--- Polycopie du cours Géostatistique ---

Master M2 Eau et Environnement

Enseignant : Pr. B. Azouzi

Partie 3

- **Introduction**
- **Variographie**
- **Estimation (Interpolation spatiale)**
- **Krigeage**
    - **(a) Krigeage simple**
    - **(b) Krigeage ordinaire**
    - **(c) Krigeage universel**
- **Conclusion**
- **TD (Exercices résolus)**

Introduction

L'étude géostatistique comprend deux étapes essentielles:

1. La caractérisation de la structure spatiale de la variable régionalisée par l'outil **Variogramme**
2. L'estimation de cette variable en utilisant un outil d'interpolation le **Krigeage**.

Première étape : Variographie (pour plus de détails voir Partie 2)

Dans la première étape on évalue la corrélation spatiale à l'aide du variogramme, qui est une mesure de la variabilité des données en fonction de la distance. On calcule le variogramme expérimental à l'aide de l'équation suivante:

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

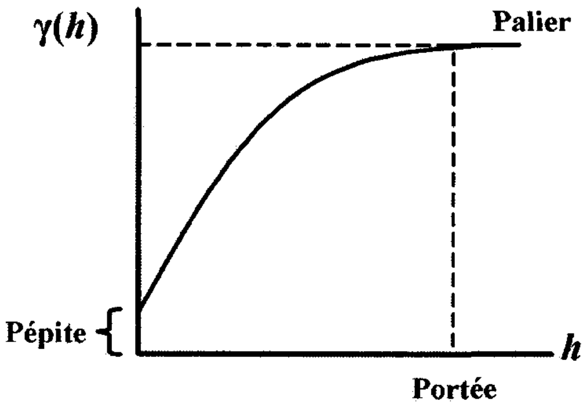
où h représente une distance donnée, z(x) est la valeur de la variable étudiée et N(h) le nombre de paires possibles pour la distance h. Ainsi, pour chaque distance h retenue, on calcule la valeur de $\gamma(h)$. Cette équation permet d'établir la différence moyenne entre deux points espacés d'une distance *h.*

Une fois le variogramme expérimental calculé, on essaie de lui ajuster un model qui rend compte des caractéristiques observées, c'est l'inférence du variogramme. Quatre principaux modèles peuvent lui être ajustés : exponentiel,gaussien, sphérique ou linéaire (pour plus de détails voir la partie 2).

Pour faire l'analyse spatiale d'une propriété, trois paramètres sont importants à considérer dans ce graphique. Ce sont **l'effet de pépite, le palier et la portée** :

- **L'effet de pépite** (nugget effect), c'est la valeur de γ (h) quand h=0, Il englobe l'incertitude due à l'échantillonnage et aux erreurs de mésure ou indique carrement une microrégionalisation (microvariabilité aux petits pas de distances). L'effet pépite doit etre estimé par extrapolation des premières valeurs du semivariogramme.
- **Le palier** (sill) est la valeur de semivariogramme qui correspond à la portée. Il correspond à la variance propre de l'échantillon. Le long du palier, γ(h) devient constant avec l'évolution de h.
- **La portée** (range) : La valeur de h pour laquelle γ(h) atteint le palier s'appelle la portée, elle indique la distance à laquelle le semivariogramme se stabilise autour d'une valeur limite. cette distance mesure la zone d'influence d'une information z(s).

En fait, au-delà de la distance h la variable aléatoire Z(s+h) est sans corrélation avec Z(s). la portée tend à augmenter avec la grandeur du jeu de donnée.

