C. 36

EXERCICES SUR LES GRANDES MAILLES

G. MATHERON

Fontainebleau Septembre 1969

EXERCICES SUR LES GRANDES MAILLES

Les trois premiers exercices constituent un test. Si vous êtes capables de comprendre la démarche des deux premiers, et d'effectuer les calculs correspondants (qui sont faciles par eux-mêmes), c'est que vous avez bien assimilé les mécanismes mentaux qui doivent être ceux du géostatisticien. Si de plus vous pouvez accepter sans être troublés les conclusions critiques du troisième exercice (non validité aux grandes mailles du principe de composition des termes de ligne et de tranche, et validité seulement très approximative du principe de composition des variances d'extension élémentaire), celà prouve que votre compréhension s'étend en profondeur, et vous permet de jauger le sens et les limites de chacune des hypothèses d'approximation qui rendent la géostatistique opératoire.

D'un point de vue épistomologique, on notera aussi qu'aux grandes mailles (c'est-à-dire supérieures à la portée) la forme exacte de la covariance ou du $\gamma(h)$ perd toute importance : le $K(h)$ n'intervient plus que par sa valeur C à l'origine, et ses premiers moments (A_0 et A_1 dans R^1, A_0, A_1, A_2 et A_3 dans R^2 ; dans R^3, il faudrait aller jusqu'à A_5). Ainsi, dans R^2 par exemple, tous les schémas se ramènent, aux grandes mailles, à un type unique dépendant seulement des 4 paramètres essentiels A_0, A_1, A_2 et A_3 et du paramètre C (qui n'intervient que comme un facteur). Aux grandes mailles, on est relativement proche des conditions de validité de la statistique classique : à la variance d'estimation C/n à laquelle conduirait cette dernière, le calcul géostatistique ajoute des termes correctifs, dont les coefficients sont précisément les A_i.

A titre de complément, j'ai ajouté un quatrième exercice destiné à montrer combien les problèmes de krigeage se simplifient aux grandes mailles, les implantationsdes différents échantillons perdant alors à peu près toute importance (seul doit être pris en compte le fait qu'un échantillon donné est intérieur ou extérieur à la surface à estimer).

Fontainebleau, Septembre 1969

G. MATHERON

Exercices sur les schémas de type transitif (grandes mailles)

Exercice 1 (1 dimension) :

On désignera par $\gamma(h) = \gamma(r)$ un demi-variogramme (isotrope) de type transitif, c'est-à-dire de la forme :

$$\gamma(r) = C - K(r) \qquad \text{avec} \qquad K(r) = 0 \text{ pour } r > \text{portée.}$$

1 - Dans l'espace à 1 dimension, pour $\ell >$ portée, on a :

$$C - F(\ell) = \frac{2}{\ell^2}\int_0^{\ell}(\ell - x)\,K(x)\,dx = \frac{2}{\ell^2}\int_0^{\infty}(\ell - x)\,K(x)\,dx$$

En déduire :

$$F(\ell) = C - \frac{A_0}{\ell} + \frac{A_1}{\ell^2}$$

$$\chi(\ell) = C - \frac{A_0}{2\ell}$$

aux

$$A_0 = 2\int_0^{\infty} K(x)\,dx, \quad A_1 = 2\int_0^{\infty} x\,K(x)\,dx$$

2 - Variance d'extension pour une maille $a >$ portée

$$\sigma_E^2 = 2\,\chi(\tfrac{a}{2}) - F(a) = C - \frac{A_0}{a} - \frac{A_1}{a^2}$$

Pour une galerie de longueur $\ell = n\,a$, la variance d'estimation est :

$$\frac{1}{n}\,\sigma_E^2 = \frac{a}{\ell}\,C - \frac{A_0}{\ell} - \frac{A_1}{a\,\ell}$$

<u>Exercice 2</u> (2 dimensions)

