

基于随机退化数据建模的设备 剩余寿命自适应预测方法

孙国玺¹,张清华¹,文成林²,段志宏¹

(1. 广东石油化工学院广东省石化装备故障诊断重点实验室,广东茂名 525000;2. 杭州电子科技大学自动化学院,浙江杭州 310018)

摘要: 针对现有剩余寿命预测研究中需要多个同类设备历史数据离线估计模型参数的问题,本文提出了一种基于退化数据建模的服役设备剩余寿命自适应预测方法.该方法,利用指数随机退化模型来建模设备的退化过程,基于退化监测数据运用 Bayesian 方法更新模型的随机参数,进而得到剩余寿命的概率分布函数及点估计.区别于现有方法,本文方法基于设备到当前时刻的监测数据,利用期望最大化算法对模型中的非随机未知参数进行在线估计,由此无需多个同类设备历史数据.最后,通过数值仿真与实例分析,验证了本文方法在剩余寿命预测时的有效性.

关键词: 寿命预测;退化;Bayesian方法;期望最大化

中图分类号: TP277 **文献标识码:** A **文章编号:** 0372-2112(2015)06-1119-08

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2015.06.1013

A Stochastic Degradation Modeling Based Adaptive Prognostic Approach for Equipment

SUN Guo-xi¹, ZHANG Qing-hua¹, WEN Cheng-lin², DUAN Zhi-hong¹

(1. Guangdong Petrochemical Equipment Fault Diagnosis Key Laboratory, Guangdong University of Petrochemical Technology, Maoming, Guangdong 525000, China; 2. School of Automation, Hangzhou Dianzi University, Hangzhou, Zhejiang 310018)

Abstract: Current prognostic studies are usually based on historical degradation data, which are collected off line from different devices in a population with the same type. However, such data are not always available in practice. Toward this end, this paper presents a degradation modeling based adaptive remaining useful life prediction method for equipments in service. In the presented method, we use an exponential-like stochastic degradation model to represent the degradation process of equipments. Then, based on the monitored data during the degradation process, Bayesian approach is applied to update the stochastic parameters in the model, so the probability distribution of the predicted remaining useful life is derived as well as its point estimation. Differing from current studies, all unknown non-stochastic parameters in the model are estimated by expectation maximization algorithm, without requiring historical degradation data of multiple devices. Finally, numerical simulations and case study results substantiate the superiority of the presented method in predicting the remaining useful life.

Key words: lifetime prognosis; degradation; Bayesian method; expectation maximization

1 引言

随着信息和传感技术的快速发展,预测与健康管理(Prognostics and Health Management, PHM)领域的研究工作引起了更多研究者的兴趣,在此基础上制定的最优维修决策对提高系统可靠性、预防系统未知失效事件发生以及减少维修费用具有重要的实际效益. PHM的核心问题^[1],就是根据传感器所获得的监测数据,有效预测

设备的剩余寿命.

传统的寿命预测方法以失效数据为基础,但是,对于可靠性高,造价昂贵的设备,通常很难获得大量的同类产品的失效数据.研究表明,退化数据比失效时间数据能提供更多的信息来预测设备的剩余寿命^[2,3].目前,基于退化建模的剩余寿命预测方法已经成为一类主要方法,文献[1]对这类方法进行了详细系统的综述.

从20世纪70年代起,国外就陆续开始了退化建模

的相关研究. Gertsbackh 和 Kordonskiy^[4]利用退化数据来对设备的可靠性进行评定. Nelson^[5]总结了 20 世纪 90 年代以前有关退化建模的方法. Lu 和 Meeker^[6]采用随机系数回归模型来模拟退化信号,并阐明了根据这些模型计算寿命分布的各种方法.该文献在退化建模领域具有重要的影响,之后沿着类似的方向,又出现了很多的扩展和变形^[7].这些研究都是对设备整体共性特征进行寿命预测,设备间的差异之处没有充分考虑.为了改进这些方面,Gebræel 等^[8~11]提出了一种 Bayesian 框架下的随机退化建模方法,通过使用带随机系数的指数模型来描述退化信号的变化过程.在获取退化监测数据后,采用 Bayesian 方法对模型的随机参数进行更新,以得到相应的后验估计,进而预测剩余寿命的概率分布.然而,对于模型中的非随机未知参数(随机参数先验分布中的参数和误差项的方差参数)却没有提出相应的估计方法,而是假设存在多个同类设备的历史退化数据,利用统计的方法来对其进行估计.然而,实际中通常很难获得足够多的同类设备的历史退化数据,尤其是对于新运行的设备而言.因此,有必要研究一种不需要多个设备历史退化数据的剩余寿命自适应预测方法.

考虑到以上问题,本文提出了一种基于退化建模的设备剩余寿命自适应预测方法,运用 Bayesian 方法和期望最大化(Expectation Maximization, EM)算法实现对服役设备的退化建模和剩余寿命预测.具体地,该方法利用指数随机退化模型来建模设备的退化过程,基于退化监测数据运用 Bayesian 方法更新模型的随机参数,进而得到剩余寿命的概率分布函数及点估计.进一步,本文方法基于设备到当前时刻的监测数据,利用期望最大化算法对模型中的非随机未知参数进行在线估计.实验分析及与已有方法的比较结果对本文方法的有效性进行了验证.

2 随机退化建模

在现有文献中,指数类随机模型作为描述轴承磨损等累积退化过程的模型,在工程实际中获得了广泛的应用并取得了很好的预测结果^[8~15].因此,本文采用这类退化轨迹模型来描述设备的退化过程.令 $S(t)$ 表示 t 时刻的退化量,设备在离散时间监测点 t_1, t_2, \dots 进行退化监测.那么,在指数型随机退化模型下,设备的在 t_k 时刻的退化量可以表示为:

$$S(t_k) = \phi + \theta \exp\left(\beta t_k + \varepsilon(t_k) - \frac{\sigma^2}{2}\right) \quad (1)$$

其中 ϕ 是固定常数; θ 和 β 是随机变量,用来描述设备间的个体差异; $\varepsilon(t_k)$ 是服从正态分布的随机误差项,且有 $\varepsilon(t_k) \sim N(0, \sigma^2)$.

