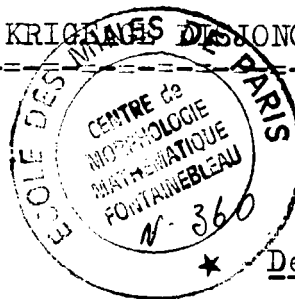


Fontainebleau

N-360

LE KRIGER DE JONCTIF



G. MATHERON



Décembre 1973

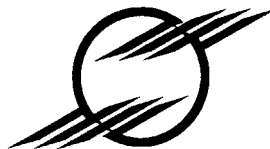
LE KRIGEAGE DISJONCTIF

G. MATHERON

Rapport N-360

Décembre 1973

CENTRE DE GÉOSTATISTIQUE
35, RUE SAINT-HONORÉ, 77305 FONTAINEBLEAU (France)



ECOLE DES MINES
DE PARIS

LE KRIGEAGE DISJONCTIF
=====

Table des Matières

1 - <u>INTRODUCTION.</u>	1
2 - <u>LES EQUATIONS DU KRIGEAGE DISJONCTIF.</u>	3
3 - <u>REPRESENTATION D'UNE LOI A DEUX VARIABLES.</u>	4
4 - <u>LES PROJECTEURS E_1 et E_2.</u>	7
5 - <u>LES VALEURS PROPRES DE $E_1 E_2$.</u>	8
6 - <u>EXEMPLE 1 : LE CAS GAUSSIEN.</u>	10
Calcul d'une espérance conditionnelle.	12
Calcul d'un krigeage disjonctif.	13
7 - <u>EXEMPLE 2 : LES LOIS GAMMA.</u>	15
Les polynomes de Laguerre	16
Krigeage disjonctif	20
8 - <u>RECHERCHE DE MODELES</u>	22
Modèles à facteurs polynomiaux.	24
Lois de Poisson	27
Lois binomiales négatives.	30
9 - <u>LE KRIGEAGE ET LA REPRESENTATION DES LOIS</u>	32
10 - <u>REPRESENTATION DES F.A.</u>	34

LE KRIGEAGE DISJONCTIF

1 - INTRODUCTION.

Entre les méthodes de Benzecri et le krigeage, il y a une sorte de parenté générale et formelle qui résulte de ce que ces deux théories manipulent des formes quadratiques, ou, si l'on préfère, travaillent dans des espaces de Hilbert. De fait, J.P. Orfeuil a montré que l'emploi d'une métrique définie par le variogramme permettait de faire rentrer le krigeage dans le cadre général du formalisme de Benzécri. Ce résultat extrêmement intéressant incite à aller plus loin, et, dans ce qui suit, je me propose de définir une version généralisée du krigeage, inspirée de la technique dite du "codage disjonctif complet", qui pourrait rendre de grands services pour étudier les phénomènes appelés (dans notre jargon) effets proportionnels.

L'idée de départ est la suivante. Pour estimer Z_0 à $Z_1, \dots, Z_n$ donnés, en toute généralité, on devrait utiliser l'espérance conditionnelle $Z_0^* = E(Z_0 | Z_1, \dots, Z_n)$, autrement dit chercher la meilleure approximation de Z_0 par une fonction $f(Z_1, \dots, Z_n)$ mesurable. En pratique, cependant, on ne connaît pas la loi simultanée des $n+1$ variables $(Z_0, Z_1, \dots, Z_n)$, et on ne peut pas calculer cette espérance conditionnelle. D'où la technique du krigeage, qui se contente de chercher la meilleure approximation de Z_0 par une combinaison linéaire $\lambda^\alpha Z_\alpha$, et dont la mise en oeuvre nécessite simplement la connaissance des covariances. Comme l'espace des combinaisons linéaires est beaucoup plus petit que l'espace des fonctions mesurables, le krigeage (sauf dans le cas gaussien)

donne une approximation moins bonne que l'espérance conditionnelle, et, vraisemblablement, d'autant moins bonne que l'effet proportionnel est plus marqué. En prenant un espace plus grand que celui des combinaisons linéaires, mais pas trop grand pourtant, pour que la mise en oeuvre de la technique correspondante ne nécessite pas la connaissance de données inaccessibles expérimentalement, on peut espérer aboutir à des méthodes pratiquement utilisables et nettement meilleures que le krigeage lorsque l'effet proportionnel est notable.

C'est ici qu'intervient le "codage disjonctif complet". Soit B_i , $i \in I$ une partition de la droite (en intervalles, ou classes, par exemple). A toute variable Y nous pouvons associer la famille Y_i , $i \in I$ définie par :

$$Y_i = 1_{B_i}(Y) = \begin{cases} 1 & \text{si } Y \in B_i \\ 0 & \text{si } Y \notin B_i \end{cases}$$

La famille des combinaisons linéaires $\sum \lambda_i Y_i$ est identique à celle des fonctions $f(Y)$, où f est mesurable pour $\sigma(B_i, i \in I)$. Lorsque la partition B_i devient de plus en plus fine, on obtient à la limite la famille des fonctions Borel-mesurables de la variable Y . Revenons à l'estimation de Z_0 . Au lieu de kriger Z_0 à l'aide des Z_α , nous pouvons le kriger sur la famille beaucoup plus riche des

$$Z_{i\alpha} = 1_{B_i}(Z_\alpha)$$

L'estimateur correspondant sera de la forme :

$$Z^* = \sum_{\alpha} f_{\alpha}(Z_{\alpha})$$

où $f_\alpha = \sum_i \lambda_{\alpha,i} 1_{B_i}$ est une fonction B_i -mesurable. Lorsque la partition B_i devient de plus en plus fine, on obtient à la limite la meilleure approximation de Z_0 par une somme $\sum f_\alpha(Z_\alpha)$, où chaque f_α est une fonction Borel-mesurable. Nous dirons que cet estimateur est le krigeage disjonctif. Son calcul effectif nécessite la connaissance de chacune des lois à deux variables $F_{0\alpha}$ (pour Z_0 et Z_α) et $F_{\alpha\beta}$ (pour Z_α et Z_β). Il s'agit donc bien d'un intermédiaire réalisable entre le krigage et l'espérance conditionnelle. Le tableau suivant résume la situation :

Méthode	Krigeage	Krigeage Disjonctif	Espérance Conditionnelle
<u>Estimateur</u> :			
Z^* est la meilleure approximation de Z_0 de la forme :	$\sum \lambda^\alpha Z_\alpha$	$\sum f_\alpha(Z_\alpha)$ f_α , fonctions mesurables d'une seule variable	$f(Z_1, \dots, Z_n)$ f , fonction mesurable de n variables.
<u>Prerequisit</u> e :	Les covariances	chacune des lois à 2 variables $F_{0\alpha}$ et $F_{\alpha\beta}$	La loi à $n+1$ variables $F(dz_0, dz_1, \dots, dz_n)$

2 - LES EQUATIONS DU KRIGEAGE DISJONCTIF.

Désignons par H_α l'espace de Hilbert constitué par les V.A. de la forme $f(Z_\alpha)$ où f est mesurable et $E[(Z_\alpha)^2] < \infty$, et par H la somme $H = H_1 \oplus \dots \oplus H_n$ (Ce n'est pas une somme directe, car la constante 1 appartient à chacun des H_α) : le krigage disjonctif Z^* est simplement

la projection de Z_0 dans H . Il est donc caractérisé par :

$$(K.D.) \quad \langle Z^*, Y \rangle = \langle Z_0, Y \rangle, \quad \forall Y \in \bigcup_{\alpha=1}^n H_{\alpha}$$

et la variance d'estimation est :

$$\|Z_0 - Z^*\|^2 = \|Z_0\|^2 - \|Z^*\|^2$$

Explicitons les équations (K.D.). Si $Y \in H_{\alpha}$, Y est de la forme $Y = f(Z_{\alpha})$ pour une fonction f mesurable. Pour toute $VA Z$, on a :

$$\langle Z, Y \rangle = E[Z f(Z_{\alpha})] = E[f(Z_{\alpha}) E_{\alpha} Z]$$

où $E_{\alpha} Z$ désigne l'espérance conditionnelle de Z en Z_{α} . Autrement dit, puisque f est une fonction mesurable quelconque, l'estimateur $Z^* = \sum f_{\alpha}(Z_{\alpha})$ est caractérisé par les relations

$$(K.D.) \quad E_{\alpha} Z_0 = \sum_{\beta} E_{\alpha} f_{\beta}(Z_{\beta}) \quad (\alpha = 1, 2, \dots, n)$$

Ce système s'explique au moyen des lois conditionnelles F_{α}^0 et F_{α}^{β} de Z_0 et de Z^{β} à Z_{α} fixé. Sa mise en oeuvre ne nécessite donc que la connaissance des lois à 2 variables $F_{0\alpha}$ et $F_{\beta\alpha}$, comme annoncé ci-dessus. Nous verrons qu'il est possible d'obtenir de manière simple des solutions approchées aussi bonnes que l'on veut. Nous allons au préalable examiner la :

3 - REPRESENTATION D'UNE LOI A DEUX VARIABLES.

Le formalisme utilisé par J.P. Benzecri pour l'analyse des correspondances se transcrit avantageusement en termes d'espaces

$L^2(\Omega, \mathcal{A}, P)$, et cette transposition est d'ailleurs nécessaire dans le cas continu (cf. Thèse Naouri).

