时间序列预测模型在压缩机零件测试中的运用

王 彬
WANG Bin

南京大学 计算机科学与技术系,南京 210093
Department of Computer Science and Technology,Nanjing University,Nanjing 210093,China
E-mail:wbighead@126.com

WANG Bin.Application in compressor parts for predication model of time series analysis.Computer Engineering and Applications,2008,44(17):228-230.

Abstract: Parts in compressors sometimes is deformed with time going.Therefore it is very important to monitor the deformation and predicate the trend to keep the performance of the compressors.In this paper,the base theories for time series predication model are introduced here.Parameter identification methods are also used.Based on the deformation data of muffle plate in compressor,a predication model for this part is established to monitor and forecast the deformation.Practical applications show that the model in this paper is reliable and useful.

Key words: time series;ARMA model;predication;parameter identification;system identification

摘 要:压缩机零件在使用时会随时间的变化发生变形,对于压缩机的性能会有很大的影响,因此在零件进行材料选择和几何尺寸设计时,对相关零件变形的检测和预报是至关重要的。介绍了时间序列分析的基本理论及其预测建模的方法;采用参数识别方法,结合压缩机消音板发生变形的动态数据,建立该压缩机零件变形的时序预测模型,对相关零件变形进行了监测和预报,具有很强的实际运用价值。

关键词:时间序列;ARMA 模型;预测;参数识别;系统辨识
DOI:10.3778/j.issn.1002-8331.2008.17.068 **文章编号:**1002-8331(2008)17-0228-03 **文献标识码:**A **中图分类号:**TP311.11

1 引言

工程实践中,人们常遇到一些相互依赖而又相互制约的变量,它们之间的关系一般可分为两类:一类是确定性的关系,即变量之间的关系可用确定性函数关系来表示,这是数学分析的内容;另一类是随机性的关系,即自变量都是随机性的,变量之间的这种关系称为相关关系。基于实测数据的统计相关分析法,是对时间(或空间)轴上的随机变量建立数学统计模型而进行的。目前的统计模型有两种形式:回归模型(回归分析法)以及在此基础上发展起来的时间序列模型(时间序列分析法,或简称为时序分析或时序方法)。时间序列在工程应用大致可分为六个方面:(1)系统辨识;(2)系统分析;(3)谱分析;(4)模式识别(特别用于工况监视、故障诊断、医疗诊断等方面的);(5)模态参数估计;(6)预测与控制。在文中主要应用了时间序列的第一个方面,即系统辨识。

时间序列是随时间改变而随机地变化的序列。时间序列分析的目的是找出它的变化规律,即线性模型,主要有三种:AR 模型(自回归模型)、MA 模型(滑动平均模型)和 ARMA 模型(自回归滑动平均模型或混合模型)。时间序列在工程中常用于做预报,如气象预报、地震预报、水文预报、电力负荷预报等。论文在时间序列理论的基础上,总结了时间序列预测/预报方法;并结合工程实际运用,成功的运用到了某大型建筑的变形预测上。实际运用表明,时间序列对于平稳随机过程的预报可靠。时间序列分析的基本理论 随着时间序列理论的发展,逐渐形成了一整套模拟、估计、建模、预测和控制的理论和方法,在动态数据的处理分析、复杂信息的加工提取、预测未来和在线控制等方面显示出传统的数理统计静态处理手段无可比拟的优越性。随着电脑的普及和计算技术的发展,时间序列方法越来越渗透到科研、生产、农业、经济及社会科学等各个领域,呈现出越来越旺盛的生命力。其中,ARMA 模型是时间序列模型中,最常用、最成熟的一种模型。

2 时间序列的建模

时间序列建模是时间序列分析的关键步骤,因为它涉及到结构损伤诊断的成功与否。建模的内容包括数据的采集与标准化处理、模型形式的选取、模型参数 φ_i,ϕ_i 的估计、模型定阶等问题。其中最关键的是模型参数的估计,其次是模型的定阶。

2.1 ARMA 模型

对于一个振动系统,某测点的输出(响应)为$\{x_t\}$,可有 n 阶自回归—m 阶滑动平均混合时序模型来描述这个振动系统,记为 ARMA(n,m)。即:

作者简介:王彬(1978-),南京大学计算机专业。
收稿日期:2007-09-17 **修回日期:**2007-12-03

$$x_t-\varphi_1 x_{t-1}-\cdots-\varphi_n x_{t-n}=a_t-\theta_1 a_{t-1}-\cdots-\theta_j a_{t-j}-\cdots-\theta_m a_{t-m}$$