为简单起见,对指数随机模型进行对数变换,可得:

$$\begin{aligned} L(t_k) &= \ln(S(t_k) - \phi) = \ln\theta - \frac{\sigma^2}{2} + \beta t_k + \varepsilon(t_k) \\ &= \theta' + \beta t_k + \varepsilon(t_k) \end{aligned} \quad (2)$$

其中 $\theta' = \ln\theta - \sigma^2/2$.

在本文中,假设随机参数 θ' 和 β 的先验分布服从联合的二元正态分布,分别具有均值 μ_0' 和 μ_1 、方差 σ_0^2 和 σ_1^2 ,且相关系数 ρ_0 .这一假设区别于文献[16]中 θ' 和 β 相互独立性的假设,因而更具一般性.

基于以上模型和参数假设,一旦得到退化的监测数据,就可以通过 Bayesian 理论来对先验分布中的参数进行更新,得到相应的后验分布.在后文中,为简单标记,记 $L_k = L(t_k)$, $L_{1:k} = \{L_1, \dots, L_k\}$.

3 基于 Bayesian 方法的模型随机参数更新

假设在 t_1, \dots, t_k 时刻有退化观测量 L_1, \dots, L_k .由于误差项 $\varepsilon(t_i)$, $i = 1, \dots, k$ 是独立同分布的随机变量,在给定 θ' 和 β 的条件下,退化样本 $L_{1:k}$ 的条件联合密度函数为:

$$f(L_{1:k} | \theta', \beta) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^k \times \exp\left(-\sum_{i=1}^k \frac{(L_i - \theta' - \beta t_i)^2}{2\sigma^2}\right) \quad (3)$$

由于 θ' 和 β 的先验分布是二元正态分布,与样本分布 $f(L_{1:k} | \theta', \beta)$ 是共轭的,故 θ' 和 β 关于 $L_{1:k}$ 的联合后验分布仍然是二元正态分布,即 $\theta', \beta | L_{1:k} \sim N(\mu_{\theta', k}, \sigma_{\theta', k}^2, \mu_{\beta, k}, \sigma_{\beta, k}^2, \rho_k)$.因此,有:

$$\begin{aligned} f(\theta', \beta | L_{1:k}) &\propto f(L_{1:k} | \theta', \beta) \cdot f(\theta', \beta) \\ &\propto \exp\left[-\sum_{j=1}^k \frac{(L_j - \theta' - \beta t_j)^2}{2\sigma^2} + \frac{(\theta' - \mu_0')^2}{\sigma_0^2}\right] \\ &\cdot \exp\left[-\frac{2\rho_0^2(\theta' - \mu_0')(\beta - \mu_1)}{\sigma_0\sigma_1} + \frac{(\beta - \mu_1)^2}{\sigma_1^2}\right] \end{aligned} \quad (4)$$

$$\begin{aligned} &\propto \frac{1}{2\pi\sigma_{\theta', k}\sigma_{\beta, k}\sqrt{1-\rho_k^2}} \cdot \exp\left[-\frac{(\theta' - \mu_{\theta', k})^2}{2(1-\rho_k^2)\sigma_{\theta', k}^2}\right] \\ &\cdot \exp\left[\rho_k \frac{(\theta' - \mu_{\theta', k})(\beta - \mu_{\beta, k})}{(1-\rho_k^2)\sigma_{\theta', k}\sigma_{\beta, k}} - \frac{(\beta - \mu_{\beta, k})^2}{2(1-\rho_k^2)\sigma_{\beta, k}^2}\right] \end{aligned}$$

经过一些代数的运算,上式对应的参数 $\mu_{\theta', k}, \sigma_{\theta', k}^2, \mu_{\beta, k}, \sigma_{\beta, k}^2, \rho_k$, 可以由下面命题给出.

命题 1^[9] 给定到 t_k 时刻的所有退化观测数据 $L_{1:k} = \{L_1, \dots, L_k\}$, θ', β 关于 $L_{1:k}$ 的联合后验分布服从二元正态分布,即:

$\theta', \beta | L_{1:k} \sim N(\mu_{\theta', k}, \sigma_{\theta', k}^2, \mu_{\beta, k}, \sigma_{\beta, k}^2, \rho_k)$, 具体参数表达式如下:

$$\begin{aligned} \mu_{\theta',k} &= \frac{\mu_0' \sigma^2 \sigma_1 (Y + Z \rho_0) - \mu_1 \sigma^2 \sigma_0 (Y \rho_0 + Z)}{\sigma_1 (XY - Z^2)} \\ &+ \frac{(1 - \rho_0^2) \sigma_0 \sigma_1 \left(Y \sigma_0 \sum_{i=1}^k L_i - Z \sigma_1 \sum_{i=1}^k L_i t_i \right)}{\sigma_1 (XY - Z^2)} \\ \mu_{\beta,k} &= \frac{\mu_1 \sigma^2 \sigma_0 (X + Z \rho_0) - \mu_0' \sigma^2 \sigma_1 (X \rho_0 + Z)}{\sigma_0 (XY - Z^2)} \\ &+ \frac{(1 - \rho_0^2) \sigma_0 \sigma_1 \left(X \sigma_1 \sum_{i=1}^k L_i t_i - Z \sigma_0 \sum_{i=1}^k L_i \right)}{\sigma_0 (XY - Z^2)} \\ \sigma_{\theta',k}^2 &= \frac{\left[(1 - \rho_0^2) \sigma_1^2 \sum_{i=1}^k t_i^2 + \sigma^2 \right] (1 - \rho_0^2) \sigma^2 \sigma_0^2}{XY - W^2} \\ \sigma_{\beta,k}^2 &= \frac{\left[k(1 - \rho_0^2) \sigma_0^2 + \sigma^2 \right] (1 - \rho_0^2) \sigma^2 \sigma_1^2}{XY - W^2} \\ \rho_k &= - \left(\frac{(1 - \rho_0^2) \sigma_0 \sigma_1 \sum_{i=1}^k t_i - \rho_0 \sigma^2}{\sqrt{\left[k(1 - \rho_0^2) \sigma_0^2 + \sigma^2 \right] \left[(1 - \rho_0^2) \sigma_1^2 \sum_{i=1}^k t_i^2 + \sigma^2 \right]}} \right) \end{aligned} \quad (5)$$

这里 $X = k(1 - \rho_0^2) \sigma_0^2 + \sigma^2$, $Y = (1 - \rho_0^2) \sigma_1^2 \sum_{i=1}^k t_i^2 + \sigma^2$, $Z = (1 - \rho_0^2) \sigma_0 \sigma_1 \sum_{i=1}^k t_i - \rho_0 \sigma^2$, $W = (1 - \rho_0^2) \sigma_0 \sigma_1 \sum_{i=1}^k t_i + \rho_0 \sigma^2$.

