

SUFFIT-IL, POUR UNE COVARIANCE, D'ETRE DE TYPE POSITIF ?

Georges MATHERON*

TABLE DES MATIERES

RESUME	52	ABSTRACT	52
A - INTRODUCTION	53	3 - Les modèles booléens	60
B - UN PROBLEME DE COMPATIBILITE ...	54	D - LES VARIOGRAMMES D'ORDRE 1 ...	61
C - LA COVARIANCE D'UN ENSEMBLE ALEATOIRE	56	E - ANNEXE : La loi lognormale est elle compatible avec le modèle sphérique ?	64
1 - Exemple des trois points ..	57	BIBLIOGRAPHIE	65
2 - La covariance d'un ensemble aléatoire	59		

ILLUSTRATIONS

Tableau 1 - Loi trivariante en tout ou rien	58
Figure 1 - Les trois points	58

Centre de Géostatistique, Ecole des Mines de Paris, 35 rue Saint-Honoré, FONTAINEBLEAU CEDEX, FRANCE.

"ETUDES GEOSTATISTIQUES IV", Séminaire CFSG - 16-17 Juin 1986 - Fontainebleau, France, *in* Sci. de la Terre, Sér. Inf., n° 26.

RESUME

Ce texte pose plus de problèmes qu'il n'en résout : n'importe quelle loi (par exemple log-normale) est-elle compatible avec un variogramme donné (par exemple sphérique) ? Quelles conditions doit vérifier un système de lois bivariées pour être compatible avec une fonction aléatoire ? Comment caractériser la classe des covariances des indicatrices et celle des variogrammes d'ordre 1 ?

MOTS-CLE : distributions bivariées - covariance d'indicatrice - ensemble aléatoire - variogramme d'ordre 1 - variogramme d'ordre α .

ABSTRACT

Is it enough, for a covariance, to be positive ?

More problems are raised than solved in this paper : is any distribution (e.g. lognormal) compatible with a given variogram (e.g. spherical) ? What conditions must a bivariate distributions system satisfy to be compatible with a random function ? How should the class of indicator covariances and that of order 1 variograms be characterized ?

KEY WORDS : bivariate distributions - indicator covariances - random set - order 1 variogram - order α variogram.

A - INTRODUCTION

Par ce titre d'apparence paradoxale, j'entends soulever un problème assez sérieux, qui concerne la cohérence interne des modèles utilisés en Géostatistique. Naturellement, chacun sait qu'une fonction $C(h)$ donnée sur $\mathbb{R}^n$ est la covariance d'une fonction aléatoire stationnaire $Y(x)$ si et seulement si c'est une fonction de type positif. La nécessité de la condition ne pose aucun problème : elle exprime simplement que toute combinaison linéaire $\sum \lambda_i Y(x_i)$ doit avoir une variance positive. Pour montrer que la condition est suffisante, on construit effectivement une fonction aléatoire gaussienne admettant cette covariance. L'apparition de la loi de Gauss n'est pas fortuite, car la construction utilise de manière essentielle une propriété très particulière des distributions multigaussiennes : à savoir que toute matrice C_{ij} définie positive est la matrice de covariance d'une telle distribution.

Le théorème n'affirme donc nullement qu'il existe toujours une F.A. lognormale (par exemple) admettant comme covariance une fonction de type positif donnée. Et ce serait faux en général. En effet, la covariance $C(h)$ d'une F.A. lognormale est nécessairement de la forme :

$$C(h) = m^2 [e^{\sigma(h)} - 1]$$

où $\sigma(h)$ est la covariance d'une F.A. gaussienne. Il ne suffit donc pas que $C(h)$ soit de type positif. Il faut une condition plus forte : c'est maintenant la fonction

$$\sigma(h) = \log \left[1 + \frac{C(h)}{m^2} \right]$$

(et non plus $C(h)$) qui doit être de type positif. Et il s'agit bien d'une condition strictement plus forte. On le voit, par exemple, en cherchant le maximum possible pour le rapport $-C(h)/C(o)$ (effet de trou). Dans le cas à une seule dimension, l'inégalité de Schwarz donne $-\sigma(h) \leq \sigma(o)$ et l'égalité peut être atteinte avec $\sigma(h)$ de la forme $A \cos \alpha h$. Il en résulte pour $C(h)$:

$$-\frac{C(h)}{C(o)} \leq \frac{1 - e^{-\sigma(o)}}{e^{\sigma(o)} - 1} = e^{-\sigma(o)} = \frac{m^2}{m^2 + C(o)}$$

et cette limite supérieure est strictement inférieure à 1. Ainsi, il n'existe pas dans $\mathbb{R}^1$ de F.A. lognormale admettant comme covariance la fonction $\cos \alpha h$, qui est pourtant de type positif.

