FONTAINEBLEAU/CG N-30/86/G

SUR LA POSITIVITE DES POIDS DE KRIGEAGE

G. MATHERON

Octobre 1986

SUR LA POSITIVITE DES POIDS DU KRIGEAGE

Table des Matières

SUR LA POSITIVITE DES POIDS DU KRIGEAGE

G. MATHERON

SUR LA POSITIVITE DES POIDS DE KRIGEAGE

0. INTRODUCTION

J'aborde ici la question de savoir s'il existe des covariances assurant automatiquement la positivité des poids du krigeage pour toutes les configurations possibles constituées par les points informés et le point à estimer. Le krigeage en question peut être le krigeage simple (KS), le krigeage ordinaire (KO) ou plus généralement universel (KU) ou intrinsèque (KI), la covariance étant, le cas échéant, remplacée par un variogramme ou une covariance généralisée. Dans le cas d'une covariance stationnaire, un problème annexe consiste à savoir si les poids de l'estimateur optimal de la moyenne sont, eux aussi, positifs pour toutes les configurations.

On reconnait là les deux problèmes classiques de la théorie du potentiel : étant donné un noyau $C(x,y)$ sur $\mathbb{R}^n \times \mathbb{R}^n$ (éventuellement de la forme $C(x-y)$) :

i/ pour tout ouvert borné B et tout point $x_o \notin \overline{B}$, existe-t-il une mesure positive $\lambda_{x_o}(d\xi)$ sur $\overline{B}$ telle que l'on ait

$$\int_{\overline{B}} \lambda_{x_o}(d\xi)\, C(\xi,y) = C(x_o,y)$$

pour tout (ou presque tout) y sur $\overline{B}$: c'est le problème du balayage.

ii/ de même, pour tout ouvert borné B, existe-t-il une mesure positive $W(d\xi)$ sur $\overline{B}$ réalisant le potentiel d'équilibre, i.e. :

$$\int_{\overline{B}} W(d\xi)\, C(\xi,y) = 1$$

pour tout, ou presque tout, $y \notin \overline{B}$: c'est le problème du potentiel d'équilibre.

La seule différence est qu'au lieu d'ensembles ouverts B, nous considérons des ensembles finis $\{x_\alpha,\ \alpha = 1,2,\dots,k\}$. Cela suffit à exclure les potentiels classiques, du moins dans le cas $n \geq 2$. Par exemple, dans $\mathbb{R}^3$, le potentiel newtonien $C(x-y) = 1/|x-y|$ ou le potentiel besselien

$C(x-y) = e^{-a|x-y|}/|x-y|$ vérifient bien les conditions i/ et ii/ ci-dessus, mais ils ne constituent pas des covariances, car C(h) devient infini en h = 0 : si l'on remplace l'ouvert B par un ensemble fini, les problèmes i/ et ii/ n'ont plus de sens. (Notons toutefois que ces noyaux, newtonien ou besselien, peuvent cependant être interprétes comme les mesures-covariances de mesures aléatoires stationnaires). Par contre, dans le cas à une seule dimension n = 1, le potentiel newtonien correspond au variogramme linéaire $|x-y|$, et le potentiel besselien à la covariance exponentielle $e^{-a|x-y|}$: ce variogramme et cette covariance assurent tous deux la positivité des poids de krigeage.

En fait, ces problèmes de théorie du potentiel admettent, comme on sait, une interprétation simple en termes de processus de Markov. En particulier, la mesure de balayage $\lambda_{x_o}(d\xi)$ d'un point x_o sur $\bar{B}$ représente la loi d'entrée d'un processus X_t dans B (loi de X_t à l'instant aléatoire T où l'on a pour la première fois $X_t \in \bar{B}$). Pour un potentiel newtonien, ce processus X_t est le mouvement brownien et, pour un potentiel besselien, le même mouvement brownien stoppé à un temps d'arrêt aléatoire à loi exponentielle. On peut naturellement envisager bien d'autres processus markoviens et leur associer leur noyau potentiel.

Cette interprétation suggère que c'est seulement dans le cas à une seule dimension que l'on peut trouver des covariances sur $\mathbb{R}^n$ assurant la positivité des poids de krigeage : car c'est seulement à une dimension qu'il y a une probabilité non nulle pour que le processus atteigne (ou, plus exactement, passe infiniment près d') un point donné en un temps fini.

1. LE CAS FINI, KRIGEAGE SIMPLE

Soit σ_{ij} une matrice N x N, symétrique et strictement définie positive (i.e. $b^i \sigma_{ij} b^j \geq 0 \ \forall b^i$) avec égalité si et seulement si $b^i = 0$), que l'on peut interpréter comme la covariance de N variables aléatoires Z_i d'espérance nulle. Nous dirons que cette matrice est fortement positive si elle vérifie de plus les deux propriétés suivantes :

i/ Pour tout sous-ensemble $A = \{\alpha, \beta \ldots\}$ de l'ensemble $I = \{1,2,..N\}$ des indices i et tout indice $u \notin A$, les poids λ_u^α du krigeage simple de Z_u sur les Z_α , $\alpha \in A$ sont ≥ 0.

ii/ De même, les coefficients λ^{α}_{D} de l'estimateur optimal $\lambda^{\alpha}_{D} Z_{\alpha}$ de l'espérance sont ≥ 0 pour tout $A \subset I$.

1.1. CRITERE DE POSITIVITE FORTE.

Si l'on désigne par $Q^{\alpha\beta}$ l'inverse de la sous-matrice $\sigma_{\alpha\beta} (\alpha, \beta \in A)$, la condition i/ équivaut à :

(1-1) $$Q^{\alpha\beta} \sigma_{\beta u} \geq 0 \qquad (\alpha \in A,\ u \notin A)$$

et la condition ii/ à :

(1-2) $$\sum_{\beta \in A} Q^{\alpha\beta} \geq 0 \qquad (\alpha \in A)$$

Pour obtenir une caractérisation plus synthétique, nous partirons de l'identité suivante, où B^{ij} désigne l'inverse de la matrice $N \times N$ σ_{ij} : quels que soient les z_i, $i = 1,2,\ldots,N$, on a :

(1-3) $$z_i B^{ij} z_j = z_\alpha Q^{\alpha\beta} z_\beta + (z_u - \lambda^{\alpha}_{u} z_\alpha) B^{uv} (z_v - \lambda^{\beta}_{v} z_\alpha)$$

Dans cette écriture, on utilise la convention de sommation sur les indices répétés et de plus : les indices i,j parcourent l'ensemble complet $I = \{1,2,\ldots,N\}$, mais les indices α, β ne parcourent qu'un sous-ensemble $A \subset I$, et les indices u, v l'ensemble complémentaire. $Q^{\alpha\beta}$ est l'inverse de la sous-matrice $\sigma_{\alpha\beta}$ (donc, en général, $Q^{\alpha\beta} \neq B^{\alpha\beta}$), et λ^{α}_{u} est le poids de Z_α dans le krigeage de Z_u.

Pour démontrer la relation (1-3), il suffit de considérer le cas où les Z_i ont une distribution multigaussienne. En écrivant que la densité $f(z_i)$ des Z_i est égale au produit de la densité $f_A(z_\alpha)$ de la loi marginale des Z_α par la densité $f(z_u | z_\alpha)$ de la loi conditionnelle des Z_u, $u \notin A$, soit :

$$f(z_i) = f_A(z_\alpha)\ f(z_u | z_\alpha)$$

et en passant aux logarithmes, on obtient, en effet, l'identité (1-3). De plus, on remarque que la matrice B^{uv} (restriction de B^{ij} aux indices $u \in A$) est <u>l'inverse de la matrice des covariances des résidus</u> $Z_u - \lambda^{\alpha}_{u} Z_\alpha$:

$$(1\text{-}4) \qquad B^{wv}\, S_{vu} = \delta^{w}_{u} \quad ; \quad S_{vu} = \sigma_{vu} - \lambda^{\alpha}_{u}\, \sigma_{\alpha\beta}\, \lambda^{\beta}_{v}$$

Incidemment, on note aussi que les quantités :

$$z^{*}_{u} = \lambda^{\alpha}_{u}\, z_{\alpha}$$

réalisent le minimum de $z_i\, B^{ij}\, z_j$ lorsque les z_α , $\alpha \in A$ sont donnés.

En identifiant les coefficients de chacun des monomes $z_i\, z_j$, on déduit de (1-3) les relations suivantes :

$$(1\text{-}5) \qquad B^{\alpha v} = -\, \lambda^{\alpha}_{u}\, B^{uv}$$

$$(1\text{-}6) \qquad B^{\alpha\beta} = Q^{\alpha\beta} + \lambda^{\alpha}_{u}\, B^{uv}\, \lambda^{\beta}_{v}$$

qui, avec (1-4) permettent de reconstituer B^{ij} à partir de $Q^{\alpha\beta}$. En sens inverse, on aura

$$(1\text{-}5') \qquad \lambda^{\alpha}_{u} = -\, B^{\alpha v}\, S_{vu}$$

$$(1\text{-}6') \qquad Q^{\alpha\beta} = B^{\alpha\beta} - B^{\alpha u}\, S_{uv}\, B^{\beta v}$$

Voici maintenant notre critère :

<u>CRITERE 1</u> - Pour qu'une matrice symétrique et régulière σ_{ij} soit fortement positive, il faut et il suffit que son inverse B^{ij} vérifie les deux conditions suivantes :

$$(1\text{-}7) \qquad B^{ij} \leqq 0 \quad \text{pour } i \neq j \qquad (i,j \in I)$$

$$(1\text{-}8) \qquad \sum_j B^{ij} \geqq 0 \qquad (i \in I)$$

Ces conditions sont nécessaires : tout d'abord, (1-8) exprime que les quantités $\lambda^{i}_{c} = \sum_j B^{ij}$, solution de $\lambda^{i}_{c}\, \sigma_{ij} = 1$ sont $\geqq 0$, i.e. la condition ii/ relative à l'ensemble de tous les indices. Ensuite, pour un indice $x \in I$ prenons pour $A = \{\alpha, \beta, \dots\}$ l'ensemble de tous les autres indices $i \neq x$. Dans la formule (1-5'), le second membre ne comprend qu'un terme $u = v = x$, et donc

$$\lambda^{\alpha}_{x} = -\, B^{\alpha x}\, \sigma^{2}_{K}(x)$$

où $\sigma^{2}_{K}(x) = 1/B^{xx}$ est la variance de krigeage, donc > 0 (strictement, puisque la matrice σ_{ij} est régulière). Donc la positivité des λ_x^α pour $\alpha \neq x$ entraine bien $B^{\alpha x} \leqq 0$.

Montrons maintenant que les conditions sont suffisantes. D'après la première partie de la démonstration, il suffit de montrer que pour tout sous-ensemble $A \subset I$ l'iverse $Q^{\alpha\beta}$ de la sous-matrice $\sigma_{\alpha\beta}$ vérifie encore les conditions (1-7) et (1-8). En procédant par récurrence, il suffit même de prouver qu'il en est ainsi lorsque l'on enlève un seul indice $x \in I$.

Soit donc $x \in I$ et $A = I \setminus \{x\}$ l'ensemble de tous les autres indices. D'après (1-6'), on a donc :

$$Q^{\alpha\beta} = B^{\alpha\beta} - \frac{B^{\alpha x}\, B^{\beta x}}{B^{xx}}$$

Comme B^{xx} est > 0, et $B^{\alpha x}$ et $B^{\beta x}$ négatifs ou nuls, on a

$$Q^{\alpha\beta} \leqq B^{\alpha\beta}$$

et donc $Q^{\alpha\beta} \leqq 0$ pour $\alpha \neq \beta$, à cause de (1-7). Ensuite, en sommant sur les $\beta \in A$:

$$\sum_\beta Q^{\alpha\beta} = \sum_{i=1}^N B^{\alpha i} - B^{\alpha x} - \frac{B^{\alpha x}}{B^{xx}} \left(\sum_i B^{ix} - B^{xx} \right)$$

et donc :

$$\sum_\beta Q^{\alpha\beta} = \lambda_c^\alpha + \lambda_x^\alpha \; \lambda_c^x \geqq 0$$

puisque $\lambda_c^i = \sum_j B^{ij} \geqq 0$ et $\lambda_x^\alpha = - B^{\alpha x}/B^{xx} \geqq 0$.

Donc $Q^{\alpha\beta}$ vérifie les conditions (1-7) et (1-8), et ceci achève la démonstration.

