

# 基于粒子群算法的岩体结构面产状模糊 C 均值聚类分析\*

宋金龙 黄润秋 裴向军

(成都理工大学 地质灾害防治与地质环境保护国家重点实验室 成都 610059)

**摘要** 结构面产状分析是进行岩体力学分析及稳定性评价的基础,玫瑰花图、等密度图等传统的图形分析方法较为粗糙,无法对产状数据进行准确分析,采用模糊聚类方法则可以得到较为准确的优势产状,但需要事先指定分组数及初始聚心,且模糊聚类算法为局部寻优算法,初始划分对最终的数据分析结果影响较大。为了得到较为客观的分组结果及优势产状,同时能够针对大量结构面产状数据进行聚类分析,提出了基于粒子群算法的岩体结构面产状模糊 C 均值聚类算法。采用粒子群算法进行模糊 C 均值聚类算法初始聚心的求解,在计算过程中可同时确定最佳分组数,避免了人为指定的主观性,克服了模糊 C 均值聚类算法易陷入局部极小值和对初始划分敏感的不足。最后,通过工程实例中不同方法的聚类效果对比分析该算法的有效性,并将该方法应用于实测结构面产状数据的分析,得到较为符合实际的优势结构面分组。

**关键词** 岩体 结构面 模糊 C 均值聚类 粒子群算法

**中图分类号:**TU45 **文献标识码:**A

## PARTICLE SWARM OPTIMIZATION ALGORITHM BASED FUZZY C-MEANS CLUSTER ANALYSIS FOR DISCONTINUITIES OCCURRENCE IN ROCK MASS

SONG Jinlong HUANG Runqiu PEI Xiangjun

(State Key Laboratory of Geological Hazard Prevention and Geological Environment Protection, Chengdu University of Technology, Chengdu 610059)

**Abstract** The analysis of occurrence of discontinuities is a basic work for further study on mechanical analysis and stability of rock mass. Traditional analysis of the plot method is insufficient and inadequate. And they are unable to analyze occurrence data of discontinuities accurately. Although the fuzzy cluster method can achieve accurate dominant occurrences, the initial cluster centroids must be given in advance. And the method uses a local optimized algorithm. Different choices of initial guesses of cluster centroids can lead to different partitions of the same data. In order to obtain objective results of classification and the dominant occurrence, a fuzzy c-means cluster analysis method based on the Particle Swarm Optimization (PSO) algorithm is proposed. This hybrid method uses the PSO algorithm to solve the problem. So it can avoid the subjectivity of the initial cluster centroids specified manually, overcome the defects of the fuzzy c-means algorithm such as the local optima and sensitivity to

\* 收稿日期: 2011-10-25; 收到修改稿日期: 2012-02-26.

**基金项目:** 国家自然科学基金(40972195), 地质灾害防治与地质环境保护国家重点实验室自由探索基金(SKLG2009Z009).

**第一作者简介:** 宋金龙, 主要从事岩体稳定及环境地质工程方面的研究. Email: boris\_song@163.com

initialization, and analyse the occurrence data efficiently, even if there are a large number of discontinuities. Meanwhile, the optimal cluster number can be determined automatically during the operational process of the algorithm. On the basis of the field measured data from the real rock mass, the proposed approach has been testified to be reliable and reasonable. And the classification and dominant occurrences are more realistic.

**Key words** Rock mass, Discontinuity, Fuzzy C-means cluster algorithm, PSO algorithm

## 1 引言

岩体中的结构面是岩体中力学强度相对薄弱的部位,它导致岩体力学性能的不连续性、不均一性和各向异性。只有掌握岩体的结构特征,查明岩体中结构面的发育特征及规律,才有可能阐明岩体在不同荷载条件下内部的应力分布和分异状况<sup>[1, 2]</sup>。而结构面的产状则是其重要的几何特征之一,对结构面产状的分析是进行结构面三维网络模拟、岩体渗流分析以及岩体稳定性评价的基础。

岩体结构面的发育具有规律性,对结构面产状的传统分析方法主要采用玫瑰花图、极点图和等密度图等,这些方法较为直观,但其分析主要依靠经验,且不能得到较准确的结构面优势产状,主观性因素较大,对操作人员的专业素养要求较高。因此需要采用科学合理的数学方法来进行结构面产状数据的分析。

Shanley 等<sup>[3, 4]</sup>提出了针对结构面产状的聚类算法,其主要通过确定不同的小球半径以寻找密度点来进行结构面聚类;而后 Hammah 和 Curran<sup>[5]</sup>将模糊 C 均值(FCM)聚类算法用于结构面的分类,能够在一定程度上得到较优的结构面分组情况。但 FCM 算法的聚类个数依赖人为指定,且其本质上为局部寻优算法,易陷入局部极小点,对初始条件较为敏感。

