N-85

N O T E   G E O S T A T I S T I Q U E   N° 91

RECHERCHE D'ESTIMATEURS UNIVERSELS OPTIMAUX

Table des Matières

Table des Matières

(Suite)

N O T E   G E O S T A T I S T I Q U E   N° 91

RECHERCHE D'ESTIMATEURS UNIVERSELS OPTIMAUX

INTRODUCTION

Nous avons en vue le problème suivant : soit $Z_\omega(x)$ $(x \in \mathbb{R}^n)$ une fonction aléatoire sur un espace $(\Omega, \mathcal{A}, P)$, dont nous savons qu'elle est de la forme :

$$Z_\omega(x) = Y_\omega(x) + a_\ell f^\ell(x)$$

(avec $a_\ell f^\ell$ au lieu de $\sum_\ell a_\ell f^\ell$, selon la convention de sommation) Les $f^\ell(x)$ $(\ell = 1,2,\ldots k)$ sont des fonctions connues de $x \in \mathbb{R}^n$, les a_ℓ des coefficients constants dont nous ne connaissons pas la valeur numérique; $Y_\omega(x)$, enfin, est une F.A. admettant une espérance nulle et une covariance $C(x,y)$:

$$E\left(Y_\omega(x)\right) = 0 \quad , \quad E\left(Y_\omega(x)\, Y_\omega(y)\right) = C(x,y)$$

On suppose connue une réalisation de $Z_\omega(x)$ pour $x \in S$, $(S \subset \mathbb{R}^n$, ensemble fini ou non de points; en général, S sera compact). On cherche :

1 - à former un estimateur A optimal et sans biais du vecteur dérive $a = (a_\ell)$ à l'aide d'opérations linéaires effectuées sur la réalisation $Z_\omega(x)$, $x \in S$.

2 - à former, dans les mêmes conditions, un estimateur

optimal et sans biais de $Z_\omega(z)$, $z \notin S$, ou, plus généralement, de $Z = \int \mu(dx)\ Y(x)$, μ étant une mesure dont le support est disjoint de S.

3 - à examiner les rapports existant entre ces deux estimateurs (en particulier, a-t-on le droit de travailler sur les résidus $Z(x) - A_\ell f^\ell(x)$, et quelles sortes de biais résultent du fait que la vraie dérive a_ℓ a été remplacée par son estimation A).

Ces estimateurs méritent le qualificatif universel, puisqu'ils doivent être utilisables quelle que soit la vraie dérive a_ℓ. Il s'agit, cependant, d'une "universalité" limitée puisque nous supposons connues les fonctions $f^\ell(x)$ et la covariance $C(x,y)$ (celle-ci, éventuellement, à un facteur multiplicatif près, comme nous le verrons).

Nous ferons les raisonnements dans le cas général où S est un ensemble quelconque, mais nous particulariserons toujours les résultats au cas où $S = \{x_\alpha,\ \alpha = 1,2,\ldots m\}$ est un ensemble fini : ce cas correspond à la plupart des applications pratiques, et entraîne d'ailleurs de grandes simplifications.

I - LES ESPACES $\overline{H}$ ET $\overline{H}_f$

Soit, sur $(\Omega, \mathcal{A}, P)$, une Fonction Aléatoire $Y_\omega(x)$ continue m.q, avec :

$E(Y_\omega(x)) = 0$ $\quad E(Y_\omega(x)\ Y_\omega(y)) = C(x,y)$ (fonction continue). Nous supposons $C(x,y)$ de type <u>strictement positif</u>, en ce sens que $\iint \mu(dx)\ \mu(dy)\ C(x,y) = 0$ pour une mesure μ entraîne $\mu = 0$.

Nous allons introduire les sous-espaces suivants de $L^2(\Omega,\mathcal{A},P)$:

~ <u>Espace H</u> des combinaisons linéaires finies $\sum \lambda_i\ Y_\omega(x_i)$ pour $x_i \in S$. Pour condenser les notations, nous désignerons par Λ l'ensemble des mesures λ <u>à support fini contenu dans S</u>. Tout $Y \in H$ est donc de la forme :

$$Y = \int \lambda(dx)\ Y(x) \qquad , \lambda \in \Lambda$$

~ <u>L'espace H_f</u> des variables aléatoires de la forme $\int \lambda(dx)\ [Y(x) + f(x)]$, $\lambda \in \Lambda$ $f(x)$ désignant une fonction donnée.

Si S est fini, H et H_f sont des sous-espaces <u>fermés</u> de $L^2\ (\Omega,\mathcal{A},P)$, donc des espaces de Hilbert, et les discussions se simplifieront beaucoup dans le cas fini. Si S est infini, H et H_f ne seront pas fermés en général, et nous devrons considérer leurs fermetures dans $L^2(\Omega,\mathcal{A},P)$:

~ <u>Les espaces $\bar{H}$ et $\bar{H}_f$</u>, fermetures dans $L^2(\Omega,\mathcal{A},P)$ des précédents : ce sont des espaces de Hilbert.

<u>Remarque</u> : Tout $Y \in \bar{H}$ est limite dans L^2 d'une suite $Y_n \in H$. Chaque Y_n est combinaison linéaire finie des $Y(x)$, $x \in S$, soit

$Y_n = \int \lambda_n(dx)\ Y(x)$. Mais la convergence $Y_n \to Y$ dans L^2 n'entraîne pas que les mesures λ_n convergent vaguement vers une mesure μ. En général, donc, un $Y \in \overline{H}$ quelconque n'admettra pas de représentation de la forme

$$(1) \qquad Y = \int \mu(dx)\ Y(x) \qquad , \qquad \mu \in M$$

pour une mesure μ bornée à support dans S (fini ou non) : l'ensemble $\overline{H}_M$ des $Y \in \overline{H}$ admettant la représentation (1) pour une mesure μ appartenant à l'ensemble M des mesures bornées à support dans S est un sous-espace de $\overline{H}$ <u>non fermé</u> dans $\overline{H}$ en général.

Il sera capital, dans les applications, de savoir si un $Y \in \overline{H}$ est ou non dans $\overline{H}_M$. En effet, si Y est un estimateur, on ne pourra calculer sa valeur numérique à partir d'une réalisation que si $Y \in \overline{H}_M$.

On trouvera toujours une suite $\lambda_n \in \Lambda$ de mesures à support fini telles que

$$|| Y - \int \lambda_n(dx)\ Y(x)|| \to 0$$

mais celà n'impliquera la convergence numérique :

$$Y_n = \int \lambda_n(dx)\ Y_\omega(x) \to Y_\omega$$

pour presque tout $\omega \in \Omega$ que si $Y \in \overline{H}_M$ (on peut, théoriquement, extraire de la suite Y_n une suite partielle convergeant presque sûrement vers Y, mais, dans les applications, on ne saura pas construire effectivement cette suite à partir d'une réalisation). Si $Y \notin \overline{H}_M$, la suite numérique des Y_n construites à partir d'une réalisation sera divergente, en général, bien que les variances des estimateurs $\acute{Y}_n$ successifs tendent en décroissant vers la variance de l'estimateur optimal Y.

Ces remarques s'appliquent évidemment aussi à H_f.

<u>L'Application Φ</u> - Considérons l'application de $L^2(\Omega, \mathcal{A}, P)$ dans lui-même définie par :

$$\Phi(Z) = Z - E(Z)$$

Elle est <u>continue</u> $\left(\|Z - E(Z)\|^2 = \|Z\|^2 \times (E(Z))^2 \leq \|Z\|^2\right)$

Φ applique H_f <u>sur</u> H, puisque tout $Y \in H$ est de la forme $Y = \int \lambda(dx)\ Y(x)$, $\lambda \in \Lambda$, donc est l'image de :

$$\Phi^{-1}(Y) = Y + \int \lambda(dx)\ f(x) \quad \in H_f$$

Elle est <u>injective</u> : $Z - E(Z) = Z' - E(Z')$ pour Z et $Z' \in H_f$ entraîne $E(Z) - E(Z') = Z - Z' \in H_f$. Si $Z \neq Z'$, on a donc $E(Z) - E(Z') \neq 0$ et $1 \in H_f$. Il existe donc une mesure $\lambda \in \Lambda$ avec $1 = \int \lambda(dx)\ Y(x) + \int \lambda(dx)\ f(x)$, d'où $0 = \int \lambda(dx)\ Y(x)$ avec $\lambda \neq 0$, ce qui contredit le fait que C est strictement de type positif.

Ainsi Φ est un isomorphisme continu de H_f sur H. En tant qu'application continue, Φ se prolonge en une application continue Φ_f de $\overline{H}_f$ dans $\overline{H}$. On a $\Phi_f(Z) = Z - E(Z)$, et les inclusions

$$H = \Phi_f(H) \subset \Phi_f(\bar{H}_f) \subset \bar{H}$$

montrent que l'espace image $\Phi_f(\bar{H}_f)$ est dense dans $\bar{H}$: mais, en général, il n'est pas fermé et ne coïncide pas avec $\bar{H}$: si une suite $\lambda_n \in \Lambda$ vérifie le critère de Cauchy dans $L^2(\Omega, \mathcal{A}, P)$ (existence de $\lim\limits_{n,m} \iint \lambda_n(dx)\ \lambda_m(dy)\ \ C(x,y)$), il n'en résulte pas nécessairement que la suite $\int \lambda_n(dx)\ f(x)$ soit convergente, de sorte que $Y_n \to Y$ n'entraîne pas que $\Phi^{-1}(Y_n)$ converge dans $\bar{H}_f$.

Cherchons une condition nécessaire et suffisante pour que l'on ait $\Phi_f(\bar{H}_f) = \bar{H}$: cette condition s'écrira $1 \notin \bar{H}_f$, et nous verrons que cette condition entraîne aussi que Φ_f est un homéomorphisme de $\bar{H}_f$ sur $\bar{H}$. Posons quelques préliminaires :

Lemme 1 - On a $Z \in \bar{H} \cap \bar{H}_f$ si, et seulement si, $Z \in \bar{H}_f$ et $E(Z) = 0$.

Si $E(Z) = 0$, $\Phi_f(Z) = Z \in \bar{H}$ dès que $Z \in \bar{H}_f$ - Inversement, soit $Z \in \bar{H} \cap \bar{H}_f$: comme $Z \in \bar{H}$, Z est limite de $Z_n \in H$ avec $E(Z_n) = 0$, et par suite $E(Z) = 0$.

Lemme 2 - $\bar{H} \cap \bar{H}_f = \overline{H \cap H_f}$

Il suffit de montrer $\bar{H} \cap \bar{H}_f \subset \overline{H \cap H_f}$, l'inclusion inverse étant toujours vraie. Soit $Z \in \bar{H} \cap \bar{H}_f$, et une suite $Z_n = \int \lambda_n(dx)\ [Y(x) + f(x)] \in H_f$ convergeant vers Z dans $\bar{H}_f$. On a $E(Z_n) \to E(Z) = 0$, c'est-à-dire $\int \lambda_n(dx)\ f(x) \to 0$. Choisissons une mesure $\mu \in \Lambda$ avec :

$$\int f(x)\ \mu(dx) = 1 \quad , \quad \iint \mu(dx)\ \mu(dy) \quad C(x,y) \ = \ a < \infty$$

Les variables $Y_n = \int(\lambda_n - E(\mathcal{Z}_n)\mu)\ \big(Y(x) + f(x)\big)$ vérifient $Y_n \in H_f$, $Y_n \in H$ $\left(\text{car} \int(\lambda_n - E(\mathcal{Z}_n)\mu)\ f(x) = 0\right)$ et $Y_n \to \mathcal{Z}$ dans L^2 : par suite $\mathcal{Z} \in \overline{H \cap H_f}$.

<u>Lemme 3</u> - Tout $\mathcal{Z} \in \overline{H}_f$ est limite d'une suite $\mathcal{Z}_n \in H_f$, avec $E(\mathcal{Z}_n) = E(\mathcal{Z})$.

Si $E(\mathcal{Z}) = 0$, celà résulte des deux lemmes ci-dessus. Soit $E(\mathcal{Z}) \neq 0$, et $\mathcal{Z} = \lim \int \lambda_n(dx)\ [Y(x) + f(x)]$ avec $\lambda_n \in \Lambda$. On a $\int \lambda_n(dx)\ f(x) \to E(\mathcal{Z})$, et la suite

$$\mathcal{Z}_n = \frac{E(\mathcal{Z})}{\int \lambda_n(dx)\ f(x)} \int \lambda_n(dx)\ [Y(x) + f(x)]$$

converge vers $\mathcal{Z}$ dans $\overline{H}_f$ et vérifie $E(\mathcal{Z}_n) = E(\mathcal{Z})$ -

<u>Proposition 1</u>. Les 4 propriétés suivantes sont équivalentes

a/ $1 \notin \overline{H}_f$

b/ La forme linéaire $\lambda \to \int \lambda(dx)\ f(x)$ est continue sur Λ munie de la norme $||\lambda||^2 = \iint \lambda(dx)\ C(x,y)\ \lambda(dy)$

c/ Il existe un élément $Y_f \in \overline{H}$ tel que $\langle Y_f, Y(x)\rangle = f(x)$ pour tout $x \in S$.

d/ L'application $\Phi_f : \mathcal{Z} \to Z - E(\mathcal{Z})$ est un homéomorphisme de $\overline{H}_f$ sur $\overline{H}$.

Montrons a $\Longleftrightarrow$ b : $1 \in \overline{H}_f$ signifie que l'on peut trouver une suite $\lambda_n \in \Lambda$ avec $\mathcal{Z}_n = \int \lambda_n(dx)\ \big(Y(x) + f(x)\big) \to 1$, et, d'après le lemme 3, on peut même supposer $\int \lambda_n(dx)\ f(x) = 1$ et

$\int \lambda_n(dx)\ Y(x) \to 0$ dans $\overline{H}$. Par suite, $1 \notin \overline{H}_f$ équivaut à : $\exists\ \varepsilon > 0$, $\forall \lambda \in \Lambda$, $\int \lambda(dx)\ f(x) = 1 \Longrightarrow \int \lambda(dx)\ \lambda(dy)\ C(x,y) \geq \varepsilon$, donc aussi à

$$\varepsilon \left| \int \lambda(dx)\ f(x) \right|^2 \leq \int \lambda(dx)\ \lambda(dy)\ C(x,y)$$

c'est-à-dire à b/

b $\Longleftrightarrow$ c : D'autre part, Λ, muni de la norme $||\lambda||$ ci-dessus, est homéomorphe à H. Si la forme linéaire : $\lambda \to \int \lambda\ f$ est continue, elle se prolonge sur $\overline{H}$, et par suite il existe $Y_f \in \overline{H}$ tel que l'on ait $\langle Y, Y_f \rangle = \int \lambda(dx)\ f(x)$ pour tout $Y = \int \lambda(dx)\ Y(x) \in H$. En particulier, on a

$$\langle Y(x), Y_f \rangle = f(x)$$

pour tout $x \in S$. Inversement, si $f(x) = \langle Y(x), Y_f \rangle$ pour un $Y_f \in \overline{H}$ et tout $x \in S$, on a $\int \lambda(dx)\ f(x) = \langle Y, Y_f \rangle$ avec $Y = \int \lambda(dx)\ Y(x)$, et cette forme linéaire est continue sur Λ pour la norme $||\ \ ||$

c $\Longleftrightarrow$ d : L'application $\Phi_f : \xi \to \xi - E(\xi)$ de $\overline{H}_f$ dans $\overline{H}$ est continue et linéaire. Elle constitue un homéomorphisme si, et seulement si, elle est bijective.