REMARQUE - Dans l'énoncé du critère 1, je n'ai pas dit que la matrice σ_{ij} ou, ce qui revient au même, son inverse B^{ij} devait être définie positive : de fait, cette propriété découle des conditions (1-7) et (1-8).

COROLLAIRE - Si une matrice B^{ij}vérifie les condiions (1-7) et (1-8), elle est définie positive (mais non nécessairement strictement).

On le démontre par récurrence sur l'ordre N, le théorème étant vrai pour l'ordre 1. Supposons-le vrai à l'ordre N, et soit B^{ij} une matrice (N+1) x (N+1) vérifiant les conditions (1-7) et (1-8). Prenons (par exemple) x = N+1 et A = {1,...,N} . D'après (1-7) et (1-8), on a :

$$B^{xx} \geq -\sum_{\alpha \in A} B^{\alpha x} \geq 0$$

Si $B^{xx} = 0$, il s'ensuit $B^{\alpha x} = 0$ pour $\alpha \in A$ (puisque les $B^{\alpha x}$ sont ≥ 0, et la matrice B^{ij} sera définie positive en même temps que $B^{\alpha\beta}$. Mais $B^{\alpha\beta}$ est justement définie positive d'après l'hypothèse de récurrence, puisqu'elle est d'ordre N et vérifie les conditions (1-7) et (1-8) (à cause de $B^{\alpha x} = B^{xx} = 0$).

Si B^{xx} est > 0 strictement, l'identité

$$z_i \, B^{ij} \, z_j = z_\alpha \, Q^{\alpha\beta} \, z_\beta + B^{xx} \left(z_x - \frac{B^{\alpha x} \, z_\alpha}{B^{xx}} \right)^2$$

montre que B^{ij} sera définie positive en même temps que la matrice N x N $Q^{\alpha\beta}$, i.e.

$$Q^{\alpha\beta} = B^{\alpha\beta} - B^{\alpha x} \, B^{\beta x} / B^{xx}$$

Mais nous avons vu, dans la démonstration du critère, que cette matrice N x N vérifie encore les conditions (1-7) et (1-8). Elle est donc bien définie positive, d'après l'hypothèse de récurrence. QED.

<u>REMARQUE</u> - Toujours avec $A = I \setminus \{x\}$ (tous les indices sauf x), on a :

$$\sum_\alpha \lambda_x^\alpha = -\frac{1}{B^{xx}} \sum_\alpha B^{\alpha x} = \frac{-1}{B^{xx}} \left(\sum_i B^{xi} - B^{xx} \right)$$

et donc $\lambda_c^x = \sum_i B^{xi}$ est donné par :

$$\lambda_c^x = B^{xx} \left(1 - \sum_\alpha \lambda_x^\alpha \right)$$

On en déduit que la condition ii/ (positivité pour tout $A \subset I$ de la mesure d'équilibre λ_A définie par $\lambda_A^\alpha \, \sigma_{\alpha\beta} = 1$) équivaut à $\sum \lambda_u^\alpha \leq 1$ pour toute configuration de krigeage.

1.2 CRITERE DE POSITIVITE DES POIDS DE KRIGEAGE SIMPLE.

Laissons maintenant tomber cette condition ii/, et cherchons une condition nécessaire et suffisante pour que les poids de krigeage soient $\geqq 0$ pour toutes les configurations.

Une condition nécessaire est :

(1-9) $$\sigma_{ij} \geqq 0 \qquad \forall\, i, j \in I$$

On le voit en prenant le krigeage de Z_i sur la seule variable Z_j. De même, pour tout $A \subset I$, la covariance des résidus S_{uv} doit vérifier :

(1-10) $$S_{uv} \geqq 0 \qquad (u, v \in I,\ u, v \notin A)$$

On le voit en comparant le krigeage de Z_v sur les Z_α et sur les Z_i, $i \in A \cup \{u\}$: Si Z_u^* et Z_v^* sont les krigeages de Z_u et Z_v sur les Z_α , celui de $Z_K - Z_K^*$ sur les Z_i, $i \in A \cup \{u\}$ est la projection de $Z_v - Z_v^*$ sur $Z_u - Z_u^*$. Donc le poids λ_v^u de Z_u dans le krigeage de Z_v sur les Z_i sera

$$\lambda_v^u = \frac{S_{uv}}{S_{uu}}$$

Il faut donc bien $S_{uv} \geqq 0$.

Soient alors b^i des coefficients <u>strictement positifs</u>. Posons

$$\varphi_i = \sigma_{ij}\, b^j$$

Sous la condition nécessaire (1-9), on aura $\varphi_i > 0$ (strictement, car la variance σ_{ii} est > 0 strictement).

Considérons alors la nouvelle matrice

$$\tilde{\sigma}_{ij} = \frac{1}{\varphi_i}\, \sigma_{ij}\, \frac{1}{\varphi_j}$$

Elle est encore définie positive (strictement). La positivité des poids de krigeage subsiste : car, avec la nouvelle matrice, le poids $\tilde{\lambda}_u^\alpha$ se déduit de l'ancien poids λ_u^α par :

(1-11) $$\tilde{\lambda}^{\alpha}_{u} = \frac{\varphi^{\alpha}}{\varphi_u} \lambda^{\alpha}_{u} \qquad \text{(sans sommation)}$$

D'autre part, l'inverse de $\tilde{\sigma}_{ij}$ est la matrice

$$\tilde{B}^{ij} = \varphi_i B^{ij} \varphi_j \qquad \text{(sans sommation)}$$

Donc

$$\sum_j \tilde{B}^{ij} = \varphi_i \sum_j B^{ij} \varphi_j = \varphi_i b^i \qquad \text{(sans sommation)}$$

Comme φ_i et b^i sont > 0, il en est de même de $\sum_j \tilde{B}^{ij}$. Autrement dit, $\tilde{B}^{ij}$ vérifie nécessairement les conditions (1-7) et (1-8) du critère 1.

Inversement, supposons que nous ayons trouvé des $\varphi_i > 0$ tels que la matrice $\tilde{B}^{ij} = \varphi_i B^{ij} \varphi_j$ (sans sommation) vérifie (1-7) et (1-8), l'inégalité (1-8) étant stricte). D'après le critère 1, les poids $\tilde{\lambda}^{\alpha}_{u}$ d'un krigeage effectué à partir de $\tilde{\sigma}_{ij} = \sigma_{ij}/\varphi_i \varphi_j$ sont positifs. Mais il en est alors de même des poids $\lambda^{\alpha}_{u} = \tilde{\lambda}^{\alpha}_{u} \times (\varphi_u/\varphi_{\alpha})$. Ainsi :

<u>CRITERE 2</u> - Pour une matrice symétrique régulière σ_{ij}, les poids du krigeage seront $\geqq 0$ pour toute configuration si et seulement si on peut trouver des $\varphi_i > 0$ tels que la matrice $\sigma_{ij}/(\varphi_i \varphi_j)$ soit forement positive.

D'après le critère 1, cela équivaut encore à :

<u>CRITERE 2'</u> - Avec les mêmes notations, les poids du krigeage seront $\geqq 0$ pour toute configuration si et seulement si la matrice $B = \sigma^{-1}$ vérifie les deux conditions :

i/ $B^{ij} \leqq 0$ pour $i \neq j$

ii/ Il existe des $\varphi_j > 0$ tels que $\sum_j B^{ij} \varphi_j > 0$ (strictement) pour tout i.

<u>REMARQUE 2</u> - Les $\varphi_i > 0$ qui figurent dans le critère 2' possèdent une propriété intéressante : comme la matrice $\tilde{\sigma}_{ij}$ est fortement positive, les poids de krigeage $\tilde{\lambda}^{\alpha}_{u}$ qui lui sont associés vérifient $\sum_{\alpha} \tilde{\lambda}^{\alpha}_{u} < 1$ pour toute configuration (Remarque 1). D'après (1-11), cela entraine :

(1-12) $$\lambda^{\alpha}_{u} \varphi_{\alpha} < \varphi_u$$

Autrement dit, le krigeage, considéré comme un interpolateur, diminue les φ_u.

Inversement, si (1-12) est vrai pour toute configuration, ou même seulement pour tout indice $x \in I$ avec $A = I \setminus \{x\}$, les φ_i vérifieront les conditions ii/ du critère 3. Plus précisément, posons

$$b^i = B^{ij} \varphi_j$$

Alors

$$\lambda_x^\alpha \varphi_\alpha = - \frac{B^{x\alpha}}{B^{xx}} \varphi_\alpha = \varphi_x - \frac{b^x}{B^{xx}} < \varphi_x$$

donne $b^x > 0$. Comme ceci a lieu pour tout $x \in I$, il en résulte $b^i > 0$. Autrement dit :

> Les φ_i pour lesquels tout krigeage est décroissant sont les éléments de la forme $\varphi_i = \sigma_{ij} b^j$, $b^j > 0$.

On note que, dans le critère 2, on ne suppose pas la matrice σ_{ij} définie positive. Mais cette propriété résulte automatiquement des hypothèses du critère (voir remarque 1).

Nous pouvons aussi laisser de côté la propriété concernant l'existence d'une mesure d'équilibre. Pour abréger la terminologie, nous dirons qu' une covariance est positivement krigeante, ou krige positivement, si eloe assure la positivité des poids du krigeage simple pour toutes les configurations. Dans le critère suivant (contrairement au critère 2), on suppose explicitement le caractère défini positif de la matrice.

CRITERE 3 - Pour qu'une matrice symétrique strictement définie positive krige positivement, il faut et il suffit que les termes non diagonaux de la matrice inverse soient négatifs ou nuls.

La condition est nécessaire, comme ci-dessus, puisque :

$$\lambda_x^\alpha = - B^{x\alpha}/B^{xx}$$

pour $x \in I$ et $A = \{\alpha, \beta ..\} = I \setminus \{x\}$. Pour montrer qu'elle est suffisante, on raisonne ici encore par récurrence sur l'ordre N. Le critère est vrai pour $N = 2$. Supposons-le vrai pour N, et soit B^{ij} une matrice $(N+1) \times (N+1)$ inverse d'une matrice σ_{ij} vérifiant les hypothèses. Soit $x \in I$ et $A = I \setminus \{x\}$. L'inverse de $\sigma_{\alpha\beta}$ est :

$$Q^{\alpha\beta} = B^{\alpha\beta} - B^{\alpha x} B^{\beta x}/B^{xx} \leqq B^{\alpha\beta} \leqq 0 \text{ pour } \alpha \neq \beta$$

Comme $\sigma_{\alpha\beta}$ est encore de type positif strict, elle est positivement krigeante d'après l'hypothèse de récurrence. Donc σ_{ij} krige positivement pour toute configuration comportant moins de N points communs. Mais elle krige aussi positivement les configurations à N points communs, puisque dans ce cas $\lambda^{\alpha x} = - B^{\alpha x}/B^{xx}$ est $\geqq 0$ (car $B^{xx} > 0$, la matrice étant définie positive, et $B^{\alpha x} \leqq 0$ par hypothèse). Donc σ_{ij} est bien positivement krigeante.

1.3 INTERPRETATION PROBABILISTE.

Soit σ_{ij} une matrice de covariance fortement positive, et B^{ij} son inverse. Nous désignerons par λ^i_c les quantités définies par :

$$\lambda^i_c = \sum_j B^{ij} \quad \text{ou} \quad \lambda^i_c \sigma_{ij} = 1$$

Elles sont positives ou nulles. Nous les supposerons <u>strictement positives</u> :

$$\lambda^i_c > 0$$

(Cela ne nuit pas à la généralité, puisque, comme on l'a vu dans le sous-paragraphe précédent, on peut toujours remplacer σ_{ij} par une matrice $\tilde{\sigma}_{ij}/(\varphi_i \varphi_j)$ qui reste fortement positive et assure la positivité stricte de λ^i_c). La matrice R définie par

$$R_{ij} = \sigma_{ij} \lambda^j_c \qquad \text{(sans sommation)} \tag{1-13}$$

est alors une matrice stochastique ($R_{ij} \geqq 0$, d'après (1-9), et $\sum_j R_{ij} = 1$ (par construction). Son inverse est la matrice $(1/\lambda^i_c) B^{ij}$. Nous poserons :

$$A_{ij} = \delta_{ij} - \frac{1}{\lambda^i_c} B^{ij} \tag{1-14}$$

D'après le critère 1, cette matrice vérifie les conditions :

$$A_{ij} \geqq 0 \text{ pour } i \neq j \quad \text{et} \quad \sum_j A_{ij} = 0$$

Elle constitue donc le <u>générateur infinitésimal</u> d'une chaîne de Markov à temps continu. Nous désignerons par $P_{ij}(t)$, $t > 0$ les probabilités de transition associées à cette chaîne, i.e. :

$$P(t) = e^{At}$$

D'après (1-14) et la définition des λ_c^i , on a aussi :

$$\sum_i \lambda_c^i A_{ij} = 0$$

Donc notre processus admet une <u>probabilité stationnaire</u>, définie par les

$$\lambda_D^i = \lambda_c^i / C \qquad (C = \sum_i \lambda_c^i)$$

(en théorie du potentiel, C s'appellerait la capacité de l'ensemble I) et la relation :

$$\lambda_D^i A_{ij} = \lambda_D^j A_{ji} \qquad \text{(sans sommation)}$$

montre qu'il s'agit d'un <u>processus réversible</u>.

