

基于动态参数线性回归模型的路堤沉降预测*

董友弟¹, 武斌敬¹, 王光辉²

(1. 中国煤田地质局水文地质局物测队, 河北 邯郸 056021; 2. 中国矿业大学 环境与测绘学院, 江苏 徐州 221008)

摘要:随着观测信息的不断丰富, 常规的线性模型的不足逐渐凸现, 影响了沉降数据分析的精度。据此, 为了克服常规沉降数据处理模型的不足提出了动态参数线性回归模型。该模型要求参与建模的数据个数保持不变, 利用新观测的数据取代旧数据, 获取新的模型参数, 进而建立动态参数线性回归模型进行沉降预测。通过实例验证, 动态对数模型预测值平均误差率为 5.4%, 动态多项式模型预测值平均误差率为 2.7%; 与其对应的常规模型相比平均预测精度分别提高了 6.4 和 2.1 个百分点, 动态模型的效益十分明显; 从建模精度 σ^2 来看, 动态对数模型 $\sigma^2 = 11.5$, 动态多项式 $\sigma^2 = 2.3$, 均小于常规模型的 σ^2 。

关键词:沉降预测; 线性回归; 动态参数模型**中图分类号:** P 258**文献标识码:** A**文章编号:** 1007-9394(2008)01-0012-04

Brief Prediction of Embankment Settlement Based on Dynamic Linear Regression Model

DONG You-di¹, WU Bin-jing¹, WANG Guang-hui²

(1. Physical Survey Team of Hydrogeology Bureau of China Coal Field Geologic Administration, Handan Hebei 056021, China; 2. Faculty of Environment and Surveying & Mapping, China University of Mining, Xuzhou Jiangsu 221008, China)

Abstract: Linear regression model is one of the most widely used methods in the prediction of embankment settlement. With the data augmented, the disadvantages of the traditional linear model stands out extremely, so the accuracy of prediction is reduced. To overcome deficiencies in the traditional model, dynamic linear regression model is put forward. This model requests that the number of data used to make model is kept unchanged, new data replaced the old data, so the new parameters are gained. Ideal result is gotten through the instance. The average error of Dynamic Logarithm Model was 5.4%, and the average error of Dynamic Polynomial Model was 2.7%. For the model accuracy σ^2 , Dynamic Logarithm Model was $\sigma^2 = 11.5$ and Dynamic Polynomial Model was $\sigma^2 = 2.3$, both of them all below general model σ^2 .

Key words: embankment settlement prediction; linear regression; dynamic parameter model

0 引言

随着中国道路建设的飞速发展, 高填方及特殊土等路堤大量出现。在建设过程中, 为确保铁路公路建设质量、指导施工和安全运营, 相应的铁路和公路路基规范都要求对路堤进行沉降观测。对于沉降数据的处理方法有很多, 如神经网络、时间序列和小波分析等。通常用建模的精度和预测结果的准确度来衡量数据处理方法的优劣, 并且从模型中找出路堤沉降规律, 及时发现问题、处理问题。

笔者在生产实践中运用动态参数线性回归模型解决了不少变形观测问题, 本文试将该模型运用到路堤的沉降预测中来, 并用实例验证了其效果。

1 动态参数线性回归模型的提出

线性回归模型是应用最为广泛的模型之一, 也是其他模型应用的坚实基础。在现实世界中, 许多变量之间具有线性或近似线性的依赖关系; 有些变量之间是非线性的, 但经过适当的变换后, 新产生的变量之间就具有了近似的线性关系; 另外线性关系容易处理, 也容易达到比较理想的结果。

1.1 线性回归模型

通常可以把线性回归模型分为多元线性回归和可化为线性回归模型的非线性回归两种, 而后者主要有幂函数模型、指数函数模型、对数函数模型、双曲函数模型、多项式函数模型, 这些模型简称为常规模型。

限于篇幅, 本文仅列出 3 种常规模型^[1]:

多元线性回归模型:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m \tag{1}$$

对数函数模型:

$$y = \beta_0 + \beta_1 \ln x \tag{2}$$

多项式函数模型:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_m x^m \tag{3}$$

1.2 动态线性回归模型的思想

动态线性回归模型是为得到更好的建模精度和预测效果而提出的。当信息比较丰富时, 如果使用常规模型, 则由于观测误差的积累和数据的异常跳跃, 模型受外界影响的程度增加, 而用于描述建模精度的单位权方差 σ^2 与残差平方和呈正相关关系, 致使建模结果不够理想, 同时其模型稳定性不能得到保障, 与观测值的走势也不能够很好地吻合, 导致预测误差偏大。另外, 常规模型不能很好的反映外界因素对沉降观测值的影响和沉降本身变化规律。因此必须设法克服常规模型的这些缺点。

