

基于粒子群算法的 Kriging 插值在区域地面沉降中的应用

岳建平 甄宗坤

(河海大学 地球科学与工程学院 江苏 南京 210098)

Application of Particle Swarm Optimization Based Kriging Interpolation Method in Regional Land Subsidence

YUE Jianping , ZHEN Zongkun

摘要: 经典 Kriging 插值算法在当区域化变量的变化呈非正态分布时,变异函数会出现明显的病态,另外,变异函数是对区域化变量间差值取平方,致使插值结果容易受区域化变量异常值的影响。采取对区域化变量取自然对数的措施,以减弱上述因素的影响,另外在变异函数模型中,引入粒子群(PSO)优化算法对变异函数模型参数进行寻优,取得了良好的效果。

关键词: Kriging; 线性无偏; 最优估计; 变异函数; 粒子群算法

一、引言

地面沉降是在自然和人为因素共同作用下,地壳表层土体压缩而导致区域性地面标高降低的一种环境地质现象。由于它生成缓慢、持续时间长、影响范围广、成因机制复杂并且防止难度大,已对城市规划建设、经济发展、人们生产生活构成了较大的威胁。而目前城市地面沉降主要采用水准测量的方法,受技术的限制,采样点数量有限,导致难以有效地反映城市地面整体的沉降情况,因此,需要对未采样地区插值以补充沉降信息。

传统的插值方法主要有距离反比插值算法、平面法和曲面法等。其中,距离反比插值算法易受数据点集群的影响,插值结果经常出现一些呈“鸭蛋”式分布的孤立点^[1];平面模型中由于只考虑了区域化变量的趋势性,它在变量变化均匀的地区拥有较好的精度,而在变量变化较快的地区精度就比较差;曲面模型也只适合区域化变量变化不大且分布连续的区域^[2]。另外,传统的插值算法也往往只是简单地考虑了空间点的集合关系,缺乏对属性内在变化的分析。而 Kriging 插值立足于属性的空间变化,建立了描述这种属性变化的完备数学关系,并能在给出估计值的同时给出预测点的估计误差。本文在经典 Kriging 插值算法的基础上,引入简化的粒子群(PSO)优化算法,对变异函数模型中的参数进行寻优,并对区域化变量取自然对数,计算其对

应的变异函数值,以有效地减小由分布问题和异常值引起的影响,从而达到提高整体插值精度的目的。

二、Kriging 插值法原理

Kriging 插值算法是研究空间变异和进行空间插值的一种线性无偏最优估计算法。假设区域化变量不是相互独立的,具有一定的随机性和结构特性,而且满足二阶平稳和内蕴平稳。

Kriging 插值法的基本数学模型为^[3]

$$Z'(X_0) = \sum_{i=1}^n \lambda_i Z(X_i) \quad (1)$$

式中 $Z'(X_0)$ 为预测点的估计值; $Z(X_i)$ 为预测点邻域内参与预测的参考点的值; λ_i 为 Kriging 权系数,它是在无偏性和最小方差性的条件下,依赖变异函数的计算结果而确定的。

1. 无偏性条件

假定随机函数 $Z(X)$ 的期望是平稳的,要使 $Z'(X_0)$ 是 $Z(X)$ 的无偏估计量,即要求

$$E[Z'(X_0) - Z(X)] = 0 \quad (2)$$

在二阶平稳及内蕴假设下,得到无偏条件

$$\sum_{i=1}^n \lambda_i = 1 \quad (3)$$

2. 最小方差条件

在满足无偏条件的前提下,使方差达到最小,即

$$E[Z'(X_0) - Z(X_0)]^2 = \sigma_k^2 = \min \quad (4)$$

按方差最小原则求解 Kriging 权系数,这是一个

收稿日期: 2011-05-19

作者简介: 岳建平(1963—),男,江苏溧阳人,博士,教授,主要研究方向为建筑物安全监控和精密工程测量。

用拉格朗日乘数法求解目标函数的条件极值问题，为此构造函数

$$F = \sigma_k^2 + 2\mu(\sum_{i=1}^n \lambda_i - 1) \quad (5)$$

式中 μ 为拉格朗日乘数。

求 F 对 λ_i 的偏导数，令其等于零，并同无偏条件联立，得到正规方程组

$$\left. \begin{aligned} \sum_{j=1}^n \lambda_j \gamma(i, j) + \mu &= \bar{\gamma}(i, p) \\ \sum_{i=1}^n \lambda_i &= 1 \end{aligned} \right\} \quad (6)$$

将 $\gamma(i, j)$ 、 $\bar{\gamma}(i, p)$ 简写为 γ_{ij} 、 $\bar{\gamma}_{ip}$ ，并将上式展开为矩阵形式

