Matemática II

Exercícios de Geoestatística

Fernanda Valente e Marta Mesquita

INSTITUTO SUPERIOR DE AGRONOMIA

- 2013/2014 -

III - Introdução à Geoestatística

1. Pretende-se estudar o teor de argila (%) na Terra Grande da Tapada da Ajuda. Representou-se a área em estudo por um rectângulo 250 m × 150 m. Escolheram-se aleatoriamente 20 pontos para recolher amostras de solo onde se determinou o teor de argila. Os resultados obtidos (x e y, coordenadas dos pontos amostrados, em m e z, teor de argila do solo, em %) estão no quadro seguinte e encontram-se disponíveis na *data frame* `argila`, no ficheiro `argila20.RData`). As distâncias de cada ponto amostrado aos pontos $A = (75, 40)$, $B = (120, 60)$ e $C = (100, 120)$ são também apresentadas.

ponto	x	y	z	distância do ponto i a		
i				A	B	C
1	210.2	63.9	47.3	137.3	90.3	123.7
2	40.8	130.6	46.3	96.8	106.1	60.1
3	175.5	34.3	42.5	100.7	61.2	114.2
4	11.8	12.1	47.6	69.1	118.3	139.4
5	61.3	114.7	42.2	75.9	80.2	39.1
6	115.5	34.9	46.1	40.8	25.5	86.5
7	52.8	101.6	43.4	65.5	79.0	50.7
8	248.6	53.3	44.3	174.1	128.8	162.9
9	173.2	110.8	44.8	121.1	73.6	73.8
10	75.0	123.8	42.9	83.8	78.1	25.3
11	116.9	144.3	46.4	112.4	84.4	29.6
12	147.2	10.3	43.9	78.1	56.7	119.4
13	166.1	15.6	45.0	94.3	64.0	123.6
14	225.8	148.8	44.1	186.0	138.1	129.1
15	36.3	66.6	43.1	47.0	84.0	83.1
16	12.6	24.0	47.9	64.4	113.3	129.8
17	18.2	50.8	46.2	57.8	102.2	107.1
18	171.2	10.0	42.8	100.8	71.6	131
19	162.5	68.6	46.2	92.1	43.4	80.9
20	167.6	100.6	45.1	110.7	62.6	70.3

(a) Efectue uma análise descritiva dos dados de teor de argila, focando os seguintes aspectos: representação gráfica, medidas de localização e de dispersão e forma da distribuição dos dados.

(b) Utilize a função `plot` do R para visualizar a localização dos pontos onde foram recolhidas as amostras de solo.

(c) Estime o teor de argila nos pontos A, B e C, utilizando os seguintes métodos deterministicos:

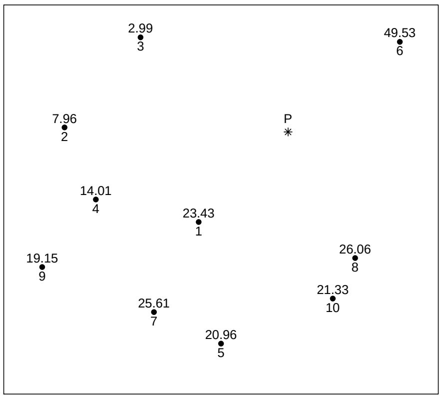
i. interpolação por vizinho mais próximo,

ii. interpolação por média simples, numa vizinhança com 50 m de raio,

iii. interpolação pelo inverso da potência das distâncias com $k = 1$ e $k = 2$, numa vizinhança com 50 m de raio.

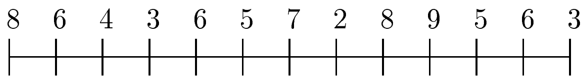
Repita as duas subalíneas anteriores considerando, desta vez, uma vizinhança com 75 m de raio. Compare e comente os resultados obtidos pelos três métodos.