而动态线性回归模型就是应用新陈代谢法则, 固定参与模型求解个数的观测值(简称等维), 即考虑增加一个新数据剔除一个离预测值最近的一个旧数据, 从而获得了模型的新参数, 预测下一个数据, 如此不断进行建模和预测。动态模型能充分顾及到事物发生发展的客观规律, 因此动态模型预测精度好, 准确度高。另外, 参与建模的数据个数也较常规模型少, 简化了计算。

2 实例研究

2.1 工程概况

正在建设中的上(海)瑞(丽)高速公路湖南境内(湘)潭邵(阳)段主要为丘陵地貌, 高填深挖路基、特殊土(如软土、膨胀

土)路基较多, 为确保路基施工安全、施工质量和研究沉降变化规律, 特别是为防止桥头跳车对路桥过渡段的桥背路堤沉降严格控制, 对路堤沉降进行了施工监测。实测数据, 见表 1^[2]。

表 1 DK39+098 桥台路堤沉降实测值

Tab. 1 Actually monitoring data of DK39+098 bridge approach embankment

| 时间/×10d | 实测/mm | 时间/×10d | 实测/mm | 时间/×10d | 实测/mm |
|---------|-------|---------|-------|---------|-------|
| 1 | 6.7 | 10 | 53.8 | 19 | 98.0 |
| 2 | 13.3 | 11 | 66.7 | 20 | 99.0 |
| 3 | 25.8 | 12 | 69.7 | 21 | 102.0 |
| 4 | 30.8 | 13 | 73.8 | 22 | 104.0 |
| 5 | 35.0 | 14 | 81.8 | 23 | 104.6 |
| 6 | 39.7 | 15 | 85.0 | 24 | 107.0 |
| 7 | 41.8 | 16 | 88.2 | 25 | 107.0 |
| 8 | 45.6 | 17 | 91.4 | 26 | 107.9 |
| 9 | 48.8 | 18 | 95.0 | 27 | 108.0 |

2.2 用(可化为)线性的回归模型预测其沉降

通常情况下, 在分析采用何种线性模型时, 首先要分析散点图, 根据表 1 数据绘制出散点图, 如图 1 所示。从图 1 中不难看出, 这段路堤沉降呈现为一种非线性关系, 另外当 $t=3, t=11$ 和 $t=14$ 时, 沉降观测值有明显的跃变。初步认为对数模型(2)和多项式函数模型(3)适合本问题。

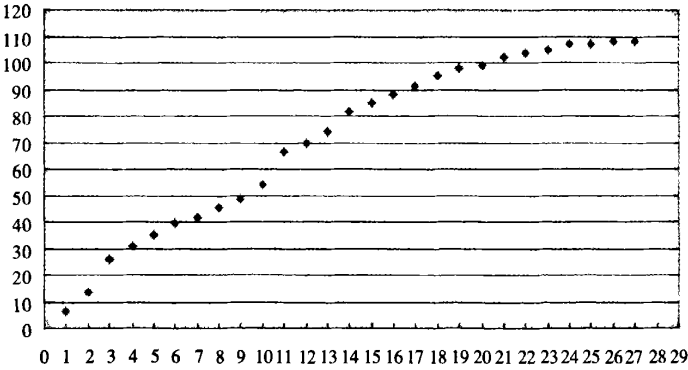


图 1 DK39+098 桥台路堤沉降实测值散点图

Fig. 1 Actually monitoring data scatter plot of DK39+098 bridge approach embankment

2.2.1 常规线性模型

为了比较建模的效果, 首先按照常规模型的算法, 获得相关的模型和预测值, 见表 2 和表 3 ($N < 11$ 时, 属于贫信息, 不宜用动态线性回归模型)。

从表 2 和表 3 中可以看出无论是对数模型还是多项式模型, 效果都不理想, 尤其是常规对数模型预测值平均误差率竟高达 11.8%, 而常规多项式模型预测值平均误差率也将近 5%。同时从反映建模精度的 σ^2 来看, 常规对数模型最大达 89.2, 最小的也有 23.0, 平均为 71.8; 常规多项式模型的 σ^2 平均值 $\bar{\sigma}^2 = 7.3$ 。很明显, 随着观测值的增多, 观测误差凸现, 模型受外

