VICTORIA UNIVERSITY OF TECHNOLOGY

DEPARTMENT OF COMPUTER AND MATHEMATICAL SCIENCES

Dispersion Control for Multivariate Processes

Pak Fai Tang and Neil Barnett

(65 EQRM 20)

February, 1996

(AMS : 62N10)

TECHNICAL REPORT

VICTORIA UNIVERSITY OF TECHNOLOGY
DEPARTMENT OF COMPUTER AND MATHEMATICAL SCIENCES
P O BOX 14428
MCMC
MELBOURNE, VICTORIA 8001
AUSTRALIA

TELEPHONE (03) 9688 4492
FACSIMILE (03) 9688 4050

Footscray Campus

DISPERSION CONTROL FOR MULTIVARIATE PROCESSES

PAK FAI TANG [1] AND NEIL BARNETT [1]
Victoria University of Technology

Summary

Some techniques for monitoring and controlling the dispersion of multivariate normal processes based on subgroup data are presented. The procedures involve use of independent statistics resulting from the decomposition of the covariance matrix. Those that do not depend on prior estimates of the process covariance matrix are particularly attractive to short-run or low volume manufacturing environments.

Key words : Dispersion control; multivariate normal processes; rational subgroups; probability integral transformation; short production runs.

1. Introduction

Over the last decade, the problem of multivariate quality control has received considerable attention in the literature (see for eg., Woodall et al.(1985), Murphy (1987), Healy (1987), Crosier (1988), Pignatiello et al.(1990), Doganaksoy et al.(1991), Sparks (1992), Tracy et al.(1992), Lowry et al.(1992), Hawkins (1991,1993), Hayter et al.(1994), Chan et al.(1994) and Mason et al.(1995)). This

[1] Dept of Computer and Mathematical Sciences,
PO Box 14428, MMC Melbourne, Victoria 8001, Australia

work has focussed on the detection of parameter changes, departures from distributional assumptions, and the identification of out-of-control variables. Most of the work is based on the assumption that the observation vectors, $\mathbf{X}_i's$ are independently and identically distributed (*i.i.d*) multivariate normal variables and that the true values of the process parameters, in particular the process variance-covariance matrix, Σ, are known. Recently, Tang (1995) has developed a procedure for monitoring the mean level of multivariate normal processes for situations where prior information about the in-control process parameters is unavailable. He demonstrated that this procedure is particularly useful when subgroup data are used.

Whilst substantial work has been devoted to the control of the process mean vector, μ, very little emphasis has been placed on the importance of monitoring and controlling Σ. In fact, the issue is a formidable one due to the complexity of the distribution theory involved. One exception is the paper by Alt et al.(1986) who proposed two control techniques for Σ; one based on the likelihood ratio principle and the other that makes use of the sample generalized variance, which is sometimes taken as a measure of dispersion or spread of multivariate processes. Although traditional multivariate control charts such as the Hotelling χ^2 or T^2 charts may signal certain shifts in Σ (see Hawkins (1991) and Tracy et al.(1992)), other particular changes in Σ will remain undetected. This is also true for the technique based on generalized variance. For instance, if Σ shifts in such a way that the resulting process region (i.e the ellipsoidal region in which almost all observations fall) is contained completely within the undisturbed one, this 'shrunken' process is unlikely to be detected by a χ^2 chart, especially when the sample size is small. In addition, Hawkins (1991) stated that 'measures based on quadratic forms (like T^2) also confound mean shifts with variance shifts and require quite extensive analysis following a signal to determine the nature of the shift'. Note that the 'T^2' term that he used actually refers to the more commonly called χ^2 statistic which uses (presumably) the true value of the process covariance matrix. When 'special' or 'assignable' causes affecting both process parameters are present, it is also possible that the effect of the mean (vector) shift is masked or 'diluted' by the accompanying change in the variance-covariance matrix.

The purpose of this paper is to present some control procedures for the dispersion of multivariate normal processes based on subgroup data. Special attention is drawn to the situations where prior information about Σ is not available as is often the case in situations of short production runs, which have become increasingly prevalent. When Σ is specified or assumed known, the proposed procedure involves the decomposition of the sample covariance matrix and uses the resulting independent components, which have meaningful interpretations, as the bases for checking the constancy of the process covariance matrix. Another possible approach is also outlined for this case. As for the case where Σ is unknown in advance of production, the proposed procedure is adapted from the step-down test of Anderson (1984, p.417) which is based on the decomposition of the likelihood ratio statistic for testing the equality of several covariance matrices. When these procedures are used together with Hotelling χ^2 or T^2-type charts, they supplement the latter by providing independent information about the stability of the process covariance matrix. Furthermore, the proposed procedures effectively replace existing competing techniques to provide enhanced detection of general shifts in Σ.

This paper is organized into three subsequent sections. In 2, the underlying methodology is presented. In 3, appropriate control statistics are given for both the cases regarding prior knowledge or lack of prior knowledge of the process covariance matrix. In a sequel to this paper, comparisons are made between the proposed techniques and various competing procedures. The total discourse is given in the context of the manufacture of discrete items.

2. Methodology

Suppose that the vectors of observations on p correlated product characteristics, $\mathbf{X}_i's$ follow a multivariate normal $N_p(\mu, \Sigma)$ distribution with mean vector μ and covariance matrix Σ when the process is operating under stable conditions. In practice, the validity of this assumption should be checked using, for example, a multivariate normal goodness-of-fit test (Gnanadesikan (1977)). The aim here is to develop control procedures for monitoring and controlling the *dispersion* of such a multivariate process based on *rational* subgroups where the sample size, n may vary.

In order to provide more flexibility, ease of implementation and better control of the false alarm rate than existing procedures as well as to facilitate the interpretation of out-of-control signals, it is suggested that the sample variance-covariance matrix be partitioned into various statistically independent components having physical interpretation and known distributions. These components are then used to indicate the stability of the process covariance matrix.

