

KRIGAGEM INDICADORA E PROBABILÍSTICA
PARA A ESTIMAÇÃO DAS
RESERVAS RECUPERÁVEIS

Este exemplar corresponde a redação
final da tese devidamente corrigida
e defendida pelo Sr. Bernardo Moisés
Lagos Alvarez e aprovada pela
Comissão Julgadora.

Campinas, 07 de Novembro de 1991


Prof. Dr. Ademir José Petenate
Orientador

Dissertação apresentada ao Instituto
de Matemática, Estatística e
Ciência da Computação, UNICAMP, como
requisito parcial para obtenção
do Título de Mestre em Estatística.

L137k

14938/BC

UNICAMP
BIBLIOTECA CENTRAL

À Violetta
e
Leonardo.

AGRADECIMENTOS

Ao Professor Dr. Ademir José Petenate, pela orientação dedicada a este trabalho.

Ao Professor Dr. Armando Zaupa Remacre do Instituto de Geociências da UNICAMP, pela valiosa informação entregue.

Ao CNPq, CAPES, pelo auxílio financeiro oferecido.

Aos Professores, Funcionários e Colegas do IMECC-UNICAMP, pelo conhecimento, experiência e consideração.

Meus sinceros agradecimentos.

SUMARIO

	págs. nº.
Capítulo 0. INTRODUÇÃO	1-24
1. Breve Descrição do Problema	01
1.1. O Problema	03
1.2. Os Objetivos	05
1.3. A Geoestatística	06
2. Noções Gerais de Processos Estocásticos	07
2.1. O Modelo	08
2.2. Hipóteses de Estacionariedades	10
2.3. Propriedades da Covariância e do Variograma	13
3. Função de Predição	15
4. A Krigagem	17
4.1. Krigagem Disjuntiva	19
4.2. Krigagem Simples	22
4.3. Krigagem Ordinaria	23
4.4. Krigagem Universal	23
Capítulo I. CONCEITOS ELEMANTARES PARA RESOLVER PROBLEMAS DE MINERAÇÃO	25-47
1. Variáveis Médias	25
1.1. Variância de Extensão	26
1.2. Variância de Dispersão	28
2. Dispersões Verdadeiras	33
2.1. Princípio de Conservação da Distribuição	35
3. Análise Estrutural	36
3.1. Modelos de Variogramas	37
Capítulo II. KRIGAGEM INDICADORA E PROBABILÍSTICA	48-68
1. Variável Indicadora	49
1.1. Transformada Indicadora	50
1.2. Funções Locais de Recuperação	53

2. Krigagem Indicadora	54
2.1 Krigagem Indicadora Simples	55
2.2 Krigagem Indicadora Ordinaria	57
2.3 Análise Estrutural	59
3. Cokrigagem Indicadora	64
3.1 Krigagem Probabilística	65
Capítulo III. PRÁTICA DA KRIGAGEM INDICADORA E PROBABILISTICA	69-85
CONSIDERAÇÕES FINAIS	86
APÊNDICE	90
REFERÊNCIAS	91

CAPÍTULO 0

INTRODUÇÃO

A natureza gera fenômenos distribuídos espacialmente, que são de interesse para diferentes áreas da ciência.

Considere particularmente a área da Geociência, onde um dos interesses é conhecer as características daqueles fenômenos observáveis que geram uma variável num depósito metalogênico.

Com a finalidade de situar o problema de interesse dentro das diferentes etapas pelas quais deve passar um reconhecimento geológico, é apresentado a seguir uma breve descrição cronológica dos problemas típicos achados na área da mineração, desde seu primeiro reconhecimento geológico até sua exploração industrial.

1. BREVE DESCRIÇÃO DOS PROBLEMAS-TÍPICOS

Esquemáticamente o que se tem é uma região metalogênica favorável e nessa região busca-se identificar a existência de possíveis depósitos.

Geralmente, essa identificação é feita usando um reconhecimento geológico sistemático e cego.

Se os resultados obtidos nessa identificação forem

indicativos de um possível depósito, então, procede-se a uma interpretação desses resultados.

Uma vez percorridos os passos acima, com o resultado da localização do depósito, só é possível conhecer suas características *globais* por meio de estimativas das *características médias globais*.

Se as características globais estão de acordo com as de um possível depósito economicamente viável então continua-se descrevendo o possível depósito, através de estimativas das *características médias locais*.

Finalizada essa fase, com a estimação local aprovada, ou seja, a economicidade do depósito de mineral é viável, um problema surge quando se pretende extrair a totalidade destes recursos *in situ*. Os fatores que impedem a extração completa dos recursos estimados são de dois tipos:

- Por razões técnicas: tamanho de equipamento, problemas de organização.
- Por razões econômicas: o custo de extração.

Dessa forma, obriga-se a fazer uma seleção de um subconjunto dos recursos *in situ* que serão efetivamente explorados. Chamam-se esses recursos selecionados de *reservas* e a esses recursos exploráveis de *reservas recuperáveis*.

Para descrever o comportamento das *reservas* em forma local, tem sido usado na prática o estudo das características médias da subárea $v(y)$ a ser selecionada como mineral (o teor médio de $v(x)$ é maior que um teor fixo) dentro da área maior $V(x)$. Chamam-se essas características médias de *funções locais de recuperação*.

O depósito é posto em exploração ao término desse estudo de *reservas*.

Esta dissertação focaliza um dos problemas-tipos. Esse problema está no estudo das *reservas*, mais especificamente, no problema das *reservas recuperáveis*.

1.1. O Problema: Reservas Recuperáveis

Quando se busca selecionar um bloco de minério do depósito, é feito um controle de qualidade do mineral recuperado ou seja um controle das condições de exploração. Mais geralmente procura-se uma otimização de uma certa função, por exemplo a função benefício, que depende das características do mineral extraído.

Considerando o exemplo da função benefício que é da forma

$$B\{ q_v(z) , t_v(z) \} = q_v(z) - wt_v(z)$$

onde $q_v(z)$ representa a quantidade de metal útil extraído, $t_v(z)$ a quantidade de metal extraído e w é um parâmetro do processo de custo de extração.

As características $q_v(z)$ e $t_v(z)$ do mineral extraído, dependem de dois parâmetros de cortes, v e z .

A cada conjunto (v,z) dos parâmetros de corte corresponde um certo mineral recuperado com as características, $q_v(z)$ e $t_v(z)$ e um certo valor de benefício $B\{ q_v(z) , t_v(z) \}$.

O problema é encontrar o conjunto de parâmetros (v^0, z^0) ótimos que maximizam a função benefício sob certas condições. A função B , das características do mineral eventualmente recuperado, chama-se *critério de seleção* e aos parâmetros ótimos da seleção de *parâmetros de cortes* (M. Alfaro, 1975).

O termo *recuperável* adotado na Geoestatística refere-se a uma situação muito particular, do tipo:

- O meio técnico da exploração pode ser resumido na definição de um volume mínimo de seleção, chamado de *bloco*;
- O meio econômico pode ser resumido na definição de uma seleção de corte aplicado ao teor médio do volume mínimo de seleção.

O passo básico de seleção é aplicado à unidade de volume de seleção, chamada de seleção por bloco.

O critério de seleção é um corte aplicado a um teor médio estimado z_v^* . Nesse caso, a seleção é chamada de *indireta*. Se o corte é aplicado ao teor médio verdadeiro z_v , a seleção é chamada de *direta*. Se o bloco é selecionado sem qualquer restrição geométrica, a seleção é dita *livre*.

Na estimação dos *recursos*, o corpo mineral é dividido em subáreas não necessariamente iguais (mas na prática geralmente são iguais) de dimensões geralmente definidas pela malha regular das amostras. Chamam-se essas subáreas de *painéis*.

Na estimação recuperável o corpo mineral também será dividido em painéis regulares de volume V , e se assume que o mineral é selecionado por:

- blocos de tamanho v ,
- sem restrição geométrica (seleção livre),
- aplicação de um critério de corte no teor do bloco $Z_v(x)$ (seleção direta).

Definindo o termo *recuperável* e a forma de *seleção* dos *recursos*, têm que se definir as ferramentas para poder fazer o estudo desses *recursos selecionados* ou *reservas*. Isso é feito a seguir.

Para cada painel de volume V definem-se as *funções de recuperação*

- Tonelagem recuperável num corte z

$$t(z) = 1 - F_V(z).$$

Na hora de apresentar os resultados práticos a tonelagem recuperável é da forma $t(z)TV$, onde TV é a tonelagem total do painel com volume V .

- Quantidade de metal recuperável num corte z

$$q(z) = E\{ [1 - I(x;z)] Z_V(x) \}$$

onde

$$F_V(z) = P(Z_V(x) \leq z)$$

e

$$I(x;z) = \begin{cases} 1 & \text{se } Z(x) \leq z \\ 0 & \text{se não.} \end{cases} \quad (1)$$

Observa-se que as *funções de recuperação* dependem diretamente da função de distribuição $F_V(z)$, interpretada como a *proporção de teores num bloco-v, acima de um corte z , num painel V* (A. Marechal, 1984).

1.2. O Objetivo

Essa proporção de teores bloco-v, num corte z , é o maior interesse na área das reservas recuperáveis da Geoestatística. Há diferentes métodos para poder estimar essa proporção que varia de acordo com as restrições feitas ao processo que modela o comportamento da formação dos teores de mineral num depósito.

Esta dissertação foi orientada de forma a fazer uma revisão de um dos métodos usados para poder estimar a função de distribuição $F_V(z)$ que é a *Krigagem Probabilística*.

- Apresentando o método
- Discutindo suas vantagens e desvantagens

- **Mostrando os passos fundamentais na prática para que seja implementado.**

1.3. A Geoestatística

Nos anos 30 na África do sul, S. H. Siechel foi o primeiro a aplicar os métodos estatísticos em problemas de *reservas recuperáveis* (reservas selecionadas para a exploração), obtendo teores de ouro notoriamente erráticos e um histograma com uma alta curtose e uma assimetria negativa, sugerindo a não normalidade dos dados. Algumas inconveniências existem nesse método: as amostras não são independentes, provêm de uma distribuição *log-normal* e não considera a dependência das localidades onde as amostras estão situadas. Nos anos 50, o engenheiro D. G. Krige também fez ajustes aos teores de ouro de suas amostras com uma distribuição *log-normal*, só que essas distribuições são funções das dimensões do suporte (volume no qual a variável está definida) das amostras (I. Clark, 1979).

Em 1951, D. G. Krige incorporou a posição e a relação espacial, estudando as regressões entre os teores verdadeiros dos blocos (unidade de mineração) já explorados e as médias móveis dos teores das amostras, definindo novos estimadores que não estão necessariamente ligados à hipótese de *log-normalidade*. Desses estimadores de D. G. Krige originou-se o nome da técnica de Krigagem.

Finalmente, em 1962, G. Matheron, matemático probabilista, traduziu toda essa experiência da escola Sul-africana numa linguagem probabilística formal, que ele chamou de *Geoestatística* e a definiu como a *aplicação do formalismo das funções aleatórias ao reconhecimento e estimação dos fenômenos naturais*. (A. G. Journal, 1978).

A Geoestatística tem como base a observação de toda variável de interesse (teor, potência, acumulação, etc) de um fenômeno metalogênico distribuída num espaço e possui uma função de autocorrelação espacial. Os teores de um metal em duas localidades estão autocorrelacionados, no sentido mais geral da covariância. Supõe-se que a autocorrelação depende do vetor diferença (módulo

e direção) entre as duas localidades. Essa autocorrelação é caracterizada por um momento de segunda ordem chamado de *variograma*, que é a ferramenta estrutural fundamental da Geostatística. Todo o trabalho reside na modelagem do variograma a partir das informações disponíveis. Um bom modelo é, em geral, produto da prática e de um bom conhecimento do fenômeno a ser estudado. É necessário notar que no momento de estimar o variograma a partir dos dados é que se coloca o questionamento das *hipóteses de estacionariedade*. Se houver uma não estacionariedade devido a uma tendência (por exemplo, espera-se que algumas áreas sejam ricas e outras pobres), essa esperança pode ser expressa por uma variação razoavelmente regular. Ao redor dessa tendência espera-se ter uma variação aleatória, ou seja, o valor em qualquer lugar do depósito está composto de um componente fixo da tendência (provavelmente desconhecida), e um outro componente, uma variável residual aleatória. Uma vez eliminada essa tendência é possível estimar o variograma do fenômeno residual estacionário (N. Cressie, 1986).

Havendo a necessidade de prever o teor $Z(x)$ de um metal numa localidade x , tendo como base uma informação conhecida de teores na vizinhança de x , $\{Z(x_i)=Z_i; i=1 \text{ a } n\}$, o procedimento de predição conhecido com o nome de *Krigagem* fornece o melhor preditor linear não-viciado (BLUP) para o teor $Z(x)$. O objetivo não é prever o teor $Z(x)$ mas a proporção do teor médio do bloco com suporte de volume v (bloco- v), acima de um valor de corte z_0 , ou seja prever uma transformação não-linear nos teores médios dos blocos. Usar o preditor de *Krigagem Linear* (A. G. Journel, 1977) é inadequado e é perigosamente viciado. Os métodos que resolvem esse problema são: *Krigagem Disjuntiva* (KD) (G. Matheron, 1975) e *Krigagem Gaussiana Multivariada* (KG), necessitando nesse caso de hipóteses em relação às distribuições bidimensionais. O método *Krigagem Probabilística* (KP), que é um método melhorado da *Krigagem Indicadora* (KI), também é utilizado nessas situações embora apresente inconveniências, mas não faz nenhuma hipótese sobre a distribuição multivariada de (Z_1, Z_2, \dots, Z_n) , nem das distribuições bidimensionais de $(Z_i, Z_j), i, j = 1 \text{ a } n, i \neq j$.

2. NOÇÕES GERAIS DE PROCESSOS ESTOCÁSTICOS

Em um fenômeno metalogênico a prática tem mostrado que uma medição do teor de um metal $Z(x)$ em uma localidade x , x em um depósito G , G

subconjunto de R^2 ou R^3 com limites bem definidos, tem um comportamento localmente errático e que por trás desse comportamento, também existe um comportamento característico ou estrutural da variabilidade espacial, ou seja, duas medições do teor em localidades (pontos) diferentes x e $x+h$, em geral não são independentes, mas estão relacionadas por uma função de correlação que caracteriza esse comportamento estrutural. Essas características de aleatoriedade estrutural dos teores do metal num depósito são fundamentais, portanto há que considerá-las no modelo que vai descrever o comportamento do processo que gera os teores de um metal num depósito G . Aparte dessas considerações outras restrições são feitas ao modelo referentes à aplicabilidade deste na prática. Quatro são as principais características que o modelo tem que satisfazer, descritas a seguir.

2.1. O Modelo

Para que o modelo utilizado tenha validade prática quatro são as principais condições a serem satisfeitas, resumidas em:

- (i). A inferência Estatística deve ser possível.
- (ii). O modelo deve ser operacional, construído para responder aos objetivos. Cada resposta particular pode não requerer todas as hipóteses imposta ao modelo em uma primeira tentativa.
- (iii). Isso está relacionado a uma terceira restrição; o modelo deve ser compatível com os dados físicos.
- (iv). Finalmente as predições devem ser verificadas. Um modelo pode diferir amplamente desde a predição. Isso significa que as características do sistema tem fugido do pesquisador (M. David, 1978).

A princípio o sistema é modelado por um *processo estocástico*. Mas na prática não é possível conhecer a lei que governa esse processo estocástico. É aqui onde se obriga a impor restrições ao modelo, não ao sistema, para poder fazer inferência, não da lei, mas de suas principais características.

A seguir são apresentadas algumas definições básicas de processos estocásticos.

Um *processo estocástico* Z é uma família de variáveis aleatórias indexadas por um conjunto apropriado

$$Z = \{Z(x); x \in G\}, G \text{ subconjunto de } \mathbb{R}^2 \text{ ou } \mathbb{R}^3.$$

O processo estocástico Z fica completamente especificado pelo conjunto das funções de distribuições conjuntas de (Z_1, Z_2, \dots, Z_n) para todo inteiro n e para todo $x_i \in G; i=1$ a n , escrita na forma

$$F = \{ F_{z_1, z_2, \dots, z_n}(z_1, z_2, \dots, z_n); n \in \mathbb{N}, x_i \in G \text{ } i=1 \text{ a } n \}$$

senda F chamado de *conjunto das distribuições finita dimensionais*.

Definem-se as seguintes funções do processo estocástico Z :

- **Função média**

$$m(x) = E\{ Z(x) \}, \forall x \in G.$$

- **Função covariância não-centrada**

$$\text{Cov}_n(x_1, x_2) = E\{ Z(x_1)Z(x_2) \}, \forall x_1, x_2 \in G.$$

- **Função covariância centrada**

$$\text{Cov}(x_1, x_2) = E\{ [Z(x_1) - m(x_1)] [Z(x_2) - m(x_2)] \}, \forall x_1, x_2 \in G.$$

- **Variância**

$$\text{Cov}(x, x) = \text{Var}\{ Z(x) \}, \forall x \in G.$$

- **Variograma**

$$2\text{Vari}(x_1, x_2) = \text{Var}\{ Z(x_1) - Z(x_2) \}, \forall x_1, x_2 \in G.$$

- A função $\text{Vari}(x_1, x_2)$ é chamada de *semi-variograma* (a metade do variograma).

- Correlograma

$$\text{Corr}(x_1, x_2) = \frac{\text{Cov}(x_1, x_2)}{\sqrt{\text{Cov}(x_1, x_1)\text{Cov}(x_2, x_2)}}, \forall x_1, x_2 \in G, \text{ se a variância}$$

existe.)

2.2. Hipóteses de Estacionariedade

Uma das características desse tipo de processo é que só há uma medição registrada de $Z(x)$ em $x \in G$, ou seja, tem-se uma realização $\{z(x); x \in G\}$ do processo estocástico Z . Porém, o conhecimento de uma só realização do processo, não permite chegar à lei do processo estocástico Z . É necessário conhecer várias realizações do processo Z , para poder inferir sua lei ou ao menos as características principais.

Na prática, não só se dispõe de apenas uma realização, mas também essa é restrita a um número finito de medições de $Z(x)$ nos pontos x_i , $i=1$ a n . Para contornar isso, certas suposições são necessárias no modelo. Essas suposições envolvem vários graus de *homogeneidade espacial* e são introduzidas sob o conceito geral de *hipóteses de estacionariedade*.

Frequentemente o fenômeno em estudo pode ser considerado como homogêneo, isto é, tendo feitas as medições $\{z(x_i)=z_i; i=1 \text{ a } n\}$ das variáveis $\{Z_i; i=1 \text{ a } n\}$, qualquer subconjunto das medições feitas provem da mesma distribuição das medições transladadas. Essa homogeneidade é introduzida ao modelo como *hipótese estacionária estrita*.

A Geoestatística postula uma homogeneidade muito menos exigente, que não está nos valores $z(x_i)$ mas nos incrementos $[z(x_i) - z(x_i+h)]$, ou seja, que os incrementos $[z(x_1) - z(x_1+h)]$, $[z(x_2) - z(x_2+h)]$, ..., $[z(x_n) - z(x_n+h)]$ para qualquer h , sem importar onde estão localizados, provêm de uma mesma distribuição. Este tipo de homogeneidade é introduzida ao modelo com o nome de *hipótese intrínseca*. Consideram-se então os incrementos como n realizações do incremento aleatório $[Z(x) - Z(x+h)]$.

Para conhecer o comportamento do teor de minério no depósito G , então constrói-se, para cada vetor distância h , um histograma do incremento aleatório $[Z(x) - Z(x+h)]$, cada histograma sendo caracterizado pela média e variância. Toda essa informação fica resumida ao considerar uma função média e uma função de variância, que depende só da distância h e que corresponde justamente à função variograma.

Hipóteses de Estacionariedade Estrita

Um processo estocástico $Z = \{Z(x); x \in G\}$ é dito *estritamente estacionário*, se a distribuição conjunta de $(Z(x_1), Z(x_2), \dots, Z(x_n))$ é igual à distribuição conjunta de $(Z(x_1+h), Z(x_2+h), \dots, Z(x_n+h))$ para todo $x_i \in G, i=1$ a n e para qualquer $h \in \mathbb{R}^2$ ou \mathbb{R}^3 talque $x_i+h \in G, i=1$ a $n, \forall n \in \mathbb{N}$.

Hipóteses de Covariância Estacionária

O processo estocástico Z é dito de *covariância estacionária*, se houver um segundo momento finito, satisfazendo

- i) $E\{Z(x)\} = m, \forall x \in G$
- ii) $Cov(x, x+h) = C(h), \forall x, x+h \in G$.

Um processo estocástico de covariância estacionária também é dito estacionário no sentido amplo, ou fracamente estacionário.

A estacionariedade da covariância implica na estacionariedade da variância e do variograma.

$$\text{Var}\{Z(x)\} = E\{[Z(x) - m]^2\} = C(0), \forall x \in G.$$

$$\begin{aligned} \text{e} \quad 2\text{Vari}(x, x+h) &= \text{Var}\{Z(x) - Z(x+h)\} \\ &= \text{Var}\{Z(x)\} + \text{Var}\{Z(x+h)\} - 2\text{Cov}(x, x+h) \end{aligned}$$

$$2\text{Vari}(x,x+h) = 2C(0) - 2C(h), \quad \text{dependente só de } h. \quad (2)$$

Logo, tem-se que $\text{Vari}(x,x+h) = C(0) - C(h) = \gamma(h)$, se a covariância existir.

