关键工况参数趋势预测详细设计

> 陈鸿煜
>
> Update date: 2023/5/25

趋势预测，即根据历史数据变化的趋势，预测未来一段时间的数据。要获取数据的趋势，本质就是提取数据的特征，即需要进行信号处理。本科阶段学习的“信号分析”讲授了一些提取信号特征的方法，比如拉普拉斯变换获取频域特征；或者直接分析时域特征（如有效值、极值）。目前也有基于深度学习等智能算法的趋势预测方法。

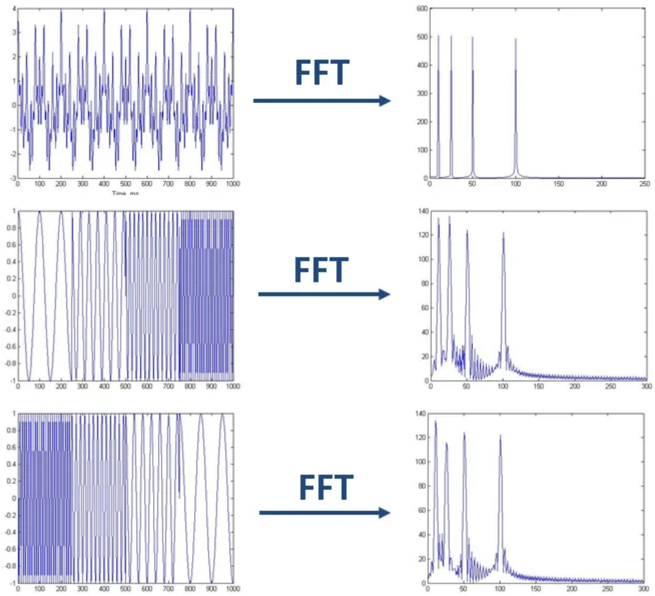
已实现的方法是通过传统信号分析方法对数据的时频域特征进行提取，进行趋势预测的。

一、基于经验模态分解的趋势提取

EMD（Empirical Mode Decomposition）是时频域的信号处理方法，小波变换也是基于时频域的处理方法。将两者进行对比。

1. 小波变换

单纯的频域分析在处理非平稳信号时存在弊端，如下图所示：



三个时域特征有明显差异的信号经过快速傅里叶变换以后，频域特征基本相似。所以对于这样的非平稳信号，只知道包含哪些频率成分是不够的，我们还想知道**各个成分出现的时间**。知道**信号频率随时间变化的情况**，各个时刻的瞬时频率及其幅值——这也就是**时频域分析。**

一个简单的时频域分析方法，就是给信号加上时间窗口，对每个时间窗口进行傅里叶变换。这种方法叫做短时傅里叶变换（Short-time Fourier Transform, STFT）。但是使用STFT存在一个问题，我们不知道该用多宽的窗函数。如果框太宽，时域分辨率低；如果框太窄，窗内的信号太短，频域分辨率低。

那么可以使用一个可以滑动的，大小可变的窗口函数，进行傅里叶变换。小波变换就是这种思路，只不过是通过基函数来实现的。

小波变换：

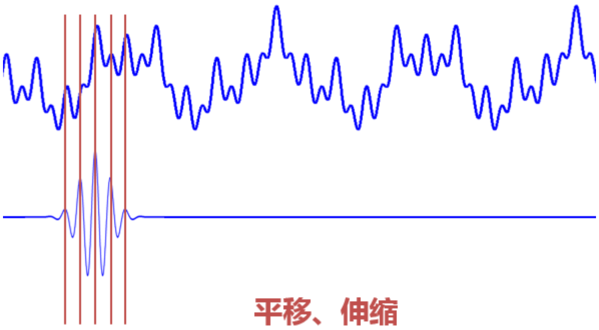
$$WT(a,\tau)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}f(t)\Psi(\frac{t-\tau}{a})dt$$

傅里叶变换：

$$F(\omega)=\int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt$$

傅里叶变换的基函数是一个时域无限长的三角函数基，小波变换的基函数是一个有限长的衰减函数。（因为只在一段区域有函数，其他地方都是0，所以可以称为有限长）由于基函数只在有限区域不为0，所以叫做“小波”。

从公式可以看出，不同于傅里叶变换，变量只有频率ω，小波变换有两个变量：尺度a（scale）和平移量 τ（translation）。**尺度**a控制小波函数的**伸缩**，**平移量** τ控制小波函数的**平移**。**尺度**就对应于**频率**（反比），**平移量** τ就对应于**时间**。这就像一个可变大小的时间窗口，当伸缩、平移到基函数和时域信号重合情况时，=相乘得到一个大的值。这**不仅可以知道信号有这样频率的成分，而且知道它在时域上存在的具体位置。**如下图所示：



[形象易懂讲解算法I——小波变换 - 知乎]

[浅谈傅里叶变换、小波变换、HHT变换_hht变换缺点]