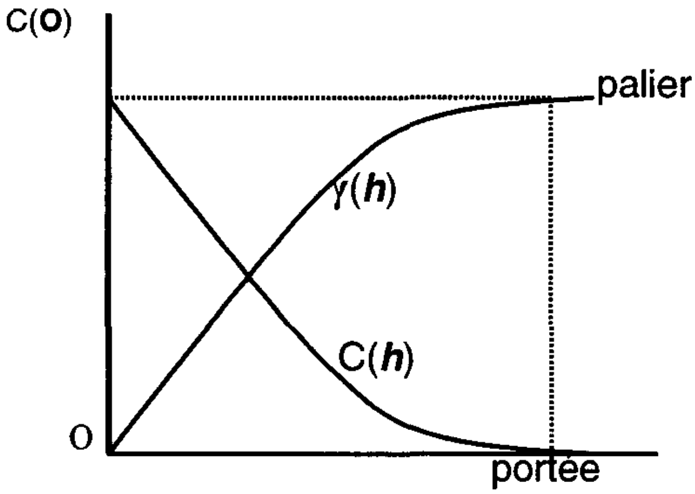


Paramètres du Variogramme

Relation entre covariance et variogramme

Lorsque le variogramme montre un palier, c'est-à-dire si $\lim_{h \to m} \gamma(h) = C$ où C est une constante, il existe une relation entre la valeur du variogramme pour la distance h et la covariance pour deux observations séparées de 1a distance h , cette relation est très importante, elle est continuellement utilisée en géostatistique

$$\gamma(h) = \sigma^2 - \text{cov}(h)$$



Covariance et variogramme sous l'hypothèse stationnaire

Deuxième étape : Estimation (Interpolation spatiale)

Dans la deuxième étape, on interpole les données observées en fonction de ce degré de dépendance (Variogrammes), processus qu'on appelle le krigeage proprement dit.

Il faut rappeler que l'interpolation spatiale se définie par la prévision (estimation) de la valeur d'une variable en un site à partir de valeurs mesurées en des sites voisins. Elle peut s'effectuer par une méthode **déterministe** ou **stochastique**.

Les polygones de Thiessen, la méthode de l'inverse des distances et les splines sont des exemples de méthodes déterministes (voir la partie 1).

Le Krigeage est une méthode d'interpolation stochastique (géostatistique) relativement récente pour étudier la variabilité spatiale. Elle est basée sur la théorie des variables régionalisées, développé par Krige (1951), Matheron (1965), Journel Huijbregts (1978), et Cressie (1991).

Le krigeage

Le Krigeage porte le nom de son précurseur, l'ingénieur minier sud-africain D.G. Krige. Dans les années 50, Krige a développé une série de méthodes statistiques empiriques afin de déterminer la distribution spatiale de minerais à partir d'un ensemble de forages. C'est cependant le français Matheron, (1963) qui a formalisé l'approche en utilisant les corrélations entre les forages pour en estimer la répartition spatiale. C'est lui qui a baptisé la méthode " Krigeage ". Il a aussi été le premier a utiliser le terme " géostatistiques " pour désigner la modélisation statistique de données spatiales.

Le krigeage est une méthode d'estimation d'un phénomène connu en un certain nombre de points et a l'avantage parmi d'autres techniques d'interpolation d'utiliser la structure spatiale de la variable pour l'estimation (Delhomme,1978). Ses avantages sont les suivants :

- Le krigeage intègre la connaissance émanat de l'analyse préalable de la structure par le variogamme ;
- Le krigeage est un interpolateur exact. Si un point d'échantillonnage coincide avec un nœud de grille,la solution du krigeage est égale à la valeur de l'échantillon ;
- Le krigeage livre une interpolation sur l'erreur d'estimation traduite par la variance de krigeage.

  Pour cela le krigeage est baptisé BLUE (Best Linear Unbaised Estimator).

Les différents types de krigeage :

Il existe trois types de krigeage: le krigeage simple, le krigeage ordinaire et le krigeage universel. En pratique le krigeage ordinaire est le plus fréquemment utilisé.

- Krigeage simple : il est utilisé pour une variable stationnaire de moyenne connue.
- Krigeage ordinaire : quand la variable stationnaire est de moyenne inconnue.
- Krigeage universel : pour une variable non-stationnaire (qui contient une tendance).

