

基于改进遗传算法的岩体结构面产状聚类分析

崔学杰, 晏鄂川, 陈 武

(中国地质大学(武汉) 工程学院, 湖北 武汉 430074)

摘 要: 根据产状对结构面进行分组是研究岩体结构的重要环节。传统分组方法通常需要依靠地质经验, 缺乏客观性, 而现有的聚类方法也存在一些缺陷。基于变长度字符串遗传算法, 提出了一种改进的 K 均值算法, 实现了岩体结构面产状的自动聚类。该方法的核心思想是使用遗传算法为 K 均值算法选择恰当的聚类中心, 克服了 K 均值(K -means)算法受初始聚类中心影响, 易收敛于局部最优解的缺陷。由于使用了变长度字符串, 该方法能够在聚类过程中自动确定最佳结构面组数, 同时提供最优的分组结果。针对产状数据, 提出了一种新的变异方法, 该方法利用 C++ 语言实现, 并被应用于浙江省某地下水封洞库结构面产状数据的分析, 得到较为合理的分组结果, 证明了该方法的有效性。

关 键 词: 岩体结构面; 产状数据; K 均值算法; 变长度字符串遗传算法

中图分类号: TU 452

文献标识码: A

文章编号: 1000-7598 (2019) 增 1-0374-07

Cluster analysis of discontinuity occurrence of rock mass based on improved genetic algorithm

CUI Xue-jie, YAN E-chuan, CHEN Wu

(Faculty of Engineering, China University of Geosciences, Wuhan, Hubei 430074, China)

Abstract: Grouping structural planes based on their occurrences is an important way in analyzing the structure characteristics of rock mass. Traditional classification methods usually greatly rely on geological experience of researchers, which is lack of objectivity. The powerful clustering methods, however, also have drawbacks. In this paper, based on variable length string genetic algorithm, an improved K -means clustering method was proposed, making the automatic clustering of discontinuity occurrence of rock mass available. The essence of the proposed method is to select appropriate initial cluster centers for K -means algorithm applying the genetic algorithm. It overcomes the limitation that the K -means algorithm is usually affected by the initial cluster center and easily converges to the local optimal solution. The application of variable length strings in the improved algorithm of classification, however, can not only automatically determine the number of the optimal structural plane groups during the clustering process, but also provide optimal grouping results. In addition, a new mutation algorithm is proposed based on the occurrence data. It is realized in C++ and is applied in analyzing the occurrence of structural planes in an underground water-sealing rock caverns located at Zhejiang province, China. Reasonable classification results are achieved, proving the applicability of the proposed approach in this paper.

Keywords: rock mass discontinuity; occurrence data; K -means algorithm; variable length string genetic algorithm

1 引 言

结构面在确定岩体的力学和水力学性质方面起着重要作用^[1-2], 为分析结构面发育的规律性, 常将具有某些共同特征的结构面归为一组。国际岩石力学协会提出了定量描述结构面的 10 项指标^[3], 包

括产状、粗糙度、粗糙度、间距和填充状况等, 目前的分组方法分为单指标和多指标分组两类。相对于其他指标, 产状的物理意义明确, 易于现场测量, 并广泛用于岩体稳定性的初步评价, 因此, 基于产状的单指标分组在实际应用中仍然占据主流。

传统的结构面分组通过分析玫瑰花图或极点

收稿日期: 2018-10-27

基金项目: 国家自然科学基金项目 (No.41172282, No.41672313)。

This work was supported by the National Natural Science Foundation of China(41172282, 41672313).