针对 FCM 算法的缺陷,蔡美峰等<sup>[6]</sup>基于遗传算法对 FCM 算法进行了改进,通过遗传算法对初始聚类中心进行求解,并采取多种模糊分类指标对聚类中心以及聚类效果进行评价,最后通过比较确定结构面的分组结果;周玉新等<sup>[7]</sup>采用模糊等价聚类和模糊软聚类相结合的方法,通过模糊等价聚类方法求得较优分组结果,以此作为模糊软聚类法的结构面组初始划分,并以模糊软聚类法所得聚类中心作为各分组的中心产状;冯羽等<sup>[8]</sup>通过对传统图形分析法、模糊等价聚类方法以及模糊 C 均值聚类算法的优缺点进行分析,提出将此 3 种方法有顺序得结合起来,综合应用以得到更为准确合理的结构面分

组及优势产状。

以上方法在一定程度上对模糊 C 均值算法进行了不同的优化,但仍无法克服人为确定结构面分组数时的主观性,尤其当结构面样本数据之间离散程度较大时,将直接影响最终的结构面分组情况。滕继东等<sup>[9]</sup>将蚁群算法用于结构面产状模糊 C 均值聚类算法的优化,将聚类中心作为蚁群算法的求解结果,可在确定优势产状的同时确定结构面分组数,但蚁群算法计算量大,求解时间长,且求解过程易出现停滞现象。

为了克服上述算法进行结构面产状分析时的主观性,并能有效进行大样本量时的数据分析,本文结合 FCM 聚类算法能够较为准确计算结构面组优势产状的特点,提出一种基于粒子群(PSO)算法的结构面产状模糊 C 聚类算法,不仅可以利用 PSO 算法的全局寻优能力克服 FCM 算法易陷入局部最优的缺陷,得到较为准确的分组结果及各组的优势产状,还可在算法运行过程中自动识别最佳分组数,提高了聚类分析的客观性及算法的自主性,而且算法简单,收敛速度快。

## 2 结构面模糊 C 聚类分析

### 2.1 结构面产状的空间表示方法

岩体结构面在忽略厚度的情况下,一般认为是在空间范围内延伸的具有一定形状的平面<sup>[10]</sup>,其产状采用倾向( $0^\circ \leq \alpha \leq 360^\circ$ )和倾角( $0^\circ \leq \beta \leq 90^\circ$ )表示。本文选取结构面单位法向矢量来表征其空间方位,采用上半球投影法,以正北方(N)为  $y$  轴正方向,正东方(E)为  $x$  轴正方向,正上方为  $z$  轴正方向,建立三维空间直角坐标系(图 1)。

则倾向为  $\alpha$ ,倾角为  $\beta$  的结构面  $\theta$  产状可由其单位法向量  $x$  表示为  $(x_1, x_2, x_3)$ ,其中,

$$\begin{cases} x_1 = \cos\alpha \sin\beta \\ x_2 = \sin\alpha \sin\beta \\ x_3 = \cos\beta \end{cases} \quad (1)$$

$$\sum_{i=1}^3 x_i^2 = 1 \quad (2)$$

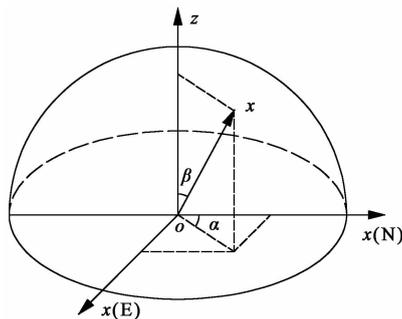


图1 坐标系及结构面的法向量

Fig. 1 Coordinate system and normal vector of discontinuity

通过对结构面产状的单位矢量化可以实现对聚类数据的归一化处理。

## 2.2 相似性度量

FCM 算法中一般采用距离作为样本间的相似性度量,而模糊数学中常采用相似系数表示向量之间的相似程度,相似程度越高,相似系数值越大,其计算方法有数量积法、夹角余弦法、相关系数法等<sup>[11]</sup>。文中采用夹角余弦法计算任意两结构面  $x_i = (x_{i1}, x_{i2}, x_{i3})$ ,  $x_j = (x_{j1}, x_{j2}, x_{j3})$  的相似系数  $r_{ij}$  为:

$$r_{ij} = \frac{\left| \sum_{k=1}^3 x_{ik} \cdot x_{jk} \right|}{\sqrt{\sum_{k=1}^3 (x_{ik})^2} \sqrt{\sum_{k=1}^3 (x_{jk})^2}} = |x_{i1} \cdot x_{j1} + x_{i2} \cdot x_{j2} + x_{i3} \cdot x_{j3}| \quad (3)$$

由  $r_{ij} \in [0, 1]$ , 本文取两结构面距离  $d(x_i, x_j)$  计算公式为:

$$d(x_i, x_j) = 1 - r_{ij} = 1 - |x_{i1} \cdot x_{j1} + x_{i2} \cdot x_{j2} + x_{i3} \cdot x_{j3}| \quad (4)$$