Concernant le variogramme sphérique, je conjecture qu'il n'est pas compatible avec la log-normalité, bien que je ne sois pas parvenu à le démontrer. On trouvera cependant en Annexe la démonstration du résultat partiel suivant : il existe une valeur finie du coefficient de variation $\sqrt{C(o)}/m$ au-dessus de laquelle il ne peut pas exister de F.A. bilognormale admettant un

variogramme sphérique. Par F.A. bilognormale, on entend que la loi de $\log Z_x$ et $\log Z_y$ est bi-gaussienne. Cette précision est essentielle, car on verra aussi en Annexe cet autre résultat : quelle que soit la valeur du coefficient de variation, il existe toujours une F.A. monolognormale admettant un variogramme sphérique. Il n'y a pas contradiction, car il s'agit cette fois de F.A. monolognormales, c'est-à-dire dont les marginales monovariées sont lognormales, (mais non les marginales bivariées). Ainsi, ce qui est incompatible avec le modèle sphérique, ce n'est pas la monolognormalité, mais, au moins pour les grandes valeurs du coefficient de variation, la bilognormalité. Cet exemple montre que ces problèmes d'incompatibilité peuvent être relativement subtils.

On a un autre exemple simple (sur lequel nous reviendrons) avec les F.A. en tout ou rien, i.e. les indicatrices d'ensembles aléatoires stationnaires. Comme ces indicatrices ne sont jamais dérivables en moyenne quadratique, on voit qu'une fonction de type positif deux fois différentiable, par exemple la fonction

$$C(h) = \exp(-a|h|^2)$$

ne peut pas être la covariance d'un ensemble aléatoire.

B – UN PROBLEME DE COMPATIBILITE

On voit ainsi se poser le problème de la compatibilité d'une covariance avec une loi marginale donnée. N'importe quelle fonction de type positif $C(h)$ n'est pas nécessairement compatible avec une loi donnée. Il y a des restrictions, et les deux exemples précédents montrent que ces restrictions seront vraisemblablement particulièrement fortes pour des lois très dissymétriques et/ou présentant des composantes atomiques importantes : en somme, toujours le problème du zéro et de l'infini.

Plus généralement, d'ailleurs (G. Matheron, 1973, M. Armstrong et G. Matheron, 1986) se pose le problème de la cohérence des systèmes de distributions bivariées : étant donnée, pour chaque couple de points (x,y) , une loi bivariée, définie, par exemple, par sa fonction de répartition :

$$F_{x,y}(z,z') = P(Z_x < z, Z_y < z')$$

existe-t-il effectivement une F.A. $Z(x)$ admettant ces marginales bivariées ? Il s'agit là d'un problème fort important pour les techniques d'estimation non linéaire comme le krigeage disjonctif, et même déjà pour le krigeage des indicatrices : dans ce dernier cas, en effet, la covariance de l'ensemble aléatoire $\{Z(x) < a\}$ est la fonction :

$$C_a(x,y) = F_{x,y}(a,a)$$

et l'on voit déjà se poser le problème de la cohérence du choix des covariances C_a pour les

différentes valeurs de a .

Ces problèmes sont très difficiles, et on ne connaît pas, à l'heure actuelle, de critères généraux. Dans la pratique, on se contente le plus souvent d'utiliser des modèles particuliers, pour lesquels la cohérence est automatiquement assurée. Citons-en trois :

- ~ Le modèle multigaussien anamorphosé, en usage depuis 15 ans (et même beaucoup plus dans le cas particulier lognormal) et ses diverses variantes (G. Matheron, 1971, 1975). Il est compatible avec n'importe quelle marginale monovariante (abstraction faite du problème de l'"anamorphose inverse" s'il y a une composante atomique) et impose relativement peu de restrictions sur le modèle de covariance. Comme on peut lui associer de bons modèles de changement de support, il reste de loin le plus utilisé encore aujourd'hui. Mais il impose un style particulier (multigaussien, précisément) aux distributions bivariées, pas toujours satisfaisant dans le cas de substances très dispersées et/ou à fort effet zéro.
- ~ Le modèle "mosaïque" (G. Matheron, 1983, 1984), compatible lui aussi avec n'importe quelle marginale monovariante et admettant de bons modèles de changement de support. Mais les distributions bivariées sont trop particulières (Z_x et Z_{x+h} égales avec la probabilité $\rho(h)$ ou indépendantes avec la probabilité $1 - \rho(h)$). C'est le seul modèle dans lequel les indicatrices ont le même corrélogramme $\rho(h)$. Mais on ne connaît pas la classe des covariances admissibles. L'exponentielle $\rho(h) = \exp(-a|h|)$ est admissible, car elle correspond à une partition en polyèdres poissonniens, mais le modèle sphérique ne l'est probablement pas.
- ~ Les modèles à lois indéfiniment divisibles, obtenus en régularisant une mesure aléatoire orthogonale par l'indicatrice d'un borélien borné B . On obtient ainsi une covariance $C(h)$ proportionnelle au covariogramme géométrique $K(h) = \int 1_B(x) 1_B(x+h) dx$ (G. Matheron, 1965, p. 165 et sq). Plus généralement d'ailleurs, par superposition, on peut obtenir comme covariance n'importe quelle combinaison linéaire (à coefficient ≥ 0) de covariogrammes géométriques, par exemple un mélange de modèles sphériques.

