

基于最大熵谱估计的工程非平稳信号的频谱分析研究与实践

吴道虎

(华中科技大学水电与数字化工程学院,湖北 武汉 430074)

摘要:大型机电设备的运行过程常常为非平稳运行过程,其动态的状态信号也常表现为非平稳性信号。非平稳过程往往包含着丰富的设备状态信息,也包含了丰富的故障信息,因此研究工程中的非平稳信号分析和特征提取方法对提高诊断的准确率具有十分重要的意义。采用基于最大熵谱估计(MESE)的方法对机电设备故障诊断中的实时工程非平稳信号进行频谱估计,以期进一步提高其分析精度和准确性,并与传统的谱分析方法进行了对比。

关键词:非平稳信号; 最大熵谱估计; 信号处理; 频谱分析

中图分类号: TP274; TN911

文献标识码: A

文章编号: 1003-4897(2006)13-0038-04

0 引言

机电设备,尤其是旋转设备,如水轮机、发电机等,在运行过程中,都要经历启、停机、增、减负荷等非平稳过程,另外,随着设备复杂程度的增加,各部件之间的关联性加强,某个零部件的异常(故障)在系统中会不断地传播,这也是非平稳过程。总之,机组在状态发展、变化时处于非平稳状态,其中也包括正常向异常的变化、故障状态趋向严重的发展,可见机电设备的非平稳过程更具有普通性和一般性。机组状态变化反映在机组的动态信号之中,非平稳过程往往包含的信息更全,包含了丰富的故障信息,因此研究工程中的非平稳信号分析和特征提取方法是提高诊断的准确率的重要手段之一。

在机电设备状态监测和故障诊断过程中,尤其是基于噪声分析和振动检测法的旋转机电设备故障诊断中,常常需要依据工程检测信号的频谱分布特征识别故障的特性。基于FFT技术的经典频谱估计由于是对信号真实功率谱的渐近无偏估计,而非一致估计。当分析样本长度增加时,频谱估计的方差、偏差和分辨率并不能同时得到改善;另外,由于实时工程检测信号的非平稳性,必然会存在非同步采样误差,加窗截断处理往往会导致频谱能量泄漏和频谱估计偏差,最终影响机电设备故障诊断的准确性。本文采用基于MESE的现代谱估计方法对机电设备故障诊断中的实时工程非平稳信号进行频谱估计,以期进一步提高其分析精度和准确性。

1 最大熵谱估计方法

最大熵谱估计^[1,2]的思路是把所观测的有限数

据以外的数据不作任何确定性的假设,而仅仅假设是随机的(经典谱估计假设以外数据的自相关序列值为零,而在实际中的信号并非如此),在信息熵最大的前提下,把自相关函数的两边外推至无穷,构造出自相关函数的相邻值,即预测出实测信号区间以外的数据,以得到一个很长的时间区域,然后再进行频域分析,求得连续的功率谱估计。由于最大熵谱的递推特性,熵谱估计所取得的信号时间序列可以在较短的样本时间序列下较为准确反应信号的时频信息。采用最大熵准则的目的就是使工程非平稳信号时间序列的随机性最大,而对预测的信号时序的约束最小,从而可以得到信号最小偏差的频谱估计,提高工程非平稳信号谱估计的精度和准确性。

1.1 最大熵谱估计算法

对一个研究对象建立数学模型是现代工程中常用的方法,一则可使所研究的对象有一个简洁的表达式,更为重要的是通过对模型的研究,可以得到更多的被研究对象的状态参数,可对所研究的对象有更深入的了解。对工程非平稳信号选定分析模型时,除了遵循“节俭(Parsimony)原理”^[1]外,主要考虑模型能否表示被表征信号的谱峰、谱谷等方面的能力。自回归模型AR、滑动平均模型MA以及自回归滑动平均模型ARMA模型是参数化谱估计最主要的参数模型。由于AR模型是一全极点的模型,系统稳定,计算过程相对简单,能够较好地描述频谱的谱峰(故障诊断中主要考核频谱的谱峰),具有良好的频谱分辨能力,因而在工程中获得了较为成功的应用^[2,3],应用AR模型的另一个原因是AR算法较为简单,可利用小样本数据分析,同时工程应用中