1 - <u>Fonction F(a,b)</u> - La valeur moyenne de $\gamma(h) = \gamma(r)$ dans le rectangle (a, b) est :

$$(1) \quad F(a,b) = \frac{4}{a^2\ b^2} \int_0^a \int_0^b (a-x)\ (b-y)\ \gamma\ (\sqrt{x^2 + y^2})\ dx\ dy$$

2 - Soit $\gamma(h) = C - K(h)$ un schéma de type transitif (i. e_i : $K(h) = 0$ pour $r = |h| > \varepsilon$, ε désignant la portée) -
<u>Aux grandes mailles</u> (a et b > portée) (1) donne :

$$C - F(a,b) = \frac{4}{a^2\ b^2} \int_0^\infty \int_0^\infty (a-x)(b-y)\ K\ (\sqrt{x^2 + y^2})\ dx\ dy$$

En déduire :

$$F(a,b) = C - \frac{\pi A_1}{ab} + \frac{2\ A_2}{ab}\left(\frac{1}{a} + \frac{1}{b}\right) - A_3 \frac{1}{a^2\ b^2}$$

avec:

$$\begin{cases} A_1 = \dfrac{1}{\pi} \displaystyle\int_{R^2} K(h)\ dh = 2 \int_0^\infty r\ K(r)\ dr \\ A_2 = 2 \displaystyle\int_0^\infty r^2\ K(r)\ dr \\ A_3 = 2 \displaystyle\int_0^\infty r^3\ K(r)\ dr \end{cases}$$

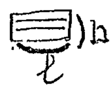
3 - En déduire $\chi(a\ ;\ b)$, $\gamma(a\ ;\ b)$; $Q(a\ ;\ b)$ (toujours pour a,b>portée)

[appliquer $\chi(a, b) = \frac{1}{a} \frac{\partial}{\partial a} \frac{a^2 F}{2}$, $\gamma(a,b) = \frac{\partial^2}{\partial a^2} (\frac{a^2}{2} F)$

$$Q(a, b) = \frac{1}{ab} \frac{\partial^2}{\partial a \, \partial b} \frac{a^2 b^2 F}{4}$$

ceci doit donner : $\chi = C - \frac{\pi A_1}{2a \; b} + \frac{A_2}{a \; b^2}$, $\gamma = C$

$$Q = C - \frac{\pi A_1}{4a \; b}$$]

4/ <u>Variance d'extension d'une galerie</u> ℓ dans $\ell \times h$  ,
ℓ, $\frac{h}{2}$ > portée

$$\sigma_E^2 = 2 \chi (\frac{h}{2}, \ell) - F(h,\ell) - F(\overset{\ell}{\cancel{h}},0) .$$

$$\sigma_E^2 = \frac{A_0}{\ell} - A_1 (\frac{\pi}{\ell h} + \frac{1}{\ell^2}) + 2 A_2 (\frac{1}{h\ell^2} - \frac{1}{h^2 \ell}) + A_3 \frac{1}{h^2 \ell^2}$$

5/ <u>Variance d'estimation</u> de S par galeries de longueurs ℓ_i > portée (équidistance h > portée) - Poser $\Sigma \ell_i = L$ N nombre de galeries

$$\frac{\Sigma \ell_i^2 \sigma_{E_i}^2}{(\Sigma \ell_i)^2} = \frac{A_0}{L} - A_1 (\frac{\pi}{Lh} + \frac{N}{L^2}) + 2 A_2 (\frac{N}{hL^2} - \frac{1}{h^2 L}) + A_3 \frac{N}{h^2 L^2}$$

$$= A_0 \frac{h}{S} - A_1 (\frac{\pi}{S} + \frac{h^2 N}{S^2}) + 2 A_2 (\frac{hN}{S^2} - \frac{1}{hS}) + \frac{A_3 N}{S^2}$$

