

Sobre Geoestatística e mapas

Paulo M. Barbosa Landim

Departamento de. Geologia Aplicada

Instituto de Geociência e Ciências Exatas

Universidade Estadual Paulista – UNESP/Rio Claro

plandim@rc.unesp.br

RESUMO *A Geologia, até há bem pouco tempo, era frequentemente considerada uma ciência baseada em interpretações puramente qualitativas dos fenômenos geológicos. Nos últimos 40 anos, porém, tem sido notável a mudança da fase descritiva para a utilização de métodos quantitativos, principalmente na área da Geologia Aplicada. Entre tantos temas relacionados com a quantificação em Geologia o assunto abordado neste artigo, que pretende ser de cunho didático, diz respeito à aplicação da Geoestatística na confecção de mapas geológicos. Na primeira parte do trabalho é feita uma introdução sobre os aspectos quantitativos na interpretação dos fenômenos geológicos, em seguida conceitos básicos sobre a metodologia geoestatística e, na última parte, uma aplicação clássica da geoestatística linear, baseada nos conceitos de estacionaridade da covariância e do variograma, para a confecção de mapas.*

PALAVRAS-CHAVE *Geoestatística, krigagem ordinária, mapeamento*

ABSTRACT *Until very time ago Geology was frequently considered a science based on purely qualitative interpretations of the geological phenomena. In the last 40 years, however, occurred a notable change from the descriptive phase to utilization of quantitative methods, mostly in the Applied Geology area. Among many related subjects to the quantification in Geology the boarded one in this paper concern the application of Geostatistics in geological mapping. In the first part of the article is introduced quantitative aspects in the interpretation of the geological phenomena, soon after basic concepts concerning geostatistical methodology and, finally, a classical application of lineal geostatistics, based on covariance and variogram stationary concepts, for spatial estimation and contouring.*

KEYWORDS

Geostatistics, ordinary kriging, spatial estimation and contouring

* Este documento deve ser referido como segue:

Landim P.M.B. 2006. Sobre Geoestatística e mapas. *Terræ Didática*, 2(1):19-33.
<<http://www.ige.unicamp.br/terraedidatica/>>

Introdução

A tentativa de aplicação de métodos quantitativos em Geologia coincide com o seu estabelecimento como ciência moderna e um exemplo marcante é a subdivisão do Terciário, feita por Lyell em 1830, baseada na proporção de espécies recentes de moluscos presentes nos diversos estratos da Bacia de Paris. Não alcançou, todavia, o mesmo estágio comparável àquele existente nas outras Ciências Naturais como a Física ou a Química. A Geologia, até a bem pouco tempo, era frequentemente considerada uma ciência qualitativa, isto é, descritiva. Nos últimos 40 anos, porém, tem sido notável a mudança da fase puramente descritiva dos fenômenos geológicos para um enfoque quantitativo. Esse processo de transformação das Ciências da Terra baseia-se em três fatores principais: a introdução do conceito de modelos matemáticos para a explicação de fenômenos geológicos, a procura da integração dos diversos fenômenos geológicos através da Teoria da Deriva Continental e a facilidade de acesso a computadores digitais.

A aceitação por parte dos geólogos tem sido bem maior no tocante à Tectônica Global, e isso pode ser facilmente constatado nos livros textos de Geologia básica recentes, que apresentam e discutem tais conceitos. O mesmo não acontece com a aplicação de modelos quantitativos, utilizando-se ou não de computadores, e isso porque os geólogos acostumados com o método das múltiplas hipóteses julgam que conseguem lidar com fenômenos naturais apenas em termos qualitativos, e acabam por criar uma barreira entre os dados obtidos e seu manuseio. Um dos motivos alegados, inclusive, é que os métodos quantitativos apresentados são tidos como extremamente complexos e numa linguagem matemática de difícil entendimento. As discussões geológicas geralmente tratam de questões sobre o que aconteceu no passado e em que ordem cronológica. Por exemplo, se numa determinada bacia ocorreu deposição de sedimentos marinhos e estes se apresentam hoje em dia erodidos na sua parte superior isso implica que após a deposição teria ocorrido uma regressão marinha. Todavia, para um estudo quantitativo, informações numéricas são exigidas sobre quão profunda era a lâmina de água na época da deposição, por quanto tempo demorou tal deposição, quando ocorreu a regressão marinha e com que taxa de soerguimento regional. Em outras palavras, após entendida a seqüência de fenômenos geológicos torna-se necessário estabelecer equações matemáticas que os representem e talvez aqui resida a grande dificuldade.

Nesse sentido, é muito importante entender que a matemática é mais uma linguagem do que propriamente uma ciência. Ela é muito mais um método de comunicação para descrever quantitativamente uma série de eventos do que um conjunto de conhecimentos sobre esses eventos.

Seja um lago que está sendo assoreado e, portanto, as lâminas mais inferiores são as mais antigas e que a taxa de sedimentação é conhecida e constante. Isto significa que a idade das lâminas é proporcional à profundidade de soterramento e tal relação pode ser matematicamente expressa por:

Idade = k.profundidade, sendo k a taxa constante de sedimentação.

Isso também pode ser representado por $I = k.p$ ou mesmo $i = k.\pi$, ou ainda por caracteres hebreus, pictogramas chineses, hieróglifos egípcios. É tudo uma questão de convenção para representar o mesmo fenômeno e nesse sentido, por tradição, as letras gregas têm a preferência dos matemáticos.

Um outro exemplo refere-se à conhecida fórmula estatística usada como medida de tendência central:

$$\bar{x} = \frac{1}{n} \left(\sum_{i=1}^n x_i \right)$$

Ela tanto expressa num texto em português, como em inglês ou alemão, que a média de valores de uma amostra (representada pela letra latina x com um travessão superposto) é igual à somatória dos valores de 1 até n dessa amostra, dividida pelo número n de valores. A fórmula, em linguagem matemática, é sempre a mesma.

Em muitas circunstâncias, a formulação matemática não oferece nada mais do que já se sabia a respeito do fenômeno, mas em muitas outras vezes o manuseio e a combinação de expressões matemáticas oferecem novas e mais claras visões sobre o processo geológico, pois são consistentes fornecendo sempre os mesmos resultados quando usados os mesmos dados. Podem, também, serem verificadas para, caso estejam corretas, serem utilizadas em previsões.

Desse modo, espera-se que o processo de atribuição de valores, de acordo com certas regras, aos dados de campo ou laboratório e o subsequente tratamento matemático dessas observações, traga uma sensível melhoria nos seguintes campos de atividade em Geologia:

- a) na amostragem, pelo fornecimento de critérios segundo os quais as amostras geológicas

- coletadas sejam representativas das populações sob estudo;
- na análise de dados, pelo registro sistemático e ordenado dos valores obtidos e pela representação gráfica que resuma os resultados, e também pela identificação de tendências, agrupamentos e correlações;
 - na comprovação de hipóteses de trabalho, pela verificação de conceitos ou modelos de processos geológicos;
 - na previsão quantitativa, quando da solução de problemas específicos que envolvam interpolações e extrapolações.

Somente desse modo é que a Geologia conseguirá, além da sua característica fundamental de ciência que estuda o presente para interpretar o passado, passar a ser também uma ciência que entende o presente para prever o futuro.

Na área mineral, com destaque para a do petróleo, onde a interpretação geológica, além de estar fundamentada em conceitos científicos, precisa ter enfoque econômico, observa-se, felizmente, uma marcante tendência quantitativa que vem possibilitando avanços importantes no uso de técnicas espaciais. Ver a propósito Houlding (2000).

Em Geologia, onde normalmente os dados são coletados segundo um plano de amostragem com coordenadas definidas, torna-se muito importante, quando de sua análise, que a configuração geométrica espacial seja considerada. Com frequência, resultados devem ser apresentados sob a forma de mapas, o que exige uma metodologia específica. Entre tantos outros temas relacionados com a quantificação em Geologia, este é o assunto a ser abordado neste artigo, que pretende ser de cunho didático e endereçado para servir de consulta inicial àqueles que desejam entender o significado da Geoestatística. Na primeira parte do trabalho é feita uma introdução sobre os aspectos quantitativos na interpretação dos fenômenos geológicos, em seguida conceitos básicos sobre a metodologia geoestatística e, finalmente, uma aplicação clássica da geoestatística linear, baseada nos conceitos de estacionaridade da covariância e do variograma, para a confecção de mapas.

O que significa “Geoestatística”?