## 2.3 结构面 FCM 聚类方法

设实测结构面样本集为  $X = \{x_1, x_2, \dots, x_n\}$ , 利用 FCM 算法将  $X$  划分为  $c$  类 ( $2 \leq c \leq n$ ):  $A_1, A_2, \dots, A_c$ , 可设  $U$  为  $X$  的  $C$ -模糊划分矩阵, 表示为:

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & & \vdots \\ u_{c1} & u_{c2} & \cdots & u_{cn} \end{pmatrix} \quad (5)$$

其中,  $u_{ij}$  为第  $j$  个样本属于第  $i$  个模糊子类  $A_i$  的隶属度, 需满足如下条件:

$$(1) 0 \leq u_{ij} \leq 1, i = 1, 2, \dots, c; j = 1, 2, \dots, n;$$

$$(2) \sum_{i=1}^c u_{ij} = 1, j = 1, 2, \dots, n;$$

$$(3) 0 < \sum_{j=1}^n u_{ij} < n, i = 1, 2, \dots, c.$$

(1) 表示每一个样本  $x_j$  属于  $A_i$  的隶属度取值范围; (2) 表示每一个  $x_j$  属于  $c$  个  $A_i$  的总和为 1; (3) 表示每一个  $A_i$  不等于  $\emptyset$  或  $X$ 。

对应划分矩阵  $U$  的聚类中心表示为  $Y = \{y_1, y_2, \dots, y_c\}$ , 则  $(U, Y)$  为  $X$  的一个聚类。若有  $X$  的聚类  $(U, Y)$  使得泛函

$$J(U, Y) = \sum_{i=1}^c \sum_{j=1}^n (u_{ij})^m d(x_j, y_i)^2 \quad (6)$$

达到最小值, 即得最优聚类。其中,  $m$  为模糊加权指数,  $m \in [1, +\infty)$ , 用于控制隶属度的分配及聚类的模糊程度, Bezdek<sup>[12]</sup> 得到了  $m=2$  时 FCM 算法的物理解释, 并与 Pal<sup>[13]</sup> 从聚类有效性角度得出  $m$  的取值范围为  $[1.5, 2.5]$ , 本文通过多次试验后取值  $m=2$ ;  $d(x_j, y_i)$  表示第  $j$  个样本  $x_j$  到第  $i$  类聚类中心  $y_i$  的距离。当  $x_j \neq y_i$  时, 运用拉格朗日乘法可得  $u_{ij}$  和  $y_i$  的更新公式为:

$$u_{ij} = 1 / \sum_{l=1}^c (d(x_j, y_l) / d(x_j, y_i))^{2(m-1)} \quad (7)$$

$$y_i = \sum_{j=1}^n (u_{ij})^m x_j / \sum_{j=1}^n (u_{ij})^m \quad (8)$$

考虑近直立结构面组中出现倾角近似, 倾向大致相差  $180^\circ$  的情况, 本文对聚心更新方法进行适当优化, 当样本与聚心夹角大于  $90^\circ$  时, 将该样本以坐标原点进行中心对称后再进行聚心更新计算。

算法在给定分组数及初始划分矩阵后, 逐步迭代, 并根据公式(7)和公式(8)更新  $U$  和  $Y$ , 计算泛函  $J(U, Y)$ , 算法收敛即获得最优划分矩阵  $U$  及对应的聚类中心集  $Y$ 。模糊分类的清晰化可采用最大隶属度原则实现<sup>[11]</sup>, 即令  $x_j$  归于隶属度最大的类。

由于 FCM 算法为局部寻优陷入局部极优点, 法, 当初始划分矩阵不合理时易且若预先给定的分组数不合理, 亦无法得到合理的聚类结果。因此, 本文尝试采用 PSO 算法解决岩体结构面模糊 C 均值聚类算法的全局寻优问题, 并在计算过程中自动完成最佳分组数的确定。

## 3 基于 PSO 的结构面 FCM 聚类

粒子群 (PSO) 优化算法<sup>[14]</sup> 是由 Eberhart 和 Kennedy 于 1995 年提出的基于种群的全局随机优化算法, 算法简洁, 可调参数少, 易于实现, 且具有较强的全局搜索能力及扩展性, 求解速度快。该算法

源自对鸟群觅食行为的研究,将每个个体看作是搜索空间内以一定速度运动的粒子,通过信息共享及个体经验不断调整速度,并判断自己位置的优劣。算法经过不断进化后收敛于全局最优的个体位置。

为了解决离散特征空间内的优化问题,Kennedy和Eberhart<sup>[15]</sup>在连续 PSO 算法的基础上发展了离散 PSO 算法。其本质区别主要体现在采用二元变量表示粒子,并将粒子的速度映射成为概率的变化,即二元变量值为 1 的可能性。