第一作者简介: 崔学杰, 男, 1990 年生, 博士研究生, 主要从事岩体稳定性与地质灾害防治研究。E-mail: 1229054139@qq.com

通讯作者: 晏鄂川, 男, 1969 年生, 博士, 教授, 主要从事岩体稳定性与地质灾害防治研究。E-mail: yecyec6970@163.com

等密度图来实现, 简单直观, 但分组结果比较依赖分析者的专业素养和经验, 往往缺乏客观性^[4-5]。另外, 分组方法无法将每一个结构面明确地划分到所属的组, 组与组之间的界限也不清晰。聚类技术的引进改善了这一情况, 1976 年 Shanley 等^[6]提出了一种基于密度的聚类算法用于结构面分组, 并于 1982 年对方法进行了改进^[7]。这种方法的聚类结果非常依赖于两个输入参数, r (邻域半径) 和 k (最小包含点数), 而这两个参数需要结合经验确定, 仍存在一定主观性。Hammah 等^[8]采用模糊 C 均值 (fuzzy C -means, 即 FCM) 算法进行结构面组数的划分, 范雷等^[5]则使用 K 均值 (K -means) 算法对结构面产状进行了聚类分析。FCM 和 K -means 都使用 Picard 迭代算法来最小化目标函数, 这两种算法的聚类结果受初始聚类中心影响, 易收敛于局部最优解^[9]。为了避免选择初始聚类中心的困难, Jimenez-Rodriguez 等^[10]提出了用于结构面产状数据的谱聚类方法 (spectral clustering method)。另外, 一些学者将 K -means 或 FCM 算法与随机优化算法结合, 开发了很多新的聚类算法诸如基于遗传算法的 FCM 算法^[4]、基于多尺度混沌优化算法的 FCM 算法^[11]、基于粒子群优化算法的 K -means 算法^[12]等。这些算法能够稳健地收敛于全局最优解, 性能优于 K -means 或 FCM 算法, 但仍需要预先指定结构面组数或通过多次试算确定。

变长度字符串遗传算法 (variable length string genetic algorithm, 即 VGA) 是传统遗传算法的改进版本, 已成功运用于卫星图像解译和基因数据分析^[13-14]。VGA 算法继承了传统遗传算法的优点, 是一种并行, 高效, 稳健的随机搜索和优化技术, 适用于搜索空间较大, 复杂且峰值较多的情况^[15]。本文结合 K -means 算法和变长度字符串遗传算法, 提出一种用于结构面分组的聚类算法, 能够在聚类分析过程中自动确定最佳结构面组数, 真正地实现全自动聚类, 用于某地下水封洞库结构面产状数据的聚类分析, 验证了其有效性和优越性。

2 结构面产状的表示方法

在工程地质分析中结构面通常被假设为空间中一个几何平面, 其产状使用倾向 α ($0 \leq \alpha \leq 360^\circ$) 和倾角 β ($0 \leq \beta \leq 90^\circ$) 表示, 也可用结构面的单位法向量表示其产状, 并与单位球面上的点一一对应, 建立如图 1 所示的空间直角坐标系。图中, x_1 轴指向北为正; x_2 轴指向东为正; x_3 轴指向上为正。结构面 j 可以表示为

$$\mathbf{x}_j = \begin{bmatrix} x_{j1} \\ x_{j2} \\ x_{j3} \end{bmatrix} = \begin{bmatrix} \cos \alpha \sin \beta \\ \sin \alpha \sin \beta \\ \cos \beta \end{bmatrix} \quad (1)$$

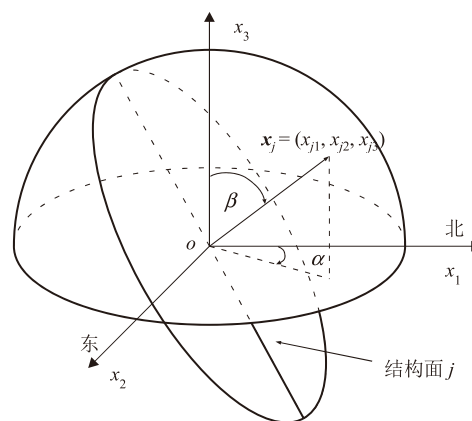


图 1 结构面产状的表示

Fig.1 Representation of the occurrence of structural plane

3 聚类分析

聚类分析依据相似性或非相似性, 将给定的 n 个数据分为 K 组 (或类), 使组内数据之间具有最大的相似性, 而不同组之间的数据具有最大的差异性^[16]。 K -means 算法是常用的聚类分析算法, 具体步骤见文献^[12, 16], 这种算法使用距离度量表征数据的非相似性, 以 K 个随机选择的初始聚类中心为起点, 迭代地计算划分矩阵和聚类中心来最小化目标函数, 实现动态聚类分析, K -means 算法的目标函数为

$$J = \sum_{k=1}^K \sum_{j=1}^n u_{kj} D^2(z_k, \mathbf{x}_j) \quad (2)$$

式中: $D(z_k, \mathbf{x}_j)$ 为数据 $\mathbf{x}_j$ 与第 k 组的聚类中心 z_k 的距离; $u_{kj} = 1$ 或 0 。当

$$D(z_k, \mathbf{x}_j) < D(z_i, \mathbf{x}_j), (i, k = 1, \dots, K; i \neq k) \quad (3)$$

