

DOI:10.3876/j.issn.1000-1980.2017.02.002

空间自回归模型在水库边坡位移预测中的应用

秦 栋,陈慧艳

(大唐环境产业集团股份有限公司大唐(北京)水务工程技术有限公司,北京 100097)

摘要: 针对传统位移监测很少考虑不同测点之间相互作用的问题,基于经济学领域空间计量学基本理论,研究了空间自回归模型在边坡位移预测中的应用。以某工程高边坡外观位移数据为例,对边坡的位移状况进行预测,并与传统的自回归积分滑动平均模型相比较。结果表明:(a)在空间自相关系数较为显著的条件下,运用空间自回归模型可以较为精确地预测边坡变形状况,且优于传统模型;(b)空间自回归模型相较于传统模型参数更加简洁、考虑的影响因素更全面,可以同时

关键词: 空间计量学;空间自回归模型;水库边坡;边坡位移预测;自回归积分滑动平均模型
中图分类号:TV698.1 **文献标志码:**A **文章编号:**1000-1980(2017)02-0104-05

Application of spatial autoregressive model in prediction of high slope deformation

QIN Dong, CHEN Huiyan

(Datang (Beijing) Water Engineering Technology Co., Ltd.,
Datang Environment Industry Group Co., Ltd., Beijing 100097, China)

Abstract: In consideration of the fact that the interaction between different monitoring points is seldom considered in traditional displacement monitoring, the application of the spatial autoregressive model in the prediction of high slope deformation was examined based on the basic theory of spatial econometrics. Using a high slope as an example, the results of displacement prediction based on the monitoring data of exterior deformation were compared with those from the traditional autoregressive integrated moving average model. The results show that, with a significant spatial autocorrelation coefficient, the spatial autoregressive model is more accurate than the traditional model in prediction of slope deformation, and that, compared with the traditional model, the spatial autoregressive model has more concise parameters, considers the factors more comprehensively, and can perform displacement prediction at different monitoring points at the same time.

Key words: spatial econometrics; spatial autoregressive model; reservoir slope; slope deformation prediction; autoregressive integrated moving average model

水库边坡安全关系到水利工程的根本安全,如果边坡出现滑坡,后果不堪设想。历史上有很多水库边坡滑坡灾害的实例,滑坡也因此被称为世界三大自然灾害之一。针对边坡的安全监测,一般主要进行外部的变形监测、内部的多点位移计变形监测、应力计应力监测等^[1-5]。因为内观仪器很容易损坏,因此最为可靠和稳定的监测数据为外观变形数据。外观变形数据是了解边坡变形状态的重要参考和依据。

为了研究外观变形数据和边坡安全状态的关系,有很多算法被用于边坡变形的预测,以此判断边坡的安全变化趋势^[6-9]。根据 Tobler 提出的著名的“地理学第一定律”:万事万物之间存在广泛的关联性,并且相近的事物之间关系更紧密。边坡作为一个整体,不同区域的变形有可能存在很强的关联性,但传统的变形预测很少考虑多测点变形值之间的相互关系。空间自回归模型是空间计量经济学领域的模型之一,主要用

来分析空间不同区域参数值之间的关联性,从而实现对数值的分析和预测^[10-15]。边坡的监测数据也有变形值、测点坐标、时间段等多个维度,本质上可以被看作一种面板数据。本文基于空间计量学基本理论,探讨这种空间自回归模型在边坡变形面板数据预测中的应用的可行性,并和传统的单测点模型的预测效果进行比较。

1 空间自回归模型

1.1 空间权重矩阵

为了衡量不同测点之间的关联性,有人提出了空间权重矩阵,这也是空间自回归模型最基础的部分,是进行空间自相关分析的基础。假设空间有 n 个测点,如果用 w_{ij} 表示测点 i 和测点 j 之间的距离,定义空间权重矩阵如下:

$$\mathbf{W} = \begin{bmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & & \vdots \\ w_{n1} & \cdots & w_{nn} \end{bmatrix} \tag{1}$$

关于 w_{ij} 的最为经典的取值方法是根据不同测点的邻近关系进行取值,公式如下:

$$w_{ij} = \begin{cases} 1 & i \text{ 和 } j \text{ 相邻} \\ 0 & i \text{ 和 } j \text{ 不相邻} \end{cases} \tag{2}$$