2. O gráfico seguinte contem a localização de 10 locais de um parque florestal urbano onde foi amostrado o teor em potássio do solo. No quadro apresenta-se o teor obtido em cada local e a distância do ponto P a cada uma desses locais. Estime o teor em potássio no ponto P utilizando o método de interpolação pelo inverso das distâncias, numa vizinhança de raio 35.



local	teor em potássio (em $g\,\mathrm{K}^+\ m^{-2}$)	distância ao ponto P
1	23.43	28.28
2	7.96	50.01
3	2.99	39.12
4	14.01	45.54
5	20.96	49.34
6	49.53	32.02
7	25.61	50.00
8	26.06	31.76
9	19.15	62.65
10	21.33	38.33

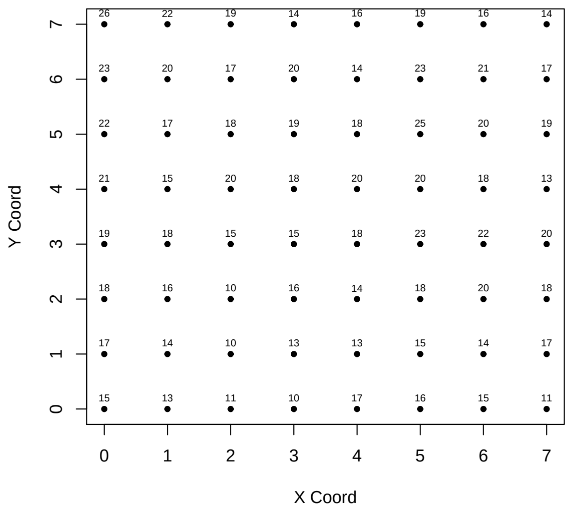
3. Considere o conjunto de 13 observações espaçadas de 5 m ao longo de uma linha (ver figura)



   (a) Determine o variograma experimental utilizando 20 m como a distância máxima entre pares de pontos.

   (b) Dê estimativas para os valores do efeito pepita e do patamar.

4. Considere os seguintes dados referentes a uma variável amostrada numa grelha regular de 1×1 unidades^2, numa área de 8×8 unidades^2 (dados disponíveis no objecto `ex3`, no ficheiro `GeoGrelha.RData`).



(a) Utilizando o R, faça uma análise exploratória destes dados.

(b) Utilizando as funções `variog` e `plot` do *package* **geoR**, determine o variograma experimental (omnidireccional) e faça a sua representação gráfica. Considere 4 unidades como a distância máxima entre pares de pontos.

(c) Considere a seguinte informação resultante do ajustamento de um modelo ao variograma experimental:

```
covariance model is: spherical
parameter estimates:
  tausq sigmasq     phi
 1.3480 10.2486  3.4418
```

i. Escreva a expressão do modelo ajustado.

ii. Indique qual a amplitude deste modelo.

iii. Utilize a função `eyefit` para representar graficamente o variograma ajustado.

(d) Calcule a covariância espacial entre o ponto amostrado $(5, 1)$ e os pontos $A = (4.5, 0.5)$ e $B = (6.5, 5.5)$. Comente.

5. Com o objectivo de estudar a estrutura espacial de um conjunto de dados, foi ajustado ao variograma experimental (omnidireccional) um modelo teórico, tendo-se obtido a seguinte informação

```
covariance model is: gaussian
parameter estimates:
  tausq sigmasq     phi
 0.3484  0.7281   ?????
Practical Range with cor=0.05 for asymptotic range: 0.952763
```

(a) Determine o parâmetro ϕ (`phi`) do modelo ajustado.

(b) Diga como se denomina e o que representa o parâmetro de valor 0.3484.

(c) Escreva a expressão da função variograma ajustada e a correspondente função covariância.

(d) Determine o coeficiente de correlação entre dois pontos que distam entre si de 0.8 unidades.

