

# 基于主成分分析的多参数岩体结构面 优势分组方法研究

董福瑞, 王述红, 侯钦宽

(东北大学 资源与土木工程学院, 辽宁 沈阳 110819)

**摘 要:** 对赋存于岩体内的结构面进行统计分析、分组聚类, 是进行岩体工程稳定性分析的基础。考虑到结构面的力学和水力学性质受产状、迹长、张开度、粗糙度和充填状态等诸多因素共同影响, 提出一种基于主成分分析的多参数岩体结构面优势分组方法。首先, 利用主成分分析法选取多参数岩体结构面优势分组的判据, 并计算其参与相似性度量的权重值。进而, 基于退火遗传算法搜寻模糊 C 均值聚类算法的全局最优初始聚类中心。最后, 以待分组结构面与聚类中心的距离加权和最小的方式建立目标函数, 实现多参数岩体结构面优势分组。对计算机模拟的 200 个结构面进行多参数优势分组并与其他方法对比, 计算结果表明, 该方法具有较高的分组精度。将该方法用于重庆市三环高速花阳隧道实测结构面多参数优势分组中, 分组结果合理、可靠, 进一步验证了该方法具有较好的工程应用价值。

**关 键 词:** 岩体结构面; 优势分组; 主成分分析; 退火遗传算法; 聚类分析

中图分类号: TU452

文献标识码: A

文章编号: 1000-7598 (2022) 09-2457-14

## Multi-parameter dominant grouping method of rock mass discontinuity based on principal component analysis

DONG Fu-rui, WANG Shu-hong, HOU Qin-kuan

(School of Resources and Civil Engineering, Northeastern University, Shenyang, Liaoning 110819, China)

**Abstract:** Statistical analysis and clustering of discontinuities existing in rock mass are the basis of rock mass engineering stability analysis. Considering that the mechanical and hydraulic properties of discontinuities are affected by many factors such as occurrence, trace length, opening, roughness, and filling state. A multi-parameter dominant grouping method of rock mass discontinuities based on principal component analysis is proposed. Firstly, the principal component analysis method is used to select the criterion of the dominant grouping and calculate the weight value of its participation in similarity measurement. Secondly, the global optimal initial clustering center of the fuzzy C-means clustering algorithm is searched using the annealing genetic algorithm. Lastly, the objective function is established by minimizing the weighted sum of the distance between the discontinuities to be grouped and the clustering center, achieving the dominant grouping of multi-parameter rock mass discontinuity. The 200 discontinuities simulated by the computer are divided into dominant groups using the multi-parameter method and compared with other methods. The results show that this method has higher grouping accuracy. The method is applied to the multi-parameter dominant grouping of the measured discontinuities of Huayang Tunnel of Chongqing Third Ring Expressway. The grouping results are reasonable and reliable, which further verify that the method has significant engineering application value.

**Keywords:** discontinuities of rock mass; dominant grouping; principal component analysis; annealing genetic algorithm; cluster analysis

## 1 引 言

自然界中的岩体是一种非均匀、非连续性介质,

由于长期的地质构造作用和其他因素扰动, 岩体内发育有大量规模和形态各异的结构面<sup>[1]</sup>。结构面是岩体内部的薄弱环节, 它的存在极大地降低了岩体

收稿日期: 2021-11-16

修回日期: 2022-05-06

基金项目: 国家自然科学基金 (No.U1602232, No.52004052); 中央高校基本科研业务费项目 (No.170108029, No.2101027)。

This work was supported by the National Natural Science Foundation of China (U1602232, 52004052) and the Fundamental Research Funds for the Central Universities (170108029, 2101027).

第一作者简介: 董福瑞, 男, 1996 年生, 博士研究生, 主要从事隧道工程施工及围岩稳定性分析方面的研究。E-mail: dongfurui123@126.com

通讯作者: 王述红, 男, 1969 年生, 博士, 教授, 博士生导师, 主要从事岩土工程相关的教学和科研工作。E-mail: shwangneu@126.com

的强度、整体性和抗渗性能,在很大程度上影响着岩体的稳定性<sup>[2-3]</sup>。因此,采取科学的手段对工程中获取的结构面数据进行统计分析、分组聚类,进而了解岩体内众多结构面的发育情况及物理力学性质,是进行岩体工程稳定性分析的重要环节。

将具有相同或类似性质的结构面归为同一组是对结构面进行统计分析的有效方法,这个过程也被称为优势分组。传统的优势分组方法是根据产状数据绘制节理玫瑰花图或极点等密度图,而后进行人工分组,这种方法虽然简单高效,但主观性较强,对工程人员的职业素养要求较高。Shanley 等<sup>[4]</sup>将一种较为成熟的聚类分析算法引入到优势分组中,算法通过求解目标函数的方式实现结构面优势组的自动化分,但是该方法中涉及的小球半径合理选择问题没有得到有效解决。Dershowitz 等<sup>[5]</sup>尝试用一种基于概率理论的迭代计算方法实现结构面优势分组,分组结果较为理想。Hammah 等<sup>[6]</sup>将一种经典的软聚类算法模糊 C 均值聚类算法(fuzzy C-means algorithm, 简称 FCM)用于结构面优势分组中,取得了较为满意的结果。该算法本质上是一种局部优化算法,本身存在一定缺陷,但提供了一种优势分组的新思路,无疑是一种有益的探索。此后,蔡美峰<sup>[7]</sup>、卢波<sup>[8]</sup>、宋腾蛟<sup>[9]</sup>、王述红<sup>[10]</sup>、侯钦宽<sup>[11]</sup>等将部分经典的现代优化算法应用到结构面优势分组中,均取得了良好的分组结果,极大地丰富了相关领域的理论成果。

但结构面的空间、几何和物理性质具有很强的随机性和复杂性,在实际勘测统计中,常用产状(倾向和倾角)、迹长、张开度、粗糙度、密度和充填状态等 10 余项指标定量或定性地描述结构面特征。仅考虑结构面的空间分布,以产状作为优势组划分的判据,无法进一步考虑结构面的几何和物理力学特性,其分组结果不够全面细致,方法存在一定的局限性。Behzad 等<sup>[12]</sup>考虑结构面多指标判据的重要性,采用 K 均值聚类算法验证了结构面多指标优势分组的优势及应用价值。Zhou 等<sup>[13]</sup>以结构面产状、张开度、粗糙度等多个变量作为分组判据,采用临近算法、K 均值聚类算法、模糊 C 均值聚类算法和向量算法进行优势分组。徐黎明<sup>[14]</sup>、宋腾蛟<sup>[15]</sup>、宋盛渊<sup>[16]</sup>等分别用变尺度混沌算法、人工蜂群算法和量子粒子群算法对结构面多指标优势分组进行了研究,并用计算机模拟数据和工程实测数据验证了算法的可行性,但在样本相似性度量方法上都认为所有指标判据对分组的贡献率相同。然而,受采集数据分布规律及选取指标数量等的影响,忽略各指标

对分组的贡献权重对结构面进行聚类分组,可能会导致分组结果不理想,也会极大地增加分组的计算量,不利于方法的工程应用。因此,合理地分配各指标在优势分组中的贡献权重,并以这些指标为判据,实现结构面优势组的有效划分具有较大的研究价值及工程实践意义。

为了克服上述问题,本文提出一种基于主成分分析的多参数岩体结构面优势分组方法。采用主成分分析法对结构面原始样本数据进行降维分析,根据累计贡献率大于 85% 的原则,选取若干主成分作为优势分组的综合指标判据,计算并为各综合指标判据赋予权重,将寻优能力更强的退火遗传算法与模糊 C 均值聚类算法结合,实现对结构面优势组的客观划分,最后通过计算机模拟结构面数据和重庆市三环高速花阳隧道实测结构面数据验证了算法的可行性。

## 2 问题的数学描述

### 2.1 结构面表征指标的选取

反映岩体结构面特性的指标众多,本文以 6 个常用指标为例,对本方法的实现过程及应用效果进行说明。选取的指标分别为:倾向  $\alpha$  ( $0^\circ \leq \alpha \leq 360^\circ$ )、倾角  $\beta$  ( $0^\circ \leq \beta \leq 90^\circ$ )、迹长 TL、张开度  $e$ 、粗糙度  $c$  和充填状态  $f$ 。前 4 个指标可以通过工程量测的方式直接获取,结构面的粗糙度通常选用 Barton 提出的粗糙度系数进行定量表征<sup>[17-18]</sup>,但其在工程应用中存在计算难度大的问题<sup>[19]</sup>。为了便于后续分析计算,本文依据现场量测资料对结构面充填状态和粗糙度进行简便量化表征,结构面充填状态和粗糙度量化表征取值见表 1 和表 2。

对于结构面产状的描述,通常把结构面假想为一个无限延伸的空间理想平面<sup>[20]</sup>,用平面的单位法向量  $\mathbf{X} = (x, y, z)$  表示其空间方位,这样可以将单位法向量  $\mathbf{X}$  投影到一个单位圆上,便于对结构面的整体方位模式进行直观分析。对于单位圆上一点  $A$ ,

表 1 结构面充填状态简便量化取值  
Table 1 Simple quantitative values of filling state of discontinuities

| 充填状态  |      | 量化值         |
|-------|------|-------------|
| 充填物类别 | 充填度  |             |
| 黏土    | 一般填充 | 0.000~0.200 |
| 黏土    | 密实填充 | 0.210~0.400 |
| 砂土    | 一般填充 | 0.410~0.600 |
| 砂土    | 密实填充 | 0.610~0.800 |
| 岩屑    | 一般填充 | 0.810~1.000 |
| 岩屑    | 密实填充 | 1.010~1.200 |

