

基于时序分析的经验模式分解法及其应用*

贾民平 凌娟 许飞云 钟秉林
(东南大学机械工程系 南京 210096)

摘要: 经验模式分解方法可以将非线性非平稳信号分解为有限的固有模式函数, 在故障诊断中这个固有模式函数常常就是故障信号。但当两侧端点不为极值点时, 会造成三次样条拟合的极值包络线大大偏离实际值, 并且随着分解的不断进行向内“污染”。提出采用时间序列建模与预测方法, 对原信号两端点进行预测, 有效地消除了端点效应。指出经验模式分解具有分解的自适应性特点。最后, 给出了齿轮箱振动信号的应用实例。

关键词: 经验模式分解 时间序列分析 预测

中图分类号: TN911.7 TH165+3

0 前言

机电设备的工况监测与故障诊断是保证机器安全运行的重要手段, 加强对工况信号的分析对提高故障诊断的精度有着十分重要的意义。

机电设备的故障信号, 如齿轮箱、滚动轴承等一般为非平稳信号, 传统的基于平稳随机信号的分析方法没有能力解析。而时频分析方法虽然有助于非平稳信号的分析, 但由于难以定量等原因, 在诊断方面的应用还没有达到实用阶段^[1]。参考文献[2]提出的经验模式分解方法是一种新的非平稳信号分析方法, 其本质是对非平稳信号进行平稳化处理, 将信号中不同尺度的波动或趋势逐级分解开来。通过分解非平稳故障信号可以把原来不易获得的微弱故障信号分解出来, 从而实现故障早期诊断。但在分解过程中构成上下包络的三次样条函数在数据序列的两端会出现发散现象, 并且随着过程的不断进行逐渐向内“污染”整个数据序列而使所得结果严重失真。在进行希尔伯特变换时, 信号的两端也会出现严重的端点效应^[2,3]。

参考文献[3]采用单层、单个神经元和线性神经网络来对原始信号进行延拓, 寻找极值点求极值包络线。然后剔除延拓的那部分数据点, 同时也消除了“端点效应”。但该方法存在训练数据选取问题, 并有可能存在网络训练不收敛的情况, 且在分解过程中需要进行多次神经网络的训练, 计算时间过长; 另外, 初值选取对网络训练影响十分大, 也是造成网络不收敛的原因之一。

现采用时间序列分析建模与预测来解决上述端点效应, 并通过实例应用验证上述经验模式分解方

法的有效性。

1 经验模式分解法及存在问题

1.1 基本原理

为了得到对非平稳信号有意义的瞬时频率, 要求它必须满足: ①信号极值点的数量与过零点的数量必须相等或最多相差一个。②在任一时间点上, 信号的局部最大值与局部最小值定义的包络的均值必须是零。满足这两个条件的信号称为固有模式函数。

由于大多数信号不是固有模式函数, 可能包含不只一个振荡模式。因此, 必须把信号分解为若干个固有模式函数。参考文献[2]提出了把信号分解为固有模式函数的算法——经验模式分解算法。其实质即为根据经验确定数据中有效信号的基本振荡模式(固有模式函数), 并据此分解数据。分解模式函数的过程, 也被称为筛选过程, 思路如下。

找到信号中的所有局部极值点, 其中所有的局部最大值被一个三次样条连接成为上包络。同理, 局部最小值产生下包络, 上下包络应将所有的数据都包含在它们之间。上下包络线的均值定义为 m_1 , 而原始信号 $x(t)$ 与 m_1 的差值定义为函数 h_1 , 则

$$x(t) - m_1 = h_1 \quad (1)$$

理想情况下, h_1 应是一个固有模式分量。然而, 实际上, 对于非线性数据, 包络均值可能不同于真实的局部均值, 因此, 一些非对称波仍可能存在。筛选过程主要有两个作用: 一是去除叠加波, 二是使波形更加对称。为了达到这个效果, 该过程可以被重复多次。在第二次过滤处理中, 分量 h_1 被当作待处理数据, 于是

$$h_1 - m_{11} = h_{11} \quad (2)$$

可以把处理过程重复 k 次, 直到 h_{1k} 是一个固有

* 国家 863 高技术研究发展计划(2001AA423240)和国家自然科学基金(59875013)资助项目。20031025 收到初稿, 20040410 收到修改稿

