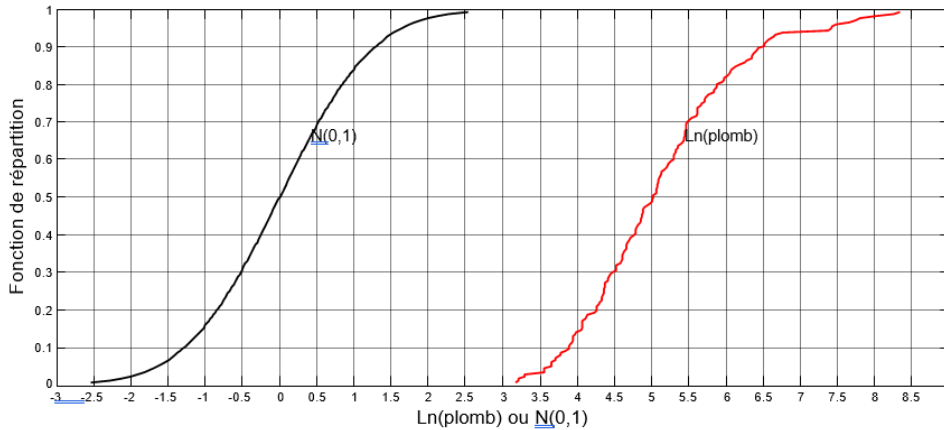


**Question 5 – Automne 2019 - SGS**

La figure 1 montre les fonctions de répartition d'une  $N(0,1)$  et du  $\ln(\text{plomb})$  obtenu avec les données de Dallas vues au TP10.



**Figure 1**

a) On a observé trois données ayant les localisations et valeurs suivantes. Compléter le tableau suivant :

Donnée	Coord. x (m)	Coord. y (m)	Pb (ppm)	Ln(Pb)	N(0,1)
1	5500	5500	245	5.5	
2	6000	5500	403	6	
3	6500	6500	1097	7	

La variable  $Z \sim N(0,1)$  correspondant aux teneurs de Pb montre un variogramme sphérique isotrope de portée 2000 met avec  $C_0=0.1$  et  $C=0.9$ .

b) Quelles sont les unités pour  $C_0$  et  $C$  ?

On effectue le krigeage simple au point  $x_0=(6000, 6000)$  et l'on obtient le système de krigeage suivant :

$$\begin{bmatrix} 1 & 0.57 & 0.105 \\ 0.57 & 1 & 0.224 \\ 0.105 & 0.224 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} 0.443 \\ 0.57 \\ 0.443 \end{bmatrix} \quad \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} 0.186 \\ 0.388 \\ 0.336 \end{bmatrix}$$

c) Quelles seront la valeur krigée et la variance de krigeage simple en ce point  $x_0$  ?

d) On tire de la distribution conditionnelle la valeur gaussienne 1.0. Quelle est la concentration de plomb que l'on simule à ce point ?

e) Sur quelle propriété particulière repose en pratique le SGS lorsqu'on l'applique dans un contexte où plusieurs milliers de données sont disponibles ?

**Question 3 – Automne 2022 - LU**

On simule une variable  $Z$  en 4 emplacements par la méthode de décomposition de Cholesky. Les emplacements sont placés dans l'ordre 1, 2, 3, 4 dans la matrice de covariance. La matrice  $L$  obtenue est la suivante :

$$L = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 1 & 3 \end{bmatrix}$$

- Quelle est la covariance entre les variables aléatoires  $Z$  correspondant aux 3<sup>e</sup> et 4<sup>e</sup> emplacements ?
- Quelle est la variance non-conditionnelle au 3<sup>e</sup> ? au 4<sup>e</sup> emplacements?
- Selon votre réponse en b), est-ce que l'hypothèse de stationnarité du second ordre est respecté ? Justifiez.
- On a observé  $Z_1=1$ . Exprimez  $Z_3$  et  $Z_4$  sous la forme  $a + b Y_2 + c Y_3 + d Y_4$ .
- Déduisez de l'expression en d) la covariance (conditionnelle à  $Z_1$ ) entre les variables aléatoires  $Z_3$  et  $Z_4$ .
- On a généré le vecteur aléatoire  $Y=[Y_2 \ Y_3 \ Y_4]=[-1.25 \ 0.48 \ -0.19]$ . Que valent  $Z_3$  et  $Z_4$  (conditionnelle à  $Z_1$ )?

**Question 6 – Automne 2022 – Krigeage d'indicateur.** Note : On va vu deux types de krigeage d'indicateur (ordinaire et simple), les deux versions sont à l'examen. Ici, un exemple de question pour le krigeage d'indicateur ordinaire.

On vous présente les observations suivantes donnant l'élévation d'un horizon repère.

Point	Coord x (m)	Coord. Y (m)	Élévation (m)
x <sub>1</sub>	30	100	29
x <sub>2</sub>	90	160	37
x <sub>3</sub>	220	80	52
x <sub>4</sub>	300	350	47
x <sub>5</sub>	40	200	$20 \leq Z(x_5) \leq 49$

En x<sub>5</sub>, on n'a aucune information sur les probabilités pour Z(x<sub>5</sub>) à l'intérieur de l'intervalle indiqué [20, 49].