需要对信号频率的成分需要大体的了解,因此,应用 AR 模型有利于提高监测系统的实时性。

AR 模型现在的输出是由自身的若干次过去的值和激励的现在值的线性组合产生:

$$x(n) = - \sum_{r=1}^p a_r x(n-k) + u(n) \quad (1)$$

其中: $u(n)$ 为方差为 σ^2 的白噪声序列, p 为 AR 模型的阶次, $a_r (r = 1, 2, \dots, p)$ 为 AR 模型的参数。对式

(1) 两端进行 z 变换,得到系统传递函数:

$$H(z) = \frac{X(z)}{W(z)} = \frac{1}{A(z)} = \frac{1}{1 + \sum_{k=1}^p a_k z^{-k}} \quad (2)$$

那么采样序列的功率谱函数为:

$$P_x(\omega) = |H(e^{j\omega})|^2 \sigma^2 = \frac{\sigma^2}{\left| 1 + \sum_{k=1}^p a_k e^{-j\omega k} \right|^2} \quad (3)$$

由此可知,工程非平稳信号的谱估计的问题就转化为模型参数的估计问题,也即工程非平稳信号谱估计过程就是根据实时检测的工程信号的有限采样值估计 AR 模型参数的过程。模型参数估计的准确性越高,模型的拟合程度就越好,频谱的分辨率以及谱保真度就越高。理论上 AR 模型参数可根据预测误差最小的原则来估计。常用的计算方法有自相关法、Burg 法、改进协方差法和递推极大似然估计法^[4~6]。在实际的工程应用中,由于改进协方差函数包含了预测系统的统计特性和动态特性,因此在工程应用中比 Burg 算法更有优势^[6]。改进协方差法的具体算法为:令 $b_p(n)$ 与 $e_p(n)$ 分别为线性预测 AR 模型的后向预测误差与前向预测误差:

$$b_p(n) = \sum_{r=0}^p a_{pr} x[n - (p - r)]$$

$$e_p(n) = \sum_{r=0}^p a_{pr} x(n - r)$$

其中: $a_{p0} = 1$, 设 P_b 和 P_e 分别为后向预测误差功率和前向预测误差功率: $P_b = [b_p]^H b_p$ 和 $P_e = [e_p]^H e_p$, 令 P_b 和 P_e 的加权平均为 P_{eb} , 则

$$P_{eb} = \frac{1}{2} (P_e + P_b) = \frac{1}{2} \left(\frac{1}{N-p} \sum_{n=p}^{N-1} |e_p(n)|^2 + \frac{1}{N-p} \sum_{n=p}^{N-1} |b_p(n)|^2 \right) \quad (4)$$

当阶次 r 由 1 至 p 时, $b_p(n)$ 与 $e_p(n)$ 有如下的递推关系:

$$e_r(n) = e_{r-1}(n) + k_r b_{r-1}(n-1) \quad (5)$$

$$b_r(n) = b_{r-1}(n-1) + k_r e_{r-1}(n) \quad (6)$$

且 $e_0(n) = b_0(n) = x(n)$, 改进的协方差方法不仅要求 P_{eb} 相对反射系数 k_r 最小,而且要求 P_{eb} 相对

$a_{ri} (i = 1, 2, \dots, r)$ 都最小。由此可得:

$$\begin{bmatrix} c_x(1,1) & c_x(2,1) & \dots & c_x(p,1) \\ c_x(2,1) & c_x(2,2) & \dots & c_x(p,2) \\ \dots & \dots & \dots & \dots \\ c_x(p,1) & c_x(p,2) & \dots & c_x(p,p) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_p \end{bmatrix} = - \begin{bmatrix} c_x(1,0) \\ c_x(2,0) \\ \dots \\ c_x(p,0) \end{bmatrix} \quad (7)$$

$$\sigma^2 = c_x(0,0) + \sum_{r=1}^p a_r c_x(0,r) \quad (8)$$

式中:

$$c_x(i,k) = \frac{1}{2(N-p)} \left[\sum_{n=p}^{N-1} x(n-i) \cdot x(n-k) + \sum_{n=0}^{N-1-p} x(n+k) \cdot x(n+i) \right]$$

