行政院國家科學委員會專題研究計畫　成果報告

製程 Profiles 之監控--應用無母數迴歸方法

計畫類別：個別型計畫
計畫編號：NSC93-2118-M-009-007-
執行期間：93 年 08 月 01 日至 94 年 07 月 31 日
執行單位：國立交通大學統計學研究所

計畫主持人：洪志真

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 94 年 10 月 31 日

中文摘要

在品管上，一般而言，製程或產品之品質特性都是一個變數。然而對某些製程而言，品質特性是由反應變數和解釋變數間之關係來界定。因此一個品質特性乃以一個函數、或是一個曲線之資料型式來呈現，稱之為profile (縱斷面或剖面)。本計畫主旨在於研究探討如何有效地監控製程profiles。在本報告中，對品質特性為一個函數或一條曲線的製程，我們提出一些新的控制圖來做監控；而除了平滑的假設外，在此不限定曲線的形式。我們提出了三個新的監控方法：(1)用無母數迴歸方法“B-spline”對曲線做配適，再利用T^2控制圖對製程做監控；(2)用典型的控制圖EWMA、EWMSD或R控制圖對樣本曲線和參考曲線間的殘差做監控；(3)用事先對樣本曲線和參考曲線間的差別所制定的測度(metrics)對品質特性曲線做監控。我們利用模擬，以平均連串長度(ARL)來比較所有方法。結果顯示，我們所提出的新方法表現都相當良好，尤以EWMA型之控制圖一般而言表現最優。

關鍵詞：EWMA管制圖、多變量T^2控制圖、EWMSD管制圖、平均連串長度、平滑方法、Spline 迴歸、B-Spline、SPC

英文摘要

In this project, we propose and study some control charts for monitoring processes in which the quality of a product item is characterized by a profile or a function. No assumptions are made on the functional form of the profiles except that profiles are smooth curves. Three approaches of monitoring schemes are considered: (i) use “B-splines” to fit each of the sample process profiles and design a T^2 chart accordingly to monitor process profiles; (ii) use typical control charts－the EWMA, EWMSD, and R chart－to monitor the “residuals” of each sample profile from the in-control reference profile; and (iii) monitor some metrics defined for measuring the deviation of each sample profile from the in-control reference profile. Performances of the proposed schemes are evaluated and compared in terms of the average run length via simulation studies. All proposed approaches appear to perform reasonably well with the EWMA-type control charts outperforming others in general.

Keywords: EWMA Control Chart, Multivariate T^2 Control Chart, EWMSD Control Chart, Average Run Length, Smoothing Techniques, Spline Regression, B-Spline, Statistical Process Control

報告內容

一、前言

Statistical process control (SPC) has been successfully proven useful for quality and productivity improvement in many domains, especially in industries. For most of SPC applications, the quality of a process or product is measured by one or multiple quality characteristics. However, some processes are better characterized by profiles or functions. Kang and Albin (2000) described an example of aspartame (an artificial sweetener), in which the quality is characterized by the amount that dissolves per liter of water at different temperatures. However, the real example considered in Kang and Albin (2000) is a semiconductor manufacturing application involving the calibration of a mass flow controller in which the performance of the process is characterized by a linear function. For other examples, see Kim *et al.* (2003) and the papers cited therein.

Some profile monitoring methods have been proposed for this type of processes in the literature. Kang and Albin (2000) and Kim *et al.* (2003) proposed several methods for monitoring the process in which the quality is characterized by a linear profile.

Note that these two papers address only *liner* profiles. For methods to adapt to more applications such as the aspartame example in Kang and Albin (2000), it is desirable to relax the restriction of linearity. To model profiles with no restriction on the functional form, it is natural to consider the nonparametric regression approach in which the unknown function is only assumed to be smooth.

二、研究目的

The main purpose of this project is to propose and study some new monitoring schemes for profiles of more flexible shapes so that the schemes can be applied to more general and practical situations. In this project, we apply spline regression for profile modeling and then develop monitoring schemes accordingly.