由于粒子群算法在收敛到全局最优解附近时,解的精确性很难进一步提高,易陷入近似局部最优<sup>[16]</sup>,因此本文根据结构面 FCM 聚类的特点,以 FCM 算法的初始划分作为粒子群算法的优化对象,利用 FCM 的局部寻优提高粒子群算法的精度,并采用改进的适应度函数表征粒子位置的优劣,将粒子群算法与 FCM 算法进行有机的结合。

### 3.1 粒子编码及更新方案

#### 3.1.1 编码方案

将粒子采用  $D$  维向量表示,则第  $i$  个粒子的位置可用二元变量表示为  $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$ ,  $x_{id} \in \{0, 1\}$ 。取  $N_c$  个结构面样本向量形成初始聚心集  $M$ ,令  $D = N_c$ ,随机初始化粒子,则各粒子每一维  $x_{id}$  的取值即代表所对应  $M$  中的结构面样本是否选作初始聚心,被选为初始聚心的样本总个数即为结构面分组数  $c$ 。因而粒子的位置可代表结构面 FCM 聚类的初始划分,而粒子的速度用以表征粒子位置的变化,可表示为  $v_i = (v_{i1}, v_{i2}, \dots, v_{id})$ ,  $v_{id} \in R$ 。通过对粒子位置的全局优化求解可获得较优的 FCM 初始划分。由于结构面分组数至少为 1,因此需避免出现粒子各维均为 0 的情况。算法运行时可以采用如下公式初始化粒子:

$$x_{id} = \begin{cases} 1, & r_d < p_{ini} \\ 0, & r_d \geq p_{ini} \end{cases} \quad (9)$$

其中,  $r_d \in (0, 1)$  为随机数;  $p_{ini}$  为初始化粒子位置的概率,  $p_{ini}$  取值越大,选定的聚类中心个数越多。

#### 3.1.2 更新方案

粒子的运动根据其速度确定,在离散 PSO 算法中,速度则表示粒子相应维度取 1 的概率,一般采用 Sigmoid 函数将速度值限定在  $[0, 1]$  范围内<sup>[17]</sup>:

$$s(v_{id}) = \frac{1}{1 + \exp(-v_{id})} \quad (10)$$

则  $t+1$  次迭代时的粒子位置更新公式为:

$$x_{id}^{(t+1)} = \begin{cases} 1, & r < s(v_{id}^{(t)}) \\ 0, & \text{其他情况} \end{cases} \quad (11)$$

其中,  $r \in (0, 1)$  为随机数。为了避免  $s(v_{id})$  过于接近 0 或 1,可使用常量  $V_{\max}$  限制  $v_{id}$  的范围,即令  $v_{id} \in [-V_{\max}, +V_{\max}]$ ,一般取  $V_{\max}$  值为 4。

粒子速度的调整公式如下:

$$v_{id}^{(t+1)} = \omega v_{id}^{(t)} + c_1 r_1 (p_{id}^{(t)} - x_{id}^{(t)}) + c_2 r_2 (p_{gd}^{(t)} - x_{id}^{(t)}) \quad (12)$$

式中,  $\omega$  为惯性权重;  $c_1$  为认知学习因子;  $c_2$  为群体学习因子,取常数且  $0 \leq c_1 + c_2 \leq 4$ ;  $r_1 \in (0, 1)$ ;  $r_2 \in (0, 1)$ ;  $p_i = (p_{i1}, p_{i2}, \dots, p_{id})$ , 为粒子的局部最优解;  $p_g = (p_{g1}, p_{g2}, \dots, p_{gd})$ , 为粒子群体的全局最优解。

由于  $\omega$  越大,算法全局搜索能力越强;反之,局部调整能力越强<sup>[18]</sup>。为提高算法性能,本文采用随时间变化的非线性函数动态调整惯性权重<sup>[16]</sup>,  $\omega$  的动态调整公式为:

$$\omega = \left\{ \frac{(\text{iter}_{\max} - \text{iter})^n}{\text{iter}} \right\} (\omega_{\text{initial}} - \omega_{\text{final}}) + \omega_{\text{final}} \quad (13)$$

其中,  $\omega_{\text{initial}}$  为初始惯性权重;  $\omega_{\text{final}}$  为终止惯性权重;  $\text{iter}_{\max}$ ,  $\text{iter}$  分别表示算法运行的最大迭代次数和当前迭代次数;  $n$  表示非线性调整指数。通过改变  $n$  的值可以调整算法全局搜索能力和局部搜索能力间的平衡。

### 3.2 适应度函数

本文以结构面样本空间内各结构面与其所属分组聚心的距离平方和作为粒子的适应度函数,如下所示:

$$F = \sum_{i=1}^c \sum_{j=1}^{m_i} d(x_{ij}, y_i) \quad (14)$$

其中,  $c$  为聚类个数;  $m_i$  为第  $i$  类  $A_i$  中样本的个数;  $d(x_{ij}, y_i)$  为分组  $A_i$  内第  $j$  个结构面样本与聚心的距离,可采用公式(4)进行计算。