Soient ξ et $\xi' \in \overline{H}_f$ avec $\xi - E(\xi) = \xi' - E(\xi') \in \overline{H}$. On en déduit $E(\xi) - E(\xi') = \xi - \xi' \in \overline{H}_f$, et $\xi \neq \xi'$ si et seulement si $E(\xi) - E(\xi') \neq 0$, ce qui entraîne $1 \in \overline{H}_f$. Donc $1 \notin \overline{H}_f$ entraîne que Φ_f est injective.

Montrons que Φ_f est surjective si c/ est vérifiée.
Soit $Y \in \bar{H}$, avec $Y = \ell im\, Y_n$, $Y_n = \int \lambda_n(dx)\ Y(x) \in H$ $(\lambda \in \Lambda)$.
La continuité du produit scalaire donne :

$$< Y,\ Y_f > \ = \ell im < Y_n,\ Y_f > \ = \ \ell im \int \lambda_n(dx)\ f(x)$$

Par suite $Y + < Y,\ Y_f > \in \bar{H}_f$: Φ_f admet donc comme inverse l'application $Y \to Y + < Y, Y_f >$. Linéaire, bijective et continue, Φ_f est un homéomorphisme.

Inversement, supposons que Φ_f soit un homéomorphisme. Comme Φ_f^{-1} est continue, la forme linéaire : $Y \to E\,[\Phi_f^{-1}(Y)]$ est continue sur $\bar{H}$. Il existe donc $Y_f \in \bar{H}$ tel que :

$$E\,[\Phi_f^{-1}(Y)] \ = \ < Y,\ Y_f >$$

pour tout $Y \in \bar{H}$. En particulier, pour $Y = Y(x)$, on a $\Phi_f^{-1}\big(Y(x)\big) = Y(x) + f(x)$, et par suite $f(x) = < Y(x),\ Y_f >$.

Remarques - 1/ Les 4 propriétés de la proposition entraînent que la fonction $f(x)$ est continue sur S si la covariance $C(x,y)$ est continue : en effet, les $Y(x)$ sont sontinues m.q., et $x_n \to x$ entraîne $Y(x_n) \to Y(x)$ dans L^2, et a fortiori $< Y(x_n), Y_f > \to < Y(x), Y_f >$, c'est-à-dire $f(x_n) \to f(x)$

D'ailleurs, si $f(x)$ n'est pas continue en $x_o \in S$, soit $x_n \to x_o$ et $\ell im\,[f(x_n) - f(x_o)] = C \neq 0$. Comme $Y(x_n) \to Y(x_o)$ dans $\bar{H}$, on a : $\lim\limits_{x_n \to x_o} [Y(x_n) - Y(x_o) + f(x) - f(x_o)] = C \in \bar{H}_f$, et $1 \in \bar{H}_f$.

2/ Toujours en supposant $C(x,y)$ continue et, de plus, S compact, la convergence étroite $\mu_n \to \mu$ d'une suite de mesures à support dans S entraîne la convergence dans chaque $\overline{\overline{H}}_f$ de $\int \mu_n(dx)\,[Y(x) + f(x)]$ vers $\int \mu(dx)\,[Y(x) + f(x)]$. On a, en effet, $\int \mu_n(dx)\, f(x) \to \int \mu(dx)\, f(x)]$ et $\int \mu_m\, C(x,y)$ tend vers la fonction $\varphi(x) = \int \mu(dy)\, C(x,y)$ continue, donc bornée sur S. On en déduit :

$$\iint \mu_n(dx)\, \mu_m(dy)\, C(x,y) = \iint \mu_n(dx)\, \big(\mu_m(dy) - \mu(dy)\big)\, C(x,y)$$
$$+ \int \mu_n(dx)\, \mu(dy) \quad C(x,y)$$

Pour m assez grand, la valeur absolue du 1er terme du 2ème membre est $\leqslant \varepsilon$, et on obtient

$$\lim_{n,m} \iint \mu_n(dx)\, \mu_m(dy)\, C(x,y) = \iint \mu(dx)\, \mu(dy)\, C(x,y)$$

c'est-à-dire le critère de Cauchy.

La réciproque n'est pas vraie : la convergence dans $\overline{\overline{H}}_M$ n'entraîne pas la convergence étroite des mesures μ_n, comme nous l'avons déjà indiqué, et les inclusions $H \subset \overline{\overline{H}}_M \subset \overline{\overline{H}}$ sont en général strictes.

En particulier, le vecteur Y_f, qui existe dès que $1 \notin \overline{\overline{H}}_f$, n'est pas forcément dans $\overline{\overline{H}}_M$.

3/ La condition $Y_f \in \overline{H}_M$, c'est-à-dire l'existence d'une mesure $\mu_f \in M$ avec

$$Y_f = \int \mu_f(dx)\; Y(x)$$

entraîne pour tout $y \in S$:

$$< Y_f, Y(y) > = \int \mu_f(dx) < Y(x), Y(y) >$$

c'est-à-dire :

$$f(y) = \int \mu_f(dx)\; C(x,y)$$

Inversement, si $f(y)$ est de cette forme, le vecteur $Y_f = \int \mu_f(dy)\; Y(x) \in \overline{H}_M$ vérifie bien la condition $< Y_f, Y(y) > = f(y)$

Dans les applications, on s'efforcera de réaliser cette condition $Y_f \in \overline{H}_M$ (moyennant laquelle l'estimation numérique de la dérive sera possible à partir d'une réalisation). Il suffira pour celà de ne choisir que des fonctions f de la forme

$$\int \mu(dx)\; C(x,y), \quad \mu \in M$$

Dans tout ce qui suit, nous supposons réalisées les 4 propriétés équivalentes de la proposition 1.

<u>Caractérisation de Y_f</u> - Soit Z la projection de 1 dans $\overline{H}_f$. On a :

$$< Z - 1, Y(x) + f(x) > = 0, \; \forall x \in S$$

c'est-à-dire $< Z, Y(x) > = \big(1 - E(Z)\big) f(x)$. Posons $X = Z - E(Z)$ $(X \in \overline{H})$, on a aussi :

$$< X, Y(x) > = \big(1 - E(Z)\big) f(x)$$

On voit que la condition $E(\mathcal{Z}) = 1$ équivaut à $1 \in \overline{H}_f$. En effet, $E(\mathcal{Z}) = 1$ entraîne $\langle X, Y(x) \rangle = 0$ pour tout $x \in S$, donc $X = 0$ puisque $X \in \overline{H}$, c'est-à-dire $\mathcal{Z} = E(\mathcal{Z}) = 1 \in \overline{H}_f$. Inversement, si $1 \in \overline{H}_f$, on a $\mathcal{Z} = 1$ et à fortiori $E(\mathcal{Z}) = 1$.

Si $E(\mathcal{Z}) \neq 1$, c'est-à-dire si $1 \notin \overline{H}_f$, on a alors $Y_f = \frac{X}{1 - E(\mathcal{Z})}$, c'est-à-dire, en désignant par Π_f le projecteur de $\overline{H}_f$:

$$Y_f = \frac{\Pi_f\, 1 - E(\Pi_f\, 1)}{1 - E(\Pi_f\, 1)}$$

II - L'ESTIMATEUR OPTIMAL DE LA DERIVE.

Soit $\mathcal{Z}(x) = Y(x) + a_\ell f^\ell(x)$ une fonction aléatoire déduite de la fonction aléatoire $Y(x)$ précédente par addition d'une dérive $a_\ell f^\ell(x)$, où les $f^\ell(x)$ sont k fonctions connues (linéairement indépendantes pour $x \in S$, et telles que $1 \notin H_{f^\ell}$) et les a_ℓ des coefficients constants, mais inconnus.

Nous nous proposons de former l'estimateur optimal de ces coefficients a_ℓ, c'est-à-dire de trouver les variables aléatoires $A_\ell \in \overline{H}_{a_\ell f^\ell}$ (appartenant à l'espace de Hilbert engendré par les combinaisons linéaires des $\mathcal{Z}(x)$, $x \in S$, qui constitue la seule réalité accessible expérimentalement) et vérifiant les conditions :

$$\begin{cases} E\,(A_\ell) = a_\ell \\ D^2\,(b^\ell A_\ell) \text{ minimum quels que soient les k nombres } b^\ell \end{cases}$$

ces conditions devant être de plus réalisées dans tous les

$H_{a_\ell f^\ell}$ à la fois (quels que soient les a_ℓ) : c'est un tel estimateur que nous appelons <u>universel</u>.

Soient $Y^\ell \in \overline{H}$ les variables aléatoires telles que :

$$< Y(x), Y^\ell > = f^\ell(x) \qquad \forall x \in S$$

Ces vecteurs de $\overline{H}$ sont linéairement indépendants (puisque les $f^\ell(x)$ le sont) et le déterminant des $< Y^i, Y^j >$ est différent de 0. La matrice des $< Y^i, Y^j >$ admet une inverse, soit g_{ij}, et les

$$Y_i = g_{ij} Y^j$$

forment, pour l'espace vectoriel engendré par les Y^ℓ, la base duale de (Y^ℓ) :

$$< Y_i, Y^j > = \delta_i^{\ j}$$

a/ Considérons, dans $\overline{H}$, la variété linéaire fermée $\overline{V}(b)$ définie par les équations

$$< Y, Y^\ell > = b^\ell \qquad (\ell = 1,2,\ldots k)$$

(c'est la fermeture dans $\overline{H}$ des $Y \in H$ de la forme $Y = \int \lambda(dx)\, Y(x)$, $\lambda \in \Lambda$ vérifiant $\int \lambda(dx)\, f^\ell(x) = b^\ell$)

A tout $X \in \overline{V}(b)$ est associé dans chaque $\overline{H}_{a_\ell f^\ell}$ l'élément

$$Z_a = X + a_\ell < X, Y^\ell > = X + a_\ell b^\ell \quad \in \overline{H}_{a_\ell f^\ell}$$

et cet élément vérifie par construction :

$$E(Z_a) = a_\ell b^\ell$$

Inversement, si une famille $(Z_a)_{a \in \mathbb{R}^k}$ d'éléments vérifiant $Z_a \in H_{a_\ell f^\ell}$ et $Z_a - E(Z_a) = Z_o$ pour tout $a \in \mathbb{R}^k$ remplit de plus la condition $E(Z_a) = a_\ell b^\ell$ pour tout $a \in \mathbb{R}^k$, on a

$$E(Z_a) = a_\ell < Z_o, Y^\ell > = a_\ell b^\ell \quad , \quad \forall a \in \mathbb{R}^k,$$

donc $\quad < Z_o, Y^\ell > = b^\ell$, et $\quad Z_o \in \overline{V}(b)$

Ainsi, la variété $\overline{V}(b)$ est constituée des éléments $X \in \overline{H}$ dont les images inverses par $\Phi_{a_\ell f^\ell}$ sont dans chaque $\overline{H}_{a_\ell f^\ell}$ des estimateurs sans biais de $a_\ell b^\ell$. Cette variété engendre donc les estimateurs universels que nous cherchons.

Si $X \in \overline{V}(b)$ est dans $\overline{H}_M$, soit

$$X = \int \mu(dx)\ Y(x) \quad , \quad \mu \in M$$

sa représentation. L'élément $Z_a = X + < X, Y^\ell > a_\ell = X + a_\ell b^\ell$ est alors de la forme

$$Z_a = \int \mu(dx)\ [Y(x) + a_\ell f^\ell(x)] = \int \mu(dx)\ Z(x)$$

et l'on voit que cet estimateur est accessible numériquement à partir de la réalisation de $Z(x)$, c'est-à-dire à partir des seules données effectivement disponibles.

b/ Parmi les estimateurs sans biais Z_a que nous venons de construire, il reste à choisir le meilleur, c'est-à-dire celui qui minimise la variance $D^2(\xi_a) = ||\xi_a - E(\xi_o)||^2$. Comme $\xi_a - E(\xi_a) = X$ ne dépend pas de a, il suffit de chercher dans $\overline{V}(b)$ l'élément $X = X(b)$ de norme minimale : son image inverse dans chaque $H_{a_\ell f^\ell}$ constituera l'estimateur optimal cherché.

Projetons donc 0 dans la variété fermée $\overline{V}(b)$: nous voyons que $X(b)$ est l'unique élément de $\overline{V}(b)$ vérifiant :

$$< Y, X(b)> = 0 \quad \text{pour tout } Y \in \overline{V}(o) \ (< Y, Y^\ell > = 0)$$

Si Y est un élément quelconque de $\overline{H}$, posons :

$$\begin{cases} < Y, Y^\ell > = C^\ell \\ C_i = g_{ij} C^j \qquad (C^\ell = C_i < Y^i, Y^\ell >) \end{cases}$$

Le vecteur $Y - C_\ell Y^\ell$ est alors orthogonal aux Y^ℓ, donc appartient à $\overline{V}(b)$, et $X(b)$ vérifie

$$< Y, X(b)> = C_\ell < Y^\ell, X(b)> \quad , \forall Y \in \overline{H}$$

Il suffit d'ailleurs de considérer les éléments $Y(x)$, $x \in S$, qui engendrent $\overline{H}$. Pour $Y(x)$, on a par définition de Y^ℓ :

$$C^\ell = < Y, Y^\ell > = f^\ell(x)$$

Ainsi $X(b)$ est l'unique élément de $\overline{V}(b)$ vérifiant :

$$< Y(x), X(b)> = < X, Y^i > \ g_{i\ell} \ f^\ell(x) \quad , \forall x \in S$$

c'est-à-dire une relation de la forme :

$$< Y(x), X(b)> = \mu_\ell \ f^\ell(x) \quad , \forall x \in S$$

Inversement, si $X \in \overline{V}(b)$ vérifie cette relation pour des constantes μ_ℓ données, on a :

$$X = \mu_\ell Y^\ell$$

et par suite $\langle X, Y^i \rangle = \mu_\ell \langle Y^\ell, Y^i \rangle$, d'où résulte bien

$$\mu_\ell = \langle X, Y^i \rangle g_{i\ell} \quad \text{et} \quad X = X(b)$$

En résumé, l'estimateur universel optimal cherché de $a_\ell b^\ell$ est l'image dans $\overline{H}_{a_\ell f^\ell}$ de l'unique élément X de $\overline{H}$ vérifiant le système :

$$\begin{cases} \langle Y(x), X(b) \rangle = \mu_\ell f^\ell(x) & \forall x \in S \\ \langle X(b), Y^\ell \rangle = b^\ell & (\ell = 1,2,\ldots k) \end{cases}$$

Les μ_ℓ sont les paramètres de Lagrange habituels, et les secondes relations permettent de les déterminer.