模式分量，于是

$$h_{1(k-1)} - m_{1k} = h_{1k} \tag{3}$$

定义

$$h_{1k} = c_1 \tag{4}$$

则 c_1 就是从原始数据中处理得到的第一个固有模式分量。它包含原始信号中最短的周期分量，即频率最高的周期分量。这时式(1)改写为

$$x(t) - c_1 = r_1 \tag{5}$$

剩余部分 r_1 仍然包含较长周期的固有模式分量，因此，把 r_1 当作新数据重复以上步骤，得到各个固有模式分量

$$r_1 - c_2 = r_2, r_2 - c_3 = r_3, \dots, r_{n-1} - c_n = r_n \tag{6}$$

当分量 c_n 或剩余分量 r_n 比预定值小，或当 r_n 变成单调函数，不能再筛选出模式函数时停止上述分解过程。此时，剩余分量 r_n 称为余项。对于有趋势的数据，其余项就应该是该趋势。这样，就把信号 $x(t)$ 分解成 n 个模式函数及一余项 r_n 。

1.2 经验模式分解法存在的主要问题

经验模式方法中十分重要的一个步骤就是获取信号曲线的极大值点和极小值点的包络线，其精度直接影响分解的模式函数和余项的精度。但在求取极值点包络过程中，由于信号两侧端点一般不是极值点，导致产生误差，而且这种误差随着“筛选”过程的不断进行逐渐向内“污染”整个数据序列而使所得结果严重失真。这称之为端点效应。

2 基于时间序列分析改进经验模式分解法的端点效应

2.1 基于时间序列预报改进端点效应的理论

在计算极大、极小值的包络时，由于两端点不是极值造成了端点效应，因此，是否可以通过某种模型预测两端的极值点，这样就可以将“端点效应”释放到原数据的外端，从而消除经验模式分解法计算区间的端点效应。

基于前向神经网络如前所述存在训练数据选取、训练不收敛和初值选取的问题，难以得到应用。为此，提出采用时间序列建模并预测的方法，效果良好，且具有结果稳定，速度快的优点。具体方法如下。

ARMA 模型是时间序列分析中最基本的模型，它可以解释动态数据本身的结构与规律，并可以进行预测。当 ARMA 模型满足可逆性时，可以转换为 AR 模型，采用最小二乘、Burg 和 Marple 等算法进行建模^[5]。假设 x_t 为连续信号 $x(t)$ 的离散值，则对于

一个时间序列 $\{x_t, t=1, 2, \dots, N\}$ ，其 AR(n)模型为

$$x_t - \varphi_1 x_{t-1} - \varphi_2 x_{t-2} - \dots - \varphi_n x_{t-n} = a_t \tag{7}$$

式中 φ_j ——自回归参数

a_t ——白噪声

n ——模型的阶数

其在 t 时刻向前第一步最佳预测值为

$$x_{t+1} = \hat{x}_t(1) = \varphi_1 x_t + \varphi_2 x_{t-1} + \dots + \varphi_n x_{t+1-n} \tag{8}$$

建模的过程即是确定阶数 n 和自回归参数 φ_j 。

文中采用 Marple 算法，模型阶数采用固定阶数为 30，其理由是采用 AIC、BIC、最小残差和 F 检验等方法固然能够提高建模精度，但建模时间延长，而且对于预测来说没有大的影响。

预测两端极值点的步骤如下。

(1) 采用 Marple 算法建立 AR 模型。

(2) 根据式(8)预测边界外的第一个延拓值 $x'_1 = x_{t+1}$ ，并以 x'_1 为数据序列 $\{x_t\}$ 右侧的新边界点。

(3) 再次利用式(8)计算出边界外的第二个延拓值 $x'_2 = x_{t+2}$ ，以此类推，根据所需延拓的数据个数 N 可得到全部延拓序列 $x'_j (j=1, 2, \dots, N)$ 。 N 要求大于所分析数据的周期。