三、文獻探討

Kang and Albin (2000) presented two approaches to monitoring linear profiles. The first approach uses a multivariate T^2 control chart to monitor the profile parameters, slope and intercept, simultaneously. The second approach treats the “residuals” of a sample profile－defined as the deviations from the reference profile at some set values (X-values)－as a rational subgroup and uses a combined EWMA/R (exponentially weighted moving average/range) chart for profile monitoring. Kim *et al.* (2003) presented another approach: for each profile, code the X-values of a profile by centering so that the estimators of the *Y*-intercept and the slope of the regression line are independent; then construct two two-sided EWMA charts to monitor the *Y*-intercept and slope separately and a one-sided EWMA chart to monitor the process variation. They combined these three charts and called this scheme EWMA_3.

Walker and Wright (2002) proposed additive models to assess the sources of variation of density profiles of particleboards. Jin and Shi (2001) used wavelets to monitor and diagnose process faults. Gardner *et al.* (1997) use some spatial signature metrics defined for measuring the deviation of the observed profile from the reference profile to diagnose the equipment faults. They reported that these metric-related charts are very powerful in detecting standard deviation shifts. In industrial applications other than process monitoring, Miller (2002) and Nair *et al.* (2002) analyzed designed experiments with responses being linear functions.

As to the profile monitoring, many nonparametric regression estimation methods are available, including, for example, the popular kernel estimation, smoothing splines, local polynomial regression, and spline regression. For nonparametric regression, readers are referred to, for example, the books by Wahba (1990), Hardle (1990, 1991), Hastie and Tibshirani (1990), Green and Silverman (1994), Simonoff (1996), Eubank (1999), and papers cited therein.

In this project, for the profile modeling, we adopt spline regression as the curve-fitting/smoothing technique for its simplicity and readiness of a direct extension from the simple linear regression used by Kang and Albin (2000) and Kim *et al.* (2003) in constructing the T^2 chart and their residuals charts.

四、研究方法

Three different approaches are proposed and studied in this study. The first approach is a T^2-chart extension of Kang and Albin (2000). The second approach is a monitoring scheme based on the “residuals” of sample profiles, similar to the combined EWMA/R scheme by Kang and Albin (2000) and the combined EWMA_3 scheme by Kim *et al.* (2003). We use separately the EWMA chart for detecting the mean shift, R chart or exponentially weighted moving standard deviation (EWMSD) chart for variation changes. The third approach is to use some metrics defined for measuring the deviation of the observed profile from the reference profile to detect process changes. In this project, for the profile modeling, we adopt spline regression as the curve-fitting/smoothing technique for its simplicity and readiness for a direct extension from the simple linear regression used by Kang and Albin (2000) and Kim *et al.* (2003) in constructing the T^2 chart and their residuals charts.

Simulation studies are conducted to investigate the effectiveness of the proposed methods and also to compare their performances in terms of the average run length (ARL). Profiles mimicking the aspartame profiles mentioned before are used as an illustrative example.

五、結果與討論

In order to extend linear profiles to smooth profiles of any shapes, a smoothing technique is needed for de-noising sample profiles. We adopt the B-spline regression method for its popularity and simplicity in this project. Simply put, the B-spline regression is just a multiple linear regression with B-spline bases. See de Boor (1978) for the definition of the B-spline basis.

Denote $B_{l,k}$ the l^{th} B-spline basis of order k, $l = 1, 2, ..., b$. $B_{l,k}$ can be defined iteratively by

$$B_{l,1}(t) = \begin{cases} 1 & \text{for } t_l \le t < t_{l+1}, \\ 0 & \text{for } t < t_l \text{ or } t \ge t_{l+1}; \end{cases}$$

$$\text{For } k \ge 2, \quad B_{l,k}(t) = \frac{t - t_l}{t_{l+k-1} - t_l} B_{l,k-1}(t) + \frac{t_{l+k} - t}{t_{l+k} - t_{l+1}} B_{l+1,k-1}(t).$$

Note that $B_{l,k}$ is nonzero only on the interval (t_l, t_{l+k}).

For B-spline regression, we consider the following linear model:

$$y_i = \sum\nolimits_{l=1}^{b} c_l B_{l,k}(x_i) + \varepsilon_i, \quad i = 1, ..., n. \qquad (1)$$

where ε_i 's are i.i.d. normal variates with mean zero and common variance $\sigma^2 > 0$.

