

Análise de Agrupamentos Aplicada à Parâmetros Geomecânicos com vistas à Avaliação da Estabilidade de Taludes de Mina

Allan Erlikhman Medeiros Santos

Universidade Federal de Ouro Preto, Ouro Preto, Brasil, allanboni@hotmail.com

Milene Sabino Lana

Universidade Federal de Ouro Preto, Ouro Preto, Brasil, milene@demin.ufop.br

Ivo Eyer Cabral

Universidade Federal de Ouro Preto, Ouro Preto, Brasil, cabralmg@uol.com.br

Tatiana Barreto dos Santos

Universidade Federal de Ouro Preto, Ouro Preto, Brasil, tati.barreto12@gmail.com

Denise de Fátima Santos da Silva

Universidade Federal de Ouro Preto, Ouro Preto, Brasil, denisefss@yahoo.com.br

RESUMO: O trabalho apresenta a aplicação de uma técnica da análise estatística multivariada a um banco de dados geotécnico em taludes de mina com o objetivo de avaliar as condições de estabilidade dos mesmos. A técnica da análise estatística multivariada aplicada para esta pesquisa é a análise de agrupamentos (cluster analysis) que pode ser utilizada quando se deseja identificar as similaridades entre indivíduos definindo-os em grupos. Para o problema a ser analisado, as condições de estabilidade dos taludes de mina, o trabalho pode auxiliar na proposição de ferramentas de análise de risco e hierarquização dos taludes de mina. O banco de dados em estudo se trata de um levantamento de parâmetros geomecânicos de oitenta e quatro taludes operacionais no mundo inteiro, sendo cinco destes no Brasil. Este banco de dados foi estruturado em dezoito variáveis relacionadas ao ambiente de localização, qualidade da rocha, propriedades do maciço, tensões in-situ, condições hidráulicas, propriedades das descontinuidades, histórico de deslizamentos, método de construção e geometria da cava. Logo, cada talude apresenta valores relativos às dezoito variáveis relacionadas aos parâmetros citados anteriormente. Uma vez que nestes taludes são conhecidas as condições reais de estabilidade, registradas também no banco de dados, a aplicação da técnica de agrupamentos pode ser comparada com a situação real resultando em uma validação da técnica utilizada. A aplicação da análise de agrupamentos a este banco de dados permite identificar os taludes que são similares e assim estabelecer grupos de acordo com a similaridade com base nos valores das dezoito variáveis. Para a pesquisa os grupos seriam relacionados às condições de estabilidade, e os indivíduos que definiriam os grupos seriam os taludes. Na aplicação da técnica foram definidos dois grupos, que seriam os taludes estáveis e os taludes instáveis, a partir disso os resultados obtidos foram comparados com a real estabilidade dos taludes. Os resultados foram de 86,36% de acerto para os taludes instáveis e 90,00% de acerto para os talude estáveis. Estes resultados são representativos, pois dos 84 casos estudados, apenas 10 casos foram classificados de forma equivocada, o que representa cerca de 88% de acerto no total. Portanto as aplicações da análise de agrupamentos com o objetivo de avaliar condições de estabilidade de taludes se mostrou eficaz obtendo uma dispersão pequena em relação ao número de dados e variáveis utilizadas no estudo.

PALAVRAS-CHAVE: Estabilidade de taludes operacionais, análise multivariada, análise de agrupamentos, banco de dados geotécnico, hierarquização de taludes.

1 INTRODUÇÃO

A análise de agrupamentos, conhecida também como análise de conglomerados, é uma técnica estatística multivariada que consiste em agrupar amostras que são similares em um mesmo grupo. A similaridade é medida de acordo com os valores das variáveis das amostras, podendo ser medida de diversas formas a depender da natureza das variáveis em estudo.

A técnica de agrupamento pode ser útil para a análise de dados em muitas situações, como a redução da dimensão de um conjunto de dados, restringindo uma ampla gama de objetos à informação do centro do seu conjunto (Linden, 2009). Além de permitir a definição e caracterização de grupos que poderão ser pré-definidos ou definidos ao se analisar os resultados da técnica.