这样就得到了初始的空间权重矩阵 $\mathbf{W}$ 。计算 w_{ij} 之后,还需要对空间权重矩阵进行“行标准化”,即用每一个 w_{ij} 分别除以 w_{ij} 所在行之和,计该值为 $\tilde{w}_{ij}$,行标准化后的空间权重矩阵为 $\tilde{\mathbf{W}}$ 。这样矩阵每一行的和都为1,如果将空间权重矩阵乘以所有测点数值的列矩阵,这样可以很方便地得到每个测点临近测点的平均值。

1.2 空间自相关系数

空间自相关系数一般被用来衡量空间上数据的聚集程度。最常用的空间自相关系数为莫兰指数 (I),其取值一般介于-1~1之间。当自相关系数大于0,说明空间数值存在高高聚集或低低聚集;当自相关系数小于0,说明空间数值存在高低聚集或者低高聚集;当自相关系数接近于0,说明空间基本不存在空间自相关关系,各测点的数值相对比较独立。莫兰指数的计算公式为

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{S^2 \sum_{i=1}^n \sum_{j=1}^n w_{ij}} \tag{3}$$

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

其中

式中: x_i 、 $\bar{x}$ ——单测点值、均值; S ——样本方差。

莫兰指数 I 的期望为

$$E(I) = \frac{-1}{n-1} \tag{4}$$

标准化的莫兰指数服从渐进标准正态分布,可以使用标准正态的临界值进行显著性检验,一般用 p 值表示。一般认为,当 $p<0.05$,莫兰指数显著成立。

1.3 空间自回归模型的构建

因为边坡不同测点的监测数据具有时间和测点2个维度,所以边坡监测数据从根本上讲是一种面板数据。因为测点的位置还具有一定空间特性,所以称这种数据为空间面板数据。如果仅仅考虑时间的一阶滞后,则一阶动态面板数据的空间自回归模型为

$$u_{mt} = \tau u_{m,t-1} + \lambda \mathbf{w}_m' u_t + \mu \mathbf{w}_m' u_{t-1} + v_m + \varepsilon_{mt} \tag{5}$$

其中

$$\mathbf{w}_m' u_t = \sum_{n=1}^k w_{mn} y_{nt}$$

式中: u_{mt} ——因变量; τ 、 λ 、 μ ——回归系数; y_{nt} ——因变量矩阵中的元素; $\mathbf{w}_m'$ —— $\mathbf{W}$ 的第 m 行; v_m ——个体效应量; ε_{mt} ——随机误差。

如果仅仅针对横截面数据,由于共时性的存在,最小二乘法的估算将不会随着样本量的增加而收敛到

它们的“真实”值。相反,在最小二乘形式中被忽略的反馈效应或依赖效应将增大,因此对空间自回归模型一般采用最大似然估计。

假设扰动项 $\varepsilon_{mt} \sim N(0, \sigma^2 I_k)$, 有样本对数似然函数如下:

$$\ln L(u|\lambda, \sigma^2, \beta) = -\frac{k}{2} \ln 2\pi - \frac{k}{2} \ln \sigma^2 + \ln(\text{abs}|A|) - \frac{1}{2\sigma^2} (Au - X\beta)' (Au - X\beta) \tag{6}$$

其中 $A \equiv I - \lambda W$ $X = \begin{bmatrix} u_{1,t-1} & w_1'u_{t-1} \\ \vdots & \vdots \\ u_{k,t-1} & w_k'u_{t-1} \end{bmatrix}$ $\beta = \begin{bmatrix} \tau \\ \mu \end{bmatrix}$

式中: $\text{abs}|A|$ —— A 的绝对值。

对式(6)的最大化估计分 2 步进行:(a)在给定 λ 的情况下,选择最优的 β, σ^2 ;(b)根据最优 β, σ^2 ,选择最优的 λ 。

对于面板数据,如果采用固定效应模型,可以先做组内离差变换,去掉 v_m ,然后采用类似横截面模型的最大似然估计得到各系数;如果采用随机效应模型,可以通过广义离差变换,然后采用最大似然估计。随机模型和固定效益模型的选择需要通过豪斯曼检验(Hausman Test)确定。这样就可以确定式(5)的参数,至此完成空间自回归模型的建立。

