

ALGORITMO PARA IDENTIFICAÇÃO AUTOMÁTICA DE AGRUPAMENTOS DE DESCONTINUIDADES BASEADO EM MÚLTIPLAS PROPRIEDADES DE CARACTERIZAÇÃO.

Luiz Filipe Rodrigues Viana¹; Arquimedes Martins Gois²; André Monteiro Klen³.

1 Luiz Filipe Rodrigues Viana, Bolsista IFMG, Engenharia Geológica, Universidade Federal de Ouro Preto, Ouro Preto - MG; luiz.viana1@aluno.ufop.edu.br.

2 Arquimedes Martins Gois, IFMG – Campus Ouro Preto, Ouro Preto – MG

3 André Monteiro Klen: Pesquisador do IFMG, Campus Ouro Preto; andre.klen@ifmg.edu.br.

RESUMO

O agrupamento de descontinuidades em famílias é uma das tarefas fundamentais quando se busca avaliar os mecanismos de ruptura de um maciço rochoso. As características das descontinuidades e o modo como as famílias se interceptam controlam a permeabilidade, diminuem a resistência da rocha, reduzem a estabilidade e controlam a queda de blocos. As descontinuidades podem ser caracterizadas por mais de dez propriedades, porém, na maioria dos casos, apenas a orientação espacial é utilizada como critério para agrupá-las e determinar suas atitudes médias. Este critério por ser indispensável, é adotado tanto no método de projeção estereográfica, quanto em diversas técnicas numéricas. Contudo, estudos demonstram que a orientação espacial explica aproximadamente 45,6% da variabilidade total do conjunto de dados, este fato indica que, em determinadas situações somente essas informações não são suficientes para classificar e compreender os grupos formados. O trabalho tem como objetivo estudar a influência dessas propriedades adicionais na definição das famílias de descontinuidades e como o algoritmo de agrupamentos K-means pode ser aplicado nesses casos. Para isso, é utilizado um banco de dados artificial que permite analisar e confrontar os resultados do K-means em relação aos métodos que adotam apenas a orientação espacial como critério de classificação, neste estudo em particular, o método de projeção estereográfica e o algoritmo Nébula desenvolvido no *software* livre *Scilab*. Para a execução do K-means, tratamento, análise dos dados e criação dos gráficos 3D é utilizado o *software* livre R. Os resultados demonstram que as propriedades adicionais são a chave para que o K-means consiga identificar corretamente as famílias, enquanto que, os demais métodos falham ao não distinguir os agrupamentos que possuem orientações espaciais semelhantes. Os resultados ainda demonstram a aplicabilidade do algoritmo K-means e a importância de se empregar a visualização em três dimensões para classificar as descontinuidades em famílias e apontam para um melhor entendimento dos mecanismos de ruptura dos maciços rochosos.

Palavras chaves: Estabilidade de maciços rochosos; Agrupamento de descontinuidades; Algoritmo K-means; Projeção estereográfica.

INTRODUÇÃO:

A definição das descontinuidades em agrupamentos, famílias, e a caracterização de suas atitudes médias é um dos aspectos mais importantes em aplicações de engenharia geotécnica, de petróleo, civil e de minas, uma vez que, essas constituem os planos de fraqueza da rocha e seus conjuntos definem um sistema que rege os mecanismos e as geometrias de ruptura dos maciços rochosos (KLEN e LANA, 2014); (SONG *et al.*, 2016).

As descontinuidades podem ser de diversos tipos como: xistosidade, dobras, fraturas, juntas, falhas, veios, planos de acamamento, planos de foliação. Os principais parâmetros que abrangem as propriedades das descontinuidades são: orientação espacial, espaçamento, alteração, persistência, abertura, preenchimento, material do preenchimento, rugosidade, entre outros (FOSSEN, 2017).

A Figura 1 ilustra as descontinuidades de um maciço rochoso que podem ser agrupadas em duas famílias distintas. As setas vermelhas indicam as descontinuidades pertencentes à família subvertical, enquanto que, as setas azuis identificam aquelas que compõem a família sub-horizontal.



Figure 1: Representação de descontinuidades em um maciço rochoso.
Fonte: Adaptado de Lana *et al.* (2019).