It is well known, that under the stable or in-control normality assumption, the sample covariance matrix, $\mathbf{S}$, multiplied by the factor $(n-1)$ follows the *Wishart* distribution with parameters $(n-1)$ and Σ, denoted by

$$(n-1)\mathbf{S} \sim W_p(n-1,\Sigma)$$

Let the sample and population covariance matrices be similarly expressed in partitioned form as follows :-

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix} = \begin{pmatrix} S_1^2 & \underset{\sim}{\mathbf{S}}_{12}^{\mathrm{T}} \\ \underset{\sim}{\mathbf{S}}_{12} & \mathbf{S}_{22} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \underset{\sim}{\Sigma}_{12}^{\mathrm{T}} \\ \underset{\sim}{\Sigma}_{12} & \Sigma_{22} \end{pmatrix}$$

where $S_1^2\ (\sigma_1^2)$, $\underset{\sim}{\mathbf{S}}_{12}\ (\underset{\sim}{\Sigma}_{12})$ and $\mathbf{S}_{22}\ (\Sigma_{22})$ denote respectively the sample (population) variance of the 1st variable, the vector of sample (population) covariances between the 1st and each of the remaining variables, and the sample (population) covariance matrix excluding the 1st variable. Next, define

$$\begin{aligned} \mathbf{S}_{22\bullet 1} &= \mathbf{S}_{22} - \mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12} \\ &= \mathbf{S}_{22} - \frac{\underset{\sim}{\mathbf{S}}_{12}\underset{\sim}{\mathbf{S}}_{12}^{\mathrm{T}}}{S_1^2}, \end{aligned}$$

then according to a well-known theorem (see for eg., theorem 6.4.1, p.120, Giri (1977), where $\Sigma_{(22)} - \Sigma_{(21)}\Sigma_{(11)}^{-1}\Sigma_{(12)}$ in (c) should be replaced by $\Sigma_{(11)} - \Sigma_{(12)}\Sigma_{(22)}^{-1}\Sigma_{(21)}$ using his notation),

(i) $\mathbf{S}_{22\bullet 1}$ is *independent* of $\left(S_1^2, \underset{\sim}{\mathbf{S}}_{12}^{\mathrm{T}}\right)$,

(ii) $(n-1)S_1^2 \sim \sigma_1^2\chi_{n-1}^2$,

(iii) $(n-1)\mathbf{S}_{22\bullet1} \sim W_{p-1}(n-2, \Sigma_{22\bullet1})$ where $\Sigma_{22\bullet1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$

$$= \Sigma_{22} - \frac{\underset{\sim}{\Sigma_{12}}\underset{\sim}{\Sigma_{12}^{T}}}{\sigma_1^2},$$

(iv) The *conditional* distribution of $\dfrac{\underset{\sim}{\mathbf{S}_{12}}}{S_1^2}$, given $S_1^2 = s_1^2$, is $N\left(\dfrac{\underset{\sim}{\Sigma_{12}}}{\sigma_1^2}, \dfrac{\Sigma_{22\bullet1}}{(n-1)s_1^2}\right)$.

Note that $\mathbf{S}_{22\bullet1}$ and $\Sigma_{22\bullet1}$ here denote respectively the conditional sample and population covariance matrices of the last (p-1) variables given the 1st. Note also that $\dfrac{\underset{\sim}{\mathbf{S}_{12}}}{S_1^2}$ and $\dfrac{\underset{\sim}{\Sigma_{12}}}{\sigma_1^2}$ represent respectively the vectors of sample and population regression coefficients when each of the last (p-1) variables is regressed on the 1st variable. Furthermore, S_1^2, $\dfrac{\underset{\sim}{\mathbf{S}_{12}}}{S_1^2}$ and $\mathbf{S}_{22\bullet1}$ may be regarded as *independent* components following the above decomposition of the sample covariance matrix $\mathbf{S}$. Further, decomposing $\mathbf{S}_{22\bullet1}$ in the same manner yields $S_{2\bullet1}^2$ (the conditional sample variance of the 2nd variable given the 1st one), $\dfrac{\underset{\sim}{\mathbf{S}_{12\bullet1}}}{S_{2\bullet1}^2}$ (the vector of regression coefficients when each of the last $(p-2)$ variables is regressed on the 2nd whilst the 1st variable is held fixed) and $\mathbf{S}_{3,4,\ldots,p\,\bullet\,1,2}$ (the conditional sample covariance matrix of the last $(p-2)$ variables given the first two) which are independently distributed as

$$S_{2\bullet1}^2 = \left(1 - R_{12}^2\right)S_2^2 \sim \frac{(1-\rho_{12}^2)\sigma_2^2}{(n-1)}\chi_{n-2}^2$$

$$\frac{\underset{\sim}{\mathbf{S}_{12\bullet1}}}{S_{2\bullet1}^2}\Bigg/ S_{2\bullet1}^2 = s_{2\bullet1}^2 \sim N\left(\frac{\underset{\sim}{\Sigma_{12\bullet1}}}{\sigma_{2\bullet1}^2}, \frac{\Sigma_{3,\ldots,p\,\bullet\,1,2}}{(n-1)s_{2\bullet1}^2}\right)$$

$$\mathbf{S}_{3,\ldots,p\bullet 1,2} \sim \frac{1}{(n-1)} W_{p-2}\left(n-3, \Sigma_{3,\ldots,p\bullet 1,2}\right)$$

where R_{12}^2 and ρ_{12}^2 are respectively the squares of the sample and population correlations between the 1st and the 2nd variables. Repeating the above procedures until further decomposition is impossible results in *p scaled* chi-square variables S_1^2, $S_{j\bullet 1,\ldots,j-1}^2$, $j = 2,\ldots,p$ and *p*-1 *conditional* (univariate or multivariate) normal variables which are independent and have meaningful interpretations. $S_{j\bullet 1,\ldots,j-1}^2$ here denotes the conditional sample variance of the *j*th variable given the first *j*-1 variables.