On a modélisé les variogrammes des différentes indicateurs et on a effectué le krigeage ordinaire au point x<sub>0</sub> de coordonnées (50,150). Les variogrammes des indicateurs étant ici tous proportionnels, les poids de krigeage sont les mêmes pour tous les seuils.

Lorsqu'on effectue le krigeage ordinaire avec les 5 points, on obtient les poids :

$$\lambda_1 = 0.35, \lambda_2 = 0.29, \lambda_3 = 0.07, \lambda_4 = 0.04, \lambda_5 = 0.25$$

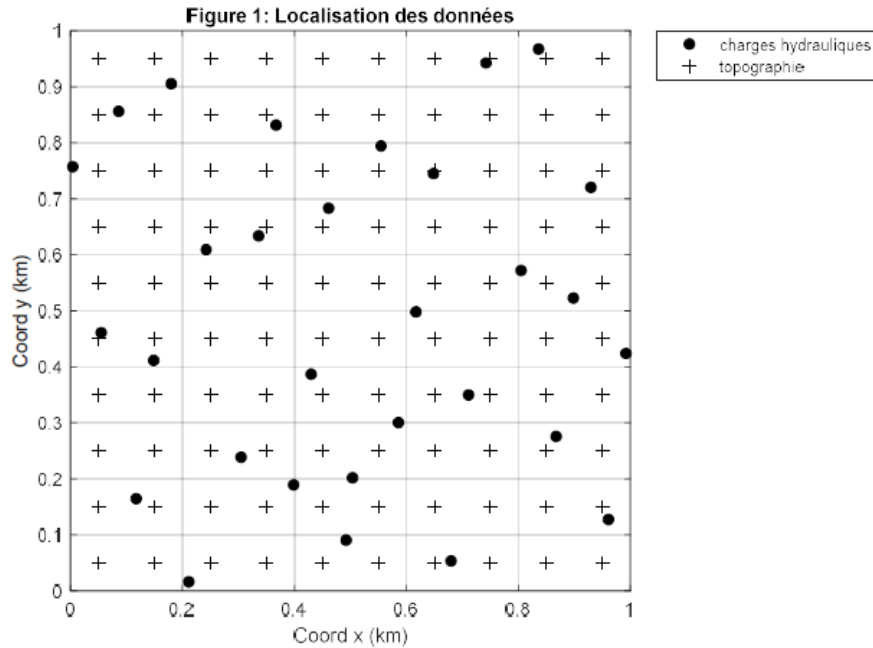
Si l'on effectue le krigeage ordinaire avec les quatre premiers points, on obtient plutôt les poids :

$$\lambda_1 = 0.31, \lambda_2 = 0.39, \lambda_3 = 0.14, \lambda_4 = 0.16$$

- a) Quelle est la probabilité que l'élévation de l'horizon repère au point x<sub>0</sub> soit inférieure à 50m?  
 b) Quelle est la probabilité que l'élévation de l'horizon repère au point x<sub>0</sub> soit inférieure à 40m?

### Inspiré de la question 3 Final 2019 - Cokrigage

La figure 1 donne la localisation de données où l'on a mesuré la charge hydraulique (variable principale Z) dans un aquifère de surface et l'élévation topographique (variable auxiliaire Y). Il y a en tout 30 valeurs de charge et 100 valeurs de topographie.



Le modèle linéaire de corégionalisation est le suivant :

$$\begin{bmatrix} C_{ZZ} & C_{YZ} \\ C_{ZZ} & C_{YY} \end{bmatrix} = \begin{bmatrix} 20 & 5 \\ 5 & 20 \end{bmatrix} \delta(h) + \begin{bmatrix} 60 & 15 \\ 15 & 70 \end{bmatrix} sph(C = 1, a = 0.2km)$$

- Vérifiez s'il s'agit d'un modèle admissible.
- Dans le contexte illustré et avec le modèle de corégionalisation anticipe-t-on que le cokrigage devrait améliorer significativement l'estimation de la charge par rapport à un krigeage ? Justifiez.
- Commentez cet énoncé : « étant donné que la covariance croisée est symétrique en h, on aurait pu calculer le variogramme croisé expérimental, avec les données disponibles, plutôt que la covariance croisée ».

