

灰色极限学习机在滚动轴承故障预测中的应用

徐 遥

(北京交通大学 电气工程学院, 北京 100044)

摘要: 针对较强噪声环境下的滚动轴承故障预测问题, 为提高轴承故障预测的精度, 提出并研究了一种新的滚动轴承预测技术; 采用将灰色模型和极限学习机 (ELM) 相结合的方法, 针对轴承运行状态值的非线性特点, 先将样本数据进行灰色处理, 解决数据的随机性和波动性问题, 然后代入学习速度快, 泛化精度高的 ELM 神经网络进行训练; 在训练完毕后, 对未来的轴承运行状态数据进行分析, 将其与轴承设备的理论诊断标准相比较以达到故障预测的目的。

关键词: 灰色理论; 极限学习机; 滚动轴承; 故障预测

Rolling Element Bearing Fault Prediction Based on Grey Sequential Extreme Learning Machine

Xu Yao

(School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China)

Abstract: Aiming at the prediction of rolling element bearing fault in the strong noise environment, a novel method of prediction for rolling element bearing is proposed to improve the bearing fault prediction accuracy. This paper presents a kind of new rolling bearing prediction technology, using grey model combined with the extreme learning machine (ELM). The sample is first grey processed to solve the randomness and volatility, and then introduced into the extremely fast learning speed and high generalization accuracy of ELM neural network training. Based on the trained model, the bearing operation state of future time points is analyzed, and the result is compared with the theoretical diagnosis standard of the bearing equipment to realize the fault prediction.

Keywords: grey theory; ELM; rolling element bearings; fault prediction

0 引言

在旋转机械中, 滚动轴承是一种使用极其广泛的易损机械零件, 其运行状况往往能够直接影响到整个机械系统的正常运作, 即便是出现一点小故障也极有可能造成重大事故。根据相关数据显示, 有三成左右的旋转机械设备故障是由于滚动轴承造成的^[1]。因此, 识别滚动轴承运行状态是否正常, 并根据反应运行态势的状态数据进行分析预测, 对防止机械事故、确保机械设备正常运作都极为重要, 同时也能避免和减少不必要的经济损失。在轴承的实际运行中, 反应其工作状态的数据具有非线性和时变的特点^[2], 灰色理论则比较擅长处理这种类型的数据^[3], 波动和不规则的原始数据通过灰色理论的方法能演变成具有一定规律性的生成数列, 在此基础上进行运算就可建立灰色预测模型对未态势进行预测^[4]。但是轴承工作时同时也会受噪声等环境因素的影响, 因此待处理的数据也会收起波动性的影响。灰色理论在处理波动性较大的数据时的拟合效果较差^[5], 因此本文引入具有高线性映射特性的 ELM^[6] 神经网络来帮助拟合数据, 减少预测误差。在工业应用中, 训练数据往往是单个或批量化获得, 而传统 ELM 必须要在获得全部的训练数据之后才可以进行。对此在传统 ELM 上进行改进, 使其能单个或者批量地学习样本

数据, 扩展其应用范围, 减少网络的学习时间。

1 基本原理与方法

1.1 灰色预测模型

在传统的时间序列拟合预测中, 通常我们需要大量的数据作为基础才能得到一个相对确切的预测值^[7]。而灰色预测模型被证明可以用于有效的处理那些数量较少, 信息贫乏的数据, 即便是在表征系统行为特征的原始数据序列较少的情况下, 它也能够通过变换原始数据序列来构建微分方程^[8]。在外部噪声较强时, 考虑到系统会受到外部环境的强烈干扰, 使轴承状态值呈现离乱状态, 将反映这种状态的数据提取出来, 即成为我们要处理的灰色数据, 灰色模型即是针对这种类型的数据建立的模型。灰色模型通常用一般微分方程表示, 其描述了某段时间内系统内部事物发展变化的过程。灰色模型的建立有三个基本的步骤: (1) 对原始数据进行累加生成 (AGO); (2) 对累加生成数据在灰色模型中进行灰色预测; (3) 将数据进行累减生成, 使得在灰色预测中得到的数据重新回到初始形态, 以此得到准确的预测值。在灰色预测模型中最常应用到的是 GM (1, 1) 模型^[9], 下面对 GM (1, 1) 模型的基本建模原理进行阐述。