这样根据式 (7) 就可以估计出 AR 模型的参数, 由此也就得到了基于 MESE 的谱估计。

1.2 AR 模型阶次的选择原理

在基于 AR 模型的谱估计时,模型阶次 p 的选择非常关键,它影响着谱估计的质量。一般而言,由于被测信号的极点未知,因此,模型的阶次不是先验的,需要在递推过程中确定。经验上用 AR 模型来拟合某一具体随机过程时,总有某一确定的最合适的阶次 p_r 。当模型阶次低于 p_r 时,对给定记录长度为 N 的时间序列的最大熵谱估计会过于平滑,就如同采用低阶曲线拟合高阶曲线,会产生平滑的结果一样,不能给出足够的分辨率,有些信号就会无法识别,结果可能仅出现被测信号中最易预测、变化最为缓慢的频率点的峰值;而阶次高于 p_r 时,则如同高阶曲线拟合低阶曲线,会产生急剧变化和振荡的结果,在一定程度上,所得的谱估计具有虚假的细节,背景噪声会被作为真实信号拟合出来,产生假峰,并使计算量加大,使频谱估计产生偏差,甚至在低噪声或无噪声时,过高阶次的 AR 模型的谱估计会产生谱线分裂现象,其计算量剧增。

目前,针对不同的应用目的,提出了多种判阶准则,但在工程中被广泛采用的有 FPE 准则(最终预测误差准则)、AIC 准则(信息论准则)和 CAT 准则(自回归传递函数准则)三种^[2,4,5,6]。试验结果表明^[8],当信号的信噪比较高时,上述三种方法确定的最佳阶次 p_r 基本一致,可以给出较佳的 p_r 。但是,当信号的信噪比较低时,这三种方法得出的结果有所偏差,且 p_r 一般偏低,其中以 FPE(r)方法较为准确。

一般情况下,根据上述三个准则计算的 p_r 值一般随采样序列长度 N 单调增加,但比 N 增加要缓慢的多,通常限定 $p_{\max} \leq N/2$ 或 $\sqrt{N}$,并有经验公式:

$$p_r = \begin{cases} (\frac{N}{3} - 1) \sim (\frac{N}{2} - 1) & 20 \leq N \leq 100 \\ (0.01 \sim 0.2)N & N > 100 \end{cases} \quad (9)$$

应当指出,如何选择最佳阶次,仍然是一个有待于进一步研究的问题,目前,针对现代谱的估计,尚无任何情况下都能给出最佳结果的判阶准则。

2 最大熵谱估计的工程实践应用

水轮机的启/停机、加/卸载过程是典型的非平稳过程,水轮机此过程的振动、运行噪声信号也是一典型的非平稳信号。本次实验数据取自水轮机尾水管检修门处的振动信号,采用精密的振动加速度传感器,并采用研华加固型壁挂式工业控制计算机 IPC-6606 和 PCI-1712/1714 高速数采卡采集此信号,以此获得水轮机运行状态特征信息,实施对水轮机的状态检测与故障诊断。

