行政院國家科學委員會專題研究計畫成果報告

模糊集群法在退化模式上之應用
Application of Fuzzy Clustering Method in Degradation Models

計畫編號: NSC 88-2118-M-032-009
執行期限: 87年8月1日至88年7月31日
主持人: 吳碩傑 淡江大學統計學系
共同主持人: 蔡宗儒 淡江大學統計學系

1 中文摘要

隨著科技的進步, 許多壽命試驗在結束時常常只收集到很少產品壽命資料, 甚至可能沒有一個測試元件發生故障. 在這種情況下, 我們便無法使用傳統的可靠度分析方法來評估產品的可靠度. 最近被討論的一種方法則是在壽命試驗的同時去測量導致產品故障的退化特徵, 而這些產品機能的退化量常包含許多可靠度的訊息. 一般說來, 我們必須建立產品退化量和時間的關係, 然後利用此關係來評估產品的可靠度. 具有隨機係數的非線性迴歸模式常被用來描述產品的退化路徑. 如果我們能得到此模式中參數的估計值, 那麼我們便可估計出產品的壽命分配. 在許多收集退化資料的試驗中, 我們常發現有少部分測試元件退化路徑的型態和大部分測試元件的退化路徑不相同. 因此, 在這個研究裡, 我們根據模糊集群法提出一個加權估計的方法, 來獲得比較穩定和準確的模式參數估計值, 並且將此種方法應用到真實的資料上.

關鍵詞: 最小平方估計式; 非線性混合效果模式; 最佳模糊集群法; 可靠度

Abstract

Some life tests are terminated with few or no failures. In such cases, a recent approach is to obtain degradation measurements of product performance which may contain some useful information about product reliability. Generally degradation paths of products are modeled by a nonlinear regression model with random coefficients. If we can obtain the estimates of parameters under the model, then the failure time distribution can be estimated. In some cases, the patterns of a few degradation paths are different from those of most degradation paths in a test. Therefore, this study develops a weighted method based on fuzzy clustering procedure to robust estimation of the underlying parameters. The method will be tested on real data.

Keywords: Least squares estimator; Nonlinear mixed-effect model; Optimal fuzzy clustering method; Reliability

2 Introduction

With today's high technology, manufacturers face increasingly intense global competition.

To remain profitable, they are challenged to design, develop test and produce high reliability products. This results in very few or no failures, so that the traditional approach of life testing with censoring is no longer effective in assessing product reliability. Products such as insulations, semiconductors and electrical devices are the cases.

To study this type of reliability data, a recent approach is to obtain degradation measurements of product performance over time. To conduct a degradation test, one has to prespecify a level of degradation, and define that failure occurs when the amount of degradation for a test unit exceeds this level. Thus, degradation data provide sample paths of degradation as a function of time. In most reliability tests, degradation data have some important practical advantages (e.g., see Nelson 1990, Chapter 11; Meeker and Escobar 1993; Meeker and Hamada 1995). In the literature, Lu and Meeker (1993) considered a nonlinear mixed-effects model and used a two-stage method to obtain estimates of the percentiles of failure time distribution. Tang and Chang (1995) modeled nondestructive accelerated degradation data from power supply units as a collection of stochastic processes. Tseng, Hamada and Chiao (1995) used a simple linear regression with random coefficients to model the luminosity degradation, which is a quality characteristic of fluorescent. They used a fractional factorial design to improve the reliability of fluorescent lamps. Lu, Park and Yang (1997) proposed a model with random regression coefficients and standard-deviation function for analyzing linear degradation data from semiconductors. Wu and Shao (1999) established the asymptotic properties of the least squares estimators for the degradation measurements under nonlinear mixed-effects model.

