RELAÇÃO FUNCIONAL ENTRE VARIOGRAMA E VARIÂNCIA AMOSTRAL

Jorge Kazuo Yamamoto - Instituto de Geociências/USP

Freqüentemente tem-se utilizado da variância amostral como nível máximo de variabilidade dos dados para ajuste do variograma. Entretanto, a variância amostral é a máxima variabilidade dos dados se for considerada a hipótese de independência dos dados, o que não é válida, pois no caso de dados para o cálculo de variogramas existe uma forte dependência espacial dos dados, dependente, inclusive, do arranjo particular do conjunto de pontos de dados, que pode modificar completamente o variograma e, conseqüentemente, o seu nível de variabilidade máxima. Barnes (1991), fazendo uma revisão sobre o assunto constatou que, segundo Journel & Huijbregts (1978, apud Barnes 1991), a variância de dispersão pode algumas vezes ser utilizada no ajuste do patamar e que, conforme David (1977, apud Barnes 1991), o patamar do variograma é igual à variância das amostras do depósito, que pode ser calculado a partir das amostras. Segundo Barnes (1991), muitas vezes este ajuste é absolutamente impróprio o que levará a modelos impróprios de variogramas. Assim, Barnes (1991), demonstrou que o valor esperado da variância amostral é igual à média aritmética do variograma entre todos os N^2 pares de valores de amostras disponíveis.

$$E[S^2] = \left[\frac{1}{N}\sum_{i=1}^{N}(X_i-\bar{X})^2\right] = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\gamma(X_i,X_j) \qquad (1)$$

Em termos geoestatísticos, segundo Barnes (1991), o valor esperado da variância amostral é uma função da configuração geográfica dos dados e de todo o variograma, não exatamente do patamar do variograma. Em geral, se um patamar é claramente apresentado no variograma experimental, seu valor deveria ser usado como uma estimativa da variância populacional e a variância amostras não deveria ser usada como uma estimativa do patamar do variograma (Barnes 1991). Em termos práticos, o trabalho de Barnes (1991), não recomenda o uso da variância amostral como valor de patamar do variograma experimental. Freqüentemente, tem-se observado que o patamar pode ser superior à variância amostral e, nesse caso, o variograma fica de difícil interpretação. Assim, este trabalho tem como objetivo procurar a relação funcional entre variograma e variância amostral, conforme se apresenta a seguir. A variância amostral (vide equação 1), segundo Rendu (1978), pode ser reescrita como:

$$S^2 = \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}(X_i-X_j)^2 \qquad (2)$$

Nestes termos, a variância amostral torna-se igual a média de diferenças ao quadrado, ou seja, a própria definição do variograma segundo a interpretação geométrica pelo momento de inércia. Tendo em vista a equação (2), pode-se demonstrar que a relação funcional entre variância e variograma para N dados dispostos ao longo de uma linha é expressa como:

$$S^2 = \sum_{i=1}^{N-1} f(i).\gamma(h=i) \qquad (3)$$

onde N é o número de pontos ao longo da linha e f(i) é o fator de correção igual a: $2(N-i)/N^2$. Portanto, a variância amostral é igual a soma dos variogramas (variâncias espaciais para intervalos de amostragem) multiplicados por um fator de correção, que depende somente do número total de pontos na linha. Para o caso de variogramas médios calculados a partir de j=1,M linhas, cada uma com N(j) pontos de dados, a variância amostral para esse arranjo particular de dados pode ser calculada como soma da variância dentro das linhas com a variância entre as linhas, conhecida como relação de aditividade de variâncias (Rendu, 1978 e Journel & Huijbregts 1978). Assim, a variância amostral pode ser escrita como:

$$S^2 = \frac{1}{M}\sum_{j=1}^{M}\frac{1}{N[j]}\sum_{i=1}^{N[j]}\left[X(i,j)-\bar{X}(j)\right]^2 + \frac{1}{M}\sum_{j=1}^{M}\left[\bar{X}(j)-\bar{X}\right]^2 \qquad (4)$$

onde o primeiro termo a direita é a variância dentro (S^2_{dentro}) das linhas e o segundo a variância entre as linhas (S^2_{entre}). Dessa forma, a relação funcional entre variância e variograma fica como:

$$S^2_{dentro} = \frac{1}{M}\sum_{j=1}^{M}\sum_{i=1}^{N[j]-1} f(j,i).\gamma(j,i) \qquad (5)$$