Para o presente trabalho a técnica é utilizada para estudo das condições de estabilidade dos taludes de um banco de dados universal, com o objetivo de agrupá-los segundo suas características (variáveis medidas). Os grupos são pré-definidos pois a informação sobre a estabilidade é conhecida, e o resultado do agrupamento é comparado com a situação real. Assim faz-se possível a validação do agrupamento a partir da porcentagem de acerto.

O banco de dados em estudo contém informações geotécnicas de 84 taludes. No total são 18 variáveis, além da classificação real do talude sobre a sua estabilidade, ou seja o banco de dados possui 1512 informações e 84 classificações de estabilidade. Estes dados foram coletados de trabalhos e estudos em taludes operacionais publicados no mundo todo, contendo taludes do Brasil, Canadá, Austrália, Iran, África do Sul, Espanha, Suécia, Estados Unidos e Nova Guiné. Nesse sentido o banco de dados se torna representativo e consistente para aplicação das técnicas estatísticas multivariadas. A Tabela 1, resume as características dos dados.

Tabela 1. Origem dos taludes do banco de dados. (modificada de Naghadehi, 2013).

Nome da mina	País	Casos
Sandsloot	África do Sul	6
Venetia	África do Sul	7
Cadia Hill	Austrália	5
Águas Claras	Brasil	5
Panda	Canadá	1
Chuquicamata	Chile	5
Escondida	Chile	7
Ujina	Chile	1
Aznalcollar	Espanha	5
La Yesa	Espanha	2
Esperanza	EUA	1
Betze-Post	EUA	4
Angooran	Iran	4
Chadormalou	Iran	5
Choghart	Iran	5
Gole-Gohar	Iran	4
Sarcheshmeh	Iran	4
Sungun	Iran	4
Ok Tedi	Nova Guiné	2
Aitik	Suécia	6

As variáveis selecionadas para o estudo são classificadas segundo Hudson (1992), que propôs uma atlas categórico de fatores gerais que afetam a estabilidade de taludes e encostas rochosas. As variáveis são tipo de rocha, precipitação, resistência da rocha intacta, RQD, alteração, regime tectônico, condições de água subterrânea, número de famílias de descontinuidade, persistência, espaçamento, orientação, abertura, rugosidade, preenchimento, altura do talude, ângulo do talude, método de desmonte e a presença de eventos de instabilidade no passado.

2 METODOLOGIA

2.1 Método de Classificação

Segundo Davis (1986) existem quatro tipos de método de classificação, sendo eles os métodos de partição, de origem arbitrária, de similaridade mútua e o hierárquico. Neste trabalho utilizaremos o método hierárquico que segundo Landim (2011) é um dos mais utilizados na geologia.

O método se inicia pelo cálculo dos coeficientes de similaridade do conjunto de

dados original, ou seja, calcula-se os coeficientes de similaridade entre os 84 taludes, obtendo-se uma matriz de ordem 84. Criada a matriz de similaridade a partir do banco de dados original, o segundo passo é a seleção dos pares com mais alta similaridade, neste caso, tem-se a similaridade baseada na distância de Manhattan, logo quanto menor a distância entre os pontos, maior será a similaridade.

Assim, em cada passo os taludes com maior similaridade são agrupados, e os grupos com maior similaridade são arranjados até que seja alcançado o número de grupos almejado. O gráfico que apresenta esses agrupamentos é conhecido como dendograma, que apresentará as relações entre os grupos encontrados.

2.2 Medida de Similaridade

Como medida de similaridade foi utilizada a distância de Manhattan que está representada na Equação 1. Observa-se que se trata de uma equação genérica onde se adota $\lambda = 1$ para se obter a equação de similaridade pela distância de Manhattan.

$$d(X_i, X_k) = \left[\sum_{i=1}^p w_i |X_{ii} - X_{ik}|^\lambda \right]^{1/\lambda} \quad (1)$$

Para a Equação 1 tem-se o w_i como o peso de ponderação das variáveis. O número de observações, ou taludes, são denominados pelos índices l a k . Os índices i a p representam as variáveis do banco de dados.

2.3 Técnica de Agrupamento

As técnicas de agrupamentos mais utilizadas são a ligação simples, ligação completa, agrupamento pareado proporcionalmente ponderado, agrupamento pareado igualmente ponderado e o método pela variância mínima.

