

# Fault Prediction of Rolling Bearing based on Frequency Domain Analysis

Jie YUAN

Chongqing Maritime Bureau, Chongqing, 400000, China

**Abstract:** A failure prognosis model which fuses the frequency domain analysis of vibration signals, the reliability analysis and BP neural network is proposed. Primarily, the frequency domain analysis of antifriction bearing vibration data is conducted by discrete Fourier transform. And then, the reliability of antifriction bearings is assessed by making use of Weibull distribution based on the maintenance data. Finally, BP neural network is trained and examined by Matlab using sample data.

**Keywords:** Failure prognosis; Frequency domain analysis; Reliability analysis; BP neural network; Antifriction bearing

## 基于频域分析的滚动轴承故障预测

袁杰

海装重庆局 重庆市 中国 400000

**摘要:** 提出了一种融合振动信号频域分析、可靠性分析和 BP 神经网络的故障预测模型。首先通过离散傅立叶变换对滚动轴承振动数据进行频域分析；其次利用维修记录根据威布尔分布对滚动轴承可靠性进行评估；最后利用 MATLAB 软件和样本数据对基于故障预测模型进行训练和检验。

**关键词:** 故障预测；频域分析；可靠性分析；BP 神经网络；滚动轴承

### 1 引言

随着机械设备结构和功能日益复杂，使用者对设备安全性和可靠性要求也越来越高。维修技术逐步由事后维修和定期维修向视情维修发展，故障预测是视情维修的基础，设备及零部件故障预测方法研究愈发受到重视。滚动轴承在机械设备中应用广泛，各种滚动轴承故障机理和故障现象存在很多相似指出，本文研究具有一定普遍性。本文以滚动轴承为例，提出了一种融合振动信号频域分析、可靠性分析和 BP 神经网络的故障预测模型。

### 2 数据采集

#### 2.1 特征参数选取

通过实验在线监测离心泵运行参数，实验所用离心泵运行时间为 18 个月，整体性能良好。离心泵滚动轴承磨损通常伴有泵体振动加剧和电机输入功率增大的现象，因此选取离心泵的流量、扬程、电机输入功率和振动信号作为离心泵特征参数。在线监测离心泵流量、扬程、电机输入功率等参数。使用便携式振

动数据采集器采集滚动轴承水平方向和垂直方向（记为 x 方向和 y 方向）的振动数据并储存。通过维持流量和扬程不变来控制实验变量，通过分析电机输入功率和振动信号的变化可以了解离心泵轴承的性能。

#### 2.2 振动信号采集

离心泵一套滚动轴承从安装运行到出现功能性故障时间大约为两年，由于实验时间和实验条件限制，不能在两年甚至更长时间内全程跟踪记录滚动轴承振动数据。采取为离心泵更换三套生产厂家相同型号相同但运行时间长短不同的滚动轴承的方法，以此模拟安装同一套滚动轴承的离心泵随运行时间增长特征参数的变化。实验所选用的三套滚动轴承分别为：一套全新滚动轴承（记为 1#轴承）、一套运行 18 个月的正常滚动轴承（记为 2#轴承）和一套运行两年磨损严重的故障滚动轴承（记为 3#轴承）。利用振动数据采集器采集振动信号的位移峰峰值。

为了控制实验变量，维持离心泵流量在 100 1m<sup>3</sup>/h 范围内，扬程在 34 0.5m 范围内。更换上述三套滚动轴承，分别记录相应情况下 x 和 y 方向

的位移峰峰值和电动机输入功率。

离心泵的轴额定转速为 3000r/min。在离心泵安装全新滚动轴承的情况下，使用振动数据采集器分别以 1000Hz、2000Hz 采样频率采集振动信号。振动数据采集器以 1000Hz、2000Hz 采样频率所采集到振动信号的时域数据和频域数据分别如图 1 所示，观察可

知两种采样情况下所采集数据的差别不大。根据 Shannon 采样定理，采样频率  $f_s$  大于等于信号中最高频率成分的频率  $f_{max}$  的 2 倍，信号中的频率成分信息就不会丢失。根据振动信号频域数据，信号的最高频率成分的频率为 150Hz，1000Hz 采样频率可以满足采集数据保留振动有效信息的要求。