A estacionariedade restrita do processo não implica num processo de covariância estacionária.

Hipótese Intrínseca

Existem muitos fenômenos físicos e processos estocásticos que têm uma capacidade de dispersão infinita, isto é, que não têm uma variância a priori, nem uma covariância finita, mas é possível que seja definido um variograma.

Se houver uma hipótese de estacionariedade e o variograma existir então acontece uma estacionariedade do variograma.

O processo estocástico Z é dito *intrínseco* quando:

- i) a esperança existe e não depende do ponto x

$$E\{ Z(x) \} = m, \quad \forall x$$

- ii) para todo h o incremento aleatório $[Z(x) - Z(x+h)]$ tem uma variância finita que não depende de x

$$2\gamma(h) = \text{Var}\{ Z(x) - Z(x+h) \}, \quad \forall x.$$

Assim, a covariância estacionária implica na hipótese intrínseca, mas a recíproca não é verdadeira. Por exemplo, seja $\{ W(t); t \geq 0 \}$ um processo estocástico de Wiener, o incremento aleatório $[W(t) - W(t+h)]$ é normalmente distribuído com média $m = 0$, variograma $2\gamma(h) = \text{Var}\{ Z(x) - Z(x+h) \} = |h|$ que é função só de h e função covariância $C(t,t+h) = \min(t,t+h)$, que não é função só de h . Assim, a classe de processos intrínsecos contém à classe de processos de covariância estacionária (D. E. Myer, 1989).

2.3. Propriedades da Covariância e do Variograma

Supondo o processo de covariância estacionária e considerando uma combinação linear $Y = \sum_{j=1}^n \lambda_j Z_j$, cuja variância é não negativa, temos:

$$\text{Var}(Y) = \sum_{j=1}^n \sum_{i=1}^n \lambda_j \lambda_i C(x_i - x_j) \geq 0.$$

Então, uma condição necessária, para que qualquer função $g(h)$ seja uma função covariância é que seja *positiva semi-definida*.

A variância escrita em termos do semi-variograma, usando (2), fica

$$\text{Var}(Y) = C(0) \sum_{i=1}^n \lambda_i \sum_{j=1}^n \lambda_j - \sum_{j=1}^n \sum_{i=1}^n \lambda_j \lambda_i \gamma(x_i - x_j).$$

Caso o processo estocástico venha a ter uma capacidade de dispersão infinita, só a hipótese intrínseca é suposta. Então, a variância da combinação linear finita Y é definida quando:

$$\sum_{i=1}^n \lambda_i = 0,$$

resultando:

$$\text{Var}(Y) = - \sum_{j=1}^n \sum_{i=1}^n \lambda_j \lambda_i \gamma(x_i - x_j) \geq 0$$

ou

$$\sum_{j=1}^n \sum_{i=1}^n \lambda_j \lambda_i \gamma(x_i - x_j) \leq 0,$$

e assim, uma condição necessária para que qualquer função $f(h)$ seja uma função semi-variograma é que seja *negativa semi-definida*, com a condição nos pesos, $\sum_{j=1}^n \lambda_j = 0$, chamada de função *negativa semi-definida condicional* (condicionada a $\sum_{j=1}^n \lambda_j = 0$).

O teorema de Bochner (Armstrong M. e Diamond P. 1984), dá a condição necessária e suficiente, para que uma função $g(h)$ seja positiva semi-definida. A função $g(h)$ é contínua e positiva semi-definida se e só se esta é a transformada de Fourier de uma medida de Borel em \mathbb{R}^2 limitada positiva μ

$$g(h) = \int_{\mathbb{R}^2} \exp(i\langle u, h \rangle) d\mu(u).$$

Para um fenômeno modelado por um processo de covariância estacionária, na prática, a correlação entre as variáveis aleatórias $Z(x)$ e $Z(x+h)$ desaparece quando a distância h cresce indefinidamente:

$$C(h) \rightarrow 0, \text{ quando } |h| \rightarrow \infty.$$

Definimos $C(h)=0$ para $|h| \geq a$, onde a é a distância para que $C(h)$ seja aproximadamente zero. Essa distância a é chamada de *alcance* e representa a transição do estado onde existe uma correlação espacial ($|h| < a$), ao estado onde não existe uma correlação espacial ($|h| \geq a$). Essa distância a também pode ser interpretada como a amplitude média do corpo mineralizado para cada direção do vetor distância considerado. O alcance dá a idéia intuitiva de uma zona de influência do fenômeno.

Pela definição do variograma tem-se que:

$$\gamma(0) = 0, \gamma(h) = \gamma(-h) \geq 0$$

Se ao considerar processos com a covariância não-negativa então ao aumentar o valor h , $\gamma(h)$ cresce até $C(0)$ a partir de seu valor inicial zero, como consequência da expressão (1), e chega a ser mais ou menos estável em volta do valor limite $\gamma(\infty)$. Esse valor limite é chamado de valor *patamar* e corresponde à *variância a priori* do processo estocástico Z :

$$\gamma(\infty) = \text{Var}\{ Z(x) \} = C(0)$$

onde se associa um alcance a . Se a variável de interesse depende não só do ponto $x \in G$ mas de um outro valor z , o patamar vai ser uma função de z .

Tais variogramas caracterizados por um patamar e um alcance são chamados de modelos de *transição*, que correspondem não só a processos estocásticos intrínsecos, mas também a processos de covariância estacionária. No caso de um variograma não ter caracterizado o patamar é porque há um componente não estacionário. Isso pode ser visto mais claramente por meio de ter caracterizado as diferentes causas que explicam a variabilidade do teor $Z(x)$, que na realidade as causas aparecem em domínios de escalas diferentes, por exemplo:

- a nível do suporte ($h \approx 0$), existe uma variabilidade devido às medições
- uma segunda variabilidade aparece devido à transição de um elemento a outro ($h < 1$ cm)
- uma terceira variabilidade pode dever-se à alternância de estratos ($h < 100$ mts)
- e uma quarta variabilidade aparece devido à distribuição dos depósitos, que é modelada por uma estrutura média determinística $m(x) = E\{Z(x)\}$ ($h < 100$ Kms).

Ao focalizar a variabilidade devido aos depósitos ou devido à grande escala (maiores que o alcance do variograma construído a escala de observação), na hora de fazer crescer a distância h e o variograma não ter estabilidade ou seja que a estrutura determinística $m(x)$ também não tem estabilidade levando a perseverar uma tendência nos limites do depósito (A. G. Journel, 1978).

3. FUNÇÃO DE PREDIÇÃO

Quando um fenômeno é observado consegue-se uma realização finita $\{z_1, z_2, \dots, z_n\}$ do processo estocástico real Z . O problema é conhecer o valor $z(x)$ da variável $Z(x)$ no ponto $x = x_0$. O melhor que se pode fazer é construir uma função real mensurável g que dependa das n variáveis aleatórias Z_1, Z_2, \dots, Z_n , usando seus valores na função g , para obtermos um valor real. Esse número é chamado de *preditor* (estimador) do valor de $Z(x_0)$ ($Z(x_0) = Z_0$).

Existem muitos conjuntos possíveis de funções de predição, dependendo das restrições feitas ao processo estocástico Z , como também da

medida de erro cometida ao predizer a variável aleatória Z_0 .

O conjunto mais geral de funções de predição é o conjunto das funções Borel mensuráveis de Z_1, Z_2, \dots, Z_n .

A seguir serão definidos alguns conceitos básicos.

É definido o *erro de predição* $R(x_0)$ como:

$$R(x_0) = g(Z_1, Z_2, \dots, Z_n) - Z(x_0).$$

$R(x_0)$ é uma função aleatória e, para obter uma medida quantificada do erro cometido na predição, simplesmente poderia ser usado o valor esperado do erro de predição $E\{R(x_0)\}$, porém é usual considerar-se uma função perda $L(R(x_0))$ e então calcular o valor esperado dessa função perda. Por conveniência, é que se considera $L(R(x_0)) = R^2(x_0)$. Então, a medida do erro de predição é o *erro quadrático médio* (MSE) ou *variância de predição*, se o preditor é não-viciado para $Z(x_0)$

$$E\{R^2(x_0)\} = E\{[g(Z_1, Z_2, \dots, Z_n) - Z(x_0)]^2\}.$$

Agora, o conjunto de funções de predição é restrito ao conjunto das funções Borel-mensuráveis de Z_1, Z_2, \dots, Z_n de quadrado integrável denotado por $L_2(\Omega, \mathfrak{F}, P)$ ou simplesmente por L_2 , onde \mathfrak{F} é a σ -álgebra gerada pelas n variáveis aleatórias Z_1, Z_2, \dots, Z_n , e P é a medida de probabilidade definida em (Ω, \mathfrak{F}) .

Uma função $g^* \in L_2$ é dita a *melhor função de predição* (BPF) de Z_0 baseada em Z_1, Z_2, \dots, Z_n , se

$$\text{MSE}(g^*) \leq \text{MSE}(g), \quad \forall g \in L_2. \quad (3)$$

Diante do exposto acima as perguntas que devem ser respondidas são:

- existe g^* ?,
- se existe, é g^* única?.

Seja o conjunto H_n de funções Borel-mensuráveis, de quadrado integrável com um produto interno definido pela covariância não-centrada

$$\langle f, g \rangle = E\{fg\} \quad \forall f, g \in H_n,$$

ou seja, que o conjunto H_n é implementado por uma estrutura geométrica ordinária, isto é, H_n é um espaço de Hilbert e L_2 é um espaço fechado de H_n . Pelo teorema da projeção, existe um elemento único $g^* \in L_2$ que satisfaz (3), isto é, um preditor de Z_0 é a projeção de Z_0 no espaço L_2 .

4. A KRIGAGEM

A palavra *krigagem* é um termo para designar o procedimento clássico de seleção do preditor com uma variância de predição mínima, em uma classe de possíveis preditores (lineares ou não-lineares). Essa variância de predição pode ser vista como a distância ao quadrado entre a variável desconhecida e seu preditor. O processo de minimização da distância poderia ser visto como a projeção da variável desconhecida no espaço onde se procura esse preditor.

Na prática, o problema é encontrar o melhor preditor possível do teor médio de um bloco, levando em consideração as variáveis amostradas, avaliadas dentro ou fora do bloco a predizer. D. G. Krige, no ano 1951, propôs um procedimento de regressão para designar pesos a cada variável avaliada. G. Matheron formalizou e generalizou esse procedimento de regressão, denominando-o de *Krigagem*, que definiu como *o procedimento probabilístico de obter o melhor* (no sentido da minimização da variância de predição) *preditor linear não-viciado de uma variável desconhecida*. Em 1975, G. Matheron fez uma maior generalização, obtendo preditores não-lineares e não-viciados, que chamou de *Krigagem Disjuntiva* (KD) (A. G. Journal, 1977).

Seja Z o processo estocástico que modela o fenômeno, tendo n variáveis avaliadas de Z , $\{Z_1, Z_2, \dots, Z_n\}$, com o objetivo de predizer uma variável $Z(x_0)$ com base nessa realização finita do processo estocástico Z . Então o procedimento é como se segue.

Considera-se o espaço vetorial H_n , implementado com um produto interno definido por:

$$\langle f, g \rangle = E\{fg\} \quad \forall f, g \in H_n.$$

Considera-se também o espaço vetorial H de todas as funções mensuráveis das $n+1$ variáveis aleatórias $\{Z_0, Z_1, Z_2, \dots, Z_n\}$ de quadrado integrável. Assim, a esperança condicional de Z_0 dado Z_1, Z_2, \dots, Z_n

$$g(Z_1, Z_2, \dots, Z_n) = E\{Z_0/Z_1, Z_2, \dots, Z_n\} = E_n\{Z_0\}$$

definida no espaço H é o projetor ortogonal de Z_0 no subespaço L_2 . (Nevéu, 1965), ou seja:

$$g^*(Z_1, Z_2, \dots, Z_n) = E_n\{Z_0\}$$

é a BPF de Z_0 , mais, ainda, a esperança condicional $E_n\{Z_0\}$ é não-viciada para Z_0

$$\begin{aligned} E\{E_n\{Z_0\} - Z_0\} &= E\{E_n\{Z_0\}\} - E\{Z_0\} \\ &= E\{Z_0\} - E\{Z_0\} \\ &= 0. \end{aligned}$$

Então, a esperança condicional fica caracterizada pelas propriedades de:

(i) não-viciabilidade

$$\langle E_n\{Z_0\} - Z_0, 1 \rangle = 0$$

(ii) e ortogonalidade

$$\langle E_n\{Z_0\} - Z_0, Y \rangle = 0, \quad \forall Y \in L_2.$$

Concluindo que a esperança condicional é o melhor preditor não-viciado, segundo a definição que enrega a Geoestatística, a esperança condicional é o preditor de Krigagem.

Para poder construir um preditor de Krigagem

$$E_n(Z_0) = g^*(Z_1, Z_2, \dots, Z_n) = g^* \in L_2$$

tem que se expressar o produto interno

$$\langle g^* - Z_0, Y \rangle = E((g^* - Z_0) Y), \quad \forall Y \in L_2.$$

Que necessita do conhecimento da distribuição das $n+1$ variáveis Z_0, Z_1, \dots, Z_n .

Na prática, a esperança condicional pode ser construída só em casos muitos particulares, por exemplo, se o processo estocástico Z é Gaussiano e estacionário, nesse caso a esperança condicional toma a forma de:

$$g^* = E_n(Z_0) = \sum_{j=1}^n \lambda_j Z_j.$$

(Graybill, 1975).

4.1. Krigagem Disjuntiva

No caso, onde se tem só uma realização finita do processo estocástico Z não é possível a inferência da distribuição conjunta das $n+1$ variáveis $Z_0, Z_1, Z_2, \dots, Z_n$, ficando a esperança condicional desconhecida.

Frequentemente, o que se faz nesses casos é reduzir o espaço onde se está projetando a variável por predizer, Z_0 ; assim, o espaço H_n é reduzido ao espaço gerado por todas as combinações lineares finitas $\sum_{j=1}^n \lambda_j Z_j$ das n variáveis avaliadas. Entretanto, a idéia do processo Krigagem Disjuntivo é de projetar Z_0 num espaço que inclua o espaço gerado pelas combinações lineares finitas, e que esteja incluído no espaço H_n .

A construção do espaço onde se quer projetar Z_0 , denota-se por D_n e é gerado pela soma de n funções mensuráveis de uma variável, isto é:

$$D_n = \{ g_1(Z_1) + g_2(Z_2) + \dots + g_n(Z_n) \}.$$

A projeção Z_{DK}^* de Z_0 em D_n é caracterizada pelas duas condições anteriores:

$$(i) \quad Z_{DK}^* \in D_n, \text{ isto é, } Z_{DK}^* = \sum_{j=1}^n g_j(Z_j)$$

(ii) o vetor $Z_0 - Z_{DK}^*$ é ortogonal para qualquer vetor $Y \in D_n$, isto é,

$$\langle Z_0 - Z_{DK}^*, Y \rangle = 0, \quad \forall Y \in D_n.$$

Desenvolvendo a condição (ii), tem-se:

$$\langle Z_0, \sum_{j=1}^n g_j(Z_j) \rangle = \langle Z_{DK}^*, \sum_{j=1}^n g_j(Z_j) \rangle; \quad \forall g_j \text{ mensurável}$$

$$\sum_{j=1}^n \langle Z_0, g_j(Z_j) \rangle = \sum_{j=1}^n \langle Z_{DK}^*, g_j(Z_j) \rangle; \quad \forall g_j \text{ mensurável}$$

equivalentemente a

$$\langle Z_0, g_j(Z_j) \rangle = \langle Z_{DK}^*, g_j(Z_j) \rangle$$

por ter que ser satisfeita para toda função g_j mensurável.

Como D_n é gerado por funções mensuráveis $g_j(Z_j)$ dependendo de uma só variável Z_j , a condição (ii) fica como:

$$\langle Z_0, f(Z_i) \rangle = \langle Z_{DK}^*, f(Z_i) \rangle; \quad \forall i=1 \text{ a } n \quad (4)$$

e para toda função f mensurável de uma só variável. Essa condição é satisfeita se e só se:

$$E\{ Z_0/Z_i \} = E\{ Z_{DK}^*/Z_i \}; \quad i=1 \text{ a } n$$

A prova dessa passagem encontra-se no apêndice, página 90.

Então as funções g_j do preditor KD são caracterizadas pelo sistema de n equações

$$\sum_{j=1}^n E\{ g_j(Z_j)/Z_i \} = E\{ Z_0/Z_i \}; \quad i=1 \text{ a } n. \quad (5)$$

Dessa forma fica completamente caracterizado o preditor KD, Z_{DK}^* . O sistema (4) é chamado de *sistema KD*.

Vê-se que o sistema KD depende só das esperanças condicionais de Z_0 dado Z_j e de Z_i dado Z_j . Para conhecer essas esperanças são necessárias as distribuições bidimensionais de (Z_0, Z_j) e de (Z_i, Z_j) ; $j=1$ a n .

Geralmente, não é possível resolver o sistema KD porque está na forma de equações integrais. É possível resolvê-lo fazendo uma discretização apropriada para obter funções aproximadas por expansões limitadas em polinômios ortogonais (G. Matheron, 1975).

Esse procedimento foi desenvolvido com o propósito de prever certas variáveis que dependem de uma forma não-linear nos dados avaliados, por exemplo, para a proporção de valores de teor acima de um dado valor de corte.

A resolução do sistema KD (4) proporciona o estimador KD Z_{DK}^* e a variância de predição corresponde a:

$$\sigma_{DK}^2 = E\{ (Z_0 - Z_{DK}^*)^2 \} = \|Z_0 - Z_{DK}^*\|^2.$$

Como o espaço D_n está contido no espaço H_n , a projeção de Z_0 em H_n está mais perto a Z_0 que a projeção de Z_0 em D_n . Em outros termos:

$$\begin{aligned} \sigma_{DK}^2 &= \|Z_0 - Z_{DK}^*\|^2 = \|Z_0 - E_n(Z_0)\|^2 + \|E_n(Z_0) - Z_{DK}^*\|^2 \\ &= \|Z_0 - E_n(Z_0)\|^2 + \|E_n(Z_0) - Z_{DK}^*\|^2 \geq \sigma_E^2 \end{aligned}$$

Se não são conhecidas as distribuições bidimensionais de (Z_0, Z_j) e de (Z_i, Z_j) , é usual restringir o espaço H_n , mais ainda do que o espaço D_n , fazendo com que as funções $f_j(Z_j)$ sejam lineares, isto é, $f_j(Z_j) = \lambda_j Z_j$. Dessa forma o espaço de procura para o melhor preditor de Z_0 , é o espaço vetorial E_n definido pelo conjunto de todas as combinações lineares finitas dos elementos do processo estocástico Z

$$E_n = \left\{ \sum_{j=1}^n \lambda_j Z_j; Z_j \in Z, \lambda_j \in R \right\}.$$

Feita essa restrição, o processo Krigagem, chamado de *Krigagem Linear*, é definido como o procedimento de projetar a variável desconhecida Z_0 em E_n ou em qualquer variedade linear $C \subset E_n$ de maneira que se possa garantir a não-viciabilidade do preditor g^* . Agora denota-se g^* por Z^* .

Condição de Não-viciabilidade

Considere-se a esperança de $Z^* \in E_n$

$$E\{Z^*\} = E\left\{\sum_{j=1}^n \lambda_j Z_j\right\} = \sum_{j=1}^n \lambda_j E\{Z_j\} = \sum_{j=1}^n \lambda_j m(x_j).$$

O preditor Z^* é não-viciado para Z_0 , se e só se, $E\{Z^*\} = E\{Z_0\}$, em termos da função média do processo estocástico Z , Z^* é não-viciado para Z_0 , se e só se,

$$\sum_{j=1}^n \lambda_j m(x_j) = m(x_0). \quad (6)$$

Dois casos são distinguidos de acordo com os valores esperados $m(x_0)$ e $m(x_j)$:

- (i) os valores esperados são conhecidos ou desconhecidos;
- (ii) a função média do processo estocástico é estacionária ou não.

4.2. Krigagem Simples

Todas as esperanças são conhecidas:

Com o conhecimento de todos os valores esperados, a condição de não-viciabilidade do preditor Z^* fica caracterizada só pela expressão (6) e no mesmo espaço de projeção que foi usado pelo procedimento Krigagem Linear E_n . O procedimento Krigagem nessas condições é chamado de *Krigagem Simples* (KS).

4.3. Krigagem Ordinária

A esperança é estacionária desconhecida:

Nas condições onde $m(x) = m$ desconhecida para todo $x \in G$, a condição de não-viciabilidade (6) é satisfeita se e só se:

$$\sum_{j=1}^n \lambda_j = 1. \quad (7)$$

Essa restrição aos pesos λ 's da combinação linear, que gera o espaço E_n , faz com que o espaço de projeção de Z_0 seja restrito à variedade linear $C_1 \subset E_n$, definida por a condição (7). Nesse caso o processo Krigagem é chamado de *Krigagem Ordinária* (KO).