A metodologia aplicada utilizou a ligação completa, pelo fato de que a similaridade entre dois conglomerados é definida pelos elementos que são “menos semelhantes”, o que implica em um processo com mais restrição na seleção dos grupos direcionando aos melhores resultado

2.4 Validação dos Resultados

A última etapa da metodologia aplicada foi a comparação dos resultados do agrupamento com a situação real dos taludes. Os taludes agrupados de maneira equivocada foram considerados como erro do agrupamento.

3 RESULTADOS E DISCUSSÃO

3.1 Análise do processo de agrupamentos

De acordo com a metodologia proposta, a análise de agrupamentos foi realizada, utilizando como medida de similaridade a distância de Manhattan, método de agrupamento hierárquico e por ligação completa.

A Tabela 2 apresenta os resultados de cada passo no agrupamento, se tratando de um agrupamento hierárquico, a análise começa com 84 clusters individualizados. No primeiro passo agrupa-se o primeiro par obtendo assim 83 clusters, com nível de similaridade de 95,29% com respectiva distância de 0,40. A medida que a distância aumenta, a similaridade tende a se reduzir, o que pode ser observado em cada passo subsequente.

Tabela 2. Processo de agrupamento

Passo	Número de clusters	Similaridade	Distância
1	83	95.29	0.40
2	82	92.94	0.60
3	81	91.76	0.70
4	80	91.76	0.70
5	79	91.76	0.70
6	78	91.76	0.70
7	77	90.59	0.80
8	76	89.41	0.90
9	75	88.24	1.00
10	74	87.06	1.10
11	73	87.06	1.10
12	72	87.06	1.10
13	71	87.06	1.10
14	70	85.88	1.20
15	69	85.88	1.20

16	68	84.71	1.30
17	67	83.53	1.40
18	66	83.53	1.40
19	65	83.53	1.40
20	64	83.53	1.40
21	63	82.35	1.50
22	62	82.35	1.50
23	61	82.35	1.50
24	60	82.35	1.50
25	59	82.35	1.50
26	58	81.18	1.60
27	57	81.18	1.60
28	56	80.00	1.70
29	55	78.82	1.80
30	54	78.82	1.80
31	53	78.82	1.80
32	52	78.82	1.80
33	51	78.82	1.80
34	50	77.65	1.90
35	49	77.65	1.90
36	48	76.47	2.00
37	47	76.47	2.00
38	46	75.29	2.10
39	45	74.12	2.20
40	44	74.12	2.20
41	43	74.12	2.20
42	42	74.12	2.20
43	41	72.94	2.30
44	40	72.94	2.30
45	39	71.76	2.40
46	38	71.76	2.40
47	37	70.59	2.50
48	36	70.59	2.50
49	35	69.41	2.60
50	34	69.41	2.60
51	33	69.41	2.60
52	32	67.06	2.80
53	31	67.06	2.80
54	30	65.88	2.90
55	29	64.71	3.00
56	28	63.53	3.10
57	27	62.35	3.20
58	26	62.35	3.20
59	25	61.18	3.30
60	24	60.00	3.40
61	23	60.00	3.40

62	22	58.82	3.50
63	21	57.65	3.60
64	20	56.47	3.70
65	19	52.94	4.00
66	18	52.94	4.00
67	17	52.94	4.00
68	16	47.06	4.50
69	15	47.06	4.50
70	14	47.06	4.50
71	13	44.71	4.70
72	12	43.53	4.80
73	11	43.53	4.80
74	10	43.53	4.80
75	9	42.35	4.90
76	8	41.18	5.00
77	7	36.47	5.40
78	6	28.24	6.10
79	5	25.88	6.30
80	4	23.53	6.50
81	3	20.00	6.80
82	2	14.12	7.30
83	1	0.00	8.50

A Tabela 3 apresenta os resultados dos números de observações em cada cluster, além das distâncias média e máxima dos centroides de cada cluster, pode-se concluir que os centroides apresentam valores similares, o que era esperado, pois o banco de dados apresenta o mesmo número de taludes estáveis e instáveis.

Tabela 3. Informações dos clusters

Cluster	Número de observações	Distância	
		Média do centroide	Máxima do centroide
1	44	0.98	1.20
2	40	0.94	1.24

3.2 Dendograma

O dendograma obtido, está apresentado na Figura 1. Observa-se os grupos que foram se formando ao longo do processo de agrupamento com suas respectivas similaridades.