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} & 1 \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} \bar{\gamma}_{1p} \\ \bar{\gamma}_{2p} \\ \vdots \\ \bar{\gamma}_{np} \\ 1 \end{bmatrix} \quad (7)$$

即

$$K\lambda = M \quad \text{或} \quad \lambda = K^{-1}M \quad (8)$$

式中，拉格朗日乘数正好等于已知点的离散方差

$$\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(i, j) = \mu \quad (9)$$

3. 估计方差

Kriging 插值法对于单个预测点的估计方差为

$$\sigma_k^2 = \sum_{i=1}^n \lambda_i \bar{\gamma}(i, p) - \bar{\gamma}(p) - \mu \quad (10)$$

三、粒子群算法

粒子群(PSO)算法的基本思想是：社会中的个体通过社会交互，不断修正自己关于搜索空间的知识。在社会交互的过程中，所有个体倾向于模仿成功的同伴，最终导致了社会规范的涌现^[4]。粒子群算法的数学描述如下^[5]。

为了讨论方便，设 $f(x)$ 为适应度函数，则第 i 个粒子的当前最优位置由下式确定

$$P_i(t+1) = \begin{cases} P_i(t) & f(x_i(t+1)) \geq f(P_i(t)) \\ X_i(t+1) & f(x_i(t+1)) < f(P_i(t)) \end{cases} \quad (11)$$

设搜索空间为 D 维，总粒子数为 n ，第 i 个粒子位置为向量 $x_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ ，第 i 个粒子迄今为止搜索到的最优位置为 $P_{best_i} = (P_{i1}, P_{i2}, \dots, P_{iD})$ ，整个粒子群迄今为止搜索到的最优位置为 $P_{g_{best}} = (P_{g1}, P_{g2}, \dots, P_{gD})$ ，第 i 个粒子的位置变化率(速度)为向量 $v_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ 。粒子的每维速度

和位置按如下公式进行变化

$$v_{id}(t+1) = \omega v_{id}(t) + c_1 \cdot r_2 \cdot (P_{id}(t) - x_{id}(t)) + c_2 \cdot r_2 \cdot (P_{gd}(t) - x_{id}(t)) \quad (12)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad 1 \leq i \leq n \quad 1 \leq d \leq D \quad (13)$$

式中 ω 为惯性权重； c_1 、 c_2 为正常数，称为加速因子， c_1 为调节粒子飞向自身最好位置方向的步长； c_2 为调节粒子向全局最好位置飞行的步长； r_1 、 r_2 为 $[0, 1]$ 之间的随机数。

四、基于粒子群算法的 Kriging 插值分析

为了更好地了解某地区的地面沉降情况，进一步做好灾害预警预报工作。本文选取地面沉降量作为区域化变量，根据在参考点上测定的沉降值进行插值计算，具体步骤^[6]如下：

1) 提取区域已知点的沉降数据 Δh_i 及相应的坐标 (x_i, y_i) ($i=1, 2, \dots, N$)，并作相应的规范化处理。

2) 计算已知数据点之间的距离 h_{ij}

$$h_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (i=1, 2, \dots, N; j=1, 2, \dots, N) \quad (14)$$

3) 对步骤 2) 中计算的距离进行分组^[2]，用 $\{h_m\}$ 表示

$$\{h_m\} = m \times \frac{\max(h_{ij}) - \min(h_{ij})}{N_H} \quad (m=1, 2, \dots, N_H) \quad (15)$$

式中 N_H 表示距离组的个数。在划分距离组时需保证变异函数中有意义的参数，至少要划分 3~4 组来计算变异函数 $\gamma(\{h_m\})$ ，即 $N_H \geq 4$ ，另外也要保证每个距离组包含足够多的数据，以便真实地反映空间分布特征。

4) 计算各距离组所对应的对数变异函数 $\gamma(\{h_m\})$ 的估计值^[7]

$$\gamma(\{h_m\}) = \frac{1}{2N(\{h_m\})} \sum_{i=1}^{N(\{h_m\})} \{ \ln[\Delta h_i(x_i, y_i)] - \ln[\Delta h_i(x_i + h, y_i + h)] \}^2 \quad (16)$$

式中 $N(\{h_m\})$ 表示相隔距离矢量 h 的所有已知点的个数。

5) 根据计算的 $\gamma(\{h_m\})$ ，选择合适的变异函数模型进行，以获得区域内任意距离 h 下的 $\gamma(h)$ 。常用的变异函数模型有球状模型、指数模型、高斯模型，本文经过比较选择了指数模型。指数模型的具体形式为

$$\gamma(h) = C_0 + C(1 - e^{-h/a}) \quad (17)$$

式中， C_0 为基台值； C 为块金值； $3a$ 为变程(因为 $h=3a$ 时 $\gamma(h) \approx C_0 + C$)。

6) 计算预测点邻域内参与预测的参考点之间及与预测点的变异函数值 $\gamma(h)$, 进而利用式(8)求解诸参考点到预测点的 Kriging 权系数 λ_i 。

