{+} + W_{tj} - r) \tag{9}$$

and

$$S_{tj}^{-} = \min(0, S_{(t-1)j}^{-} + W_{tj} + r), \tag{10}$$

where $S_{0j}^{+} = S_{0j}^{-} = 0$ and r is the reference value. Here S_{tj}^{+} and S_{tj}^{-} are designed to detect, respectively, increases and decreases in the variance of the jth component of Z_t, since $(\sigma^{1/2} - 1) > 0$ if $\sigma^2 > 1$ and < 0 if $\sigma^2 < 1$. An out-of-control signal is detected on the multiple CUSUM charts as soon as

$$\max_{1 \leq j \leq p}\{\max(S_{tj}^{+}, -S_{tj}^{-})\} > h,$$

where h is the decision value.

8. A CUSUM Chart Based on Projection Pursuit

The idea of using projection pursuit to develop CUSUM charts for monitoring the covariance matrix (Chan and Zhang [7]) is predicated on the following two important observations:

(i) $\Sigma = \Sigma_0$, i.e., the covariance matrix remains in-control, if and only if $a'_{\max}\Sigma_0^{-1/2}X$ and $a'_{\min}\Sigma_0^{-1/2}X$ have unit variance, where a_{max} and a_{min} are the eigenvectors that correspond to, respectively, the largest and smallest eigenvalues of the matrix $\Sigma_0^{-1/2}\Sigma\Sigma_0^{-1/2}$; and

(ii) a_{max} and a_{min} give the maximum and minimum signed difference between the variance of $a'\Sigma_0^{-1/2}X$ and 1, respectively.

Therefore, the projection pursuit CUSUM chart can be developed first by selecting a univariate CUSUM chart for monitoring changes in variance (from unit variance), and then applying the univariate CUSUM chart to both $\hat{a}'_{\max}\Sigma_0^{-1/2}X$ and $\hat{a}'_{\min}\Sigma_0^{-1/2}X$, where $\hat{a}_{max}$ and $\hat{a}_{min}$ are some estimates of a_{max} and a_{min}, respectively.

The authors proposed using the univariate CUSUM chart developed in Johnson and Leone [16] for monitoring $\hat{a}'_{\max}\Sigma_0^{-1/2}X$ and $\hat{a}'_{\min}\Sigma_0^{-1/2}X$. Moreover, they discussed ways of obtaining $\hat{a}_{max}$ and $\hat{a}_{min}$ in two cases: when $n=1$ and when $n>1$. When individual observations are collected, assuming that $\mu_0 = 0$ and $\Sigma_0 = I_p$, denote respectively by λ_{tj}^{max} and λ_{tj}^{min}, $1 \le j \le t$, the largest and smallest eigenvalues of the sample matrix

$$X_j X_j' + X_{j+1} X_{j+1}' + \cdots + X_t X_t'.$$

Also define $Q_{tj}^{+} = \lambda_{tj}^{max} - (t-j+1)r_{+}$ and $Q_{tj}^{-} = \lambda_{tj}^{min} - (t-j+1)r_{-}$, where r_{+} and r_{-} are two reference values. Now define

$$Q_t^{+} = \max\{0, Q_{t1}^{+}, Q_{t2}^{+}, \ldots, Q_{tt}^{+}\} \tag{11}$$

and

$$Q_t^{-} = \min\{0, Q_{t1}^{-}, Q_{t2}^{-}, \ldots, Q_{tt}^{-}\}, \tag{12}$$

where $Q_0^{+} = Q_0^{-} = 0$. The projection pursuit CUSUM chart signals as soon as either $Q_t^{+} > h_{+}$ or $Q_t^{-} < h_{-}$, where h_{+} and h_{-} are decision values.

When $n>1$, the only modification is that the sample matrix from which λ_{tj}^{max} and λ_{tj}^{min} are computed is replaced by the subgroup sample matrix

$$\sum_{i=j}^{t} \frac{1}{n-1} \sum_{k=1}^{n} X_{ik} X_{ik}'.$$

Note that in the projection pursuit CUSUM chart, Σ_0 is either assumed to be known or that it can be estimated by S. Furthermore, it is also assumed that the process mean μ_0 stays unchanged since the chart is sensitive to shifts in process mean. Based on Monte-Carlo simulations, designs for the projection pursuit based CUSUM chart were provided for p = 2, 3, and 4, and n = 1, 2, 5, and 10.

PART III: MULTIVARIATE EWMA CONTROL CHARTS

9. Multiple EWMA Charts Based on Regression Adjusted Variables

The multiple CUSUM charts discussed in Section 7, Hawkins [11, 13], can easily be adapted to multiple EWMA charts. For $t \ge 1$ and $j = 1, 2, \ldots, p$, one calculates

$$E_{tj} = \lambda W_{tj} + (1-\lambda) W_{(t-1)j}, \tag{13}$$

where $0 < \lambda < 1$ is a smoothing constant, $E_{0j} = 0$ and W_{tj} was defined in equation (8). For a given λ, an out-of-control signal is detected on the multiple EWMA charts as soon as

$$\max_{1 \le j \le p} \{ |E_{tj}| \} > L \times \sqrt{\frac{\lambda}{2-\lambda}},$$

where L is a pre-determined value which depends on λ and the in-control ARL (ARL_0). Some designs of the multiple EWMA charts were also discussed in the references cited in Section 7.

10. An EWMA Chart Based on Probability Integral Transformation

Extending the idea proposed in Yeh and Lin [39] (see Section 5), Yeh *et al.* [40] developed a multivariate EWMA control chart based on the probability integral transformation. Specifically, based on the statistic v_t (Equation (6)), define, for $t \ge 1$,

$$J_t = \lambda \times (v_t - .5) + (1-\lambda) \times J_{t-1}, \tag{14}$$

where $0 < \lambda < 1$ is a smoothing constant and $J_0 = 0$. For any given t, $E(J_t) = 0$ and $Var(J_t) = [\lambda/12(2-\lambda)](1-(1-\lambda)^{2t})$. Therefore, the UCL and LCL can be determined by

$$UCL = L \times \sqrt{\frac{1}{12}(\frac{\lambda}{2-\lambda})[1-(1-\lambda)^{2t}]} \tag{15}$$

$$LCL = -L \times \sqrt{\frac{1}{12}(\frac{\lambda}{2-\lambda})[1-(1-\lambda)^{2t}]}, \tag{16}$$

where L is chosen to control the ARL_0 of the control chart. The authors called the proposed chart the V-chart.