En effet, soit (X, Y) une V.A. à valeurs dans un ensemble produit $I \times J$ fini, $i \in I$, $j \in J$ et :

$$f_{ij} = P(\{X = i\} \cap \{Y = j\})$$

$$f_i = P(\{X = i\}) \quad , \quad f_j = P(Y = j)$$

$$f_i^j = f_{ij}/f_j = P(X = i | Y = j)$$

(on suppose $f_i, f_j > 0$ pour tout i et tout j). Si p est une probabilité sur I définie par $p_i = p(\{i\})$, la norme de p pour la métrique euclidienne du χ^2 est :

$$\|p\|^2 = \sum_i (p_i)^2 / f_i$$

Posons $\varphi^i = p_i / f_i$, de sorte que la fonction $i \rightarrow \varphi^i$ est la densité φ de p relativement à la loi f de X . On trouve :

$$\|p\|^2 = \sum_i (\varphi^i)^2 f_i = E(\varphi^2)$$

Autrement dit : la norme de p pour la métrique du χ^2 est identique à la norme de sa densité φ pour la métrique hilbertienne de $L^2(I, f)$.

Plus généralement, soient $(\Omega_1, \mathcal{A}_1)$ et $(\Omega_2, \mathcal{A}_2)$ deux espaces mesurables, $(\Omega, \mathcal{B})$ l'espace produit ($\Omega = \Omega_1 \times \Omega_2$, $\mathcal{B} = \mathcal{A}_1 \otimes \mathcal{A}_2$) et P une probabilité sur $(\Omega, \mathcal{B})$. Si $\Omega_1 = \Omega_2 = \mathbb{R}$, P est simplement une loi à deux variables $P(dx, dy)$. Pour garantir l'existence des probabilités conditionnelles, nous supposerons plus généralement

que Ω_1 et Ω_2 sont deux espaces LCD munis de leurs tribus boréliennes.

Désignons par X et Y les projections ($X(x,y) = x$, $Y(x,y) = y$ pour tout $(x,y) \in \Omega = \Omega_1 \times \Omega_2$) et par ϖ_1 et ϖ_2 les lois de ces V.A., i.e. les lois marginales. Pour ϖ_1 -presque tout $x \in \Omega_1$, il existe une loi conditionnelle P_x sur Ω_2 (loi de Y à $X = x$ fixé). Nous admettrons de plus qu'il existe une fonction φ sur Ω telle que $P_x = \varphi(x, \cdot) \varpi_2$, ou explicitement :

$$(1) \quad P(dx, dy) = \varphi(x, y) \varpi_1(dx) \varpi_2(dy)$$

En particulier, pour ϖ_1 -presque tout $x \in \Omega_1$ (resp. pour ϖ_2 -presque tout $y \in \Omega_2$) $\varphi(x, \cdot) \varpi_2$ (resp. $\varphi(\cdot, y) \varpi_1$) est la loi de Y (resp. de X) conditionnelle en $X = x$ (resp. en $Y = y$).

Nous écrirons souvent φ_x pour $\varphi(x, \cdot)$, φ_y pour $\varphi(\cdot, y)$ et nous sous-entendrons les incantations fastidieuses : ϖ_1 -p.s. sur Ω_1 etc... On notera que $x \rightarrow \varphi_x$ et $y \rightarrow \varphi_y$ peuvent aussi être interprétés comme des fonctions aléatoires, et nous aurons à examiner le sens d'un krigeage effectué à l'aide de ces F.A. très particulières.

Enfin, nous supposerons essentiellement $\varphi \in L^2(\Omega, \mathcal{B}, P)$, ce qui entraîne $\varphi_x \in L^2(\Omega_1, \mathcal{A}_1, \varpi_1)$ et $\varphi_y \in L^2(\Omega_2, \mathcal{A}_2, \varpi_2)$. Pour abréger les notations, on posera :

$$H = L^2(\Omega, \mathcal{B}, P)$$

$$H_1 = L^2(\Omega_1, \mathcal{A}_1, \varpi_1)$$

$$H_2 = L^2(\Omega_2, \mathcal{A}_2, \varpi_2)$$

d'où $H_1 \subset H$ et $H_2 \subset H$. C'est, pour l'essentiel, cet espace de Hilbert

H et ses deux sous-espaces H_1 et H_2 qui vont constituer notre cadre de travail.

4 - LES PROJECTEURS E_1 ET E_2 .

Pour tout $f \in H$, on désignera par $\bar{E}_1 f$ (resp. $\bar{E}_2 f$) l'espérance conditionnelle de f en X (resp. en Y), soit

$$E_1 f(x) = \int f(x,y) \varphi(x,y) \varpi_2(dy)$$

avec $\bar{E}_1 f \in H_1$. On sait que $\bar{E}_1 f$ est la projection de f dans H_1 , de sorte $\bar{E}_1$ est le projecteur associé à ce sous-espace.

Il sera commode de désigner par E_1 (non surligné) la restriction du projecteur $\bar{E}_1$ au sous-espace H_2 , de sorte que E_1 applique H_2 dans H_1 . De même, E_2 , restriction de $\bar{E}_2$ à H_1 , applique H_1 dans H_2 . Ces opérateurs sont adjoints l'un de l'autre, soit :

$$E_1^* = E_2, \quad E_2^* = E_1$$

En effet, soit $f \in H_1$ et $g \in H_2$. On trouve :

$$\langle f, E_1 g \rangle = \langle f, g \rangle = \langle E_2 f, g \rangle$$

donc $E_2 = E_1^*$.

Les opérateurs $E_1 E_2$ et $E_2 E_1$.

L'opérateur $E_2 E_1$ applique H_1 dans lui-même. Il est auto-adjoint

et ≥ 0 , (car $E_2 = E_1^*$, et $E_1^* E_1$ est ≥ 0 , i.e. $\langle E_1^* E_1 f, f \rangle = \|E_1 f\|^2 \geq 0$ pour tout $f \in H_1$).

Explicitement, si $f \in H_1$ et $g \in H_2$, on a :

$$(2) \quad \begin{cases} E_2 f(y) = \int f(x) \varphi(x,y) \varpi_1(dx) = \langle f, \varphi_y \rangle \\ E_1 g(x) = \int g(y) \varphi(x,y) \varpi_2(dy) = \langle g, \varphi_x \rangle \end{cases}$$

Par suite :

$$E_1 E_2 f(x) = \iint f(\xi) \varphi(\xi,y) \varphi(x,y) \varpi_1(d\xi) \varpi_2(dy)$$

c'est-à-dire :

$$(3) \quad \begin{cases} E_1 E_2 f(x) = \int f(\xi) C_1(x,\xi) \varpi_1(d\xi) \\ C_1(x,\xi) = \int \varphi(x,y) \varphi(\xi,y) \varpi_2(dy) = \langle \varphi_x, \varphi_\xi \rangle = (E_1 \varphi_x)(\xi) \end{cases}$$

Ainsi, $E_1 E_2$ est l'opérateur intégral associé à la covariance C_1 de la fonction aléatoire φ_x . De même, $E_2 E_1$ est l'opérateur intégral sur H_2 associé à la covariance C_2 de la F.A. φ_y .

5 - LES VALEURS PROPRES DE $E_1 E_2$.

L'opérateur $E_1 E_2 = E_2^* E_2$ sur H_1 est auto-adjoint et positif (car $\langle E_1 E_2 f, f \rangle = \|E_2 f\|^2$). Ses valeurs propres sont donc réelles et positives. Elles sont ≤ 1 , car E_1 et E_2 sont des projecteurs. La fonction 1 est fonction propre associée à la valeur propre 1, et c'est en général la seule (à moins qu'il n'existe des fonctions f et g mesurables et non constantes telles que $f(X) = g(Y)$ p.s.).

Nous nous intéresserons surtout aux valeurs propres $\lambda > 0$.

Pour $\lambda > 0$, si $f \in H_1$ vérifie $E_1 E_2 f = \lambda f$, on en déduit $E_2 E_1 (E_2 f) = \lambda E_2 f$, de sorte que $E_2 f \in H_2$ est fonction propre de $E_2 E_1$ pour la valeur propre λ , et, λ n'étant pas nul, $E_2 f \neq 0$. Posons $f_\lambda = f/\|f\|$ et $g_\lambda = E_2 f_\lambda / \|E_2 f_\lambda\|$. On a :

$$\|E_2 f_\lambda\|^2 = \langle E_1 E_2 f_\lambda, f_\lambda \rangle = \lambda$$

et par suite :

$$(4) \quad g_\lambda = \frac{1}{\sqrt{\lambda}} E_2 f_\lambda ; \quad f_\lambda = \frac{1}{\sqrt{\lambda}} E_1 g_\lambda$$

et ces relations établissent une correspondance biunivoque entre les fonctions propres normées des opérateurs $E_1 E_2$ et $E_2 E_1$ associés à la valeur propre $\lambda > 0$.

Nous nous intéresserons surtout au cas où l'opérateur $E_1 E_2$ a un spectre discret et où par suite :

$$(5) \quad C_1(x, \xi) = \sum \lambda f_\lambda(x) f_\lambda(\xi)$$

Dans ce cas, la densité φ admet la représentation :

$$(6) \quad \varphi(x, y) = \sum \sqrt{\lambda} f_\lambda(x) g_\lambda(y)$$

En effet, on trouve d'abord :

$$\varphi(x, y) = \sum \langle \varphi_x, g_\lambda \rangle g_\lambda(y)$$

Mais, d'après (2) et (4), $\langle \varphi_x, g_\lambda \rangle = E_1 g_\lambda(x) = \sqrt{\lambda} f_\lambda(x)$, d'où (6).