c/ Comme les Y^ℓ vérifient $\langle Y(x), Y^\ell \rangle = f^\ell(x)$, on a

$$\begin{cases} X(b) = \mu_\ell Y^\ell \\ \mu_\ell = g_{\ell i} b^i \end{cases}$$

et $X(b) = g_{\ell i} b^i Y^\ell$ dépend linéairement des constantes arbitraires b^i. Introduisons donc le vecteur aléatoire Y_i défini par

$$Y_i = g_{\ell i} Y^\ell$$

Ce vecteur est l'unique solution du système :

$$(I) \qquad \begin{cases} \langle Y(x), Y_i \rangle = \mu_{i\ell} f^\ell(x) & , \quad x \in S \\ \langle Y_i, Y^\ell \rangle = \delta_i^\ell \end{cases}$$

et, quelles que soient les constantes b^ℓ, on a $X(b) = b^i Y_i$.
On note aussi que les paramètres de Lagrange :

$$\mu_{ij} = \langle Y_i, Y_j \rangle = g_{ij}$$

constituent la <u>matrice des covariances</u> des Y_i. L'image inverse de Y_i dans chaque $\overline{H}_{a_\ell f^\ell}$ constitue <u>l'estimateur optimal universel</u> A_i du vecteur dérive a_i : on a

$$A_i = Y_i + a_\ell \langle Y_i \; Y^\ell \rangle$$

c'est-à-dire, compte tenu de la seconde relation (I)

$$A_i = Y_i + a_i$$

Dans chaque $\overline{H}_{a_\ell f^\ell}$, le vecteur A vérifie donc

$E(A) = a$ et $D^2(A_i b^i)$ minimale $\forall b \in \mathbb{R}^k$

<u>Cas Particuliers</u>. Supposons que les Y^i ou, ce qui revient au même, les Y_i, soient dans $\overline{H}_M$, autrement dit qu'il existe des mesures $\lambda_i \in M$ avec $Y_i = \int \lambda_i(dx)\, Y(x)$. Ces mesures λ_i constituent l'unique solution du système :

$$(\mathrm{I}') \qquad \begin{cases} \int \lambda_i(dy)\, C(x,y) = \mu_{i\ell}\, f^\ell(x) \quad , \quad x \in S \\ \int \lambda_i(dx)\, f^\ell(x) = \delta_i^\ell \end{cases}$$

et les constantes de Lagrange μ_i^ℓ sont données par :

$$\mu_{ij} = \iint \lambda_i(dx)\, C(x,y)\, \lambda_j(dy)$$

Dans chaque $\overline{H}_{a_\ell f^\ell}$, l'estimateur universel A_i admet alors la représentation :

$$A_i = \int \lambda_i(dx)\ Z(x)$$

et se calcule numériquement à partir de la réalisation $Z(x) = Y(x) + a_\ell f^\ell(x)$ effectivement disponible.

De plus, la <u>matrice des covariances des A_i</u> est identique à la matrice μ_{ij} des paramètres de Lagrange du système I' :

$$E\left[(A_i - a_i)(A_j - a_j)\right] = \mu_{ij}$$

Dans le cas général où les Y_i ne sont pas dans $\overline{H}_M$, A_i est cependant limite dans L^2 d'une suite :

$$A^n{}_i = \int \lambda^n{}_i(dx)\ Z(x) \qquad , \qquad \lambda^n{}_i \in \Lambda$$

de combinaisons linéaires des $Z(x)$. Les $A^n{}_i$ constituent une suite d'estimateurs sans biais dont les variances, décroissantes, tendent vers la variance minimale; mais cette suite n'est pas nécessairement convergente numériquement pour une réalisation donnée de $Z(x)$.

<u>Dans le cas fini</u>, les Y_i sont toujours dans $\overline{H}_M$, et le système I' a toujours une solution. Pour $S = \{x_\alpha,\ \alpha = 1,2,\dots m\}$ ce système s'écrit :

$$(\mathrm{I}'') \qquad \begin{cases} \lambda_i^\beta\ \sigma_{\alpha\beta} = \mu_{i\ell}\ f_\alpha^\ell \\ \lambda_i^\beta\ f_\beta^\ell = \delta_i^\ell \end{cases}$$

$[\alpha,\beta = 1,\dots m \; ; \; i, \ell = 1,2,\dots k \; ; \; f_\alpha^\ell = f^\ell(x_\alpha), \; \sigma_{\alpha\beta} = C(x_\alpha,x_\beta)]$

C'est un système de m k + k^2 équations où figurent les m k inconnues λ_i^β et les k^2 paramètres de Lagrange $\mu_{i\ell}$ (d'ailleurs nécessairement symétriques : $\mu_{ij} = \mu_{ji}$)

L'estimateur universel de la dérive est :

$$A_i = \lambda_i^\alpha Z_\alpha \qquad \left(Z_\alpha = Z(x_\alpha)\right)$$

Il vérifie :

$$\begin{cases} E(A_i) = a_i \\ E(A_i A_j) - a_i a_j = \mu_{ij} = \lambda_i^\beta \lambda_j^\alpha \sigma_{\alpha\beta} \end{cases}$$

<u>Exemple</u> : Soit $Z(x)$ une fonction aléatoire, d'espérance $af(x)$ (a constante inconnue, $f(x)$ fonction connue) et de covariance $C(x,y)$ (avec $1 \notin \bar{H}_f$). Par exemple, si $f(x) = 1$ et $C(x,y) = C(x-y)$, le problème posé est celui de <u>l'estimation de l'espérance</u> $a = E[Z(x)]$ d'une fonction aléatoire stationnaire d'ordre 2 dont on connaît à priori la covariance $C(x,y)$.

L'estimateur optimal universel A a pour image $A - E(A) = A - a$ dans $\bar{H}$ l'élément X vérifiant :

$$\begin{cases} \langle Y(x), X \rangle = \mu f(x) \\ \langle X, X_f \rangle = 1 \end{cases}$$

Par conséquent $X = \dfrac{X_f}{||X_f||^2}$ et $\mu = \dfrac{1}{||X_f||^2} = ||X||^2 = D^2(X)$

Si $X_f = \int \lambda_f(dx)\ Y(x)$ pour $\lambda_f \in M$, on a

$$A = \int \lambda(dx)\ Z(x)$$

dans chaque $\overline{H}_{af}$, avec une mesure $\lambda \in M$ vérifiant :

$$\int \lambda(dy)\ C(x,y) = \mu\ f(x) \qquad x \in S$$
$$\int \lambda(dx)\ f(x) = 1$$

et de plus $E(A) = a,\ \forall a \in \mathbb{R}$ et $D^2(A) = \mu = \iint \lambda(dx)\ C(x,y)\ \lambda(dy)$.

La condition pour qu'il en soit ainsi est que la fonction $f(x)$ soit de la forme

$$f(x) = \int \lambda_f(dy)\ C(x,y) \qquad x \in S$$

autrement dit soit le potentiel d'une mesure λ_f pour le noyau potentiel $C(x,y)$. En particulier, pour $f = 1$, l'estimateur universel de l'espérance $E(Z) = a$ est de la forme $\int \lambda(dx)\ Z(x)$ si et seulement si on peut trouver une mesure $\lambda_1 \in M$ telle que :

$$1 = \int \lambda_1(dx)\ C(x,y) \quad , \forall\ y \in S$$

autrement dit si le noyau $C(x,y)$ admet un potentiel d'équilibre sur S. La capacité de S est alors $\int \lambda_1(dx) = \frac{1}{\mu} = \frac{1}{D^2(A)}$

III - COMPARAISON AVEC LE MAXIMUM DE VRAISEMBLANCE DANS LE CAS GAUSSIEN

Montrons que la méthode du maximum de vraisemblance conduit, dans le cas gaussien fini, au même estimateur de la dérive.

Soient Z_α ($\alpha = 1,2,\ldots m$) m variables gaussiennes, $\sigma_{\alpha\beta}$ leurs covariances, $B^{\alpha\beta}$ la matrice inverse des $\sigma_{\alpha\beta}$, et

$$E(Z_\alpha) = a_i \, f_\alpha{}^i \qquad (i = 1,2,\ldots k)$$

leurs espérances. On connaît les $\sigma_{\alpha\beta}$ et les $f^i{}_\beta$, mais non les a_i. Formons le logarithme de la densité de probabilité Ψ des Z_α :

$$\log \Psi = - m \log 2\pi - \frac{1}{2} \log \left(\mathrm{Det} \, (\sigma_{\alpha\beta})\right) - \frac{1}{2} B^{\alpha\beta}(z_\alpha - a_i f^i_\alpha)(z_\beta - a_j f^j_\beta)$$

et annulons les dérivées partielles en a_i ($i = 1,2,\ldots k$) : l'estimateur A_i du maximum de vraisemblance est solution du système :

$$(\mathrm{I'''}) \qquad A_j \, f^j{}_\beta \, f^i{}_\alpha \, B^{\alpha\beta} = z_\beta \, f^i{}_\alpha \, B^{\alpha\beta}$$

(k équations, k inconnues). La solution de ce système est de la forme :

$$A_i = \Lambda_i{}^\alpha \, z_\alpha$$

Montrons que cette matrice $\Lambda_i{}^\alpha$ ne diffère pas de la matrice $\lambda_i{}^\alpha$, solution de (II") : la première relation II' donne :

$$\lambda_i{}^\beta = \mu_{i\ell} \, f^\ell_\alpha \, B^{\alpha\beta}$$

et la deuxième exprime que la matrice $\mu_{i\ell}$ vérifie :

$$\mu_{i\ell} \, f^\ell_\alpha \, B^{\alpha\beta} \, f^j{}_\beta = \delta_i{}^j$$

(I"'), de son côté, donne :

$$\Lambda^\beta_i \, f^i{}_\alpha \, f^j{}_\gamma \, B^{\alpha\gamma} = f^j{}_\alpha \, B^{\alpha\beta}$$

Ainsi, les systèmes (I") et (I"') sont équivalents : dans le cas gaussien fini, l'estimateur universel de la dérive est identique à l'estimateur du maximum de vraisemblance.

Remarque - Comme les systèmes (I") et (I"') sont équivalents, on peut calculer directement A_i par le système (I"') qui ne comporte que k équations à k inconnues (au lieu de $m\,k + k^2$ pour l'autre système). Mais il faut ensuite calculer directement les covariances $E(A_i A_j)$. Le système (I"') donne ces covariances à titre de sous-produit, puisqu'elles coïncident avec les paramètres de Lagrange.

IV - ESTIMATION DE LA DERIVE DANS LE CAS INTRINSEQUE

Plaçons-nous maintenant dans le cas intrinsèque, c'est-à-dire supposons que Y(x) soit une fonction aléatoire vérifiant :

$$\begin{cases} E\,[Y(x+h) - Y(x)] = 0 \\ \frac{1}{2}\,D^2\,[Y(x+h) - Y(x)] = \gamma(h) \end{cases}$$

et considérons les fonctions aléatoires $Z(x) = Y(x) + a_\ell f^\ell(x)$. On peut reprendre les raisonnements précédents en prenant comme espace $\overline{H}$ et $\overline{H}_f$ les espaces de Hilbert engendrés par les combinaisons linéaires finies $\int \lambda(dx)\, Y(x)$ et $\int \lambda(dx)\, Z(x)$ avec $\lambda \in \Lambda$ vérifiant de plus la condition $\int \lambda(dx) = 0$ - soit $\lambda \in \Lambda_0$, ensemble des mesures à support fini $\subset S$ et de somme nulle -

On peut aussi se ramener au cas précédent en remplaçant $Y(x)$ par $Y(x) - Y(x_o)$, $x_o \in S$: cette nouvelle fonction aléatoire a une espérance nulle et admet la covariance :

$$C(x,y) = -\gamma(x-y) + \gamma(x-x_o) + \gamma(y-x_o)$$

Moyennant la condition $1 \notin \bar{H}_{f\ell}$, on introduit les variables aléatoires $Y^{\ell} \in \bar{H}_f$ vérifiant :

$$< Y(x) - Y(x_o), Y^{\ell} > = f^{\ell}(x) - f^{\ell}(x_o) \quad \forall x \in S$$

et le système (I) devient

$$\begin{cases} < Y(x) - Y(x_o), Y_i > = \mu_{i\ell} \left(f^{\ell}(x) - f^{\ell}(x_o)\right) \\ < Y_i, Y^{\ell} > = \delta_i^{\ell} \end{cases}$$

L'estimateur universel $A_i = Y_i + a_{\ell} < Y_i, Y^{\ell} > = Y_i + a_i$ vérifie alors $E(A_i) = a_i$ et $D^2 (b^i A_i)$ minimal $\forall b \in \mathbb{R}^k$. Si les $Y_i \in \bar{H}_M$, on a :

$$A_i = \int \lambda_i(dx)\, Z(x)$$

avec des mesures $\lambda_i \in M$ vérifiant :

$$\left| \begin{array}{l} \int \lambda_i(dy)\, [-\gamma(x-y) + \gamma(x-x_o) + \gamma(y-y_o)] = \mu_{i\ell} \left(f^{\ell}(x) - f^{\ell}(x_o)\right) \\ \quad \int \lambda_i(dx)\, f^{\ell}(x) = \delta_i^{\ell} \\ \quad \int \lambda_i(dx) = 0 \end{array} \right.$$

et ce système équivaut au suivant :

$$-\int \lambda_0(dy)\,\gamma(x-y) = \mu_{i0} f^{\ell}(x) + \mu_0$$

$$(II')\quad \begin{cases} -\int \lambda_i(dy)\,\gamma(x-y) = \mu_{i\ell}\, f^{\ell}(x) + C_i \quad , \quad \forall x \in S \\ \int \lambda_i(dx)\, f^{\ell}(x) = \delta_i^{\ell} \\ \int \lambda_i(dx) \quad = 0 \\ \int \lambda_0(dx) = 1 \quad , \quad \int f^{\ell}(x)\,\lambda_0(dx) = 0 \quad (\ell > 0) \end{cases}$$

Il ne diffère de (I') que par l'addition des paramètres de Lagrange C_i et des conditions correspondantes $\int \lambda_i = 0$. On a encore

$$E\,(A_i) = a_i$$

$$E\,(A_i A_j) - a_i a_j = \mu_{ij} = -\int \lambda_i(dx)\,\gamma(x-y)\,\lambda_j(dy)$$

Dans le <u>cas fini</u>, (I") est remplacé par :

$$(II'')\quad \begin{cases} -\lambda_i^{\beta}\ \gamma_{\alpha\beta} = \mu_{i\ell}\, f_{\alpha}^{\ell} + C_i \\ \lambda_i^{\beta}\ f_{\beta}^{\ell} = \delta_i^{\ell} \\ \sum_{\beta} \lambda_i^{\beta} = 0 \end{cases}$$

avec toujours :

$$\begin{cases} A_i = \lambda_i^{\alpha}\, z_{\alpha} \\ E(A_i) = a_i \\ E(A_i A_j) - a_i a_j = \mu_{ij} = -\lambda_i^{\beta}\,\lambda_j^{\alpha}\,\gamma_{\alpha\beta} \end{cases}$$

Il n'y a pas de difficultés particulières à vérifier que l'estimateur du maximum de vraisemblance du cas gaussien fini conduit encore aux mêmes résultats.