表 2 结构面粗糙度简便量化取值

Table 2 Simple quantitative values of surface morphology of discontinuities

| 表面形态 |      | 量化值         |
|------|------|-------------|
| 起伏形态 | 粗糙程度 |             |
| 平直   | 光滑   | 0.000~0.500 |
| 平直   | 稍粗   | 0.510~1.000 |
| 平直   | 粗糙   | 1.010~1.500 |
| 波状   | 光滑   | 1.510~2.000 |
| 波状   | 稍粗   | 2.010~2.500 |
| 波状   | 粗糙   | 2.510~3.000 |
| 锯齿状  | 光滑   | 3.010~3.500 |
| 锯齿状  | 稍粗   | 3.510~4.000 |
| 锯齿状  | 粗糙   | 4.010~4.500 |

其空间位置可以由倾向 $\alpha$ 和倾角 $\beta$ 唯一确定：

$$\left. \begin{aligned} x_A &= \cos \alpha \sin \beta \\ y_A &= \sin \alpha \sin \beta \\ z_A &= \cos \beta \\ x_A^2 + y_A^2 + z_A^2 &= 1 \end{aligned} \right\} \quad (1)$$

综上所述，采用上述 6 个指标可以将某一特定结构面客观地表述为  $P = [X, TL, e, c, f]$ 。

## 2.2 主成分分析

存在于同一地理区域，由相同地质构造作用形成的结构面，其诸多性质是相近甚至是相同的，把表征这些性质的指标作为优势组划分判据的意义不大。主成分分析法<sup>[21]</sup>是解决复杂多元问题的一种有效手段，能够实现高维数据特征降维的目的，从而节省大量的计算时间和成本。

大量研究及工程实践表明，产状（倾向和倾角）数据在实现结构面优势分组中起到了不可替代的作用<sup>[22-27]</sup>。因此，本文没有将结构面产状指标纳入主成分分析的数据范围。同时，为了消除量纲的影响，进行主成分分析前，首先要对数据进行标准化处理。将除产状外的其余 4 项指标的原始数据集作为主成分分析待处理数据，利用下式对原始数据集进行标准化处理：

$$x_{la} = \frac{x'_{la} - \bar{x}_a}{\sigma_a} \quad (a=1, 2, 3, 4 \quad l=1, 2, \dots, n) \quad (2)$$

式中： $x'_{la}$  为标准化前第  $l$  个结构面的第  $a$  项指标； $x_{la}$  为标准化后第  $l$  个结构面的第  $a$  项指标； $\bar{x}_a$ 、 $\sigma_a$  分别为样本第  $a$  项指标的均值和标准差； $n$  为结构面的样本总数。

标准化处理后，结构面用上述 4 项指标可表述为  $P = [x_1 \ x_2 \ x_3 \ x_4]$ ，其协方差矩阵为  $R$ ：

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \\ r_{41} & r_{42} & r_{43} & r_{44} \end{bmatrix} \quad (3)$$

$$r_{a_1 a_2} = \frac{1}{n-1} \sum_{l=1}^n x_{la_1} x_{la_2} \quad (4)$$

式中： $x_{la_1}$  和  $x_{la_2}$  分别为第  $l$  个结构面的第  $a_1$  和  $a_2$  项指标； $r$  为协方差矩阵  $R$  的元素。

$\lambda = [\lambda_1 \ \lambda_2 \ \lambda_3 \ \lambda_4]$  和  $v = [v_1 \ v_2 \ v_3 \ v_4]$  分别为协方差矩阵  $R$  的特征值和特征向量。其中， $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 \geq 0$ 。则在新坐标空间下，结构面综合指标判据  $y$  可表示为

$$y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} v_{11} & v_{12} & v_{13} & v_{14} \\ v_{21} & v_{22} & v_{23} & v_{24} \\ v_{31} & v_{32} & v_{33} & v_{34} \\ v_{41} & v_{42} & v_{43} & v_{44} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \quad (5)$$

$$y_k = v_{k1}x_1 + v_{k2}x_2 + v_{k3}x_3 + v_{k4}x_4 \quad (k=1, 2, 3, 4) \quad (6)$$

式中： $y_a$  为新的坐标空间中第  $a$  项综合指标判据；当前  $t$  ( $1 \leq t \leq 4$ ) 项综合指标判据的累计贡献率  $\alpha_t \geq 85\%$  时，认为这些综合指标判据的数据信息代表了结构面原始数据集的主要信息。因此，在保证数据信息完整性的前提下，选取前  $t$  项综合指标  $y_1, \dots, y_t$  及产状指标  $X$  作为优势分组的指标判据。需要注意的是，结构面的产状数据由其单位法向量表示，因此无需重复进行标准化操作。贡献率  $b_a$  及累计贡献率  $\alpha_t$  按下式计算：

$$b_a = \frac{\lambda_a}{\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4} \quad (a=1, 2, 3, 4) \quad (7)$$

$$\alpha_t = \sum_{a=1}^t b_a \quad (t \leq 4) \quad (8)$$

## 2.3 相似性度量

自然界岩体中存在着倾角接近  $90^\circ$  的高陡倾角结构面，当它们的倾向彼此相差约为  $180^\circ$  时，也应划分为同一组（如图 1 所示）。传统的欧氏距离无法很好地解决上述高陡倾角结构面分组问题<sup>[28]</sup>。鉴于此，本文选取两结构面夹角（锐角）正弦平方距离来衡量样本间产状的相似性，采用欧氏距离衡量样本综合指标判据  $y_a$  的相似性。对于任意两结构面  $P_i (X_i, y_{i1}, y_{i2}, \dots, y_{it})$  和  $P_j (X_j, y_{j1}, y_{j2}, \dots, y_{jt})$ ，其中  $X_i = (x_i, y_i, z_i)$ ， $X_j = (x_j, y_j, z_j)$  分别为结构面  $P_i$  和结构面  $P_j$  对应的单位法向量，两结构面所夹锐角为  $\gamma$ ，则两结构面产状的相似性距离和综合指标判据的相

似性距离  $d$  分别为

$$d^2(X_i, X_j) = \sin^2 \gamma = 1 - (X_i \cdot X_j^T)^2 \quad (9)$$

$$d^2(y_{it}, y_{jt}) = (y_{it} - y_{jt})^2 \quad (10)$$

式中:  $y_{it}$ 、 $y_{jt}$  分别为第  $i$  和第  $j$  个结构面的第  $t$  项综合指标判据;  $d$  为两结构之间的相似距离。

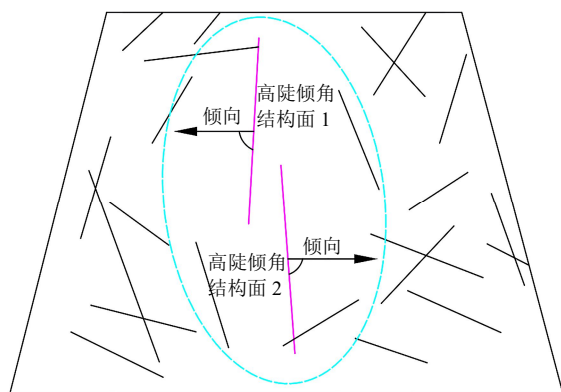


图 1 高陡倾角结构面示意图

Fig.1 Schematic diagram of extremely steep dip discontinuities

提出一种基于客观赋权的多参数优势分组相似性度量方法, 考虑到结构面产状判据在优势组划分中起不可替代的作用, 且结构面产状判据不参与主成分分析, 在默认结构面产状判据权重为 1 的基础上, 按贡献率的大小分别赋予各综合指标判据不同的权重  $m_a$  ( $1 \leq a \leq t$ ),  $0 \leq m_a \leq 1$ 。最后, 以各指标判据的距离加权和计算两结构面样本的相似性:

$$d^2(P_i, P_j) = D \cdot M^T = d^2(X_i, X_j) + m_1 d^2(y_{i1}, y_{j1}) + L + m_a d^2(y_{ij}, y_{jj}) + L + m_t d^2(y_{it}, y_{jt}) \quad (11)$$

式中:  $D$  为不考虑权重的两结构面相似性距离;  $M = [1 \ m_1 \ \cdots m_a \ \cdots m_t]$  为结构面产状和前  $t$  项综合指标判据的相似性度量权重矩阵。各综合指标判据权重  $m_a$  的计算方法为

$$m_a = \frac{b_a}{\alpha_t} \quad (1 \leq a \leq t) \quad (12)$$

## 2.4 目标函数

假设  $n$  个待分组结构面  $P_l$  ( $l = 1, 2, \cdots, n$ ), 将其划分为  $c$  ( $2 \leq c \leq n$ ) 个互不相交的组,  $C = [c_1 \ c_2 \ \cdots c_c]^T$  为  $c$  个组所对应的聚类中心矩阵,  $U$  为其隶属度矩阵, 结构面分组效果的好坏通过目标函数  $J(U, C)$  体现, 目标函数定义为每个结构面与各聚类中心的距离加权之和:

$$\min J(U, C) = \min \left\{ \sum_{k=1}^c \sum_{l=1}^n (\mu_{kl})^b d(P_l, c_k) \right\} \quad (13)$$

$$\sum_{k=1}^c \mu_{kl} = \sum_{k=1}^c \mu_{kl} = 1 \quad (l = 1, 2, L, n)$$

式中:  $\mu_{kl}$  为对于任一结构面  $P_l$  隶属于第  $k$  个组的隶属度值, 其取值范围为  $0 \leq \mu_{kl} \leq 1$ ;  $d(P_l, c_k)$  为结构面  $P_l$  与第  $k$  个聚类中心  $c_k$  之间的距离;  $b \in [1, +\infty)$  为模糊加权指数,  $b$  的取值决定样本数据的隶属度关系,  $b$  的取值越大, 隶属度关系越模糊, 本文取  $b = 2$ 。

## 3 多参数岩体结构面优势分组

### 3.1 退火遗传算法

遗传算法 (genetic algorithm, 简称 GA) 是一种现代群智能优化算法<sup>[29]</sup>, 凭借着计算速度快、鲁棒性强等优势在包括岩土工程在内的诸多领域得到广泛应用, 但算法存在的“早熟”现象一直未能得到有效解决。模拟退火算法 (simulated annealing, 简称 SA) 也是一种在工程上被大量应用的优化算法, Metropolis 准则允许算法以一定的概率接受一个非局部最优解, 这一操作大幅增加了算法的局部搜索能力。遗传算法和模拟退火算法在工程上的应用已经十分成熟, 基本原理本文不再详细赘述。将模拟退火算法与遗传算法相结合, 提出一种高效、稳定的全局寻优退火遗传算法, 并用以优化传统模糊 C 均值聚类算法的初始聚类中心是本文的工作重点。