Note that in the V-chart, it is assumed that Σ_0 can be estimated by $\overline{S}$ which is derived from k training samples each of size n, collected when the process was in control. It is also assumed that $n > p$ to ensure that S_t has full rank.

11. An EWMA Chart Based on Likelihood Ratio Test

Yeh *et al.* [38] treated the problem as a two-sample problem of testing $H_0: \Sigma = \Sigma_0$ v.s. $H_a: \Sigma \neq \Sigma_0$, with one sample coming from the training data and the other sample coming from the repeatedly drawn samples when process monitoring begins. For any given $t \geq 1$, an unbiased test derived in Sugiura and Nagao [28] can be performed and the test is based on a modified likelihood ratio

$$L_t = \frac{|k(n-1)\overline{S}|^{\frac{1}{2}(kn-1)} |(n-1)S_t|^{\frac{1}{2}(n-1)}}{|k(n-1)\overline{S} + (n-1)S_t|^{\frac{1}{2}(kn+n-2)}}.$$

The testing procedure is typically performed by computing

$$\begin{aligned} r_t &= -2ln(L_t) \\ &= (kn+n-2)ln|k(n-1)\overline{S} + (n-1)S_t| \\ &\quad -(kn-1)ln|k(n-1)\overline{S}| - (n-1)ln|(n-1)S_t|, \end{aligned}$$

and H_0 is rejected if $r_t > c_\alpha$, where c_α is a critical value determined by α. Based on this r_t, the authors proposed computing the EWMA of r_t. Specifically, define the EWMA statistics as

$$R_t = \lambda r_t + (1-\lambda) R_{t-1}, \tag{17}$$

where $0 < \lambda < 1$ is a smoothing constant and $R_0 = r_1$. Note that the initial value R_0 is set to be equal to r_1, instead of the conventional $E(R_t)$ because $E(R_t)$ is unknown and needs to be estimated. However, by doing so, the variances of R_t, when $R_0 = r_1$ and when $R_0 = E(R_t)$, differ only up to a constant. The proposed chart is called the exponentially weighted moving likelihood ratio (EWMLR) chart.

Since R_t is based on the EWMA of the logarithms of the likelihood ratios, the chart will signal if $R_t > UCL$. Based on Monte-Carlo simulations, the authors provided the UCL's which produced an ARL_0 of approximately 370, for numerous cases such as different numbers of training samples and different sample sizes. Note that in the EWMLR chart, Σ_0 need not be known, but training samples need to be available and $n > p$ to ensure that S_t has full rank.

12. An EWMA Chart for Individual Observations

When $n = 1$, the sample covariance matrix is not readily available. For any given individual observation X_t, $t \geq 1$, the matrix $(X_t - \mu_0)(X_t - \mu_0)'$ still provides an unbiased estimator of Σ_0, however. Yeh *et al.* [37] proposed taking the EWMA of the running matrices $(X_t - \mu_0)(X_t - \mu_0)'$'s by defining

$$W_t = \lambda(X_t - \mu_0)(X_t - \mu_0)' + (1-\lambda)W_{t-1}, \tag{18}$$

where $0 < \lambda < 1$ is a smoothing constant and $W_0 = (X_1 - \mu_0)(X_1 - \mu_0)'$.

It can easily be shown that $E(W_t) = \Sigma_0$ and that W_t is positive-definite with probability one when $t \geq p$. Without loss of generality, let $\mu_0 = 0$ and $\Sigma_0 = I_p$. The authors proposed first separating the diagonal and upper off-diagonal elements of W_t and comparing them separately with diagonal and upper off-diagonal elements of I_p based on the Euclidean distance between two vectors. The two statistics are then combined to derive the statistic used for monitoring changes in the covariance matrix. Specifically, define

$$W_{t_v} = (w_{t(11)}, w_{t(22)}, \ldots, w_{t(pp)})'$$

and

$$W_{t_c} = (w_{t(12)}, w_{t(13)}, \ldots, w_{t(ij)}, \ldots, w_{t((p-1)p)})', \text{ for all } i < j$$

where W_{t_v} is a $p \times 1$ vector consisting of the p diagonal elements of W_t, and W_{t_c} is a $p(p-1)/2 \times 1$ vector consisting of the upper off-diagonal elements of W_t. The vector W_{t_v} is a natural estimator of the p population variances, while the vector W_{t_c} can be used to estimate the vector of $p(p-1)/2$ population covariances. To measure the deviation of W_{t_v} and W_{t_c} from the population parameter vectors, the sum of squared errors can be used by defining

$$D_{t1} = (W_{t_v} - 1_p)'(W_{t_v} - 1_p) \tag{19}$$

and

$$D_{t2} = W_{t_c}' W_{t_c}, \tag{20}$$

where 1_p is a $p \times 1$ vector of 1's. Furthermore, D_{t1} and D_{t2} can be combined to obtain

$$MaxD_t = \max\left[\frac{D_{t1} - E(D_{t1})}{\sqrt{Var(D_{t1})}}, \frac{D_{t2} - E(D_{t2})}{\sqrt{Var(D_{t2})}}\right]. \tag{21}$$

When the monitoring begins, $MaxD_t$ is calculated and plotted against t. The proposed chart signals as soon as the value of $MaxD_t$ exceeds a pre-determined UCL. The proposed chart is called the maximum multivariate exponentially weighted moving variability (MaxMEWMV) chart.

The authors derived the asymptotic expected value and variance of both D_{t1} and D_{t2}. They also provided, based on Monte-Carlo simulations, UCL's for p = 2, 3 and different values of λ when the ARL_0 is set to equal to approximately 370, equivalent to a 3σ Shewhart control chart.

It is important to note that the MaxMEWMV chart is specifically designed for individual observations, although it can easily be extended to the case when $n > 1$. Here, Σ_0 is assumed to be known and μ_0 is assumed unchanged.