Para muitos geólogos “geoestatística” significa simplesmente a aplicação de métodos estatísticos em Geologia, o que não corresponde à definição do termo. Na África do Sul, o engenheiro de minas Daniel G. Krige e o estatístico H.S. Sichel, desenvol-

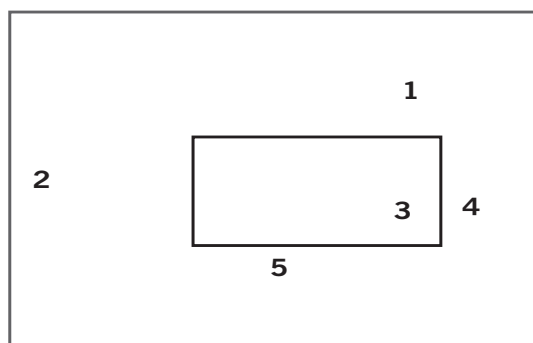


Figura 1 – Geometria do bloco a ser estimado a partir de 5 amostras

veram empiricamente uma técnica própria de estimativa para o cálculo de reservas minerais, a qual posteriormente recebeu tratamento formal por G. Matheron, nos início dos anos 60 do século passado, na França com o nome *Geoestatística*, para o estudo das chamadas *variáveis regionalizadas*, ou seja, variáveis com condicionamento espacial. O trabalho de Krige foi publicado em 1951 e os primeiros trabalhos de Matheron foram realizados após sua viagem à África do Sul quando conheceu Krige. A primeira publicação, descrevendo os resultados de Krige foi em co-autoria com Duval e Levy em 1955. Nos anos seguintes vieram as obras mais completas de Matheron, já com o arcabouço completo da Geoestatística Linear (Matheron 1962, 1963 e 1965). Em Chilés e Delfiner (1999) é apresentada uma revisão histórica sobre a Geoestatística com uma síntese sobre o desenvolvimento de suas ferramentas.

A estimativa de reservas sempre se constituiu parte fundamental do planejamento mineiro. O conhecimento geológico que se tenha a respeito do bem mineral a ser explorado, as eficientes instalações na boca da mina, os detalhes precisos sobre a *commoditie* em questão são variáveis importantes, mas o que vai decidir se o empreendimento será lucrativo ou não, dentro de um intervalo de tempo considerado, é a reserva medida com precisão.

No planejamento mineiro tal estimativa é realizada por meio de blocos onde as reservas minerais são amalgamadas para produzir reservas globais e curvas de teor/tonelagem. O processo todo, e para cada bloco, é baseado na coleta de amostras pontuais, ou seja, com volumes menores que os blocos. As amostras, com teores conhecidos, é que irão estimar o teor médio dos blocos, a duas ou a três dimensões. Seja, portanto, um bloco a ser estimado a partir de 5 amostras (Fig. 1).

Supondo que ocorra uma relação espacial entre os teores, ou seja, os valores serão muito

próximos em dois pontos vizinhos e progressivamente mais diferentes à medida que os pontos vão ficando mais distantes, é intuitivo esperar que o teor da amostra 3 seja similar, porém não necessariamente idêntico, ao teor médio do bloco. Pode-se esperar que as amostras 1, 4 e 5 também apresentem teores similares ao valor médio do bloco, mas não tanto como o teor em 3. Finalmente, com relação à amostra 2, situada mais distante, seria necessário um conhecimento melhor sobre a disposição espacial de valores no depósito para decidir se ela tem, ou não, relação com o valor médio do bloco. Em outras palavras, amostras situadas perto do bloco deverão apresentar teores altamente relacionados com ele e poderão, portanto, serem utilizadas para estimar o seu valor médio, e à medida que se situem a distâncias maiores o seu relacionamento diminui até se tornar independente. O peso da influência de cada amostra é, pois, inversamente correspondente à distância e essa noção pode ser aplicada para a estimativa do valor médio do bloco utilizando para tanto amostras com valores conhecidos, mas situadas a distâncias julgadas “convenientes”. Quanto mais próximas estiverem maior será o seu peso no processo de estimativa.

Nesta estimativa surgem, evidentemente, algumas questões: Até que distâncias devem ser consideradas as amostras? Quantas devem ser usadas? Aquela eventualmente colocada no centro do bloco terá um peso maior que as demais? Se amostras formarem grupos, qual a influência desses agrupamentos? Como evitar que os resultados sejam sub ou super estimados? A relação espacial, em termos geométricos, entre as amostras estimadoras e o bloco a ser estimado, tem importância? Essa técnica de estimativa pode ser utilizada indistintamente para depósitos do tipo cobre porfirítico, lateritas níquelíferas, veios de cassiterita, depósitos de urânio e outros?

Para responder a essas questões é que surgiu a geoestatística, a preocupar-se com o entendimento, por meio de análise matemática, da gênese e leis naturais que governam fenômenos interpretados como regionais. Isso traz como consequência direta a estimativa das variáveis regionais usando informações e relações a partir de um conjunto discreto de amostras, juntamente com a avaliação dos erros de estimativa, para estabelecer o grau de segurança em previsões e os padrões ótimos de amostragem, que assegure que um erro máximo de estimativa não seja excedido.

Inicialmente a aplicação era apenas para situações em geologia mineira na lavra e prospecção e,

como exemplos, podem ser citados o livro clássico sobre o assunto de Journel e Huijbregts (1978) e as obras em português de Valente (1982) e Yamamoto (2001). Posteriormente se estendeu para outros campos, especialmente nesses últimos anos, com aplicação em agricultura de precisão, cartografia, climatologia, geologia ambiental, geotecnia, hidrogeologia, pedologia, entre outros. Praticamente todas as últimas versões de *softwares* para Confecção de Mapas ou Sistemas de Informações Georreferenciadas apresentam métodos geoestatísticos.

Atualmente o termo *Geoestatística* acha-se consagrado como um tópico especial da estatística aplicada que trata de problemas referentes às *variáveis regionalizadas*, as quais têm um comportamento espacial mostrando características intermediárias entre as variáveis verdadeiramente aleatórias e as totalmente determinísticas. Nesse sentido pode-se afirmar que tal metodologia representa a grande contribuição da Geologia para a Estatística Aplicada.

As variáveis regionalizadas são constituídas por um duplo aspecto contraditório. Pela sua característica “aleatória” apresenta irregularidades e variação imprevisível de um ponto para outro e pela sua característica “estrutural” apresenta relações existentes entre os pontos no espaço motivadas pela sua gênese. Em outras palavras: é impossível prever com exatidão o teor do minério num determinado ponto da jazida (aspecto aleatório), mas é provável que se encontre minério rico perto de minério rico (aspecto estrutural). No estudo do comportamento das variáveis regionalizadas duas são as ferramentas fundamentais dos métodos geoestatísticos: o *semivariograma* e a *krigagem*.

O semivariograma

Seja uma variável regionalizada $x(i)$ coletada em diversos pontos i regularmente distribuídos por certa região. O valor de cada ponto está relacionado de algum modo com valores obtidos a partir de pontos situados a certa distância, sendo razoável pensar que a influência é tanto maior quanto menor for a distância entre os pontos. Para expressar essa relação é definido o vetor de distância $\Delta \vec{h}$, o qual tem uma orientação específica. O grau de relação entre pontos numa certa direção pode ser expresso pela covariância e, embora a covariância exista entre todas as distâncias possíveis ao longo de h , pode ser estipulado que somente sejam considerados valores entre pontos regularmente espaçados por múltiplos inteiros de Δh .

A covariância entre valores encontrados nessas distâncias separadas por Δh ao longo de h é

$$C(h) = C(\Delta h) = \frac{1}{n} \sum x_i x_{i+h} - m^2$$

onde m é a média da variável regionalizada $x(i)$.

Isso significa que a covariância é igual à média dos produtos-cruzados dos valores $x(i)$ encontrados nos pontos i pelos valores $x(i+h)$ nos pontos $i+h$, distantes a um intervalo Δh , subtraídos do quadrado da média da variável regionalizada $x(i)$; e n representa o número de pares de valores comparados. Como a covariância depende do tamanho do vetor h , se $h=0$, $C(h)$ passará a representar a variância, representada por $C(0)$.

$$C(0) = E[X^2] - E[X]^2 = Var[X]$$

Desse modo, pode-se calcular uma função, denominada *semivariância*, definida como metade da variância das diferenças.

$$\gamma(\vec{h}) = \gamma(h) = \frac{1}{2n} \sum (x_{i+h} - x_i)^2$$

Lembrando que $Var[X] = E[X^2] - E[X]^2$ pode-se representar $\gamma(h)$ por:

$$\gamma(h) = \left[\frac{1}{2n} \sum (x_{i+h} - x_i)^2 \right] - \left[\frac{1}{2n} \sum (x_{i+h} - x_i) \right]^2$$

Como a média da variável regionalizada $x(i)$ é também a média da variável regionalizada $x(i+h)$, pois se trata da mesma variável, apenas tomadas em lugares i e em $i+h$, tem-se que:

$$\left[\frac{1}{2n} \sum (x_{i+h} - x_i) \right]^2 = 0$$

e desenvolvendo o 1º termo:

$$\gamma(h) = \frac{1}{2n} \sum x_{i+h}^2 - 2 \frac{1}{2n} \sum x_{i+h} x_i + \frac{1}{2n} \sum x_i^2 = \frac{1}{n} \sum x_i^2 - m^2 - C(h)$$

e isso significa que $\gamma(h) = C(0) - C(h)$.