As descontinuidades de um maciço rochoso são consideradas como estruturas planares, identificadas por suas orientações espaciais e retratadas por seus polos. A orientação espacial é expressa em termos de dip direction, direção de mergulho, ($0^\circ \leq \alpha \leq 360^\circ$) e dip angle, mergulho, ($0^\circ \leq \beta \leq 90^\circ$). O dip é o ângulo de maior declividade do plano com a horizontal. Enquanto que, o dip direction é a direção desta declividade em relação ao norte geográfico (FOSSSEN 2017).

Conforme Jimenez e Sitar (2006) e Gao *et al.* (2019), a definição dos agrupamentos envolve a projeção estereográfica das descontinuidades, a partir da qual, se constrói o diagrama de frequência de polos e as curvas de isofrequência. Essa técnica, amplamente empregada, não é satisfatória em alguns casos, pois se baseia na interpretação visual, e por isso é influenciada por fatores subjetivos, os quais podem conduzir a resultados divergentes. Os principais fatores são: a experiência profissional, o tamanho da célula de contagem, a presença de sobreposição e *outliers* entre os conjuntos e a definição do número de agrupamentos.

Com o intuito de minimizar esses fatores, diversos algoritmos apoiados em técnicas numéricas e estatísticas foram desenvolvidos para a identificação automática dos conjuntos de descontinuidades. Dentre esses a análise de *cluster* é a ferramenta mais aplicada. Essa técnica é capaz de agrupar dados em conjuntos sem informações *a priori* de suas estruturas. Para classificar as descontinuidades é necessário um algoritmo de agrupamento e os mais utilizados são o *K-means* e sua variação *Fuzzy K-means*, principalmente após os estudos de Hammah e Curran (1998).

Porém, o método de projeção estereográfica e a grande maioria dos algoritmos descritos na literatura científica, utilizam apenas a orientação espacial como critério para identificar as famílias de descontinuidades. A adoção de apenas esta propriedade, em alguns casos, pode não ser suficiente para compreender o sistema de agrupamento e, conseqüentemente, os mecanismos de estabilidade do maciço rochoso.

Diante da relevância da aplicação das técnicas numéricas para classificação das descontinuidades e a necessidade de uma maior adequação entre os dados observados em campo e aqueles alcançados por essas técnicas, se torna fundamental a adição de outras propriedades para se caracterizar as famílias de descontinuidades.

A introdução de novas propriedades aumenta a complexidade da classificação, uma vez que, esta tarefa deixa de ser bidimensional e se torna multidimensional. Este fato tem forte implicação, pois mesmo para dados multidimensionais, a projeção estereográfica deve ser utilizada para se interpretar os grupos formados e estudar como estes afetam a estabilidade do maciço rochoso.

Nesse ponto é que o K-means é importante, pois ele é um algoritmo de classificação adequado para dados multidimensionais. Após a sua aplicação os agrupamentos podem ser visualizados através de projeções estereográficas e gráficos 3D. Para maiores informações sobre o algoritmo K-means e sua aplicação sugere-se os trabalhos de Tokhmechi *et al.* (2011) e Li *et al.* (2015).

METODOLOGIA:

Para avaliação e melhor entendimento do problema apresentado o K-means tem seus resultados testados em relação à técnica de projeção estereográfica e ao algoritmo de agrupamento Nébulas desenvolvido por Klen e Lana (2014) e aprimorado por Klen e Carvalho (2022). Este algoritmo é elaborado para a classificação de descontinuidades, tendo como critério a orientação espacial.

Ele é baseado no método de classificação conhecido como Fuzzy K-means (FKM), que é uma variação do algoritmo K-means, sendo que, a principal característica é o fato dele ser menos sensível à presença de sobreposição e *outliers* entre os agrupamentos. Como ponto de referência para análise do método FKM e ao algoritmo Nébulas, sugere-se os trabalhos supracitados e de Hammah e Curran (1998).

Para esses testes são geradas quatro famílias artificiais de juntas com 50 elementos cada, e quatro propriedades de caracterização: (dip, dip direction, tipo de material do preenchimento e porcentagem de preenchimento). Conforme Fossen (2017), as juntas são fraturas constituídas por duas paredes separadas por alguns centímetros.