Krigeage simple

La théorie du krigeage a d'abord été développée dans un cadre stationnaire de second ordre. Sous cette hypothèse, le krigeage le moins complexe est celui dans lequel la stationnarité postulée est de deuxième ordre et l'espérance de la fonction aléatoire étudiée est supposée connue et constante sur tout le champ. Il s'agit du krigeage simple (Matheron, 1970).
Donc, quand on connaît la moyenne "m" d'un champ à estimer, on utilise le Krigeage simple comme un estimateur sans biais minimisant la variance d'estimation.

$$Zo = \sum \lambda\mathrm{i}\, Zi + \left(1 - \sum \lambda\mathrm{i}\right) m$$

Sous forme matricielle le système d'équation du KS s'écrit en termes de variogrammes :

$$\begin{vmatrix} \gamma_{11} & \gamma_{12} & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \gamma_{2n} \\ \gamma_{n1} & \gamma_{n2} & \gamma_{nn} \end{vmatrix} \begin{vmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_n \end{vmatrix} = \begin{vmatrix} \gamma_{1\,0} \\ \gamma_{2\,0} \\ \gamma_{n\,0} \end{vmatrix}$$

La variance d'estimation du krigeage simple :

$$\boldsymbol{\sigma^2 = \mathrm{Var}(z_0^* - z_0) = E|z_0^* - z_0| = Var(Zi) - \sum \Lambda_i\, \gamma(x_i - x_0)}$$ **variance d'estimation**

$$\boldsymbol{\sigma = \sqrt{var(zi) - \sum \Lambda_i\, \gamma(x_i - x_0)}}$$ **Ecart type d'estimation du krigeage simple**

Avec Var(Zi)= variance observée = Palier

Krigeage ordinaire

Dans le cadre de l'hypothèse de stationnarité, la méthode du krigeage stationnaire à moyenne inconnue dit **krigeage ordinaire** consiste à chercher le vecteur des coefficients λi , $i_{=1.......n}$ pour l'estimateur otimal où la variable spatiale Z^*_0 au point X_0 est estimée par la somme pondérée des autres valeurs échantillonnées Zi situées aux points X_i:

tel que : $$Z^*_0 = \sum \lambda_i \; Zi$$

Le krigeage ordinaire s'effectue dans deux conditions :

- la condition de non biais où l'espérance mathématique est supposée être constante
- la condition d'optimalité qui consiste à minimiser le mieux possible la variance d'estimation en ayant une espérance quadratique moyenne **E $[Z^*_0\text{-}Z_0]^2$** minimale

1- Condition de non biais : qui implique $\sum \lambda_i$=1

m=E $[Z^*_0]$= E $[Z_0]$ donc E $[Z^*_0\text{-}Z_0]$=0 , remplaçons Z^*_0 par sa valeur $Z^*_0 = \sum \lambda_i \; Zi$

E $[\sum \lambda_i \; Zi]$= m ======> $\sum \lambda_i$ **E [Zi]** =m

On a donc $\sum \lambda_i$ **m** = **m** ======> $\boxed{\sum \boldsymbol{\lambda_i} = \mathbf{1}}$

2- Condition d'optimale : E $[Z^*_0\text{-}Z_0]^2$ minimale

En développons ses 2 conditions on trouve le système d'équations linéaire du krigeage ordinaire suivante :

$$\sum \boldsymbol{\lambda}\Lambda_{\boldsymbol{i}} = \boldsymbol{\gamma} \;\; (\boldsymbol{x_i} - \boldsymbol{x_j}) + \boldsymbol{u} = \gamma(\boldsymbol{x_i} - \boldsymbol{x_0})$$

$$\sum \Lambda_{\boldsymbol{i}} = \mathbf{1}$$

sous forme matricielles, ce système d'équations sera:

$$\begin{vmatrix} \gamma_{11} & \gamma_{12} & \gamma_{1n} & 1 \\ \gamma_{21} & \gamma_{22} & \gamma_{2n} & 1 \\ \gamma_{n1} & \gamma_{n2} & \gamma_{nn} & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix} \begin{vmatrix} h_1 \\ h_2 \\ h_n \\ u \end{vmatrix} = \begin{vmatrix} \gamma_{1\,0} \\ \gamma_{2\,0} \\ \gamma_{n\,0} \\ 1 \end{vmatrix}$$