成立时, $u_{kj} = 1$, 表示 $\mathbf{x}_j$ 是否属于第 k 组, 否则, $u_{kj} = 0$ 。

距离度量决定聚类结果是否能真实地反映数据的结构^[8]。对于结构面产状数据的聚类分析, Hammah 等^[17]建议使用结构面法向量的夹角的正弦平方值表示结构面的距离, 即

$$D^2(\mathbf{x}_i, \mathbf{x}_j) = 1 - (\mathbf{x}_i^T \mathbf{x}_j)^2 \quad (4)$$

式中: 上标 “T” 表示矩阵转置。

文献[8, 12]研究表明, 使用一组结构面产状数据的算术平均值表示其聚类中心(即平均产状)是错误的, 并建议使用 Markland 提出的特征分析方法^[18]计算平均产状。该方法分为两步, 以计算第 k 组结构面的平均产状为例说明。

(1) 构造第 k 组结构面的产状矩阵

$$\left. \begin{aligned} S_k &= [S_{klm}]_{3 \times 3} (l, m = 1, 2, 3), \\ \text{其中 } S_{klm} &= \sum_{j=1}^n u_{kj} x_{jl} x_{jm} \end{aligned} \right\} \quad (5)$$

(2) 计算矩阵 S_k 的特征值 λ_{k1} 、 λ_{k2} 、 λ_{k3} 和归一化特征向量 τ_{k1} 、 τ_{k2} 、 τ_{k3} 。

如果 $\lambda_{k1} < \lambda_{k2} < \lambda_{k3}$, 则最大特征值 λ_{k3} 对应的特征向量 τ_{k3} 即为第 k 组的聚类中心, 即

$$z_k = \tau_{k3} \quad (6)$$

字符串 1	0.433	0.250	0.866	0.500	0.500	0.707	0.354	0.612	0.707	0.000	0.866	0.500
字符串 2	0.171	0.296	0.940	0.000	0.643	0.766	0.750	0.433	0.500			

(a) 父代字符串

字符串 1	0.433	0.250	0.866	0.500	0.500	0.707	0.354	0.612	0.707	0.000	0.866	0.500
字符串 2	0.171	0.296	0.940	0.000	0.643	0.766	0.750	0.433	0.500			

(b) 交叉点位置

子代字符串 1	0.433	0.250	0.866	0.500	0.500	0.707	0.354	0.612	0.707	0.000	0.643	0.766	0.750	0.433	0.500
子代字符串 2	0.171	0.296	0.940	0.000	0.866	0.500									

(c) 子代字符串

图 2 字符串的交叉运算

Fig.2 Crossover operation of two strings

对于初始种群中任一字符串 i , 需要先产生一个随机整数 $K_i \in [K_{\min}, K_{\max}]$, 然后在结构面产状数据中随机选择 K_i 个数据作为字符串 i 中的聚类中心, $K_{\min}$ 一般设置为 2。实际工程中岩体结构面组数一般为 3~5 组, $K_{\max}$ 设置为 6 是恰当的。如果种群中包含 P 个字符串, 上述步骤将重复 P 次以完成种群初始化。

4.2 适应度函数

适应度值定量表示字符串中聚类中心所对应的聚类方案的优劣, 在相应的搜索空间内聚类方案越接近最优解, 其适应度值越高。在遗传算法及其他进化算法中常用的适应度函数包括 Xie-Beni 指标^[13]、Davies-Bouldin 指标^[20]、Average Silhouette Width 准则^[21]等。这些指标(或准则)最初都是使用欧拉度量定义的, 并不适用于产状或其他球形数

4 基于 VGA 的 K-means 算法

变长度字符串遗传算法与传统遗传算法相同, 需要将可能的聚类结果编码为字符串格式, 然后通过遗传算子引导种群进化, 从而获取最佳聚类结果。本节重点阐述该算法中的核心内容, 即编码方法与种群初始化, 适应度函数和遗传算子。有关遗传算法的原理和具体步骤见文献[4, 15, 19], 这里不再赘述。

4.1 编码方法与种群初始化

对于结构面产状的聚类分析, 可将聚类中心作为遗传信息时编码为字符串, 字符串长度较短, 可提高计算效率。如果一个字符串中包含 K 个聚类中心, 则该字符串的长度为 $3K$, 如图 2(a)中的字符串 1、2 分别包含 4、3 个聚类中心, 长度分别为 12、9。