由于 $\theta', \beta | L_{1:k} \sim N(\mu_{\theta',k}, \sigma_{\theta',k}^2, \mu_{\beta,k}, \sigma_{\beta,k}^2, \rho_k)$ 是二元正态分布的, 根据二元正态分布的性质, 有如下推论成立.

命题 2 令 $\cdot | L_{1:k}$ 表示条件随机变量, $E(\cdot | L_{1:k})$ 表示条件期望, 则:

$$\begin{aligned} &\theta' | L_{1:k} \sim N(\mu_{\theta',k}, \sigma_{\theta',k}^2), \\ (1) &\beta | L_{1:k} \sim N(\mu_{\beta,k}, \sigma_{\beta,k}^2), \\ &\beta | \theta', L_{1:k} \sim N(\mu_{\beta|\theta',k}, \sigma_{\beta|\theta',k}^2); \\ &E(\theta' \beta | L_{1:k}) = \rho_k \sigma_{\theta',k} \sigma_{\beta,k} + \mu_{\theta',k} \mu_{\beta,k}, \\ (2) &E(\theta' | L_{1:k}) = \mu_{\theta',k}, \\ &E(\beta | L_{1:k}) = \mu_{\beta,k}, E(\theta'^2 | L_{1:k}) = \mu_{\theta',k}^2 + \sigma_{\theta',k}^2, \\ &E(\beta^2 | L_{1:k}) = \mu_{\beta,k}^2 + \sigma_{\beta,k}^2. \end{aligned}$$

其中:

$$\begin{aligned} \mu_{\beta|\theta',k} &= \mu_{\beta,k} + \rho_k \sigma_{\beta,k} (\theta' - \mu_{\theta',k}) / \sigma_{\theta',k}, \\ \sigma_{\beta|\theta',k}^2 &= \sigma_{\beta,k}^2 (1 - \rho_k^2). \end{aligned}$$

4 剩余寿命预测

在得到 θ' 和 β 的后验估计后, 对于给定的失效阈值 ω , 要实现剩余寿命的预测, 首先要基于到当前时刻 t_k 的退化观测数据 $L_{1:k}$ 预测 $t + t_k$ 时退化量. 由此, 剩余寿命的预测问题就转化为预测退化量达到失效阈值 ω 的

时间的问题. 对于将来时刻 $t + t_k$ 的退化量 $L(t + t_k)$, 在给定 $L_{1:k}$ 的条件下, 有以下命题成立.

命题 3 给定到 t_k 时刻的所有退化观测数据 $L_{1:k}$, 预测的 $t + t_k$ 时刻对应的退化量 $L(t + t_k)$ 服从正态分布, 且具有的均值和方差分别为:

$$\begin{aligned} \tilde{\mu}(t + t_k) &= \mu_{\theta',k} + \mu_{\beta,k}(t + t_k) - \frac{\sigma^2}{2}, \\ \tilde{\sigma}^2(t + t_k) &= \sigma_{\theta',k}^2 + \sigma_{\beta,k}^2(t + t_k) + \sigma^2 \\ &+ 2\rho_k(t + t_k) \sigma_{\theta',k} \sigma_{\beta,k} \end{aligned} \quad (6)$$

命题 3 是基于所建模型式(12)得到的, 因为正态随机变量的线性组合依然是正态分布的, 因此只需要确定 $L(t + t_k)$ 的期望和方差就可以得到 $L(t + t_k)$ 的分布. 而 $L(t + t_k)$ 的期望和方差可以基于式(12)得到, 即 $L(t + t_k) = \theta' + \beta(t + t_k) + \varepsilon(t + t_k)$. 因此, 利用对参数的 θ' 和 β 后验估计结果, 即命题 2, 就可得到命题 3 的结论.

根据以上结果, 设 T 为设备在 t_k 时刻的剩余寿命, 即 T 满足 $L(t + t_k) = \ln \omega$. 于是, 在给定 $L_{1:k}$ 的条件下, 预测的剩余寿命的条件概率分布函数 $F_{T|L_{1:k}}(t)$ 为:

$$\begin{aligned} F_{T|L_{1:k}}(t) &= \Pr(T \leq t | L_{1:k}) \\ &= \Pr(S(t + t_k) \geq \ln \omega | L_{1:k}) \\ &= \Pr\left(Z \geq \frac{\ln \omega - \tilde{\mu}(t + t_k)}{\sqrt{\tilde{\sigma}^2(t + t_k)}}\right) = \Phi(g(t)) \end{aligned} \quad (7)$$

其中 Z 服从标准正态分布, $\Phi(\cdot)$ 为标准正态随机变量的累积分布函数, 且有 $g(t) = (\tilde{\mu}(t + t_k) - \ln \omega) / \sqrt{\tilde{\sigma}^2(t + t_k)}$.

注意到 $\lim_{t \rightarrow -\infty} g(t) = -\mu_{\beta,k} / \sigma_{\beta,k}$, 所以 T 的值域为 $(-\infty, \infty)$. 因此, 考虑到 T 表示设备的寿命, 是非负的实数, 由此得到在 $(0, \infty)$ 上截断的累积分布函数作为设备的剩余寿命预测结果, 表示为:

$$\begin{aligned} F_{T|L_{1:k}, T \geq 0}(t) &= \Pr(T \leq t | L_{1:k}, T \geq 0) \\ &= \frac{\Pr(0 \leq T \leq t | L_{1:k})}{\Pr(T \geq 0 | L_{1:k})} \\ &= \frac{\Phi(g(t)) - \Phi(g(0))}{1 - \Phi(g(0))} \end{aligned} \quad (8)$$

基于以上的结果, 剩余寿命预测的条件概率密度函数为:

$$f_{T|L_{1:k}, T \geq 0}(t) = \frac{dF_{T|L_{1:k}, T \geq 0}(t)}{dt} = \frac{\varphi(g(t))}{1 - \Phi(g(0))} \cdot g'(t) \quad (9)$$

其中 $\varphi(\cdot)$ 为标准正态随机变量的概率密度函数.