6 - EXEMPLE 1 : LE CAS GAUSSIEN.

Le cas gaussien est traité dans la thèse de Naouri, et les résultats essentiels figurent déjà dans le Cramer.

Rappelons d'abord la définition des polynomes d'Hermite H_n :

$$(7) \quad e^{-\frac{x^2}{2}} H_n(x) = \frac{d^n}{d x^n} e^{-x^2/2} \quad (n = 0, 1, \dots)$$

Ils vérifient :

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} H_n(x) H_m(x) e^{-x^2/2} dx = \delta_{nm} n!$$

et les polynomes :

$$\eta_n = \frac{1}{\sqrt{n!}} H_n$$

constituent une base hilbertienne orthonormée de $L^2(\mathbb{R}, \frac{1}{\sqrt{2\pi}} e^{-x^2/2})$.
Notons d'abord le développement suivant :

$$(8) \quad e^{\lambda x} = e^{\lambda^2/2} \sum_{n=0}^{\infty} (-1)^n \frac{\lambda^n}{n!} H_n(x)$$

Cette formule résulte de :

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int e^{\lambda x} H_n(x) e^{-x^2/2} dx &= \frac{1}{\sqrt{2\pi}} \int e^{\lambda x} \frac{d^n}{d x^n} e^{-x^2/2} = \\ &= (-1)^n \frac{\lambda^n}{\sqrt{2\pi}} \int e^{\lambda x} e^{-x^2/2} dx = (-1)^n \lambda^n e^{\lambda^2/2} \end{aligned}$$

et du fait que les $\eta_n = (1/\sqrt{n!}) H_n$ constituent une base orthonormée.

Soit maintenant la loi de Gauss à 2 variables X et Y admettant la densité

$$\varphi(x,y) \propto \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$$

avec

$$\varphi(x,y) = \frac{1}{\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)} (\rho^2 x^2 + \rho^2 y^2 - 2\rho xy)}$$

On a le développement :

$$(9) \quad \varphi(x,y) = \sum_{n=0}^{\infty} \frac{\rho^n}{n!} H_n(x) H_n(y) = \sum \rho^n \eta_n(x) \eta_n(y)$$

Pour établir rapidement cette formule, partons de la transformée de Laplace de la loi de X à $Y = y$ fixé, soit :

$$E[e^{\lambda X}|Y=y] = e^{\lambda \rho y + (1-\rho^2) \frac{\lambda^2}{2}}$$

En utilisant le développement (8) pour exprimer $e^{\lambda \rho y}$, nous trouvons :

$$E(e^{\lambda X}|Y=y) = e^{\lambda^2/2} \sum (-1)^n \frac{\rho^n \lambda^n}{n!} H_n(y)$$

D'autre part, en exprimant directement $e^{\lambda x}$ à l'aide du développement (8), nous trouvons

$$\begin{aligned} E(e^{\lambda X}|Y=y) &= \int e^{\lambda x} \varphi(x,y) e^{-x^2/2} \frac{dx}{\sqrt{2\pi}} = \\ &= e^{\lambda^2/2} \sum (-1)^n \frac{\lambda^n}{n!} \int H_n(x) \varphi(x,y) e^{-x^2/2} \frac{dx}{\sqrt{2\pi}} \end{aligned}$$

En identifiant ces deux expressions de $E(e^{\lambda x}|Y)$, nous trouvons :

$$\frac{1}{\sqrt{2\pi}} \int H_n(x) \varphi(x,y) e^{-x^2/2} dx = \rho^n H_n(y)$$

at par suite $\langle \eta_n, \varphi_y \rangle = \rho^n \eta_n(y)$, d'où (9).

Ainsi, dans le cas d'une loi de Gauss à 2 variables, les

"facteurs" sont les polynomes d'Hermite normalisés η_n associés aux valeurs propres ρ^{2n} , $n = 0, 1, \dots$. On en déduit l'expression explicite de l'opérateur $E_1 E_2$ que l'on peut aussi vérifier par un calcul direct :

$$C_1(x, y) = \sum \rho^{2n} \eta_n(x) \eta_n(y) = \frac{1}{\sqrt{1-\rho^4}} e^{-\frac{1}{2(1-\rho^4)} (\rho^4 x^2 + \rho^4 y^2 - 2\rho^2 xy)}$$

Voyons quelques applications.

Calcul d'une espérance conditionnelle.

Ce cas correspond, par exemple, à une F.A. stationnaire qui se laisse ramener, par anamorphose gaussienne, à une F.A. à loi spatiale gaussienne. Le problème consiste alors à calculer l'espérance conditionnelle de $f(Z_0)$ connaissant Z_α , $\alpha = 1, 2, \dots, n$, f étant une fonction donnée et les Z_0 , Z_α étant des V.A. normales de variances unité admettant les coefficients de corrélation $\rho_{0\alpha}$, $\rho_{\alpha\beta}$.

A $Z_1, \dots, Z_n$ fixés, Z_0 est une gaussienne d'espérance $\lambda^\alpha Z_\alpha$ et de variance σ_K^2 (les λ^α et σ_K^2 étant calculés par résolution du système de krigeage). On peut donc calculer directement l'espérance et la variance de $f(Z_0)$. Il est peut-être plus rapide (si l'on doit traiter un grand nombre de configurations différentes, pour lesquelles les λ^α et les σ_K^2 ne sont évidemment pas constantes) d'utiliser le développement de f en polynomes d'Hermite normalisés :

$$f = \sum f_n \eta_n$$

$$f_n = \langle f, \eta_n \rangle = \frac{1}{\sqrt{2\pi}} \int f(x) \frac{H_n(x)}{\sqrt{n!}} e^{-x^2/2} dx$$

En effet, posons $Z_K = \lambda^\alpha Z_\alpha$ et $Y = Z_K / \sigma_K$. Z_0 et Y sont des gaussiennes de variance unité admettant le coefficient de corrélation :

$$\rho = \langle Z_0, Y \rangle = \frac{1}{\sigma_K} \|Z_K\|^2 = \sigma_K$$

A $Z_1, \dots, Z_n$ fixé, Z_0 est une gaussienne de densité $\varphi(z_0, y) e^{-z_0^2/2} 1/\sqrt{2\pi}$, $\varphi(z_0, y)$ étant donné par (9) ($y = \lambda^\alpha z_\alpha / \sigma_K$).
On a :

$$E(\eta_p(Z_0)/y) = \rho^p \eta_p(y)$$

et par suite :

$$E(f(Z_0) | Z_1, \dots, Z_n) = \sum_p \rho^p f_p \eta_p \left(\frac{\lambda^\alpha Z_\alpha}{\sigma_K} \right)$$

Cette formule doit permettre un calcul rapide.

Calcul d'un krigeage disjonctif.

Bien souvent, on pourra vérifier que l'anamorphose gaussienne transforme la F.A. initiale en une F.A. $Z(x)$ telle que pour deux points d'appui quelconques x et y la loi à 2 variables $(Z(x), Z(y))$ soit gaussienne (où à peu près) sans qu'il soit cependant possible de vérifier que la loi spatiale de $Z(x)$ est normale dans son intégralité (i.e. lorsque plus de deux points sont en jeu). On doit alors faire un krigeage disjonctif de $f(Z_0)$ à l'aide des Z_α .

Notons comme ci-dessus $\rho_{0\alpha} = E(Z_0 Z_\alpha)$, $\rho_{\alpha\beta} = E(Z_\alpha Z_\beta)$ et $f = \sum f_p \eta_p$. Le krigeage disjonctif est $f^* = \sum f_\alpha(Z_\alpha)$ avec les conditions $E_\alpha f = E_\alpha f^*$, $\alpha = 1, 2, \dots, n$. On posera donc :

$$f_{\alpha} = \sum f_{\alpha}^p \eta_p$$

et on notera :

$$E_{\alpha}(\eta_p(Z_0)) = \rho_{0\alpha}^p \eta_p(Z_{\alpha}) \quad ; \quad E_{\alpha}(\eta_p(Z_{\beta})) = \rho_{\alpha\beta}^p \eta_p(Z_{\beta})$$

On obtient donc le krigeage disjonctif $\sum f_{\alpha}(Z_{\alpha})$ en résolvant pour chaque p le système suivant :

$$(10) \quad \sum_{\alpha} f_{\alpha}^p \rho_{\alpha\beta}^p = \rho_{0\beta}^p$$

qui n'est autre que le système de krigeage simple écrit avec la matrice $\sigma_{\alpha\beta} = \sigma_{\alpha\beta}^p$ (ce qui garantit une décroissance assez rapide des f_{α}^p lorsque p augmente).

Variances d'estimation.

Pour apprécier l'estimation, calculons dans les deux cas ci-dessus la variance $E((f-f^*)^2)$ (qui n'est pas la variance conditionnelle). Dans le premier cas, on trouve immédiatement :

$$E((f-f^*)^2) = E(f^2) - E(f^{*2}) = \sum (1 - \rho^{2p}) f_p^2$$

Dans le cas du krigeage disjonctif :

$$(f-f^*)^2 = \sum_{p,q} [f_p \eta_p(Z_0) - \sum_{\alpha} f_{\alpha}^p \eta_p(Z_{\alpha})] [f_q \eta_q(Z_0) - \sum_{\beta} f_{\beta}^q \eta_q(Z_{\beta})]$$

Or, pour deux gaussiennes X, Y centrées, normées, de coefficient de corrélation ρ :

$$E(\eta_p(X) \eta_q(Y)) = \delta_{pq} \rho^p$$

Par suite :

$$E[(f-f^*)^2] = \sum_p f_p^2 - 2 \sum_{\alpha p} f_p f_p^\alpha \rho_{0\alpha}^p + \sum_{p, \alpha, \beta} f_p^\alpha f_p^\beta \rho_{\alpha\beta}^p$$

Compte tenu de (10), ceci se réduit à :

$$E[(f-f^*)^2] = \sum_p f_p^2 - \sum_{\alpha, p} f_p f_p^\alpha \rho_{0\alpha}^p$$

7 - EXEMPLE 2 : LOIS GAMMA.