6/ Estimation de S par galeries (équidistance h > portée) échantillonnées à maille a supérieur à portée :

$$\sigma^2_{E_{st}} = \frac{ah}{S} C - A_1 \left(\frac{h}{a\,S} + \frac{\pi}{S} + \frac{h^2 N}{S^2}\right) + 2\,A_2 \left(\frac{hN}{S^2} - \frac{1}{hS}\right) + \frac{A_3 N}{S^2}$$

7/ Variance d'extension d'un sondage $\boxed{\cdot}$

$$\sigma^2_E = 2\,Q\left(\frac{a}{2}, \frac{a}{2}\right) - F(a,a) = C - \pi \frac{A_1}{a^2} - 4\,\frac{A_2}{a^3} + \frac{A_3}{a^4}$$

d'où variance d'estimation de S :

$$\frac{1}{n}\,\sigma^2_E = C\,\frac{a^2}{S} - \pi\,\frac{A_1}{S} - 4\,\frac{A_2}{aS} + \frac{A_3}{a^2 S}$$

Exercice 3 - (critique)

1/ Jeter un regard critique : pour h = a, la formule de 6/ n'est pas compatible avec le résultat trouvé en 7/ (à part le premier terme $\frac{C}{n}$, qui est l'approximation de la statistique classique). Des deux principes d'approximation :

~ composition des termes de tranche et des termes de ligne (utilisé en 6/)

~ composition des variances d'extensions élémentaires (utilisé en 7/)

l'un au moins est inutilisable dans le cas des grandes mailles.

A la réflexion, il apparait que la variance d'estimation associée à une maille rectangulaire (a,b), dès que a et b sont supérieurs à la portée, ne doit dépendre que de la surface ab du rectangle de maille, et non de son allongement (a/b). Or la formule 6/ fait intervenir explicitement ce rapport a/h. On peut conclure que le principe de composition des termes de tranches n'est pas applicable aux grandes mailles, et que la formule 7/ doit être préférée.

2/ Pour le vérifier de manière rigoureuse, partir de la formule générale :

$$\sigma^2_{E_{st}} = \frac{1}{S^2} \int_S \int_S K(x-y)\, dx\, dy - \frac{2}{NS} \sum_i \int_S K(x_i - y) dy + \frac{1}{N^2} \sum_{ij} K(x_i - x_j)$$

Dès que les distances $|x_i - x_j|$ entre sondages, ainsi que les distances des sondages à la frontière de S sont supérieures à la portée, ceci se réduit à la formule (rigoureuse)

$$\sigma^2_{E_{st}} = C - F(S) + \frac{C}{N} - \frac{2\pi A_1}{S}$$

formule valable indépendamment de la forme de la maille (et applicable, en particulier, même à des mailles irrégulières)

C - F(S) ne dépend que de la surface S (et non de la maille) et a comme partie principale (pour S grand) : $\pi A_1/S$, d'où la formule d'approximation

$$\sigma^2_{E_{st}} \simeq \frac{C}{N} - \pi \frac{A_1}{S}$$

3/ Examinons si le principe de composition des variances d'extension élémentaires donne de meilleurs résultats. On trouve :

$$\frac{1}{N}\left[2\,Q\left(\frac{a}{2},\frac{b}{2}\right) - F(a,b)\right] = \frac{C}{N} - \pi\frac{A_1}{S} - \frac{2\,A_2}{S}\left(\frac{1}{a}+\frac{1}{b}\right) + \frac{A_3}{Sab}$$

Les deux premiers termes sont bons. Les termes d'ordre supérieur (en A_2 et A_3) ne sont pas acceptables, puisqu'ils dépendent encore de a et b, alors que la formule exacte ne fait intervenir que la géométrie de S. Néanmoins, cette formule est meilleure que 6/ (dont le premier terme seul est bon).

4/ Dans le même esprit qu'au 2/ ci-dessus, montrer que la variance d'estimation d'un segment de longueur L par une maille a > portée est (rigoureusement) $C - F(L) + \frac{C}{N} - 2\,\frac{A_0}{L}$.