界影响的程度增加。加之受异常跳跃数据的影响, 致使预测精度降低, 误差率居高不下。

2.2.2 动态线性回归模型

对于动态线性回归模型, 首先要探讨用于建模观测值的维数 N , 从散点图中可知, 当 $N = 11$ 时, 模型基本稳定, 规律比较明显, 另外当 $N > 11$ 时, 对 σ^2 的影响不大。

多项式阶数的确定, 依据文献[3]中的 F 检验进行, 则

$$F = \frac{S_{\varepsilon_{m-1}} - S_{\varepsilon_m}}{S_{\varepsilon_m} / (n - m)} \tag{4}$$

表 2 常规对数模型

Tab. 2 Conventional logarithm model

| 时间 / $\times 10d$ | 实测 /mm | 预测值 /mm | 误差值 /mm | 误差率 /% | σ^2 |
|----------------------|-----------|------------|------------|-----------|------------|
| 12 | 69.7 | 57.3 | -12.4 | -17.7 | 23.0 |
| 13 | 73.8 | 61.6 | -12.2 | -16.5 | 33.1 |
| 14 | 81.8 | 65.7 | -16.1 | -19.7 | 41.2 |
| 15 | 85.0 | 70.2 | -14.8 | -17.4 | 55.9 |
| 16 | 88.2 | 74.3 | -13.9 | -15.8 | 66.0 |
| 17 | 91.4 | 78.1 | -13.3 | -14.6 | 73.2 |
| 18 | 95.0 | 81.6 | -13.4 | -14.1 | 78.6 |
| 19 | 98.0 | 85.0 | -13.0 | -13.2 | 83.5 |
| 20 | 99.0 | 88.3 | -10.7 | -10.8 | 87.3 |
| 21 | 102.0 | 91.1 | -10.9 | -10.7 | 88.2 |
| 22 | 104.0 | 93.9 | -10.1 | -9.7 | 89.1 |
| 23 | 104.6 | 96.5 | -8.1 | -7.7 | 89.2 |
| 24 | 107.0 | 98.8 | -8.2 | -7.7 | 87.8 |
| 25 | 107.0 | 101.0 | -6.0 | -5.6 | 86.6 |
| 26 | 107.9 | 103.0 | -4.9 | -4.5 | 84.2 |
| 27 | 108.0 | 104.8 | -3.2 | -2.9 | 81.6 |

表 3 常规多项式模型

Tab. 3 Conventional polynomial model

| 时间 / $\times 10d$ | 实测 /mm | 预测值 /mm | 误差值 /mm | 误差率 /% | σ^2 |
|----------------------|-----------|------------|------------|-----------|------------|
| 12 | 69.7 | 77.5 | 7.8 | 11.2 | 2.7 |
| 13 | 73.8 | 83.5 | 9.7 | 13.2 | 3.9 |
| 14 | 81.8 | 86.7 | 4.9 | 6.0 | 5.9 |
| 15 | 85.0 | 93.6 | 8.6 | 10.1 | 6.0 |
| 16 | 88.2 | 96.7 | 8.5 | 9.6 | 7.4 |
| 17 | 91.4 | 98.3 | 6.9 | 7.6 | 8.7 |
| 18 | 95.0 | 99.6 | 4.6 | 4.9 | 9.3 |
| 19 | 98.0 | 101.4 | 3.4 | 3.4 | 9.2 |
| 20 | 99.0 | 103.0 | 4.0 | 4.1 | 9.0 |
| 21 | 102.0 | 103.2 | 1.2 | 1.2 | 8.9 |
| 22 | 104.0 | 104.3 | 0.3 | 0.3 | 8.4 |
| 23 | 104.6 | 105.4 | 0.8 | 0.7 | 8.0 |
| 24 | 107.0 | 105.5 | -1.5 | -1.4 | 7.7 |
| 25 | 107.0 | 106.5 | -0.5 | -0.5 | 7.4 |
| 26 | 107.9 | 106.4 | -1.5 | -1.4 | 7.0 |
| 27 | 108.0 | 106.4 | -1.6 | -1.4 | 6.8 |

根据前 11 个数据分别得到了二阶、三阶和四阶的残差平方和分别为:112.558 1,24.279 1,12.025 1。计算得 $F = 6.11 < F_{0.025}(1,6) = 8.81$,接受原假设,即认为三阶与四阶没有显著区别,故取三阶。同时用三阶的多项式分别计算 $N = 8 \sim 15$ 时 σ 的值,如表 4。可以看出 $N = 11$, σ 变化已经很小。