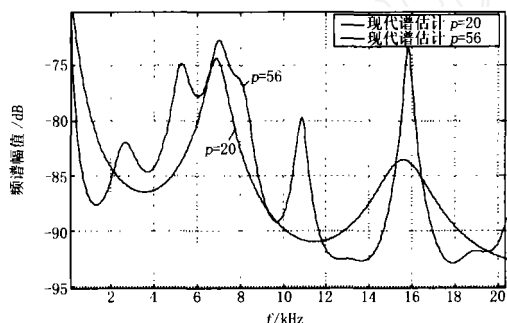


图 1 阶次 $p=20$ 和 $p=56$ 的谱图比较

Fig 1 Spectral comparison between $p=20$ and $p=56$

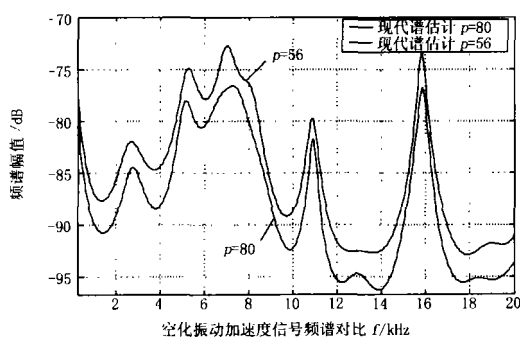


图 2 阶次 $p=80$ 和 $p=56$ 的谱图比较

Fig 2 Spectral comparison between $p=80$ and $p=56$

前已说明,基于最大熵谱分析法的工程非平稳信号谱估计的关键是根据实际监测得到的工程信号的时间序列信号的特征确定递推算法,并选取合适

的阶次。本次数据样本数据为在水轮机尾水检修门采集的振动加速度信号,采集频率 $F_s = 250$ k samples/s,采集的时间为 2.4 s(约两个水轮机转频周期,采样样本为 720 000)。采用 AR 时,取数据样本为 4 096,模型阶次上界与样本长度 N 之间的关系可以按 $(0.01 \sim 0.2)N$ 确定,确定下限为 $p_r = 41$,试验发现:阶次 $p=56 \sim 70$ 的频谱估计结果相近似,为减少计算量,取其阶次为 $p=56$ 。现特取两个极限值 $p=20$ 和 $p=80$ 与 $p=56$ 的谱图进行对比,分别见图 1 和图 2 所示。为便于在一个图形中对比,特将 $p=56$ 的谱图提高 3 dB。在图 1 中,当 $p=20$ 时,由于模型阶次较低,谱线过于平滑,在谱图中的 2.4 kHz、4.6 kHz、10.8 kHz 的谱峰无法表征,同时 16 kHz 的谱峰发生了畸变,整个谱图均发生了变化,基本上不能较为准确地表征谱分布。而当 $p=80$ 时,谱图基本上与 $p=56$ 相近似,但在 13 kHz 附近却出现了一个明显的伪峰,同时计算时间较 $p=56$ 的长。试验还发现,当 $p=100$ 时,谱图中 13 kHz 附近的伪峰能量加大,同时出现了其他的伪峰,使谱图变得模糊起来,某些局部甚至会出现急剧变化和振荡的现象,在一定程度上,所得的谱估计具有更多的虚假细节,反而降低了谱估计的准确性。

3 最大熵谱与经典谱估计的对比

为了充分比较最大熵谱估计与经典谱估计的差异,经典谱估计采用 Welch 周期图法,并选取 Hanning 窗,窗函数的宽度为 4 096 点,重叠率为 50%。最大熵谱估计的阶次取 $p=56$,样本长度取 4 096。图 3 为经典谱估计的频谱图与阶次为 $p=56$ 的最大熵谱估计的频谱图,同样为便于在一个图形中对比,特将 $p=56$ 的谱图提高 3 dB,结果发现两者之间有着惊人的相似之处,尤其是波峰的位置,但最大熵谱估计谱图清晰平滑,谱峰突出且分辨率较高。基于 FFT 的经典谱估计具有运算简单、物理意义明确,理论比较成熟和完善的优点,尤其在长数据记录情况下可靠性高,能较为稳定和正确地反映系统的频率分布情况。由此推断,可用大样本数据的经典谱图来验证 AR 模型的阶次^[9]的准确性。

经典谱估计与最大熵谱估计是信号分析领域的两种重要分析方法,二者相辅相成,对于大样本信号,经典谱估计系统稳定,通过合理选择谱估计方法、窗函数,并通过一定的频谱修正算法可获得较为满意的频谱估计结果。基于周期图法的分段数越多,谱估计的方差性能可得到改善,谱线较光滑,但

频谱估计的分辨率却会下降,偏差也会变大。目前还没有一种能使谱估计的方差、偏差以及分辨率同时得以改善的窗函数,这是其应用的一大弊端。最大熵谱估计均假设叠加在信号上的噪声为高斯白噪声,其关键在于如何根据观测到的有限长数据通过预测模型估计相邻的数据,并使预测模型的输出功率最小,不同的预测模型就形成了不同的算法。在基于 AR 预估模型时,AR 系数是由噪声方差的特征矢量决定的,对噪声极为敏感,在实践中,为方便分析,均假设噪声是具有高斯特性的随机序列,在较短的数据样本的情况下可以获得与超大样本的经典谱估计相似的分析结果。但有时叠加在信号上的噪声(如发电机的励磁杂波)往往是非高斯分布的有色噪声,这种情况下此方法在一定的程度上就受到了限制,必须采用更为先进的算法,如 MSE (Mean - Square Errors) 和 SVD (奇异值分解) 方法等。总之,信号分析的算法研究永远是信号处理领域中的核心。