La situation précédente, où une anamorphose gaussienne restitue une F.A. à loi spatiale gaussienne, n'est absolument pas générale. C'est pourquoi il n'est pas sans intérêt d'étudier au moins un exemple où cette simplification ne se produit pas.

Partons d'une mesure aléatoire μ sur $\mathbb{R}^n$, stationnaire, orthogonale et à loi gamma, i.e. pour tout borélien $B \subset \mathbb{R}^n$ de volume $V(B) = V < \infty$, on a

$$\begin{cases} E[e^{-\lambda \mu(B)}] = e^{-\psi(\lambda)} \\ \psi(\lambda) = \theta V \log(a + \lambda) \end{cases}$$

Désignons par $Z(x)$ la régularisée de μ par un compact B donné, i.e.

$$Z(x) = \int \mu(d\xi) 1_B(x+\xi)$$

et par $K(h) = \int 1_B(x) 1_B(x+h)$ le covariogramme de B . On trouve alors :

$$E[e^{-\lambda Z(x) - \mu Z(x+h)}] = e^{-\theta K(h) \psi(\lambda + \mu) - \theta [K(0) - K(h)](\psi(\lambda) + \psi(\mu))}$$

En effet :

$$\begin{cases} Z(x) = \mu(B_x \cap B_{x+h}) + \mu(B_x \setminus B_x \cap B_{x+h}) \\ Z(x+h) = \mu(B_x \cap B_{x+h}) + \mu(B_{x+h} \setminus B_x \cap B_{x+h}) \end{cases}$$

Ainsi, les variables $X = Z(x)$ et $Y = Z(x+h)$ sont de la forme :

$$\begin{cases} X = X_1 + Z \\ Y = Y_1 + Z \end{cases}$$

où X_1 , Y_1 et Z sont trois variables gamma indépendantes dont les deux premières admettent le même paramètre $\alpha = \theta[K(o)-K(h)]$ et la dernière $\alpha = \theta K(h)$.

X et Y sont elles-mêmes des V.A. (non indépendantes) admettant le même paramètre

$$\alpha = \theta K(o) = \theta V(B)$$

En posant

$$\rho = K(h)/K(o)$$

(coefficient de corrélation de X et Y), on voit que Z a le paramètre $\rho\alpha$ et X_1 et Y_1 le paramètre $(1-\rho)\alpha$. Pour simplifier, nous supposons que le paramètre d'échelle a est égal à 1 et nous nous proposons de déterminer la représentation de la loi gamma à deux paramètres ainsi définie.

Les Polynômes de Laguerre.

Pour un paramètre $\alpha > 0$ donné, le polynôme de Laguerre $L_n^\alpha = L_n$ est défini par :

$$\frac{d^n}{dx^n} x^{n+\alpha-1} e^{-x} = (-1)^n n! x^{\alpha-1} e^{-x} L_n(x)$$

Ces polynomes sont orthogonaux, car on trouve en intégrant par parties :

$$\frac{1}{\Gamma(\alpha)} \int_0^\infty L_n(x) x^{\alpha-1} e^{-x} dx = \delta_{n,m} C_n$$

$$C_n = \Gamma(n+\alpha)/n! \Gamma(\alpha)$$

et les polynomes $\ell_n = L_n / \sqrt{C_n}$ constituent une base orthonormée de $L^2(\mathbb{R}_+, \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x})$.

Calculons le développement de $e^{-\lambda x}$ en polynomes de Laguerre.

On trouve :

$$\begin{aligned} \frac{1}{\Gamma(\alpha)} \int_0^\infty e^{-\lambda x} L_n(x) x^{\alpha-1} e^{-x} dx &= \frac{(-1)^n}{n!} \frac{1}{\Gamma(\alpha)} \int_0^\infty e^{-\lambda x} \frac{d^n}{dx^n} (x^{\alpha+n-1} e^{-x}) dx \\ &= \frac{(-1)^n}{\Gamma(\alpha)} \frac{\lambda^n}{n!} \int_0^\infty x^{\alpha+n-1} e^{-(1+\lambda)x} dx = (-1)^n \frac{\lambda^n}{n!} \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)} \frac{1}{(1+\lambda)^{\alpha+n}} \end{aligned}$$

c'est-à-dire :

$$\langle e^{-\lambda x}, \ell_n \rangle = (-1)^n \sqrt{C_n} \frac{\lambda^n}{(1+\lambda)^{n+\alpha}}$$

D'où le développement :

$$(11) \quad e^{-\lambda x} = \sum (-1)^n \sqrt{C_n} \frac{\lambda^n}{(1+\lambda)^{n+\alpha}} \ell_n = \sum (-1)^n \frac{\lambda^n}{(1+\lambda)^{n+\alpha}} L_n(x)$$

Montrons, maintenant, que la fonction associée ^φ à la loi de (X,Y) , (i.e. (X,Y) admet la densité :

$$(12) \quad \left(\frac{1}{\Gamma(\alpha)} \right)^2 \varphi(x, y) x^{\alpha-1} y^{\alpha-1} e^{-x-y}$$

se met sous la forme :

$$(12') \quad \varphi(x, y) = \sum u_n \ell_n(x) \ell_n(y)$$

d'où résultera que les polynômes normés ℓ_n sont les facteurs associés à cette loi, et les u_n les valeurs propres correspondantes. Pour cela, notons que si φ est de la forme (12'), la loi de densité (12) admet la transformée de Laplace :

$$E[e^{-\lambda X - \mu Y}] = \sum u_n \langle e^{-\lambda x}, \ell_n \rangle \langle e^{-\mu y}, \ell_n \rangle$$

soit, d'après (11) :

$$(13) \quad E[e^{-\lambda X - \mu Y}] = \sum c_n u_n \frac{\lambda^n}{(1+\lambda)^{n+\alpha}} \frac{\mu^n}{(1+\mu)^{n+\alpha}}$$

Il suffit donc de calculer directement cette transformée de Laplace et de montrer qu'elle admet un développement de la forme (13) pour établir la validité de (12) et obtenir l'expression explicite de u_n . Or on trouve directement :

$$\begin{aligned} E[e^{-\lambda X - \mu Y}] &= E[e^{-\lambda X_1 - \mu Y_1 - (\lambda + \mu) Z_1}] = \\ &= \frac{1}{(1+\lambda)^{(1-\rho)\alpha}} \frac{1}{(1+\mu)^{(1-\rho)\alpha}} \frac{1}{(1+\lambda+\mu)^{\rho\alpha}} = \\ &= \frac{1}{(1+\lambda)^\alpha (1+\mu)^\alpha} \left(\frac{(1+\lambda)(1+\mu)}{1+\lambda+\mu} \right)^{\rho\alpha} \end{aligned}$$

D'après le développement :

$$\left(\frac{(1+\lambda)(1+\mu)}{1+\lambda+\mu} \right)^{\rho\alpha} = \left(1 - \frac{\lambda \mu}{(1+\lambda)(1+\mu)} \right)^{-\rho\alpha} = \sum \frac{\Gamma(\rho\alpha+n)}{n! \Gamma(\rho\alpha)} \left(\frac{\lambda \mu}{(1+\lambda)(1+\mu)} \right)^n$$

nous trouvons bien :

$$E(e^{-\lambda X - \mu Y}) = \sum \frac{\Gamma(\rho\alpha + n)}{n! \Gamma(\rho\alpha)} \frac{\lambda^n \mu^n}{(1+\lambda)^{n+\alpha} (1+\mu)^{n+\alpha}}$$

En comparant avec (13), nous concluons que la représentation (12') est valable, et que les valeurs propres u_n sont données par

$$u_n = \frac{1}{c_n} \frac{\Gamma(\rho\alpha + n)}{n! \Gamma(\rho\alpha)} = \frac{\Gamma(\rho\alpha + n)}{\Gamma(\rho\alpha)} \frac{\Gamma(\alpha)}{\Gamma(\alpha + n)}$$

Contrairement au cas gaussien, on n'a pas $u_n = \rho^n$, mais $u_n = E(\Pi^n)$ où Π est une variable Beta ($\rho\alpha$, $(1-\rho)\alpha$) d'espérance ρ . En particulier (mais c'était bien évident au départ) une anamorphose gaussienne ne transforme pas cette loi en loi de Gauss à deux variables.

A titre de curiosité, on peut se demander s'il existe une loi à 2 variables pour laquelle la fonction ϕ serait

$$\phi(x, y) = \sum \rho^n \ell_n(x) \ell_n(y)$$

Il suffit pour cela que ϕ soit ≥ 0 .