Note that the ordering of the variables is *not* unique. In fact, there are $p!$ possible permutations each of which results in $(2p-1)$ terms in the decomposition. If all of these $p! \times (2p-1)$ variables are used as the control statistics, there will be a multitude of control charts even when *p* is quite small. For instance, when $p = 3$, there will be $p!(2p-1) = 30$ terms in total that can be obtained from the various decompositions. When $p = 8$, this number increases to $p!(2p-1) = 604800$ which clearly renders the approach impractical ! Furthermore, there are component variables in common to the various partitionings and terms that reflect essentially the same information. Therefore, one particular arrangement of the variables is deemed to be adequate for the purpose of decomposition. It is suggested that the choice of this should reflect the relative importance of the variables involved. In particular, the variables should be arranged in decreasing order of importance from 1 to *p*. For the case of a 'cascade' process as described by Hawkins (1993), the variables should be arranged from the most 'upstream' (being the 1st) to the most 'downstream' one (being the last) so that a shift in a variable will not be masked by the accompanying change in the downstream variables.

If Σ is specified or assumed known in advance of production, the statistics obtained in the above manner can of course be used separately to monitor the dispersion of the multivariate process. However, due to independence, these statistics can be combined into a single aggregate-type control statistic as considered in the next section. In practice, if the latter approach is adopted, it is recommended that the values of the individual statistics be retained for post-signal analysis.

To illustrate the above idea, consider the case of $p = 3$ product characteristics. Using conventional notation, the sample covariance matrix of n observations on these variables is given by

$$\mathbf{S} = \begin{pmatrix} S_1^2 & R_{12}S_1S_2 & R_{13}S_1S_3 \\ R_{12}S_1S_2 & S_2^2 & R_{23}S_2S_3 \\ R_{13}S_1S_3 & R_{23}S_2S_3 & S_3^2 \end{pmatrix}$$

Letting $\mathbf{S}_{11} = S_1^2$, $\mathbf{S}_{12} = (R_{12}S_1S_2,\ R_{13}S_1S_3)$, $\mathbf{S}_{22} = \begin{pmatrix} S_2^2 & R_{23}S_2S_3 \\ R_{23}S_2S_3 & S_3^2 \end{pmatrix}$ and proceeding as previously, we have

$$\begin{aligned}(n-1)\mathbf{S}_{22\bullet1} &= (n-1)\left[\mathbf{S}_{22} - \mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\right] \\ &= (n-1)\begin{pmatrix} S_2^2(1-R_{12}^2) & S_2S_3(R_{23}-R_{12}R_{13}) \\ S_2S_3(R_{23}-R_{12}R_{13}) & S_3^2(1-R_{13}^2) \end{pmatrix} \sim W_2\left(n-2, \Sigma_{22\bullet1}\right)\end{aligned}$$

independently distributed of $(\mathbf{S}_{11}, \mathbf{S}_{12}) = \left(S_1^2,\ R_{12}S_1S_2,\ R_{13}S_1S_3\right)$ and

$$(n-1)S_1^2 \sim \sigma_1^2\chi^2(n-1) \tag{1}$$

$$\mathbf{S}_{21}\mathbf{S}_{11}^{-1} \mid \mathbf{S}_{11} = \mathbf{s}_{11} \equiv \left(\frac{R_{12}S_2}{S_1}, \; \frac{R_{13}S_3}{S_1} \right)^{\mathrm{T}} \Big/ S_1^2 = s_1^2 \; \sim N\left(\frac{\Sigma_{21}}{\sigma_1^2}, \; \frac{\Sigma_{22\bullet 1}}{(n-1)s_1^2} \right)$$
$$\sim N\left(\left(\frac{\rho_{12}\sigma_2}{\sigma_1}, \; \frac{\rho_{13}\sigma_3}{\sigma_1} \right)^{\mathrm{T}}, \; \frac{\Sigma_{22\bullet 1}}{(n-1)s_1^2} \right) \tag{2}$$

where

$$\Sigma_{22\bullet 1} = \left[\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \right]$$
$$= \begin{pmatrix} \sigma_2^2(1-\rho_{12}^2) & \sigma_2\sigma_3(\rho_{23}-\rho_{12}\rho_{13}) \\ \sigma_2\sigma_3(\rho_{23}-\rho_{12}\rho_{13}) & \sigma_3^2(1-\rho_{13}^2) \end{pmatrix}$$

Note that $\mathbf{S}_{21}\mathbf{S}_{11}^{-1} = (R_{12}S_2 / S_1, \; R_{13}S_3 / S_1)^{\mathrm{T}}$ represents the vector of regression coefficients when each of the 2nd and 3rd variables is regressed on the 1st variable. Further, decomposing $\mathbf{S}_{22\bullet 1}$ in the same manner yields independent components,

$$\left(S_{2\bullet 1}^2 \;, \; R_{32\bullet 1}S_{2\bullet 1}S_{3\bullet 1} \right) = \left(S_2^2(1-R_{12}^2), \; S_2S_3(R_{23}-R_{12}R_{13}) \right)$$

and

$$S_{3\bullet 1,2}^2 = S_3^2\left(1 - R_{3(1,2)}^2\right)$$

where

$$(n-1)S_2^2\left(1-R_{12}^2\right) \sim \sigma_2^2\left(1-\rho_{12}^2\right)\chi_{n-2}^2 \tag{3}$$

$$\frac{S_3(R_{23}-R_{12}R_{13})}{S_2\left(1-R_{12}^2\right)} \Big/ S_2^2\left(1-R_{12}^2\right) = s_2^2(1-r_{12}^2) \sim N\left(\frac{\sigma_3(\rho_{23}-\rho_{12}\rho_{13})}{\sigma_2(1-\rho_{12}^2)}, \; \frac{\sigma_3^2(1-\rho_{3(1,2)}^2)}{(n-1)s_2^2(1-r_{12}^2)} \right) \tag{4}$$

$$(n-1)S_3^2\left(1-R_{3(1,2)}^2\right) \sim \sigma_3^2\left(1-\rho_{3(1,2)}^2\right)\chi_{n-3}^2 \tag{5}$$

and $R_{3(1,2)}^2$ and $\rho_{3(1,2)}^2$ denote respectively the sample and population multiple R^2 when the 3rd variable is regressed on the first two variables. Note that

$\dfrac{S_3(R_{23}-R_{12}R_{13})}{S_2\left(1-R_{12}^2\right)}$ is an unbiased estimate of the slope coefficient for the regression of the 3rd variable on the 2nd variable whilst the first variable is held fixed.