La <u>résolvante</u> de ce processus, transformée de Laplace de la matrice de transition, soit

$$R_{ij}(\mu) = \int_0^\infty e^{-\mu t} P_{ij}(t)\, dt \qquad (\mu > 0)$$

est, comme on sait, l'inverse de la matrice

$$\mu\, \delta_{ij} - A_{ij}$$

Pour $\mu = 1$, $\delta_{ij} - A_{ij} = (1/\lambda_c^i)\, B^{ij}$ est l'inverse de la matrice R_{ij} définie en (1-13), soit $R_{ij}(1) = R_{ij}$.

Ainsi, notre <u>matrice de covariance</u> fortement positive σ_{ij} est liée à la <u>résolvante</u> $R_{ij} = R_{ij}(1)$ du processus par :

$$\sigma_{ij} = R_{ij} / \lambda_c^j$$

Réciproquement, soit A_{ij} le générateur infinitésimal d'une chaine de Markov, vérifiant donc les conditions :

$$\sum_j A_{ij} = 0 \quad \text{et} \quad A_{ij} \geqq 0 \text{ pour } i \neq j$$

et admettant une probabilité stationnaire $W_i > 0$ par rapport à laquelle le processus est réversible, i.e. :

(1-15) $$W_i\, A_{ij} = W_j\, A_{ji}$$

Posant, pour une valeur de $\mu > 0$:

(1-16) $$\sigma_{ij}(\mu) = R_{ij}(\mu)/W_j$$

où $R(\mu) = [\mu I - A]^{-1}$ est la résolvante du processus, nous voyons que l'inverse de $\sigma_{ij}(\mu)$ est la matrice :

$$B^{ij}(\mu) = \mu\, W_i \delta_{ij} - W_i\, A_{ij}$$

Elle est symétrique, d'après (1-15), et vérifie les conditions :

$$B^{ij} = -\, W_i\, A_{ij} \leqq 0 \quad \text{pour } i \neq j$$

$$\sum_j B^{ij}(\mu) = \mu\, W_i > 0$$

Par suite, $\sigma_{ij}(\mu)$ est fortement positive (critère 1).

Il y a donc identité entre la classe des covariances fortement positives et la classe des résolvantes des chaînes de Markov stationnaires et réversibles.

Voici maintenant l'interprétation probabiliste du krigeage pour la matrice de covariance (1-16). Soit $A = \{\alpha, \beta \ldots\}$ un sous-ensemble de l'ensemble I des indices, et un indice $u \notin A$. Le processus étant supposé dans l'état u au temps initial $t = 0$, soit $X_o = u$, désignons par T_A l'instant aléatoire de première entrée du processus dans la classe A (T_A peut être fini ou non, mais nous verrons dans un instant qu'il est presque sûrement fini si la chaîne est irréductible). Nous poserons :

$$\lambda_u^\alpha(\mu) = E\left[e^{-\mu T_A} 1_{\{X_{T_A} = \alpha\}} \middle| X_o = u\right]$$

transformée de Laplace en t de la mesure $F_{u,\alpha}(dt)$ représentant la probabilité pour que l'entrée en A se fasse au temps t et au point α. Comme le processus ne peut se trouver en $\beta \in A$ au temps t que si l'entrée en A a lieu antérieurement à t en un $\alpha \in A$, on a

$$P_{u\beta}(t) = \int_o^t \sum_\alpha F_{u,\alpha}(d\tau)\; P_{\alpha\beta}(t-\tau)$$

et, en prenant la transformée de Laplace :

$$R_{u\beta}(\mu) = \sum_\alpha \lambda_u^\alpha(\mu)\; R_{\alpha\beta}(\mu)$$

Divisant par $W_\beta > 0$, il vient d'après (1-16) :

$$\lambda_u^\alpha\, \sigma_{\alpha\beta} = \sigma_{u\beta}$$

Autrement dit, les $\lambda_u^\alpha(\mu)$ qui définissent la loi d'entrée dans A coïncident avec les coefficients du krigeage pour la matrice $\sigma_{ij}(\mu)$: d'où la positivité de ces coefficients.

Cherchons maintenant une interprétation probabiliste pour la matrice S_{uv} des covariances des résidus. Pour cela, notons $\tilde{P}_{uv}(t)$ la probabilité pour que le processus, partant de $u \notin A$ au temps 0, se trouve au temps $t > 0$ en $v \notin A$ sans être passé par aucun état $\alpha \in A$ dans l'intervalle (o,t). On a évidemment :

$$P_{uv}(t) = \tilde{P}_{uv}(t) + \sum_\alpha \int_o^t F_{u,\alpha}(d\tau)\; P_{\alpha v}(t-\tau)$$

Désignons par $\tilde{R}_{uv}(\mu)$ la transformée de Laplace de $\tilde{P}_{uv}(t)$. Il vient alors :

$$R_{uv}(\mu) = \tilde{R}_{uv}(\mu) + \sum_\alpha \lambda_u^\alpha(\mu)\; R_{\alpha v}(\mu)$$

et donc

$$\tilde{R}_{uv} = (\sigma_{uv} - \lambda_u^\alpha\, \sigma_{\alpha v})\; W_v$$

On reconnait l'expression de la covariance des résidus S_{uv}. Ainsi :

(1-17) $$S_{uv} = \tilde{R}_{uv}/W_v$$

IRREDUCTIBILITE - Compte tenu de la relation de symétrie (1-15), tous les états $i \in I$ sont récurrents : partant d'un état i, le processus repassera preque sûrement une infinité de fois par cet état. Mais il se peut qu'il y ait plusieurs classes C_1, C_2 ... d'états récurrents sans communication entre elles ($A_{\alpha\beta} = A_{\beta\alpha} = 0$ pour $\alpha \in C_k$, $\beta \in C_{k'}$, $k \neq k'$). La matrice symétrique B^{ij}, ou, aussi bien $\sigma_{ij}(\mu)$ est alors décomposée en blocs carrés. En termes de krigeage, cela signifie que les variables Z_α et Z_β sont orthogonales pour $\alpha \in C_k$ et $\beta \in C_{k'}$ ($k \neq k'$) : le problème du krigeage se décompose en autant de sous-problèmes concernant les variables de chaque classe. De la même manière, le processus se décompose en sous-processus possibles. Si l'état initial est choisi dans une classe C, le processus reste indéfiniment dans cette classe. On peut donc restreindre à C l'ensemble des états, et prendre le générateur $A_{\alpha\beta}$ ($\alpha, \beta \in C$).

On ne nuira donc pas à la généralité en supposant qu'il y a une seule classe d'états, i.e. que la chaîne est irréductible.

1.4 PROCESSUS A TEMPS D'ARRET

Dans le cas où les $\lambda_c^i = \sum_j B^{ij}$ ne sont pas toutes strictement positives, il n'est pas indispensable de faire la transformation $\sigma_{ij} \rightarrow \tilde{\sigma}_{ij} = \sigma_{ij}/\varphi_i \varphi_j$, mais on peut aussi raisonner directement sur σ_{ij} :

Soit, en effet, σ_{ij} une matrice régulière, fortement positive et irréductible. D'après le critère 1, son inverse B^{ij} vérifie $B^{ij} \leq 0$ pour $i \neq j$. Choisissons arbitrairement des quantités W_i telles que

$$W_i > 0 \quad \text{strictement} \quad (\text{par exemple } W_i = 1)$$

et posons :

$$A_{ij} = - B^{ij}/W_i$$

On a $\sum_j A_{ij} \leq 0$, mais on ne peut avoir $\sum_j A_{ij} = 0$ pour tout i (sinon, B^{ij} ne serait pas régulier). Par suite, A_{ij} est le générateur infinitésimal d'un processus de Markov à temps d'arrêt. Ou, si l'on préfère, on peut ajouter un état ω avec :

$$\begin{cases} A_{i\omega} = - \sum_j A_{ij} \\ A_{\omega\omega} = A_{\omega i} = 0 \end{cases}$$

Compte tenu de l'hypothèse d'irréductibilité et du fait que l'on a $A_{i\omega} > 0$ pour un i au moins, cet état supplémentaire est absorbant, et, partant d'un étét $j \in I$ quelconque, le processus est absorbé en en un instant T presque sûrement fini.

L'interprétation des poids du krigeage subsiste entièrement. Soit $C = \{\alpha, \beta \ldots\} \subset I$ et $u \notin C$. Désignons par λ_u^α la probabilité pour que le processus entre en C au point α lorsqu'il part de u à l'instant initial. On a, ici encore :

$$\lambda_u^\alpha = - \frac{1}{A_{uu}} \left(A_{u\alpha} + \sum_{\substack{v \notin C \\ v \neq u}} A_{uv} \lambda_v^\alpha \right)$$

et donc

$$B^{uv} \lambda_v^\alpha = - B^{u\alpha}$$

Il s'agit bien des poids du krigeage.

2. LE CAS FINI, KRIGEAGE ORDINAIRE

De prime abord, on pourrait envisager un krigeage universel ou intrinsèque d'ordre plus élevé que le KO. Mais nous allons voir que dans ce cas il ne peut exister de covariance généralisée garantissant la positivité des coefficients de krigeage.

2.1 NECESSITE DE SE LIMITER AU KS ET AU KO.

Soit $I = \{1, 2, \ldots, N\}$ un ensemble d'indices, f^ℓ, $\ell = 0, 1, \ldots k < N-1$ des fonctions linéairement indépendantes sur I et K_{ij} une covariance généralisée sur I que nous supposerons de type positif conditionnel strict sur I, i.e. :

$$b^i f_i^\ell = 0 \quad \text{entraine} \quad b^i K_{ij} b^j \geq 0 \quad \text{et} \quad = 0 \text{ ssi } b^i = 0$$

A ce modèle, on peut associer une fonction aléatoire généralisée (FAG) qui est l'analogue d'une FAI, à ceci près qu'il n'y a pas de clause de stationnarité (une telle clause serait dépourvue de sens sur un ensemble I quelconque), i.e. une application linéaire $\tilde{Z}$ de la classe des combinaisons linéaires autorisées λ $(\lambda^i f_i^\ell = 0)$ dans un espace de VA d'espérance nulle et de variance finie, avec

$$\mathrm{Var}(\tilde{Z}(\lambda)) = \lambda^i K_{ij} \lambda^j$$

On peut aussi adopter le point de vue du krigeage universel, en raisonnant sur une représentation Z_i de la FAG $\tilde{Z}$. Cela revient à replacer la covariance généralisée K_{ij} par une vraie covariance $\sigma_{ij} = \langle Z_i\, Z_j \rangle$ nécessairement de la forme

$$\sigma_{ij} = K_{ij} - a_{\ell i} f_j^\ell - a_{sj} f_i^s + T_{\ell s} f_i^\ell f_j^s$$

Cela ne modifie pas la variance des combinaisons linéaires autorisées $\tilde{Z}(\lambda) = \lambda^i Z_i$ $(\lambda^i f_i^\ell = 0)$ et n'altère pas, par conséquent, les poids du krigeage.

Avec les hypothèses faites (K_{ij} de type positif conditionnel strict et $C_\ell f_i^\ell = 0 \Rightarrow C_\ell = 0$) on sait que la matrice du système de krigeage :

$$\begin{bmatrix} K_{ij} & f_i^\ell \\ f_j^s & 0 \end{bmatrix}$$

est régulière. Nous écrirons son inverse sous la forme :

$$\begin{bmatrix} B^{ij} & C_\ell^i \\ C_s^j & \nu_{\ell s} \end{bmatrix}$$

Explicitement :

$$(2\text{-}1) \qquad \begin{cases} B^{ik} K_{kj} + C_\ell^i f_j^\ell = \delta_j^i & ; \quad B^{ij} f_j^s = 0 \\ C_\ell^i K_{ij} + \nu_{\ell s} f_j^s = 0 & ; \quad C_\ell^i f_i^s = \delta_s^\ell \end{cases}$$

On sait aussi que la variance des combinaisons linéaires autorisées ne dépend que de la matrice B^{ij}. On peut donc, sans changer les poids de krigeage, choisir arbitrairement les C_ℓ^i et les $\nu_{\ell s}$ (par exemple en utilisant la représentation $Y_i = Z_i - f_i^\ell\, C_\ell^j\, Z_j$ et en remplaçant K_{ij} par $\langle Y_i\, Y_j\rangle$).

