

邬春明,齐森南,王迪. 基于柯西变异动态粒子群优化的克里金插值算法[J]. 湖南科技大学学报(自然科学版), 2021, 36(3):97-103. doi:10.13582/j.cnki.1672-9102.2021.03.014

WU C M, QI S N, WANG D. Kriging Interpolation Algorithm Based on Cauchy Mutation Dynamic Particle Swarm Optimization [J]. Journal of Hunan University of Science and Technology (Natural Science Edition), 2021, 36(3):97-103. doi:10.13582/j.cnki.1672-9102.2021.03.014

基于柯西变异动态粒子群优化的 克里金插值算法

邬春明¹, 齐森南^{1*}, 王迪²

(1. 东北电力大学 现代电力系统仿真控制与绿色电能新技术教育部重点实验室, 吉林 吉林 132012;
2. 国网吉林省电力有限公司 四平供电公司, 吉林 四平 136000)

摘要:针对克里金插值算法中变差函数拟合曲线误差过大、插值精度低等问题,通过基于线性动态变化因子结合柯西变异粒子群算法对变差函数的拟合模型参数进行最优化估计,同时在适应度函数中引入克里金地理权重来增强变量的空间相关性,最后与基于约束粒子群算法的克里金插值进行比较实验.仿真实验结果表明:改进算法使基台值误差减少近 75%,获得的变差函数值拟合曲线更接近实际情况,插值精度更高.

关键词:克里金插值;变差函数;数据拟合;柯西变异;粒子群算法

中图分类号:TP301.6; TE319 文献标志码:A 文章编号:1672-9102(2021)03-0097-07

Kriging Interpolation Algorithm Based on Cauchy Mutation Dynamic Particle Swarm Optimization

WU Chunming¹, QI Sennan¹, WANG Di²

(1. Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology,
Ministry of Education, Northeast Electric Power University, Jilin 132012, China;

2. Siping Power Supply Company, State Grid Jilin Electric Power Supply Company, Siping 136000, China)

Abstract: In the Kriging interpolation algorithm, the curve fitting error of variogram was too large and the interpolation accuracy was low. The parameters of the fitting model of the variogram were estimated optimally based on the linear dynamic change factor and the Cauchy mutation particle swarm optimization algorithm. At the same time, Kriging geographical weight was introduced into the fitness function to enhance the spatial correlation of the variables. Finally, a comparative experiment was carried out with Kriging interpolation based on the constrained particle swarm optimization algorithm. The simulation results show that the improved algorithm reduce the base station error by nearly 75%. The fitting curve of the variogram value obtained is closer to the actual situation and the interpolation accuracy is higher.

Keywords: Kriging interpolation; variation function; data fitting; cauchy variation; particle swarm optimization

收稿日期:2019-12-30

基金项目:国家自然科学基金资助项目(61901102)

*通信作者, E-mail: neepu_qsn0521@163.com

在变差函数基础上发展起来的克里金插值法(Kriging),可以对分布于空间内的离散或连续数据进行线性最优化处理、无偏内插求解.变异函数的建模是大多数地质统计学方法的关键步骤^[1-2],同时也是地质统计学的核心组成部分^[3-4].该算法在求解过程中,变差函数的拟合方法及参数选取都对插值精度和效果有直接的影响^[5-6].然而,大多数流行的基于变差函数的解的设计都没有充分考虑插值过程对其应用的影响.为改善插值效果,国内外学者们提出了多种改进方法.在确定变差函数参数过程中,采用最小二乘法^[7]、群搜索优化算法^[8]、遗传算法^[9-10]、粒子群算法^[11]等方法,上述方法虽解决传统人工拟合方法中存在的人为误差问题,但这些算法均有不同的缺点,例如,颜华等采用的最小二乘法,在确定拟合参数时会出现正负号问题;陈华等采用的群搜索算法,其追随者模式较单一,会使其陷入局部最优;岳建平学者将变差函数模型框架看作一个具有2个子目标的非线性优化问题,在给定参数的条件下,优化实验值与理论值的拟合优度.虽然该方法提高了数据拟合精度,但是因设定参数过多导致算法执行效率下降.