En déduire que la variance d'estimation à une dimension (Ex. 1/, parag. 2/) est en réalité :

$$\sigma^2_{E_{st}} = \frac{C}{N} - \frac{A_0}{L} - \frac{A_1}{L^2}$$

Ainsi, à une dimension, le principe de composition des variances d'extension donne la valeur exacte des deux premiers termes, mais non celle du troisième.

5/ Conclure du paragraphe précédent que seuls les deux premiers termes de la formule du parag. 5/ Ex.2 sont valables (variance d'estimation de S par des galeries parfaitement échantillonnées)

Pour reconstituer la formule exacte, on supposera que S est le rectangle $\ell \times H$, et on raisonnera directement sur le variogramme γ' déduit de γ par montée sous puissance ℓ. Pour h supérieur à la portée, la fonction F'(h) de ce schéma transitif à 1 dimension est de la forme :

$$F'(h) = C' - \frac{A'_0}{h} + \frac{A'_1}{h^2}$$

(cf. Ex. 1) - Montrer, par un raisonnement direct, que la nouvell constante C' est :

$$C' = C - F(\ell, o) = \frac{A_0}{\ell} - \frac{A_1}{\ell^2}$$

Pour déterminer les nouveaux coefficients A'_0 et A'_1, procéder par identification avec $F(\ell,h)$ (Ex.2, par. 2/), ce qui donne

$$A'_0 = \pi \frac{A_1}{\ell} - 2\frac{A_2}{\ell^2}$$

$$A'_1 = 2\frac{A_2}{\ell} - \frac{A_3}{\ell^2}$$

Montrer, en appliquant le résultat du 4/ ci-dessus, que la variance d'estimation pour des galeries de longueur ℓ équidistantes de $h = \frac{H}{N}$ est :

$$\sigma^2_{E_{st}} = \frac{C'}{N} - \frac{A'_0}{H} - \frac{A'_1}{H^2}$$

$$= \frac{A_0}{N\ell} - A_1\left(\frac{1}{N\ell^2} + \frac{\pi}{S}\right) + 2\frac{A_2}{S}\left(\frac{1}{\ell} - \frac{1}{H}\right) + \frac{A_3}{S^2}$$

Comparer avec Ex. 2, par.5/-

Exercice 4 - Krigeage aux grandes mailles.

On veut kriger une surface S à l'aide de n+k échantillons, dont k sont prélevés à l'intérieur de S et n à l'extérieur, en supposant que les distances entre deux échantillons distincts, et la distance de chaque échantillon à la frontière de S sont toutes supérieures à la portée.

1/ Poser les équations du Krigeage. Montrer qu'elles se réduisent à

$$C \lambda_i = \sigma_{si} + \mu$$

$$\Sigma \lambda_i = 1$$

avec $\sigma_{si} = 0$ pour les échantillons i extérieurs et

$\sigma_{si} = \pi \frac{A_1}{S}$ pour les échantillons intérieurs.

En déduire que les k échantillons intérieurs ont le même poids λ et les n échantillons extérieurs le même poids λ' avec :

$$\lambda = \frac{1}{n+k} + \frac{n}{n+k} \pi \frac{A_1}{CS}$$

$$\lambda' = \frac{1}{n+k} - \frac{k}{n+k} \pi \frac{A_1}{CS}$$

2/ Calculer la variance de ce Krigeage :

$$\left(\sigma_k^2 = C - F(S) - \frac{2\,k}{n+k} \pi \frac{A_1}{S} - \frac{n\,k}{n+k} \left(\frac{\pi A_1}{S}\right)^2 \frac{1}{C} + \frac{C}{n+k}\right)$$

En particulier, si S est grand, on a la formule approchée :

$$\sigma_k^2 \simeq \frac{C}{n+k} + \frac{n-k}{n+k} \frac{\pi A_1}{S}$$