4.4. Krigagem Universal

A esperança não é estacionária nem conhecida:

Nesta situação não é possível caracterizar a condição de não-viciabilidade (7). É preciso dar uma forma à função média $m(x)$, por exemplo, a forma de uma combinação linear de L funções conhecidas, isto é,

$$m(x) = \sum_{i=1}^L a_i f_i(x);$$

os L parâmetros a_i são desconhecidos. A condição de não-vício (6) torna-se:

$$\sum_{i=1}^L a_i \sum_{j=1}^n \lambda_j f_i(x_j) = \sum_{i=1}^L a_i f_i(x_0)$$

Essa relação é satisfeita se e só se,

$$\sum_{j=1}^n \lambda_j f_i(x_j) = f_i(x_0); \quad i=1 \text{ a } L.$$

Essas L condições restringem o espaço E_n a uma variedade linear C_L de dimensão $n-L$ e o preditor $Z^* \in C_L \subset C_1$. Esse procedimento de Krigagem nessas condições de não-vício é chamado de *Krigagem Universal* (KU).

A variância de predição para cada processo de Krigagem é menor, se o espaço de projeção é maior, isto é,

$$\sigma_E^2 \leq \sigma_{K0}^2 \leq \sigma_{K0}^2 \leq \sigma_{KS}^2 \leq \sigma_{KU}^2.$$

CAPÍTULO I

CONCEITOS ELEMENTARES PARA RESOLVER PROBLEMAS DE MINERAÇÃO

1. VARIÁVEIS MÉDIAS

Não são, em geral, de interesse as variáveis pontuais, mas sim as variáveis médias, médias das variáveis pontuais $Z(x)$. O teor médio de um bloco com volume v centrado no ponto x é

$$Z_v(x) = \frac{1}{v} \int_{v(x)} Z(u) du. \quad (1)$$

Na prática esse teor médio é aproximado por o valor médio dos valores realizados do teor pontual

$$Z_v(x) \approx \frac{1}{n} \sum_{j=1}^n z(x_j).$$

O processo estocástico de interesse é dado por

$$Z_v = \{Z_v(x); x \in G\}, \quad G \subset \mathbb{R}^2 \text{ ou } \mathbb{R}^3.$$

A função covariância do processo estocástico Z_v é dada por

$$C_{Z_v}(x-y) = \text{Cov}\{Z_v(x), Z_v(y)\} = \frac{1}{v^2} \int_{v(x)} \int_{v(y)} C(x'-y') dx' dy', \quad (2)$$

denotada por $\bar{C}(v,v)$ e representa o valor médio da covariância $C(h)$ quando um extremo do vetor distância h descreve o domínio $v(x)$ e o outro extremo descreve independentemente o domínio $v(y)$. A integral estocástica (1) existe se a integral de Riemann (2) existe.

A função média do processo estocástico Z_v é

$$m_{Z_v}(x) = E \left\{ \frac{1}{V} \int_{v(x)} z(u) du \right\} = m.$$

Se o processo estocástico original Z é estacionário ou de covariância estacionária então o processo estocástico Z_v também é estacionário ou covariância estacionária.

1.1. Variância de Extensão

É de interesse conhecer quanto é a diferença que existe entre o teor médio numa área $v(x)$ com volume v e o teor médio de uma área maior $V(y)$ com volume V , $v \ll V$. Essa medida para a diferença é a variância do erro $Z_v(y) - Z_v(x)$, chamada *variância de extensão de V por v* e é dada por

$$\begin{aligned} \sigma_E^2(V,v) &= \text{Var} \{ Z_v(y) - Z_v(x) \} \\ &= \text{Var} \{ Z_v(y) \} + \text{Var} \{ Z_v(x) \} - 2\text{Cov} \{ Z_v(y), Z_v(x) \} \\ &= \frac{1}{V^2} \int_{v(x)} \int_{v(y)} C(x' - y') dx' dy' + \frac{1}{V^2} \int_{v(x)} \int_{v(y)} C(x' - y') dx' dy' - \\ &\quad - \frac{2}{VV} \int_{v(x)} \int_{v(y)} C(x' - y') dx' dy'. \end{aligned}$$

As integrais são em duas ou três dimensões de acordo com as dimensões do domínio G . Para simplificar as notações, escreve-se $\sigma_E^2(V,v)$ por

$$\sigma_E^2(v,V) = \bar{C}(v,v) + \bar{C}(V,V) - 2\bar{C}(v,V). \quad (3)$$

Quando a covariância $C(h)$ existe o semi-variograma $\gamma(h)$ também existe, e por (0.1) a expressão (3) pode ser escrita em termos do variograma

e do semi-variograma onde o termo constante desaparece, sendo

$$\sigma_E^2(v,V) = 2\bar{\gamma}(v,V) - \bar{\gamma}(v,v) - \bar{\gamma}(V,V). \quad (4)$$

onde $\bar{\gamma}(v,V)$ representa o valor médio de $\gamma(h)$ quando um extremo do vetor distância h descreve o domínio $V(x)$ e o outro extremo independentemente descreve o domínio $v(y)$.

A relação (4) pode mostrar-se válida quando o processo estocástico Z é só intrínseco (a função covariância do processo, $C(h)$ não existe).

Observações:

(a) As relações (3) e (4) são independentes dos domínios v e V . Em particular, o domínio V pode ser formado por dois blocos diferentes $V=V_1+V_2$ e o domínio da unidade de seleção pode ser formado por diferentes amostras $v = \sum_{j=1}^n v_j$. Algumas podem estar localizadas dentro do domínio V .

(b) Nas relações (3) e (4) há quatro fatos essenciais e intuitivos que condicionam toda a predição.

(i) a geometria do domínio $V(x)$: como o semi-variograma $\gamma(h)$ é uma função crescente de h , $\bar{\gamma}(V,V)$ aumenta com o tamanho do volume V . Se o tamanho de V é fixo, o termo $\bar{\gamma}(V,V)$ é conseqüentemente a variância de extensão $\sigma_E^2(v,V)$ que depende também da geometria de V .

(ii) a distância entre o domínio V e do suporte v : na medida que a distância entre o domínio de V e de v aumenta, assim aumenta o valor de $\bar{\gamma}(v,V)$ e também o valor da variância de extensão.

(iii) a geometria do domínio v : como a dimensão do domínio v aumenta, o valor de $\bar{\gamma}(v,v)$ também aumenta e conseqüentemente a variância de extensão decresce.

(iv) a função estrutural, semi-variograma ou covariância: a qualidade da estimação depende das características estruturais (anisotropias, grau de regularidade) do fenômeno em estudo.

1.2. Variância de Dispersão

Nos problemas de seleção é muito importante ter uma medida de dispersão (ou de variabilidade) dos teores de produção unitários v (bloco) dentro de uma sub-área (painel).

Seja um domínio centrado no ponto x , $V(x) = \bigcup_i v(x_i)$ onde os domínios $v(x_i)$ são todos iguais por translação. Quando o volume v do domínio $v(x_i)$ é muito menor em relação ao volume V do domínio de dispersão $V(x)$, $v \ll V$, o desvio quadrático médio $S^2(x)$ é uma integral sob o campo de dispersão V

$$S^2(x) = \frac{1}{V} \int_{v(x)} [Z_v(x) - Z_v(u)]^2 du.$$

Sob a hipótese de estacionariedade de covariância estacionaria ou intrínseca, a *variância de dispersão das unidades v dentro de V* é definida como a esperança estacionária de $S^2(x)$

$$D^2(v/V) = E\{S^2(x)\} \quad v \ll V.$$

Calculando, tem-se

$$\begin{aligned} E\{S^2(x)\} &= E\left\{\frac{1}{V} \int_{v(x)} [Z_v(x) - Z_v(u)]^2 du\right\} \\ &= \frac{1}{V} \int_{v(x)} \sigma_E^2(v(y), V(x)) dy, \quad v \ll V. \end{aligned} \quad (4)$$

De acordo com (3), tem-se

$$\sigma_E^2(v(y), V(x)) = \bar{C}(v(y), v(y)) + \bar{C}(V(x), V(x)) - 2\bar{C}(v(y), V(x)),$$

e como o processo é de covariância estacionária, $\bar{C}(v(y), v(y))$ e $\bar{C}(V(x), V(x))$ não dependem das posições y e x dos domínios $v(y)$ e $V(x)$ respectivamente, mas das geometrias dos domínios, isto é,

$$\bar{C}(v(y), v(y)) = \bar{C}(v, v), \quad \forall y \in G$$

e

$$\bar{C}(V(x), V(x)) = \bar{C}(V, V), \quad \forall x \in G.$$

Então esses termos são invariantes no campo $V(x)$ ficando

$$\frac{1}{V} \int_{v(x)} \{ \bar{C}(v(y),v(y)) + \bar{C}(V(x),V(x)) \} dy = \bar{C}(v,v) + \bar{C}(V,V).$$

e a condição de ser v muito menor que V , leva a que o deslocamento de $v(y)$ em $V(x)$ é equivalente ao deslocamento de um ponto y no domínio $V(x)$, assim

$$\frac{1}{V} \int_{v(x)} \bar{C}(v(y),V(x)) dy = \bar{C}(V(x),V(x)) = \bar{C}(V,V), \quad v \ll V.$$

A condição $v \ll V$ esta significando deslocar um ponto dentro de V . Quando o domínio $V(x)$ está constituido por um número finito de $v(y)$ a condição $v \ll V$ não é precisada porque o deslocamento de $v(y)$ dentro de V significa deslocar um ponto dentro de V .

A variância de dispersão (4) fica

$$D^2(v/V) = \bar{C}(v,v) - \bar{C}(V,V), \quad (5)$$

o que pode ser escrito em termos do semi-variograma, substituindo $C(h)$ por $-\gamma(h)$

$$D^2(v/V) = \bar{\gamma}(V,V) - \bar{\gamma}(v,v). \quad (6)$$

Pode-se mostrar que esta relação (em termos do semi-variograma) é válida se o processo estocástico Z satisfaz só a hipótese intrínseca.

Observações:

(a) A variável aleatória $S^2(x)$ depende da posição particular x do ponto central do domínio $V(x)$ e, sob a hipóteses de covariância estacionária ou intrínseca, $D^2(v/V)$ não depende da posição particular x , mas da geometria dos domínios $v(y)$ e $V(x)$ e da função estrutural, covariância $C(h)$ ou semi-variograma $\gamma(h)$.

(b) Para $v = 0$, isto é, a unidade $v(y)$ é um ponto. Nesse caso $v \ll V$ ficando a variância de dispersão de uma variável pontual dentro de V como

$$D^2(0/V) = \bar{\gamma}(V,V) - \bar{\gamma}(0,0).$$

(c) A relação de Krige: A propriedade de aditividade da variância de dispersão é estabelecida como uma consequência da relação (6)

$$\begin{aligned}
D^2(v/G) &= \bar{\gamma}(G,G) - \bar{\gamma}(v,v) \\
&= \bar{\gamma}(V,V) - \bar{\gamma}(v,v) + \bar{\gamma}(G,G) - \bar{\gamma}(V,V) \\
&= D^2(v/V) + D^2(V/G), \quad v \subset V \subset G. \tag{7}
\end{aligned}$$

Isto é, a dispersão da unidade v no depósito G é igual à soma das dispersões da unidade v no painel V e a dispersão do painel V no depósito G.

Essa relação foi encontrada experimentalmente por D. G. Krige usando dados de um depósito de ouro.

(d) A variância a priori: Sob a hipótese de covariância estacionária e supondo que

$$\lim_{V \rightarrow \infty} \int_{v(x)} \int_{v(x)} C(u-w) du dw < \infty,$$

o teor médio $Z_v(x)$ num campo de dispersão muito grande V, ($V \rightarrow \infty$), é igual à esperança estacionária do processo estocástico Z, isto é,

$$Z_v(x) \rightarrow E\{ Z(x) \} = m \quad \text{quando } V \rightarrow \infty,$$

ou seja, que Z é um processo estocástico ergódico.

Tendo-se também

$$\begin{aligned}
D^2(0/\infty) &= \lim_{V \rightarrow \infty} D^2(0/V) \\
&= \lim_{V \rightarrow \infty} \frac{1}{V} \int_{v(x)} E\{ [Z(y) - m]^2 \} dy \\
&= \text{Var}\{ Z(y) \} \\
&= C(0).
\end{aligned}$$

Se o processo estocástico Z é ergódico de covariância estacionária, a variância a priori é o limite da variância de dispersão da variável pontual $Z(y)$ num campo muito grande de variabilidade.

Sob as mesmas hipóteses e das expressões (5) e (6)

$$\begin{aligned} D^2(0/\infty) &= C(0) - \bar{C}(\infty, \infty) = C(0) \\ &= \bar{\gamma}(\infty, \infty) - \gamma(0) = C(0), \end{aligned}$$

onde $\bar{C}(\infty, \infty) = C(\infty) = 0,$

$$\bar{\gamma}(\infty, \infty) = \gamma(\infty) = C(0),$$

e $\gamma(0) = 0.$

Cálculo do valor médio do variograma.

As principais operações em Geoestatística (estimação da variância de extensão e de dispersão) precisam do uso do valor médio $\bar{\gamma}$ do semi-variograma pontual $\gamma(h)$ que é escrito por

$$\bar{\gamma}(v, w) = \frac{1}{vw} \int_{v(x)} \int_{w(x)} \gamma(x - y) dx dy.$$

Essa expressão representa o valor médio do semi-variograma pontual $\gamma(h)$ quando um extremo do vetor distância h descreve o domínio v e o outro extremo independentemente descreve o domínio w . Ver Fig. 1.

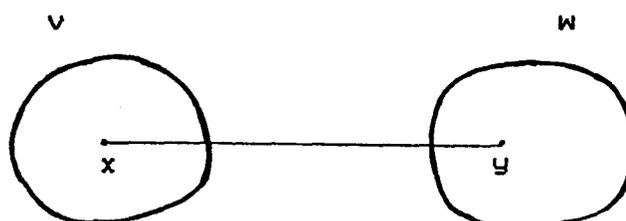


Fig. 1. x percorre v e y percorre w .

Existem duas formas de se obter $\bar{\gamma}(v,w)$:

1._ A primeira é calcular $\bar{\gamma}$ numericamente. Essa solução não faz restrições nos domínios v e w , só tem que se discretizar por uma malha regular de pontos centrados $(x_i; i=1$ a $N)$ para v e $(y_j; j=1$ a $M)$ para o domínio w e o valor médio $\bar{\gamma}(v,w)$ é aproximado por uma dupla soma discreta

$$\bar{\gamma}(v,w) \simeq \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M \gamma(x_i - y_j).$$

Observações:

(i) Outros métodos para a aproximação de $\bar{\gamma}(v,w)$ mais precisos podem ser usados, de tal forma que não levem muito tempo nos cálculos.

(ii) O efeito zero: Essa aproximação é perigosa ao considerar um dos domínios como subconjunto do outro, por exemplo $w=v$, nesse caso podendo ter pontos coincidentes nas discretizações $(x_i; i=1$ a $N)$ e $(y_j; j=1$ a $M)$ dando maior importância ao valor $\gamma(0) = 0$. Esse efeito zero faz que o valor médio do semi-variograma seja subestimado sistematicamente e, conseqüentemente, que o valor médio \bar{C} da correspondente covariância seja sobrestimado sistematicamente por $C(h) = C(0) - \gamma(h)$. O efeito zero pode ser evitado usando um modelo, que movimente levemente cada ponto de uma das malhas, independentemente de cada ponto da outra malha não movimentada, ou usando outro método mais sofisticado como o método de Gauss-Cauchy.

(iii) Densidade da aproximação discreta. A principal condição para a aproximação discreta é a densidade, que deve satisfazer todas as relações de ordem do variograma e variância de dispersão, ou seja se $v > w$, então $\bar{\gamma}(v,v) \geq \bar{\gamma}(w,w)$ e os estimadores devem também satisfazer $\bar{\gamma}^*(v,v) \geq \bar{\gamma}^*(w,w)$ e $D^{2*}(v/w) < 0$. Além disso, densidade da aproximação discreta pode ser achada iterativamente por aumentos progressivos nos números N e M até algum aumento que produza melhoria não significativa.

2._ A segunda forma para conseguir o valor médio $\bar{\gamma}(v,w)$, é que a geometria dos domínios v e w sejam simples. Foram construídas para isso quatro funções, chamadas *funções auxiliares básicas* e denotadas por α , X , F e H , definidas para

domínios de forma retangular em uma ou duas dimensões. Ao considerar modelos semi-variogramas $\gamma(h)$ isotrópicos simples (modelos linear, logaritmico, esférico, exponencial e potencia), obtém-se valor médio $\bar{\gamma}(v,w)$ para as diferentes dimensões dos domínios simples v e w . Esses valores são apresentados em forma de ábacos (cartas) (A. G. Journel, 1978).

2. DISPERSÕES VERDADEIRAS

O conjunto do mineral contido no depósito G , chamado de *recurso in situ*, é representado por histogramas de dispersões. Assim, o histograma para teores $z(x)$ de suporte pontual, supondo que seja representativo da dispersão dos teores no depósito G , tem as seguintes características:

(i) uma média experimental $m^*(G;0)$, que é supostamente igual à média verdadeira do depósito $m = E\{ Z(x) \}$.

(ii) uma variância de dispersão, $D^2(0/G) = E\{ [Z(x) - m]^2 \}$

(iii) e uma forma, por exemplo assimétrica que pode ser ajustada geralmente por uma distribuição lognormal ou gamma.

Supondo que os teores verdadeiros $z_v(x)$ de todas as unidades mineiras de suporte v sejam conhecidos, o histograma de dispersão dos teores $z_v(x)$ então será construído com as características, a seguir:

(i) uma média igual à média dos teores pontuais, $m = E\{ Z_v(x) \}$

(ii) uma variância de dispersão, $D^2(v/G)$, onde geralmente o suporte v é maior que o suporte pontual. Note que $D^2(v/G) < D^2(0/G)$, pela relação (7).

(iii) uma forma, por exemplo, assimétrica.

Essas características são mostradas na Fig. 2.

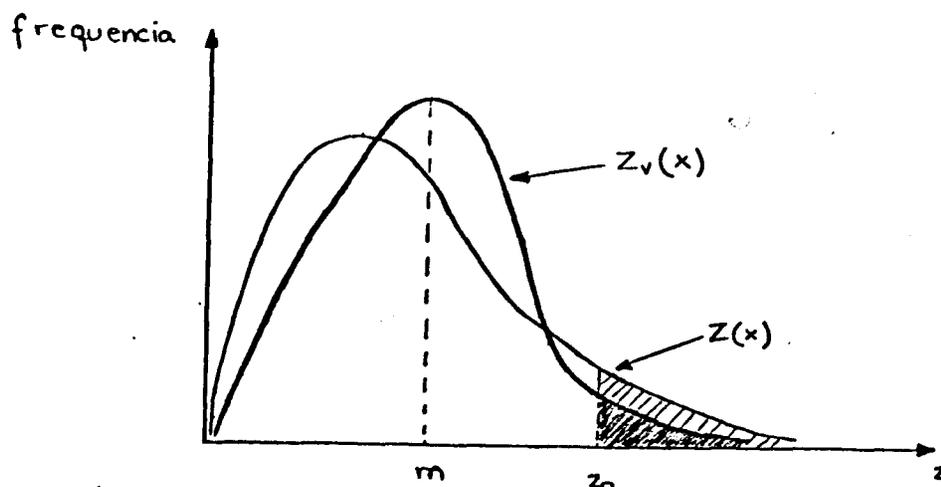


Fig. 2. Densidade do teor pontual e do bloco-v.

Assim, os dois histogramas são diferentes, levando a proporção de teores pontuais $z(x)$, maiores ou iguais a um valor de corte z_0 (região riscada, Fig. 2), a ser diferente da proporção de teores médios $Z_v(x)$, maiores ou iguais ao valor de corte z_0 (região preta). Se existir uma proporção não desprezível de testemunhas de teor $Z(x) \geq z_0\%$, não haverá nenhum ou poucos painéis de várias toneladas de teor médio $Z_v(x) \geq z_0\%$. Para altos teores de corte $z_0 > m$, a área riscada do histograma de teores testemunhas $z(x)$ pode seriamente sobrestimar o possível mineral recuperado no teor médio e na tonelagem média e correspondentemente subestimar a tonelagem do metal estéril, isto é, subestimar a área correspondente aos painéis com teores $z_v(x) < z_0$. Vê-se que a noção de suporte da variável é essencial.

Até aqui é possível conhecer as estimativas dos dois parâmetros que caracterizam uma função de distribuição dos teores $Z_v(x)$ num suporte v do depósito G como segue.

Conhecida ou estimada a distribuição do teor pontual $Z(x)$ no depósito G , $\phi^*(G, z)$, a esperança estimada de $Z_v(x)$ é dada por

$$m^* = m^*(v/G) = m^*(0/G) = \int_0^{\infty} u d\phi^*(G; u)$$

Usando a relação de D. G. Krige, expressão (8), encontrada experimentalmente por ele e formalizada por G. Matheron temos:

$$\begin{aligned} D^2(v/G) &= \bar{\gamma}(G,G) - \bar{\gamma}(v,v) \\ &= D^2(O/G) + \bar{\gamma}(O,O) - \bar{\gamma}(v,v) \\ &= D^2(O/G) - \bar{\gamma}(v,v). \end{aligned}$$

A variância de dispersão estimada é dada por

$$D^{2*}(v/G) = D^{2*}(O/G) - \bar{\gamma}(v,v),$$

onde

$$D^{2*}(O/G) = \int_0^{\infty} [u - m^*]^2 d\phi^*(G;u)$$

e $\bar{\gamma}(v,v)$ é obtido por um dos procedimentos expostos no cálculo do valor médio do semi-variograma.