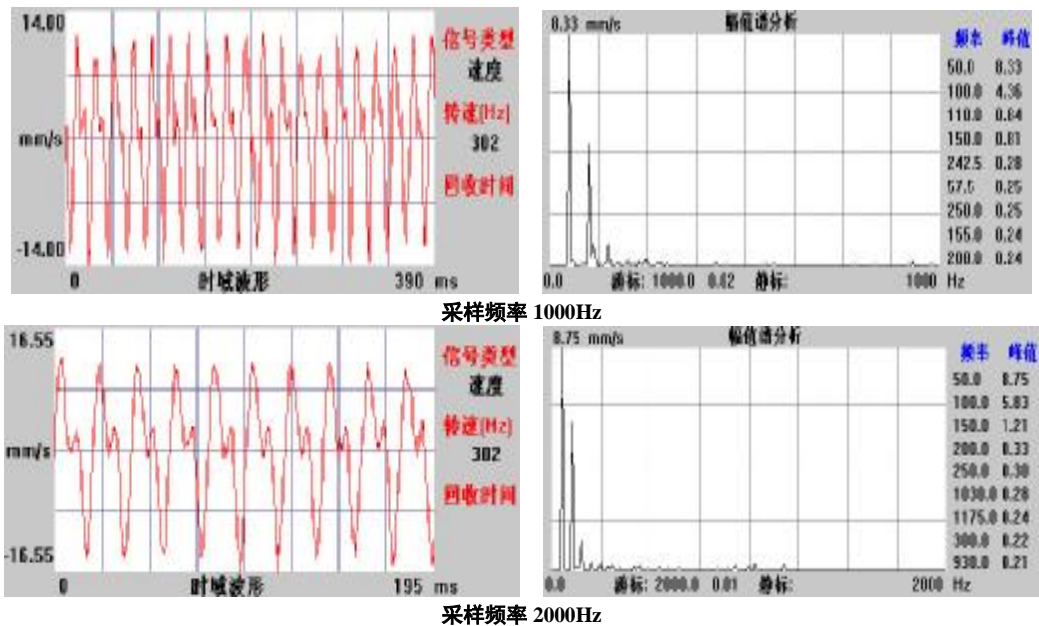
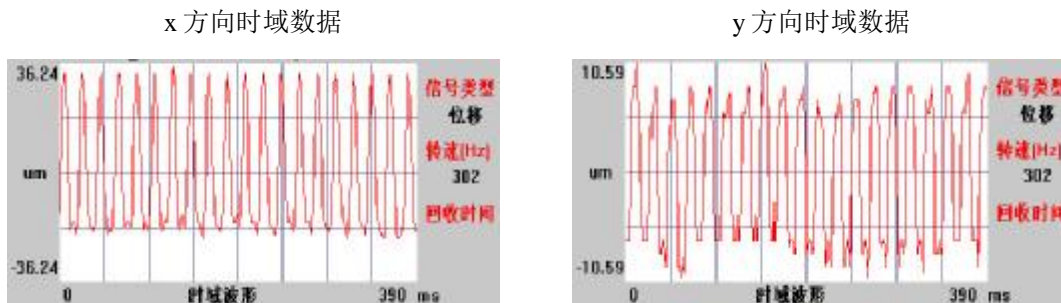


图 1. 不同采用频率对比图

振动数据采集器采集振动数据时，为了提高采集数据的精度，探头采用磁座固定方式，并且保持探头固定位置不变。在离心泵安装 1#、2#和 3#轴承的情况下，分别采集离心泵振动信号 x 和 y 方向的位移峰峰值。每次更换滚动轴承再次启动离心泵后，使离心泵稳定运行 3-5 分钟再进行数据采集，每组实验振动

数据各采集 5 次，每两次采集间隔 1 分钟。离心泵安装 3#轴承时，为防止轴承过热，每运行 15-20 要停机散热 2-3 分钟。通过对采集到的振动信号时域数据进行观察，每组实验所采集到的 5 组时域数据图像差别不大。图 2 显示各组实验采集到的振动数据中的第 1 组数据。



