

引用格式:刘帅,栾元重,马艳贺.基于插值的非等间距GM(1,1)建模方法[J].山东科技大学学报(自然科学版),2018,37(3):20-25.

LIU Shuai, LUAN Yuanzhong, MA Yanhe. Modeling methods of non-equidistant GM(1,1) model based on interpolation[J]. Journal of Shandong University of Science and Technology(Natural Science), 2018, 37(3):20-25.

基于插值的非等间距GM(1,1)建模方法

刘 帅,栾元重,马艳贺

(山东科技大学 测绘科学与工程学院, 山东 青岛 266590)

摘要:非等间距GM(1,1)模型因其对非等间距观测数据的处理能力,在变形观测的预测方面发挥着越来越重要的作用。为了更加准确地进行变形预测,提高其实用性,对非等间距GM(1,1)模型进行改进。首先分析不同插值方法对非等间距序列的适应性,选取合适的插值方式。通过插值将非等间距序列构造成等间距序列,建立插值优化背景值和不对背景值进行优化两种模型。对该等间距序列进行预测,最后再进行一次插值还原为原始非等间距序列的预测值。将预测的结果与其他方法进行比较,验证了改进的模型预测精度明显提高。

关键词:灰色模型;非等间距;插值;背景值;预测

中图分类号:TU196

文献标志码:A

文章编号:1672-3767(2018)03-0020-05

DOI: 10.16452/j.cnki.sdkjzk.2018.03.003

Modeling Methods of Non-equidistant GM(1,1) Model Based on Interpolation

LIU Shuai, LUAN Yuanzhong, MA Yanhe

(College of Geomatics, Shandong University of Science and Technology, Qingdao, Shandong 266590, China)

Abstract: Non-equidistant GM(1,1) model plays an increasingly important role in the prediction of deformation observation because of its ability to deal with non-equidistant sequence. In order to predict the deformation more accurately and improve its practicality, this paper makes an improvement on non-equidistant GM(1,1) model. First, a suitable interpolation method was chosen based on the analysis of the adaptability of different interpolation methods to non-equidistant sequence. Then an interpolation model with optimized background model and an interpolation model with non-optimized background value were established by constructing an equidistant sequence from the non-equidistant sequence by means of interpolation. Predictions of the equidistant sequence was made and the predicted values were then changed back to the predicted values of the original non-equidistant sequence by interpolation. The predicted values obtained by the proposed method was compared with other methods and the results verified the marked improvement in prediction accuracy of the new model.

Key words: gray model; non-equidistant sequence; interpolation; background value; forecast

因灰色系统具有通过“贫信息”就能对未知数据进行挖掘的能力^[1-3],其问世的30多年来,灰色预测模型(grey model, GM)已经在各个领域得到了广泛的应用^[4]。但是由于均值GM(1,1)模型仅适用于等间距序

收稿日期:2017-01-18

基金项目:地铁施工围岩与地表建筑物变形轨迹动态监测及损伤评估项目(2017GSF220010)

作者简介:刘 帅(1991—),男,山东青岛人,硕士研究生,主要从事变形监测与数据处理方面的研究. E-mail: liush612@163.com

栾元重(1963—),男,山东烟台人,教授,博士生导师,主要从事“三下”开采设计与变形监测研究,本文通信作者.

E-mail:lyz6615@163.com

列,而在实际应用中得到的数据序列往往都是不等间距的。为了解决这个问题,众多学者针对不等间距 GM(1,1)模型进行了研究^[5-13],并将其应用从单点或单一变量拓展到多点或多个变量^[14-16]。李斌等^[5]将非等间距序列根据时间间隔构造出等间距序列进行预测,再根据时间间隔还原为原始序列的预测值,验证了 GM(1,1)模型用于沉降数据分析的实用性、有效性和正确性。Shen^[6]通过梯形公式和中心差商法得到灰导数的白化值,提出一种基于最小误差平方和准则的非等间距序列建模方法,结果表明该模型的拟合与预测精度显著提高。李翠凤等^[10]通过插值优化非等间距模型的背景值,增强背景值的适应性,提高非等间距序列的预测精度。尹晖等^[14]将彼此关联的多个变形点进行整体建模,采用非等间距等距化处理的改进方法,验证了非等间距多点变形预测模型的可行性和有效性。

