Chapter 5

Identification of the Out-of-Control Variable When a Multivariate Control Chart Signals

5.1 Introduction

Multivariate control charts are a powerful tool in Statistical Process Control for identifying an out-of-control process. Woodall and Montgomery (1999) emphasized the need for much more research in this area since most of the processes involve a large number of variables that are correlated. As Jackson (1991) notes, any multivariate quality control procedure should fulfill four conditions 1) Single answer to the question "Is the process in-control?" 2) An overall probability for the event "Procedure diagnoses an out-of-control state erroneously" must be specified 3) The relationship among the variables must be taken into account and 4) Procedures should be available to answer the question "If the process is out-of-control, what is the problem?". The last question has proven to be an interesting subject for many researchers in the last years. Woodall and Montgomery (1999) state that although there is difficulty in interpreting the signals from multivariate control charts more work is needed on data reduction methods and graphical techniques.

In this chapter we present the available solutions for the problem of identification and additionally we propose a new method based on principal components analysis (PCA), for detecting the out-of-control variable, or variables, when a multivariate control chart for individual observations signals. Section 5.2 describes the use of univariate control charts for solving the above stated problem, whereas Section 5.3 gives the use of an elliptical control region. In Section 5.4 a $T^{\mathbf{2}}$ decomposition is presented. Section 5.5 summarizes the methods based on principal components analysis. A presentation of the new method, is given in the Section 5.6 with some interesting points and discussion on the performance and application of the new method. Moreover, a comparative study evaluates the performance of the proposed method in relation to the existing methods that use PCA. Finally, graphical techniques that attempt to solve the problem under investigation are presented in Section 5.7.

5.2 The Use of Univariate Control Charts

5.2.1 Univariate Control Charts with Standard Control Limits

The use of p univariate control charts, gives a first evidence for which of the p variables are responsible for an out-of-control signal. However, there are some problems in using p univariate control charts in place of X^2-Chart. These problems are that, the overall probability of the mean plotting outside the control limits if we are in-control is not controlled and the correlations among the variables are ignored. The problem of ignoring the correlations among the variables cannot be solved. The problem of controlling the overall probability of the mean plotting outside the control limits if we are in-control can be solved by using p univariate control charts with Bonferroni limits.

5.2.2 Using Univariate Control Charts with Bonferroni Control Limits

The use of Bonferroni control limits, was proposed by Alt (1985). Bonferroni control limits can be used to investigate which of the p variables are responsible for an out-of-control signal. Using the Bonferroni method the following control limits are established

$$\begin{aligned} UCL &= \mu_i + Z_{1-\alpha/2p}\frac{\sigma_i}{\sqrt{n}} \\ LCL &= \mu_i - Z_{1-\alpha/2p}\frac{\sigma_i}{\sqrt{n}}. \end{aligned}$$

Thus, p individual control charts can be constructed, each with probability of the mean plotting outside the control limits if we are in-control, equal to α/p and not α. However, as Alt (1985) states, this does not imply that p univariate control charts should be used in place of X^2-Chart.

5.2.3 Hayter and Tsui's Interpretation Method

Hayter and Tsui (1994) extended the idea of Bonferroni type control limits by giving a procedure for exact simultaneous control intervals for each of the variable means. The control procedure operates as follows. For a known variance-covariance matrix $\mathbf{\Sigma}$ and a chosen probability of the mean plotting outside the control limits if we are in-control α, the experimenter first evaluates the critical point $C_{R,\alpha}$ where $\mathbf{R}$ is the correlation matrix obtained from $\mathbf{\Sigma}$. Then, following any observation $\mathbf{x}^t = (X_1, X_2, ..., X_p)$, the experimenter constructs confidence intervals

$$(X_i - \sigma_i C_{R,\alpha}, X_i + \sigma_i C_{R,\alpha})$$

for each of the p variables. This procedure ensures that an overall probability of the mean plotting outside the control limits if we are in-control α is achieved. The process is considered to be in-control as long as each of these confidence intervals contains the respective standard value μ_{0i}. The process is considered to be out-of-control if any of these confidence intervals does not contain the respective standard value μ_{0i}. The variable or variables whose confidence intervals do not contain μ_{0i}, are identified as those responsible for the out-of-control signal.

This procedure signals when

$$M = max \left| X_i - \mu_{0i} \right| / \sigma_i > C_{R,\alpha}.$$

Hayter and Tsui (1994) give guidance and various tables for choosing the critical point $C_{R,\alpha}$.

5.3 Using an Elliptical Control Region

The second method uses an elliptical control region. This method is discussed by Alt (1985) and Jackson (1991) and can be applied only in the special case of two quality characteristics.

The simplest case in multivariate statistics is when the vector $\mathbf{x} = \binom{x_1}{x_2}$ has a bivariate normal distribution, where X_i is distributed normally with mean μ_i, standard deviation σ_i, i= 1, 2 and ρ is the correlation coefficient between the two variables. In this case an elliptical control region, can be constructed. This elliptical region is centered at $\boldsymbol{\mu}_0^t = (\mu_1, \mu_2)$ and can be used in place of the X^2-chart. All points lying on the ellipse would have the same value of X^2. While, X^2-Chart gives a signal every time the process is out-of-control, the elliptical region is useful in indicating which variable led to the out-of-control signal.

Therefore, a $100\,(1-\alpha)\,\%$ elliptical control region can be constructed by applying the following equation as given by Jackson (1991)

$$Q = \left\{ \frac{1}{1-\rho^2} \times \left[\left(\frac{X_1 - \mu_1}{\sigma_1} \right)^2 + \left(\frac{X_2 - \mu_2}{\sigma_2} \right)^2 - \frac{2\rho\,(X_1 - \mu_1)\,(X_2 - \mu_2)}{\sigma_1 \sigma_2} \right] \right\} = X^2_{2,1-\alpha}.$$

A unique ellipse is defined for given values of $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ and α. Points on the perimeter of the ellipse may be determined by setting X_1 equal to some constant and solving the resulting quadratic equation for X_2.

Mader et al. (1996) presented the use of the elliptical control region for power supply calibration.

5.4 Using T^2 Decomposition

The third method is the use of T^2 decomposition, which is proposed by Mason, Tracy and Young (1995,1997). The main idea of this method (MYT) is to decompose the T^2 statistic into independent parts, each of which reflects the contribution of an individual variable. This method is developed for the case of individual observations, but according to the authors it can be applied also with a few modification for the case of rational subgroups.

In this section we also present, the methodologies of Roy (1958), Murphy (1987), Doganaksoy et al. (1991), Hawkins (1991, 1993), Timm (1996) and Runger and Montgomery (1996), which are included in the MYT partitioning of T^2.

5.4.1 Mason, Tracy and Young's T^2 Decomposition

Mason et al. (1995) presented the following interpretation method of an out-of-control signal. The T^2 statistic can be broken down or decomposed into p orthogonal components. One form of the MYT decomposition is given by

$$T^2 = T_1^2 + T_{2\cdot 1}^2 + T_{3\cdot 1,2}^2 + ... + T_{p\cdot 1,2,...,p-1}^2 = T_1^2 + \sum_{j=1}^{p-1} T_{j\cdot 1,2,...,j-1}^2.$$

The first term of this decomposition, T_1^2, is an unconditional Hotelling's T^2 for the first variable of the observation vector $\mathbf{x}$,

$$T_1^2 = \left(\frac{X_1 - \overline{X}_1}{s_1}\right)^2,$$

where $\overline{X}_1$ and s_1 is the mean and standard deviation of variable X_1, respectively.

The general form of the other terms, referred to as conditional terms, is given as

$$T^2_{j \cdot 1,2,...,j-1} = \frac{(X_j - \overline{X}_{j \cdot 1,2,...,j-1})^2}{s^2_{j \cdot 1,2,...,j-1}}, \text{ for } j = 1, 2, ..., p,$$

where

$$\overline{X}_{j \cdot 1,2,...,j-1} = \overline{X}_j + \mathbf{b}_j^t \left(\mathbf{X}_i^{(j-1)} - \overline{\mathbf{X}}^{(j-1)}\right)$$

and $\mathbf{X}_i^{(j-1)}$ is the $(j-1)th$ vector excluding the jth variable, $\overline{X}_j$ is the sample mean of the jth variable, $\mathbf{b}_j = [\mathbf{S}_{\mathbf{xx}}^{-1}\mathbf{S}_{X\mathbf{x}}]$ is a $(j-1)th$ dimensional vector estimating the regression coefficients of the jth variable regressed on the first $(j-1)$ variables,

$$s^2_{j \cdot 1,2,...,j-1} = s_X^2 - \mathbf{S}_{X\mathbf{x}}^t \mathbf{S}_{\mathbf{xx}}^{-1} \mathbf{S}_{X\mathbf{x}}$$

and

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{\mathbf{xx}} & \mathbf{S}_{X\mathbf{x}} \\ \mathbf{S}_{X\mathbf{x}}^t & s_X^2 \end{bmatrix}.$$

Consequently, the $T^2_{j \cdot 1,2,...,j-1}$ value is the square, of the jth variable adjusted by the estimates of the mean and standard deviation of the conditional distribution of X_j given $X_1, X_2, ..., X_{j-1}$ and its exact distribution is as follows

$$T^2_{j \cdot 1,2,...,j-1} \sim \frac{n+1}{n} F\left(1, n-1\right).$$

Thus, this statistic can be used to check whether the jth variable is conforming to the relationship with other variables as established by the historical data set, since the adjusted observation is more sensitive to changes in the covariance structure.

