INSTITUTO SUPERIOR DE AGRONOMIA
Universidade de Lisboa

Departamento de Ciências e Engenharia de Biossistemas
Agrupamento de Matemática

Matemática II

Introdução à Geostatística

2013/2014

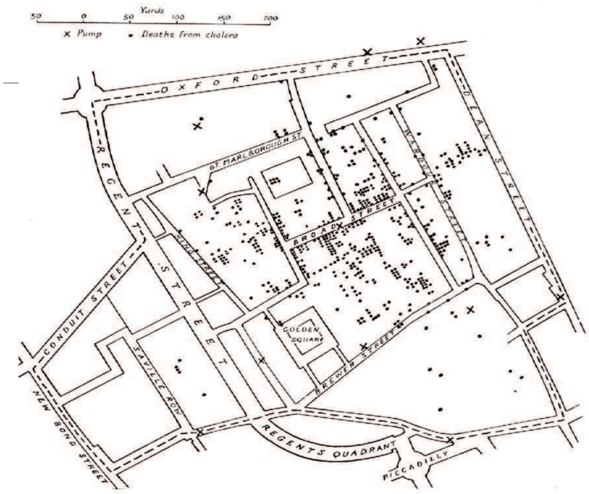
(F. Valente e M. Mesquita)

Análise espacial de dados

- Estudar e modelar fenómenos que se distribuem no espaço.
- Obter modelo inferencial que considere o relacionamento espacial presente no fenómeno em estudo.
- **Objectivo**: caracterização da distribuição espacial de variáveis que apresentam uma certa estrutura no espaço e quantificação da incerteza ligada ao fenómeno espacial em estudo.

Análise espacial de dados

- **Exemplos:**
  - modelação de recursos geológicos (concentração de metais em jazidos minerais, qualidade de águas subterrâneas),
  - estudos pedológicos (distribuição superficial de características do solo),
  - ecologia (distribuição de espécies vegetais ou animais, padrão espacial de aves de migratórias),
  - ambiente (ordenamento do território, estudos de poluição do solo, da água ou do ar).
- Exemplo pioneiro:
  - Estudo efectuado por John Snow em meados do século XIX, durante uma epidemia de cólera em Londres, para avaliar a relação entre a distribuição dos óbitos por cólera e a localização das bombas de água que abasteciam a cidade.



Tipos de dados em análise espacial

- **Padrões pontuais** – dados provenientes de fenómenos que se localizam em pontos no espaço.
  - Exemplos: localização de espécies vegetais, ocorrência de doenças;
- **Superfícies contínuas (dados geoestatísticos)** – dados provenientes de fenómenos que se distribuem de forma contínua no espaço.
  - Exemplos: dados geológicos, topográficos, ecológicos, climatológicos, pedológicos, hidrológicos.
- **Dados de áreas** – dados associados a unidades de análise, habitualmente delimitadas por polígonos fechados onde se supõe haver homogeneidade.
  - Exemplos: dados de censos, estatísticas de saúde.

Objectivos específicos

- Análise de padrões de pontos – o objectivo é estudar a distribuição espacial (localização) dos acontecimentos estudados (padrão aleatório, aglomerado ou regularmente distribuído).
- Análise de dados geoestatísticos – o objectivo é reconstruir a superfície contínua com base num conjunto finito de dados observados na região de estudo.
- Análise de dados de áreas – o objectivo é modelar o padrão espacial dos dados (áreas diferenciadas vs. continuidade espacial) e estabelecer associações com outras variáveis.

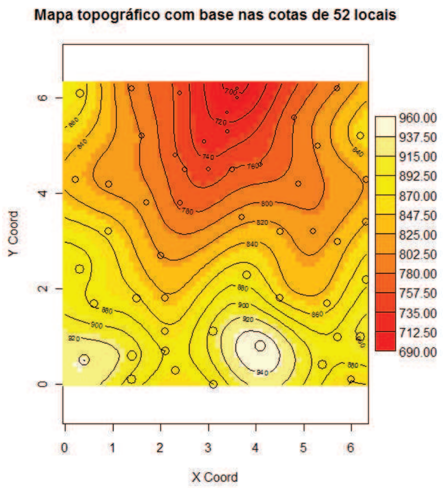
Dados geoestatísticos

- Tipicamente, os dados geoestatísticos podem ser representados por

$$(x_i, z_i), \quad i=1,\ldots,n$$

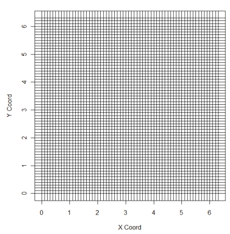
em que x_i identifica uma localização espacial (a 1, 2 ou 3 dimensões) e z_i é o valor da variável em estudo (variável resposta) associado à localização x_i.