Estas paredes podem ser preenchidas ou não, sendo que, o tipo de preenchimento tem papel importante no comportamento do maciço rochoso. As juntas e suas propriedades são fundamentais quando se analisa um maciço, visto que, elas controlam a permeabilidade, diminuem a resistência da rocha, reduzem a estabilidade e controlam a queda de blocos.

A metodologia empregada para gerar as variáveis, dip, dip direction é baseada nos estudos de Hammah e Curran (1998) e utiliza a distribuição de probabilidades de Fisher (1958) com a aplicação do algoritmo proposto por Zheng *et al.* (2014). Os parâmetros da distribuição de Fisher são descritos na Tabela 1. Já as demais variáveis seguem as proposições de Tokhmechi *et al.* (2011). A Tabela 2 apresenta as características do conjunto de dados e mostra que as famílias 2 e 3 possuem orientações semelhantes, contudo são distintas em relação as demais propriedades. As famílias são representadas pelas cores indicadas.

Tabela 1: Parâmetros da distribuição de Fisher para geração das famílias de fraturas.

Família de Juntas	Atitude média Dip/Direction	K*
1	57/47	21
2	60/205	41,5
3	69/217	32,5
4	87/312	21,5

k* Parâmetro de concentração da distribuição de Fisher

Tabela 2: Características do conjunto de dados

Propriedades das famílias de Juntas	1 ●	2 ●	3 ●	4 ●
Dip (graus): min-max	27-89	37-80	41-85	58-90
Dip direction (graus): min-max	25-232	186-237	42-311	111-346
Material do preenchimento *	1	0	1	0,5
Porcentagem de preenchimento**	0,5	0	1	0,5

*1 junta preenchida com calcita, 0,5 com argila e 0 sem preenchimento. **1 total preenchimento da junta, 0,5 apenas metade e 0 ausência de preenchimento.

Para identificar os agrupamentos de juntas, aplica-se o algoritmo de agrupamento K-means através do *software* livre R versão 4.3.1 e a biblioteca *factoextra*. O R também é utilizado para a criação dos gráficos e análises estatísticas. Para a projeção estereográfica aplica-se o *software* Dips da Rocscience versão 5.0, e para execução do Nébula o *software* livre *Scilab* versão 2023.1.0.

Para que o algoritmo K-means seja utilizado para identificar o conjunto de juntas é preciso que os dados sejam adaptados. Basicamente três etapas devem ser seguidas. Na primeira, toda variável dip direction que tem valor maior ou igual a 180° deve ser diminuída de 180°. Isso possibilita que o K-means identifique corretamente as famílias subverticais mesmo utilizando a distância euclidiana. Para informações adicionais sobre esse assunto sugere-se o trabalho de Hammah e Curran (1998).

A segunda etapa consiste em representar as variáveis dip, dip direction através de seus vetores unitários. A orientação de uma descontinuidade i qualquer, pode ser representada por um vetor normal unitário, o qual é determinado pelo cosseno diretor $e_i = (x_i, y_i, z_i)$. A Figura 2, mostra o vetor normal unitário da descontinuidade i no sistema de coordenadas cartesianas. O eixo positivo x é horizontal à direção norte, o eixo positivo y é horizontal a direção oeste e o eixo positivo z é direcionado para cima (GAO *et al.*, 2019); (LI *et al.*, 2014).

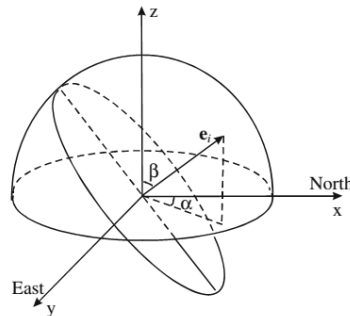


Figura.2: Representação da orientação da descontinuidade.
Adaptado de Li *et al.* (2014).

Onde: α , representa o dip direction e β o dip.

$$e_i = \begin{cases} x_i = \cos\alpha \sin\beta \\ y_i = \sin\alpha \sin\beta \\ z_i = \cos\beta \end{cases}$$

Já a terceira etapa consiste na padronização das variáveis, porém esta deve ser aplicada quando as variáveis apresentam discrepância na ordem de grandeza. No caso em estudo não é necessário a padronização.

A Tabela 3 apresenta um resumo do conjunto de dados de entrada, X, Y e Z, representam os vetores unitários das coordenadas dip e dip direction de cada uma das juntas, MP representa o material do preenchimento e PP% a porcentagem de preenchimento.