Soit alors $A = \{\alpha, \beta \ldots\}$ un sous-ensemble de I. Le krigeage sur A sera possible si et seulement si les fonctions f^ℓ sont linéairement indépendantes sur A : cela est évidemment suffisant. Inversement, si $C_\ell\, f_\alpha^\ell = 0$, $\alpha \in A$ et si le krigeage est possible pour chaque $u \notin A$, les conditions $\lambda_u^\alpha\, f_\alpha^\ell = f_u^\ell$ donnent $C_\ell\, f_u^\ell = 0$, et donc $C_\ell\, f_i^\ell = 0$, $i \in I$. Mais cela entraine $C_\ell = 0$ par hypothèse.

Supposons donc remplie cette condition ($C_\ell\, f_\alpha^\ell = 0 \Rightarrow C_\ell = 0$). La matrice du système de krigeage admet alors un inverse que nous noterons :

$$\begin{bmatrix} Q^{\alpha\beta} & \tilde{C}_\ell^\alpha \\ \tilde{C}_s^\beta & \tilde{\nu}_{\ell s} \end{bmatrix} = \begin{bmatrix} K_{\alpha\beta} & f_\alpha^\ell \\ f_\beta^s & 0 \end{bmatrix}^{-1}$$

Ainsi le système :

$$\begin{cases} \lambda_u^\alpha\, K_{\alpha\beta} + \mu_{\ell u}\, f_\beta^\ell = K_{\beta u} \\ \lambda_u^\alpha\, f_\alpha^\ell = f_u^\ell \end{cases}$$

a pour solution unique

$$\begin{cases} \lambda_u^\alpha = Q^{\alpha\beta}\, K_{\beta u} + \tilde{C}_\ell^\alpha\, f_u^\ell \\ \mu_{\ell u} = \tilde{C}_\ell^\beta\, K_{\beta u} + \tilde{\nu}_{\ell s}\, f_u^s \end{cases}$$

La <u>covariance des résidus</u> S_{uv}, de son côté, est donnée par :

$$\lambda_u^\alpha\, K_{\alpha v} + \mu_{\ell u}\, f_v^\ell = K_{uv} - S_{uv}$$

Il est plus intéressant de considérer l'identité suivante, qui généralise (1-3) et se démontre facilement : pour tout système z_i, $i \in I$ et R^ℓ, $\ell = 0,1 \ldots k$, on a :

$$(2\text{-}2)\quad \begin{cases} z_i\, B^{ij}\, z_j + 2\, z_i\, C_\ell^i\, R^\ell + R^\ell \nu_{\ell s}\, R^s = z_\alpha\, Q^{\alpha\beta}\, z_\beta + 2\, z_\alpha\, \tilde{C}_\ell^\alpha\, R^\ell + R^\ell \tilde{\nu}_{\ell s}\, R^s \\ \qquad + (z_u - \lambda_u^\alpha\, z_\alpha - \mu_{\ell u}\, R^\ell)\, B^{uv}\, (z_v - \lambda_v^\beta\, z_\beta - \mu_{sv}\, R^s) \end{cases}$$

Ici encore, B^{uv} est <u>l'inverse de la covariance des résidus</u> S_{uv} (on montre sans peine que B^{uv} est régulière si et seulement si les f^{ℓ} sont linéairement indépendants sur A). Avec $R^{\ell} = 0$, cette formule (2-2) met en évidence les propriétés du krigeage comme interpolateur : le minimum de $z_i\, B^{ij}\, z_j$ à z_α donné est obtenu pour $z_u^* = \lambda_u^\alpha\, z_\alpha$. En identifiant terme à terme les coefficients des 2 membres de (2-2), on obtient finalement :

$$B^{uv}\, S_{vw} = \delta_w^u$$

$$B^{\alpha u} = - \lambda_v^\alpha\, B^{vu} \quad ; \quad C_\ell^u = - \mu_{\ell v}\, B^{vu}$$

$$B^{\alpha\beta} = Q^{\alpha\beta} + \lambda_u^\alpha\, B^{uv}\, \lambda_v^\beta \qquad (2\text{-}3)$$

$$C_\ell^\alpha = \tilde{C}_\ell^\alpha + \lambda_u^\alpha\, B^{uv}\, \mu_{\ell v}$$

$$\nu_{\ell s} = \tilde{\nu}_{\ell s} + \mu_{\ell u}\, B^{uv}\, \mu_{\ell v}$$

La relation qui nous intéresse le plus ici est celle qui exprime les coefficients du krigeage en fonction de la matrice B :

$$\lambda_u^\alpha = - B^{\alpha v}\, S_{vu} \qquad (2\text{-}4$$

Si nous prenons $A = I \setminus \{x\}$ pour un indice $x \in I$ quelconque, il reste :

$$\lambda_x^\alpha = - B^{\alpha x}\, S_{xx}$$

où $S_{xx} = 1/B^{xx}$ est la variance de krigeage de x, et donc $S_{xx} > 0$ (strictement : sinon K_{ij} ne serait pas de type ≥ 0 conditionnel strict). Ainsi $\lambda_x^\alpha \geq 0$ équivaut à $B^{\alpha x} \leq 0$: les termes non diagonaux de la matrice B^{ij} doivent être ≥ 0 pour assurer la positivité des poids de krigeage.

Cette condition nécessaire est aussi suffisante, pourvu seulement que la fonction constante 1 figure parmi les fonctions f^{ℓ} ou leurs combinaisons linéaires. Nous supposerons toujours :

$$f_i^o = 1$$

Alors $B^{ij}\, f_j^{\ell} = 0$ donne avec $\ell = 0$

$$\sum_j B^{ij} = 0 \qquad (2\text{-}5)$$

Dans ces conditions, si l'on a de plus $B^{ij} \leq 0$ pour $i \neq j$, on montre par récurrence (exactement comme pour le critère 1) que les coefficients du krigeage seront ≥ 0 pour toutes les configurations. Donc :

En KU ou KI, avec $f^o = 1$, les coefficients du krigeage seront positifs pour toutes les configurations si et seulement si les termes non diagonaux de la matrice B^{ij} sont ≤ 0.

Mais cela est-il possible s'il y a d'autres fonctions de base que $f^o = 1$? Supposons $B^{ij} \leq 0$ pour $i \neq j$. Alors, compte tenu de (2-5), la matrice

$$A_{ij} = - B^{ij}$$

est le <u>générateur infinitésimal</u> d'une <u>chaîne de Markov réversible</u> (avec la probabilité stationnaire $W_i = 1/N$). Cette chaîne peut être irréductible ou non.

Supposons d'abord la chaîne <u>irréductible</u> (i.e. : il n'est pas possible de trouver une partition de I en deux classes C et C' avec $B^{\alpha u} = 0$ pour $\alpha \in C$ et $u \in C'$). Dans ce cas il n'existe pas d'autre fonction f invariante (i.e. vérifiant $\sum_j A_{ij} f_j = 0$), que les fonctions constantes. Autrement dit, les fonctions f^ℓ, qui vérifient $B^{ij} f_j^\ell = 0$ ne peuvent être que des constantes. Comme on les a supposées linéairement indépendantes, on voit qu'<u>il ne peut y avoir qu'une seule fonction de base</u>, à savoir $f^o = 1$. Le krigeage universel est donc simplement un krigeage ordinaire.

Si la chaîne n'est pas irréductible, les chose sont un peu plus complexes. Désignons par C_1, C_2 ... les classes d'états récurrents pour le processus de générateur $A_{ij} = - B^{ij}$ (tous les états sont récurrents à cause de la symétrie). Dans ce cas, à chaque classe C_r est associée une fonction invariante, à savoir l'indicatrice 1_{C_r} de cette classe, et inversement toute fonction invariante est combinaison linéaire d'indicatrices de classe. Il y a donc exactement autant de fonctions de base f^ℓ que de classes d'états récurrents, et l'on peut remplacer les f^ℓ par les indicatrices de classe 1_{C_ℓ}, soit :

$$f_i^\ell = \begin{cases} 1 \text{ si } i \in C_\ell \\ 0 \text{ si } i \notin C_\ell \end{cases}$$

Le problème de krigeage se décompose alors en sous-problèmes, classe par classe, exactement comme le processus : une combinaison linéaire des variables de la classe C_ℓ , soit :

$$\sum_{\alpha \in C_\ell} \lambda^\alpha Z_\alpha$$

est autorisée si et seulement si $\sum_{\alpha \in C_\ell} \lambda^\alpha = 0$, et les combinaisons linéaires autorisées de la classe C_ℓ sont orthogonales à celles de la classe C_s , $s \neq \ell$. Ainsi le problème du KU se décompose, classe par classe, en problèmes de KO. Si $u \in C_\ell$, le krigeage de Z_u sur les Z_α ,$\alpha \in A$ se réduit au krigeage simple de Z_u sur $A \cap C_\ell$, pourvu que $A \cap C_\ell$ ne soit pas vide. Si $A \cap C_\ell = \emptyset$, on est dans le cas où le krigeage est impossible.

Ainsi est justifiée notre assertion initiale : c'est seulement en KS ou en KO qu'il est possible d'assurer la positivité des poids du krigeage pour toutes les configurations.

2.2 CRITERE DE POSITIVITE POUR LE KO ET INTERPRETATION

Soit K_{ij} une covariance généralisée de type positif conditionnel strict à l'ordre 0 ($b^i K_{ij} b^j \geq 0$ si $\sum b^i = 0$, et $b^i K_{ij} b^j = 0 \Rightarrow b^i = 0$) (On peut évidemment prendre, par exemple, $K_{ij} = -\gamma_{ij}$, où γ_{ij} est un variogramme, ou, plus généralement, $K_{ij} = C - a_i - a_j - \gamma_{ij}$). La condition d'indépendance linéaire de f^ℓ sera toujours vérifiée sur tout sous-ensemble non vide d'indices $A \subset I$, de sorte que le problème aura toujours une solution. Posons

$$\begin{bmatrix} K_{ij} & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} B^{ij} & \lambda_D^i \\ \lambda_D^j & \nu \end{bmatrix}$$

D'après ce que l'on vient de voir, on a

$$\sum_j B^{ij} = 0$$

et, plus précisément :

$$B^{ij} f_j = 0 \Leftrightarrow f_j = C^{ste}$$

On a vu ci-dessus que les poids du krigeage ordinaire sont ≥ 0 si et seulement si les termes non diagonaux de la matrice B^{ij} sont ≤ 0. Choisissant des $W_i > 0$ quelconques tels que $\sum W_i = 1$, la matrice

$$A_{ij} = -\frac{1}{W_i} B^{ij}$$

et alors le générateur infinitésimal d'un processus admettant la probabilité stationnaire W_i. Ce processus est réversible (à cause de la symétrie de la matrice B^{ij}) et irréductible (puisqu'il n'y a pas d'autre fonction invariante que les constantes). Enfin, pour tout $A \subset I$ et $u \notin A$, le krigeage $\lambda_u^\alpha Z_\alpha$ de Z_u est déterminé par l'équation

(2-6) $$\lambda_u^\alpha B^{uv} + B^{\alpha v} = 0$$

Cette relation (2-6) détermine les poids λ_u^α (car nous savons que la matrice B^{uv} est régulière, son inverse étant la matrice S_{uv} des covariances des résidus) et entraine $\sum_\alpha \lambda_u^\alpha = 1$.

Montrons que λ_u^α est la probabilité pour que, partant de u à l'instant initial, l'entrée en A du processus se fasse au point α. En effet, désignons par $\varpi_{u\alpha}$ cette probabilité. L'évènement en question se réalise ou bien dès le premier changement d'état, ou bien après que le premier changement d'état ait conduit à un état $v \notin A$ (différent de u). Donc

$$\varpi_{u\alpha} = \frac{A_{u\alpha}}{-A_{uu}} + \sum_{v \neq u} \frac{A_{uv}}{-A_{uu}} \varpi_{v\alpha}$$

Ce qui se réécrit immédiatement

$$\sum_{v \notin A} B^{uv} \varpi_{v\alpha} = -B^{u\alpha}$$

Comparant avec (2-6), il vient bien $\varpi_{v\alpha} = \lambda_v^\alpha$.