D'après (13), la transformée de Laplace de ϕ est :

$$\Phi(\lambda, \mu) = \sum \rho^n \frac{\Gamma(n+\alpha)}{n! \Gamma(\alpha)} \frac{\lambda^n \mu^n}{(1+\lambda)^{\alpha+n} (1+\mu)^{\alpha+n}}$$

On reconnaît le développement en série de :

$$\Phi(\lambda, \mu) = \frac{1}{[(1+\lambda)(1+\mu)]^\alpha} \left(1 - \frac{\rho \lambda \mu}{(1+\lambda)(1+\mu)} \right)^{-\alpha} = \frac{1}{[1+\lambda+\mu+(1-\rho)\lambda\mu]^\alpha}$$

Il reste à montrer qu'il s'agit effectivement de la transformée

d'une loi à deux variables (d'ailleurs indéfiniment divisible).

Cela résulte du développement suivant :

$$\begin{aligned}\Phi(\lambda, \mu) &= \frac{(1-\rho)^\alpha}{(1+\lambda(1-\rho))^\alpha (1+\mu(1-\rho))^\alpha} \left[1 - \frac{\rho}{(1+\lambda(1-\rho))(1+\mu(1-\rho))} \right]^{-\alpha} \\ &= \frac{(1-\rho)^\alpha}{\Gamma(\alpha)} \sum_n \frac{\Gamma(\alpha+n)}{n!} \frac{\rho^n}{(1+\lambda(1-\rho))^{\alpha+n} (1+\mu(1-\rho))^{\alpha+n}}\end{aligned}$$

où l'on reconnaît (pour $\rho \geq 0$) la transformée de Laplace de la loi de densité

$$f(x, y) = \frac{1}{\Gamma(\alpha)} \sum_n \frac{1}{n! \Gamma(\alpha+n)} \frac{\rho^n}{(1-\rho)^{\alpha+2n}} (x y)^{\alpha+n-1} e^{-\frac{x+y}{1-\rho}}$$

Krigeage Disjonctif.

Dans les deux exemples ci-dessus, la fonction φ associée à la loi $\omega(dx) \varphi(x, y) \omega(dy)$ est de la forme

$$(14) \quad \varphi(x, y) = \sum u_n \chi_n(x) \chi_n(y)$$

où les χ_n constituent une base orthonormée de $L^2(\mathbb{R}, \omega)$, et où $u_n = u_n(\rho)$ ne dépend que du coefficient ρ .

Supposons qu'une F.A. Z vérifie la propriété : pour tous points α, β la loi $F_{\alpha\beta}$ de (Z_α, Z_β) est $\varphi_{\alpha\beta}(x, y) \omega(dx) \omega(dy)$ où $\varphi_{\alpha\beta}$ est de la forme (14). On posera $u_n(\rho_{\alpha\beta}) = u_{\alpha\beta}^n$. Le krigeage disjonctif de $f(Z_0)$ à l'aide des Z_α , $\alpha = 1, 2, \dots, n$ est alors aisé à réaliser. On part du développement

$$f = \sum f^p \chi_p$$

d'où l'on tire $E_{\alpha} f = \sum f^p u_{o\alpha}^p \chi_p$. Le krigeage disjonctif est $f^* = \sum f_{\alpha}$, f_{α} fonction mesurable de Z_{α} , avec :

$$E_{\alpha} f = \sum_{\beta} E_{\alpha} f_{\beta}$$

Donc, en utilisant les développements $f_{\beta} = \sum_p f_{\beta}^p \chi_p$, on obtient pour chaque p le système :

$$\sum_{\beta} u_{\alpha\beta}^p f_{\beta}^p = f^p u_{o\alpha}^p$$

Autrement dit, on krige séparément chacun des facteurs.

La résolution effective de ce système ne pose pas de problème. On est ainsi amené à chercher des modèles de fonctions aléatoires $Z(x)$ possédant les propriétés précédentes, à savoir :

a) $Z(x)$ est stationnaire et, pour un x donné, la loi $\omega(dz)$ de la V/A. $Z(x)$ admet un système dénombrable complet de fonction χ_n orthogonales pour ω (i.e. les χ_n sont une base hilbertienne de $L^2(\mathbb{R}, \omega)$).

b) pour deux points d'appui quelconques x et y , la loi de $(Z(x), Z(y))$ est $\varphi(z, z') \omega(dz) \omega(dz')$ avec :

$$\varphi(z, z') = \sum u_n(x, y) \chi_n(z) \chi_n(y)$$

On s'intéressera en particulier au cas où les facteurs χ_n sont des polynomes.

8 - RECHERCHE DE MODELES.

Pour chercher des F.A. vérifiant les 2 conditions ci-dessus, plusieurs problèmes se posent. Une densité de probabilité $f_{\alpha\beta}(z, z')$ joue, dans le krigeage disjonctif, le même rôle qu'une fonction de covariance dans le krigeage ordinaire. On aurait donc besoin d'une généralisation du théorème de Bochner, pour caractériser les familles de lois à 2 variables qui peuvent être associées à des F.A., i.e. :

Etant donnée une famille $F(dz, dz'; x_1, x_2)$ ($x_1, x_2 \in \mathbb{R}^n$) de lois à 2 variables, à quelle condition existe-t-il une F.A. $Z(x)$ sur $\mathbb{R}^n$ telle que $F(dz, dz'; x_1, x_2)$ soit la loi de $(Z(x_1), Z(x_2))$. Evidemment, pour tout $x \in \mathbb{R}^n$, la probabilité $\omega(.) = F(., R; x, x')$ doit être indépendante de x' .

Formons une condition nécessaire, par analogie avec la condition suivant laquelle une covariance doit être de type positif. Soient $x_1, \dots, x_k$ k points dans $\mathbb{R}^n$ et $f_i \in L^2(\mathbb{R}, \omega_i)$ ($i = 1, 2, \dots, k$). On doit avoir

$$E\left[\left(\sum_i f_i(Z(x_i))\right)^2\right] \geq 0$$

c'est-à-dire :

$$(15) \quad \sum_{i,j} \int f_i(z) f_j(z') F(dz, dz' ; x_i, x_j) \geq 0$$

Mais cette condition nécessaire n'est probablement pas suffisante. En effet, si B_k est une partition de R , les fonctions aléatoires en tout ou rien $1_{B_k}(Z(x))$ sont associées à une partition

aléatoire de $\mathbb{R}^n$, soit A_k ($1_{A_k}(x) = 1_{B_k}(Z(x))$), et

$$(16) \quad P(x \in A_k, x' \in A_k) = F(B_k, B_k; x, x')$$

La condition (15) exprime que les $C_{k,k}(x, x') = F(B_k, B_k; x, x')$ constituent la matrice des covariances d'une F.A. vectorielle $(Y_1(x), \dots, Y_k(x))$, mais rien ne prouve que cette F.A. vectorielle soit l'indicatrice d'une partition aléatoire.

On peut penser par contre que la condition suivante serait nécessaire et suffisante : pour toute partition dénombrable B_k , il existe une partition aléatoire A_k vérifiant (16).

Ceci attire l'attention sur le problème suivant : A quelle condition une famille $C_{k,k}(x, x')$ est-elle la matrice des covariances d'une partition aléatoire, avec le cas particulier simple :

A quelle condition une fonction $C(x, x')$ est-elle la covariance d'un ensemble aléatoire?

Ce problème n'est pas trivial, car une covariance du type $\exp \{-|x-x'|^2\}$ ne peut pas être associée à un ensemble aléatoire.

Ces problèmes semblent assez coriaces, et nous ne les abordons pas ici. Il y a, toutefois, un procédé qui conduit automatiquement à des lois à deux variables vérifiant les conditions voulues, et qui consiste à régulariser une mesure aléatoire orthogonale stationnaire μ , i.e. :

$$Z(x) = \int k(x+\xi) \mu(d\xi)$$

où k est donnée. Supposons que la loi (indéfiniment divisible) associée à μ soit définie par :

$$E[e^{-\lambda \mu(V)}] = e^{V \psi(\lambda)}$$

Alors :

$$E[e^{-\lambda Z(x) - \mu Z(x+h)}] = e^{\int \psi[\lambda k(\xi) + \mu k(\xi+h)] d\xi}$$

Si $k = 1_B$ et $K = k \check{*} k$, en particulier, on retrouve comme ci-dessus :

$$E[e^{-\lambda Z(x) - \mu Z(x+h)}] = e^{[K(o) - K(h)][\psi(\lambda) + \psi(\mu)] + K(h) \psi(\lambda + \mu)}$$

En posant $\Phi = e^{K(o)\psi}$, $\rho = K(h)/K(o)$ ceci s'écrit :

$$(17) \quad \Phi(\lambda, \mu) = \Phi(\lambda)^{1-\rho} \Phi(\mu)^{1-\rho} \Phi(\lambda + \mu)^\rho$$

Modèles à facteurs polynomiaux.

Cherchons quelles sont les lois de la forme (17) dont les facteurs χ_n sont des polynomes. Nous avons déjà vu que la loi de Gauss et la loi gamma possèdent cette propriété. Nous allons voir qu'il en est de même de la loi de Poisson et de la loi binomiale négative, et que ces quatre lois sont les seules lois non triviales de la forme (17) possédant cette propriété.

Notons d'abord que si, pour une loi $\omega(dx)$ il existe une base hilbertienne χ_n , $n = 0, 1, \dots$ où χ_n est un polynome de degré n , les deux propriétés suivantes sont équivalentes pour la variable (X, Y)

admettant une loi symétrique $\varphi(x,y) \omega(dx) \omega(dy)$:

$$1 - \varphi \text{ est de la forme } \varphi(x,y) = \sum u_n \chi_n(x) \chi_n(y)$$

2 - Pour tout $n \geq 0$, $E(X^n|Y)$ est un polynome de degré n en Y .