The ordering of the p components is not unique and the one given above represents only one of the possible $p!$ different ordering of these components. Each ordering generates the same overall T^2 value, but provides a distinct partitioning of T^2 into p orthogonal terms. If we exclude redundancies, there are $p \times 2^{p-1}$ distinct components among the $p \times p!$ possible terms that should be evaluated for potential contribution to signal.

Similarly, the p unconditional T^2 terms based on squaring a univariate t statistic can be computed and then be compared to the appropriate F distribution. Moreover, the distances $D_i = T^2 - T_i^2$ can be computed and also be compared to the F distribution.

5.4.2 Mason, Tracy and Young's Out-of-Control Variable Selection Algorithm

The following is a sequential computational scheme that has the potential of further reducing the computations to a reasonable number when the overall T^2 signals, as was proposed by Mason, Tracy and Young (1997).

Step 0: Conduct a T^2 test with a specified nominal confidence level α. If an out-of-control condition is signaled then continue with the step 1.

Step 1: Compute the individual T^2 statistic for every component of the $\mathbf{x}$ vector. Remove variables whose observations produce a significant T_i^2. The observations on these variables are out of individual control and it is not necessary to check how they relate to the other observed variables. With significant variables removed we have a reduced set of variables. Check the subvector of the remaining k variables of a signal. If you do not receive a signal we have located the source of the problem.

Step 2: Optional: Examine the correlation structure of the reduced set of variables. Remove any variable having a very weak correlation (0.3 or less) with all the other variables. The contribution of a variable that falls in this category is measured by the T_i^2 component.

Step 3: If a signal remains in the subvector of k variables not deleted, compute all $T_{i \cdot j}^2$ terms. Remove from the study all pairs of variables, (X_i, X_j), that have a significant $T_{i \cdot j}^2$

term. This indicates that something is wrong with the bivariate relationship. When this occurs it will further reduce the set of variables under consideration. Examine all removed variables for the cause of the signal. Compute the T^2 terms for the remaining subvector. If no signal is present, the source of the problem is with the bivariate relationships and those variables that were out of individual control.

Step 4: If the subvector of the remaining variables still contains a signal, compute all $T^2_{i \cdot j,k}$ terms. Remove any triple, (X_i, X_j, X_k), of variables that show significant results and check the remaining subvector for a signal.

Step 5: Continue computing the higher order terms in this fashion until there are no variables left in the reduced set. The worst case situation is that all unique terms will have to be computed.

Generally, the T^2 statistic associated with an observation from a multivariate problem is a function of the residuals taken from a set of linear regressions among the various process variables. These residuals are contained in the conditional T^2 terms of the orthogonal decomposition of the statistic. Mason and Young (1999) showed that a large residual in one of these fitted models can be due to incorrect model specification. By improving the model specification at the time that the historical data set is constructed, it may be possible to increase the sensitivity of the T^2 statistic to signal detection. Also, they showed that the resulting regression residual, can be used to improve the sensitivity of the T^2 statistic to small but consistent process shifts, using plots that are similar to cause-selecting charts.

The productivity of an industrial processing unit often depends on equipment that changes over time. These changes may not be stable, and, in many cases, may appear to occur in stages. Although changes in the process levels within each stage may appear insignificant, they can be substantial when monitored across the various stages. Standard process control procedures do not perform well in the presence of these step-like changes, especially when the observations from stage to stage are correlated. Mason et al. (1996) present an alternative control procedure for monitoring a process under these conditions, which is based on a double decomposition of Hotelling's T^2 statistic.

5.4.3 Doganaksoy, Faltin and Tucker's Out-of-Control Variable Selection Algorithm

The method that is presented in this subsection was proposed by Doganaksoy et al. (1991). The main idea of this method is the use of the univariate t ranking procedure and it is based on p unconditional T^2 terms. The statistic used is

$$t = \frac{\overline{X}_{i,new} - \overline{X}_{i,ref}}{\sqrt{s_{ii}\left(\frac{1}{n_{new}} + \frac{1}{n_{ref}}\right)}},$$

where $\overline{X}_{i,new}$ is the mean of the new sample, $\overline{X}_{i,ref}$ is the mean of the reference sample, s_{ii} is the estimate of the variance of the ith variable from the reference sample, n_{new} is the size of the new sample and n_{ref} is the size of the reference sample. The steps of this algorithm are the following

Step 1: Conduct a T^2 test with a specified nominal significance level α. If an out-of-control condition is signalled then continue with step 2.

Step 2: For each variable calculate the smallest significance level α^{ind} that would yield an individual confidence interval for $(\mu_{i,ref} - \mu_{i,new})$ that contains zero, where $\mu_{i,new}$ and $\mu_{i,ref}$ are the mean vectors of the populations from which the reference and new samples are drawn, respectively. For this α^{ind}, let t be the calculated value of the univariate t statistic for a variable and $T(t, d)$ be the cumulative distribution function of the t distribution with d degrees of freedom. Then $\alpha^{ind} = [2T(t; n_{ref} - 1) - 1]$.

Step 3: Plot α^{ind} for each variable on a 0-1 scale. Note that variables with larger α^{ind} values are the ones with relatively larger univariate t statistic values which require closer investigation as possible being among those components which have undergone a change. If indications of highly suspect variables are desired then continue.

Step 4: Compute the confidence interval α^{bonf} that yields the desired nominal confidence interval α^{sim} of the Bonferroni type simultaneous confidence intervals for $\mu_{i,ref} - \mu_{i,new}$. Here, $\alpha^{bonf} = [(p + \alpha^{sim} - 1)/p]$.

Step 5: Components having $\alpha^{ind} > \alpha^{bonf}$ are classified as being those which are most likely to have changed.

Furthermore, the authors give guidance for the choice of the α^{sim}.

5.4.4 Murphy's Out-of-Control Variable Forward Selection Algorithm (FSA)

This method is proposed by Murphy (1987). It is a subcase of the T^2 decomposition method, which was proposed by Mason et al. (1995) and stems from the field of discriminant analysis. It uses the overall T^2 value and compares it to a T_*^2 value based on a subset of variables.

The diagnostic approach is triggered by an out-of-control signal from a T^2-Chart. Murphy (1987) partitioned the sample mean vector $\overline{\mathbf{x}}$ into two subvectors $\overline{\mathbf{x}}_{*1}$ and $\overline{\mathbf{x}}_{*2}$, where the p_1 dimensional vector $\overline{\mathbf{x}}_{*1}$ is the subset of the $p = p_1 + p_2$ variables, which is suspect for the out-of-control signal. Then

$$T_p^2 = n\left(\overline{\mathbf{x}} - \boldsymbol{\mu}_0\right)^t \boldsymbol{\Sigma}_0^{-1}\left(\overline{\mathbf{x}} - \boldsymbol{\mu}_0\right)$$

is the full squared distance and

$$T_*^2 = n\left(\overline{\mathbf{x}}_{*1} - \boldsymbol{\mu}_{01}\right)^t \boldsymbol{\Sigma}_{01}^{-1}\left(\overline{\mathbf{x}}_{*1} - \boldsymbol{\mu}_{01}\right)$$

is the reduced distance corresponding to the subset of the p variables that is suspect for the out-of-control signal.

Finally, the following difference is calculated

$$D = T_p^2 - T_*^2.$$

It is proved that, under the null hypothesis, D follows a Chi-Square distribution with p_1 degrees of freedom and the subvector $\overline{\mathbf{x}}_{*1}$ follows a p_1-dimensional distribution with mean $\boldsymbol{\mu}_{01}$ and variance-covariance matrix $\boldsymbol{\Sigma}_{01}$. Murphy (1987) gave a forward selection algorithm.

The steps of this algorithm are the following; For each $\overline{\mathbf{x}} = \left[\overline{X}_1, ..., \overline{X}_p\right]$,

Step 1: Conduct a T^2 test with a specified nominal significance level α. If an out-of-control condition is signalled then continue with step 2.

Step 2: Calculate the p individual $T_1^2(X_i)$, equivalent to looking at p individual charts, and calculate the p differences $D_{p-1}(i) = \left[T_p^2 - T_1^2(X_i)\right]$. Choose the $min\left(D_{p-1}(i)\right) = D_{p-1}(r)$ and test this minimum difference.

If $D_{p-1}(r)$ is not significant then the rth variable only requires attention.

If $D_{p-1}(r)$ is significant then continue with step 3.

Step 3: Calculate the $p-1$ differences $D_{p-2}(r,j) = \left[T_p^2 - T_2^2(X_r, X_j)\right]$, $1 \leq r,\, j \leq p$ and $r \neq j$. Choose the $min\left(D_{p-2}(r,j)\right) = D_{p-2}(r,s)$ and test this minimum difference.