La variance d'estimation du krigeage ordinaire est :

$$\mathbf{Var}(\boldsymbol{z_0^*} - \boldsymbol{z_0}) = \boldsymbol{E}|\boldsymbol{z_0^*} - \boldsymbol{z_0}| = \sum \Lambda_{\boldsymbol{i}}\, \gamma(\boldsymbol{x_i} \; - \boldsymbol{x_0}\,) + \boldsymbol{u}$$

L'écart type d'estiamtion du krigeage ordinaire est:

$$\boldsymbol{\sigma} \;=\; \sqrt{\sum \Lambda_{\boldsymbol{i}}\, \gamma(\boldsymbol{x_i} \; - \boldsymbol{x_0}\,) + \boldsymbol{u}}$$

Krigeage universel

L'hypothese de stationnarité sur laquelle repose les deux types de krigeage présentés précédemment peut souvent etre mise en doute. En particulier,il semble souvent erroné de postuler que l'espérance de la fonction aléatoire étudiée reste constante ou quasi-constante sur le champ D. En krigeage avec un modèle de tendance,aussi appelé krigeage en présence d'une dérive, l'espérance est une fonction des coordonnées spatiales ou de variables régionalisées auxiliaires connues exhaustivement. La majorité des auteurs en géostatistique parlent alors de krigeage universel (Matheron, 1969 ; Cressie, 1993,) et de krigeage avec dérive externe (Goovaerts, 1997 ; Wackernagel, 2003,) respectivement.

Le modèle de base du krigeage universel est :

$$Z(s) = \sum_{j=0}^{p} f_i(s)\beta_j + \delta(s)\,, \quad \mathrm{s} \in \mathrm{D}$$

f_i(s) fonctions de la position S(x,y), β_j paramètres inconnus et $\delta(.s)$ fonction aléatoire stationnaire intrinsèque d'espérance nulle et de structure de dépendance connue. Les f_i(s) sont déterminés par l'utilisateur.

Conclusion

De la revue présentée dans cette partie, le krigeage est une méthode d'interpolation intéressante, de plus, puisqu'il s'agit d'une méthode stochastique, le krigeage permet d'estimer des erreurs de prévisions, et tenir compte de la structure de dépendance spatiale des données. Ainsi, on peut s'attendre à ce que le krigeage génère les prévisions spatiales les plus justes. De plus, l'estimation des erreurs qu'il produit est plus fiable que celles produites par les autres méthodes stochastiques, car les postulats de base du krigeage modélisent mieux la réalité pour des données à référence spatiale.

TD : Exercices Résolus:

Exercice 1

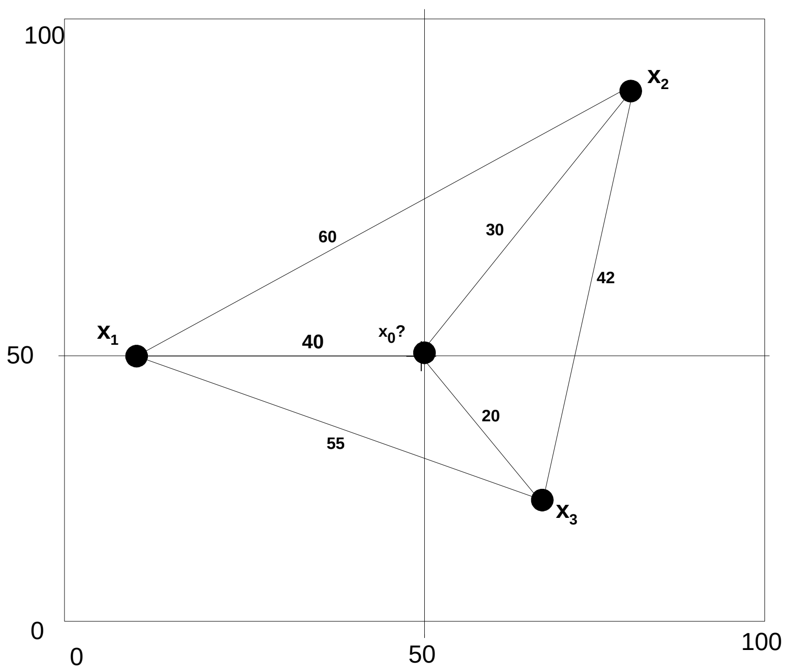
Calculez le variogramme expérimental des 10 valeurs suivantes d'une variable mesurées tous les 10 mètres le long d'une ligne:

1 2 0 0 1 0 0 1 2 3

Exercice 2

On veut estimer par krigeage la variable inconnue Z_0 au niveau du point X_0 en utilisant les autres variables Z_1, Z_2 et Z_3 mesurées aux points X_1, X_2 et X_3 comme indiqué sur le schéma. Le Variogramme étant sphérique avec une pépite C_0= 0.50, un palier C=1 et une portée a=60m. On donne Z_1=4, Z_2= 6 , Z_3 = 5

1. **Quel type de krigeage doit-on choisir ? justifiez.**
2. **Ecrire le système de ce krigeage sous forme d'équations et sous forme matricielle.**
3. **Calculez Z_0 ainsi que son intervalle de confiance.**
4. **Combien vaut Z_0 si on utilise la méthode de l'inverse des carrés des distances**
5. **Comparez les deux résultats.**



Exercice 3

Dans un problème à une seule dimension (1D), on veut estimer le point X_0 à l'aide des points X_1 et X_2. Le tableau suivant donne les coordonnées de ces points :

Point	Coordonnée x
x0	0
x1	-5
x2	3

Le modèle de variogramme est de type gaussien:

$\delta(h)= 10 \exp(-h^2/8^2)$ ($\delta(h)= 10\ e^{(-h2/82)}$)

1) quels sont les traits structuraux du variogramme ?
2) quelle est la valeur de la variance observée ?
3) Faites le croquet et dites de quel Azimut s'agit il ?
4) Construisez le système de krigeage ordinaire.
5) Que vaut la valeur estimée Z_0 et la variance de Krigeage ordinnaire si Z_1= 4 et Z_2 = 5 ?
6) Ecrire l'intervalle de confiance de Z_0 à 95%.
7) Connaissant la moyenne m=4.5, quelle serait la valeur de Z_0 avec le krigeage simple ?
8) quelle serait la valeur de Z_0 si on l'estime avec la méthode de l'inverse des carrées des distances ?
9) Comparer et interprétez les trois méthodes.

Corrigé des Exercices

Corrigé de l'exercice 1

1 2 0 0 1 0 0 1 2 3

Le variogramme à distance 10 mètres (d'ailleurs égal à celui à distance −10 mètres), est égal à:

$$\gamma(10)=(1/2)[(1-2)^2+(2-0)^2+(0-0)^2+(0-1)^2+(1-0)^2+(0-0)^2+(0-1)^2+(1-2)^2+(2-3)^2]/9$$

et on trouve ainsi:

distance	variogramme	nombre de paires
0	0.	10
10	0.56	9
20	1.00	8
30	1.07	7
40	1.25	6
50	1.40	5
60	1.88	4
70	1.50	3
80	0.50	2
90	2.00	1

Corrigé de l'exercice 2

1. Etant donné que la moyenne est inconnue On choisi le krigeage ordinaire.
2. Puisque les distances sont données sur le schéma, on procède directement au calcul des variogrammes en utilisant la formule du modèle sphérique isotropique avec les caractéristiques structurales (Pépite, palier et portée) proposées par l'exercice :

$\gamma(1,1) = \gamma(2,2) = \gamma(3,3) = 0.5$
$\gamma(1,2)\ ) = \gamma(2,1) = \gamma(60) = 1.5$
$\gamma(1,3) = \gamma(3,1) = \gamma(55) = 1.41$
$\gamma(2,3) = \gamma(3,2) = \gamma(42) = 1.37$
$\gamma(1,0) = \gamma(0,1) = \gamma(40) = 1.35$
$\gamma(2,0) = \gamma(0,2) = \gamma(30) = 1.18$
$\gamma(3,0) = \gamma(0,3) = \gamma(20) = 0.98$