Ainsi les poids λ_u^α du krigeage ordinaire de Z_u sur les Z_α, $\alpha \in A$ s'identifient à la loi de probabilité d'entrée du processus dans A en partant de u. D'où leur positivité.

2.3 COMPARAISON DU KS ET DU KO

Soit σ_{ij} une covariance fortement positive, B^{ij} son inverse et

$$\lambda_c^i = \sum_j B^{ij}$$

la mesure d'équilibre solution de $\lambda_c^i\ \sigma_{ij} = 1$. On a vu qu'on pouvait, sans nuire à la généralité, supposer

$$\lambda_c^i > 0$$

Posons :

$$\begin{bmatrix} \sigma_{ij} & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \tilde{B}^{ij} & \lambda_D^i \\ \lambda_D^j & \nu \end{bmatrix}$$

Il vient immédiatement

$$\lambda_D^i = \lambda_c^i / C \qquad\qquad C = \sum_j \lambda_c^j$$

$$\tilde{B}^{ij} = B^{ij} - C\ \lambda_D^i\ \lambda_D^j$$

$$\nu = -\ 1/C$$

A la matrice B^{ij} (inverse de σ_{ij}) nous avons associé, au chapitre 1, le générateur infinitésimal

$$A_{ij} = \delta_{ij} - \frac{B^{ij}}{\lambda_c^i}$$

De même, à la matrice $\tilde{B}^{ij}$ (qui vérifie $\sum_j \tilde{B}^{ij} = 0$ et $\tilde{B}^{ij} \leqq 0$ pour $i \neq j$), associons un autre générateur infinitésimal :

$$\tilde{A}_{ij} = -\ \frac{\tilde{B}^{ij}}{\lambda_c^i}$$

Compte tenu de l'expression ci-dessus de $\tilde{B}^{ij}$, nous trouvons la relation suivante entre ces générateurs infinitésimaux :

(2-7) $$\tilde{A}_{ij} = A_{ij} - \delta_{ij} + \lambda^j{}_D$$

Si $\tilde{X}_t$ et X_t désignent les processus associés à ces générateurs $\tilde{A}$ et A respectivement, cette relation (2-7) s'interprète comme suit : le processus $\tilde{X}_t$ coïncide avec X_t jusqu'à un temps d'arrêt S_1 (de loi exponentielle e^{-s}).

En $T_1 = S_1$ un nouvel état initial est tiré au sort selon la loi stationnaire λ_D^i (indépendamment du passé) et le premier processus X_t repart (avec cet état initial) jusqu'à un nouveau temps d'arrêt $T_2 = S_1 + S_2$ (S_2 de loi exponentielle e^{-s}) et ainsi de suite.

Ainsi $\tilde{X}_t$ s'identifie par morceau au processus X_t, avec des ruptures et redémarrage en l'état stationnaire aux points T_1 , T_2 ... d'un processus de Poisson. En ce qui concerne les matrices de transition :

$$\tilde{P}_{ij}(t) = e^{-t} P_{ij}(t) + (1-e^{-t}) \lambda_D^j$$

Comparons les poids du krigeage ordinaire $\tilde{\lambda}_u^\alpha$ et du krigeage simple λ_u^α, pour un ensemble d'indices $A \subset I$ et $u \notin A$. $\tilde{\lambda}_u^\alpha$ est la probabilité pour que l'entrée en A du processus $\tilde{X}_t$ partant de u ait lieu au point α, λ_u^α est de son côté la probabilité pour que cette même entrée ait lieu <u>avant le premier temps d'arrêt</u> S_1. Par conséquent, on a :

$$\tilde{\lambda}_u^\alpha = \lambda_u^\alpha + \left(1 - \sum_\beta \lambda_u^\beta\right) \left[\lambda_D^\alpha + \sum_{v \notin A} \lambda_D^v \tilde{\lambda}_v^\alpha\right]$$

Ceci n'est autre que <u>la relation habituelle d'additivité</u>. Car les quantités :

$$\lambda_A^\alpha = \lambda_D^\alpha + \sum_{v \notin A} \lambda_D^v \tilde{\lambda}_V^\alpha$$

sont manifestement les coefficients de l'estimateur optimal de la moyenne à l'aide des Z_α , $\alpha \in A$: ils vérifient bien, en effet, le système

$$\begin{cases} \lambda_A^\alpha \sigma_{\alpha\beta} = C^{ste} \\ \sum_\alpha \lambda_A^\alpha = 1 \end{cases}$$

2.4 SOMMES DE COVARIANCES FORTEMENT POSITIVES

Il n'est probablement pas vrai, en général, qu'une somme de covariances fortement positives soit encore fortement positive. Pourtant, si σ_{ij} est fortement positive, on s'attend à ce que la covariance

$$C_i \delta_{ij} + \sigma_{ij} \qquad (C_i \geq 0)$$

obtenue en ajoutant un effet de pépite soit encore fortement positive. Nous allons voir que cela, au moins, est vrai, et définir des classes de covariances à l'intérieur desquelles la positivité forte se conserve par addition. (Ces classes constituerons donc des cônes convexes de covariances fortement positives).

Soit, selon nos notations habituelles, σ_{ij} une covariance fortement positive et irréductible, et B^{ij} son inverse. Considérons, comme en 1-4, le processus à temps d'arrêt X_t défini par le générateur

$$A_{ij} = - B^{ij}/W_i$$

($W_i > 0$ arbitrairement choisis, par exemple $W_i = 1$). Pour $t > 0$, la matrice

$$P(t) = e^{At}$$

est donc sous-markovienne :

$$P_{ij}(t) \geqq 0 \quad \text{et} \quad \sum_j P_{ij}(t) < 1$$

(l'inégalité est stricte à cause de l'irréductibilité). Introduisons la résolvante

$$R_{ij}(\mu) = \int_o^\infty P_{ij}(t)\, e^{-\mu t}\, dt \geqq 0 \qquad (\mu \geqq 0)$$

On a

$$R(\mu) = [\mu\, I-A]^{-1}$$

et, en particulier, pour $\mu = 0$, $R(o) = - A^{-1}$, c'est-à-dire

$$R_{ij}(o) = \sigma_{ij}\, W_j$$

La matrice $R_{ij}(\mu)/W_j$ est l'inverse de la matrice

$$\mu\, W_i\, \delta_{ij} + B^{ij}$$

qui vérifie les conditions du critère 1. Ainsi :

> Pour tout $\mu \geqq 0$, $\sigma_{ij}(\mu) = R_{ij}(\mu)/W_j$ est une covariance fortement positive, et, en particulier, $R_{ij}(o)/W_j = \sigma_{ij}$.

Pour étudier cette famille, il sera commode d'utiliser une <u>représentation isofactorielle</u>. L'opérateur A_{ij} est auto-adjoint dans $L^2(W)$, et ses valeurs propres sont toutes négatives (strictement). Nous les désignerons par $-\lambda_n$, pour avoir $\lambda_n > 0$, et nous désignerons par χ_n un système de vecteurs propres orthonormés, avec

$$A \chi_n = -\lambda_n \chi_n$$

On a donc :

$$(2\text{-}8) \qquad \begin{cases} A_{ij} = -\sum\limits_n \lambda_n \chi_n(i) \ \chi_n(j) \ W_j \\ R_{ij}(\mu) = \sum\limits_n \dfrac{1}{\lambda_n+\mu} \chi_n(i) \ \chi_n(j) \ W_j \end{cases}$$

En ce qui concerne la <u>covariance fortement positive</u>, $R_{ij}(\mu)/W_j$, nous la désignerons par $\sigma_{ij}(\mu)$:

$$\sigma_{ij}(\mu) = \sum_n \frac{1}{\lambda_n+\mu} \chi_n(i) \ \chi_n(j)$$

<u>THEOREME</u> - Pour toute mesure p(dx) positive et bornée sur $\mathbb{R}_+$, la covariance

$$\tilde{\sigma}_{ij} = \int_0^\infty \sigma_{ij}(\mu) \ p(d\mu)$$

est fortement positive.

Comme toute limite de covariances fortement ≥ 0 est encore fortement positive, il suffit de démontrer le théorème dans le cas des combinaisons linéaires finies à coefficients > 0. Soient donc

$$p_k > 0 \quad (k = 1,2,\ldots K) \quad \text{et} \quad \mu_k \geq 0$$

et considérons la matrice

$$\tilde{R}_{ij} = \sum_k p_k \ R_{ij}(\mu_k)$$

Sous forme isofactorielle, on a :

$$\tilde{R}_{ij} = \sum_n \left(\sum_k \frac{p_k}{\lambda_n+\mu_k} \right) \chi_n(i) \ \chi_n(j) \ W_j$$

de sorte que l'inverse de $\tilde{K}_{ij}$ est :

(2-9)
$$\tilde{S}_{ij} = \sum_n \left(\frac{1}{\sum_k \frac{p_k}{\lambda_n + \mu_k}} \right) \chi_n(i)\, \chi_n(j)\, W_j$$

Tout l'intérêt se concentre sur la fonction

$$H(\lambda) = \sum_k \frac{p_k}{\lambda + \mu_k}$$

($H(\lambda)$ est une fraction rationnelle, avec un numérateur de degré K-1 et un dénominateur de degré K. Son inverse $1/H(\lambda)$ est donc de la forme :

(2-10)
$$\frac{1}{H(\lambda)} = C\,\lambda + D - \sum_{p=1}^{K-1} \frac{a_p}{C_p + \lambda}$$

Les quantités $-C_p$ sont les racines de $H(\lambda)$, au nombre de K-1. Il est immédiat que l'on a :

$$\mu_1 < C_1 < \mu_2 \ldots < \mu_{K-1} < C_{K-1} < \mu_K$$

Le coefficient C est donné par :

$$C = \lim_{\lambda \to \infty} \frac{1}{\lambda\, H(\lambda)} = \frac{1}{\sum p_k}$$

Les coefficients a_p, de leur côté, sont :

$$a_p = -\frac{1}{H'(-C_p)} = \frac{1}{\sum_k \frac{p_k}{(\mu_k - C_p)^2}}$$

Par conséquent, ces coefficients a_p, C_p et C sont positifs. C'est là le point essentiel. On obtient enfin D en faisant $\lambda = 0$ dans (2-10). Il vient :

$$D = \frac{1}{\sum \frac{p_k}{\mu_k}} + \sum_{p=1}^{K-1} \frac{a_p}{C_p} > 0$$

Nous prendrons :

$$D_o = \frac{1}{\sum \frac{p_k}{\mu_k}} \quad , \quad D = D_o + \sum_{p=1}^{K-1} \frac{a_p}{C_p}$$

Finalement, (2-10) s'écrit :

$$(2\text{-}11) \qquad \frac{1}{H(\lambda)} = C\,\lambda + D_o + \sum_{p=1}^{K-1} \frac{a_p}{C_p} - \sum_{p=1}^{K-1} \frac{a_p}{C_p+\lambda}$$

avec C, D_o, a_p et $C_p > 0$. Reportant ce résultat dans l'expression (2-9) de $\tilde{R}_{ij}$ et tenant compte de (2-8), il vient :

$$\tilde{S}_{ij} = \left(D_o + \sum \frac{a_p}{C_p}\right)\delta_{ij} - C\,A_{ij} - \sum_p a_p\, R_{ij}(C_p)$$

Les termes non diagonaux sont négatifs (car $R_{ij}(\mu) \geqq 0$ pour tout $\mu \geqq 0$, et $A_{ij} = - B^{ij}/W_i$).

Pour évaluer la somme $\sum_j \tilde{S}_{ij}$, remarquons que, pour tout $\mu \geqq 0$, on a :

$$\sum_j R_{ij}(\mu) \leqq \frac{1}{\mu}$$

(car $\sum_j P_{ij}(t) \leqq 1$). On trouve donc :

$$\sum_j \tilde{S}_{ij} \geqq D_o - C \sum_j A_{ij} = D_o + \frac{C}{W_i} \sum_j B^{ij} > 0$$

Ainsi, la matrice symétrique $W_i\,\tilde{S}_{ij}$, inverse de $\tilde{\sigma}_{ij} = \tilde{R}_{ij}/W_j$, vérifie les conditions du critère 1, et par suite $\tilde{\sigma}_{ij}$ est une covariance fortement positive.

