

# 基于萤火虫算法的岩体结构面优势产状聚类分析

宋腾蛟, 陈剑平, 张 文, 宋盛渊

(吉林大学 建设工程学院, 吉林 长春 130026)

**摘 要:** 传统的图形分析方法较为粗糙,且主观因素影响较大,因此提出一种基于萤火虫算法的岩体结构面产状数据划分的新方法.用结构面单位法向量表示结构面的产状,通过单位法向量所夹锐角正弦值来衡量结构面间的距离,并建立模糊目标函数,然后运用萤火虫优化算法搜索最优聚类中心,同时确定组间边界.由于萤火虫算法具有出色的全局寻优能力和较快的收敛速度,使得新算法在解决结构面产状划分问题时非常省时高效.最后应用此算法对怒江松塔坝址区平洞实测结构面产状数据进行优势组划分,结果显示,此方法合理有效,优势明显.

**关 键 词:** 岩石力学;结构面;产状数据划分;聚类方法;萤火虫算法

**中图分类号:** TU 452 **文献标志码:** A **文章编号:** 1005-3026(2015)02-0284-05

## Clustering Analysis of Dominative Attitudes of Rock Mass Structural Plane Based on Firefly Algorithm

SONG Teng-jiao, CHEN Jian-ping, ZHANG Wen, SONG Sheng-yuan

(College of Construction Engineering, Jilin University, Changchun 130026, China. Corresponding author: CHEN Jian-ping, E-mail: chenjpwwq@126.com)

**Abstract:** Traditional graphic analysis is inadequate and greatly influenced by subjective experience. A new method based on firefly algorithm is proposed to partition the discontinuity orientation data. The sine of the acute angle between the unit normal vector of two discontinuities is used to quantify the distance between the orientation of these discontinuities. Then, the fuzzy objective function is established and firefly algorithm is chosen to search the optimal clustering centers. At the same time, the boundaries between different sets are determined. Since the firefly algorithm has the outstanding ability of global optimization and a fast convergence speed, the new algorithm has the advantages of time saving and high efficiency. Finally, this new method is applied in discontinuity orientation data partition at Songta dam site, Nujiang River. The results indicate that the new approach is reasonable and has obvious superiority.

**Key words:** rock mechanics; structural plane; attitude data partition; clustering method; firefly algorithm

岩体中广泛分布的不连续结构面使岩体具有强烈的不连续性、非均质性和各向异性.结构面对岩体力学性质的影响主要取决于结构面的发育状况.如岩性完全相同的两种岩体,由于结构面的空间方位、密度、规模、宽度、表面起伏形态及其组合关系不同,它们的力学性质会有很大的差异.因此,对岩体中的结构面各项参数进行实地测量和统计分析对于研究岩体的变形规律及破坏模式是

至关重要的.

目前,对岩体结构面空间方向组合规律的研究主要是对结构面按照其产状进行归类,即岩体结构面产状优势分组.传统方法是利用玫瑰花图或极点密度等值线图进行目测分析,结合分析者的工程经验人为确定分组结果.显然,这种方法带有很强的主观性,缺少严格的理论依据,因此,有必要采用客观有效的数学方法对结构面产状进行

优势分组. Shanley 等<sup>[1-2]</sup>首次运用聚类算法分析结构面方向数据,该方法的结果受小球半径取值的影响很大,而如何确定合理的小球半径的问题一直没有得到有效解决. Hammah 等<sup>[3]</sup>提出利用 fuzzy C-means(FCM)的结构面组数划分方法,但此方法是一种局部搜索寻优法,且计算结果受初始聚类中心的影响较大. 因此,蔡美峰等<sup>[4]</sup>提出应用遗传算法来克服模糊 C 均值聚类算法的不足,但此方法需要编码和解码,操作较为复杂. 卢波等<sup>[5]</sup>采用小生境 Pareto 遗传算法,将结构面产状优势分组问题转化为多目标优化问题进行求解,由于遗传算法计算过程复杂,使得此方法也存在一定的缺陷.