- Exemplos:
  - cotas de 52 locais numa dada região; objectivo: construir um mapa de altitudes para toda a região;
  - conteúdo em cálcio de 178 locais numa dada região; objectivo: construir um mapa da variação espacial do conteúdo de cálcio no solo nesta região.



Representação computacional de dados geográficos

- Como construir este tipo de mapas?
- Num computador, a representação de uma superfície contínua pode ser feita por discretização numa malha regular.



A cada célula vai estar associado um valor numérico da variável em estudo.

Como obter esse valor?

Modelação espacial de superfícies

- A superfície que representa o fenómeno em estudo pode ser obtida por modelação/interpolação espacial através de:

1) modelos determinísticos

- **de efeitos locais**
  admitindo que predominam os efeitos locais, cada ponto da superfície é estimado por interpolação das observações mais próximas;
- **de efeitos globais**
  supondo que a variação em larga escala é dominante, a superfície é aproximada por um ajustamento polinomial aos dados;

2) modelos estatísticos (krigagem),
cada ponto da superfície é estimado através de um estimador estatístico;

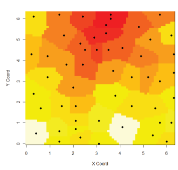
Modelos determinísticos locais

- O valor a estimar no ponto x_0 é calculado por combinação linear dos valores observados numa sua vizinhança.

- **Fórmula geral de interpolação:**
  $z(x_0)$ é o valor a estimar no ponto x_0,
  $z(x_j)$ é o valor observado em x_j
  w_j é um factor de ponderação

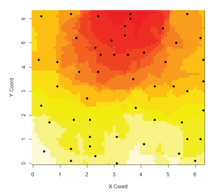
$$z(x_0) = \frac{\sum_{j=1}^{m} w_j\, z(x_j)}{\sum_{j=1}^{m} w_j}$$

- O valor a atribuir a w_j depende do método utilizado:
  interpolação por vizinho mais próximo
  (ou método dos polígonos de influência)
  factor de ponderação 1 ($w_j = 1$)
  para a observação $z(x_j)$ mais próxima de x_0
  e zero para as restantes observações.

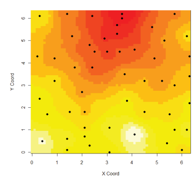


Modelos determinísticos locais (cont.)

interpolação por média simples
o valor de $z(x_0)$ é dado pela média aritmética dos valores $z(x_j)$ de m observações vizinhas (factor de ponderação 1 para as m observações vizinhas e zero para as restantes).

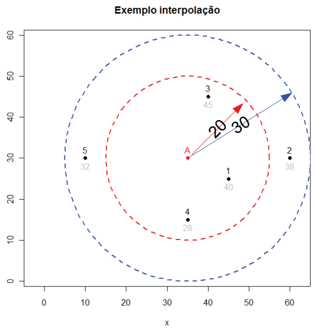


interpolação pelo inverso da potência das distâncias
o factor de ponderação é dado por $w_j = (1/d_j)^k$, em que d_j a distância euclideana do ponto x_j ao ponto x_0, para m observações vizinhas, e zero para as restantes.



Exemplo:

Ponto	x	y	z	distância
1	45	25	40	11.18
2	60	30	38	25.00
3	40	45	45	15.81
4	35	15	28	15.00
5	10	30	32	25.00
A	35	30	?	0.00



Principais vantagens e desvantagens dos modelos determinísticos

- Vantagens:
  - fáceis de implementar (estão disponíveis em muitas aplicações computacionais);
  - razoavelmente fieis aos valores observados;
  - úteis para uma rápida visualização dos dados.
- Desvantagens:
  - métodos que têm por base critérios estritamente geométricos que podem ser contraditórios com a estrutura espacial do fenómeno (por exemplo, existência de direcções privilegiadas);
  - não estimam valores da variável em estudo fora da gama de valores observados;
  - não avaliam a incerteza associada à caracterização do fenómeno espacial.

