Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 51 (2007) 3862–3870
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

The quality control chart for monitoring multivariate autocorrelated processes

Jeffrey E. Jarrett[a,*], Xia Pan[b]

[a] University of Rhode Island, USA
[b] University of Illinois at Springfield, USA

Received 13 July 2005; received in revised form 15 January 2006; accepted 17 January 2006
Available online 30 March 2006

Abstract

Previously, quality control and improvement researchers discussed multivariate control charts for independent processes and univariate control charts for autocorrelated processes separately. We combine the two topics and propose vector autoregressive (VAR) control charts for multivariate autocorrelated processes. In addition, we estimate AR(p) models instead of ARMA models for the systematic cause of variation. We discuss the procedures to construct the VAR chart. We examine the effects of parameter shifts and by example present procedures to show the feasibility of VAR control charts. We simulate the average run length to assess the performance of the chart.

Keywords: Quality control and improvement; VAR control charts; Multivariate autocorrelated processes

1. Introduction and purpose

Shewhart (1931) developed statistical process control (SPC) techniques for the improvement and assurance of product and service quality. The assumptions of these methods include serial independence of successive samples taken from a process. Alwan and Roberts (1988) proposed modeling and filtering to monitor residuals by a one-step-forecast of autocorrelated process possessing systematic patterns. English and Sastri (1990) applied Kalman filter to monitoring univariate autoregressive (AR) (1) process. Wardell et al. (1994) derived the run length distribution of Shewhart chart on residuals. Montgomery and Mastrangelo (1991), Lu and Reynolds (1999) discussed some SPC charts for autocorrelated data. Runger and Willemain (1995) compared Shewhart charts based on several process transformations to eliminate autocorrelation. Snoussi et al. (2005) indicated that residuals control charts provide much better shift detection properties. Testik (2005) considered using an AR (1) process to modify control limits for residual based charts.

Loredo et al. (2002) provided a method for monitoring autocorrelated processes based on Hawkins (1991, 1993) regression adjustment. They examined the performance of residual based quality control charts in terms of the average run length to observation based control charts via Monte Carlo simulations. Their results indicate the superiority of residual based control in comparison with the observation based control charts when detecting a mean shift in short run

* Corresponding author. Tel.: +1 401 8744169; fax: +1 401 8744312.
E-mail addresses: jejarrett@mail.uri.edu (J.E. Jarrett), xpan1@uis.edu (X. Pan).

0167-9473/$ - see front matter
doi:10.1016/j.csda.2006.01.020

autocorrelated data. This is important because the i.i.d. assumptions of standard Shewhart control charts do not hold for large amounts of historical data.

Kalgonda and Kulkarni (2003, 2004) considered using the Hotelling T^2 statistic to monitor several quality characteristics. Due to difficulties of this statistic in determining the factors responsible for out of control signals when considering changes in the process mean vector and/or process covariance matrix, they proposed a new diagnostic, "the D-technique". Their analysis including a dummy variable multiple linear regression provided some evidence of usefulness of constructing multivariate control charts. Their second study (Kalgonda and Kulkarni, 2004) extended the usefulness of multivariate control charts in detecting an out of control status and identifying variable(s) responsible for the out of control situation.

Another study, Jaerkaporn et al. (2003) corroborated the usefulness of using average run length methods to the standard Shewhart control charts for individuals. Their method showed promise that average run length control charts are superior for detecting shifts in the mean than the traditional charts. Sliwa and Schmid (2005) extended the use of average run length to multivariate time series. Their study focused on the autocovariance and the cross-covariance structure of financial assets. Hence, monitoring the process control for new multivariate time series resulted from their approach. Finally, Cheng and Thaga (2005) produced a CUSUM control chart capable of detecting changes in the mean and standard deviation of autocorrelated data. Based on the average run length, they claim that their new control chart (MCAP) is useful for monitoring modern production processes where one produces high quality goods with a tiny fraction of nonconforming output units.

We propose to continue the study of multivariate autocorrelated processes by proposing a vector autoregressive (VAR) control chart based on ARL for precision manufactured products that are now produced *en masse* in high technology oriented industries. One example of such production is optical communication products manufacturing where the output is a multivariate autocorrelated time series. One bases the production of fiber optic on SiO_2 rod made from condensation of silicon and oxygen gases. The preparation of the SiO_2 rod needs to control temperature, pressure and the concentrations of different components. We find similar processes in semiconductors where they prepare materials. In these processes, correlated variables are plentiful and the addition of multivariate analysis is necessary to monitor the process. In addition, these variables are cross-correlated with time leads and lags. Univariate quality control charts are effective for monitoring individually correlated variables but are not effective for services, which are cross-correlated. We propose to introduce (VAR) control charts designed to monitor multivariate processes of the type described.

Shewhart control charts do not emphasize parameter estimation. This lack of concern emanates from using the rough criterion of average run length (ARL) to determine effectiveness. For autocorrelated processes, estimation is the key procedure for control chart construction. Some of the proposed control charts for autocorrelation are also not applicable if we ignore estimation. Hence, we will consider estimation issues associated with ARIMA modeling.

We arrange the remaining parts of this paper as follows: In Section 2, we describe the principles and design of VAR control charts. In Section 3, we examine the effects of parameter shifts. In Section 4, we show an example VAR chart construction for a multiple variable AR (3) process. In turn, we examine the ARL. Last, in Section 5, we summarize.

2. VAR control chart

We propose using a VAR control chart (another title is Multivariate Autoregression (MAR) control chart). The process is required either stationary or we can achieve stationarity with the appropriate filtering process, i.e. differencing. For a multivariate autoregressive process of n variables, denote $\vec{y}_t = (y_{1t}, y_{2t}, \ldots, y_{nt})'$ as a $(n \times 1)$ vector. These cross-correlated variables contain autocorrelation of order p. The process follows a VAR model

$$\vec{y}_t = \vec{c} + \left(\Phi_1 L + \Phi_2 L^2 + \cdots + \Phi_p L^p\right)\vec{y}_t + \vec{e}_t = \vec{c} + \Phi(L)\vec{y}_t + \vec{e}_t, \tag{1}$$

where L is the backshift operator, $\vec{c} = (c_1, c_2, \ldots, c_n)'$ is the constant vector, and $\vec{e}_t = (e_{1t}, e_{2t}, \ldots, e_{nt})'$ is the vector for the error term. Each Φ_j is a $(n \times n)$ coefficient matrix for lag j. $\vec{e}_t$ contain no autocorrelation but does contain cross-sectional correlation. That is, $E\left(\vec{e}_t\vec{e}_t'\right) = \Omega$ is invariant about time but may not be a diagonal $(n \times n)$ matrix. Some elements of the matrix Φ_j are zero when variables are uncorrelated.

In this model, the AR coefficient matrix, $\Phi = \left(\Phi_1\Phi_2 \cdots \Phi_p\right)$, and the constant term $\bar{c}$ represent systematic patterns. The error term vector e_t represents nonsystematic disturbances. We monitor the process to filter the systematic element. In turn, we monitor the special-cause. To filter out and estimate the systematic element for a process that is in-control