Mas para conhecer a distribuição do bloco-v não é suficiente determinar só os parâmetros da média e da variância, também é necessário conhecer a forma dessa distribuição.

2.1. Princípio de Conservação da Distribuição

Diversos trabalhos usam o *princípio da conservação da distribuição*, que consiste em ajustar o histograma experimental das amostras por uma distribuição teórica (lognormal, gamma, etc.). Logo, postula o princípio de conservação da distribuição dos teores blocos-v, e que leva a distribuição dos teores bloco-v terem mesma forma da distribuição dos teores pontuais. Em particular, o postulado de conservação lognormalidade é muito frequente. E fazendo uma simples correção para a dispersão, é possível derivar a distribuição dos blocos-v qualquer seja sua forma, tamanho e localização.

Alguns dos problemas que surgem sobre esse enfoque são:

- não existe nenhuma justificativa teórica do postulado de conservação da

distribuição, ao contrário, numerosos exemplos experimentais o rejeitam.

— as amostras nunca são exatamente distribuídas segundo uma lognormal. O modelo lognormal impõe um comportamento particular nas caudas da distribuição que não é ajustado, em geral, ao comportamento experimental.

○ Esse postulado pode ser usado na prática, se o fator de correção de variância relativa for pequeno, isto é,

$$\frac{|D^2(O/G) - D^2(v/G)|}{D^2(O/G)} < 30\%.$$

Satisfeita essa restrição no fator de correção de variância relativa, a distribuição bloco-v é dada por

$$\phi_v(G; z) = \phi(G; y),$$

onde

$$y = m + \frac{D(O/G)}{D(v/G)} (z - m), \quad \forall z \in G.$$

3. A ANÁLISE ESTRUTURAL OU VARIOGRAFIA

A análise estrutural de um fenômeno distribuído espacialmente consiste em construir um modelo do variograma que caracteriza de uma forma operacional as principais características da dependência espacial do fenômeno. Essa modelagem exige um bom conhecimento físico do fenômeno e dos objetivos a serem estudados, como também possuir uma prática no ajuste de modelos geoestatísticos.

O estudo geoestatístico compreende dois passos principais:

1. Uma análise espacial que resulta na decisão essencial de estacionariedade (homogeneidade na área de estudo) e num modelo para a dependência espacial sob a área.

2._ A implantação de um dos métodos a definir no próximo capítulo com o nome de Krigagem Indicadora ou Krigagem Probabilística, para determinar completamente a função distribuição estimada.

A *análise estrutural* é importante e compreende duas fases: uma é a *análise exploratória dos dados*, onde se consegue a familiaridade com os dados e o problema, que não é considerada nesta dissertação; a segunda, é conseguir modelar o variograma. Os resultados da análise estrutural, em particular a decisão de estacionariedade, e, como consequência, os modelos do variograma, condicionam completamente os resultados do próximo passo da Krigagem e a estimação da função distribuição.

A modelagem do variograma é feita com base no *variograma experimental*

$$2\hat{\gamma}(h) = \sum_{N(h)} \frac{[z(x_i) - z(x_j)]^2}{|N(h)|}, \quad h \in \mathbb{R}^2 \text{ ou } \mathbb{R}^3. \quad (8)$$

Onde $N(h) = \{ (x_i, x_j); |x_i - x_j| = h \}$ e $|N(h)|$ é o número de elementos distintos em $N(h)$.

Plotados os pontos $\hat{\gamma}(h)$ para cada h , agora tem que se ajustar o melhor modelo semi-variograma teórico $\gamma(h)$ ao semi-variograma experimental $\hat{\gamma}(h)$.

Para dados espacialmente irregulares, $N(h)$ é usualmente modificado para o conjunto $\{ (x_i, x_j); |x_i - x_j| = T(h) \}$, onde $T(h)$ é uma região de tolerância de G referenciada por h (A. G. Journel, 1988).

Outros estimadores para o variograma são dados por Cressie e Hawkins (1980) e uma comparação entre diferentes estimadores do variograma, incluindo o do método dos momentos (8), é dada por D. L. Zimmerman e M. B. Zimmerman (1991).

3.1. Modelos dos Variogramas

Não nos ocuparemos nesta dissertação do primeiro passo do estudo geoestatístico, estaremos restringidos a expor somente os modelos mais comuns dos variogramas, deixando de lado a análise exploratória dos dados, para

isso ver A. G. Journel (1978).

Condição positiva semi-definida.

Já foi dito anteriormente que a covariância $C(h)$ de um processo estocástico de covariância estacionária deve ser uma função positiva semi-definida. Se o processo estocástico é intrínseco com o semi-variograma $\gamma(h)$, então $\gamma(h)$ deve ser uma função negativa semi-definida condicional. Condicional porque a soma dos coeficientes da combinação linear dos elementos do processo deve ser igual a zero.

Como a propriedade positiva semi-definida de uma função $g(h)$ está relacionada às dimensões do espaço onde está definido o vetor distância h , então, uma função que é positiva semi-definida num espaço unidimensional, pode não ser positiva semi-definida em um outro espaço n -dimensional com $n > 1$.

Mas, toda função que é positiva semi-definida em um espaço n -dimensional também é positiva semi-definida em um outro espaço de dimensão menor (M. Armstrong e P. Diamond, 1984).

Essas propriedades também são obtidas para funções negativas semi-definidas condicionais, e todos os modelos teóricos expostos nessa seção, são positivos semi-definidos ou negativos semi-definidos condicionais em três dimensões; portanto, podem ser usados em uma ou duas dimensões.

A geoestatística, por meio de duas propriedades das funções positivas semi-definidas, tem preferido funções que satisfaçam as propriedades abaixo descritas para construir uma família suficientemente grande de modelos de covariâncias e variogramas.

Propriedades:

1. Qualquer combinação linear de covariâncias ou semi-variogramas é uma covariância ou semi-variograma.

$$C(h) = \sum_{j=1}^K \lambda_j^2 C_j(h), \text{ é uma covariância.}$$

$$\gamma(h) = \sum_{j=1}^K \lambda_j^2 \gamma_j(h), \text{ é um semi-varograma.}$$

2._ Qualquer produto de covariâncias e uma função covariância

$$C(h) = \prod_{j=1}^K \lambda_j^2 C_j(h), \text{ é uma covariância.}$$

Como qualquer covariância é associada necessariamente a um semi-variograma, $\gamma(h) = C(0) - C(h)$, tem-se que

$$\gamma(h) = \prod_{j=1}^K \lambda_j^2 C_j(0) - \prod_{j=1}^K \lambda_j^2 C_j(h).$$

esta expressão é uma função de semi-variograma aceitável.

Modelos Teóricos Básicos

Seja $\gamma(h)$ o semi-variograma do processo estocástico Z . As duas características principais de um semi-variograma estacionário são:

- 1._ O comportamento na origem: parabólico, linear e efeito pepita.
- 2._ A presença ou ausência de um patamar, isto é, $\gamma(h) = \text{constante}$, $|h| > a$.

Comportamento na origem.

Para explicar mais claramente o comportamento do variograma na origem, é preciso ter presente dois conceitos:

(i) Um processo estocástico Z é dito contínuo no ponto $x \in G$, em média quadrática, se

$$\lim_{h \rightarrow 0} E\{ [Z(x+h) - Z(x)]^2 \} = 0. \quad (9)$$

Supondo que o processo é intrínseco ou de covariância estacionária, então a expressão (9) é equivalente a:

$$\lim_{h \rightarrow 0} \gamma(h) = 0 = \gamma(0), \quad \forall x \in G.$$

É assim que a continuidade (em média quadrática) espacial de um processo estocástico Z está relacionada à continuidade (em \mathbb{R}) da função variograma no ponto zero.

(ii) Um processo estocástico Z é dito diferenciável no ponto $x \in G$ em média quadrática, se existe um processo estocástico W , de segunda ordem, tal que

$$\lim_{h \rightarrow 0} E \left\{ \left| \frac{Z(x+h) - Z(x)}{h} - W(x) \right|^2 \right\} = 0. \quad (10)$$

Nesse caso $W(x)$ é denotado por $Z'(x)$.

Supondo que o processo seja intrínseco ou estacionário, a expressão (10) pode ser escrita como:

$$\lim_{h \rightarrow 0} \frac{\gamma(h)}{|h|^2} = E \{ |W(x)|^2 \}.$$

Esse conceito leva a relacionar o variograma numa vizinhança da origem $h=0$ com a regularidade (diferenciabilidade) espacial do processo estocástico Z , que modela o fenômeno.

Os conceitos de continuidade e regularidade nos dão o conhecimento do comportamento do fenômeno. Agora, se o fenômeno é razoavelmente homogêneo, de tal forma que se possa considerá-lo intrínseco ou de covariância estacionária então o estudo do comportamento do fenômeno se reduz ao estudo do variograma numa vizinhança na origem. Em ordem decrescente de regularidade, quatro tipos principais de comportamento na origem podem ser distinguidos :

1. Comportamento parabólico.

O comportamento parabólico do variograma é do tipo $\gamma(h) \sim A|h|^2$ quando $h \rightarrow 0$ (\sim quer dizer que para $h \rightarrow 0$, $\gamma(h)$ pode não ter forma parabólica). Ou seja que $\gamma(h)$ é duas vezes diferenciável na origem e o processo também é diferenciável. Esse tipo de comportamento é característico de uma variabilidade espacial muito regular e o processo também é contínuo.

2._ Comportamento linear.

O comportamento linear é do tipo $\gamma(h) \sim A|h|$ quando $h \rightarrow 0$, assim que $\gamma(h)$ não é diferenciável na origem, mas é contínuo em $h=0$. Assim, o processo é contínuo mas não é diferenciável.

3._ Comportamento de descontinuidade na origem ou efeito pepita.

Esse tipo de comportamento é descrito pelo fato de que $\gamma(h)$ não converge a zero quando se faz tender h ao zero, embora por definição $\gamma(0) = 0$, isto é,

$$\lim_{h \rightarrow 0} \gamma(h) = C_0.$$

A descontinuidade de um variograma na origem é chamado de *efeito pepita*.

Assim o processo descrito com um efeito pepita não é contínuo.

4._ Comportamento de efeito pepita puro.

Esse comportamento do variograma na origem é o caso limite, quando $\gamma(h)$ aparece só com descontinuidade na origem e constante fora da origem, isto é,

$$\gamma(h) = \begin{cases} 0 & h = 0 \\ C_0 & h > \epsilon \end{cases}$$

onde ϵ é muito pequeno em comparação à distância de observação, para qualquer distância de observação tão pequena quanto possível. Esse comportamento quer dizer que $Z(x)$ e $Z(x+h)$ são incorrelacionados.

Estos quatro tipos de comportamento podem ser mostrados na Fig. 3.

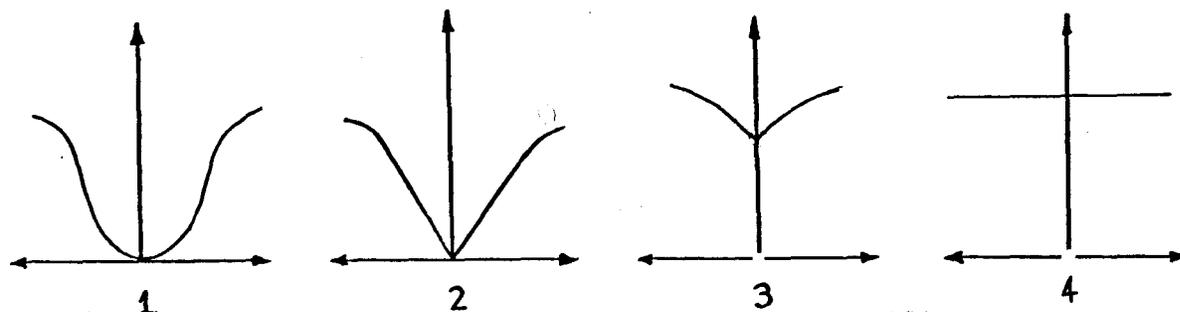


Fig. 3. Os quatro tipos principais de comportamento do variograma na origem.

Assim, os modelos mais usados podem ser classificados como:

1._ Modelos com patamar (ou modelos de transição)

(i) e um comportamento linear na origem

- modelo esférico
- modelo exponencial

(ii) e um comportamento parabólico na origem

- modelo gaussiano.

2._ Modelos sem patamar (nesse caso o processo é só intrínseco).

- modelo potência
- modelo logarítmico.

3._ Modelos com Efeito Pepita. Esses modelos apresentam uma aparente discontinuidade na origem do semivariograma, isto é, uma constante pepita C_0 , que pode ser interpretada como uma estrutura de transição, alcançando seu patamar no valor de um alcance muito pequeno C_0 comparado com a distância avaliada de observação.

4._ Modelos com Efeito Buraco. Esses modelos são os que não têm um patamar, isto é,

$$\gamma(h) = \begin{cases} \gamma_1(h) & h \leq a \\ f(h) & h > a. \end{cases}$$

onde $f(h)$ geralmente é uma função que oscila em torno do valor alcance a , por exemplo uma função senoidal.

1. Modelos com patamar ou modelos de transição.

Os modelos apresentados são normados, isto é, esses modelos correspondem a processos estocásticos com variância a priori igual a um

$$\text{Var}\{Z(x)\} = C(0) = \gamma(\infty) = 1.$$

Para obter um modelo com um patamar $C(0)=C \neq 1$, é suficiente multiplicar a expressão $\gamma(h)$ ou $C(h)$ por um valor constante C .

(i) Comportamento linear na origem. Este é o mais frequente tipo de comportamento na prática da mineração (variograma dos teores) acompanhado de um efeito pepita.

- Modelo esférico

$$\gamma(h) = \begin{cases} \frac{3}{2} \frac{|h|}{a} - \frac{1}{2} \frac{|h|^3}{a^3} & \forall |h| \in [0, a] \\ 1 = \text{patamar} & \forall |h| \geq a. \end{cases}$$

- Modelo exponencial

$$\gamma(h) = 1 - \exp\left(-\frac{|h|}{a}\right) \quad \forall |h| \geq 0.$$

Esse modelo, diferentemente do modelo esférico, não alcança seu patamar em uma distância finita a ; o modelo exponencial alcança seu patamar só assintoticamente, embora as flutuações experimentais do variograma não façam diferença na prática entre um patamar alcançado assintoticamente ou não. Para o modelo exponencial o alcance prático a' pode ser:

$$a' = 3a \text{ e } \gamma(a') = 1 - e^{-3} = 0.95 \approx 1.$$

A diferença entre o modelo esférico e o exponencial é a distância na qual a tangente na origem intercepta o patamar

$$|h| = \frac{2a}{3} \text{ para o modelo esférico,}$$

$$|h| = a = \frac{a'}{3} \text{ para o modelo exponencial.}$$

(ii) Comportamento parabólico na origem.

O modelo Gaussiano tem um comportamento parabólico na origem. Esse modelo caracteriza um fenômeno com escala de continuidade espacial extremamente curta. Esse modelo não é usado para modelar variogramas de funções espaciais com descontinuidade tal como a variável indicadora $I(x; z)$, definida em (0.1).

2._ Modelos sem patamar.

Os fenômenos com modelos sem patamar correspondem a processos com uma capacidade ilimitada de dispersão espacial; não podendo ser definida a variância a priori nem a covariância. Os processos são só intrínsecos.

(i) Modelo potência.

$$\gamma(h) = |h|^\theta \quad |h| \geq 0 \text{ e } \theta \in (0,2).$$

Na prática, só o modelo linear é frequentemente usado

$$\gamma(h) = w|h|, \quad w \text{ a inclinação na origem.}$$

3._ Modelos com Efeito Buraco.

Um modelo variograma tem um efeito buraco quando o crescimento não é monótono. O efeito buraco pode aparecer em modelos com ou sem patamar.

O modelo com efeito buraco usual é dado por:

$$C(h) = \frac{\text{sen}|h|}{|h|}$$

$$\gamma(h) = 1 - \frac{\text{sen}|h|}{|h|}$$

A amplitude relativa do efeito buraco é definida pelo mínimo valor da covariância

dividida por seu valor patamar $C(0)$

$$\alpha = \frac{|\min C(|h|)|}{C(0)}$$

5. Modelos de Anisotropia.

Freqüentemente o vetor distância h não é definido em um espaço unidimensional, em geral é definido em um espaço de duas dimensões (\mathbb{R}^2), onde o fenômeno em uma direção pode ter uma variabilidade espacial diferente de outra direção.

(i) Isotropia.

Um fenômeno é dito isotrópico se sua função variograma depende só do valor módulo $|h|$ do vetor distância h , ou seja, que para qualquer direção das distâncias observadas a variabilidade espacial do fenômeno é a mesma.

(ii) Anisotropia.

Um fenômeno é dito anisotrópico se sua variabilidade espacial não é a mesma para todas as direções, ou seja, que a função estrutural do fenômeno que caracteriza a variabilidade espacial (covariância ou variograma) não só depende do módulo do vetor distância, mas também da sua direção.

Na prática, uma anisotropia real corresponde à existência de direções preferenciais na formação do fenômeno estudado. Essas direções preferenciais são as vezes conhecidas a priori, onde suas variabilidades serão quantificadas pelo variograma.

(iii) Anisotropias e estruturas imbricadas

Sabe-se que a variabilidade espacial entre as variáveis $Z(x+h)$ e $Z(x)$, caracterizada pelo variograma $2\gamma(h) = E\{ [z(x+h) - Z(x)]^2 \}$, deve-se a muitas causas que aparecem sob um domínio de escalas diferentes, ou seja, $Z(x)$ pode ser da forma

$$Z(x) = \mu(x) + W(x) + \eta(x) + \epsilon(x), \quad (11)$$

supondo que

(i) $\mu(x)$ é o processo médio determinístico, chamado de **variação a grande escala**,

(ii) $W(x)$ é um processo intrínseco contínuo em média quadrática, com média zero e o alcance de seu variograma (se existe) maior do que $\min\{\|x_i - x_j\|; 1 \leq i < j \leq n\}$. Chama-se esse processo de **variação a pequena escala**,

(iii) $\eta(x)$ é um processo intrínseco com média zero, independente de W , com alcance do variograma menor que $\min\{\|x_i - x_j\|; 1 \leq i < j \leq n\}$. Chama-se esse processo de **variação de micro escala**,

(iv) e $\epsilon(x)$ é um processo ruído com média zero, independente de η e W , ϵ é chamado de **erro de medida** ou simplesmente ruído.

Supondo $\mu(x) = m$ constante, o processo Z definido por (11) é intrínseco, podendo ser uma combinação de processos isotrópicos e anisotrópicos, ou seja, o processo variação de micro escala pode ser isotrópico e o processo de variação de pequena escala pode ser anisotrópico.

N. Cressie (1988) discutiu que se $Z(x)$ tem um componente de erro de medida, então uma versão regularizadora é predita, considerando um modelo de estruturas imbricadas. E concluiu que as interpretações nos parâmetros do variograma, comuns na literatura, vão ser modificadas.

Todos os modelos para o variograma dados anteriormente, são modelos isotrópicos; dependem só do módulo $|h|$ do vetor distância. Esses modelos são aproveitados para modelar as anisotropias, fazendo uma correção na sua isotropia.

São duas as apresentações de anisotropias que podem ser corrigidas:

(a) **Anisotropia Geométrica.**

Se os variogramas experimentais em duas ou mais direções têm uma mesma forma (o mesmo modelo) e patamares iguais, fala-se em **anisotropia geométrica**. Fazendo-se uma transformação linear nas coordenadas do vetor

distância h é possível levar um variograma anisotrópico a um isotrópico, isto é,

$$\gamma(h) \text{ anisotrópico} = \gamma(h') \text{ isotrópico,}$$

onde $h' = Ah$ e A é uma matriz de produtos de rotações e uma transformação de afinidade.

(b) _ Anisotropia Zonal.

O modelo de anisotropia zonal é outra opção para o uso na prática. Existem variogramas experimentais onde sua anisotropia não é possível corrigir por uma mera transformação linear nas coordenadas, em particular, esses variogramas experimentais que apresentam valores patamares diferentes, em direções diferentes e onde sua forma pode ser até diferente para cada direção.

Um componente de variabilidade especificada em uma direção α , pode ser modelada por um variograma $\gamma(h_\alpha)$, dependendo só da interdistância h_α , medida ao longo da direção α , isto é, em \mathbb{R}^2 o vetor $h = (h_x, h_y)$

$$\gamma(h_x, h_y) = \gamma(h_\alpha), \text{ independente de } h_{\alpha+90^\circ}$$

onde $(h_\alpha, h_{\alpha+90^\circ})$ são dados por uma rotação do vetor h em um ângulo α .

Um modelo de anisotropia que depende só de um número reduzido de dimensões do espaço original \mathbb{R}^2 ou \mathbb{R}^3 , é dito ter uma anisotropia zonal.

As correções dessas duas anisotropias são dadas por A.G. Journel (1978).

CAPÍTULO II

KRIGAGEM INDICADORA E PROBABILÍSTICA

Em fenômenos razoavelmente bem comportados, com coeficientes de variação não maior do que um (por exemplo, depósitos do tipo porfíricos e sedimentados), são bem sucedidos os métodos mostrados anteriormente (Krigagem linear), que proporcionam a quantificação do risco e estimativas locais das características médias tais como os teores blocos.

