DERIVE EXTERNE ET REGRESSION LINEAIRE

Etienne CASTELIER
CENTRE DE GEOSTATISTIQUE
35, rue Saint-Honoré
77 305 Fontainebleau (France)

RESUME La méthode de la dérive externe est utilisée en géostatistique depuis 1979. C'est un cas particulier du krigeage universel. Elle permet d'estimer une variable Z connue en peu de points en s'aidant d'une autre variable f connue en tout point.

On montre ici que la méthode de la dérive externe contient implicitement une régression linéaire de la variable Z en fonction de la variable f .

De plus cette méthode implique que la fonction f introduite dans la dérive doit posséder une régularité supérieure à celle des résidus.

ABSTRACT The external drift method, which has been used as a geostatistical tool since 1979, is a particular case of universal kriging. It allows us to estimate a variable Z known only on a small set of points given another variable f known everywhere.

Here we show that the external drift method implicitly contains a linear regression of the variable Z on the variable f .

Furthermore this method compels the function f used in the drift to be more regular than the residual Y .

Table des Matières

RAPPEL SUR LA RÉGRESSION LINÉAIRE	48
Définition probabiliste	48
Point de vue des moindres carrés	49
Ecriture matricielle	49
KRIGEAGE DE LA DERIVE	49
Krigeage de la moyenne	49
Présentation de la méthode de dérive externe	50
Equations	51
Calcul des coefficients a et b de la dérive	51
Comparaison avec une régression linéaire	53
INFLUENCE DES DIFFÉRENTES VARIABLES	54
Poids d'un couple	55
Résumé	57
CONCLUSION	58
ANNEXE : APPLICATION	58

Avertissements

Une des difficultés de la note est le passage entre le krigeage vu comme une projection sur l'espace des variables aléatoires (c'est-à-dire le krigeage dont on a l'habitude en géostatistique) et le krigeage vu comme une simple opération algébrique sur l'espace euclidien $\mathbb{R}^N$.

Pour que ce passage se fasse le mieux possible, convenons des notations suivantes :

Dans l'espace des variables aléatoires $\mathcal{V}$

$\mathbb{R}^d$ $d = 1..3$ est l'espace géographique de travail.

x symbolise un point de $\mathbb{R}^d$.

Une fonction aléatoire (FA) $Y: \mathbb{R}^d \rightarrow \mathcal{V}$ sera toujours désignée par une majuscule. Une réalisation de cette FA sera notée $y(x)$. Il s'agit d'une fonction réelle.

On appelle points de mesure x_α $\alpha = 1..N$ les points où $y(x)$ est connue. On note, pour abrégé, $y(x_\alpha) = y_\alpha$.

Dans l'espace euclidien $\mathbb{R}^N$

Les N valeurs y_α $\alpha = 1..N$ de $y(x)$ aux points de mesure x_α forment un vecteur de $\mathbb{R}^N$.

Dans $\mathbb{R}^N$ on conviendra de noter en gras avec une majuscule une matrice M , en gras avec une minuscule un vecteur v et le scalaire s en minuscule normale.

Un vecteur v a pour composantes v_α $\alpha = 1..N$. Il s'écrit $v = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix}$. Son transposé est noté $v^T = (v_1 \dots v_N)$. Le produit scalaire des vecteurs v et w se note $v^T w$.

Tout krigeage qui suppose connues les mesures y_α est une forme linéaire sur le vecteur y de composantes y_α . On peut ainsi regarder un krigeage uniquement comme une forme linéaire de $\mathbb{R}^N$ et appliquer ce krigeage à tout vecteur de $\mathbb{R}^N$ indépendamment de tout sens probabiliste.

1 Rappel sur la régression linéaire

Dans la suite on identifiera le krigeage de la dérive à une régression linéaire. Rappelons la forme d'une régression linéaire.

1.1 Définition probabiliste

Soient Y et Z deux variables aléatoires (VA) d'espérance et de variance finies. On désigne par $E\{\}$ l'espérance d'une VA. Ainsi $E\{Y\}$ est l'espérance de Y .

Définition La régression linéaire de Z sur Y est la projection de Z sur l'espace vectoriel des VA de la forme $aY + b$ avec $a, b \in \mathbb{R}$.

La projection de Z s'écrit $a^*Y + b^*$ où a^* et b^* minimisent $E\{(Z - aY - b)^2\}$. On trouve, tous calculs faits :

$$a^* = \frac{C_{yz}}{C_y} \quad b^* = E\{Z\} - a^*E\{Y\} \quad (1.1)$$

avec $C_y = E\{Y^2\} - (E\{Y\})^2$ et $C_{yz} = E\{YZ\} - E\{Y\}E\{Z\}$.

1.2 Point de vue des moindres carrés

L'écriture ci-dessus n'est pas pratique. On ne dispose pas de variables aléatoires Y et Z , mais de N réalisations de ces variables, soit 2 vecteurs de $\mathbb{R}^N$ $\mathbf{y}$ et $\mathbf{z}$.