对于原数据序列另外一端的处理与此相同，在此不加赘述。

2.2 仿真例子

仿真为一由两个线性调频信号叠加而成的信号

$$x(t) = \sin[2\pi(10 + 2t)t] + \sin[2\pi(20 + 2t)t] \tag{9}$$

对式(9)中的信号 1 024 点进行分析，图 1 为采用时间序列建模预测后改进端点效应的效果图。图 2 为用经验模式分解法进行分解的结果。

对比改进前后的包络线情况，用时间序列分析来延拓数据所求包络线的形势和量值都比较满意。

比较改进前和改进后的分解结果，可以发现以下两点。

(1) 式(9)中的信号包含两个振动模式，每个振动模式都是比较标准的谐波信号。图 2b 中的两个模式函数幅值均匀，看起来就是两个谐波波形，与实际情况一致。而图 2a 中的两个模式函数边界点处幅值是渐弱的，这是因为没有采用合适的消除“端点效应”的措施。

(2) 式(9)信号不存在趋势变化，其余项应该是一个常量，但由于一些实际的处理，如：包络均值并不是严格的瞬时均值，包络曲线的精度等等使得实际情况与理想状况有些出入。因此，考查余项实际情况的误差成为判断包络精度的一个指标。为此，提出残差平方和这一指标