En effet, $1/ \Rightarrow 2$, car si $1/$ est vérifiée on trouve

$$E(X^n|Y=y) = \sum_p u_p \chi_p(y) \int x^n \chi_p(x) \omega(dx)$$

Comme $\langle x^n, \chi_p \rangle = 0$ pour $p > n$, il reste :

$$E(X^n|Y=y) = \sum_{p=0}^n u_p \langle x^n, \chi_p \rangle \chi_p(y)$$

et cette expression est un polynome de degré $\leq n$ (en fait, de degré strictement égal à n , car $\langle x^n, \chi_n \rangle$ ne peut être nul puisque

$$x^n = \sum_{p=0}^n \langle x^n, \chi_p \rangle \chi_p(x)$$

Inversement, supposons que 2 soit vérifiée. Il faut montrer que les polynomes orthogonaux χ_n vérifient $E(\chi_n(x)|y) = u_n \chi_n(y)$. Or l'espace $\mathcal{P}_{n+1}$ de dimension $n+1$ constitué des polynomes de degré $\leq n$ est invariant par l'opérateur $E(X|Y)$. De plus, $E(X|Y)$ applique $\mathcal{P}_{n+1}$ sur lui-même, puisqu'il conserve le degré des polynomes. Cet opérateur symétrique admet donc $n+1$ vecteurs propres orthogonaux dans $\mathcal{P}_{n+1}$ qui coïncident nécessairement avec les χ_k , $k = 0, 1, \dots, n$. La propriété 1 en résulte.

Cherchons maintenant à quelle condition une loi de la forme (17) admet des facteurs polynomiaux. D'après la proposition ci-dessus, cette loi doit vérifier la propriété 2. De la relation :

$$\int e^{-\lambda y} E(X^n | y) \omega(dy) = (-1)^n \frac{\partial^n}{\partial \mu^n} \Phi(\lambda, \mu) \Big|_{\mu=0}$$

on déduit que pour tout n la transformée de Laplace Φ doit vérifier :

$$(\Phi(\lambda))^{1-\rho} \frac{\partial^n}{\partial \lambda^n} (\Phi(\lambda))^\rho = \sum_{k=0}^n A_{n_k} \frac{\partial^k}{\partial \lambda^k} \Phi(\lambda)$$

Pour $n = 1$, cette relation est de la forme $\rho\Phi' = a\Phi + b\Phi'$. Elle est toujours vérifiée (avec $a = 0$, $b = \rho$). Pour $n = 2$ et en posant

$$\Phi = e^\psi$$

on trouve une équation différentielle de la forme

$$\psi'' = a \psi'^2 + b \psi' + c$$

Si $a = b = 0$, ψ est un polynome du second degré, et on trouve la loi de Gauss. Si $a = 0$, $b \neq 0$, on trouve $\psi = A e^{b\lambda} + B$, et la loi correspondante est une loi de Poisson. Supposons $a \neq 0$. Si le trinome $a x^2 + bx + c$ a une racine double α , on a :

$$\psi'' = a(\psi' - \alpha)^2$$

et la solution est $\psi = \alpha \lambda - (1/a) \log(1 + c\lambda)$. Il s'agit d'une loi gamma (éventuellement transposée et translatée). Supposons que le trinome n'ait pas de racines réelles, soit :

$$\psi'' = a [(\psi' + \alpha)^2 + b^2]$$

on trouve $\psi(\lambda) = (1/a) \log \cos(ab\lambda + c) - \alpha\lambda - c'$, et il n'existe pas de loi de probabilité associée à e^ψ . Il reste enfin le cas où

le trinome a deux racines nulles, soit :

$$\psi'' = a (\psi' - \alpha)(\psi' - \beta)$$

On en tire

$$\frac{\psi' - \alpha}{\psi' - \beta} = b e^{a(\beta - \alpha)\lambda}$$

$$\psi'(\lambda) = \frac{\alpha - \beta b e^{a(\beta - \alpha)\lambda}}{1 - b e^{a(\beta - \alpha)\lambda}}$$

En supposant $a(\beta - \alpha) > 0$ (ce qui est loisible, puisque l'on peut permuter α et β), on obtient la solution suivante :

$$\psi(\lambda) = \beta\lambda - \frac{1}{a} \log \left(\frac{b - e^{-a(\beta - \alpha)\lambda}}{b - 1} \right)$$

associée (si $|b| > 1$) à une loi binomiale négative (toujours à une transformation linéaire près).

Lois de Poisson.

Il reste à vérifier que les 4 lois ainsi trouvées satisfont aux conditions voulues pour $n > 2$. Les lois de Gauss et les lois gamma conviennent effectivement, comme on l'a déjà vu. Examinons la loi de Poisson. Au lieu de la transformée de Laplace, nous utiliserons la fonction génératrice $G(s)$, soit

$$G(s) = e^{\theta(s-1)}$$

La loi à deux variables associée à (17) est définie par :

$$G(s,t) = \sum P_{nm} s^n t^m = (G(s))^{1-\rho} (G(t))^{1-\rho} (G(st))^\rho$$

soit

$$\begin{aligned} \log G(s,t) &= \theta \rho(st-1) + \theta(1-\rho)(s-1) + \theta(1-\rho)(t-1) = \\ &= \theta(t-1) + \theta(1-\rho t)(s-1) \end{aligned}$$

Par suite pour n fixé :

$$\sum_m P_{nm} t^m = \frac{\theta^n}{n!} e^{-\theta} (1-\rho+\rho t)^n e^{\theta(1-\rho)(t-1)}$$

La fonction génératrice $G_n(t)$ de Y à $X = n$ fixé est donc :

$$G_n(t) = (1-\rho+\rho t)^n e^{\theta(1-\rho)(t-1)}$$

Il suffit alors d'appliquer la règle de dérivation de Leibniz pour obtenir les moments conditionnels :

$$E[Y(Y-1)\dots(Y-k)|n] = \theta^{k+1} + C_{k+1}^1 \rho \theta^k n + \dots + \rho^{k+1} n(n-1)\dots(n-k)$$

Il en résulte bien que $E(Y^n|X)$ est un polynôme de degré n en X , et l'on note que la valeur propre u_n associée au polygone χ_n est $u_n = \rho^n$ comme dans le cas de la loi de Gauss.

Pour être complet, il convient de donner l'expression des polynômes orthogonaux associés à la loi de Poisson. Partons de la formule suivante (intégration par parties) :

$$u_0 v_0 + \sum_1^n u_n [v_n - v_{n-1}] = - \sum_0^\infty (u_{n+1} - u_n) v_n$$

Désignons par Δ l'accroissement d'ordre 1 ($(\Delta u)_n = u_{n+1} - u_n$)

et par δ le rétro-accroissement $(\delta u)_n = u_n - u_{n-1}$ si $n \geq 1$ et $(\delta u)_0 = u_0$. La formule précédente s'écrit :

$$\sum u_n (\delta v)_n = - \sum (\Delta u)_n v_n$$

et reste valable si l'on remplace δ et Δ par leurs itérés. Posons pour x entier ≥ 0 :

$$\frac{\theta^x}{x!} P_n(x) = \delta^n \frac{\theta^x}{x!}$$

Il est alors immédiat que les polynomes P_n sont orthogonaux pour la loi de Poisson. Explicitement :

$$(18) \quad P_n(x) = 1 - C_n^1 \frac{x}{\theta} + C_n^2 \frac{x(x-1)}{\theta^2} - \dots + (-1)^n \frac{x(x-1)\dots(x-n+1)}{\theta^n}$$

Il reste à normer ces polynomes P_n . Pour cela, calculons :

$$\sum (P_n(x))^2 \frac{\theta^x}{x!} e^{-\theta} = e^{-\theta} \sum_x P_n(x) \delta^n \frac{\theta^x}{x!} = - e^{-\theta} \sum_x \frac{\theta^x}{x!} \Delta^n P_n(x)$$

mais, d'après (18), $\Delta^n P_n(x) = (-1)^n \theta^{-n} \Delta^n x^n = (-1)^n n!/\theta^n$. Donc :

$$\sum_{x=1}^{\infty} (P_n(x))^2 \frac{\theta^x}{x!} e^{-\theta} = \frac{n!}{\theta^n}$$

et les polynomes

$$\varpi_n(x) = \sqrt{\frac{\theta^n}{n!}} P_n(x)$$

sont les polynomes orthonormés cherchés.

A titre de vérification, calculons le développement de s^x :

$$\langle s^x, P_n \rangle = \sum_x s^x e^{-\theta} \delta^n \frac{\theta^x}{x!} = (-1)^n e^{-\theta} \sum \frac{\theta^x}{x!} \Delta^n s^x$$

Or $\Delta^n s^x = s^x (s-1)^n$. Il vient donc :

$$\langle s^x, P_n \rangle = (1-s)^n e^{\theta(s-1)}$$

et

$$s^x = \sum_n \frac{\theta^n}{n!} (1-s)^n e^{\theta(s-1)} P_n(x)$$

On retrouve bien :

$$\begin{aligned} E(s^X t^Y) &= \sum_n \rho^n \frac{\theta^n}{n!} \langle s^x, P_n \rangle \langle s^y, P_n \rangle = \\ &= \sum_n \rho^n \frac{\theta^n}{n!} (1-s)^n (1-t)^n e^{\theta(s-1)+\theta(t-1)} \\ &= e^{\theta(1-\rho)[s-1+t-1] + \theta\rho(st-1)} \end{aligned}$$

Loi Binomiale négative.