L'espérance $E\{V\}$ d'une VA V est remplacée par la moyenne $\frac{1}{N} \sum_{\alpha=1}^N v_\alpha$. Le problème est donc de trouver les coefficients a^* et b^* qui minimisent :

$$\frac{1}{N} \sum_{\alpha=1}^N (z_\alpha - a y_\alpha - b)^2$$

C'est la forme habituelle d'un problème de moindres carrés. La solution est semblable à celle du paragraphe précédent (1.1):

$$a^* = \frac{C_{yz}}{C_y} \quad b^* = E\{\mathbf{z}\} - a^* E\{\mathbf{y}\}$$

avec $C_y = \frac{1}{N} \sum_{\alpha=1}^N y_\alpha^2 - \left(\frac{1}{N} \sum_{\alpha=1}^N y_\alpha\right)^2$ et $C_{yz} = \frac{1}{N} \sum_{\alpha=1}^N y_\alpha z_\alpha - \left(\frac{1}{N} \sum_{\alpha=1}^N y_\alpha\right) \left(\frac{1}{N} \sum_{\alpha=1}^N z_\alpha\right)$.

1.3 Ecriture matricielle

On peut aussi ne s'intéresser qu'à la forme algébrique de ces expressions. Soient $\mathbf{y}$ et $\mathbf{z}$ deux vecteurs quelconques de $\mathbb{R}^N$. Soit $\mathbf{1}$ le vecteur dont les composantes sont égales à 1.

Construisons une moyenne sur $\mathbb{R}^N$ en posant :

$$E_N\{\mathbf{y}\} = \frac{\mathbf{1}^\top \mathbf{y}}{N} = \frac{1}{N} \sum_{\alpha=1}^N y_\alpha$$

Puis une covariance :

$$C_N\{\mathbf{y}, \mathbf{z}\} = \frac{\mathbf{y}^\top \mathbf{z}}{N} - E_N\{\mathbf{y}\} E_N\{\mathbf{z}\} = \frac{1}{N} \sum_{\alpha=1}^N y_\alpha z_\alpha - \left(\frac{1}{N} \sum_{\alpha=1}^N y_\alpha\right) \left(\frac{1}{N} \sum_{\alpha=1}^N z_\alpha\right)$$

La régression linéaire du vecteur $\mathbf{z}$ sur le vecteur $\mathbf{y}$ donne les coefficients :

$$a^* = \frac{C_N\{\mathbf{y}, \mathbf{z}\}}{C_N\{\mathbf{y}, \mathbf{y}\}} \quad b^* = E_N\{\mathbf{z}\} - a^* E_N\{\mathbf{y}\} \quad (1.2)$$

2 Krigeage de la dérive

2.1 Krigeage de la moyenne

Soit $Y(x)$ une fonction aléatoire stationnaire (FASt) d'espérance $E\{Y(x)\} = m$ inconnue et de fonction de covariance $E\{(Y(x+h) - m)(Y(x) - m)\} = C(h)$. On suppose $y(x)$ connue en N points de mesures x_α . On montre (cf [4]) que l'espérance m peut être estimée par :

$$M = \sum_{\alpha=1}^N \lambda^\alpha Y_\alpha$$

où les coefficients λ^α sont donnés par le système d'équations :

$$\begin{cases} \sum_{\beta} C_{\alpha\beta} \lambda^\beta + \mu = 0 & \forall \alpha \\ \sum_{\beta} \lambda^\beta = 1 \end{cases} \quad \text{où } C_{\alpha\beta} = C(x_\alpha - x_\beta) \quad (2.1)$$

Devant la linéarité des équations on s'intéresse à la forme algébrique de ces équations. Soit $\mathbf{y}$ le vecteur d'une réalisation de $Y(x)$ dans $\mathbb{R}^N$. Il lui correspond une réalisation m^* de M :

$$m^* = \sum_{\alpha=1}^N \lambda^\alpha y_\alpha$$

Posons :

$\mathbf{C} = [C_{\alpha\beta}]$ matrice de variance-covariance de taille $N \times N$ employée pour le krigeage simple.

$\mathbf{K}$ matrice de krigeage ordinaire : $\mathbf{K} = \begin{pmatrix} \mathbf{C} & \mathbf{1} \\ \mathbf{1}^\top & 0 \end{pmatrix}$

λ vecteur des λ^α définis par (2.1)

Les équations (2.1) s'écrivent $\mathbf{K} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$ d'où :

$$m^* = (\mathbf{y}^\top \ 0) \mathbf{K}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$$

Cette relation définit une forme linéaire sur $\mathbb{R}^N$. On pose alors sur $\mathbb{R}^N$:

$$\begin{aligned} M_N : \mathbb{R}^N &\rightarrow \mathbb{R} \\ \mathbf{y} &\mapsto M_N\{\mathbf{y}\} = (\mathbf{y}^\top \ 0) \mathbf{K}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \end{aligned} \quad (2.2)$$

qu'on appelle krigeage de la moyenne. Cette définition est indépendante de tout contenu probabiliste du vecteur $\mathbf{y}$ bien qu'elle utilise la fonction de covariance $C(h)$.

2.2 Présentation de la méthode de dérive externe

Z est une fonction aléatoire dont on connaît la réalisation z_α en un certain nombre de points de l'espace x_α ($\alpha = 1 \dots N$).

f est une fonction déterministe connue en tout point.

On suppose que Z est la somme d'un résidu Y et d'une dérive d .

$$Z(x) = Y(x) + d(x)$$

où Y est une fonction aléatoire d'ordre 2 stationnaire (FAST) d'espérance nulle de fonction de covariance $C(h)$:

$$E[Y(x)] = 0 \quad \text{et} \quad \text{Cov}(Y(x+h), Y(x)) = C(h) \quad \forall x, h$$

et $d(x) = af(x) + b$ où a et b sont des coefficients inconnus.

Z vérifie donc :

$$\begin{aligned} E[Z(x)] &= af(x) + b & \forall x \\ \text{Cov}(Z(x+h), Z(x)) &= C(h) & \forall x, h \end{aligned}$$

$d(x) = af(x) + b$ est la dérive qu'on distingue de f qui est la fonction de dérive.