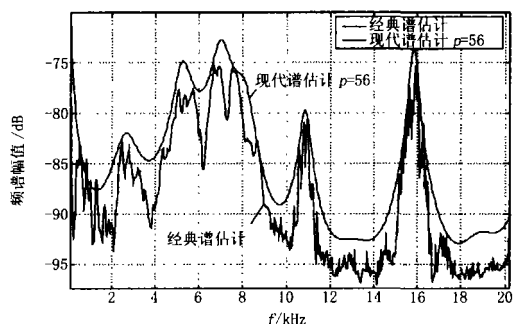


图 3 经典谱与最大熵谱 ($p=56$) 比较

Fig 3 Spectral comparison between Welch and MESE ($p=56$)

4 结束语

通过对实际的工程非平稳性信号的最大熵谱估计分析,在不改变信号的采样频率的情况下,采用较小的样本数据,通过 AR 预测模型对所观测的数据进行一致性的数据外推,可以获得比经典谱估计更高的频率分辨率和谱估计的准确性。但最大熵谱估

计的质量会受到预测模型阶次的极大影响,且随着所选的阶次 p 增大,模型预测的计算量越大,同时还伴有伪峰出现。在最大谱估计过程中,由于所对应的系统往往是未知的,模型的阶次确定比较困难,具有一定的试探性和经验,因此,模型阶次的确定成为最大熵谱估计应用的核心,同时,该算法对处理器的计算速度有一定的要求。

参考文献:

- [1] 张贤达,等. 现代信号处理 [M]. 北京:清华大学出版社, 2001.
ZHANG Xian-da, et al Modern Signal Processing [M]. Beijing: Beijing University Press, 2001.
- [2] Keith A, et al A Comparison of Autoregressive Spectral Estimation Algorithms and Order Determination Methods in Ultrasonic Tissue Characterization [J]. IEEE Trans on Ultrasonic, 1995, 42 (4): 709-716
- [3] 杨叔子,等. 时间序列分析的工程应用 [M]. 武汉:华中理工大学出版社, 1994
YANG Shu-zi, et al Engineering Application of ARMA Analysis [M]. Wuhan: Huazhong University of Science & Technology Press, 1994
- [4] Danthez J M. The Auto Coherent Spectral: a Useful Spectral Estimation for Vibration Analysis of Rotating Machinery [J]. Mechanical Systems and Signal Processing, 1998, (12): 753-768
- [5] Pineyro J, et al Effectiveness of New Spectral Tools in the Anomaly Detection of Rolling Element Bearings [J]. Journal of Alloys and Compounds, 2000, 31 (1): 276-279
- [6] Gabriele, et al Application of Autoregressive Methods to Multiage Spectral Analysis [J]. IEEE Trans on Ultrasonic in Med & Biol, 2000, 26 (4): 585-592

收稿日期: 2005-11-23; 修回日期: 2006-04-30

作者简介:

吴道虎 (1968 -), 男, 博士生, 主要从事过程控制、信号处理等方面的工作。E-mail: daohu_wu@sina.com

Spectral analysis research & application of engineering non-stationary signals based on maximum entropy spectrum estimation

WU Dao-hu

(Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract: The running process of large-scale electromechanical machine is usually non-stationary one, and its dynamic status signals are also non-stationary signals. The non-stationary running process of machine includes plenty of machine status and malfunction information. It is important to improve the fault-diagnosis accuracy to research the methods of engineering dynamic non-stationary signal.