$$s = \sum (r_{ni} - w_i)^2 \tag{10}$$

式中 r_{ni} ——余项的第 i 个分量
 w_i ——该信号趋势 w 的第 i 个分量

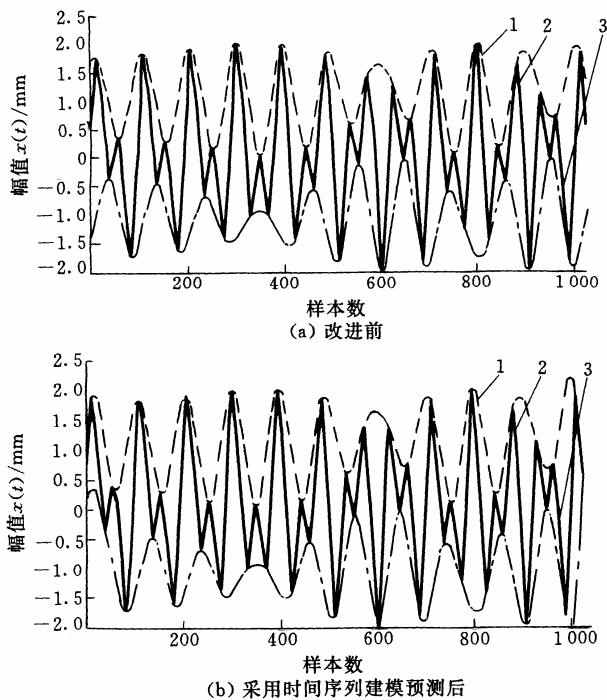


图1 式(9)信号及其上下包络线图
 1. 上包络线 2. 原始数据 3. 下包络线

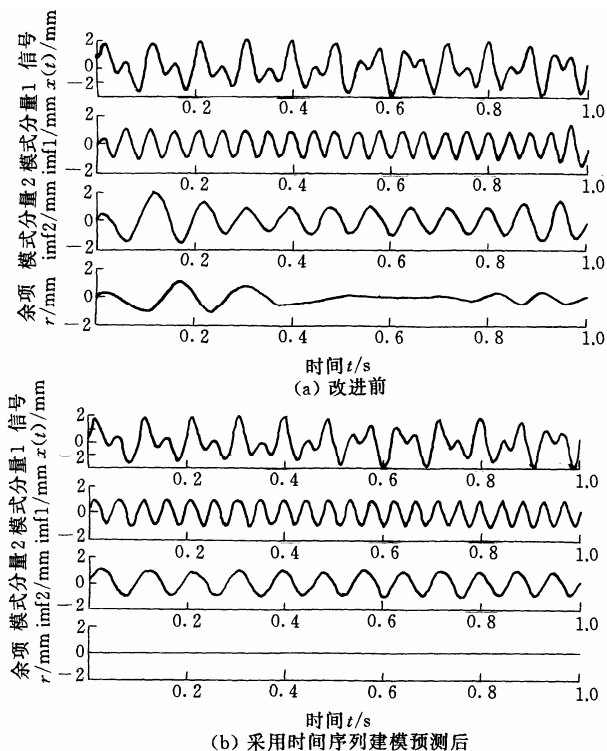


图2 式(9)信号的经验模式分解图

图 2a 中的余项波动比较大, 残差平方和高达 682; 而图 2b 中的余项波动很小, 残差平方只有 0.334 9, 两者相差 106 倍。

3 经验模式分解在齿轮箱工况分析中的应用

某厂 R05 齿轮箱共有 5 根轴组成, 内部有 4 对齿轮啮合, 在它的输入端是由一变频电动机驱动, 其转速为 1 017 r/min 左右, 经过 4 级减速后, 输出的转速约为 5 r/min。在其上装有 4 个振动加速度传感器。机组测试工况如下: 测试数据为 2002'11'12'14'20'48(数据文件名代表了当前数据发生的时间); 电动机转速为 1 017 r/min; 采样频率为 704 Hz; 采样长度为 8 192 点。测点 3 位于第 5 根轴的齿轮箱壳体垂直方向, 它测得的主要是 VI 轴与 V 轴一对啮合齿轮的振动信号, 这对齿轮的啮合频率为 3.69 Hz。由于种种原因 VI 轴与 V 轴及其齿轮对经常发生故障, 严重时断齿、断轴。下面对测点 3 的数据进行分析比较。

图 3 为对测点 3 的数据进行经验模式分解的结果。其中, 图 3a 为分解后的各模式函数的时域波形图。图中从上至下依次为信号及其前 6 个固有模式分量; 图 3b 为对应的各信号进行幅值谱分析的结果。

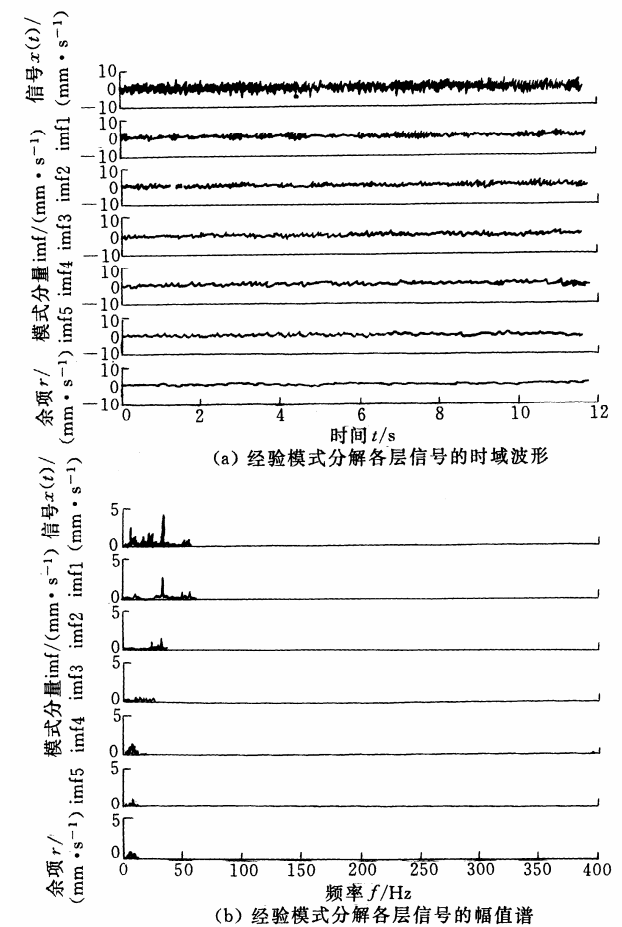


图3 某齿轮箱测点3信号的经验模式分解结果

从图3中可以看出, 经验模式分解方法把原始信号按照振动模式进行了分解, 分解后的模式函数: imf1的主要频率是33.8 Hz, 反映了2倍电动机转频的振动模式; imf4的主要频率成分是11 Hz和7.4 Hz, 为啮合频率3.69 Hz的3倍频和2倍频; imf5的主要频率成分是7.4 Hz, 反映的是啮合频率的2倍频的振动模式。还可以看到, 在imf4和imf5中主要频率都受到一些其他频率成分的频率调制, 反映了齿轮啮合过程中因齿轮轮齿破损或负载波动引起的齿轮啮合冲击。将图3与小波分析比较可以发现以下问题。