通过上面的建模过程可以发现, 在获得新的退化测量数据后, 就可以根据命题 1 重新计算随机参数 θ' 和 β 的后验分布, 由此获得 $\mu_{\theta',k}, \sigma_{\theta',k}, \mu_{\beta,k}, \sigma_{\beta,k}, \rho_k$ 的新估计. 然后, 根据这些更新的参数, 通过更新函数 $g(t)$ 中 $\tilde{\mu}(t + t_k)$ 和 $\tilde{\sigma}^2(t + t_k)$ 的值可以实现对预测的

剩余寿命分布的更新.

本文采用如下方法得到剩余寿命的点估计:根据退化失效的定义有 $L(T + t_k) = \ln\omega$, 于是利用预测的 $t + t_k$ 时刻退化量的均值 $\tilde{\mu}(t + t_k)$ 替代 $L(T + t_k)$, 使得 $\tilde{\mu}(t + t_k) = \ln\omega$, 由此得到 t_k 时刻剩余寿命的点估计 RUL_k 如下:

$$RUL_k = \frac{\ln\omega - \mu'_{\theta, k} + \sigma^2/2}{\mu_{\beta, k}} - t_k \quad (10)$$

在以上的退化建模和剩余寿命预测过程中, 模型参数 $\sigma^2, \mu'_0, \mu_1, \sigma_0^2, \sigma_1^2, \rho_0$ 是未知的, 现有的剩余寿命预测研究是通过多个同类设备的历史数据对这些参数估计, 然而实际中通常很难获得大量同类产品的历史退化数据. 为解决这一问题, 下节主要讨论退化模型未知参数的估计问题.

5 基于 EM 算法的模型参数估计

令 $\Theta = [\sigma^2, \mu'_0, \mu_1, \sigma_0^2, \sigma_1^2, \rho_0]$ 表示模型未知参数. 基于极大似然估计的方法, 在 t_k 时刻的监测数据 L_k 得到后, 计算关于 $L_{1:k}$ 的对数似然函数:

$$\ell_k(\Theta) = \log[p(L_{1:k} | \Theta)] \quad (11)$$

其中 $p(L_{1:k} | \Theta)$ 表示退化数据 $L_{1:k}$ 的联合密度函数. 因此, Θ 的极大似然估计 $\hat{\Theta}_k$ 可以通过由下式得到:

$$\hat{\Theta}_k = \arg \max_{\Theta} \ell_k(\Theta) \quad (12)$$

然而在退化模型式(2)中, 参数 θ' 和 β 是随机的, 因此难以采用式(12)直接优化得到 $\hat{\Theta}_k$. 为此, 本文采用 EM 算法实现对 Θ 的估计. EM 算法对 Θ 的估计可以通过迭代以下两步实现:

(1) E-step: 计算条件期望

$$\ell_k(\Theta | \hat{\Theta}_k^{(i)}) = E_{\theta', \beta | L_{1:k}, \hat{\Theta}_k^{(i)}} \{ \log p(L_{1:k}, \theta', \beta | \Theta) \} \quad (13)$$

其中 $\hat{\Theta}_k^{(i)}$ 表示基于退化数据 $L_{1:k}$ 进行估计时第 i 次迭代的结果.

(2) M-step: 最大化条件期望

$$\hat{\Theta}_k^{(i+1)} = \arg \max_{\Theta} \ell(\Theta | \hat{\Theta}_k^{(i)}) \quad (14)$$

在以上两步的迭代过程中, 通常在 $\hat{\Theta}_k^{(i)}$ 和 $\hat{\Theta}_k^{(i+1)}$ 的差小于一个比较小的数时, 终止迭代, 将最后一次估计的结果作为 t_k 时刻最终的参数估计结果. 下面利用上述 EM 算法对未知参数 Θ 进行估计.

为了表示估计的参数依赖于到当前时刻的所有退化监测数据, 将基于到 t_k 时刻的所有退化监测数据 $L_{1:k}$ 估计的参数表示为 $\Theta_k = [\sigma_k^2, \mu'_{0,k}, \mu_{1,k}, \sigma_{0,k}^2, \sigma_{1,k}^2, \rho_{0,k}]$.

令 EM 算法中第 i 次迭代得到的估计为 $\hat{\Theta}_k^{(i)} = [\sigma_k^{2(i)}, \mu'_{0,k}{}^{(i)}, \mu_{1,k}{}^{(i)}, \sigma_{0,k}^{2(i)}, \sigma_{1,k}^{2(i)}, \rho_{0,k}{}^{(i)}]$, 那么完全对数似然函数可以表示为

$$\begin{aligned} \ln p(L_{1:k}, \theta', \beta | \Theta_k) &= \ln p(L_{1:k} | \theta', \beta, \Theta_k) + \ln p(\theta', \beta | \Theta_k) \\ &= -\frac{k+2}{2} \ln 2\pi - \frac{k}{2} \ln \sigma_k^2 - \sum_{j=1}^k \frac{(L_j - \theta' - \beta t_j)^2}{2\sigma_k^2} \\ &\quad - \ln \sigma_{0,k}^2 - \ln \sigma_{1,k}^2 - \ln \sqrt{1 - \rho_{0,k}^2} - \frac{1}{2(1 - \rho_{0,k}^2)} \left[\frac{(\theta' - \mu'_{0,k})^2}{\sigma_{0,k}^2} \right. \\ &\quad \left. - \frac{2\rho_{0,k}(\theta' - \mu'_{0,k})(\beta - \mu_{1,k})}{\sigma_{0,k}\sigma_{1,k}} + \frac{(\beta - \mu_{1,k})^2}{\sigma_{1,k}^2} \right] \end{aligned} \quad (15)$$