随着岩体工程规模的逐渐扩大,需要分析的结构面数量急剧增加,寻找一种简单高效的结构面优势产状分析方法已经成为工程技术人员与相关学者亟须解决的问题. 鉴于以上方法存在的不足,本文应用一种全新的智能算法——萤火虫算法来搜索最佳的聚类中心,进而获得合理的岩体结构面产状优势分组结果. 萤火虫算法具有出色的全局搜索能力和较快的收敛速度,其结果要明显优于遗传算法和粒子群优化算法. 最后将该算法应用于松塔电站坝址区平硐实测结构面产状数据的分组分析.

# 1 问题的数学描述及计算原理

## 1.1 结构面产状数据的表达及相似性度量

在现场调查统计结构面数据时,结构面的产状通常用倾向( $\alpha$ )和倾角( $\beta$ )表示, $\Theta = (\alpha, \beta)$ ,其中, $0^\circ \leq \alpha < 360^\circ, 0^\circ \leq \beta \leq 90^\circ$ . 在结构面优势产状分析过程中,通常不考虑其厚度及起伏程度而将其简化为一空间平面,并用对应的单位法向量  $\mathbf{X} = (x, y, z)$  来表示结构面的空间方位:

$$\left. \begin{aligned} x &= \cos\alpha \cdot \sin\beta, \\ y &= \sin\alpha \cdot \sin\beta, \\ z &= \cos\beta. \end{aligned} \right\} \quad (1)$$

岩体结构面产状数据划分问题就是将空间方向相近的结构面分别归类分组,找出其优势方向,因此,首先需要选择合适的数学指标来度量结构面产状的相似性. 在岩体结构面产状优势组数划分问题中存在一种特殊情况,即倾向相差约  $180^\circ$  的两个高陡倾角结构面是相似的,应该划分为同一组. 因此本文采用两结构面所夹锐角的正弦值来判断两结构面方向的接近程度. 单位法向量分别为  $\mathbf{X}_1 = (x_1, y_1, z_1)$  和  $\mathbf{X}_2 = (x_2, y_2, z_2)$  的两结构

面所夹锐角为

$$\theta = \arccos |\mathbf{X}_1 \cdot \mathbf{X}_2^T|, \quad (2)$$

则两结构面产状间距离

$$d(\mathbf{X}_1, \mathbf{X}_2) = \sin\theta = \sqrt{1 - (\mathbf{X}_1 \cdot \mathbf{X}_2^T)^2}. \quad (3)$$

## 1.2 目标函数

给定  $N$  个结构面  $P_j (j = 1, 2, \dots, N)$ , 划分组数  $C$ , 每组的聚类中心为  $V_i (i = 1, 2, \dots, C)$ . 定义  $u_{ij}$  为第  $j$  个结构面属于第  $i$  个聚类中心的隶属度<sup>[6]</sup>:

$$u_{ij} = \frac{1}{d^2(P_j, V_i) \left[ \sum_{k=1}^C \left( \frac{1}{d^2(P_j, V_k)} \right) \right]^{-1}}, \quad (4)$$

则模糊目标函数

$$J_m = \sum_{j=1}^N \sum_{i=1}^C u_{ij}^2 d^2(P_j, V_i). \quad (5)$$

式(5)即为本文中岩体结构面产状数据划分方法的目标函数. 当分组数  $C$  给定时,使模糊目标函数  $J_m$  最小的聚类中心的组合即为要搜索的结构面优势方向. 为此,本文应用一种全新的智能算法——萤火虫算法来解决此问题. 由于分组数太多容易使人感到困惑且不便使用,通常限定  $C$  的取值范围为  $2 \sim 10$ . 待分别计算出对应不同  $C$  值的分组结果后,根据聚类有效性指标最终确定最佳分组数.

## 1.3 算法原理与计算过程

萤火虫算法(firefly algorithm, FA)是由剑桥大学学者 Yang<sup>[7]</sup>在 2008 年提出的一种模拟萤火虫群体行为的随机优化算法. 其计算原理可理解为:用萤火虫个体表示问题的解,根据其所处位置的亮度来度量解的优劣,利用亮度大的萤火虫吸引亮度小的萤火虫向其移动的过程来模拟问题的解向最优解逼近的过程,通过种群的不断更新最终获得问题的最优解<sup>[8]</sup>.