It is suggested that, if Σ is known, the statistics given in (1), (2), (3), (4) and (5) should all be used to provide protection against changes in the process covariance matrix Σ. It is advocated using all these components instead of only S_1^2, $S_{2\bullet1}^2$ and $S_{3\bullet1,2}^2$ because $\left(R_{12}S_2/S_1,\ \ R_{13}S_3/S_1\right)^{\mathrm{T}}$ and $\dfrac{S_3(R_{23}-R_{12}R_{13})}{S_2\left(1-R_{12}^2\right)}$ may reveal some changes in Σ that may not be reflected by the former statistics. For instance, if the 3rd quality characteristic is independent of others i.e $\rho_{3(12)}^2=\rho_{13}^2+\dfrac{(\rho_{12}\rho_{13}-\rho_{23})^2}{1-\rho_{12}^2}=0$, and $(\sigma_2,\ \rho_{12})$ shifts to $(\sigma_{2new},\ \rho_{12new})$ such that $\dfrac{\sigma_{2new}^2(1-\rho_{12new}^2)}{\sigma_2^2(1-\rho_{12}^2)}=1$, then this change is unlikely to be detected when only S_1^2, $S_{2\bullet1}^2$ and $S_{3\bullet1,2}^2$ are used because their respective distributions are not distorted under these circumstances. However, this change induces a shift in the slope coefficient for the regression of the 2nd quality characteristic on the first. Therefore, it is possible to 'pick up' such a change if the vector of population regression coefficients $\left(\rho_{12}\sigma_2/\sigma_1,\ \rho_{13}\sigma_3/\sigma_1\right)^{\mathrm{T}}$ is also monitored based on the corresponding vector of sample regression coefficients $\left(R_{12}S_2/S_1,\ \ R_{13}S_3/S_1\right)^{\mathrm{T}}$ which is known to be bivariate normal for fixed S_1^2 under the in-control and normality assumption (see (2)). If the traditional Hotelling χ^2 chart based on these coefficient vectors is used, it is readily seen that its statistical performance depends on the noncentrality parameter

$$\lambda = \frac{(n-1)S_1^2\left(\pm\sqrt{\sigma_{2new}^2 - \sigma_2^2(1-\rho_{12}^2)} - \rho_{12}\sigma_2\right)^2}{\sigma_1^2\sigma_2^2(1-\rho_{12}^2)}$$

where the +ve sign is used when $\rho_{12new} > 0$ and the −ve sign otherwise. Thus, whilst the use of the control statistics S_1^2, $S_{2\bullet 1}^2$ and $S_{3\bullet 1,2}^2$ are unlikely to register the change, it is clear that the probability of detection by the Hotelling χ^2 chart may increase depending on the value of S_1 for the current subgroup, the sample size, n, as well as the dispersion parameters. The same is true if the aggregate-type control statistic as given in the next section is used.

As an alternative, the following method of decomposition may be employed. Let $\mathbf{S}_j$ and Σ_j be the upper left-hand square submatrices of $\mathbf{S}$ and Σ respectively, of order j. Also, let $\mathbf{S}_{[j]}$ and $\Sigma_{[j]}$ denote respectively the sample and population covariance matrices of the jth, 1st, 2nd, ... and (j-1)th variable in the order indicated. In addition, let $\underset{\sim}{\mathbf{S}}_j$ and $\underset{\sim}{\sigma}_j$ be such that

$$\mathbf{S}_{[j]} = \begin{pmatrix} S_{jj} & \underset{\sim}{\mathbf{S}}_j^{\mathrm{T}} \\ \underset{\sim}{\mathbf{S}}_j & \mathbf{S}_{j-1} \end{pmatrix} \quad \text{and} \quad \Sigma_{[j]} = \begin{pmatrix} \sigma_{jj} & \underset{\sim}{\sigma}_j^{\mathrm{T}} \\ \underset{\sim}{\sigma}_j & \Sigma_{j-1} \end{pmatrix}, \qquad j = 2,\ldots,p.$$

where $S_{jj} = S_j^2$ and $\sigma_{jj} = \sigma_j^2$. Repeatedly applying theorem 6.4.1 of Giri (1977) to $\mathbf{S}_{[j]}$, starting with $j = p$ and decreasing in steps of 1, results in the following $2p-1$ (conditionally) independent statistics :

$$S_{j\bullet 1,\ldots,j-1}^2 \sim \frac{\sigma_{j\bullet 1,\ldots,j-1}^2}{(n-1)}\chi_{n-j}^2, \qquad j = 1,\ldots,p.$$

$$\mathbf{S}_{j-1}^{-1}\underset{\sim}{\mathbf{S}}_j \sim N_{j-1}\left(\Sigma_{j-1}^{-1}\underset{\sim}{\sigma}_j,\ \frac{\sigma_{j\bullet 1,\ldots,j-1}^2}{(n-1)}\mathbf{S}_{j-1}^{-1}\right), \qquad j = 2,\ldots,p.$$

where $S^2_{j\bullet 1,\ldots,j-1} = S_{jj} - \underset{\sim}{\mathbf{S}}_j^{\mathrm{T}} \mathbf{S}_{j-1}^{-1} \underset{\sim}{\mathbf{S}}_j$ and $\sigma^2_{j\bullet 1,\ldots,j-1} = \sigma_{jj} - \underset{\sim}{\sigma}_j^{\mathrm{T}} \Sigma_{j-1}^{-1} \underset{\sim}{\sigma}_j$ are respectively the conditional sample and population variances of the jth variable given the first $j-1$ variables. Note that $\mathbf{S}_{j-1}^{-1} \underset{\sim}{\mathbf{S}}_j$ is the $(j-1)$ dimensional vector estimating the regression coefficients of the jth variable regressed on the first $(j-1)$ variables (see Mason et al.(1995)). Note also that $S^2_{1\bullet 0}$ and $\sigma^2_{1\bullet 0}$ are taken to be S_1^2 and σ_1^2 respectively.