Modelos geoestatísticos

- Procuram ultrapassar as limitações dos métodos determinísticos
- Têm em consideração:

  1) a <u>estrutura espacial</u> da grandeza em estudo,

  em geral, os fenómenos distribuem-se no espaço de uma forma não aleatória;

  2) a <u>avaliação da incerteza</u> associada à caracterização do fenómeno espacial,

  a informação disponível (o conjunto de observações) é em geral escassa pelo que a modelação probabilística vai permitir a inferência espacial em pontos não amostrados e a quantificação da incerteza associada a essa estimação.

Conceitos básicos

- **Dependência espacial** – a maior parte das ocorrências apresentam entre si uma relação que depende da distância entre elas.

  Exemplos: se existe poluição num dado local é provável que locais próximos também estejam poluídos; se a presença de uma árvore adulta inibe o desenvolvimento de outras, essa inibição irá diminuir com a distância.
- **Autocorrelação espacial** – medida que avalia a dependência espacial; o prefixo *auto* indica que a correlação é avaliada para a mesma variável, mas medida em locais distintos.

  Exemplo de indicadores de autocorrelação espacial: covariância, coeficiente de correlação e variograma.

Conceitos básicos (cont.)

- **Variável aleatória $Z(x_i)$:**

  O conjunto de dados observados nos locais x_i, $i = 1, 2, \ldots, n$, $\{z(x_1), z(x_2), \ldots, z(x_n)\}$, é interpretado como uma realização de um conjunto de variáveis aleatórias $\{Z(x_1), Z(x_2), \ldots, Z(x_n)\}$ correlacionadas entre si.
  - Cada uma destas variáveis tem

    média $E[Z(x_i)] = \mu(x_i)$

    variância $Var[Z(x_i)] = E[Z^2(x_i)] - \mu^2(x_i)$
  - Entre quaisquer duas variáveis aleatórias pode definir-se

    covariância $C[Z(x_i), Z(x_j)] = E[Z(x_i)Z(x_j)] - \mu(x_i)\mu(x_j)$

    coeficiente de correlação $\rho = \dfrac{C[Z(x_i), Z(x_j)]}{\sqrt{Var[Z(x_i)]Var[Z(x_j)]}}$

    variograma $\gamma[Z(x_i), Z(x_j)] = \frac{1}{2} Var[(Z(x_i) - Z(x_j))]$

Conceitos básicos (cont.)

- Como do conjunto de v.a.s $\{Z(x_i)\}_{i=1}^{n}$ só se conhece uma realização $\{z(x_i)\}_{i=1}^{n}$, o conjunto de dados experimentais, é impossível determinar qualquer parâmetro estatístico.
- <u>Solução geoestatística</u>: assumir várias hipóteses que permitam fazer inferência, com a (pouca) informação disponível:

  1) Hipótese de estacionariedade da média – admite-se que todas as v.a. têm a mesma média $E[Z(x_i)] = \mu, \forall i.$

  2) Hipótese de estacionariedade da covariância e do variograma – a covariância e o variograma entre duas variáveis aleatórias só depende do vector distância $\boldsymbol{h}$ que as separa e é independente da sua localização,

  $$C[Z(x_i), Z(x_j)] = C[Z(x_i), Z(x_i + \boldsymbol{h})] = C(\boldsymbol{h})$$
  $$\gamma[Z(x_i), Z(x_j)] = \gamma[Z(x_i), Z(x_i + \boldsymbol{h})] = \gamma(\boldsymbol{h})$$

Consequências da estacionaridade

- A variância das v.a.s $Z(x_i)$ é constante,

$$Var[Z(x_i)] = C(0) = \sigma^2, \forall i$$

- O coeficiente de correlação também só depende do vector distância $\boldsymbol{h}$,

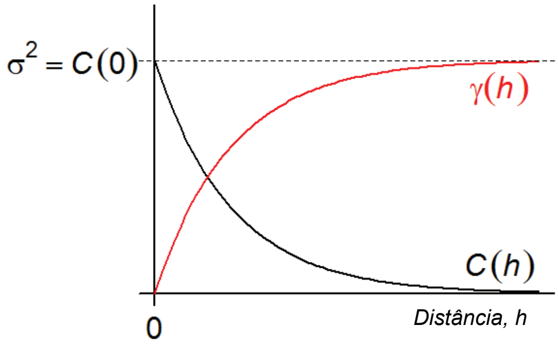
$$\rho(\boldsymbol{h}) = \frac{C(\boldsymbol{h})}{\sigma^2} = \frac{C(\boldsymbol{h})}{C(0)}$$