### 3.2 模糊 C 均值聚类算法

模糊 C 均值聚类算法是目前较为常见的一种软聚类算法。它最早由 Ruspini<sup>[30]</sup> 和 Zadeh<sup>[31]</sup> 提出, 后经 Bezdek<sup>[32]</sup> 在硬 C 均值 (hard C-means, 简称 HCM) 算法的基础上不断改进形成。算法因设计简单, 计算和收敛速度较快等优点在诸多领域取得了较为成功的应用。但从本质上讲, 该算法仍属于局部优化算法, 需要提前给定分组数目和初始聚类中心, 如果初值选取不当, 算法容易陷入局部最小点。

算法具体实施过程为: 运用拉格朗日乘数法求目标函数 (13) 的最小值, 通过给定的初始聚类中心, 不断交替迭代计算新的聚类中心矩阵  $C$  和隶属度矩阵  $U$ , 直到算法收敛, 最后按照隶属度最大原则实现数据集的自动分组。聚类中心矩阵和隶属度矩阵更新公式为

$$\mu_{kl} = \frac{1}{\sum_{j=1}^c \left( \frac{d(P_l, c_k)}{d(P_l, c_j)} \right)^{\frac{2}{b-1}}} \quad (14)$$

$$c_k = \frac{\sum_{l=1}^n \mu_{kl}^b P_l}{\sum_{l=1}^n \mu_{kl}^b} \quad (15)$$

综上所述, 本文提出一种基于主成分分析的多参数岩体结构面优势分组方法 (FCM algorithm optimized by principal component analysis and genetic simulated annealing algorithm, 简称 PGS-FCM), 方法执行流程如下:

(1) 采集并整理表征岩体结构面性质的 6 项指标数据, 对除倾向和倾角外其他指标数据进行标准化处理, 并对标准化后的数据进行 KMO (Kaiser-Meyer-Olkin) 和 Bartlett 球形检验, 满足检验条件则执行步骤 (2), 否则在计算指标贡献率并赋予权值后直接执行步骤 (3);

(2) 对标准化后数据进行主成分分析, 选取累计贡献率  $\alpha_i \geq 85\%$  的前  $t$  项主成分代替原有指标判据作为综合指标判据, 并按贡献率大小赋予权值;

(3) 给定最大迭代次数  $L$ 、初始温度  $T_0$ 、冷却系数  $q$ 、终止温度  $T_{\text{end}}$ 、种群大小  $S$ 、最大遗传代数  $G_{\text{max}}$ 、选择概率  $p_s$ 、交叉概率  $p_c$ 、变异概率  $p_m$  和目标函数终止容限  $\Delta J$ ;

(4) 初始化种群, 这里指随机生成  $S$  个初始聚类中心矩阵  $C$ , 每个初始聚类中心矩阵中含有  $c$  个聚类中心的坐标  $\{c_k\}$ , 对于每个聚类中心按照式 (14) 和式 (13) 分别计算样本隶属度矩阵  $U$  和个体目标函数值  $J_l$ , 其中  $l = 1, 2, \dots, S$ ;

(5) 对旧个体进行选择、交叉、变异运算产生新个体, 对于新个体用式 (15) 和式 (14) 更新聚类中心矩阵  $C_{\text{new}}$  和样本隶属度矩阵  $U_{\text{new}}$ , 并计算每个新个体的目标函数值  $J'_l$ ;

(6)  $\Delta J = J'_l - J_l$ , 若  $\Delta J \leq 0$ , 则以新个体替换旧个体, 否则按 Metropolis 准则接受新个体;

(7) 若种群当前代数  $\text{gen} < G_{\text{max}}$ , 则  $\text{gen} = \text{gen} + 1$ , 执行步骤 (5), 否则执行步骤 (8);

(8) 若当前温度  $T_j < T_{\text{end}}$ , 则输出当前最优解, 执行步骤 (9), 否则执行降温操作  $T_{j+1} = qT_j$ , 执行步骤 (5);

(9) 将最优初始聚类中心矩阵  $C_{\text{best}}$  作为模糊 C 均值聚类算法的初始聚类中心, 结构面产状数据和主成分分析选取综合指标判据后的数据作为新的待分组数据集, 应用模糊 C 均值聚类算法进行优势分组。

需要说明的是, 工程上常以结构面产状作为优势分组依据, 为满足工程习惯和施工需要, 按照本

文的方法进行结构面多指标优势分组时, 建议分两阶段具体实施。首先, 取权重向量  $M = [1 \ 0 \ 0 \ 0 \ 0]$ , 即先按产状指标进行优势分组; 而后, 基于主成分分析法计算各综合判据指标权重值, 取权重向量  $M = [0 \ m_1 \ m_2 \ m_3 \ m_4]$ , 在按产状分组的基础上, 进一步考虑结构面几何、物理和力学特性对分组结果进行细化, 最终分组结果更加符合客观实际。

### 3.3 聚类有效性评价

没有预先明确分组数时, 需要根据有效性指标合理的选取最佳分组数, 通常采用组内紧致度和组间离散度两个参数来表征分组有效性指标<sup>[24]</sup>。预先给定不同的划分组数 (一般取 2~10), 当有效性指标取值最优时相应的分组结果就是最佳分组结果。本文选取分离度  $V_{\text{PC}}$ 、分离熵  $V_{\text{PE}}$ 、Xie-Beni 指标  $V_{\text{XB}}$  和 Fukuyama-Sugeno 指标  $V_{\text{FS}}$  这 4 种评价指标对优势组划分的有效性进行检验, 进而确定最佳组数。4 种评价指标的计算公式如下:

$$V_{\text{PC}} = \frac{1}{n} \sum_{k=1}^c \sum_{l=1}^n \mu_{kl}^b \quad (16)$$

$$V_{\text{PE}} = -\frac{1}{n} \sum_{k=1}^c \sum_{l=1}^n \mu_{kl} \ln \mu_{kl} \quad (17)$$

$$V_{\text{XB}} = \frac{\sum_{k=1}^c \sum_{l=1}^n \mu_{kl}^b d(P_l, c_k)}{n \min_{j \neq k} d(c_j, c_k)} \quad (18)$$

$$V_{\text{FS}} = \sum_{k=1}^c \sum_{l=1}^n \mu_{kl}^b [d(P_l, c_k) - d(c_k, \bar{c})] \quad (19)$$

式中:  $\bar{c}$  为所有聚类中心的平均值;  $d(c_k, \bar{c})$  为第  $k$  个聚类中心与所有聚类中心的平均值  $\bar{c}$  的距离。有效性指标  $V_{\text{PC}}$  的值越大, 说明优势分组效果越好, 有效性指标  $V_{\text{PE}}$ 、 $V_{\text{XB}}$  和  $V_{\text{FS}}$  的值越小, 说明优势分组效果越好, 综合考虑上述 4 个有效性评价指标, 可以确定出最佳的结构面优势分组数。

## 4 算法检验

将文献[14]中计算机模拟的 5 组共计 200 个结构面数据作为测试数据, 验证 PGS-FCM 方法的准确性。模拟数据中, 结构面倾向和倾角服从双变量正态分布, 迹长服从正态分布, 张开度服从均匀分布, 粗糙度为给定量值。此外, 本文在文献[14]的基础上, 加入 200 个给定量值的充填状态数据。计算机模拟数据详细参数见表 3。

根据计算机模拟结构面数据, 绘制结构面等密度图和极点图如图 2 和图 3 所示。

表 3 结构面模拟参数表

Table 3 Parameters table of simulated distributions

| 组号 | 双正态分布  |     |         |        |    |       | 正态分布 |      |           | 均匀分布   |           | 粗糙度 | 充填状态 | 个数 |
|----|--------|-----|---------|--------|----|-------|------|------|-----------|--------|-----------|-----|------|----|
|    | 倾向/(°) |     |         | 倾角/(°) |    |       | 迹长/m |      |           | 张开度/cm |           |     |      |    |
|    | 均值     | 方差  | 范围      | 均值     | 方差 | 范围    | 均值   | 方差   | 范围        | 均值     | 范围        | 量化值 | 量化值  |    |
| 1  | 135    | 80  | 115~164 | 55     | 20 | 47~67 | 0.5  | 0.10 | 0.23~1.21 | 1.0    | 0.8~1.18  | 1.5 | 0.4  | 70 |
| 2  | 300    | 120 | 281~326 | 30     | 24 | 19~41 | 2.0  | 0.40 | 0.63~3.50 | 1.2    | 1.11~1.30 | 2.0 | 0.6  | 40 |
| 3  | 30     | 50  | 11~44   | 70     | 40 | 50~86 | 4.0  | 0.45 | 2.36~5.52 | 0.5    | 0.31~0.70 | 1.0 | 0.2  | 40 |
| 4  | 55     | 120 | 29~83   | 65     | 40 | 56~81 | 1.0  | 0.18 | 0.33~1.69 | 2.0    | 1.81~2.09 | 2.5 | 0.8  | 30 |
| 5  | 255    | 40  | 238~271 | 40     | 30 | 28~51 | 6.0  | 0.40 | 5.26~7.51 | 0.2    | 0.15~0.25 | 3.0 | 1.0  | 20 |