The hypothesis $\mathrm{H}_0 : \Sigma = \Sigma_0$ may be tested based on these statistics for each subgroup in a sequential or step-down manner. At the jth step, the component hypothesis $\sigma^2_{j\bullet 1,\ldots,j-1} = (\sigma^2_{j\bullet 1,\ldots,j-1})_0$ is tested at α_j significance level by means of a chi-square test based on

$$\frac{S^2_{j\bullet 1,\cdots,j-1}}{(\sigma^2_{j\bullet 1,\ldots,j-1})_0} \tag{6}$$

If there is failure to reject this sub-hypothesis, then $\underset{\sim}{\sigma}_j = \left(\underset{\sim}{\sigma}_j\right)_0$ (or $\Sigma_{j-1}^{-1} \underset{\sim}{\sigma}_j = \left(\Sigma_{j-1}^{-1}\right)_0 \left(\underset{\sim}{\sigma}_j\right)_0$) is tested at significance level δ_j on the assumption that $\Sigma_{j-1} = \left(\Sigma_{j-1}\right)_0$. The test statistic,

$$\left(\mathbf{S}_{j-1}^{-1} \underset{\sim}{\mathbf{S}}_j - \left(\Sigma_{j-1}^{-1}\right)_0 \left(\underset{\sim}{\sigma}_j\right)_0\right)^{\mathrm{T}} \frac{(n-1)\mathbf{S}_{j-1}}{\left(\sigma^2_{j\bullet 1,\ldots,j-1}\right)_0} \left(\mathbf{S}_{j-1}^{-1} \underset{\sim}{\mathbf{S}}_j - \left(\Sigma_{j-1}^{-1}\right)_0 \left(\underset{\sim}{\sigma}_j\right)_0\right) \tag{7}$$

is a χ^2_{j-1} variable if the component hypothesis is true. If there is failure to reject this component hypothesis, then the (j+1)th step is taken. The hypothesis $\mathrm{H}_0 : \Sigma = \Sigma_0$ is accepted provided there is failure to reject all the $2p-1$ component hypotheses. The overall significance level of this test for each subgroup is then given by

$$1-\prod_{j=1}^{p}(1-\alpha_j)\prod_{j=2}^{p}(1-\delta_j).$$

Anderson (1984, p.417-418) has presented such an approach for testing the equality of covariance matrices as an alternative to the standard maximum likelihood ratio procedure, with the unknown parameters replaced by appropriate estimates based on previous subgroups and other suitable adjustments made. The resulting statistics for all successive subgroups follow Snedecor-*F* distributions and were shown by this author to be stochastically independent (Anderson (1984), theorem 10.4.2, p.414). Although this method is not proposed in the context of SPC, it can be used for monitoring the stability of the process covariance matrix for which the true in-control value is *unknown* and cannot be reliably estimated. Following the conventional approach, however, a *single* control chart based on all these statistics is considered instead of using them separately. This control technique, which is particularly useful for short production runs and low volume manufacturing, is discussed in detail in the next section.

3. Monitoring the Dispersion of Multivariate Processes

The techniques now presented involve use of the *probability integral transformation* in order to produce sequences of independent chi-square variables (see Quesenberry (1991)). The suggested approach permits the monitoring of various components resulting from the decomposition of the covariance matrix on a single chart.

For uniformity of notation and ease of presentation, define $\mathbf{S}_{*k}$ (Σ_{*k}) and $\underset{\sim}{\mathbf{S}}_{v,u}$ $(\underset{\sim}{\sigma}_{v,u})$ respectively as the sample (population) covariance matrix of the *last* k variables and the vector of sample (population) covariances between the vth variable and each of the *first* u variables. Accordingly, the sample and population covariance matrices are expressible as

$$\mathbf{S}=\left(\begin{array}{c|c}\mathbf{S}_{j-1} & \underset{\sim}{\mathbf{S}}_{j,j-1},\ldots,\underset{\sim}{\mathbf{S}}_{p,j-1}\\ \hline \underset{\sim}{\mathbf{S}}^{\mathrm{T}}_{j,j-1} & \\ \vdots & \mathbf{S}_{*p-j+1}\\ \underset{\sim}{\mathbf{S}}^{\mathrm{T}}_{p,j-1} & \end{array}\right) \quad \text{and} \quad \Sigma=\left(\begin{array}{c|c}\Sigma_{j-1} & \underset{\sim}{\sigma}_{j,j-1},\ldots,\underset{\sim}{\sigma}_{p,j-1}\\ \hline \underset{\sim}{\sigma}^{\mathrm{T}}_{j,j-1} & \\ \vdots & \Sigma_{*p-j+1}\\ \underset{\sim}{\sigma}^{\mathrm{T}}_{p,j-1} & \end{array}\right)$$

where $\mathbf{S}_j$ (Σ_j) denotes the sample (population) covariance matrix of the *first* j variables and $\underset{\sim}{\mathbf{S}}_{j,j-1}=\underset{\sim}{\mathbf{S}}_j$ $(\underset{\sim}{\sigma}_{j,j-1}=\underset{\sim}{\sigma}_j)$ as defined in the preceding section. The conditional sample variance of the jth variable given the first $j-1$ variables, is then given by

$$S^2_{j\bullet 1,\ldots,j-1}=S^2_j-\underset{\sim}{\mathbf{S}}^{\mathrm{T}}_{j,j-1}\mathbf{S}^{-1}_{j-1}\underset{\sim}{\mathbf{S}}_{j,j-1} \tag{8}$$