- O variograma pode escrever-se como

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

$$\gamma(\boldsymbol{h}) = \sigma^2(1 - \rho(\boldsymbol{h})) = C(0) - C(\boldsymbol{h})$$

Variograma e covariância espacial



Estacionaridade

- Com as hipóteses de estacionaridade, a média, o variograma e a covariância passam a ser independente da localização x_i (e x_j) e podem ser estimadas pelos estimadores

$$\mu* = \frac{\sum_{i=1}^{n} Z(x_i)}{n}$$

$$\gamma^*(\boldsymbol{h}) = \frac{\sum_{k=1}^{n(\boldsymbol{h})} [Z(x_k) - Z(x_k + \boldsymbol{h})]^2}{2n(\boldsymbol{h})}$$

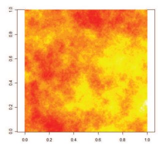
$$C^*(\boldsymbol{h}) = \frac{\sum_{k=1}^{n(\boldsymbol{h})} [Z(x_k) Z(x_k + \boldsymbol{h}) - \mu*(x_k)\mu*(x_k + \boldsymbol{h})]}{n(\boldsymbol{h})}$$

em que $\mu*(x_k) = \frac{1}{n(\boldsymbol{h})} \sum_{k=1}^{n(\boldsymbol{h})} Z(x_k)$, $\mu*(x_k + \boldsymbol{h}) = \frac{1}{n(\boldsymbol{h})} \sum_{k=1}^{n(\boldsymbol{h})} Z(x_k + \boldsymbol{h})$

e $n(\boldsymbol{h})$ é o número de pares de pontos (x_k, x_k+$\boldsymbol{h}$) para cada valor $\boldsymbol{h}$.

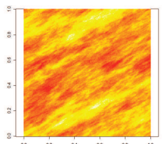
Isotropia

- Quando para além da estacionaridade, a covariância e o variograma têm o mesmo comportamento em todas as direcções diz-me que há **isotropia.** Isto é, C($\boldsymbol{h}$) e γ($\boldsymbol{h}$) só dependem de $h = ||\boldsymbol{h}||$, a distância euclediana entre as localizações.
- Caso contrário, quando a estrutura de correlação varia, quer com a distância h, quer com a direcção, o fenómeno é **anisotrópico**.



Isotropia

anisotropia



Modelos geoestatísticos: aspectos fundamentais

- Os valores da amostra são realizações de variáveis aleatórias localizadas espacialmente numa região A.
- A correlação entre estas v.a.s, medida pela covariância ou variograma, não depende da sua localização mas unicamente do vector $\boldsymbol{h}$ e avalia a continuidade espacial/dispersão da variável em estudo na região A.
- Um valor não amostrado em x_0 é também considerado uma v.a. $Z(x_0)$ que poderá ser estimado através de um novo estimador $Z^*(x_0)$.

Fases de um estudo geoestatístico

- **Análise exploratória dos dados**
  - representação visual dos dados em gráficos e mapas;
  - estatística descritiva: medidas de localização e dispersão, histograma, caixa-de-bigodes, "outliers";
  - identificação de padrões de dependência espacial no fenómeno em estudo, procura de regiões não homogéneas.
- **Análise estrutural dos dados**
  - estudo da continuidade espacial: determinação do variograma experimental e sua modelação.
- **Estimação**
  - determinação da variável em estudo em pontos não amostrados: krigagem

Análise exploratória dos dados

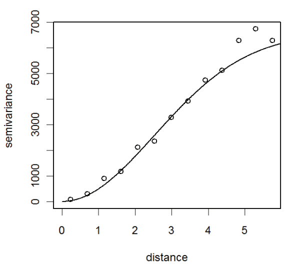
- Descrição univariada – com o objectivo de avaliar a dispersão da variável (atributo) em estudo;
- Descrição bivariada – com o objectivo de estudar o comportamento conjunto de dois atributos, ou o mesmo atributo medido em localizações distintas;
- Descrição espacial – com o objectivo de visualizar o modo como o atributo se dispersa no espaço:
  - localização espacial dos valores extremos (valores extremos distribuem-se na região em estudo vrs. valores muito altos/baixos concentrados);
  - anisotropia, anomalias, descontinuidades,...