Il est intéressant de noter que l'on a là une classe assez générale de modèles de covariances (le cône convexe engendré par les covariogrammes géométriques) compatibles avec n'importe quelle loi indéfiniment divisible, même discrète (Poisson, binomiale négative, Sichel, etc...). Nous retrouverons cette classe un peu plus loin. De plus, dans le cas des lois gamma, binomiale négative, ou Poisson, les lois bivariées sont du type isofactoriel à facteurs polynomiaux, d'où une grande facilité pour former des modèles de changement de support (G. Matheron, 1973, 1983), M. Armstrong et G. Matheron, 1986). En particulier, les modèles isofactoriels à loi gamma sont maintenant d'utilisation courante (Ch. Lantuéjoul, 1986).

Concernant, enfin, la classe très générale des modèles isofactoriels que j'ai présentée dans G. Matheron (1985), la recherche de ces conditions de cohérence est encore à faire.

Faute d'avoir trouvé des critères généraux, je me contente dans ce qui suit d'examiner le cas des ensembles aléatoires (cas où la F.A. ne prend que les 2 valeurs 0 et 1). Même dans ce cas très simple, on ne dispose pas de résultats généraux. Il semble d'ailleurs que si l'on

pouvait trouver la condition nécessaire et suffisante que doit vérifier une fonction $C(h)$ pour être la covariance d'un ensemble aléatoire, on aurait déjà avancé d'un grand pas vers la solution du problème général.

C - LA COVARIANCE D'UN ENSEMBLE ALEATOIRE STATIONNAIRE

Si A est un ensemble aléatoire stationnaire dans $\mathbb{R}^n$, et $1_A(x)$ son indicatrice aléatoire, nous poserons dans ce qui suit :

$$p = E [1_A(x)] \quad ; \quad C(h) = E [1_A(x) 1_A(x+h)]$$

$C(h)$ désigne donc la covariance non centrée, ou, ce qui revient au même, la probabilité d'avoir à la fois $x \in A$ et $x+h \in A$. Le variogramme de l'indicatrice est évidemment

$$\gamma(h) = p - C(h)$$

Comme il s'agit d'indicatrices, on a toujours

$$(1) \quad [1_A(x+h) - 1_A(x)]^2 = |1_A(x+h) - 1_A(x)|$$

Autrement dit, le variogramme d'une indicatrice est identique à son variogramme d'ordre 1. D'où une première condition nécessaire (mais non suffisante) que doivent vérifier $\gamma(h)$ ou $C(h)$:

THEOREME 1 - Le variogramme $\gamma(h)$ d'une indicatrice vérifie l'inégalité triangulaire :

$$(2) \quad \gamma(a+b) \leq \gamma(a) + \gamma(b)$$

et, de même, sa covariance vérifie :

$$(3) \quad |C(h) - C(h')| \leq C(o) - C(h-h')$$

En effet, on a toujours :

$$|1_A(x+a+b) - 1_A(x)| \leq |1_A(x+a+b) - 1_A(x+b)| + |1_A(x+b) - 1_A(x)|$$

d'où (2), compte tenu de (1). Avec $h = a + b$ et $h' = b$, il vient

$$\gamma(h) - \gamma(h') \leq \gamma(h - h')$$

et, en échangeant h et h' , $\gamma(h') - \gamma(h) \leq \gamma(h - h')$, d'où (3).

Voici une conséquence simple : si, dans une direction donnée; $\gamma(h)$ a un comportement du type $\gamma(h) \sim A|h|^\lambda$ pour $|h|$ petit, alors nécessairement $\lambda \leq 1$.

En effet, d'après (2), $\gamma(2h) \leq 2 \gamma(h)$ donne, pour $|h|$ petit :

$$2^\lambda |h|^\lambda \leq 2 |h|^\lambda$$

et donc $\lambda \leq 1$.

En particulier, l'indicatrice $1_A(x)$ n'est jamais dérivable en moyenne quadratique (il faudrait $\lambda = 2$). Le cas $\lambda = 1$ correspond à un ensemble dont la frontière admet une surface spécifique d'espérance finie. Pour $\lambda < 1$ strictement, on a une frontière de type "fractal", i.e. dont la dimension de Hausdorff est supérieure à $n - 1$.

Ces inégalités (2) et (3), qui limitent déjà sérieusement les candidats possibles, ne sont cependant pas suffisantes. Voici un exemple.