13. Multivariate Extensions of Univariate EWMS and EWMV Charts

MacGregor and Harris [23] introduced two univariate control charts for monitoring process variance with individual observations. One is based on the EWMA of the mean squared deviations of observations, called the exponentially weighted moving mean square deviation (EWMS) chart, and the other is based on the EWMA of the updated variances, called the exponentially weighted moving variance (EWMV) chart. Expanding on the idea of MaxMEWMV chart (see Section 12), Huwang *et al.* [15] extended the two univariate control charts of MacGregor and Harris [23] to multivariate processes.

Based on W_t (Equation (18)), the EWMA of the running matrices $(X_t - \mu_0)$ $(X_t - \mu_0)'$, the authors proposed using $trace(W_t)$ to detect changes in the covariance matrix. Assuming that $\Sigma_0 = I_p$, it can be shown that $E(trace(W_t)) = p$ and $Var(trace(W_t))$ $= 2p[\lambda/2 - \lambda + (2 - 2\lambda/2 - \lambda)(1 - \lambda)^{2(t-1)}]$. Therefore, the control limits of the proposed chart are given by

$$p \pm L_s \sqrt{2p\left[\frac{\lambda}{2-\lambda} + \frac{2-2\lambda}{2-\lambda}(1-\lambda)^{2(t-1)}\right]}, \tag{22}$$

where L_s depends on λ and the desired ARL_0. Based on Monte-Carlo simulations, the authors provided values of L_s for which the ARL_0 is approximately equal to 370 for p = 2, 3 and $\lambda = .1, .2, \ldots, .9$. The proposed chart is called the multivariate exponentially weighted moving mean square error (MEWMS) chart.

The MEWMS chart assumes that μ_0 is known and does not change. However, if the mean also shifts, the MEWMS chart will be affected in such a way that the false alarm rate generally increases. It was demonstrated through numerical examples that if the mean shifts but Σ_0 remains unchanged, the MEWMS chart can no longer maintain its ARL_0. Moreover, if both μ_0 and Σ_0 change simultaneously, the out-of-control ARL's of MEWMS chart are smaller than those obtained when μ_0 stays unchanged.

In order to tackle the problem of potential mean shifts, the authors proposed the following modification to W_t, for $t \geq 1$,

$$V_t = \lambda(X_t - Y_t)(X_t - Y_t)' + (1-\lambda)V_{t-1}, \tag{23}$$

where $0 < \lambda < 1$ and $V_0 = (X_1 - Y_1)(X_1 - Y_1)'$. Here, Y_t is some estimate of the process mean, which in the paper was taken to be the multivariate exponentially weighted moving average (MEWMA) of X_t (Lowry *et al.* [22])

$$Y_t = w X_t + (1-w) Y_{t-1}, \tag{24}$$

where $0 < w < 1$ is a smoothing constant and $Y_0 = \mu_0$. The modified chart which uses $trace(V_t)$ to detect changes in covariance matrix is called the multivariate exponentially weighted moving variance (MEWMV) chart. Note that when $p=1$, MEWMS and MEWMV charts reduce to, respectively, the univariate EWMS and EWMV charts of MacGregor and Harris [23].

The control limits of the MEWMV chart are given by

$$E(trace(V_t)) \pm L_v \sqrt{Var(trace(V_t))}$$

$$= p\sum_{i=1}^{t} q_{ii} \pm \sqrt{2p\sum_{i=1}^{t}\sum_{j=1}^{t} q_{ij}^2}\,, \tag{25}$$

where q_{ij}, $i,j=1,2,\dots,t$, is the ith row and jth column element of a $t \times t$ matrix Q such that

$$Q=(I_t - M)'C(I_t - M).$$

Here $C = diag((1-\lambda)^{t-1}, \lambda(1-\lambda)^{t-2}, \dots, \lambda(1-\lambda), \lambda)$ and

$$M=\begin{pmatrix} w & 0 & \cdots & 0 \\ w(1-w) & w & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ w(1-w)^{t-1} & w(1-w)^{t-2} & w(1-w) & w \end{pmatrix}.$$

Based on Monte-Carlo simulations, the authors provided values of L_v which produce ARL_0 approximately equal to 370 for $p=2,3$ and λ, w = .1, .2, .3 and .4. It was also demonstrated that when there is a mean shift but Σ_0 remains unchanged, the ARL_0 of the MEWMV chart maintains at 370. Furthermore, when the mean and covariance matrix both change, the out-of-control ARL's of MEWMV chart are approximately the same as those obtained when only the covariance matrix changes.

14. A New EWMA Chart Based on Generalized Variance

A Shewhart chart, generally referred to as the $|S|$-chart, which is based on the sample generalized variance $|S_t|$ was developed in Alt and Smith [3]. A number of Shewhart charts discussed earlier also rely on $|S_t|$ or some function of it, such as the conditional entropy chart (Section 2) and the probability integral transformation based chart (Section 5). A good reference of the statistical properties of the sample generalized variance in the context of the $|S|$-chart can be found in Aparisi *et al.* [5].

If the objective is to detect changes in generalized variance, it is fairly easy to develop a multivariate EWMA chart. Specifically, it is known that (see for example Anderson [4]) if the process is in control (i.e., $X_t \sim N_p(\mu_0, \Sigma_0)$) then the distribution of

$$Y_t = \sqrt{\frac{n-1}{2p}}\, ln \frac{|S_t|}{|\Sigma_0|}$$

is asymptotically distributed as $N(0,1)$. Furthermore, if Σ_0 changes to Σ, then Y_t is asymptotically distributed as $N(ln|\Sigma|/|\Sigma_0|,1)$. In other words, the change in generalized variance from $|\Sigma_0|$ to $|\Sigma|$ in the original p-dimensional quality characteristics of interest now translates into a mean shift in Y_t. Therefore, one can devise an univariate

EWMA chart for detecting mean shifts in Y_t. Assuming Σ_0 is known, define, for $t \geq 1$,

$$G_t = \lambda Y_t + (1-\lambda) Y_{t-1}, \tag{26}$$

where $0 < \lambda < 1$ is a smoothing constant and $G_0 = 0$. The control limits of the EWMA chart are given by

$$\pm L \times \sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2t}]}. \tag{27}$$