粒子群算法(Particle Swarm Optimization, PSO)是群智能算法(Swarm Intelligence, SI)中的一种,可以自适应地进行有效局部区域的搜索,模型操作简单容易实现,调节参数不多.现今已被广泛应用于神经网络训练^[12-13]、模式识别等多个领域中^[14-15].基于此,本文提出一种改进的克里金插值方法,采用柯西约束^[16]与克里金地理权重相结合的粒子群算法优化变差函数拟合参数,使插值的变差函数模型估计更合理,提高变差函数的拟合精度,使最终插值的结果更精确.

1 克里金插值法概述

克里金插值是基于变差函数及其模型建立起来的插值法,在考虑插值点位置与已知数据点位置的相关性基础上,还考虑了变量的空间相关性.在计算时,首先要知道数据样本点的总数 n ;周围所有样本点到待插值点的空间距离 h ,即滞后距;每个样本点的空间属性 $Z(x)$.设空间区域上的一些观测点 $x_1, x_2, \dots, x_n, Z(x_1), \dots, Z(x_n)$ 为其相应观测值,则区域化变量在未知待插值点 x_0 处的观测值 $Z^*(x_0)$ 由式(1)得出.

$$Z^*(x_0) = \sum_{i=1}^n \lambda_i Z(x_i). \quad (1)$$

式中: λ_i 为克里金权重系数.

条件无偏性、估计方差最小及拉格朗日乘数因子求得由变差函数组成的克里金方程组,其变差函数定义如式(2),记为 $\gamma(h)$ ^[17].

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^n [Z(x_i) - Z(x_i + h)]^2. \quad (2)$$

式中: $Z(x_i)$ 和 $Z(x_i + h)$ 为区域化变量 $Z(x)$ 在距离 x_i 和 $x_i + h$ 的真实值.为求得在任意距离下的变差函数,以便后续数据插值或预测,需要采用基本模型对 $\gamma(h)$ 进行拟合.

2 标准粒子群算法及其改进

2.1 标准粒子群算法原理

在标准粒子群算法中,初始化时是一些随机粒子,用2个矩阵 $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})^T$, $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})^T$ 分别表示粒子 i 的位置向量和速度向量,其中, D 是空间维数,进行多次迭代,得到最优解,在迭代中,粒子通过跟踪“ $P_{\text{best}}, G_{\text{best}}$ ”进行自我更新^[18]. P_{best} 为粒子 i 历史最优位置, G_{best} 为整个群体最优位置.找到最优值后,通过式(3)和式(4)更新.

$$V_{iD}^{k+1} = w(t) V_{iD}^k + c_1 \text{rand}(P_{\text{best}iD}^k - x_{iD}^k) + c_2 \text{rand}(G_{\text{best}iD}^k - x_{iD}^k); \quad (3)$$

$$x_{iD}^{k+1} = x_{iD}^k + V_{iD}^{k+1}. \quad (4)$$

式中: $i = 1, 2, \dots, M, M$ 为该粒子群中粒子数目总和; V_i 为粒子的速度;rand为介于(0,1)之间的随机数; x_i 为粒子的当前位置; c_1, c_2 为学习因子,取值为2; $w(t)$ 为惯性权重因子, $w(t)$ 由大到小线性递减,引入使

粒子群算法可以同时调整全局和局部搜索能力,性能有很大提高^[19].

2.2 改进的粒子群算法

2.2.1 柯西变异

标准 PSO 算法存在局部最优和后期收敛速度过慢的缺点,应用到克里金插值中,会使变差函数拟合曲线误差过大,最终导致插值结果不理想^[20].经研究表明,其缺点的根本原因之一是粒子多样性的逐渐降低,鉴于以上原因,本文提出基于柯西变异的动态粒子群算法(Dynamic Particle Swarm Optimization, DPSO)对以上问题进行改进.

在迭代时,每经5次迭代,则按照式(5)进行随机选取粒子,然后按式(6)进行柯西变异.式(5)使变异的粒子数量呈线性逐渐增加,让粒子群算法在全局上可以保持粒子的多样性.在陷于局部最优时,较大的粒子多样性能使粒子跳出,当粒子接近收敛并搜索最优解时,较小的粒子多样性则能够加速粒子的收敛;柯西变异其密度函数具有较长的尾巴,使个体逃离局部最优的概率较高,同时其函数的中心点处有较小的峰值,这表明柯西变异将花费更少的时间来搜索邻域空间,具有更好的调节能力^[16].