<u>COROLLAIRE</u> - Pour toute covariance σ_{ij} fortement positive et tout choix de coefficients $C_i \geqq 0$, la matrice

$$C_i\,\delta_{ij} + \sigma_{ij}$$

est encore une covariance fortement positive.

En effet, avec les notations ci-dessus, la matrice $\sigma_{ij} + \mu\,\sigma_{ij}(\mu)$ est fortement positive. Pour $\mu \to \infty$, on voit sur (2-8) que $\mu\,R_{ij}(\mu)$ tend vers δ_{ij}, et donc $\mu\,\sigma_{ij}(\mu)$ vers $(1/W_i)\,\delta_{ij}$. Ainsi $\sigma_{ij} + (1/W_i)\,\delta_{ij}$ est fortement positive. Comme les $W_i > 0$ sont arbitraires, cela établit le corollaire dans le cas où les C_i sont strictement > 0. Le cas $C_i \geqq 0$ s'en déduit par continuité.

3. PASSAGE AU CAS INFINI

3.1 ARGUMENT HEURISTIQUE

Examinons maintenant le cas "continu". Compte tenu de la difficulté de ce problème, je me limiterai ici à chercher à caractériser la classe des covariances stationnaires $\sigma(h)$ fortement positives sur $\mathbb{R}^n$, et simultanément celle des variogrammes $\gamma(h)$ assurant la positivité des poids du KO. La démarche suivie sera heuristique plutôt que rigoureuse.

L'étude du cas fini suggère qu'une covariance fortement positive sur $\mathbb{R}^n$ doit pouvoir être interprétée comme la résolvante d'un processus de Markov à valeurs dans $\mathbb{R}^n$. Du fait qu'il s'agit d'une covariance stationnaire, les probabilités de transition du processus doivent être invariantes par translation dans $\mathbb{R}^n$. C'est dire qu'il s'agit obligatoirement d'un processus à accroissements indépendants et stationnaires. De plus, la condition de réversibilité du processus entraine la symétrie des lois de distribution. Nous avons donc affaire à un processus $X_t = (X_1(t), X_2(t),\dots,X_n(t))$ vérifiant :

$$E\left[e^{i\sum u_k(X_k(t)-X_k(o))}\right] = e^{-t\,\psi(u)}$$

avec une fonction réelle et symétrique $\psi(u) = \psi(u_1,\dots,u_n)$ donnée par la formule de Lévy Khintchin (dans le cas symétrique):

$$\psi(u) = Q(u) + \psi_1(u) \tag{3-1}$$

où $Q(u)$ est une forme quadratique positive (éventuellement nulle), qui représente une composante brownienne, et où $\psi_1(u)$ est de la forme :

$$\psi_1(u) = \int_{\mathbb{R}^n} \frac{1-\cos(\langle u\xi\rangle)}{|\xi|^2}\,\chi(d\xi) \tag{3-1'}$$

où $\chi(d\xi)$ est une mesure positive, symétrique, sur $\mathbb{R}^n$, sans atome à l'origine et vérifiant

$$\int \frac{\chi(d\xi)}{1+|\xi|^2} < \infty$$

(autrement dit : $\psi_1(u)$ est un variogramme centré). Il en résulte déjà une majoration de la forme

(3-2) $$\psi(u) \leqq a + b|u|^2$$

sur $\mathbb{R}^n$. La <u>résolvante</u> de ce processus est la mesure :

$$G_\mu(.) = \int_0^\infty e^{-\mu t} F_t(.)\, dt$$

où F_t est la loi de $X_t - X_o$. Nous caractériserons cette résolvante par sa <u>transformée de Fourier</u>

$$\tilde{g}_\mu(u) = \int G_\mu(dx)\, e^{i<ux>} = \int_0^\infty e^{-t\psi(u)-\mu t}\, dt$$

donc

(3-3) $$\tilde{g}_\mu(u) = \frac{1}{\mu + \psi(u)}$$

Noter que l'on a bien $g_\mu(u) \geqq 0$: cette fonction $g_\mu(u)$ pourra être considérée comme la mesure spectrale d'une <u>covariance</u>, pourvu seulement qu'elle soit <u>sommable</u>. Mais c'est là la condition nécessaire et suffisante pour que la résolvante $G_\mu(dx)$ admette une <u>densité</u> $g_\mu(x)$ finie en $x = 0$: $g_\mu(o) < \infty$. Sous cette condition, $g_\mu(h)$ est une <u>covariance stationnaire</u> (théorème de Bochner). Si cette condition n'est pas remplie, la mesure $G_\mu(dh)$ est de type positif. Elle peut être considérée comme la mesure-covariance d'une mesure aléatoire stationnaire, mais non d'une fonction aléatoire stationnaire, puisque $g_\mu(o) = \infty$. En résumé :

Sous la condition nécessaire et suffisante :

(3-4) $$\int \frac{du}{\mu + \psi(u)} < \infty$$

la résolvante G_μ admet une densité $g_\mu(h)$, donnée par

(3-5) $$g_\mu(h) = \frac{1}{(2\pi)^n} \int \frac{\cos(<uh>)\, du}{\mu + \psi(u)}$$

qui est alors une fonction continue de type positif, i.e. une covariance stationnaire continue.

Or la condition (3-4) n'est jamais remplie dans $\mathbb{R}^n$ pour $n \geqq 2$. D'après (3-2), en effet, on a :

$$\int_{\mathbb{R}^n} \frac{du}{\mu + \psi(u)} \geqq C_n \int_0^\infty \frac{r^{n-1}\, dr}{\mu + a + br^2}$$

et cette minoration diverge pour $n > 1$. C'est donc exclusivement dans le cas 1-D que notre problème peut admettre des solutions. Nous supposons donc, à partir de maintenant : $n = 1$.

Soit donc $g_\mu(h)$ défini par (3-5), avec $n = 1$, la condition (3-4) étant supposée satisfaite. Il reste à montrer que cette covariance $g_\mu(h)$ est fortement positive.

On rencontre ici une difficulté assez sérieuse : c'est que, contrairement au cas fini, la notion de premier passage du processus par un point x donné n'est pas immédiatement utilisable. Pour la plupart des processus (sauf le mouvement brownien, à trajectoire presque sûrement continue, et quelques autres) , on trouvera, en effet, en général, une probabilité nulle pour que le processus passe effectivement par un point donné x. Par contre, si on considère les voisinages $V_n(x) = \{y : |y-x| < 1/n\}$ du point x, on trouve en général une probabilité > 0 pour que le processus atteigne $V_n(x)$ en un temps fini, et on peut (moyennant la résolution de certaines difficultés techniques sur lesquelles je n'insiste pas ici) définir une V.A. T_n représentant l'instant d'entrée dans V_n. Posant $T = \text{Sup}\,(T_n)$, on obtient alors l'instant aléatoire T (fini ou non) où le processus passe infiniment près du point x (rencontre tous les voisinages V_n). Intuitivement, on dira alors que le processus ne peut être en x à l'instant t donné que si cet instant t est postérieur à T, c'est-à-dire si le processus est déjà passé infiniment près de x antérieurement à t. Ceci conduit à la relation :

(3-6) $$g_\mu(x) = \varpi(x;\mu)\ g_\mu(o)$$

où $\varpi(x;\mu)$ est la transformée de Laplace de la variable T, soit

$$\varpi\,(x;\mu) = g_\mu(x)/g_\mu(o)$$

On comprend ainsi pourquoi la condition $g_\mu(o) < \infty$ nous est apparue comme nécéssaire.

De même, si l'on considère un ensemble $A = \{x_\alpha, x_\beta \ldots\}$ fini ou non, on peut définir l'instant T_A où, pour la première fois, le processus (partant d'un point $x \notin A$) passe infiniment près d'un point de A. Posant :

$$\lambda_x^\alpha(\mu) = E\left[e^{-\mu T_A}\ 1_{\{X_{T_A} = x_\alpha\}}\right]$$

il vient immédiatement :

$$g_\mu(x_\alpha - x) = \sum_\beta \lambda_x^\beta(\mu)\, g_\mu(x_\alpha - x_\beta) \tag{3-7}$$

Cette relation montre que les $\lambda_x^\alpha(\mu)$ sont les poids du krigeage de x sur A pour la covariance $g_\mu(h)$. D'après leur interprétation probabiliste, ils vérifient :

$$\lambda_x^\alpha(\mu) \geqq 0 \quad ; \quad \sum_\alpha \lambda_x^\alpha(\mu) \leqq 1$$

Donc $g_\mu(h)$ est bien fortement positive.

3.2. LES COVARIANCES STATIONNAIRES FORTEMENT $\geqq 0$

Ce qui précède n'est pas une démonstration rigoureuse dans le cas général, sauf toutefois dans le cas d'un mouvement brownien (où la continuité p.s. de la trajectoire nous garantit que le processus atteint effectivement pour la première fois un point donné en un instant bien défini), ou encore dans le cas d'un processus dont la trajectoire ne contient p.s. qu'un nombre fini de discontinuités sur tout intervalle de temps fini : tel est le cas d'un processus X_t déduit d'un mouvement brownien en surimposant des sauts aléatoires à des instants poissoniens. Pour un tel processus, la fonction $\psi(u)$ est de la forme :

$$\psi(u) = \frac{c}{2} u^2 + a\left[1 - \Phi(u)\right] \tag{3-8}$$

où $\Phi(u)$ est la fonction caractéristique d'une distribution symétrique (loi des sauts).

Si $\psi(u)$ est de la forme (3-8), avec $c > 0$, la condition (3-4) est toujours remplie, et les considérations heuristiques ci-dessus deviennent démonstratives, de sorte que $g_\mu(h)$ est réellement une covariance fortement positive. Comme toute fonction ψ de la forme (3-1) est limite d'une suite de fonctions ψ_n de type (3-8), et que les relations (3-7) passent à la limite au moins dans le cas où l'ensemble $A = \{x_\alpha \ldots\}$ est fini, on arrive à la conclusion suivante :

|| Sous réserve toujours de la condition (3-4), la densité $g_\mu(h)$ de la résolvante d'un processus à accroissements indépendants stationnaires et

symétriques est une covariance fortement positive, en ce sens qu'elle assure la positivité des poids λ_x^α du krigeage et la relation $\sum_\alpha \lambda_x^\alpha \leq 1$ pour toute configuration finie.

Cas du Variogramme. Avec les mêmes notations que ci-dessus, faisons tendre μ vers 0. Si $g_\mu(o)$ reste fini, i.e. si

$$\int \frac{du}{\psi(u)} < \infty$$

(ce qui est possible, par exemple avec $\psi(u) = c\, u^2 + b|u|^\alpha$, $0 < \alpha < 1$))
$g_o(h)$ existe, et constitue une covariance fortement positive. Mais, même si l'on a

$$\int \frac{du}{\psi(u)} = \infty$$

le variogramme

$$\gamma_u(h) = g_\mu(o) - g_\mu(h) = \frac{1}{2\pi} \int \frac{1 - \cos uh}{\mu + \psi(u)}\, du$$

admet une limite finie :

$$\gamma(h) = \frac{1}{2\pi} \int \frac{1 - \cos uh}{\psi(u)}\, du \tag{3-9}$$

En effet, $\psi(u)$ est elle-même un variogramme. Par suite $\psi(u)/u^2$, pour u tendant vers 0, a une limite > 0 (finie, mais non nulle, si $\psi(u)$ est deux fois différentiable, ou sinon infinie), et l'intégrale (3-9) est toujours convergente. Pour $\mu > 0$, $g_\mu(h)$ est fortement positive, donc $\gamma_\mu(h)$ assure la positivité des poids du KO. Passant à la limite $\mu \downarrow 0$, il en résulte que $\gamma(h)$ respecte lui aussi cette positivité (au moins pour les configurations finies).