Exemple du Variogramme Linéaire - Prenons $\gamma(h) = |h|$ dans l'espace à une dimension, $S = [0, L]$ intervalle fermé de longueur $L < \infty$, et comme fonctions $f^{\ell}(x)$ les monomes x^{ℓ}, $\ell = 1,2,\ldots k$ (il n'y a aucun intérêt à introduire $x^0 = 1$, puisqu' $Y(x)$, en tant que fonction intrinsèque, n'est en réalité définie qu'à une constante près et qu'on travaille toujours sur des accroissements $Y(x+h) - Y(x)$).-

Les mesures de Dirac δ et δ_L placées en 0 et en L vérifient :

$$\int \delta \; |x-y| = y \quad , \int \delta_L(dx)|x-y| = L - y$$

D'où aussi :

$$\text{(a)} \quad \begin{cases} \displaystyle\int \frac{1}{L} (\delta(dx) + \delta_L(dx) \; |x-y| = 1 \\ \displaystyle\int \Big(\delta_L(dx) - \delta(dx)\Big) |x-y| = L - 2y \end{cases}$$

De même, la mesure de densité x^p sur $[0, L]$ vérifie :

$$\int_0^L x^p \; |x-y| \; dx = \int_0^y x^p(y-x) \; dx + \int_y^L x^p(x-y) \; dx$$

c'est-à-dire :

$$\text{(b)} \quad \int_0^L x^p|x-y| \; dx = \frac{2}{(p+1)(p+2)} \; y^{p+2} - y \frac{L^{p+1}}{p+1} + \frac{L^{p+2}}{p+2}$$

L'estimateur optimal de la dérive $\sum_1^k a_n x^n$ est le polynôme $\sum_1^k A_n x^n$ dont les coefficients A_n sont donnés par :

(c) $$A_n = \int_0^L \lambda_n(dx)\, \check{z}(x)$$

avec des mesures λ_n vérifiant le système (II') :

(d) $$\begin{cases} -\int \lambda_n(dx)\ |x-y| = \sum_{\ell} \mu_{n_\ell}\ y^{\ell} + C_n \\ \int \lambda_n(dx)\ x^{\ell} = \delta_i^{\ell} \\ \int \lambda_n(dx) = 0 \end{cases}$$

Il résulte de (a) et (b) que toute combinaison linéaire des δ, δ_L et des x^p, $p \leqslant n-2$ vérifiera la première relation (d). On pourra donc toujours trouver n combinaisons linéaires de ces mesures vérifiant les deux autres équations (d).

- Effectuons explicitement ce calcul dans le cas d'une dérive quadratique $a_1\, x + a_2\, x^2$ - On prendra

$$\begin{cases} \lambda_1(dx) = A\ (\delta_L - \delta) + B\ (dx - \frac{L}{2}\ (\delta+\delta_L) \\ \lambda_2(dx) = C\ (\delta_L - \delta) + D\ (dx - \frac{L}{2}\ (\delta+\delta_L) \end{cases}$$

de sorte que $\int \lambda_1(dx) = \int \lambda_2(dx) = 0$ - La condition $\int x\ \lambda_2(dx) = 0$ donne $C = 0$ et $\int x^2\ \lambda_2(dx) = 1$ conduit à : $D = -\frac{6}{L^3}$. Les deux conditions analogues en λ_1 donnent le système :

$$\begin{cases} \int \lambda_1(dx)\ x = A\ L = 1 \\ \int x^2\ \lambda_1(dx) = A\ L^2 - B\frac{L^3}{6} = 0 \end{cases}$$

D'où $$A = \frac{1}{L} \quad \text{et } B = \frac{6}{L^2}$$

On a ainsi la solution cherchée :

$$\begin{cases} \lambda_1(dx) = \frac{1}{L}(\delta_L - \delta) + \frac{6}{L^2}\left(dx - \frac{L}{2}(\delta + \delta_L)\right) \\ \lambda_2(dx) = -\frac{6}{L^3}\left(dx - \frac{L}{2}(\delta + \delta_L)\right) \end{cases}$$

Les estimateurs universels A_1 et A_2 des coefficients a_1 et a_2 de la dérive quadratique sont donc :

$$\begin{cases} A_1 = \frac{1}{L}(Z(L) - Z(o)) + \frac{6}{L}\left[\bar{Z} - \frac{1}{2}\left(Z(L) + Z(o)\right)\right] \\ A_2 = -\frac{6}{L^2}\left[\bar{Z} - \frac{1}{2}\left(Z(L) + Z(o)\right)\right] \end{cases}$$

$\left(\bar{Z} = \frac{1}{L}\int_o^L Z(x)\,dx\right)$ - Pour calculer la matrice des covariances des $A_i A_j$, il suffit de remarquer que l'on a :

$$-\int \lambda_1(dx)\,|x-y| = \frac{2y}{L} - 1 - \frac{6}{L^2}\left(y^2 - y\,L + \frac{L^2}{2} - \frac{L^2}{2}\right)$$

$$= -1 + 8\frac{y}{L} - 6\frac{y^2}{L^2}$$

$$-\int \lambda_2(dx)\,|x-y| = \frac{6}{L^3}\left(y^2 - yL + \frac{L^2}{2}\right) - \frac{3}{L^2}\,L = -6\frac{y}{L^2} + \frac{6}{L^3}y^2$$

La matrice des covariances des A_1, A_2 dont les composantes sont les paramètres de Lagrange, est donc :

$$\begin{cases} \mu_{11} = \frac{8}{L} \\ \mu_{12} = \mu_{21} = -\frac{6}{L^2} \\ \mu_{22} = \frac{6}{L^3} \end{cases}$$

Cas Discret - Soit, de même, à estimer les coefficients a_1 et a_2 de la dérive quadratique à partir de $f(i\ h)$, $i = 0,1,2,\ldots n$ ($nh = L$) - La mesure ν_1 : $\int \nu_1 f = f(nh) - f(o)$ et la mesure ν_2 :

$$\int \nu_2 f = \sum_{i=o}^{n} f(i\ h) - \frac{n+1}{2} (f(nh)+f(o))$$

vérifient les relations :

$$\begin{cases} \int \nu_1(dx)\ |x-x_j| = (n-2j)\ h \\ \int \nu_2(dx)\ |x-x_j| = -j(n-j)\ h \end{cases}$$

On prendra :

$$\begin{cases} \lambda_1 = A\ \nu_1 + B\ \nu_2 \\ \lambda_2 = C\ \nu_1 + D\ \nu_2 \end{cases}$$

La condition $\int x\lambda_2(dx) = 0$ donne $C = 0$, $\int x^2\ \lambda_2(dx)$ donne $D = \frac{6}{n(1-n^2)\ h^2}$. Pour λ_1, on trouve le système

$$\begin{cases} \int x\ \lambda_1(dx) = A\ n\ h = 1 \\ \int x^2\ \lambda_2(dx) = h^2\ (A\ n^2 + B\ \frac{n(1-n^2)}{6} = 0 \end{cases}$$

D'où

$$A = \frac{1}{n\ h}, \quad B = \frac{6}{h(n^2-1)}$$

On a donc :

$$\begin{cases} \lambda_1 = \frac{1}{n\ h}\ \nu_1 + \frac{6}{h(n^2-1)}\ \nu_2 \\ \lambda_2 = \frac{6}{n(1-n^2)h^2}\ \nu_2 \end{cases}$$

D'où les estimateurs cherchés :

$$\begin{cases} A_1 = \dfrac{1}{n\,h}\,(Z_n - Z_o) + \dfrac{6}{(n-1)h} \left(\bar{Z} - \dfrac{1}{2}\,(Z_n + Z_o)\right) \\[2ex] A_2 = \dfrac{6}{h^2\,n(n-1)} \left(\bar{Z} - \dfrac{1}{2}\,(Z_n + Z_o)\right) \end{cases}$$

Pour obtenir la matrice des covariances, calculons les expressions

$$-\int \lambda_1(dx)\ |x - jh| = -\frac{n-2j}{n} + \frac{6}{n^2-1}\ j\ (n-j) =$$

$$\doteq -1 + \frac{8n^2-1}{n(n^2-1)}\ j - \frac{6\ j^2}{n^2-1}$$

$$-\int \lambda_2(dx)\ |x - jh| = -\frac{6}{n^2-1}\ \frac{j}{h} + \frac{6}{n(n^2-1)}\ \frac{j^2}{h}$$

On en déduit :

$$\begin{cases} \mu_{11} = \dfrac{8\ n^2-1}{n\ (n^2-1)}\ \dfrac{1}{h} \\[2ex] \mu_{12} = \mu_{21} = -\dfrac{6}{(n^2-1)}\ \dfrac{1}{h^2} \\[2ex] \mu_{22} = \dfrac{6}{n\ (n^2-1)}\ \dfrac{1}{h^3} \end{cases}$$

Remarque : Si l'on multiplie $\mathcal{C}(x,y)$ ou $\gamma(x-y)$ par une constante arbitraire ϖ, les systèmes (I) ou(II) ne sont pas modifiés, à ceci près que les paramètres de Lagrange doivent être multipliés par ϖ. L'expression de l'estimateur optimal A_i en fonction de $Z(x)$ n'est donc pas non plus modifiée, seules les covariances des A_i, A_j sont multipliées par ϖ.

Cette remarque est capitale pou les applications. Il suffit, en effet, de savoir que la fonction aléatoire $Z(x) - a_i f^i(x)$ admet un variogramme $\gamma(h) = \varpi\, \gamma_0(h)$ pour qu'il soit possible d'appliquer les résultats précédents, même si le facteur ϖ est indéterminé, pourvu que γ_0 soit connu. L'universalité de notre estimateur relativement aux coefficients a de la dérive s'étend même à un facteur arbitraire ϖ affectant le variogramme.

Beaucoup de variables régionalisées, de par leur nature physique, peuvent être assimilées <u>localement</u>, et moyennant une correction convenable de dérive, à des réalisations de fonctions aléatoires intrinsèques possédant un variogramme isotrope et linéaire au voisinage de l'origine. Pour estimer cette dérive, on doit procéder localement, puisque cette assimilation n'a de valeur que locale, c'est-à-dire n'utiliser que des points expérimentaux voisins du point x où l'on veut calculer $A_i\, f^i(x)$. Dans ce même voisinage, le variogramme pourra souvent lui-même être remplacé par sa tangente à l'origine. Onnsera alors exactement dans les conditions voulues (Fonction aléatoire intrinsèque de variogramme $\varpi|h|$ avec un facteur ϖ inconnu) pour que les résultats précédents soient applicables.

Naturellement, il faudra s'assurer que ces diverses approximations sont légitimes. La méthode consistera à construire le variogramme expérimental γ^* de la variable régionalisée $Z(x)$ corrigée de la dérive estimée, et de le comparer avec son espérance mathématique $E[\gamma^*(h)]$ (calculée dans l'hypothèse où $Z(x)$ corrigée de la vraie dérive admet le variogramme $\gamma(h)$.

Il ne faut pas croire que l'on a $E[\gamma^*(h)] = \gamma(h)$. Du fait que l'on a remplacé la vraie dérive a par son estimateur optimal A, on doit s'attendre à l'apparition de biais irréductibles.

Le Variogramme des Résidus.

Dans les hypothèses précédentes, cherchons l'espérance du variogramme expérimental :

$$2\,\gamma^*(h) = \frac{1}{K(h)} \int k(x)k(x+h)\,[Z(x+h) - A_\ell f^\ell(x+h) - Z(x) + A_\ell f^\ell(x)]^2$$

de la variable régionalisée $Z(x)$ corrigée de la dérive estimée $Z(x)$ - $\Big(k(x)$ et $K(h)$, indicatrice et covariogramme de $S\Big)$

Quitte à remplacer $Z(x)$ par $Y(x) = Z(x) - a_\ell f^\ell(x)$, on peut toujours supposer $E(A_\ell) = a_\ell = 0$. Evaluons, avec cette simplification, l'espérance de

$$[\, Z(x+h) - Z(x) - A_\ell \Big(f^\ell(x+h) - f^\ell(x)\Big)\,]^2 = [\, Z(x+h) - Z(x)]^2$$

$$- 2\,[\, Z(x+h) - Z(x)]\, A_\ell \Big(f^\ell(x+h) - f(x)\Big)$$

$$+ A_\ell\, A_s\ \ [f^\ell(x+h) - f(x)][\, f^s(x+h) - f^s(x)]$$

Le premier et le troisième terme donnent $2\gamma(h)$ et

$$E(A_\ell A_s)\ [\ f^\ell(x+h) - f^\ell(x)]\ [\ f^s(x+h) - f^s(x)]$$

$\left(\text{avec } E(A_\ell A_s) = \mu_{\ell s}\right)$ - Pour calculer le deuxième, nous devons calculer :

$$E\left(A_\ell\ \ [Z(x+h) - Z(x)]\right)$$

Or $A_\ell = \mu_{\ell s}\ \ Y^s$, avec un vecteur Y^s vérifiant

$$< Y,\ Y^s > = \int \lambda(dx)\ f^s(x)$$

pour tout $Y = \int \lambda(dx)\ Z(x)$ avec $\lambda \in \Lambda$ et $\int \lambda(dx) = 0$. En particulier, on a :

$$E\ [A_\ell\ \left(Z(x+h) - Z(x)\right)\ ] = \mu_{\ell s} < Z(x+h) - Z(x),\ Y^s >$$

$$= \mu_{\ell s}\ [f^s(x+h) - f(x)]$$

Il en résulte que le second terme, au signe près, est le double du troisième, et il vient :

$$E\ [Z(x+h) - Z(x) - A_\ell\ \left(f^\ell(x+h) - f^\ell(x)\right)\ ]^2 = 2\ \gamma(h) -$$

$$- \mu_{\ell s}\ [f^\ell(x+h) - f^\ell(x)]\ [f^s(x+h) - f(x)]$$

En ce qui concerne le variogramme expérimental, on trouve donc :

$$E\ [2\gamma^*(h)] = 2\ \gamma(h) - \mu_{\ell s}\ \ T^{\ell s}(h)$$

avec

$$T^{\ell s}(h) = \frac{1}{K(h)} \int k(x)\ k(x+h)\ [f^{\ell}(x+h) - f^{\ell}(x)]\ [f^{s}(x+h) - f^{s}(x)]dx$$

Ce biais $\mu_{\ell s}\ T^{\ell s}$ a pour effet d'aplatir $\gamma^*(h)$ dès que h n'est plus très petit.

Exemple - Soit $\gamma(h) = |h|$ et une dérive quadratique sur le segment $[0, L]$ - Calculons les $T^{\ell s}$:

$$\begin{cases} T^{11} = \frac{1}{L-h} \int_0^{L-h} h^2\ dx = h^2 \\ T^{12} = T^{21} = \frac{1}{L-h} \int_0^{L-h} h(2hx + h^2)\ dx = h^3 + h^2(L-h) = L\ h^2 \\ T^{22} = \frac{1}{L-h} \int_0^{L-h} (2hx + h^2)^2 dx = \frac{4}{3} h^2(L-h)^2 + 2\ h^3(L-h) + h^4 \end{cases}$$

$$= \frac{h^2}{3}(h^2 - 2\ hL + 2\ L^2)$$

Les $\mu_{\ell s}$ ont été calculés plus haut ($\mu_{11} = \frac{8}{L}$, $\mu_{12} = -\frac{6}{L^2}$ et $\mu_{22} = \frac{6}{L^3}$). Tous calculs faits, il vient :

$$\mu_{\ell s}\ T^{\ell s} = 4\frac{h^2}{L} - 4\frac{h^3}{L^2} + 2\frac{h^4}{L^3}$$

et finalement :

$$E\,[\gamma^*(h)] = h - 2\frac{h^2}{L} + 2\frac{h^3}{L^2} - \frac{h^4}{L^4}$$

Le graphe correspondant est une parabole aplatie, symétrique autour de $h = \frac{L}{2}$. En particulier $E[\gamma^*(L)] = 0$ (contre

$\gamma(L) = L$) et $E[\gamma^*(\frac{L}{2})] = \frac{3}{16}\, L$ (contre $\gamma(\frac{L}{2}) = \frac{L}{2}$). Par contre, au voisinage de l'origine, $E(\gamma^*(h)]$ a le même comportement que $\gamma(h) = h$. C'est à cette parabole aplatie que l'on doit comparer le variogramme des résidus expérimentaux pour tester l'hypothèse de départ.