对于非等间距 GM(1,1)模型的改进,无论是单一序列预测或多个相关序列预测,目的都在于提高模型的预测精度。本文基于非等间距的变形观测序列,建立两个模型,采用插值将其构造成等间距序列,其中一个模型采用文献[10]的方法对新序列背景值进行优化,另一个模型采用均值 GM(1,1)模型的背景值选取方式,分别预测出新序列的预测值,最后再进行一次插值还原为原始非等间距序列的预测值。为验证该改进模型的有效性,采用文献[5]中的实例进行对比分析,结果证明本文中的变形预测改进模型的预测精度更高。

1 均值 GM(1,1)模型

1.1 均值 GM(1,1)模型的建立

设有非负等间距非负序列:

$$\mathbf{X}(0) = \{x^{(0)}(k) \mid k = 1, 2, \dots, n\} \quad (1)$$

作 $\mathbf{X}(0)$ 的 1-AGO 序列得到:

$$\mathbf{X}(1) = \{x^{(1)}(k) \mid k = 1, 2, \dots, n\}, \quad (2)$$

其中, $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), k = 1, 2, \dots, n$ 。

通过 $\mathbf{X}(1)$ 构造该模型的背景值:

$$\mathbf{Z}(1) = \{z(k) \mid k = 1, 2, \dots, n\}, \quad (3)$$

其中, $z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k-1)), k = 2, 3, \dots, n$ 。

$$x^{(0)}(k) + az^{(1)}(k) = b, \quad (4)$$

称为 GM(1,1)模型的均值形式。其白化微分方程为:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b. \quad (5)$$

在式(4)中,通过最小二乘法求解参数向量 $\hat{\mathbf{a}} = [a, b]^T$ 得:

$$\hat{\mathbf{a}} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{Y}, \quad (6)$$

式中, \mathbf{Y}, \mathbf{B} 分别为:

$$\mathbf{Y} = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}. \quad (7)$$

借助白化微分方程的解构造均值 GM(1,1)的时间响应式:

$$\hat{x}^{(1)}(k) = (x^{(0)}(1) - \frac{b}{a})e^{-a(k-1)} + \frac{b}{a}, k = 1, 2, \dots, n. \quad (8)$$

再作累减,还原预测值:

$$x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1), k = 2, 3, \dots, n. \quad (9)$$

1.2 模型检验

通过均值 GM(1,1)模型得到预测值序列:

$$\hat{\mathbf{X}}(0) = \{\hat{x}^{(0)}(k) \mid k = 1, 2, \dots, n\}。 \quad (10)$$

构成残差序列:

$$\mathbf{e} = \{e(k) \mid k = 1, 2, \dots, n\}, \quad (11)$$

其中, $e(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$ 。

残差表示预测值相对于原始数据的偏离程度, 偏离程度越小, 说明预测的精度越高, 反之预测的精度越低。根据残差序列计算每一期数据预测的相对误差, 进而得出预测的平均相对误差:

$$\Delta_k = \frac{|e(k)|}{x^{(0)}(k)}, \quad (12)$$

$$\Delta = \frac{1}{4} \sum_{k=2}^n \Delta_k. \quad (13)$$

相对误差表示每一期预测的误差大小, 相对误差越小, 精度越高, 反之越低; 平均相对误差表示该预测序列的误差大小, 平均相对误差越小, 预测精度越高, 反之越低。

2 非等间距 GM(1,1)建模改进

主要采用插值的方式对 GM(1,1)模型进行改进, 建立两种不同的模型进行对比试验。但插值方法种类繁多, 各种插值方法对特定序列的插值准确程度差异显著, 故在对某一特定序列进行灰色预测, 应先对各个插值模型进行试验对比, 选择适合该序列的方法进行预测。