1 - EXEMPLE DES TROIS POINTS

Considérons simplement le cas de trois points, c'est-à-dire de trois variables X_1, X_2, X_3 en tout ou rien. On a ici :

$$C_{ij} = E(X_i X_j) = P(X_i = X_j = 1)$$

Comme l'espace des événements distincts comprend $2^3 = 8$ éléments, la loi de probabilité trivariante comprend $8 - 1 = 7$ paramètres. En plus des 6 valeurs C_{ij} , introduisons la quantité

$$\omega = P(X_1 = X_2 = X_3 = 1)$$

La loi est alors ainsi définie (Tableau 1). En écrivant que ces 8 quantités doivent être comprises entre 0 et 1, on obtient 8 inégalités du type $\omega \leq f_k(C_{ij})$ et 8 autres du type $\omega \geq g_k(C_{ij})$ et la condition

$$\sup_k g_k(C_{ij}) \leq \inf_k f_k(C_{ij})$$

est donc nécessaire et suffisante pour qu'il existe une loi trivariante en tout ou rien admettant la covariance C_{ij} . C'est là une condition strictement plus forte que de dire que C_{ij} doit être définie positive. Pour le voir, considérons le cas particulier où l'on a :

$$C_{11} = \frac{1}{2} \quad ; \quad C_{12} = C_{23} = \frac{1+r}{4} \quad ; \quad C_{13} = \frac{1-r}{4}$$

TABLEAU 1 - Loi trivariable en tout ou rien				
Valeurs de	X_1	X_2	X_3	Probabilités
	1	1	1	ϖ
	1	1	0	$C_{12} - \varpi$
	1	0	1	$C_{13} - \varpi$
	0	1	1	$C_{23} - \varpi$
	1	0	0	$C_{11} - C_{12} - C_{13} + \varpi$
	0	1	0	$C_{22} - C_{23} - C_{21} + \varpi$
	0	0	1	$C_{33} - C_{31} - C_{32} + \varpi$
	0	0	0	$1 - C_{11} - C_{22} - C_{33} + C_{12} + C_{23} + C_{31} - \varpi$

En termes d'ensemble aléatoire stationnaire : $p = 1/2$, et le point x est le milieu du segment (x_1, x_3) . Dans ce cas, on trouve $\rho \leq 1$, $r \leq 1$ et de plus :

$$1 + 2\rho + r \geq 0 ; 1 - 2\rho + r \geq 0$$

Sur la Figure 1, on voit qu'en axe (ρ, r) , le point représentatif, évidemment intérieur au carré de sommets $(\pm 1, \pm 1)$, doit de plus se trouver au-dessus des deux droites figurées en pointillé.

De leur côté, les conditions pour que la matrice C_{ij} soit définie positive, s'écrivent :

$$\rho^2 \leq 1 ; r^2 \leq 1 ; 1 + r - 2\rho^2 \geq 0$$

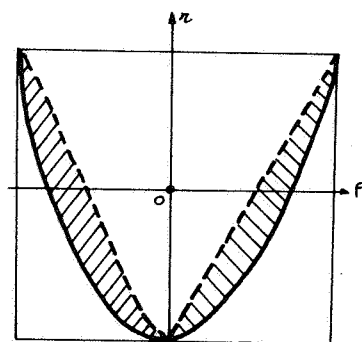


Figure 1 : Les 3 points.

En axes (ρ, r) , le point représentatif, intérieur au carré, doit être au-dessus des 2 droites en pointillé dans le cas en tout ou rien, et au-dessus de la parabole, dans le cas gaussien. L'aire hachurée représente le domaine autorisé dans le cas gaussien, et interdit dans le cas en tout ou rien.

Autrement dit, le point (ρ, r) , toujours intérieur au carré, doit cette fois se trouver au-dessus de la parabole d'équation $r = 2\rho^2 - 1$. Ces conditions sont notablement plus faibles que les précédentes. Par exemple, pour $r = 0$, on trouve :

$$|\rho| \leq 1/2 \quad \text{dans le cas en tout ou rien}$$

$$|\rho| \leq 1/\sqrt{2} = 0,707 \quad \text{dans le cas gaussien.}$$

Remarquons que la condition $1 - 2\rho + r \geq 0$, i.e. $(1-r) \leq 2(1-\rho)$ s'identifie à $\gamma(2h) \leq 2\gamma(h)$, cas particulier de (2). L'autre condition $1 + 2\rho + r \geq 0$ est plus spécifique (elle est liée à la valeur particulière $p = 1/2$). Mais cet exemple suffit pour voir que la relation (2) n'est pas une condition suffisante.

Examinons quelques exemples encore.