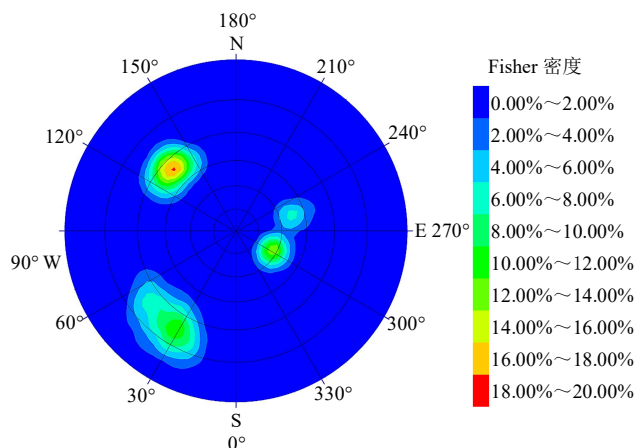


图 2 计算机模拟结构面等密度图

Fig.2 Contour of pole density of artificial discontinuities

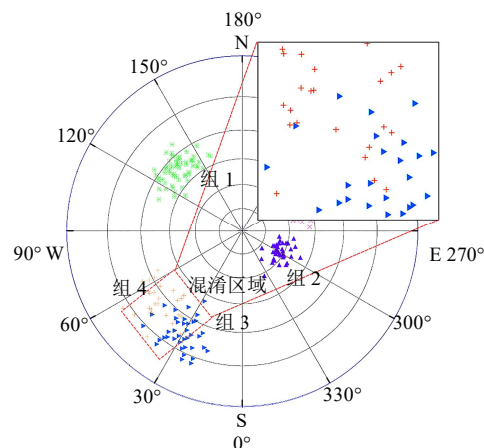


图 4 组间混淆区域示意图

Fig.4 Schematic diagram of intergroup confusion area

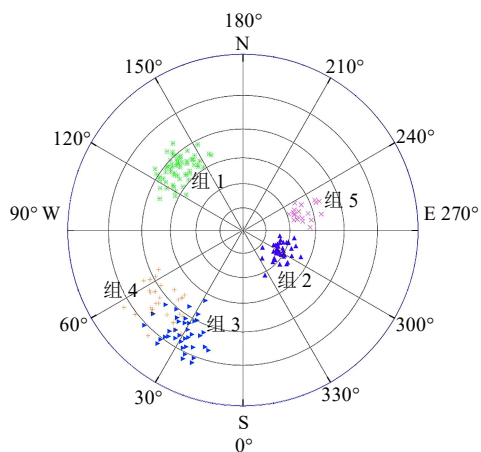


图 3 计算机模拟结构面极点图

Fig.3 Pole diagram of artificial discontinuities

观察可知,依据产状指标判据,容易将 200 个结构面划分为 3 组界限分明的优势组。由图 3 可知,在已知优势分组数为 5 组的情况下,这种仅以产状指标作为判据的优势分组方法与实际不符。此外,由图 4 可以看出,计算机模拟数据组间交界处存在混淆区域,仅以产状作为判据难以对混淆区域内的结构面进行正确分组,这也说明了结构面多参数优势分组的必要性。

根据本文提出的方法,用式(2)对迹长、张开度、粗糙度和充填状态 4 项指标判据的原始数据进行标准化处理。而后,对计算机模拟结构面除产状外的指标数据进行主成分分析。经计算, KMO 检验系数为 0.677 ( $>0.5$ ), Bartlett 球形度检验的统计值显著性概率 sig 为 0.000 ( $<0.05$ ),各指标判据在一定程度上相互独立,可以采用主成分分析法进行数据降维及特征选择。如图 5 所示,计算机模拟结构面数据中,张开度和迹长 2 项指标判据在优势分组中作用较为明显,粗糙度和充填状态 2 项指标判据作用较弱,但仍具有一定的参考意义。

进一步,当综合指标判据数取 2 时,其特征值为 1.488 ( $\geq 1$ ),累计方差贡献率为 92.866% ( $>85\%$ ),故选取前 2 项综合指标判据数据代替全部结构面数据参与计算。从综合指标判据矩阵图 6 可以清晰地看出,以前 2 项综合指标作为结构面优势分组的判据,分组效果较好。

PGS-FCM 方法相关参数设置为:最大迭代次数  $L=20$ ,初始温度  $T_0=1\ 000\ ^\circ\text{C}$ ,冷却系数  $q=0.9$ ,终止温度  $T_{\text{end}}=0.1\ ^\circ\text{C}$ ,种群大小  $S=10$ ,最大遗传代数  $G_{\text{max}}=10$ ,选择概率  $p_s=0.9$ ,交叉概率  $p_c=0.7$ ,

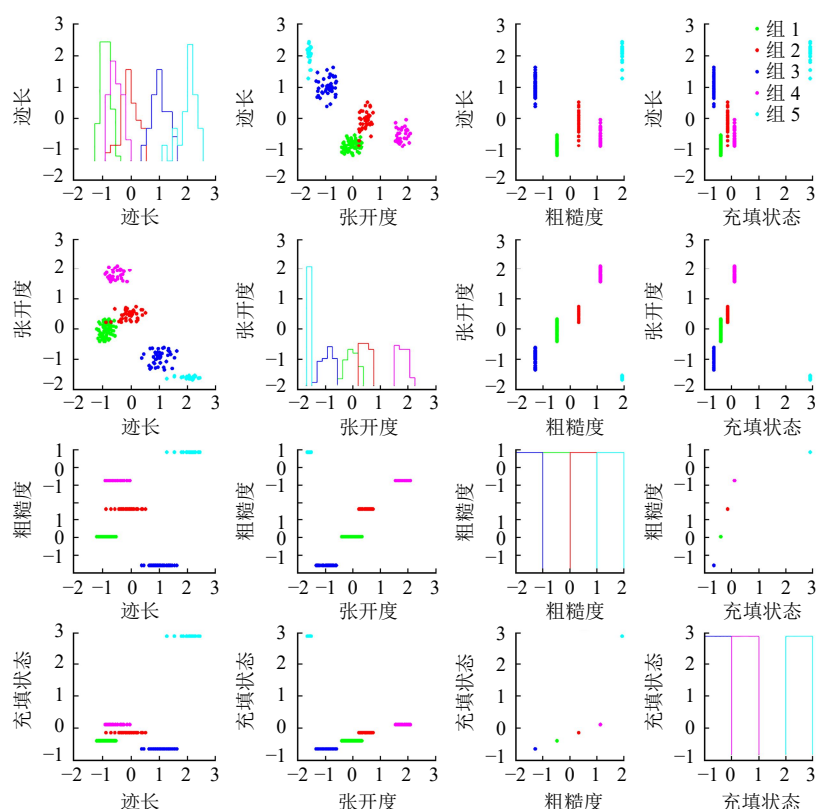


图5 指标判据矩阵图

Fig.5 Matrix diagram of index criterion

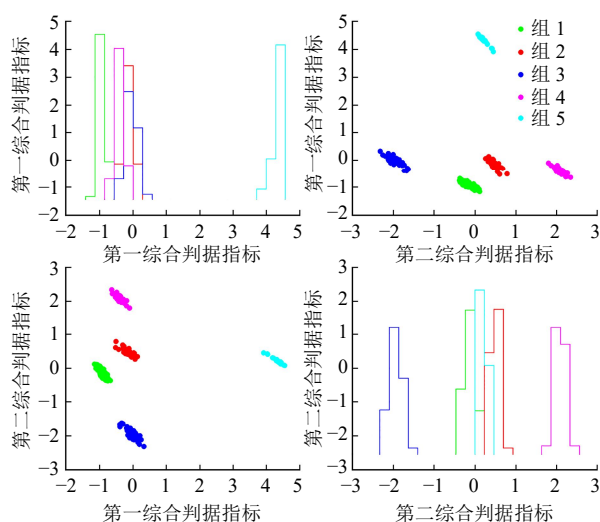


图6 综合指标判据矩阵图

Fig.6 Matrix diagram of comprehensive index criterion

变异概率  $p_m = 0.01$ , 目标函数终止容限  $\Delta J = 1 \times 10^{-6}$ 。取相似性度量权重矩阵  $M = [1 \ 0 \ 0]$ , 此时仅考虑以结构面产状作为优势分组的判据, 由图 2 可知, 计算机模拟结构面测试数据按产状可以分为 3 类界限明显的优势组, 因此取分组数为 3, 不再进行评价指标试算。PGS-FCM 方法计算结果如图 7 所示。图中 1[70] 表示组名为 1 的结构面, 该组结构面数量

为 70, 以此类推。

由图 7 可知, 以结构面产状为判据可将计算机模拟结构面划分为 3 个优势组, 将 3 个优势组数据提取分别作为新的待分组结构面数据, 由式 (12) 计算第一综合指标判据和第二综合指标判据的权重  $m_1$  和  $m_2$  分别为 0.599 和 0.401, 取权重向量  $M = [0 \ 0.599 \ 0.401]$ , 由计算机模拟结构面数据已知分组结果可知, 优势分组数目分别取 1、2、2, 其他参数不变, 采用 PGS-FCM 方法对 3 组待分组结构面进行优势分组, 最终计算结果如图 8 所示。为进一步说明本方法的优势, 采用不考虑综合指标判据权重的 PGS-FCM 方法 (PGS-FCM method without considering weight, 简称 WCPGS-FCM, 即取  $M = [1 \ 1 \ 1]$ ) 和仅以产状为判据的传统 FCM 方法分别对 200 个计算机模拟结构面进行优势分组, 计算结果如图 9 和 10 所示。

200 个计算机模拟结构面采用 PGS-FCM 方法进行优势分组, 分组结果较为理想, 且该方法对第 3 组和第 4 组组间混淆区域内的结构面实现了准确划分, 200 个结构面分组结果与已知结果完全相同。WCPGS-FCM 方法在优势分组时没有赋予指标判据合理的权值, 编号为 98 的结构面被错分到了第 1 组, 整体分组误差为 0.5%。查看原数据可知, 编号