3.3. SOMME DE COVARIANCES FORTEMENT POSITIVES

Avec les mêmes notations que ci-dessus, considérons une covariance g(h) de la forme :

$$g(h) = \sum_{k=1}^{K} p_k\, g_{\mu_k}(h) \qquad (p_k\ ,\ \mu_k > 0)$$

Montrons qu'elle est encore fortement positive.
Pour cela, considérons sa transformée de Fourier :

$$\tilde{g}(u) = \sum \frac{p_k}{\mu_k + \psi(u)}$$

Montrons (ce qui établira notre proposition) que l'on a :

$$\tilde{g}(u) = \frac{1}{\tilde{\mu} + \tilde{\psi}(u)}$$

pour un $\tilde{\mu} > 0$ et une fonction $\tilde{\psi}(u)$ du type Lévy-Khintchin. Il faut prendre :

$$\tilde{\mu} + \tilde{\psi}(u) = \frac{1}{\sum \frac{p_k}{\mu_k + \psi(u)}} = \frac{1}{H[\psi(u)]}$$

avec les notations du paragraphe 2.4. D'après (2-11), il vient donc :

$$\tilde{\mu} + \tilde{\psi}(u) = C\,\psi(u) + D_o + \sum_{p=1}^{K-1} \frac{a_p}{C_p}\,\frac{\psi(u)}{C_p + \psi(u)}$$

avec C, D_o, a_p et $C_p > 0$. En faisant $u = 0$, il vient $D_o = \tilde{\mu} > 0$, et il reste :

$$\tilde{\psi}(u) = C\,\psi(u) + \sum_p b_p\,\frac{\psi(u)}{C_p + \psi(u)}$$

avec C, b_p et $C_p > 0$: il en résulte bien que $\tilde{\psi}$ est encore du type Lévy-Khintchin . Comme toute limite de covariance fortement $\geqq 0$ est encore une covariance fortement positive, on en déduit :

THEOREME - Avec les mêmes notations, pour toute mesure $p(dx)$ positive et bornée sur $\mathbb{R}^+$, la fonction

$$g(h) = \int_o^\infty g_\mu(h)\; p(d\mu)$$

est encore une covariance fortement positive sur $\mathbb{R}$.

EXEMPLE - Le mouvement brownien correspond à $\psi(u) = c\,u^2$. A un facteur près, la résolvante est l'exponentielle

$$g_\mu(h) = e^{-|h|\sqrt{\mu/c}}$$

Ainsi toute combinaison linéaire à coefficients $\geqq 0$ de covariances exponentielles est encore une covariance fortement positive. Plus généralement :

COROLLAIRE - Pour toute probabilité p(dx) sur $\mathbb{R}^+$, la fonction

$$g(h) = \int_0^\infty e^{-a|h|}\, p(da)$$

est une covariance fortement positive sur $\mathbb{R}$ et, de même, le variogramme

$$\gamma(h) = A|h| + 1 - g(h) \qquad (A \geq 0)$$

assure la positivité des poids du KO.

Ainsi, l'exponentielle itérée $\exp(e^{-a|h|})$, ou encore la fonction

$$\left(\frac{1}{1 + a|h|}\right)^{\alpha} \qquad (a, \alpha > 0)$$

sont des covariances fortement positives sur $\mathbb{R}$. (Mais elles ne le sont nullement sur $\mathbb{R}^2$ ou $\mathbb{R}^3$).

4. DEUX AUTRES INTERPRETATIONS PROBABILISTES

Dans ce qui précède, on voit coexister deux fonctions aléatoires sans lien apparent entre elles : d'un côté, un processus de Markov X_t , dont le générateur infinitésimal se déduit de l'inverse de la covariance σ_{ij}, et de l'autre une fonction aléatoire Z_i ou $Z(x)$, admettant la covariance σ_{ij} ou $\sigma(h)$. Il serait intéressant de les relier l'une à l'autre et, par exemple, de construire explicitement, à partir du processus X(t), une F.A. admettant cette covariance.

4.1. LA DENSITE DU TEMPS DE SEJOUR

Considérons d'abord le cas d'une chaîne de Markov irréductible finie X(t), éventuellement à temps d'arrêt. Soit A_{ij} son générateur, $R_{ij}(\mu)$ sa résolvante et W_i sont (unique) probabilité stationnaire avec $W_i > 0$ strictement.

Pour interpréter la résolvante $R_{ij}(\mu)$, pour un $\mu > 0$ donné, on introduit classiquement un temps d'arrêt S_μ, indépendant du processus, admettant la loi exponentielle de moyenne $1/\mu$. Alors $\mu R_{ij}(\mu)$ représente la distribution de $X(S_\mu)$ (état j du processus au temps aléatoire $t = S_\mu$) lorsque l'état initial est i. On peut aussi introduire le temps de séjour T_j du processus dans l'état j antérieurement au temps d'arrêt S_μ. Lorsque l'état initial est i, la variable T_j admet l'espérance :

(4-1) $$E(T_j | i) = \int_o^\infty P_{ij}(t)\, e^{-\mu t}\, dt = R_{ij}(\mu)$$

ce qui donne une deuxième interprétation de la résolvante. Si l'état initial est lui-même aléatoire avec comme distribution la loi stationnaire W_i, le temps de séjour T_j dans l'état j admettra donc l'espérance :

$$E(T_j) = \sum_i W_i\, R_{ij}(\mu) = \frac{W_j}{\mu}$$

Cette espérance est, comme il se doit, proportionnelle à la probabilité stationnaire W_j. En vue de permettre le passage au cas continu, il vaut mieux introduire la densité du temps de séjour par rapport à la mesure W_i. Nous poserons

(4-2) $$Z_i = \frac{\mu}{W_i}\, T_i$$

de manière à avoir $E(Z_i) = 1$.

Cherchons maintenant la covariance de cette fonction aléatoire Z_j (avec W_i comme distribution initiale). Pour cela, nous allons, en premier lieu, calculer la covariance conditionnelle :

$$S_{u;i_o,j_o} = E(T_{i_o}\, T_{j_o} | X(o) = u)$$

A cette fin, désignons par

$$\Phi_u(\lambda) = E\left[e^{-\sum \lambda_i T_i} | X(o) = u\right]$$

la transformée de Laplace de la loi des T_i lorsque l'état initial est u.

Partant de l'état u, X(t) subit un premier changement d'état à l'issue d'un temps T exponentiel avec

$$E\left[e^{-\lambda_u T}\right] = \frac{\mu - A_{uu}}{\mu + \lambda_u - A_{uu}}$$

Ce temps T correspond à un séjour dans u, et doit donc être comptabilisé dans T_u. Le changement d'état qui se produit à l'instant T peut être l'arrêt du processus ($T=S_\mu$), ce qui a lieu avec la probabilité $\mu/[\mu-A_{uu}]$, ou bien correspondre à une transition en un état i, avec la probabilité $A_{ui}/[\mu-A_{uu}]$, et le processus repart alors avec l'état initial i. On en déduit :

$$\Phi_u(\lambda) = \frac{\mu - A_{uu}}{\mu + \lambda_u - A_{uu}} \left[\frac{\mu}{\mu - A_{uu}} + \sum_{i \neq u} \frac{A_{ui}}{\mu - A_{uu}} \Phi_i \right]$$

ce qui se récrit :

$$(\mu + \lambda_u)\, \Phi_u - \sum_i A_{ui}\, \Phi_i = \mu \tag{4-3}$$

Dérivons en λ_{i_o}. Il vient :

$$\delta_{ui_o}\, \Phi_u + (\mu + \lambda_u) \frac{\delta \phi_u}{\delta \lambda_{i_o}} - \sum_i A_{u_i} \frac{\delta \phi_i}{\delta \lambda_{i_o}} = 0 \tag{4-4}$$

En $\lambda = 0$, $\delta\phi_i/\delta\lambda_{i_o}$ devient $-E(T_{i_o}/ i)$, que nous savons être égal à $-R_{ij}(\mu)$ d'après (4-1), et de fait (4-4) se réduit à

$$\delta_{ui_o} = \sum_i \left[\mu\, \delta_{u_i} - A_{u_i}\right] R_{ii_o}(\mu)$$

ce qui est bien la définition de la résolvante. Dérivant (4-4) en λ_{j_o} avant de faire $\lambda = 0$, nous allons obtenir une relation entre les covariances conditionnelles

$$S_i;_{i_o j_o} = \left. \frac{\delta^2\, \Phi_i}{\delta \lambda_{i_o}\, \delta \lambda_{j_o}} \right|_{\lambda=o}$$

En tenant compte aussi de

$$\frac{\delta\, \phi_i}{\delta\, \lambda_k} = -\, R_{ik}(\mu)$$

cela donne :

$$\sum_i [\mu\, \delta_{ui} - A_{ui}]\, S_i\, ;\, i_o j_o = \delta_{ui_o}\, R_{uj_o} + \delta_{uj_o}\, R_{ui_o}$$

En multipliant à gauche par la matrice $R(\mu)$, inverse de $\mu I-A$, il vient ainsi :

(4-5) $$S_{i;i_o j_o} = R_{i\,i_o}\, R_{i_o j_o} + R_{i\,j_o}\, R_{j_o i_o}$$

De cette expression (remarquablement simple) de la covariance conditionnelle, nous pouvons déduire celle de la covariance non conditionnelle, qui correspond au cas où la distribution initiale est la probabilité stationnaire W_i . De

$$S_{i_o j_o} = \sum_i W_i\, S_{i;i_o j_o}$$

et compte tenu de la relation

$$\sum_i W_i\, R_{ij}(\mu) = W_j/\mu$$

nous déduisons ainsi de (4-5) l'expression cherchée :

(4-6) $$S_{ij} = \frac{W_i\, R_{ij} + W_j\, R_{ji}}{\mu}$$

Il s'agit donc de la version symétrisée de la matrice $W_i\, R_{ij}$.

En fait, le cas qui nous intéresse est celui d'un processus <u>réversible</u> (i.e. : $W_i\, R_{ij} = W_j\, R_{ji}$), puisque nous voulons que la matrice

$$\sigma_{ij}(\mu) = \mu\, R_{ij}/W_j$$

soit une covariance, donc soit symétrique. On a ainsi :

$$S_{ij} = \frac{2\, W_i\, \sigma_{ij}(\mu)\, W_j}{\mu^2}$$

Avec les variables <u>densité de temps de séjour</u> Z_i, définies en (4-1), il vient donc

(4-7) $$E(Z_i\, Z_j) = 2\, \sigma_{ij}(\mu)$$

Ainsi est achevée la construction annoncée : <u>les variables "densité de temps de séjour"</u>, au facteur 2 près, <u>admettent bien la covariance $\sigma_{ij}(\mu)$</u>.

4.2. LE MODELE DES FEUILLES MORTES

Rappelons rapidement la définition du modèle de partition aléatoire sur $\mathbb{R}^n$ connu sous le nom de "feuilles mortes". On part d'un compact aléatoire A et d'un processus de Poisson spatio-temporel de densité $\theta\ dx \times dt$ sur $\mathbb{R}^n \times \mathbb{R}_+$. A chaque point (X_i, T_i) de ce processus de Poisson, on associe le dépôt, dans $\mathbb{R}^n$, au temps T_i d'un compact aléatoire A_{X_i} (isomorphe au translaté de A par la translation X_i), ces différents compacts aléatoires étant évidemment pris indépendamment les uns des autres. On a donc un modèle de dépôts successifs de feuilles mortes A_{X_i} à des instants successifs T_i.

Pour passer à un modèle de partition aléatoire, on peut regarder le tas de feuilles mortes d'en haut, ou d'en bas. Dans le premier cas (tas de feuilles mortes vu d'en haut) chaque feuille morte A_{X_i} tombée au temps T_i cache les portions de feuilles mortes tombées antérieurement qu'elle vient recouvrir, et on s'intéresse à la situation limite où $t \to \infty$ (ou plutôt : à t fixe, on rejette à $-\infty$ l'origine des temps). L'espace $\mathbb{R}^n$ est alors entièrement recouvert, et les classes de la partition sont constituées des portions restées apparentes des feuilles non entièrement recouvertes par les suivantes.

Dans le deuxième cas, qui nous intéresse ici, on regarde au contraire le tas d'en bas. Le processus partant de $t = 0$, on ne retient de chaque nouvelle feuille qui tombe que la portion qui ne recouvre pas la réunion des feuilles antérieurement tombées, et on fait tendre t vers l'infini. Ici donc pour qu'un compact ou un ouvert borné B soit contenu dans une classe unique de la partition, il faut que la première feuille ayant rencontré B ait recouvert entièrement ce compact. La probabilité pour que ceci ait lieu est, comme on sait :

$$Q(B) = \frac{E\left[V(A \ominus \check{B})\right]}{E\left[V(A \oplus \check{B})\right]}$$

(cette probabilité a d'ailleurs la même expression dans le modèle où le tas est vu d'en haut : il s'agit en fait du même modèle). Nous nous intéresserons surtout au cas où le compact B est constitué de deux points x et x+h. Alors, si nous désignons par K(h) l'espérance du covariogramme géométrique du compact aléatoire A,

$$K(h) = E \int 1_A(x)\ 1_A(x+h)\ dx$$

on trouve pour la fonction

$$\rho(h) = Q(\{ x, x+h \})$$

qui donne la probabilité pour que deux points x et x+h appartiennent à la même classe de la partition aléatoire, l'expression

(4-8) $$\rho(h) = \frac{K(h)}{2K(o) - K(h)}$$

REMARQUE - A partir d'un modèle de partition aléatoire, on peut toujours construire un modèle d'ensemble aléatoire. Il suffit d'affecter à chacune des classes de la partition des variables X_i en tout ou rien, indépendantes les unes des autres. Le corrélogramme de cet ensemble aléatoire est encore $\rho(h)$, de sorte que la fonction définie en (4-8) est aussi le corrélogramme d'un ensemble aléatoire stationnaire.