Análise estrutural dos dados

- **Objectivo**: caracterização e quantificação do modo como se dispersa espacialmente a variável em estudo.
  - Esta etapa é a base dos processos de inferência/estimação espacial.
- **Ferramenta básica**: variograma
  - Representa a taxa média de variação do fenómeno em estudo com a distância;
  - Descreve o padrão de variação espacial;
  - Resume a informação num ponto, a partir dos valores observados em pontos próximos.

Variograma

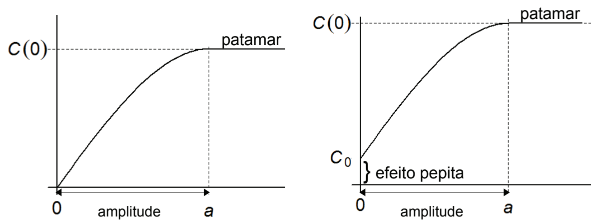
- **Variograma experimental –** construído com os dados da amostra
- **Variograma teórico –** modelo teórico que relaciona a variância entre dois pontos distanciados de *h* unidades



- $\gamma^*(h)=\frac{\sum_{k=1}^{n(h)}\left[z(x_k)-z(x_k+h)\right]^2}{2n(h)}$
- modelo

Características do variograma

- Para fenómenos estacionários, o variograma apresenta as seguintes características:
  1) $\gamma(0) = 0$, mas pode ser descontínuo na origem;
  2) é uma função crescente;
  3) quando $h \rightarrow \infty$, $\gamma(h) \rightarrow C(0)$



Patamar, amplitude e efeito pepita

- **Patamar –** representa a variabilidade da variável *Z*.
- **Amplitude –** mede a distância a partir da qual os valores da variável *Z* deixam de estar correlacionados.
- **Efeito pepita –** mede fundamentalmente duas parcelas da variabilidade total do fenómeno em estudo:
  - variação espacial numa escala inferior à distância mínima entre pontos amostrados;
  - variabilidade devida a erros de medição.

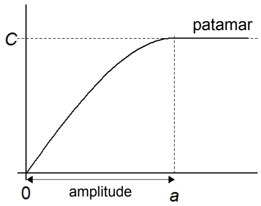
Modelos de variogramas

- Antes do variograma poder ser utilizado, tem de ser ajustado aos dados um modelo teórico (uma função matemática).
  - A função a ajustar tem de satisfazer certas condições.
- Na prática, a modelação do variograma limita-se ao uso de um conjunto restrito de funções que cobrem a generalidade das situações de dispersão dos fenómenos espaciais que nos interessam estudar:
  - **modelo esférico**,
  - **modelo exponencial**,
  - **modelo Gaussiano**.

Modelo esférico

- Um dos modelos mais usuais em geoestatística.

$$\gamma(h)=\begin{cases} C\left[1.5\frac{h}{\phi}-0.5\left(\frac{h}{\phi}\right)^3\right], & 0\le h\le\phi \\ C\ , & h>\phi. \end{cases}$$

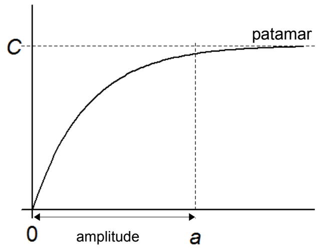


Função de dois parâmetros:
C - patamar
$\phi = a$ - amplitude

Modelo exponencial

- Função de 2 parâmetros: C - patamar e $\phi = a/3$.
  Neste caso, o valor da amplitude a é a distância em que o modelo atinge 95% do patamar, isto é, $\gamma(a) = 0.95C(0)$.

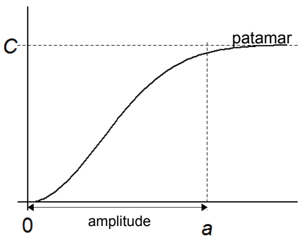
$$\gamma(h)=C\left(1-e^{-h/\phi}\right),\quad h\ge 0$$



Modelo Gaussiano

- Função de 2 parâmetros: C - patamar e $\phi = a/\sqrt{3}$.
  Também aqui, o valor da amplitude a é a distância em que o modelo atinge 95% do patamar, isto é, $\gamma(a) = 0.95C(0)$.

$$\gamma(h)=C\left(1-e^{-(h/\phi)^2}\right),\quad h\ge 0$$



Contrariamente aos modelos anteriores, este modelo tem um crescimento lento junto à origem.