2.1 方法一

均值 GM(1,1)适用于等间距序列, 但对于非等间距的序列预测精度较差。如图 1 所示, 由于间距 $k_{i+1} - k_i \neq k_{i+2} - k_{i+1}$, 故其不适用于均值 GM(1,1)模型。对图中曲线进行插值, 便可得到新的等间距序列(14), 以满足通过均值 GM(1,1)模型进行预测的要求。

$$\mathbf{x}'(0) = \{x^{(0)}(k') \mid k' = 1, 2, \dots, n\}. \quad (14)$$

通过均值 GM(1,1)模型对等间距序列(14)进行预测, 得到该序列的预测值, 进而再对该预测序列值进行插值, 还原得到原始非等间距序列的预测序列。

2.2 方法二

对于非等间距序列, 如果背景值构造得当, 亦可通过均值 GM(1,1)模型进行预测^[10]。如图 2 所示, 根据均值 GM(1,1)模型背景值选取方式, $Z^{(1)}(k' + \frac{1}{2})$ 即为模型的背景值。但从图中可以看出, 该值仅为实际背景值的近似值, 而 $Z^{(1)}(k' + \frac{1}{2})$ 为其精确值。在曲线斜率较小时, 两者差异不大, 但斜率较大时, 就必须考虑背景值的取值对预测结果造成的影响。在方法一的基础上构造背景值, 并运用方法一进行预测。

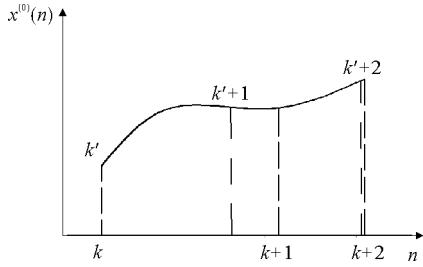


图 1 等间距插值

Fig. 1 Equidistant interpolation

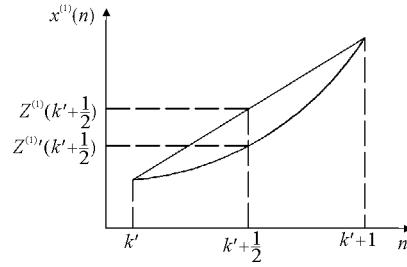


图 2 背景值插值

Fig. 2 Background value interpolation

3 实例分析

为对两个预测模型进行检验, 采用文献[5]中的观测数据进行试验对比。观测点的布设如图 3 所示。

选取南侧地表 D15 点的沉降值作为时间序列进行预测检验, 该点共有 10 期非等间距的沉降观测数据。

选取该观测数据中的前8期为已知数据进行建模,对后两期数据进行预测,分析对比预测精度。D15点的观测资料如表1所示。

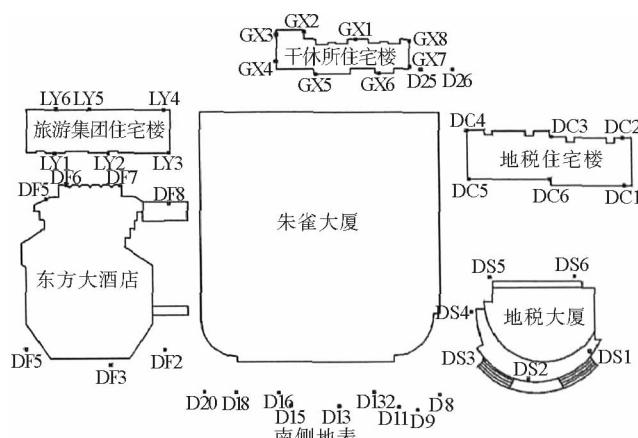


图3 地表沉降点点位布设图

Fig. 3 Location map of surface subsidence

表1 D15点观测资料

Tab. 1 D15 point observation data

序号	观测时间	时间间隔/d	累计时间间隔/d	累计沉降量/mm
1	2005-10-22	0	0	1.11
2	2005-11-06	15	15	1.45
3	2005-11-22	16	31	1.93
4	2005-12-04	12	43	2.08
5	2005-12-18	14	57	2.52
6	2006-01-01	14	71	2.92
7	2006-01-14	13	84	3.20
8	2006-02-26	43	127	3.96
9	2006-03-11	13	140	4.38
10	2006-03-25	14	154	4.68