设原始灰色数据为 $x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)$, 记为:

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$$

对其作累加, 得到:

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)), k = 1, 2, \dots, n$$

收稿日期: 2017-03-30; 修回日期: 2017-04-24。

作者简介: 徐 遥 (1987-), 男, 湖南桃江人, 硕士研究生, 主要从事机器学习, 故障检测方向的研究。

其中: $x^{(1)}(k) = \sum_{i=1}^k (x^{(0)}(i))$

这个过程称之为 AGO (Accumulated Generating Operation), AGO 通过将原始数列转化为递增数列的方式来解决随机性和波动性问题, 由于其有规律可循, 故而很好的提高了构建基于微分方程形式的预测模型的精度。

建立如下 GM (1, 1) 模型:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u \quad (1)$$

灰色预测模型即方程 (1) 所示, 其中采用最小二乘法拟合便可求得 a, u 两个常数的值:

$$\begin{pmatrix} a \\ u \end{pmatrix} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{Y}_n$$

其中:

$$\mathbf{B} = \begin{bmatrix} -\frac{1}{2}(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -\frac{1}{2}(x^{(1)}(2) + x^{(1)}(3)) & 1 \\ \dots & \dots \\ -\frac{1}{2}(x^{(1)}(n-1) + x^{(1)}(n)) & 1 \end{bmatrix},$$

$$\mathbf{Y}_n = (x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n))$$

微分方程 (1) 的解 (定义其为时间响应函数) 如下:

$$\hat{x}^{(1)}(k+1) = (x^{(0)}(1) - \frac{u}{a})e^{-ak} + \frac{u}{a} \quad (2)$$

数列的预测公式如 (2) 式所示, 可以根据公式 (3) 对原始数列的还原预测值进行求解, 其中式 (2) 是对一次累加生成数列的预测值, 这个过程也叫做 IAGO (Inverse Accumulated Generating Operation)。

$$\hat{x}^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1) \quad (3)$$

1.2 ELM 算法

灰色理论与传统神经网络相结合组成的预测模型已在实际中取得应用^[10-12], 但传统 BP 神经网络存在学习速率固定, 收敛速度慢, 训练时间长的缺点^[13]。针对其存在的问题, Huang 等人提出了 ELM 算法^[14], 该算法区别于传统的应用迭代求解的学习算法, 相较于传统算法运行速度更快, 同时具备很高的泛化性能。

假设有 N 个任意的不同的输入样本 $(x_i, t_i) \in \mathbf{R}^n \times \mathbf{R}^m$, 在这里 x_i 是一个 $n \times 1$ 的输入向量, t_i 为 $m \times 1$ 的目标向量。我们把 SLFN 的隐含层节点数定义为 \tilde{N} 个, 则可以得到如下的输出表达式:

$$\sum_{i=1}^{\tilde{N}} \beta_i G(a_i \cdot x_j + b_i) = f_{\tilde{N}}(x_j) \quad (4)$$

$x \in \mathbf{R}^n, a_i \in \mathbf{R}^n, \beta_i \in \mathbf{R}^m$

在此式中, $a_i = [a_{i1}, a_{i2}, \dots, a_{in}]^T$ 是权值向量, 链接第 i 个隐含层节点与输入层, b_i 代表第 i 个隐含层节点偏差; $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ 代表第 i 个隐含层节点的输出权值向量; $a_i \cdot x$ 表示向量 a_i 与输入 x_j 之间的内积。

若在误差为零的情况下此 SLFN 能够逼近 N 个输入样本, 此种情况等效于 $\sum_{j=1}^{\tilde{N}} \|f_{\tilde{N}}(x_j) - t_j\| = 0$, 则存在 a_i, β_i, b_i 使得:

$$\begin{aligned} \sum_{i=1}^{\tilde{N}} \beta_i G(a_i \cdot x_1 + b_i) &= t_1, \\ &\vdots \\ \sum_{i=1}^{\tilde{N}} \beta_i G(a_i \cdot x_N + b_i) &= t_N \end{aligned} \quad (5)$$

现在将式 (5) 合并矩阵格式为:

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T} \quad (6)$$

式 (6) 的 \mathbf{H} 矩阵定义为隐含层输出矩阵。 \mathbf{H} 矩阵的 i 列与第 i 个隐含层节点关于 x_1, x_2, \dots, x_N 的输出向量相对应, 隐含层关于 x_j 的输出则于 \mathbf{H} 矩阵的 j 列相对应。

在 ELM 算法里, 当输入权值及隐含层节点偏差被取随机值设定后, 隐含层矩阵 \mathbf{H} 就被确定为一具体的矩阵, 在这种条件下通过求解输出权值矩阵获得其最小二乘范数解便能起到训练前馈神经网络的作用, 最小二范数可以通过对 $\|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\| = \min_{\boldsymbol{\beta}} \|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\|$ 的求解得到。进一步简化此线性系统的最小二乘范数解为:

$$\bar{\boldsymbol{\beta}} = \mathbf{H}^+ \mathbf{T} \quad (7)$$

其中: \mathbf{H}^+ 为隐含层输出矩阵 \mathbf{H} 的广义逆矩阵。

2 轴承故障预测模型

2.1 针对轴承故障预测的改进 ELM 算法

轴承的工作数据在实际应用中通常是连接到达的^[15], 为使 ELM 能够对这种类型的样本进行学习, 可在原算法基础上作如下优化:

假设一开始到达的训练样本集为 \mathbf{X}_0 , 将其表示为:

$$\mathbf{X}_0 = \{(x_i, t_i)\}_{i=1}^{\tilde{N}}, \mathbf{N}_0 \geq \tilde{N} \quad (8)$$

根据 ELM 算法, 应当将 $\|\mathbf{H}_0 \boldsymbol{\beta} - \mathbf{T}_0\|$ 进行最小化, 通过下式可以计算出 $\|\mathbf{H}_0 \boldsymbol{\beta} - \mathbf{T}_0\|$ 的最小化解:

$$\begin{cases} \boldsymbol{\beta}^{(0)} = \mathbf{K}_0^{-1} \mathbf{H}_0^T \mathbf{T}_0 \\ \mathbf{K}_0 = \mathbf{H}_0^T \mathbf{H}_0 \end{cases} \quad (9)$$

假设又收到一组训练样本, 设该组训练样本为 $\mathbf{X}_1 = \{(x_i, t_i)\}_{i=1}^{\tilde{N}_1}, \mathbf{N}_1$ 为样本的长度, 于是问题便可转化为求解下式的最小化解:

$$\left\| \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix} \boldsymbol{\beta} - \begin{bmatrix} \mathbf{T}_0 \\ \mathbf{T}_1 \end{bmatrix} \right\| \quad (10)$$

对这两组训练样本和进行综合分析, 此时输出权值的表达式变为

$$\boldsymbol{\beta}^{(1)} = \mathbf{K}_1^{-1} \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix}^T \begin{bmatrix} \mathbf{T}_0 \\ \mathbf{T}_1 \end{bmatrix} \quad (11)$$

其中,

$$\mathbf{K}_1 = \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix}^T \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix} = \mathbf{K}_0 + \mathbf{H}_1^T \mathbf{H}_1 \quad (12)$$

为了满足顺序学习的要求, 将 $\boldsymbol{\beta}^{(1)}$ 的表达式变换为与 $\boldsymbol{\beta}^{(0)}$ 、 \mathbf{P}_1 、 \mathbf{H}_1 及 \mathbf{T}_1 相关的函数式:

$$\begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix}^T \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix} \boldsymbol{\beta}^{(0)} = (\mathbf{P}_1 - \mathbf{H}_1^T \mathbf{H}_1) \boldsymbol{\beta}^{(0)} \quad (13)$$

如此便可以将式 (11) 的表达式写为如下形式:

$$\boldsymbol{\beta}^{(1)} = \boldsymbol{\beta}^{(0)} + \mathbf{K}_1^{-1} \mathbf{H}_1^T (\mathbf{T}_1 - \mathbf{H}_1 \boldsymbol{\beta}^{(0)}) \quad (14)$$

通过分析上述过程, 每当有新的样本到达网络, 一般的递归最小二乘法与对最小二乘解的更新的递归算法基本一致^[16]。因此当第 $k+1$ 组数据达到时, 便可用下式表达:

$$\mathbf{X}_{k+1} = \{(x_i, t_i)\}_{i=1}^{k+1} \left(\sum_{i=0}^k N_i \right) + 1 \quad (15)$$