Mas as dificuldades começam com fenômenos altamente variantes que contêm uma proporção de dados com valores extremos muito pequenos apresentando uma distribuição de caudas compridas com um coeficiente de variação no domínio de 2 até 5 (por exemplo, depósitos de urânio, ouro, diamantes, etc.). Isto leva o variograma a ser extremamente sensível, fazendo-o distorcido, difícil de modelar e sem utilidade alguma (B. M. Davis, 1984).

As saídas para essa dificuldade são:

1. Usar aquilo que a indústria mineira está usando hoje, tratar os valores extremos sobre alguma superfície geológica (tipo de rocha, grau de alteração) ou probabilístico (os dados são reritos a uma distribuição de normalidade). Esse tratamento pode tender a serios erros de estimação das reservas (usualmente subestima a quantidade total de metal), especialmente quando o tamanho da amostragem não é suficiente.

2._ Os métodos usados pelos geoestatísticos no passado, regularizando os dados por meio de transformações (por exemplo, raiz quadrada, logaritmo natural ou transformada score normal). Essas transformações, com a característica de não linearidade, trazem métodos não-lineares para predizer uma variável não-linear, tais como KD, MG e Krigagem Lognormal (A. G. Journel, 1980). Infelizmente todos esses métodos têm que se impor à suposição de normalidade da distribuição multivariada do processo estocástico Z .

A idéia da Krigagem Indicadora (KI) desenvolvido por A. G. Journel no ano de 1982, que foi inspirada pelos resultados obtidos por G. Matheron no ano de 1972, para o KD, é basicamente dividida em duas partes:

1._ Tendo um fenômeno altamente variável, é desejável registrar a valiosa informação dos dados, sem ter que ajustá-lo a alguma transformação regularizadora. É assim que Journel pensou transformar os dados originais em novos dados indicadores.

2._ É preciso que este método a apresentar entregue as respostas, com suas dificuldades (Armando Z. Remacre, 1989 e G. Matheron, 1982), ao problema da mineração de predizer as reservas recuperáveis, o que se traduz em predizer a distribuição espacial dos valores $z(x)$, em uma área bem definida. A forma de fazer isso é basicamente discretizar o domínio da variável teor $Z(x)$ em um número suficiente de classes e interpolar, separadamente para cada classe, as estimativas da distribuição espacial dos teores $Z(x)$. Nessa etapa ainda se está lidando com variáveis com suporte pontual, onde na verdade o interesse está nas variáveis com suporte de um bloco- v , $Z_v(x)$. Nessa etapa de passar de uma distribuição pontual para uma distribuição bloco- v é que nasce outra dificuldade, que outros métodos não lineares não apresentam por supor uma distribuição normal. Para isso é feita uma correção para o efeito do suporte na distribuição estimada pontual.

1. VARIÁVEL INDICADORA

Considere o processo estocástico $Z = \{Z(x); x \in G\}$ que modela o fenômeno metalogênico num depósito G com limites bem definidos, a mineralização

em G razoavelmente homogênea, de tal forma a considerar o processo estocástico, estacionário em G , e uma realização finita do processo Z , $\{z(x_j) = z_j; j=1 \text{ a } n\}$, que são os dados avaliados.

1.1. Transformada Indicadora

Define-se a *transformada indicadora*

$$I(Z(x);z) = I(x;z) = \begin{cases} 1 & \text{se } Z(x) \leq z \\ 0 & \text{se não,} \end{cases}$$

para cada $x \in G$ e z um valor de corte no domínio de variabilidade do teor $Z(x)$

Uma realização de uma variável aleatória indicadora $I(x_0; z)$, $i(x_0; z)$, no ponto x_0 , pode ser interpretada como a probabilidade da variável aleatória $Z(x_0)$ dado $Z(x_0) = z_0$, como segue-se

$$\begin{aligned} P(Z(x_0) \leq z / Z(x_0) = z_0) &= \frac{P(Z(x_0) \leq z , Z(x_0) = z_0)}{P(Z(x_0) = z_0)} \\ &= \begin{cases} 1 & \text{se } Z(x_0) = z_0 \leq z \\ 0 & \text{se } Z(x_0) = z_0 > z \end{cases} \\ &= i(x_0; z) \quad \square \end{aligned}$$

Fazendo essa transformação temos o processo estocástico estacionário

$$I_z = \{I(x;z); x \in G\}, \quad z \text{ variando no domínio de } Z(x).$$

Assim é possível definir as seguintes funções do processo I_z .

- função média.

$$\begin{aligned} E\{ I(x;z) \} &= 1 \cdot P(Z(x) \leq z) + 0 \cdot P(Z(x) > z) \\ &= P(Z(x) \leq z) \end{aligned}$$

$$= F(z), \quad \forall z \in G.$$

- função covariância cruzada não-centrada.

$$\begin{aligned} K_f(h; z, z') &= E\{ I(x+h; z) I(x; z') \} \\ &= P(Z(x+h) \leq z, Z(x) \leq z') \\ &= F(h; z, z') \end{aligned}$$

= função distribuição conjunta (f.d.c) de $Z(x)$ e $Z(x+h)$.

Note que

$$K_f(h; z, z') = K_f(-h; z', z) \neq K_f(-h; z, z'),$$

e

$$K_f(0; z, z') = F(\min(z, z')).$$

- função covariância cruzada centrada.

$$C_f(h; z, z') = K_f(h; z, z') - F(z)F(z').$$

Note que

$$\begin{aligned} C_f(h; z) &= C_f(h; z, z) = \text{Cov}\{ I(x+h; z) I(x; z) \} \\ &= K_f(h; z, z) - F^2(z) \\ &= K_f(h; z) - F^2(z) \end{aligned}$$

e

$$C_f(0; z) = \text{Var}\{ I(x; z) \} = F(z) [1 - F(z)].$$

$$\begin{aligned} 2\gamma_f(h; z, z') &= E\{ [I(x+h; z) - I(x; z)] [I(x+h; z') - I(x; z')] \} \\ &= 2K_f(0; z', z) - K_f(h; z, z') - K_f(h; z', z). \end{aligned}$$

Note que

$$\begin{aligned}\gamma_1(h; z, z') &= \gamma_1(h; z', z) \\ &= F(\min(z, z')) - F(z)F(z') - \frac{1}{2} \left\{ C_1(h; z, z') + C_1(h; z', z) \right\}.\end{aligned}$$

e

$$\begin{aligned}\gamma_1(h; z) &= \gamma_1(h; z, z) = \frac{1}{2} E \left\{ [I(x; z) - I(x+h; z)]^2 \right\} \geq 0 \\ &= C_1(0; z) - C_1(h; z) \\ &= F(z) - K_1(h; z).\end{aligned}$$

- função correlograma cruzado.

$$\rho_1(h; z, z') = \frac{C_1(h; z, z')}{\sqrt{C_1(0; z) C_1(0; z')}} \in [-1, 1].$$

Note que

$$\rho_1(h; z) = \rho_1(h; z, z') = \frac{C_1(h; z)}{C_1(0; z)} = 1 - \frac{\gamma_1(h; z)}{C_1(0; z)}.$$

A proporção exata dos teores $Z(y)$, menores que um valor de corte z , em uma sub-área $V(x)$ de G , centrada no ponto x com volume V , é dada por:

$$\Phi(V(x); z) = \Phi(V; z) = \frac{1}{V} \int_{V(x)} I(y; z) dy.$$

Essa função aleatória é não-viciada para a f.d. de $Z(y)$, ou seja

$$\begin{aligned}E\{ \Phi(V; z) \} &= E \left\{ \frac{1}{V} \int_{V(x)} I(y; z) dy \right\} \\ &= \frac{1}{V} E \left\{ \int_{V(x)} I(y; z) dy \right\} = \frac{1}{V} \int_{V(x)} E\{ I(y; z) \} dy \\ &= F(z) \quad \square\end{aligned}$$

A covariância não-centrada do processo estocástico $\Phi(V(x);z)$ é dada por

$$\begin{aligned} \text{Cov}\{ \Phi(V; z), \Phi(V; z') \} &\stackrel{\circ}{=} \frac{1}{V^2} \int_{V(x)} \int_{V(x)} K_1(u - v; z, z') du dv - F(z) F(z') \\ &= \frac{1}{V^2} \int_{V(x)} \int_{V(x)} C_1(u - v; z, z') du dv. \end{aligned}$$

E sua variância é

$$\text{Var}\{ \Phi(V; z) \} = \frac{1}{V^2} \int_{V(x)} \int_{V(x)} C_1(u - v; z, z') du dv \leq 1.$$

Ao considerar uma sub-área $V(x)$ com volume V muito grande ($V \rightarrow \infty$), com a $\text{Var}\{ \Phi(V; z) \}$ finita, é que se dá a condição suficiente para que $\Phi(V; z)$ convirja à esperança estacionária

$$E\{ I(x; z) \} = F(z).$$

Vendo $\Phi(V; z)$ como função de z , é possível interpretar uma realização de $\Phi(V; z)$, $\phi(V; z)$, como a f.d., $F(z)$, da variável aleatória $Z(y)$, para qualquer $y \in G$.

1.2. Funções Locais de Recuperação.

Determinada a função distribuição $\phi(V; z)$, é possível determinar as funções recuperáveis pontuais em V .

- Tonelagem recuperável.

$$1 - \phi(V; z) = \text{Proporção}\{ z(y) > z; y \in V(x) \}.$$

- Quantidade de metal recuperável.

$$a(V; z) = E\{ [1 - I(x; z)] Z(x) \} = \int_z^\infty u d\phi(V; u). \quad (1)$$

- Teor médio de mineral num corte z.

$$m(V;z) = \frac{q(V;z)}{1 - \phi(V;z)}$$

Observação:

Fazendo $z = 0$, o teor médio in-situ em V é escrito como

$$m(V;0) = q(V;0) = \int_0^{\infty} u d\phi(V;u).$$

Se a área $V(x)$ é reduzida a um ponto x

$$\phi(x;z) = i(x;z),$$

e se $z = 0$, o teor médio in-situ é o valor realizado da variável $Z(x)$

$$m(x;0) = z(x) = \int_0^{\infty} u di(x;u).$$

A expressão (1), é uma integral de Stieltjes com densidade $d\phi(V;z)$. Embora a função $\phi(V;z)$ não seja derivável em z , essa integral sempre é definida na prática, onde $z(x) < \infty \forall x \in G$.

Assim, a função a ser predita é a proporção exata $\phi(V;z)$, para poder obter estimativas das distribuições espaciais locais. Em termos da mineração, é estimar as reservas recuperáveis locais pontuais.

2. Krigagem Indicadora

Na presença da continuidade espacial entre os dados, $\{z_j; j=1, \dots, n\}$, considera-se que cada dado z_j pode ter diferente peso, dependendo da sua configuração espacial. E o preditor que será considerado a princípio é da forma:

$$\hat{\phi}(V;z) = \lambda_0 + \sum_{j=1}^n \lambda_j(z) I(x_j; z); \quad x_j \in G,$$

onde os pesos λ_j 's ficam caracterizados pelo procedimento de minimização, exposto

no capítulo introdutorio seção 4, só que agora o procedimento é desenvolvido usando as variáveis indicadoras. É necessário então achar os pesos λ_j 's em decorrência da minimização do erro quadrático médio

$$\| \Phi(V;z) - \hat{\Phi}(V;z) \|^2 = E\{ [\Phi(V;z) - \hat{\Phi}(V;z)]^2 \}, \quad (2)$$

sujeito a que o preditor $\hat{\Phi}(V;z)$ seja não-viciado para $\Phi(V;z)$, isto é,

$$E\{ \Phi(V;z) - \hat{\Phi}(V;z) \} = 0.$$

Em termos geométricos o preditor $\hat{\Phi}(V;z)$, que minimiza o erro quadrático médio denotado por $\Phi^*(V;z)$, é a projeção da função proporcional desconhecida $\Phi(V;z)$ no espaço gerado pelas combinações lineares

$$\lambda_0 + \sum_{j=1}^n \lambda_j(z) I(x_j; z).$$

Essa projeção é única, e é caracterizada pela ortogonalidade do vetor $[\Phi(V;z) - \Phi^*(V;z)]$ e cada dos $(n+1)$ elementos geradores do espaço das combinações lineares, e também é caracterizado pela condição de não-viciabilidade, isto é,

$$E\{ \Phi(V;z) - \hat{\Phi}(V;z) \} = 0$$

$$E\{ [\Phi(V;z) - \hat{\Phi}(V;z)] I(x_j; z) \} = 0; \quad j=1 \text{ a } n.$$

2.1. Krigagem Indicadora Simples

Considerando o processo I_z estacionário com média conhecida, expressa-se a caracterização em termos da média e a covariância não-centrada,

$$\lambda_0 + \sum_{j=1}^n \lambda_j(z) F(z) = F(z)$$

$$\lambda_0 F(z) + \sum_{j=1}^n \lambda_j(z) K_I(x_i - x_j; z) = \bar{K}_I(V, x_j; z); \quad j=1 \text{ a } n.$$

Multiplicando a primeira equação da caracterização por $F(z)$ e subtraindo a primeira equação da $(j+1)$ -ésima equação, obtém-se

$$\sum_{i=1}^n \lambda_i(z) \left(K_i(x_i - x_j; z) - F^2(z) \right) = \bar{C}_i(V, x_j; z) - F^2(z); \quad j=1 \text{ a } n.$$

ou em termos da covariância

$$\sum_{i=1}^n \lambda_i(z) C_i(x_i - x_j; z) = \bar{C}_i(V, x_j; z); \quad j=1 \text{ a } n. \quad (3)$$

Esse sistema é chamado de *sistema krigagem indicadora simples*, que é resolvido independentemente para cada valor de corte z .

Depois de resolvido o sistema (3) para obter os pesos λ_j 's, o peso λ_0 é dado pela primeira equação da caracterização de $\hat{\Phi}(V; z)$. Logo, o preditor chamado de *Krigagem Indicador (KI)* é dado por

$$\hat{\Phi}^*(V; z) = F(z) \left(1 - \sum_{j=1}^n \lambda_j(z) \right) + \sum_{j=1}^n \lambda_j(z) I(x_j; z).$$

E o preditor centrado $\hat{\Phi}^*(V; z) - F(z)$ aparece como uma combinação linear da variáveis indicadoras:

$$\hat{\Phi}^*(V; z) - F(z) = \sum_{j=1}^n \lambda_j(z) \left(I(x_j; z) - F(z) \right).$$

Em termos geométricos, $\hat{\Phi}^*(V; z) - F(z)$ é a projeção da variável desconhecida $\Phi(V; z) - F(z)$ no espaço vetorial gerado por as combinações das n variáveis $I(x_j; z) - F(z)$.

Variância de Predição.

A variância de predição $\text{Var}\left\{ \hat{\Phi}(V; z) - \hat{\Phi}^*(V; z) \right\}$ do preditor KI $\hat{\Phi}^*(V; z)$ é denotado por σ_{KI}^2 e escrito por:

$$\sigma_{KI}^2 = \bar{C}_i(V, V; z) - \sum_{j=1}^n \lambda_j(z) \bar{C}_i(V, x_j; z).$$

Observações:

(a) O sistema KI simples pode ser escrito em termos do variograma usando $\gamma(h) - C(0)$, ficando como:

$$C_i(0; z) \left(1 - \sum_{j=1}^n \lambda_j(z) \right) = \bar{\gamma}_i(V, x_i; z) - \sum_{j=1}^n \lambda_j(z) \gamma_i(x_i - x_j; z); \quad i=1 \text{ a } n.$$

(b) E a variância pode ser escrita como:

$$\sigma_{KI}^2 = C_f(0; z) \left(1 - \sum_{j=1}^n \lambda_j(z) \right) + \bar{\gamma}_f(V; V; z) + \sum_{j=1}^n \lambda_j(z) \bar{\gamma}_f(V; x_j; z).$$

Na hora de ter uma estimativa $\hat{\phi}^*(V; z)$ da proporção $\phi^*(V; z)$ devemos ter valores realizados $i(x_j; z)$, das variáveis indicadoras $I(x_j; z)$, dessa forma podemos ter o valor $\hat{\phi}^*(V; z)$

$$\hat{\phi}^*(V; z) = F(z) + \sum_{j=1}^n \lambda_j(z) i(x_j; z).$$

Vemos que ainda não está completamente determinado o valor $\hat{\phi}^*(V; z)$, por não estar $F(z)$ determinada. Para poder determiná-lo é preciso conhecer $F(z)$. Uma alternativa para estimar o valor $F(z)$, para o cálculo de $\hat{\phi}^*(V; z)$, é pela *estimação direta de $F(z)$* , que será exposta mais em frente.

2.2. Krigagem Indicadora Ordinária

Ao considerar o processo Z estacionário, com esperança desconhecida, vão acontecer mudanças na forma do preditor $\hat{\phi}(V; z)$ para poder satisfazer a condição de não-viciabilidade para $\phi(V; z)$. Esse preditor é da mesma forma do considerado anteriormente, onde os pesos λ_j 's são aqueles que minimizam o erro quadrático médio (2). Esses pesos vão definir o preditor que minimiza o erro médio quadrático, que também tem que ser não-viciado para $\phi(V; z)$, isto é,

$$E(\hat{\phi} - \phi) = \lambda_0 + \sum_{j=1}^n \lambda_j(z) F(z) - F(z) = 0.$$

Equivalentemente, $\sum_{j=1}^n \lambda_j(z) = 1$ e $\lambda_0 = 0$.

O preditor $\hat{\phi}(V; z)$ é um vetor do espaço vetorial gerado pelas combinações lineares

$$\lambda_0 + \sum_{j=1}^n \lambda_j(z) I(x_j; z)$$

e restrito a

$$\sum_{j=1}^n \lambda_j(z) = 1 \text{ e } \lambda_0 = 0.$$

ou seja, $\hat{\phi}(V; z)$ pertence à variedade linear de uma dimensão menor do que o espaço gerado pelas combinações lineares. É assim que o preditor $\hat{\phi}(V; z)$ não-viciado que minimiza (2), denotado por $\hat{\phi}^*(V; z)$ e chamado de *preditor de krigagem indicador ordinario* (KIO), está caracterizado pelas condições de não-viciabilidade e ortogonalidade (por ser uma projecção), isto é,

$$E\{ \hat{\phi}^* - \phi \} = 0 \quad (4)$$

e
$$E\{ [\hat{\phi}^* - \phi] [I(x_j; z) - Y] \} = 0,$$
 com $I(x_j; z)$ e Y pertence à variedade linear.

Supondo o processo I , estacionário e com média desconhecida, a caracterização (4) pode ser escrita:

$$\lambda_0 = 0$$

$$\sum_{j=1}^n \lambda_j(z) = 1 \quad (5)$$

$$\sum_{j=1}^n \lambda_j(z) K_I(x_i - x_j; z) - \mu = \bar{K}_I(V, x_j; z); \quad j=1 \text{ a } n,$$

onde
$$\mu = E\{ [\hat{\phi}^* - \phi] Y \}.$$

Essa caracterização é chamada de *sistema krigagem indicadora ordinária*.

Variância de Predição.

A variância de predição para o preditor KIO, denotada por σ_{KIO}^2 , é dada por

$$\sigma_{KIO}^2 = \bar{K}_I(V, V; z) - \sum_{j=1}^n \lambda_j(z) \bar{K}_I(V, x_j; z) + \mu.$$

Em termos do semi-variograma fica

$$\sigma_{KIO}^2 = \sum_{j=1}^n \lambda_j(z) \bar{\gamma}_I(V, x_j; z) - \bar{\gamma}_I(V, V; z) + \mu.$$

É assim que a variância de predição da KIS é menor que a variância de predição da KID em uma quantidade $\mu = E\{ [\hat{\Phi}^* - \Phi] Y \}$. Isso quer dizer que o preço pago por não conhecer uma característica do processo, neste caso, a esperança $E\{ I(x; z) \} = F(z)$, é o aumento da variância de predição σ_{KID}^2 .

2.3. Análise Estrutural

Feita a caracterização do preditor $\hat{\Phi}^*(V; z)$, o passo fundamental para obter a solução dos sistemas da Krigagem é fazer uma análise estrutural das variáveis indicadoras, estudar as limitações do variograma indicador $\gamma_i(h; z)$ e conseqüentemente as do preditor KI.

Ao plotar os valores $z(x+h)$ versus $z(x)$ se obtém o seguinte

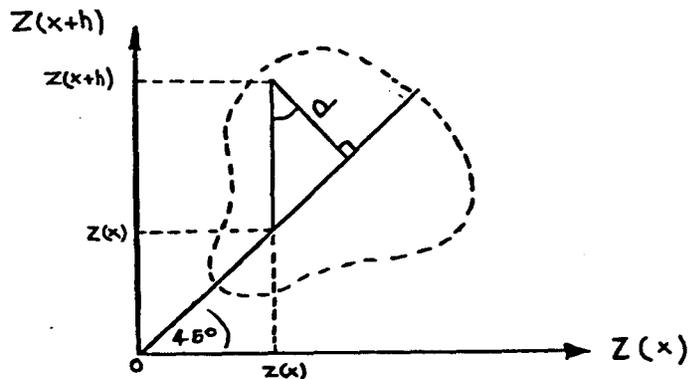


Fig. 4. Plot $z(x+h)$ versus $z(x)$.

sendo

$$d = [z(x+h) - z(x)] \cos(45^\circ)$$

$$d^2 = [z(x+h) - z(x)]^2 \left(\frac{1}{\sqrt{2}} \right)^2$$

e

$$E\{ D^2 \} = \frac{1}{2} E\{ [Z(x+h) - Z(x)]^2 \} = \gamma(h)$$

= momento de inércia em relação ao bissetor.