1#轴承

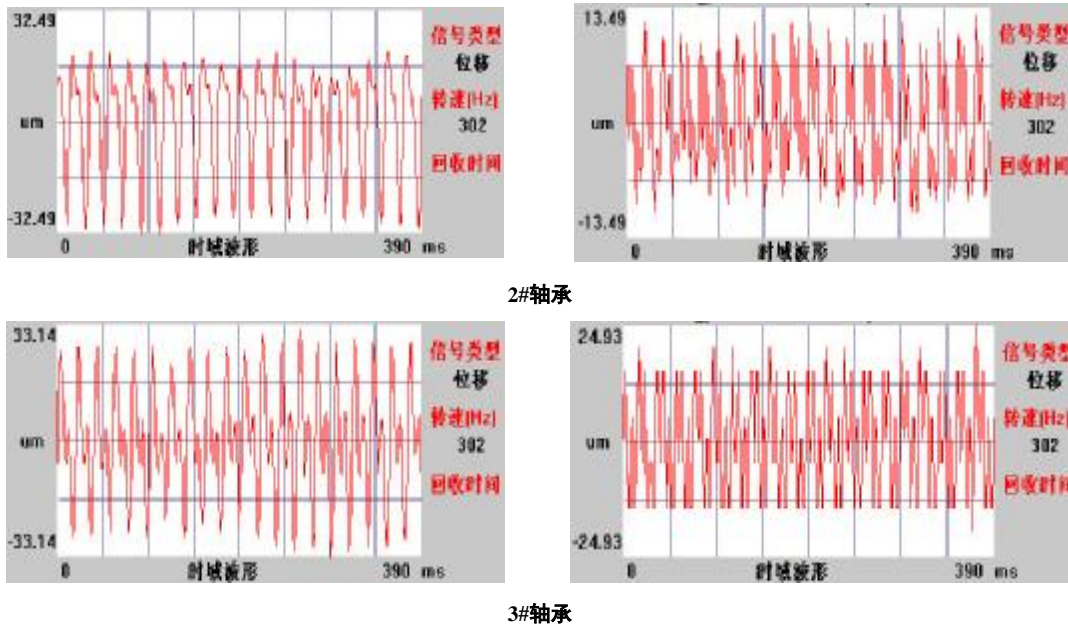


图 2. 离心泵振动信号位移峰峰值时域数据

比较采集时域数据可以发现：随着轴承磨损程度加深，离心泵振动信号位移峰峰值数据的振幅有所增大，而且数据波动的不规则性加剧，y 方向相比 x 方向振动不规则性更为明显。

离心泵电动机输入功率是实时变化的，但在同组实验中离心泵电动机输入功率变化不大，每次振动数据采集开始时记录离心泵电动机输入功率。

### 3 数据处理

#### 3.1 离散傅里叶变换

时域分析能够直观明了地观察离心泵振动信号的幅值大小以及幅值变化规律，而且当轴承磨损到一定程度后，通过时域分析可以较为准确地确定故障设备。但是时域分析不能有效地提供振动信号中可能包含的简谐信号、周期信号或脉冲信号等关键信息，而这些振动信息对故障预测更为有效<sup>[1-2]</sup>。

假设原始的振动信号为  $x(t)$ ，根据傅里叶变换公式，振动信号  $x(t)$  的傅里叶变换为

$$X(f) = \int_{-\infty}^{+\infty} x(t)e^{-j2\pi ft} dt \quad (1)$$

振动数据采集过程中，虽然原始振动信号是连续的，但是振动数据采集是间断的。在采样频率

1000Hz 的情况下，振动数据采集间隔  $\Delta t$  为 0.001s。

假设  $i$  为所有整数，一个频率为 1000Hz 的单位脉冲采样函数为

$$z(t) = \sum_{i=-\infty}^{+\infty} d(t+i\Delta t) \quad (2)$$

单位脉冲采样函数  $z(t)$  的傅里叶变换为

$$Z(f) = \frac{1}{\Delta t} \sum_{i=-\infty}^{+\infty} d(f - \frac{i}{\Delta t}) \quad (3)$$

采集到的振动信号  $x'(t)$  可以看作是实际振动信号  $x(t)$  乘以单位脉冲采样函数  $z(t)$

$$x'(t) = x(t) z(t) \quad (4)$$

根据傅里叶变换乘积定理可知，采集信号  $x'(t)$  的傅里叶变换为

$$X'(f) = X(f) * Z(f) = \Delta t \sum_{i=-\infty}^{+\infty} x(i\Delta t)e^{-j2\pi fi\Delta t} \quad (5)$$