In a degradation test, performance is obtained as it degrades over time and different units may have different performances. Thus, the general approach is to model the degradations of the individual units using the same functional form and differences between individual units using random effects. The model is:

$$\begin{aligned} y_{ij} &= \eta(t_{ij};\boldsymbol{\alpha},\boldsymbol{\beta}_i) + \varepsilon_{ij}, \\ & \qquad i = 1,\ldots,n, \quad j = 1,\ldots,m_i, \quad (1) \end{aligned}$$

where y_{ij} is the degradation amount of the ith unit at time t_{ij}; t_{ij} is time of the jth measurement for the ith unit; $\boldsymbol{\alpha}$ is a $p \times 1$ vector of fixed effects which describe population characteristics; $\boldsymbol{\beta}_i$ is an $r \times 1$ vector of the ith unit random effects which represent an individual unit's characteristic; ε_{ij} is the random error. Assume that ε_{ij} are $i.i.d.$ with mean 0 and variance σ_ε^2, $\boldsymbol{\beta}_i$'s are independently distributed as $\pi(\boldsymbol{\beta}|\boldsymbol{\phi})$, with a known function π and an unknown $q \times 1$ parameter vector $\boldsymbol{\phi}$, and $\{\varepsilon_{ij}\}$ and $\{\boldsymbol{\beta}_i\}$ are independent. Under degradation model, one has to get the estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\phi}$ in order to estimate the percentiles of failure time distribution. One of the estimation methods is the two-stage method proposed by Lu and Meeker (1993). However, the estimation results of two-stage method may be inadequate in the following two kinds of data in degradation tests. (1) The patterns of some sample degradation paths are different from those of most degradation paths (e.g., see the data of metal film resistor in Wu and Shao (1999)). (2) Some failures occur in a test, but the degradation amounts of these failed units can not be observed any more.

In this study, we are going to analyze the above two kinds of degradation data. One possible approach is to delete the data of these test units which failed and can not be observed any more during degradation test, or which have different patterns of degradation paths. However, if the number of test units is small, we may lose some information by deleting data. In order to avoid losing information, we may treat these special test units as outliers in some senses. Thus, we can use the optimal fuzzy clustering method (e.g. see Gath and Geva 1989; Van Cutsem and Gath 1993; Wu, Jang and Tsai 1996) to obtain the weights for these units, and then we can get the weighted estimates of the parameters in the degradation model. This procedure may provide more accurate results than the two-stage method. We will examine the performance of proposed procedure in real data problems.

3 The Optimal Fuzzy Clustering Analysis Method

In a degradation test, the relationship between degradation measurement and time is modeled by a known function, and a random effect term is used to describe an individual product unit's characteristic. Hence, the sample degradation paths are expected to have the same pattern. However, in some situations, the patterns of some sample degradation paths are different from the patterns of most sample degradation paths. The metal film resistor example in Wu and Shao (1999) is the case. In that case only a few sample paths of test units have different patterns from the sample paths of the other units. We may treat these test units as outliers in some senses. That is, sample degradation paths with similar pattern can be grouped together. To avoid undesirable effects of different patterns on computing estimates of model parameters, a weighted estimation method is considered.

Gath and Geva (1989) provided an efficient clustering analysis method to identify groups among observations. This method is called the optimal fuzzy clustering analysis method. The advantages of this method are that each observation belongs to one group with a possibility between 0 and 1, and it provides a more stable clustering results than the fuzzy k-means clustering method. Van Cutsem and Gath (1993) provided a robust estimation based on the optimal fuzzy clustering analysis method. The simulation results in their paper showed that the fuzzy-weighted estimators are less sensitive to the influence of outliers. We will use the optimal fuzzy clustering analysis method to modify the two-stage estimation method proposed by Lu and Meeker (1993).

Under model (1), let $\hat{\boldsymbol{\theta}}_i = \begin{bmatrix} \hat{\boldsymbol{\alpha}}_i \\ \hat{\boldsymbol{\beta}}_i \end{bmatrix}$ be the estimates of $\begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta}_i \end{bmatrix}$, $i = 1, \ldots, n$ computed by using the first stage in Lu and Meeker (1993). Assume that $\hat{\boldsymbol{\theta}}_i$, $i = 1, \ldots, n$ can be grouped into q clusters, where $1 \leq q \leq n$. Let $\boldsymbol{V}_k$ be the kth cluster center and u_{ki} be the degree of membership that $\hat{\boldsymbol{\theta}}_i$ is in the kth cluster, $k = 1, \ldots, q$. Let n_k be the number of observations in the kth cluster. The steps of optimal fuzzy clustering analysis method are given as follows (Gath and Geva (1989)):

1. Carry out unsupervised tracking

of the initial set of cluster centers, $(\boldsymbol{V}_1, \ldots, \boldsymbol{V}_q)$.