1.4 模型运作流程

综上,可以发现空间自回归模型仅仅适用于空间自相关效应比较明显的情况。根据这点,可以简单将空间自回归模型的运作流程归纳如下:(a)收集并整理变形数据,剔除错误数据,插补缺失数据。(b)计算变形数据的空间自相关系数,并检验其显著性。(c)如果存在显著的空间自相关性质,则采用空间自回归模型;如果不存在显著的空间自相关性质,则采用传统预测模型。(d)回代各项回归系数,检验模型的预测精度。

2 工程实例

2.1 工程概况

我国西南某省某电站下游为满足工程布置需要,开挖形成长约 1 530 m,最大坡高为 230 m 的人工高边坡。为保证工程运行安全,需要对高边坡进行安全监测。外观变形监测共设有 16 个监测剖面,一共布置了 94 个变形测点,实际观测 81 个测点。测点布置情况如图 1 所示。因为 L11 断面和 L12 断面相距较远,可以认为它们之间基本没有相互影响。以 L11 断面为界,将边坡分为 L1 ~ L11 和 L12 ~ L16 两部分。为方便起见,只研究测点分布均匀且损坏率较低的 L1 ~ L11 部分。

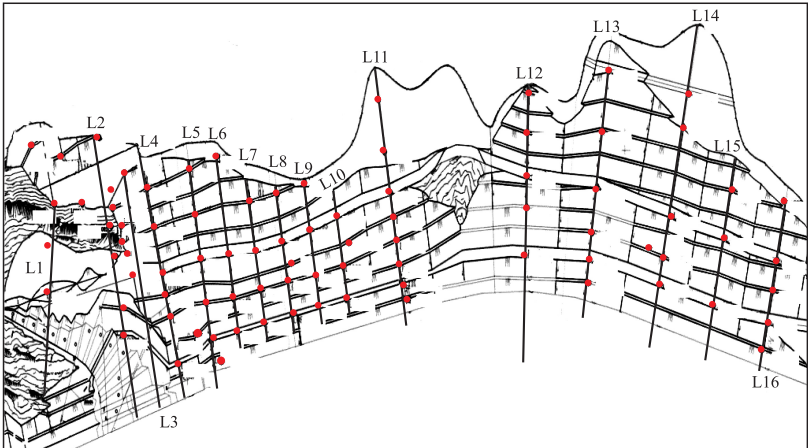


图 1 某工程下游人工高边坡外观变形测点布置

Fig.1 Layout of measuring points for exterior deformation of an artificial high slope downstream of a project

L1 ~ L11 断面共有 49 个有效测点,将这 49 个测点单独提取出来并对所有测点进行编号。这里考虑测点在水平和竖直 2 个方向上的临近关系,如果两点有直接的临近关系,则认为这两点“相邻”,反之认为“不

相邻”;同时对所有测点之间的相互影响做出假设:即认为测点对临近测点的影响程度与空间权重矩阵成正比(空间自回归模型的基本假设)。如果用连接的直线表示测点之间的临近关系,那不同测点之间的临近拓扑关系如图 2 所示。

根据图 2 的拓扑关系图,可以根据式(2)得到相应的空间权重矩阵 W 。进一步,对空间权重矩阵进行“行标准化”操作,便可以得到标准化的空间权重矩阵。

2.2 自相关性检验

为验证空间自回归模型的适用性,还需要了解边坡不同测点测值的空间自相关程度。本文截取了边坡 49 个测点从 2010 年 2 月到 2013 年 12 月共 47 个时间截面沿左右岸方向(X 轴方向)的变形测值,并根据式(3)分别计算每个截面的空间自相关系数。其中部分空间自相关系数如表 1 所示,47 个时间截面的莫兰指数和 P 的均值分别为 0.284、0.088。由表 1 可知,大部分时间截面都表现出显著的空间自相关性质($P < 0.05$),所有时间截面平均莫兰指数为 0.284,表明 L1 ~ L11 断面表现出较强的空间正自相关性质,可以采用空间自回归模型对边坡的变形进行预测。