If $D_{p-2}(r,s)$ is not significant then the rth and the sth variables only require attention.

If $D_{p-2}(r,s)$ is significant then continue with step 4.

Step 4: Similar to step 3.

Step .: Similar to steps 3,4.

Step p: If the final $D_{p-(p-1)}$ test is significant then all p variables will require attention.

Murphy (1987) recommends that in tests of $D_{p-i} \backsim X_{p-i}^2$ a significance level in the interval $0.1 \leq \alpha \leq 0.2$ be used. From a practical point of view the applicability of this approach is severely limited when the number of quality or process variables is moderately large.

5.4.5 Chua and Montgomery's Out-of-Control Variable Selection Method

The method that is presented in this subsection was proposed by Chua and Montgomery (1992). This method, like the method of Murphy (1987), uses the overall T^2 value and compares it to a T_*^2 value based on a subset of variables. It is quite similar to Murphy's method, but uses a backward selection algorithm (BSA). The operations of the control scheme can be described as follows;

1) A multivariate observation is fed into the multivariate EWMA control chart.

2) If the observation is in-control, the control operation loops back to the beginning and checks the next observation. Otherwise, it checks the number of process variables.

3) If the number of process variables is greater than five, bypass the BSA and use the hyper-plane method directly. Otherwise, feed the observation into BSA.

4) The BSA will select the out-of-control variable set and feed it into the hyper-plane method.

5) The hyper-plane method will generate the necessary elliptical control charts for diagnosis.

6) Based on the diagnosis, corrective actions are then taken and the control operation loops back to the beginning and checks for the next observation.

Chua and Montgomery (1992) describe all the needed actions for the application of their method. The algorithm divides the p variables into two subsets of size p_1 and p_2, where $p_1 + p_2 = p$. Also, the sample covariance matrix is divided into two parts with p_1 and p_2 variables respectively. For all the p variables the T_p^2 is calculated. Similarly, for the p_1 variables the T_{p1}^2 is calculated. Then a difference value $D = T_p^2 - T_{p1}^2$ is calculated, which is distributed as a Chi-Square distribution with p_2 degrees of freedom.

The steps of this algorithm are the following; For each $\overline{\mathbf{x}} = \left[\overline{X}_1, ..., \overline{X}_p\right]$,

Step 1: Conduct a T^2 test with a specified nominal confidence level α. If an out-of-control condition is signalled then continue with step 2.

Step 2: Perform T_{p-1}^2 tests and calculate the differences D_{p-1}=T_p^2-T_{p-1}^2. If D_{p-1}

is not significant then the union of variable sets for all sets which are denoted as not significant is the new criterion. If $D_{p-1}(r)$ is significant then keep the criterion from step 1, denote the $min\,(D_{p-1}) = D(p)$ and continue to the next step.

Step 3: Perform T^2_{p-2} tests and calculate the differences D_{p-2}=T^2_p-T^2_{p-2}. If D_{p-2} is not significant then the union of variable sets for all sets which are denoted as not significant is the new criterion. If D_{p-2} is significant then keep the criterion from step 2, denote the $min\,(D_{p-2}) = D(p-1)$ and continue to the next step.

Step 4: Similar to step 3.

Step .: Similar to step 3, 4.

Step p: Perform T^2_1 tests and calculate the differences D_1=T^2_p-T^2_1. If D_1 is not significant then the union of variable sets for all sets which are denoted as not significant is the new criterion. If D_1 is significant then keep the criterion from step p-1, denote the $min\,(D_1) = D(1)$ and continue to the next step.

Step p+1: If all the tests from steps 2 to p are significant, exit. All p variables are out-of-control. If $D(i) \leq D(j)$, where $i < j$, $i = 1, 2, ..., p-1$, and $j = i+1, ..., p$ and the variable set for $D(i)$ is a subset of the variable set for $D(j)$, then the variable set for $D(i)$ is the out-of-control variable set. Otherwise, increment i by 1 and repeat the condition check.

The hyper-plane method is an extension of the elliptical control chart. Chua and Montgomery (1992) present the hyper-plane method graphically. The mathematical background of the hyper-ellipsoid method follows.

The equation for the hyper-ellipsoid control region is $(\mathbf{x}-\overline{\mathbf{x}})\mathbf{S}^{-1}(\mathbf{x}-\overline{\mathbf{x}}) = T^2_{1-\alpha}$. Obviously, if the left-hand side of the equation is less than $T^2_{1-\alpha}$, the observation will be inside the hyper-ellipsoid, and thus in-control. On the other-hand, if it is greater than $T^2_{1-\alpha}$, the observation will be outside the hyper-ellipsoid, and thus out-of-control.

Moreover, the vector equation of a hyper-plane is $\mathbf{d}\cdot(\mathbf{x}-\mathbf{b}) = 0$, where $\mathbf{d}$ is the direction vector of the hyperplane, $\mathbf{x}$ is the observation which lies in the hyperplane. In order to obtain a hyperplane which has a direction vector parallel to one of the co-ordinate axes, it is necessary that the direction vector $\mathbf{d}$ be $\mathbf{d}^t = (0, 0, ..., i, 0, 0, 0)$, where $i = 1$ denotes the desired co-ordinate axis. To cut the hyper-ellipsoid with the hyperplane, it is equivalent to solve the p-1 simultaneous equations

$$\begin{aligned}(\mathbf{x}-\overline{\mathbf{x}})^t\mathbf{S}^{-1}(\mathbf{x}-\overline{\mathbf{x}}) &= T^2_{1-\alpha}\\ \mathbf{d}(\mathbf{x}-\mathbf{b}) &= 0\end{aligned}$$

for $i = 1, 2, 3, ..., p$, excluding the two targeted control variables.

After the substitution, the remaining equation is a second-degree polynomial equation with two unknown variables. In fact, the idea here is to reduce the $p-1$ simultaneous equations into a second-degree polynomial equation. Then the polynomial equation is diagonalized, so that the cross-product term is eliminated. The remaining equation is further reduced to the standard equation of an ellipse by completing the squares. Under normal conditions, the hyperplane method will provide $p(p-1)/2$ elliptical control charts.

5.4.6 Roy's Interpretation Algorithm

Another method based on the T^2 is the step-down procedure of Roy (1958). It assumes that there is an a priori ordering among the means of the p variables and it sequentially tests subsets using this ordering to determine the sequence. The test statistic has the form

$$F_j = \frac{T_j^2 - T_{j-1}^2}{1 + \left[T_{j-1}^2/\left(n-1\right)\right]},$$

where the T_j^2 represents the unconditional T^2 for the first j variables in the chosen group. In the setting of a multivariate control chart, F_j would be the charting statistic, which under the null has the following distribution

$$\frac{\left(n_f-1\right)j}{n_f-j}F\left(p_j, n_f-j\right).$$

This procedure can be considered as an alternative to the regular T^2 chart and not only as supplement. Moreover, it can be shown that the enumerator of F_j is a conditional T^2 value

$$T_j^2 - T_{j-1}^2 = T_{j \cdot 1,2,...,j-1}.$$

5.4.7 Timm's Interpretation Algorithm

Another method based on the T^2 is the step-down procedure of Timm (1996), using Finite Intersections Tests (FIT). It assumes that there is an a priori ordering among the means of the p variables and it sequentially tests subsets using this ordering to determine the sequence.

Although T^2 is optimal for finding a general shift in the mean vector, it is not optimal for shifts that occur for some subset of variables, a variable a time. Timm (1996) states that when this occurs, the optimal procedure is to utilize a finite intersection test.

A process is in-control if each hypothesis

$$H_i : \mu_i = \mu_{0i}, \qquad i = 1, 2, ..., p$$

or equivalently the intersection of the H_i

$$H_0 : \bigcap_{i=1}^{p} H_i$$

is simultaneously true, where μ_i is the mean of the ith variable. To test each hypothesis, we may use the FIT procedure. To construct a FIT of H_0, we may define a likelihood ratio test statistic for each elementary hypothesis H_i and determine the joint distribution of the test statistics. Let the test statistics be defined as

$$Z_i^2 = \frac{(X_i - \mu_{0i})^2}{\sigma_{ii}}.$$

The joint distribution of Z_i^2 follows a non-central multivariate X^2 distribution with 1 degree of freedom. The known variance-covariance matrix $\mathbf{\Sigma} = \sigma^2 \mathbf{\Omega}$, where $\mathbf{\Omega} = [p_{ij}]$ is the correlation matrix of the accompanying multivariate normal. A multivariate quality control process is in-control if $Z_i^2 < X^2_{p,1-\alpha}$ where $P(Z_i^2 < X^2_{p,1-\alpha}/H_0) = 1-\alpha$ and $X^2_{p,1-\alpha}$ is the $1-\alpha$ percentage value of the multivariate X^2 distribution with 1 degree of freedom. The process is out-of-control if $maxZ_i^2 > X^2_{p,1-\alpha}$, where $P(maxZ_i^2 > X^2_{p,1-\alpha}) = \alpha$. Since a non-central multivariate X^2 distribution with 1 degree of freedom is a special case of the multivariate t distribution with infinite degrees of freedom, the process may alternatively be judged as out-of-control if $P(max\,|T_i| = |X_i - \mu_{0i}|\,/\sqrt{\sigma_{ii}} > T^2_{p,1-\alpha}) = \alpha$. Hence,

$$P(|T_i| \leq T^2_{p,1-\alpha}/H_0) = 1 - \alpha.$$

The $1-\alpha$ level simultaneous confidence sets are easily established for each variable

$$X_i - T^2_{p,1-\alpha}\sqrt{\sigma_{ii}} \leq \mu_i \leq X_i + T^2_{p,1-\alpha}\sqrt{\sigma_{ii}}$$

The process is said to be out-of-control if the confidence sets d_i do not contain μ_{0i} for $i = 1, 2, ..., p$. Timm (1996) gave a step down FIT procedure for the case that the variance-covariance matrix $\mathbf{\Sigma}$ is unknown.