Similarly, the corresponding population parameter is

$$\sigma^2_{j\bullet 1,\ldots,j-1}=\sigma^2_j-\underset{\sim}{\sigma}^{\mathrm{T}}_{j,j-1}\Sigma^{-1}_{j-1}\underset{\sim}{\sigma}_{j,j-1} \tag{9}$$

In terms of variances and multiple correlation coefficients, these are expressible as $S^2_{j\bullet 1,\ldots,j-1}=S^2_j(1-R^2_{j(1,\ldots,j-1)})$ and $\sigma^2_{j\bullet 1,\ldots,j-1}=\sigma^2_j(1-\rho^2_{j(1,\ldots,j-1)})$. The conditional sample and population covariance matrices of the last $p-j+1$ variables given the remaining $j-1$ variables are respectively

$$\mathbf{S}_{j,\dots,p\bullet 1,\dots,j-1} = \mathbf{S}_{*p-j+1} - \left(\underset{\sim}{\mathbf{s}}_{j,j-1} \;\cdots\; \underset{\sim}{\mathbf{s}}_{p,j-1}\right)^{\mathrm{T}} \mathbf{S}_{j-1}^{-1} \left(\underset{\sim}{\mathbf{s}}_{j,j-1} \;\cdots\; \underset{\sim}{\mathbf{s}}_{p,j-1}\right) \tag{10}$$

and

$$\Sigma_{j,\dots,p\bullet 1,\dots,j-1} = \Sigma_{*p-j+1} - \left(\underset{\sim}{\sigma}_{j,j-1} \;\cdots\; \underset{\sim}{\sigma}_{p,j-1}\right)^{\mathrm{T}} \Sigma_{j-1}^{-1} \left(\underset{\sim}{\sigma}_{j,j-1} \;\cdots\; \underset{\sim}{\sigma}_{p,j-1}\right) \tag{11}$$

Apart from these, let $\underset{\sim}{d}_j$ and $\underset{\sim}{\theta}_j$ (j = 2,..., p) denote respectively the vectors of sample and population regression coefficients when each of the last p-j+1 variables is regressed on the (j-1)th variable whilst the remaining j-2 variables are held fixed. Then, these are given by the following expressions :-

$$\underset{\sim}{d}_j = \frac{\left\{\left(s_{j-1,j} \;\cdots\; s_{j-1,p}\right) - \underset{\sim}{\mathbf{s}}_{j-1,j-2}^{\mathrm{T}} \mathbf{S}_{j-2}^{-1} \left(\underset{\sim}{\mathbf{s}}_{j,j-2} \;\cdots\; \underset{\sim}{\mathbf{s}}_{p,j-2}\right)\right\}^{\mathrm{T}}}{s_{j-1,j-1} - \underset{\sim}{\mathbf{s}}_{j-1,j-2}^{\mathrm{T}} \mathbf{S}_{j-2}^{-1} \underset{\sim}{\mathbf{s}}_{j-1,j-2}} \tag{12}$$

and

$$\underset{\sim}{\theta}_j = \frac{\left\{\left(\sigma_{j-1,j} \;\cdots\; \sigma_{j-1,p}\right) - \underset{\sim}{\sigma}_{j-1,j-2}^{\mathrm{T}} \Sigma_{j-2}^{-1} \left(\underset{\sim}{\sigma}_{j,j-2} \;\cdots\; \underset{\sim}{\sigma}_{p,j-2}\right)\right\}^{\mathrm{T}}}{\sigma_{j-1,j-1} - \underset{\sim}{\sigma}_{j-1,j-2}^{\mathrm{T}} \Sigma_{j-2}^{-1} \underset{\sim}{\sigma}_{j-1,j-2}} \tag{13}$$

Note that $\underset{\sim}{d}_2$ and $\underset{\sim}{\theta}_2$ should be interpreted as the vectors of unconditional sample and population regression coefficients when each of the last p-1 variables is regressed on the 1st variable and these are given by

$$\underset{\sim}{d}_2 = \frac{\left(s_{12} \;\cdots\; s_{1p}\right)^{\mathrm{T}}}{s_{11}} \quad \text{and} \quad \underset{\sim}{\theta}_2 = \frac{\left(\sigma_{12} \;\cdots\; \sigma_{1p}\right)^{\mathrm{T}}}{\sigma_{11}}$$

respectively.

In addition to the above, the following notation will be used :-

$\Phi^{-1}(\bullet)$: inverse of the standard normal distribution function

$\chi_{\nu}^{2}(\bullet)$: distribution function of a chi-square variable with ν degrees of freedom

$F_{\nu1;\nu2}(\bullet)$: distribution function of an F variable with $\nu1$ numerator and $\nu2$ denominator degrees of freedom

In the following subsections, control statistics for monitoring the stability of the process covariance matrix are presented for the case where either Σ is known or unknown. In order to specify the chronological order of the subgroups, the sample statistics are indexed with an additional subscript enclosed within a bracket.