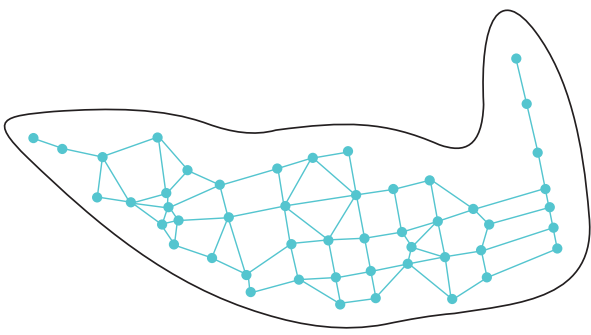


图 2 L1 ~ L11 断面区域测点拓扑关系

Fig. 2 Topology diagram of measuring points in sections L1 through L11

表 1 不同时间截面的莫兰指数

Table 1 Moran's I for different time sections							
序号	日期	莫兰指数	p	序号	日期	莫兰指数	p
1	2010-02-01	0.257	0.011	12	2011-01-01	0.048	0.520
2	2010-03-01	0.582	0	⋮	⋮	⋮	⋮
3	2010-04-01	0.063	0.445	40	2013-05-01	0.276	0.007
4	2010-05-01	0.434	0	41	2013-06-01	0.402	0
5	2010-06-01	0.378	0	42	2013-07-01	0.483	0
6	2010-07-01	0.062	0.446	43	2013-08-01	0.298	0.003
7	2010-08-01	0.334	0.001	44	2013-09-01	0.137	0.148
8	2010-09-01	0.335	0.001	45	2013-10-01	-0.192	0.112
9	2010-10-01	0.515	0	46	2013-11-01	0.026	0.665
10	2010-11-01	0.162	0.096	47	2013-12-01	0.151	0.108
11	2010-12-01	0.324	0.002				

2.3 拟合精度分析

47 组数据中,用前 42 组数据进行拟合建模,对后 5 组数据进行预测并与实测值进行比较。为验证模型的拟合精度,选用传统的自回归积分滑动平均模型(autoregressive integrated moving average model, ARIMA)对所有测点的变形进行拟合,并与空间自回归模型的拟合精度相比较。

不同模型的回归结果及拟合精度见表 2。因为空间自回归模型在边坡监测领域的应用较少,这里分别采用了随机效应模型和固定效应模型进行建模。可以看出,空间自回归模型(包含随机效应模型和固定效应模型)的拟合精度要明显高于传统时间序列模型。

表 2 空间自回归模型和 ARIMA 模型拟合精度比较

Table 2 Comparison of fitting accuracy of spatial autoregressive model and ARIMA model								
模型	d_1	R	W_1	W_2	S	R_{SQ}	A_{IC}	B_{IC}
随机效应模型	0.876 2	-0.092 3	-0.434 5	-0.145 9	0.577 5	0.920 4	6 057.297	6 103.061
固定效应模型	0.813 5		-0.395 1	-0.126 5	0.578 9	0.920 3	5 914.754	5 949.077
ARIMA 模型						0.910 8		

注: d_1 、 W_1 、 W_2 为模型系数; R 为相关系数; R_{SQ} 为残差平方和; A_{IC} 、 B_{IC} 为信息准则; S 为 Spatial 空间自相关系数。

2.4 预测精度分析

运用随机效应模型、固定效应模型、ARIMA 模型对变形数值预测进行,预测时如果没有实测值可以将预测值作为实测值代入下一步预测之中。模型的最后预测结果见表3。从表3可知,随着时间的推移,3种模型的精度都出现了下降的情况;但是整体来看,随机效应模型和固定效应模型的精度相当,其中固定效应模型的精度要更高一些,但都明显高于传统单测点的 ARIMA 模型。

3 结 语

基于空间计量学基本理论,将空间自回归模型引入边坡位移预测的研究当中,计算简单,原理清晰。研究表明:对于整体性较好的边坡,可以用空间自相关系数验证其整体性;在空间自相关系数较为显著的条件下,运用空间自回归模型可以较为精确地预测边坡变形状况,且优于传统单测点模型;空间自回归模型相较于传统模型参数更加简洁,可以同时空间所有测点进行估计;空间自回归模型还可以在在一定程度上揭示边坡的空间变形性质,值得进一步研究和推广。

参考文献:

[1] 吴中如. 水工建筑物安全监控理论及其应用[M]. 北京:高等教育出版社,2003.

[2] 吴中如,顾冲时. 重大水工结构隐患病害检测与健康诊断[M]. 北京:高等教育出版社,2005.

[3] 吴中如,顾冲时. 大坝原型反分析及其应用[M]. 南京:江苏科学技术出版社,2002.