首先进行插值方法的选取。分别用分段线性插值、邻近点插值、三次样条插值和三次多项式插值对前8期数据进行插值,插值点间隔为10,插值点区间从0到300。各个插值方法的插值曲线如图4所示。

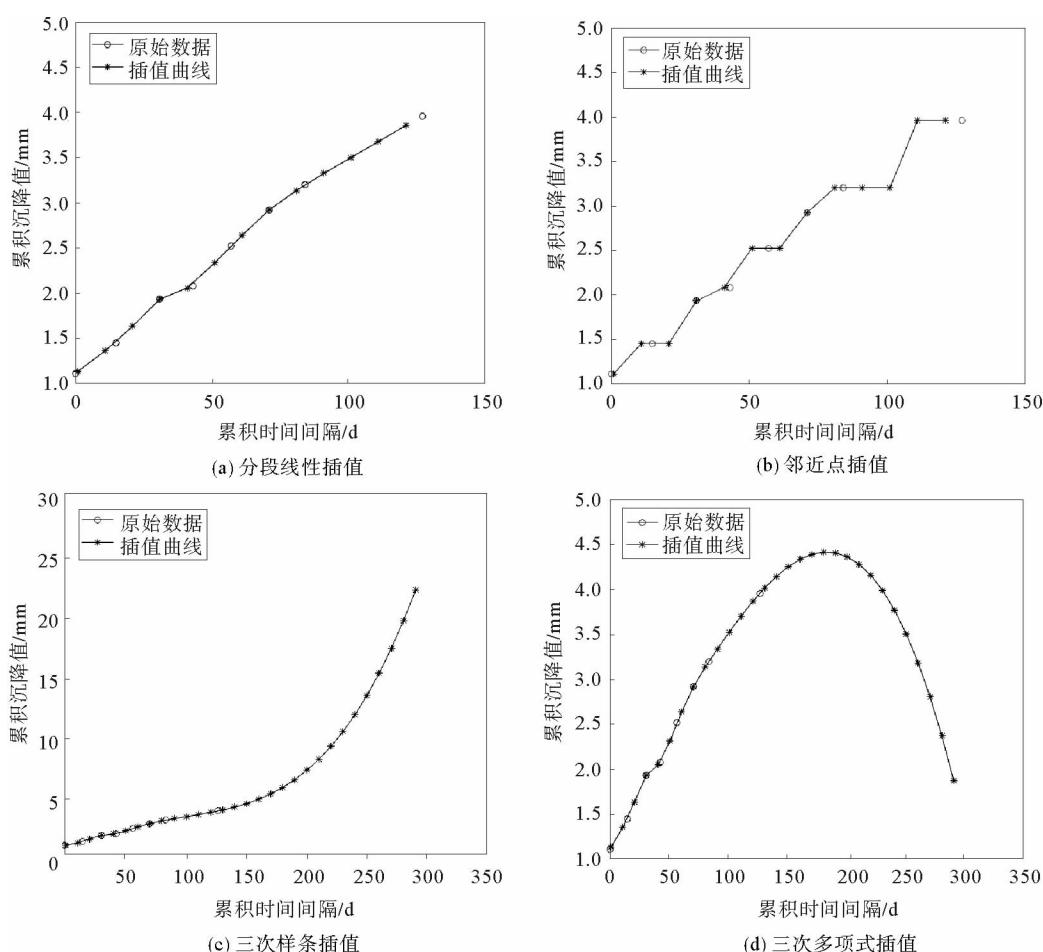


图4 插值曲线图

Fig. 4 The interpolation curve

从图中可以看出,分段线性插值并不圆滑,但对较为稠密的序列适用,在序列外的点不能得到其函数值;邻近点插值显然并不适用于该序列的插值计算;三次样条插值和三次多项式插值曲线比较圆滑,但对于沉降数据而言,累计沉降值总是在不断增大,显然三次多项式插值对于沉降观测数据而言并不适用,故选取三次样条插值方法。