#### 3.2 振动数据离散傅里叶变换

对离心泵振动信号位移峰峰值时域数据进行离散傅里叶变换，得到振动信号位移峰峰值频域数据如图 3 所示。

x 方向频域数据

y 方向频域数据

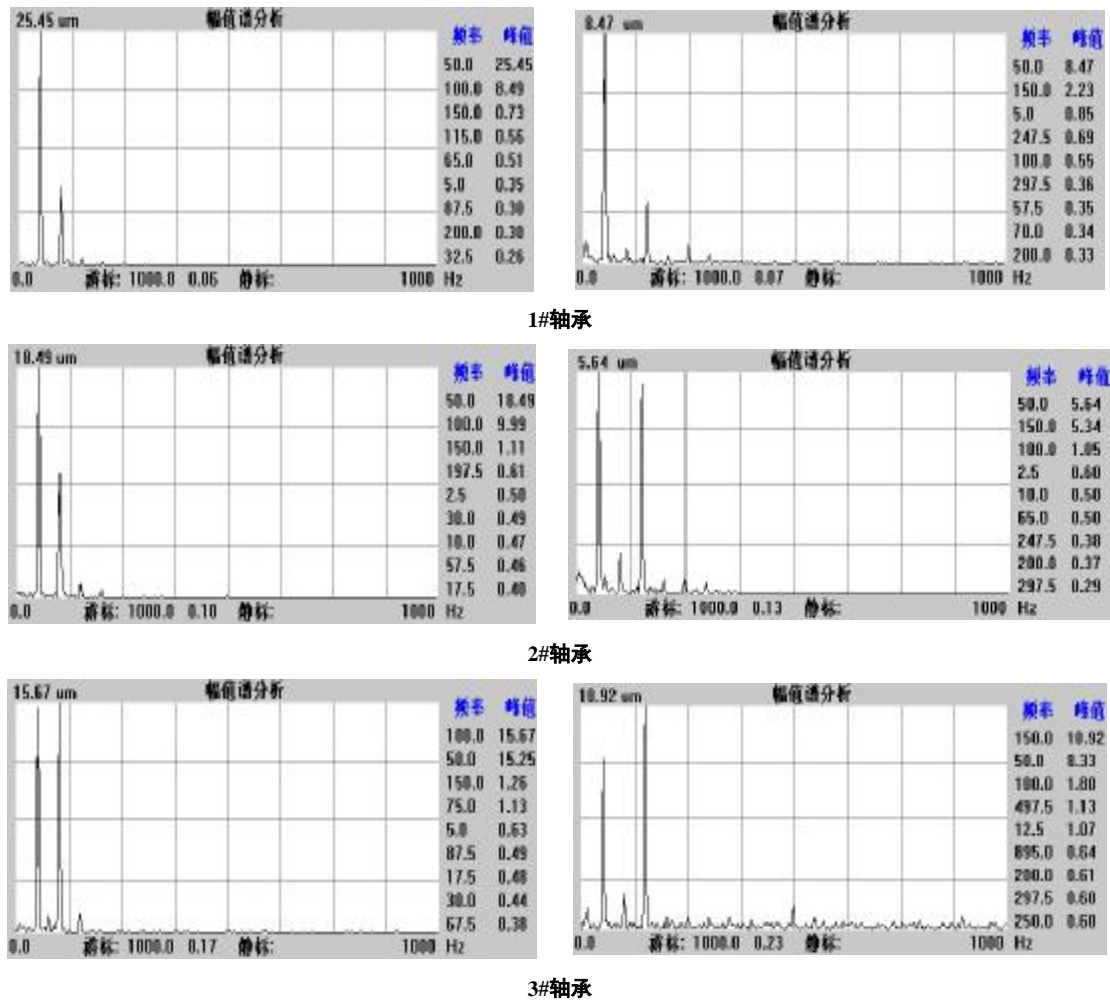


图 3. 离心泵振动信号移峰峰值频域数据

分析比较离心泵的振动信号频域数据可以发现：每组实验 5 次采集到的数据经过离散傅里叶变换频域数据的变化规律相似；随着滚动轴承磨损程度加深，x 方向的位移峰峰值频域数据 1 倍频（50Hz）幅值明显减小，2 倍频（100Hz）幅值明显增大；y 方向的位移峰峰值频域数据 1 倍频幅值明显减小，2 倍频和 3 倍频（150Hz）幅值明显增大；位移峰峰值 x 方向 2 倍频幅值和 y 方向 3 倍频幅值分别是 3#轴承位移峰峰值 x 方向和 y 方向频谱数据中幅值极大值。根据文

献资料<sup>[3-5]</sup>表明：滚动轴承磨损在频域数据上通常反映为 2 倍频幅值增大，同时伴随着 3 倍频到高频幅值增大。本文振动信号频域分析也印证了上述结论。

### 3.3 数据标准化处理

选取电动机输入功率、振动信号 x 方向位移峰峰值频域数据 2 倍频幅值和振动信号 y 方向位移峰峰值频域数据 3 倍频幅值作为 BP 神经网络的输入值，离心泵参数数据如表 1 所示。

表 1. 离心泵参数数据

轴承编号	实验次数	电动机输入功率 (W)	x 方向位移峰峰值频域数据 2 倍频幅值 (μm)	y 方向位移峰峰值频域数据 3 倍频幅值 (μm)
1#轴承	1	19538	9.49	2.23
	2	19546	9.31	2.36
	3	19572	9.72	2.15
	4	19586	9.58	2.29



	5	19578	9.45	2.43
2#轴承	1	19876	9.99	5.34
	2	19893	10.18	4.88
	3	19849	9.81	5.21
	4	19868	9.75	5.56
	5	19873	9.94	5.32
3#轴承	1	21126	15.67	10.92
	2	20982	15.98	10.27
	3	20999	14.92	10.54
	4	21114	15.36	10.71
	5	20973	15.17	10.66

电动机输入功率和振动信号频域数据幅值不但单位不同，而且数值最大相差约 4 个数量级,因此需要数据进行离差标准化处理。某参数数据 s 的离差标准化转换函数为

$$s' = \frac{s - s_{\min}}{s_{\max} - s_{\min}} \quad (6)$$

$s_{\max}$  表示 s 参数数据最大值， $s_{\min}$  表示 s 参数数据最小值， $s'$  表示参数数据 s 的标准化形式。标准化参数数据如表 2 所示，数据均精确到 0.01。