[4] QIN Dong,ZHENG Xueqin,LIU Hezhi,et al. Back analysis of high concrete arch dams' general elastic modulus range [J]. Journal of Food Agriculture& Environment,2013,11(1):1182-1186.

[5] GU Chongshi,QIN Dong,LI Zhanchao,et al. Study on semi-parametric statistical model of safety monitoring of cracks in concrete dams [J]. Mathematical Problems in Engineering,2013,2013(2):831-842.

[6] 刘大有,陈慧灵,齐红,等. 时空数据挖掘研究进展[J]. 计算机研究与发展,2013,50(2):225-239. (LIU Dayou, CHEN Huiling,QI Hong, et al. Study on spatiotemporal data mining [J]. Journal of Computer Research and Development,2013,50(2):225-239. (in Chinese))

[7] 温廷新,张波. 露天煤矿边坡稳定性的随机森林预测模型[J]. 科技导报,2014,32(4/5):105-109. (WEN Tingxin,ZHANG Bo. Stochastic forest forecasting model for slope stability of opencast coal mine [J]. Acta Scientia Sinica,2014,32(4/5):105-109. (in Chinese))

[8] 张豪,罗亦泳. 基于人工免疫算法的边坡稳定性预测模型[J]. 煤炭学报,2012,37(6):911-917. (ZHANG Hao, LUO Yiyong. Prediction model of slope stability based on artificial immune algorithm [J]. Acta Coal Science,2012,37(6):911-917. (in Chinese))

[9] 宋马林,王舒鸿,汝慧萍. 一种新的考虑时间和空间的相关系数及其算例[J]. 数量经济技术经济研究,2010(7):142-152. (SONG Malin,WANG Shuhong,RU Huiping. A new correlation coefficient considering time and space and its calculations [J]. Quantitative Technology and Economic Research,2010(7):142-152. (in Chinese))

[10] 王艳霞. 模糊数学在边坡稳定分析中的应用[J]. 岩土力学,2010,31(9):170-173. (WANG Yanxia. Application of fuzzy mathematics in slope stability analysis [J]. Rock and Soil Mechanics,2010,31(9):170-173. (in Chinese))

[11] 霍家平,傅中志,刘若星. 水电站边坡坍塌堆积体边坡稳定与变形分析[J]. 水利水电科技进步,2012,32(增刊2):98-101. (HUO Jiaping,FU Zhongzhi,LIU Ruoxing. Stability and deformation analysis of slope slope stabilized slope of hydropower station [J]. Advances in Water Resources and Hydropower Engineering,2012,32(Sup2):98-101. (in Chinese))

[12] 姜德义,朱合华,杜云贵. 边坡稳定性分析与滑坡防治[M]. 重庆:重庆大学出版,2005.

[13] 高玉峰,王迪,张飞. 三维土质边坡稳定性分析方法研究现状与展望[J]. 河海大学学报(自然科学版),2015,43(5):456-464. (GAO Yufeng,WANG Di,ZHANG Fei. Studies and prospects of three-dimensional soil slope stability analysis methods [J]. Journal of Hohai University (Natural Sciences),2015,43(5):456-464. (in Chinese))

[14] 狄圣杰,李晓敏,魏强. GRNN 在边坡稳定预测分析中的应用[J]. 水利水电科技进步,2011,31(3):80-83. (DI Shengjie,LI Xiaomin,WEI Qiang. Application of GRNN in slope stability prediction and analysis [J]. Advances in Water Science and Technology of Water Conservancy,2011,31(3):80-83. (in Chinese))

[15] 王海军,涂凯,闫晓荣. 基于果蝇优化算法的 GRNN 模型在边坡稳定预测中的应用[J]. 水电能源科学,2015,33(1):3000-3004. (WANG Haijun,TU Kai,YAN Xiaorong. Application of GRNN model based on Drosophila optimization algorithm in slope stability prediction [J]. Power and Hydropower Engineering,2015,33(1):3000-3004. (in Chinese))

表3 不同模型预测精度比较

日期	预测精度		
	随机效应模型	固定效应模型	ARIMA 模型
2013-08-01	0.982 1	0.982 3	0.991 2
2013-09-01	0.965 1	0.964 3	0.952 3
2013-10-01	0.965 3	0.962 2	0.953 2
2013-11-01	0.921 1	0.932 2	0.912 8
2013-12-01	0.832 1	0.846 6	0.754 6