据。

本文对 Davies-Bouldin (DB) 指标进行了修正, 作为评价产状数据聚类结果适应度的函数。DB 指标被定义为组内方差与组间距离的比值的函数^[22]。为了将 DB 指标用于产状数据, 将式(4)代入计算 DB 指标的原始公式, 获得修正后的式(7)~(10)。当 DB 指标取最小值时获得最佳聚类结果。因此, 适应度函数取 DB 指标的倒数, 见式(11)。

$$v_i = \sum_{j=1}^n u_{ij} \left[1 - (z_i^T x_j)^2 \right] / \sum_{j=1}^n u_{ij} \quad (7)$$

$$d_{ij} = 1 - (z_i^T z_j)^2 \quad (8)$$

$$R_i = \max \left\{ \frac{v_i + v_j}{d_{ij}} \right\}, \quad (j=1, \dots, K; j \neq i) \quad (9)$$

$$DB = \frac{1}{K} \sum_{i=1}^K R_i \quad (10)$$

$$f = 1 / DB \quad (11)$$

式中： v_i 为第 i 组数据的方差； d_{ij} 为第 i 、 j 组数据聚类中心的距离； R_i 为考虑了组内方差与组间距离的综合指标。

对于某个字符串，需要以该字符串内编码的数据作为初始聚类中心，按照 K -means 算法对结构面产状进行聚类，然后根据聚类结果，用式 (7) ~ (11) 计算适应度值。

4.3 遗传算子

选择、交叉和变异是 3 种基本遗传算子，3 种算子共同引导种群的进化。

选择算子：选择算子模拟达尔文进化论中的自然选择和适者生存过程。适应度越高的个体，有更大的机会繁殖后代，传递遗传信息。常用的选择方法有赌轮盘选择法，锦标赛选择法，随机抽样选择法等。本文选择易于实现的赌轮盘选择法。

交叉算子：交叉算子交换父代个体的遗传信息，产生与父代不同的后代。常用的交叉方法均不适用于由可变长度字符串组成的种群，Maulik 等^[13]对单点交叉方法进行了修正，扩展了其适用范围。该算子的交叉概率设置为固定值 μ_c ，在交叉过程中每一个聚类中心都视为不可分割的，交叉点只能在两个聚类中心之间。假设两个父代字符串 1 和 2 分别包含 K_1 和 K_2 个聚类中心，则字符串 1 的交叉点 η_1 通过式 (12) 确定，而字符串 2 的交叉点 η_2 由式 (13) ~ (15) 确定。

$$\eta_1 = \text{rand}() \text{ Mod } K_1 \quad (12)$$

$$\eta_{2l} = \min \{ 2, \max [0, 2 - (K_1 - \eta_1)] \} \quad (13)$$

$$\eta_{2u} = [K_2 - \max (0, 2 - \eta_1)] \quad (14)$$

$$\eta_2 = \eta_{2l} + \text{rand}() \text{ Mod } (\eta_{2u} - \eta_{2l}) \quad (15)$$

式中： $\text{rand}()$ 产生一个随机正整数； Mod 为取模运算； η_{2u} 、 η_{2l} 分别为 η_2 取值的上、下限。

如果字符串 1、2 如图 2(a)所示，即 $K_1 = 4$ ， $K_2 = 3$ 。假设式 (14) 中的随机整数取 15，则 $\eta_1 = 3$ ，同时由式 (15)、(16) 可以确定 $\eta_2 \in [1, 3]$ ，假设式 (17) 中的随机整数为 50， $\eta_2 = 1$ 。交叉点的位置在图 2(b)中以“||”表示，交叉后产生的子代字符串如图 2(c)所示。由这个实例可知，子代字符串的长度不受父代的影响。

变异算子：变异算子的作用是通过产生新的基因值来保持种群的多样性，防止收敛于局部最优解。结构面产状是定义在单位球面上的数据，传统的遗传算子可能产生无意义的字符串。因此，本文设计一种新的变异方法。在变异过程中每个聚类中心发生变异的概率均为 μ_m ，假设第 k 组结构面当前的聚类中心为 z_k ，变异后的聚类中心为 z'_k ，如图 3 所示。从图中可以看出，可以认为 z'_k 是 z_k 绕原点 o 旋转角度 δ 产生的，其构成了一个圆锥面。要获得唯一的 z'_k ，还需指明旋转方位角 θ 。为保证变异的随机性， θ 从均匀分布 $U(0, 2\pi)$ 随机产生。变异算子的数学表达为

$$z'_k = R_{x3}(\alpha_k) R_{x2}(\beta_k) \begin{bmatrix} \cos \theta \sin \delta \\ \sin \theta \sin \delta \\ \cos \delta \end{bmatrix} \quad (16)$$