<u>Estimation du facteur ϖ</u> - Le variogramme réel pourrait bien être $\varpi\,\gamma(h)$ au lieu de $\gamma(h)$ - dans l'exemple précédent $\varpi|h|$ au lieu de $|h|$ - avec un facteur ϖ indéterminé. L'estimateur de la dérive n'est pas affecté par cette indétermination. Par contre, les covariances μ_{ij} des A_i, A_j et le $E[\gamma^*(h)]$ calculé ci-dessus sont multipliés par ϖ. Il peut donc y avoir intérêt à estimer ce facteur ϖ. On peut évidemment le faire en ajustant par affinité le $\gamma^*(h)$ expérimental à l'expression théorique de son espérance. On peut aussi utiliser la variance expérimentale des résidus :

$$s^2 = \frac{1}{S} \int_S [Z(x) - A_\ell f^\ell(x) - m]^2 dx$$

$$m = \frac{1}{S} \int_S [Z(x) - A_\ell f^\ell(x)]\, dx$$

Effectuons le calcul de l'espérance de s^2, en supposant, comme d'habitude, $E(A_\ell) = 0$. En posant :

$$\bar{Z} = \frac{1}{S} \int_S Z(x)\, dx, \text{ et } \quad \bar{f}^\ell = \frac{1}{S} \int_S f^\ell(x)\, dx$$

on trouve :

$$E(s^2) = \frac{1}{S}\int_S E\left(Z(x) - \bar{Z}\right)^2 dx - \frac{2}{S}\int_S \left(f^\ell(x)-\bar{f}^\ell\right) E[A_\ell(Z(x)-\bar{Z})]\, dx$$
$$+ \frac{2}{S}\ E(A_\ell A_s) \int_S \left(f^\ell(x)-\bar{f}^\ell\right)\left(f^s(x)-\bar{f}^s\right) dx$$

D'après les propriétés de l'estimateur optimal A_ℓ, on a justement

$$E\ [A_\ell\left(Z(x)-\bar{Z}\right)] = \mu_{\ell s}\left(f^s(x)-f^{\bar{s}}\right)$$

d'où par suite :

$$E(s^2) = \frac{1}{S}\int_S E[\left(Z(x)-\bar{Z}\right)^2]dx - \frac{\mu_{\ell s}}{S}\int_S [f^\ell(x)-\bar{f}^\ell][f^s(x)-\bar{f}^s]\ dx$$

On en tire :

$$E(s^2) = \frac{1}{S^2}\int K(h)\ \gamma(h)\ dh - \mu_{\ell s}\left[\frac{1}{S}\int_S f^\ell(x)\ f^s(x)\ dx - \bar{f}^\ell\bar{f}^s\right]$$

Le premier terme représente la variance de $Y(x)$ dans S en l'absence de dérive). Le deuxième terme est un biais, et provient du fait que les a_ℓ ont été remplacés par leurs estimations A_ℓ. Reportons-nous à l'expression de $E[\gamma^*(h)]$. On vérifie sans peine la relation :

$$\frac{1}{S^2}\int K(h)\ T^{\ell s}(h)\ dh = \frac{2}{S}\int_S f^\ell(x)\ f^s(x)\ dx - 2\ \bar{f}^\ell\bar{f}^s$$

Par suite, on trouve :

$$E(s^2) = \frac{1}{S^2}\int K(h)\ E\left(\gamma^*(h)\right)\ dh$$

Cette relation montre que l'on calcule $E(s^2)$ comme s'il s'agissait de la variance dans S d'une fonction aléatoire intrinsèque dont le variogramme serait $E(\gamma^*(h))$. Elle permet d'estimer ϖ, en comparant la variance expérimentale s^2 à l'expression obtenue dans le cas $\varpi = 1$:

$$\varpi^* = \frac{s^2}{\frac{1}{S^2}\int K(h)\, E\big(\gamma^*(h)\big)\, dh}$$

<u>Méthode du Maximum de vraisemblance</u>. Dans le cas gaussien fini, reportons-nous à l'expression de la densité log Ψ :

$$\log \Psi = -\, m \log 2\pi - \frac{1}{2} \log \ \mathrm{Det}\, \sigma - \frac{m}{2} \log \varpi - \frac{1}{2\varpi} B^{\alpha\beta}[Z_\alpha - A_i\, f^i{}_\alpha][Z_\beta - A_j\, f^j{}_\beta]$$

(avec ici $Z_\alpha = Z(x_\alpha) - Z(x_o)$, $f^i{}_\alpha = f^i(x_\alpha) - f^i(x_o)$ pour un point x_o choisi parmi les m + 1 points expérimentaux, et

$$\sigma_{\alpha\beta} = -\gamma(x_\alpha - x_\beta) + \gamma(x_o - x_\alpha) + \gamma(x_o - x_\beta)$$

En annulant la dérivée en ϖ, on trouve :

$$\varpi^* = \frac{1}{m} B^{\alpha\beta}[Z_\alpha - A_i\, f^i{}_\alpha][Z_\beta - A_j\, f^j{}_\beta]$$

C'est un estimateur sans biais. Comme les A_i vérifient (I''), ϖ^* se met aussi sous la forme :

$$\varpi^* = \frac{1}{m} B^{\alpha\beta} Z_\alpha Z_\beta - \frac{1}{m} A_i\, f^i{}_\alpha\, A_j\, f^j{}_\beta \ B^{\alpha\beta}$$

V - KRIGEAGE EN PRESENCE D'UNE DERIVE

Reprenons les notations des deux premiers chapitres, et soit $Y_o \in L^2(\Omega, \mathcal{A}, P)$ un vecteur d'espérance nulle et n'appartenant pas à $\overline{H}$; et soient b^ℓ ($\ell = 1,2,\ldots k$) k nombres donnés. A Y_o, nous faisons correspondre la variable aléatoire :

$$Z_a = Y_o + a_\ell b^\ell$$

pour chaque $(a_\ell) \in \mathbb{R}^k$. On a donc $E(Z_a) = a_\ell b^\ell$. Le problème du krigeage en présence de la dérive $a_\ell f^\ell(x)$ consiste à trouver un élément $Y^* \in \overline{H}$ tel que $Z_a^* = Y^* + a_\ell b^\ell \in \overline{H}_{a_\ell f^\ell}$ quel que soit le choix des a_ℓ, et minimisant la variance :

$$D^2 [Z_a^* - Z_a] = E[(Y^* - Y_o)^2$$

La première condition $(Y^* + a_\ell b^\ell \in \overline{H}_{a_\ell f^\ell}\ \forall (a_\ell) \in \mathbb{R}^k)$ entraîne, comme nous l'avons vu au paragraphe II, que l'on a :

$$Y^* \in \overline{V}(b)$$

où $\overline{V}(b)$ est la variété fermée engendrée par les $Y = \int \lambda(dx)\ Y(x) \in H$ vérifiant $\lambda \in \Lambda$ et $\int \lambda(dx)\ f^\ell(x) = b^\ell$. Par suite Y^* est la projection de Y_o dans $\overline{V}(b)$: il est univoquement défini par la condition :

$$< Y_o - Y^*, Y > = 0 \qquad \forall Y \in \overline{V}(0)$$

En procédant exactement comme nous l'avons fait pour établir le système (I), on en déduit que Y^* est l'unique solution du système :

$$
(\mathrm{III})\begin{cases} \langle Y(x),\ Y^* \rangle = \langle Y(x),\ Y_o \rangle + \mu_\ell f^\ell(x) & ,\ \forall\ x \in S \\ \langle Y^*,\ Y^\ell \rangle = b^\ell & (\ell = 1,2,\ldots k) \end{cases}
$$

(les Y^ℓ sont les vecteurs vérifiant $\langle Y(x),\ Y^\ell \rangle = f^\ell(x), \forall x \in S$)

Si $Y^* \in \overline{H}_M$, on a $Y^* = \int \lambda(dx)\ Y(x)$, avec une mesure λ vérifiant le système :

$$
(\mathrm{III}')\quad \begin{cases} \int \lambda(dy)\ C(x,y) = \langle Y(x),\ Y_o \rangle + \mu_\ell f^\ell(x) & ,\ \forall\ x \in S \\ \int \lambda(dx)\ f^\ell(x) = b^\ell \end{cases}
$$

et l'estimateur universel $Z^*_a \in \overline{H}_{a_\ell f^\ell}$ s'écrit pour tout a :

$$
Z^*_a = \int \lambda(dx)\ Z(x)
$$

La seconde relation (III') montre que l'on a pour tout a :

$$
Z^*_a = \int \lambda(dx)\ Y(x) + a_\ell b^\ell
$$

Dans le <u>cas fini</u>, on a toujours $Y^* \in \overline{H}_M$, et la mesure λ, c'est-à-dire les poids λ^α attribués à chaque x_α $(\alpha = 1,2,\ldots m)$ est l'unique solution du système:

$$
(\mathrm{III}'')\qquad \begin{cases} \lambda^\beta\ \sigma_{\alpha\beta} = C_\alpha + \mu_\ell f^\ell_\alpha \\ \lambda^\alpha\ f^\ell_\alpha = b^\ell \end{cases}
$$

(avec $C_\alpha = \langle Y(x_\alpha),\ Y_o \rangle$, covariance de la variable à estimer avec $Y(x_\alpha)$)

Cas particulier : Krigeage ponctuel - Si la variable Z_a à estimer est $Z(z)$, valeur en $z \notin S$ de la fonction aléatoire $Z(x)$, on a : $Z(z) = Y(z) + a_\ell f^\ell(z)$, et :

$$\langle Y(x), Y_o \rangle = C(x,z)$$

$$b^\ell = f^\ell(z)$$

Le système (III'), par exemple, est alors remplacé par :

$$\text{(III''')} \quad \begin{cases} \int \lambda(dy)\ C(x,y) = C(x,z) + \mu_\ell f^\ell(x), & \forall x \in S \\ \int \lambda(dx)\ f^\ell(x) = f^\ell(z) \end{cases}$$

Relation avec l'estimation de la dérive

En l'absence de dérive, le meilleur estimateur de Y_o est la projection de Y_o dans $\bar{\bar{H}}$. C'est l'élément unique $Y_K \in \bar{\bar{H}}$ vérifiant les équations du krigeage en l'absence de dérive :

$$\text{(IV)} \qquad \langle Y(x), Y_K \rangle = \langle Y(x), Y_o \rangle \quad , \forall x \in S$$

Posons alors

$$Y^* = Y_K + Y_D$$

et substituons dans (III) : L'élément Y_D représente la correction que l'on doit effectuer en présence d'une dérive (dont les coefficients a_ℓ sont inconnus) pour obtenir un estimateur universel. Il doit vérifier :

$$\text{(IV')} \quad \begin{cases} \langle Y(x), Y_D \rangle = \mu_\ell f^\ell(x) & , \forall x \in S \\ \langle Y_D, Y^\ell \rangle = b^\ell - \langle Y_K, Y^\ell \rangle = b^\ell - \langle Y_o, Y^\ell \rangle \end{cases}$$

La première relation (IV') montre que Y_D est de la forme :

$$Y_D = \mu_\ell Y^\ell$$

les Y^ℓ désignant toujours les éléments de $\bar{H}$ vérifiant

$$< Y(x),\ Y^\ell > = f^\ell(x) \qquad , \forall x \in S$$

Substituons dans les secondes relations (IV'). Il vient :

$$\mu_i < Y^i,\ Y^\ell > = b^\ell - < Y_K,\ Y^\ell >$$

d'où:

$$\mu_\ell = \mu_{i\ell} \left[b^i - < Y_K,\ Y^i > \right]$$

$\mu_{i\ell}$, matrice inverse des $< Y^i,\ Y^\ell >$, est aussi la matrice des covariances du vecteur A_i, estimateur optimal de la dérive :

$$\mu_{i\ell} = E\ (A_i\ A_\ell) - a_i\ a_\ell$$

Ce sont les mêmes $\mu_{i\ell}$ que les paramètres de Lagrange du système (I). Ainsi, $Y^* = Y_K + Y_D$ se met sous la forme :

$$Y^* = Y_K - \mu_{i\ell} < Y_K,\ Y^i > Y^\ell + b^i\ \mu_{i\ell}\ Y^\ell$$

Prenons l'image inverse de Y^* dans $H_{a_\ell f^\ell}$. C'est l'estimateur universel de Z_a, soit :

$$Z^*_a = Y^* + a_\ell < Y^*,\ Y^\ell > = Y^* + a_\ell\, b^\ell$$

De leurs côtés, Y_K et Y^ℓ ont comme images inverses :

$$Z_K = Y_K + a_\ell < Y_K,\ Y^\ell >$$
$$Z^\ell = Y^\ell + a_i < Y^\ell,\ Y^i >$$

Or, l'estimateur universel A_i de la dérive a pour expression :

$$A_i = \mu_{i\ell}\, Z^{\ell} = \mu_{i\ell}\, Y^{\ell} + a_i$$

(puisque $\mu_{i\ell}$ et $< Y^i, Y^{\ell} >$ sont inverses) - On a donc

$$\text{(IV")} \qquad Z_a^* = Z_K - A_i < Y_K, Y^i > + A_i\, b^i$$

Or, si l'on connaissait les vraies valeurs des a_i, le meilleur estimateur de Z_a serait :

$$Y_K + a_{\ell} b^{\ell} = Z_K - a_i < Y_K, Y^i > + a_{\ell} b^{\ell}$$

On voit la signification de la relation (IV") : elle exprime que l'on a le droit de calculer Z_a^* comme si $A_{\ell} f^{\ell}(x)$ représentait la véritable dérive $a_{\ell} f^{\ell}(x)$, et non une simple estimation de celle-ci. Autrement dit, le <u>krigeage universel coïncide avec le krigeage ordinaire effectué sur les résidus $Y(x) - A f(x)$</u>.