采用三次样条插值法结合前文两种方法对文献[5]中的前 8 期观测数据进行拟合并预测出后两期数据,并将其与文献[5]中的预测结果进行对比,如表 2 所示。计算每种预测方法的相对误差及平均相对误差,如表 3 所示。并分别绘制预测值对比图与残差值对比图,如图 5、图 6 所示。

表 2 预测结果对比
Tab. 2 Comparison of predicted results

序号	累积时间 间隔/d	实测值 /mm	方法一		方法二		文献[5]方法	
			预测值/mm	残差/mm	预测值/mm	残差/mm	预测值/mm	残差/mm
1	0	1.11	1.11	0.00	1.11	0.00	1.301 5	-0.191 5
2	15	1.45	1.71	-0.26	1.72	-0.27	1.498 5	-0.048 5
3	31	1.93	1.98	-0.05	1.99	-0.06	1.741 8	0.188 2
4	43	2.08	2.16	-0.08	2.17	-0.09	1.949 7	0.130 3
5	57	2.52	2.43	0.09	2.44	0.08	2.224 3	0.295 7
6	71	2.92	2.71	0.21	2.72	0.20	2.536 8	0.383 2
7	84	3.20	3.02	0.18	3.03	0.17	2.866 5	0.333 5
8	127	3.96	4.26	-0.30	4.29	-0.33	4.294 3	-0.334 3
9	140	4.38	4.74	-0.36	4.76	-0.38	4.852 5	-0.472 5
10	154	4.68	5.30	-0.62	5.32	-0.64	5.535 0	-0.855 0

表 3 误差对比
Tab. 3 Comparison of errors

序号	累积时间 间隔/d	实测值 /mm	方法一		方法二		文献[5]方法	
			相对误差/%	平均相对 误差/%	相对误差/%	平均相对 误差/%	相对误差/%	平均相对 误差/%
1	0	1.11	0.00		0.00		17.25	
2	15	1.45	17.93		18.62		3.34	
3	31	1.93	2.59		3.11		9.75	
4	43	2.08	3.85		4.33		6.26	
5	57	2.52	3.57		3.17		11.73	
6	71	2.92	7.19	7.76	6.85	8.01	13.12	10.24
7	84	3.20	5.63		5.31		10.42	
8	127	3.96	7.58		8.33		8.44	
9	140	4.38	8.22		8.68		10.79	
10	154	4.68	13.25		13.68		18.27	

从表 2 和图 5 中可以看出,相比于文献[5]中的方法,运用本方法一和方法二进行预测,所得的预测值更加接近实测值,并且图 6 中的残差图以及表 3 中的平均误差与相对误差也很好地说明了这一点,证明了本文两种预测模型的有效性及实用性。

对比方法一和方法二可以看出,虽然其预测精度非常接近,但不通过插值进行背景值的优化比采用插值进行背景值优化的误差更小,说明了在该观测序列下,多次进行插值处理会在一定程度上降低预测的精度。

另外,图 6 直观地表明了无论是方法一、方法二还是文献[5]方法,再进行第二期预测时,其偏差程度明显增大。这说明了灰色模型的可预测长度非常短,进行过长的预测将与实际值产生很大的偏差。

