

Krigage dual

Il est intéressant de reformuler le krigage sous une forme différente que l'on appelle krigage dual. Considérons le système de krigage habituel sous forme matricielle :

$$\begin{bmatrix} K_s & F \\ F' & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} k_s \\ f \end{bmatrix}$$

où

K_s est la matrice de covariance (nxn) entre les points d'observation,

k_s est le vecteur nx1 des covariances points d'observation- point ou bloc à estimer,

F est la matrice (nxc) des fonctions de base décrivant la dérive évaluée aux points observations,

f est le vecteur nc x 1 des fonctions de base évaluées au point ou bloc à estimer,

0 est la matrice ncxc de 0,

λ et μ sont les poids de krigage et les multiplicateurs de Lagrange.

On peut écrire la valeur estimée comme : $Z^* = \begin{bmatrix} Z' & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix}$

Or, on a :

$$\begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} K_s & F \\ F' & 0 \end{bmatrix}^{-1} \begin{bmatrix} k_s \\ f \end{bmatrix}$$

Donc,

$$Z^* = \begin{bmatrix} Z' & 0 \end{bmatrix} \begin{bmatrix} K_s & F \\ F' & 0 \end{bmatrix}^{-1} \begin{bmatrix} k_s \\ f \end{bmatrix} = \begin{bmatrix} a' & b' \end{bmatrix} \begin{bmatrix} k_s \\ f \end{bmatrix}$$

avec

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} K_s & F \\ F' & 0 \end{bmatrix}^{-1} \begin{bmatrix} Z \\ 0 \end{bmatrix}$$

Cette dernière équation montre que le krigage peut être vu comme un interpolateur composé de 2 termes, un premier terme (a et k_s) décrit la composante aléatoire et un 2^e terme (b et f) décrit la composante déterministe du phénomène. Ces deux composantes de l'interpolation correspondent à la décomposition $Z(x)=Y(x)+m(x)$ où $Y(x)$ est aléatoire (mais structuré selon le variogramme) et $m(x)$ est déterministe.

Avantages et désavantages du krigage dual

Avantages :

- Rapidité pour le calcul d'une nouvelle valeur estimée
- Possibilité de formuler des contraintes

Désavantages :

- Pas possible d'obtenir la variance de krigage

Rapidité : Considérons un krigage global, on a alors une seule matrice à inverser pour les 2 approches. L'estimation d'un point demande par krigage « primal » de multiplier cette inverse par un vecteur, ce qui nécessite ($O((n+nc)^2)$) multiplications afin de déterminer les poids de krigage. Par la suite, on doit multiplier ce vecteur de poids par les valeurs observées ($O(n)$). Au contraire pour le krigage dual, on obtient directement l'estimé avec un simple produit scalaire ($O(n+nc)$), d'où un gain appréciable, surtout si l'on désire estimer un très grand nombre de points (ou blocs).

Imposer des contraintes : Comme les valeurs estimées, sous forme duale, présentent une forme analytique, on peut imposer facilement des contraintes liant les estimations en différents points ainsi que des transformations linéaires de ces estimations.

Exemple :

- Imposer que 2 points distincts aient la même valeur estimée

$$[a' \ b'] [k_1; f_1] = [a' \ b'] [k_2; f_2]$$

ou

$$[a' \ b'] [k_1 - k_2; f_1 - f_2] = [0]$$

Cette dernière équation est une ligne supplémentaire dans la matrice de gauche et le membre de droite du système dual. Cette ligne correspond à la contribution dans l'estimation du couple de v.a. $Z_1 - Z_2$. Chaque élément de la ligne sera donc $Cov(Z_1 - Z_2, Z_i)$. On doit aussi ajouter une colonne correspondante pour le multiplicateur de Lagrange associé à cette nouvelle équation (autrement il n'y aurait pas de solution au système). Bref, comme pour tous les krigages, la matrice de gauche doit toujours être conservée symétrique. L'élément sur la diagonale n'est autre que $Var(Z_1 - Z_2)$.

- Imposer que dans une zone donnée la moyenne des estimations soit égale à une certaine valeur « t » supposée connue.

$$[a' \ b'] \frac{1}{v} \int_v \begin{bmatrix} k(x) \\ f(x) \end{bmatrix} dx = t$$

- Contraintes sur les dérivées de la fonction, etc...