3.1. Case (i) : Σ known

In practice, the process parameters, in particular the true value of the process covariance matrix Σ, is never known exactly. Instead, it is estimated based on a presumably large enough set of relevant data that have been collected during the period in which the production process is assumed to be stable or in control. It will be assumed for current purposes that Σ is known precisely prior to production. In this case, the appropriate control statistic is

$$T_k = \sum_{j=1}^{2p-1} Z_{j(k)}^2 \qquad k = 1,2,\ldots \tag{14}$$

where

$$Z_{1(k)} = \Phi^{-1}\left\{\chi_{n_k-1}^2\left[\frac{(n_k-1)S_{1(k)}^2}{\sigma_1^2}\right]\right\}$$

$$Z_{j(k)} = \Phi^{-1}\left\{\chi_{n_k-j}^2\left[\frac{(n_k-1)S_{j\bullet1,\ldots,j-1(k)}^2}{\sigma_{j\bullet1,\ldots,j-1}^2}\right]\right\} \qquad j = 2,\ldots,p.$$

$$Z_{p+1(k)} = \Phi^{-1}\left\{\chi^2_{p-1}\left[(n_k - 1)S^2_{1(k)}\left(\underset{\sim}{d}_{2(k)} - \underset{\sim}{\theta}_2\right)^{\mathrm{T}} \Sigma^{-1}_{2,\ldots,p\bullet 1}\left(\underset{\sim}{d}_{2(k)} - \underset{\sim}{\theta}_2\right)\right]\right\}$$

$$Z_{p+j-1(k)} = \Phi^{-1}\left\{\chi^2_{p-j+1}\left[(n_k - 1)S^2_{j-1\bullet 1,\ldots,j-2(k)}\left(\underset{\sim}{d}_{j(k)} - \underset{\sim}{\theta}_j\right)^{\mathrm{T}} \Sigma^{-1}_{j,\ldots,p\bullet 1,\ldots,j-1}\left(\underset{\sim}{d}_{j(k)} - \underset{\sim}{\theta}_j\right)\right]\right\} \qquad j = 3,\ldots,p.$$

It is readily seen from the foregoing discussion, that under the in-control and normality assumption, $\{Z_{j(k)}\}$, $j = 1,\ldots,2p-1$ are sequences of independently and identically distributed (*i.i.d*) standard normal variables, whence T_k's are independent χ^2_{2p-1} variables. Although control charts may be constructed based on the transformed statistics $Z_{j(k)}$'s, this is not considered a viable option due to the proliferation of charts that results even when p is fairly small. Besides, it is found that combining the $Z_{j(k)}$'s in the proposed manner results in better control performance for certain shifts in Σ.

Note that since the arguments of the normalizing transformation are independent chi-square variables, a single aggregate-type control statistic may be obtained by summing them. Similarly, the sum may be taken over the transformed statistics $Z_{j(k)}$'s giving a sequence of independent $N(0,\ 2p-1)$ variables. In either case, however, certain deviations of the process covariance matrix from the specified Σ are likely to be missed by the resulting techniques. In particular, if Σ shifts in such a way that the values of some $Z_{j(k)}$'s tend to be larger whilst others tend to be smaller than that attributable to common causes, then this type of change is unlikely to be detected by the resulting charts. To provide protection against such changes, it is suggested that the $Z_{j(k)}$'s be squared before summation as in formula (14), and only an upper control limit is then necessary. It should be noted that a similar technique can also be developed based on the alternative partitioning method outlined in the foregoing section. This is, however, not considered further because it is found that the proposed technique always performs better.

3.2. Case (ii) : Σ unknown

In the absence of prior information about the process parameters, a natural solution is to estimate the various components resulting from the decomposition of Σ sequentially from the data stream of the current production. The resulting estimates, together with the corresponding observations from the next sample are then used to test whether or not Σ remains constant.

Before proceeding, define the quantities :-

$$N_{j,k} = \sum_{i=1}^{k} (n_i - j)$$

$$S^2_{j,k\,(pooled)} = \frac{1}{N_{j,k}} \sum_{i=1}^{k} (n_i - 1) S^2_{j\bullet 1,\ldots,j-1(i)} \qquad j = 1,\ldots,p.$$

with $S^2_{1\bullet 0(i)} = S^2_{1(i)}$,

$$\underset{\sim}{\mathbf{U}}_{j,k} = \frac{1}{k^2} \sum_{i=1}^{k} (n_i - 1)^{-1} \mathbf{S}^{-1}_{j(i)} \qquad j = 2,\ldots,p.$$

$$\underset{\sim}{\mathbf{V}}_{j,k} = \frac{1}{k} \sum_{i=1}^{k} \mathbf{S}^{-1}_{j-1(i)} \underset{\sim}{\mathbf{S}}_{j,j-1(i)} \qquad j = 2,\ldots,p.$$

These values can be updated through the following recursive equations :-

$$S^2_{j,k+1(pooled)} = \frac{1}{N_{j,k+1}} \left[N_{j,k} S^2_{j,k\,(pooled)} + (n_{k+1} - 1) S^2_{j\bullet 1,\ldots,j-1(k+1)} \right]$$

$$\underset{\sim}{\mathbf{U}}_{j,k+1} = \frac{1}{(k+1)^2}\left[k^2\,\underset{\sim}{\mathbf{U}}_{j,k} + (n_{k+1}-1)^{-1}\mathbf{S}^{-1}_{j(k+1)}\right]$$

$$\underset{\sim}{\mathbf{V}}_{j,k+1} = \frac{1}{k+1}\left[k\,\underset{\sim}{\mathbf{V}}_{j,k} + \mathbf{S}^{-1}_{j-1(k+1)}\,\underset{\sim}{\mathbf{S}}_{j,j-1(k+1)}\right]$$

The appropriate control statistic for this case is then given by

$$T_k = \sum_{j=1}^{2p-1} Z^2_{j(k)} \qquad k = 2,\ldots \qquad (15)$$

where

$$Z_{j(k)} = \Phi^{-1}\left[F_{n_k - j;N_{j,k-1}}\left[\frac{S^2_{j\bullet 1,\ldots,j-1(k)}}{S^2_{j,k-1\ (pooled)}}\right]\right] \qquad j = 1,\ldots,p.$$

and

$$Z_{p+j-1(k)} = \Phi^{-1}\left[F_{j-1;N_{j,k}}\left[\frac{\left(\mathbf{S}^{-1}_{j-1(k)}\underset{\sim}{\mathbf{S}}_{j,j-1(k)} - \underset{\sim}{\mathbf{V}}_{j,k-1}\right)^{\mathrm{T}}\left((n_k-1)^{-1}\mathbf{S}^{-1}_{j-1(k)} + \underset{\sim}{\mathbf{U}}_{j-1,k-1}\right)^{-1}\left(\mathbf{S}^{-1}_{j-1(k)}\underset{\sim}{\mathbf{S}}_{j,j-1(k)} - \underset{\sim}{\mathbf{V}}_{j,k-1}\right)}{(j-1)S^2_{j,k\ (pooled)}}\right]\right]$$