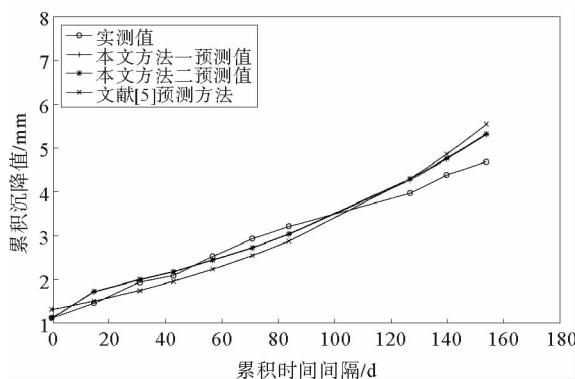


图 5 预测值对比图

Fig. 5 Comparison of predicted values

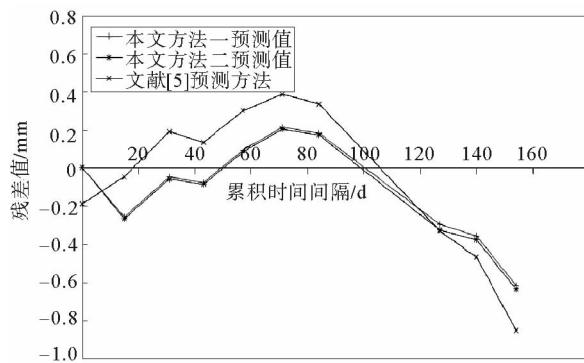


图 6 残差值对比图

Fig. 6 Comparison of residual error

4 结束语

针对非等间距 GM(1,1)建模问题,提出了两种基于插值的建模方法:方法一通过插值的方式将非等间距序列构造成等间距序列进行预测;方法二在方法一的基础上运用插值对背景值进行更新构造。因不同的插值方式对不同序列预测的准确程度不一样,提出对非等间距 GM(1,1)模型的基于插值的改进需要注重对插值方式的选取。方法一和方法二的不同在于是否采用插值法对背景值进行优化。预测结果显示在对数据进行预测时,方法一比方法二更加准确,原因是多次插值增大了误差。但是当背景值引起的误差大于插值引起的误差时,方法二比方法一更加准确。工程实例证明了两种建模方法可以得到很好的不等间距沉降预测结果,对评判建筑物沉降具有一定参考意义。

参考文献:

- [1] DENG J L. Control problems of grey systems[J]. Systems & Control Letters, 1982, 1(5): 288-294.
- [2] 刘思峰, 谢乃明. 灰色系统理论及其应用[M]. 6 版. 北京: 科学出版社, 2013: 13-14.
- [3] 党耀国, 王正新, 钱吴永, 等. 灰色预测技术方法[M]. 北京: 科学出版社, 2015: 1-2.
- [4] 谢乃明, 张可. 离散灰色预测模型及其应用[M]. 北京: 科学出版社, 2016: 10-11.
- [5] 李斌, 朱健. 非等间隔灰色 GM(1,1)模型在沉降数据分析中的应用[J]. 测绘科学, 2007, 32(4): 52-55.
- LI Bin, ZHU Jian. Application of unequal interval grey model in analysis of settlement data[J]. Science of Surveying and Mapping, 2007, 32(4): 52-55.
- [6] SHEN C G. Grey models for non-equidistant sequence based on trapezoid formula and central difference quotient[J]. Journal of Grey System, 2010, 22(1): 17-24.
- [7] 郑艳琳, 刘保东. 非等间距 GM(1,1)模型的模糊优化[J]. 山东科技大学学报(自然科学版), 2004, 23(4): 75-77.
- ZHENG Yanlin, LIU Baodong. Fuzzy optimization of non-equidistant GM(1,1) model[J]. Journal of Shandong University of Science and Technology(Natural Science), 2004, 23(4): 75-77.
- [8] 尹晖, 周晓庆, 张晓鸣. 非等间距 GM(1,1)建模方法对比分析及应用[J]. 测绘工程, 2017, 26(2): 1-4.
- YIN Hui, ZHOU Xiaoqing, ZHANG Xiaoming. Non-equidistant GM(1,1) modeling comparison and application[J]. Engineering of Surveying and Mapping, 2017, 26(2): 1-4.
- [9] 姜峰, 郑运虎, 梁瑞. 改进的非等间距 GM(1,1)模型预测腐蚀立管剩余寿命[J]. 中国安全科学学报, 2014, 24(6): 57-61.
- JIANG Feng, ZHENG Yunhu, LIANG Rui. Method for corrosion risers remaining life prediction based on improved non-equidistant GM(1,1) model[J]. China Safety Science Journal, 2014, 24(6): 57-61.
- [10] 李翠凤, 戴文战. 非等间距 GM(1,1)模型背景值构造方法及应用[J]. 清华大学学报(自然科学版), 2007, 47(增 2): 1729-1732.
- LI Cuifeng, DAI Wenzhan. Determinator of the background level in the non-equidistant GM(1,1) model[J]. Journal of Tsinghua University(Science and Technology), 2007, 47(S2): 1729-1732.