2 - EXEMPLE DE L'INDICATRICE GAUSSIENNE

Ici $Y(x)$ est une F.A. gaussienne stationnaire, de moyenne nulle et de variance unité, et on désigne sa covariance (= corrélogramme) par $\rho(h)$. Pour un nombre a donné, on considère l'ensemble aléatoire A défini comme l'ensemble des points x où l'on a $Y(x) \geq a$:

$$A = \{x : Y(x) \geq a\}$$

On désigne par g la densité de la loi de Gauss réduite, et par G sa fonction de répartition. Alors :

$$p = P(Y(x) \geq a) = 1 - G(a)$$

Pour la covariance $C(h)$ (non centrée), il vient :

$$C(h) = P(Y(x+h) \geq a \text{ et } Y(x) \geq a)$$

On l'exprime commodément à l'aide des polynômes d'Hermite usuels :

$$(4) \quad C(h) = p^2 + \sum_{n \geq 1} \frac{\rho^n(h)}{n!} [H_{n-1}(a) g(a)]^2$$

En particulier, dans le cas $p = 1/2$, i.e. $a = 0$, on a le résultat classique

$$C(h) = 1/4 + \text{Arcsin } \rho / 2\pi$$

Pour mettre en évidence le comportement de $C(h)$ au voisinage de $h = 0$, il est plus commode d'utiliser un développement en $(1-p)$. On l'obtient en remarquant que l'évènement $Y_x \geq a$ et $Y_y \geq a$

s'obtient de deux manières (incompatibles et également probables) :

$$Y_x \geq Y_y \geq a \quad \text{ou} \quad Y_y \geq Y_x \geq a$$

En posant

$$Y = \frac{Y_x - Y_y}{\sqrt{2(1-\rho)}}$$

il vient donc $C(h) = 2 P(Y \geq 0 \text{ et } Y_y \geq a)$. Comme le coefficient de corrélation de Y et Y_y est $-\sqrt{(1-\rho)/2}$, il vient :

$$C(h) = 2 \left[\frac{p}{2} + g(o) g(a) \sum_{n \geq 1} \frac{(-1)^n}{n!} \left(\frac{1-\rho}{2} \right)^{n/2} H_{n-1}(o) H_{n-1}(a) \right]$$

Compte tenu de

$$H_{2n+1}(o) = 0 \quad ; \quad H_{2n}(o) = (-1)^n \frac{(2n)!}{n!} 2^{-n}$$

il vient finalement pour le variogramme $\gamma(h) = p - C(h)$:

$$(5) \quad \gamma(h) = g(a) \frac{1}{\sqrt{\pi}} \sqrt{1-\rho} \sum_{n \geq 0} \left(\frac{1-\rho}{4} \right)^n \frac{H_{2n}(a)}{(2n+1)n!}$$

Ainsi, quel que soit le seuil a , pour $\rho = \rho(h)$ voisin de 1, le variogramme $\gamma(h)$ de l'indicatrice se comporte comme la racine carrée du variogramme $1 - \rho(h)$ de la gaussienne. En particulier, c'est seulement dans le cas où $Y(x)$ est différentiable en moyenne quadratique que l'ensemble aléatoire A a une surface spécifique d'espérance finie. Dans tous les autres cas, sa frontière est du type "fractal". On voit la conséquence pratique : il ne serait pas cohérent d'attribuer à la gaussienne et à l'indicatrice un variogramme du même type, par exemple sphérique. Si le variogramme de la gaussienne se comporte comme $|h|$ au voisinage de l'origine, celui de l'indicatrice doit se comporter en $|h|^{1/2}$. On notera aussi sur ces formules (4) ou (5) que les variogrammes des indicatrices, pour les différentes valeurs de a , ne sont nullement indépendants les uns des autres.

3 - LES MODELES BOOLEENS

Ces modèles d'ensembles aléatoires sont étudiés en détail dans G. Matheron (1975) et J. SERRA (1978). Ils consistent à partir d'un processus de Poisson dans $\mathbb{R}^n$ de densité θ , dont les points seront appelés germes. Puis, en chacun de ces germes, indépendamment des autres, on plante des compacts aléatoires appelés grains primaires. On trouve pour le complémentaire de l'union des grains primaires :

$$p = e^{-\theta K(o)} \quad ; \quad C(h) = p^2 e^{\theta K(h)}$$

où $K(h)$ est l'espérance du covariogramme géométrique du grain primaire. Si le grain primaire est convexe, ce modèle possède d'agréables propriétés "semi-markoviennes". Dans le cas général (grain primaire quelconque), $\theta K(h)$ appartient à une classe que nous avons déjà rencontrée plus haut : celle du cône convexe engendré par les covariogrammes géométriques, dont les éléments sont des covariances (centrées) compatibles avec n'importe quelle loi indéfiniment divisible. En exponentiant les éléments de cette classe, on obtient les covariances (non centrées) des modèles booléens (au facteur p^2 près).