$$j = 2,\ldots,p.$$

Note that when $k = 2$, the argument of $Z_{p+j(k)}$ is

$$\frac{\left(\mathbf{S}^{-1}_{j-1(2)}\underset{\sim}{\mathbf{S}}_{j,j-1(2)} - \mathbf{S}^{-1}_{j-1(1)}\underset{\sim}{\mathbf{S}}_{j,j-1(1)}\right)^{\mathrm{T}}\left((n_1-1)^{-1}\mathbf{S}^{-1}_{j-1(1)} + (n_2-1)^{-1}\mathbf{S}^{-1}_{j-1(2)}\right)^{-1}\left(\mathbf{S}^{-1}_{j-1(2)}\underset{\sim}{\mathbf{S}}_{j,j-1(2)} - \mathbf{S}^{-1}_{j-1(1)}\underset{\sim}{\mathbf{S}}_{j,j-1(1)}\right)}{(j-1)S^2_{j,2\ (pooled)}}$$

where

$$S^2_{j,2(pooled)} = \frac{(n_1-1)S^2_{j\bullet 1,\ldots,j-1(1)} + (n_2-1)S^2_{j\bullet 1,\ldots,j-1(2)}}{(n_1+n_2-2j)}.$$

This is different from that given by Anderson (1984, p.418, expression (21)) which apparently contains a typographical error.

When the process covariance matrix is constant, $\{Z_{j(k)}\}$, $j = 1,\ldots,2p-1$ are independent sequences of $i.i.d$ $N(0,1)$ variables (see Theorem 10.4.2, p.414 of Anderson(1984)). Thus, the T_k's here are again distributed as χ^2_{2p-1} variables. Note that although the arguments of the $Z_{j(k)}$'s have different degrees of freedom, the control limit for the resulting technique remains constant for successive subgroups. Note also that, using this technique, process monitoring can begin with the second subgroup without having to wait until considerable process performance data have been accumulated for computation of the unknown Σ.

4. Concluding Remarks

An underlying methodology and appropriate control statistics for dispersion control for the cases where Σ is known and unknown initially have been provided. In a sequel to this paper, comparisons are made between the proposed techniques and various competing procedures and the proposed techniques are demonstrated to be superior.

References

ALT, F. B. and BEDEWI, G. E. (1986). 'SPC of Dispersion for Multivariate Data'. *ASQC Quality Congress Transaction - Anaheim.* American Society for Quality Control, pp. 248-254.

CHAN, L. K. and LI, GUO-YING. (1994). 'A Multivariate Control Chart for Detecting Linear Trends'. *Communications in Statistics - Simulation and Computation* 23, pp.997-1012.

CROSIER, R. B. (1988). 'Multivariate Generalization of Cumulative Sum Quality-Control Schemes'. *Technometrics* 30, pp. 291-303.

DOGANAKSOY, N.; FALTIN, F. W. and TUCKER, W. T. (1991). 'Identification of Out Of Control Quality Characteristics in a Multivariate Manufacturing Environment'. *Communications in Statistics - Theory and Methods* 20, pp. 2775-2790.

GIRI, N. C. (1977). *Multivariate Statistical Inference*, Academic Press, New York, NY.

GNANADESIKAN, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations*, John Wiley & Sons, New York, NY.

HAWKINS, D. M. (1991). 'Multivariate Quality Control Based on Regression-Adjusted Variables'. *Technometrics* 33, pp. 61-75.

HAWKINS, D. M. (1993). 'Regression Adjustment for Variables in Multivariate Quality Control'. *Journal of Quality Technology* 25, pp. 170-182.

HAYTER, A. J. and TSUI, K. (1994). 'Identification and Quantification in Multivariate Quality Control Problems'. *Journal of Quality Technology* 26, pp. 197-208.

HEALY, J. D. (1987). 'A Note on Multivariate CUSUM Procedures'. *Technometrics* 29, pp.409-412.

LOWRY, C. N.; WOODALL W. H.; CHAMP, C. W. and RIGDON, S. E. (1992). 'A Multivariate Exponentially Weighted Moving Average Control Chart'. *Technometrics* 34, pp. 46-53.

MASON, R. L.; TRACY, N. D. and YOUNG, J. C. (1995). 'Decomposition of T^2 for Multivariate Control Chart Interpretation'. *Journal of Quality Technology* 27, pp.99-108.

MURPHY, B. J. (1987). 'Selecting Out of Control Variables with the T^2 Multivariate Quality Control Procedure'. *The Statistician* 36, pp.571-583.

PIGNATIELLO, J. J. Jr and RUNGER, G. C. (1990). 'Comparisons of Multivariate CUSUM Charts'. *Journal of Quality Technology* 22, pp.173-186.

QUESENBERRY, C. P. (1991). 'SPC Q Charts for Start-Up Processes and Short or Long Runs'. *Journal of Quality Technology* 23, pp.213-224.

SPARKS, R. S. (1992). 'Quality Control with Multivariate Data'. *Australian journal of Statistics* 34, pp.375-390.

TANG, P. F. (1995). 'Mean Control for Multivariate Processes with Specific Reference to Short Runs'. *International Conference on Statistical Methods and Statistical Computing for Quality and Productivity Improvement - Seoul.* The International Statistical Institute (ISI), pp.579-585.

TRACY, N. D.; YOUNG, J. C. and MASON, R. L. (1992). 'Multivariate Control Charts for Individual Observations'. *Journal of Quality Technology* 24, pp.88-95.

WOODALL, W. H. and NCUBE, M. M. (1985). 'Multivariate CUSUM Quality-Control Procedures'. *Technometrics* 27, pp.285-292.