式中，φ_i 为自回归系数，($i=1,2,\cdots,n$)；θ_j 为滑动平均系数($j=1,2,\cdots,m$)；a_t 为白噪声序列，a_t 服从 $NID(0,\sigma_0^2)$，σ_0^2 为白噪声方差。用线性后移算子 B，上式可表示为：

$$\varphi(B)x_t=\theta(B)a_t$$

$$\varphi(B)=1-\varphi_1 B-\varphi_2 B^2-\cdots-\varphi_n B^n$$

$$\theta(B)=1-\theta_1 B-\theta_2 B^2-\cdots-\theta_m B^m$$

特殊地，若 $n=0$，模型称为纯滑动平远近模型，记为 MA(m)；若 $m=0$，模型称为纯自回归模型，记为 AR(n)；若 $n=m=0$，模型(1)退化为 $x_t=a_t$，即$\{x_t\}$为白噪声序列。

2.2 ARMA 模型的定阶与参数估计

对时间序列的$\{x_t\}$($t=1,2,\cdots N$)，首先要进行相关性分析。相关性分析的任务是计算序列$\{x_t\}$的样本自相关函数和样本偏自相关函数，并由他们的截尾性来进行模型类别的判断。可根据表 1 进行模型结构的初选。

表 1 $ARMA(n,m)$模型的序列特征表

	$AR(n)$	$MA(m)$	$ARMA(n,m)$
自相关函数	拖尾	截尾$k=m$ 处	拖尾
偏自相关函数	截尾 $k=n$ 处	拖尾	拖尾

(1)自相关函数 m 步截尾的序列，选用模型。

(2)偏相关函数 n 步截尾的序列，选用模型。

(3)自、偏相关函数全部指数衰减的序列，选用 ARIMA 模型(阶数需进一步判断)。

(4)自相关函数成周期规律的序列，可选用季节性乘积模型。

(5)自相关函数规律复杂的序列，可能需要作非线性模型拟合。

2.3 ARMA 模型的参数估计

ARMA(n,m)中的 n,m 参数的确定：n,m 并不能直接确定，而是需要先假定一组值，一般是从(1,1)开始，建立模型，然后逐步升高 n,m 的值，求出一系列模型，并根据 loss function，AIC 等准则，找出一个最优模型。这里介绍一种粗略的判断方法：n 的选择是看平稳序列的偏自相关图。选择有三：落入随机区间外的偏自相关个数；有效偏自相关的时滞；自相关函数的震荡周期。m 选择要看平稳序列的自相关图。选择有二：显著不为 0 的自相关数目；自相关函数从 $k=m_0$ 开始迅速衰减，则 $m=m_0$。在实际应用中，可以将以上方法结合起来，即先根据自相关图和偏自相关图粗略判断 n,m 的值，然后在该组值附近求出一系列模型，从中选出一个最优的。

时序模型的建立需要离散的时间序列$\{x_t\}$。对于时间序列$\{x_t\}$，当其取值过大或过小时，为保证计算精度、减少舍入误差、避免溢出，可对$\{x_t\}$进行标准化处理。记所得的时序为$\{x_t\}$，当$\{x_t\}$满足均值为$\hat{\mu}_x$、方差为$\hat{\sigma}_x^2$的正态分布时，对$\{x_t\}$中各数据进行如下标准化处理：

$$y_t=\frac{x_t-\hat{\mu}_x}{\hat{\sigma}_x}$$

注意：为了方便起见，本文将经过预处理的时间序列仍然记为$\{x_t\}$。

同时也要注意，模型形式的选择，是对经过预处理的新序列$\{x_t\}$来进行。虽然自、偏相关函数截尾的序列虽少见，但也不可忽视，充分利用此类特性，可大大减少建模工作量。对于自、偏相关函数指数衰减的序列，用高阶自回归模型拟合比 $ARMA(m,n)$模型拟合有其方便优越之处；模型形式的选择，要充分利用数据点图的直观规律及数据物理背景中蕴含的信息。通常在模型阶数给定的情形下，进行模型参数的估计。模型参数需要用一个样本做估计。必须指出这里白噪声方差也作为一个模型参数。在工程计算中，只要用样本自协方差函数 r_k 或样本自相关函数 ρ_k 中的一部分数值。

2.4 平稳时间序列的预报

在经过模型类型识别，参数估计和模型检验，得一个较为满意的时间序列模型之后，剩下的问题就是如何利用这个模型进行预报了。也就是说，根据包括现在的和以往的所有观察资料，对未来时刻的取值作出估计。