Il nous reste enfin à vérifier que la loi binomiale négative satisfait également aux conditions annoncées. Cette loi admet une fonction génératrice de la forme :

$$G(s) = \left(\frac{1-\alpha}{1-\alpha s} \right)^\beta = (1-\alpha)^\beta \sum_{n=0}^{\infty} \alpha^n \frac{\Gamma(\beta+n)}{n! \Gamma(\beta)} s^n$$

($0 < \alpha < 1$, $\beta > 0$). La loi correspondante à 2 variables X et Y est définie par

$$G(s, t) = E(s^X t^Y) = G(s)^{1-\rho} G(t)^{1-\rho} G(st)^\rho$$

Cherchons la loi conditionnelle de Y en X fixé. Pour cela, partons de $X = X_1 + Z$, $Y = Y_1 + Z$ où X_1 , Y_1 , Z sont indépendantes et :

$$E(s^{X_1}) = E(s^{Y_1}) = G(s)^{1-\rho}, \quad E(s)^Z = G(s)^\rho$$

X_1 et Z ont des lois binomiales négatives, et :

$$P(X_1 = k, Z = p) = (1-\alpha)^\beta \alpha^{k+p} \frac{\Gamma(\rho\beta+p)}{p! \Gamma(\rho\beta)} \frac{\Gamma((1-\rho)\beta+k)}{k! \Gamma((1-\rho)\beta)}$$

On en déduit la loi de Z à $X = X_1 + Z = n$ fixé :

$$\begin{aligned} P(Z = p | X = n) &= \frac{P(Z = p, X_1 = n-p)}{P(X = n)} = \\ &= C_n^p \frac{\Gamma(\beta)}{\Gamma(\rho\beta) \Gamma((1-\rho)\beta)} \frac{\Gamma(\rho\beta+p) \Gamma((1-\rho)\beta+n-p)}{\Gamma(\beta+n)} \end{aligned}$$

avec, évidemment, $0 \leq p \leq n$. Ceci est une loi binomiale randomisée par une loi Beta, car :

$$P(Z = p | X = n) = C_n^p \frac{\Gamma(\beta)}{\Gamma(\rho\beta) \Gamma((1-\rho)\beta)} \int_0^1 x^p (1-x)^{n-p} x^{\rho\beta-1} (1-x)^{(1-\rho)\beta-1} dx$$

On en déduit la fonction génératrice conditionnelle de Z :

$$E(s^Z | X = n) = \frac{\Gamma(\beta)}{\Gamma(\rho\beta) \Gamma((1-\rho)\beta)} \int_0^1 [1-x+sx]^n x^{\rho\beta-1} (1-x)^{(1-\rho)\beta-1} dx$$

et par suite celle de $Y = Z + Y_1$:

$$E(s^Y | X = n) = \left(\frac{1-\alpha}{1-\alpha s} \right)^{(1-\rho)\beta} E(s^Z | n)$$

Compte tenu de la règle de dérivation de Leibniz, il suffit de vérifier que $E(Z^p|X=n)$ est un polynome de degré p en n pour qu'il en soit de même pour $E(Y^p|X=n)$. Or on trouve, en dérivant la fonction génératrice conditionnelle $E(s^Z|X=n)$:

$$E[Z(Z-1)\dots(Z-p+1)|n] = n(n-1)\dots(n-p+1) \int_0^1 x^p f_\beta(x) dx$$

Par suite, la loi binomiale négative répond aux conditions voulues. Nous ne donnerons pas ici l'expression explicite des polynomes orthogonaux. Notons seulement, d'après la relation ci-dessus, l'expression des valeurs propres :

$$u_p = \int_0^1 x^p f_\beta(x) dx = \frac{\Gamma(\rho\beta+p)}{\Gamma(\rho\beta)} \frac{\Gamma(\beta)}{\Gamma(\beta+p)}$$

9 - LE KRIGEAGE ET LA REPRESENTATION DES LOIS A 2 VARIABLES.

Reprenant, maintenant, les notations du paragraphe 3, nous allons montrer comment la représentation de la loi $\varphi(x,y)\omega_1(dx)\omega_2(dy)$ peut s'interpréter en termes de krigeage.

Notons d'abord que φ_x (c'est-à-dire l'application $x \rightarrow \varphi_x$ de Ω_1 dans $L^2(\Omega_2, \mathcal{A}_2, \omega_2) = H_2$) et $y \rightarrow \varphi_y$ sont des fonctions aléatoires. On désignera par $H'_2 \subset H_2$ l'espace de Hilbert engendré par les combinaisons linéaires $\sum \lambda_i \varphi_{x_i}$, et de même par $H'_1 \subset H_1$ l'espace de Hilbert engendré par les φ_y , $y \in \Omega_2$. Π_1 et Π_2 désigneront les projecteurs associés à H'_1 et H'_2 , de sorte que pour toute V.A. Y ,

π_1 Y est le krigeage de Y à l'aide de la F.A. φ_Y . Rappelons aussi l'expression des opérateurs d'espérance conditionnelle E_1 et E_2 . Si $f \in H_2$, $E_1 f \in H_1$ est défini par :

$$(9-1) \quad E_1 f(x) = \int \varphi(x,y) f(y) \varpi_2(dy) = \langle f, \varphi_x \rangle$$

Mais, par définition, la fonction $E_1 f = \int \varphi_Y f(y) \varpi_2(y)$ est dans H_1' (et non pas seulement dans H_1), puisqu'elle s'exprime linéairement à l'aide de la F.A. φ_Y . Il en résulte que E_1 coïncide avec la restriction à H_2 du projecteur π_1 , et $E_1 f$ est le krigeage de f :

$$(9-2) \quad \begin{cases} E_1 f = \pi_1 f & f \in H_2 \\ E_2 g = \pi_2 g & g \in H_1 \end{cases}$$

Comme $E_1 f$ est le krigeage de f , on a, pour tout $y \in \Omega_2$, $\langle E_1 f, \varphi_y \rangle = \langle f, \varphi_y \rangle$. D'après (9-1), on a aussi $\langle E_1 f, \varphi_y \rangle = E_2 E_1 f(y)$, d'où :

$$E_2 E_1 f(y) = \langle E_1 f, \varphi_y \rangle = \langle f, \varphi_y \rangle = \int \langle \varphi_y, \varphi_{y'} \rangle f(y') \varpi_2(dy')$$

Toutefois, ces interprétations prêtent à équivoque, du fait du double rôle imparti à la fonction φ : d'une part, elle représente la densité de la loi $\varphi(x,y) \varpi_1(dx) \varpi_2(dy)$ sur $\Omega_1 \times \Omega_2$, d'autre part, pour chaque x , φ_x est considéré comme un élément de H_2 , et $x \rightarrow \varphi_x$ comme une fonction aléatoire. En particulier, $\varphi(x,y) \varpi_2(dy)$ est la loi conditionnelle en x , d'où les relations :

$$\int \varphi_x \varpi_1(dx) = 1 \quad , \quad \int \varphi_y \varpi_2(dy) = 1$$

qui expriment que φ_x (ou φ_y), en tant que F.A., est entachée de dégénérescence (la covariance centrée $\sigma(x, x') = \langle \varphi_x, \varphi_{x'} \rangle - 1$ vérifie $\int \varpi(dx) \sigma(x, x') \varpi(dx') = 0$). Il y a donc intérêt à généraliser légèrement le problème, de manière à distinguer nettement les deux rôles joués par la fonction φ .

10 - REPRESENTATION D'UNE F.A.

Soit Ω et X deux espaces, P une probabilité sur Ω et ϖ une probabilité sur X , et enfin soit Z une fonction appartenant à $L^2(\Omega \times X, P \otimes \varpi)$. Bien que les deux espaces Ω et X jouent des rôles parfaitement symétriques, il sera commode de considérer les $Y \in L^2(\Omega, P)$ comme des V.A. et les $f \in L^2(X, \varpi)$ comme des fonctions. En particulier, l'application $x \rightarrow Z_x$ de X dans $L^2(\Omega, P)$ sera interprétée comme une F.A. Pour abréger, on posera :

$$\begin{cases} V = L^2(\Omega, \mathcal{A}, P) & \text{(espace des V.A.)} \\ F = L^2(X, \varpi) & \text{(espace des fonctions)} \end{cases}$$

Considérons alors les opérateurs A et A^* définis comme en (9-1) : A applique V dans F , selon la formule :

$$A Y = \int Z(., \omega) Y(\omega) P(d\omega)$$

soit explicitement :

$$(9-3) \quad A Y(x) = E(Z_x Y) = \langle Z_x, Y \rangle$$

Son adjoint A^* applique F dans V selon la relation :

$$(9-4) \quad A^* f = \int f(x) Z_x \omega(dx)$$

Si nous supposons $\sup_x \|Z_x\| < \infty$, A^* est continu et défini sur F entier. Sous la même hypothèse, $|A Y(x)| \leq \|Z_x\| \|Y\| \leq B \|Y\|$, de sorte que $\int |A Y(x)|^2 \omega(dx) \leq B^2 \|Y\|^2$, et A est lui aussi partout défini et continu sur V .