表 2. 离心泵参数标准化数据

轴承编号	实验次数	电动机输入功率	x 方向位移峰值频域数据 2 倍频幅值	y 方向位移峰值频域数据 3 倍频幅值
1#轴承	1	0.00	0.03	0.01
	2	0.01	0.00	0.02
	3	0.02	0.06	0.00
	4	0.03	0.04	0.02
	5	0.03	0.02	0.03
2#轴承	1	0.21	0.10	0.36
	2	0.22	0.13	0.31
	3	0.20	0.07	0.35
	4	0.21	0.07	0.39
	5	0.21	0.09	0.36
3#轴承	1	1.00	0.95	1.00
	2	0.91	1.00	0.93
	3	0.92	0.84	0.96
	4	0.99	0.91	0.98
	5	0.90	0.88	0.97

## 4 滚动轴承健康程度评估

### 4.1 最小二乘法估算威布尔分布参数

威布尔分布又称 Weibull 分布，是可靠性理论的基础。大量的研究及工程实践证明<sup>[6-9]</sup>，滚动轴承累积故障分布函数服从二参数 Weibull 分布

$$F(t) = 1 - \exp\left(-\left(\frac{t}{h}\right)^b\right) \quad (7)$$

其中 t 为时间变量，η 和 β 是待定参数。由 Weibull 累积故障分布函数得到的滚动轴承可靠性函数为

$$R(t) = \exp\left(-\left(\frac{t}{h}\right)^b\right) \quad (8)$$

对二参数 Weibull 分布函数进行对数转换得到

$$\ln \ln\left(\frac{1}{1-F(t)}\right) = b \ln(t) - b \ln(h) \quad (9)$$

令  $Y = \ln \ln(1/(1-F(t)))$ ,  $X = \ln(t)$ ,  $B = -b \ln(h)$ ，则二参数 Weibull 分布函数对数式可以写成

$$Y = bX + B \quad (10)$$

假设对 n 套滚动轴承进行可靠性实验，n 套滚动轴承失效时间分别是  $t_1, t_2, \dots, t_n$ ，其中  $t_1 < t_2 < \dots < t_n$ 。取 i 为 [1,n] 区间内任意自然数，根据 GB/T24607-2009，可以用中位秩法估算滚动轴承  $t_i$  时刻对应的累积故障分布函数近似为

$$F(t_i) = \frac{i-0.3}{n+0.4} \quad (11)$$

通过  $t_i$  和  $F(t_i)$  可以得到相应  $Y_i$  和  $X_i$  的值。利用

最小二乘法对参数  $\beta$  和  $B$  进行估算。将式 (10) 转换为关于参数  $\beta$  和  $B$  的二元函数  $G(\beta, B)$ ，最小二乘法即求得使函数  $G$  值最小的参数  $\beta$  和  $B$  的值。

$$G(b, B) = \sum_{i=1}^n (Y_i - bX_i - B)^2 \quad (12)$$

令函数  $G$  分别对参数  $\beta$  和  $B$  求偏导数

$$\frac{\partial G}{\partial B} = -2 \sum_{i=1}^n (Y_i - bX_i - B)X_i \quad (13)$$

$$\frac{\partial G}{\partial B} = -2 \sum_{i=1}^n (Y_i - bX_i - B) \quad (14)$$

令式 (13) 和式 (14) 等于 0，求得参数  $\beta$  和  $B$  的最优解  $b$  和  $B$  为

$$\hat{b} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} \quad (15)$$

$$B = \bar{Y} - b\bar{X} \quad (16)$$

其中

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (17)$$

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \quad (18)$$

### 4.2 基于威布尔分布的滚动轴承健康程度评估

根据泵厂相同型号离心泵滚动轴承更换记录，计算最近 10 台离心泵从出厂/上一次更换滚动轴承到再次更换滚动轴承的时间，即滚动轴承寿命。

对滚动轴承进行健康程度评估基于以下四点假设：第一，10 台离心泵的运行环境相似，泵的其它零部件性能良好；第二，10 台离心泵的运行强度相似，即泵运行的流量和扬程时相似，每天运行时间也相近；第三，10 台离心泵更换滚动轴承时，滚动轴承性能相似；第四，轴承出现故障后，泵厂及时派维修人员进行更换。

将 10 台泵滚动轴承寿命由短到长分别记作  $t_1, t_2, \dots, t_{10}$ 。令  $i$  代表  $[1, 10]$  中任一自然数，根据公式 3-5 估算滚动轴承故障概率密度  $F(t_i)$ ，根据  $X$  和  $Y$  的定义式分别计算  $X_i$  和  $Y_i$ ，计算结果如表 3 所示，表中  $F(t_i)$ 、 $X_i$  和  $Y_i$  数值均精确到小数点后两位。