Given a set of data $\{(x_i, y_i), i = 1, ..., n\}$, the spline regression method finds the best spline approximation via the following least squares regression:

$$\min_{\mathbf{c}} \sum_{i=1}^{n} \{y_i - \sum_{l=1}^{b} c_l B_{l,k}(x_i)\}^2, \qquad (2)$$

where $\mathbf{c} = (c_1, ..., c_b)'$. The least squares estimator of $\mathbf{c}$ is

$$\hat{\mathbf{c}} = (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{y}, \qquad (3)$$

where $\mathbf{y} = (y_1, ..., y_n)'$, $\hat{\mathbf{c}} = (\hat{c}_1, ..., \hat{c}_b)'$, and $\mathbf{B}$ is the $n \times b$ design matrix with the $(i, l)^{\text{th}}$ element $B_{l,k}(x_i)$, $l = 1, ..., b$, $i = 1, ..., n$. Then, under model (1), $\hat{\mathbf{c}}$ has a multivariate normal distribution with mean vector $\mathbf{c}$ and variance-covariance matrix $\mathbf{\Sigma} = \sigma^2(\mathbf{B}'\mathbf{B})^{-1}$.

In Phase II of process monitoring, it is usually assumed that the in-control reference profile is *known* (or has been estimated from a set of historical data and treated as *known*). Denote it by $f(x)$. We now establish a B-spline representation (or approximation) of the reference profile. First obtain $\{f(x_i), i = 1, ..., n\}$, the values of the reference profile at the n set X-values $\{x_i, i = 1, ..., n\}$. Let $\mathbf{c}$ be the least squares solution of (2) with y_i replaced by $f(x_i)$. More specifically, $\mathbf{c} = (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{f}$, where $\mathbf{f} = (f(x_1), ..., f(x_n))'$. We then treat this vector $\mathbf{c}$ as the "true" in-control B-spline coefficient vector. For each newly observed sample profile $\{(x_i, y_i), i = 1, ..., n\}$, compute its B-spline coefficient vector $\hat{\mathbf{c}}$ by Equation (3). Since σ^2 is also assumed "known" in Phase II, the T^2 statistic of the sample is given by $T^2 = (\hat{\mathbf{c}} - \mathbf{c})'\mathbf{\Sigma}^{-1}(\hat{\mathbf{c}} - \mathbf{c})$. When the process is in control, T^2 is distributed as the Chi-square distribution with b degrees of freedom. Thus, the upper control limit in Phase II is $\chi^2_{b,\alpha}$.

In Phase I, assume that a set of historical data containing k sample profiles $\{(x_i, y_{ij}), i = 1,...,n, j = 1,...,k\}$ is available. Let $\mathbf{y}_j = (y_{1j}, ..., y_{nj})', j = 1,...,k.$ For $j = 1,...,k,$ perform the least squares regression on the jth sample profile to obtain $\hat{\mathbf{c}}_j$ and $MSE_j = \left\| \mathbf{y}_j - \mathbf{B}\hat{\mathbf{c}}_j \right\|^2 /(n-b)$, where $\left\| \bullet \right\|$ denotes the Euclidean norm of R^n. Then the $n \times 1$ vector **f** representing the reference curve can be estimated by $\hat{\mathbf{y}} = \mathbf{B}\tilde{\mathbf{c}}$ with

$$\tilde{\mathbf{c}} = \sum\nolimits_{j=1}^{k} \hat{\mathbf{c}}_j / k \quad \text{and} \quad \hat{\sigma}^2 = MSE = \sum\nolimits_{j=1}^{k} MSE_j / k. \qquad (4)$$

The T^2 statistic of the jth sample profile is then modified by

$$T^2_{0j} = \frac{k}{k-1}(\hat{\mathbf{c}}_j - \tilde{\mathbf{c}})'\mathbf{S}^{-1}(\hat{\mathbf{c}}_j - \tilde{\mathbf{c}}),$$

where $\mathbf{S} = \hat{\sigma}^2 (\mathbf{B}'\mathbf{B})^{-1}$ is an unbiased estimator of $\mathbf{\Sigma}$. It can be easily shown that T^2_{0j}/b has F distribution with degrees of freedom b and $(n-b)k.$ Thus, the upper control limit of the T^2 chart in Phase I is $bF_{b,(n-b)k,\alpha}$.