7) 根据求得的 Kriging 权系数 λ_i 利用式(1)计算出预测点 Kriging 插值的估计值。

8) 重复步骤6)、7), 计算各预测点的地面沉降量。

五、应用实例

为了测试 PSO-Kriging 插值算法的准确性和可

靠性, 本文以某地区沉降监测的实测数据为算例进行分析。该测区共有 25 个控制点, 控制点分布均匀且地面沉降变化缓慢。本次测试中选择 16 个控制点作为参考点进行学习, 其余 9 个点作为检核点。首先按式(15)给出的选取距离组的原则, 将距离值划分为 9 个分组 $\{h_m\}$, 求出对应距离组的变异函数的估计量 $\gamma'(h_m)$ (如表 1 所示); 然后通过变异函数, 计算 Kriging 权系数, 进而对检核点的插值结果进行检验, 结果如表 2 所示。

表 1 距离组 $\{h_m\}$ 、变异函数估计值 $\gamma'(h_m)$

	组号								
	1	2	3	4	5	6	7	8	9
$\{h_m\}$	2.809	3.943	2.616	3.726	7.667	13.147	10.276	16.200	15.937
$\gamma'(h_m)$	1.271	1.906	2.542	3.177	3.813	4.448	5.084	5.719	6.355

表 2 检核点插值精度检验结果

序号	ΔH	PSO-Kriging 插值	残差 Δ	Kriging 插值	残差 Δ	二次曲面插值	残差 Δ
1	8.4	8.3	-0.1	8.3	-0.1	6.2	-2.2
2	8.3	8.9	0.6	9.8	1.5	9.6	1.3
3	10.9	11.4	0.5	11.6	0.7	11.1	0.2
4	8.2	8.4	0.2	8.3	0.1	7.5	0.7
5	11.4	10.4	-1.0	12.1	2.0	13.0	1.6
6	10.7	10.3	-0.4	10.3	-0.4	10.3	-0.4
7	9.1	9.2	0.1	9.5	0.4	8.7	-0.4
8	9.7	10.5	0.6	10.7	1.0	10.1	0.4
9	10.3	11.6	1.3	11.8	1.5	12.1	1.9
插值误差			± 0.7		± 1.1		± 1.3

由表 2 可知, 在沉降变化平缓的地区, 二次曲面插值就可以取得较好的效果, 其最大残差为 2.2 mm, 误差为 ± 1.3 mm; Kriging 插值的效果也比较理想, 其最大残差为 2.0 mm, 误差为 ± 1.1 mm。PSO-Kriging 插值的精度最高, 其最大残差为 1.3 mm, 误差为 ± 0.7 mm。从图 1 和图 2 可以直观地看出, PSO-Kriging 插值的精度明显要高, 且残差值变化比较平缓。

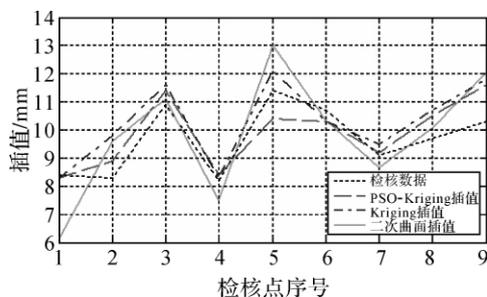


图 1 PSO-Kriging、Kriging 和二次曲面插值效果图

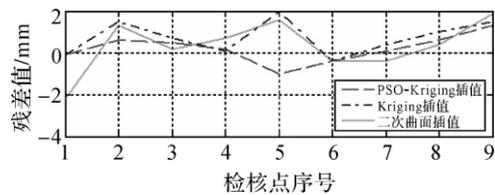


图 2 PSO-Kriging、Kriging 和二次曲面插值残差图

六、结论

本文主要研究了在局部地区应用 Kriging 插值方法分析研究区域地面沉降的问题。结合工程实例, 对比分析了二次曲面插值、Kriging 插值和 PSO-Kriging 插值的效果, 得出以下结论:

1) 在小区域单个沉降漏斗地区, 二次曲面插值就可以取得较好的精度, 相比于 Kriging 插值和 PSO-Kriging 插值, 其方法简单、计算量小, 精度也相对较高。然而, 在地面沉降变化大且分布不连续的