Nous pouvons alors considérer l'opérateur autoadjoint, positif et continu AA^* sur F (ou, aussi bien, A^*A sur V). Il sera commode d'introduire la covariance (non centrée) :

$$C(x, x') = \langle Z_x, Z_{x'} \rangle$$

En particulier, en prenant $Y = Z_x$ dans (9-3) :

$$A Z_x = C_x = C(x, \cdot)$$

Compte tenu de (9-4), ceci entraîne :

$$AA^* f = \int f(x) C_x \omega(dx)$$

de sorte que AA^* est l'opérateur associé au noyau intégral $C(x, x')$. Si Z_x est continu m.q. et si le support S de ω est compact, on sait que l'opérateur AA^* est compact. Il existe alors un système orthonormé de vecteurs propres χ_n associés aux valeurs propres $\lambda_n > 0$ tel que :

$$C(x, x') = \sum_{n=0}^{\infty} \lambda_n \chi_n(x) \chi_n(x')$$

Il est alors facile de voir que les formules :

$$Y_n = \frac{1}{\sqrt{\lambda_n}} A^* \chi_n, \quad \chi_n = \frac{1}{\sqrt{\lambda_n}} A Y_n$$

mettent en bijection les systèmes de vecteurs propres orthonormés associés aux valeurs propres > 0 pour les deux opérateurs A^*A et AA^* . En particulier

$$\begin{aligned} \langle Z_x, Y_n \rangle &= \frac{1}{\sqrt{\lambda_n}} \langle Z_x, A^* \chi_n \rangle = \frac{1}{\sqrt{\lambda_n}} \langle A Z_x, \chi_n \rangle \\ &= \frac{1}{\sqrt{\lambda_n}} \int C(x, x') \chi_n(x') \varpi(dx') = \sqrt{\lambda_n} \end{aligned}$$

On en déduit la représentation de la F.A. Z_x :

$$(9-5) \quad Z_x(\omega) = \sum \sqrt{\lambda_n} Y_n(\omega) \chi_n(x)$$

La seule différence avec la représentation des lois vient de ce que 1 n'est plus nécessairement valeur propre associée à la fonction constante. D'après (9-5), les Y_n forment une base orthonormée de l'espace de Hilbert $H \subset V$ engendré par les Z_x , $x \in X$. Le krigeage d'une V.A. Y quelconque est donc :

$$(9-6) \quad \Pi Y = Y_K = \sum_n \langle Y_n, Y \rangle Y_n$$

De fait, pour tout x , on trouve, d'après (9-5), $\langle Y, Z_x \rangle = \sum_n \sqrt{\lambda_n} \langle Y, Y_n \rangle \chi_n(x)$. Mais $\chi_n = (1/\sqrt{\lambda_n}) A Y_n$, et

$$\chi_n(x) = \frac{1}{\sqrt{\lambda_n}} \langle Y_n, Z_x \rangle$$

Donc $\langle Y, Z_x \rangle = \sum \langle Y, Y_n \rangle \langle Y_n, Z_x \rangle$, et (9-6) en résulte.

Le krigeage n'est donc pas autre chose que la représentation de Y en élément supplémentaire.

Dans tout ceci, le choix de la mesure ω sur X reste arbitraire. (Celui de la probabilité P sur Ω , au contraire, est imposé si Z_x est une F.A. donnée). Montrons qu'il est (parfois) possible de choisir ω de manière à ce que 1 soit fonction propre pour l'opérateur AA^* , soit :

$$\chi_0 = 1$$

En effet, $\chi_0 = 1$ est fonction propre pour AA^* si :

$$\int C(x,y) \omega_0(dy) = \lambda_0$$

$$\int \omega_0(dx) = 1$$

Ces relations expriment que $m^* = \int \omega(dx) Z_x$ est l'estimateur optimal d'une dérive $a_0 = m$. On sait que cet estimateur existe toujours dans $L^2(\Omega, P)$, mais n'admet pas forcément (sauf dans le cas fini) une représentation intégrale au moyen d'une mesure ω_0 . D'autre part, même s'il admet une telle représentation intégrale, la mesure ω_0 n'est pas forcément positive (et alors ω_0 n'est pas une probabilité). Néanmoins, on montre en théorie du potentiel que pour certaines covariances, comme l'exponentielle $\exp \{-|x-y|\}$, ω_0 existe et est une probabilité. Lorsqu'il en est ainsi, il est naturel de munir l'espace X de ω_0 et de travailler dans $L^2(X, \omega_0)$. De plus, quitte à remplacer C par C/λ_0 , i.e. Z_x par $Z_x/\sqrt{\lambda_0}$, on peut supposer que la valeur propre correspondante λ_0 est égale à 1. Le vecteur propre A_0 correspondant pour A^*A est alors

$$A_0 = \int \omega(dx) Z_x$$

c'est-à-dire, justement, l'estimateur optimal m^* .

On peut alors obtenir une interprétation du krigeage à espérance inconnue. Supposons, en effet, que $Z(x)$ vérifie $E(Z_x) = 0$, et que la fonction aléatoire étudiée soit de la forme $X(x) = Z(x) + m$ avec m inconnu (tandis que la covariance centrée $C(x, x')$ est supposée connue). Comme $\int \chi_n(x) \varpi(dx) = 0$ pour $n \geq 1$, on a $\sqrt{\lambda_n} Y_n = \int \chi_n(x) Z_x \varpi(dx) = \int \chi_n(x) X(x) \varpi(dx)$, et les Y_n forment une base orthonormée de l'espace des combinaisons linéaires autorisées. D'autre part $\int \varpi(dx) X(x) = m + Y_0$ est l'estimateur optimal m^* de m .

Ceci étant, soit x_0 un point supplémentaire (i.e. n'appartenant pas au support de la mesure ϖ) et $X(x_0) = Z_{x_0} + m$. On peut représenter Z_{x_0} en "élément supplémentaire", soit

$$Z_{x_0} = \sum \langle Z_{x_0}, Y_n \rangle Y_n$$

ce qui conduit à prolonger χ_n en posant :

$$\chi_n(x_0) = \frac{1}{\sqrt{\lambda_n}} \langle Z_{x_0}, Y_n \rangle$$

et, plus généralement, toute $f \in L^2(X, \varpi)$, en posant

$$f(x_0) = \sum \frac{1}{\sqrt{\lambda_n}} \langle f, \chi_n \rangle \langle Z_{x_0}, Y_n \rangle$$

En particulier, supposons que $f(x)$ soit la réalisation de $X(x)$ pour un $\omega = \omega_0$, soit $f(x) = Z_x(\omega_0) + m$. Pour $n \geq 1$, on trouve

$$\begin{aligned} \langle f, \chi_n \rangle &= \int Z_x(\omega_0) \chi_n(x) \varpi(dx) = \int Z_x(\omega_0) \chi_n(x) \varpi(dx) \\ &= \sqrt{\lambda_n} Y_n(\omega_0) \end{aligned}$$

et pour $n = 0$

$$\langle f, \chi_0 \rangle = m + \int Z_x(\omega_0) \varpi(dx) = m + Y_0(\omega_0)$$

Le prolongement ci-dessus donne donc

$$f(x_0) = m + Y_0 + \sum_{n \geq 1} \langle Z_{x_0}, Y_n \rangle Y_n(\omega_0)$$

Comme $m + Y_0 = m^*$, cette expression est identique à celle que l'on obtient en krigeant sous la condition habituelle

$$\int \lambda(dx) = 1.$$

On peut encore présenter ces mêmes résultats sous une forme plus symétrique (équivalente à celle qu'a utilisée J.P. Orfeuill en partant de la métrique du variogramme). On prend :

$$\varphi(x, y) = C(x, y)$$

covariance de la F.A. ci-dessus, la probabilité ϖ définie par

$$\begin{aligned} \int \varpi(dx) C(x, y) &= 1 \\ \int \varpi(dx) &= 1 \end{aligned}$$

et la loi $\varphi(x, y) \varpi(dx) \varpi(dy)$ sur l'espace produit $X \times Y$. On est cette fois exactement dans les conditions du paragraphe 3, et on trouve

$$C(x, y) = \sum_{n=0}^{\infty} \lambda_n \chi_n(x) \chi_n(y)$$

avec $\lambda_0 = 1$, $\chi_0 = 1$.

Si x_0 est un point supplémentaire, on représente la fonction $C_{x_0} = C(x_0, \cdot)$ en "élément supplémentaire" :

$$C_{x_0} = \sum \langle C_{x_0}, \chi_n \rangle \chi_n$$

et on prolonge χ_n par $\chi_n(x_0) = 1/\lambda_n \langle C_{x_0}, \chi_n \rangle$. Pour une fonction $f(x)$, de même :

$$f(x_0) = \sum \frac{1}{\lambda_n} \langle f, \chi_n \rangle \langle C_{x_0}, \chi_n \rangle$$

Supposons que $f(x) = m + Z_x(\omega_0)$. Pour $n \geq 1$, on a encore $\langle f, \chi_n \rangle = \sqrt{\lambda_n} Y_n(\omega_0)$ et

$$\langle C_{x_0}, \chi_n \rangle = \int \langle Z_{x_0}, Z_x \rangle \chi_n(x) \varpi(dx) = \sqrt{\lambda_n} \langle Z_{x_0}, Y_n \rangle$$

Pour $n = 0$, $\langle f, \chi_n \rangle = m + Y_0 = m^*$. Par suite :

$$f(x_0) = m + \sum_{n \geq 1} \langle Z_{x_0}, Y_n \rangle Y_n(\omega_0)$$

C'est, à nouveau, l'expression du krigeage sous la contrainte $\int \lambda(dx) = 1$.