2. Calculate the $q \times n$ weighted matrix $\boldsymbol{U}$ with entry

$$u_{ki} = \frac{1/d^2(\hat{\boldsymbol{\theta}}_i, \boldsymbol{V}_k)}{\sum_{l=1}^{q}[1/d^2(\hat{\boldsymbol{\theta}}_i, \boldsymbol{V}_l)]},$$

where $d^2(\hat{\boldsymbol{\theta}}_i, \boldsymbol{V}_k)$ is the Euclidean distance between $\hat{\boldsymbol{\theta}}_i$ and $\boldsymbol{V}_k$, $k = 1, \ldots, q$, $i = 1, \ldots, n$.

3. Compute the new set of cluster centers $(\hat{\boldsymbol{V}}_1, \ldots, \hat{\boldsymbol{V}}_q)$, where

$$\hat{\boldsymbol{V}}_k = \frac{\sum_{i=1}^{n} u_{ki}^2 \hat{\boldsymbol{\theta}}_i}{\sum_{i=1}^{n} u_{ki}^2},$$

$k = 1, \ldots, q$, and update the entries u_{ki}'s in $\boldsymbol{U}$ to $\hat{u}_{ki}$'s, according to step 2.

4. If $\max_{i,k} |u_{ki} - \hat{u}_{ki}| < e$ stop, else go to step 3. e is a number between 0 and 1.

For step 1, we need two statistics to do unsupervised fuzzy partition. In order to decide the initial subgroups and the optimal numbers of clusters for $\hat{\boldsymbol{\theta}}_i$, $i = 1, \ldots, n$, Gath and Geva (1989) proposed the performance measures – fuzzy hypervolume F_{HV} and average partition density D_{PA}. The fuzzy hypervolume is defined as

$$F_{HV} = \sum_{k=1}^{q} \sqrt{det(\boldsymbol{F}_k)},$$

where $det(\cdot)$ is the determinant of a square matrix and

$$\boldsymbol{F}_k = \frac{\sum_{i=1}^{n} u_{ki}(\hat{\boldsymbol{\theta}}_i - \boldsymbol{V}_k)(\hat{\boldsymbol{\theta}}_i - \boldsymbol{V}_k)^T}{\sum_{i=1}^{n} u_{ki}}.$$

The average partition density is defined as

$$D_{PA} = \frac{1}{q} \sum_{k=1}^{q} \frac{S_k}{\sqrt{det(F_k)}},$$

where $S_k = \sum_{\{i:\hat{\boldsymbol{\theta}}_i \in \chi_k\}} u_{ki}$ is called the sum of central members, and $\chi_k = \{\hat{\boldsymbol{\theta}}_i | (\hat{\boldsymbol{\theta}}_i - \boldsymbol{V}_k)^T \boldsymbol{F}_k^{-1} (\hat{\boldsymbol{\theta}}_i - \boldsymbol{V}_k) < 1\}$. Gath and Geva (1989) suggested that the optimal number of clusters in a data set is the number corresponding to global minimum in F_{HV} and global maximum in D_{PA}.

When the optimal fuzzy clustering method is performed, the degrees of membership in each classified cluster can be normalized to be the fuzzy weights w_{ki}, $i = 1, \ldots, n$, $k = 1, \ldots, q$, such that $\sum_{i=1}^{n} w_{ki} = 1$. After obtaining the fuzzy weights, we can use them to calculate the weighted estimates of model parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\phi}$.

4 Metal Fatigue-crack-growth Example

In this section, we will discuss the example of metal fatigue-crack-growth data from Lu and Meeker (1993). The nonlinear mixed-effect degradation model is:

$$y_{ij} = -\frac{1}{\theta_{2i}} \log(1 - 0.9^{\theta_{2i}} \theta_{1i} \theta_{2i} t_j) + \varepsilon_{ij}, \quad i = 1, \ldots, n, \;\; j = 1, \ldots, m_i,$$

where y_{ij} = log(observed crack length at time $t_j/0.9$), t_j is the measurement time (in million cycles) and $n = 21$ test units. The random effects (θ_1, θ_2) are from a bivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

We apply the proposed method in Section 3 to this data set. In this example, two clusters are identified. We can obtain the fuzzy weighted covariance matrices $S_{w1} = \sum_{i=1}^{n} w_{1i}(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\mu})(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\mu})^T$ and $S_{w2} = \sum_{i=1}^{n} w_{2i}(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\mu})(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\mu})^T$. Let $pw1 = \frac{det(S_{w1})}{det(S_{w1}) + det(S_{w2})}$, $pw2 =$