Em outras palavras, o vetor Δh apresentando-se infinitamente pequeno faz com que a variância seja mínima e a covariância máxima. Haverá um valor Δh para o qual ambas podem apresentar valores aproximadamente iguais, porém, à medida que Δh aumenta a covariância diminui enquanto a variância aumenta, porque ocorre progressivamente maior independência entre os pontos a distâncias cada vez maiores.

A semivariância distribui-se assim de zero, quando $h=0$, até um valor igual à variância das observações para um alto valor de h , se os dados forem estacionários, isto é, não ocorrer a presença de tendência nos valores. Essas relações são mostradas quando a função $\gamma(h)$ é colocada em gráfico contra Δh para originar o semivariograma. A distância segundo a qual $\gamma(h)$ atinge um patamar, denominado soleira ou patamar (*sill*), igual à variância a priori dos dados, é chamada de alcance

ou amplitude (*range*). Geralmente a soleira é representada por C e o alcance por a . A semivariância não é apenas igual à média das diferenças ao quadrado entre pares de pontos espaçados as distâncias h , mas também é igual à variância dessas diferenças.

O semivariograma mostra a medida do grau de dependência espacial entre amostras ao longo de um suporte específico e, para sua construção, são usados simplesmente as diferenças ao quadrado dos valores obtidos, assumindo-se uma estacionaridade nos incrementos. Isso significa que o semivariograma é uma medida da variabilidade geológica condicionada pela distância. Tal variabilidade pode ser bastante diferente quando consideradas diferentes direções. Por exemplo, em estratos sedimentares com inclinação ocorre maior correlação de valores na direção das camadas do que no sentido do mergulho das mesmas.

Para construir um semivariograma é necessário, portanto, dispor de um conjunto de valores obtidos a intervalos regulares dentro de um mesmo suporte geométrico. Sendo $x(1), x(2), \dots, x(i), \dots, x(n)$, valores de uma variável regionalizada a seguinte fórmula fornece uma estimativa não tendenciosa da semivariância:

$$\gamma(h) = \frac{1}{2n} \sum (x_{i+h} - x_i)^2$$

O estudo é feito em uma direção ao longo de uma linha ou ao longo de uma série de linhas paralelas, utilizando n possíveis diferenças a intervalos Δh ou múltiplos de Δh .

Em lugar do termo semivariograma é muito comum o uso mais simplificado da expressão variograma, porém, para o cálculo é sempre levado em consideração a divisão por $2n$.

Em Clark 1979 é apresentada a construção de um semivariograma a partir de uma rede regular, com espaçamento entre os pontos de 100 pés (Fig. 2). Trata-se de um depósito estratiforme de ferro com valores em porcentagem por peso.

Para o cálculo dos semivariogramas, em diversas direções, são encontradas as somatórias dos quadrados das diferenças e posterior divisão por duas vezes o número dessas diferenças. Assim para a direção Leste-Oeste inicia-se com o menor intervalo possível, ou seja, 100 pés da seguinte maneira:

$$\begin{aligned} \gamma^*(100) = & [(40 - 42)^2 + (42 - 40)^2 + (40 - 39)^2 + \\ & (39 - 37)^2 + (37 - 36)^2 + (43 - 42)^2 + (42 - \\ & 39)^2 + (39 - 39)^2 + (39 - 41)^2 + (41 - 40)^2 + \\ & (40 - 38)^2 + (37 - 37)^2 + (37 - 37)^2 + (37 - \\ & 35)^2 + (35 - 38)^2 + (38 - 37)^2 + (37 - 37)^2 + \\ & (37 - 33)^2 + (33 - 34)^2 + (35 - 38)^2 + (35 - \\ & 37)^2 + (37 - 36)^2 + (36 - 36)^2 + (36 - 35)^2 + \\ & (36 - 35)^2 + (35 - 36)^2 + (36 - 35)^2 + (35 - \\ & 34)^2 + (34 - 33)^2 + (33 - 32)^2 + (32 - 29)^2 + \\ & (29 - 28)^2 + (38 - 37)^2 + (37 - 35)^2 + (29 - \\ & 30)^2 + (30 - 32)^2] / [2 \times 36] = \mathbf{1,46} \end{aligned}$$

Para o intervalo de 200 pés:

$$\begin{aligned} \gamma^*(200) = & [(44 - 40)^2 + (40 - 40)^2 + (42 - 39)^2 + \\ & (40 - 37)^2 + (39 - 36)^2 + (42 - 43)^2 + (43 - \\ & 39)^2 + (42 - 39)^2 + (39 - 41)^2 + (39 - 40)^2 + \\ & (41 - 38)^2 + (37 - 37)^2 + (37 - 35)^2 + (37 - \\ & 38)^2 + (35 - 37)^2 + (38 - 37)^2 + (37 - 33)^2 + \\ & (37 - 34)^2 + (38 - 35)^2 + (35 - 36)^2 + (37 - \\ & 36)^2 + (36 - 35)^2 + (36 - 36)^2 + (35 - 35)^2 + \\ & (36 - 34)^2 + (35 - 33)^2 + (34 - 32)^2 + (33 - \\ & 29)^2 + (32 - 28)^2 + (38 - 35)^2 + (35 - 30)^2 + \\ & (30 - 29)^2 + (29 - 32)^2] / [2 \times 33] = \mathbf{3,30} \end{aligned}$$

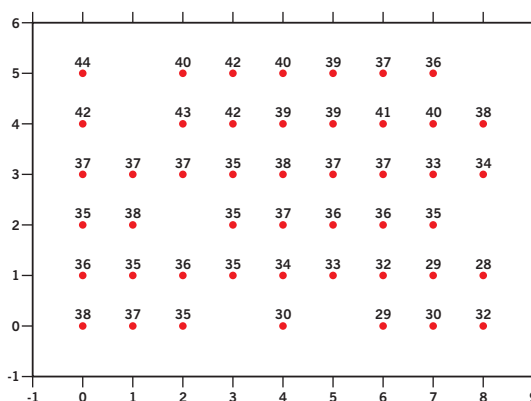


Figura 2 – Mapa de localização dos teores de ferro (Clark 1979)

Tabela 1 – Valores da função semivariograma calculados para as direções Leste-Oeste e Norte-Sul

Direção	Distância	Semivariograma	Nº Pares
Leste-Oeste	100	1,46	36
	200	3,30	33
	300	4,31	27
	400	6,70	23
Norte-Sul	100	5,35	36
	200	9,87	27
	300	18,88	21

E assim por diante, tanto para esta direção como para a Norte-Sul. O resultado é apresentado na Tabela 1.

Estes resultados permitem a construção dos semivariogramas experimentais (Fig. 3), nas duas direções consideradas, e o que se pode perceber é que há uma distinta diferença na estrutura dos dados, ou seja, a presença de uma anisotropia. Na direção Norte-Sul os valores aumentam muito mais rapidamente, sugerindo uma maior continuidade na direção Leste-Oeste.

Os semivariogramas expressam o comportamento espacial da variável regionalizada e mostram:

- o tamanho da zona de influência em torno de uma amostra, pois toda amostra cuja distância ao ponto a ser estimado for menor ou igual ao alcance, fornece informações sobre o ponto;
- a anisotropia, quando os semivariogramas mostram diferentes comportamentos para diferentes direções de linhas de amostragem e de estudo da variável; neste caso a anisotropia pode ser geométrica quando o alcance varia de acordo com as diversas direções consideradas, mantendo constante a soleira e zonal quando

- o alcance permanece constante e a soleira varia conforme for modificada a direção;
- c) continuidade, pela forma do semivariograma, em que para $h \equiv 0$, $\gamma(h)$ já apresenta algum valor. Essa situação é conhecida como efeito pepita (*nugget effect*) e é representada por C_0 . O efeito pepita pode ser atribuído a erros de medição ou ao fato de que os dados não foram coletados a intervalos suficientemente pequenos para mostrar o comportamento espacial subjacente do fenômeno em estudo.