$$g(t) = [(g_{\max} - g_{\min})/T_{\max}]t - g_{\max}. \quad (5)$$

式中: T_{\max} 为最大迭代数; t 为当前迭代数; $g(t)$ 为维持粒子多样性的线性递增权重因子, g_{\max} 和 g_{\min} 为最大或最小取值,取 $g_{\max} = 0.9, g_{\min} = 0.4$.

$$\mathbf{x}_{id}^{k+1} = \mathbf{x}_{id}^k + \eta C(0,1). \quad (6)$$

式中: \mathbf{x}_{id}^{k+1} 为柯西变异后的位置向量; \mathbf{x}_{id}^k 为柯西变异前的位置向量; η 为调节柯西变步步长的常数,本文设为 0.1; $C(0,1)$ 为由 $t = 1$ 的柯西分布函数产生的随机数.

将变异前后的适应度值进行相互比较,若变异后的适应度值比变异前的更优,则进行替换.柯西密度函数较长的尾巴使算法可以避免局部最优,但也会使经过变异后得出的位置向量超出定义区间(位置 \mathbf{x}_{id}^{k+1} 大于 \mathbf{X}_D^{\max} 或小于 \mathbf{X}_D^{\min}),此时,按照式(7)进行适当处理.

$$\mathbf{x}_{id}^k = \mathbf{X}_D^{\min} + \text{rand} \times (\mathbf{X}_D^{\max} - \mathbf{X}_D^{\min}). \quad (7)$$

式中: \mathbf{X}_D^{\max} 与 \mathbf{X}_D^{\min} 为粒子位置向量的上界和下界;rand 为介于(0,1)的随机数.

2.2.2 克里金地理权重

在大多数情况下,选取的适应度函数由式(8)给出.

$$F(i) = \sum_{j=1}^n [\gamma(h_{i,j}) - \gamma^*(h_{i,j})]^2. \quad (8)$$

在克里金插值中,为加强适应度函数与空间相关性和地质特征之间的关联程度,进而优化后期插值效果,引入克里金地理权重 δ_i , 对适应度函数进行改进.改进后适应度函数如式(9),地理权重 δ_i 定义如式(10).

$$F(i) = \delta_i \sum_{j=1}^n [\gamma(h_{i,j}) - \gamma^*(h_{i,j})]^2; \quad (9)$$

$$\delta_i = \frac{h_i}{\bar{h}} \left(1 - \frac{n_i}{n}\right) \frac{\overline{\gamma^*(h)}}{\gamma^*(h_i)}. \quad (10)$$

式中: $F(i)$ 为粒子 i 的适应度值函数; $\gamma(h_{i,j})$ 与 $\gamma^*(h_{i,j})$ 为粒子 i 在第 j 个滞后距下的变差函数值与理论变差函数值; δ_i 为克里金地理权重; h_i 为粒子 i 的滞后距; \bar{h} 为滞后距平均值; n_i 为粒子 i 的观测样本点个数; n 为观测样本点总个数; $\overline{\gamma^*(h)}$ 为理论变差函数平均值.

3 克里金插值优化算法

本文选取指数模型的表达式为

$$\gamma(h) = \lambda_i (C_c + C_0) \left[1 - \exp\left(-\frac{3h}{a}\right)\right]. \quad (11)$$

式中: a 为变程,在变程范围 a 内,数据散点具有相关性,在变程范围外,则不相关; C_c 为拱高,在有效数据内,可对变异性幅度大小进行观测; C_0 为块金值,在 C_0 范围内,空间变异性很大,且即使 h 很小,与 $\gamma(h)$ 都不会相关; C_0, C_c 之和为基台值,用来观测变量滞后距 h 在空间范围上总变异性发展情况.

结合改进的粒子群算法进行拟合参数最优值估计,理论变差函数中共有 3 个未知参数 C_c, C_0, a , 所以将这 3 个参数打包为一个粒子—— (C_c, C_0, a) , 以下为基于改进的粒子群算法拟合过程:

Step 1:初始化及设定参数.群体规模 $size = 30$; 维数 $Dimension = 20$; 总迭代次数 $T_{max} = 800$; 未知参数取值范围: $C_c = [0, 10]$, $C_0 = [0, 5]$, $a = [5, 10]$; 初始化位置和速度.

Step 2:计算克里金地理权重 δ_i . 根据粒子滞后距、观测样本点个数和式(10)计算.

Step 3:计算适应度 $F(i)$. 按照改进后的适应度函数式(9)求出种群中各粒子的 $F(i)$.