2.3 Equations

La méthode de la dérive externe est une application du krigeage universel. Le krigeage universel est décrit dans le chapitre 4 de **La théorie des variables régionalisées et ses applications** de G. Matheron[4]. On y trouve démontrés les résultats suivants : les estimateurs du krigeage universel sont des combinaisons linéaires des Z_α . Ils permettent d'estimer les quantités suivantes :

- Z_x^* estimateur de Z_x : $Z_x^* = \sum_{\alpha \in A} \lambda_Z^\alpha(x) Z_\alpha$
- Y_x^* estimateur du résidu Y_x : $Y_x^* = \sum_{\alpha \in A} \lambda_Y^\alpha(x) Z_\alpha$
- d_x^* estimateur de la dérive d_x : $d_x^* = \sum_{\alpha \in A} \lambda_d^\alpha(x) Z_\alpha$
- a^* estimateur du coefficient a : $a^* = \sum_{\alpha \in A} \lambda_a^\alpha Z_\alpha$
- b^* estimateur du coefficient b : $b^* = \sum_{\alpha \in A} \lambda_b^\alpha Z_\alpha$

Les coefficients $\lambda_{(\dots)}^\alpha$ de ces combinaisons linéaires sont obtenus en inversant un système linéaire d'équations (2.3):

$$\begin{bmatrix} \ddots & & & 1 & \vdots \\ & C_{\alpha\beta} & & \vdots & f_\beta \\ & & \ddots & 1 & \vdots \\ 1 & \dots & 1 & 0 & 0 \\ \dots & f_\alpha & \dots & 0 & 0 \end{bmatrix} \times \begin{bmatrix} \vdots \\ \lambda_{(\dots)}^\alpha \\ \vdots \\ \mu_0(\dots) \\ \mu_1(\dots) \end{bmatrix} = \begin{bmatrix} \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{bmatrix}$$

où $C_{\alpha\beta}$ est la covariance de Y_α et de Y_β soit :

$$\begin{aligned} C_{\alpha\beta} &= cov(Y_\alpha Y_\beta) = cov(Z_\alpha Z_\beta) \\ &= C(x_\alpha - x_\beta) \end{aligned}$$

Dans ces équations la matrice de gauche est toujours la même, seul change le second membre en fonction de l'estimateur cherché. Le second membre pour chaque estimateur est donné ci-dessous :

$$Z_x^* \Rightarrow \begin{bmatrix} \vdots \\ C_{\alpha x} \\ \vdots \\ 1 \\ f_x \end{bmatrix} \quad Y_x^* \Rightarrow \begin{bmatrix} \vdots \\ C_{\alpha x} \\ \vdots \\ 0 \\ 0 \end{bmatrix} \quad d_x^* \Rightarrow \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ f_x \end{bmatrix} \quad a^* \Rightarrow \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad b^* \Rightarrow \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad (2.4)$$

Estimation à distance

Lorsque le point x s'éloigne de l'ensemble des points de mesure x_α les covariances $C_{\alpha x}$ tendent vers 0. En comparant les seconds membres donnés par (2.4) on voit que

$$\lim_{\|x\| \rightarrow \infty} Z^*(x) = d^*(x).$$

Donc, pour les grandes distances, on peut assimiler $Z^*(x)$ et $d^*(x)$.

Nous voyons qu'il est important de comprendre à quoi correspond le terme de dérive. Or nous allons aboutir à une expression simple de ce terme.

2.4 Calcul des coefficient a et b de la dérive

De nouveau, on se place dans $\mathbb{R}^N$ où les différents krigeages sont des formes linéaires. On évacue tout contenu probabiliste et on considère le vecteur z comme étant quelconque. Soit f le vecteur de la fonction de dérive aux points x_α : $f^T = (f_1 \dots f_N)$.

Les équations précédentes, (2.3) et (2.4), donnent :

$$a^* = (z^T \ 0 \ 0) \begin{pmatrix} \mathbf{K} & \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix} \\ (\mathbf{f}^T \ 0) & 0 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0} \\ 0 \\ 1 \end{pmatrix} \quad b^* = (z^T \ 0 \ 0) \begin{pmatrix} \mathbf{K} & \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix} \\ (\mathbf{f}^T \ 0) & 0 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \\ 0 \end{pmatrix} \quad (2.5)$$

Ces expressions sont inutilisables sous cette forme. Il faut les transformer au moyen d'un calcul qu'on retrouvera sous forme analogue dans presque toutes les démonstrations. Il s'appuie sur les propriétés très puissantes des matrices de variance-covariance. Il consiste à poser :

$$\begin{pmatrix} \mathbf{K} & \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix} \\ (\mathbf{f}^T \ 0) & 0 \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A} & \mathbf{v} \\ \mathbf{v}^T & w \end{pmatrix} \quad (2.6)$$

et à écrire $\begin{pmatrix} \mathbf{K} & \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix} \\ (\mathbf{f}^T \ 0) & 0 \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{v} \\ \mathbf{v}^T & w \end{pmatrix} = \begin{pmatrix} \mathbf{I}_{N+1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}$ où $\mathbf{I}_{N+1}$ est la matrice identité de taille $N + 1$.