Assim, o variograma da variável original $Z(x)$ é só o momento de inercia do plot $\{ (z(x+h), z(x)); x \in G \}$ em relação ao bissetor.

Agora, o variograma indicador $2\gamma_i(h; z)$ é por definição:

$$2\gamma_i(h; z) = E\{ [I(x+h; z) - I(x; z)]^2 \},$$

onde

$$| I(x+h; z) - I(x; z) | = \begin{cases} 1 & \text{se } Z(x+h) \leq z \text{ e } Z(x) > z \text{ ou } Z(x+h) > z \text{ e } Z(x) \leq z \\ 0 & \text{se } Z(x+h) \leq z \text{ e } Z(x) \leq z \text{ ou } Z(x+h) > z \text{ e } Z(x) > z. \end{cases}$$

Então, o variograma indicador pode ser interpretado como a distribuição bivariada:

$$2\gamma_i(h; z) = P\{ Z(x+h) \leq z, Z(x) > z \} + P\{ Z(x+h) > z, Z(x) \leq z \}$$

ou o semi-variograma interpretado por

$$\gamma_i(h; z) = F(z) - K_i(h; z) = P\{ Z(x+h) \leq z, Z(x) > z \}$$

Graficamente se vê que, plotando os valores $\{(z(x+h), z(x)); x \in G\}$, a região riscada na fig. 5 está informando a distribuição bivariada $P\{Z(x+h) > z, Z(x) \leq z\}$.

Essa interpretação está diretamente relacionada ao preditor KI, que tem muito mais informação da distribuição bivariada do que um variograma $\gamma(h)$ das variáveis originais $Z(x)$.

Estimação Direta de $F(z)$

Há vários métodos para estimar o valor $F(z)$ da função distribuição univariada do processo Z . O método a apresentar está na base de um valor patamar do semi-variograma indicador $\gamma_i(h; z)$.

Considerando que a variância da variável indicadora:

$$\text{Var}\{ I(x; z) \} = F(z) [1 - F(z)]$$

existe, logo o patamar de um semi-variograma indicador, denotado por $S^2(z)$, e definido por

$$S^2(z) = \gamma_{\infty}(z) = C_1(0; z) = F(z) - F^2(z)$$

também existe.

Esse patamar pode ser estimado desde o semi-variograma indicador experimental

$$\hat{\gamma}(h; z) = \sum_{N(h)} \frac{[i(x_i; z) - i(x_j; z)]^2}{2|N(h)|}, \quad h \in \mathbb{R}^2 \text{ ou } \mathbb{R}^3,$$

onde $N(h) = \{ (x_i, x_j) \mid |x_i - x_j| = h \}$ e $|N(h)|$ é o número de elementos distintos em $N(h)$.

Dessa forma é possível estimar diretamente o valor $F(z)$ pela resolução da equação de segundo grau

$$F^2(z) - F(z) + S^2(z) = 0.$$

As seguintes relações de ordem valem para o patamar $S^2(z)$:

- (a) $S^2(z)$ é máximo e igual a 0.25 para z igual à mediana z_M da distribuição univariada do processo Z , $F(z)$,
- (b) $S^2(z)$ aumenta em $[0, 0.25]$ quando z cresce para z_M ($F(z)$ cresce a 0.5),
- (c) $S^2(z)$ decresce de 0.25 a 0 quando z decresce para z_M .

Na prática z_M é escolhido de forma a que o patamar $S^2(z)$ seja máximo e perto de 0.25; então $\hat{F}(z_M) = 0.5$

$$\hat{F}(z_M) = 0.5$$

$$\hat{F}(z) = \frac{1 + \epsilon(z) \sqrt{1 - 4\hat{S}^2(z)}}{2}$$

onde

$$\epsilon(z) = \text{signo}(z - z_M).$$

Efeito Desestruturação

Ao considerar valores de z extremos ($z \rightarrow \infty$), os valores do conjunto de dados indicadores $\{ i(x_j; z), x_j \in G; j=1 \text{ a } n \}$ vão tender todos a um mesmo valor, 1 ou 0, dessa forma a diferença:

$$| i(x_i; z) - i(x_j; z) |$$

vai tender a ser igual ao valor 0, e dessa forma, na diferença $| i(x_i; z) - i(x_j; z) |$, desaparece a dependência do vetor distância $x_i - x_j = h$, levando o semi-variograma experimental a um modelo do semi-variograma com um *efeito pepita puro*. Por exemplo, um estudo de casos é dado por K. Fytas e outros (1990).

Para sair dessa dificuldade é recomendado definir o semi-variograma para quantis z numa vizinhança da mediana z_M da distribuição $F(z)$, que também evita o problema das relações de ordem descrita a seguir.

Dificuldades do KI

O preditor KI tem três dificuldades

(a) Se K valores de corte são considerados ($z_k; k=1 \text{ a } K$) para discretizar o domínio dos valores das variáveis $Z(x)$, então K sistemas do tipo (3) ou (4) precisam ser resolvidos, onde cada sistema, com um modelo próprio de variograma, dá lugar a um sistema de pesos $\lambda_j(z_k)$'s, dependendo só do valor de corte z_k , o que independe dos outros sistemas pesos. Consequentemente, não é possível assegurar as relações de ordem próprias de uma função de distribuição;

$$\phi^{\#}(V; z) \in [0, 1]; \quad \forall \text{ área } V \text{ e } \forall z,$$

$$\phi^{\#}(V; z) \leq \phi^{\#}(V; z') \quad \forall z \leq z'.$$

Na hora de estimar $\phi(V; z)$ e a estimativa $\phi^{\#}(V; z)$ não satisfaz relações de ordem, uma aproximação $\phi^{\#\#}(V; z)$ para $\phi^{\#}(V; z)$ consiste em corrigir

sequencialmente o problema. Uma das correções a fazer pode ser:

$$\phi^{**}(V; z_1) = \max\{ \phi^*(V; z_1), 0 \}$$

$$\phi^{**}(V; z_k) = \max\{ \phi^{**}(V; z_{k-1}), \phi^*(V; z_k) \}$$

$$\phi^{**}(V; z_k) = \min\{ \phi^{**}(V; z_k), 1 \}$$

com a série de valores ordenados $\{ z_1 < z_2 < \dots < z_k < \dots < z_K \}$ dos K valores de corte considerados.

Uma solução ao problema das relações de ordem é abordado por I. C. Lemmer (1986), que mostra a existência de um teor de corte particular, chamado de valor de corte *monodal*, onde a covariância da variável indicadora e da variável original do teor, satisfazendo certos, critérios, é proporcional um ao outro com um alto grau de exatidão.

A expressão teórica para a covariância indicadora é dada em termos de uma dupla série de polinômios ortogonais com a função densidade como função peso. Uma aproximação dessa expressão simplifica drasticamente o uso do KI para estimar a distribuição de teor pontual local e conseqüentemente evita o problema das relações de ordem. O enfoque não supõe transformar o modelo numa função densidade marginal padrão. Altos termos nas séries, para covariância indicadora, alteram a proporcionalidade estrita. Embora esses termos mostrem amplitudes de dependência do valor do corte, é possível escolher seletivamente o teor de corte. Essa característica é usada para identificar o teor de corte *monodal*.

Resultados referentes às relações de ordem são expostas por A. G. Journel e D. Possa (1990). Eles estabelecem, diretamente das propriedades da função de distribuição bivariada, às relações de ordem que devem ser verificadas nos modelos de variogramas indicadores para múltiplos valores de cortes, chegando a um máximo efeito buraco.

(b) Outra dificuldade do preditor KI está no variograma indicador, que não informa a maneira como a massa de dados está distribuída na região G, mostrada na

figura abaixo. Isso quer dizer que uma distribuição, que é estreita no bissetor, poderia ter um valor do variograma indicador igual ao de uma outra massa diferente, proveniente de uma outra distribuição bivariada espalhada em volta do bissetor.

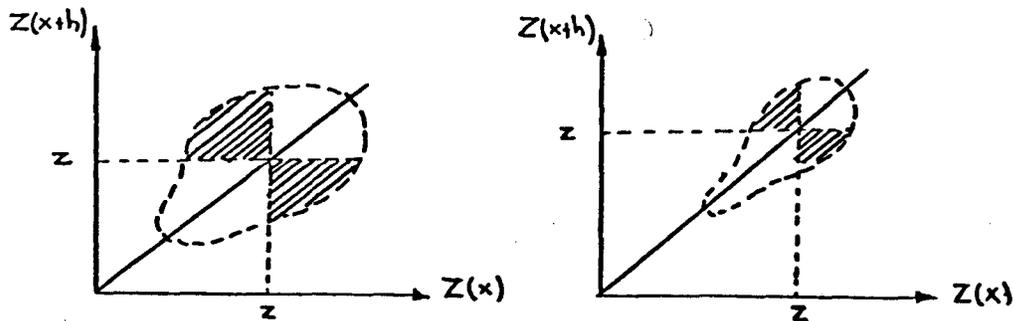


Fig. 5. Distribuição bivariada espalhada em volta do bissetor.

(c) e o efeito desestruturação, exposto anteriormente.

3. Cokrigagem Indicadora

Se se considera o valor z , já não mais fixo mas variável, de maneira a cobrir toda a região da distribuição bivariada, então para cada z tem-se n variáveis indicadoras e para cada par (z_1, z_2) tem-se um variograma indicador cruzado $\gamma_i(h; z_1, z_2)$; $z_1 \neq z_2$. Como não é possível ter esses valores para cada z e cada par (z_1, z_2) ; $z_1 \neq z_2$, por não conhecer a lei que modela o fenômeno, discretiza-se o domínio de variabilidade de z em $K+1$ classes por K valores de corte z_k , obtendo o preditor da forma:

$$\hat{\phi}(V; z_{k_0}) = \sum_{k=1}^K \sum_{j=1}^n \lambda(x_j; z_k) I(x_j; z_k); \quad x_k \in G,$$

onde k_0 é um dos $k=1$ a K .

Este preditor krigagem indicador completo, chamado *preditor de cokrigagem indicador* (CoKI), faz uso de K^2 variogramas cruzados $\gamma_i(h; z_k, z_k)$ que

levam a conhecer a função distribuição bivariada:

$$F(h; z_k, z_k) = P(Z(x+h) \leq z_k, Z(x) \leq z_k).$$

Conseguindo de essa forma que o preditor tenha um maior uso da dependência espacial das variáveis aleatórias $Z(x_j)$. Essa dependência está sujeita a uma muito boa modelagem dos K^2 variogramas cruzados, que não é fácil de obter pelo trabalho computacional nem pelas relações de ordem a serem satisfeitas.

3.1. Krigagem Probabilística

De maneira de diminuir o trabalho da modelagem dos K^2 variogramas indicadores cruzados $\gamma_i(h; z_k, z_k)$ e a melhorar o preditor KI, tem-se o preditor da função distribuição dado por:

$$\hat{\Phi}^*(V; z) = \lambda_0 + \sum_{j=1}^n \lambda_j(z) I(x_j; z) + \sum_{j=1}^n \beta_j(z) Z(x_j).$$

Acontece que as unidades das variáveis indicadoras $I(x; z)$ e das variáveis originais $Z(x)$, são muito diferentes, de maneira que as estimativas poderão ser não-estáveis.

O que se faz para melhorar essa instabilidade das estimativas é transformar a variável teor $Z(x)$ numa variável definida no domínio $[0,1]$. Essa transformação é definida por

$$U(x) = F(Z(x)),$$

onde $F(z) = P(Z(x) \leq z)$ e recebe o nome de *transformada integral*.

Ficando o preditor de $\hat{\Phi}(V; z)$ da forma

$$\hat{\Phi}(V; z) = \lambda_0 + \sum_{j=1}^n \lambda_j(z) I(x_j; z) + \sum_{j=1}^n \beta_j(z) U(x_j), \quad (6)$$

Esse preditor fica totalmente caracterizado pelas condições de não-viciabilidade e ortogonalidade. Equivalentemente o preditor é caracterizado pela minimização da $\text{Var}\{ \hat{\Phi}(V; z) - \hat{\Phi}(V; z) \}$ com respeito aos pesos λ_j 's e β_j 's, sujeito à condição de não-viciabilidade.

Se a esperança do processo I_z é conhecida, a condição de não-viciabilidade é equivalente a:

$$\lambda_0 = F(z) \left[1 - \sum_{j=1}^n \lambda_j(z) \right]$$

$$\text{e } \sum_{j=1}^n \beta_j(z) = 0$$

de maneira que o preditor fica

$$\hat{\Phi}(V; z) = F(z) \left[1 - \sum_{j=1}^n \lambda_j(z) \right] + \sum_{j=1}^n \lambda_j(z) I(x_j; z) + \sum_{j=1}^n \beta_j(z) U(x_j),$$

onde os pesos λ 's e β 's decorrem da minimização da variância de predição sujeito a $\sum_{j=1}^n \beta_j(z) = 0$

Usando o método dos multiplicadores de Lagrange obtem-se o sistema

$$\sum_{j=1}^n \lambda_j(z) C_I(x_k - x_j; z) + \sum_{j=1}^n \beta_j(z) C_{IU}(x_k - x_j; z) = \bar{C}_I(V, x_k; z); \quad k=1 \text{ a } n.$$

$$\sum_{j=1}^n \beta_j(z) C_U(x_k - x_j; z) + \sum_{j=1}^n \lambda_j(z) C_{IU}(x_k - x_j; z) + \mu = \bar{C}_{IU}(V, x_k; z); \quad k=1 \text{ a } n. \quad (7)$$

$$\sum_{j=1}^n \beta_j(z) = 0.$$

Este sistema com $2n+1$ incógnitas e $2n+1$ equações, chamado de *sistema da Krigagem Probabilística*. Este sistema não muda se é escrito nos termos do variograma. Na hora de resolvê-lo é preciso modelar três funções variogramas; o variograma do indicador $\gamma_I(h; z)$, o variograma da variável uniforme $\gamma_U(h)$ e o variograma cruzado da variável indicadora e uniforme. Logo depois de resolver o sistema e obter os pesos λ_j 's e β_j 's é que se tem o preditor de $\Phi(V; z)$, denotado por $\hat{\Phi}^*(V; z)$ e chamado por preditor da *Krigagem Probabilística Simple* (KPS).

Variância de Predição.

A variância de predição do preditor KPS é obtido usando os pesos do sistema (7)

$$\sigma_{KPS}^2 = \bar{C}_I(V, V; z) - \sum_{j=1}^n \lambda_j(z) \bar{C}_I(V, x_j; z) + \sum_{j=1}^n \beta_j(z) \bar{C}_{IU}(V, x_j; z).$$

Considerar agora o caso mais realista, ou seja, onde a esperança do processo I_z não é conhecida. Neste caso, o preditor (6) vai ficar caracterizado pelos pesos λ_j 's e β_j 's que minimizen o erro medio quadrático de predição sujeito à condição de não-viciabilidade

$$\lambda_0 + F(z) \left[1 - \sum_{j=1}^n \lambda_j(z) \right] + \frac{1}{2} \sum_{j=1}^n \beta_j(z) = 0,$$

equivalentemente

$$\lambda_0 = 0$$

$$1 - \sum_{j=1}^n \lambda_j(z) = 0$$

$$\sum_{j=1}^n \beta_j(z) = 0$$

Para determinar completamente o preditor $\hat{\Phi}(V; z)$ dado por a relação (6) têm-se que minimizar a função

$$\text{Var} \left\{ \hat{\Phi}(V; z) - \Phi(V; z) \right\} + 2\mu(z) \left\{ \sum_{j=1}^n \lambda_j(z) - 1 \right\} + 2\nu(z) \sum_{j=1}^n \beta_j(z).$$

Utilizando multiplicadores de Lagrange obtem-se o sistema

$$\sum_{j=1}^n \lambda_j(z) C_I(x_k - x_{j; z}) + \sum_{j=1}^n \beta_j(z) C_{IU}(x_k - x_{j; z}) + \mu = \bar{C}_I(V, x_k; z); \quad k=1 \text{ a } n.$$

$$\sum_{j=1}^n \beta_j(z) C_U(x_k - x_{j; z}) + \sum_{j=1}^n \lambda_j(z) C_{IU}(x_k - x_{j; z}) + \nu = \bar{C}_{IU}(V, x_k; z); \quad k=1 \text{ a } n. \quad (8)$$

$$\sum_{j=1}^n \lambda_j(z) = 1.$$

$$\sum_{j=1}^n \beta_j(z) = 0.$$

Uma vez resolvido este sistema, chamado de sistema *Krigagem Probabilística Ordinária*, é que se origina o preditor que minimiza o erro quadrático medio, denotado por $\hat{\Phi}^*(V; z)$ e chamado de *preditor de Krigagem Probabilística Ordinária*.

Para poder resolver esse sistema de $2(n+1)$ equações e de

$2(n+1)$ incógnitas é preciso modelar, também como no sistema KPS o semi-variograma indicador $\gamma_I(h; z)$, o semi-variograma da variável uniforme $\gamma_U(h)$ e o semi-variograma cruzado da variável indicadora e da variável uniforme $\gamma_{IU}(h; z)$.

Variância de Predição.

A variância de predição do preditor KPS é obtida usando os pesos do sistema KPS que fica na forma

$$\sigma_{KP}^2 = \bar{C}_I(V, V; z) - \sum_{j=1}^n \lambda_j(z) \bar{C}_I(V, x_j; z) + \sum_{j=1}^n \beta_j(z) \bar{C}_{IR}(V, x_j; z) - \mu.$$

Ao fazer uma discretização no intervalo de variabilidade de $Z(x)$, em K valores de cortes z_k ; $k=1$ a K , acontece que K sistemas krigagem probabilística têm que ser resolvidos. Isso exige modelar $2K+1$ funções semi-variogramas K semi-variogramas indicadores $\gamma_I(h; z_k)$ $k=1$ a K , um semi-variograma da variável rango $\gamma_R(h)$ e K semi-variogramas cruzados indicador-rango $\gamma_{IR}(h; z_k)$; $k=1$ a K .

Note que, como na estimativa do KI, na estimativa KP também surge o problema de não satisfazer as relações de ordem, próprias de uma função de distribuição. As correções são feitas como nas do KI.

CAPÍTULO III

PRÁTICA DA KRIGAGEM INDICADORA E PROBABILÍSTICA

Krigagem Indicadora (KI) e Krigagem Probabilística (KP) são dois métodos não-paramétricos para obter uma estimativa da função distribuição:

$$F(z) = P(Z(x) \leq z) = E\{ \phi(V;z) \},$$

estimativa que esta modelada por uma combinação linear de valores indicadores e valores de uma transformação uniforme

$$\phi^*(V;z) = \lambda_0 + \sum_{j=1}^n \lambda_j(z) i(x_j; z), \text{ no caso do KI}$$

$$\phi^*(V;z) = \lambda_0 + \sum_{j=1}^n \lambda_j(z) i(x_j; z) + \sum_{j=1}^n \beta_j(z) u(x_j), \text{ no caso do KP,}$$

onde os pesos λ 's, para o caso de KI são determinados pelo sistema KI e para o caso do KP os pesos λ 's e β 's são determinados pelo sistema KP.

A prática do metodo KI e KP é dividida em nove passos, como se segue:

1º. Determinação do Histograma Desagrupado.

Este passo consiste em ter um conhecimento inicial do

comportamento do fenômeno, resumindo-o em um *histograma de desagrupamento*.

É usual que as localizações $x_i \in G$, dos dados $z(x_i)$, sejam posicionadas em áreas onde se espera que os valores dos teores $Z(x_i)$ sejam grandes ou em áreas onde é possível tirar as amostras facilmente, ou seja, onde as localizações dos dados são preferenciais, portanto esse posicionamento não é regular nem aleatório. Desse tipo de configuração dos dados diz-se que é *agrupado* ou que existe um *agrupamento*.

Esse agrupamento leva a que as estimativas:

$$z_n(x) = \hat{z}_v(x) = \frac{1}{n} \sum_{j=1}^n z(x_j),$$

$$\phi_n(V; z) = \hat{\phi}(V; z) = \frac{1}{n} \sum_{j=1}^n I(x_j; z)$$

das variáveis médias:

$$Z_v(x) = \frac{1}{V} \int_{v(x)} Z(u) du,$$

$$\Phi(V; z) = \frac{1}{V} \int_{v(x)} I(u; z) du,$$

estimem incorretamente os valores dessas variáveis médias, por ter uma amostra com valores muito grandes, não representativa da área de interesse. Esses preditores seriam úteis se a variabilidade do teor $Z(x)$ fosse homogênea e isotrópica sob a área $V(x)$.

Um método alternativo para dar pesos adequados aos valores dados $I(x_j; z)$ é o método empírico proposto por A. G. Journel (1982); onde a idéia é ponderar os dados dentro de cada grupo em relação aos dados do próprio grupo. Daqui é que provem o nome de métodos de *desagrupamentos*. O desagrupamento é uma forma de relacionar a redundância da informação dentro dos grupos de dados.

Técnica de Desagrupamento por Celulas.