(下转第 66 页 continued on page 66)

- tion of Electric Power Systems, 1999, 23 (8): 56-60.
- [4] 张雁靖. 分层分布开放式计算机监控系统的设计 [J]. 水利水电技术, 1997, 28 (5): 36-39.
ZHANG Yan-jing The Design of Hierarchy Distributed Open-Type Computer Supervisory and Control Systems [J]. Water Conservancy and Hydroelectric Source, 1997, 28 (5): 36-39.
- [5] 东方电子 DF8002集控系统技术说明书 [Z]. 东方电子有限公司.
Dongfang Electronics DF8002 Integrated Control Technical Introduction [Z]. Dongfang Electronics Information Industry Co., Ltd
- [6] 金午桥. 变电站自动化的发展策略 [J]. 电力系统自动化, 1999, 23 (22): 58-62.
JIN Wu-qiao Development Strategies of Substation Automation System [J]. Automation of Electric Power Systems, 1999, 23 (22): 58-62.
- [7] 何克忠, 李伟. 计算机控制系统 [M]. 北京: 清华大学出版社, 1998.
HE Ke-zhong, LI Wei Computer-based Control System [M]. Beijing: Tsinghua University Press, 1998.
- [8] 潘爱民. COM 原理与应用 [M]. 北京: 清华大学出版社, 2001.
PAN Ai-min COM Theory and Its Application [M]. Beijing: Tsinghua University Press, 2001.
- [9] Eddon G, Eddon H. Inside COM + Base Services [M]. 北京: 希望电子出版社, 2004.
Eddon G, Eddon H. Inside COM + Base Services [M]. Beijing: Publishing House of Hope Electronics, 2004.
- [10] Cadmen C M. COM /DCOM 编程指南 [M]. 北京: 清华大学出版社, 2004.
Cadmen C M. COM /DCOM Programming Guide [M]. Beijing: Tsinghua University Press, 2004.
- [11] 黄晓莉. 面向 21 世纪的国家电力数据通信网络 [J]. 电力系统自动化, 1999, 23 (23): 45-49.
HUANG Xiao-li 21th Century Oriented State Power Data Network [J]. Automation of Electric Power Systems, 1999, 23 (23): 45-49.
- [12] 郑健超. 电力技术发展趋势浅议 [J]. 电网技术, 1997, 21 (11): 4-10.
ZHENG Jian-chao The Discussion on the Power Source Development Trend [J]. Power Network Technology, 1997, 21 (11): 4-10.
- [13] 于尔铿, 刘广一, 周京阳, 等. 能量管理系统 (EMS) —— 第 14 讲能量管理系统 (EMS) 技术展望 [J]. 电力系统自动化, 1998, 22 (2): 68-71.
YU Er-keng, LIU Guang-yi, ZHOU Jing-yang, et al Energy Management Systems (EMS) [J]. Automation of Electric Power Systems, 1998, 22 (2): 68-71.
- [14] 韩祯祥, 薛禹胜, 张伯明. 1996 年 CIGRE 第 36、37、38 组简介 [J]. 电力系统自动化, 1997, 21 (1): 44-48.
HAN Zhen-xiang, XUE Yu-sheng, ZHANG Bo-ming 1996 CIGRE Session [J]. Automation of Electric Power Systems, 1997, 21 (1): 44-48.

收稿日期: 2005-11-15

作者简介:

徐 蓉 (1969 -), 女, 讲师, 主要从事电力系统自动化方面的教学及科研工作。E-mail: powercome@163.com

D istributed structure-based main-sub concentrate control center groups

XU Rong

(Dept of Electric Engineering, Shaanxi Institute of Technology, Hanzhong 723003, China)

Abstract: New independent monitoring systems are built without information exchange between the systems. When one monitoring system is destroyed, many substations will lose the monitoring which damages the operation of the whole power grid and badly affects its recovery. So, this paper provides a new method that builds the distributed structure of main-sub concentrate control center group, it achieves the change of data between all concentrate control center, its aim is standby with each other in the main-sub monitoring systems, which the management of operating is concentrated and the maintenance is integrated.

Key words: concentrate control center groups; distributed structure; main control; sub-control; middleware

(上接第 41 页 continued from page 41)

analysis and feature extraction. In the paper, the spectral analysis technique is discussed on real-time engineering non-stationary signal of the running process of machine based on maximum entropy spectrum estimation (MESE) method. MESE may improve its analysis precision and accuracy at machine fault-diagnosis. A comparison between MESE and traditional spectral analysis is presented.

Key words: non-stationary signals; MESE; signals processing; spectral analysis