表 3. 健康程度相关计算值

离心泵	1#	2#	3#	4#	5#	6#	7#	8#	9#	10#
轴承寿命 (月)	23.5	23.8	23.8	24.3	24.5	24.5	24.5	26.3	26.8	28.3
F(ti)	0.07	0.16	0.26	0.36	0.45	0.55	0.64	0.74	0.84	0.93
Yi	-2.67	-1.72	-1.2	-0.82	-0.51	-0.23	0.03	0.30	0.59	0.99
Xi	3.16	3.17	3.17	3.19	3.20	3.20	3.20	3.27	3.29	3.34

计算得到  $\bar{X}$  和  $\bar{Y}$  分别为 -0.52 和 3.22，进而根据公式 3-9、3-10 得到  $\eta$  和  $\beta$  分别为 15.63 和 -50.83，最后根据  $B$  的定义式得到  $\eta$  为 25.84。

综上所述由 Weibull 故障分布函数得到的离心泵滚动轴承健康程度函数为

$$R(t) = \exp\left(-\left(\frac{t}{25.84}\right)^{15.63}\right) \quad (19)$$

分别将  $t=0$ 、 $t=18$  和  $t=24$  分别代入公式 3-13，计算得到 1#、2#和 3#轴承的可靠性分别为 1、0.99 和 0.64。

## 5 基于 BP 神经网络滚动轴承故障预测模型

### 5.1 BP 神经网络概述

人工神经网络 (Artificial Neural Network, ANN) 简称神经网络，是一种由生物神经网络结构

简化而来的具有信息处理分析功能的数学模型。人工神经网络类似于生物神经网络，是由人工神经元（简称为神经元或节点）以不同的拓扑结构连接而成。神经网络具备神经元、拓扑结构和学习算法三大特征要素。误差反向传播 (Back Propagation, BP) 神经网络简称 BP 神经网络，是一种应用广泛的多层前向神经网络。BP 神经网络输出值前向传播，其误差反向传播。

BP 神经网络中神经元常用的传递函数有 sigmoid 型函数 (包括 log-sigmoid 函数、tan-sigmoid 函数等) 和线性函数 purelin，上述三种传递函数如图 5 所示。log-sigmoid 函数输出值在  $[0, 1]$  区间内；tan-sigmoid 函数输出值在  $[-1, 1]$  区间内；purelin 线性函数输出值就没有限制。

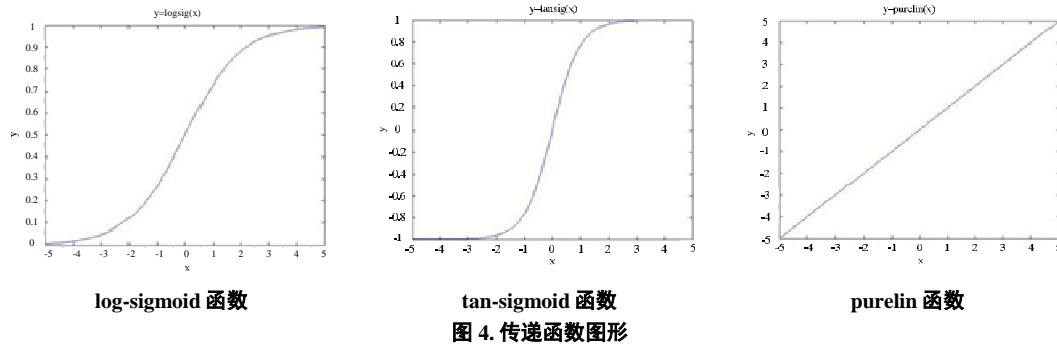


图 4. 传递函数图形

BP 神经网络的优点是在隐层数目和隐层节点数目合适的情况下，其可以逼近任意非线性映射关系。但是如果隐层数目和隐层节点数目不合理，BP 神经网络收敛速度较慢，而且容易陷入局部最小。具有一个隐层的 BP 神经网络结构可如图 6 所示。

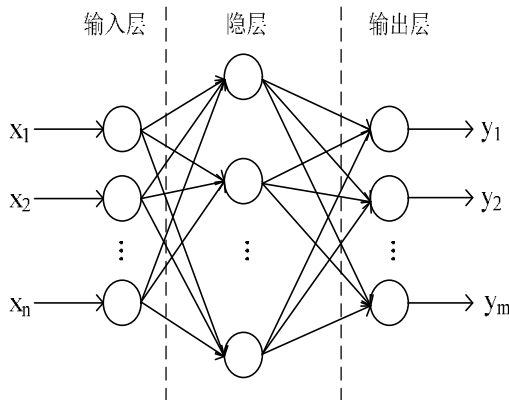


图 5. BP 神经网络结构

BP 神经网络采用有监督学习算法，训练中需要向 BP 神经网络提供输入值向量和期望值向量。BP 神经网络通常采用 BP 算法或其改进算法训练，根据误差性能调整 BP 神经网络的权值和偏差，最终实现均方误差最小化。BP 学习算法是利用误差函数的梯度，沿着误差减小最快的方向对权值进行调整。