(1) 小波分析可以看成为滤波器, 为不同频率段滤波器的叠加。而经验模式分解方法是按不同的振动模式对信号进行分解, 每一次分解得到一个振动模式分量, 有着明确的物理意义。

(2) 一旦选定小波函数, 小波分析在分解过程中不能根据实际情况进行相应的更改, 从这个角度来说, 小波变换不是自适应的。而经验模式分解方法则完全是自适应的, 不需要事先确定分解基。信号不同, 极值包络线的均值就不同, “筛选”过程就不同。

4 结论

针对经验模式分解过程中的“端点效应”现象, 提出了基于时间序列建模预测的改进方法, 给出了该方法在某齿轮箱振动信号上的应用, 并比较了经验模式方法与小波分析的不同点。

(1) 针对经验模式分解过程中出现的“端点效应”现象, 提出基于时间序列方法的改进方法。比较发现该方法效果良好, 具有结果稳定、速度快的优点, 可以大大的改善分解过程中的端点效应问题。

(2) 提出了采用残差平方和作为检验经验模式分解方法分解精度的指标。

(3) 将经验模式分解方法应用于实际的齿轮箱数据, 表明该方法可以从复杂信号中提取出系统固有的振动模式分量, 从而为更好的监测机械设备的运行状况和故障诊断提供了有力的手段。

(4) 与小波分析相比较, 经验模式分解方法是自适应的, 即分解过程能够根据信号不同自动选择极值包络线进行筛选, 给出该信号的各个模式分量, 且具有明确的物理意义, 而小波分析则不具有这个自适应特性。

参 考 文 献

- 1 钟秉林, 黄仁. 机械故障诊断学. 第2版. 北京: 机械工业出版社, 2002
- 2 Norden E H, Zheng S, Steven R L, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lond. A, 1998, 454: 903~995
- 3 邓拥军, 王伟, 钱成春, 等. EMD方法及Hilbert变换中边界问题的处理. 科学通报, 2001, 46(3): 257~263
- 4 盖强, 马孝江. 几种局域波分解方法的比较研究. 系统工程与电子技术, 2002, 24(2): 57~59
- 5 Robert A Y, Monnie M G. Introduction to Time Series Analysis and Forecasting. New York: Academic Press, 2000

EMPIRICAL MODE DECOMPOSITION BASED ON TIME SERIES ANALYSIS AND ITS APPLICATION

Jia Minping Ling Juan Xu Feiyun Zhong Binglin
(Department of Mechanical Engineering, Southeast
University, Nanjing 210096)

Abstract: The nonlinear and non-stationary data set can be decomposed into a finite and often small number of ‘intrinsic mode functions’ by empirical mode decomposition(EMD). The intrinsic mode functions usually are the fault signal in fault diagnosis problem. The most serious problem of EMD method is the end effects due to the spline fitting at the data ends, i.e. the envelope curve fitted may have wide swings at the data ends if the ends are not the extremum. The decomposition quality would be polluted further along with the decomposition. An improved empirical mode decomposition based on the time series analysis is derived. The envelope curve can be well fitted with the predicted extremum at the data ends from the time series model. It’s useful to eliminate the end effects. Compared with the wavelet analysis, EMD method is adaptive. An example from the gearbox signal is given to demonstrate the power of the proposed method.

Key words: Empirical mode decomposition
Time series analysis Forecast

作者简介: 贾民平, 男, 1960年出生, 教授, 博士。主要研究方向为机械故障诊断、测试技术。

E-mail: mpjia@seu.edu.cn