Em destaque, observam-se os dados que o K-means utiliza para identificar os agrupamentos. Estes dados, assim como o script do R estão disponíveis no link: <https://github.com/andreklein/KM-in-R.git>.

Tabela 3: Resumo do conjunto de dados.

Dip	Direction	X	y	z	MP	PP%
66	49	0,689461509	0,599339746	0,406736643	1	0,5
75	27	0,438521149	0,860646213	0,258819045	1	0,5
50	63	0,682550597	0,3477769	0,64278761	1	0,5
53	56	0,662098845	0,44659131	0,601815023	1	0,5
62	60	0,764655046	0,441473796	0,469471563	1	0,5

RESULTADOS E DISCUSSÕES:

A Figura 3 mostra os agrupamentos obtidos, do conjunto de dados artificial, pelos diferentes métodos. A Figura 3 (b) apresenta a projeção estereográfica dos dados artificiais e as diferentes famílias. Note que, as famílias 2 e 3 são representadas em grupos distintos, porém que se sobrepõem devido a valores comuns de orientação espacial.

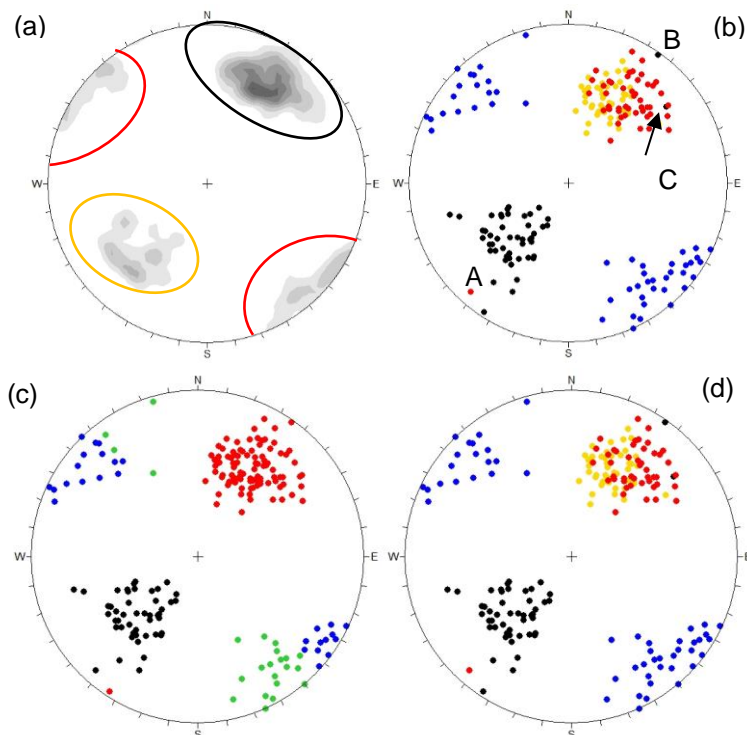


Figura 3: Projeções estereográficas do conjunto de dados artificial. (a) Curvas de isofrequência. (b) Identificação das famílias. (c) Resultado do Nébulas. (d) Resultado do K-means. As projeções estereográficas são representadas no hemisfério inferior com igual área e tamanho da célula de contagem igual 0.5%.

A Figura 3 (a) demonstra a projeção estereográfica e as curvas de isofrequências do conjunto de descontinuidades. A interpretação visual revela apenas três agrupamentos, e que não é possível separar as

famílias 2 e 3. Esse resultado se deve ao fato do método utilizar apenas as orientações espaciais das juntas como critério de particionamento.

O mesmo ocorre com os grupos definidos pelo algoritmo Nébulas, Figura 3 (c), ao indicar para o algoritmo que o conjunto de dados possui 4 famílias, ele procura por agrupamentos que sejam heterogêneos entre si e homogêneos internamente e, por isso, não separa as famílias 2 e 3. Por outro lado, o Nébulas divide a família 4 em duas, justamente por ela ser mais dispersa que as demais.

Quando as famílias 2 e 3 são classificadas no mesmo agrupamento, implica em um entendimento diverso do maciço rochoso, pois essas famílias interferem de maneira oposta na estabilidade do mesmo. Por exemplo, as propriedades MP e PP% da família 2, mostram que este é um grupo de juntas abertas e sem preenchimento.