<u>Exemple : Krigeage Ponctuel</u> - Soit à estimer $Z(z)$ en $z \notin S$ -

Le système (III) donne :

$$\begin{cases} < Y(x), Y^* > = C(x,z) + \mu_{\ell} f^{\ell}(x) \quad , \forall x \in S \\ < Y, Y^{\ell} > = b^{\ell} \end{cases}$$

En l'absence de dérive, l'estimateur du krigeage Y_K est la solution de :

$$< Y(x), Y_K > = C(x,z) \quad , \quad \forall x \in S$$

et l'estimateur universel (IV") s'écrit :

$$Z^*(z) = Z_K - < Y_K, Y^i > A_i + A_\ell f^\ell(z)$$

Plaçons-nous dans le cas où A_i et $Y_K \in \bar{H}_M$:

$$\left| \begin{array}{l} A_i = \int \lambda_i(dx) \, [Y(x) + a_\ell f^\ell(x)] = \int \lambda_i(dx) \, Y(x) + a_i \\ \int \lambda_i(dx) \, f^\ell(x) = \delta_i^\ell \\ Y_K = \int \lambda_K(dx) \, Y(x) \quad , \quad \lambda_K \in M \end{array} \right.$$

avec une mesure λ_K vérifiant :

$$\int \lambda_K(dy) \, C(x,y) = C(x,z) \quad , \quad \forall x \in S$$

On voit que la condition $Y_K \in \bar{H}_M$ exprime que le problème du <u>balayage</u> sur S est soluble pour le noyau potentiel $C(x,y)$. L'estimateur cherché se met sous la forme

$$\text{(IV"')} \qquad Z^*(z) = A_\ell f^\ell(z) + \int \lambda_K(dx) \, [Z(x) - A_i f^i(x)]$$

Cette relation exprime que <u>l'on peut prolonger l'estimation $A_\ell f^\ell(x)$ de la dérive à l'extérieur de S</u>, <u>et kriger les résidus (à l'aide du variogramme $\gamma(h)$) comme s'il n'y avait pas de dérive</u>.

Mais dans les applications, il sera en général inutile de déterminer séparément A_ℓ et Z_K. On cherchera directement un estimateur de la forme

$$Z^* (z) = \int \lambda_z(dy)\ Z(y)$$

avec une mesure λ_z vérifiant (III"')

Variance du Krigeage.

Evaluons d'abord la variance σ_K^2 du krigeage (en l'absence de dérive) de Y_o : d'après (IV), on a

$$\text{(V)} \qquad \sigma_K^2 = < Y_K - Y_o,\ Y_K - Y_o > \ = \|Y_o\|^2 - \|Y_K\|^2 = \|Y_o\|^2 - <Y_K,\ Y_o>$$

Par exemple dans le cas du krigeage ponctuel, et pour $Y_K \in \bar{H}_M$

$$\text{(V')}\ \sigma_K^2 = D^2(Z(z)) - D^2(Z_K) = C(z,z) - \int \lambda_K(dx)\ C(x,z)$$

Soit maintenant σ_u^2 la variance du krigeage universel (en présence d'une dérive). On a :

$$Y^* - Y_o = Y_K - Y_o + \mu_{i\ell}\ (b^i - \beta^i)\ Y^\ell$$

$$(\beta^i = < Y_K,\ Y^i > \ = \ <Y_o,\ Y^i >)$$

D'autre part $Y^\ell \in \bar{H}$ est orthogonal à $Y_K - Y_o$, d'où :

$$\sigma_u^2 = D^2\ (Y^* - Y_o) = \sigma_K^2 + \mu_{i\ell}\ \mu_{js}\ (b^i-\beta^i)(b^j-\beta^j) < Y^\ell,\ Y^s >$$

Enfin, $\mu_{js} = < A_j-a_j,\ A_s-a_s >$ étant l'inverse de $< Y^\ell,\ Y^s >$, on trouve :

$$(V'') \qquad \sigma_u^2 = \sigma_K^2 + \mu_{i\ell} (b^i - \beta^i)(b^\ell - \beta^\ell)$$

Dans le cas du krigeage ponctuel, on a

$$b^i = f^i(z)$$

$$\beta^i = \int \lambda_K(dx)\, f^i(x)$$

La relation (V") <u>montre que la variance du krigeage universel est la somme de la variance du krigeage ordinaire et de la variance de l'expression</u> $A_i\ (b^i - \beta^i)$ - (correction de dérive). Dans les applications, on résoudra en général directement le système (III). La variance σ_u^2 se calcule alors directement par :

$$\sigma_u^2 = \|Y^* - Y_o\|^2 = < \|Y^*\|^2 + \|Y_o\|^2 - 2 < Y^*, Y_o >$$

Mais, d'après (III), Y* vérifie :

$$< Y^*, Y > = < Y_o, Y > + \mu_\ell < Y^\ell, Y > \quad , \forall Y \in \bar{\bar{H}}$$

En particulier :

$$\|Y^*\|^2 = < Y^*, Y_o > + \mu_\ell < Y^\ell, Y^* > = < Y^*, Y_o > + \mu_\ell b^\ell$$

D'où :

$$(VI) \qquad \sigma_u^2 = \|Y_o\|^2 + \mu_\ell b^\ell - < Y^*, Y_o >$$

Dans le cas du krigeage ponctuel, et pour $Y^* = \int \lambda_z(dx)\, Y(x) \in \bar{\bar{H}}_M$, on trouve :

$$(VI') \qquad \sigma_u^2 = C(z,z) + \mu_\ell f^\ell(z) - \int \lambda_z(dx)\, C(x,z)$$

Cas Intrinsèque.

Dans le cas intrinsèque, les résultats précédents restent valables. On le voit en remplaçant $Y(x)$ par $Y(x) - Y(x_0)$, $C(x,y)$ par $-\gamma(x-y) + \gamma(x - x_0) + \gamma(y-y_0)$ et $f^{\ell}(x)$ par $f^{\ell}(x) - f^{\ell}(x_0)$ ($x_0 \in S$ fixé). En particulier, on a encore le droit de prolonger l'estimation $A_{\ell}f^{\ell}(x)$ de la dérive à l'extérieur de S et de kriger les résidus $Y(x) - A_{\ell}f^{\ell}(x)$ comme s'il n'y avait pas de dérive. Explicitons seulement les équations du krigeage ponctuel dans le cas $Y^* \in \bar{H}_M$: l'estimateur de $Z(z)$ est de la forme

$$Z^*(z) = \int \lambda_z(dx)\ Z(x)$$

avec une mesure λ_z vérifiant :

$$\text{(VII)} \quad \left| \begin{array}{rcl} \int \gamma(x-y)\ \lambda_z(dy) & = & \gamma(x-z) - \mu_{\ell}f^{\ell}(x) - C \\ \int \lambda_z(dx)\ f^{\ell}(x) & = & f^{\ell}(z) \\ \int \lambda_z(dx) & = & 1 \end{array} \right.$$

et la variance de cette estimation est

$$\sigma_u^2 = 2\int \lambda_z(dx)\ \gamma(x-z) - \iint \lambda(dx)\ \lambda(dy)\ \gamma(x-y)$$

Soit :

$$\text{(VII')} \quad \left| \ \sigma_u^2 = \int \lambda_z(dx)\ \gamma(x-z) + \mu_{\ell}f^{\ell}(z) + C \right.$$

Exemple : Reprenons l'exemple du segment $[0, L]$ et du variogramme $\gamma(h) = |h|$. Soit, par exemple, à kriger $Z(z)$ en $z > L$ - En l'absence de dérive, on a simplement $Z_K = Z(L)$ - En présence d'une dérive quadratique $a_1\ x + a_2\ x^2$, on peut utiliser (IV'), qui donne simplement :

$$Z^*(z) = A_1\ z + A_2\ z^2 + Z(L) - A_1\ L - A_2 L^2$$

avec les valeurs de A_1 et A_2 calculées au chapitre II.

Pour calculer la variance du krigeage, utilisons (VII'). On a ici :

$$\sigma_K^2 = 2\ (z-L)$$

$$b^1 = z,\ b^2 = z^2$$

$$\beta^1 = L,\ \beta^2 = L^2$$

$$\mu_{11} = \frac{8}{L}\ ,\ \mu_{12} = -\frac{6}{L^2}\ ,\ \mu_{22} = \frac{6}{L^3}$$

Il vient donc :

$$\sigma_u^2 = 2(z-L) + \frac{8}{L}\ (z-L)^2 - \frac{12}{L^2}\ (z-L)(z^2-L^2) + \frac{6}{L^3}\ (z^2-L^2)^2$$

Soit, en désignant par $h = z - L > 0$ la distance du point à kriger à l'extrémité droite du segment :

$$\sigma_u^2 = 2\ h + 8\ \frac{h^2}{L} + 12\ \frac{h^3}{L^2} + 6\ \frac{h^4}{L^3}$$

On voit que cette variance augmente très vite avec la distance h, et que l'extrapolation ne peut pas être poussée très loin.

Cas Fini - Dans le cas fini, désignons par Z_α les valeurs de $Z(x)$ aux points $x_\alpha \in S$ $(\alpha = 1,2,\ldots m)$; et posons $f^\ell(x_\alpha) = f^\ell_\alpha$ - L'estimateur de $Z(z)$ est :

$$Z^*(z) = \lambda^\alpha Z_\alpha$$

avec des coefficients λ^α solution du système :

$$\text{(VIII)} \qquad \begin{cases} \lambda^\beta \lambda_{\alpha\beta} = \gamma_\alpha(z) - \mu_\ell f^\ell_\alpha - C \\ \lambda^\alpha f^\ell_\alpha = f^\ell(z) \\ \sum_\alpha \lambda^\alpha = 1 \end{cases}$$

$(\gamma_{\alpha\beta} = \gamma(x_\alpha - x_\beta)$, $\gamma_\alpha(z) = \gamma(x_\alpha - z)$ - La variance de ce krigeage est alors

$$\text{(VIII')} \qquad \sigma_u^2 = \lambda^\alpha \gamma_\alpha(z) + \mu_\ell f^\ell(z) + C$$

Exemple : Soit toujours un variogramme linéaire, une dérive quadratique, et le segment $[0, 1]$. On donne $Y^i = Y(\frac{i}{n})$, $i = 0,1,\ldots n$ - En l'absence de dérive, l'estimateur $Z_K(i+x)$ $(0 \le x \le \frac{1}{n})$ de $Z(i+x)$ est $Z_K(i+x) = x\, Y_{i+1} + (1-x)\, Y_i$, avec la variance $2\, \varpi\, x(1-x)$ (pour $\gamma(h) = \varpi\, |h|$. Soient A_1 et A_2 les estimations optimales des coefficients a_1 et a_2 de la dérive $a_1 x + a_2 x^2$ - L'estimateur universel est $Z_u(i+x) = x\,(Y_{i+1} - A_1(i+1) - A_2(i+1)^2)$ $+ (1-x)\,[Y_i - A_1 i - A_2(i+1)^2] + A_1(i+x) + A_2(i+x)^2$ - On trouve :

$$\mathcal{Z}_u(i+x) = x\ Y_{i+1} + (1-x)\ Y_i - A_2\ x(1-x)$$

$$\sigma_u^2 \doteq 2\ \varpi\ x(1-x) + x^2(1-x)^2\ D^2\ (A_2)$$

VI - APPLICATION POSSIBLE A LA CARTOGRAPHIE AUTOMATIQUE.

On suppose que la cote $\mathcal{Z}(x,y)$ au point de coordonnées (x,y) est assimilable localement à une réalisation d'une fonction aléatoire intrinsèque de variogramme $\gamma(h) = \gamma(r) = \varpi\ r$, moyennant correction d'une dérive quadratique ($f^1(x,y) = x$, $f^2(x,y) = y$, $f^3(x,y) = x^2$, $f^4x,y) = x\ y$, $f^5(x,y) = y^2$)

Pour évaluer $\mathcal{Z}(M)$ en un point M, on part de m points M_α où les valeurs Z_α sont connues expérimentalement($M_\alpha = (x_\alpha, y_\alpha)$ - $M = (x,y)$. On résoud le système :

$$\left|\begin{array}{l}
\sum_\beta \lambda_\beta\ r_{\alpha\beta} = r_\alpha(M) - \mu_1 x_\alpha - \mu_2 y_\alpha - \mu_3 x_\alpha^2 - \mu_4 x_\alpha y_\alpha - \mu_5 y_\alpha^2 - C \\
\quad \sum_\alpha \lambda_\alpha\ x_\alpha = x \\
\quad \sum_\alpha \lambda_\alpha\ y_\alpha = y \\
\quad \sum_\alpha \lambda_\alpha\ x_\alpha^2 = x^2 \\
\quad \sum_\alpha \lambda_\alpha\ x_\alpha y_\alpha = x\ y \\
\quad \sum_\alpha \lambda_\alpha\ y_\alpha^2 = y^2 \\
\quad \sum_\alpha \lambda_\alpha = 1
\end{array}\right.$$

qui comporte m + 6 équations à m inconnues λ_α et 6 paramètres de Lagrange (les μ_i et C). L'estimateur cherché est :

$$Z^*(M) = \sum_\alpha \lambda_\alpha(M)\, Z_\alpha$$

et la variance de krigeage est

$$\sigma_u^2 = \varpi^* \left[\sum_\alpha \lambda_\alpha\, r_\alpha(M) + \mu_1 x + \mu_2 y + \mu_3 x^2 + \mu_4 xy + \mu_5 y^2 + C \right]$$

avec un estimateur ϖ^* de ϖ défini par les méthodes du paragraphe IV.

Vérification Expérimentale

Ce procédé d'interpolation par la méthode du krigeage universel n'est, évidemment, légitime que si l'on a pu vérifier la validité des hypothèses faites :

- dérive quadratique
- variogramme $\gamma(r) = \varpi\, r$ isotrope et linéaire (au moins en première approximation et dans le voisinage du point à kriger où se trouvent les points expérimentaux utilisés : dans ce même voisinage, $Z(x)$ corrigé de la dérive quadratique peut être assimilé à une fonction aléatoire intrinsèque).

La vérification suivante est facile à effectuer, si les points expérimentaux sont donnés sur des courbes à peu près rectilignes : On prendra sur ces courbes un certain nombre de segments de même longueur L et de même direction θ. Sur chacun de ces segments, on effectuera la correction de dérive quadratique comme dans l'exemple du chapitre III, et on construira le variogramme des résidus. On prendra la moyenne $\gamma^*(h)$ des variogrammes expérimentaux relatifs à une même direction θ, et on comparera $\gamma^*(h)$ avec son espérance :

$$E\left(\gamma^*(h)\right) = \varpi \left(h - 2\frac{h^2}{L} + 2\frac{h^3}{L^2} - \frac{h^4}{L^3}\right)$$

Il faudra vérifier que le coefficient ϖ ainsi ajusté ne varie pas trop lorsque l'on fait varier la direction θ. Enfin, on éxécutera cette opération pour diverses valeurs croissantes de L : la valeur de L_o à partir de laquelle l'accord cesse d'être acceptable donne le diamètre maximum du cercle à l'intérieur duquel les hypothèses faites sont admissibles. Les points expérimentaux M_α que l'on a le droit d'utiliser pour estimer $Z(M)$ sont ceux qui tombent à l'intérieur du cercle de centre M et de diamètre L_o

ANNEXE I

COMPARAISON AVEC LES METHODES USUELLES DE MOINDRE CARRE

Pour estimer les coefficients a_ℓ d'une dérive $a_\ell f^\ell(x)$, où les fonctions connues $f^\ell(x)$ sont, par exemple, des polynomes, on a souvent recours à une méthode d'ajustement par moindres carrés. Nous allons établir les formules générales de ce procédé, et en comparer l'efficacité avec celle des estimateurs universels. Nous illustrerons les résultats généraux que nous obtiendrons en explicitant les calculs dans le cas simple d'un processus défini sur la droite réelle et admettant (dérive exclue) un variogramme linéaire $\gamma(h) = |h|$.