On obtient les identités suivantes :

$$\begin{cases} \mathbf{KA} + \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix} \mathbf{v}^T = \mathbf{I}_{N+1} \\ \mathbf{Kv} + \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix} w = \mathbf{0} \\ (\mathbf{f}^T \ 0)\mathbf{A} = \mathbf{0}^T \\ (\mathbf{f}^T \ 0)\mathbf{v} = 1 \end{cases}$$

$$\text{Posons } \mathcal{K}_N : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R} \quad (2.7)$$

$$(\mathbf{y}, \mathbf{z}) \mapsto \mathcal{K}_N\{\mathbf{y}, \mathbf{z}\} = (\mathbf{y}^T \ 0) \mathbf{K}^{-1} \begin{pmatrix} \mathbf{z} \\ 0 \end{pmatrix}$$

On trouve les égalités :

$$w = \frac{-1}{\mathcal{K}_N\{\mathbf{f}, \mathbf{f}\}} \quad \mathbf{v} = \frac{\mathbf{K}^{-1} \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix}}{\mathcal{K}_N\{\mathbf{f}, \mathbf{f}\}} \quad \mathbf{A} = \mathbf{K}^{-1} - \frac{\mathbf{K}^{-1} \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix} \left(\mathbf{K}^{-1} \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix} \right)^T}{\mathcal{K}_N\{\mathbf{f}, \mathbf{f}\}} \quad (2.8)$$

Nous pouvons maintenant obtenir de nouvelles expressions pour a^* et b^* au moyen de (2.5), (2.6) et (2.8).

$$a^* = (z^T \ 0 \ 0) \begin{pmatrix} \mathbf{A} & \mathbf{v} \\ \mathbf{v}^T & w \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ 0 \\ 1 \end{pmatrix} = (z^T \ 0 \ 0) \begin{pmatrix} \mathbf{v} \\ w \end{pmatrix} = (z^T \ 0) \mathbf{v} = \frac{(z^T \ 0) \mathbf{K}^{-1} \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix}}{\mathcal{K}_N\{\mathbf{f}, \mathbf{f}\}}$$

$$\boxed{a^* = \frac{\mathcal{K}_N\{\mathbf{z}, \mathbf{f}\}}{\mathcal{K}_N\{\mathbf{f}, \mathbf{f}\}}} \quad (2.9)$$

De même :

$$b^* = (z^T \ 0 \ 0) \begin{pmatrix} \mathbf{A} & \mathbf{v} \\ \mathbf{v}^T & w \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ 1 \\ 0 \end{pmatrix} = (z^T \ 0) \mathbf{A} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$$

$$= (z^T \ 0) \mathbf{K}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} - (z^T \ 0) \frac{\mathbf{K}^{-1} \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix} \left(\mathbf{K}^{-1} \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix} \right)^T}{\mathcal{K}_N\{\mathbf{f}, \mathbf{f}\}} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$$

On reconnaît dans $(\mathbf{z}^\top \ 0)\mathbf{K}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$ (resp. $(\mathbf{f}^\top \ 0)\mathbf{K}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$) le krigeage de la moyenne (2.2) de $\mathbf{z}$ (resp. $\mathbf{f}$). D'où :

$$\boxed{b^* = M_N\{\mathbf{z}\} - a^* M_N\{\mathbf{f}\}} \quad (2.10)$$

2.5 Comparaison avec une régression linéaire

Les coefficients a^* et b^* de la régression linéaire du vecteur $\mathbf{z}$ sur le vecteur $\mathbf{f}$ sont donnés par (1.2) :

$$a^* = \frac{C_N\{\mathbf{f}, \mathbf{z}\}}{C_N\{\mathbf{f}, \mathbf{f}\}} \quad b^* = E_N\{\mathbf{z}\} - a^* E_N\{\mathbf{f}\}$$

Pour la dérive on trouve (2.9) et (2.10) :

$$a^* = \frac{\mathcal{K}_N\{\mathbf{f}, \mathbf{z}\}}{\mathcal{K}_N\{\mathbf{f}, \mathbf{f}\}} \quad b^* = M_N\{\mathbf{z}\} - a^* M_N\{\mathbf{f}\}$$

L'analogie entre les deux expressions est frappante. Montrons qu'en fait il s'agit d'une même expression en comparant successivement les termes exprimant une moyenne, puis ceux exprimant une covariance. Pour cela il faut transformer les expressions de M_N et $\mathcal{K}_N$ au moyen du calcul suivant :

Calcul préliminaire

On sait que $\mathbf{K} = \begin{pmatrix} \mathbf{C} & \mathbf{1} \\ \mathbf{1}^\top & 0 \end{pmatrix}$. Posons¹ $\mathbf{K}^{-1} = \begin{pmatrix} \mathbf{A} & \mathbf{v} \\ \mathbf{v}^\top & w \end{pmatrix}$ et écrivons $\mathbf{K}\mathbf{K}^{-1} = \mathbf{I}_{N+1}$. Nous obtenons les égalités suivantes :

$$\begin{cases} \mathbf{C}\mathbf{A} + \mathbf{1}\mathbf{v}^\top & = \mathbf{I}_N \\ \mathbf{C}\mathbf{v} + \mathbf{1}w & = \mathbf{0} \\ \mathbf{1}^\top\mathbf{A} & = \mathbf{0}^\top \\ \mathbf{1}^\top\mathbf{v} & = 1 \end{cases}$$

d'où :

$$w = \frac{-1}{\mathbf{1}^\top\mathbf{C}^{-1}\mathbf{1}} \quad \mathbf{v} = \frac{\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}^\top\mathbf{C}^{-1}\mathbf{1}} \quad \mathbf{A} = \mathbf{C}^{-1} - \frac{\mathbf{C}^{-1}\mathbf{1}(\mathbf{C}^{-1}\mathbf{1})^\top}{\mathbf{1}^\top\mathbf{C}^{-1}\mathbf{1}}$$