基于式(13), 可以得到 $\ell_k(\Theta | \hat{\Theta}_k^{(i)})$ 如下:

$$\begin{aligned} \ell_k(\Theta | \hat{\Theta}_k^{(i)}) &= E_{\theta', \beta | L_{1:k}, \hat{\Theta}_k^{(i)}} \{ \log p(L_{1:k}, \theta', \beta | \Theta) \} \\ &= -\frac{k+2}{2} \ln 2\pi - \frac{k}{2} \ln \sigma_k^2 - \ln \sigma_{0,k}^2 - \ln \sigma_{1,k}^2 - \ln \sqrt{1 - \rho_{0,k}^2} \\ &\quad - \sum_{j=1}^k \frac{1}{2\sigma_k^2} \left[L_j^2 - 2t_j(\mu'_{\theta, k} + \mu_{\beta, k} t_j) + \mu_{\theta, k}^2 + \sigma_{\theta, k}^2 \right. \\ &\quad \left. + 2t_j(\rho_{0,k} \sigma_{\theta, k} \sigma_{\beta, k} + \mu'_{\theta, k} \mu_{\beta, k}) + t_j^2(\mu_{\beta, k}^2 + \sigma_{\beta, k}^2) \right] \\ &\quad - \frac{1}{2(1 - \rho_{0,k}^2)} \left[\frac{\mu_{\theta, k}^2 + \sigma_{\theta, k}^2 - 2\mu'_{0,k} \mu'_{\theta, k} + \mu_{0,k}^2}{\sigma_{0,k}^2} \right. \\ &\quad \left. - \frac{2\rho_{0,k}(\rho_{0,k} \sigma_{\theta, k} \sigma_{\beta, k} + \mu'_{\theta, k} \mu_{\beta, k} - \mu_{1,k} \mu'_{\theta, k} - \mu'_{0,k} \mu_{\beta, k} + \mu'_{0,k} \mu_{1,k})}{\sigma_{0,k}\sigma_{1,k}} \right. \\ &\quad \left. + \frac{\mu_{\beta, k}^2 + \sigma_{\beta, k}^2 - 2\mu_{1,k} \mu_{\beta, k} + \mu_{1,k}^2}{\sigma_{1,k}^2} \right] \end{aligned} \quad (16)$$

令 $\frac{\partial \ell(\Theta_k | \hat{\Theta}_k^{(i)})}{\partial \Theta_k} = 0$ 可以得到第 $i+1$ 步的参数估计

$\hat{\Theta}_k^{(i+1)}$ 为:

$$\begin{aligned} \hat{\sigma}_k^{2(i+1)} &= \frac{1}{k} \sum_{j=1}^k (L_j^2 - 2L_j(\mu'_{\theta, k} + \mu_{\beta, k} t_j) + \mu_{\theta, k}^2 + \sigma_{\theta, k}^2 \\ &\quad + 2t_j(\rho_{0,k} \sigma_{\theta, k} \sigma_{\beta, k} + \mu'_{\theta, k} \mu_{\beta, k}) + t_j^2(\mu_{\beta, k}^2 + \sigma_{\beta, k}^2)) \\ \hat{\mu}'_{0,k}{}^{(i+1)} &= 2 \frac{\sigma_{0,k}}{\sigma_{1,k}} \cdot \rho_{0,k} (\mu_{1,k} - \mu_{\beta, k}) + \mu'_{\theta, k} \\ \hat{\mu}'_{1,k}{}^{(i+1)} &= 2 \frac{\sigma_{1,k}}{\sigma_{0,k}} \cdot \rho_{0,k} (\mu_{0,k} - \mu'_{\theta, k}) + \mu_{\beta, k} \\ \hat{\sigma}_{0,k}^{2(i+1)} &= \frac{1}{4} \left(-\frac{1}{2} B + \sqrt{\frac{1}{4} B^2 + 4A} \right)^2 \\ \hat{\sigma}_{0,k}^{2(i+1)} &= \frac{1}{4} \left(-\frac{1}{2} D + \sqrt{\frac{1}{4} D^2 + 4C} \right)^2 \\ \hat{\rho}_{0,k}^{(i+1)} &= \sqrt{1 - 2E} \end{aligned} \quad (17)$$

其中:

$$\begin{aligned} A &= \frac{\mu_{\theta, k}^2 + \sigma_{\theta, k}^2 - 2\mu'_{0,k} \mu'_{\theta, k} + \mu_{0,k}^2}{2(1 - \rho_{0,k}^2)} \\ B &= \frac{\rho_{0,k}(\rho_{0,k} \sigma_{\theta, k} \sigma_{\beta, k} + \mu'_{\theta, k} \mu_{\beta, k} - \mu_{1,k} \mu'_{\theta, k} - \mu'_{0,k} \mu_{\beta, k} + \mu'_{0,k} \mu_{1,k})}{\sigma_{1,k}(1 - \rho_{0,k}^2)} \\ C &= \frac{\mu_{\beta, k}^2 + \sigma_{\beta, k}^2 - 2\mu_{1,k} \mu_{\beta, k} + \mu_{1,k}^2}{2(1 - \rho_{0,k}^2)} \end{aligned}$$

$$D = \frac{\rho_{0,k}^2 (\rho_0 \sigma_{\theta,k} \sigma_{\beta,k} + \mu_{\theta,k} \mu_{\beta,k} - \mu_{1,k} \mu_{\theta,k} - \mu_{0,k} \mu_{\beta,k} + \mu_{0,k} \mu_{1,k})}{\sigma_{0,k} (1 - \rho_{0,k}^2)}$$

$$E = \frac{\rho_0 \sigma_{\theta,k} \sigma_{\beta,k} + \mu_{\theta,k} \mu_{\beta,k} - \mu_{1,k} \mu_{\theta,k} - \mu_{0,k} \mu_{\beta,k} + \mu_{0,k} \mu_{1,k}}{\sigma_{0,k} \rho_{1,k}}$$

上述的参数估计方法是基于到 t_k 时刻的所有退化监测数据 $L_{1:k}$ 对参数 Θ 估计的结果. 由于 $L_{1:k}$ 对监测时刻的依赖关系, 在任一监测时刻的数据得到后都可以采用上面的方法进行参数估计. 另外, M 步只可以得到显示的最优解, 因而易于计算. 以上的参数估计方法的理论性质和收敛性分析可以参考文献[17~22], 限于篇幅, 这里不再赘述.

