Chapitre 4 : Théorie du krigeage

4.1. Le Krigeage

Le principal objectif de la géostatistique est la prédiction spatiale, encore appelée krigeage, consistant à prédire une variable régionalisée d'intérêt sur un domaine d'étude, à partir des données observées à certains emplacements. Le krigeage repose fondamentalement sur la modélisation et l'estimation de la structure de dépendance spatiale. La description de cette dernière se fait couramment à l'aide d'outils statistiques tels que le variogramme ou la covariance, calculés sur l'ensemble du domaine d'intérêt et sous une hypothèse de stationnarité.

Seule la méthode d'interpolation par krigeage repose sur une méthode statistique satisfaisante et permet d'obtenir la vraie variance d'estimation.

Puisqu'on peut calculer la variance d'estimation pour tout estimateur linéaire, pourquoi ne pas choisir celui qui assure la variance d'estimation minimale? C'est précisément ce qu'effectue le krigeage. Dans le cas stationnaire, on en reconnaît 2 types principaux, selon que la moyenne du processus est connue ou non, soit le krigeage simple et le krigeage ordinaire. Ce dernier est, de loin, le plus fréquemment utilisé.

4.2. Type de Krigeage

4.2.1. Krigeage stationnaire à moyenne inconnue (krigeage linéaire ordinaire)

Le krigeage ordinaire est une opération qui est répétée en chaque noeud x_0 d'une grille régulière recouvrant le domaine étudié. Pour un ensemble de n points de données x d'un voisinage centré autour d'un point x_0 de la maille d'estimation, on peut construire, en minimisant la variance d'estimation. Ce système permet donc de retrouver les N pondérateurs λi . Ces derniers, dit pondérateurs de krigeage, donnent la variance d'estimation la plus petite possible et elle est appelée variance de Krigeage. Où les λ_i sont des pondérateurs à affecter aux points de données et où μ est un paramètre de Lagrange qui intervient pour des raisons algébriques. Le membre gauche du système contient les covariances entre les points de données, tandis que le membre droit contient les covariances entre chaque point de donnée et le point d'estimation x_0. Ce système, une fois résolu, permet de transférer au point x_0 de l'information en provenance des points de données voisins, par le calcul d'une moyenne pondérée

En effectuant la multiplication matricielle, on peut réécrire le système sous la forme :

$$\sum_{i=1}^{N} \lambda_i \, C(x_\alpha - x_i) + \mu = C(x_\alpha - x_0), \qquad \alpha = 1 \text{ à } N \text{ (\textit{Équation 30})}$$

$$\sum_{i=1}^{N} \lambda_i = 1$$

Le système de krigeage ordinaire suivant :

$$\begin{vmatrix} C(x_1 - x_1) & C(x_1 - x_2) & \ldots.. & C(x_1 - x_n) & 1 \\ C(x_2 - x_1) & \ldots.. & \ldots.. & C(x_2 - x_n) & 1 \\ \ldots.. & \ldots.. & \ldots.. & \ldots.. & \ldots.. \\ C(x_n - x_1) & C(x_n - x_2) & \ldots.. & C(x_n - x_n) & 1 \\ 1 & 1 & 1 & 1 & 0 \end{vmatrix} \begin{vmatrix} \lambda_1 \\ \lambda_2 \\ \ldots.. \\ \lambda_n \\ \mu \end{vmatrix} = \begin{vmatrix} C(x_1 - x_0) \\ C(x_2 - x_0) \\ \\ C(x_N - x_0) \\ 1 \end{vmatrix}$$

$$Z_v^*(x_0) = \sum_{i=1}^{N} \lambda_i Z(x_i) \text{(\textit{Équation 31})}$$

La variance d'estimation en krigeage ordinaire est :

$$\sigma_{ord}^2 = C(0) - \sum_{i=1}^{N} \lambda_i C(x_0, x_i) - \mu \text{ (\textit{Équation 32})}$$

4.2.2. Krigeage stationnaire à moyenne connue (krigeage linéaire simple)

Parfois on connaît la moyenne "µ" du champ à estimer ou du moins on en possède un estimé fiable. On peut alors former un estimateur sans biais :

$$Z_v^* = \sum_{i=1}^{N} \lambda_i Z(x_i) + \left(1 - \sum_{i=1}^{N} \lambda_i\right)\mu \quad \text{(\textit{Équation 33})}$$