Para aplicar esta técnica deve-se ter os n dados em $V(x)$ que tendem a reagrupar-se em L grupos e feita uma divisão da área $V(x)$ em L sub-áreas, correspondendo aos L grupos e dando um sistema de pesos iguais a os dados para

cada área, resultando uma estimativa de $\hat{\phi}(V;z)$, chamada de *histograma acumulado desagrupado*, da seguinte forma:

$$\hat{\phi}(V;z) = \frac{1}{V} \sum_{i=1}^L \int_{v_i} i(x;z) dx = \frac{1}{V} \sum_{i=1}^L v_i \phi(v_i;z),$$

onde $V = \sum_{i=1}^L v_i$, $n = \sum_{i=1}^L n_i$; n_i = número de observações feitas em cada área v_i , e

$$\hat{\phi}(v_i;z) = \frac{1}{n_i} \sum_{j=1}^{n_i} i(x_j;z), \quad x_j \in v_i(x).$$

O procedimento da construção do histograma acumulado desagrupado pode ser resumido em

- 1_ Sobrepor na área $V(x)$ uma malha regular com células de tamanho d .
- 2_ Contar o número de dados n_i que estejam dentro de cada célula d_i ,
- 3_ Ponderar cada dado dentro da células d_i por $\frac{1}{n_i}$,

ficando o histograma acumulado desagrupado como:

$$\hat{\phi}(V;z) = \frac{1}{L} \sum_{i=1}^L \sum_{j=1}^{n_i} \frac{1}{n_i} i(x_j;z).$$

Construído o histograma acumulado desagrupado, com seu correspondente histograma, é possível fazer uma análise monomodal. Se o histograma apresenta modas claramente separadas será checado o mapa das localizações dos dados. Tomar cuidado com o histograma multimodal pois a existência de varias modas não é suficiente para justificar diferentes populações.

2º_ Escolha de K Valores de Cortes.

Neste passo trata-se de fazer a melhor discretização do domínio de variabilidade do teor $Z(x)$, a princípio sem considerar os valores de corte de interesse na prática.

Primeiro deve fazer-se uma análise de diferenças entre

classes definidas por um número fixo de valores de corte $K' = 9$ da forma:

$$\hat{\phi}(V; z_{k'}) = \frac{k'}{10}; \quad k' = 1 \text{ a } 9,$$

ou seja que

$$z_{k'} = \frac{k'}{10} - \text{quantil do histograma acumulado desagrupado.}$$

Esses 9 decis, $z_{k'}$; $k' = 1$ a 9, dividem o histograma acumulado desagrupado em 10 intervalos, $[z_{\min}, z_1], (z_1, z_2), \dots, (z_8, z_9), (z_9, z_{\max}]$, de frequência aproximadamente igual a $\frac{1}{10}$. O valor mínimo e o máximo são determinados de acordo com o problema.

É possível detectar alguma diferença ou semelhança entre as 10 classes, cujo número talvez possa ser reduzido a um número menor, se fossem semelhante, por meio de uma análise dos variogramas indicadores, $\gamma_i(h; z_{k'})$.

Se cada variograma indicador, $\gamma_i(h; z_{k'})$, mostra ser um modelo claramente diferente (não-proporcional) aos próximos variogramas indicadores, $\gamma_i(h; z_{k'-1})$ e $\gamma_i(h; z_{k'+1})$, então as 10 classes são razoavelmente diferentes e os nove valores de corte, definidos pelos nove decis, são mantidos ($K' = K$).

Agora se os variogramas indicadores, $\gamma_i(h; z_{k'})$ e $\gamma_i(h; z_{k'+1})$ têm uma equivalência, que geralmente é proporcional, ou seja, a diferença que se produz é só por um fator constante de proporcionalidade:

$$\gamma_i(h; z_{k'}) \approx \gamma_i(h; z_{k'+1}); \quad \forall h,$$

então, os dois sistemas de pesos, entregados pelo sistema KI para os dois semi-variogramas, $\gamma_i(h; z_{k'})$ e $\gamma_i(h; z_{k'+1})$, são proporcionais, ou seja, que se tem só um sistema. Consequentemente, com os limites correspondentes, entre as classes não há diferenças, levando à redução do número de classes K' , num número igual de variogramas equivalentes.

Considere que não existem diferenças entre as classes $(z_{k'-1}, z_{k'})$ e $(z_{k'}, z_{k'+1})$, ou seja, que na realidade essas duas classes se fundem só em uma. Então para que isso aconteça se transformam os dois limites superiores de

cada classe $z_{k'}$ e $z_{k'+1}$ em um só, por meio de seu quantil médio, isto é, z_p tal que:

$$\hat{\phi}(V; z_p) = \frac{k' + 0.5}{10}.$$

Havendo algum valor de corte, z_{k_0} de interesse, diferente dos outros valores de corte dos já escolhidos ao início (no primeiro passo), é possível escolhê-lo como limite de alguma classe, mudando ligeiramente um valor decil z_k , que possa identificar o valor de corte de interesse, de maneira que a frequência correspondente a cada intervalo

$$\hat{\phi}(V; z_{k+1}) - \hat{\phi}(V; z_k)$$

não mude drasticamente.

A escolha dos $K'=9$ valores de corte iniciais, depende de que hajam pelo menos 10 dados em cada classe, o que é razoável para poder ter uma descrição da distribuição do teor $Z(x)$. Pode-se escolher um maior número de classes $K > 9$, se for possível obter um conjunto de dados de cardinalidade suficientemente grande, de maneira a que cada classe tenha pelo menos 10 dados.

3^o - Modelos para K variogramas indicadores.

O passo anterior, foi uma análise dos semi-variogramas indicadores $\gamma_i(h; z_k)$; $k=1$ a K , com a finalidade de esclarecer se dois semi-variogramas indicadores contíguos são ou não são claramente equivalentes. Note-se que essa análise foi feita uma vez designado o modelo teórico do semi-variograma experimental. Este terceiro passo preocupa-se com essa modelagem. Essa modelagem é produto de K semi-variogramas experimentais $\hat{\gamma}_i(h; z_k)$; $k=1$ a K , onde cada variograma experimental é construído a partir do mesmo conjunto de dados indicadores $\{i(x_j; z_k); j=1$ a n e $k=1$ a $K\}$, isto é, de cada conjunto de dados $\{i(x_j; z_k); j=1$ a $n\}$ é construído $\hat{\gamma}_i(h; z_k)$; $k=1$ a K . Como os K conjuntos indicadores provêm de uma mesma origem, de um mesmo fenômeno, existe uma continuidade entre esses K conjuntos de dados indicadores, mesmo que variogramas contíguos não sejam equivalentes (independentes), existindo uma estrutura que relaciona esses variogramas. Esse relacionamento é aproveitado na prática para construir K

modelos variogramas indicadores desde uma fonte comum, de L estruturas básicas, por exemplo, L modelos esféricos, com pesos dependentes do valor de corte z_k , para cada estrutura que forma a estrutura comum, isto é:

$$\gamma_i(h; z_k) = \sum_{j=1}^L C_j(z_k) \text{Sph}_{a_j}(h).$$

onde a_j denota o alcance do modelo.

4º - Modelo para as K matrizes variogramas.

Este passo é específico para a KPO. O sistema KP, (II.8), agora dado na forma matricial

$$\begin{bmatrix} \gamma(z) & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{b}^T \\ \mu \\ \nu \end{bmatrix} = \begin{bmatrix} \bar{\gamma} \\ \mathbf{1} \\ \mathbf{0} \end{bmatrix}$$

onde

$$\gamma(z) = \begin{bmatrix} \gamma_i(h; z) & \gamma_{iU}(h; z) \\ \gamma_{iU}(h; z) & \gamma_U(h) \end{bmatrix}, \quad \mathbf{1} = \begin{bmatrix} \mathbf{1}^T & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{1}^T \end{bmatrix}, \quad \mathbf{0}_2 = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

$$\mathbf{b} = [\lambda_1(z), \lambda_2(z), \dots, \lambda_n(z), \beta_1(z), \beta_2(z), \dots, \beta_n(z)], \quad \mathbf{1} = [\mathbf{1}, \mathbf{1}, \dots, \mathbf{1}]_{1 \times n}$$

$$\mathbf{0} = [\mathbf{0}, \mathbf{0}, \dots, \mathbf{0}]_{1 \times n}$$

$$\bar{\gamma} = \begin{bmatrix} \bar{\gamma}_i(V, x; z) \\ \bar{\gamma}_{iU}(V, x; z) \end{bmatrix}, \quad \text{cada } \bar{\gamma}_{..}(V, x; z) = \begin{bmatrix} \bar{\gamma}_{..}(V, x_1; z) \\ \bar{\gamma}_{..}(V, x_n; z) \end{bmatrix}$$

$$\gamma_{..}(h; z) = \begin{bmatrix} \gamma_{..}(0; z) & \dots & \gamma_{..}(x_1 - x_n; z) \\ \gamma_{..}(x_2 - x_1; z) & \dots & \gamma_{..}(x_2 - x_n; z) \\ \vdots & \ddots & \vdots \\ \gamma_{..}(x_n - x_1; z) & & \gamma_{..}(0; z) \end{bmatrix}$$

Vê-se que para resolver este sistema de $2(n+1)$ incógnitas e $2(n+1)$ equações, é preciso considerar

- a não existência de dados totalmente redundantes, i. e.,

$$x_i \neq x_j \text{ se } i \neq j, \quad i, j = 1 \text{ a } n.$$

- um modelo para a matriz de funções variogramas γ , onde é preciso modelar os dois variogramas diretos e um variograma cruzado, baseado na condição positiva semi-definida da matriz de covariância.

Assim, como qualquer função $g(h)$ não pode ser considerada como covariância ou semi-variograma de um processo de covariância estacionária ou intrínseco, qualquer conjunto de funções $\{g_{j,j'}(h); j, j' = 1 \text{ a } K\}$ não pode ser considerado como matriz de funções covariâncias ou semi-variogramas de um conjunto de processos estocásticos correlacionados.

A seguir é apresentada uma forma de construir matrizes de funções covariâncias ou semi-variogramas do tipo γ :

Considerar L processos estocásticos independentes, intrínsecos, Y_l ; $l=1$ a L , cada um com semi-variograma $\gamma_l(h)$ (função positiva semi-definida condicional);

Construir dois processos correlacionados Z_j ; $j=1,2$ por meio de uma combinação linear dos L processos, Y_l ; $l=1$ a L , isto é:

$$Z_j(x) = \sum_{l=1}^L a_{jl} Y_l(x); \quad j=1,2.$$

A forma de poder construir uma matriz de semi-variogramas se baseia fundamentalmente em que os coeficientes da combinação linear $Z_j(x)$, satisfaçam as condições necessárias para que cada função de covariância cruzada

$$C_{j,j'}(h) = \text{Cov}\{ Z_j(x), Z_{j'}(x+h) \}$$

seja um elemento de uma matriz de covariância. Considerar o variograma cruzado entre os dois processos Z_j ; $j=1,2$, que é dado por:

$$\begin{aligned} 2\gamma_{j,j'}(h) &= \text{Cov}\{ Z_j(x) - Z_j(x+h), Z_{j'}(x) - Z_{j'}(x+h) \} \\ &= \sum_{i=1}^L \sum_{i'=1}^L a_{ji} a_{j'i'} \text{Cov}\{ Y_i(x) - Y_i(x+h), Y_{i'}(x) - Y_{i'}(x+h) \} \\ &= \sum_{i=1}^L a_{ji} a_{j'i} 2\gamma_i(h); \quad j,j'=1,2. \end{aligned} \quad \text{pela independência dos processos.}$$

Portanto o semi-variograma fica:

$$\gamma_{j,j'}(h) = \sum_{i=1}^L a_{ji} a_{j'i} \gamma_i(h); \quad j,j'=1,2. \quad (1)$$

Podendo ser negativo, ou seja, um crescimento positivo numa variável j' corresponde em média a um decréscimo na outra variável j .

Veja que toda estrutura $\gamma_i(h)$ usada na modelagem dos variogramas cruzados $\gamma_{j,j'}(h)$ também tem que estar presente na modelagem do variograma direto $\gamma_{j,j}(h) = \gamma_j(h)$.

Se os processos são de covariância estacionária, pode escrever-se (1) em termos da covariância:

$$\begin{aligned} C_{j,j'}(h) &= \sum_{i=1}^L a_{ji} a_{j'i} C_i(h); \quad j,j'=1,2. \\ &= \sum_{i=1}^L b_{j,j'} C_i(h); \quad j,j'=1,2. \end{aligned} \quad (2)$$

Para que cada elemento $C_{j,j'}(h)$; $j,j'=1,2$ seja um elemento de uma matriz de funções de covariâncias, uma condição necessária e suficiente é que a

matriz (2) seja positiva semi-definida, ou seja, para qualquer vetor de elementos não-nulos $d=[d_1, d_2]$ a forma quadrática

$$do \left[C_{jj'}(h) \right] o d^T,$$

tem que ser não-negativa.

Cálculos:

$$\begin{aligned} do \left[C_{jj'}(h) \right] o d^T &= \sum_{j=1}^2 \sum_{j'=1}^2 d_j C_{jj'}(h) d_{j'} \\ &= \sum_{l=1}^L \sum_{j=1}^2 \sum_{j'=1}^2 d_j b_{jj'}^{(l)} C_l(h) d_{j'} \\ &= \sum_{l=1}^L \sum_{j=1}^2 \sum_{j'=1}^2 d_j b_{jj'}^{(l)} C_l(x_j - x_{j'}) d_{j'}; \text{ tal que } x_j - x_{j'} = h. \end{aligned}$$

Como $C_l(h)$ é a função covariância do processo Y_l ; $l=1$ a L , logo $C_l(h)$ é uma função positiva semi-definida; $\forall l=1$ a L então

$$\sum_{l=1}^L C_l(x_j - x_{j'}) \sum_{j=1}^2 \sum_{j'=1}^2 d_j b_{jj'}^{(l)} d_{j'} \geq 0$$

o que ocorre se só se

$$C_l(h) \sum_{j=1}^2 \sum_{j'=1}^2 d_j b_{jj'}^{(l)} d_{j'} \geq 0; \forall l=1 \text{ a } L$$

equivalentemente

$$\sum_{j=1}^2 \sum_{j'=1}^2 d_j b_{jj'}^{(l)} d_{j'} \geq 0; \forall l=1 \text{ a } L,$$

se $C_l(h) > 0$, que é o caso considerado. O que significa que a matriz de coeficientes $b_{jj'}^{(l)}$; $j, j'=1, 2$ é positiva semi-definida para todo l ; $l=1$ a L .

Observação:

A matriz

$$\left[b_{jj'}^{(l)} \right] \tag{3}$$

é positiva semi-definida se todos seus valores próprios são não-negativos e reais.

Na prática, não é preciso analisar os valores próprios da matriz (3), porque como máximo são considerados dois ou três processos correlacionados. Como a matriz é simétrica, então uma matriz é definida semi-positiva, se seu determinante principal e todos os determinantes menores são não-negativos, ocupando muito menos tempo nos cálculos dos determinantes que na procura da raiz de um polinômio de terceiro grau.

Assim é que a expressão (1) é um modelo para uma matriz de funções semi-variogramas, se a matriz

$$\left[b_{jj'}^{(l)} \right]; \text{ com } b_{jj'}^{(l)} = a_{j1} a_{j'1} ,$$

é definida semi-positiva, para todo l ; $l=1$ a L .

Esta modelagem foi feita para um específico valor de corte z .

Ao considerar K valores de cortes z_k ; $k=1$ a K , $2K+1$ variogramas são necessários de modelar; K variogramas indicadores $\gamma_k(h; z_k)$; $k=1$ a K , K variogramas cruzados $\gamma_{kl}(h; z_k)$; $k=1$ a K e um variograma direto da variável uniforme $\gamma_U(h)$ que é repetido nas K matrizes variograma $\gamma(z_k)$; $k=1$ a K . Assim obtendo-se como consequência, todas as estruturas usadas na modelagem dos K variogramas cruzados $\gamma_{kl}(h; z_k)$ devem estar presentes como estruturas na modelagem do variograma uniforme $\gamma_U(h)$.

5º - Validação Cruzada.

Neste passo utiliza-se a *distribuição composta*, exposta a seguir:

O valor indicador elementar $i(x_j; z)$ pode ser interpretado como a probabilidade condicional da variável $Z(x_j)$ dado que $Z(x_j) = z_j$,

$$P(Z(x_j) \leq z / Z(x_j) = z_j) = i(x_j; z).$$

Consequentemente, o valor $\hat{i}(x_j; z)$ aparece como uma estimativa da probabilidade

condicional desconhecida:

$$P(Z(x) \leq z / (\text{informação})).$$

Similarmente, o valor

$$\hat{\phi}(V; z) = \frac{1}{V} \int_{V(x)} \hat{i}(u; z) du$$

aparece como uma estimativa da *distribuição condicional composta*:

$$\frac{1}{V} \int_{V(x)} P(Z(x) \leq z / (\text{informação})) dx. \quad (4)$$

Uma validação cruzada para o KI/KP, consiste da comparação da distribuição (ou proporção) estimada composta (4), sob um número n_1 de dados indicadores ao correspondente histograma acumulado dos n_1 dados. Mais explicitamente:

1._ O conjunto de n_1 dados $\{z(x_j; j); j=1 \text{ a } n_1\}$ é diferenciado desde o comêço de toda krigagem, de forma que seja representativo do painel V.

2._ Obtidos os n_1 dados, eventualmente distribuídos e representativos do painel V, o histograma desses n_1 dados é:

$$F_{n_1}(z) = \frac{1}{n_1} \sum_{j=1}^{n_1} i(x_j; z)$$

e para cada valor de corte é

$$F_{n_1}(z_k) = \frac{1}{n_1} \sum_{j=1}^{n_1} i(x_j; z_k); k=1 \text{ a } K. \quad (5)$$

3._ Usar os restantes $n-n_1$ dados avaliados sob V para determinar por KI/KP os n_1 valores indicadores estimados, $\hat{i}(x_j; z); j=1 \text{ a } n_1$, interpretados como a probabilidade condicional $P(Z(x_j) \leq z / (n-n_1))$. E construir a distribuição condicional composta estimada para cada valor de corte

$$\hat{\phi}(V; z_k) = \frac{1}{n_1} \sum_{j=1}^{n_1} \hat{i}(x_j; z_k); k=1 \text{ a } K. \quad (6)$$

4._ Construir os plots para

- os pontos $(F_{n_1}(z_k), \hat{\phi}(V; z_k))$,
- um número de p-quantis do histograma $F_{n_1}(z_k)$ e o correspondente p-quantil da distribuição condicional composta estimada $\hat{\phi}(V; z_k)$. Por exemplo, os nove decis.

Se os plots estão alinhados no bissetor, então o KI/KP não é invalidado. A invalidação do KI/KP dá-se quando os plots estão muito desviados do bissetor.

Idealmente, o exercício de validação cruzada através das expressões (5) e (6) é repetida, usando-se os mesmos modelos variogramas, mas diferentes conjuntos de n_1 dados.

6^o_ Apresentação do KI/KP.

Como toda produção da Krigagem não é só representada pelo resultado da estimativa de $\phi^*(V; z)$ ou $i^*(x; z)$, é necessário também apresentar uma referência de

- a variância da Krigagem $\sigma_{KI}^2(x; z)$ ou $\sigma_{KP}^2(x; z)$. Se há uma variância negativa ou próxima de zero, então existem problemas de quasi-singularidade no sistema Krigagem.
- o número de dados usados para a Krigagem e os valores correspondentes.

Essas referências são as que vão permitir julgar o resultado da Krigagem, como interessante ou duvidoso.

7^o_ Correção das relações de ordem.

As correções das relações de ordem da proporção $\phi^*(V; z)$, podem ser encaminhadas pelas correções apresentadas no capítulo anterior.

8^o - Interpolação entre os K valores de $\phi^*(V; z)$.

Lembramos que o que se busca é conhecer uma estimativa $\phi^*(V; z)$ da função distribuição $\phi(V; z)$ no domínio todo, não só em alguns valores do domínio como se tem feito até agora, por meio da KI (passos 1, 2, 3, 5, 6, e 7) ou por a KP (passos 1, 2, 3, 4, 5, 6 e 7). Então procede-se para completar essa estimativa através de uma interpolação qualquer. Qualquer interpolação, satisfazendo as relações de ordem nos intervalos $(z_k, z_{k+1}]$; $k=1$ a K , é assumida como sendo a estimativa da distribuição $\phi(V; z)$ dentro das classes. Um exemplo de interpolação é apresentado a seguir:

- Ajustar qualquer modelo de distribuição paramétrica, por exemplo de dois parâmetros, identificando os limites da classe z_k e z_{k+1} como seus p_1 e p_2 -quantis, isto é,

$$p_1 = \phi^*(V; z_k)$$

$$p_2 = \phi^*(V; z_{k+1}).$$

Então qualquer distribuição $T(z)$, com a expressão para os quantis conhecidas como função dos dois parâmetros, é aceitável para a interpolação. Por exemplo, o modelo *lognormal*: a variável X , com média $E(X) = \alpha$ e variância $\text{Var}(X) = \beta^2$, tem uma distribuição *lognormal*, se a variável $Y = \log X$ tem uma distribuição normal, com média $E(Y) = \mu$ e variância $\text{Var}(Y) = \sigma^2$.

Aqui a expressão log-quantil é $\log z_p = \mu + \sigma T^{-1}(p)$, onde $T(z)$ é a função normal padrão e $T^{-1}(p)$ tem uma boa aproximação polinomial (Abramovitz and Stegun, 1965).

Os dois parâmetros, μ e σ , são determinados resolvendo-se o sistema linear

$$\log z_k = \mu + \sigma T^{-1}(p_1)$$

$$\log z_{k+1} = \mu + \sigma T^{-1}(p_2),$$

com $p_1 = \phi^*(V; z_k)$

e $p_2 = \phi^*(V; z_{k+1})$.