由于上述适应度函数不适用于比较具有不同分组数的划分的优劣,因此本文采用一种融入启发式的适应度函数<sup>[19]</sup>,在上述适应度函数的基础上乘一个关于分组数  $c$  的幂次函数,得下式:

$$F = a^{(c+1)} \cdot \sum_{i=1}^c \sum_{j=1}^{m_i} d(x_j, y_i) \quad (15)$$

式中,  $a$  为常数,当  $a$  取不同的值时可以调整对粒子位置的适应度的判断,本文中取  $a$  为 1.9。

### 3.3 实现过程

(1)从结构面样本集  $X$  中排除相同数据后随机选取  $N_c$  个样本组成初始聚心集  $M$ ,初始化粒子群体

S, 并给定算法中各参数取值;

(2) 根据每个粒子的位置计算初始划分矩阵, 而后利用结构面产状的 FCM 聚类算法寻优, 得到对应的结构面样本的分组情况及各组的优势产状;

(3) 计算各粒子的当前适应度, 更新群体全局最优解及个体局部最优解, 同时更新群体 S 中每个粒子的位置和速度, 若 PSO 算法迭代至最大步数, 则运算停止, 否则转至第(2)步继续求解;

(4) 从 X 中随机选取新的样本, 并结合全局最优解对应的结构面聚心更新 M, 对新的 M 根据(2)、(3)步求解计算, 重复此过程至算法收敛停止;

(5) 使用 FCM 算法对 PSO 算法搜索到的最优解进行优化, 得到结构面产状聚类分析的最终结果。

笔者根据前述原理及设计方法流程, 采用 VC++ 6.0 平台编制了相应的结构面产状分组程序, 以利于实际应用。

### 3.4 聚类有效性评价

目前已提出很多评价聚类有效性的指标<sup>[20]</sup>, 本文根据结构面产状数据聚类的特点, 采用划分系数  $V_{PC}$  及分类熵  $V_{CE}$  两项指标对比评价不同模糊聚类方法的聚类有效性。相应计算公式如下:

$$V_{PC}(U, c) = \frac{1}{n} \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^2 \quad (16)$$

$$V_{CE}(U, c) = -\frac{1}{n} \sum_{k=1}^n \sum_{i=1}^c u_{ik} \log_a(u_{ik}) \quad (17)$$

式中, 对数的底数  $a \in (1, +\infty)$ , 且规定当  $u_{ik} = 0$  时  $u_{ik} \log_a(u_{ik}) = 0$ , 本文取自然对数。划分系数  $V_{PC}$  越接近 1, 分类熵  $V_{CE}$  越接近 0, 聚类效果越好; 反之, 聚类效果越差。

## 4 实例分析

为了验证基于粒子群算法的结构面产状模糊 C 均值聚类方法的有效性和可靠性, 本文取两处工程实例进行结构面产状的聚类分析。算法中取初始种群个数为 20, 初始聚心集个数为 25, 粒子位置初始化概率  $p_{ini} = 0.8$ , 粒子速度更新公式中的各种参数设置如下:  $n = 1.2$ ,  $\omega_{initial} = 0.2$ ,  $\omega_{initial} = -0.3$ ,  $iter_{max} = 300$ ,  $c_1 = c_2 = 2$ 。

(1) 京珠高速鄂北某路段岩质边坡结构面数据实例数据取自文献[10]第 8 章工程实例一, 共 337 组结构面数据, 其赤平投影下的产状极点及密度等值线图如图 2 所示。

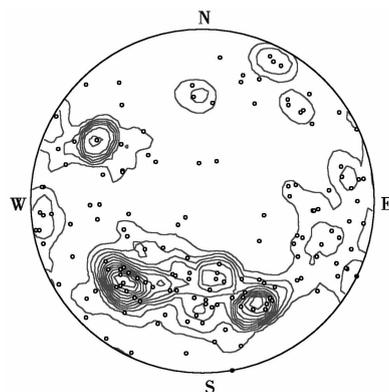


图 2 结构面产状极点及密度等值线图

Fig. 2 Pole and contour plot of discontinuity occurrences

从图中可以看出, 该处结构面的优势结构面组间界限并不明显。文献中根据玫瑰花图并结合实际情况人工判断结构面分组数为 5, 后采用 K-means 动态聚类算法划分结构面组。本文针对此结构面数