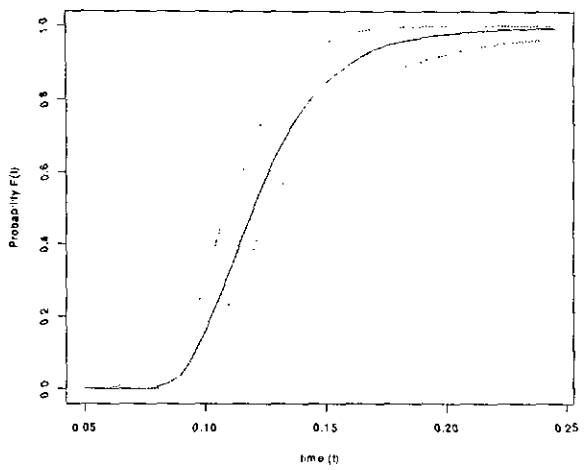

Figure 1: The point estimates and 90% confidence intervals for percentiles of the failure time distribution.

$\frac{det(S_{w2})}{det(S_{w1})+det(S_{w2})}$, $M_{bw1} = \sum_{i=1}^{n} w_{1i} Var(\hat{\boldsymbol{\theta}}_i)$, and $M_{bw2} = \sum_{i=1}^{n} w_{2i} Var(\hat{\boldsymbol{\theta}}_i)$. Therefore, we can estimate $\boldsymbol{\Sigma}$ by using $M_{aw} - M_{bw}$, where $M_{aw} = pw1 * S_{w1} + pw2 * S_{w2}$ and $M_{bw} = pw1 * M_{bw1} + pw2 * M_{bw2}$. The estimates of the model parameters are:

$$\hat{\boldsymbol{\mu}} = \begin{bmatrix} 3.732 \\ 1.571 \end{bmatrix}$$

and

$$\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} 0.60188 & -0.08029 \\ -0.08029 & 0.08701 \end{bmatrix}.$$

Figure 1 shows the point estimate of $F_T(t)$ and pointwise two-sided 90% bootstrap confidence intervals for $F_T(t)$. The confidence intervals were obtained by using the bootstrap simulation with $B = 1000$ and $N_B = 10000$.

References

Gath, I. and Geva, A. B. (1989). Fuzzy clustering for the estimation of the parameters of mixtures of normal distributions, *Pattern Recognition Letters*, **9**, 77-86.

Lu, C. J. and Meeker, W. Q. (1993). Using degradation measures to estimate a time-to-failure distribution, *Technometrics*, **35**, 161-174.

Lu, J.-C., Park, J. and Yang, Q. (1997). Statistical inference of a time-to-failure distribution derived from linear degradation data, *Technometrics*, **39**, 391-400.

Meeker, W. Q. and Escobar, L. A. (1993). A review of recent research and current issues in accelerated testing, *International Statistical Review*, **61**, 147-168.

Meeker, W. Q. and Hamada, M. (1995). Statistical tools for the rapid development & evaluation of high-reliability products, *IEEE Transactions on Reliability*, **44**, 187-198.

Nelson, W. (1990). *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*. Wiley, New York.

Tang, L. C. and Chang, D. S. (1995). Reliability prediction using nondestructive accelerated-degradation data: Case study on power supplies, *IEEE Transactions on Reliability*, **44**, 562-566.

Tseng, S.-T., Hamada, M. and Chiao, C.-H. (1995). Using degradation data to improve fluorescent lamp reliability, *Journal of Quality Technology*, **27**, 363-369.

Van Cutsem, B. and Gath, I. (1993). Detection of outliers and robust estimation using fuzzy clustering, *Computational Statistics & Data Analysis*, **15**, 47-61.

Wu, J.-W., Jang, J.-B. and Tsai, T.-R. (1996). Fuzzy weighted scaled coefficients in semi-parametric model, *Ann. Inst. Statist. Math.*, **48**, 97-110.

Wu, S.-J. and Shao, J. (1999). Reliability analysis using the least squares method in nonlinear mixed-effect degradation models, *Statistica Sinica*, **9**, 855-877.