### 5.2 BP 神经网络训练

利用 MATLAB 软件建立 BP 神经网络故障预测模型。因为离心泵的健康程度是模糊概念，所以健康程度数值精确到 0.01 就可以满足实际需求。用标准化处理后的 15 组样本数据中的 12 组对 BP 神经网络进行训练，训练数据如表 4 所示。

表 4. BP 神经网络训练样本

样本序号	电动机输入功率	x 方向位移峰峰值频域数据 2 倍幅值	y 方向位移峰峰值频域数据 3 倍幅值	健康程度
1	0.00	0.03	0.01	1.00
2	0.01	0.00	0.02	1.00
3	0.02	0.06	0.00	1.00
4	0.03	0.04	0.02	1.00
5	0.21	0.10	0.36	0.99
6	0.22	0.13	0.31	0.99
7	0.20	0.07	0.35	0.99
8	0.21	0.07	0.39	0.99
9	1.00	0.95	1.00	0.64
10	0.91	1.00	0.93	0.64
11	0.92	0.84	0.96	0.64
12	0.99	0.91	0.98	0.64

采用 Fletcher-Reeves 共轭梯度算法作为 BP 神经网络的学习算法[10]。Fletcher-Reeves 共轭梯度算法沿着误差共轭梯度方向对权值进行调整，是改进 BP

学习算法的其中一种，其收敛速度一般比标准 BP 学习算法更快。

分别对三层和四层 BP 神经网络进行训练，选择

其中最优方案。对三层 BP 神经网络进行训练时，隐层分别设置有 3、4、5、6 个节点。对四层 BP 神经网络进行训练时，第一隐层设置 5 个节点，第二隐层分别设置 4、5、6、7 个节点。BP 神经网络隐层采用 log-sigmoid 函数作为传递函数，输出层采用 purelin

函数作为传递函数，学习步长为 0.03。利用训练样本分别对上述 8 种 BP 神经网络进行训练，设置训练目标误差为  $10^{-5}$ ，训练效果如图 7 所示。第一隐层和第二隐层均设 5 个节点的四层 BP 神经网络预测效果最好。

