

用 AR 模型建立变形预测模型的研究*

许国辉¹, 张新长²

(1. 广州大学土木工程学院, 广东 广州 510405;
2. 中山大学城市与资源规划系, 广东 广州 510275)

摘要:在用时间序列分析方法建立变形测量的预测模型时, 基于动态数据系统方法的传统 F 检验定阶法, 由于样本的随机性可能过早地退出对模型的循环检验, 从而不能找到合适的预测模型。在探讨如何用 AR 模型建立变形测量预测模型的基础上, 提出了一种改进的建模方法, 并在实例的应用中取得了较好的效果。

关键词: AR 模型; 时间序列; 预测; 阶次

中图分类号: P228 **文献标识码:** A **文章编号:** 0529-6579 (2004) 01-0107-03

建立变形体的预测模型并对其动态变化趋势进行预测, 是变形测量的一项重要工作内容。自回归模型 (简称为 AR 模型) 是建立变形预测模型的一种常用模型^[1], 对于 AR 模型的运用, 模型识别与检验的核心问题是确定模型的阶次。在变形测量中, 一般采用传统的 F 检验定阶法^[2], 但在样本数有限的情况下, 样本的随机性可能使 F 检验定阶法在找到合适的模型阶次前就退出了循环检验, 因此有必要对传统的 F 检验定阶法进行改进。

1 AR 模型及建模方法

AR 模型用于描述一个平稳的随机过程, 其一般表达式为^[2,3]:

$$x(t) = a_1 x(t-1) + a_2 x(t-2) + \cdots + a_p x(t-p) + \varepsilon(t) \quad (1)$$

式中, $x(t)$ 为信号时间序列, p 为模型阶次, a_i 为模型参数, $\varepsilon(t)$ 为白噪声序列。

由式(1)可以看出, AR 模型的形式完全由模型阶次 p 所决定, 一般称式 (1) 为 AR (p) 模型。设有 t 个样本观测值分别为 $x(0), x(1), \cdots, x(t)$, 若预测模型的阶次为 p , 由式(1)可得 $t-p$ 个等式

$$\begin{cases} x(p) = a_1 x(p-1) + a_2 x(p-2) + \cdots + a_p x(0) + \varepsilon(p) \\ x(p+1) = a_1 x(p) + a_2 x(p-1) + \cdots + a_p x(1) + \varepsilon(p+1) \\ \vdots \\ x(t) = a_1 x(t-1) + a_2 x(t-2) + \cdots + a_p x(t-p) + \varepsilon(t) \end{cases} \quad (2)$$

令:

$$X(t) = \begin{bmatrix} x(p) \\ x(p+1) \\ \cdots \\ x(t) \end{bmatrix}, A = \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_p \end{bmatrix},$$

$$B(t) = \begin{bmatrix} x(p-1) & x(p-2) & \cdots & x(0) \\ x(p) & x(p-1) & \cdots & x(1) \\ \cdots & \cdots & \cdots & \cdots \\ x(t-1) & x(t-2) & \cdots & x(t-p) \end{bmatrix},$$

$$\Delta(t) = \begin{bmatrix} \varepsilon(p) \\ \varepsilon(p+1) \\ \cdots \\ \varepsilon(t) \end{bmatrix}$$

则式 (2) 的矩阵形式为:

$$X(t) = B(t)A + \Delta(t) \quad (3)$$

设 $\hat{A}^T = [\hat{a}_1 \quad \hat{a}_2 \quad \cdots \quad \hat{a}_p]$ 为 A^T 的估计值, 根据最小二乘原理, 其最佳估计为:

$$\hat{A} = (B^T(t)B(t))^{-1} B^T(t)X(t) \quad (4)$$

预测模型则为

$$\hat{x}(t) = \hat{a}_1 x(t-1) + \hat{a}_2 x(t-2) + \cdots + \hat{a}_p x(t-p) \quad (5)$$

预测残差为:

$$\hat{\varepsilon}(t) = \hat{x}(t) - x(t) \quad (6)$$

式中, $x(t), \hat{x}(t)$ 为时间序列的当前值和预测值, $x(t-i)$ 为时间序列的过去值。

2 F 检验定阶法

设存在 AR(p) 模型与 AR($p+1$) 模型, 对于同

* 收稿日期: 2003-04-08

基金项目: 测绘遥感信息工程国家重点实验室开放基金资助项目 (WKL (020) 0302)