If Σ_0 is not known, it can be estimated by $\overline{S}$, obtained from k training samples each of size n. The statistic Y_t needs to be modified to

$$Y_t^* = \sqrt{\frac{k(n-1)}{2p(k+1)}} ln \frac{|S_t|}{|\overline{S}|}.$$

If the process is in control, the Y_t^* is asymptotically distributed as $N(0,1)$. On the other hand, if Σ_0 changes to Σ, the Y_t^* is distributed asymptotically as $N(\sqrt{k/k+1} ln |\Sigma| / |\Sigma_0|, 1)$. In this case, the EWMA statistic is given by

$$G_t^* = \lambda Y_t^* + (1-\lambda) Y_{t-1}^*. \tag{28}$$

Even though Y_t and Y_t^* both are asymptotically normally distributed, the exact distribution could be quite skewed especially when sample size n is small to moderate. Furthermore, in many industrial applications large samples may not be readily available. Therefore, it is also of interest to use the proposed chart when n is small. For $p = 2$, we provide in Table 1 the LCL and UCL of the proposed EWMA chart based on G_t, i.e., when Σ_0 is assumed to be equal to I_p. The UCL's (similarly the LCL's), ignoring the term $(1-\lambda)^{2t}$, were obtained based on Monte-Carlo simulations such that the ARL_0 is approximately equal to 740. The standard errors of the simulations are all within 1% of the simulated ARL_0's. The LCL's and the UCL's are given for n ranging from 4 to 300 and $\lambda = .05, .1, .15$ and .2. For $p = 3$, the LCL's and the UCL's are given in Table 2.

Table 1. The LCL and UCL of the EWMA chart based on generalized variance ($p = 2$).

n	λ = .05		λ = .10		λ = .15		λ = .20	
	LCL	UCL	LCL	UCL	LCL	UCL	LCL	UCL
4	-3.582	-1.225	-4.310	-0.704	-4.925	-0.310	-5.491	0.022
5	-2.965	-0.831	-3.598	-0.345	-4.129	0.024	-4.609	0.355
8	-2.248	-0.338	-2.796	0.114	-3.244	0.464	-3.646	0.762
10	-2.036	-0.183	-2.558	0.263	-2.989	0.610	-3.370	0.904
15	-1.756	0.034	-2.255	0.475	-2.658	0.814	-3.016	1.109
20	-1.610	0.155	-2.095	0.590	-2.487	0.934	-2.834	1.226
40	-1.362	0.365	-1.826	0.802	-2.199	1.144	-2.532	1.438
60	-1.258	0.457	-1.719	0.894	-2.084	1.236	-2.411	1.529
80	-1.201	0.509	-1.657	0.946	-2.021	1.288	-2.341	1.585
100	-1.162	0.544	-1.616	0.982	-1.980	1.326	-2.291	1.623
150	-1.110	0.602	-1.554	1.046	-1.915	1.391	-2.229	1.690
200	-1.070	0.633	-1.513	1.080	-1.874	1.428	-2.188	1.726
300	-1.030	0.673	-1.469	1.120	-1.825	1.470	-2.140	1.774

Table 2. The LCL and UCL of the EWMA chart based on generalized variance ($p = 3$).

n	λ = .05		λ = .10		λ = .15		λ = .20	
	LCL	UCL	LCL	UCL	LCL	UCL	LCL	UCL
4	-10.524	-5.592	-12.110	-4.534	-13.460	-3.755	-14.727	-3.098
5	-7.330	-3.598	-8.448	-2.741	-9.377	-2.087	-10.227	-1.548
8	-4.890	-1.857	-5.747	-1.129	-6.443	-0.565	-7.077	-0.085
10	-4.285	-1.390	-5.093	-0.683	-5.750	-0.131	-6.337	0.328
15	-3.532	-0.782	-4.292	-0.098	-4.909	0.436	-5.435	0.888
20	-3.157	-0.465	-3.889	0.206	-4.483	0.733	-5.002	1.190
40	-2.529	0.082	-3.232	0.745	-3.795	1.261	-4.284	1.717
60	-2.281	0.309	-2.972	0.969	-3.516	1.488	-4.010	1.938
80	-2.133	0.442	-2.822	1.100	-3.368	1.619	-3.849	2.069
100	-2.040	0.530	-2.722	1.191	-3.266	1.708	-3.746	2.161
150	-1.891	0.669	-2.570	1.326	-3.113	1.847	-3.584	2.301
200	-1.806	0.750	-2.484	1.410	-3.021	1.928	-3.492	2.382
300	-1.705	0.845	-2.378	1.508	-2.920	2.026	-3.384	2.478

PART IV: CHART PERFORMANCE AND DIAGNOSTICS

15. Performance Comparisons

In this section, we will discuss the performance comparisons among different control charts that exist in the literature discussed so far. The performance comparisons are based on the out-of-control ARL's of competing charts.

(i) Among the Shewhart charts, Tang and Barnett [30] compared their proposed S_t-decomposition based chart (Section 3) with that of the $|S|$-chart and the Shewhart chart based on the likelihood ratio for testing $H_0 : \Sigma = \Sigma_0$ v.s. $H_a : \Sigma \neq \Sigma_0$ in the Σ_0 known and unknown cases (see for example Alt and Smith [3]). They found that the S_t-decomposition based chart is far more sensitive to covariance matrix changes considered in the paper than are the other two competing charts. Yeh and Lin [39] (Section 5) compared their probability integral transformation based Shewhart chart with the $|S|$-chart and found that, although the $|S|$-chart generally has slightly smaller out-of-control ARL's, these two charts have very comparable performance from a practical standpoint. The conclusion is that, among these Shewhart charts designed for the case when $n > p$, the S_t-decomposition based chart has the best performance and is recommended.

(ii) Among the CUSUM charts, Chan and Zhang [7] (Section 8) compared the performance of their proposed projection pursuit based CUSUM chart with that of the $|S|$-chart, likelihood ratio based Shewhart chart and another Shewhart chart derived from the Roy's maximum and minimum eigenvalues of sample covariance matrices (see for example Anderson [4]). They found that the projection pursuit based CUSUM chart generally produces smaller out-of-control ARL's than the other three competing Shewhart charts. The likelihood ratio based chart has better performance than the $|S|$-chart.