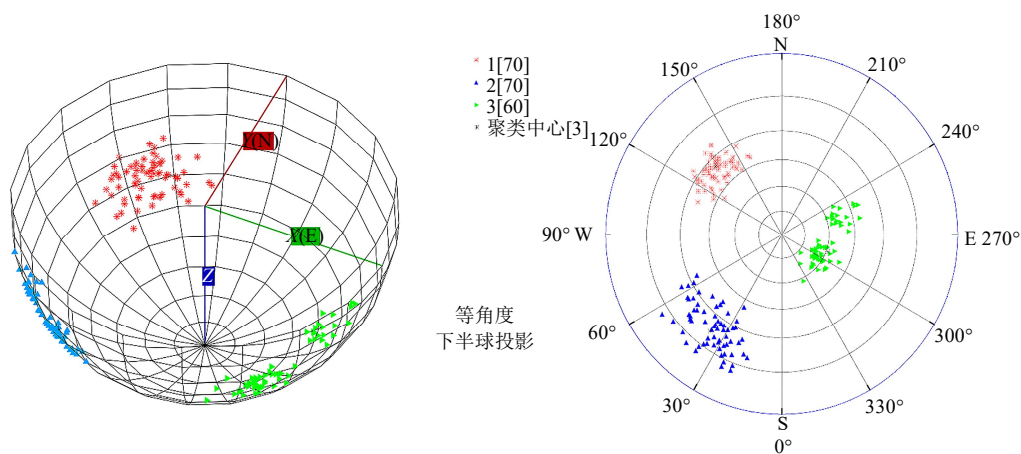


图 7 计算机模拟结构面按产状判据分组 Schmidt 下半球投影和极点图

Fig.7 Schmidt lower hemisphere projection map and pole diagram map grouped according to the occurrence discrimination basis for artificial discontinuities

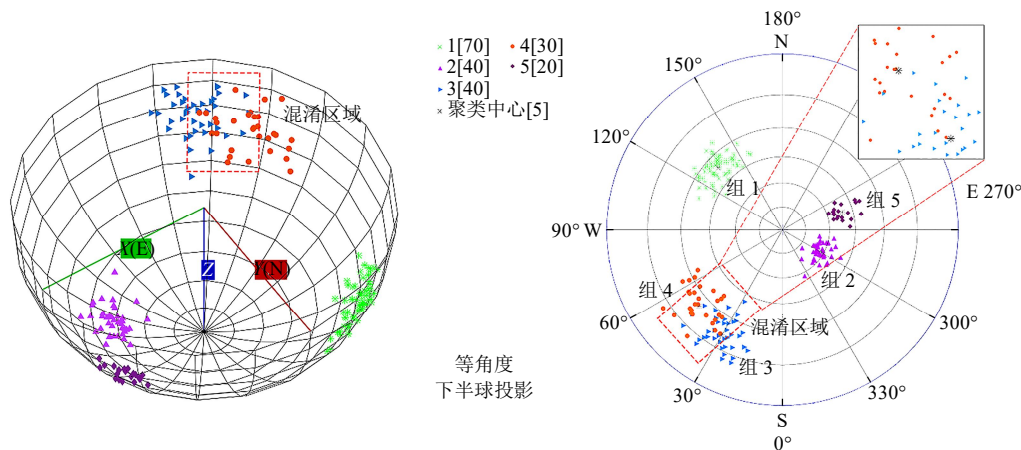


图 8 计算机模拟结构面采用 PGS-FCM 方法分组 Schmidt 下半球投影和极点图

Fig.8 Schmidt lower hemisphere projection map and pole diagram map grouped by PGS-FCM method for artificial discontinuities

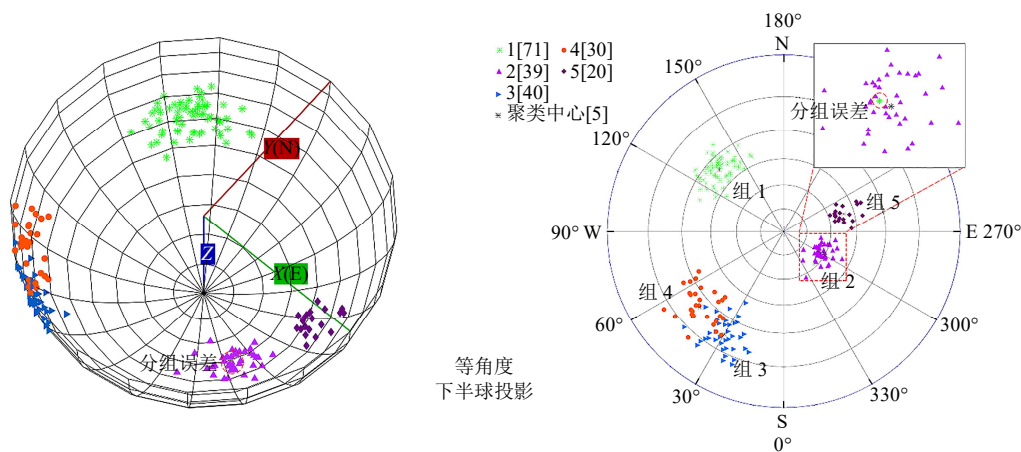


图 9 计算机模拟结构面采用 WCPGS-FCM 方法分组 Schmidt 下半球投影和极点图

Fig.9 Schmidt lower hemisphere projection map and pole diagram map grouped by WCPGS-FCM method for artificial discontinuities

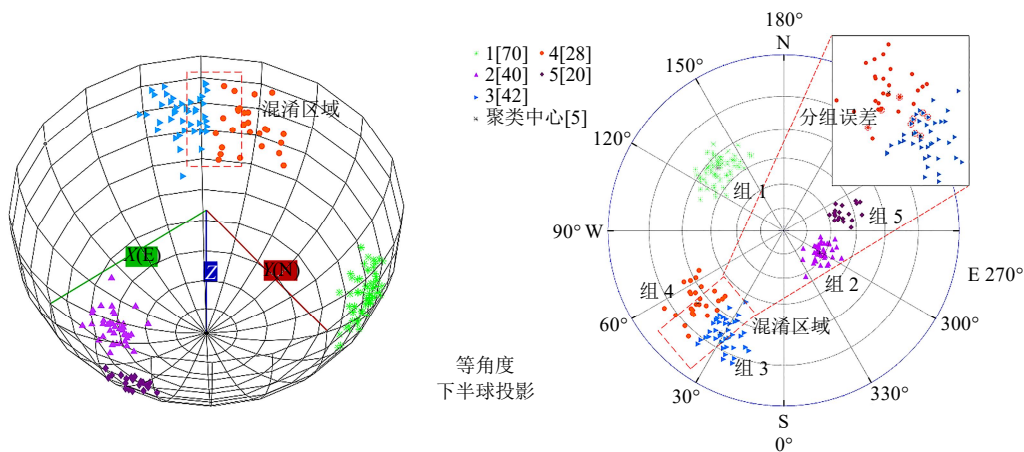


图 10 计算机模拟结构面采用 FCM 方法分组 Schmidt 下半球投影和极点图

Fig.10 Schmidt lower hemisphere projection map and pole diagram map grouped by FCM method for artificial discontinuities

为 98 的结构面可以表征为  $P_{98} = [300.452 \ 26.350 \ 0.511 \ 1.120 \ 2.000 \ 0.600]$ 。在不考虑各指标权重时,该结构面与第 1 组聚类中心的迹长、张开度、粗糙度和充填状态 4 项指标均较为接近,因此被错分到第 1 组。需要说明的是,由于计算机模拟 5 组数据组间间隔较为明显,因此在不考虑指标权重的情况下进行多参数优势分组结果较为理想,若数据组间界限较为模糊,则分组结果将出现较大误差。以产状为判据的传统 FCM 方法分组误差为 4.5%, FCM

方法不能对第 3 组和第 4 组组间混淆区域内的结构面进行准确区分,这也表明以产状单一指标为判据进行优势分组无法进一步考虑结构面的几何和物理特性,方法存在一定缺陷。

此外,对比 3 种方法,PGS-FCM 方法计算出的聚类中心与已知聚类中心吻合度最好, WCPGS-FCM 方法次之, FCM 方法计算出的聚类中心误差较大,这也验证了本方法的可靠性。3 种优势分组方法的详细分组结果见表 4。

表 4 模拟参数与分组结果对比

Table 4 Comparison of simulation parameters and grouping results

| 组别 | 聚类中心      | 倾向/(°)  | 倾角/(°) | 迹长/m  | 张开度/cm | 粗糙度   | 填充状态  | 聚类数量 |
|----|-----------|---------|--------|-------|--------|-------|-------|------|
| 1  | 已知聚类中心    | 135.000 | 55.000 | 0.500 | 1.000  | 1.500 | 0.400 | 70   |
|    | PGS-FCM   | 134.524 | 54.997 | 0.500 | 1.061  | 1.503 | 0.401 | 70   |
|    | WCPGS-FCM | 134.488 | 54.945 | 0.372 | 0.996  | 1.503 | 0.397 | 71   |
|    | FCM       | 134.684 | 55.726 | —     | —      | —     | —     | 70   |
| 2  | 已知聚类中心    | 300.000 | 30.000 | 2.000 | 1.200  | 2.000 | 0.600 | 40   |
|    | PGS-FCM   | 299.997 | 28.643 | 2.000 | 1.222  | 2.000 | 0.600 | 40   |
|    | WCPGS-FCM | 298.487 | 28.589 | 2.002 | 1.256  | 1.983 | 0.601 | 39   |
|    | FCM       | 302.053 | 27.591 | —     | —      | —     | —     | 40   |
| 3  | 已知聚类中心    | 30.000  | 70.000 | 4.000 | 0.500  | 1.000 | 0.200 | 40   |
|    | PGS-FCM   | 30.650  | 69.274 | 4.064 | 0.518  | 1.000 | 0.200 | 40   |
|    | WCPGS-FCM | 30.611  | 68.833 | 4.053 | 0.526  | 1.001 | 0.204 | 40   |
|    | FCM       | 32.064  | 69.173 | —     | —      | —     | —     | 42   |
| 4  | 已知聚类中心    | 50.000  | 65.000 | 1.000 | 2.000  | 2.500 | 0.800 | 30   |
|    | PGS-FCM   | 48.902  | 65.005 | 1.113 | 1.988  | 2.500 | 0.800 | 30   |
|    | WCPGS-FCM | 48.362  | 65.105 | 1.123 | 1.957  | 2.495 | 0.792 | 30   |
|    | FCM       | 50.728  | 65.977 | —     | —      | —     | —     | 28   |
| 5  | 已知聚类中心    | 255.000 | 40.000 | 6.000 | 0.200  | 3.000 | 1.000 | 20   |
|    | PGS-FCM   | 254.000 | 40.256 | 5.953 | 0.193  | 2.999 | 1.000 | 20   |
|    | WCPGS-FCM | 253.825 | 39.587 | 5.913 | 0.199  | 2.997 | 1.000 | 20   |
|    | FCM       | 259.053 | 37.391 | —     | —      | —     | —     | 20   |

