



Identification of rock discontinuity sets based on a modified affinity propagation algorithm

Jie Liu, Xing-Dong Zhao*, Zeng-he Xu

School of Resources & Civil Engineering, Northeastern University, Shenyang, China 110819

ARTICLE INFO

Keywords:

Rock discontinuity sets
Cluster analysis
Affinity propagation algorithm
Orientation analysis

ABSTRACT

Identification of rock discontinuity sets is an important foundation for stability analysis of rock engineering applications. A modified affinity propagation (AP) algorithm is proposed for identifying rock discontinuity sets based on discontinuity orientations. The method considers all data points as potential clustering centers simultaneously, which could avoid the hard selection on initial clustering centers as well as achieve the global optimization. Euclidean distance measure in the original algorithm is not suitable for the clustering of discontinuity orientations and therefore a similarity measure based on the negative sine-squared value of the acute angle between discontinuity unit normal vectors was used. Moreover, the Silhouette validity index was introduced to determine the optimal clustering number. The validity of the new method was tested by using artificial data, and in-situ data compiled from the literature. Finally, the proposed method was applied to discontinuity grouping in an underground water-sealed oil storage cavern in Liaoning Province, China. The results show that the new method can effectively filter noisy data and achieve good clustering results with stronger robustness than other methods.

1. Introduction

Discontinuities developed in rock masses during wide varieties of geological processes and these discontinuities at different times overlapped with each other to form fracture networks. The stability of rock slopes, tunnels and underground caverns is strongly influenced by discontinuities. Discontinuity orientation (dip direction and dip angle) is one of the most important geometrical properties. Discontinuity classification based on the orientations is the basis for developing discrete fracture modeling technique and assessing rock mass stability.^{1,2}

Most discontinuity measurements present a mixture of many sets. Among problems faced by rock and civil engineers, one of the important challenges is identification and description of discontinuity sets. The conventional graphical method employed in identifying discontinuity sets is stereographic plot projection of discontinuity poles. This method is intuitive and simple, but interpretation of the stereonets is very subjective due to personal bias and depends greatly on the engineer's experience. Discontinuity grouping cannot be achieved with the conventional graphical method when discontinuity poles are scattered and the class number is enormous. Given the complexity of the problem, one of the most promising approaches is to develop alternative techniques for automatic identification of rock

discontinuity sets.³ Cluster analysis is an exploratory tool of data analysis, which is capable of grouping data into classes without a priori information. Since Shanley and Mahtab⁴ first used the clustering algorithm for identifying discontinuity sets, several different methods such as fuzzy *C*-means (FCM) cluster,^{5–7} *K*-means cluster^{8,9} and spectrum cluster^{3,10} have been widely applied to the statistical analysis of discontinuities. Among of them, FCM and *K*-means belong to dynamic cluster, which are sensitive to initial cluster centers and are easily trapped in a local optimum, and also the number of clusters must be pre-specified. For the spectrum cluster, it is not easy to construct proper similarity matrix and choose reasonable scale parameters, and meanwhile this method needs a large amount of numerical calculations.

In order to obtain a global optimum, Xu et al.¹¹ proposed FCM based on the mutative scale chaos optimization algorithm for classifying discontinuity sets. Ma et al.¹² proposed an enriched *K*-means clustering method with improved initial cluster centers based on a gathering degree function and a hierarchical clustering method. Meanwhile, many researchers introduced various clustering algorithms based on artificial intelligence to implement discontinuity grouping. Cai et al.¹³ and Li et al.¹⁴ proposed FCM based on the genetic algorithm for automatic identification of joint sets, but the genetic algorithm tends to produce premature convergence. Li et al.¹⁵ utilized the ant

* Corresponding author.

E-mail address: zhaoxingdong@mail.neu.edu.cn (X.-D. Zhao).

colony algorithm to cluster discontinuity data successfully. Unlike the conventional ant colony algorithm, Song et al.^{16,17} proposed the effective discontinuity grouping method based on the firefly algorithm and artificial bee colony algorithm, respectively. However, the swarm intelligence methods are complicated in theory and need some input parameters whose reasonable values are not easily determined.

Therefore, it is necessary to seek a clustering method with rapid convergence, good global search capability, simple and convenient implementation for engineering applications. Fortunately, the affinity propagation (AP) algorithm has been successfully employed in applications including face recognition, gene discovery, text mining and image segmentation.^{18–20} The AP algorithm was originally proposed by Freyh and Dueck.²¹ Compared with conventional clustering methods, this method is insensitive to initial cluster centers and is able to achieve a global optimum. In this paper, a modified AP algorithm was proposed for grouping discontinuity sets based on characteristics of discontinuity orientations. First, the new method was described. Then the performance of this algorithm was evaluated by using artificial data and in-suit data compiled from a literature. Furthermore, the method was applied to data sets in a practical engineering and the clustering results were compared with the other clustering methods.

2. AP algorithm

AP algorithm is an unsupervised learning algorithm based on message-passing techniques and it aims at finding optimal clustering centers by information exchange between pairs of data points to make the maximum sum of similarities between all points and the nearest centers. AP algorithm takes the similarities between pairs of data points as input. All data points are simultaneously considered as potential clustering centers, which is able to avoid poor solutions caused by unlucky initializations. Each data point is regarded as a node in the network. Messages transmit along the node connection recursively. AP algorithm searches for clusters through an iterative process until a high-quality set of exemplars and corresponding clusters emerge. In this method, the clustering centers selected from actual data points are called exemplars.

