INDIVIDUAL PROJECT

Laura M Williams, RN, CLNC, MSN
IET603: Statistical Quality Assurance in Science and Technology
Morehead State University
Dr. Ahmad Zargari
4 April 2014
CHAPTER 11
Multivariate Process Monitoring and Control
LEARNING OBJECTIVES

1. Understand why applying several univariate control charts simultaneously to a set of related quality characteristics may be an unsatisfactory monitoring procedure.
2. How the multivariate normal distribution is used as a model for multivariate process data.
3. Know how to estimate the mean vector and covariance matrix from a sample of multivariate observations.
4. Know how to set up and use a chi-square control chart.
5. Know how to set up and use the Hotelling $T^2$ control chart.
6. Know how to set up and use the multivariate exponentially weighted moving average (MEWMA) control chart.
7. Know how to use multivariate control charts for individual observations.
8. Know how to find the phase I and phase II limits for multivariate control charts.
9. Use control charts for monitoring multivariate variability.
10. Understand the basis of the regression adjustment procedure and know how to apply regression adjustment in process monitoring.
11. Understand the basis of principal components and know how to apply principal components in process monitoring.
This chapter will present control charts that can be regarded as the multivariate extensions of some univariate charts in previous chapters.

These multivariate control charts work well when the number of process variables is not too large—say, 10 or fewer.

As the number of variables grows; however, traditional multivariate control charts lose efficiency with regard to shift detection.

A popular approach in these situations is to reduce the dimensionality of the problem using principal components.
11.1 The Multivariate Quality- Control Problem

Simultaneous Monitoring
Use of multiple independent control charts distorts the simultaneous monitoring of the averages.

These control charts indicate that both processes are in control.

*Figure 11.1* Control charts for inner (\( \bar{x}_1 \)) and outer (\( \bar{x}_2 \)) bearing diameters.
Each point is plotted inside the joint control region.

The control ellipse determines if both processes together are in control.

When monitored simultaneously, the control ellipse shows the process is out of control.
Disadvantages:

- Time sequence of plotted points is lost
- Not as useful when there are more than two variables to monitor
11.2 Description of Multivariate Data

- Normal Distribution
- Sample Mean Vector
- Covariance Matrix
Normal Distribution

describes the behavior of a continuous quality characteristic

- **Univariate**
  - Parameters:
    - mean $\mu$
    - variance $\sigma^2$
  - Expressed in Standard Deviation units
  - One Dimensional
  - Graphically represented by

- **Multivariate**
  - Parameters:
    - mean vector, $\mu$
    - covariance matrix, $\Sigma$
  - Each element has a univariate normal distribution
  - Graphically represented by
This is a set of 5 observations which measure 3 variables:

Each row vector is a single observation.

Each column represents the variables being monitored.
Example: Length, Width, and Height.

The mean vector consists of the means of each variable.
Covariance Matrix

The variance-covariance matrix consists of the variances of the variables along the main diagonal and the covariances between each pair of variables in the other matrix positions.

The formula for computing the covariance of the variables X and Y:

\[ COV = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{n - 1} \]

The Results:

\[
\begin{bmatrix}
0.025 & 0.0075 & 0.00175 \\
0.0075 & 0.0070 & 0.00135 \\
0.00175 & 0.00135 & 0.00043 \\
\end{bmatrix}
\]
Estimating $\mu$ and $\Sigma$

- In practice, it is usually necessary to estimate $\mu$ and $\Sigma$ from the analysis of preliminary samples of size $n$, taken when the process is assumed to be in control.

- The average of the sample covariance matrices $\mathbf{S}$ is an unbiased estimate of $\Sigma$ when the process is in control.
Charts for Process Monitoring

Hotelling $T^2$ and Chi-Squared

- Shewhart-type control charts
- Uses information only from current sample
- Relatively insensitive to small and moderate shifts in the mean vector
MONITORS MEAN VECTOR OF THE PROCESS

- Its ability to detect a shift in the mean vector only depends on the magnitude of the shift, not in its direction.

- A direct analog of the univariate Shewhart $x$ Chart.

- A directionally invariant control chart.

- Data can consist of subgroups or individual observations.

- There are two phases for this control chart.
Phase I (retrospective analysis)

Establishes statistical control in the preliminary samples and calculates the upper control limits for Phase II.

Formula for individual observation

\[ T^2 = \frac{n(\bar{x} - \bar{\bar{x}})'S^{-1}(\bar{x} - \bar{\bar{x}})} \]

(11.19)

\[ T^2 = \frac{(x - \bar{x})'S^{-1}(x - \bar{x})} \]

(11.23)

\[ UCL = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1} \]
Phase II
(monitors future production)

\[
UCL = \frac{p(m+1)(n-1)}{mn - m - p + 1} F_{\alpha, p, mn - m - p + 1}
\]

LCL = 0

(11.21)

Note that the UCL in equation 11.21 is just the UCL in equation 11.20 multiplied by \( (m + 1)/(m - 1) \).
In general, the higher the $T^2$ value, the more distant the observation is for the mean.
Chi-Squared Control Chart

Time sequence is preserved (unlike the control ellipse)

The test statistic plotted on the chi-square control chart for each sample is

\[ \chi^2_0 = n(\bar{x} - \mu)' \Sigma^{-1}(\bar{x} - \mu) \]  \hspace{1cm} (11.12)

where \( \mu' = [\mu_1, \mu_2, \ldots, \mu_p] \) is the vector of in-control means for each quality characteristic and \( \Sigma \) is the covariance matrix. The upper limit on the control chart is

\[ \text{UCL} = \chi^2_{\alpha,p} \]  \hspace{1cm} (11.13)
Time sequence of the data is preserved by this control chart, so that runs or other nonrandom patterns can be investigated.