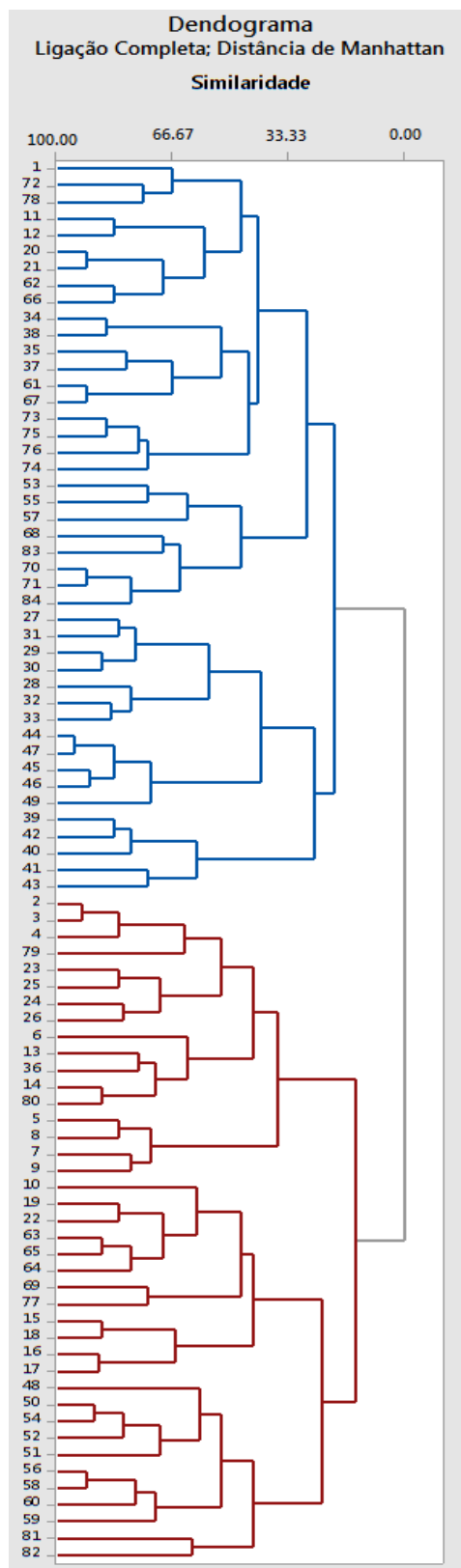


Figura 1. Dendrograma para o Agrupamento dos Taludes.

Com os resultados obtidos do dendrograma observou-se a frequência dos tipos de taludes que foram enquadrados em cada grupo e assim os dois grupos foram identificados. No

dendrograma o grupo azul foi denominado como o grupo dos taludes instáveis e o grupo vermelho como o grupo dos taludes estáveis.

3.3 Comparação com os resultados reais

Concluída a análise do dendrograma é possível realizar a comparação da situação real e do grupo de cada talude, sendo que se o grupo corresponde a situação real o talude é classificado corretamente. A Tabela 4 apresenta as relações descritas.

Tabela 4. Relação da condição real e do grupo de cada talude.

Obs.	Caso Real	Grupo	Obs.	Caso Real	Grupo
1	Instável	Instável	2	Estável	Estável
11	Instável	Instável	3	Estável	Estável
12	Instável	Instável	4	Estável	Estável
20	Instável	Instável	5	Estável	Estável
21	Instável	Instável	6	Estável	Estável
27	Instável	Instável	7	Estável	Estável
28	Estável	Instável	8	Estável	Estável
29	Instável	Instável	9	Estável	Estável
30	Instável	Instável	10	Estável	Estável
31	Instável	Instável	13	Instável	Estável
32	Estável	Instável	14	Estável	Estável
33	Estável	Instável	15	Estável	Estável
34	Instável	Instável	16	Estável	Estável
35	Instável	Instável	17	Estável	Estável
37	Instável	Instável	18	Estável	Estável
38	Instável	Instável	19	Estável	Estável
39	Estável	Instável	22	Estável	Estável
40	Estável	Instável	23	Estável	Estável
41	Instável	Instável	24	Estável	Estável
42	Estável	Instável	25	Instável	Estável
43	Instável	Instável	26	Estável	Estável
44	Instável	Instável	36	Estável	Estável
45	Instável	Instável	48	Estável	Estável
46	Instável	Instável	50	Estável	Estável
47	Instável	Instável	51	Estável	Estável
49	Instável	Instável	52	Estável	Estável
53	Instável	Instável	54	Estável	Estável
55	Instável	Instável	56	Estável	Estável
57	Instável	Instável	58	Estável	Estável
61	Instável	Instável	59	Estável	Estável
62	Instável	Instável	60	Estável	Estável