The second approach we consider in this project is to use the EWMA and the R (or EWMSD) chart to monitor the "residual" average and the range (or standard deviation), respectively. The regression "residuals" of the jth sample profile are defined as

$$e_{ij} = y_{ij} - f(x_i), \quad i = 1, 2, ..., n.$$

Then the EWMA and the R charts for Phase II can be constructed by the same way as that of Kang and Albin (2000). Similarly, to construct the EWMA and R charts for Phase I, just replace the σ in the control limits by its estimate $\sqrt{MSE}$ given in (4).

We can also use exponentially weighted moving standard deviation (EWMSD) control chart to monitor process variation. The exponentially weighted moving averaging feature makes this chart very sensitive in detecting shifts in σ, especially when shifts are relatively small. Define the sample standard deviation of the residuals of the jth sample profile as

$$s_j = \left\| e_j \right\| / \sqrt{n-b} = \left\| \mathbf{y}_j - \hat{\mathbf{y}}_j \right\| / \sqrt{n-b},$$

where $\hat{\mathbf{y}}_j = \mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{y}_j, \quad j = 1, 2, ...$

The EWMSD statistic is given by

$$v_j = \theta s_j + (1-\theta)v_{j-1},$$

where θ $(0 < \theta \le 1)$ is a smoothing constant and the initial value v_0 is the in-control value of $c_5\sigma$, where $c_5 = \frac{\Gamma((n-b+1)/2)}{\Gamma((n-b)/2)}\sqrt{\frac{2}{n-b}}$. In Phase II, σ is assumed known. The control limits for the EWMSD chart are $c_5\sigma \pm L'\sigma\sqrt{\frac{\theta}{2-\theta}(1-c_5^2)}$ and L' is again a constant chosen to give a pre-specified in-control ARL. In Phase I, the control limits are modified by substituting σ with $\sqrt{MSE}$.

The third approach is to use some metrics to monitor the process. For our profile-monitoring problem, we select the following two metrics to study:

$$M1 = \frac{1}{n}\sum_{i=1}^{n} | g(x_i) - f(x_i) | \quad \text{(averaged-absolute-value-deviation metric),}$$

$$M2 = \frac{1}{n}\sum_{i=1}^{n} \{g(x_i) - f(x_i)\}^2 \quad \text{(averaged-squared-error metric),}$$

where g is the fitted B-spline of a newly observed profile and f is the reference profile. The charts constructed based on these two metrics will be called M1 and M2 charts, respectively. To construct the control charts, we need to find the critical value of the null distribution of each metric.

Since the distributions of these metrics are difficult to obtain, we use a set of historical data of k sample profiles to characterize the in-control process during Phase I. We fit a B-spline to each of the k sample profiles. Obtain the average of the k regression estimates of the B-spline coefficient vector, $\tilde{\mathbf{c}} = (\tilde{c}_1, ..., \tilde{c}_b)'$, and MSE by (4). Let the estimated reference profile be $\sum_{l=1}^{b} \tilde{c}_l B_{l,k}(x)$ and $\hat{\sigma}^2 = MSE$. Simulate M sets of n observations from the following model:

$$y_i = \sum\nolimits_{l=1}^{b} \tilde{c}_l B_{l,k}(x_i) + \varepsilon_i, \quad i = 1, ..., n, \tag{5}$$

where ε_i are i.i.d. $N(0, \hat{\sigma}^2)$. For each metric, compute the metric value for each of the M simulated profiles and let the $100(1-\alpha)^{\text{th}}$ percentile of these M metric values be the critical value. The process is claimed out of control when the metric of the newly observed profile is greater than the critical value. We choose M = 50,000 in our study.