Estas aberturas facilitam a percolação da água e o crescimento de raízes ao longo de suas paredes, levando a ocorrência de ações que são desfavoráveis à estabilidade do maciço rochoso como: fenômenos de natureza erosiva, físico químicas e deslocamento de blocos. Por outro lado, a família 3 confere ao maciço características de maior estabilidade, justamente pelo fato de possuir as aberturas totalmente preenchidas com calcita.

A Figura 3 (d), mostra a projeção estereográfica dos resultados do método K-means e as Figuras 4 (a, b) os resultados em 3D. A Figura 3 (d) mostra que o K-means, identifica os 4 agrupamentos e, sobretudo, distingue as famílias 2 e 3. Observe nas Figuras 4 (a, b), que apesar dessas famílias possuírem orientações similares o algoritmo as separa, tendo como critério as propriedades MP e PP%.

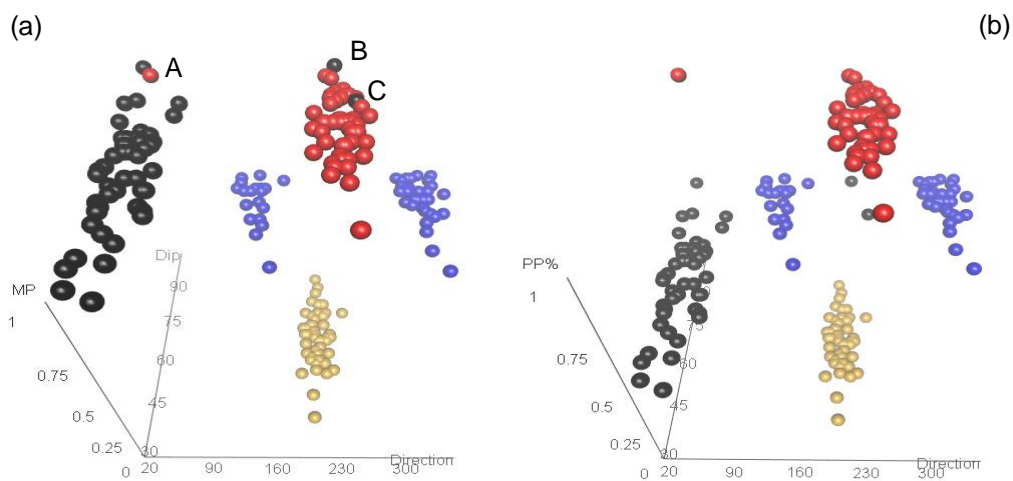


Figura 4: Resultados do K-means em 3D. (a) Agrupamentos em relação ao MP. (b) Agrupamentos em relação ao PP%.

Os gráficos em 3D, são muito importantes no entendimento da relação entre as propriedades e a identificação das famílias. Veja como é possível verificar como as famílias 2 e 3 possuem orientações próximas. Outro exemplo, são as amostras identificadas pelas letras A, B e C, Figura 3 (b). A amostra A é agrupada corretamente na família 3, Figura 3 (d), porém ao se observar as Figuras 4 (a, b) ela se mostra próxima ao grupo 1.

Isso ocorre porque, a amostra A possui o dip direction semelhante aos valores da família 1. No entanto, ela compartilha as demais propriedades com família 3. A mesma explicação se estende as amostras B e C. Note que os demais métodos não conseguem classificar adequadamente essas amostras. É importante destacar que, muito provavelmente, a justa distinção ou não dessas amostras tenha pouca influência na prática, porém neste estudo, demonstra a capacidade do algoritmo de classificar corretamente os dados.

CONCLUSÕES:

A tarefa de classificar as descontinuidades em famílias ou agrupamentos é uma das atividades principais para a compreensão do comportamento e estabilidade de um maciço rochoso. Nem sempre essa prática é simples e diversas metodologias são empregadas. O método de projeção estereográfica é o mais utilizado e recentemente seu uso vem acompanhado da aplicação de algoritmos de agrupamentos.

Contudo, a projeção estereográfica e a maioria dos métodos numéricos são bidimensionais e utilizam como critério de classificação apenas a orientação espacial das descontinuidades. Este critério, em algumas situações, conduz a erros de interpretação que podem alterar a análise do maciço rochoso.