Ajustar um modelo ao variograma experimental

- Processo que envolve várias tentativas e no qual a experiência e o conhecimento do fenómeno é muito importante.
- Pode fazer-se o ajustamento de forma manual, analisando o resultado visualmente, ou utilizar técnicas de ajustamento automático.
- A forma analítica de um modelo não é muito importante desde que as principais características do fenómeno sejam respeitadas, nomeadamente:
  - o efeito pepita,
  - o declive na origem,
  - a amplitude,
  - o patamar,
  - as anisotropias.

Variograma ajustado

- Com o comando `eyefit` do *package* **geoR** é possível visualizar no variograma experimental alguns modelos teóricos, nomedamente, os modelos referidos anteriormente.

```
> eyefit(exemplo.variog)
cov.model  sigmasq  phi tausq kappa kappa2 practicalRange
exponential   9.52 1.06  1.19  <NA>   <NA>          3.18
```

Neste exemplo, o modelo escolhido foi o exponencial com parâmetros $C = 9.52$ (`sigmasq`) e $\phi = 1.06$ (`phi`), efeito pepita $C_0 = 1.61$ (`tausq`) e amplitude $a = 3.18$ (`practicalRange`)
O variograma resultante pode ser descrito por

$$\gamma_{ajust}(h)=C_0+\gamma(h),\quad h>0$$

Estimação

- **Krigagem:**
  – método de estimação que deve o seu nome ao trabalho pioneiro de D.G. Krige (1951).
  – método que utiliza o "melhor" estimador linear não-enviezado (centrado).

- A krigagem pode assumir diversas formas. A mais usual é designada por **krigagem normal** ou ordinária (*ordinary kriging*): assume-se que o fenómeno em estudo é estacionário mas de média μ desconhecida.

Estimação

- **Objectivo:** estimação do valor da variável Z num ponto x_0 não amostrado, $Z(x_0)$.
- **Estimador a utilizar:** $Z^*(x_0)$, que deverá
  1) ser combinação linear dos valores conhecidos $Z(x_i)$, $i = 1, \ldots, n$
  $$Z^*(x_0)=\sum_{i=1}^{n} \lambda_i Z(x_i)$$
  2) com média do erro de estimação (ε) nula
  $$E[\varepsilon(x_0)]=E[Z^*(x_0)-Z(x_0)]=0$$
  3) com variância do erro de estimação mínima
  $$\min \quad Var[\varepsilon(x_0)]=E[(Z^*(x_0)-Z(x_0))^2]$$

Estimação

- Considerando as hipóteses de estacionaridade e a primeira condição,
$$E[Z^*(x_0)]=E[Z(x_0)]=\mu$$
$$E[Z^*(x_0)]=E[\sum_{i=1}^{n} \lambda_i Z(x_i)]=\mu\sum_{i=1}^{n} \lambda_i$$
a segunda condição é satisfeita se $\boxed{\sum_{i=1}^{n} \lambda_i=1}$

e a terceira pode ser reescrita em termos da covariância ou do variograma
$$\min \quad Var[\varepsilon(x_0)]=C(0)+\sum_i\sum_j \lambda_i\lambda_j C(x_i,x_j)-2\sum_i \lambda_i C(x_i,x_0)$$
$$\min \quad Var[\varepsilon(x_0)]=2\sum_i \lambda_i \gamma(x_i,x_0)-\sum_i\sum_j \lambda_i\lambda_j \gamma(x_i,x_j)$$