(iii) Among the EWMA charts, Yeh *et al.* [40] (Section 10) compared the probability integral transformation based EWMA chart with the $|S|$-chart in terms of the changes in generalized variance as expressed by $|\Sigma|/|\Sigma_0|$. They found that their proposed V-chart outperforms the $|S|$-chart, especially when a small smoothing constant is used in constructing the V-chart. Yeh *et al.* [38] (Section 11) compared their EWMLR chart with the multiple CUSUM charts (Section 7) and the multiple EWMA charts (Section 9). It was found that the EWMLR chart generally outperforms both multiple CUSUM and EWMA charts. The improvement in performance of the EWMLR chart is particularly noticeable when a small smoothing constant is used and when there exist moderate to strong correlations among variables. It was also found that the multiple CUSUM and EWMA charts could produce undesirable outcomes when only correlations change while variances remain unchanged in that the out-of-control ARL's of these two charts could be larger than ARL_0. Among these EWMA charts designed for the case when $n > p$, the EWMLR chart is recommended since it does not have the drawback that the generalized variance based EWMA charts do, i.e., different covariance matrices can produce the same generalized variance.

We next focus the discussion on monitoring multivariate individual observations, i.e., $n = 1$. Yeh *et al.* [37] compared their proposed MaxMEWMV chart (Section 12) with the multiple CUSUM and EWMA charts. They found that the MaxMEWMV chart outperforms the other two competitors in cases when (i) variances in variables increase with or without accompanying changes in correlations and (ii) when only correlations change but variances stay unchanged. It was also found that the performance of all three charts will be affected by the presence of mean shifts in such a way that the out-of-control ARL's decrease when mean shifts also occur.

The MEWMV and MEWMS charts (Huwang *et al.* [15], Section 13) were also compared to the multiple CUSUM and EWMA charts. When the process mean remains in control, both MEWMS and MEWMV charts outperform multiple CUSUM and EWMA charts, with the MEWMS chart performing slightly better than the MEWMV chart. Furthermore, when the process mean and covariance matrix change simultaneously, all of the MEWMS, multiple CUSUM and multiple EWMA charts will be affected by producing smaller out-of-control ARL's than they would if process mean stayed in control. The MEWMV chart, on the other hand, was not found to be affected by mean shifts, thus making it more robust than the other three charts to be used in detecting changes in a covariance matrix. Our conclusion is that, if the process mean remains unchanged, the MEWMS chart is recommended among the EWMA charts designed for the case when $n = 1$, since it has the best overall performance. However, if the process mean also shifts, the MEWMV chart is the only chart that is unaffected, and thus recommended.

Table 3. A summary of the recommended charts by sample size and chart type.

Sample Size	Chart Type	Shewhart	CUSUM	EWMA	
$n > p$	Recommended Chart	S_t-decomposition(3)[1]	Projection Pursuit (8)	EWMLR (11)	
	Comments	Not affected by mean shifts.	1. μ_0 is assumed known. 2. Affected by mean shifts.	Not affected by mean shifts.	
$n = 1$	Recommended Chart	Not Applicable	No Existing Performance Comparison	MEWMS (13)	MEWMV (13)
	Comments			1. μ_0 is assumed known. 2. Affected by mean shifts.	Not affected by mean shifts.

1. The number in parentheses indicates the section number in which the control chart is discussed.

Table 3 summarizes the recommended charts by sample size ($n > p$ and $n = 1$) and chart type. As pointed out by one referee, there are several features based on which control charts can be compared. Our focus in the current paper is to compare control charts based on how sensitive they are, in terms of ARL, to changes in the population covariance matrix.

Table 4. A summary of multivariate Shewhart control charts.

Requirement	Conditional Entropy (2)[1]	S_t Decomp. (3)	Two-Sample Test (4)	Prob. Int. Trans. (5)	Moving Ranges (6)
$n = 1$	no	no[2]	no	no	yes
$n > p$	required	required	required	required	not extended
developed assuming Σ_0 is known	yes	yes	not required	not required	yes
can use $\bar{S}$ as an estimate	yes	yes	required	required	no
μ_0 is known	not required	not required	not required	not required	not required
affected by mean shifts	no	no	no	no	yes
key feature	$\sum_{i=1}^{p} \ln\left(\frac{s_i^2}{\sigma_{io}^2}\right)$	decomposing S_t	likelihood ratio of $H_0: \Sigma = \Sigma_0$ v.s. $H_a: \Sigma \neq \Sigma_0$	i.i.d. $U(0,1)$'s	$(X_{t+1} - X_t)'\Sigma_0^{-1}(X_{t+1} - X_t)$

1. The number in parentheses indicates the section number in which the control chart is discussed.
2. The decomposition can be extended to the case when 1 < n < p.

Table 5. A summary of multivariate CUSUM control charts.

Requirement	Multiple CUSUM (7)[1]	Projection Pursuit (8)
$n = 1$	yes	yes
$n > p$	can be extended[2]	yes
developed assuming Σ_0 is known	yes	yes
can use $\bar{S}$ as an estimate	yes	yes
μ_0 is known	required	required
affected by mean shifts	yes	yes
key feature	regression adjusted variables and p univariate CUSUM's	eigenvectors corresponding to smallest and largest eigenvalues

1. The number in parentheses indicates the section number in which the control chart is discussed.
2. The chart can be extended to the case when $n > p$ (this was not discussed in the original paper).

Another important consideration is how sensitive a control chart is to changes in the population mean vector, for which the chart is not designed to detect. Unfortunately, there was very little discussion of this issue in the existing literature except in our earlier discussion in this section and Tables 4, 5 and 6. It is also important to consider how sensitive a control chart is to the violation of the underlying multivariate normality assumption. Investigation of this particular issue is, however, largely absent in the literature and therefore deserves further attention for future research.

Table 6. A summary of multivariate EWMA control charts.