Comparaison des moyennes

Soit $\mathbf{y}$ un vecteur quelconque de $\mathbb{R}^N$. D'après (2.2), puis (1.2) :

$$\begin{aligned} M_N\{\mathbf{y}\} &= (\mathbf{y}^\top \ 0)\mathbf{K}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} = (\mathbf{y}^\top \ 0) \begin{pmatrix} \mathbf{A} & \mathbf{v} \\ \mathbf{v}^\top & w \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} = \mathbf{y}^\top\mathbf{v} \\ &= \frac{\mathbf{y}^\top\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}^\top\mathbf{C}^{-1}\mathbf{1}} \\ E_N\{\mathbf{y}\} &= \frac{\mathbf{y}^\top\mathbf{1}}{N} = \frac{\mathbf{y}^\top\mathbf{1}}{\mathbf{1}^\top\mathbf{1}} \end{aligned}$$

Ces deux expressions sont équivalente si on remplace le produit scalaire canonique de $\mathbb{R}^N$ par le produit scalaire induit par la matrice définie positive $\mathbf{C}^{-1}$.

1. $\mathbf{A}$, $\mathbf{v}$ et w sont des intermédiaires de calcul. Il sont différents de ceux du paragraphe 2.4

Comparaison des variances

Soient $\mathbf{y}$ et $\mathbf{z}$ deux vecteurs quelconques de $\mathbb{R}^N$. D'après (2.7), puis (1.2).

$$\begin{aligned} \mathcal{K}_N\{\mathbf{y}, \mathbf{z}\} &= (\mathbf{y}^\top \ 0) \mathbf{K}^{-1} \begin{pmatrix} \mathbf{z} \\ 0 \end{pmatrix} = (\mathbf{y}^\top \ 0) \begin{pmatrix} \mathbf{A} & \mathbf{v} \\ \mathbf{v}^\top & w \end{pmatrix} \begin{pmatrix} \mathbf{z} \\ 0 \end{pmatrix} = \mathbf{y}^\top \mathbf{A} \mathbf{z} \\ &= \mathbf{y}^\top \mathbf{C}^{-1} \mathbf{z} - \frac{(\mathbf{y}^\top \mathbf{C}^{-1} \mathbf{1})(\mathbf{z}^\top \mathbf{C}^{-1} \mathbf{1})}{\mathbf{1}^\top \mathbf{C}^{-1} \mathbf{1}} \\ &= \mathbf{1}^\top \mathbf{C}^{-1} \mathbf{1} \left(\frac{\mathbf{y}^\top \mathbf{C}^{-1} \mathbf{z}}{\mathbf{1}^\top \mathbf{C}^{-1} \mathbf{1}} - \frac{(\mathbf{y}^\top \mathbf{C}^{-1} \mathbf{1})(\mathbf{z}^\top \mathbf{C}^{-1} \mathbf{1})}{\mathbf{1}^\top \mathbf{C}^{-1} \mathbf{1}} \right) \\ &= \mathbf{1}^\top \mathbf{C}^{-1} \mathbf{1} \left(\frac{\mathbf{y}^\top \mathbf{C}^{-1} \mathbf{z}}{\mathbf{1}^\top \mathbf{C}^{-1} \mathbf{1}} - M_N\{\mathbf{y}\} M_N\{\mathbf{z}\} \right) \end{aligned}$$

$$C_N\{\mathbf{y}, \mathbf{z}\} = \frac{\mathbf{y}^\top \mathbf{z}}{N} - E_N\{\mathbf{y}\} E_N\{\mathbf{z}\} = \frac{\mathbf{y}^\top \mathbf{z}}{\mathbf{1}^\top \mathbf{1}} - E_N\{\mathbf{y}\} E_N\{\mathbf{z}\}$$

Au facteur $\mathbf{1}^\top \mathbf{C}^{-1} \mathbf{1}$ près, les deux expressions sont équivalentes, pourvu qu'on remplace le produit scalaire canonique de $\mathbb{R}^N$ par le produit scalaire induit par la matrice définie positive $\mathbf{C}^{-1}$. Quant au facteur $\mathbf{1}^\top \mathbf{C}^{-1} \mathbf{1}$, il est sans incidence sur le résultat car les covariances n'interviennent dans l'expression de a^* qu'à un facteur multiplicatif près.

Autres résultats

On vient de montrer l'équivalence entre les coefficients d'une régression linéaire dans $\mathbb{R}^N$ et l'estimation des coefficients de la dérive dans la méthode de la dérive externe.

Nous pouvons désormais harmoniser les notations. Soit $\mathbf{y}, \mathbf{z}$ deux vecteurs de $\mathbb{R}^N$. On pose :

$$\begin{aligned} E_N\{\mathbf{y}\} &= (\mathbf{y}^\top \ 0) \mathbf{K}^{-1} \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \end{pmatrix} & \mathcal{K}_N\{\mathbf{y}, \mathbf{z}\} &= (\mathbf{y}^\top \ 0) \mathbf{K}^{-1} \begin{pmatrix} \mathbf{z} \\ 0 \end{pmatrix} \\ C_N\{\mathbf{y}, \mathbf{z}\} &= \mathbf{1}^\top \mathbf{C}^{-1} \mathbf{1} \mathcal{K}_N\{\mathbf{y}, \mathbf{z}\} \end{aligned} \quad (2.11)$$

Notons que lorsque le résidu est un effet de pépite pur, cette régression linéaire redonne la régression linéaire plus classique du paragraphe (1.3).