Assim a função distribuição estimada intra-classe é

$$\phi^*(V; z) = T\left(\frac{\log z - \mu}{\sigma}\right), \quad \forall z \in (z_k, z_{k+1}].$$

Outros exemplos encontram-se em A. G. Journel (1988).

9º - Estimação da função distribuição bloco-v.

Até agora conseguiu-se a estimação da distribuição da variável teor $Z(x)$, com suporte pontual, dentro de uma área $V(x)$ (ou painel). Na prática não é de interesse a distribuição dos teores com suporte pontual, mas sim a distribuição do teor médio ou do teor com suporte bloco-v sob um painel $V(x)$, ou seja conhecer a distribuição da variável aleatória

$$Z_v(x) = \frac{1}{V} \int_{v(x)} Z(u) du, \quad v(x) \subset V(y)$$

denotada por $\phi_v(V; z)$ e definida pela proporção de teores blocos, menores ou iguais ao valor de corte z

$$\phi_v(V; z) = \frac{1}{V} \int_{v(x)} I_v(x; z) dx,$$

com
$$I_v(x; z) = \begin{cases} 1 & \text{si } Z_v(x) \leq z \\ 0 & \text{si não.} \end{cases}$$

Note que

$$I_v(x; z) = \frac{1}{V} \int_{v(x)} I(y; z) dy, \quad \text{média sob } v(x) \text{ da variável indicadora pontual.}$$

Agora, o problema é prever a integral estocástica $\phi_v(V; z)$ ou a variável indicadora do teor médio sob $v(x)$, $I_v(x; z)$, a partir da informação fornecida pelas variáveis com suporte pontual. No entanto, essa predição pode ser dada por KI ou KP.

Como a variável indicadora bloco- v , $I_v(x;z)$, não é uma média das variáveis indicadoras pontuais, a variável aleatória indicadora bloco- v é de natureza diferente da variável indicadora pontual. Veja então, que não é razoável definir como preditor de $\Phi_v(V;z)$,

$$\hat{\Phi}(V;z) = \sum_{j=1}^n \lambda_j(z) I(x_j;z),$$

onde os pesos λ_j 's, ficam determinados resolvendo o sistema KI (simples ou ordinário), uma vez que o preditor $\hat{\Phi}(V;z)$ é viciado para $\Phi_v(V;z)$ e é preciso resolver o sistema

$$\sum_{j=1}^n \lambda_j C_1(x_i - x_j; z) = \bar{K}_{1, I_v}(V, x_i; z); \quad j=1 \text{ a } n$$

$$\sum_{j=1}^n \lambda_j = 1,$$

onde aparece a covariância cruzada não-centrada média $\bar{K}_{1, I_v}(V, x_i; z)$, produto da covariância cruzada não-centrada

$$K_{1, I_v}(h; z) = E[I(x+h; z) I_v(x; z)] = P(Z(x+h) \leq z \text{ e } Z_v(x) \leq z),$$

que tem que ser modelada.

Como infelizmente não existe informação das variáveis bloco $Z_v(x)$, conseqüentemente também, não existem dados indicadores blocos da variável $I_v(x;z)$. Faltando essa informação não é possível a inferência da covariância cruzada não-centrada $K_{1, I_v}(h; z)$, nem do variograma cruzado das variáveis pontual e bloco $\gamma_{1, I_v}(h; z)$.

Três alternativas são propostas para resolver a predição de $\Phi_v(V;z)$:

1. Operar com um modelo para a covariância indicadora pontual-bloco para obter a estimativa $\hat{\Phi}_v^*(V;z)$ da função distribuição $\Phi_v(V;z)$. Este é o enfoque considerado pelo método KD, cuja técnica produz bom resultado na prática, porém muito dependente de extensas expansões polinomiais, qualquer que sejam os modelos de covariâncias.

2._ Assumir desde o início que o processo estocástico tem uma distribuição multivariada conveniente, por exemplo, uma distribuição relacionada à Normal, seguindo-se que a variável média $Z_v(x)$ seja também distribuída segundo uma Normal.

3._ Estimar a função distribuição pontual $\phi^*(V;z)$ pelas métodos apresentados KI ou KP. Logo, corrigir essa estimativa que produz o efeito suporte por meio de uma redução da variância para uma estimativa da distribuição bloco. Esta alternativa foi exposta no capítulo Conceitos Elementais para a Resolução de Problemas de Mineração, que será válida só se o tamanho do bloco-v for pequeno em consideração ao painel de volume V.

Correção do efeito suporte para as funções locais de recuperação

Logo de ter feito a estimativa da função distribuição espacial local $\phi^*(V;z)$ sobre uma subarea $V \subset G$, é preciso transformar essa estimativa em uma estimativa da distribuição espacial bloco-v, $\phi_v^*(V;z)$, um fator de correção de variância relativa pode ser aplicado

$$K^2(v/V) = \frac{D^{2*}(0/V) - D^{2*}(v/V)}{D^{2*}(0/V)} = \frac{\bar{\gamma}(v,v)}{D^{2*}(0/V)}$$

onde

$$D^{2*}(0/V) = \int_0^\infty [u - m_v^*]^2 d\phi^*(V;u)$$

e

$$m_v^* = \int_0^\infty u d\phi^*(V;u)$$

são a média e variância da distribuição pontual estimada.

A correspondente correção de afinidade da variância entrega a estimativa para a distribuição espacial bloco-v dentro do painel V, ficando da forma

$$\phi_v^*(V;z) = \phi^*(V;z')$$

onde

$$z' = m_V^* + \frac{1}{\sqrt{[1-K^2(v/V)]}}(z - m_V^*)$$

para todo z , todo $v \ll V$ e $V \subset G$.

Dessa forma é possível obter as estimativas das funções de recuperação locais bloco- v .

- Tonelagem recuperável.

$$t_V^*(V; z) = 1 - \phi_V^*(V; z)$$

- Quantidade de metal recuperável.

$$q_V^*(V; z) = \int_z^\infty u d\phi_V^*(V; u). \quad (7)$$

Essa integral de Stieljes na prática é aproximada por uma soma discreta

$$q_V^*(V; z) \simeq \sum_{k=1}^{K(z)} z_k' [\phi_V^*(V; z_{k+1}) - \phi_V^*(V; z_k)],$$

onde z_k ; $k=1$ a $K(z)$ são os valores que discretizam o intervalo de integração $[z, \infty]$ e z_k' é um valor central do intervalo $[z_{k+1}, z_k]$

- Teor médio de mineral num corte z .

$$m_V^*(V; z) = \frac{q_V^*(V; z)}{t_V^*(V; z)}$$

CONSIDERAÇÕES FINAIS

A metodologia descrita nessa dissertação resulta muito útil para descrever o comportamento de fenômenos muito variáveis, típico dos depósitos com altos teores, por exemplo, depósitos de urânio, ouro, etc.. A utilidade é fundamentada pelo tipo de dados considerados, dados de forma ordenada. A informação desses dados é guardada nos variogramas diretos $\gamma_i(h; z)$, $\gamma_U(h)$ e no variograma cruzado $\gamma_{UV}(h; z)$ definido pelo primeiro momento.

O método inicialmente considera uma discretização do domínio da variável teor, de forma a representar a variável teor. Nasce aqui o aspecto descontínuo da distribuição local estimada. Assim ao considerar o painel como uma simples localidade x e tendo feito uma pequena amostragem na vizinhança de x , o estimador da distribuição local corresponde de fato a uma distribuição discreta com poucos valores (os valores amostrados, observados na vizinhança da localidade x).

Uma solução clássica para regularizar esse histograma, mesmo sendo suficiente o tamanho da amostragem, é por meio de qualquer ajuste paramétrico que satisfaz as relações próprias de uma função de distribuição

$F(z)$ é uma função não decrescente em z ,

$F(z)$ é um valor em $[0,1]$.

O fenômeno metalogênico a estudar tem que ser razoavelmente homogêneo, de forma a considerar o processo que modela o fenômeno como estacionário. Sem essa hipótese é difícil a aplicação da técnica Geoestatística. E ao transformar o processo original num processo indicador é que é possível encontrar um ambiente adequado para a Geoestatística, ou seja, que se cumpre a hipótese de estacionariedade de segunda ordem ou de covariância estacionária, podendo conseguir uma relação biunívoca entre a covariância e o variograma. Com isso o processo tem uma função média estacionária que depende só do valor de corte. Conhecendo ou não essa função média do processo transformado obtêm-se estimadores diferentes para a função de distribuição. A diferença está só num coeficiente que torna o estimador não-viciado. Porém a importância do

conhecimento da função média do processo faz o estimador ter uma variância de estimação menor que o estimador obtido de um processo com uma função média desconhecida.

No capítulo III foi exposta parcialmente a Krigagem Indicadora completa, chamada de Cokrigagem Indicadora, argumentando que é grande o esforço de cálculo e modelagem para fazer uso dela. O verdadeiro argumento está só em duas situações:

(i) Funções indicadoras $I(x; z)$ são não-correlacionadas para varios valores de cortes. A covariância indicadora cruzada é nula só se as variáveis teores $Z(x+h)$ e $Z(x)$ são independentes para qualquer h . Por outro lado, sempre existe uma forte correlação positiva entre as funções indicadoras em vários cortes, especialmente para funções indicadoras na mesma localização x .

(ii) A Cokrigagem é reduzida à Krigagem devido às correlações intrinsecas existentes entre as funções aleatórias em diferentes cortes $I(x; z)$ e $I(x'; z')$. Matheron (1982) mostrou que esta situação ocorre só num modelo muito especial de distribuições bivariadas, chamado por ele, de Modelo Mosaico, onde o correlograma cruzado e direto da variável indicadora $I(x; z)$ são todos iguais.

Uma vez escolhidos os valores dos teores que irão discretizar o domínio em estudo, se procede à estimação da função distribuição nesses teores. As estimativas são independentes umas das outras, isso faz com que as estimativas não necessariamente satisfazam as relações de ordem.

Esse problema encontra-se em todos os estimadores que são projeção de uma incógnita no subespaço do maior espaço possível de definir (espaço das funções Borel-mensurável das n variável), mas o estimador esperança condicional verifica as relações de ordem.

Esse problema das relações de ordem está ligado à não-convexidade do processo da Krigagem: o estimador de Krigagem não necessariamente pertence ao intervalo definido pelo menor e maior valor dos dados avaliados. Na Krigagem Ordinaria a estimativa não-convexa está ligado aos pesos negativos.

A nível de construção de modelos que possam assegurar as relações de ordem, uma alternativa rápida é supor um modelo de dependência espacial intrínseco para a matriz variograma indicadora [$\gamma_i(h; z_k, z_k)$]. Mais explicitamente, K variogramas diretos [$\gamma_i(h; z_k)$; $k=1$ a K] são requeridos para a KI, supondo que sejam os K variogramas indicadores proporcionais uns com os outros, conseqüentemente um só variograma direto precisa ser inferido. O variograma que é escolhido é o variograma indicador correspondente ao valor de corte $z_k = M = F^{-1}(0.5)$, reduzindo K sistemas de pesos a um só, esta restrição é equivalente se a distribuição bivariada tem um modelo Mosaico (Matheron, 1982).

Uma outra alternativa é considerar um relacionamento entre os variogramas indicadores menos exigente, por exemplo um modelo do tipo exposto no cap.III, 9º. passo (Journel, 1988).

E por último, é através da modelagem da covariância indicadora que, ao usar polinômios ortogonais, se chega a uma proporcionalidade entre o variograma indicador e o variograma da variável teor sob certas restrições (Lemmer, 1986).

Note-se que a primeira e a última alternativas são basicamente as mesmas.

Por outro lado, essas relações de ordem podem ser corrigidas depois de se ter feito as estimativas da função de distribuição pontual. Uma correção desse tipo foi apresentada no cap.II (dificuldade da KI), usada também na KP. Uma engenhosa forma de corrigir as estimativas e usar uma técnica de programação quadrática. (J. Sullivan, 1984)

A definição da variável indicadora, que depende de um valor de corte, leva a um efeito de desestruturação dos altos teores, i.e., na medida em que é aumentado o valor do teor de corte, maior é a quantidade de valores indicadores que vão tomar o valor zero, chegando a uma diminuição considerável de dados indicadores para conseguir uma inferência razoável do variograma indicador e do variograma cruzado entre a variável indicadora e uniforme.

O limite da Krigagem Indicadora e Probabilística é encontrado

ao estimar a função distribuição bloco-v. Não é possível estimar a distribuição bloco-v por não se ter dados indicadores bloco-v $i_v(x; z)$, tal como definido no cap. III, 9º. passo. Dessa forma, não sendo possível a inferência das funções estruturais definidas na variável indicadora bloco-v $I_v(x; z)$, chega-se assim ao limite do método que tentou conservar a informação fornecida só pelos dados.

Nesta fase é onde se deve usar a técnica, já muito usada desde o início da Geoestatística, chamada correção de afinidade da variância, para poder corrigir o efeito produzido na estimativa pontual ao considerar variáveis de tipo médias, variáveis bloco-v.

Essa técnica que ajuda a corrigir o efeito suporte consiste em supor que as distribuições pontuais e bloco-v são idênticas, exceto para suas variâncias, i.e., as distribuições padronizadas são iguais

$$P\left\{ \frac{Z_v(x) - m_v^*}{D^*(v/V)} \leq z \right\} = P\left\{ \frac{Z(x) - m_v^*}{D^*(0/V)} \leq z \right\}, \quad \forall z,$$

onde $D^*(\cdot/V) = \sqrt{D^{2*}(\cdot/V)}$.

Como todo método, antes de ser aplicado é fundamental que exista uma coerência entre as retrições do método e os dados com que se vai trabalhar, evitando estimativas não-interpretáveis.

Em todo método baseado exclusivamente nos dados (sem impor nenhuma hipótese à distribuição do processo), como é o caso aqui, a amostragem é feita de forma a ser eventualmente representativa do painel onde se faz a estimativa das reservas recuperáveis locais.

Seria de grande interesse que se continue trabalhando na busca da relação entre a distribuição do teor e representação da covariância indicadora associada, porque dessa forma é possível interpretar o variograma do teor usando o variograma indicador, no caso de difícil modelagem quando se tem uma porção muito pequena de teores com valores muito altos e assim também evitar as relações de ordem.

APÊNDICE

A condição $\langle Z_0, f(Z_i) \rangle = \langle Z_{DK}^*, f(Z_i) \rangle \quad \forall i=1 \text{ a } n$ e para toda função mensurável f é satisfeita, se e só se, o sistema

$$E\{ Z_0/Z_i \} = E\{ Z_{DK}^*/Z_i \}; \quad i=1 \text{ a } n.$$

é satisfeito

Seja d_i o espaço vetorial gerado por todas as funções mensuráveis de uma só variável $f(Z_i)$. A esperança condicional de Z_0 dado Z_i , $E_i\{ Z_0 \} = E\{ Z_0/Z_i \}$ é a definição da projeção de Z_0 em d_i . Assim o vetor $Z_0 - E_i\{ Z_0 \}$ é ortogonal a qualquer vetor $f(Z_i)$ que pertença ao espaço d_i , isto é:

$$\langle Z_0 - E_i\{ Z_0 \}, f(Z_i) \rangle = 0; \quad \forall f(Z_i) \in d_i.$$

Essa ortogonalidade pode ser escrita como

$$\langle Z_0, f(Z_i) \rangle = \langle E_i\{ Z_0 \}, f(Z_i) \rangle; \quad \forall f(Z_i) \in d_i \quad (1)$$

O mesmo procedimento é realizado ao considerar a projeção de Z_{DK}^* em d_i

$$\langle Z_{DK}^*, f(Z_i) \rangle = \langle E_i\{ Z_{DK}^* \}, f(Z_i) \rangle; \quad \forall f(Z_i) \in d_i \subset D_n.$$

Com isso a expressão (1.3) pode ser escrita como

$$\langle Z_0, f(Z_i) \rangle = \langle E_i\{ Z_{DK}^* \}, f(Z_i) \rangle; \quad i=1 \text{ a } n$$

e a expressão (1) é escrita como:

$$\langle E_i\{ Z_0 \}, f(Z_i) \rangle = \langle E_i\{ Z_{DK}^* \}, f(Z_i) \rangle; \quad i=1 \text{ a } n \text{ e } \forall f \in d_i$$

que é satisfeita, se e só se, $E_i\{ Z_0 \} = E_i\{ Z_{DK}^* \}; \quad i=1 \text{ a } n \quad \square$

REFERÊNCIAS

- 1._ Abramovitz, M. and Stegun, I.A. (1965). *Handbook of Mathematical: Dover Publishers, New York, 1046 p.*
- 2._ Alfaro, M. (1975). *Aplicaciones de la Teoria de la Variable Regionalizada: Catedra de Matemáticas, 3, Esc. Tec. Sup. de Ingenieros de Minas, Madrid.*
- 3._ Armstrong, M. and Diamond, P. (1984). *Testing Variograms Positive-Definiteness: Mathematical Geology, Vol. 16, No. 4, pp.*
- 4._ Clark, I. (1979). *Practical Geostatistics: Applied Science Publishers, London,129p.*
- 5._ Cressie, N. and Hawkins, D.M. (1980). *Robust Estimation of the Variogram: Journal of the International Association for Mathematical Geology, Vol. 12, pp.115-125.*
- 6._ Cressie N. (1986). *Kriging Nonstationary Data: Journal of the American Statistical Association, Vol. 81, No. 395, pp. 625-634.*
- 7._ Cressie N. (1988). *Spatial Prediction and Ordinary Kriging: Mathematical Geology, Vol. 20, No. 4, pp. 405-421.*
- 8._ Davis, M. (1978). *Sampling and Estimation Problems for Three Dimensional Spatial Stationary and Non Stationary Stochastic Processes as Encountered in the Mineral Industry: Journal of Statistical Planning and Inference, Vol. 2, pp.221-244.*
- 8._ Davis, B.M. (1984). *Indicator Krigins as Applied to an Alluvial Gold Deposit: G. Verly et-al. (eds.), Geostatistic for Natural Resource Characterization, Part 1, pp.337-348, Reidel Publ. Co.*
- 9._ Graybill, F.A. (1976). *Theory and Application of the Linear Model: Duxbury Press, North Scituate, Massachusetts.*
- 10._ Journel, A.G. (1977). *Kriging in Terms of Projections: Mathematical Geology, Vol. 9, No. 6, pp. 563-586.*
- 11._ Journel, A.G. and Huijbregts, Ch. J. (1978). *Mining Geostatistics: Academic Press, London, 600 p.*
- 12._ Journel, A.G. (1980). *The Lognormal Approach to Predicting Local Distributions of Selective Minig Unit Grades: Journal of the International Association for Mathematical Geology, Vol. 12, No. 4, pp. 283-303.*
- 13._ Journel, A.G. (1983). *Nonparametric Estimation of Spatial Distributions: Mathematical Geology, Vol. 15, No. 3, pp. 445-468.*
- 14._ Journel, A.G. (1988). *Geostatistics for the Environmental Sciences: United States Environmental Protection Agency, Las Vegas.*

- 15._ Journel, A.G. and Posa, D. (1990). *Characteristic Behavior and Order Relations for Indicator Variograms*: *Mathematical Geology*, Vol. 22, No. 8, pp. 1011–1025.
- 16._ Kytas, K., Chaouai, N.E. and Lavigne, M. (1990). *Gold Deposits Estimation Using Indicator Kriging*: *Bulletin CIM*, Vol. 83, No. 934, pp. 77–83.
- 17._ Lemmer, I.C. (1986). *Monodal Indicator Variography, Part I, II*: *Mathematical Geology*, Vol. 18, No. 7, pp. 589–623.
- 18._ Marechal, A. (1984). *Recovery Estimation. A Review of Models and Methods*: G. Verly et al. (eds.). *Geostatistics for Natural Resources Characterization, NATO-ASI, Part 1*, pp. 385–420, Reidel Publ. Co., Dordrecht, Netherland.
- 19._ Matheron, G. (1976). *A Simple Substitute for Conditional Expectation, Disjunctive Kriging*: Guarascio et al. (eds.), *Geostatistics for Natural Resources Characterization, NATO-ASI*, pp. 221–236, Reidel Publ. Co., Dordrecht, Netherlands.
- 20._ Matheron, G. (1982). *La Destructuration des Hautes Teneurs et le Krigage des Indicatrices*: *Fontainebleau Series*, No. 761.
- 21._ Myers, D.J. (1989). *To Be or Not to Be... Stationary? That is the Question*: *Mathematical Geology*, Vol. 21, No. 3, pp. 347–362.
- 22._ Nevéu, J. (1965). *Mathematical Foundations of the Calculus of Probability*: Holden-Day, Inc., 223 p.
- 23._ Remacre, A.Z. (1989). *Uniform Conditioning versus Indicator Kriging*: M. Armstrong (eds.). *Geostatistics*, Vol. 2. pp. 947–960. Kwwer Acad. P.
- 24._ Sullivan, J. (1984). *Conditional Recovery Estimatio Through Probability Kriging. Theory and Practice*: G. Verly et al. (eds.), *Geostatistics for Natural Resources Characterization*, pp. 365–384, Reidel Dordrecht, Holland.
- 25._ Zimmerman, D.L. and Zimmerman, M.B. (1991). *A Comparison of Spatial Semivariogram Estimators and Corresponding Ordinary Kriging Predictor*: *Technometrics*, Vol. 33, No. 1, pp. 77–91.