通常有两种预报方式：Box 的平稳线性最小方差预报方法和新息预报法。其中平稳线性最小方差预报方法设$\{w_k\}$是平稳、正态、且均值为零的序列，$\hat{w}_k(l)$为由 k 时刻及以前数据对 w_{k+l} 作的 l 步平稳线性最小方差预报，可以表示为：

$$\hat{w}_k(l)=c_0^* w_k+c_1^* w_{k-1}+\cdots$$

则序列$\{c_j^*\}$必须满足预报误差 $e_k(l)=w_{k+l}-\hat{w}_k(l)$的均方值达到最小值。

2.5 模型残差检验

对序列通过了平稳性检验，并建立了相应的 ARMA 模型之后，为考核所建模性的优劣，一般还需对 ARMA 模型残量 a_1，a_2，…，a_n 检验是不是白噪声。也就是说，如果经检验确是白噪声序列，则可认为模型是合理的，否则，就应当进一步改进模型。

3 基于时间序列分析的压缩机壳体零件预测

大型压缩机壳体在工作时的变形也可以看作一组随机时间序列值，本文拟采用时间序列的方法来对壳体的变形进行分析；并建立对变形的预测模型。

3.1 时间序列预测建模的过程

如图 1 所示是利用时间序列分析方法，对壳体的变形进行预测建模方法的流程图。首先通过传感器采集数据，得到壳体相对变形的位置，然后通过预处理，主要是去均值和去趋势，得到平稳时间序列；接着对模型的类型进行确定，并确定模型参数的数值；最后通过残差比较确定最终的模型是否合格，否则从新进行选择适用模型。

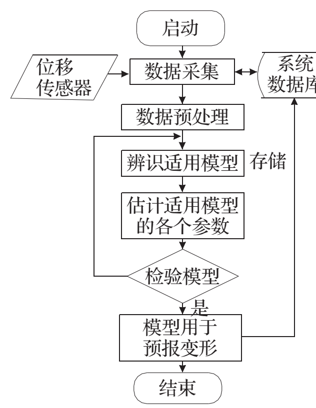

图 1 时间序列预测建模方法流程

3.2 案例描述

图 2 所示是某种大型商用压缩机的照片，该压缩机为 15HP 涡旋压缩机，高低压腔有消音板隔开，分成下侧的低压腔和上侧的高压腔，该压缩机直径为 226 mm，高度为 500 mm，消音板厚度为 5 mm，由于在使用过程中的高低压腔的压力变化，在壳体和消音板上有持续的交变应力作用，发生形变，特别是消音板的尺寸及形状变化对压缩机的性能，能效以及压缩机的噪音产生较大影响，因此结合时间序列分析，对该零件的尺寸，形状设计进行优化。节约成本，降低不良。

图 3 显示是对该压缩机进行测点分布的示意图，以变形和受力较为集中的点作为研究对象，重点监测消音板的变化值。

图 2　某种大型商用压缩机

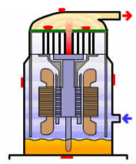

图 3　测点分布

以压缩机测试时间为单位对该点过去大约三个月中的数值进行统计，测量值如表 2 所示。

表 2　监测数据一览

序号	监测值（×10−2mm）	序号	监测值（×10−2mm）	序号	监测值（×10−2mm）
1	0.47	13	2.40	25	4.07
2	0.47	14	2.73	26	4.17
3	0.47	15	3.07	27	4.27
4	0.80	16	3.30	28	4.37
5	1.13	17	3.53	29	4.47
6	1.13	18	3.67	30	4.50
7	1.13	19	3.80	31	4.53
8	1.27	20	4.07	32	4.57
9	1.40	21	4.33	33	4.60
10	1.73	22	4.17	34	4.67
11	2.07	23	4.00	35	4.73
12	2.23	24	4.03		