Dans [1] on montre que les résultats s'étendent au cas de plusieurs fonctions de dérive externe. [1] traite également du cas où le résidu $Y(x)$ n'est plus modélisé par une fonction aléatoire stationnaire mais par une fonction aléatoire intrinsèque. Dans ce cas le coefficient b^* , quoiqu'existant algébriquement, perd son sens. Le coefficient a^* reste inchangé.

3 Influence des différentes variables

La méthode de la dérive externe traite différemment la fonction aléatoire $Z(x)$ et la fonction de dérive f . On connaît des mesures de ces deux variables en N points x_α , ce qui nous fournit N couples (f_α, z_α) . La différence de nature de ces deux variables se traduit par le fait que l'estimation de la dérive est une régression linéaire de z sur f . Une analyse de la forme $\mathcal{K}_N\{\}$ permet d'apprécier l'influence relative de ces variables.

3.1 Poids d'un couple

Soient $\mathbf{y}, \mathbf{z}$ deux vecteurs de $\mathbb{R}^N$. Par commodité de notation on place en première position la $\alpha^{\text{ième}}$ composante. Ainsi $\mathbf{y}^\top = (y_\alpha, \mathbf{y}_\alpha^\top)$ où $\mathbf{y}_\alpha$ est le vecteur de $\mathbb{R}^{N-1}$ contenant les composantes de $\mathbf{y}$ autres que y_α : $\mathbf{y}_\alpha = (y_1 \dots y_{\alpha-1} y_{\alpha+1} \dots y_N)$.

$\mathbf{K}$, matrice de krigeage ordinaire, se décompose en $\mathbf{K} = \begin{pmatrix} c_\alpha & \mathbf{k} \\ \mathbf{k}^\top & \mathbf{K}_\alpha \end{pmatrix}$. Posons $\mathbf{K}^{-1} = \begin{pmatrix} w & \mathbf{v}^\top \\ \mathbf{v} & \mathbf{A} \end{pmatrix}$ et écrivons $\mathbf{K}\mathbf{K}^{-1} = \mathbf{I}_{N+1}$. Nous obtenons les égalités suivantes :

$$\begin{cases} c_\alpha w + \mathbf{k}^\top \mathbf{v} = 1 \\ c_\alpha \mathbf{v}^\top + \mathbf{k}^\top \mathbf{A} = \mathbf{0}^\top \\ \mathbf{k} w + \mathbf{K}_\alpha \mathbf{v} = \mathbf{0} \\ \mathbf{k} \mathbf{v}^\top + \mathbf{K}_\alpha \mathbf{A} = \mathbf{I}_N \end{cases}$$

d'où :

$$w = \frac{1}{c_\alpha - \mathbf{k}^\top \mathbf{K}_\alpha^{-1} \mathbf{k}} \quad \mathbf{v} = \frac{-\mathbf{K}_\alpha^{-1} \mathbf{k}}{c_\alpha - \mathbf{k}^\top \mathbf{K}_\alpha^{-1} \mathbf{k}} \quad \mathbf{A} = \mathbf{K}_\alpha^{-1} + \frac{\mathbf{K}_\alpha^{-1} \mathbf{k} (\mathbf{K}_\alpha^{-1} \mathbf{k})^\top}{c_\alpha - \mathbf{k}^\top \mathbf{K}_\alpha^{-1} \mathbf{k}}$$

On en déduit :

$$\begin{aligned} \mathcal{K}_N\{\mathbf{y}, \mathbf{z}\} &= (\mathbf{y}^\top \ 0) \mathbf{K}^{-1} \begin{pmatrix} \mathbf{z} \\ 0 \end{pmatrix} \\ &= y_\alpha c_\alpha z_\alpha + (\mathbf{y}_\alpha^\top \ 0) \mathbf{v} z_\alpha + y_\alpha \mathbf{v}^\top \begin{pmatrix} z_\alpha \\ 0 \end{pmatrix} + (\mathbf{y}_\alpha^\top \ 0) \mathbf{K}_\alpha^{-1} \begin{pmatrix} \mathbf{z}_\alpha \\ 0 \end{pmatrix} \\ &= w(y_\alpha - (\mathbf{y}_\alpha^\top \ 0) \mathbf{K}_\alpha^{-1} \mathbf{k})(z_\alpha - (\mathbf{z}_\alpha^\top \ 0) \mathbf{K}_\alpha^{-1} \mathbf{k}) + \mathcal{K}_{N-1}^\alpha\{\mathbf{y}_\alpha, \mathbf{z}_\alpha\} \end{aligned}$$

où $\mathcal{K}_{N-1}^\alpha\{\mathbf{y}_\alpha, \mathbf{z}_\alpha\}$ est une forme bilinéaire sur $\mathbb{R}^{N-1}$ induite par $\mathbf{K}_\alpha$ comme pour (2.7) et appliquée aux vecteurs $\mathbf{y}_\alpha$ et $\mathbf{z}_\alpha$.