Nous allons montrer que tout corrélogramme fortement positif et stationnaire sur $\mathbb{R}$ est du type "feuille morte" défini ci-dessus en (4-8).

On peut aussi établir le même résultat dans le cas fini pour une covariance fortement positive σ_{ij}, pourvu seulement que la variance σ_{ii} soit une constante indépendante de i : on peut toujours se ramener à ce cas en prenant

$$\rho_{ij} = \sigma_{ij}/\sqrt{\sigma_{ii}\ \sigma_{jj}}$$

Dans le cas fini, le processus spatio-temporel poissonien dans $I \times \mathbb{R}_+$ aura comme "densité" la mesure $W_i \times dt$, où W_i est la solution de $\sum_i W_i\ \rho_{ij} = 1$. Dans le cas stationnaire, on prendra évidemment la densité $dx \times dt$ sur $\mathbb{R} \times \mathbb{R}_+$.

Raisonnons, par exemple, dans le cas stationnaire, et adoptons les notations du paragraphe 3. Soit X_t un processus à accroissements indépendants; stationnaires et symétriques, défini par sa fonction $\psi(u) \geq 0$ du type Lévy-Khintchine. En toute rigueur, il conviendrait de se limiter, dans un premier temps, au cas où $\psi(u)$ est du type

$$(4\text{-}9) \qquad \psi(u) = \frac{c}{2} u^2 + b \left[1 - \Phi(u)\right]$$

$\Phi(u)$ représentant la transformée de Fourier d'une loi symétrique : X_t est alors un mouvement brownien interrompu par des sauts de loi $\Phi(u)$ survenant à des instants poissoniens. C'est, en effet, pour ces processus que la notion de premier passage par un point donné est immédiatement définissable.

La probabilité $\rho(h)$ pour que, partant d'un point donné x, le processus atteigne le point x+h avant le temps d'arrêt S_μ (temps d'arrêt indépendant du processus, de loi exponentielle et d'espérance $1/\mu$) est alors, comme on l'a vu, égale à :

$$\rho(h) = g_\mu(h)/g_\mu(o)$$

où $g_\mu(h)$ est la densité de la résolvante. La transformée de Fourier de $g_\mu(h)$ est elle-même, on l'a vu :

$$\tilde{g}_\mu(u) = \frac{1}{\mu + \psi(u)}$$

Pour passer au modèle des feuilles mortes, prenons comme compact aléatoire :

$$A = \{X_t,\ 0 \leqq t \leqq S_\mu\} \qquad \text{(avec } X(o) = 0\text{)}$$

Si $\psi(u)$ est de la forme (4-9), cette définition ne pose pas de problème, bien qu'elle fasse intervenir une infinité non dénombrable d'instants t, du fait que la trajectoire X_t, pour $0 \leqq t \leqq S_\mu$, ne comporte p.s. qu'un nombre fini de discontinuité (Dans le cas général, la définition de A serait plus complexe et devrait, entre autre, faire intervenir la séparabilité du processus, mais je n'insisterai pas sur ces difficultés techniques).

Avec ce compact aléatoire et le modèle de partition aléatoire correspondant au tas de feuilles mortes vu d'en bas, la probabilité pour que deux points donnés x et x+h appartiennent à la même classe est encore égale à $\rho(h)$.

En effet, considérons la première feuille rencontrant l'ensemble $\{x,x+h\}$. Elle a germé en un point X_i et en un temps T_i du processus spatio-temporel qui a servi d'origine à un processus du type X_t (noter que l'on utilise ici une propriété markovienne forte), et ce processus X_t a ensuite atteint pour la première fois l'ensemble $\{x, x+h\}$ en un instant aléatoire T avec $S_\mu \geq T > T_i$. A cet instant T, il se trouve, soit en x, soit en x+h ($X_T = x$ ou $x+h$), et la propriété conditionnelle pour que X_t atteigne ensuite l'autre point (x+h ou x) avant S_μ est bien égale à $\rho(h) = g_\mu(h)/g(o)$: ce qui achève de démontrer le résultat annoncé.

Voici, entre autres, deux conséquences :

En premier lieu, conformément à une remarque faite plus haut, on voit qu'à un facteur près, toute covariance stationnaire fortement positive sur $\mathbb{R}$ est la covariance d'un ensemble aléatoire stationnaire.

En second lieu, l'espérance du covariogramme géométrique du compact aléatoire A (ensemble des points atteints par le processus X_t avant le temps d'arrêt S_μ), soit K(h), est reliée à $\rho(h)$ par la relation (4-8). On a donc, inversement :

$$\frac{K(h)}{K(o)} = \frac{2\,\rho(h)}{1+\rho(h)}$$

Comme K(o), espérance de la mesure de A, est manifestement donné par :

$$K(o) = \int_{-\infty}^{+\infty} \rho(h)\, dh$$

on voit que l'on reconstitue sans peine K(h) à partir de $\rho(h)$.

Considérons le cas particulier du mouvement brownien, i.e. $\psi(u) = \frac{c}{2} u^2$. On a alors, comme on sait :

$$\rho(h) = e^{-|h|\sqrt{\frac{2\mu}{c}}} = e^{-a|h|}$$

(covariance exponentielle) et donc, en posant $a = \sqrt{2\mu/c}$:

$$K(h) = \frac{4}{a}\ \frac{e^{-a|h|}}{1 + e^{-a|h|}}$$

Mais, d'un autre côté, la trajectoire du mouvement brownien étant p.s. continue, le compact aléatoire A est nécessairement un intervalle fermé (compact convexe dans $\mathbb{R}$). Désignons par L la longueur (aléatoire) de cet intervalle fermé A. On a, d'après la définition du covariogramme géométrique :

$$K(h) = E\,[(L - |h|)_+]$$

La distribution F(dx) de L vérifie donc :

$$\int_x^\infty [\xi - x]\, F(d\xi) = K(x) \qquad (x \geq 0)$$

Elle admet donc la moyenne :

$$E(L) = K(o) = \frac{2}{a}$$

et une fonction de répartition F(x) donnée par

$$F(x) = 1 - K'(x) = \frac{1}{1 + e^{-ax}} \tag{4-10}$$

On voit ainsi cette conclusion paradoxale : un ensemble aléatoire stationnaire admettant une covariance exponentielle est nécessairement un processus de Markov à deux états (a et 1). Il est néanmoins possible de construire ce processus markovien de manière indirecte, à partir du modèle des feuilles mortes, en prenant pour feuilles mortes des intervalles fermés de longueur aléatoire L avec la loi (4-10).

5. REMARQUES TERMINALES

On peut se demander si ce qui précède constitue une caractérisation exhaustive des covariances fortement positives. Dans le cas fini, la réponse est indubitablement affirmative, puisque nous avons établi des critères donnant les conditions nécessaires et suffisantes que doit vérifier une matrice (finie) de covariance pour être fortement positive. Il y a réellement identification de cette classe de covariances avec la classe des résolvantes des chaînes de Markov réversibles. Dans le cas infini, les choses sont sans doute plus complexes, et la démarche semi-heuristique que nous avons suivie ne nous permet pas d'être affirmatif. De fait, nous avons postulé, par analogie avec le cas fini, qu'une covariance stationnaire $\sigma(h)$ fortement positive devait s'identifier à la résolvante d'un processus sur $\mathbb{R}$ (nécessairement à accroissements indépendants et stationnaires, à cause de l'invariance par translation). Mais, pour démontrer qu'il en est réellement ainsi, il aurait fallu procéder à un passage à la limite délicat : considérant une famille croissante $I_n = \{x_1, \ldots, x_n\}$ de points de $\mathbb{R}$(ou $\mathbb{R}^k$) dont la réunion constituerait un ensemble dénombrable dans $\mathbb{R}$ (ou $\mathbb{R}^k$), il faudrait étudier le comportement des matrices B_n (inverse de la matrice $\sigma_{ij} = \sigma(x_i - x_j)$, i et $j \in I_n$), et montrer que le passage à la limite conduit effectivement au générateur infinitésimal d'un processus à accroissements indépendants et stationnaires.

Or il n'en est certainement pas ainsi sans quelques hypothèses restrictives sur $\sigma(h)$. Le procédé utilisé au Chapitre 3 nous a conduit, en effet, à une classe de covariances admettant une densité spectrale (par rapport à la mesure de Lebesgue). Il s'agit donc de covariances $\sigma(h)$ continues et (au moins) faiblement mélangeantes. Or il est indéniable qu'il existe des covariances fortement positives mais non continues : par exemple, un effet de pépite à l'état pur, ou plus généralement, dans le cas 1-D, une covariance du type a $\theta(h) + \sigma(h)$, où $\sigma(h)$ est du type étudié au Chapitre 3, et $c\theta$ un effet de pépite (il s'agit d'un effet de pépite fini, non d'un Dirac : la fonction θ est définie par $\theta(h) = 1$ pour $h = 0$, $\theta(h) = 0$ pour $h \neq 0$). Les résultats du paragraphe (2-4) montrent bien qu'il s'agit encore d'une covariance fortement positive et stationnaire, bien qu'elle ne soit plus continue. De même encore, pour $n > 1$, il existe des covariances fortement positives, bien qu'on ne puisse plus les associer à des processus. Par exemple : un effet de pépite, une covariance constante, ou encore une somme du type

$$\sigma(h) = c\;\theta(h) + q$$

où $c\;\theta(h)$ est un effet de pépite pur, et a une constante ≥ 0. Néanmoins, le caractère un peu pathologique de ces contre-exemples incite à conjecturer qu'il n'y a pas de covariances stationnaires, continues, ergodiques et centrées (i.e. admettant une mesure spectrale sans atome en $u = 0$) pour $n > 1$, et pas d'autres, dans le cas 1-D, que celles que nous avons trouvées au Chapitre 3 : mais ceci reste conjectural, pour l'instant du moins.

La deuxième remarque concerne le cas infini dénombrable. Il existe indéniablement des covariances fortement positives σ_{ij} dans le cas infini dénombrable, par exemple la classe des résolvantes des chaînes infinies dénombrables. En particulier, il existe certainement sur l'espace $\mathbb{Z}^n$, version discrétisée de $\mathbb{R}^n$, des covariances stationnaires et fortement positives, même pour $n > 1$: à savoir les résolvantes des processus à accroissements indépendants, stationnaires et symétriques sur $\mathbb{Z}^n$. Pour un tel processus X_t, en effet, on aura :

$$E\left[e^{iu\, X_t} \middle| X_o = 0 \right] = e^{-at[1-\Phi(u)]}$$

où $\Phi(u)$ est la fonction caractéristique de la loi des sauts (loi symétrique sur $\mathbb{Z}^n$) soit

$$\Phi(u) = \sum_{j \in Z^n} p_j \; e^{i<ju>}$$

La résolvante $R_{ij}(\mu)$ est alors de la forme

$$R_{ij}(\mu) = G_{j-i}(\mu)$$

avec

$$G_j(\mu) = \frac{1}{(2\pi)^n} \int_o^{2\pi} \cdots \int_o^{2\pi} \frac{e^{-i<ju>}}{\mu + a[1-\Phi(u)]} \, du_1 \ldots du_n$$

(intégrale toujours convergente pour $\mu > 0$).

Par construction, la matrice $\sigma_{ij} = G_{|i-j|}$ est bien une covariance fortement positive et stationnaire sur $\mathbb{Z}^n$.

On peut, par exemple, considérer le mouvement brownien sur $\mathbb{Z}^n$, qui correspond au cas

$$\Phi(u) = \frac{\cos u_1 + \cos u_2 + \ldots + \cos u_n}{n}$$

Les $G(\mu)$ sont alors les coefficients de Fourier de la fonction périodique

$$\frac{1}{\mu + \mu(1-\Phi)} = \frac{1}{\mu + a - \frac{a}{n} \sum \cos u_k}$$

Mathématiquement, il ne semble pas possible d'obtenir une expression explicite simple.