GEOESTATÍSTICA

- **Estatística clássica**
  - **variáveis independentes**
  - **sem continuidade espacial**

- **Estatística espacial**
  - **valores associados à localização no espaço e/ou no tempo**
  - **distribuição contínua dos valores**
  - **processos de estimativa para pontos não amostrados**

- **Geoestatística**
- **variáveis regionalizadas: fenômeno natural**

- **Fenômeno natural**
  - **aspecto estrutural (determinístico)**
  - **aspecto errático (casual)**
  - **correlação espacial**

- **Variograma**
  - **quantificação da continuidade espacial**

- **Procedimentos em geoestatística**
  - **análise exploratória dos dados**
  - **calculo do variograma experimental**
  - **modelagem**
  - **krigagem: estimativa e interpolação**
  - **simulação**

VARIOGRAMA

- **Ferramenta básica, que permite descrever quantitativamente a variação no espaço de um fenômeno regionalizado.**

- **A natureza estrutural de um conjunto de dados, assumido pela variável regionalizada, é definida a partir da comparação de valores tomados simultaneamente em dois pontos, segundo uma determinada direção.**

- **A função variograma $2\gamma(h)$ é definida como sendo a esperança matemática do quadrado da diferença entre os valores de pontos no espaço, separados por uma distância h.**

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

$$2\gamma(h) = \frac{1}{n}\sum_{i=1}^{n} [Z(x+h) - Z(x)]^2$$

- **A interpretação do variograma permite obter parâmetros que descrevem o comportamento espacial das variáveis regionalizadas.**

- **Uma feição resultante da análise dos parâmetros do variograma experimental é a zona de influência: qualquer valor de Z(x) estará correlacionado com outros valores Z(x+h) que estiverem dentro de um raio "a" de x. Esta correlação, ou a influência de um valor em outro, decresce conforme Z(x+h) aproxima-se de "a".**

- **O variograma é utilizado para calcular os valores de semivariância, para uma dada distância, os quais são necessários para a organização do sistema de equações de krigagem.**

- **Necessidade de ajustar uma função matemática que descreva continuamente a variabilidade ou correlação espacial existente nos dados. O variograma experimental não serve para esse fim, porque há necessidade de interpolação e os pontos apresentar-se-ão com uma certa dispersão, principalmente para distâncias grandes, quando o número de pares de amostras diminui.**

- **O variograma substitui a distância euclidiana "h" pela distância "$2\gamma(h)$", atributo específico do local em estudo. A distância dada pelo variograma mede o grau médio de similaridade entre um valor não amostrado e um valor conhecido vizinho.**

A geoestatística estuda o comportamento das chamadas variáveis regionalizadas e fundamentalmente baseia-se nos seguintes pressupostos:

Ergodicidade: a esperança referente à média de todas as possíveis realizações da variável é igual a média de uma única realização dentro de um certo domínio;

Estacionariedade: na região em que se pretende fazer estimativas, o fenômeno é descrito como homogêneo dentro desse espaço;

Hipótese íntrinsica: as diferenças entre valores apresentam fraco incremento, isto é, as diferenças são localmente estacionárias.

Para a obtenção de um variograma, portanto, é suposto que a variável regionalizada tenha um comportamento fracamente estacionário, onde os valores esperados, assim como sua covariância espacial, sejam os mesmos por uma determinada área.
Assume-se, desse modo, que os valores dentro da área de interesse não apresentem tendência que possam afetar os resultados.

- variável regionalizada x(i) coletada em pontos i

- x(i) e x(i+h): valores de uma variável regionalizada obtidos nos pontos i e i+h, separados entre si por múltiplos da distância $\vec{h}$, vetor com direção específica, em um espaço a uma, duas ou três dimensões

- valor de cada ponto está relacionado com valores obtidos à partir de pontos situados a uma certa distância; a influência será tanto maior quanto menor for a distância entre os pontos

- grau de relação entre valores numa certa direção pode ser expresso pela covariância ou pela variância;

- covariância e variância entre valores de uma mesma variável, porem obtidos em pontos x e x+h.

- para qualquer deslocamento h, os dois primeiros momentos da diferença [x(i)-x(i+h)] são independentes da localização de x e função apenas de h
  média = m = E[x(i)-x(i+h)] = E[Z(x]
  variância = E{[x(i)-x(i+h)]-m]2} = E{[Z(x) – m]2}

- covariância = C(h) = E[Z(x+h).Z(x)] – m2
  covariância depende do tamanho de h
  h = 0, C(h) passa a representar a variância = C(0)

- semivariância: metade da variância das diferenças x(i+h) - x(i)]
  $\gamma(\vec{h}) = \gamma(h) = \Sigma[Z(x+h) – Z(x)] / 2$

- variância de X = $[\Sigma x^2 /n] – [(\Sigma x/n)^2]$
  $\gamma(h) = [1/2\ \Sigma(x(i+h) - x(i))^2]/n - [1/2\ \Sigma\ (x(i+h) - x(i))\ /n^2\ ]$
  como $[\{ 1/2\ \Sigma\ (x(i+h) - x(i)) \} / n] = 0$,
  $\gamma(h) = 1/2[\{\Sigma x(i+h)^2\}/n + \{\Sigma\ x(i)^2 \}/n] - \{\Sigma x(i+h)x(i)\}/n$,
  o que significa que
  $\gamma(h) = C(0) - C(h)$

- sendo x(1), x(2), ... x(i), ... x(n), realizações de uma variável regionalizada, satisfazendo a hipótese intrínseca, a estimativa não tendenciosa da semivariância é dada por
  $\gamma(h) = 1/2n\ \Sigma\{x(i+h) - x(i)\}^2$

- relações são mostradas quando a função $\gamma(h)$ é colocada em gráfico contra h para originar o semivariograma

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n} [x(i+h) - x(i)]^2$$

- semivariância distribui-se de 0, quando h=0, até um valor igual a variância das observações para um alto valor de h

- a distância, segundo a qual $\gamma(h)$ atinge um patamar (soleira/sill), igual a variância dos dados, é chamada de alcance (range).