3.3 时间序列预测建模

在论文的模型中，把表 2 中前 30 组数据作为建模数据；把后 5 组数据作为检验数据；来对模型进行分析。观测测量数据，由于有变形的累积作用。数据有明显的增长趋势。首先对模型去趋势化处理，通常都是去除线性量：

$$d'(t)=0.14t+0.537$$

则去除趋势以后的序列模型可以表示为：

$$x'(t)=x(t)-d'(t)$$

其中 $d'(t)$ 表示趋势项；$x(t)$ 和 $x'(t)$ 分布表示原来序列和去趋势后的序列。

通过对序列样本的自相关性和偏自相关性性分析；得到上面函数图（见图 4 和图 5）。从上面可以看出自相关函数呈明显的拖尾变化，没有明显的截尾变化。通过上述分析，初步判断模型为 AR 模型，基于 $AR(n)$ 模型的建模与识别，对于时间序列 $\{x_t\}$，$(t=1,2,\cdots,N)$，其中 $AR(n)$ 为：

$$x_t=\varphi_1 x_{t-1}+\cdots+\varphi_n x_{t-n}+a_t,\ a_t\text{服从 }NID(0,\sigma_0^2)$$

参数估计就是按照一定的方法估计出 $\varphi_1,\cdots,\varphi_n,\sigma_0^2$ 这 $n+1$ 个参数。由于有：

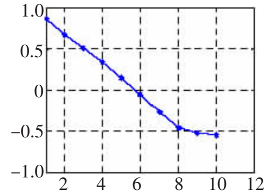

图 4　自相关函数图

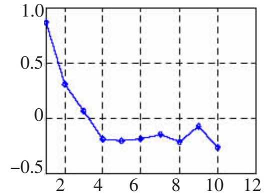

图 5　偏自相关函数图

$$a_t=x_t-\varphi_1 x_{t-1}-\cdots-\varphi_n x_{t-n}$$

$$\sigma_0^2=\frac{1}{N-n}\sum_{t=n+1}^{N}(x_t-\varphi_i x_{t-i})^2$$

于是，一旦估计出 φ_i，就可以按照式(3)估计出 σ_0^2，因此 AR(n)时序模型的参数估计即是指对 $\varphi_1,\cdots,\varphi_n(t=1,2,\cdots,n)$ 这 n 个参数的估计。参数估计的方法分为直接法和间接法两类：直接法包括最小二乘法、解 Yule-Walker 方程法、Ulrych-Clayton 法等；间接法包括 LUD 法、BSMF 法、Burg 法等。上述方法中，用最小二乘法进行参数估计非常简单，参数估计无偏，精度高，可表示为如下方程组：$\boldsymbol{Y}=\varphi\boldsymbol{X}+\boldsymbol{a}$，式中：

$$\boldsymbol{Y}=(x_{n+1}\,x_{n+2}\cdots x_N)^{\mathrm{T}}$$

$$\boldsymbol{\varphi}=(\varphi_1\,\varphi_2\cdots\varphi_n)^{\mathrm{T}}$$

$$\boldsymbol{a}=(a_{n+1}\,a_{n+2}\cdots a_N)^{\mathrm{T}}$$

$$\boldsymbol{X}=\begin{pmatrix} x_n & x_{n-1} & \cdots & x_1 \\ x_{n+1} & x_n & \cdots & x_2 \\ \vdots & \vdots & & \vdots \\ x_{N-1} & x_{N-2} & \cdots & x_{N-n} \end{pmatrix}$$

则 ϕ 的最小二乘估计为：$\boldsymbol{\varphi}=(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}^{-1})\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$。

模型的适用性检验准则有白噪声检验准则、参余平方和检验准则、Akaike 信息检验准则等。本文以直接最小二乘估计和 Akaike 信息检验准则的 FPE（Final Prediction Error）、AIC（An-Information Criterion）、BIC 准则结合具体的实例进行讨论。FPE 准则函数：

$$FPE(n)=\frac{N+n}{N-n}\sigma_a^2$$

AIC 准则函数：

$$AIC(n)=\mathrm{N}\ln\sigma_a^2+2n$$

BIC 准则函数：

$$BIC(n)=\mathrm{N}\ln\sigma_a^2+n\ln n$$

在各自的准则函数取得最小值时的模型为适用模型。

本例通过 FPE 定阶准则进行模型阶次的选择。计算得到：$FPE(1)=0.053\,8$；$FPE(2)=0.051\,6$；$FPE(3)=0.067\,2$。

所以模型的最佳阶次为 2 次，同时计算得到的方差值为 σ^2。则得到的模型方程为：

$$x_t=\varphi_1 x_{t-1}+\varphi_2 x_{t-2}+a_t$$

通过采用最小二乘法估计，得到：$\varphi=1.123$，$\varphi_2=-0.301$。

3.4 结果讨论

利用上面得到的预测模型，就可以对该消音板的变形进行预测。这样用后 5 组数据来检验上述预测模型，表 2 所示是该模型监测值、预报值和预报残差值表。

（下转 237 页）