式中： α_k 、 β_k 分别为聚类中心 z_k 的倾向和倾角； $R_{x2}()$ 、 $R_{x3}()$ 分别为绕 x_2 、 x_3 轴旋转的旋转矩阵，定义如下：

$$R_{x2}(\beta_k) = \begin{bmatrix} \cos \beta_k & 0 & \sin \beta_k \\ 0 & 1 & 0 \\ -\sin \beta_k & 0 & \cos \beta_k \end{bmatrix} \quad (17)$$

$$R_{x3}(\alpha_k) = \begin{bmatrix} \cos \alpha_k & -\sin \alpha_k & 0 \\ \sin \alpha_k & \cos \alpha_k & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (18)$$

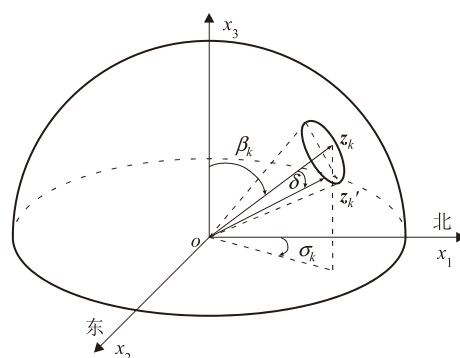


图3 变异算子示意图

Fig.3 Sketch of mutation operator

4.4 算法的实现和收敛准则

Visual Studio Community 是一款免费的编程软件，功能强大，支持多达 36 种编程语言。本文基于该平台，利用 C++ 语言实现了上述算法。算法运行至种群中最大适应度值稳定时认为收敛，停止运行。本文使用了精英模型保证最佳字符串不会因遗传算子的随机性而被遗漏^[21-23]。

5 工程实例应用

为了验证本文提出的聚类算法的有效性,将其运用于浙江省某地下水封洞库的结构面分组。该洞库位于浙江省东部沿海地带,主要岩性为上侏罗统高坞组(J_3g)花岗斑岩。本文所用的 279 个结构面产状数据来自于其中 3 个钻孔影像解译结果,其等面积,上半球极点等密度图如图 4 所示。由图可知,陡倾角结构面极点形成了 2 个高密度区,中等倾角结构面极点形成 1 个中~低密度区,还有少量缓倾角结构面极点构成了 1 个低密度区。根据经验判断应将结构面分为 4 组。

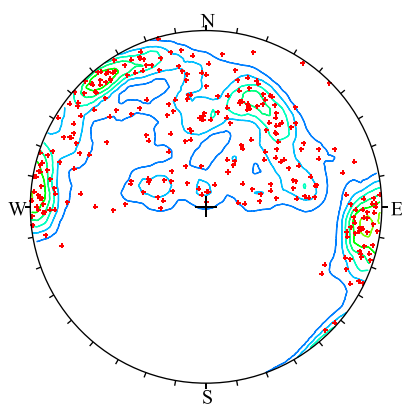
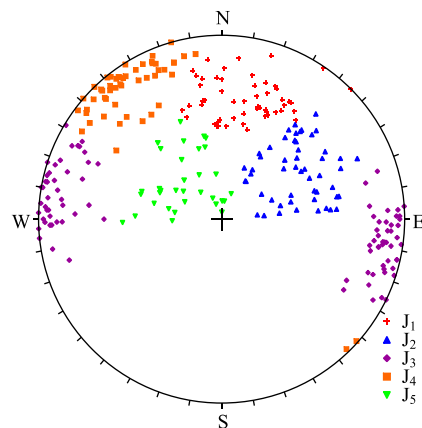


图 4 结构面极点等密度图

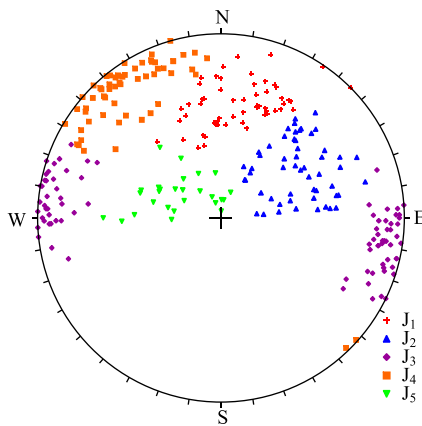
Fig.4 The isodensity contour of the poles of structural planes