- para a utilização do semivariograma as seguintes suposições básicas são requeridas:

a) as diferenças entre pares de valores de amostras são determinadas apenas pela orientação espacial relativa dessas amostras;

b) o interesse é enfocado apenas na média e na variância das diferenças, significando que esses dois parâmetros dependem unicamente da orientação (hipótese intrínseca);

c) por conveniência assume-se que os valores da área de interesse não apresentam tendência que possa afetar os resultados e assim a preocupação será apenas com a variância das diferenças entre valores das amostras.

- Semivariograma mostra, pela análise estrutural, o comportamento espacial da variável regionalizada ou de seus resíduos, quando na presença de tendência:

- tamanho da zona de influência em torno de uma amostra; toda amostra cuja distância ao ponto a ser estimado for menor ou igual ao alcance, fornece informações sobre o ponto

- anisotropia, quando os semivariogramas se mostram diferentes para diferentes direções de linhas de amostragem;

- continuidade, pela forma do variograma quando para *h=0* $\gamma(h)$ já apresenta algum valor (*efeito pepita/nugget*); pode ser atribuído à erros de medição ou ao fato de que os dados não foram coletados a intervalos suficientemente pequenos para mostrar o comportamento espacial subjacente do fenômeno em estudo.

1. **semivariograma experimental**

- **mínimo de 30 pares**
- **remoção de valores anômalos**
- **maior Δh, a metade da maior distância existente entre os pontos.**
- **grau de casualidade dos dados, E = Co/C**

  E<0,15: componente aleatória pequena

  0,15 < E < 0,30: componente aleatória significante

  E > 0,30: componente aleatória muito significativa

  - **extremo do grau de casualidade é o modelo de pepita pura, onde não ocorre covariância entre os valores e a análise semivariográfica não se aplica**
- **iniciar com semivariograma omnidirecional**

2. **semivariograma teórico**

- **modelagem: processo que envolve várias tentativas e no qual a experiência pesa muito**
- **pode-se optar por um ajuste manual, mais sujeito à erros, ou com o auxílio de algoritmos**

3. ajuste do variograma experimental a um modelo variográfico teórico

- **Comparação visual**

- **Técnicas de ajuste automático:**
  - **Método dos mínimos quadrados**
  - **Critério AIC (Akaike Information Criterion)**
  - **Critério Cressie**
  - **Critério Variowin, etc.**

  - **Validação cruzada**

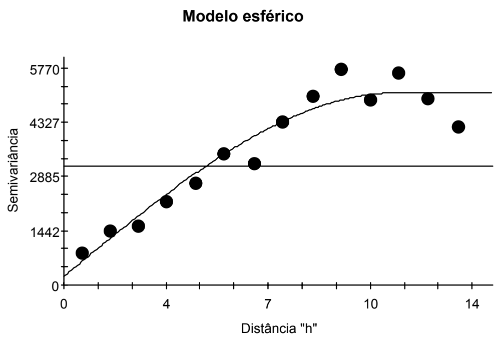
Modelos teóricos de variogramas

- **Modelos com soleira**
  - **modelo esférico:**

$\gamma(h) = C[((3/2)(h/a)) - ((1/2)(h^3/a^3))]$, **quando *h* < *a***

$\gamma(h) = C$, **quando *h* >= *a*,**

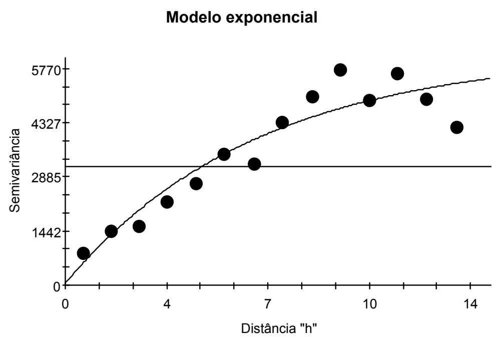
a inclinação da tangente junto a origem *(h=0)* é *3C/2a*; modelo mais comum



  - **modelo exponencial**

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

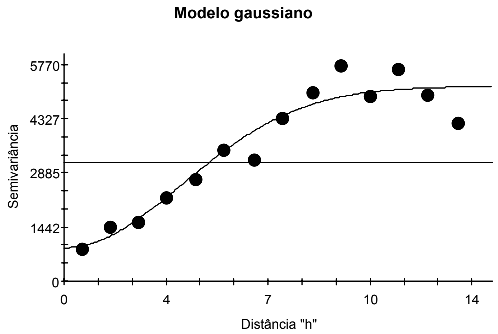
a inclinação da tangente junto a origem é *C/a; C* é a assíntota de uma curva exponencial e pode ser equalizada junto à soleira.



- **modelo gaussiano**

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

curva parabólica junto a origem e a tangente nesse ponto é horizontal; indica pequena variabilidade para curtas distâncias



- ***Modelos sem soleira***
  - **modelo a potência**

$\gamma\,(h)=Ch^n$, com n entre 0 e 2;

quando n = 1: modelo linear;

o modelo mais simples.