2.1. Representation of discontinuity orientations

The discontinuity orientation in field geological survey is normally represented by dip direction α_d and dip angle β_d . Suppose discontinuities can be regarded as planar structures, the orientation of a discontinuity can be expressed by a pole on the surface of a unit sphere (see Fig. 1a). The pole is the point that the downward unit vector normal to the discontinuity plane intersects with the reference sphere of unit radius. In the right-handed Cartesian coordinate system as shown in Fig. 1, the positive x_1 -axis is horizontal to the north; the positive x_2 -axis is horizontal to the east and the positive x_3 -axis is vertically downward. The trend α_n of the normal vector is the angle between the positive x_1 -axis and the projection of the dip direction vector on the x_1 - x_2 plane, measured in clockwise rotation from north. The plunge β_n is the angle between the pole vector and x_1 - x_2 plane. α_n and β_n are related to α_d and β_d as follows: $\alpha_n = \alpha_d \pm 180^\circ$ ($0^\circ \leq \alpha_n \leq 360^\circ$) and $\beta_n = 90^\circ - \beta_d$ ($0^\circ \leq \beta_n \leq 90^\circ$). The orientation $\Theta = (\alpha_n, \beta_n)^T$ of a pole vector can be expressed by its three-dimensional Cartesian coordinates $\mathbf{X} = (x_1, x_2, x_3)^T$, where

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

In this paper, the lower hemisphere projection was adopted to represent the orientation of structural planes. In this projection method, the unit vector of the planar structure is downward and therefore the endpoint A is located in the lower hemisphere. A line is

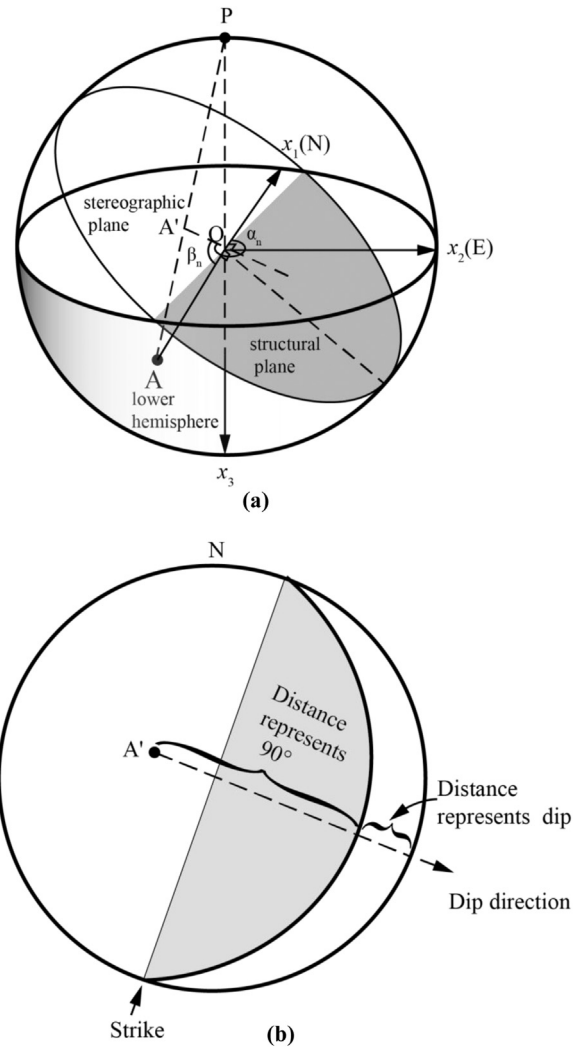


Fig. 1. Representation of discontinuity orientations (N denotes North; E denotes East): (a) spherical projection; (b) plane view of hemispherical projection (modified from Klose et al.²² and Borradaile²³).

drawn from the upper pole P to the point A , intersecting the x_1 - x_2 plane (stereographic plane) at the point A' . The projection A' on the x_1 - x_2 plane is called a polar stereographic, or hemispherical plot (Fig. 1). To plot the points representing discontinuities is a relatively good method for large numbers of discontinuity orientations.

2.2. Similarity measures

The selection of a distance (similarity) between discontinuity orientations is a key step for successfully partitioning data into groups. In the original AP algorithm, the similarity between data points is set to a negative squared error (Euclidean distance), indicating that the similarity is larger if the distance between points is closer. However, similarity can be set depending on the specific issues without meeting Euclidean space constraints in actual applications. For example in Fig. 2, a pair of discontinuities with both steeply inclined angles and dip directions 180° apart makes a small acute angle with each other, which should belong to the same set. If the Euclidean distance or spherical distance between discontinuity unit normal vectors is adopted as the similarity, the distance A_1A_2 is very larger and two discontinuities may be mistaken as two different sets.

To solve the above problem, the negative sine-squared value of the acute angle between discontinuity unit normal vectors is chosen as the similarity. The sin-based similarity measure has some advantages that

Download English Version:

<https://daneshyari.com/en/article/5020203>

Download Persian Version:

<https://daneshyari.com/article/5020203>

[Daneshyari.com](https://daneshyari.com)