## 5 工程实例

重庆三环高速公路合川至长寿段第 HC04 标段花阳隧道位于重庆市合川区, 隧道区属于侵蚀剥蚀台地地貌区, 地形起伏较大, 斜坡较陡, 隧址区第四系覆盖层主要为更新统坡积成因粉质黏土, 分布不均匀, 部分地段基岩出露, 出露地层岩性为侏罗纪中统上沙溪庙组泥岩和砂岩等。隧址区砂岩和泥岩节理裂隙发育, 岩体破碎, 含少量裂隙水, 浅部连通性较好, 赋水条件一般。应用本文提出的方法, 对花阳隧道隧址区 158 个结构面现场实测数据进行多参数优势分组。图 11 为花阳隧道实测结构面叠覆等值线极点图。

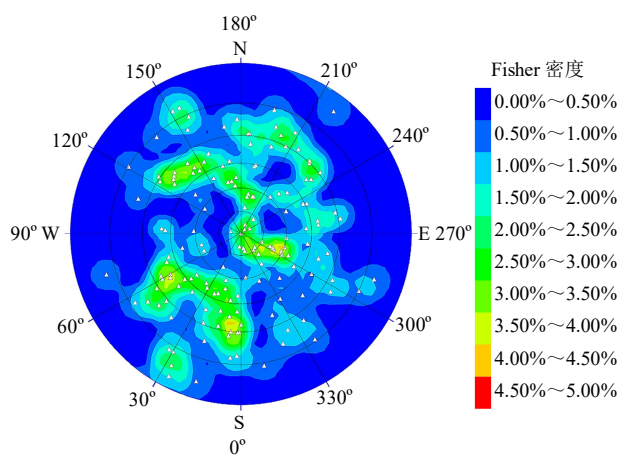


图 11 花阳隧道实测结构面叠覆等值线极点图  
Fig.11 Pole diagram of overlapping isoline of measured discontinuities of Huayang Tunnel

观察可知, 仅依据 158 个实测结构面的产状难以确定最佳分组数, 且结构面样本的迹长、张开度、粗糙度和充填状态之间也存在较大差异, 对分组结

果势必产生较大影响。按照本文所提方法, 对花阳隧道实测结构面除产状外的数据进行标准化处理, KMO 检验系数为 0.775 ( $>0.5$ ), Bartlett 球形度检验的统计值显著性概率 sig 为 0.000 ( $<0.05$ ), 符合主成分分析标准。对标准化后数据进行主成分分析可知, 当选取综合指标判据数目为 2 时, 其特征值为 1.224 ( $\geq 1$ ), 累计方差贡献率为 87.604% ( $\geq 85\%$ ), 因此选取前 2 项综合指标判据作为优势分组判据。取分组数为 2、3、4、5、6, 权重向量  $M = [1 \ 0 \ 0]$ , 其他参数不变, 分别计算不同分组下的有效性评价指标  $V_{PC}$ 、 $V_{PE}$ 、 $V_{XB}$  和  $V_{FS}$ 。由图 12 可知, 花阳隧道以产状为判据的最佳分组数为 4, 图 13 为花阳隧道以产状为判据进行结构面优势分组的计算结果。

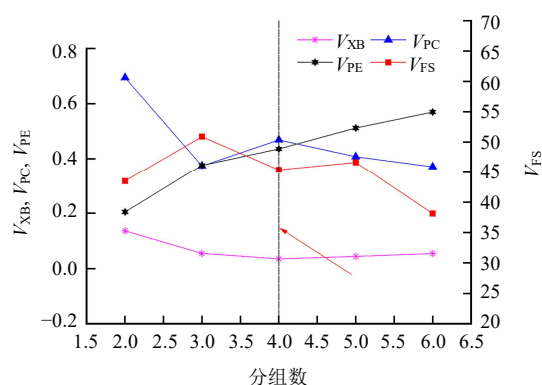


图 12 实测结构面数据聚类有效性指标  
Fig.12 Clustering validity index of measured discontinuities

分别提取前述以产状为判据划分的 4 组结构面数据, 并将此作为新的待分组结构面数据。由式(12)计算第一综合指标判据和第二综合指标判据的权重  $m_1$  和  $m_2$  分别为 0.809 和 0.191, 取分组数为 2、3、

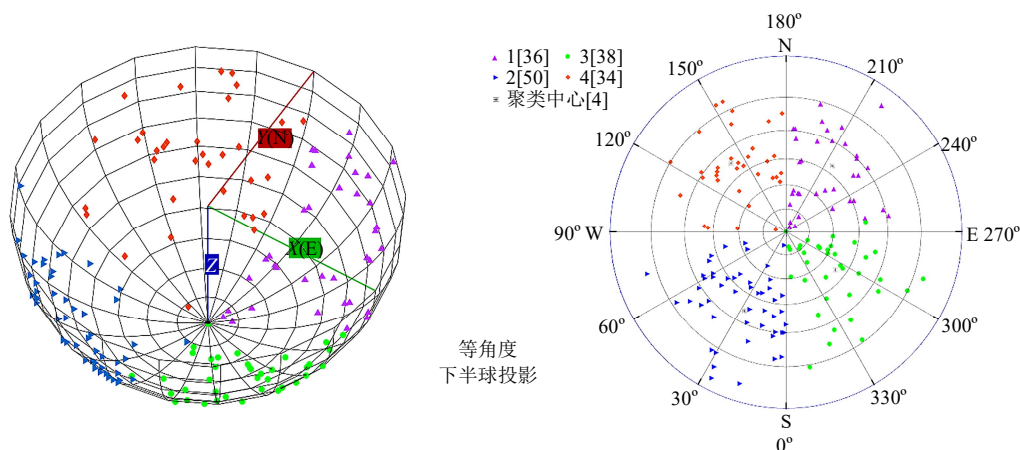


图 13 花阳隧道实测数据按产状判据分组 Schmidt 下半球投影和极点图  
Fig.13 Schmidt lower hemisphere projection map and pole diagram map of measured data of Huayang Tunnel grouped according to the occurrence

4、5, 权重向量  $M = [0.0809 \ 0.191]$ , 其他参数不变, 分别计算不同分组下的有效性指标  $V_{PC}$ 、 $V_{PE}$ 、 $V_{XB}$  和  $V_{FS}$ 。由计算结果可知, 各组结构面数据最佳分组数均为 2。给定分组数为 2, 应用本方法对 4 组待分组结构面进行优势组划分, 分组过程如图 14 所示, 最终分组结果如图 15 所示。

表 5 为花阳隧道实测结构面分组结果, 结合图 15 分析可知, 实测 158 个岩体结构面可清晰地划分为 8 组, 8 组结构面的倾角均较为平缓, 其中, 第 5 组和第 6 组结构面产状相近, 但迹长、张开度、粗糙度和充填状态 4 项指标均存在较大差异, 第 7 组和第 8 组结构面产状、粗糙度和充填状态特征相近, 但迹长和张开度 2 项指标差异明显。查看花阳隧道结构面原始数据可知, 结构面各指标数据均在不同程度上存在一定的随机性, 在不考虑各指标对优势组划分贡献权重的条件下进行多参数优势分组, 势必导致分组结果存在大量交叉。本文提出的

岩体结构面多参数优势分组方法, 在优先考虑以结构面产状为判据进行优势分组的基础上, 各综合指标判据按贡献率参与优势分组, 对原始数据的适应性较强, 可在传统按产状进行优势分组的基础上, 进一步区分结构面的几何和物理特性, 分组结果合理可靠, 可为实际施工提供指导。

为进一步说明 PGS-FCM 方法分组结果的合理性, 本文继续采用 WCPGS-FCM 方法对花阳隧道实测结构面数据进行多参数优势分组, 分组结果如图 16 所示。相比于本文 PGS-FCM 方法, WCPGS-FCM 方法在不考虑各指标判据对优势组划分贡献权重的条件下进行多参数优势分组, 分组结果存在大量交叉, 各优势组产状的空间分布杂乱无章, 这种结果显然是不理想、不具备工程实践意义的, 进一步凸显了 PGS-FCM 方法的分组结果具有较强的合理性和工程应用价值。