式中， \mathbf{N}_{k+1} 代表第 $k+1$ 组的数据长度。此时的输出权值 β 便可用下面的公式表示：

$$\beta^{(k+1)} = \beta^{(k)} + \mathbf{H}_{k+1}^T (\mathbf{T}_{k+1} - \mathbf{H}_{k+1} \beta^{(k)}) \quad (16)$$

值得注意的是，为了使得修改后的模型达到与原 ELM 同等的学习性能，不损失其原有的识别准确率，隐含层节点的数目和 \mathbf{H}_0 的秩应当保持一致，这就要求在初始化数据时应当合理界定其数量，使其不小于隐含层节点数目。

2.2 灰色 ELM 组合预测模型

从上面介绍的方法我们可以看到，GM (1.1) 灰色预测模型在运用最小二乘法对微分方程进行求解，很难反映出复杂的非线性映射的灰色数列的动态过程。因此我们引入一个简单的解决方案，运用 ELM 神经网络来拟合非线性系统下的输入输出映射。这样一来，灰色预测模型和 ELM 神经网络的结合能显著的减弱数据的不确定性带来的影响，同时也能为组合模型良好的收敛速度和非线性等优点。二者的组合预测模型如图 1 所示。

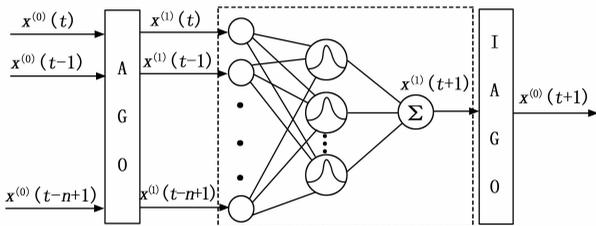


图 1 灰色 elm 预测模型

在封闭的虚线内的 ELM 神经网络包含如下三层：输入层直接向隐藏层传递输入信号；隐藏层对输入层产生响应，在本文中，高斯函数被应用为隐藏层的激活函数，输出层将隐藏层的各个节点线性的链接起来。AGO 和 IAGO 模块分别对应灰色预测中的累加生成和累减生成。

训练数据经过累加生成后进入 ELM 神经网络，通过对训练数据的学习，各层的连接系数在 ELM 算法的帮助下进行了不断的修正，最后将神经网络输出的数据作累减生成处理，得到最终的预测数据。

3 轴承故障预测实例

要进行故障预测，如何确定诊断标准是首要解决的问题，本文运用测振仪对发电机轴承的振动加速度值进行监测，在积累一定的历史数据后，以概率统计为依据，结合实践经验并根据轴承实际运行情况，制订出中频发电机的诊断标准如表 1 所示。

表 1 轴承故障诊断标准

状态区间	诊断值/g	评价	应对措施
小于 $(\mu + \sigma)$	小于 1.55	正常	无
$(\mu + \sigma) \sim$	1.55 ~	有待	允许
$(\mu + 2.5\sigma)$	2.36	观察	工作
$(\mu + 2.5\sigma) \sim$	2.36 ~	重点	状态
$(\mu + 4\sigma)$	1.37	观察	维修
大于 $(\mu + 4\sigma)$	大于 3.17	故障	停机

表 1 中 μ 为轴承设备在一定的正常工作时间内运行所得值的平均值， σ 为该时间内正常值的标准差， σ 的系数可根据轴承设备的重要性和实际应用情况进行确定，轴承在不同运行条件下 σ 系数应结合现场的具体情况进行修正。

选取一定量的监测数据作为训练数据，采用 MATLAB 作为数据处理平台，应用灰色 ELM 神经网络模型进行训练，步骤如下：

- 1) 建立灰色 AGO 模型对训练数据进行累加生成处理，以达到弱化数据随机性的目的。
- 2) 将经过灰色处理的数据代入到 ELM 神经网络模型中，在对网络输入层的权值和隐含层的偏差取随机值后，便可得到隐含层的具体矩阵。
- 3) 根据优化后的 ELM 算法，求解输出权值矩阵，在这个求解过程中神经网络各层的系数被不断的进行修正，直到在得到输出权值矩阵的最小二乘范数解后修正过程才结束，完成对 ELM 神经网络训练。
- 4) 神经网络输出的数据最终流入到灰色 IAGO 模型中，经过累减生成处理最终将数据还原到灰色处理以前的形态。