Step 4:更新位置、速度、适应度值.位置和速度分别按照式(3)和式(4)计算更新,计算 $w(t)$. 将 Step 3 中计算出的各个粒子的 $F(i)$ 与经过的最佳位置 P_{best} 的 $F(i)$ 两者相互比较取最优值更新 P_{best} ; 将 Step 3 中计算出的各个粒子的适应度值与全局经过的最佳位置 G_{best} 进行比较取得最优值来更新 G_{best} .

Step 5:若速度、位置不在范围内,则在搜索区域内为该粒子重新取值.

Step 6:柯西变异.若当前迭代次数 t 是 5 的倍数,即 $\text{if } \text{mod}(t, 5) == 0$, 则随机选取粒子总数的 $g(t)$ 倍进行柯西变异,并按 Step 4 计算更新其适应度值,使其跳出局部最优.

Step 7:结束条件.拟合迭代过程满足结束条件(到达最大迭代次数),则输出最优解,如果未满足则转向 Step 2,算法继续,直到满足结束条件.

4 试验与分析

4.1 拟合曲线比较

为验证变差函数拟合曲线精度,本实验中涉及的变差函数 $\gamma(h)$ 、粒子 i 的滞后距 h 、粒子 i 的观测样本点个数 n_i 等均来自文献[20]中某已知矿钻孔中 V_2O_5 品位的变差函数值.由式(10)计算得到理论克里金地理权重 δ_i . 具体参数及实验结果见表 1,其中权重值保留到小数点后 3 位.

表 1 变差函数拟合值及克里金地理权重参数

h_i	n_i	$\gamma(h_i)$	δ_i	h_i	n_i	$\gamma(h_i)$	δ_i
0.5	19	0.73	0.316	4.5	14	2.25	0.947
1.0	18	1.16	0.400	5.0	14	2.30	1.029
1.5	18	1.44	0.484	5.5	13	2.44	1.072
2.0	17	1.65	0.565	6.0	13	2.53	1.128
2.5	17	1.79	0.652	6.5	12	2.51	1.237
3.0	16	2.00	0.703	7.0	12	2.59	1.291
3.5	15	2.07	0.796	7.5	12	2.64	1.357
4.0	15	2.18	0.864	4.5	14	2.25	0.947

分别使用最小二乘法(LSM)、基于约束粒子群算法(Constraint Particle Swarm Optimization, CPSO)^[3]和动态粒子群算法对 $\gamma(h)$ 拟合参数进行最优估计计算.得到 a, C_c, C_0 理论值;得到基台值如表 2.

表 2 3 种不同算法结果

算法	a	C_c	C_0	$C_c + C_0$	$ \Delta(C_c + C_0) $
LSM	10.17	2.27	0.69	2.96	0.32
CPSO	9.19	2.11	0.64	2.76	0.12
DPSO	8.32	2.24	0.43	2.67	0.03

V_2O_5 品位变差函数值的散点分布和基于 3 种算法对理论变差函数拟合曲线如图 1 所示.

图 1 中,五角星代表 V_2O_5 品位变差函数的散点;方形虚线是使用 LSM 算法得出的理论变差函数曲线;

菱形虚线是用 CPSO 算法得出的理论变差函数曲线;三角形虚线是用 DPSO 算法得到的理论变差函数曲线.

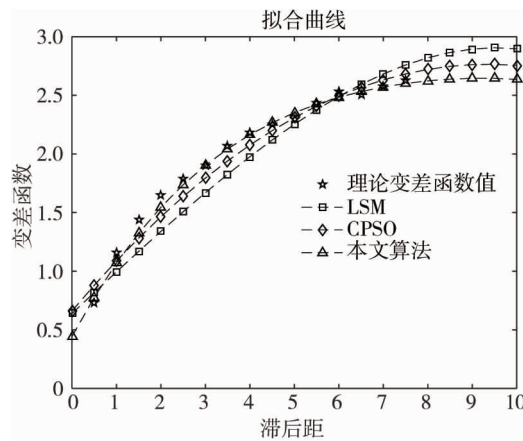


图 1 算法拟合曲线比较

观察图 1 可知:采用 DPSO 算法拟合的曲线在整体上与实际变差函数值相近,尤其是在前 9 个散点处;DPSO 算法与 CPSO 算法相比,更加接近实际值;在滞后距较大的地方,变差函数的值则更加接近实际基台值.结合表 1、表 2 中的数据也可得知:相较于 CPSO 算法,经过 DPSO 算法计算出的基台值更接近实际 V_2O_5 品位变差函数稳定后的基台值,拟合更加精准,误差更小,其中基台值误差减少 75%.