6 实验分析

6.1 数值仿真

为了验证本文提出的方法, 本节给出一个数值例子并与文献[9]中的方法进行比较. 为了产生仿真的退化数据, 本文利用文献[9]中的退化模型来产生一组用于验证本文方法. 这里采用的模型为 $X(t) = \varphi + \theta \exp(\beta t + \varepsilon(t) - \sigma^2/2)$. 令模型的参数为: $\varphi = 0.1$, $\ln \theta \sim N(0.2, 2 \times 10^{-8})$, $\sigma^2 = 1 \times 10^{-6}$, $\beta \sim N(0.005, 1 \times 10^{-8})$, $\varepsilon(t) \sim N(0, \sigma^2)$, $\rho_0 = 0$. 基于此模型产生仿真的退化测量数据, 在数据仿真过程中假定采样间隔为 4, 失效阈值为 $w = 15.59$ 时. 基于此, 得到了仿真的退化轨迹如图 1 所示, 共得到 125 个监测数据, 对应的超过失效阈值的时间约为 496 个仿真步长.

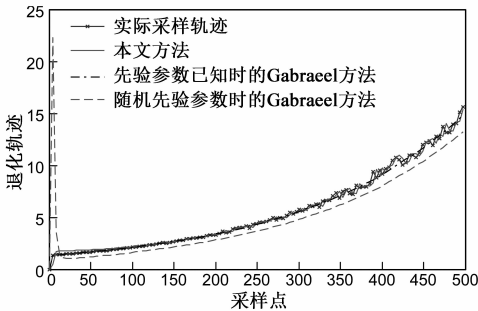


图1 仿真的退化轨迹及预测的退化轨迹

为了比较模型对仿真数据的拟合程度, 本文提出方法中的初始参数通过随机产生的方式确定, 而对于文献[9]中的方法, 采用两种不同的初始参数选择方法: 一种是采用上面参数仿真数据的真实参数直接初始化模型, 第二种也是随机的产生模型的初始参数. 图 1 给出了在不同情况下本法方法和文献[9]中的方法对退化量预测的结果. 在所有数据点预测的均方误差分别为 0.0746, 0.0624, 1.3852. 可以发现不合理的初始参数会对文献[9]中的方法的结果产生很大的影响.

为进一步验证本文提出的模型参数估计方法对初

始参数选择的鲁棒性, 对本文方法和文献[9]中方法在初始参数随机选择的情况下进行了 10000 次仿真, 得到了实际退化数据和预测值的均方误差的直方图, 见图 2.

通过图 2 中的结果可以发现, 本文提出的自适应方法对于模型初始参数的选择具有较好的鲁棒性, 而文献[9]中的方法对初始参数很敏感. 因此, 当在实际应用中, 没有足够多的历史退化数据时, 本文的方法具有很大的优越性.

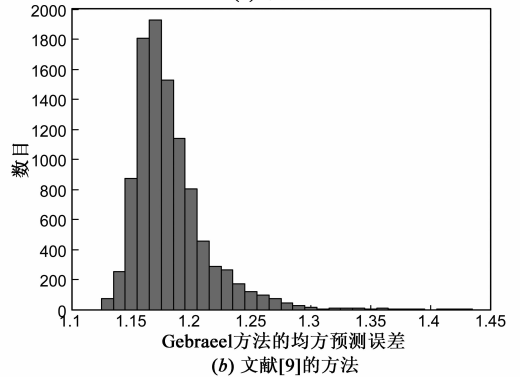
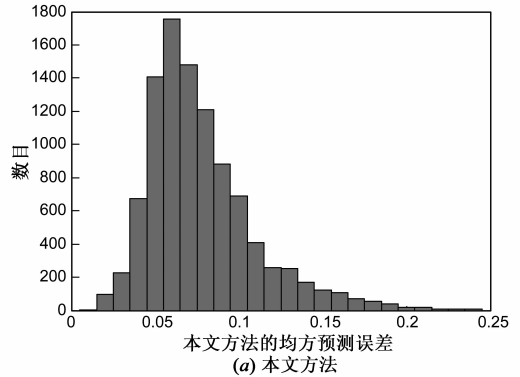


图2 退化预测均方误差

对于剩余寿命预测的结果, 图 3 和图 4 分别给出了本文方法和文献[9]中的方法在不同退化监测点预测的剩余寿命的概率密度曲线, 其中“*”表示实际的剩余寿命. 这里考虑到文献[9]中的剩余寿命预测方法对初始参数很敏感, 选择用于产生仿真数据的真实参数作为初始模型参数.

通过图 3 和图 4 的对比可以发现, 采用本文的剩余寿命自适应预测方法, 剩余寿命预测的结果随着数据的积累, 预测不确定性不断减小, 因为剩余寿命的概率密度函数变窄. 然而, 文献[9]中方法得到的预测结果的不确定性几乎是不变的, 主要原因在于文献[9]中的方法是基于历史数据对模型的确定性参数进行估计, 一旦估计得到后不再利用实时数据进行更新, 而由于 $\varepsilon(t)$ 的存在, 使得剩余寿命预测的不确定性基本上被 $\varepsilon(t)$ 的不确定性主导. 但是, 刻画 $\varepsilon(t)$ 不确定性的参数