Donc le système de krigeage ordinaire s'écrit sous forme d'équations :

$$\begin{aligned}
0.5\lambda_1 + 1.5\lambda_2 + 1.41\lambda_3 + \mu &= 1.35\\
1.5\lambda_1 + 0.5\lambda_2 + 1.37\lambda_3 + \mu &= 1.18\\
1.41\lambda_1 + 1.37\lambda_2 + 0.5\lambda_3 + \mu &= 0.98\\
\lambda_1 + \lambda_2 + \lambda_3 + 0 &= 1
\end{aligned}$$

Il s'écrit sous forme matricielle :

$$\begin{vmatrix} 0.5 & 1.5 & 1.41 & 1 \\ 1.5 & 0.5 & 1.37 & 1 \\ 1.41 & 1.37 & 0.5 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix} \begin{vmatrix} \lambda_1 \\ \lambda_1 \\ \lambda_1 \\ \mu \end{vmatrix} = \begin{vmatrix} 1.35 \\ 1.18 \\ 0.98 \\ 1 \end{vmatrix}$$

3. Calcul de Zo

On calcule les déterminants de la matrice ci-dessus en utilisant la règle de Cramer :

D=-2.5584002 D1=-0.41700003 D2=-0.7982001 D3=-1.3432 D4=-1.1541279

λ_1=D1/D= 0.163 λ_2=D2/D=0.312 λ_3=D3/D=0.525 μ =D4/D=0.451

On remarque $\Sigma\lambda i = \lambda_1 + \lambda_2 + \lambda_3 = 1$ donc la condition de non biais est très bien respectée.

On a Z1=4 Z2=6 Z3=5 , donc **Zo**= 0.163*4+0.312*6+0.525*5 = **5.149**

$\sigma_k^2 = 0.451 + (0.163*1.35+0.312*1.18+0.525*0.98) = 1.55$ $\sigma_k = 1.24$

IC = [3.9 6.38]

4. Calcul de Zo par l' Inverse des carrés des distances :

Z	d	d au carré	1/dcarré	Zi * 1/dcarré
4	40	1600	0.000625	**0.0025**
6	30	900	0.001111	**0.00666**
5	20	400	0.0025	**0.0125**
		Somme =	0.004236	**0.02166**
			Zo =0.0216/0.00423	**8.6666**
			Donc Zo = 8.66	

5. Comparaison :

On remarque bien que 8.66>>5.149

8.66 est donnée sans aucune précision sur l'erreur (Pas d'écart type donc pas d'intervalle de confiance, par contre la valeur du krigeage est estimée avec un écart type d'où son intervalle de confiance. Donc le krigeage est le meilleur estimateur que la méthode de l'Inverse des carrés des distances.

Corrigé de l'exercice 3 :

Pépite = 0 , Palier= 10 , Portée = 8 et la variance = Palier = 10

1- Croquet : X_1--- ---- ---- ---- ----X_0 ---- ---- ---- X_3
2- Azimut = 90° (Direction Est-West)

3- Distance

Distance	x0	x1	x2
x0	0	5	3
x1	5	0	8
x2	3	8	0

4- Variogram

Variogram	x0	x1	x2
x0	10	6.77	8.68
x1	6.77	10	6.68
x2	8.68	3.68	10

5- Matrice :

10	3.68	1	λ_1	6.77
3.68	10	1	λ_1	8.68
1	1	0	µ	1

D=-12.64 D1=-9.41 D2=-8.23 D3=-97.64 $\lambda_1 = 0.74$
$\lambda_1 = 0.65$ µ=7.72

Z_0=6.21 varianceK= 18.36 (ecarttype=4.2) cad entre 2 et 10.4

6 - Avec le krigeage simple
La matrice est :

10	3.68	1	λ_1	6.77
3.68	10	1	λ_1	8.68

D= 86.45 D1=35.75 D2=61.88 $\lambda_1 = 0.41$ $\lambda_1 = 0.71$

Z_0= 0.41*4+0.71*5 +(1-(0.41+0.71))*4.5 = 4.65

Avec Inverse de distance on trouve Z0 =4.73

	1/d	1/d carree
z1,0	0,2	0,04
z2,0	0,33333333	0,11
Somme		0,15