On reconnaît :

- l'expression algébrique² du krigeage ordinaire de y_α par les $N-1$ points y_β ($\beta \neq \alpha$) : $(\mathbf{y}_\alpha, 0)^\top \mathbf{K}_\alpha^{-1} \mathbf{k}$
- l'erreur algébrique de ce krigeage : $\Delta(y_\alpha|\mathbf{y}) = y_\alpha - (\mathbf{y}_\alpha^\top \ 0) \mathbf{K}_\alpha^{-1} \mathbf{k}$
- la variance de ce krigeage : $\sigma_\alpha^2 = \frac{1}{w} = c_\alpha - \mathbf{k}^\top \mathbf{K}_\alpha^{-1} \mathbf{k}$

on obtient :

$$\boxed{\mathcal{K}_N\{\mathbf{y}, \mathbf{z}\} = \frac{1}{\sigma_\alpha^2} \Delta(y_\alpha|\mathbf{y}) \Delta(z_\alpha|\mathbf{z}) + \mathcal{K}_{N-1}^\alpha\{\mathbf{y}_\alpha, \mathbf{z}_\alpha\}} \quad (3.1)$$

Cette formule sépare dans la forme $\mathcal{K}_N\{\}$ la part du couple (y_α, z_α) des autres couples de données.

2. Ici le krigeage n'a aucun sens probabiliste.

Poids d'un couple sur l'ajustement

Le coefficient a^* de la dérive externe s'écrit (2.9) $a^* = \frac{\mathcal{K}_N\{\mathbf{z}, \mathbf{f}\}}{\mathcal{K}_N\{\mathbf{f}, \mathbf{f}\}}$. Le numérateur exprime une relation linéaire pour $\mathbf{z}$ et pour $\mathbf{f}$.

Le dénominateur vaut $\mathcal{K}_N\{\mathbf{f}, \mathbf{f}\}$, qui est une forme bilinéaire positive. Dans cette expression les différentes composantes f_α de $\mathbf{f}$ ont un poids différent. Or ce poids est évalué par la formule précédente (3.1). Il vaut $\frac{\Delta(f_\alpha|\mathbf{f})^2}{\sigma_\alpha^2}$.

Plus ce facteur est important, plus le couple (f_α, z_α) a d'importance dans l'ajustement de la dérive. L'importance d'un couple sur l'ajustement de la dérive dépend donc essentiellement de la valeur de la fonction de dérive f_α ; il ne dépend de z_α que linéairement.

Régularité de la fonction de dérive

Nous allons montrer que pour qu'un ajustement de dérive ait un sens physique, cela impose des conditions de régularité à la fonction de dérive.

Supposons dans un premier temps que nous souhaitions avoir pour l'ajustement de la dérive des poids équivalents pour chaque couple. Cela impose d'avoir :

$$\forall \alpha = 1 \dots N \quad \frac{\Delta(f_\alpha|\mathbf{f})^2}{\sigma_\alpha^2} = k \quad (3.2)$$

où k est une constante. On reconnaît dans l'expression $\frac{\Delta(f_\alpha|\mathbf{f})^2}{\sigma_\alpha^2}$ le résultat d'une validation croisée des données f_α pour la covariance $C(h)$. L'égalité (3.2) signifie donc que, au coefficient k près, les valeurs de $f(x)$ aux points x_α constituent une "bonne" réalisation du résidu $Y(x)$ de covariance $C(h)$. Ce résultat est surprenant car il impose à $f(x)$ la même régularité qu'au résidu $Y(x)$.

L'exemple qui suit montre qu'en fait, il n'est pas souhaitable que les couples aient le même poids lors de l'ajustement de la dérive. Il montre en plus que la fonction de dérive doit être plus régulière que le résidu pour que l'ajustement garde un sens intuitif.

Exemple

On étudie dans cet exemple le poids d'un bipoint, c'est-à-dire de deux points de mesure très proches, sur l'ajustement de la dérive.

$x_2(z_2, f_2)$	x	$x_2(z_2, f_2)$	Prenons trois points x_1, x_2 et x_3 tels
	x		que x_2 et x_3 sont très proches et x_1 loin
			des deux autres. On connaît en chacun
		$x_3(z_3, f_3)$	de ces points des valeurs de Z et de f .

On suppose que leur matrice de covariance s'écrit

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 - \epsilon \\ 0 & 1 - \epsilon & 1 \end{pmatrix} \quad \text{où } \epsilon \text{ dépend de la distance entre } x_2 \text{ et } x_3.$$

On montre que :

$$\mathcal{K}_3\{\mathbf{z}, \mathbf{f}\} = \frac{1}{\epsilon} \frac{(f_2 - f_3)(z_2 - z_3)}{2} + \frac{2}{4 - \epsilon} \left(f_1 - \frac{f_2 + f_3}{2} \right) \left(z_1 - \frac{z_2 + z_3}{2} \right)$$

Faisons tendre x_2 vers x_3 . Considérons deux cas :

1. f est plus régulière qu'une Fonction Aléatoire de covariance $C(h)$. On a alors :

$$\lim_{x_2 \rightarrow x_3} \frac{(f_2 - f_3)^2}{\epsilon} = 0$$

Dans ce cas :

$$a^* \approx \frac{(z_1 - \frac{z_2+z_3}{2})}{(f_1 - \frac{f_2+f_3}{2})}$$

Le bipoint formé par x_2 et x_3 ne compte plus que comme un seul point portant les valeurs moyennes $\frac{z_2 + z_3}{2}$ et $\frac{f_2 + f_3}{2}$.

2. f est moins régulière qu'une F.A. de covariance $C(h)$. On a alors :

$$\lim_{x_2 \rightarrow x_3} \frac{(f_2 - f_3)^2}{\epsilon} = \infty$$

Dans ce cas :

$$a^* \approx \frac{(z_3 - z_2)}{(f_3 - f_2)}$$

Le bipoint formé par x_2 et x_3 devient prépondérant pour l'ajustement de la dérive et annule l'influence du point x_1 . Cet effet est indésirable.