6. A um variograma experimental foi ajustado um modelo teórico esférico com patamar igual a 0.8 e amplitude igual a 1 unidade. Sabendo que o valor estimado do variograma entre pontos que distam entre si 0.3 unidades é 0.53, determine:

(a) o efeito pepita;

(b) a covariância entre pontos que distam entre si 0.3 unidades.

7. Considere o variograma $\gamma(\mathbf{h})$ e a covariância $C(\mathbf{h})$ de um fenómeno espacial estacionário Z, com $\mathbf{h}$ o vector distância entre dois pontos. Utilizando a definição de variograma, mostre que $\gamma(\mathbf{h}) = C(0) - C(\mathbf{h})$.

8. Considere os pontos $x_1 = (0,0)$ e $x_2 = (100,0)$ e $Z(x_i)$, $i = 1,2$ variáveis aleatórias satisfazendo as hipóteses de estacionaridade com $Var[Z(x_i)] = 2.5$, $i = 1,2$. Admita que o variograma (omnidireccional)de $Z(x)$ é modelado pelo modelo exponencial com $\phi = 200$. Considere o estimador de krigagem normal $Z^*(x) = \lambda_1 Z(x_1) + \lambda_2 Z(x_2)$.

   (a) Calcule a variância de Z^*, em função de λ_1 e λ_2.

   (b) Mostre que a solução do sistema de krigagem normal é dada por

   $$\lambda_1 = \frac{1}{2} + \frac{\gamma_{20} - \gamma_{10}}{2\gamma_{12}}, \quad \lambda_2 = \frac{1}{2} + \frac{\gamma_{10} - \gamma_{20}}{2\gamma_{12}}, \quad \alpha = \frac{\gamma_{10} + \gamma_{20} - \gamma_{12}}{2},$$

   em que γ_{i0} é o variograma entre os pontos x_i e x_0 $(i = 1,2)$ e γ_{12} é o variograma entre os pontos observados x_1 e x_2.

   (c) Sabendo que $z(x_1) = 5$, $z(x_2) = 15$ e $x_0 = (70,0)$, calcule a estimativa de $Z(x_0)$, dada pelo estimador de krigagem normal, e a estimativa da variância do erro de estimação associado.

   (d) Repita a alínea anterior considerando $x_0 = (30,0)$.

   (e) Repita as alíneas (c) e (d) considerando agora o modelo esférico (com $\phi = 200$) para o variograma.

Soluções de alguns Exercícios

1. Ver script `ExercíciosGeoEst.R`

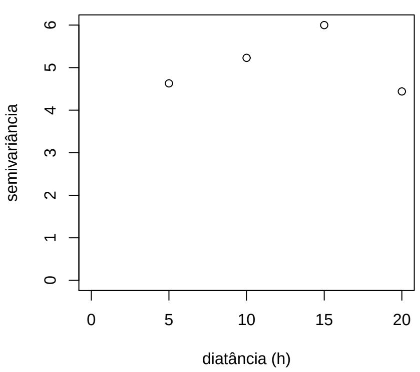
2. Os pontos 1, 6 e 8 são os únicos que se encontram numa vizinhança de raio 35 do ponto P. O valor estimado do teor em potássio no ponto P utilizando o método de interpolação pelo inverso das distâncias é

$$\widehat{\texttt{teor}}(P) = \frac{\frac{1}{d_1} \times \texttt{teor}_1 + \frac{1}{d_6} \times \texttt{teor}_6 + \frac{1}{d_8} \times \texttt{teor}_8}{\frac{1}{d_1} + \frac{1}{d_6} + \frac{1}{d_8}}$$

Logo,

$$\widehat{\texttt{teor}}(P) = \frac{\frac{1}{28.28} \times 23.43 + \frac{1}{32.02} \times 49.53 + \frac{1}{31.76} \times 26.06}{\frac{1}{28.28} + \frac{1}{32.02} + \frac{1}{31.76}} = 32.58g \text{ K}^+ \ m^{-2}$$

3. (a) $\gamma(5) = 4.63, \quad \gamma(10) = 5.23, \quad \gamma(15) = 6.0 \quad \gamma(20) = 4.44$



(b) O efeito pepita poderá ser cerca de 4 ou ter um valor inferior. Pela análise do variograma, o patamar pode ser um valor entre 5 e 6. Admitindo que o fenómeno em estudo é estacionário, uma estimativa do patamar pode também ser obtida pela variância dos valores observados $s^2 = 4.6$.