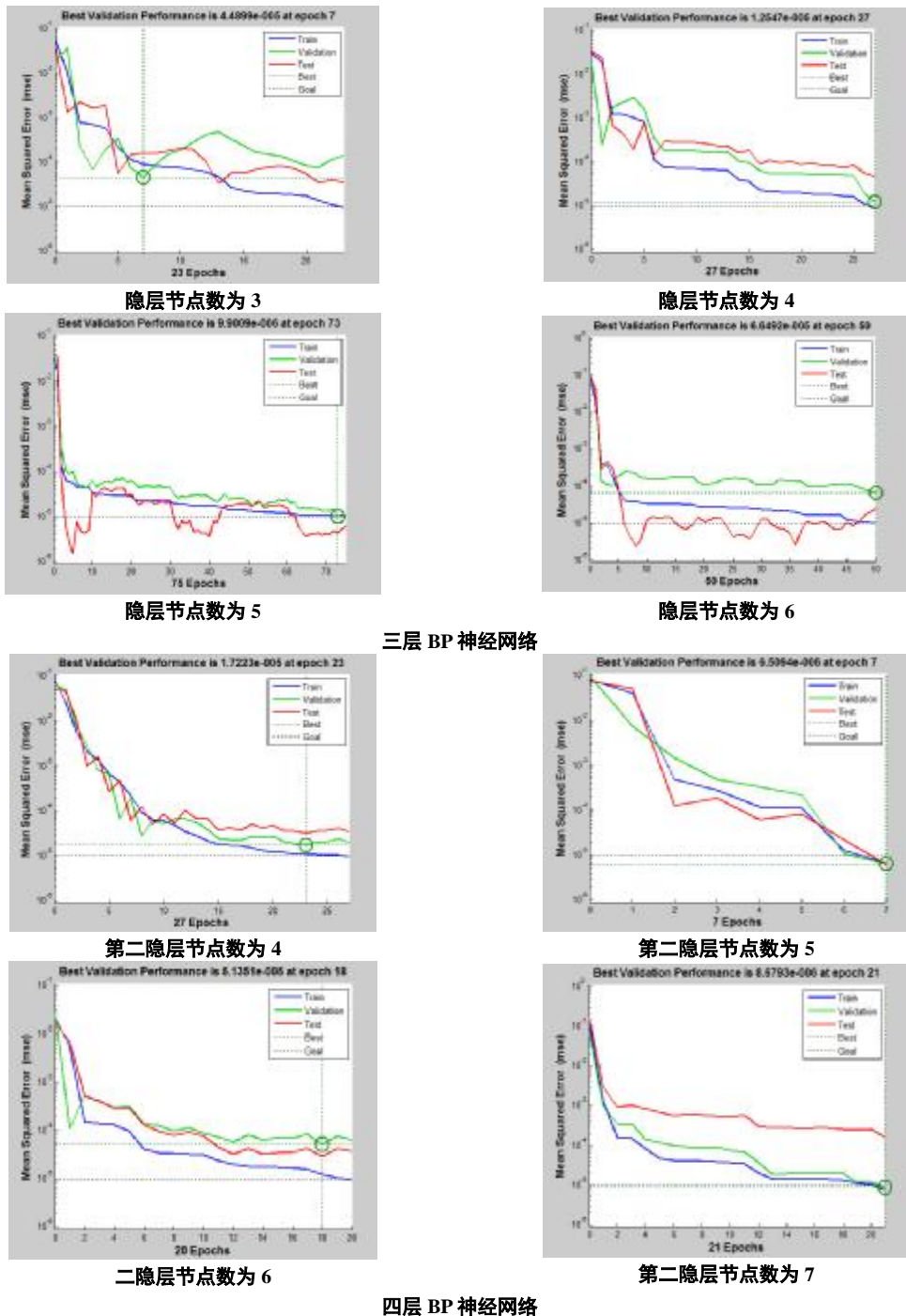


图 6. 不同隐层节点数训练效果



### 5.3 BP 神经网络验证

利用剩余 3 组样本数据对训练后的 BP 神经网络进行验证, BP 神经网络的输出值与轴承健康程度可靠性对比如表 5 所示。为了检验 BP 神经网络输出值是否在规定取值区间[0,1]内, 加入了 4 号和 5 号样本

数据。由检验结果可知: BP 神经网络的输出值范围在规定区间[0,1]内, 而且输出值与轴承健康程度期望值误差不超过 0.5%。说明通过样本数据训练 BP 神经网络达到了既定的预测效果, 由 BP 神经网络得到的离心泵轴承可靠性较为可信。

表 5. 期望值与输出值对比

样本序号	电动机输入功率	x 方向位移峰峰值频域数据 2 倍频幅值	y 方向位移峰峰值频域数据 3 倍频幅值	健康程度期望值	神经网络输出值
1	0.03	0.02	0.03	1.00	1.00
2	0.21	0.09	0.36	0.99	0.99
3	0.90	0.88	0.97	0.60	0.60
4	0.00	0.00	0.00	1.00	1.00
5	1.00	1.00	1.00	0.60	0.60

### 5.4 滚动轴承维修临界值估算

上述 BP 神经网络输出值是滚动轴承可靠性, 要实现滚动轴承故障预测需要设定可靠性维修临界值。

离心泵滚动轴承寿命最短为 23 个月零两周。考虑到滚动轴承损坏速度逐步加快, 将  $t=23$  代入公式 3-13, 得到  $R=0.85$ , 该滚动轴承可靠性维修临界点设定为 0.85。利用训练后的 BP 神经网络对该离心泵滚动轴承进行故障预测, 一旦 BP 神经网络输出值达到 0.85 就说明该离心泵滚动轴承需要维修更换。

### References (参考文献)

- [1] 何正嘉,陈进,王太勇,褚福磊.机械故障诊断理论及应用[M].北京:高等教育出版社,2010.
- [2] Suehiro N. Fast algorithms for the DFT and other sinusoidal

transforms [J].IEEE Trans. on ASSP, 1986, ASSP-34(3) : 642-644.

- [3] 袁周,黄志坚.工业泵常见故障及维修技巧[M].北京:化学工业出版社,2008.
- [4] 张权,周盼,孙晓东,率志军,陈宝星.基于振动测试数据的多级离心泵故障诊断[J].水泵技术, 2013, 3: 40-43.
- [5] 黄志坚,高立新,廖一凡.机械设备振动故障监测与诊断[M].北京:化学工业出版社,2010.
- [6] Weibull W. Statistical theory of the strength of materials ingenious vetenskaps akademiens handlingar [J]. Royal Swedish institute for engineering research,1939,153:1-45.
- [7] Weibull W. A statistical distribution of wide applicability [J]. Jomalof Applied Mechanics, 1951, 18:293-297.
- [8] 但召江.无失效数据条件下滚动轴承的寿命与可靠性评价[D].杭州:中国计量学院,2013.
- [9] 徐现昭.滚动轴承可靠性评价方法研究[D].杭州:中国计量学院,2012.
- [10] 朱凯,王正林.精通 MATLAB 神经网络[M].北京:电子工业出版社,2010.