表 4 $N = 8 \sim 15$ 对应 σ 的值

Tab. 4 The σ of $N = 8 \sim 15$

| N | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|----------|---------|---------|---------|---------|---------|---------|---------|---------|
| σ | 1.106 2 | 1.630 7 | 1.652 8 | 1.895 0 | 1.890 0 | 1.839 5 | 1.821 5 | 1.890 1 |

选取 11 个数据参与建模,分别采用对数模型和多项式模型进行计算,结果如表 5 和表 6 所示。

表 5 动态对数模型

Tab. 5 Dynamic Logarithm Model

| 时间 / $\times 10d$ | 实测 /mm | 预测值 /mm | 误差值 /mm | 误差率 /% | σ^2 |
|----------------------|-----------|------------|------------|-----------|------------|
| 12 | 69.7 | 57.3 | -12.4 | -17.7 | 23.0 |
| 13 | 73.8 | 64.0 | -9.8 | -13.3 | 24.3 |
| 14 | 81.8 | 69.3 | -12.5 | -15.3 | 30.0 |
| 15 | 85.0 | 76.6 | -8.4 | -9.9 | 32.0 |
| 16 | 88.2 | 83.3 | -4.9 | -5.6 | 26.2 |
| 17 | 91.4 | 89.4 | -2.0 | -2.2 | 17.7 |
| 18 | 95.0 | 94.9 | -0.1 | -0.1 | 8.3 |
| 19 | 98.0 | 99.4 | 1.4 | 1.4 | 4.6 |
| 20 | 99.0 | 102.9 | 3.9 | 3.9 | 3.9 |
| 21 | 102.0 | 104.6 | 2.6 | 2.6 | 4.7 |
| 22 | 104.0 | 105.7 | 1.7 | 1.6 | 1.7 |
| 23 | 104.6 | 107.8 | 3.2 | 3.1 | 1.9 |
| 24 | 107.0 | 108.8 | 1.8 | 1.7 | 2.1 |
| 25 | 107.0 | 109.6 | 2.6 | 2.5 | 0.6 |
| 26 | 107.9 | 110.6 | 2.7 | 2.5 | 1.1 |
| 27 | 108.0 | 111.1 | 3.1 | 2.9 | 1.3 |

表 6 动态多项式模型

Tab. 6 Dynamic Polynomial Model

| 时间 / $\times 10d$ | 实测 /mm | 预测值 /mm | 误差值 /mm | 误差率 /% | σ^2 |
|----------------------|-----------|------------|------------|-----------|------------|
| 12 | 69.7 | 77.5 | 7.8 | 11.2 | 2.7 |
| 13 | 73.8 | 85.0 | 11.2 | 15.2 | 3.8 |
| 14 | 81.8 | 83.8 | 2.0 | 2.4 | 4.8 |
| 15 | 85.0 | 88.8 | 3.8 | 4.5 | 4.4 |
| 16 | 88.2 | 88.8 | 0.6 | 0.7 | 3.9 |
| 17 | 91.4 | 88.0 | -3.4 | -3.7 | 3.1 |
| 18 | 95.0 | 90.3 | -4.7 | -5.0 | 3.4 |
| 19 | 98.0 | 95.0 | -3.0 | -3.0 | 3.6 |
| 20 | 99.0 | 100.9 | 1.9 | 1.9 | 2.4 |
| 21 | 102.0 | 102.3 | 0.3 | 0.3 | 2.4 |
| 22 | 104.0 | 102.5 | -1.5 | -1.5 | 1.1 |
| 23 | 104.6 | 106.3 | 1.7 | 1.6 | 0.6 |
| 24 | 107.0 | 106.9 | -0.1 | -0.1 | 0.7 |
| 25 | 107.0 | 107.6 | 0.6 | 0.6 | 0.2 |
| 26 | 107.9 | 107.7 | -0.2 | -0.2 | 0.2 |
| 27 | 108.0 | 108.3 | 0.3 | 0.3 | 0.2 |

动态对数模型预测值平均误差率为5.4%,动态多项式模型预测值平均误差率为2.7%;与其对应的常规模型相比平均预测精度分别提高了6.4和2.1个百分点,可见,动态模型的效益十分明显。从建模精度 σ^2 来看,动态对数模型为 $\sigma^2 = 11.5$,动态多项式 $\sigma^2 = 2.3$,均小于常规模型的 σ^2 。如果仅从建模精度来考虑的话,常规对数模型误差太大,不适合本问题,但使用了动态对数模型,随着可选择的数据增多,它也能取得较好的效果。换句话说,动态模型的建模精度有了大幅度的提高。