Requirement	multivariate EWMA (9)[1]	V-chart (10)	EWMLR (11)	MaxMEWMV (12)	MEWMS (13)	MEWMV (13)	EWMA of $\|S_t\|$ (14)
$n=1$	yes	no	no	yes	yes	yes	no
$n>p$	can not extended[2]	required	required	was discussed[3]	can not extended[2]	can not extended[2]	required
developed assuming Σ_0 is known	yes	no	no	yes	yes	yes	yes
can use $\overline{S}$ as an estimate	yes	required	required	yes	yes	yes	yes
μ_0 is known	required	no	no	required	required	no	no
affected by mean shifts	yes	no	no	yes	yes	no	no
key feature	reg. adj. var. and p univariate EWMA's	EWMA of i.i.d. $U(0,1)$'s	EWMA of log-likeliho od ratio	EWMA of $(X-\mu_0)(X-\mu_0)'$ and L_2 norm	EWMA of $(X-\mu_0)(X-\mu_0)'$ and *trace*	EWMA of $(X-\hat{\mu})$ $(X-\hat{\mu})'$ and *trace*	EWMA of generalized var.

1. The number in parentheses indicates the section number in which the control chart is discussed.
2. The chart can be extended to the case when n > p (this was not discussed in the original paper).
3. It was discussed in the original paper that the MaxMEWMV chart can be extended to the case when n > p.

16. Possible Diagnostics After Detecting an Out-of-Control Signal

Another important problem is that of determining which parameters of the covariance matrix have actually changed when a control chart detects an out-of-control signal. Unlike the case of the process mean with p parameters, there are a total of $p(p+1)/2$ parameters in the covariance matrix that could change individually or in combination which could potentially trigger an out-of-control signal.

For control charts that are derived based on the sample generalized variance, an out-of-control signal is interpreted as a change in the generalized variance, i.e., an increase or a decrease in the determinant of the covariance matrix. When the process is in control, $|\Sigma_0|$ is proportional to the square of the volume of the ellipsoid generated by $\{X \in R^p, (X-\mu_0)'\Sigma_0^{-1}(X-\mu_0) \le C^2\}$, which is the form of the confidence region for the mean vector under normality assumption. Therefore, an increase or a decrease in the generalized variance is also associated with an increase or a decrease in the volume of the confidence region of the mean vector. However, a major limitation for the generalized variance is that different matrices can produce the same determinant.

In developing the conditional entropy chart (Section 2), Guerrero-Cusumano [9] argued that since the chart is essentially based on the ratios of sample variance over population variance for each of the p variables, when the chart gives an out-of-control signal, one can proceed to find out which of these variances are out-of-control. It was suggested that one use the Bonferroni probability inequality to set up confidence intervals for each of the p population variances. One obvious limitation of such an approach is that it is not capable of capturing changes in correlations among variables.

The multiple CUSUM and multiple EWMA charts (Sections 7 and 9) have the same advantage as the conditional entropy chart. Since they are essentially p univariate CUSUM or EWMA charts, one can monitor each of the p univariate charts. When the overall chart detects an out-of-control signal, one can proceed to determine which of these univariate charts also signal. The multiple CUSUM and EWMA charts share the same drawback as does the conditional entropy chart, namely that they lack the ability to capture changes in correlations.

For the charts that are based on the likelihood ratio of testing $H_0: \Sigma = \Sigma_0$ v.s. $H_a: \Sigma \neq \Sigma_0$, the problem becomes a one-sample problem if Σ_0 is assumed known, or a two-sample problem if Σ_0 is assumed unknown. When the chart signals one might consider performing a series of hierarchical likelihood ratio testing procedures proposed by Manly and Rayner [24] (also see Section 8.3 of Wierda [33]). These testing procedures are designed to test, in a series of steps, whether (1) Σ and Σ_0 differ only in correlations; (2) Σ and Σ_0 differ only in variances; and (3) $\Sigma = c\Sigma_0$ where $c > 0$, $\neq 1$ is a constant. If the result of test (1) is significant, it is concluded that Σ and Σ_0 differ only in correlations. If the result of test (1) is not significant, proceed to test (2). If the result of test (2) is significant, it is concluded that Σ and Σ_0 differ only in variances, whereas the correlations are equal. If the result of test (2) is not significant, perform test (3). If the result of test (3) is significant, it is concluded that $\Sigma = c\Sigma_0$, $c > 0$, $\neq 1$. If the result of test (3) is not significant, it is then concluded that $\Sigma = \Sigma_0$.

In the MaxMEWMV chart (Section 12), the statistic is based on the maximum of D_{t1} and D_{t2} (Equations (19) and (20)), where D_{t1} and D_{t2} are the squared errors for the sample variances and the sample covariances, respectively. It is suggested that one also monitor D_{t1} and D_{t2}. When the MaxMEWMV chart signals, depending on whether D_{t1} or D_{t2} (or both) signal, it can be interpreted as the variances or the correlations (or both) are out of control. One limitation of such an approach is that it is not capable of showing exactly which of the variances or which of the correlations are out of control.

In a recent study, Apley and Shi [6] proposed a diagnostic technique for identifying root causes for changes in process variability which closely resembles factor analysis. However, the objective is not to identify which of the variances of the p variables or correlations among the p variables have changed. Rather, it is based on a fault model which assumes that there are m fault factors ($m < p$) acting independently and that each of the p correlated quality characteristics is affected by some linear combination of the m uncorrelated fault factors. Under such a framework, the problem of diagnosing is transformed into one whose objective is to estimate the number of faults m that are contributing to process variability, as well as the linear combinations by which each of the p correlated quality characteristics is affected.

Chen and Hong [8] developed another diagnostic technique which is based on decomposing S_t (via Barlett or Cholesky decomposition) and turning S_t into a matrix T. When the process is in control, the squares of the diagonal elements of T are i.i.d. χ^2 with various degrees of freedom and the squares of the off-diagonal elements of T are also i.i.d. χ_1^2. By linking the changes in each of the variances and correlations in the covariance matrix to the changes one might expect to observe from T, the authors developed diagnostic rules by observing the patterns of changes from T.