4.2 插值效果比较

在 100×100 和 200×200 这 2 种区域范围内分别对 200 个样本点和 400 个样本点进行插值计算,并对结果进行对比.使用在 4.1 节中得到的 2 种拟合参数,对变差函数进行拟合,得到的 2 种理论变差函数对待插值点做插值处理,图 2 和图 3 所示为 CPSO 算法和 DPSO 算法的插值效果对比图,其中图 2a、图 2b、图 3a、图 3b 为 200 个样本点数实验结果,图 2c、图 2d、图 3c、图 3d 为 400 个样本点数实验结果.

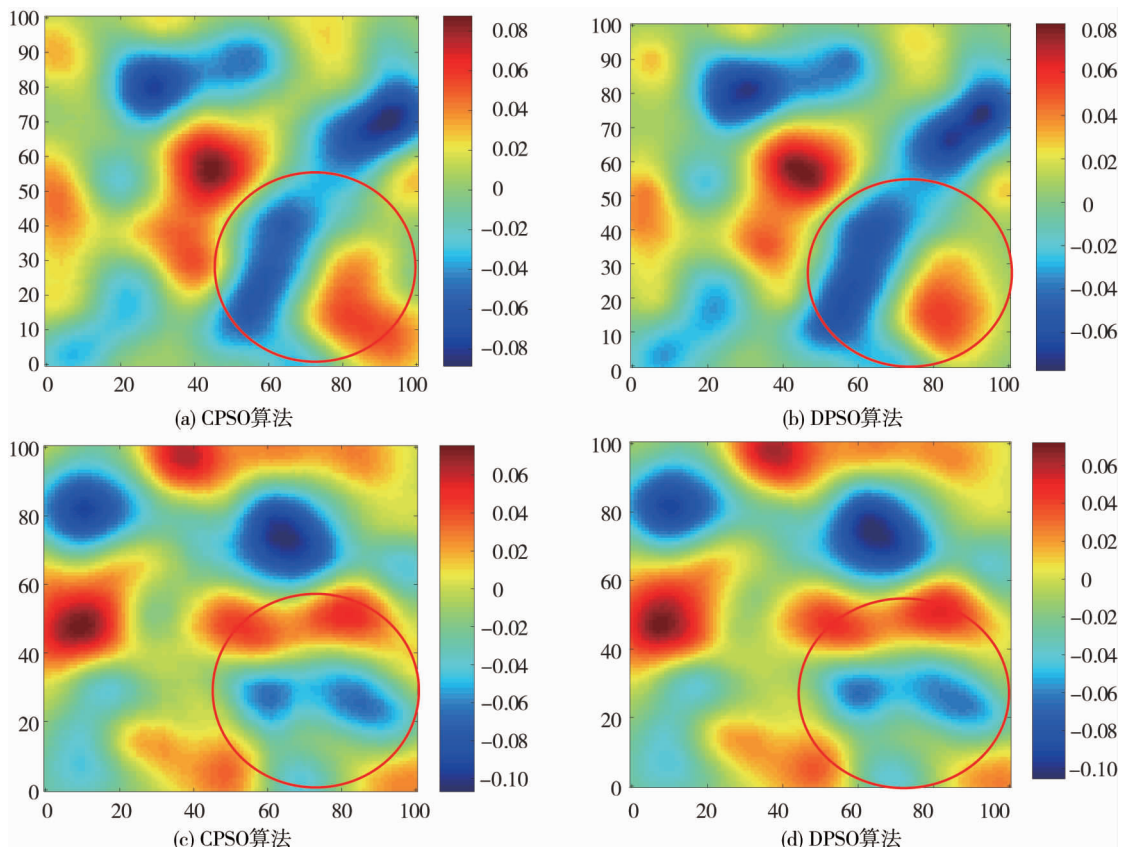


图 2 不同优化算法的插值效果(范围 100×100)