为直观起见,方便比较,将表1、2、3、5、6绘制成图,如图2。

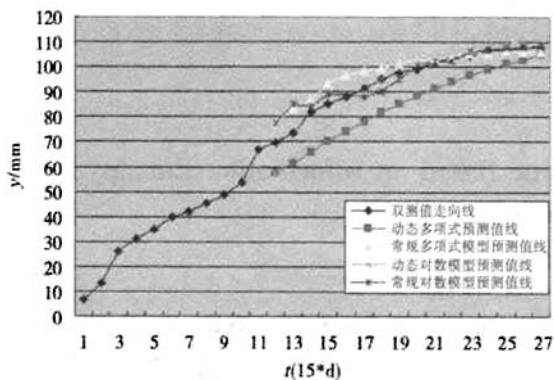


图2 4种模型预测值与观测值走势对比

Fig. 2 The predictive value of four models and the comparison of predictive value

综上所述,可选取三阶动态多项式模型为主要模型来研究潭邵段高速公路路堤的沉降规律。

3 结论

线性动态模型固定了参与模型求解的观测值个数,顾及到路堤动态变化的客观事实,能有效地预测在一个时间段内路堤沉降量,并且计算量相对常规模型小,建议今后的生产推广使用该模型。由于本文在确定参与建模的数据个数时,所用的方法还显粗糙,同时,在前文实例验证中也不难发现,动态模型抗畸变能力还有些偏弱,尚不能完全消除,因此在这些问题上还需要人们进一步进行研究。

【参考文献】

- [1] 刘大杰,陶本藻.实用测量数据处理方法[M].北京:测绘出版社,2003.
- [2] 肖武权,冷伍明.陆地沉降的时间序列动态预测方法[J].铁道工程学报,2004,12(4):52~56.
- [3] 王新洲,陶本藻,等.高等测量平差[M].北京:测绘出版社,2006.
- [4] 郝红伟. Matlab6 实用教程[M].北京:中国电力出版社,2001.
- [5] 邹德强,王桂尧.遗传算法在公路堤沉降预测中的应用[J].长沙交通学院学报,2004,20(1):19~24.
- [6] 张仪萍,王士金,张士乔.沉降预测的多层递阶时间序列模型研究[J].浙江大学学报(工学版),2005,39(7):983~986.
- [7] 田其煌.基于小波技术的软土路基沉降数据分析方法研究[D].南京:河海大学,2007.

作者简介:董友弟(1969~),男,湖北黄冈人,工程师,现主要从事三维地震勘探测量工作。

(上接第2页)

过了一块沙盘控制电路板的输出控制范围,则考虑增加第二块控制电路板,所需要的工作是在网络结构上多增加一个网络节点,在控制命令上要针对两块控制电路分别编码,GIS控制模块需要多增加一个连接,每次传递控制命令的时候需要对目标进行判断,向对应的控制电路发送目标编码。当需要扩展的目标数量超过两块或更多控制电路板输出范围的时候,一方面通过增加控制电路板的输出数量来满足要求,另一方面可以通过增加控制电路板的数量来实现扩充目的,当然控制电路板的数量要在局域网容许的节点范围之内,如果超过这一范围可以通过扩充局域网来实现。

4 结束语

综合沙盘系统将实物沙盘和电子沙盘有机的结合,充分发挥了二者的优势,为人们进行空间认知提供了更加有利的工具。

本文提出了一种基于局域网的综合沙盘灯光控制方法,并利用一个实验系统给出了具体的实现方案。实验证明,该方法具有良好的可扩展性,利用该方法能够有效的保证实物沙盘和电子沙盘之间的联动控制。

【参考文献】

- [1] 任波,陈纳新,李利军.城市规划电子沙盘系统设计[J].计算机与数字工程,2006,(10):124~126.
- [2] 石旭.GIS在福州城建沙盘演示控制系统中的应用研究[J].测绘信息与工程,2006,(1):45~47.
- [3] Paul Della Maggiora, Jim Doherty. Cisco Networking Simplified[M].北京:人民邮电出版社,2004.

作者简介:张威巍(1981~),男,河南信阳人,博士研究生,主要研究方向:战场环境仿真。