**Question 4 Final 2018 a) – Applications de la géostatistique**

Vous voulez estimer le volume contaminé par des hydrocarbures sur le site d'une centrale électrique utilisant des génératrices au diesel. Tout sol excédant la norme 10 ppm (huiles et graisses) doit être décontaminé. La contamination progresse à partir de la surface. Pour décontaminer le site, deux méthodes seront utilisées. La première consiste à excaver toute contamination sur les 4 premiers mètres et à traiter ex-situ à l'aide de procédés biologiques. La seconde méthode sera utilisée pour la contamination excédant 4 mètres et consiste en un bio-traitement et ventilation in-situ. Des analyses dans des forages ont montré que la contamination n'atteignait pas la nappe phréatique et qu'un sol contaminé en profondeur impliquait également une contamination allant de la surface jusqu'à la profondeur considérée. Suite à ces observations, vous décidez, pour simplifier le problème, de définir une nouvelle variable représentant l'épaisseur de sol contaminé au-delà de la norme 10 ppm (en un point donné en surface). Vous devez fournir des estimations des volumes contaminés qui soient réalistes et indiquer à votre employeur la précision des estimations des volumes de sol à excaver et de sol à traiter in-situ afin de planifier un budget d'intervention raisonnable.

a) Expliquez comment vous pourriez utiliser les simulations conditionnelles des épaisseurs contaminées pour fournir ces estimations de volume de sols contaminés à excaver, de volumes de sols contaminés à traiter in-situ, ainsi que les intervalles de confiance sur ces quantités.

**Question 6 final 2019 – Applications de la géostatistique**

À la suite d'une étude de faisabilité basée sur des milliers de données de carottes sur un certain gisement, on décide de l'exploiter en fosse. Un design de la fosse est effectué donnant les limites ultimes de la fosse. Le design de la fosse est considéré fixe. On voudrait construire un intervalle de confiance sur le profit que va générer cette fosse. On suppose qu'au moment du minage des différents blocs, leur vraie teneur sera quasiment connue (les erreurs sur les teneurs de blocs seront alors négligeables, ce qui n'est pas le cas actuellement). Tous les blocs de la fosse ultime doivent être remontés en surface. Si leur valeur en métal couvre les coûts de traitement au concentrateur ils seront traités, sinon ils seront disposés sur la halde de stérile.

Suggérez une méthode géostatistique permettant d'atteindre l'objectif de construire l'intervalle de confiance sur la valeur de la fosse. Expliquez les principales étapes d'application de la méthode et la façon dont vous l'utiliserez pour fournir l'intervalle de confiance souhaité.

**Question 6 Final 2018 (Seulement a, e et f) - Vrai ou Faux**

Discutez les énoncés suivants (vrai, partiellement vrai, ou faux). Justifiez votre réponse en maximum trois phrases.

a) Dans la méthode SGS, il est requis que les données suivent une loi normale. L'utilisateur doit donc fournir le variogramme de la variable après transformation vers la loi normale (ex. transformation graphique). La simulation reproduit le variogramme de la variable normale (en moyenne). Après transformation inverse pour revenir au domaine initial, la simulation reproduit (en moyenne) le variogramme de la variable originale si le champ transformé est initialement (multi)gaussien.

e) On pourrait utiliser le recuit simulé pour calibrer automatiquement un champ de transmissivité de sorte qu'il reproduise presque parfaitement les charges hydrauliques observées dans quelques piézomètres. La fonction objectif pourrait être la somme des différences au carré entre les charges hydrauliques observées et les charges hydrauliques retournées par un simulateur d'écoulement. Par contre, le facteur « temps de calcul » empêche d'envisager sérieusement cette solution lorsqu'une seule solution du simulateur d'écoulement prend plusieurs minutes à obtenir.

f) L'on peut calculer la variance conditionnelle représentant l'étalement de la distribution conditionnelle estimée par KI ordinaire. La valeur de cette variance dépend en bonne partie de la similitude des teneurs aux points utilisés dans le krigeage d'indicatrices. Plus les observations montrent des valeurs semblables, moins forte est la variance conditionnelle.

**Question vrai ou faux.**

Indiquez si les énoncés suivants sont vrais ou faux. Justifiez lorsque la réponse est fausse.

- a) Dans un cokrigeage, les poids associés à la variable secondaire n'ont jamais d'unités.
- b) Dans la méthode SGS, on doit faire l'hypothèse que la distribution de la variable transformée est multinormale. Il n'est pas suffisant que la distribution marginale soit normale.
- c) Dans le SGS, même si l'on doit transformer la variable pour qu'elle soit normale, on doit utiliser le variogramme de la variable originale dans le krigage propre à cette méthode.
- d) Dans un cokrigeage, les poids de cokrigeage associés à la variable principale (c.-à-d. celle que l'on estime) n'ont jamais d'unités.
- e) La variance de cokrigeage est plus élevée là où localement les valeurs observées varient beaucoup.
- f) Dans un recuit simulé on rejette toujours une modification qui fait croître la fonction objectif que l'on cherche à minimiser.
- g) Dans le krigage d'indicatrices, le calcul de la variance conditionnelle n'implique pas les variances de krigage des indicatrices correspondant aux divers seuils.
- h) Dans un krigage d'indicatrices, la variance conditionnelle augmente généralement avec la variabilité des valeurs observées dans le voisinage du point à estimer.
- i) Dans un krigage d'indicatrices effectué avec un nombre limité de seuils, l'espérance conditionnelle calculée en un point où se trouve une donnée coïncide exactement avec la valeur de la donnée.
- j) On peut construire des intervalles de confiance sur les ressources d'un gisement en effectuant plusieurs réalisations conditionnelles aux données disponibles.