作者简介: 许国辉 (1960 年生), 男, 高级工程师; 通讯联系人: 张新长; E-mail: eeszxc@zsu.edu.cn

一时间序列,按最小二乘法求得的模型参数分别为 $\hat{A}(p)$ 和 $\hat{A}(p+1)$,与之相应的残差分别为 $\varepsilon_p(t)$ 、 $\varepsilon_{p+1}(t)$,残差的平方和分别为 $RSS(p) = \sum_i \varepsilon_p^2(i)$ 、

$RSS(p+1) = \sum_i \varepsilon_{p+1}^2(i)$,构造统计量 F

$$F = \frac{(RSS(p) - RSS(p+1))(n-p)}{RSS(p+1)} \quad (7)$$

则 F 服从 $F(1, n-p)$ 分布,其中 n 为时间序列的样本数。对于给定的显著性水平 α ,作如下假设检验:

H_0 : AR(p) 模型是合适的(即 $a_{p+1} = 0$);

H_1 : AR(p) 模型不合适(即 $a_{p+1} \neq 0$);

当 $F > F_\alpha$ 时,拒绝原假设 H_0 ,即 AR(p) 与 AR($p+1$) 模型存在显著差异,AR(p) 模型不是合适的模型;当 $F < F_\alpha$ 时,接受原假设 H_0 ,即 AR(p) 模型是合适的,模型的阶次为 p ;

为了确定模型的阶次,一般从 AR(1) 模型开始,逐次递增模型的阶次,循环检验,直至检验原假设 H_0 被接受为止。

但在变形测量中,由于时间序列的样本数十分有限,样本的随机性可能导致这样的情况:合适的预测模型是 AR(p^*) 模型,并且 AR(p^*) 模型与 AR(p^*-1) 模型之间存在显著的差异,但由于样本的随机性使得 AR(p^*-1) 模型与 AR(p^*-2) 模型之间的差异并不显著,这时用上述 F 检验定阶法得到的模型将是 AR(p^*-2) 模型,而不是合适的 AR(p^*) 模型。

3 改进后的 F 检验定阶法

为了避免上述情况,可对 F 检验定阶法作如下改进:第一次检验不是从最低阶 $p=1$ 开始,而是从某个初始值 p_0 开始,并且第一次检验不是确定模型的最合适阶次 p^* ,而是判别 p^* 的方向。具体的方法为:首先取 $p=p_0$,同时拟合 AR(p) 模型与 AR($p+1$) 模型,并按式(7)构造统计量 F ,这时假设检验为:

$H_0: p^* \leq p_0$,即合适的阶次低于或等于 p_0 ;

$H_1: p^* > p_0$,即合适的阶次高于 p_0 ;

对于给定的显著性水平 α ,当 $F > F_\alpha$ 时,拒绝原假设 H_0 ,即合适的阶次应高于 p_0 ,以后则按传统的 F 检验定阶法确定模型的阶次;当 $F < F_\alpha$ 时,接受原假设 H_0 ,即合适的模型阶次低于或等于 p_0 ,以后按逐次递减阶次的方法确定模型的阶次;即先拟合 AR($p-1$) 模型,再检验 AR(p) 模型与 AR($p-1$) 模型是否存在显著的差异,其检验统计量为

$$F = \frac{(RSS(p-1) - RSS(p))(n-p+1)}{RSS(p)} \quad (8)$$

假设检验为:

H_0 : AR(p) 模型不合适(即 $p > p^*$);

H_1 : AR(p) 模型是合适的(即 $p = p^*$);

统计量 F 服从 $F(1, n-p+1)$ 分布,对于给定的显著性水平 α ,当 $F > F_\alpha$ 时,拒绝原假设 H_0 ,即 AR(p) 模型是合适的,模型的阶次为 p ;当 $F < F_\alpha$ 时,接受原假设 H_0 ,即 AR(p) 模型不合适,再拟合 AR(p_0-2) 模型,并对 AR(p_0-1) 模型与 AR(p_0-2) 模型进行检验,依次循环,直至找到合适的模型。

对于变形测量而言,合适的模型阶次一般不是很高,模型阶次的初始值可取 5~8。

4 应用实例及预测效果

本文仅以深圳市某大厦的沉降测量为例说明按新方法建模的预测效果。大厦沉降测量的监测期为 16 个月,布设了 18 个测点,共观测了 42 期沉降。我们用其中一点的前 26 期观测值建模,为了说明新的 F 定阶法与传统的 F 定阶法之间的差异,表 1 列出了阶次不大于 6 的各阶模型参数 $\hat{a}_i$ 及残差平方和 RSS ,表中最后两栏为相邻阶次模型的检验统计量 F 及检验临界值 $F_\alpha(1, n-p)$ 。