Variance de krigage

Comme les poids λ ne sont pas disponibles dans la formulation duale, l'on ne peut calculer la variance de krigage.

Interprétation des coefficients duaux

Les poids « b » associés aux composantes de « f » sont les coefficients de la régression (généralisée) pour l'estimation de la dérive. Dans le cas du KO, il s'agit simplement de m^* , dans le cas d'une dérive linéaire en 1D, ce sont l'ordonnée à l'origine et la pente de la droite de régression, etc... Si la covariance utilisée est un effet de pépité pur, on retrouve exactement les coefficients de la régression habituelle.

Du système de krigeage dual, il résulte aussi : $a'F = 0$, ce qui indique que les coefficients « a » se comportent comme des résidus dans une régression (i.e. ils sont orthogonaux à la matrice de design F). En fait ce sont effectivement les résidus de la régression généralisée au facteur multiplicatif C_0+C près (une régression généralisée est une régression classique dans laquelle on tient compte des covariances entre les résidus du modèle). On constate que si le résidu est faible, i.e. si la valeur est bien estimée par la dérive, le poids « a_i » correspondant est faible et la contribution du terme aléatoire liée à ce point est faible. L'inverse se produit lorsque le résidu est élevé.

Bref, le krigeage peut être vu comme une régression généralisée que l'on complète, en vue d'améliorer l'estimation, par un terme aléatoire dont la contribution augmente avec l'imprécision (locale) de la régression, prise isolément, et augmente aussi avec le degré de corrélation spatiale.

Discontinuité de l'interpolation aux points échantillons

La formulation duale permet d'écrire sous une forme un peu plus explicite:

$$Z_0^* = \sum_{i=1}^n a_i Cov(Z_i, Z_0) + \sum_{k=1}^{nc} b_k f_k(x_0)$$

Ce qui illustre bien qu'au point échantillon x_i , la fonction interpolée effectue un saut de $a_i C_0$ où C_0 est l'effet de pépité.

Transformations

Soit $L(Z(x))$ une transformation linéaire de $Z(x)$ (exemples : intégration sur un support fini, moyenne pondérée, dérivée de tout ordre, Laplacien, etc...). Le système de krigeage primal de cette transformation linéaire peut s'écrire :

$$\begin{bmatrix} K_s & F \\ F' & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = L \left(\begin{bmatrix} k_s \\ f \end{bmatrix} \right)$$

L'estimation de la transformation s'obtient en appliquant la transformation au membre de droite. Sous l'écriture duale, on aura :

$$L(Z_0)^* = \sum_{i=1}^n a_i L(Cov(Z_i, Z_0)) + \sum_{k=1}^{nc} b_k L(f_k(x_0)) = L(Z_0^*)$$

En mots, l'estimation d'une transformation linéaire à partir des données Z_i est égale à la transformation linéaire appliquée aux valeurs estimées. Plus concrètement par exemple, l'estimation par krigeage de la dérivée est égale à la dérivée du krigeage, etc...

Krigage avec dérive externe

On a vu que le krigage avec dérive correspondait à une décomposition $Z(x)=W(x)+m(x)$ où $W(x)$ est aléatoire et $m(x)$ est déterministe. Jusqu'à maintenant, on a supposé que l'on pouvait exprimer $m(x)$ comme :

$$m(x) = \sum_{k=1}^{nc} a_k f_k(x)$$

où les fonctions f sont des fonctions de base, le cas le plus courant étant simplement de prendre des monômes selon les coordonnées d'un faible degré (0, 1 ou 2).