2. EMD

回顾小波变换，小波变换的基函数（$\Psi(\frac{t-\tau}{a})$）需要根据想要得到的性质来推导得出。选取的基函数即使能满足全局最优，也不能保证局部最优。

EMD克服了以上问题。EMD最显著的特点，就是其克服了基函数无自适应性的问题。**EMD的分解基来自原信号本身，因此它具有自适应性，而不象傅里叶变换或者小波变换那样把信号分解为固定的函数。**

固有模态函数（Intrinsic Mode Functions, IMF）就是原始信号被EMD分解之后得到的各层信号分量。EMD的提出人黄锷认为，任何信号都可以拆分成若干个固有模态函数之和。而固有模态函数有两个约束条件：

1）在整个数据段内，极值点的个数和过零点的个数必须相等或相差最多不能超过一个。

2）在任意时刻，由局部极大值点形成的上包络线和由局部极小值点形成的下包络线的平均值为零，即上、下包络线相对于时间轴局部对称。

为什么IMF一定要满足这两个条件呢？经黄锷等人的研究，满足这两个条件的信号都是单组分的，相当于序列的每一个点只有一个瞬时频率，无其他频率组分叠加。这就为后续的**希尔伯特变换**铺平了道路，也使得瞬时频率有了意义。

EMD的步骤如下：

1. **求极值点** 通过Find Peaks算法获取信号序列的全部`极大值`和`极小值`
2. **拟合包络曲线** 通过信号序列的`极大值`和`极小值`组，经过`三次样条插值法`获得两条光滑的波峰/波谷拟合曲线，即信号的`上包络线`与`下包络线`
3. **均值包络线** 将两条极值曲线平均获得`平均包络线`
4. **中间信号** 原始信号减均值包络线，得到`中间信号`

5. **判断本征模函数（IMF）** IMF需要符合两个条件：
   1）在整个数据段内，极值点的个数和过零点的个数必须相等或相差最多不能超过一个。
   2）在任意时刻，由局部极大值点形成的上包络线和由局部极小值点形成的下包络线的平均值为零，即上、下包络线相对于时间轴局部对称。

总而言之，EMD将非平稳信号分解为多个平稳信号。

这篇文章能让你明白经验模态分解（EMD）——基础理论篇 - 知乎

经验模态分解 (Empirical Mode Decomposition) - 简书

经验模态分解（EMD）介绍 - 知乎

EMD方法基本基本知识

3. 希尔伯特黄变换HHT

经过EMD分解出的IMF分量包含了信号的不同时间尺度的局部特征信号。

再经过Hilbert变换，最终得到信号瞬时频率和瞬时幅值。

1. HHT的结果反映的是信号的时频特征，即信号的频域特征随时间变化的规律。相对于傅里叶变化得到的是信号的频率组成，HHT还可以获取频率成分随时间的"变化"。比如我们要分析的信号代表的是一个性能迅速退化的发动机（假设信号表征是某些IMF分量的频率逐渐升高），使用HHT就可以对该现象进行很好的捕捉。
2. HHT可以对局部特征进行反映，这点主要得益于EMD的作用。EMD可以自适应地进行时频局部化分析，有效提取原信号的特征信息。
3. "分解"往往可以对应着"重构"，从HHT结果中选择出满足要求的特征分量并重组信号，有利于将关注的特征从复杂的混合信号中分离出来。

希尔伯特-黄变换

希尔伯特黄变换(Hilbert-Huang)原理、HHT求时频谱、边际谱，及MATLAB实现

4. 结合已实现的方法

可以结合正常工况下信号的时频特征，对比分析信号在即将出现停堆时的特征，进行趋势预测。

二、 基于自适应指数模型的关键工况参数趋势预测

1. 自适应指数模型

一个比较易懂的曲线拟合方法，关键在于时间窗口的选择。目前使用的方法是：

在上述算法中，为使训练模型的数据窗口能够充分反映趋势信息而非局部变化，要 求数据窗口所包含的信号具有充分大的变化。为此，窗口长度s的选取方式如下：设定 阈值 ξ_1 ，随着时间不断采集新的数据，当采集的最新数据y^{new} 与历史数据中的最初时刻数据y^1的差值大于阈值ξ_1时，确定数据窗口，再次索引历史数据中与最新数据y^{new}差值 大于阈值ξ_1的最短窗口长度S。其中初始窗口长度设定为采集到的第一个数据与报警阈值的差值的六分之一，滑动窗口长度设定为初始窗口长度的 2 倍。

关于报警阈值的选择并没有分析方法，或许可以通过第一张的EMD提取数据特征，看看是否和报警阈值的选择有关。

三、基于于自适应线性模型的关键工况参数趋势预测

第二章使用指数拟合，适用于高阶非线性数据，但是给出的数据接近线性，所以可以使用线性拟合模型。

方法于第三章相同，只不过使用了线性拟合。