取 F 检验的显著性水平 $\alpha = 0.05$,根据传统的 F 检验定阶法,模型的阶次为 2,按新的 F 检验定阶法,取模型阶次的初始值 $P_0 = 5$,得模型阶次为 4。

表 1 深圳市某大厦沉降预测模型的阶次及参数

Tab.1 The rank and parameters of forecasting model for building subsidence in Shenzhen city

AR(p)	模型参数						RSS	F	$F_\alpha(1, n-p)$
	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{a}_5$	$\hat{a}_6$			
AR (1)	1.0391						461.49		
AR (2)	1.9082	-0.9045					108.16	81.66	4.24
AR (3)	2.1024	-1.3414	0.2458				100.53	1.82	4.26
AR (4)	2.0620	-1.3393	0.3391	-0.0539			81.03	5.53	4.28
AR (5)	1.9724	-1.1250	-0.0313	0.4229	-0.2334		73.85	2.13	4.30
AR (6)	1.9185	-1.0015	-0.0650	0.1257	0.2843	-0.2600	68.25	1.72	4.32

表 2 预测结果比较

Tab.2 Comparison of forecasting results

期号	实测	传统方法预测结果			新方法预测结果		
	沉降值/mm	沉降值/mm	预测误差/mm	预测相对误差	预测值/mm	预测误差/mm	预测相对误差
27	138	138.0	0.0	0.01	138.1	0.1	0.08
28	141	143.0	2.0	1.45	143.2	2.2	1.56
29	147	144.2	-2.8	1.88	144.1	-2.9	1.96
30	155	153	-2.0	1.31	153.9	-1.1	0.71
31	161	162.8	1.8	1.13	163.1	2.1	1.31
32	163	167.0	4.0	2.47	166.6	3.6	2.23
33	164	165.4	1.4	0.86	165.1	1.1	0.68
34	166	165.5	-0.5	0.29	166.1	0.1	0.06
35	171	168.4	-2.6	1.50	169.2	-1.8	1.03
36	181	176.2	-4.8	2.67	177.1	-3.9	2.15
37	188	190.7	2.7	1.45	191.7	3.7	1.94
38	193	195.0	2.0	1.05	194.3	1.3	0.66
39	200	198.2	-1.8	0.88	198.3	-1.7	0.83
40	206	207.1	1.1	0.52	207.9	1.9	0.93
41	210	212.2	2.2	1.05	212.2	2.2	1.06
42	212	214.4	2.4	1.13	214.5	2.5	1.20
RRS		94.05			84.35		

将传统方法所确定的模型与新方法所确定的模型分别用于后期的预测，后 16 期的预测结果如表 2 所示。与实测结果相比较，用传统方法所确定的 AR (2) 模型进行预测，其最大相对误差为 2.67%，残差平方和为 94.05 mm²；用新方法所确定的 AR (4) 模型进行预测，其最大相对误差为 2.23%，残差平方和为 84.35 mm²，AR (4) 模型优于 AR (2) 模型。

参考文献：

- [1] 杨元喜, 崔先强. 动态定位有色噪声影响函数[J]. 测绘学报, 2003(1): 6-10.
- [2] 陈玉祥, 张汉亚. 预测技术与应用[M]. 北京: 机械工业出版社, 1985.
- [3] 阎凤文. 测量数据处理方法[M]. 北京: 原子能出版社, 1990.
- [4] 潘国荣, 王穗辉. 建筑物动态变形的模型辨识与预测[J]. 测绘学报, 1999(4): 340-344.

A Study of the Establishment of Deformation Forecasting Model with AR Model

XU Guo-hui¹, ZHANG Xin-chang²

(1. School of Civil Engineering, Guangzhou University, Guangzhou 510405, China)

(2. Department of City and Resources Program, Sun Yat-sen University, Guangzhou 510275, China)

Abstract: When time series analysis method is used to establish forecasting model of deformation survey, the traditional *F*-test rank determining method which is based on DDS method may withdraw from the cycle of testing model too early to find the appropriate model because of the random characteristics of the sample. Based on the establishment of forecasting model of deformation survey with AR model, a improved modeling method is proposed.

Key words: AR model; time series; forecast; rank