在萤火虫算法中,亮度与吸引度是两个非常重要的概念. 特定位置  $s$  的萤火虫的亮度  $I(s)$  是由解空间对应位置的目标函数值决定的. 萤火虫的吸引度  $\eta$  是随着个体之间的距离  $r$  的变化而变化的,同时也与萤火虫所处环境的介质密切相关. 萤火虫的吸引度  $\eta$  由下式确定:

$$\eta = \eta_0 \times \exp(-\gamma r^2). \quad (6)$$

式中: $\eta_0$  是最大吸引度,即  $r = 0$  时的吸引度; $\gamma$  是介质对光的吸收因子.

萤火虫  $i$  受到比其明亮的萤火虫  $k$  的吸引而向  $k$  的位置移动,萤火虫  $i$  移动后的位置由下式决定:

$$s_i = s_i + \eta(s_k - s_i) + \xi(\text{rand} - 0.5). \quad (7)$$

式中: $s_i, s_k$  表示萤火虫  $i$  和  $k$  所处的位置; $\xi$  为移

动步长;rand 为[0,1]上的随机数.

萤火虫算法具有操作简单、鲁棒性强、易于实现等特点. 与遗传算法和粒子群算法相比,其搜索全局最优解的能力更强,收敛速度更快<sup>[9]</sup>,目前已被应用到非线性问题优化<sup>[10]</sup>、图像压缩处理<sup>[11]</sup>等方面.

将萤火虫算法应用到岩体结构面优势方向的搜索,式(5)作为目标函数,约束条件为倾向  $\alpha \in [0^\circ, 360^\circ]$ ,倾角  $\beta \in [0^\circ, 90^\circ]$ . 通常意义上的萤火虫算法是要寻找亮度最大的个体,而结构面优势分组问题是要寻找当分组数给定时使目标函数最小时的聚类中心,为此,将  $s$  处萤火虫的亮度  $I(s)$  定义为

$$I(s) = \frac{1}{J_m + 1}.$$

(8)

应用萤火虫算法进行岩体结构面产状数据划分的计算流程如图 1 所示.

1.4 聚类有效性检验

在寻找到对应于不同分组数  $C$  的一系列最优解后,通过聚类有效性函数来评价这些聚类结果的优劣,进而确定最佳聚类数. 划分系数  $v_{PC}$  和 Xie - Beni 指标  $v_{XB}$  是常用的衡量聚类有效性的函数: $v_{PC}$  越大,聚类结果越好; $v_{XB}$  越小,聚类效果越好. 计算公式如下:

$$v_{PC} = \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^C u_{ij}^2.$$

(9)

$$v_{XB} = \frac{\sum_{i=1}^C \sum_{j=1}^N u_{ij}^2 d^2(P_j, V_i)}{N \cdot (\min_{i \neq k} d^2(V_i, V_k))}.$$

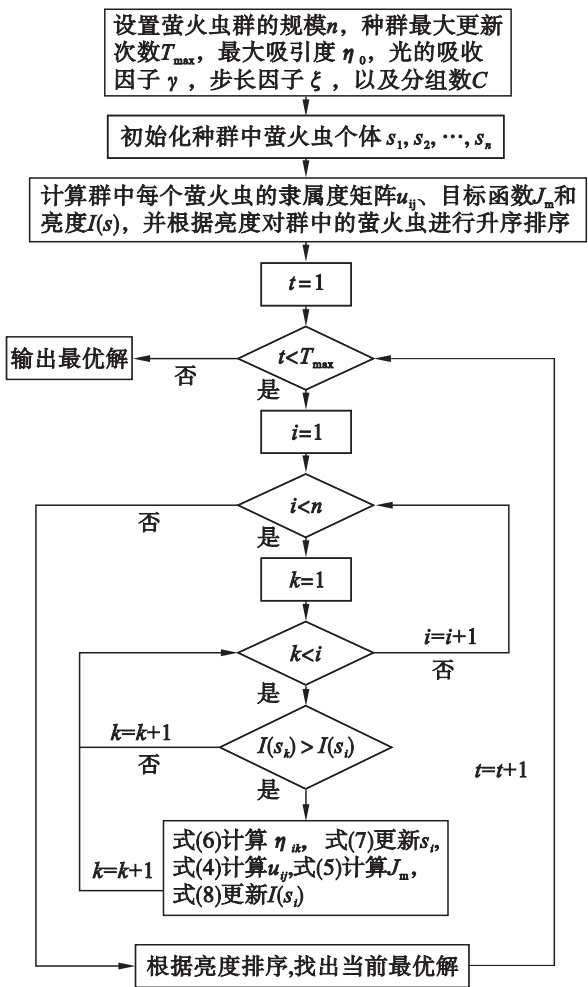
(10)

