

基于高斯过程的变形监测沉降预测模型应用

王申波

广东省核工业地质局二九一大队

DOI:10.32629/gmsm.v2i4.249

[摘要] 高斯过程处理高维数、小样本等非线性复杂问题具有高自适应性,文中深入学习高斯过程(Gaussian Process)的基本原理,并将其应用到变形监测的建模中,通过优选协方差函数训练样本和检验样本得到预测结果。通过实例验证分析,与传统GM(1,1)预测模型进行比较分析,结果表明高斯过程模型预测精度较高,值得应用推广,是变形监测沉降预测的一种好的方法。

[关键词] 变形监测; GM(1,1); 高斯过程

引言

针对变形监测数据的分析及预测,研究学者做了大量的研究工作,提出建立了许多好的预测模型,如灰色理论、时间序列分析、神经网络等,建模预测精度不断提高^[1-2]。由于变形体形变的随机性和复杂性,不同方法均有一定的适用性,需要对其不断进行改进。近年来“核学习”是机器领域的研究热点,高斯过程则是新兴机器方法中较好的具有原则性和实用性的一种算法。高斯过程实质是随机变量的集合,并且集合中的任意随机变量组合均服从高斯分布,由协方差函数和均值函数来确定^[3]。高斯过程模型在数学上等同于很多众所周知的模型,比如贝叶斯线性模型、神经网络和样条模型等^[4]。相比于其他算法如BP神经网络、最小支持向量机等,其优点主要在于可以利用先验概率形式表述过程中的先验信息,进一步提高过程模型的性能。其不仅仅可以对未知输入做预测输出,同时还可以给出预测的参数精度,明显减少了模型参数,使得参数优化变得容易,更易于收敛^[5]。文中尝试将高斯过程应用于变形监测数据处理中,并且将这种方法与GM(1,1)预测模型进行比较分析,结果证明高斯过程预测模型在变形监测时间序列预测精度更高。

1 高斯过程预测模型建立

1.1 高斯过程原理

高斯过程是近几年发展的较为热门的一种机器学习技术,其主要由均值函数和协方差函数决定。考虑回归问题模型^[6-8]:

$$y = f(x) + \varepsilon \quad (1)$$

式中, y 为观测值, x 表示输入向量, 噪声服从正态分布 $\varepsilon \sim N(0, \sigma_n^2)$ 。利用贝叶斯后验概率公式导出预测值 y' 的期望和方差,

$$\mu_{y'} = C(x', X)[C(X, X) + \sigma_n^2 I_n]^{-1} y \quad (2)$$

$$\sigma_{y'}^2 = C(x', x') - C(x', X)[C(X, X) + \sigma_n^2 I_n]^{-1} C(X, x) \quad (3)$$

GP算法的协方差函数即为核函数,常用的协方差函数为平方指数协方差函数,式中 σ^{ij} , 为克洛内克尔符号, l 表示

方差尺度, $M = \text{diag}(l^{-2})$ 为超参数的对角矩阵,表示协方差函数的信号方差。令表示超参数,最优超参数利用对数函数极大似然法进行解算。负对数似然函数 $L(\theta)$ 和超参数 θ 的偏导数如下:

$$C(x_i, x_j) = \sigma_j^2 \exp[-\frac{1}{2}(x_i - x_j)^T M (x_i - x_j)] + \sigma_n^2 \sigma^{ij} \quad (4)$$

$$L(\theta) = \frac{1}{2} y^T C^{-1} y + \frac{1}{2} \log |C| + \frac{n}{2} \log 2\pi \quad (5)$$

$$\frac{\partial L(\theta)}{\partial \theta_i} = \frac{1}{2} \text{tr}[(\alpha \alpha^T - C^{-1}) \frac{\partial C}{\partial \theta_i}] \quad (6)$$

式中, $\alpha = C^{-1} y$ 。

1.2 模型建立过程

对于沉降观测预测模型的建立,其主要是为了探究变形体随着时间的推移的变形规律,在历史观测数据的基础上建立符合变形体变形规律的预测模型。高斯过程预测模型的协方差函数通常选择 CovNNone 神经网络核函数,均值函数采用 MeanZero。对于沉降观测序列 $\{X = (t_i, X_i (i=1, 2, \dots, N))\}$, 采用前M个变形值进行拟合训练建模, $\{X = (t_i, X_i (i=M+1, M+2, \dots, N))\}$ 作为检验样本。为了便于比较分析,采用平均绝对误差 MAE 和平均绝对误差百分比 MAPE 为评价指标。

其中, $\text{MAE} = (\sum |e|) / n$, $\text{MAPE} = (\sum (|e| / y)) / n$, 为残差值,为数据长度。

2 应用实例分析

实例 1: 利用文献[4]中的算例某炼钢厂钢炉沉降观测点监测的10期沉降变形观测数据, $x^{(0)} = \{2.79, 4.34, 5.17, 6.97, 7.63, 8.4, 9.13, 9.74, 10.09, 10.46\}$ 。以前七期实测值为原始数据序列,用传统的GM(1,1)模型,PGM(1,1)模型,利用高斯过程进行拟合并比较分析(高斯过程算法均值函数采用 MeanZero, 协方差函数选择精度较高的 CovNNone 神经网络

核函数),以 $\{X = (t_i, X_i(i = 1, 2, \dots, 7))\}$ 为高斯过程 GP 训练样本, $\{X = (t_i, X_i(i = 8, 9, 10))\}$ 作为检验样本。同时利用 GM(1, 1)和 GP 算法两种预测模型来计算 8~10 期的预测变形值,与相应的实测值进行比较分析。具体建模结果见表 1 和表 2。

