ATLANTIS PRESS
Advances in Engineering Research, volume 93
International Symposium on Mechanical Engineering and Material Science (ISMEMS 2016)

Reliability adaptive modeling method for CNC Machine Tools based on kernel function

W. Jili[1,a*], Y. Dingyu[1, b], Z. Xiaocui[1, c], Y. Zhaojun[1, d], L. Guofa[1, e]

[1]College of Mechanical Science and Engineering, Jilin University, Changchun 130022, China

[a]wangjili1100@126.com, [b]yangdy15@mails.jlu.edu.cn,[c]zxc@jlu.edu.cn, [d]yangzj@126.com, [e]ligf@jlu.edu.cn

Keywords: Reliability of CNC machine tools, Kernel function distribution, Adaptive modeling, Gaussian distribution, Cross-validation method, Traversal search method, Goodness of fit test

Abstract. A reliability adaptive modeling method for CNC Machine Tools based on kernel function is proposed. Choosing Gaussian distribution as the kernel function, the optimal model with least integral square error as the objective is established, and cross-validation method and traversal search method are used to solve the bandwidth parameter of kernel function. In examples, we use kernel function distribution model to approximate different life distributions, verifying that the kernel function distribution model has good approximation property and can replace traditional distribution model. In the reliability modeling process of one type of CNC machine tools during the early failure period, comparing the kernel function distribution model with the traditional Weibull distribution model, we verify that the kernel function distribution model is more suitable for reliability modeling of CNC machine tools by goodness of fit test.

Introduction

At present, for CNC machine tools which is a complex system, reliability and method of simple life distribution or mixing distribution are commonly used [1-3]. Simple life distribution mainly includes exponential distribution, logarithmic normal distribution and Weibull distribution. By contrast, the shape of probability density curve of Weibull distribution is varied, thus it has good fitting performance and is commonly applied to reliability modeling of CNC machine tools. Due to the fact that the reliability of CNC machine tools is obviously influenced by factors such as use condition, work environment, maintenance condition and so on, and the failure modes are diverse, the distribution regularity of reliability data is complex, so it's easy to arise the problem of under-fitting using single distribution model for reliability modeling. Mixing distribution model includes mixing normal distribution, mixing Weibull distribution, etc. Compared to simple distribution model, the shape of probability density curve of mixing distribution model is more complex as well as various, so its fitting performance is better. But because that mixing distribution model has more model parameter, it's difficult to determine the mixing multiplicity and easy to arise the problem of over-fitting. Both simple distribution model and mixing distribution model have the problem of poor generalization performance, so it's difficult to confirm whether the model we choose can reflect the real distribution regularity of reliability data. But the reliability modeling method based on kernel function has the ability of adaptability. This paper use adaptive method to determine reasonable bandwidth parameter by choosing reasonable kernel function, and apply to reliability modeling of one type of CNC machine tools during the early failure period.

Kernel Function Distribution Model

Form of Kernel Function. Kernel function distribution model is a kind of important distribution model for fitting distribution regularity of complex data, which is mainly applied to fields such as pattern recognition analysis, loading spectrum analysis and so on [5, 6]. For one-dimensional reliability data sample, the general form of kernel function distribution model is shown as Eq. 1.

$$\begin{cases} f(x;h) = \frac{1}{n}\sum_{i=1}^{n} H(x-\hat{x}_i) \\ H(x) = h^{-\frac{1}{2}} G(h^{-\frac{1}{2}} x) \end{cases} \tag{1}$$

In Eq.1, $f(x;h)$ is the kernel function distribution model; $H(x)$ is the kernel function; $G(x)$ is the specific form of kernel function; n denote the sum of data sample $\hat{x}$; h denote the bandwidth of the kernel parameter.

Type of Kernel Function. During the data fitting process of kernel function distribution model, there is no need for traditional distribution test. But during the modeling process, reasonable type of kernel function should be chosen according to experience and the characteristics of data. Common standard kernel functions includes rectangular kernel function, trigonal kernel function, gaussian kernel function , logarithmic kernel function , quadratic kernel function, fourth-order kernel function and sixth-order kernel function [7]. The shape of distribution curve of different kernel functions are also largely different. The shape of distribution curves of above standard kernel functions is shown in Fig. 1.