Cela montre que la fonction de dérive f doit être au moins aussi régulière que le modèle de fonction aléatoire dont on se sert pour modéliser les résidus Y . Ceci correspond à l'intuition qu'on peut avoir d'une dérive qui porte les basses fréquences d'un phénomène alors que les hautes fréquences sont représentées par le résidu.

3.2 Résumé

Finalement, on peut retenir trois points importants concernant la régularité de la fonction de dérive.

1. Pour que les poids les plus forts ne soient pas accordés à des bipoints (qui en théorie apportent une information redondante) il faut veiller à ce que **la fonction de dérive soit plus régulière que la fonction aléatoire des résidus**. Donc n'importe quelle variable f connue dans tout l'espace n'est pas une bonne fonction de dérive.
2. Puisque la dérive est plus régulière que les résidus, **les couples (Z_α, f_α) auront une importance inégale**.
3. Il n'est pas souhaitable qu'un ou deux couples seulement aient autant de poids que les autres réunis. Pour cela, il faudra **se méfier des fortes valeurs de f** , veiller à ce qu'elles ne soient pas isolées.

Les fonctions polynômes, par exemple, sont de bonnes dérivées pour ce qui est de la régularité. Mais pour les employer, il faut se méfier des points isolés qui risquent d'imposer la dérive.

4 Conclusion

La méthode de la dérive externe décompose le krigeage en un krigeage simple des résidus et un krigeage de la dérive. Le krigeage de la dérive se trouve être une régression linéaire de la variable inconnue Z par la fonction de dérive f .

- Une régression linéaire introduit une dissymétrie sur les deux variables mises en jeu. Des conditions de régularité sont requises pour la variable f , ce qui n'est pas le cas de la variable Z .
- Ces conditions demandent à f d'être plus régulière que la fonction aléatoire modélisant les résidus.
- Pour contrôler que lors de l'ajustement certains points ne s'imposent pas trop il faut contrôler les valeurs de f s'écartant de la moyenne. (L'annexe propose une méthode simple pour contrôler l'influence des couples de données sur l'ajustement de la dérive.)

Références

- [1] E. Castelier. (1992) *La dérive externe vue comme une régression linéaire*. Note N-34/92/G, Centre de Géostatistique. Ecole des Mines de Paris.
- [2] P. Delfiner, J.P. Delhomme, J. Pelissier Combescure. (1983) *Application of Geostatistical Analysis to the Evaluation of Petroleum Reservoirs with Well Logs*. 24th Annual Login Symposium Calgary, Canada.
- [3] J.P. Delhomme. (1979) *Etude de la géométrie du réservoir de Chemery*. Rapport interne, Centre d'Informatique Géologique. Ecole des Mines de Paris.
- [4] G. Matheron. (1970) *La théorie des variables régionalisées et ses applications*. Fascicule 5 des *Cahiers du Centre de Morphologie Mathématique de Fontainebleau*. Ecole des Mines de Paris.

Annexe : Application

La dérive est ajustée uniquement à partir de la matrice de covariance $C = [C_{\alpha\beta}]$ et des N couples de données (z_α, f_α) . L'importance d'un couple sur l'ajustement dépend uniquement de la valeur de f_α par rapport aux autres valeurs f_β de f .

Cette importance peut se mesurer si on calcule et compare $\forall \alpha$ les termes $\frac{\Delta(f_\alpha|\mathbf{f})^2}{\sigma_\alpha^2}$,
où

$\Delta(f_\alpha|\mathbf{f})$ est l'erreur de krigeage ordinaire de f_α par les valeurs de f aux autres points x_β et pour la covariance $C(h)$.

σ_α^2 est la variance de ce krigeage.

Un couple aura d'autant plus d'importance dans l'ajustement de la dérive que $\frac{\Delta(f_\alpha|\mathbf{f})^2}{\sigma_\alpha^2}$ sera grand.

Or on peut accéder simplement à ces termes au moyen de l'inverse de la matrice

$$\mathbf{K} = \begin{pmatrix} \ddots & & & 1 \\ & C_{\alpha\beta} & & \vdots \\ & & \ddots & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix}.$$

On montre dans [1] que :

$$\mathbf{K}^{-1} \begin{pmatrix} \vdots \\ f_\alpha \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \vdots \\ \frac{\Delta(f_\alpha|\mathbf{f})}{\sigma_\alpha^2} \\ \vdots \\ m^* \end{pmatrix}$$

D'autre part $\frac{1}{\sigma_\alpha^2}$ est le terme diagonal de $\mathbf{K}^{-1}$ pour la composante α .

$$\frac{\Delta(f_\alpha|\mathbf{f})^2}{\sigma_\alpha^2} = \left(\frac{\Delta(f_\alpha|\mathbf{f})}{\sigma_\alpha^2} \right)^2 \sigma_\alpha^2$$

Le calcul de $\mathbf{K}^{-1}$ permet donc de quantifier l'importance relative des couples (z_α, f_α) lors de l'ajustement de la dérive.

Exemple

Si $C(h)$ est un effet de pépité, l'ajustement de la dérive se ramène à la régression linéaire habituelle.

On trouve dans ce cas :

$$\forall \alpha = 1 \dots N \quad \frac{\Delta(f_\alpha|\mathbf{f})^2}{\sigma_\alpha^2} = \frac{\left(f_\alpha - \frac{1}{N} \sum_{\beta=1}^N f_\beta \right)^2}{N-1}$$