Para ajudar na solução dessa questão o trabalho propôs a utilização de outras propriedades de caracterização das descontinuidades, além da orientação espacial, para identificar as famílias de descontinuidades. Para isso, foi aplicado o algoritmo de agrupamentos K-means a um banco de dados artificiais e seus resultados foram comparados com os métodos de projeção estereográfica e o algoritmo Nébula.

Os agrupamentos encontrados pelo K-means foram condizentes com o banco de dados artificial, enquanto os demais métodos não diferenciaram os agrupamentos que compartilham orientações espaciais semelhantes. Além disso, o emprego do K-means permitiu análise, em três dimensões, das características dos grupos formados e de seus elementos. Isto ajuda os geólogos e engenheiros em suas análises durante o desenvolvimento de projetos nas áreas de construção civil, mineração e extração de petróleo.

Esse projeto de pesquisa proporcionou um avanço no que tange esse tema e para que os resultados sejam mais próximos da realidade é importante à aplicação do método a um banco de dados real com um número maior de propriedades como: espaçamento, alteração, persistência, abertura, preenchimento, material do preenchimento, rugosidade etc. Esta é a próxima etapa do projeto de pesquisa, uma vez que este ainda não foi finalizado e os resultados são parciais.

REFERÊNCIAS BIBLIOGRÁFICAS:

FISHER, R. A. Dispersion on a sphere. **Proceedings of the Royal Society of London**, v. A, n. 217, p. 295-305, 1958.

FOSSEN, H. **Geologia estrutural**. 2. Ed. Sao Paulo: Oficina de textos, 2017.

GAO, Feng et al. A Fast-Clustering Method for Identifying Rock Discontinuity Sets. **KSCE Journal of Civil Engineering**, v. 23, n. 2, p. 556-566, 2019.

HAMMAH, R. E.; CURRAN, J. H. Fuzzy cluster algorithm for the automatic identification of joint sets. **International Journal of Rock Mechanics and Mining Sciences**, v. 35, n. 7, p. 889-905, 1998.

JIMENEZ-RODRIGUEZ, R.; SITAR, N. A spectral method for clustering of rock discontinuity sets. **International Journal of Rock Mechanics and Mining Sciences**, v. 43, n. 7, p. 1052-1061, 2006.

KLEN, A. M.; LANA, M. S. Fuzzy algorithm of discontinuity sets. **REM: R. Esc. Minas**, Ouro Preto, 67(4), p. 439-445, oct. dec. 2014.

KLEN, A., M.; CARVALHO, P., G. **Algoritmo para agrupamento automatico de descontinuidades baseado no metodo Fuzzy K-means**. 17º Congresso Brasileiro de Geologia de Engenharia e Ambiental. 2022. Disponível em https://schenautomacao.com.br/cbge2022/envio/files/trabalho1_36.pdf Acesso em 05/07/2023.

LANA, MILENE SABINO; ALVES, J. M. P.; KLEN, A. M. **Definition of discontinuity sets by cluster analysis: A case study In: Definition of discontinuity sets by cluster analysis: A case study**. 1 ed. London: A Balkema Book, v.6, p. 637-644, 2019.

LI, Y.; WANG, Q.; CHEN, J.; XU, L.; SONG, S. K-means algorithm based on particle swarm optimization for the identification of rock discontinuity sets. **Rock mechanics and Rock Engineering**, 48 p. 375-385, 2015.

SONG, S.; WANG, Q.; CHEN, J.; LI, Y.; ZHANG, W.; RUAN, Y. Fuzzy C-means clustering analysis based on quantum particle swarm optimization algorithm for the grouping of rock discontinuity sets. **KSCE Journal of Civil Engineering**, 21(4), p. 556-566, 2016.

TOKHMECHI, B.; MEMARIAN, H.; MOSHIRI, B.; RASOULI, V.; NOUBARI, H., A. Investigating the validity of conventional join set clustering methods. *Engineering Geology*, 118, p.75-81, 2011.

ZHENG, J.; DENG, J.; YANG, X.; WEI, J. A improved Monte Carlo simulation method for discontinuity orientations based on Fisher distribution and its program implementation. **Comp. Geotech.**, v. 62, p.266-276, 2014.