The “state” of the process is characterized by a single number. This is particularly helpful when there are two or more quality characteristics of interest.
Developed to provide more sensitivity to small shifts in univariate cases, but can be extended to multivariate quality control problems (a Phase II procedure).

Since the MEWMA with \( l = 1 \) is equivalent to the \( T^2 \) (or chi-square) control chart, the MEWMA is more sensitive to smaller shifts.

Defined as follows:

\[
Z_i = \lambda x_i + (1 - \lambda)Z_{i-1} 
\]  
(11.30)
Prabhu and Runger (1997) have provided a thorough analysis of the average run length performance of the MEWMA control chart, using a modification of the Brook and Evans (1972) Markov chain approach. They give tables and charts to guide selection of the upper control limit $UCL = H$ for the MEWMA.

Table 11.3 contains ARL performance for MEWMA for various values of $\lambda$ for $p = 2, 4, 6, 10, \text{and } 15$ quality characteristics.
Regression adjustments are often made to experimental data. Since randomization does not justify the models, almost anything can happen.

Each subject has two potential responses, one if treated and the other if untreated. Only one of the two responses is observed.

Regression estimates are generally biased, but the bias is small with large samples.

Adjustment may improve precision, or make precision worse; standard errors computed according to usual procedures may overstate the precision, or understate, by quite large factors. Asymptotic expansions make these ideas more precise.

David A Freedman. Statistics Department, University of California, Berkeley CA
**Monitoring multivariate processes** requires attention on two levels: process mean vector \( m \) and process variability.

Process variability is summarized by the \( p \times p \) covariance matrix \( \Sigma \).

The main diagonal elements of this matrix are the variances of the individual process variables, and the off-diagonal elements are the covariances.
In practice, $\Sigma$ usually will be estimated by a sample covariance matrix $S$, based on the analysis of preliminary samples.

$$\begin{align*}
\text{UCL} &= \left| \Sigma \right| \left( b_1 + 3b_2^{1/2} \right) \\
\text{CL} &= b_1 \left| \Sigma \right| \\
\text{LCL} &= \left| \Sigma \right| \left( b_1 - 3b_2^{1/2} \right)
\end{align*}$$

(11.36)
Representative Control Chart

Figure 11.13 presents the control chart. The values of $I S_i$ for each sample are shown in the last column of panel (c) of Table 11.1.

![Control Chart](image)

**FIGURE 11.13**  A control chart for the sample generalized variance, Example 11.2.
11.7 Latent Structure Methods

Methods for discovering the subdimensions in which the process moves about

- Principal Components
- Partial Least Squares
The principal components of a set of process variables $x_1, x_2, \ldots, x_p$ are just a particular set of linear combinations of these variables—say,

$$
z_1 = c_{11}x_1 + c_{12}x_2 + \cdots + c_{1p}x_p \\
z_2 = c_{21}x_1 + c_{22}x_2 + \cdots + c_{2p}x_p \\
\vdots \\
z_p = c_{p1}x_1 + c_{p2}x_2 + \cdots + c_{pp}x_p
$$

(11.37)

where the $c_{ij}$’s are constants to be determined. Geometrically, the principal component variables $z_1, z_2, \ldots, z_p$ are the axes of a new coordinate system obtained by rotating the axes of the original system (the $x$’s). The new axes represent the directions of maximum variability.
Shown here are two original variables $x_1$ and $x_2$, and two principal components: $z_1$ and $z_2$.

Note that the principal component $z_1$ accounts for most of the variability in the two original variables.
This graph illustrates three original process variables.

Most of the variability or “motion” in these two variables is in a plane, so only two principal components have been used to describe them.

In this picture, once again $z_1$ accounts for most of the variability, but a nontrivial amount is also accounted for by the second principal component $z_2$. 
Partial Least Squares

- Classifies the variables into x’s (inputs) and y’s (outputs).

- The goal is to create a set of weighted averages of the x’s and y’s that can be used for prediction of the y’s or linear combinations of the y’s.

- The procedure maximizes covariance in the same fashion that the principal component directions maximize variance.

- The most common applications of partial least squares today are in the chemometrics field.
In practice, many if not most, process monitoring and control scenarios involve several related variables.

Although applying univariate control charts to each individual variable is a possible solution, this is inefficient and can lead to erroneous conclusions.

Multivariate methods that consider the variables jointly are required.
References