5.4.8 Contributors to a Multivariate SPC Chart Signal

The contribution of a variable to a control signal can be measured by the minimum value of the chi-squared statistic that can be obtained by changing a single variable. Variables that incur large changes are important to the signal. Runger et al. (1996) propose the following diagnostics.

Let $\mathbf{x}$ be the p dimensional vector of an observation, with known mean vector $\mathbf{0}$ and known variance-covariance matrix $\mathbf{\Sigma}$. Let $\mathbf{e}_i$ be the unit vector in the direction of the ith coordinate axis. To measure the contribution of X_i to X^2, we determine c_i to minimize

$$\left(\mathbf{x} - \frac{c_i\mathbf{e}_i}{\sqrt{\mathbf{e}_i^t\mathbf{\Sigma}^{-1}\mathbf{e}_i}}\right)^t \mathbf{\Sigma}^{-1} \left(\mathbf{x} - \frac{c_i\mathbf{e}_i}{\sqrt{\mathbf{e}_i^t\mathbf{\Sigma}^{-1}\mathbf{e}_i}}\right)$$

The expression $\mathbf{e}_i^t \mathbf{\Sigma}^{-1}\mathbf{e}_i$ is a scale factor that is used so that c_i can be interpreted as a measure of Euclidean distance. This is discussed further below. If c_i is large relative to the others c_i's, then a large modification to X_i is required to minimize X^2 and this indicates that X_i is an important contributor to X^2.

Also, let

$$\mathbf{z} = \mathbf{\Sigma}^{-1/2}\mathbf{x} \tag{5.1}$$

and

$$\mathbf{v}_i = \mathbf{\Sigma}^{-1/2}\frac{\mathbf{e}_i}{\sqrt{\mathbf{e}_i^t \mathbf{\Sigma}^{-1}\mathbf{e}_i}}.$$

Then, $\mathbf{v}_i$ is a unit vector in the direction $\mathbf{\Sigma}^{-1/2}\mathbf{e}_i$. Now, (5.1) can be written as

$$(\mathbf{z} - c_i\mathbf{v}_i)^t(\mathbf{z} - c_i\mathbf{v}_i)$$

and c_i is interpreted as the Euclidean distance along the unit vector $\mathbf{v}_i$ that minimizes X^2. The value of c_i that minimizes (5.1) can be determined as the slope of the regression of $\mathbf{z}$ on $\mathbf{v}_i$. Therefore,

$$c_i = \mathbf{v}_i^t\mathbf{z} = \frac{\mathbf{e}_i^t \mathbf{\Sigma}^{-1}\mathbf{x}}{\sqrt{\mathbf{e}_i^t \mathbf{\Sigma}^{-1}\mathbf{e}_i}}$$

and obviously c_i is proportional to $\mathbf{e}_i^t \mathbf{\Sigma}^{-1}\mathbf{x}$.

Interestingly, c_i equals the recommended control statistic to detect a shift of the process mean along the vector $\mathbf{e}_i$ as described by Healy (1987), Pignatiello and Runger (1990), and Hawkins (1993). A simple multivariate control is to compare the relative magnitudes of c_i^2 for $i = 1, 2, 3, ..., p$. If c_i^2 is large, the conclusion is that the component measurement X_i is distant from the bulk of the historical data and X_i is a primarily contributor to X^2. Geometrically, c_i^2 is the length of the orthogonal projection of the vector $\mathbf{z}$ onto $\mathbf{v}_i$. That is

$$c_i^2 = \|\mathbf{P}_p\mathbf{z}\|,$$

where the orthogonal projection matrix onto the vector $\mathbf{v}_i$ is denoted as $\mathbf{P}_p$. The geometric analogies are facilitated by considering $\mathbf{Z}$ and c_i because, after transforming to these variables, X^2 is just the squared Euclidean distance of $\mathbf{Z}$ from the zero vector.

A second approach to the problem is the following. Consider the minimum value of X^2 that is obtained when X_i is changed by c_i. Define

$$X_i^2 = \left(\mathbf{x} - \frac{c_i \mathbf{e}_i}{\sqrt{\mathbf{e}_i^t \mathbf{\Sigma}^{-1} \mathbf{e}_i}}\right)^t \mathbf{\Sigma}^{-1} \left(\mathbf{x} - \frac{c_i \mathbf{e}_i}{\sqrt{\mathbf{e}_i^t \mathbf{\Sigma}^{-1} \mathbf{e}_i}}\right) = (\mathbf{z} - c_i \mathbf{v}_i)^t (\mathbf{z} - c_i \mathbf{v}_i). \tag{5.2}$$

Then, (5.2) can be interpreted as the residual sum of squares in a regression model. Therefore,

$$X_i^2 = \mathbf{z}^t(\mathbf{I} - \mathbf{P}_i)\mathbf{z} = \mathbf{z}^t\mathbf{z} - \mathbf{z}^t\mathbf{v}_i\mathbf{v}_i^t\mathbf{z} = \mathbf{z}^t\mathbf{z} - c_i^2. \tag{5.3}$$

Thus, X_i is a major contributor to X^2, if the value of X^2 can be substantially reduced by a modification to X_i. Consequently, if the metric $D_i = X^2 - X_i^2$ is large then X_i is a major contributor to X^2. Since $X^2 = \mathbf{x}_i^t \mathbf{\Sigma}^{-1} \mathbf{x}_i = \mathbf{z}^t \mathbf{z}$ and because of relation (5.3) we have that $D_i = c_i^2$ therefore the metric D_i is equivalent to c_i^2. Finally, a third approach to develop a diagnostic is the union-intersection principle of multivariate hypothesis tests.

5.4.9 Cause-Selecting Control Chart

Wade and Woodall (1993) consider a two step process in which the steps are not independent. The first step of a cause-selecting control chart is to chart variable X_1 (first step of the process) and then monitor the outgoing quality X_2 (second step of the process) after adjusting for the incoming quality. This method uses a relation between X_2 and X_1, where a simple regression model appears to be very useful. To be more precise, a chart for X_1 and a chart for $Z = X_2 - \widehat{X}_2$, where $\widehat{X}_2$ is the estimate for X_2 based on the regression line, are used. Thus, the Z_i's are independent normal random variables if X_1 and X_2 are normally distributed variables. If controllable assignable causes are present in the process the distribution of Z_i's shifts from normality for some values of i.

5.4.10 Hawkins' Interpretation Method

Hawkins (1991, 1993), as we have already mentioned, defined a set of regression-adjusted variables in which each variable was regressed on all others. In a first approach Hawkins defined the set of regression-adjusted variables using the vector $\mathbf{Z} = \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu})$. If we replace $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ with their sample estimates and expanding the left-hand side of the previous relation Hawkins showed that the jth component of $\mathbf{Z}$ is equal to

$$Z_j = \frac{T_{j \cdot 1,2,\ldots,j-1,j+1,\ldots,P}}{s_{j \cdot 1,2,\ldots,j-1,j+1,\ldots,P}}.$$

Then, Z_j is the standardized residual when the jth variable is regressed on the remaining $p-1$ variables in $\mathbf{X}$. This regression statistic is useful in the interpretation of a T^2 signal because its value is directly connected with the value of the T^2 statistic although it is only one of the several different conditional T^2 values.

An additional approach presented by Hawkins is based on the decomposition of the T_j^2 statistics, using the standardized residuals from the jth variable on the first $j-1$ variables. This is defined by $\mathbf{Y} = \mathbf{C}(\mathbf{X} - \boldsymbol{\mu})$ where $\mathbf{C}$ is the Cholesky lower triangular root of $\boldsymbol{\Sigma}^{-1}$. The jth component of vector $\mathbf{Y}$ is given by

$$Y_j = \frac{T_{j \cdot 1,2,\ldots,j-1}}{s_{j \cdot 1,2,\ldots,j-1}}$$

As in the first approach Y_j is one of the conditional T^2 values.

5.4.11 Minimax Control Chart

Sepuldveda and Nachlas (1997) presented a new control chart. This chart is called the Minimax control chart because it is based on monitoring the maximum standardized sample mean and the minimum standardized sample mean of samples taken from a multivariate process. It is assumed that the data are normally distributed and that the variance-covariance matrix is known and constant over time.