Parfois, l'on dispose d'une autre variable observée (Y) (ou plusieurs) que l'on sait reliée au niveau moyen de notre variable Z . On pourrait vouloir utiliser cette variable pour prédire le niveau de $Z(x)$. Un exemple pourrait être la contamination d'un sol (Z) et la topographie (Y); le niveau de la nappe phréatique (Z) et la topographie (Y); la concentration d'un élément (Z) et le type de roche (Y); la profondeur d'un horizon (Z) et les temps sismiques (Y); la profondeur d'un horizon et les temps radar (Y). On va supposer que :

$$m(x) = b_1 + b_2 Y(x)$$

On ne sera pas étonné de constater qu'en appliquant la même démarche que pour le krigage avec dérive, l'on sera conduit à un système de krigage identique:

$$\begin{bmatrix} K_s & F \\ F' & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} k_s \\ f \end{bmatrix}$$

La seule différence étant que les fonctions de base « $f(x)$ » sont remplacées par la (ou les) fonction(s) $Y(x)$. Ainsi, avec 1 seule variable $Y(x)$, on aura :

$$\begin{bmatrix} K & 1 & Y \\ 1' & 0 & 0 \\ Y' & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} k \\ 1 \\ y_0 \end{bmatrix}$$

Interprétation

$Y(x)$ doit normalement être une fonction assez régulière puisqu'elle est liée à la partie variant très lentement dans $Z(x)$ (i.e. $m(x)$). Comme on sait que le krigage effectue implicitement la régression avant d'effectuer un krigage simple avec les résultats de cette régression, il suit que l'information de $Y(x)$ sera prise en compte pour l'estimation de $m(x)$ et donc de $Z(x)$. Toutefois, on sait aussi que le terme de dérive influence peu l'estimation lorsque la corrélation spatiale est forte et que l'on dispose d'un nombre suffisant de données. Donc, à proximité des données, l'influence de $Y(x)$ sera très faible. Par contre, à plus grande distance, la surface interpolée aura tendance à suivre les fluctuations exprimées par $Y(x)$.

On remarquera qu'il faut que l'on connaisse la variable $Y(x)$ en tout point où l'on a observé $Z(x)$ et aussi en tout point où l'on désire effectuer l'estimation. Lorsque $Y(x)$ n'est pas connu en certains points, l'on devra au préalable l'estimer par krigage ou une autre méthode adéquate.

Un dernier point à considérer est le fait que la covariance devrait être calculée sur les résidus de la régression de $Z(x)$ par $Y(x)$. Comme on estime les paramètres de la régression, on introduit automatiquement un certain biais dans le variogramme ou la fonction de covariance que l'on calcule des résidus estimés. Ce biais sur le variogramme est généralement faible à petite distance. Les méthodes et trucs utilisés pour le cas du krigage avec dérive peuvent être utilisés ici également (ex. calculer le variogramme de $Z(x)$ pour des $Y(x)$ à peu près constant, utiliser le principe de validation croisée, utiliser des approches itératives).

La formulation duale peut bien sûr s'appliquer et elle permet de retrouver aisément les coefficients de la régression ainsi que les résidus de la régression.

Lien linéaire entre $Z(x)$ et $Y(x)$

La régression de $Z(x)$ sur $Y(x)$ doit présenter un lien linéaire. On peut transformer $Y(x)$ à loisir pour tenter de linéariser la régression (incluant codages). On peut, entre autres, inclure des composantes quadratiques ou périodiques en $Y(x)$ si ceci s'avère approprié. Il est toutefois beaucoup plus délicat de transformer $Z(x)$ et, autant que possible, on doit l'éviter.

Règle générale, l'influence de $Y(x)$ dans l'estimation est plutôt modérée sauf si l'on se retrouve en extrapolation. Si l'on désire inclure une plus grande contribution de $Y(x)$ dans l'estimation, le cokrigage s'avère généralement une meilleure méthode.

Variance de krigage

Dans le cas du krigage avec dérive externe, la présence de la dérive implique des contraintes supplémentaires dans le système de krigage, ce qui résulte, toute chose égale, en une variance de krigage supérieure à celle du krigage ordinaire. Il peut sembler paradoxal que l'ajout d'une information ($Y(x)$) fasse accroître la variance de krigage. Tout s'explique si l'on réalise que l'information $Y(x)$ sert qu'à moduler l'espérance de $Z(x)$. Normalement on devrait avoir une fonction de covariance différente dans les deux cas. En effet $Y(x)$ tient compte d'un élément de variation de grande échelle qui devrait affecter le variogramme utilisé pour le krigage, ce qui pourrait en réalité mener à des variances de krigage inférieures. C'est l'expérience ou des techniques comme la validation croisée qui peuvent permettre de décider des performances relatives de chaque modèle.