本文提出的多参数岩体结构面优势分组方法,

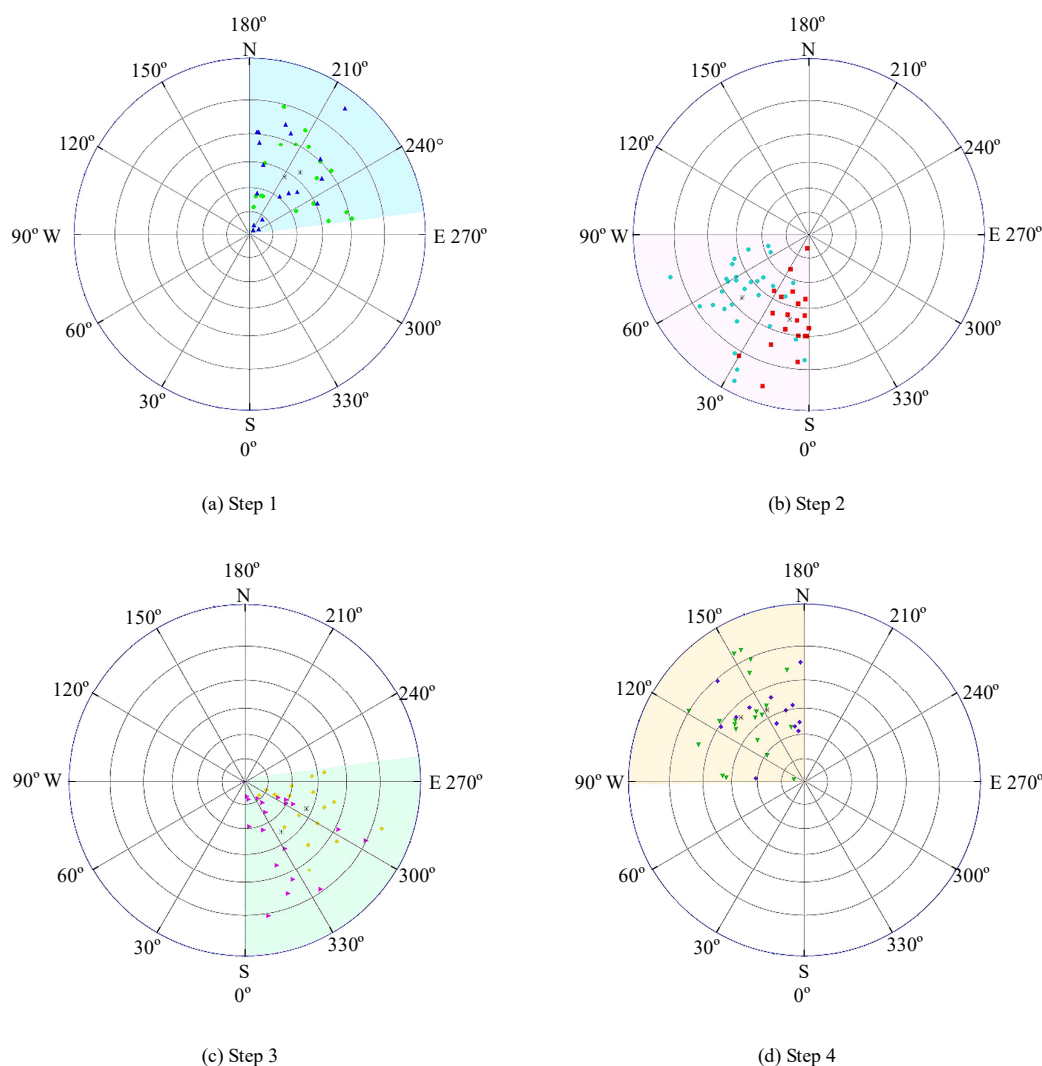


图 14 花阳隧道实测数据采用 PGS-FCM 方法细分过程示意图

Fig.14 Schematic diagram of further refining grouping process of measured data of Huayang Tunnel by PGS-FCM method

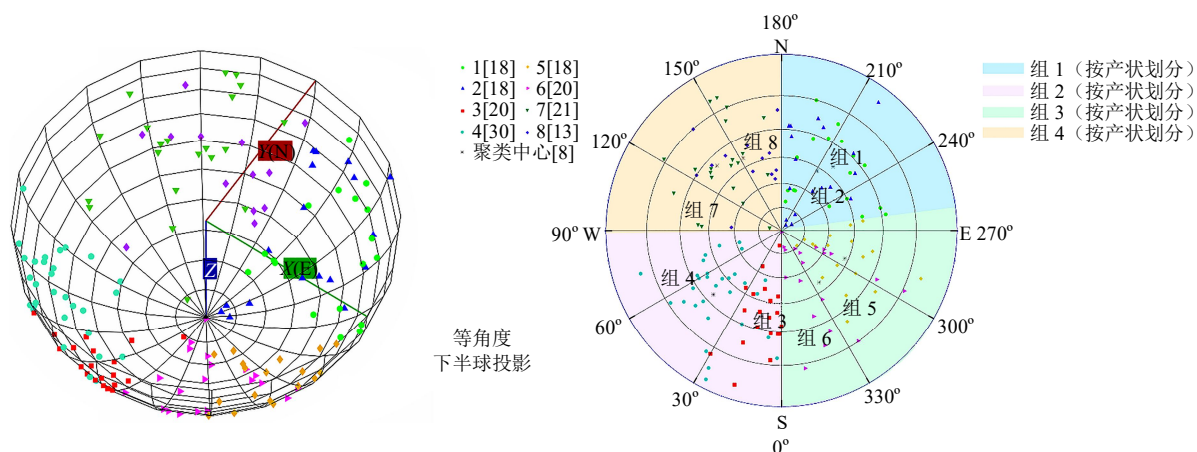


图 15 花阳隧道实测数据采用 PGS-FCM 方法分组 Schmidt 下半球投影和极点图

Fig.15 Schmidt lower hemisphere projection map and pole diagram map of measured data of Huayang Tunnel grouped by PGS-FCM method

表 5 花阳隧道实测结构面数据分组结果

Table 5 Grouping results of measured discontinuities data of Huayang Tunnel

| 聚类中心 | 倾向/(°)  | 倾角/(°) | 迹长/m  | 张开度/cm | 粗糙度   | 充填状态  | 聚类数量 |
|------|---------|--------|-------|--------|-------|-------|------|
| 1    | 219.333 | 49.469 | 3.039 | 0.805  | 3.140 | 0.514 | 18   |
| 2    | 211.244 | 42.854 | 4.011 | 1.737  | 3.153 | 0.374 | 18   |
| 3    | 13.693  | 52.212 | 0.517 | 0.608  | 1.112 | 0.946 | 20   |
| 4    | 47.264  | 55.664 | 2.526 | 1.060  | 2.401 | 0.663 | 30   |
| 5    | 294.392 | 42.751 | 1.406 | 0.479  | 1.821 | 0.863 | 18   |
| 6    | 324.239 | 39.358 | 5.646 | 1.912  | 3.909 | 0.113 | 20   |
| 7    | 135.061 | 54.524 | 2.919 | 0.484  | 2.772 | 0.473 | 21   |
| 8    | 152.600 | 49.128 | 3.540 | 1.783  | 3.164 | 0.398 | 13   |

基于退火遗传算法的原理,将模拟退火算法局部搜索能力强和遗传算法全局搜索能力强的特点有效结合,克服了模糊 C 均值聚类算法受初始聚类中心影响容易陷入局部极小值的缺点,提高了方法的分组效率和计算精度。此外,基于主成分分析法的原始数据前处理过程,可以根据原始数据特点,在保证数据完整性的前提下大幅减少计算工作量。同时,也为相似性度量问题中,各指标判据权重的合理选取问题提供了客观的解决方案,这是其他结构面优势分组方法没有做到的。

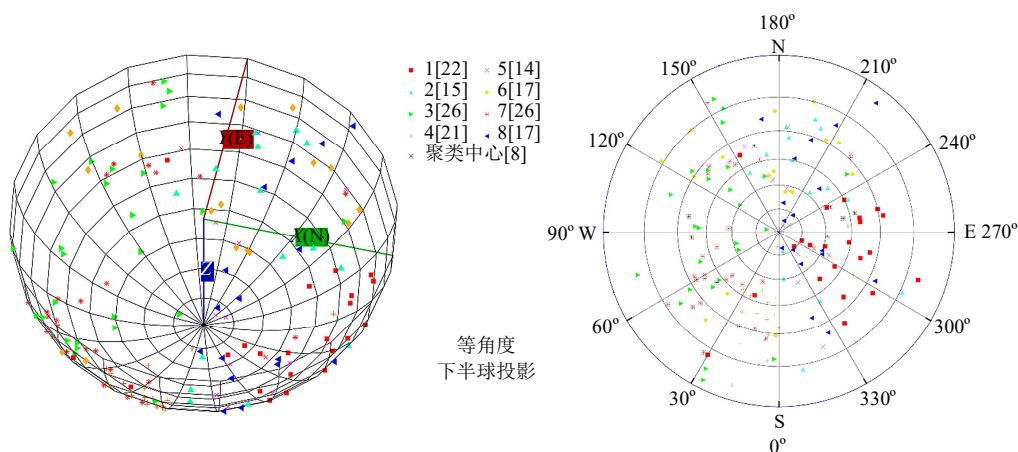


图 16 花阳隧道实测数据采用 WCPGS-FCM 方法分组 Schmidt 下半球投影和极点图

Fig.16 Schmidt lower hemisphere projection map and pole diagram map of measured data of Huayang Tunnel grouped by WCPGS-FCM method

## 6 结 论

传统的结构面优势分组方法仅以产状作为分组依据,没有考虑结构面几何、物理和力学特性的影

响,且多参数结构面优势分组方法存在计算量大,相似性度量权重选取主观性强等缺陷。为此,本文提出一种基于主成分分析的多参数岩体结构面优势分组方法(PGS-FCM)。主要结论如下:

(1) 利用主成分分析法在较小的数据损失范围内, 将迹长、张开度、粗糙度和充填状态 4 项指标转化为 2 项综合判据指标, 减少了计算工作量, 同时为相似性度量问题中指标判据权重的合理选取提供了客观的解决方案, 避免了人工主观赋权造成的分组误差。

(2) 将模拟退火算法和遗传算法结合, 优化模糊 C 均值聚类算法的初始聚类中心, 克服了模糊 C 均值聚类算法容易陷入局部最优解的缺陷, 保证了本文方法具有更高的计算精确度和稳定性。

(3) 应用本文方法对计算机模拟的 5 组共 200 个结构面进行优势分组, 对比不考虑指标权重的 WCPGS-FCM 方法和传统以产状作为优势分组判据的 FCM 方法, 本文提出的 PGS-FCM 方法能够对组间混淆区域内的结构面进行准确分组, 且整体分组正确率最高, 计算出的聚类中心与已知聚类中心吻合度最好, 从而验证了本方法的可靠性。