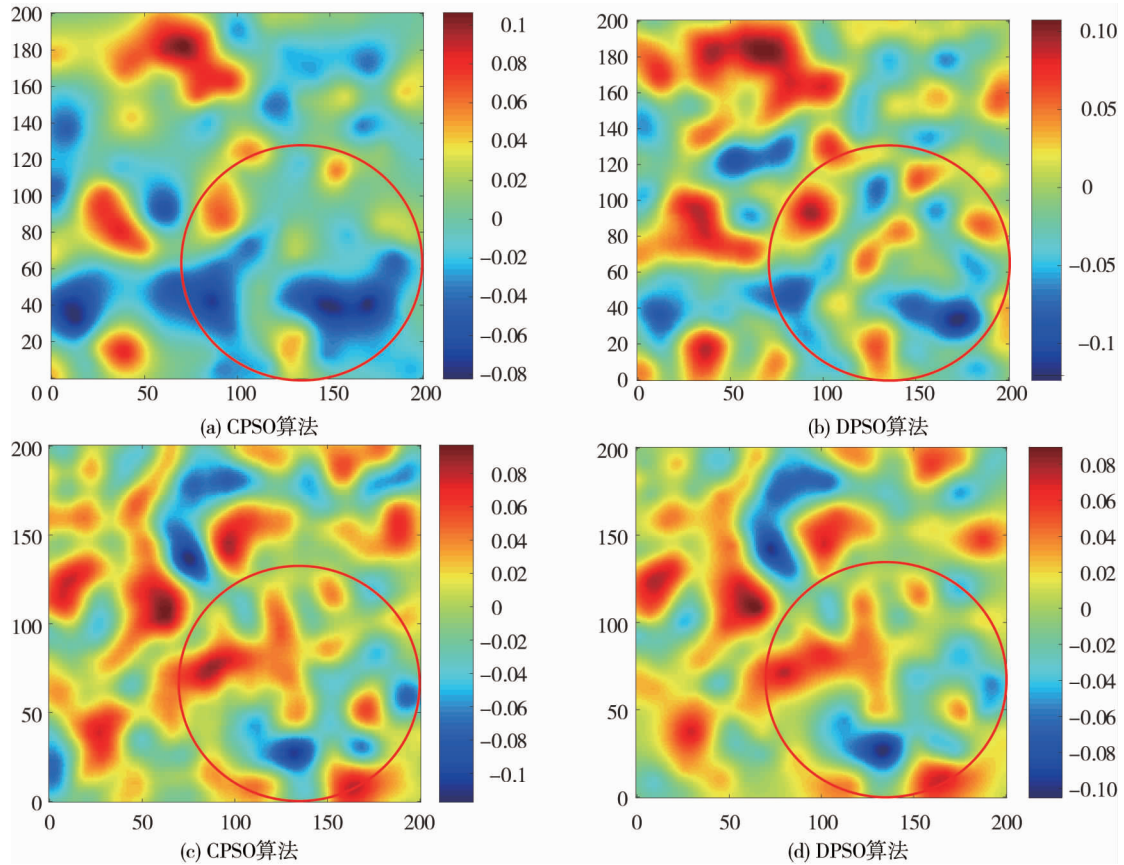
图3 不同优化算法的插值效果(范围 200×200)

图2和图3中红色部分代表海拔高位点,颜色越偏向蓝色说明空间海拔越低.对比图2和图3可以看出,在空间细节变化的特征上,图2b、图2d和图3b、图3d的插值结果要比图2a、图2c和图3a、图3c的更加明显,高点处更为清晰(图中红圆圈处),且范围越大越明显;与此同时图2a、图2c和图3a、图3c插值空间变化较为平缓,对空间细节的变化有一定程度上的忽视.并且,在相同范围大小时,样本点数越多,插值后的空间细节越明显;在样本点数相同时,范围越大,需要的细节则越多.CPSO算法和DPSO算法在实验中对不同数目样本点的插值时间对比如表3所示.

表3 CPSO和DPSO算法插值时间

区域范围	100×100		200×200	
样本点数	200	400	200	400
CPSO 插值时间/s	1.686	2.212	1.986	2.757
DPSO 插值时间/s	1.463	2.073	1.758	2.356

由表3中的数据可得知:在相同区域范围内2种理论变差函数对样本点做插值处理时,相比较于CPSO算法,改进的DPSO算法优势更加明显,其能在更少的时间里达到更好的插值效果,当插值样本点数增加时这一优势依然存在.

5 结论

1) 结合考虑粒子群算法的优势及克里金地理因素的特点,提出一种改进的粒子群算法,对变差函数中的参数选择问题进行改进.