La variance d'estimation en krigeage simple est :

$$\sigma_{simp}^2 = C(0) - \sum_{i=1}^{N} \lambda_i C(x_0, x_i) \text{ (\textit{Équation 34})}$$

4.3. Propriétés du krigeage

Les principales propriétés et caractéristiques associées au krigeage sont: Linéaire, sans biais, à variance minimale, par construction. Interpolateur exact. : Si l'on estime un point connu, on retrouve la valeur connue (le point de point connue sera égale 1 et le poids des autres point sera égal à 0). Présente un effet d'écran : les points les plus près reçoivent les poids les plus importants. Cet effet d'écran varie selon la configuration et selon le modèle de variogramme utilisé pour le krigeage. Plus l'effet de pépite est important, moins il y a d'effet d'écran. Tient compte de la taille du champ à estimer et de la position des points entre eux. Par l'utilisation du variogramme, tient compte de la continuité du phénomène étudié (effet de pépite, anisotropie, etc.). Effectue généralement un lissage, i.e. les estimations sont moins variables que les teneurs réelles (point ou bloc) que l'on cherche à estimer. Presque sans biais conditionnel. Ceci signifie que lorsqu'on applique une teneur de coupure à des valeurs estimées, on récupérera approximativement la teneur prévue. C'est une propriété très importante pour les mines. Cette propriété implique que l'estimateur utilisé soit plus lisse que la valeur qu'il cherche à estimer, ce qui est le cas pour le krigeage. Si l’on observe en un point une valeur coïncidant avec la valeur krigée pour ce point, alors les valeurs krigées en d'autres points ne sont pas modifiées par l'inclusion de ce nouveau point dans les krigeage. Par contre les variances de krigeage, elles, sont diminuées. De même, si l’on krige un certain nombre de points et que l’on utilise les valeurs krigées comme si c’étaient de nouvelles données, alors les krigeages subséquents ne s’en trouvent pas modifiés (sauf pour la variance de krigeage).

4.4. Cokrigeage

Souvent l'on a plusieurs variables mesurées, soit aux mêmes points échantillons, soit en des points différents. Sans perte de généralité, l'on va considérer le cas où une des variables est identifiée comme prioritaire (variable principale Z), et les autres sont des variables secondaires. Pour simplifier l'écriture, on va considérer que l'on a une seule variable secondaire (Y). Toutefois l'extension à plusieurs variables est immédiate et ne pose aucun problème théorique particulier.

4.4.1. Cokrigeage ordinaire

On veut former une estimation linéaire de la variable principale Z(x) (avec nz observations) à partir d'observations de la variable principale et de la variable secondaire Y(x) (avec : ny observations).

Les moyennes de Z(x) et Y(x) sont inconnues :

$$Z_0^* = \sum_{i=1}^{nz} \lambda_i Z(x_i) + \sum_{i=1}^{ny} \alpha_i Y(x_i) \quad (\textit{Équation 35})$$

L'estimateur doit être sans biais, ceci est assuré en imposant :

$$\sum_{i=1}^{nz} \lambda_i = 1 , \sum_{i=1}^{ny} \alpha_i = 0$$

La variance d'estimation s’écrit :

$$Var\left(Z_0 - Z_0^*\right) = Var(Z_0) + \sum_{i=1}^{nz}\sum_{j=1}^{nz} \lambda_i \lambda_j Cov\left(Z_i, Z_j\right) + \sum_{i=1}^{ny}\sum_{j=1}^{ny} \alpha_i \alpha_j Cov\left(Y_i, Y_j\right) + 2\sum_{i=1}^{nz}\sum_{j=1}^{ny} \lambda_i \alpha_j Cov\left(Z_i, Y_j\right) - 2\sum_{i=1}^{nz} \lambda_i Cov(Z_0, Z_i) - 2\sum_{i=1}^{ny} \alpha_i Cov(Z_0, Y_i) \quad (\textit{Équation 36})$$