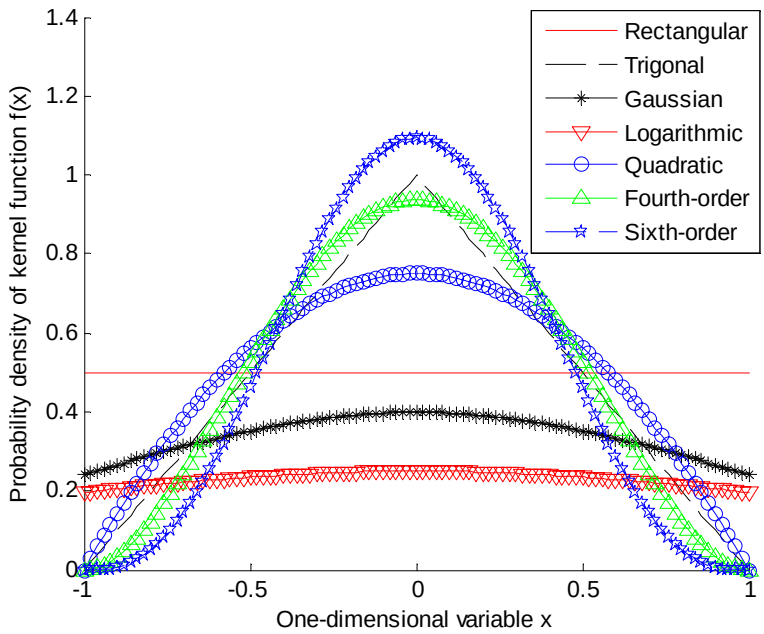

Fig. 1 Contrastive Analysis of Standard Kernel Functions.

Solving Adaptive Bandwidth Parameter

As shown in Eq. 1, the adaptive bandwidth parameter of kernel distribution model controls the shape of probability density curve of distribution model, so determining reasonable adaptive parameter is the most crucial factor of kernel distribution model. The solving of adaptive bandwidth parameter directly determines the fitting effect and influences the generalization performance.

Solving Method for Bandwidth. At present, the most common solving methods for bandwidth are adaptive solving methods such as cross-validation method, insertion method, etc [8, 9]. In order to generalization performance of kernel function distribution model, generally, an optimal model for determination of optimal bandwidth parameters is established with the minimum integral square error as the objective , which is shown as Eq. 2.

$$\begin{aligned} Min\ \ ISE\{\hat{f}(x;h)\} &= \int \left(\hat{f}(x;h) - f(x)\right)^2 dx \\ &= \int \hat{f}(x;h)^2 dx - 2\int \hat{f}(x;h) f(x) dx + \int f(x)^2 dx \end{aligned} \tag{2}$$

In Eq. 2, $\hat{f}(x;h)$ is the kernel distribution model; $f(x)$ is the theoretical probability density distribution function of sample data which is a constant and unique function model. The goal function $ISE\{\hat{f}(x;h)\}$ can be decomposed into three parts: $\int \hat{f}(x;h)^2 dx$, $2\int \hat{f}(x;h) f(x) dx$ and $\int f(x)^2 dx$. $\int \hat{f}(x;h) dx$ can be solved by numerical integration method; $\int \hat{f}(x;h) f(x) dx$ is the equivalent of the expectation of $\hat{f}(x;h)$; $\int f(x)^2 dx$ is a constant, and it is irrelevant to bandwidth parameter h, so it can be neglected when optimal solving h.