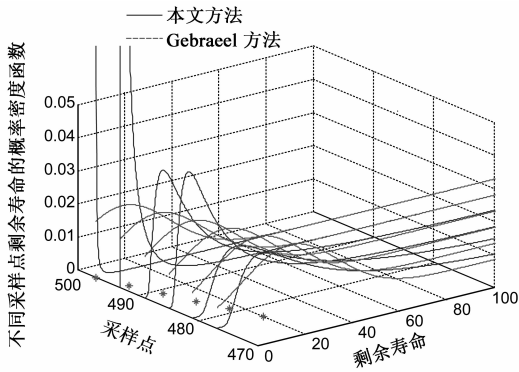


图3 不同采样点本文方法与文献[9]中方法剩余寿命预测结果

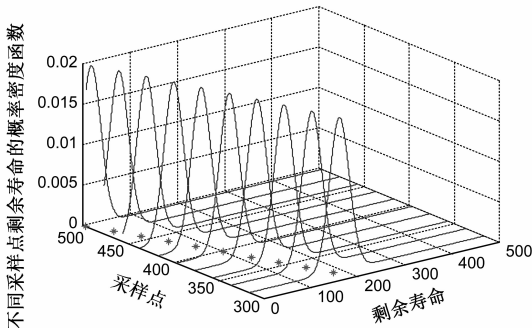


图4 文献[9]中方法剩余寿命预测结果

σ^2 是离线确定的.这就造成了剩余寿命预测的不确定性不随退化过程信息增加而减小.这其实也验证了文献[9]的方法受初始参数选择及初始知识的约束很大.相比之下,本文的方法,由于引进了基于EM算法的参数自适应更新机制,使得模型的所有参数都随着服役设备实时数据的积累不断更新,因而预测的结果更能反映设备的实际运行情况,达到减小预测不确定性的目的,具有实际应用的潜在价值.

6.2 实例验证

锂电池是许多电子设备和复杂系统的关键部件,已被广泛用于航空航天领域以及军事领域.许多卫星和航天飞行器的故障都是由电池失效导致的.因此,有效建模电池性能退化指标、精确预测电池的剩余寿命能够有效降低失效风险,提高系统可靠性、安全性.

下面以 NASA 提供的锂电池数据为例,验证所提方法的效果.锂电池电池的寿命包括使用寿命、贮存寿命和循环寿命.其中,循环寿命的应用最为广泛,研究价值最为重要.因此,本实验考虑电池的循环剩余寿命.锂电池的使用过程是反复充放电的过程.随着充放电次数的增加,电池的容量会逐渐衰减,循环寿命即为电池容量下降到额定容量之前,可以反复充电的次数.电池的容量变化表征了电池的退化状态,因此,将容量作为电池实际的退化状态.NASA 研究中心联合美国能源部国家实验室,对锂电池的循环寿命开展了大量的实

验性研究.所得数据集主要来源于研究中心搭建的锂电池测试床.测试过程中,综合考虑多种工况条件,反复对锂电池充电、放电,加速电池老化,导致电池容量衰减,当电池容量下降到大约额定容量的 30%,即认为电池循环寿命结束,停止实验.在每一个充放电周期都记录对应的电池容量和其他相关的内部参数如阻抗测量等.

对于 NASA 提供的电池数据集的介绍,可参考文献[23].与文献[23]类似,本文将测试电容的容量作为电池监测量,利用 5 号电池的数据验证方法的效果,其失效时间约为 125 个循环周期.为了与文献[16]比较,说明本文方法的有效性,在 30%、60%、90% 寿命分位点分别计算了剩余寿命预测的点估计值,具体结果见表 1.

表 1 不同寿命分位点的剩余寿命预测点估计

	30% 寿命分位点	60% 寿命分位点	90% 寿命分位点
文献[15]的方法	102.3	56.3	14.6
本文方法	82.4	48.2	11.9
实际剩余寿命	87.5	50.0	12.5

通过表 1 的比较结果,可以发现本文提出的剩余寿命预测方法能够有效地提高电池剩余寿命预测的准确性.此外,文献[16]中的方法容易造成对剩余寿命的高估,因而增加了电池运行失效的风险,不利于及时的采取健康管理措施.相比之下,本文提出的方法可以有效地避免这些问题,从而验证了本文方法的有效性.

7 结论

针对服役设备的剩余寿命预测问题,本文提出了一种基于退化数据建模的服役设备剩余寿命自适应预测方法,该方法最大不同之处在于不需要多个同类设备历史数据来初始化退化模型.具体地,本文方法利用指数随机退化模型来建模设备的退化过程,基于退化监测数据运用 Bayesian 方法更新模型的随机参数,进而得到了剩余寿命的概率分布函数及点估计.进一步,利用期望最大化算法对模型中的非随机未知参数进行在线估计,使得模型参数和剩余寿命分布都能随数据不断更新.实例分析结果表明:本文提出的方法能够充分考虑个体间的差异,退化建模精度和剩余寿命预测都得到了提升.

参考文献

- [1] Si X S, Wang W B, Hu C H, et al. Remaining useful life estimation—A review on the statistical data driven approaches[J]. European Journal of Operational Research, 2011, 213(1): 1–14.
- [2] Meeker W Q, Escobar L A. A review of recent research and