We evaluate the performances of these approaches in terms of ARL via simulation studies. Assume the underlying reference profile is known in Phase II. Denote the in-control ARL value by ARL_0. All charts are designed to have the same ARL_0 = 200, which corresponds to $\alpha = 0.005$. The smoothing constant θ may affect the ARL performances of the EWMA and EWMSD charts. Details can be seen in Lucas and Saccucci (1990). For simplicity, in this study, the smoothing constant is set to 0.2.

We consider the following model for the in-control process profiles:

$$Y = I_0 + M_0 e^{-N_0(x-1)^2} + \varepsilon, \quad x \in [0,4],, \tag{6}$$

where $\varepsilon \sim N(0,\sigma^2)$. In our simulation study, $\sigma = 1$ and the in-control reference profile is $f(x) = 1 + 15e^{-(x-1)^2}$. Four different types of shifts are considered in the simulation study: *I*-shift, *M*-shift, *N*-shift, and error-variance-increase.

For smoothing data, we use the equidistant sequence (-1.2, -0.8, -0.4, 0, 0.4, 0.8, 1.2, 1.6, 2, 2.4, 2.8, 3.2, 3.6, 4, 4.4, 4.8, 5.2) as the knots with order k = 4 so that the number of B-spline bases is b = 13. We choose equidistant x_i-values of 0, 0.08,…, 3.92 for n = 50 in our simulation. Approximate the in-control exponential profile by spline regression with these 13 B-spline bases.

It is well known that boundary effects are a potential problem in smoothing methods. Loosely speaking, boundary effects mean that the fittings at the neighboring area of the boundaries usually are not as good as the fittings at the interior points. We use a simple simulation study to illustrate the boundary effect that we encounter in our study. Let $n = 50, b = 13, k = 4,$ and $\sigma = 1$. Generate 50,000 in-control profiles from model (6) with $I_0 = 1, M_0 = 15,$ and $N_0 = 1$. Compute the B-spline coefficients $\hat{c}_l$'s for each profile. For $l = 1,...,b,$ average the 50,000 replications of $\hat{c}_l$ and denote the average by $\bar{\hat{c}}_l$. The simulation is repeated three times and the results show that the two largest deviations occur at the last (i.e., 13th) and the first coefficient in all three simulations. This apparently demonstrates the boundary effects in our context. Unfortunately this causes a big problem: the poor estimation at the boundaries leads to a potentially larger T^2 statistic for each of the generated profiles and thus results in a smaller ARL_0 than expected. In order to achieve $ARL_0 = 200$, we thus choose to omit the first and the last coefficients in constructing the T^2 statistic. Consequently, the control limit of the T^2 chart is adjusted to $\chi^2_{11,0.005}$.

We now evaluate the performances of the charts proposed in this project. For mean shifts, such as *I*, *M*, or *N*-shifts, we evaluate and compare the T^2, EWMA, M1, and M2 charts. For shifts in the error variance, we compare the T^2, R, M1, M2, and EWMSD charts. These ARL values are calculated based on 100,000 replications of simulated run lengths. Results show that the EWMA chart outperforms the others for all shifts in process "mean", while the EWMSD chart outperforms the others for shifts in process "variation". This also demonstrates that the EWMA-type charts are indeed more sensitive than others, especially for smaller shifts.

To investigate the effect of the number of set points n and the number of B-spline bases b, a simulation study with n = 20, 30, 40, 50 and b = 5, 9, 13 is conducted. We observe that the detecting power of all the charts increases as the number of set points n increases, which is expected. It is interesting to note that the detecting power of the T^2 chart increases as parameter b decreases. However, it is found that the B-spline estimate does not approximate the reference profile well enough with b = 5. Thus, b = 5 may not be a good choice for the profile function under study. For practical applications, we simply suggest selecting an appropriate *b* from a set of candidates by visual evaluation of the fittings of the in-control reference profile. Choose a small *b* with a reasonable fit.