(4) 将本方法应用于重庆三环高速公路合川至长寿段第 HC04 标段花阳隧道隧址区 158 个结构面现场实测数据中, 并与 WCPGS-FCM 方法进行对比。分析表明本文方法分组结果清晰、合理、可靠, 可以为工程应用提供依据, 进一步表明本文方法具有较好的工程应用价值。

### 参考文献

- [1] 洪陈杰, 黄曼, 夏才初, 等. 岩体结构面各向异性变异系数的尺寸效应研究[J]. 岩土力学, 2020, 41(6): 2098-2109.  
HONG Chen-jie, HUANG Man, XIA Cai-chu, et al. Study of size effect on the anisotropic variation coefficient of rock joints[J] **Rock and Soil Mechanics**, 2020, 41(6): 2098-2109.
- [2] WANG S, HUANG R, NI P, et al. Advanced discretization of rock slope using block theory within the framework of discontinuous deformation analysis[J]. **Geomechanics & Engineering**, 2017, 12(4): 723-738.
- [3] 蒋水华, 欧阳苏, 冯泽文, 等. 基于结构面参数概率分布更新的节理岩质边坡可靠性分析[J]. 岩土力学, 2021, 42(9): 2589-2599.  
JIANG Shui-hua, OUYANG Su, FENG Ze-wen, et al. Reliability analysis of jointed rock slope using on updated probability distribution of structural plane parameters[J] **Rock and Soil Mechanics**, 2021, 42(9): 2589-2599.
- [4] 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.
- [5] DERSHOWITZ W, GEIER J, UCHIDA M, et al. A stochastic approach for fracture set definition[C]//Boston Golder Associates 2nd North American Rock Mechanics Symposium. Rotterdam: A. A. Balkema, 1996: 1809-1813.
- [6] 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.
- [7] 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 Mei-feng, WANG Peng, ZHAO Kui, et al. Fuzzy C-means cluster analysis based on genetic algorithm for automatic identification of joint sets[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2005, 24(3): 371-376.
- [8] 卢波, 丁秀丽, 邬爱清. 岩体随机不连续面产状数据划分方法研究[J]. 岩石力学与工程学报, 2007, 26(9): 1809-1816.  
LU Bo, DING Xiu-li, WU Ai-qing. Study on method of orientation data partitioning of randomly distributed discontinuities of rocks[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2007, 26(9): 1809-1816.
- [9] 宋腾蛟, 陈剑平, 张文, 等. 基于萤火虫算法的岩体结构面优势产状聚类分析[J]. 东北大学学报(自然科学版), 2015, 36(2): 284-287, 304.  
SONG Teng-jiao, CHEN Jian-ping, ZHANG Wen, et al. Cluster analysis of dominant attitudes of rock mass structural plane based on firefly algorithm[J]. **Journal of Northeastern University (Natural Science)**, 2015, 36(2): 284-287, 304.
- [10] 王述红, 任艺鹏, 陈俊智, 等. 一种改进鱼群聚类算法在结构面分组中的应用[J]. 东北大学学报(自然科学版), 2019, 40(3): 420-424.  
WANG Shu-hong, REN Yi-peng, CHEN Jun-zhi, et al. An improved fish swarm clustering algorithm for structural grouping[J]. **Journal of Northeastern University (Natural Science)**, 2019, 40(3): 420-424.
- [11] 侯钦宽, 雍睿, 杜时贵, 等. 基于编网算法的岩体结构面优势产状聚类分析[J]. 岩石力学与工程学报, 2020, 39(增刊1): 2871-2881.  
HOU Qin-kuan, YONG Rui, DU Shi-gui, et al. A method for cluster orientation data of discontinuities of rock mass based on netting algorithm[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2020, 39(Suppl.1): 2871-2881.
- [12] BEHZAD T, HOSSEIN M, BEHZAD M, et al. Investigating the validity of conventional joint set clustering methods[J]. **Engineering Geology**, 2011, 118(3): 75-81.
- [13] ZHOU W, MAERZ N H. Implementation of multivariate clustering methods for characterizing discontinuities data from oriented boreholes[J]. **Computers & Geosciences**, 2002, 28(7): 827-839.

- [14] 徐黎明, 陈剑平, 王清. 多参数岩体结构面优势分组方法研究[J]. 岩土力学, 2013, 34(1): 189-195.  
XU Li-ming, CHEN Jian-ping, WANG Qing. Study of method for multivariate parameter dominant partitioning of discontinuities rock mass[J]. **Rock and Soil Mechanics**, 2013, 34(1): 189-195.
- [15] 宋腾蛟, 陈剑平, 张文, 等. 基于人工蜂群算法的岩体结构面多参数优势分组研究[J]. 岩土力学, 2015, 36(3): 861-868.  
SONG Teng-jiao, CHEN Jian-ping, ZHANG Wen, et al. A method for multivariate parameter dominant partitioning of discontinuities of rock mass based on artificial bee colony algorithm[J]. **Rock and Soil Mechanics**, 2015, 36(3): 861-868.
- [16] 宋盛渊, 王清, 陈剑平, 等. 岩体结构面的多参数优势分组方法研究[J]. 岩土力学, 2015, 36(7): 2041-2048.  
SONG Sheng-yuan, WANG Qing, CHEN Jian-ping, et al. A method for multivariate parameter dominant partitioning discontinuities of rock mass[J]. **Rock and Soil Mechanics**, 2015, 36(7): 2041-2048.
- [17] BARTON N. A relationship between joint roughness and joint shear strength[C]//Rock Fracture Process International symposium on Rock Mechanics. Nancy: [s. n.], 1971.
- [18] ROCHES M. The shear strength of rock joints in theory and practice[J]. **Rock Mechanics**, 1977, 10(1-2): 1-54.
- [19] 陈曦, 曾亚武. 基于Grasselli模型的一个新的岩石节理三维粗糙度指标[J]. 岩土力学, 2021, 42(3): 700-712.  
CHEN Xi, ZENG Ya-wu. A new three-dimensional roughness metric based on Grasselli's model[J]. **Rock and Soil Mechanics**, 2021, 42(3): 700-712.
- [20] 陈剑平, 石丙飞, 王清. 工程岩体随机结构面优势方向的表示法初探[J]. 岩石力学与工程学报, 2005, 24(2): 241-245.  
CHEN Jian-ping, SHI Bing-fei, WANG Qing. Study on the dominant orientations of random fractures of fractured rock masses[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2005, 24(2): 241-245.
- [21] 丁敬国, 郭锦华. 基于主成分分析协同随机森林算法的热连轧带钢宽度预测[J]. 东北大学学报(自然科学版), 2021, 42(9): 1268-1274, 1289.  
DING Jing-guo, GUO Jin-hua. Prediction of rough rolling width based on principal component analysis collaborated with random forest algorithm[J]. **Journal of Northeastern University (Natural Science)**, 2021, 42(9): 1268-1274, 1289.
- [22] LIU J, ZHAO X D, XU Z H. Identification of rock discontinuity sets based on a modified affinity propagation algorithm[J]. **International Journal of Rock Mechanics and Mining Sciences**, 2017, 94: 32-42.
- [23] LI Y Y, WANG Q, CHEN J P, et al. K-means algorithm based on particle swarm optimization for the identification of rock discontinuity sets[J]. **Rock Mechanics and Rock Engineering**, 2015, 48(1): 375-385.
- [24] 杨杰. 岩体结构面的优势产状模糊聚类分析及其有效性研究[D]. 哈尔滨: 哈尔滨工业大学, 2017.  
YANG Jie. Fuzzy cluster analysis of dominant occurrence of rock mass discontinuities and its effectiveness[D]. Harbin: Harbin University of Technology, 2017.
- [25] 李宁, 王李管, 贾明涛, 等. 改进遗传算法和支持向量机的岩体结构面聚类分析[J]. 岩土力学, 2014, 35(增刊2): 405-411.  
LI Ning, WANG Li-guan, JIA Ming-tao, et al. Application of improved genetic algorithm and support vector machine to clustering analysis of rock mass structural plane[J]. **Rock and Soil Mechanics**, 2014, 35(Suppl.2): 405-411.
- [26] 张奇, 王清, 阙金声, 等. 基于凝聚层次聚类分析法的岩体随机结构面产状优势分组[J]. 岩土工程学报, 2014, 36(8): 1432-1437.  
ZHANG Qi, WANG Qing, QUE Jin-sheng, et al. Dominant grouping of discontinuities of rock mass based on AGNES[J]. **Chinese Journal of Geotechnical Engineering**, 2014, 36(8): 1432-1437.
- [27] 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 Feng-shan, GONG Cheng-cheng, et al. Data analysis method for optimized and dominant orientations of joints in rock mass[J]. **Journal of Engineering Geology**, 2011, 19(6): 887-892.
- [28] 王述红, 朱宝强, 王鹏宇. 模拟退火聚类算法在结构面产状分组中的应用[J]. 东北大学学报(自然科学版), 2020, 41(9): 1328-1333.  
WANG Shu-hong, ZHU Bao-qiang, WANG Peng-yu. Application of simulated annealing clustering algorithm in grouping of discontinuity orientation[J]. **Journal of Northeastern University (Natural Science)**, 2020, 41(9): 1328-1333.
- [29] 王述红, 董福瑞, 朱宝强, 等. 山岭隧道围岩参数智能反演及稳定性分析[J]. 应用基础与工程科学学报, 2021, 29(5): 1171-1185.  
WANG Shu-hong, DONG Fu-rui, ZHU Bao-qiang, et al. Parameters intelligent inversion and stability analysis of surrounding rock of mountain tunnel[J]. **Journal of Basic Sciences and Engineering**, 2021, 29(5): 1171-1185.
- [30] RUSPINI E H. A new approach to clustering[J]. **Information and Control**, 1969, 15(1): 22-32.
- [31] ZADEH L A. Fuzzy logic[J]. **Computer**, 1988, 21(4): 83-93.
- [32] BEZDEK J C. Pattern recognition with fuzzy objective function algorithms[M]. New York: A Division of Plenum Publishing Corporation, 1981.