2) 应用该算法于变差函数,得出最佳参数,得到最优拟合模型,将其应用到克里金插值中,并与CPSO算法的结果进行实验比较.仿真实验结果表明:所提方法可以获得更好的拟合曲线,且基台值误差减少75%左右,最终的插值效果有所提高.

参考文献:

- [1] 聂志红, 阚常壮, 谢扬. 连续压实质量检测参数单点异常值识别及处理[J]. 哈尔滨工业大学学报, 2019, 51(3): 150-157.
- [2] 高文武, 姜燕, 赵晋陵. 基于协同克里金插值法的土壤锰元素含量预测[J]. 地理与地理信息科学, 2018, 34(3): 119-124.
- [3] 贾雨, 邓世武, 姚兴苗, 等. 基于约束粒子群优化的克里金插值算法[J]. 成都理工大学学报(自然科学版), 2015, 42(1): 104-109.
- [4] Hundelshausen R, Costa J F C L, Marques D M, et al. Localised kriging parameter optimisation based on absolute error minimisation[J]. Applied Earth Science, 2018, 127(4): 153-162.
- [5] Golmohammadi A, Khaninezhad M R M, Jafarpour B. Pattern-based calibration of complex subsurface flow models against dynamic response data[J]. Advances in Water Resources, 2018, 120: 48-59.
- [6] Tiwari K, Goyal R, Sarkar A. GIS-based spatial distribution of groundwater quality and regional suitability evaluation for drinking water[J]. Environmental Processes, 2017, 4(3): 645-662.
- [7] 颜华, 李欣, 王善辉. 基于最小二乘法和克里金插值的三维温度场重建[J]. 沈阳工业大学学报, 2014, 36(3): 303-307.
- [8] 陈华, 张艺丹, 葛新民. 基于改进群搜索优化算法的变差函数拟合[J]. 物探化探计算技术, 2016, 38(4): 566-570.
- [9] Li Z L, Zhang X L, Clarke K C, et al. An automatic variogram modeling method with high reliability fitness and estimates[J]. Computers and Geosciences, 2018, 120(2018): 48-59.
- [10] Ma F, Han L, Zhou Y, et al. Multi-island genetic algorithm and Kriging model-based design of vehicle product comprising multi-material[J]. IEEE Access, 2018, 6: 53397-53408.
- [11] Xiang T, Wang H. Research on distributed 5G signal coverage detection algorithm based on PSO-BP-Kriging[J]. Sensors, 2018, 18(12): 4390-4408.
- [12] Tang K, Huang C, Lin C. Automatic text recognition in natural scene using neural network classifier with dynamic-group-based hybrid particle swarm optimization[J]. Journal of Information Science and Engineering, 2019, 35(3): 555-575.
- [13] Dang T L, Hoshino Y. Hardware/Software Co-design for a neural network trained by particle swarm optimization algorithm[J]. Neural Processing Letters, 2018, 49(5): 1-25.
- [14] Logesh R, Subramaniaswamy V, Vijayakumar V, et al. A hybrid quantum-induced swarm intelligence clustering for the urban trip recommendation in smart city[J]. Future Generation Computer Systems, 2018, 83: 653-673.
- [15] Huang K W, Wu Z X, Peng H W, et al. Memetic particle gravitation optimization algorithm for solving clustering problems[J]. IEEE Access, 2019, 7: 80950-80968.
- [16] Peng G, Fang Y W, Chai D, et al. Multi-objective particle swarm optimization algorithm based on sharing-learning and cauchy mutation[C] // IEEE Control Conference, Chengdu, China, 2016, 3: 9155-9160.
- [17] 刘金辉, 王京彬, 李峰, 等. 矿产资源估算中定量邻域分析与优化研究[J]. 中国矿业, 2016, 25(3): 154-157.
- [18] 王宏志, 姜方达, 周明月. 基于遗传粒子群优化算法的认知无线电系统功率分配[J]. 吉林大学学报(工学版), 2019, 49(4): 1363-1368.
- [19] Xu G, Yu G S. Reprint of: On convergence analysis of particle swarm optimization algorithm[J]. Journal of Computational and Applied Mathematics, 2018, 340: 709-717.
- [20] 梁昔明, 肖晓芳. 基于 PSO 算法的变差函数球状模型参数拟合[J]. 计算机工程, 2011, 37(14): 155-157.