D - LES VARIOGRAMMES D'ORDRE 1

Nous avons déjà remarqué que le variogramme d'ordre 1 d'une indicatrice coïncide avec son variogramme usuel (d'ordre 2). D'une façon générale, si $Z(x)$ est une F.A. à accroissements stationnaires, on définit son variogramme d'ordre α en posant :

$$\gamma_{\alpha}(h) = \frac{1}{2} E \left(|Z(x+h) - Z(x)|^{\alpha} \right)$$

On s'attend donc à ce qu'il y ait un rapport étroit entre la classe des variogrammes d'ordre 1 et celle des variogrammes (d'ordre 1 ou 2) des indicatrices. En fait, nous allons montrer la proposition suivante :

THEOREME 2 - La classe des variogrammes d'ordre 1 est identique au cône convexe fermé engendré par les variogrammes d'indicatrices.

Il est facile de vérifier que la classe des variogrammes d'ordre 1 est fermée (i.e. toute limite de variogramme d'ordre 1 est encore un variogramme d'ordre 1). Il suffit donc de montrer que tout variogramme d'ordre 1 est limite de combinaisons linéaires à coefficients positifs de variogrammes d'indicatrice. Comme toute fonction aléatoire intrinsèque est limite de F.A. stationnaire, il suffit même de démontrer la propriété dans le cas stationnaire. Or cela résulte de la propriété suivante :

THEOREME 3 - Si $Z(x)$ est une F.A. stationnaire admettant le variogramme d'ordre 1 :

$$\gamma_1(h) = \frac{1}{2} E [|Z_{x+h} - Z_x|]$$

on a

$$(6) \quad \gamma_1(h) = \int_{-\infty}^{+\infty} \gamma(h; z) dz$$

où $\gamma(h; z)$ est le variogramme de l'indicatrice $1_{Z(x) \geq z}$

En effet, par définition de $\gamma(h; z)$ et compte tenu de (1), on a :

$$\gamma(h; z) = \frac{1}{2} E \left[|1_{Z_{x+h} > z} - 1_{Z_x > z}| \right]$$

L'expression figurant sous le signe E prend la valeur 1 lorsque z est compris entre Z_{x+h} et Z_x , et 0 dans tous les autres cas. Intégrons en z. D'après le théorème de Fubini, il est légitime d'intervertir les symboles d'intégration et d'espérance. Si nous intégrons sous le symbole E, nous trouvons d'après la remarque ci-dessus :

$$\int |1_{Z_{x+h} > z} - 1_{Z_x > z}| dz = |Z_{x+h} - Z_x|$$

Par suite, il vient bien $\int \gamma(h; z) dz = \gamma_1(h)$.

Ce résultat très simple mérite réflexion. Il semble indiquer que si l'on trouvait un critère simple pour caractériser les variogrammes d'ordre 1, on ne serait pas loin de pouvoir aussi caractériser la classe des variogrammes des indicatrices. Malheureusement, je ne connais pas ce critère. Voici cependant une condition nécessaire :

CRITERE 1 - Tout variogramme d'ordre 1 est un variogramme d'ordre 2 (i.e. de type négatif conditionnel) et vérifie l'inégalité triangulaire :

$$\gamma_1(a + b) \leq \gamma_1(a) + \gamma_1(b)$$

Ceci est une simple conséquence du Théorème 2 et de la relation (2). On peut aussi remarquer que $\gamma_1(h)$ est la norme $\|Z_{x+h} - Z_x\|$ au sens de l'espace L_1 , d'où l'inégalité triangulaire.

Voici maintenant une condition suffisante :

CRITERE 2 - Si $\gamma(h)$ est un variogramme, alors $\gamma_1 = \sqrt{\gamma}$ est un variogramme d'ordre 1.

Il suffit, en effet, de considérer une F.A. intrinsèque gaussienne admettant le variogramme : son variogramme d'ordre 1 est $\sqrt{\gamma/\pi}$.

Cette condition suffisante n'est d'ailleurs nullement nécessaire. Par exemple, dans l'espace à une dimension, il est facile de construire un ensemble aléatoire dont l'indicatrice admet le variogramme (d'ordre 1 et 2 à la fois)

$$\gamma_1(h) = \frac{1}{4} \text{Min}(1, |h|)$$

Mais le carré de γ_1 , soit, à un facteur près, $\text{Min}(1, |h|)^2$ ne peut pas être un variogramme : car c'est une fonction deux fois différentiable en $|h| = 0$. Donc, si une F.A. admettait ce variogramme, elle serait dérivable en moyenne quadratique. Par suite, son variogramme serait encore deux fois différentiable en $h = 1$, ce qui n'est pas le cas.