式(10)中,  $\min_{i \neq k} d^2(V_i, V_k)$  表示聚类中心之间的最小距离.

2 工程应用

怒江松塔水电站位于西藏自治区察隅县察瓦龙乡境内,坝址区属高山峡谷地貌,岩性主要为燕山晚期黑云二长花岗岩,节理裂隙比较发育. 为了进行围岩稳定性评价,首先要对结构面的发育规律进行分析研究,确定结构面优势方向. 本文选取坝址区右岸某高程相邻的四条平洞,实测 1 439 条结构面产状数据,图 2 为结构面的极点等密度图. 应用文中提出的算法对这些结构面进行优势分组. 算法参数设置为:萤火虫群的规模  $n = 20$ ,种群最大更新次数  $T_{max} = 50$ ,  $\eta_0 = 1$ ,  $\gamma = 1$ ,  $\xi = 0.1$ . 计算不同的分组数对应的聚类有效性检验结

果,见表 1.



图中  $t, i$  和  $k$  均为循环变量. 其中,  $t$  表示种群当前更新次数,  $i$  和  $k$  表示萤火虫个体

图 1 算法流程图

Fig. 1 Flow chart of the algorithm

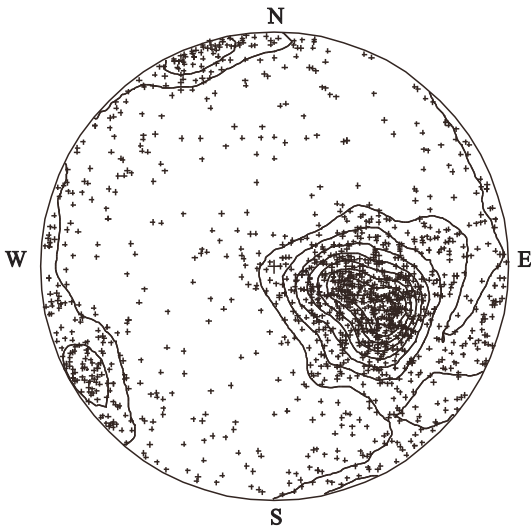


图 2 结构面极点等密度图

Fig. 2 Iso-density plot of discontinuity poles

从表 1 中可以看出,划分系数  $v_{PC}$  和 Xie-Beni 指标  $v_{XB}$  同时显示,当分组数为 3 时聚类效果最

好,所以综合确定最优分组数为 3,此时各组的中心产状及各组的结构面数目见表 2,详细分组结果见图 3a.

表 1 聚类结果有效性检验

Table 1 Validity test of clustering results

指标	分组数 $C$									
	2	3	4	5	6	7	8	9	10	
$v_{PC}$	0.62	0.64	0.56	0.52	0.48	0.45	0.43	0.41	0.37	
$v_{XB}$	0.21	0.11	0.32	0.29	0.34	0.31	0.33	0.32	0.29	