Samples of size n are taken from a p-dimensional process. The process is assumed to have sudden shifts in the mean such that the new mean μ_{1i}, $i = 1, 2, ..., p$, for the ith changing variable is given by $\mu_{1i} = \mu_{0i} + \delta_i \sigma_{0i}$ and consequently $\delta_i = (\mu_{1i} - \mu_{0i})/\sigma_{0i}$. Whenever $\delta_i = 0$, the process is said to be in-control. To decide whether the process is in-control or not, the minimax control chart is used as a method to test in each sample $H_0 : \boldsymbol{\delta} = \mathbf{0}$ against $H_1 : \boldsymbol{\delta} \neq \mathbf{0}$.

The principal idea behind the minimax control chart is to standardize all p means and to monitor the maximum Z^{mx} and the minimum Z^{mn} of those standardized sample means. Note that Timm (1996) monitors only the maximum. Let

$$\begin{aligned} Z^{mn} &= \min(Z_i),\ i = 1, 2, 3, ..., p \\ Z^{mx} &= \max(Z_i),\ i = 1, 2, 3, ..., p, \end{aligned}$$

where

$$Z_i = \frac{\overline{X}_i - \mu_{0i}}{\sigma_{ii}/\sqrt{n}}.$$

Therefore, by monitoring the maximum and the minimum standardized sample mean, an out-of-control signal is directly connected with the corresponding out-of-control variable. Sepuldveda and Nachlas (1997), also discussed the statistical properties and the ARL performance of the minimax control chart.

5.5 Using Principal Components

Another method used for the identification problem is principal components analysis (PCA). This method was first proposed by Jackson (1991) and further discussed later by Pignatiello and Runger (1990), Kourti and MacGregor (1996). A $p \times p$ symmetric, non-singular matrix, such as the variance-covariance matrix $\boldsymbol{\Sigma}$, may be reduced to a diagonal matrix $\mathbf{L}$ by premultiplying and postmultiplying it by a particular orthonormal matrix $\mathbf{U}$ such that $\mathbf{U}^\intercal \boldsymbol{\Sigma} \mathbf{U} = \mathbf{L}$. The diagonal elements of $\mathbf{L}$, $l_1, l_2, ..., l_p$ are the characteristic roots, or eigenvalues of $\mathbf{\Sigma}$. The columns of $\mathbf{U}$ are the characteristic vectors, or eigenvectors of $\mathbf{\Sigma}$. Based on the previous result the method of PCA was developed (see e.g., Jackson (1991)). The PCA transforms p correlated variables $x_1, x_2, ..., x_p$ into p new uncorrelated ones. The main advantage of this method is the reduction of dimensionality. Since the first two or three PCs usually explain the majority of the variability in a process, they can be used for interpretation purposes instead of the whole set of variables.

5.5.1 Jackson's Approach

Principal components can be used to investigate which of the p variables in a multivariate control chart are responsible for an out-of-control signal. The most common practice is to use the first k most significant principal components, if a T^2 control chart gives an out-of-control signal, for further investigation.

The basic idea is that the first k principal components can be physically interpreted, and named. Consequently, if the T^2 chart gives an out-of-control signal and, for example, the second principal component chart also gives an out-of-control signal, then from the interpretation of this component, a direction to the variables which are suspect to be out-of-control can be deduced (Jackson (1991)).

The practice just mentioned transforms and considers the variables as a set of attributes. The discovery of the assignable cause of the problem, with this method, demands a further knowledge of the process itself, from the practitioner. The basic problem is that the principal components do not always have a physical interpretation.

5.5.2 Bivariate Control Chart for Paired Measurements

As we already said one of the ways to use PCA in the problem under question is to chart them. However, if the components are not easily interpreted the problem remains. Tracy et al. (1995) expanded previous work and provided an interesting bivariate setting in which the principal components have meaningful interpretations. When monitoring a process with paired measurements (for example two testing laboratories) on a single sample, the principal components of the correlation matrix actually represent the characteristics of interest for process control. The correlation coefficient between the original variables is the only additional information needed to describe the condition of the process.

5.5.3 Kourti's and MacGregor's Approach

Kourti and MacGregor (1996) provided a newer approach based on principal components analysis. The T^2 statistic is expressed in terms of normalized principal components scores of the multinormal variables. When an out-of-control signal is received, the normalized scores with high values are detected, and contribution plots are used to find the variables responsible for the signal. A contribution plot indicates how each variable involved in the calculation of that score contributes to it. The computation of variable contributions showed that principal components can actually have physical interpretation. This approach is particularly applicable to large ill-conditioned data sets, due to the use of principal components.

5.6 A New Method

Let $\mathbf{x}_i = (x_{1i}, x_{2i}, ..., x_{pi})^\intercal$ denote the observation (vector) i for the p variables of a process. Assume that $\mathbf{x}_i$ follows a p-dimensional normal distribution $N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, where $\boldsymbol{\mu}_0$ is the vector $(p \times 1)$ of known means and $\boldsymbol{\Sigma}_0$ is the known $(p \times p)$ variance-covariance matrix. We want to keep this process under control. For this purpose we use a X^2 control chart given by the formula $X_i^2 = (\mathbf{x}_i - \boldsymbol{\mu}_0)^\intercal \, \boldsymbol{\Sigma}_0^{-1} \, (\mathbf{x}_i - \boldsymbol{\mu}_0)$. If the value of this statistic plots above $X^2_{p,1-a}$ we get an out-of-control signal, where $X^2_{p,1-a}$ is the chi-square distribution with p degrees of freedom and a is the probability of plotting outside the control limits if we are in-control. The next problem is to detect which variable is the one that caused the problem.

The typical form of a PCA model is the following: $Z_k = u_{1k}X_1 + u_{2k}X_2 + u_{3k}X_3 + ... +$

$u_{pk}X_p$, where Z_k is the k PC, $(u_{1k}, u_{2k}, u_{3k}, ..., u_{pk})^\intercal$ is the corresponding k eigenvector and $X_1, ..., X_p$ are the process variables. The score for vector $\mathbf{x}_i$ in PC k is $Y_{ki} = u_{1k}x_{1i} + u_{2k}x_{2i} + ... + u_{pk}x_{pi}$. Assuming that the process variables follow a multivariate normal distribution the PCs are also normally distributed.

Our purpose is to use PCA, when we have an out-of-control signal in the X^2 control chart, to identify the variable or variables that are responsible. For this objective two different methodologies are developed one for the case that the covariance matrix has only positive correlations and the second one for the case that we have both positive and negative ones (Maravelakis, Bersimis, Panaretos, Psarakis (2002)).

5.6.1 Covariance matrix with positive correlations

Assume that using one of the existing methods for choosing PCs (see, e.g. Jackson (1991)), Runger and Alt (1996)) we choose $d \leq p$ significant PCs. The proposed method in this case is based on ratios of the form

$$r_{ki} = \frac{(u_{k1} + u_{k2} + ... + u_{kd})x_{ki}}{Y_{1i} + Y_{2i} + ... + Y_{di}}, \tag{5.4}$$

where x_{ki} is the ith value of variable X_k, Y_{ji}, $j = 1, .., d$ is the score of the ith vector of observations in the jth PC (Bersimis (2001)). In this ratio, the numerator corresponds to the sum of the contributions of variable X_k in the first d PCs in observation (vector) i, whereas the denominator is the sum of scores of observation (vector) i in the first d PCs. Since we have assumed that the variables follow a multivariate normal distribution the ratios are ratios of two correlated normal variables.

The rationale of this method is to compute the impact of each of the p variables on the out-of-control signal by using its contribution to the total score. It is obvious that the use of only the first d PCs excludes pieces of information. However, a multivariate chart is used when there is at least moderate and usually large correlation between the variables. Under such circumstances the first d PCs account for the largest part of the process variability. The main disadvantage of using PCs in process control, as reported by many authors (see e.g., Runger and Alt (1996), Kourti and McGregor (1996)), is the lack of physical interpretation. The proposed method eliminates much of that criticism.

Since we have a ratio of two correlated normals its distribution can be computed using the analytical result of Hinkley (1969). Specifically, if X_1, X_2 are normally distributed random variables with means μ_i, variances σ_i^2 and correlation coefficient ρ the distribution function of $R = X_1/X_2$ is given by the formula

$$F(r) = L\left\{\frac{\mu_1 - \mu_2 r}{\sigma_1\sigma_2\alpha(r)}; -\frac{\mu_2}{\sigma_2}; \frac{\sigma_2 r - \rho\sigma_1}{\sigma_1\sigma_2\alpha(r)}\right\} + L\left\{\frac{\mu_2 r - \mu_1}{\sigma_1\sigma_2\alpha(r)}; \frac{\mu_2}{\sigma_2}; \frac{\sigma_2 r - \rho\sigma_1}{\sigma_1\sigma_2\alpha(r)}\right\}, \tag{5.5}$$

where $L(h;k;\gamma) = \frac{1}{2\pi\sqrt{1-\gamma^2}} \int_h^\infty \int_k^\infty \exp\left\{-\frac{x^2 - 2\gamma xy + y^2}{2(1-\gamma^2)}\right\} dxdy$ is the standard bivariate normal integral.