1 - LES VARIOGRAMMES D'ORDRE α

Ces propriétés se généralisent comme suit. Conditions nécessaires :

Tout variogramme d'ordre $\alpha > 0$ quelconque vérifie l'inégalité :

$$[\gamma_\alpha(a+b)]^{1/\alpha} \leq [\gamma_\alpha(a)]^{1/\alpha} + [\gamma_\alpha(b)]^{1/\alpha}$$

Cela résulte du fait que $[\gamma_\alpha(h)]^{1/\alpha}$ s'identifie à la norme $\|Z_{x+h} - Z_x\|$ au sens de l'espace L_α

THEOREME 4 - Tout variogramme d'ordre α , avec $0 < \alpha < 2$, est un variogramme d'ordre 2 (i.e. de type négatif conditionnel).

La démonstration, un peu plus longue, figure dans G. Matheron (1982). J'en rappelle les principales étapes :

Partant d'une fonction $f(x)$ quelconque, on définit une F.A. $Z_1(x)$ en posant :

$$Z_1(x) = f^N(x) \quad ; \quad N : \text{poissonienne d'espérance } \theta$$

Divisant ensuite $Z_1(x)$ par son écart type, on obtient une F.A. $Z_2(x)$ admettant la covariance :

$$C(x,y) = e^{-\frac{\theta}{2} [f(x) - f(y)]^2}$$

L'étape suivante consiste à remplacer $f(x)$ par une F.A. stationnaire ou intrinsèque quelconque, et à passer à l'espérance. Il apparaît ainsi que l'expression

$$E \left[e^{-a(Z_{x+h} - Z_x)^2} \right]$$

est une covariance pour tout $a > 0$. On remplace ensuite a par une variable aléatoire T admettant la transformée de Laplace $E(e^{-\lambda T}) = e^{-b\lambda^{\alpha/2}}$ (loi stable de paramètre $\alpha/2$). On trouve alors que

$$E \left[e^{-b|Z_{x+h} - Z_x|^\alpha} \right]$$

est une covariance pour tout $b > 0$, et donc

$$E \left[(1 - e^{-b|Z_{x+h} - Z_x|^\alpha})/b \right]$$

est un variogramme. Faisant ensuite tendre b vers 0, on obtient le résultat voulu : γ_α est un variogramme pour $0 < \alpha < 2$.

Enfin, une condition suffisante :

Si γ est un variogramme d'ordre 2, $\gamma^{\alpha/2}$ est un variogramme d'ordre α ($\alpha > 0$).

En effet, pour une F.A. gaussienne admettant le variogramme γ , on trouve sans peine que $E[|Z_{x+h} - Z_x|^\alpha]$ est proportionnel à $(\gamma(h))^{\alpha/2}$

Comme corollaire : si γ est un variogramme d'ordre 2, γ^β est encore un variogramme d'ordre 2 pour $0 < \beta < 1$.

En effet, γ^β est alors un variogramme d'ordre 2β , d'après ce qui précède, donc aussi un variogramme d'ordre 2, puisque $0 < 2\beta < 2$.

E - ANNEXE : LA LOI LOGNORMALE EST-ELLE COMPATIBLE AVEC LE MODELE SPHERIQUE ?

Considérons une fonction réelle $R(h)$ définie positive et continue sur $\mathbb{R}^n$, et telle que $R(0) = 1$ (un "modèle de covariance"). D'après le théorème de Bochner, $R(h)$ est la fonction caractéristique (transformée de Fourier) d'une loi de probabilité symétrique sur $\mathbb{R}^n$. En général, cette loi n'est pas indéfiniment divisible. On a alors le théorème suivant :

THEOREME - Si $R(h)$ n'est pas la fonction caractéristique d'une loi indéfiniment divisible sur $\mathbb{R}^n$, alors il existe une valeur finie du coefficient de variation $C(0)^{1/2}/m$ au-dessus de laquelle il ne peut pas exister de F.A. bilognormale admettant une covariance centrée $C(h)$ proportionnelle à $R(h)$.

En effet, pour toute F.A. stationnaire $Z(x)$ bilognormale, i.e. telle que la loi de (Z_x, Z_y) soit lognormale bivariable pour tout x et tout y dans $\mathbb{R}^n$, la covariance centrée $C(h)$ est de la forme

$$C(h) = m^2 [e^{\sigma(h)} - 1]$$

où $\sigma(h)$ est elle-même une covariance. Considérons la fonction :

$$\phi(h) = e^{[\sigma(h) - \sigma(0)]} = \frac{C(h) + m^2}{C(0) + m^2}$$

Si $C(h)$ est continue, il en est de même de $\phi(h)$ et de $\sigma(h)$. Comme $\sigma(h)$ est de type positif, il résulte des théorèmes classiques que $\phi(h)$ est la fonction caractéristique d'une loi indéfiniment divisible (I.D.).