According to the Least squares cross-validation method proposed by researchers Redemo and Bowman, the key of solving goal function lies in the calculation of the expectation of $\hat{f}(x;h)$. Take the -th sample from the sample set as the observation sample, and take the rest samples as training sample to approximate , which is shown as Eq. 3.

$$\hat{f}\left(x_i, h\right) = \frac{1}{(n-1)h} \sum_{j \neq i}^{n} G\left(\frac{x_i - x_j}{h}\right) \tag{3}$$

According to Eq. 3, we can obtain the observed value of the expectation of $f\left(x_i\right)$, which is shown as Eq. 4.

$$\hat{E}\left(\hat{f}\left(x,h\right)\right) = \sum_{i=1}^{n} \hat{f}\left(x_i, h\right) \tag{4}$$

Analysis of Examples. In order to verify the effect of applying kernel function distribution to life distribution, this paper choose the illustrations of logarithmic normal distribution, three-parameter Weibull distribution, gamma distribution and two-folds Weibull distribution to verify the effect. Sample data with a sample size of 100 are generated respectively by inverse cumulative probability distribution method. The Gaussian distribution is chosen as the kernel function, and the traversal search algorithm is used to solve the bandwidth parameters in Eq. 2, so we can obtain theoretical value of parameters of each distribution model and the optimal bandwidth parameter, which are shown in Table 2. The corresponding results of illustration analysis are shown in Fig. 2.

Table 2. Illustration analysis table under different life distribution.

Illustration	Theoretical distribution	parameters of theoretical distribution	bandwidth parameter
Illustration 1	Logarithmic normal distribution	μ=2,σ =0.1	*h*=0.39
Illustration 2	Three-parameter Weibull distribution	γ=4,η=2,β=3	*h*=0.35
Illustration 3	Gamma distribution	η=3,β=2	*h*=1.29
Illustration 4	Two-folds Weibull distribution	ω1=0.4,η1=2,β1=3 ω2=0.6,η2=5,β2=3	*h*=0.54

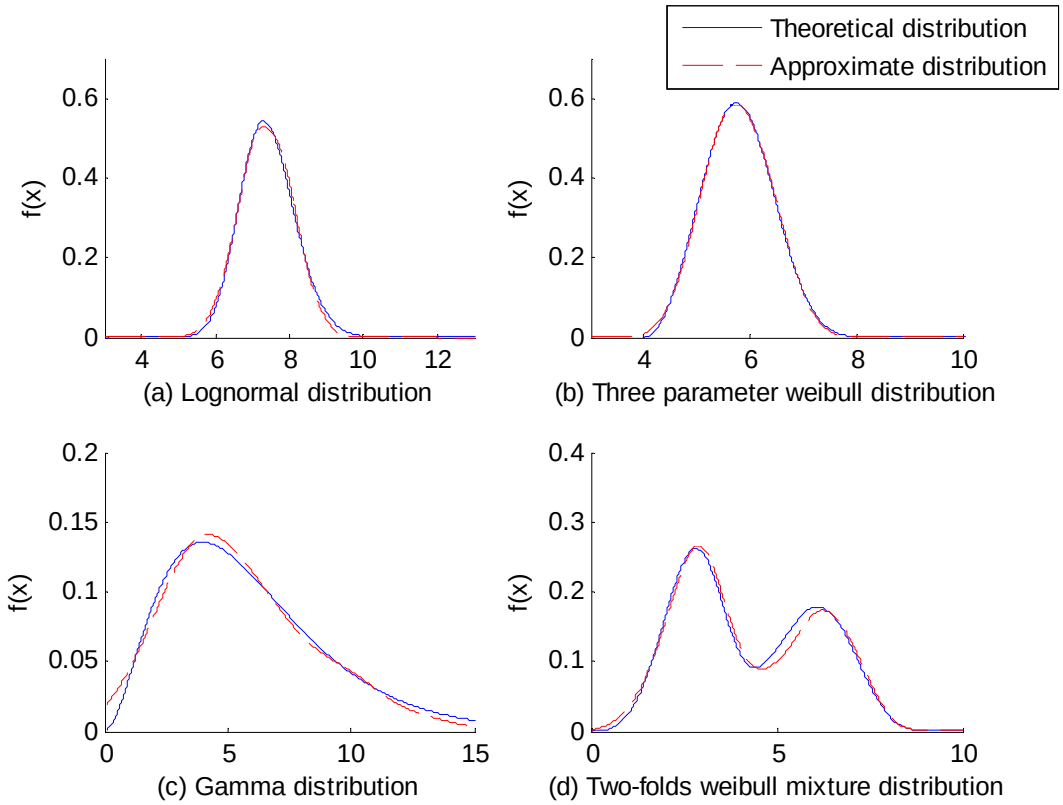

Fig. 2 Common approximate distribution figure of kernel function distribution model.