17. Concluding Remarks

In the preceding sections, we reviewed numerous multivariate control charts, developed between 1990 and 2005, which are designed to detect changes in process variability as measured by the covariance matrix. As previously mentioned, the review focused on Phase II control charts designed for multivariate normal processes, assuming that independent subgroups of observations or independent individual observations are being collected as the process monitoring proceeds. Therefore, we did not discuss other types of control charts such as the nonparametric procedures designed to detect changes in covariance matrix as defined by changes of some function of the matrix (Hawkins [12]), the data depth based nonparametric control charts for non-normal processes (Liu [20]), the principal component analysis and dissimilarity index based control charts designed for multivariate time-series data (see, for example, Kano *et al.* [17] and references therein), and the likelihood ratio based preliminary Shewhart chart designed for use in Stage 1 (retrospective stage) of the Phase I control (Sullivan and Woodall [29]).

Tables 4, 5 and 6 summarize, respectively, the Shewhart charts, the CUSUM charts and the EWMA charts in terms of the sample size requirement, the population parameter assumptions, and whether the charts will be affected by the presence of mean shifts. Some observations emerge from the control charts discussed herein.

(i) When the sample covariance matrix can be computed and has full rank, the approaches typically rely on either the sample generalized variance or the likelihood ratio associated with testing the equality of two matrices (Sections 4, 5, 10,11 and 14). In this context, Σ_0 is either known or can be estimated from in-control training samples. Exceptions include the conditional entropy control chart (Section 2) which relies on the sum of the logarithms of the ratios of sample variance over population variance for each of the p variables, the control chart (Section 3) which relies on the sum of $2p-1$ independent χ_1^2 derived from decomposing the sample covariance matrix, and the control chart developed based on projection pursuit method (Section 8). These charts (Sections 2, 3, 4, 5, 10, 11 and 14) typically are not affected by the presence of mean shifts, with the exception of the projection pursuit control chart. Among the Shewhart charts, we recommend using the S_t-decomposition based chart since it generally outperforms both the likelihood ratio and generalized variance based charts. The projection pursuit CUSUM chart also outperforms both the likelihood ratio and generalized variance based Shewhart charts. Additional research is needed to determine whether the S_t-decomposition Shewhart chart or the projection pursuit CUSUM has a better performance. Among the EWMA charts, we recommend using the EWMLR chart since it can detect changes in a covariance matrix in which the out-of-control covariance matrix has the same generalized variance as the in-control covariance matrix.

(ii) Numerous control charts have been developed for use when only individual observations are available. These include the control chart based on moving ranges (Section 6), the multiple CUSUM and EWMA charts (Sections 7 and 9), the projection pursuit based chart (Section 8), the MaxMEWMV chart (Section 12), and the MEWMS and MEWMV charts (Section 13). All, except the MEWMV chart, implicitly assume that the process mean μ_0 stays in control during process monitoring, and therefore the performance of these charts will be affected if mean shifts also take place. The out-of-control ARL's of these charts in general are smaller when the process mean also shifts than when the mean stays in control, leading to increased false alarms. Therefore, we recommend using the MEWMV chart when $n=1$ and when the process is subject to both mean shifts and covariance matrix changes.

As previously discussed, the performance comparisons that exist in the literature were scattered and limited in their scopes. Therefore, one important concern worthy of future investigation is to compare all the existing charts in a systematic, organized and thorough manner. Another important area of potential future research is diagnostic techniques. Such a task is more complicated than with the multivariate process mean due to the complexity of the covariance matrix. Because so many parameters are contained in the covariance matrix and that changes in one or some of the parameters can trigger an out-of-control signal, it is of eminent importance to be able to further pinpoint which of these parameters are out-of-control.

Acknowledgments

We would like to thank two referees for their numerous suggestions, which improved the presentation of the paper.

References

1. Alt, F. B. (1984). Multivariate quality control, *The Encyclopedia of Statistical Sciences*, Kotz, S., Johnson, N. L. and Read, C. R., (eds) John Wiley, New York, 110-122.
2. Alt, F. B. and Bedewi, G. E. (1986). SPC for dispersion for multivariate data. *ASQC Quality Congress Transactions*, 248-254.
3. Alt, F. B. and Smith, N. D. (1988). Multivariate process control, *Handbook of Statistics*, Krishnaiah, P. R. and Rao, C. R., (eds) Elsevier Science Publishers, New York, 333-351.
4. Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd edition. John Wiley & Sons, New York.
5. Aparisi, F., Jabaloyes, J. and Carrión, A. (1999). Statistical properties of the $|S|$ multivariate control chart. *Communications in Statistics - Theory and Methods*, 28, 2671-2686.
6. Apley, D. W. and Shi, J. (2001). A factor-analysis method for diagnosing variability in multivariate manufacturing processes. *Technometrics*, 43, 84-95.
7. Chan, L. K. and Zhang, J. (2001). Cumulative sum control charts for the covariance matrix. *Statistica Sinica*, 11, 767-790.
8. Chen, A. and Hong, H.-J. (2004). *Fault detection and classification by sample covariance matrix and its application to plasma etcher.* Preprint, National Taiwan University, Taiwan.
9. Guerrero-Cusumano, J.-L. (1995). Testing variability in multivariate quality control: A conditional entropy measure approach. *Information Sciences*, 86, 179-202.
10. Hawkins, D. M. (1981). A cusum for a scale parameter. *Journal of Quality Technology*, 13, 228-231.
11. Hawkins, D. M. (1991). Multivariate quality control based on regression-adjusted variables. *Technometrics*, 33, 61-75.
12. Hawkins, D. M. (1992). Detecting shifts in functions of multivariate location and covariance parameters. *Journal of Statistical Planning and Inference*, 33, 233-244.
13. Hawkins, D. M. (1993). Regression adjustment for variables in multivariate quality control. *Journal of Quality Technology*, 25, 170-182.
14. Healy, J. D. (1987). A note on multivariate CUSUM procedures. *Technometrics*, 29, 409-412.
15. Huwang, L., Yeh, A. B. and Wu, C.-W. (2005). Monitoring multivariate process variability for individual observations. Preprint (submitted for publication), National Tsing Hua University, Taiwan.
16. Johnson, N. L. and Leone, F. C. (1962). Cumulative sum control charts. *Industrial Quality Control*, 18, 15-21; 19, 22-36.