- current issues in accelerated testing[J]. *International Statistical Review*, 1993, 61(1): 147 – 168.
- [3] 司小胜, 胡昌华, 周东华. 带测量误差的非线性退化过程建模与剩余寿命估计[J]. *自动化学报*, 2013, 39(5): 590 – 601.
Si X S, Hu C H, Zhou D H. Nonlinear degradation process modeling and remaining useful life estimation subject to measurement error[J]. *Acta Automatica Sinica*, 2013, 39(5): 590 – 601. (in Chinese)
- [4] Gertsbach I B, Kordonskiy K B. *Models of Failure*[M]. New York: Springer-Verlag, 1969.
- [5] Nelson W. *Accelerated Testing: Statistical Models, Test Plans and Data Analyses*[M]. New York: Wiley, 1990.
- [6] Lu C J, Meeker W Q. Using degradation measures to estimate a time-to-failure distribution[J]. *Technometrics*, 1993, 35(2): 161 – 174
- [7] 周东华, 魏慕恒, 司小胜. 工业过程异常检测、寿命预测与维修决策的研究进展[J]. *自动化学报*, 2013, 39(6): 711 – 722.
Zhou D H, Wei M H, Si X S. A survey on anomaly detection, life prediction and maintenance decision for industrial processes [J]. *Acta Automatica Sinica*, 2013, 39(6): 711 – 722.
- [8] Gebraeel N Z, Lawley M A, Li R, et al. Residual-life distributions from component degradation signals: A Bayesian approach [J]. *IIE Transactions*, 2005, 37(6): 543 – 557.
- [9] N Gebraeel. Sensory-updated residual life distributions for components with exponential degradation patterns[J]. *IEEE Transactions on Automation Science and Engineering*, 2006, 3(4): 382 – 393.
- [10] Gebraeel N Z, Elwany A, Pan J. Residual life predictions in the absence of prior degradation knowledge[J]. *IEEE Transactions on Reliability*, 2009, 58(1): 106 – 117.
- [11] Elwany A, Gebraeel N Z, Maillart L. Structured replacement policies for systems with complex degradation processes and dedicated sensors[J]. *Operations Research*, 2011, 59(3): 684 – 695.
- [12] Si X S, Wang W B, Hu C H, Zhou D H, Pecht M. Remaining useful life estimation based on a nonlinear diffusion degradation process [J]. *IEEE Transactions on Reliability*, 2012, 61(1), 50 – 67.
- [13] Park C, Padgett W J. Stochastic degradation models with several accelerating variables[J]. *IEEE Transactions on Reliability*, 2006, 55(2): 379-390.
- [14] Tseng S T, Tang J, Ku L H. Determination of optimal burn-in parameters and residual life for highly reliable products[J]. *Naval Research Logistics*, 2003, 50(1): 1 – 14.
- [15] 程龙, 冯静, 孙权, 周经伦, 蔡永超. 非连续运行设备贮存 – 工作联合退化模型及其应用[J]. *电子学报*, 2012, 40(12): 2549 – 2552.
Cheng L, Feng J, Sun Q, Zhou J L, Cai Y C. Storage-work joint degradation model for discontinuous-working devices and its application[J]. *Acta Electronica Sinica*, 2012, 40(12): 2549 – 2552. (in Chinese)
- [16] 司小胜, 胡昌华, 李娟, 陈茂银. Bayesian 更新与 EM 算法协作下退化数据驱动的剩余寿命估计方法[J]. *模式识别与人工智能*, 2013, 26(4): 357 – 365.
Si X S, Hu C H, Li J, Chen M Y. Degradation data-driven remaining useful life estimation approach under collaboration between Bayesian updating and EM algorithm [J]. *Pattern Recognition and Artificial Intelligence*, 2013, 26(4): 357 – 365. (in Chinese)
- [17] 周鑫. 基于 EM 算法的 GO 分布参数最大似然估计[J]. *电子学报*, 2013, 41(1): 178 – 184.
Zhou X. An EM algorithm based maximum likelihood parameter estimation method for the GO distribution[J]. *Acta Electronica Sinica*, 2013, 41(1): 178 – 184. (in Chinese)
- [18] Dempster A P, Laird N M, Rubin D B. Maximum likelihood from incomplete data via the EM algorithm[J]. *Journal of the Royal Statistical Society Series B*, 1977, 39(1): 1 – 38.
- [19] Wu C F. On the convergence property of the EM algorithm [J]. *Annals of Statistics*, 1983, 11(1): 95 – 103.
- [20] 彭宝华, 周经伦, 冯静, 刘学敏. 金属化膜脉冲电容器剩余寿命预测方法研究[J]. *电子学报*, 2011, 39(11): 2674 – 2679.
Peng B H, Zhou J L, Feng J, Liu X M. Residual lifetime prediction of metallized film pulse capacitors[J]. *Acta Electronica Sinica*, 2011, 39(11): 2674 – 2679. (in Chinese)
- [21] 王小林, 郭波, 程志君. 融合多源信息的维纳过程性能退化产品的可靠性评估[J]. *电子学报*, 2012, 40(5): 977 – 982.
Wang X L, Guo B, Cheng Z J. Reliability assessment of products with Wiener process degradation by fusing multiple information[J]. *Acta Electronica Sinica*, 2012, 40(5): 977 – 982. (in Chinese)
- [22] 钟强晖, 张志华, 李大伟. 基于性能退化的电子产品筛选试验设计[J]. *电子学报*, 2013, 41(9): 1788 – 1793.
Zhong Q H, Zhang Z H, Li D W. Screening test design for electronics based on performance degradation[J]. *Acta Electronica Sinica*, 2013, 41(9): 1788 – 1793. (in Chinese)
- [23] Saha B, Goebel K, Poll S, Christophersen J. Prognostics methods for battery health monitoring using a Bayesian framework [J]. *IEEE Transactions on Instrumentation Measurement*, 2009, 58(2): 291 – 296.

作者简介



孙国玺 男,1972年10月出生,黑龙江友谊县人,副教授,硕士生导师.2006年毕业于华南理工大学信号与系统专业,获工学博士学位,现为广东石油化工学院广东省石化装备故障诊断重点实验室副主任,主要从事石化装备的寿命预测、故障诊断和健康管理等方向的研究.

E-mail: sguoxi@126.com



张清华 男,1965年3月出生,广东梅州人,教授、博士生导师.1995年在华南理工大学获工业自动化专业硕士学位,2004年获控制理论与控制工程专业博士学位.现任广东省石化装备故障诊断重点实验室主任,主要致力于石化生产过程和生产装备的监测与故障诊断、复杂系统优化控制与仿真等方面的应用研究.

E-mail: fengliangren@tom.com



文成林 男,1963年出生,博士(后),教授、博士生导师.1999年西北工业大学控制理论与控制工程专业毕业,获工学博士学位;2002年清华大学控制科学与工程博士后流动站出站.现任杭州电子科技大学系统科学与控制工程研究所所长,是上海交通大学和上海海事大学博士生导师,并与浙江大学、河海大学、南京航空航天大学联合培养博士学位研究生.主要研究方向:多源信息融合与目标检测识别和跟踪,复杂系统故障诊断与健康管理、及安全评估与安全运行理论和方法.

E-mail: wencel@hdu.edu.cn



段志宏 男,1969年3月出生,四川省大竹县人,工程师.毕业于武汉理工大学机械工程硕士专业.现为广东石油化工学院广东省石化装备故障诊断重点实验室办公室主任.主要从事石化装备故障诊断、油液分析、转子动平衡等方向的研究.

E-mail: 13727781119@163.com