Na construção do semivariograma, as somatórias necessárias para o cálculo de $\gamma(h)$ devem ser constituídas por um número suficiente de pares, que tornem o resultado consistente. Como regra prática adota-se para tanto um mínimo de 30 pares, o que pode ser conseguido se for escolhido como maior Δh , a metade da maior distância existente entre os pontos. Isto significa que, para uma análise geoestatística, exige-se que o número mínimo de pontos amostrados seja razoável, por volta de 30 a 40.

Uma outra consideração importante a ser feita é determinar o grau de aleatoriedade presente nos dados pela fórmula $E = C_0/C$ (Guerra 1988):

- $E < 0,15$: componente aleatória pequena;
 $0,15 \leq E \leq 0,30$: componente aleatória significativa;
 $E > 0,30$: componente aleatória muito significativa.

O extremo dessa situação é o modelo de pepita pura, onde não ocorre covariância entre os valores e, portanto, a análise semivariográfica não se aplica, sendo sugerido o uso de outros métodos de interpolação.

De posse do semivariograma experimental é necessário ajustá-lo a um modelo teórico, ou seja,

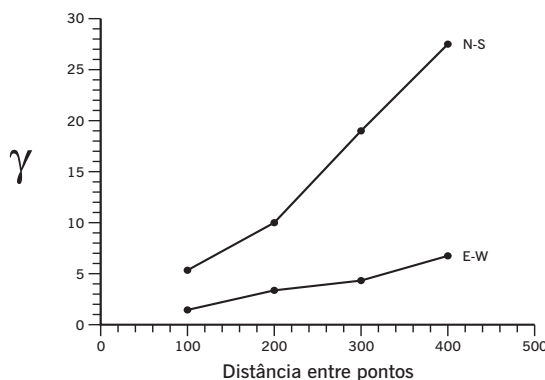


Figura 3 – Semivariogramas experimentais calculados para as direções Leste-Oeste e Norte-Sul

a uma função que definirá os parâmetros do semivariograma, “efeito pepita”, “alcance” e “patamar”. Entre os modelos disponíveis estes são os mais utilizados:

a) Modelos com patamar

a.1) Modelo esférico

$$\gamma(h) = C \left[\frac{3}{2} \left(\frac{h}{a} \right) - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right] \text{ para } h < a$$

$$\gamma(h) = C, \text{ para } h \geq a,$$

neste modelo a inclinação da tangente junto à origem ($h \equiv 0$) é $3C/2a$; é o modelo mais comum, podendo-se afirmar que equivale à função de distribuição normal da estatística clássica.

a.2) Modelo exponencial

$$\gamma(h) = C[1 - e^{-3h/a}]$$

neste modelo a inclinação da tangente junto à origem é C/a ; C é a assíntota de uma curva exponencial e pode ser equalizada junto à soleira; “a” corresponde ao alcance prático igual à distância segundo a qual 95% da soleira foi alcançada.

a.3) Modelo gaussiano

$$\gamma(h) = C[1 - e^{(-3h/a)^2}]$$

a curva é parabólica junto à origem e a tangente nesse ponto é horizontal, o que indica pequena variabilidade para curtas distâncias; “a” corresponde ao alcance prático igual à distância segundo a qual 95% da soleira foi alcançada.

b) Modelos sem patamar

b.1) Modelo potencial

$\gamma(h) = Ch^\alpha$, com a potência α assumindo valores entre zero e próximo a 2;

quando $\alpha = 1$ o modelo torna-se linear;

$\gamma(h) = ph$, sendo p a inclinação da reta;

é o modelo mais simples e representado por uma reta passando pela origem do gráfico.

Para dados que estão irregularmente distribuídos no espaço bidimensional não é possível, em princípio, encontrar pares de amostras suficientes com exatamente o mesmo espaçamento Δh para o cálculo

em uma determinada direção, como feito com dados dispostos em malha regular. Para contornar essa situação define-se uma distância de tolerância $\Delta'h$ para o espaçamento Δh entre os pares de amostras de um ângulo de tolerância $\Delta'\alpha$ para a direção α considerada. Assim, para o cálculo do semivariograma de uma distribuição irregular de pontos ao longo de uma determinada direção α , consideram-se todas as amostras que se encontram no ângulo $\alpha \pm \Delta'\alpha$, e, em seguida, classificam-se os pares de amostras em classes de distância $\Delta h \pm \Delta'h$, $2\Delta h \pm \Delta'h$, onde Δh é a distância básica (Fig. 4). As direções consideradas e seus respectivos ângulos de tolerância devem cobrir a área toda.

Para a estimativa do semivariograma experimental, não se tendo certeza se o fenômeno sob estudo é isotrópico ou anisotrópico no espaço bidimensional, inicialmente consideram-se quatro direções, E-W, N-S, NE-SW e NW-SE, com um ângulo de abertura com tolerância de 45° . Quando se constata uma direção bem marcante de anisotropia deve-se adotar tal direção com um pequeno ângulo de tolerância ($\alpha^0 \pm \Delta'\alpha^0$) para estimar o semivariograma nessa direção. Melhores estimativas são obtidas quando os modelos são baseados em semivariogramas experimentais que apresentam a menor razão “efeito pepita/patamar” e, também, o maior alcance.

Em resumo, para a utilização do semivariograma as seguintes suposições básicas são requeridas:

- as diferenças entre pares de valores de amostras são determinadas apenas pela orientação espacial relativa dessas amostras;
- o interesse é focado apenas na média e na variância das diferenças, significando que esses dois parâmetros dependem unicamente da orientação;
- por conveniência assume-se que os valores da área de interesse não apresentam tendência que

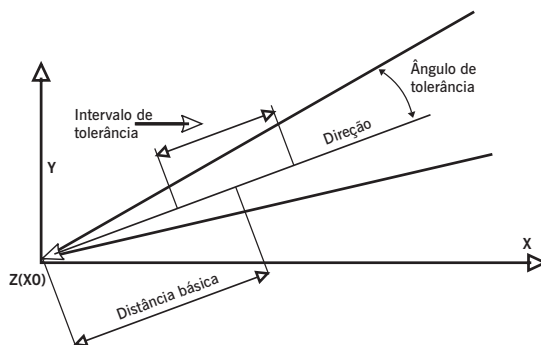


Figura 4 – Esquema para localização de pontos em distribuições irregulares para cálculo de variogramas experimentais

possa afetar os resultados e, assim, a preocupação é apenas com a variância das diferenças entre valores das amostras.

A modelagem, ou seja, o ajuste de um variograma experimental a uma função é um passo fundamental na análise variográfica, sendo um processo que envolve várias tentativas e na qual a experiência pesa muito. Pode-se optar por um ajuste manual por comparação visual, mais sujeito a erros, ou, com o auxílio de algoritmos, para ajustes automáticos como apresentado, entre outros, em Pannatier (1996). Acrescentar, em seguida, a essa verificação a “validação cruzada”. Nessa análise, depois de obtido o modelo variográfico, cada valor original é removido do domínio espacial e, usando-se os demais, um novo valor é estimado para esse ponto. Desse modo, um gráfico pode ser construído mostrando a relação entre valores reais e estimados. A validação cruzada, porém, não prova que o modelo escolhido é o mais correto, mas sim que o mesmo não é inteiramente incorreto. A melhor verificação, então, é aquela resultante do confronto entre os valores estimados e a realidade de campo.

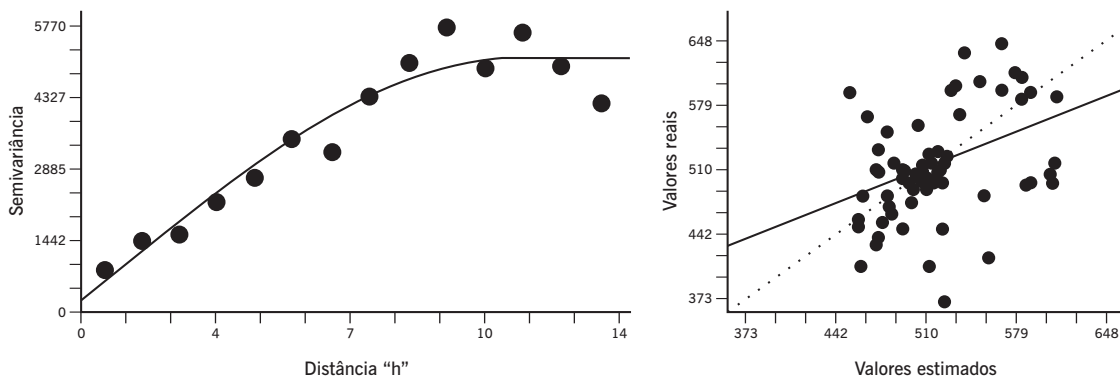


Figura 5 – Ajuste ao modelo esférico (0,548) e diagrama de dispersão da validação cruzada

Uma comparação entre os ajustes a um modelo esférico e a um modelo linear é apresentada a seguir e o ajuste é indicado, no gráfico, pela relação entre a reta a 45° e a obtida pela análise (Figs. 5 e 6). Um valor igual a 1,0 significa a indicação de ajuste perfeito.