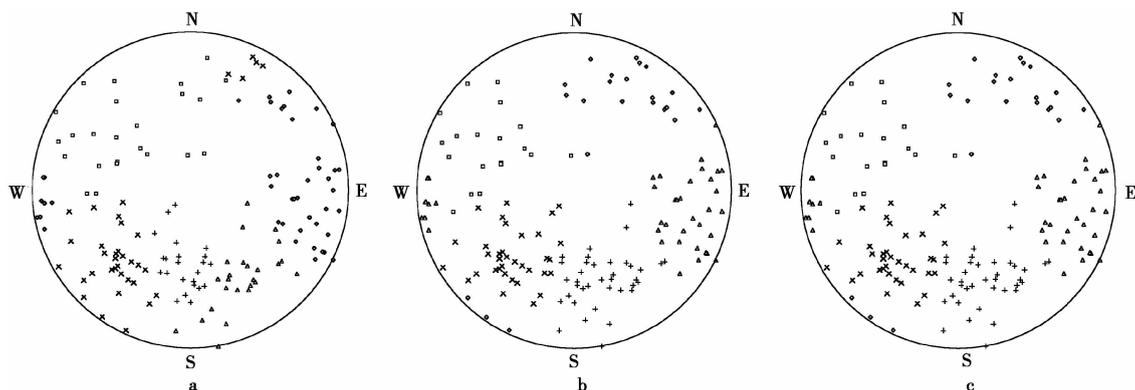


图 3 不同算法聚类结果对比

Fig. 3 Comparison of clustering results for different algorithms

a. 模糊 C 均值法; b. 模糊 C 均值法(优化初始划分); c. 基于粒子群的模糊 C 均值法

表 1 不同算法所得优势结构面组具体参数对比

Table 1 Comparison of basic parameters for the dominant discontinuity sets obtained by different algorithms

组别	模糊 C 均值法		模糊 C 均值法(优化初始划分)		基于 PSO 的模糊 C 均值法	
	优势产状/(°)	结构面数量/条	优势产状/(°)	结构面数量/条	优势产状/(°)	结构面数量/条
1(×)	223.6∠58.2	96	221.5∠51.2	101	221.4∠51.2	101
2(+)	184.8∠37.6	58	161.3∠51.0	93	161.2∠51.0	93
3(□)	305.2∠56.1	59	298.8∠58.2	49	298.9∠58.1	49
4(Δ)	151.9∠58.7	59	97.0∠72.0	60	97.0∠72.1	60
5(◇)	87.7∠73.7	65	24.4∠68.2	34	24.5∠68.2	34

据分别采用 FCM 法和基于 PSO 的 FCM 法进行聚类分析。其中,FCM 法需指定分组数为 5,同时将文献中所示优势产状及分组情况作为优化的初始划分进行 FCM 聚类计算并与一般的 FCM 法进行聚类效果比较。

不同算法的聚类结果对比见图 3 及表 1。可以看出,3 种方法所得结构面的聚类分组均符合文献中实例情况的介绍及实际判断,结果可靠,但对比各方法的具体分组可看出,优化初始划分的 FCM 法与基于 PSO 的 FCM 法所得结构面聚类分组结果一致,仅优势产状的计算有细微差别,而一般 FCM 法所得分组与上述 2 种方法有一定差别。通过对比表 2 中的聚类有效性指标,可以看出相比于一般的 FCM 算法,基于 PSO 的 FCM 法在聚类效果上得到了明显的提升,并且与优化初始划分的 FCM 法相比在两项指标上更优。相比之下,优化初始划分的 FCM 法需要事先判断出正确的分组数,继而才能进一步根据分组数采取其他方法进行 FCM 法初始划分的优化,此过程无法避免人为操作的主观性,与基于 PSO 的 FCM 法相比操作更加繁琐,且易受干扰而直接影响最终的聚类效果。

表 2 不同算法聚类有效性对比

Table 2 Comparison of different algorithms' cluster validity

算法	划分系数 $V_{PC}$	分类熵 $V_{CE}$
模糊 C 均值法	0.7848	0.4232
模糊 C 均值法(优化初始划分)	0.8342	0.3268
基于 PSO 的模糊 C 均值法	0.8344	0.3266

通过上述对比分析表明,基于 PSO 算法的 FCM 法能够客观地对结构面产状数据进行聚类分析,避免了主观性因素的影响,并且具有良好的聚类效果,可以快速准确地识别优势组数,得到合理有效的分组结果。

(2) 锦屏水电站某处边坡局部岩体结构面数据

取锦屏一级水电站 C Ⅲ 标左岸拱间槽下游坡局部实测 413 组结构面数据进行分析,图 4 为结构面统计的极点图及密度等值线图。从图 4 中可以较直观地发现该处主要发育 4 组优势结构面,其优势产状范围分别是:  $280^{\circ} \sim 315^{\circ} \angle 33^{\circ} \sim 41^{\circ}$ ;  $130^{\circ} \sim 155^{\circ} \angle 50^{\circ} \sim 65^{\circ}$ ;  $180^{\circ} \sim 200^{\circ} \angle 80^{\circ} \sim 90^{\circ}$ ;  $80^{\circ} \sim 100^{\circ} \angle 45^{\circ} \sim 60^{\circ}$ 。