However, the proposed method has a correlation problem since the ratios of different variables are interrelated. A simulation study presented in Section 5.6.4 is implemented to test the effect on the performance of the proposed procedure. In the following we present the method as a stepwise procedure.

- Step 1. Calculate the X^2 statistic for the incoming observation. If we get an out-of-control signal continue with Step 2.

- Step 2. Calculate ratios for all the variables using relation (5.4). Calculate as many ratios for each variable as the number of observations from the beginning of the process. If the proposed process is not used for the first time, calculate as many ratios for each variable, as the number of observations from the last out-of-control signal till the out-of-control signal we end up with in Step 1. Alternatively, calculate ratios for only the (last) observation that caused the out-of-control signal (see Section 5.6.4).

- Step 3. Plot the ratios for each variable in a control chart. Compute the a and $1-a$ percentage points of distribution (5.5) with suitable parameters and use them as lower control limit (LCL) and upper control limit (UCL), respectively.

- Step 4. Observe which variable, or variables, issue an out-of-control signal
- Step 5. Fix the problem and continue with Step 1.

In the case where all the variables are positively correlated as Jackson (1991) indicates, the first PC is a weighted average of all the variables. Consequently, we can use only this PC for inferential purposes.

5.6.2 Covariance matrix with positive and negative correlations

In this case we propose the computation of ratios of the form

$$r^*_{ki} = \frac{(u_{k1} + u_{k2} + ... + u_{kd})x_{ki}}{\overline{Y}_1 + \overline{Y}_2 + ... + \overline{Y}_d}, \tag{5.6}$$

where x_{ki} is the ith value of variable X_k and $\overline{Y}_j$, $j = 1, ..., d$ is the score of the jth PC using, in place of each X_k, their in-control values. The subscript d stands for the number of significant principal components as in the preceding case.

These ratios are the sum of the contributions of variable X_k in the first d PCs in observation (vector) i, divided by the sum of the in-control scores of the first d PCs. Since the denominator of this statistic is constant we actually compute the effect of each of the p variables on the out-of-control signal. The numerator of the ratios is normally distributed, as already stated, whereas the denominator is just a constant. Therefore the ratios (5.6) are normally distributed.

Since the variables are correlated the statistic proposed in (5.6) for the k different variables may exhibit a correlation problem. As in the previous case a simulation study is presented in Section 5.6.4 to test for the effect of the correlation on the control limits performance of the proposed procedure. The proposed method in steps is as follows:

- Step 1. Calculate the X^2 statistic for the incoming observation. If we get an out-of-control signal continue with Step 2.

- Step 2. Calculate ratios for all the variables using relation (5.6). Calculate as many ratios for each variable as the number of observations from the beginning of the process. If the proposed process is not used for the first time, calculate as many ratios for each variable, as the number of observations from the last out-of-control signal till the out-of-control signal we end up with in Step 1. Alternatively, calculate ratios for only the (last) observation that caused the out-of-control signal (see Section 5.6.4).

- Step 3. Plot the ratios for each variable in a control chart. Compute the a and $1 - a$ percentage points of the normal distribution with suitable parameters and use them as LCL and UCL, respectively.

- Step 4. Observe which variable, or variables, issue an out-of-control signal

- Step 5. Fix the problem and continue with Step 1.

We have to mention that this procedure can not be applied when we have standardized values since the denominator of the ratios in (5.6) equals zero.

5.6.3 Illustrative examples

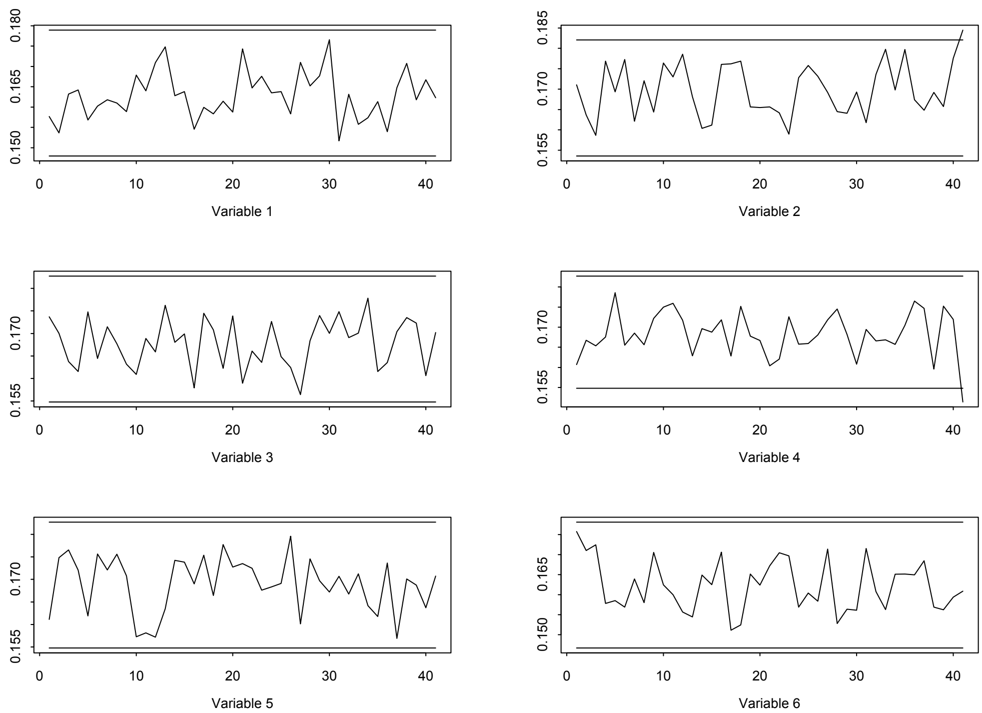
Two examples, one for each case are presented in the sequel.

Example 1. Assume that we have a process with known covariance matrix

$$\begin{bmatrix} 100 & & & & & \\ 70 & 100 & & & & \\ 80 & 80 & 100 & & & \\ 75 & 85 & 75 & 100 & & \\ 75 & 80 & 80 & 80 & 100 & \\ 75 & 72 & 75 & 75 & 75 & 100 \end{bmatrix}$$

and in-control vector of means $(100, 100, 100, 100, 100, 100)^\intercal$. We simulated 40 in-control observations from a multivariate normal distribution with the preceding parameters. Then, we simulated out-of-control ones with the same covariance matrix but now with vector of means $(100, 115, 100, 85, 100, 100)^\intercal$, until we get an out-of-control signal in the X^2 test. The shift is 1.5σ in the means of variables 2 and 4. We get a signal on the first out-of-control observation and we plot each of the variables in a control chart (Figure 5.1) with the control limits from distribution (5.5) using $a = 0.05$. Note that we used the average root method for simplicity and we ended up with one significant principal component (see, e.g. Jackson (1991)). It is obvious from Figure 5.1 that the out-of-control variables were identified and additionally the direction of the shift was also revealed.

Figure 5.1. Control charts for positive covariance matrix

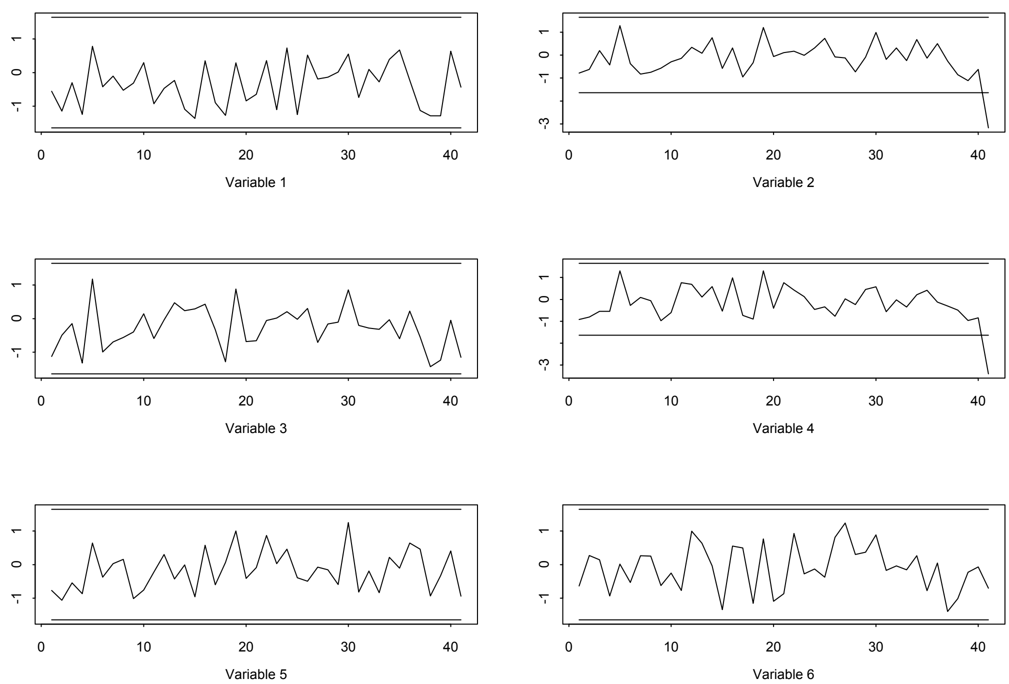


Example 2. Assume that we have a process with known covariance matrix

$$\begin{bmatrix} 100 & & & & & \\ -70 & 100 & & & & \\ 80 & -80 & 100 & & & \\ 75 & -85 & 75 & 100 & & \\ 75 & -80 & 80 & 80 & 100 & \\ 75 & -72 & 75 & 75 & 75 & 100 \end{bmatrix}$$

and in-control vector of means $(100, 100, 100, 100, 100, 100)^\intercal$. We simulated 40 in-control observations from a multivariate normal distribution with the same parameters as in the previous example. Then, we simulated out-of-control ones with vector of means