为验证上述轴承预测模型的效果，选取一定量接近重点观察区域的监测数据作为测试样本，代入到训练好的模型中进行预测，同时为起到对比参照的作用，将其也代入到灰色 GM (1.1) 模型中进行预测，仿真结果如图 2 和图 3 所示，通过比较可以看到，灰色 GM (1.1) 预测模型基本反映了轴承从正常工作阶段到重点观测阶段的走势，但有些点的误差较大；而灰色 ELM 神经网络模型的预测值更贴合实际值，精确的贴合了数据变化的趋势，显著提高了预测精度。

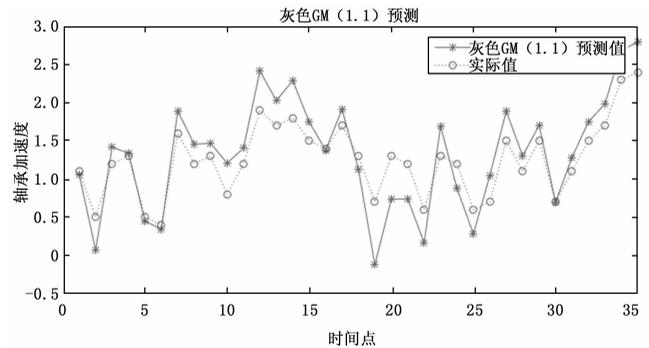


图 2 灰色预测仿真结果

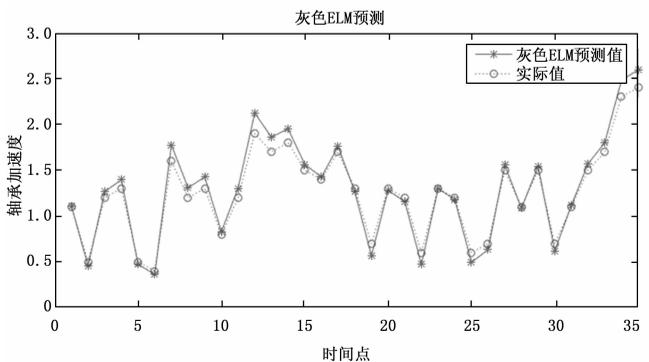


图 3 灰色 ELM 预测仿真结果

确。至此,问题解决。

4) 说明。

地面试验是多个环节配合完成,实时数据处理软件采用自主研制,试验中必须严格按照软件要求解析:参数名长度不大于24字节,个数不大于500个;试验中服务器与客户端必须连接交换机,组成局域网(服务器端数据是通过网络分发给客户端的),且端口号服务器端与客户端必须一致;007卡卡槽与实时/预处理设备连接可能会出现虚连接,导致数据无法传输。

5 结束语

文章针对飞行载荷实测的海量参数难以测试问题,提出了一种通道复用的方案。针对通道复用过程中的技术难点,采用点、线、面的方法进行排查和分析,验证了该方案的可行性,良好的解决了海量数据难以实时测量问题。该方案现已成功运

(上接第65页)

在将图3中灰色ELM神经网络得到的预测值和制定的故障诊断标准作分析比较后,可以发现在时刻 $k=34$ 时候轴承加速度的预测值已经达到 2.42 g ,处于重点观察区域,为避免事故发生减少不必要的损失,应当针对实际情况采用有效的维修手段,杜绝其工作到停机线附近的可能。

ELM跟传统的神经网络相比训练速度和泛化能力都很优秀,而在实际训练过程中ELM因左侧权值是随机确定的,因此在稳定性和模型精度上还有待改进。ELM和灰色理论的结合,是利用了组合预测方法结合单项预测方法的优势,综合应用单项预测所提供的信息,从而有效的减少了预测误差,提高了预测精度。在轴承故障预测领域中,灰色理论和ELM神经网络的结合所建立的轴承故障预测模型,有效的解决了轴承故障类别多样化所引发的单一模型预测精度有限的问题,拓宽了使用范围。在原始样本数据有限的情况下依然能够通过获取高精度的拟合数据来提高预测的精度,实现对轴承的运行趋势的预测,在实际的工业生产中有着现实的意义。