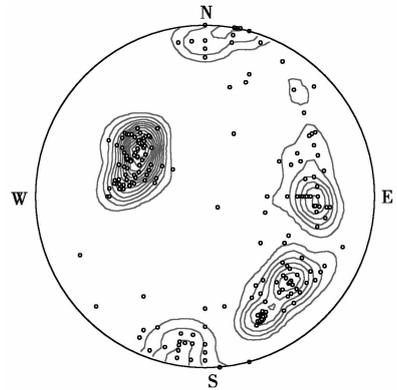


图 4 结构面产状极点及密度等值线图

Fig. 4 Pole and contour plot of discontinuity occurrences

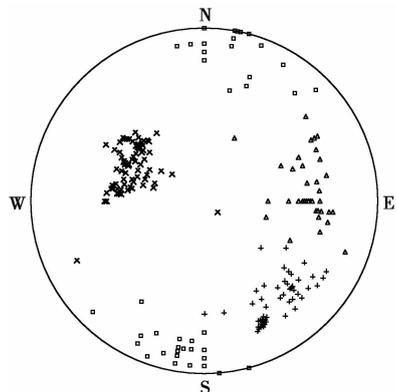


图 5 结构面分组结果

Fig. 5 Results of cluster analysis for discontinuities

对此结构面数据运用基于 PSO 的 FCM 方法进行产状的聚类分析,计算结果如图 5 及表 3 所示。

表 3 优势结构面组的具体参数

Table 3 Basic parameters of the dominant discontinuity sets

组号	优势产状		结构面数量/条
	倾向/(°)	倾角/(°)	
1(×)	298.5	38.7	166
2(+)	140.9	61.2	98
3(□)	189.1	89.0	77
4(Δ)	86.4	53.0	72

从图 5 可以看出,结构面的分组划分结果与其产状的密度等值线图所反映的分组区域一致。聚类有效性指标计算结果为:划分系数  $V_{PC}$ :0.9534,分类熵  $V_{CE}$ :0.0958,表明基于 PSO 的 FCM 方法所得聚类效果较好,计算得到的各结构面优势组所对应的优势产状准确,算法收敛速度快,便于实际应用。

## 5 结 论

本文在结构面产状的模糊 C 均值聚类分析的基础上,引入了人工智能研究中的粒子群(PSO)算法,将二者进行有机结合,提出了基于粒子群算法的岩体结构面产状模糊 C 均值聚类分析方法。通过采用二元变量编码,将结构面的最佳分组数及 FCM 算法所需的初始划分转化为 PSO 算法的寻优目标,针对分组数不同时聚类效果的判断采用一种融入了启发式的适应度函数。避免了人为给定分组数的主观性,克服了直接采用 FCM 算法进行结构面分组时对初始划分敏感、易陷入局部最优点的缺陷,可以给出较为客观的分组结果及各组优势产状。同时编制程序并结合工程实例进行分析,分析结果表明该方法能够准确识别最佳分组数,分组结果合理,求解精度较高,收敛速度快,可以为岩体的力学分析、数值模拟及稳定性分析提供可靠的依据。

## 参 考 文 献

[1] 张倬元,王士天,王兰生. 工程地质分析原理[M]. 北京:地质出版社,1994.  
Zhang Zhuoyuan, Wang Shitian, Wang Lansheng. Principles of Engineering Geological Analysis. Beijing: Geological Publishing House, 1994.

[2] 伍法权. 岩体工程地质动力学基本原理[J]. 工程地质学报, 2011,19(3): 304~316.  
Wu Faquan. Principles of engineering geological dynamics of rockmass. Journal of Engineering Geology, 2011,19(3): 304~316.

[3] Shanley RJ, Mahtab MA. Delineation and analysis of clusters in orientation data. Math. Geol., 1976,8(3): 9~23.

[4] Mahtab MA, Yegulalp TM. A rejection criterion for definition of clusters in orientation data // Proceedings of the 22nd Symposium on Rock Mechanics. New York: American Institute of Mining Metallurgy and Petroleum Engineers, 1982, 116~123.

[5] Hammah RE, Curran JH. Fuzzy cluster algorithm for the automatic identification of joint sets. International Journal of Rock Mechanics and Mining Sciences, 1998,35(7): 889~905.

[6] 蔡美峰,王鹏,赵奎,等. 基于遗传算法的岩体结构面的模糊 C 均值聚类方法[J]. 岩石力学与工程学报, 2005,24(3): 371~376.  
Cai Meifeng, Wang Peng, Zhao Kui, et al. Fuzzy C-means cluster analysis based genetic algorithm for automatic identification of joint sets. Chinese Journal of Rock Mechanics and Engineering, 2005,24(3): 371~376.