As shown in Fig. 2, comparing theoretical probability density curve of each distribution model with approximate probability density curve of kernel function distribution model, we can see that kernel function distribution model can well approximate the illustration model, so it indicates that kernel function distribution model can replace the common life distribution and apply to reliability data analysis.

Application Case of CNC Machine Tools

The life cycle of CNC machine tools is generally divided into early failure period, random failure period and late failure period. The regularity of early failure period is complex, so simple distribution model is difficult to reflect the actual distribution regularity of early failure data. This paper adopt kernel function distribution model to analyze the failure data of a certain kind of new domestic CNC machine tools. We obtain the complete sample data of time between failures of 10 CNC machine tools which are in early failure period by reliability field test. The concrete data are shown in Table 3.

Table 3. Values of Time between failures of a certain kind of CNC machine tools.

number	time between failures/h			
1	184	197	603	——
2	567	114	380	319
3	239	125	505	241
4	510	30	476	27
5	187	472	372	——
6	466	194	101	——
7	944	86	240	143
8	445	460	——	——
9	578	175	358	——
10	330	598	208	193

Solving Kernel Function Distribution Model. For the complete sample data of early failure period of CNC machine tools, taking Gaussian distribution as the kernel function, adopting traversal search algorithm for the solving of bandwidth parameter in formulate (2), wo obtain the bandwidth parameter. The reliability distribution regularity of early failure period of CNC machine tools is shown as the solid line in Fig. 3a. The imaginary line in Fig. 3a is Weibull distribution curve fitted according to sample data, and we can see that the probability density curves obtained by two distribution models are obviously different. Fig. 3 exhibits the reliability curve under kernel function model fitted according to sample data.

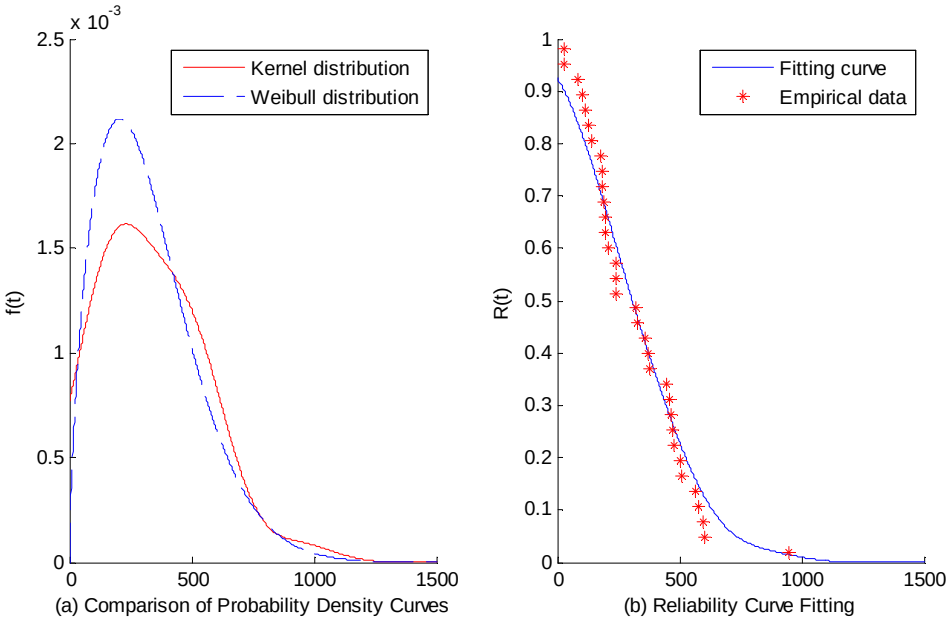

Fig. 3 Reliability distribution regularity analysis figure of CNC machine tools.