地区,难以达到满意的精度,从而制约了其实际应用。

2) Kriging 插值算法通过引入以距离为自变量的变异函数来计算权值,同时考虑了区域化变量的随机特性和结构特性,有效地消除了模型误差,提高了精度。改进的 PSO-Kriging 插值算法在应对地面沉降分布不均和局部沉降异常时有了较大的提高,另外在变异函数中引入粒子群(PSO)算法,提高了变异函数的精度,在一定程度上也提高了最终的插值精度。

3) 如何智能地选择合适距离分组是 Kriging 插值提高精度的一个关键问题。目前,一般通过人工对比各种距离分组下的插值结果,根据偏差均差最小确定分组情况,一方面致使工作量增加;另一方面也使得结果具有不确定性。

4) 由于 Kriging 插值立足于区域变化量的空间自相关性,并借助于变异函数描述这种相关程度,因此,变异函数能否准确地表达这种相关性将直接影响到最终的插值结果。然而,目前变异函数模型比较复杂、易失稳且计算量大,对于如何建立简单有效的变异函数模型,也需要进一步研究。

(上接第 58 页)

GPS 测量的基本原理是距离后方交会。利用测点与空中 4 个以上的 GPS 卫星(其位置是已知的)的距离求得其在空间位置,进而利用已知点坐标通过约束平差求得其在相应坐标系统中的平面坐标和大地高,其精度只与 GPS 观测精度和已知点精度有关。因为 GPS 的观测精度已经很高了,在已知点精度满足要求且分布大致合理的情况下,其平面坐标和大地高精度一般都能满足工程测量要求。而在高程测量中,因采用正常高系统,虽然大地高精度一般能达到精度要求,但影响高程测量精度的主要是高程异常。一旦已知点的分布与密度不理想或地形起伏较大,则 GPS 拟合高程就很难达到精度要求,且一般商用平差软件很难将这种误差准确表现出来。

四、结 论

1) GPS 拟合高程的误差主要来源于高程异常模型的确定。在已知点的分布和密度不合理或地形起伏较大时,会大大降低 GPS 拟合高程的精度,在 20~30 km 的范围内,其误差可能会超过 1 m。在地形平缓地区,已知点密度足够且分布合理,则 GPS

参考文献:

- [1] 彭楠峰. 距离反比插值算法与 Kriging 插值算法的比较[J]. 大众科技 2008(5): 57-58.
- [2] 李明,高星伟,文汉江,等. Kriging 方法在 GPS 水准中的应用[J]. 测绘科学 2009,39(1): 106-107.
- [3] 郭炳岐. 基于 Kriging 方法的 GPS 高程模型及其应用研究[D]. 西安: 西安科技大学 2008.
- [4] 陈金环,王冠,王东云. 粒子群算法及特性的研究[J]. 中原工业学报 2007,18(4): 13-15.
- [5] EBERHART R C, SHI Y. Particle Swarm Optimization: Developments, Applications and Resources [C]// Proc. Congress on Evolutionary Computation 2001. Piscataway, NJ: IEEE, 2001: 81-86.
- [6] 刘承香,阮双琛. 基于 Kriging 插值的数字地图生成算法研究[J]. 深圳大学学报理工版,2004,21(4): 295-299.
- [7] 黄玉琴. 基于 SAR 图像的城市形态时空变化的研究[D]. 北京: 中国科学院遥感应用研究所 2006.
- [8] 欧阳艾嘉,刘利斌,乐光学,等. 求解非线性方程组的混合粒子群算法[J]. 计算机工程与应用,2011,47(9): 33-36.

拟合高程还是可用的,反之,则要相当慎重。

2) 目前我国中东部地区大都已完成了地区大地水准面精化工作,分辨率一般达到 $2.5' \times 2.5'$,精度达到 ± 5 cm 以内,可以满足大部分的工程测量工作。在条件许可的情况下,应尽量收集、利用这方面的资料。

3) GPS 平面测量中对网形要求不甚严,相比于三角网和导线网网形,其要求要低得多,在已知点精度满足要求且分布大致合理的情况下,其平面坐标一般都能满足工程测量要求。

参考文献:

- [1] 宁津生,陈俊勇,李德仁,等. 测绘学概论[M]. 武汉: 武汉大学出版社 2008.
- [2] 张高兴,陈建水,郭达志. 山区城市 GPS 控制测量及其高程精度分析[J]. 测绘通报 2006(11): 29-31.
- [3] 国家测绘局. GB/T 18314—2009 全球定位系统(GPS) 测量规范[S]. 北京: 中国标准出版社 2009.
- [4] KAPLAN E D, HEGARTY C J. GPS 原理与应用[M]. 寇艳红,译. 北京: 电子工业出版社 2007.
- [5] 孔祥元,郭际明,刘宗泉. 大地测量学基础[M]. 武汉: 武汉大学出版社 2005.