4. (a) Ver script `ExercíciosGeoEst.R`

(b) Ver script `ExercíciosGeoEst.R`

(c) i. $\gamma(h) = \begin{cases} 0 & , h = 0 \\ 1.3484 + 10.2486 \left[1.5 \frac{h}{3.4418} - 0.5 \left(\frac{h}{3.4418} \right)^3 \right] & , 0 < h \leq 3.4418 \\ 11.597 & , h > 3.4418 \end{cases}$

ii. 3.4418

iii. Ver script `ExercíciosGeoEst.R`

(d) $C(\| (5,1) - A \|) = 7.134723 \quad C(\| (5,1) - B \|) = 0$

5. (a) 0.550078

   (b) efeito pepita - representa a variabilidade devida a erros de medição e a variabilidade espacial numa escala inferior à distância mínima entre pontos amostrados.

   (c) $\gamma(h) = \begin{cases} 0 & , \ h = 0 \\ 0.3484 + 0.7281 \left[1 - e^{-\left(\frac{h}{0.55}\right)^2}\right] & , \ h > 0 \end{cases}$

   $C(h) = \begin{cases} 1.0765 & , \ h = 0 \\ 0.7281 \, e^{-\left(\frac{h}{0.55}\right)^2} & , \ h > 0 \end{cases}$

   (d) $\rho(0.8) = 0.0815$

6. (a) 0.1808 (b) 0.4508

7. Sendo Z um fenómeno espacial estacionário, quer a média quer a variância são constantes na região em questão, isto é,

$$E[Z(x_i)] = \mu, \ \forall i \quad \text{e} \quad Var[Z(x_i)] = C[Z(x_i), Z(x_i + \mathbf{h})] = C(0), \ \forall i,$$

e o variograma e a covariância só dependem do vector $\mathbf{h}$ entre os pontos. Logo,

$$\gamma(\mathbf{h}) \underset{(1)}{=} \frac{1}{2} Var[Z(x_i) - Z(x_i + \mathbf{h})] \underset{(2)}{=} \frac{1}{2}\Big(Var[Z(x_i)] + Var[Z(x_i + \mathbf{h})] - 2Cov[Z(x_i), Z(x_i + \mathbf{h})]\Big)$$

$$= \tfrac{1}{2}\Big(C(0) + C(0) - 2C(\mathbf{h})\Big) = C(0) - C(\mathbf{h})$$

(1) Definição de variograma (no formulário)
(2) Propriedade da variância (no formulário)

8. (a) $Var[Z^*(x)] = 2.5(\lambda_1^2 + \lambda_2^2) + 5e^{-0.5}\lambda_1\lambda_2$

   (c) $z^*(x_0) \simeq 11.98$, $\hat{Var}[\varepsilon(x_0)] \simeq 0.517$, com $\varepsilon(x_0) = Z^*(x_0) - Z(x_0)$ o erro de estimação no ponto x_0.

   (d) $z^*(x_0) \simeq 8.02$, $\hat{Var}[\varepsilon(x_0)] \simeq 0.517$

   (e) $x_0 = (70, 0):$ $z^*(x_0) \simeq 12.04$, $\hat{Var}[\varepsilon(x_0)] \simeq 0.815$
   $x_0 = (30, 0):$ $z^*(x_0) \simeq 7.96$, $\hat{Var}[\varepsilon(x_0)] \simeq 0.815$