Figure 5.2. Control charts for positive negative covariance matrix



$(100, 115, 100, 85, 100, 100)^\intercal$, the same covariance matrix till we get an out-of-control signal in the X^2 test. The shift is again 1.5σ in the means of variables 2 and 4. We plot each of the variables in a control chart (Figure 5.2) with the control limits from a normal distribution using $a = 0.05$. As in example 1, we have one significant PC using the average root method again (see, e.g. Jackson (1991)). We have to indicate that in Figure 5.2 the ratios are standardized and the control limits are properly modified. However, it is not necessary to do this when using this technique. From Figure 5.2, we deduce that the out-of-control variables were identified but the direction of the shift was not.

5.6.4 Further Investigation

One may observe that the ratios in both methods are interrelated. This fact may affect the control limits of the charts. In order to examine this possible correlation we performed a simulation study. In particular, we simulated 100000 in-control ratios from the known covariance matrices and vector of means of the two examples and we computed the theoretical control limits as proposed in Section 5.6.3 with $a = 0.05$. Then, we checked if each ratio is in or out of these limits and recorded it. We used this information in order to approximate the probability of plotting outside the control limits if we are in-control of our limits and compare it with the theoretical one. The results are presented in Table 5.1.

Table 5.1. Probability of plotting outside the control limits if we are in-control

Variable	1	2	3	4	5	6
example 1	5040	5006	5007	4968	5000	5096
example 2	4941	4983	4898	4882	4906	5005

It should be noted that although the two examples are specific cases, a large number of other cases revealed the same performance. Consequently, we may draw the conclusion that the interrelation does not affect either of the proposed processes. However, we have to point out that after careful examination the procedure proposed for positive correlations can be used only when we have positive values for the in-control means. If a process does not have positive in-control means, which is a rare event, we may use instead the statistic (5.6) that is not affected in any case.

Another point which has to be checked is the performance of the processes in identifying the out-of-control variables and the direction of the shift. For this purpose a simulation study was conducted. Using the covariance matrices and the in-control means of the examples we computed in each case the theoretical control limits. Then, we simulated observations from the out-of-control mean vector used in the examples until we got a signal from the X^2 test using $a = 0.05$. Next, we computed the ratios for each variable and plotted them in a chart with the corresponding control limits. We checked each ratio if it is in or out of the control limits for every variable and recorded which variable, or variables, have given an out-of-control signal and in which direction. We repeated the whole process 10000 times and the results are presented in Tables 5.2 and 5.3. In the first row of the tables (U), we have the number of times the generated ratios crossed the UCL for each of the variables, in the second row (L) we have the corresponding number of times the generated ratios crossed the LCL for each of the variables and in the last line (Total) we have the number of times the generated ratios crossed UCL or LCL for each of the variables. One may observe that in some variables there is an inconsistency, since the sum of rows U and L does not equal the total. This happens because in one iteration we may generate more than one observations (vectors) until we get an out-of-control signal in the X^2 test -although this test is sensitive for such shifts- and after computing the ratios for each variable it is possible that for one variable the first ratio crosses UCL and the second ratio crosses LCL. Therefore, we record one value in row U and one in row L but only one in row total.

Table 5.2. Out of control performance for example 1

Variable	1	2	3	4	5	6
U	283	9371	228	0	275	233
L	240	0	299	9429	231	245
Total	523	9371	527	9429	506	478

From Table 5.2, we observe that the statistic used is very informative since it is able to detect the out-of-control variables with very high precision and also to identify the direction of the shift with absolute success. This kind of behavior is similar in other examples also, keeping in mind the limitation about the positive in-control means. Note also that the total times the other variables gave a false signal almost coincides with the type I error rate of the constructed limits.

Table 5.3. Out of control performance for example 2

Variable	1	2	3	4	5	6
U	364	1	357	2	366	350
L	353	4627	394	4603	357	371
Total	715	4628	748	4605	717	720

In Table 5.3, we see that the statistic used detected the out-of-control situation but not with the same degree of success as in the previous case. Moreover, the direction of the shift was not identified. The false alarm rate of the in-control variables is not significantly different from the theoretical one.

We already said that the X^2 test is sensitive in the sense that it gives a quick signal when we have out-of-control observations. In order to examine the ability of the last generated observation (the one that gives the signal on the X^2 test) to identify the shifted variable we checked their performance on the previous simulation study. The results are displayed in Tables 5.4 and 5.5.

Table 5.4. Out of control performance of the last observation for example 1

Variable	1	2	3	4	5	6
U	283	9370	228	0	275	233
L	240	0	299	9429	231	245
Total	523	9370	527	9429	506	478

From Table 5.4, we conclude that the performance of the statistic (5.4) is almost totally explained by the ratios of the last observation meaning that the ratios of the last observation are sufficient to draw a conclusion about the out-of-control variable. On the other hand, from Table 5.5 we observe that this does not happen for the statistic (5.6). One may argue that since the X^2 test is sensitive we produce a small number of observations in each iteration hence the performance in both cases is a result of this fact. When we have small shifts, where we produce more observations, the last observation is not that informative.

Table 5.5. Out of control performance of the last observation for example 2

Variable	1	2	3	4	5	6
U	333	1	316	2	324	315
L	312	4279	343	4251	310	333
Total	645	4280	659	4253	634	648

The proposed procedures are valid under the assumption of known variance-covariance matrix. However, this is not a case usually met in practice. Tracy et al. (1992) examined the performance of multivariate control charts for individual observations when the covariance matrix is known and unknown. They showed that the test statistics used in either case perform the same for a number of observations that depends on the variables involved. This number of observations is small enough, for instance when we have five variables we need 100 observations (vectors) for the two statistics to give a close number.

As Woodall and Montgomery (1999) note in today's industry we have huge data sets therefore such a number of observations should not be a problem. From this number of observations we estimate the mean vector and the covariance matrix used in our process. Although the control limits computed using the procedures developed in subsections 5.6.1 and 5.6.2 will not be exact under the estimation process, we expect them to have a satisfactory performance if we use the required number of observations.

5.6.5 A Comparison

As we already stated in section 5.5, the competitive methods that use principal components for the specific problem are Jackson's (1991), Tracy et al.'s (1995) and Kourti and MacGregor's (1996). Kourti and MacGregor (1996) provide an improved method in relation to Jackson's (1991) and Tracy et al. (1995) have the disadvantage that their method is applied to a bivariate case. Therefore, the rival of the proposed method is the one by Kourti and MacGregor (1996).

In order to check the performance of the two antagonistic methods we perform a simulation study. We apply the method of Kourti and MacGregor (1996) in the data of examples 1 and 2 of Section 5.6.3. As Kourti and MacGregor (1996) propose, we use Bonferroni limits on the normalized scores and we calculate the contributions of the variables with the same sign as the score, since contributions of the opposite sign does not add anything to the score, in fact they make it smaller. In the paper of Kourti and MacGregor (1996) there is not a specific rule on how to decide whether a contribution is significant or not. Since in the two examples of Section 5.6.3 we have two variables shifted, we choose to record the first and the second larger contributions in each iteration of the simulation study if they exist. The simulation study was conducted 10000 times in order to have a credible estimate of the ability of Kourti and MacGregor's (1996) method to identify the out-of-control variables. In Tables 5.6 and 5.7 we have the results of this simulation for examples 1 and 2 respectively.

Table 5.6. Performance of Kourti and MacGregor's method for example 1

Variable	1	2	3	4	5	6
Largest contribution	2	5204	16	4717	0	0
Second largest contribution	1511	2864	1865	3655	30	10
Total	1513	8068	1881	8372	30	12

Table 5.7. Performance of Kourti and MacGregor's method for example 2

Variable	1	2	3	4	5	6
Largest contribution	23	191	10	2	51	2
Second largest contribution	64	63	47	74	19	12
Total	87	254	57	76	70	14

From the results in Table 5.6 for Example 1, we observe that the method of Kourti and MacGregor does not succeed in identifying the out-of-control variables as many times as our proposed method does (see Tables 5.2 and 5.4). Moreover, the inherent inability of the method to point if there is an upward or a downward shift is also present. Things are even worse in Table 5.7 for example 2. Specifically, the method of Kourti and MacGregor leads to recordable contributions very rarely, a fact that may lead the practitioner to assume that the signal on the multivariate chart is due to the probability of plotting outside the control limits if we are in-control. The ability of the new method to operate more effectively is obvious (see Tables 5.3 and 5.5).