Num estudo geoestatístico, a parte fundamental refere-se à determinação do semivariograma. Isso é importante e todo o cuidado deve ser tomado na análise variográfica para que possa obter uma criteriosa análise geoestatística.

Krigagem

Krigagem é um processo de estimativa de valores de variáveis distribuídas no espaço, e/ou no tempo, a partir de valores adjacentes enquanto considerados como interdependentes pelo semivariograma. Trata-se, em último caso, de um método de estimativa por médias móveis. O termo, tradução do francês *krigeage*, e do inglês *kriging*, foi cunhado pela escola francesa de geoestatística em homenagem ao engenheiro de minas sul-africano e pioneiro na aplicação de técnicas estatísticas em avaliação mineira, Daniel G. Krige.

A krigagem pode ser usada, como algoritmo estimador, para:

- previsão do valor pontual de uma variável regionalizada em um determinado local dentro do campo geométrico; é um procedimento de interpolação exato que leva em consideração todos os valores observados, o qual pode ser a base para cartografia automática por computador quando se dispõe de valores de uma variável regionalizada dispostos por uma determinada área;
- cálculo médio de uma variável regionalizada para um volume maior que o suporte geométrico como, por exemplo, no cálculo do teor

médio de uma jazida a partir de informações obtidas de testemunhas de sondagens;

Em todas essas situações o método fornece, além dos valores estimados, o “erro” associado a tal estimativa, o que o distingue dos demais algoritmos à disposição. A krigagem usa informações a partir do semivariograma para encontrar os pesos ótimos a serem associados às amostras que irão estimar um ponto, um área ou um bloco. Como o semivariograma é uma função da distância entre locais de amostragens, mantendo o mesmo número de amostras, os pesos são diferentes de acordo com o seu arranjo geográfico. O uso do semivariograma para a estimativa por krigagem não exige que os dados tenham distribuição normal, mas a presença de distribuição assimétrica, com muitos valores anômalos, deve ser considerada, pois a krigagem é um estimador linear.

Krigagem ordinária

A metodologia geoestatística apresenta diversas técnicas de estimativas disponíveis e a mais usual é a krigagem ordinária

Seja um ponto que se deseja estimar, sendo o valor real desconhecido representado por V . O valor estimado (V^*) é calculado, utilizando n amostras localizadas segundo coordenadas conhecidas, com valores $x_1, x_2, x_3, \dots, x_n$ (conjunto S), de forma linear, como por exemplo, através da técnica da ponderação pelo inverso das distâncias.

$$V^* = p_1 x_1 + p_2 x_2 + p_3 x_3 + \dots + p_n x_n, \text{ onde os } p_i \text{ são os pesos atribuídos a cada amostra } i.$$

É evidente que existe associado a esse estimador um erro $\epsilon = V - V^*$ e que se, teoricamente, diversas estimativas forem feitas a média de erros é zero. Se os erros, portanto, apresentarem valores próximos a zero, o estimador é de confiança e isso pode ser

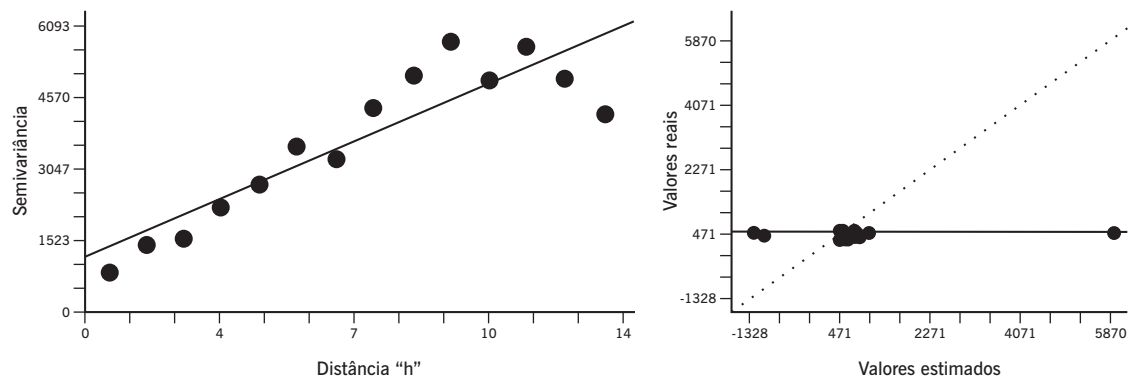


Figura 6 – Ajuste ao modelo linear (0,004) e diagrama de dispersão da validação cruzada

verificado pela distribuição desses valores. A maneira mais simples de medir estatisticamente tal distribuição é via o desvio padrão ou a variância. No caso em questão, porém, a variância não pode ser obtida porque não se conhece o valor real que se esta estimando e, portanto, também não se sabe qual o erro associado.

$$\text{Variância dos erros} = \sigma_{\epsilon}^2 = \text{desvios ao quadrado em relação ao erro médio} = \text{média de } (V-V^*)^2.$$

Para encontro da variância pode-se, porém, utilizar o semivariograma, em que são medidas as diferenças ao quadrado. Num semivariograma, previamente calculado, dada uma distância h entre os pontos, pode-se estimar a variância simplesmente lendo o valor no eixo dos γ 's e multiplicando-o por 2

$$\sigma_{\epsilon}^2 = 2\gamma(h)$$

Desse modo, para o processo de estimativa de um ponto utilizando o método da krigagem, procede-se da seguinte maneira:

$$V^* = p_1x_1 + p_2x_2 + p_3x_3 + \dots + p_nx_n$$

Se a soma dos pesos for igual a 1 e não ocorrer tendência local dos valores, esse estimador é o melhor e não tendencioso, pois a partir dos pesos atribuídos a cada amostra, minimiza a estimativa da variância.

$$\partial\sigma_{\epsilon}^2 / \partial\lambda_i = 0, \quad i = 1, 2, 3, 4, \dots, n$$

Isso é obtido construindo-se um sistema de n equações com n incógnitas ($\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$) e havendo a restrição de que $\sum\lambda_i = 1$, passa-se a $n+1$ equações. Como se tem apenas n incógnitas desconhecidas, introduz-se uma outra, também desconhecida, para balancear o sistema, ou seja, o chamado multiplicador de Lagrange, μ .

$$\partial_{\epsilon}^2 - \lambda(\sum\lambda_i - 1) = 0, \text{ se } \sum\lambda_i - 1 = 0$$

O objetivo da krigagem é procurar pelo conjunto ótimo de ponderadores de modo que a variância do erro de estimativa seja a menor possível. Para tanto, é organizado um sistema de equações com $n+1$ incógnitas, para a estimativa de um ponto (S_0):

$$\begin{aligned} \lambda_1 \bar{\gamma}(S_1, S_1) + \lambda_2 \bar{\gamma}(S_1, S_2) + \lambda_3 \bar{\gamma}(S_1, S_3) + \dots + \mu &= \bar{\gamma}(S_1, S_0) \\ \lambda_1 \bar{\gamma}(S_2, S_1) + \lambda_2 \bar{\gamma}(S_2, S_2) + \lambda_3 \bar{\gamma}(S_2, S_3) + \dots + \mu &= \bar{\gamma}(S_2, S_0) \\ \lambda_1 \bar{\gamma}(S_n, S_1) + \lambda_2 \bar{\gamma}(S_n, S_2) + \lambda_3 \bar{\gamma}(S_n, S_3) + \dots + \mu &= \bar{\gamma}(S_n, S_0) \\ \lambda_1 + \lambda_2 + \lambda_3 + \dots + 0 &= 1 \end{aligned}$$

Estas equações constituem equações normais a $n+1$ incógnitas, as quais podem ser resolvidas, para a obtenção dos coeficientes, por cálculo matricial, segundo:

$$[X][A] = [Y]$$

Multiplicando ambos os termos da equação pelo inverso de $[X]$, isto é, $[X]^{-1}$:

$$[X]^{-1} \cdot [X] \cdot [A] = [X]^{-1} [Y];$$

como $[X]^{-1} \cdot [X] = [I]$ (matriz de identidade) e $[I] \cdot [A] = [A]$,

$$[A] \cdot [X]^{-1} = [Y].$$

Em notação matricial:

$$\begin{bmatrix} \bar{\gamma}(S_1, S_1) & \bar{\gamma}(S_1, S_2) & \dots & \bar{\gamma}(S_1, S_n) & 1 \\ \bar{\gamma}(S_2, S_1) & \bar{\gamma}(S_2, S_2) & \dots & \bar{\gamma}(S_2, S_n) & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \bar{\gamma}(S_n, S_1) & \bar{\gamma}(S_n, S_2) & \dots & \bar{\gamma}(S_n, S_n) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} \bar{\gamma}(S_1, S_0) \\ \bar{\gamma}(S_2, S_0) \\ \vdots \\ \bar{\gamma}(S_n, S_0) \\ 1 \end{bmatrix}$$

$[S_i, S_i] \qquad \qquad \qquad [\lambda_i] \qquad \qquad [S_i, S_0]$

A matriz $[S_i, S_i]$ contém os valores obtidos no semivariograma referentes às distâncias entre as amostras estimadoras; o vetor $[S_i, S_0]$ contém os valores obtidos no semivariograma referentes às distâncias entre cada amostra e o ponto (S_0) a ser estimado e o vetor $[\lambda_i]$ contém os ponderadores a serem calculados.