4 结论

本文分别应用灰色预测模型与灰色ELM神经网络模型,对提取的反应轴承振动加速度的少量样本数据进行分析预测,从仿真结果可以看出,灰色ELM神经网络模型与灰色预测模型相比,预测效果更好,精度更高,跟实际数据更为贴合。对轴承特征量加速度的研究,不仅为工作状态中的故障的预测提供一种可行有效的方法,也能为轴承寿命及可靠性等性能指标的提供一定的参考价值;同时预测选取的特征量也不局限于加速度,亦可将采用其他特征量,运用到灰色ELM神经网络预测模型的中实现其故障预测。

参考文献:

- [1] 陈进. 机械设备振动监测与故障诊断 [M]. 上海: 上海交通大学出版社, 1999.
- [2] 邵毅敏, 涂文兵. 深沟球轴承三维非线性时变振动特性研究 [J].

用于在某型号飞行载荷海量数据实测中,并取得了预期的效果。

参考文献:

- [1] 周自全. 飞行力学与飞行试验 [J]. 飞行力学, 2002, 20 (2): 212-217.
- [2] 中国飞行试验研究院. 飞行试验指南 [M]. 北京: 航空工业出版社, 2010.
- [3] 霍朝晖, 覃杨森, 祈春. 飞行试验机载关键参数快速处理系统设计 [J]. 现代电子技术, 2012, 36 (5): 121-124.
- [4] 袁炳南, 张建琳. PCM与网络数据采集系统技术分析 [J]. 测控技术, 2009, 28 (4): 29-31.
- [5] Range Commanders Council. IRIG Standard 106-01 [S]. New Mexico: Range Commanders Council, 2001, 88002-5110.
- [6] 王亮, 滕克难, 吕卫民, 等. 基于粒子群算法的非线性时变参数离散灰色预测模型 [J]. 统计与决策, 2015, (12): 16-19.
- [7] Lin Y H, Chiu C C, Lee P C, et al. Applying fuzzy grey medication model on inflow forecasting [J]. Eng. Appl. Artif. Intell, 2012 (25): 734-743.
- [8] 张文斌, 郭德伟, 普亚松, 等. 谐波窗分解样本熵与灰色关联度在转子故障识别中的应用 [J]. 中国电机学报, 2013 (21): 132-137.
- [9] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: theory and applications [J]. Neurocomputing, 2006 (70): 489-501.
- [10] 魏二虎, 李智强, 龚光裕, 等. 极移时间序列模型的拟合与预测 [J]. 武汉大学学报(信息科学版), 2013, (12): 1420-1424.
- [11] 王大鹏. 灰色预测模型及中长期电力负荷预测应用研究 [D]. 武汉: 华中科技大学, 2013.
- [12] 刘思峰, 曾波, 刘解放, 等. GM(1, 1)模型的几种基本形式及其适用范围研究 [J]. 系统工程与电子技术, 2014 (3): 501-508.
- [13] Alvisi S, Franchini M. Grey neural networks for river stage forecasting with uncertainty [J]. Phys. Chem. Earth, 2012, 42-44, 108-118.
- [14] 陈帅, 王勇, 吕丰, 等. 基于灰色理论和神经网络的短期电力负荷预测 [J]. 上海电力学院学报, 2013 (6): 527-531.
- [15] 廉书林. 基于灰色理论与神经网络的油液污染和机械磨损状况研究 [D]. 郑州: 河南工业大学, 2014.
- [16] 刘春艳, 凌建春, 寇林元, 等. GA-BP神经网络与BP神经网络性能比较 [J]. 中国卫生统计, 2013 (2): 173-176.
- [17] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: theory and applications [J]. Neurocomputing, 2006, 489-501.
- [18] 朱可恒. 滚动轴承振动信号特征提取及诊断方法研究 [D]. 大连: 大连理工大学, 2013.
- [19] Edwin K, Stanislaw H Z. An introduction to optimization [J]. Antennas and Propagation Magazine IEEE, 2001, 38 (2): 1-40.