66	Instável	Instável	63	Estável	Estável
67	Instável	Instável	64	Estável	Estável
68	Instável	Instável	65	Estável	Estável
70	Instável	Instável	69	Estável	Estável
71	Instável	Instável	77	Estável	Estável
72	Instável	Instável	79	Estável	Estável
73	Instável	Instável	80	Instável	Estável
74	Instável	Instável	81	Estável	Estável
75	Instável	Instável	82	Instável	Estável
76	Instável	Instável			
78	Instável	Instável			
83	Instável	Instável			
84	Instável	Instável			

No grupo dos taludes estáveis a técnica de agrupamentos classificou um total de 40 taludes, sendo que apenas 4 taludes foram classificados erroneamente, ou seja 90% de acerto. Já no grupo dos taludes instáveis a técnica agrupou 44 taludes no total, com 6 taludes classificados erroneamente, ou seja a técnica apresentou 86% de acerto. A Tabela 5 apresenta um resumo dos resultados, em relação as classificações

Tabela 5. Resumo dos resultados da técnica.

Situação Real	Agrupamentos	
	Estáveis	Instáveis
Estáveis	36	6
Instáveis	4	38
Total	40	44

Em todo o banco de dados a técnica agrupou 84 taludes, sendo 74 taludes agrupados corretamente e 10 taludes agrupados equivocadamente. Para um banco de dados de 18 variáveis para cada talude a técnica apresenta eficácia devido ao tamanho do banco de dados.

4 CONCLUSÃO

O banco de dados em estudo se trata de um banco de dados universal, contendo 84 casos de taludes no mundo inteiro, com cada talude apresentando 18 variáveis. Nesse contexto o banco de dados é representativo tanto em volume de dados quanto em variabilidade de

tipos de taludes.

A técnica de agrupamentos aplicada ao banco de dados geotécnico se mostrou eficaz, respondendo através da variabilidade os diferentes estados dos taludes que estão no banco de dados, separando os em grupos, definidos em taludes estáveis e instáveis.

A técnica apresentou 90% e 86% de acerto nos agrupamentos para os taludes estáveis e instáveis, respectivamente. Para todo o banco de dados a técnica agrupou os 84 taludes, sendo 74 taludes agrupados corretamente e 10 taludes agrupados equivocadamente. Assim a aplicação da técnica se mostrou eficaz para o estudo proposto.

AGRADECIMENTOS

Os autores agradecem o apoio da FAPEMIG no desenvolvimento deste trabalho, pelas concessões de bolsas de pesquisa.

ACESSO AO BANCO DE DADOS

O banco de dados utilizado está disponível em <http://dx.doi.org/10.1016/j.ijrmms.2013.01.012>.

REFERÊNCIAS

- Davis, J.C. (1986) *Statistics and Data Analysis in Geology*, 3rd ed., New York, 656 p.
- Hudson, J.A. (1992). *Rock engineering systems, theory and practice*. Chichester: Ellis Horwood; 185 p.
- Landim, P.M.B. (2011). *Análise Estatística de Dados Geológicos Multivariados*, 1nd ed., São Paulo, p. 66-83.
- Linden, R. (2009). *Técnicas de Agrupamento*, *Revista de Sistemas de Informação*, FSMA, p. 18-36
- Mingoti, S.A. (2013). *Análise de Dados Através de Métodos de Estatística Multivariada*, 2nd ed., Belo Horizonte, p. 155-211.
- Naghadehi, M.Z. Jimenez, R. KhaloKakaie, R. Jalali, S.M.E. (2013). *A new open-pit mine slope instability index defined using the improved rock engineering systems approach*, *International Journal of Rock Mechanics & Mining Sciences*, Vol. 61, p. 1-14.