Resolvido o sistema de equações, obtém-se os pesos λ_i e o multiplicador de Lagrange, μ , segundo:

$$[\lambda_i] = [S_i, S_i]^{-1} \cdot [S_i, S_0]$$

Para o ponto S_0 a ser estimado, obtém-se uma combinação linear dos valores dos pontos vizinhos e respectivos pesos

$$S_0 = \sum\lambda_i S_i$$

Para o cálculo da variância (σ^2) associada ao valor S_0 obtido por estimativa usa-se a expressão:

$$\sigma^2 = \sum\lambda_i \bar{\gamma}(S_i, S_0) + \mu = [\lambda_i]' [S_i, S_0],$$

sendo $[\lambda_i]'$ = vetor transposto com os pesos λ_i e $[S_i, S_0]$ = vetor com os valores obtidos no semivariograma referentes às distâncias entre cada amostra e o ponto (S_0) a ser estimado.

Interpolação de pontos para a confecção de mapas de contornos

Existe à disposição diversos algoritmos para a estimativa de pontos que são utilizados na interpolação para a confecção de mapas, como o inverso ao quadrado da distância, curvatura mínima, vizinho mais próximo, regressão polinomial, entre outros, além da krigagem. A diferença entre os diversos métodos está em como, a partir de valores conhecidos, podem ser estimados os nós da rede regular. Além disso, a krigagem se destaca por levar em consideração a variabilidade espacial dos dados.

Normalmente nesse tipo de estudo parte-se de uma série de pontos irregularmente distribuídos com valores conhecidos e a partir deles procura-se por construir uma rede regular de pontos interpolados (Fig. 7).

Seja uma situação hipotética (Fig. 8) em que 5 pontos foram amostrados para teores de U_3O_8 e se pretende a partir deles estimar o valor no ponto A (Clark 1979).

Aplicando, inicialmente, o método do inverso da distância, onde os pesos são corrigidos para obter uma soma igual a um, é obtido o resultado apresentado na Tabela 2.

Aplicando-se a equação da média ponderada, tem-se:

$$A = (0,319 \cdot 400) + (0,137 \cdot 380) + (0,217 \cdot 450) + (0,229 \cdot 280) + (0,098 \cdot 320) = 372,8 \text{ ppm}$$

Neste caso, para estimar um valor desconhecido a partir de valores conhecidos importa apenas as distâncias euclidianas para a atribuição dos pesos e isso é válido para todos os outros algoritmos, com exceção da krigagem. Na krigagem há necessidade de encontrar, preliminarmente, a variabilidade espacial dos dados o que é conseguido por uma análise variográfica. Em outras palavras, antes da aplicação da krigagem é preciso calcular um semivariograma experimental e em seguida modelá-lo. Tendo isso sido feito encontrou-se um modelo esférico, com os seguintes parâmetros: amplitude de influência (a),

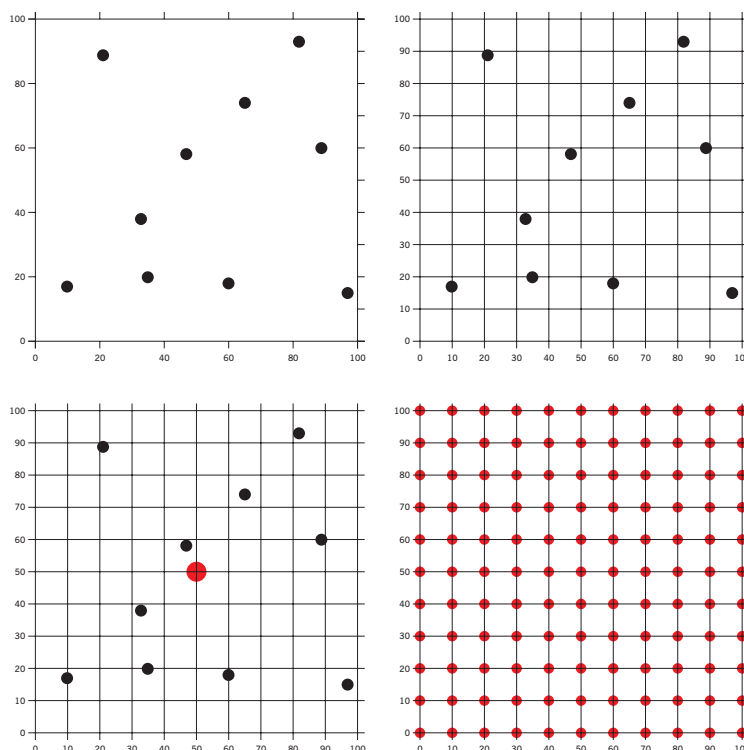


Figura 7 – Pontos de dados amostrais (superior esquerdo); malha regular superposta (superior direito); estimativa do bloco no ponto de coordenadas em vermelho (inferior esquerdo) e estimativa dos pontos da malha regular (inferior direito)

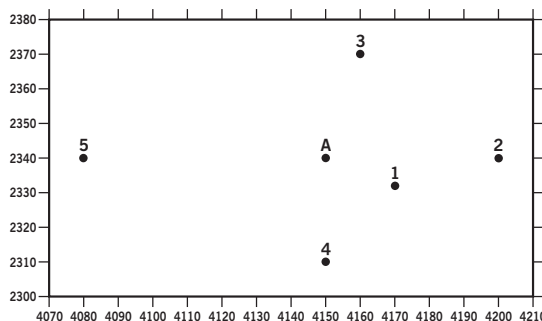


Figura 8 – Configuração de pontos de amostragem para estimativa no ponto A

Tabela 2 – Pesos calculados pelo inverso da distância para estimativa do ponto A

Pontos	X	Y	U_3O_8	Distâncias	1/D	Pesos
A	4150	2340	?	00,00		
1	4170	2332	400	21,54	0,0464	0,319
2	4200	2340	380	50,00	0,0200	0,137
3	4160	2370	450	31,62	0,0316	0,217
4	4150	2310	280	30,00	0,0333	0,229
5	4080	2340	320	70,00	0,0143	0,098

100 pés; valor da soleira (C), 700 ppm e um efeito pepita(C0) da ordem de 100 ppm (Fig. 9).

Na krigagem para a estimativa de A são necessários não apenas os valores das distâncias de cada um dos 5 pontos em relação a A, mas também as distâncias entre os pontos. Não são as distâncias euclidianas, porém, que são utilizadas, mas sim os valores resultantes da análise espacial pelo variograma modelado. Isso pode ser obtido graficamente colocando os valores no semivariograma ou analiticamente utilizando a fórmula de um modelo, no caso, esférico. A Figura 10 mostra as relações entre amostras (em preto) e entre as amostras e o ponto a ser estimado (em vermelho).

Assim a distância euclidiana entre o ponto 1, o mais próximo, e A é de 21,54 pés, a qual corresponde, pelo variograma modelado, ao valor 322,7 (Fig. 11).