[7] 周玉新,周志芳,孙其国. 岩体结构面产状的综合模糊聚类分析[J]. 岩石力学与工程学报, 2005,24(13): 2283~2287.  
Zhou Yuxin, Zhou Zhifang, Sun Qiguo. Synthetic fuzzy clustering analysis for joints occurrence of rock mass. Chinese Journal of Rock Mechanics and Engineering, 2005,24(13): 2283~2287.

[8] 冯羽,马凤山,巩城城,等. 节理岩体结构面优势产状确定方法研究[J]. 工程地质学报, 2011,19(6): 887~892.  
Feng Yu, Ma Fengshan, Gong Chengcheng, et al. Data analysis method for optimized and dominant orientations of joints in rock mass. Journal of Engineering Geology, 2011,19(6): 887~892.

[9] 滕继东,徐光黎,申艳军. 基于蚁群算法的结构面产状模糊 C 均值聚类分析[J]. 安全与环境工程, 2010,17(4): 114~117.  
Teng Jidong, Xu Guangli, Shen Yanjun. Fuzzy C-means cluster analysis based on ant colony algorithm for orientations of joint. Safety and Environmental Engineering, 2010,17(4): 114~117.

[10] 贾洪彪,唐辉明,刘佑荣,等. 岩体结构面三维网络模拟理论与工程应用[M]. 北京:科学出版社,2008.  
Jia Hongbiao, Tang Huiming, Liu Yourong, et al. Theory and Engineering Application of 3-d Network Modeling of Discontinuities in Rockmass. Beijing: Science Press, 2008.

[11] 梁保松,曹殿立. 模糊数学及其应用[M]. 北京:科学出版社,2007.  
Liang Baosong, Cao Dianli. Fuzzy Mathematics and Applications. Beijing: Science Press, 2007.

[12] Bezdek JC. A physical interpretation of fuzzy ISODATA. IEEE Trans. Syst. Man Cybern., 1987,17(5): 873~877.

[13] Pal NR, Bezdek JC. On cluster validity for the fuzzy c-means model. IEEE Trans. Fuzzy System, 1995,3(3): 370~379.

[14] Eberhart RC, Kennedy J. A new optimizer using particle swarm theory // Proceedings of the 6th International Symposium on Micro Machine and Human Science. Piscataway, NJ: IEEE Service Center, 1995, 39~43.

[15] Kennedy J, Eberhart RC. A discrete binary version of the particle swarm algorithm // Proceedings of IEEE Conference on Systems, Man and Cybernetics. Piscataway, NJ: IEEE Press, 1997, 4104~4108.

- [16] Chattejee A, Siarry P. Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization. *Computers & Operations Research*, 2006, **33**(3): 859 ~ 871.
- [17] 张长胜. 求解规划、聚类 and 调度问题的混合粒子群算法研究 [博士学位论文][D]. 长春: 吉林大学, 2009.  
Zhang Changsheng. Research on Hybrid Particle Swarm Algorithms for Programming, Clustering and Scheduling Problems. Changchun: Jilin University, 2009.
- [18] Shi Y, Eberhart R. A modified particle swarm optimizer // IEEE International Conference on Evolutionary Computation Proceedings. Piscataway, NJ: IEEE Press, 1998, 69 ~ 73.
- [19] Lee C-Y, Antonsson EK. Dynamic partitioned clustering using evolution strategies // the 26th Annual Conference of the IEEE Industrial Electronics Society. Piscataway, NJ: IEEE Press, 2000, 2716 ~ 2721.
- [20] 唐明会, 杨燕. 模糊聚类有效性的进展研究[J]. 计算机工程与科学, 2009, **31**(9): 122 ~ 124.  
Tang Minghui, Yang Yan. Research and development of fuzzy clustering validity. *Computer Engineering and Science*, 2009, **31**(9): 122 ~ 124.

## 新书推荐

# 中国地质灾害区域预警方法与应用

刘传正 刘艳辉 温铭生 唐 灿 薛群威 著

## 内 容 提 要

本书概述了国内外地质灾害区域预警预报理论研究和应用服务发展动态, 阐述了中国地质环境和地质灾害分布特点, 提出了地质灾害区域预警划分为隐式统计预警、显式统计预警和动力预警原理。简单总结了采用临界降雨判据方法(隐式统计)建立的中国第一代国家级地质灾害预警系统及 2003 ~ 2007 年期间的应用经验。通过系统建立包括 30 个图层的全国地质环境信息系统、因子分析确定权重、多元统计和人工神经网络模型对比, 对中国大陆分区建立了显式统计预警模型, 研发了第二代国际级地质灾害预警系统, 经过 2008 年汛期地质灾害区域预警预报服务检验证明效果良好。概括提出了地质灾害区域预警预报的工作程序及其基本要求。

本书可供从事地质灾害防治和防灾减灾公共管理、科学技术研究、应用服务和教学培训人员参考使用。