新算法的优势在于使用了遗传算法选择恰当的初始聚类中心,从而保证算法能够收敛于全局最优解,本文使用 3 种聚类方法对上述 279 个结构面产状数据进行了聚类分析,即本文提出的新方法、谱聚类方法^[10]和 K -means 算法,并将 3 种方法的分组结果进行对比。

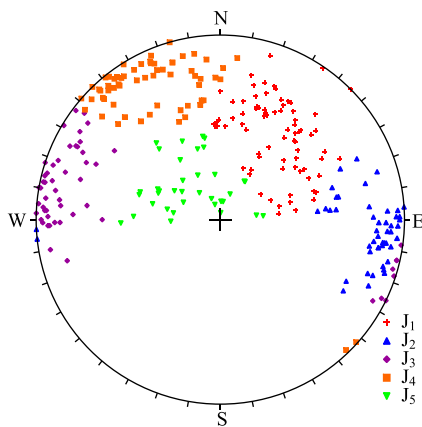
首先使用本文提出的新方法对结构面进行分组,参数见表 1。新方法的聚类结果如图 5(a)所示和见表 2。新方法将 279 个结构面按产状分为 5 组,其中, J_3 、 J_4 为陡倾角结构面; J_1 、 J_2 为中等倾角结构面,而 J_5 为缓倾角结构面。新方法的分组结果与



(a) 新方法



(b) 谱聚类算法



(c) K -means 算法

图 5 聚类结果的极点图

Fig.5 Pole map of clustering results

表 2 新算法的结构面聚类结果

Table 2 Clustering results of structural planes using the new algorithm proposed in this paper

编号	最终聚类中心产状 (倾向/ 倾角)/(°)	结构面数量 /个
J_1	11.08/56.84	53
J_2	58.44/41.15	52
J_3	99.50/89.08	89
J_4	324.69/79.46	52
J_5	313.47/23.51	33

表 1 新算法的参数

Table 1 Parameters of the new algorithm

参数	取值
种群大小 P /个	10
最大组数 $K_{\max}$ /组	6
交叉概率 μ_c	0.80
变异概率 μ_m	0.01
变异角 δ /(°)	1.00

经验分析基本一致，区别在于中等倾角结构面被分为两组。由表 2 可知，尽管 J_1 、 J_2 的平均倾角相差不大，但平均倾向差别较大，不宜划分为一组。新方法的分组结果是合理的。对比图 5(a)和图 5(b)可知，新方法和谱聚类方法的分组结果基本一致，证明了新方法的有效性。

K -means 算法对初始聚类中心十分敏感，初始聚类中心不同，聚类结果往往也不同，如果初始聚类中心初始聚类中心选择不当，算法则无法收敛于全局最优解。图 5(c)和表 3 展示了 K -means 算法的一种聚类结果，由图 5(c)可知， K -means 算法能够识别 J_4 、 J_5 ，分组结果与前两种方法基本一致。不同之处在于， K -means 算法将大部分中等倾角结构面分为一组 (J_1)，同时将少量中等倾角结构面和陡倾结构面分为一组 (J_2)，这是错误的。由表 3 可知， K -means 算法的最终聚类中心与初始聚类中心的差别（向量夹角）较小，正是该算法受初始聚类中心影响的体现。另外，新方法能够在聚类过程中自动确定恰当的组数，而谱聚类方法和 K -means 算法都需要预先指定结构面组数，或通过试算确定。

表 3 K -means 算法的结构面聚类结果

Table 3 Clustering results of structural planes using the K -means algorithm

编号	初始聚类中心产状 (倾向/倾角)/(°)	最终聚类中心产状 (倾向/倾角)/(°)	结构面数量 /个
J_1	29.20/59.64	33.95/47.31	77
J_2	99.17/77.31	93.88/74.04	49
J_3	107.08/81.92	284.11/81.43	54
J_4	332.26/75.96	332.35/74.93	63
J_5	287.29/37.61	317.23/20.12	36

6 结 语

K -means 算法是常用的动态聚类分析算法之一，但存在诸多缺陷。本文基于可变长度字符串遗传算法，对 K -means 算法进行了改进，提出了一种用于产状数据分组的新方法。

改进后 K -means 该算法的分组结果与结构面产状的等密度图相吻合，但其结果更加精确，分类更加合理，优点在于不受初始聚类中心影响，是一种全局最优方法。新方法无需指定预期分组数，也不需要试算，能够节约大量时间，其性能优于 K -means 算法与谱聚类算法。

参 考 文 献

[1] 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 Yu-xin, ZHOU Zhi-fang, SUN Qi-guo. Synthetical fuzzy clustering analysis for joints occurrence of rock mass[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2005, 24(13): 2283—2287.