Esse valor também pode ser encontrado segundo:

$$\gamma(h) = C \left[\frac{3}{2} \left(\frac{h}{a} \right) - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right] + C_0$$

(modelo esférico)

$$\gamma(21,54) = 700 \left[\frac{3}{2} \left(\frac{21,54}{100} \right) - \frac{1}{2} \left(\frac{21,54}{100} \right)^3 \right] + 100 = 322,7$$

De modo idêntico são encontrados todos os demais valores (Tab. 3):

Tabela 3 – Distâncias e valores das funções semivariograma para estimativa do ponto A na Figura 8

Pontos	Distâncias	$\Delta(h)$
p1-A	21,54	322,67
p2-A	50	581,25
p3-A	31,62	420,94
p4-A	30	405,55
p5-A	70	714,95
p1-p2	31,05	415,53
p1-p3	39,29	491,35
p1-p4	29,73	402,99
p1-p5	90,35	790,55
p2-p3	50,00	581,25
p2-p4	58,31	642,86
p2-p5	120,00	800
p3-p4	60,83	659,92
p3-p5	85,44	778,82
p4-p5	76,16	745,06

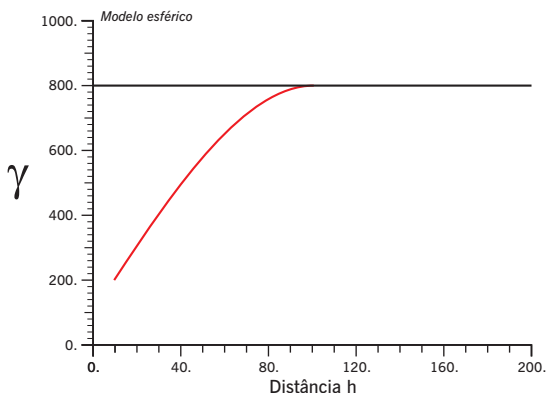


Figura 9 – Modelo de variograma para os dados da Figura 8

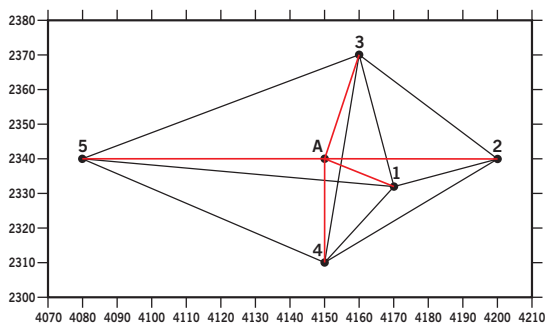


Figura 10 – Inter-relações entre amostras e ponto a ser estimado para cálculo das equações de krigagem

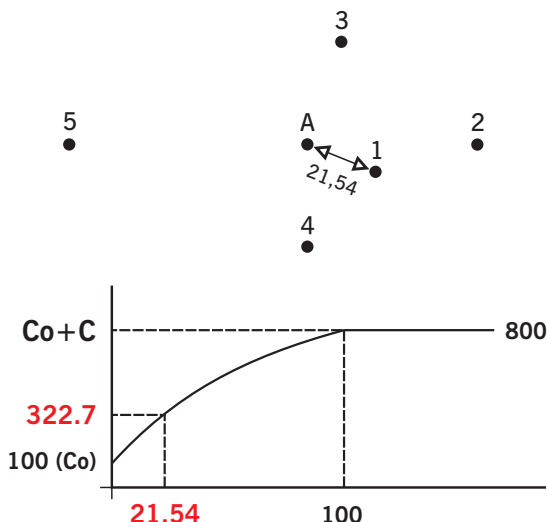


Figura 11 – Esquema mostrando a transformação da distância euclidiana (21,54 pés) em valor da função semivariograma (322,7 ppm)

De posse desses valores, pode-se organizar o sistema de equações normais para o cálculo dos pesos λ_i e do multiplicador de LaGrange μ :

$$\begin{array}{c}
 \begin{bmatrix}
 0 & 415,5 & 491,4 & 403,0 & 790,5 & 1 \\
 415,5 & 0 & 581,3 & 642,9 & 800,0 & 1 \\
 491,4 & 581,3 & 0 & 659,9 & 778,8 & 1 \\
 403,0 & 642,9 & 659,9 & 0 & 745,1 & 1 \\
 790,5 & 800,0 & 778,8 & 745,1 & 0 & 1 \\
 1 & 1 & 1 & 1 & 1 & 0
 \end{bmatrix}
 \begin{bmatrix}
 \lambda_1 \\
 \lambda_2 \\
 \lambda_3 \\
 \lambda_4 \\
 \lambda_5 \\
 \mu
 \end{bmatrix}
 =
 \begin{bmatrix}
 322,7 \\
 581,3 \\
 420,9 \\
 405,5 \\
 714,9 \\
 1
 \end{bmatrix} \\
 [A] \qquad \qquad \qquad [\lambda] \qquad \qquad [B]
 \end{array}$$

$$A^{-1} = \begin{bmatrix}
 -0,002246 & 0,000802 & 0,000535 & 0,000911 & -0,000002 & 0,038234 \\
 0,000802 & -0,001476 & 0,000343 & 0,000085 & 0,000246 & 0,216190 \\
 0,000535 & 0,000343 & -0,001314 & 0,000148 & 0,000288 & 0,215872 \\
 0,000911 & 0,000148 & 0,000154 & -0,001501 & 0,000357 & 0,214851 \\
 -0,000002 & 0,000288 & 0,000278 & 0,000357 & -0,000889 & 0,314853 \\
 0,038234 & 0,215872 & 0,223193 & 0,214851 & 0,314853 & -531,3827
 \end{bmatrix}$$

$$[\lambda] = [A]^{-1}[B] = \begin{bmatrix}
 0,3760 \\
 -0,0347 \\
 0,3017 \\
 0,2668 \\
 0,0901 \\
 10,0816
 \end{bmatrix}$$

Para o cálculo do teor em A:

$$A = (0,3760 \cdot 400) + (-0,0347 \cdot 380) + (0,3017 \cdot 450) + (0,2668 \cdot 280) + (0,0901 \cdot 320) = 376,55 \text{ ppm}$$

Este resultado é bastante próximo ao obtido pelo método do inverso da distância, mas ambos os valores estão a indicar simplesmente aproximações com relação ao verdadeiro valor de A, havendo necessidade do estabelecimento de um intervalo de confiança.

A variância associada à estimativa pela krigagem é:

$$S_k^2 = 0,3760(322,7) - 0,0347(581,3) + 0,3017(420,9) + 0,2668(405,5) + 0,0901(714,9) + 10,0816 = 410,8867$$

$$S_k = 20,27034$$

Usando este desvio padrão para estabelecer um intervalo de confiança e supondo que a distribuição dos valores da estimativa apresente distribuição normal em torno do valor real e que, portanto, 95% dessa distribuição estão no intervalo de mais ou menos 1,96 desvios padrão, tem-se que o intervalo de confiança é da ordem de $\pm 19,77 \cdot 1,96 = 39,73$ (Fig. 12).

O verdadeiro valor do ponto A deve estar, portanto entre 336,82 e 416,28 ppm. Esse resultado para o intervalo de confiança já era esperado tendo em vista que no modelo variográfico, no qual a krigagem foi baseada, o valor para o efeito pepita é de

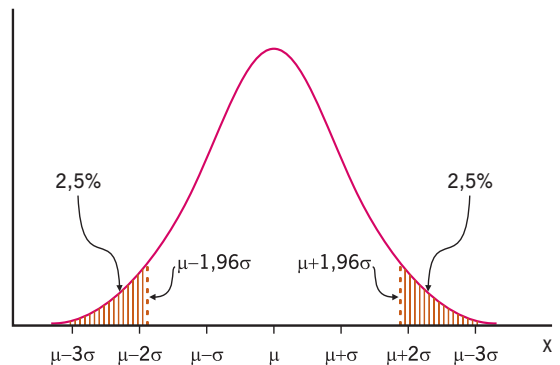


Figura 12 – Intervalo de confiança em torno do valor estimado por krigagem ordinária

100 pés. Pela equação do semivariograma é de se esperar que o valor de λ_0 seja zero, mas quando isso não ocorre é indicio de descontinuidade para distâncias menores do que a menor distância entre os pontos e que parte dessa descontinuidade pode ser devida a erros de medição ou a variabilidade numa escala menor do que aquela amostrada. Quanto menor a proporção do efeito pepita para o patamar, menor a variância da estimativa e, portanto, maior a confiança que se pode ter na estimativa.

Esta medida do erro associado à estimativa é que distingue o método da krigagem em relação aos demais algoritmos e, desse modo, quando da sua aplicação sempre resultam dois mapas: um com a distribuição dos valores interpolados e outro com a variâncias, ou o desvios padrão, referentes aos valores estimados. Nestes a localização dos menores valores coincide com as áreas com maior densidade de pontos.

Um exemplo é apresentado a seguir para uma distribuição de 60 pontos (Fig. 13).

A análise variográfica indicou que o melhor modelo ajustado seria o exponencial com $C = 11,84$; $a = 129,58$ e $C_0 = 4,8$ (Fig. 14). Baseado nesse modelo, a krigagem ordinária forneceu o mapa com valores estimados (Fig. 15) e o respectivo mapa com os valores de desvios-padrão da krigagem (Fig. 16).