Dans le cas le plus général, $Z(x)$ désignera une fonction aléatoire à accroissements (non nécessairement stationnaires) d'ordre 2, possédant une dérive a(x) :

$$\begin{cases} E\,[Z(y) - Z(x)] = a(y) - a(x) \\ \frac{1}{2}\,D^2\,[Z(y) - Z(x)] = \gamma(x;y) \end{cases}$$

On admettra souvent que la dérive est de la forme

$$a(x) = a_\ell f^\ell(x)$$

avec des fonctions $f^\ell(x)$ connues ($\ell = 0,1,2,\ldots k$) et des coefficients a_ℓ à estimer. La fonction $f^0(x)$ sera prise égale à la constante 1 : $f^o(x) = 1$, et l'estimation de a_o équivaudra à une moyenne arithmétique.

1/ <u>Théorie des Moindres Carrés</u>

Soit $Z_\omega(x)$, $x \in S$ une réalisation sur S de la fonction aléatoire $Z(x)$. On se donne k fonctions $f^\ell(x)$ ($\ell = 1,2,\dots k$) et $f^o(x) = 1$. On détermine k + 1 nombres b_o, $b_1, \dots b_k$ de manière à minimiser l'intégrale:

$$I(\omega) = \int_S [\ Z_\omega(x) - b_\ell f^\ell(x)]^2 \ dx$$

En annulant les dérivées en b_ℓ, on obtient le système :

$$\int_S f^\ell(x)\ [Z_\omega(x) - b_j f^j(x)]\ dx = 0$$

Posons

$$T^{\ell j} = \frac{1}{S} \int_S f^\ell(x)\ f^j(x)\ dx$$

$$t^\ell = \frac{1}{S} \int_S f^\ell(x)\ dx \qquad (= T^{o\ell})$$

$$T_o^{\ell j} = T^{\ell j} - t^\ell t^j = \frac{1}{S} \int_S (f^\ell(x) - t^\ell)\ (f^j(x) - t^j)\ dx$$

Les b_ℓ sont solution du système :

$$b_j\ T^{\ell j} = \frac{1}{S} \int_S f^\ell(x)\ Z(x)\ dx$$

Pour $\ell = 0$, $f^o(x) = 1$ et on trouve

$$b_j\ t^j = \frac{1}{S}\int_S z(x) = \bar{z}\ \text{(sommation de } j = 0 \text{ à } j = k)$$

A partir de maintenant, nous ferons varier ℓ de 1 à k, et nous réserverons un traitement spécial à b_o. On obtient :

$$b_o = \frac{1}{S}\int_S Z(x) - b_j\ t^j \qquad (j = 1,2,\ldots k)$$

et le système précédent s'écrit :

$$b_o\ t^\ell + b_j\ T^{\ell j} = \frac{1}{S}\int f^\ell(x)\ z(x)\ dx$$

c'est-à-dire :

$$b_j\ T^{\ell j} = \frac{1}{S}\int_S f^\ell(x)\ z(x)\ dx - b_o\ t^\ell$$

Remplaçons b_o par l'expression obtenue plus haut :

$$b_j\ T^{\ell j} = \frac{1}{S}\int_S (f^\ell(x) - t^\ell)\ z(x) + b_j\ t^j\ t^\ell$$

d'où finalement :

$$(1) \qquad b_j\ (T^{\ell j} - t^\ell t^j) = \frac{1}{S}\int_S \left(f^\ell(x) - t^\ell\right)\ z_\omega^*(x)\ dx$$

Nous admettons que la matrice $T_o^{\ell j} = T^{\ell j} - t^\ell t^j$ admet une inverse, soit $S_{\ell j}$, et nous obtenons alors :

(1')
$$b_j = \frac{S_{\ell j}}{S} \int_S [f^{\ell}(x) - t^{\ell}] \, z_{\omega}(x) \, dx$$

En tant qu'intégrale stochastique, cette expression est bien définie, puisque la mesure de densité $f^{\ell}(x) - t^{\ell}$ sur S a une intégrale égale à 0. L'estimateur de la dérive auquel conduit la méthode des moindres carrés est donc le vecteur aléatoire B défini par

(2)
$$B_j = S_{\ell j} \frac{1}{S} \int_S [f^{\ell}(x) - t^{\ell}] \, z(x) \, dx$$

Calculons son <u>espérance</u> : on trouve

$$E(B_j) = S_{\ell j} \int_S [f^{\ell}(x) - t^{\ell}] \, a(x) \, dx$$

Si la dérive est $a(x) = a_o + b_j f^j(x)$, on obtient :

$$E(B_j) = S_{\ell j} (T^{\ell s} - t^{\ell} t^s) \, a_s = a_j$$

puisque $S_{\ell j}$ et $(T^{\ell s} - t^{\ell} t^s)$ sont inverses : B_j est un <u>estimateur sans biais</u> (lorsque la dérive réelle est effectivement de la forme choisie). B_j est donc un estimateur universel, au sens des paragraphes précédents, mais il diffère de l'estimateur universel optimal A_j.

Calculons la matrice des covariances des B_i - On trouve :

$$\beta_{ij} = E(B_i B_j) - b_i b_j =$$

$$= S_{\ell i} \, S_{sj} \frac{1}{S^2} \, E\left[\int_S \int_S [f^{\ell}(x) - t^{\ell}][f^s(y) - t^s] \, z(x) \, z(y) dx \, dy\right]$$

D'où :

$$\beta_{ij} = - S_{\ell i}\, S_{sj}\, \frac{1}{S^2} \int_S \int_S [f^\ell(x) - t^\ell][f^s(x) - t^s]\, \gamma(x,y)\,dx\,dy$$

Exemple : $\gamma(h) = |h|$ sur le segment $[0, L]$ et dérive quadratique $a_1 x + a_2 x^2$. On a

$$t_1 = \frac{1}{L} \int_0^L x\,dx = \frac{L}{2} \quad , \quad t_2 = \frac{1}{L} \int_0^L x^2\,dx = \frac{L^2}{3}$$

$$T^{11} = \frac{1}{L} \int_0^L x^2 dx = \frac{L^2}{3} \quad , \quad T^{12} = T^{21} = \frac{1}{L} \int_0^L x^3 dx = \frac{L^3}{4}$$

$$T^{22} = \frac{1}{L} \int_0^L x^4 dx = \frac{L^4}{5}$$

Puis :

$$T_o^{11} = t^{11} - t^1 t^1 = \frac{L^2}{12} \quad , \quad T_o^{12} = T_o^{21} = + \frac{L^3}{12} \quad , \quad T_o^{22} = \frac{4}{45} L^4$$

La matrice des S_{ij} est :

$$(S_{ij}) = \begin{pmatrix} \frac{192}{L^2} & - \frac{180}{L^3} \\ - \frac{180}{L^3} & \frac{180}{L^4} \end{pmatrix}$$

On en déduit les estimateurs :

$$\begin{cases} B_1 = \frac{192}{L^3} \int_0^L x\, z(x)dx - \frac{180}{L^4} \int_0^L x^2\, z(x)dx - \frac{36}{L}\, \bar{z} \\ B_2 = \frac{180}{L^4} \int_0^L x\, z(x)dx + \frac{180}{L^5} \int_0^L x^2\, z(x)dx + \frac{30}{L^2}\, \bar{z} \end{cases}$$

En ce qui concerne la matrice des covariances, on trouve

$$\begin{cases} \beta_{11} = \dfrac{384}{35} \quad \dfrac{1}{L} \\[2ex] \beta_{12} = \beta_{21} = - \dfrac{60}{7} \quad \dfrac{1}{L^2} \\[2ex] \beta_{22} = \dfrac{60}{7} \quad \dfrac{1}{L^3} \end{cases}$$

(Remarque : la relation $\beta_{12} + L\,\beta_{22} = 0$ est vérifiée par tous les estimateurs sans biais. Celà se voit par symétrie)

En particulier, $B_1 L + B_2 L^2$, dérive entre les points L et 0, a pour variance $\frac{12}{5}$ L (au lieu de 2 L dans le cas de l'estimateur optimal A_i).

Condition pour que l'estimateur des moindres carrés soit optimal

L'estimateur des moindres carrés est caractérisé par les deux propriétés suivantes :

1 - Il est sans biais $\quad E(B_i) = a_i \quad , \forall a \in \mathbb{R}^k$

2 - Les B_i sont combinaisons linéaires des $\int_S [f^\ell(x) - t^\ell] Z(x)dx$ (vérification immédiate à partir de la relation (2)).

Cet estimateur est optimal si, et seulement si, on peut trouver des combinaisons linéaires des mesures $[f^\ell(x) - t^\ell]$, soit

$$\lambda_i(dx) = \frac{S_{ij}}{S} [f^j(x) - t^j] \, dx$$

vérifiant le système (II'). La troisième relation (II') est vérifiée, par définition des t^j. La deuxième exprime que les $\frac{S_{ij}}{S}$ sont les composantes de la matrice inverse de T^{ij}. La première se met sous la forme :

$$(3) \qquad - \frac{S_{ij}}{S} \int_S [f^j(y) - t^j] \, \gamma(x-y) \, dy = \mu_{i\ell} \, f^\ell(x) + C_i$$

Pour éliminer les constantes C_i, intégrons en x dans S :

$$- \frac{S_{ij}}{S^2} \int_S dx \int_S [f^j(y) - t^j] \, \gamma(x-y) \, dy = \mu_{i\ell} \, t^\ell + C_i$$

et retranchons, en posant

$$\frac{1}{S} \int_S \gamma(x-y) \, dx = \bar{\gamma}(y)$$

Il vient :

$$(3') \qquad - \frac{S_{ij}}{S} \int_S \left(f^j(y) - t^j\right) \left(\gamma(x-y) - \bar{\gamma}(y)\right) dy = \mu_{i\ell} (f^\ell(x) - t^\ell)$$

Comme $\int_S \left(f^j(y) - t^j\right) dy = 0$, introduisons le noyau symétrique :

$$(4) \qquad C(x,y) = - \gamma(x-y) + \bar{\gamma}(x) + \bar{\gamma}(y) - \frac{1}{S} \int \bar{\gamma}(x) \, dx$$

$\left(C(x,y)\right.$ est la covariance de la fonction aléatoire $\left.Y(x) - \bar{Y}\right)$.

Il vient :

$$\int_S \left(f^j(y) - t^j\right) C(x,y) \, dy = T_o^{\,ji} \, \mu_{i\ell} \, [f^\ell(x) - t^\ell]$$

Si ϖ_j est un vecteur propre de $T_o^{ji}\,\mu_{i\ell}$, de valeur propre ϖ, on voit que la fonction

$$\varpi_j\,[f^j(y) - t^j] = \varphi(y)$$

est une fonction propre pour le noyau $C(x,y)$:

$$(5) \qquad \int_S \varphi(y)\;C(x,y)\;dy = \varpi\;\varphi(x)$$

On peut donc trouver k combinaisons linéaires (indépendantes) des $f^j(x) - t^j$ correspondant à k fonctions propres distinctes $\varphi^i(x)$. Par suite, les $f^\ell(x) - t^\ell$ sont elles-mêmes des combinaisons linéaires de ces fonctions propres :

$$(6) \qquad f^\ell(x) - t^\ell = F^\ell_i\;\varphi^i(x)$$

Inversement, soient $\varphi^i(x)$ des fonctions propres distinctes de $C(x,y)$ correspondant à des valeurs propres non nulles. Il suffit d'intégrer (5) en x, compte tenu de l'expression (4) du noyau, pour vérifier que toute fonction propre (de valeur propre ϖ différente de 0) a une intégrale nulle. Par suite, toute combinaison linéaire de type (6) vérifiera (3'), donc aussi (3), et, pour de telles combinaisons linéaires, l'estimateur de moindre carré coïncidera avec l'estimateur optimal.

Ainsi, pour que l'estimateur de moindre carré coïncide avec l'estimateur optimal, il faut et il suffit que les fonctions $f^\ell(x) - t^\ell$ $\left(t^\ell = \frac{1}{S}\int_S f^\ell(x)\,dx\right)$ soient des

<u>des combinaisons linéaires de k fonctions propres φ^{ℓ} du noyau $C(x,y)$ correspondant à des valeurs propres non nulles</u>. Ce noyau n'est autre que la covariance de la fonction aléatoire $Z(x) - \frac{1}{S}\int_S Z(x)\,dx = Z(x) - \bar{Z}$.

<u>Exemple</u> - Pour $\gamma(h) = |h|$ sur le segment $[0, L]$, on a

$$C(x,y) = \frac{2}{3} L + \frac{x^2+y^2}{L} - 2 \operatorname{Sup}(x,y)$$

En dérivant deux fois la relation :

$$g(x) = \int C(x,y)\,\varphi(y)dy = \frac{2}{3} L \int_0^L \varphi(y)dy + \frac{x^2}{L}\int_0^L \varphi(y)dy +$$

$$+ \frac{1}{L}\int_0^L y^2\varphi(y)dy - 2x\int_0^x \varphi(y)dy - 2\int_x^L y\,\varphi(y)dy$$

on obtient :

$$g''(x) = \frac{2}{L}\int_0^L \varphi(y)dy + 2\,\varphi(x)$$

Les fonctions propres doivent de plus vérifier $\int_0^L \varphi(y)dy = 0$.

Elles sont donc de la forme $Sh\, a\,(x - \frac{L}{2})$, $\sin a\,(x - \frac{L}{2})$ (a quelconque) ou $\cos 2k\pi\frac{x}{L}$. Mais tout $g(x)$ de la forme ci-dessus vérifie $g'(o) = 0$. Les fonctions propres sont donc nécessairement de la forme $\cos k\pi\frac{x}{L}$, (k pair ou impair), et on vérifie directement que les $\cos k\pi\frac{x}{L}$ conviennent. Ainsi, l'estimateur de moindres carrés est optimal pour une dérive de la forme :

$$\sum_{\ell=1}^{n} a_k \cdot \cos k \pi \frac{x}{L}$$

Ainsi se trouve légitimé l'usage des moindres carrés et de l'interpolation trigonométrique dans le cas particulier où :

- les données occupent un segment de droite
- le variogramme, dérive exclue, est linéaire

Ce résultat ne subsiste pas si les données sont dans un espace à deux ou trois dimensions, ou même dans l'espace à une dimension, si le variogramme n'est pas linéaire.

ANNEXE II

PROBLEME D'INFERENCE STATISTIQUE POUR LES FONCTIONS ALEATOIRES

On a souvent l'habitude d'admettre, sans examen approfondi, que les variables régionalisées que l'on étudie peuvent être considérées comme des réalisations de fonctions aléatoires stationnaires d'ordre 2. Il se trouve que les procédés d'inférence statistique que l'on utilise ont systématiquement pour effet, par suite des biais qu'ils entraînent, de rendre cette hypothèse apparemment plausible, même lorsqu'elle est tout-à-fait fausse.