Krigagem

- O problema de determinar $\lambda_1, \lambda_2, \ldots, \lambda_n$ que
$$\min \quad Var[\varepsilon(x_0)] \quad \text{s.a} \quad \sum_{i=1}^{n} \lambda_i=1$$
pode ser resolvido através do sistema linear de $(n+1)$ equações a $(n+1)$ incógnitas – **sistema de krigagem**
$$\begin{cases}\lambda_1 C(x_1,x_1)+\lambda_2 C(x_1,x_2)+\ldots+\lambda_n C(x_1,x_n)+\alpha=C(x_1,x_0)\\ \lambda_1 C(x_2,x_1)+\lambda_2 C(x_2,x_2)+\ldots+\lambda_n C(x_2,x_n)+\alpha=C(x_{2,}x_0)\\ \vdots\\ \lambda_1 C(x_n,x_1)+\lambda_2 C(x_n,x_2)+\ldots+\lambda_n C(x_n,x_n)+\alpha=C(x_n,x_0)\\ \lambda_1+\lambda_2+\ldots+\lambda_n=1\end{cases}$$

em que α é mais uma incógnita (o parâmetro de Lagrange), introduzida pelo método utilizado na resolução deste problema (método dos multiplicadores de Lagrange).

Krigagem

- O sistema de krigagem pode também ser descrito em função do variograma
$$\begin{cases}\lambda_1 \gamma(x_1,x_1)+\lambda_2 \gamma(x_1,x_2)+\ldots+\lambda_n \gamma(x_1,x_n)-\alpha=\gamma(x_1,x_0)\\ \lambda_1 \gamma(x_2,x_1)+\lambda_2 \gamma(x_2,x_2)+\ldots+\lambda_n \gamma(x_2,x_n)-\alpha=\gamma(x_2,x_0)\\ \vdots\\ \lambda_1 \gamma(x_n,x_1)+\lambda_2 \gamma(x_n,x_2)+\ldots+\lambda_n \gamma(x_n,x_n)-\alpha=\gamma(x_n,x_0)\\ \lambda_1+\lambda_2+\ldots+\lambda_n=1\end{cases}$$

- O valor mínimo da variância do erro de estimação (ε) é

Variância de krigagem
$$Var[\varepsilon(x_0)]=C(0)-\sum_{i=1}^{n} \lambda_i C(x_i,x_0)-\alpha$$
$$Var[\varepsilon(x_0)]=\sum_{i=1}^{n} \lambda_i \gamma(x_i,x_0)+\alpha$$

Sistema de krigagem: notação matricial

- O sistema anterior pode ser escrito sob a forma matricial $\boldsymbol{K}\,\boldsymbol{x} = \boldsymbol{M}$, em que
  - $\boldsymbol{K}$ é uma matriz quadrada de ordem $(n+1)$ cujos elementos são $\gamma(x_i, x_j) = \gamma_{ij}$, o variograma entre os pontos observados x_i e x_j,
  - $\boldsymbol{x}$ é o vector das incógnitas $(\lambda_1, \lambda_2, \ldots, \lambda_n, \alpha)$,
  - $\boldsymbol{M}$ é um vector cujos elementos são $\gamma(x_i, x_0) = \gamma_{i0}$, o variograma entre o ponto observado x_i e o ponto a estimar x_0.

$$\boldsymbol{K}\,\boldsymbol{x}=\boldsymbol{M}$$

$$\begin{bmatrix}\gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} & 1\\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} & 1\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} & 1\\ 1 & 1 & \ldots & 1 & 0\end{bmatrix}\begin{bmatrix}\lambda_1\\ \lambda_2\\ \cdots\\ \lambda_n\\ \alpha\end{bmatrix}=\begin{bmatrix}\gamma_{10}\\ \gamma_{20}\\ \cdots\\ \gamma_{n0}\\ 1\end{bmatrix}$$

Estimação e erro

- Obtida a solução do sistema de krigagem,
  – a estimativa do valor da variável Z num ponto x_0 não amostrado, $z^*(x_0)$, é calculada com base no conjunto de dados experimentais $\{z(x_1), z(x_2), \ldots, z(x_n)\}$ como
  $$z^*(x_0)=\sum_{i=1}^{n} \lambda_i z(x_i)$$
  – a estimativa da variância do erro de estimação é calculada como
  $$Var[\varepsilon(x_0)]=\sum_{i=1}^{n} \lambda_i \gamma(x_i,x_0)+\alpha$$

Bibliografia:

- Armstrong, M. (1998) Basic Linear Geostatistics, Springer-Verlag.
- Diggle, P. e Ribeiro Jr., P (2007) Model-based Geostatistics, Springer.
- Soares, A. (2000) Geoestatística para as Ciências da Terra e do Ambiente, IST Press.