A krigagem, entendida como um estimador que se baseia em técnicas de análise de regressão, procura

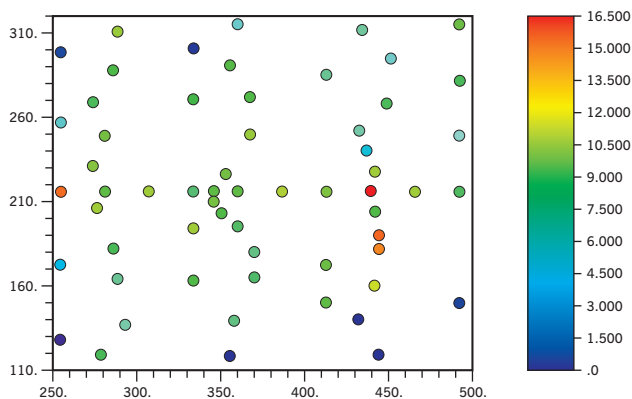


Figura 13 – Mapa de localização de pontos de amostragem

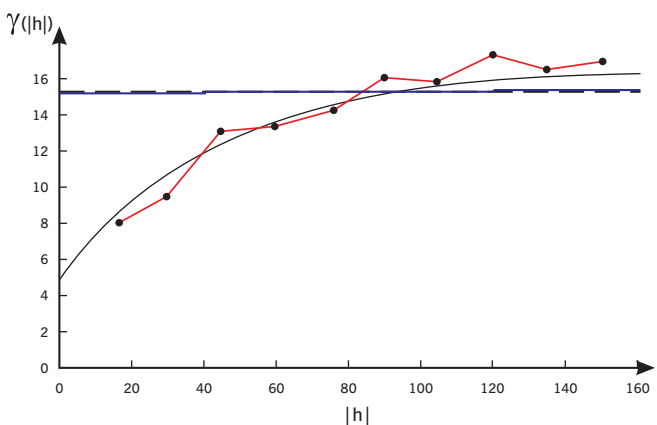


Figura 14 – Modelo de variograma ajustado ao variograma experimental para os dados do exemplo da Figura 13

ra minimizar a variância estimada a partir de um modelo prévio que leva em conta a dependência estocástica entre os dados distribuídos no espaço. Com a krigagem é calculada a melhor estimativa possível para locais não amostrados, pela minimização da variância do erro.

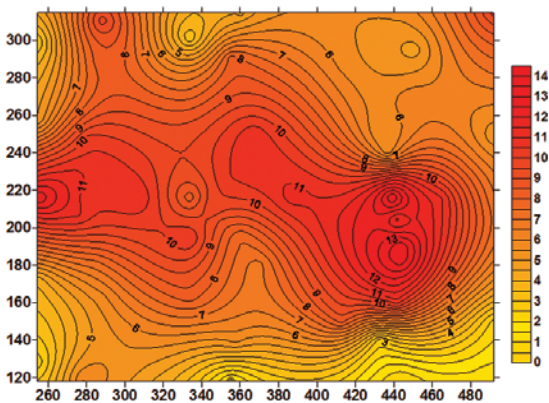


Figura 15 – Mapa estimado pelo método da krigagem ordinária para os dados do exemplo da Figura 13

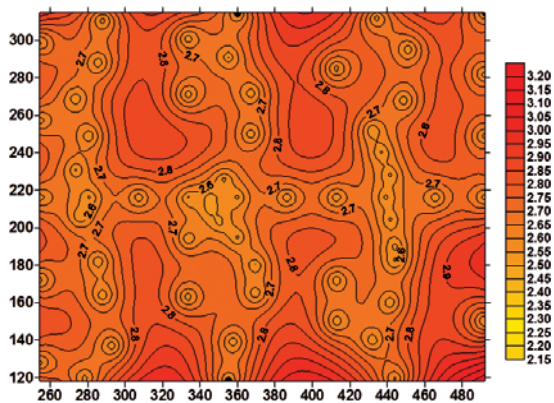


Figura 16 – Mapa de desvios-padrão da krigagem associados aos valores estimados da Figura 15

Considerações finais

Neste texto foi dado ênfase a uma das possíveis aplicações da Geoestatística, ou seja, na elaboração de mapas de contornos, os quais mostram linhas que representam medidas interpoladas com iguais valores. Valores regionalizados esses que mudam de local para local e apresentam uma variabilidade espacial. Quando de um levantamento topográfico, o cartógrafo tem condições de ver o terreno e decidir coletar mais pontos nas regiões mais acidentadas, ou seja, com maior variabilidade, e menos nas regiões planas, com menor variabilidade, com a intenção de melhor representar no mapa, em duas dimensões, a realidade em três dimensões. É possível que, nesta circunstância, a krigagem nem seja de grande valia. Ela se torna, porém, importante para o mapeamento de superfícies outras, que procuram modelar a distribuição de variáveis geoquímicas, geofísicas, hidrogeológicas, geotécnicas, etc.

Mesmo sendo a krigagem o melhor estimador linear não tendencioso, o método apresenta, porém, limitações.

Isso porque o semivariograma mostra a medida do grau de dependência espacial entre valores e é uma medida da variabilidade em relação à distância. A krigagem usa essas informações para encontrar os pesos ótimos a serem associados às amostras que irão estimar um ponto e a variância da krigagem é independente dos valores dos pontos usados para obter os estimadores Z_i^* e mede somente a configuração espacial dos dados.

Além disso, não há garantia que o mapa obtido pela krigagem tenha o mesmo histograma, a mesma variância e o mesmo semivariograma dos dados originais, pois trata-se, pela própria natureza do método, de um mapa com valores suavizados, sendo a suavização inversamente proporcional à densidade dos pontos amostrados.

A Geoestatística, porém, não é apenas krigagem, como aqui apresentado. Existem também à disposição as ferramentas da simulação, que permitem infinitas realizações de mapas, cada qual com aproximadamente o mesmo semivariograma e a mesma variância que os dados originais. Teoricamente a média de um grande número de mapas simulados deve fornecer resultados mais reais e, conseqüentemente, mais confiáveis para predições. Numa frase final: a simulação tenta atingir realismo e a estimativa por krigagem, acurácia.

Agradecimentos

Aos dois relatores que com suas críticas construtivas melhoraram, e muito, a qualidade deste texto.

Referências

- Chilés J.P.; Delfiner P. 1999. *Geostatistics : modeling spatial uncertainty*. New York: John Wiley and Sons. 695p.
- Clark I. 1979. *Practical Geostatistics*. London: Applied Science Publishers Ltd. 129p.
- <<http://uk.geocities.com/drisobelclark/practica.html>>. Último acesso: 24/12/2006.
- Duval R.; L'Évy R.; Matheron G. 1955. Travaux de D. G. Krige sur l'évaluation des gisements dans les mines d'or Sud-africaines. *Annales des Mines*, Paris, **12**:3-49.
- Krige D.G. 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, Dec. 1951:119-139.
- Guerra P.A.G. 1988. *Geoestatística operacional*. Rio de Janeiro: Departamento Nacional da Produção Mineral.
- Houlding S.W. 2000. *Practical Geostatistics : modeling and spatial analysis*. Berlin: Springer Verlag. 159p. +CD.
- Journel A.G., Huijbregts C.J. 1978. *Mining Geostatistics*. London: Academic Press. 600p.
- Matheron G. 1962. *Traité de Géostatistique Appliquée, Tome I : Mémoires du Bureau de Recherches Géologiques et Minières*. Paris: Editions Technip, vol. 14.
- Matheron G. 1963. *Traité de Géostatistique Appliquée, Tome II : Mémoires du Bureau de Recherches Géologiques et Minières*. Paris: Editions Technip, vol. 24.
- Matheron G. 1965. *Les variables régionalisées et leur estimation: une application de la théorie des fonctions aléatoires aux sciences de la nature*. Paris: Masson et Cie.
- Pannatier Y. 1996. *Variowin*. Software for Spatial Data Analysis in 2D. New York: Springer-Verlag. 91p + Disquete 3,5".
- <<http://www-sst.unil.ch/research/variowin/index.html>>. Último acesso: 24/12/2006.
- Valente J.M.G.P. 1982. *Geomatemática*. Lições de Geoestatística: Ed. Fundação Gorceix, vol. I-VIII.
- Yamamoto J.K. (ed.) 2001. *Avaliação e classificação de reservas minerais*. São Paulo: Edusp.

Submetido em 23 de fevereiro de 2006.

Aceito em 18 de julho de 2006.