Donnons un exemple :

Mouvement brownien sans dérive. Soit $Y(x)$ une réalisation, sur le segment $[0, L]$, d'un mouvement brownien sans dérive, ou plus généralement d'une fonction aléatoire intrinsèque vérifiant :

$$E\,[Y(x+h) - Y(x)] = 0 \quad , \quad D^2\,[Y(x+h) - Y(x)] = h$$

Dans la pratique usuelle, on prend

$$\bar{\bar{Y}} = \frac{1}{L}\int_0^L Y(x)\,dx,$$

on calcule les valeurs

$$C^*(x,y) = \left(Y(x) - \bar{\bar{Y}}\right)\left(Y(y) - \bar{\bar{Y}}\right) \quad (x,y \in [0, L])$$

et on en déduit la "covariance" expérimentale :

$$C^*(h) = \frac{1}{L-h} \int_0^{L-h} C^*(x+h,x)\, dx$$

Calculons $E\left(C^*(h)\right)$. On a vu ci-dessus :

$$E\left(C^*(x,y)\right) = \frac{2}{3} L + \frac{x^2+y^2}{L} - 2\ \mathrm{Sup}\ (x,y)$$

Pour h positif, donc,

$$E\left(C^*(x+h,y)\right) = \frac{2}{3} L + \frac{x^2+(x+h)^2}{L} - 2x - 2h$$

et, en intégrant en x, on trouve

$$E\left[C^*(h)\right] = \frac{1}{3} L - \frac{4}{3} h + \frac{2}{3} \frac{h^2}{L}$$

C'est une parabole, de pente $-\frac{4}{3}$ à l'origine.

Ainsi, bien qu'il n'existe pas, en réalité, de covariance $C(h)$ stationnaire, mais seulement un variogramme linéaire, les biais qu'introduit cette méthode d'estimation ont pour effet apparent de confirmer notre hypothèse sur l'existence de $C(h)$. De plus, la structure du phénomène est très déformée : non seulement la droite est remplacée par une parabole, mais même la pente à l'origine se trouve affectée ($\frac{4}{3}$ au lieu de 1)

Plus généralement, soit $Y(x)$ une fonction aléatoire intrinsèque sans dérive, de variogramme $\gamma(h)$, dans l'espace à n

dimensions, et soit S l'ensemble des points où l'on connaît la réalisation. Posant

$$\begin{cases} \bar{Y} = \frac{1}{S} \int_S Y(x)\, dx \\ C^*(x,y) = [Y(x) - \bar{Y}][Y(y) - \bar{Y}] \end{cases}$$

on trouve :

$$C(x,y) = E\,[C^*(x,y)] = -\gamma(x,y) + \bar{\gamma}(x) + \bar{\gamma}(y) - \frac{1}{S} \int_S \bar{\gamma}(x)\, dx$$

et la "covariance" expérimentale C*(h) admet l'espérance :

$$E\,[C^*(h)] = \frac{1}{K(h)} \int k(x)\; k(x+h)\; C(x+h,x)\, dx$$

($k(x)$, indicatrice de S, $K = k * \check{k}$). Ici encore le procédé utilisé fournit une confirmation apparente de l'hypothèse (fausse) de départ.

Sur un segment $[0, L]$, et en introduisant les fonctions

$$\chi(h) = \frac{1}{h} \int_0^h \gamma(x)dx\,, \qquad F(h) = \frac{2}{h^2} \int_0^h x\chi(x)\, dx$$

on trouve :

$$E\,[C^*(h)] = -\gamma(h) - F(L) + \frac{1}{L(L-h)} \left[\int_0^L dz \int_0^{L-h} \gamma(x-z)\, dx + \right.$$

$$\left. + \int_0^L \int_h^L \gamma(x-z)\, dx \right]$$

Supposons que $\gamma(h)$ admette une pente à l'origine (dérivée à droite) $\gamma_+'(o)$ - On trouve pour E $[C^*(h)]$ la pente :

$$\frac{d^+}{dh} \; E\,[C^*(h)]\Big|_{h=o} = -\,\gamma_+'(o) + \frac{2}{L}\,F\,(L) - \frac{2}{L}\,\chi\,(L)$$

Ainsi, pour une fonction aléatoire dérivable m.q.($\gamma(h)$ parabolique en $h = 0$), on trouve un E $[C^*(h)]$ qui présente une pente oblique à l'origine et fait conclure à la non-dérivabilité m.q. : la structure du phénomène est profondément défigurée.

La soi-disant indépendance des résidus

On justifie souvent la méthode d'interpolation par les moindres carrés en montrant que les résidus (après correction de la dérive estimée) présentent des corrélations expérimentales non significativement différentes de 0. Cette absence apparente de corrélation ne doit pas nous inciter à des conclusions trop hâtives. Et elle résultera souvent d'un effet de biais analogue à celui que nous avons déjà rencontré.

Nous n'écrirons pas les formules générales, assez lourdes, qui ne présentent pas d'intérêt par elles-mêmes une fois que l'on a compris le mécanisme générateur de ces biais. Nous raisonnerons sur notre exemple habituel : variogramme $\gamma(h)$, segment $[0, L]$, et dérive quadratique.

Nous considérons cette fois la variable :

$$Z(x) = Y(x) - \overline{Y} - B_1 \left(x - \frac{L}{2}\right) - B_2 \left(x^2 - \frac{L^2}{3}\right)$$

et nous posons

$$C^*(x,y) = Z(x)\, Z(y)$$

Calculons $E\,[C^*(x,y)]$:

$$E\,[C^*(x,y)] = T_1 - 2\,T_2 + T_3$$

Le premier terme a déjà été calculé :

$$T_1 = \frac{2}{3} L + \frac{x^2+y^2}{L} \quad - 2 \,\mathrm{Sup}\,(x,y)$$

Pour T_3, nous trouvons

$$T_3 = \beta_{11}\left(x - \frac{L}{2}\right)\left(y - \frac{L}{2}\right) + \beta_{12}\left(x - \frac{L}{2}\right)\left(y^2 - \frac{L^2}{3}\right) + \beta_{12}\left(y - \frac{L}{2}\right)\left(x^2 - \frac{L^2}{3}\right)$$

$$+ \beta_{22}\left(x^2 - \frac{L^2}{3}\right)\left(y^2 - \frac{L^2}{3}\right)$$

(les β_{ij} sont les covariances des B_i, B_j. On les a calculées plus haut). Enfin, le deuxième terme est de la forme :

$$2\,T_2 = \left(x - \frac{L}{2}\right) E\,[B_1(Y(y) - \overline{Y})] + \left(y - \frac{L}{2}\right) E\,[B_1(Y(x) - \overline{Y})]$$

$$+ \left(x^2 - \frac{L^2}{3}\right) E\,[B_2(Y(y) - \overline{Y})] + \left(y^2 - \frac{L^2}{3}\right) E\,[B_2(Y(x) - \overline{Y})]$$

Posons pour abréger

$$H(x,y) = \frac{2}{3} L + \frac{x^2+y^2}{L} - 2 \operatorname{Sup}(x,y) = T_1$$

En se reportant aux expressions de B_1 et B_2, on trouve :

$$E\left(B_1 \left[Y(x) - \bar{Y}\right]\right) = \frac{192}{L^3} \int_0^L z\, H(x,z)dz - \frac{180}{L^4} \int_0^L z^2\, H(x,z)dz$$

$$E\left(B_2 \left[Y(x) - \bar{Y}\right]\right) = - \frac{180}{L^4} \int_0^L z\, H(x,z) + \frac{180}{L^5} \int_0^L z^2\, H(x,z)dz$$

On trouve

$$\begin{cases} \displaystyle\int_0^L z\, H(x,z)dz = - \frac{1}{12} L^3 + \frac{1}{2} L\, x^2 - \frac{1}{3} x^3 \\ \displaystyle\int_0^L z^2\, H(x,z)dz = - \frac{7}{90} L^4 + \frac{1}{3} x^2\, L^2 - \frac{1}{6} x^4 \end{cases}$$

d'où

$$E\left[B_1(Y(x) - \bar{Y})\right] = -2 + 36 \frac{x^2}{L^2} - 64 \frac{x^3}{L^3} + 30 \frac{x^4}{L^4}$$

$$E\left[B_2(Y(x) - \bar{Y})\right] = \left(+1 - 30 \frac{x^2}{L^2} + 60 \frac{x^3}{L^3} - 30 \frac{x^4}{L^4}\right) \frac{1}{L}$$

On évalue ensuite les intégrales

$$R_{n,m}(x) = \frac{1}{1-x} \int_0^{1-x} \left(y^n - \frac{1}{n+1}\right) \left(y^m - \frac{1}{m+1}\right) dy \qquad \left(x = \frac{h}{L}\right)$$

(pour $n = 1,2$ et $m = 1,2,3,4$), ce qui donne :

$$R_{11}(x) = \frac{1}{12} - \frac{1}{6}x - \frac{1}{6}x^2$$

$$R_{12} + R_{21} = \frac{1}{6} - \frac{1}{3}x - \frac{1}{3}x^2$$

$$R_{13} + R_{31} = \frac{3}{20} - \frac{7}{20}x - \frac{1}{10}x^2 - \frac{1}{10}x^3 - \frac{1}{10}x^4$$

$$R_{22} = \frac{4}{45} - \frac{17}{90}x - \frac{17}{90}x^2 + \frac{1}{30}x^3 + \frac{1}{30}x^4$$

$$R_{41} + R_{14} = \frac{2}{15} - \frac{11}{30}x + \frac{2}{15}x^2 - \frac{1}{5}x^3 - \frac{1}{5}x^4$$

$$R_{23} + R_{32} = \frac{1}{6} - \frac{5}{12}x - \frac{1}{6}x^2$$

$$R_{24} + R_{14} = \frac{16}{105} - \frac{47}{105}x + \frac{3}{35}x^2 - \frac{4}{35}x^3 - \frac{4}{35}x^4 + \frac{2}{105}x^5 + \frac{2}{105}x^6$$

Il vient alors :

$$\frac{1}{L-h}\int_0^{L-h} 2\,T_2(h)dh = 36(R_{12}+R_{21}) - 64(R_{13}+R_{31}) + 30(R_{14}+R_{41}) - 60\,R_{22} + 60(R_{23}+R_{32}) - 30(R_{24}+R_{42}) =$$

$$= \left(\frac{52}{105} - \frac{88}{105}x - \frac{298}{105}x^2 + \frac{64}{35}x^3 + \frac{64}{35}x^4 - \frac{4}{7}x^5 - \frac{4}{7}x^6 \right) L$$

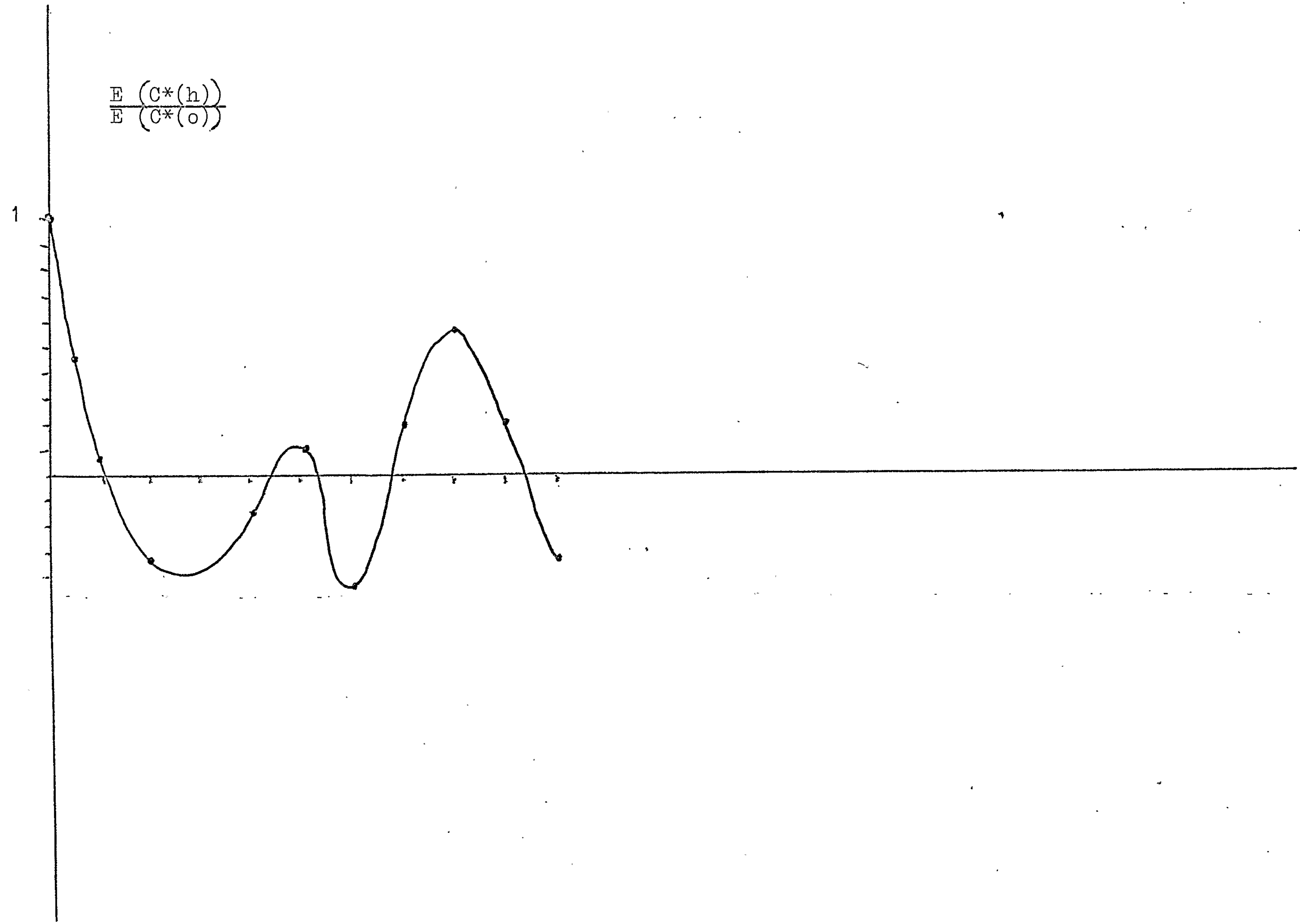
De même :

$$\frac{1}{L-h}\int_0^{L-h} T_3(h)dh = \beta_{11}\,R_{11} + \beta_{12}\,(R_{12}+R_{21}) + \beta_{22}\,R_{22} =$$

$$= \left(\frac{26}{105} - \frac{62}{105}x - \frac{62}{105}x^2 + \frac{2}{7}x^3 + \frac{2}{7}x^4 \right) L$$

et finalement : (voir figure ci-jointe)

$$E\,[C*(h)] = L\left[\frac{3}{35} - \frac{38}{35}\frac{h}{L} + \frac{102}{35}\frac{h^2}{L^2} - \frac{54}{35}\frac{h^3}{L^3} - \frac{54}{35}\frac{h^4}{L^4} + \frac{4}{7}\frac{h^5}{L^5} + \frac{4}{7}\frac{h^6}{L^6}\right]$$



$\frac{E\left(C^*(h)\right)}{E\left(C^*(o)\right)}$
1