Supposons que nous ayons

$$C(h)/C(0) = R(h)$$

Il vient alors :

$$\phi(h) = \frac{R(h) + 1/V^2}{1 + 1/V^2}$$

où $V^2 = C(0)/m^2$ est le carré du coefficient de variation. S'il existe une suite de F.A. bilognormales dont les coefficients de variation V_n tendent vers l'infini et dont les covariances centrées $C_n(h)$ sont proportionnelles à $R(h)$, alors la suite

$$\phi_n(h) = \frac{R(h) + 1/V_n^2}{1 + 1/V_n^2}$$

converge vers $R(h)$, continue par hypothèse. Comme chaque $\phi_n(h)$ est la fonction caractéristique d'une loi I.D., il résulte alors des théorèmes classiques que $R(h)$ est elle-même la fonction caractéristique d'une loi I.D. D'où le théorème.

On sait aussi, d'après les théorèmes classiques, que la transformée de Fourier d'une loi I.D. ne s'annule pas. Donc le théorème s'applique à toute covariance $R(h)$ non strictement positive pour tout h , et, en particulier au modèle sphérique (et, plus généralement, aux combinaisons linéaires finies de covariogrammes géométriques).

REMARQUE - On sait aujourd'hui que la loi lognormale est indéfiniment divisible (Bondesson, 1979). Donc, comme on l'a vu, il est toujours possible, à partir d'une mesure aléatoire auto-orthogonale, de construire une F.A. monolognormale admettant une moyenne et une variance arbitraire et une covariance proportionnelle à un covariogramme géométrique $K(h)$ donné, par exemple sphérique. Mais cette F.A. monolognormale n'est pas bilognormale, de sorte que ce résultat ne contredit pas le théorème.

En effet, pour toute F.A. du type envisagé ici, c'est-à-dire

$$Z(x) = \mu(B_x)$$

où B_x est le translaté par x d'un borélien borné B et μ une mesure aléatoire auto-orthogonale à loi I.D., on a une règle de régression linéaire :

$$E[Z_x | Z_y] = \rho Z_y + (1-\rho)m$$

Et cela est incompatible avec la bilognormalité, qui entraîne une régression en fonction puissance :

$$E[Z_x | Z_y] = A(Z_y)^\alpha$$

BIBLIOGRAPHIE

- ARMSTRONG, M. et MATHERON, G. (1986) - "Disjunctive Kriging Revisited". (A paraître dans : Mathematical Geology).
- BONDESSON, L. (1979) - "A General Result on Infinite Divisibility". The Annals of Probability, Vol. 7, n° 6, pp. 965-979.
- LANTUEJOUL, Ch., LAJAUNIE, Ch., RIVOIRARD, J. (1986) - "Testing Several Change of Support Models Used to Estimate Global Reserves". (A paraître dans : Geostatistical Case Studies, Reidel Publishing Co., Dordrecht, Netherlands.
- MATHERON, G. (1965) - "Les Variables Régionalisées et leur Estimation". Masson, Paris.
- MATHERON, G. (1974) - "Les Fonctions de Transfert des Petits Panneaux". Rapport interne, Centre de Géostatistique, Fontainebleau, n° N-395.
- MATHERON, G. (1975-a) - "Forecasting Block Grade Distributions : the Transfer Functions". Dans Proc., NATO A.S.I. : Advanced Geostatistics in the Mining Industry, pp. 237-251, D. Reidel Publishing Co., Dordrecht, Netherlands, 1976.
- MATHERON, G. (1975-b) - "Random Sets and Integral Geometry".- J. Wiley and Sons, New York.
- MATHERON, G. (1973) - "Le Krigeage Disjonctif". Rapport Interne, Centre de Géostatistique, Fontainebleau, n° N-360.
- MATHERON, G. (1982) - "La Destructuration des Hautes Teneurs et le Krigeage des Indicatrices". Rapport Interne, Centre de Géostatistique, Fontainebleau, n° N-761.
- MATHERON, G. (1983) - "Isotensorial Models and Change of Support". Dans : Proc., 2nd NATO A.S.I., "Geostatistics for Natural Resources Characterization", Part 1, pp. 449-467.- Reidel Publishing Co., Dordrecht, Netherlands, 1984.
- MATHERON, G. (1984) - "Changement de Support en Modèle Mosaïque". Sciences de la Terre, Série Informatique Géologique, n° 20, Nov. 1984, pp. 435-454. Colloque : Computers in Earth Sciences for Natural Resources Characterization, Nancy, 9-13 Avril 1984.
- MATHERON, G. (1985) - "Une Méthodologie Générale pour les Modèles Isotensoriels Discrets". Sciences de la Terre, Série Informatique Géologique, n° 21, Nov. 1984, pp. 1-64. Séminaire C.F.S.G.-CESMAT : "Etudes Géostatistiques", 14-15 Juin 1984, Fontainebleau.
- SERRA, J. (1982) "Image Analysis and Mathematical Morphology". Academic Press, London.