[2] DUNCAN C W, CHRISTOPHER W M. Rock slope engineering: Civil and mining(4th ed.)[M]. New York: Spon Press, 2004.

[3] ISRM. International society for rock mechanics commission on standardization of laboratory and field tests: Suggested methods for the quantitative description of discontinuities in rock masses[J]. **International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts**, 1978, 15(6): 319—368.

[4] 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.

[5] 范雷, 王亮清, 唐辉明. 节理岩体结构面产状的动态聚类分析[J]. 岩土力学, 2007, 28(11): 2405—2408.

FAN Lei, WANG Liang-qing, TANG Hui-ming. Dynamic cluster analysis of discontinuity orientations of jointed rock mass[J]. **Rock and Soil Mechanics**, 2007, 28(11): 2405—2408.

[6] SHANLEY R J, MAHTAB M A. Delineation and analysis of clusters in orientation data[J]. **Mathematical Geology**, 1976, 8(1): 9—23.

[7] 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: [s. n.], 1982: 116—123.

[8] 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.

- [9] BEZDEK J C. Pattern recognition with fuzzy objective function algorithms[M]. New York: Plenum, 1981.
- [10] JIMENEZ-RODRIGUEZ R, SITAR N. A spectral method for clustering of rock discontinuity sets[J]. **International Journal of Rock Mechanics and Mining Sciences**, 2006, 43(7): 1052—1061.
- [11] 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 of rock mass[J]. **Rock and Soil Mechanics**, 2013, 34(1): 189—195.
- [12] LI Yan-yan, WANG Qing, CHEN Jian-ping, 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
- [13] MAULIK U, BANDYOPADHYAY S. Fuzzy partitioning using a real-coded variable-length genetic algorithm for pixel classification[J]. **IEEE Transactions on Geoscience and Remote Sensing**, 2003, 41(5): 1075—1081.
- [14] HRUSCHKA E R, CAMPELLO R J G B, CASTRO L N D. Evolving clusters in gene-expression data[J]. **Information Sciences**, 2006, 176(13): 1898—1927.
- [15] MAULIK U, BANDYOPADHYAY S, MUKHOPADHYAY A. Multiobjective genetic algorithms for clustering[M]. Berlin: Springer, 2011.
- [16] JAIN A K, MURTY M N, FLYNN P J. Data clustering: A review[J]. **ACM Computing Surveys**, 1999, 31(3): 264—323.
- [17] HAMMAH R E, CURRAN J H. On distance measures for the fuzzy K-means algorithm for joint data[J]. **Rock Mechanics and Rock Engineering**, 1999, 32(1): 1—27.
- [18] MARKLAND J. The analysis of principal components of orientation data[J]. **International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts**, 1974, 11(5): 157—163.
- [19] 谢全敏, 夏元友. 基于遗传算法的边坡稳定性评价的动态聚类法[J]. 岩土力学, 2002, 23(2): 170—172.
XIE Quan-min, XIA Yuan-you. A dynamic clustering method for evaluation of slope stability based on genetic algorithms[J]. **Rock and Soil Mechanics**, 2002, 23(2): 170—172.
- [20] BANDYOPADHYAY S, MAULIK U. Genetic clustering for automatic evolution of clusters and application to image classification[J]. **Pattern Recognition**, 2002, 35(6): 1197—1208.
- [21] HRUSCHKA E R, EBECKEN N F F. A genetic algorithm for cluster analysis[J]. **Intelligent Data Analysis**, 2003, 7(1): 15—25.
- [22] DAVIES D L, BOULDIN D W. A cluster separation measure[J]. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 1979, PAMI-1(2): 224—227.
- [23] MURTHY C A, CHOWDHURY N. In search of optimal clusters using genetic algorithms[J]. **Pattern Recognition Letters**, 1996, 17(8): 825—832.