Observe que agora, no caso de distribuição de dados ao longo de linhas, o variograma é uma fração da variância dentro das linhas, que por sua vez é uma componente da variância total (amostral). Assim, pelos motivos expostos, na interpretação prática do patamar do variograma deve-se utilizar de um valor de variância média, para a qual os pontos tendem a se estabilizar.

A título de ilustração do uso das equações (2) e (3), considere o seguinte conjunto de dados numéricos dispostos ao longo de uma linha e separados por uma distância h: (1,3,5,7,9,8,6,4,2,0). Para cálculo da variância amostral (equação 2) os dados são dispostos em forma matricial, conforme mostrado na Figura 1.

A variância amostral, segundo a equação (2), é simplesmente a soma de todas as diferenças ao quadrado dividido por $2N^2$, que dá:

$S^2 = 1650/2.100 = 8,25$

Para verificar a relação existente entre a soma de quadrados e o variograma, observe que as diagonais paralelas à principal representam justamente as diferenças ao quadrado entre amostras separadas por um número de intervalos de amostragem conforme a sua localização em relação a principal. Assim a primeira diagonal representa as diferenças ao quadrado para amostras situadas a um intervalo, a segunda para amostras a dois intervalos e, assim por diante.

***Figura 1** - Disposição matricial dos dados para cálculo da variância amostral conforme a equação (2). Os elementos da matriz são iguais a diferença ao quadrado entre os valores das respectivas linhas e colunas. As setas indicam a somatória das diferenças ao quadrado das diagonais paralelas a diagonal principal.*

	1	3	5	7	9	8	6	4	2	0	
1	0	4	16	36	64	49	25	9	1	1	
3	4	0	4	16	36	25	9	1	1	9	1
5	16	4	0	4	16	9	1	1	9	25	10
7	36	16	4	0	4	1	1	9	25	49	35
9	64	36	16	4	0	1	9	25	49	81	84
8	49	25	9	1	1	0	4	16	36	64	165
6	25	9	1	1	9	4	0	4	16	36	212
4	9	1	1	9	25	16	4	0	4	16	179
2	1	1	9	25	49	36	16	4	0	4	106
0	1	9	25	49	81	64	36	16	4	0	33
	1	10	35	84	165	212	179	106	33	0	

REFERÊNCIAS

BARNES, R.J. 1991. The variogram sill and the sample variance. Math. Geol., 23(4): 673-678.

DAVID, M. 1977. Geostatistical ore reserve estimation. Amsterdam, Elsevier. 364 p.

JOURNEL, A.G. & HUIJBREGTS, C.J. 1978. Mining geostatistics. London, Academic Press. 600 p.

RENDU, J.M. 1978. An introduction to geostatistical methods of mineral evaluation. Johannesburg, South Afr. Inst. Min. Metall. 84 p.

A BASE DE DADOS GEOQ E O ATLAS GEOQUÍMICO DA CPRM - 1ª ETAPA

Carlos Alberto C. Lins - Companhia de Pesquisa de Recursos Minerais - CPRM-Recife

O acervo do SEAG (Sistema Estatístico de Amostragem Geoquímica) da Companhia de Pesquisa de Recursos Minerais, constituido de dados analíticos de cerca de 300.000 amostras (sedimento de corrente, concentrado de bateia, rocha, solo, água, minério, etc.) distribuidas em 263 projetos, exigia uma base de dados que servisse como seu gerenciador e que contivesse, de forma resumida, as informações disponíveis.

Dentro do programa de controle de qualidade implantado pela CPRM, foi desenvolvido pela Coordenação Nacional de Geoquímica do Programa de Levantamentos Geológicos Básicos (PLGB), um projeto de elaboração de um Atlas Geoquímico da CPRM.

Este projeto será executado em três etapas, sendo que a