5.6.6 A Graphical Technique

The charts proposed in Sections 5.6.1 and 5.6.2 are Shewhart type. Therefore, they have the ability to identify large shifts quickly but they are not that good for small shifts. An alternative way to plot these statistics is as a Cumulative Sum (CUSUM) chart. The definition of the CUSUM chart for detecting upward shifts is

$$\begin{aligned} S_0^+ &= 0 \\ S_n^+ &= \max(0, S_{n-1}^+ + (X_{pn} - k)), \end{aligned}$$

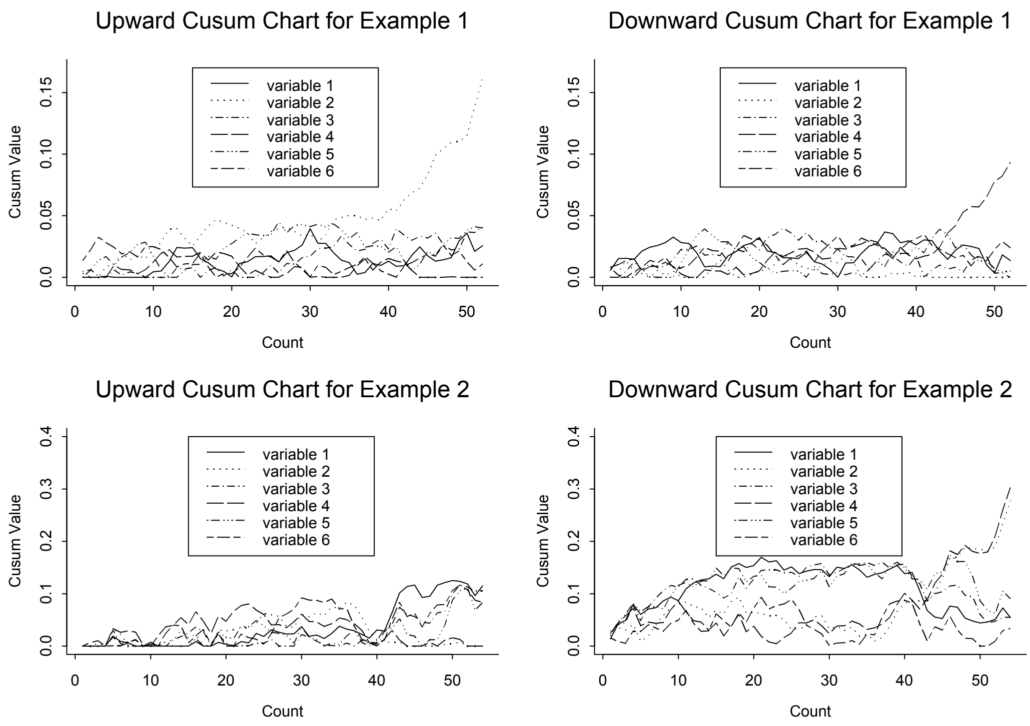
where X_{pn} is the nth observation of variable X_p and k is called the reference value. The corresponding CUSUM chart for detecting downward shifts is

$$\begin{aligned} S_0^- &= 0 \\ S_n^- &= \max(0, S_{n-1}^- + (k - X_{pn})), \end{aligned}$$

In the usual concept of CUSUM charts we evaluate an optimal value of k depending on the distributional assumption (see Hawkins (1992) and Hawkins and Olwell (1998)). This value of k is used along with the value h, which is the control limit, to characterize the ARL performance of a CUSUM chart. In our case the application of this theory for CUSUM charts is cumbersome due to the underlying distribution. However, we can use the previously defined statistics for upward and downward shifts as a graphical technique solely. The only quantity remaining unknown is the value k we have to use. A straightforward selection for k is to use in each case of statistics (5.4) and (5.6) their in-control counterparts. Specifically, for statistic (5.4) we use in place of each x_{ki} its in-control value both in the numerator and the denominator. A similar action takes place in statistic (5.6) but only in the numerator this time.

To study the performance of these statistics in practice we applied them to the examples of Section 5.6.3. We used the same 40 in-control observations but now we generated out-of-control ones with the same covariance matrix as in Section 5.6.3, and vector of means $(100, 105, 100, 95, 100, 100)^{\mathsf{T}}$, in both examples until we get an out-of-control signal in the X^2 test. The shift is 0.5σ in the means of variables 2 and 4. In example 1, we simulated 12 observations till the out-of-control signal and 14 in example 2. We computed the CUSUM values for the 52 and 54 values for all six variables in examples 1 and 2 respectively, and we plotted them in the chart given in Figure 5.3.

Figure 5.3. Control charts for CUSUM values



From Figure 5.3 we easily deduce that the charts give a clear indication of the out-of-control variables in both examples. As in the Shewhart type charts of Section 5.6.3, the statistic (5.4) used in example 1 detected also the direction of the shift something that did not happen with statistic (5.6) in example 2. We have to mention here that the effectiveness of the CUSUM as a graphical device for shifts less than 0.5σ is questionable. However, it is an easily interpreted method that can give an indication.

Summarizing, we note that the charts proposed are an easily applied alternative to most of the existing methods since the computational effort is diminished. Furthermore, we try to give an answer to the problem under a control charting perspective giving operational control limits or design strategies that are not difficult for a practitioner to apply.

5.7 Graphical Techniques

5.7.1 Multivariate Profile Charts

Fuchs and Benjamini (1994) presented a method (multivariate profile chart) for simultaneous control of a process and interpretation of an out-of-control signal. The multivariate profile chart (MP chart) is a scatterplot with symbols. Specifically, symbols are used for data of individual variables whereas the location of the symbol on the scatterplot is used for providing information about the group. Each group of observations is displayed by one symbol and this symbol of the profile plot enables the user to get a clear view of the size and the sign of each variable from its reference value.

The first step in the construction of an MP chart is to draw a horizontal base line for each symbol and then calculate sequentially a bar for every variable. This bar plots either above (below) the base line if the deviation is positive (negative). The size of the bar depends on the size of the deviation. Let $d_{ij} = \frac{\overline{x}_{ij} - m_i}{v_i}$, where $\overline{x}_{ij} = 1/n \sum_l x_{ijl}$, v_i is a scale factor, m_i is the ith reference value and x_{ijl} is the value of the lth observation on the ith variable in the jth subgroup. Then, the size of the bar is proportional to d_{ij} up to value of 4. If the standardized deviation exceeds 2 the corresponding bar is painted gray. If the standardized deviation exceeds 3 the corresponding bar is painted black. If all variables means are equal to their standard values the symbol is actually the baseline.

The symbols are plotted in sequential order along a horizontal time axis. In the vertical axis the location of each symbol is determined by the multivariate deviation of the sample mean vector from the standard mean vector as measured by T^2. The critical value of this axis equal to 0.997 is at a symbol size distance from the top and it is chosen to be equivalent to three standard deviations control limit of the univariate control charts. Additionally, a dashed line runs across the chart at the $T^2_{0.997}$ level. The critical value corresponding to 0 is placed half a symbol size distance from the bottom.

A symbol's baseline, which is placed horizontally, is located at the appropriate vertical location when it is less than the T^2 critical value. Those observations with a T^2 value higher than the critical value are placed at the top of the chart stack at the ceiling and completely beyond the dashed line.

The numerical value of the T^2 statistic is not as informative as in the univariate charts the deviations from standard values that are measured by the standard errors. For this reason Fuchs and Benjamini (1994) present on the vertical scale of the chart, the corresponding tail probabilities instead of the actual T^2 values.

If the process is in-control and at the nominal values, the chart appears as a simple horizontal line. If the process is out-of-control, the MP chart gives three visual warnings. The first one is that the symbol for the group that gives an out-of-control signal is located above the line of the observations that are under control. Secondly, as the size of the deviations gets large so does the size of the symbol. The last one is that the symbol gets darker, attracting more visual attention. All three visual warnings hold regardless if we have an upward or a downward signal. Even if the shift is constantly appearing in one direction the visual warnings will be there.

Finally, the MP chart can be used to identify easily the cause of a shift. Since all the individual variables are displayed on a common scale within a group, the MP chart gives us the ability to detect visually a change in their interrelationships. Additionally, when they deviate from their standard, our attention is drawn by the behavior of individual variables, by the height and darkness of the corresponding bars.

5.7.2 Dynamic Biplots

Sparks et al. (1997) presented a graphical method for monitoring multivariate process data based on the Gabriel biplot. This method uses reduction to two dimensions for identifying the in or out-of-control state. However, we use all the data for the decision if we are in or out-of-control even though the method is a dimension reduction one. This method can be used as a control chart and also if we are out-of-control to detect the reason that led to this problem. In particular with this approach a practitioner is able to detect changes in location, variation, and correlation structure.