17. Kano, M., Nagao, K., Hasebe, S., Hashimoto, I., Ohno, H., Strauss, R. and Bakshi, B. R. (2002). Comparison of multivariate statistical process monitoring methods with applications to the Eastman challenge problem. *Computers and Chemical Engineering*, 26, 161-174.

18. Khoo, M. B. and Quah, S. H. (2003). Multivariate control chart for process dispersion based on individual observations. *Quality Engineering*, 15, 639-642.

19. Levinson, W., Holmes, D. S. and Mergen, A. E. (2002). Variation charts for multivariate processes. *Quality Engineering*, 14, 539-545.

20. Liu, R. Y. (1995). Control charts for multivariate processes. *Journal of the American Statistical Association*, 90, 1380-1387.

21. Lowry, C. A. and Montgomery, D. C. (1995). A review of multivariate control charts. *IIE Transactions in IIE Research*, 27, 800-810.

22. Lowry, C. A., Woodall, W. H., Champ, C. W. and Rigdon, S. E. (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, 34, 46-53.

23. MacGregor, J. F. and Harris, T. J. (1993). The exponentially weighted moving variance. *Journal of Quality Technology*, 25, 106-118.

24. Manly, B. F. J. and Rayner, J. C. W. (1987). The comparison of sample covariance matrices using likelihood ratio tests. *Biometrika*, 74, 841-847.

25. Mason, R. L., Champ, C. W., Tracy, N. D., Wierda, S. J. and Young, J. C. (1997). Assessment of multivariate process control techniques. *Journal of Quality Technology*, 29, 140-143.

26. Montgomery, D. C. (2001). *Introduction to Statistical Quality Control*, 4th edition. John Wiley & Sons, New York.

27. Stoumbos, Z. G., Reynolds, M. R., Jr., Ryan, T. P. and Woodall, W. H. (2000). The state of statistical process control as we proceed into the 21st century. *Journal of the American Statistical Association*, 95, 992-998.

28. Sugiura, N. and Nagao, H. (1968). Unbiasedness of some test criteria for the equality of one or two covariance matrices. *The Annals of Mathematical Statistics*, 30, 1686-1692.

29. Sullivan, J. H. and Woodall, W. H. (2000). Change-point detection of mean vector or covariance matrix shifts using multivariate individual observations. *IIE Transactions on Quality and Reliability Engineering*, 32, 537-549.

30. Tang, P. F. and Barnett, N. S. (1996a). Dispersion control for multivariate processes. *The Australian Journal of Statistics*, 38, 235-251.

31. Tang, P. F. and Barnett, N. S. (1996b). Dispersion control for multivariate processes – Some comparisons. *The Australian Journal of Statistics*, 38, 253-273.

32. Vargas N., J. A. (2003). Robust estimation in multivariate control charts for individual observations. *Journal of Quality Technology*, 35, 367-376.

33. Wierda, S. J. (1994). *Multivariate Statistical Process Control*, Wolters-Noordhoff, Groningen, The Netherlands.

34. Woodall, W. H. (2000). Controversies and contradictions in statistical process control (with discussions). *Journal of Quality Technology*, 32, 341-378.

35. Woodall, W. H. and Montgomery, D. C. (1999). Research issues and ideas in statistical process control. *Journal of Quality Technology*, 31, 376-386.

36. Woodall, W. H. and Ncube, M. M. (1985). Multivariate CUSUM quality control procedures. *Technometrics*, 27, 285-292.

37. Yeh, A. B., Huwang, L. and Wu, C.-W. (2005). A multivariate EWMA control chart for monitoring process variability with individual observations. *IIE Transactions on Quality and Reliability Engineering*, 37, 1023-1035.

38. Yeh, A. B., Huwang, L. and Wu, Y.-F. (2004). A likelihood ratio based EWMA control chart for monitoring variability of multivariate normal processes. *IIE Transactions on Quality and Reliability Engineering*, 36, 865-879.

39. Yeh, A. B. and Lin, D. K.-J. (2002). A new variables control chart for simultaneously monitoring multivariate process mean and variability. *International Journal of Reliability, Quality and Safety Engineering*, 9, 41-59.

40. Yeh, A. B., Lin, D. K.-J., Zhou, H. and Venkataramani, C. (2003). A multivariate exponentially weighted moving average control chart for monitoring process variability. *Journal of Applied Statistics*, 30, 507-536.

Authors' Biographies:

Arthur B. Yeh received Ph.D. in statistics from Rutgers University in 1993. He is a Professor of Statistics in the Department of Applied Statistics and Operations Research at Bowling Green State University. Dr. Yeh has conducted and published research in several areas of industrial statistics, including optimal experimental designs, univariate and multivariate control charts, multivariate process capability indices, and multivariate run-by-run process control. He has also worked as a consultant for various local and international companies in both traditional and modern high-tech manufacturing environments. He had served in the past as the President of the Northwest Ohio Chapter of the American Statistical Association, and the Chair of the Toledo Section of the American Society for Quality. He currently serves as an Associate Editor for *The American Statistician* and as a director for the Toledo Section of the American Society of Quality. Dr. Yeh is a member of the American Statistical Association and the Institute of Mathematical Statistics, and a senior member of the American Society for Quality.

Dennis K. J. Lin is a University Distinguished Professor of Supply Chain and Statistics at the Penn State University. His research interests are quality engineering, industrial statistics, data mining and response surface. He has published over 100 papers in a wide variety of journals. He serves (or served) as associate editor for *Statistica Sinica, American Statisticians, Journal of Data Science, Quality Technology & Quality Management, Journal of Quality Technology*; and *Taiwan Outlook.* Dr. Lin is a Fellow of the American Statistical Association (ASA), an elected member of International Statistical Institute (ISI), a Fellow of American Society of Quality (ASQ), a lifetime member of International Chinese Statistical Association (ICSA), and a Fellow of the Royal Statistical Society. He is an honorary chair professor for various universities, including National Chengchi University (Taiwan), Remin University of China and XiAn Statistical Institute.

Richard N. McGrath is an Associate Professor in the Department of Applied Statistics and Operations Research at Bowling Green State University. His research interests include all areas of industrial statistics with a special interest in analysis of dispersion. Dr. McGrath is a member of the American Statistical Association and the Institute of Mathematical Statistics, and a senior member of the American Society for Quality.