表 1 两种模型拟合和预测值对比(单位:mm)

评价指标	GM(1,1)		高斯过程	
	拟合值	预测值	拟合值	预测值
MAE	0.33	2.47	0.12	0.23
MAPE	5.34%	24.19%	2.23%	0.75%

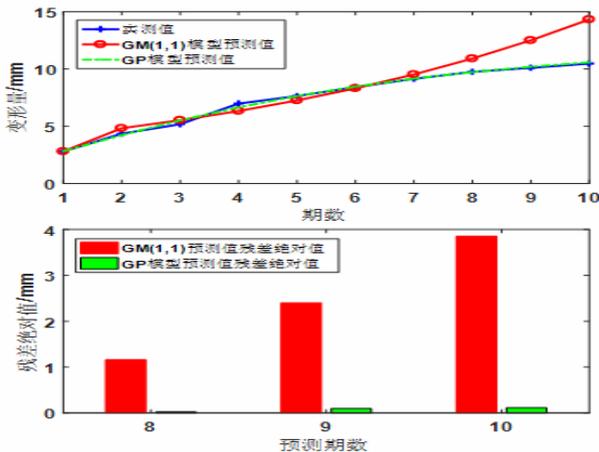


图 1 两种模型预测值及残差绝对值对比

由表1和图1给出的两种模型的拟合和预测结果发现,GM(1, 1)模型可能由于受初始条件、背景值等因素影响,其拟合预测效果不太理想。高斯过程预测模型因其处理高维数、小样本、非线性等复杂问题的高适应性特点,其拟合和预测精度得到了很大的提高,其拟合平均绝对误差为0.12 mm,而平均绝对百分比误差仅为2.23%;其预测平均绝对误差为0.23 mm,而平均绝对百分比误差仅为0.75%,平均绝对误差和平均绝对百分比误差低于GM(1, 1)模型,拟合和预测精度较高,效果较好。

实例 2: 数据来源于文献[6],原始数列 $x^{(0)} = \{0.2, 4.2, 5.0, 6.2, 9.8, 9.8, 12.6, 10.3, 15.9, 15.4, 18.1, 21.3, 20.1, 22.0, 22.6, 21.4\}$ 。利用高斯过程进行拟合并比较分析(高斯过程算法均值函数采用 MeanZero, 协方差函数选择精度较高的 CovNNone 神经网络核函数),以 $\{X = (t_i, X_i(i = 1, 2, \dots, 10))\}$ 为高斯过程 GP 训练样本, $\{X = (t_i, X_i(i = 11, 12, \dots, 16))\}$ 作为检验样本,即基于前 10 期数据利用传统的 GM(1, 1)和高斯过程模型对后 6 期数据进行预测。高斯过程预测模型的 MAE 和 MAPE 分别为 1.59mm 和 7.59%,而传统的 GM(1, 1)模型分别为 7.50mm 和 34.87%。实验结果进一步验证了高斯过程对于小样本训练预测的高适应性,其预测精度明显优于传统的预测模型,再次显示了高斯过程用于变形监测预测的高自适应性,具体预测结果见图 2。

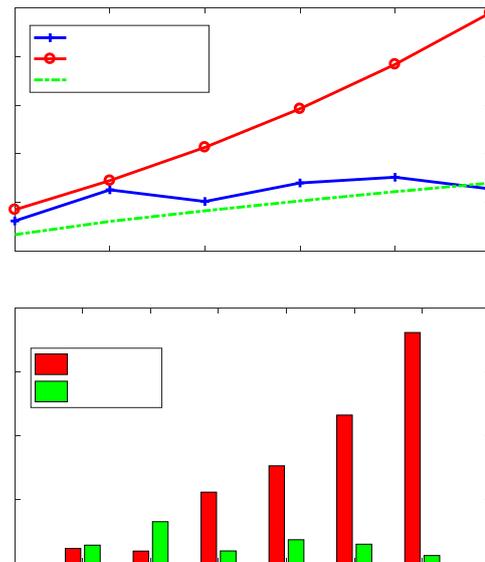


图 2 两种模型预测值及残差绝对值对比

3 结论

针对变形监测沉降数据的非线性非平稳的特点,作者深入分析高斯过程算法的原理,运用高斯过程算法和 GM(1, 1)两种模型对实际工程案例进行预测建模,探讨分析了高斯过程应用于沉降变形预测的可行性。GM(1, 1)因为受到背景值、初始条件等因素制约,建模预测效果较差。GP 算法鉴于其处理非线性非平稳信号的高适应性特点,预测和建模精度明显高于灰色预测模型,效果较好,值得应用推广。但针对不同工程案例高斯过程算法的超参数求解方法选择问题,需要作进一步探究。

【参考文献】

- [1]潘国荣.地铁隧道变形的神经网络法预测[J].大地测量与地球动力学,2007,27(1):80-84.
- [2]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.
- [3]王奉伟.改进的局部均值分解和高斯过程在变形监测数据处理中的应用研究[D].东华理工大学,2016,(11):87.
- [4]苏国韶.基于高斯过程机器学习的冲击地压危险性预测[J].辽宁工程技术大学学报:自然科学报,2009,28(5):762-765.
- [5]周昀琦,王奉伟,周世健,等.顾及邻近点变形因素的高斯过程建模及预测[J].测绘科学,2018,43(04):114-121.
- [6]周世健,赖志坤,藏德彦.加权灰色预测模型及其计算实现[J].武汉大学学报(信息科学版),2002,27(5):451-455.

作者简介:

王申波(1985--),男,山东济宁人,汉族,大学本科,测绘工程师,从事测绘基础知识研究及新技术、新理论探索。