表 2 分组结果

Table 2 Clustering results of discontinuity data

组号	1	2	3
倾向/(°)	246.61	155.43	109.36
倾角/(°)	84.27	88.97	34.40
结构面数量	300	298	841

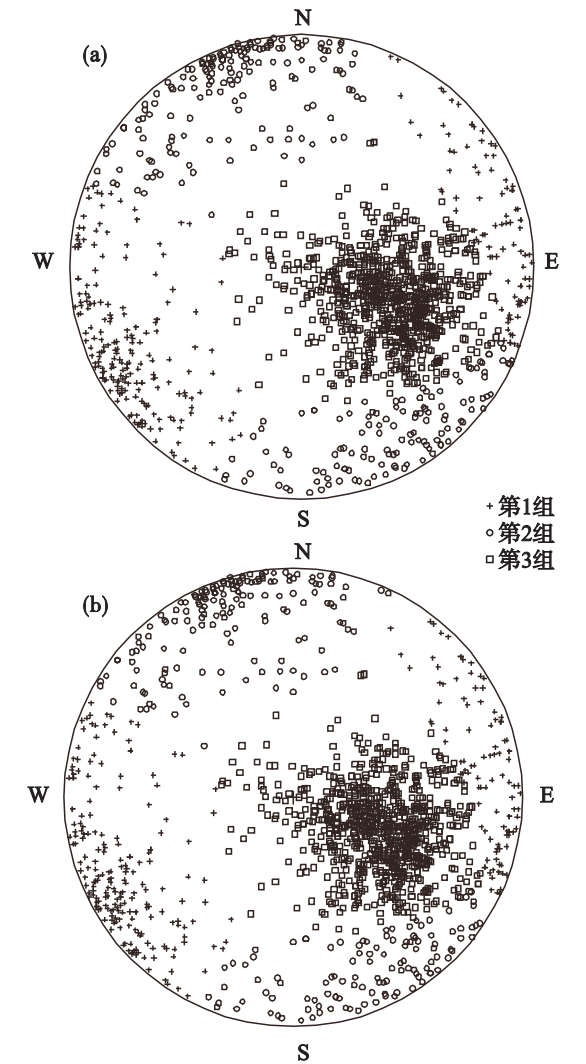


图 3 分组极点图

Fig. 3 Pole map of clustering result

(a)—新方法; (b)—文献[12]方法.

对比图 2 与图 3a 不难看出,新方法的分组结

果是符合实际情况的,此方法不但准确搜索出各组的聚类中心,还给出了清晰的组间边界,这是传统方法做不到的. 利用 Xu 等<sup>[12]</sup>提出的变尺度混沌优化模糊聚类方法计算得出的分组结果如图 3b 所示,可以看出,两种方法的分组结果是一致的,这也进一步说明了新方法的有效性. 但由于后者采用混沌搜索具有一定的盲目性,计算时间很长(约 90 min),而新方法的计算时间不足 1 min,说明新方法具有较高的计算效率,优势明显. 图 4 表示种群的更新过程,可以看出,算法的收敛速度是非常快的.

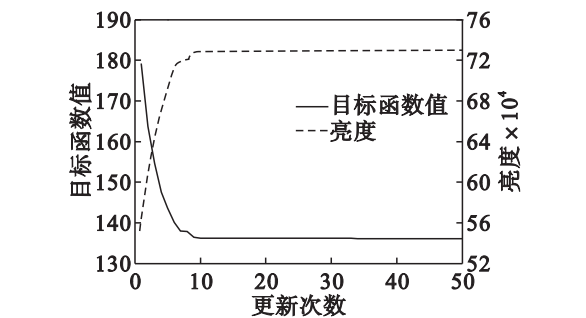


图 4 群更新过程

Fig. 4 Population updating process

### 3 结 论

- 1) 基于萤火虫算法的岩体结构面产状数据优势组数划分的新方法能够自动识别各组的聚类中心以及清晰合理的组间边界,有效避免了传统的图形分析法主观性强的缺点.
- 2) 与文献[12]方法的分组结果进行对比,进一步证明了新方法的正确性. 但是,值得注意的是,新方法耗时远远低于文献[12]方法,这一优点在工程应用中是非常重要的. 因此,新方法具有较大的实际意义.

#### 参考文献:

[1] Shanley R J, Mahtab M A. Delineation and analysis of clusters in orientation data[J]. *Journal of the International Association for Mathematical Geology*, 1976, 8(1): 9–23.

[2] Mahtab M A, Yegulalp T M. A rejection criterion for definition of clusters in orientation data[C]// *Proceedings of the 22nd Symposium on Rock Mechanics*. New York: American Institute of Mining Metallurgy and Petroleum Engineers, 1982: 116–123.

[3] Hammah R E, Curran J H. Fuzzy cluster algorithm for the automatic identification of joint sets [J]. *International Journal of Rock Mechanics and Mining Sciences*, 1998, 35(7): 889–905.