To see if the nice flexibility property enjoyed by our methods would need to pay some price for it, we compare our methods with the methods proposed by Kang and Albin (2000) and Kim *et al.* (2003) when monitoring linear profiles. The linear profile model used by Kang and Albin (2000) is $y_{ij} = 3 + 2x_i + \varepsilon_{ij}$, where ε_{ij} are i.i.d. $N(0,1)$ with fixed x_i -values of 2, 4, 6, and 8. Our ARL comparisons show that, even for linear profiles, our nonparametric residuals charts are more effective than the methods of Kang and Albin (2000) and Kim *et al.* (2003) in detecting shifts for most of cases and stay fairly competitive for the rest of cases. And most importantly, our methods are more general in applications than theirs. The main reason for this is that our nonparametric regression model can model linear profiles quite well with only a slight tradeoff of efficiency when compared to parametric regression methods.

This study extends the framework of statistical process control to more general applications. However, for many applications, some models other than the model studied in this project may be more appropriate. For example, in-control profiles need not have the same mean function. Loosely speaking, some variations are allowed as part of the common-cause variation. How to develop suitable profile monitoring schemes for such processes is currently under study.

The profile monitoring is a useful SPC technique and a promising area of research. More statistical methods, models, and ideas are needed. Curve data analysis techniques given in Ramsay and Silverman (1997) may be useful here.

參考文獻

1. de Boor, C. (1978). *A Practical Guide to Splines*. Springer, Berlin.
2. Eubank, R. (1999). *Nonparametric Regression and Spline Smoothing* 2nd Edition. Marcel Dekker, New York .
3. Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications.* Chapman & Hall, New York.
4. Gardner, M. M., Lu, J. C., Gyurcsik, R. S., Wortman, J. J., Hornung, B. E., Heinisch, H. H., Rying, E. A., Rao, S., Davis, J. C., and Mozumder, P. K. (1997). “Equipment Fault Detection Using Spatial Signatures”. *IEEE Transactions on Components, Packaging, and Manufacturing Technology – Part C,* 20, 295-304.

5. Green, P. J. and Silverman, B. W. (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach.* New York: Chapman and Hall, New York.

6. Hardle, W. (1990). *Applied Nonparametric Regression.* Cambridge University Press, New York.

7. Hardle, W. (1991). *Smoothing Techniques with Implementation in S.* Springer-Verlag, New York.

8. Hastie, T. and Tibshirani, R. (1990). *Generalized Additive Models.* Chapman and Hall, New York.

9. Jin, J. and Shi, J. (2001). "Automatic Feature Extraction of Waveform Signals for In-Process Diagnostic Performance Improvement". *Journal of Intelligent Manufacturing* 12, pp. 257-268.

10. Kang, L. and Albin, S. L. (2000). "On-Line Monitoring When the Process Yields a Linear Profile". *Journal of Quality Technology* 32, pp. 418-426.

11. Kim, K., Mahmoud, M. A., and Woodall, W. H. (2003). "On The Monitoring of Linear Profiles". *Journal of Quality Technology* 35, pp. 317-328.

12. Lucas, J. M. and Saccucci, M. S. (1990). "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements". *Technometrics* 32, pp. 1-29.

13. Miller, A. (2002). "Analysis of Parameter Design Experiments for Signal-Response Systems". *Journal of Quality Technology* 34, pp. 139-151.

14. Nair, V. N., Taam, W., and Ye, K. Q. (2002). "Analysis of Functional Responses from Robust Design Studies". *Journal of Quality Technology* 34, pp. 355-370.

15. Ramsay, J. and Silverman, B. (1997) Functional Data Analysis. Springer-Verlag, New York.

16. Simonoff, J. S. (1996). *Smoothing Methods in Statistics*. Springer-Verlag, New York.

17. Wahba, G. (1990). *Spline Models for Observational Data.* CBMS-NSF Series. SIAM, Philadelphia.

18. Walker, E. and Wright, S. P. (2002). "Comparing Curves Using Additive Models". *Journal of Quality Technology* 34, pp. 118-129.

計劃成果自評

本計劃之執行相當順利，結果亦相當豐碩，三位碩士班學生和兩位博士班學生在工業統計上得到相當不錯的訓練。研究內容與原計畫相符程度極高，亦已達成預期目標。研究成果預計將有一篇論文可以在國際知名期刊發表，目前已撰寫完成即將投稿。本研究所提出之 profile 控制圖在學術上和應用上均有貢獻。