Goodness of Fit Test. In order to further verify the effectiveness of the kernel function modeling method, test the goodness of fit of kernel function model by ks test. Goodness of fit test is a statistical method to test whether a distribution of data from a population is consistent with a theoretical distribution. ks test centralizes on the inspection of statistic $D_{\max}$, that is:

$$D_{\max} = \max\left|S_n(x) - F_0(x)\right| \tag{5}$$

In the process of ks test, we establish the assumption first:

$$\begin{aligned} &H_0:\ S_n(x) = F_0(x) \\ &H_1:\ S_n(x) \neq F_0(x) \end{aligned} \tag{6}$$

Then we calculate statistic $D_{\max}$. Based on the given significance level α and the number of samples n, we obtain the critical value d_α referring to the statistical table of single sample ks test. If $D_{\max} < d_\alpha$, accept H_0 on the level of α; otherwise, reject H_0. Referring to the statistical table of single sample ks test, we obtain the critical value $d_{0.05}=0.2270$ when significance level $\alpha=0.05$. According to the fitting result of sample data, we obtain that the test statistic based on kernel function distribution model is $d_H = 0.0838$ and the test statistic based on Weibull distribution model is $d_W = 0.0922$. Although both distribution models are considered to be subjected to acceptance condition, the statistic d_H based on kernel function distribution model is obviously less than the statistic d_W based on Weibull distribution model, which indicate that kernel function distribution

model has better goodness of fit and can better reflect the early regularity of reliability distribution of CNC machine tools.

Conclusion

(1) The kernel function is applied to the reliability modeling analysis of CNC machine tools, and the Gaussian distribution is chosen as the kernel function. The results show that the kernel function distribution model can approximate the complex life distribution well, and can be used in modeling analysis of data with complicated regularity.

(2) An optimal model is established to determine the optimal bandwidth parameters with the least integral square error as the objective. The optimal bandwidth parameters are solved by least squares cross-validation and ergodic search. The results show that the model and algorithm have adaptive characteristics and can simplify the process of model selection in reliability modeling.

(3) The kernel function distribution model and the Weibull distribution model are respectively used to analyze the failure data of a certain kind of domestic CNC machine tools at the early failure period. The results show that the kernel function distribution model has better goodness of fit and can better reflect that the reliability of this kind of CNC machine tools has the regularity of skewed distribution during the early failure period.

Acknowledgement

This work is supported by the project of Jilin province education department of China (No. [2016]414) .

References

[1] J. Ginebra, A. Sen, Minimax Approach to Accelerated Life Tests. IEEE Trans. Reliab. 47(3) (1999) 30-37.

[2] Y. Z. Zhang, G. X. Shen, S. Wu, et al. 3-parameter Weibull distribution for random truncated NC machine tool fault data. J. Jilin University, 39(2) (2009) 378-381.

[3] J. L. Wang, Z. J. Yang, G. F. Li, et al. Reliability Modeling of finite Weibull distributions Based on Improved EM Algorithm. J. Jilin University, 44(4) (2014) 1010-1015.

[4] T. Bucar, M. Nagode, M. Fajdiga. Reliability approximation using finite Weibull mixture distributions, Reliab. Eng. Syst. Safe. 84 (2004) 241-251.

[5] B. Scholkopf, A. Smola, K. R. Muller. Nonlinear component analysis as a kernel eigenvalue Problem. Neur. Comput. (3) (2011) 1299-1319.

[6] J. X. Wang, H. B. Chen, Y. Li, Y. Q. Wu, Y. S. Zhang. A Review of the Extrapolation Method in Load Spectrum Compiling. J. Mech. Eng. 62(1) (2016) 60-75.

[7] D. W. Scott. Multivariate density estimation: theory, practice, and visualization. John Wiley & Sons, 2015.

[8] N. B. Heidenreich, A. Schindler, S. Sperlich. Bandwidth selection for kernel density estimation: a review of fully automatic selectors. Adv. Statist. Anal. 97(4) (2013) 403-433.

[9] A. W. Bowman. An alternative method of cross-validation for the smoothing of density estimates. Biometrika, 71(2) (1984) 353-360.