Fontainebleau/G. N-539

NOTE GEOSTATISTIQUE N° 135

PEUT-ON IMPOSER DES CONDITIONS D'UNIVERSALITE AU KRIGEAGE DISJONCTIF ?

-=-=-=-=-=-=-

G. MATHERON Octobre 1977

NOTE GEOSTATISTIQUE N° 135

PEUT-ON IMPOSER DES CONDITIONS D'UNIVERSALITE AU KRIGEAGE DISJONCTIF ?

-=-=-=-=-=-

Dans l'état actuel des choses, le K.D. semble tributaire d'une hypothèse de stationnarité assez stricte. Je voudrais cependant - en partant des indications de P. Switzer sur l'estimation de l'histogramme - suggérer une voie qui permettrait de soumettre au K.D. certains types de phénomènes non stationnaires. Cette approche est, en un sens, inspirée de la théorie des FAI-k. De fait, on verra que, du point de vue numérique, tout se ramènera, finalement, à imposer les conditions d'universalité usuelles aux coefficients $\lambda^{\alpha,n}$ du krigeage de la composante hermitienne d'ordre n. Mais le but de ces conditions ne sera plus, ici, de filtrer une dérive du type $a_\ell f^\ell(x)$ affectant la V.R. z(x). Il résulte, d'ailleurs, clairement des analyses de P. Delfiner qu'un estimateur de type disjonctif ne peut pas posséder la propriété d'invariance à l'égard d'une dérive de cette forme. Par conséquent, ce qui dérivera dans l'espace sera non pas la V.A. Z(x) elle-même, mais son histogramme $F_x(z) = P(Z(x) \leq z)$ ou, si l'on préfère, l'anamorphose gaussienne Φ_x telle que $Y_x = \Phi_x^{-1}(Z(x))$ soit (censée être) une F.A. gaussienne et stationnaire.

L'idée de faire dériver dans l'espace l'anamorphose Φ_x n'est pas nouvelle. Mais le problème est de procéder à son estimation. De même que l'on impose à la dérive $m(x) = a_\ell f^\ell(x)$ d'appartenir à un espace vectoriel de dimension peu élevée, sans quoi l'estimation des a_ℓ est impossible, de même, ici, nous devons faire choix, pour le domaine de Φ_x, d'un espace fonctionnel pas trop riche, et surtout dont la structure soit en quelque sorte préadaptée au problème de l'estimation. On peut espérer y parvenir en prenant $\Phi_x^{-1} = G^{-1}\ F_x$ avec pour F_x un modèle de la forme

$$F_x(z) = \varpi_x^\ell \quad F_\ell(z)$$

où les ϖ_x^ℓ sont donnés (jouent le rôle des $f^\ell(x)$, avec cependant les conditions $\varpi_x^\ell \geq 0$, $\sum_\ell \varpi_x^\ell = 1$) et où les F_ℓ sont des fonctions de répartition inconnues, à estimer (jouant le rôle des coefficients a_ℓ). Commençons par le cas stationnaire, c'est-à-dire le cas où il y a une seule composante $F_\ell = F$ (et $\varpi_x^\ell = 1$).

1 - LE CAS STATIONNAIRE.

Dans le cas stationnaire, il s'agit d'estimer $F(z)$ à partir des données $z_\alpha = z(x_\alpha)$. P. Switzer propose d'utiliser un estimateur de la forme :

$$(1-1) \qquad F^*(z) = \sum_\alpha \lambda^\alpha \, 1_{z_\alpha \leq z}$$

avec des poids $\lambda^\alpha = \lambda^\alpha(z)$ vérifiant la condition de non biais :

$$(1-1') \qquad \sum \lambda^\alpha = 1$$

et minimisant la variance de $F^*(z)$. (Il s'agit d'une procédure itérative, et la variance que l'on minimise est celle de $F^*(z)$ calculée dans un premier modèle).

Ici, nous utiliserons une classe d'estimateurs plus large que (1-1). Pour estimer $F(z)$, il n'y a pas de raison de se limiter aux seules indicatrices $1_{z_\alpha \leq z}$, on peut espérer améliorer les choses en utilisant aussi les autres indicatrices $1_{z_\alpha \leq z'}$ pour z' différent de z. Mais cela revient à introduire un estimateur disjonctif, du type :

$$(1-2) \qquad F^*(z) = \sum_\alpha f_\alpha(z_\alpha)$$

La condition de non biais (1-1') doit alors être remplacée par une condition d'universalité exprimant que l'on doit avoir $E[F^*(z)] = F(z)$, quelle que soit la loi (inconnue) $F(dz)$, donc :

$$\sum_\alpha \int f_\alpha(\eta)\ F(d\eta) = F(z) \quad , \quad \forall\ F$$

En prenant $F(d\eta) = \delta_{z_o}(d\eta)$, on voit sans peine que cette condition equivaut à :

$$\sum_\alpha f_\alpha(z_o) = 1_{z \geq z_o} \quad , \quad \forall\ z_o \in R$$

Pour plus de clarté, formulons ce problème dans un cadre un peu plus général. Soit φ une fonction quelconque, et soit à estimer l'espérance $E[\varphi(z)] = \int \varphi(z)\ F(dz)$, qui est une caractéristique de la loi (inconnue) F. Nous cherchons un estimateur de la forme

$$E^*_\varphi = \sum_\alpha f_\alpha(Z_\alpha)$$

vérifiant la condition $E(E^*_\varphi) = E(\varphi)$ quelle que soit la loi F, soit :

$$\sum_\alpha \int f_\alpha(z)\ F(dz) = \int \varphi(z)\ F(dz) \quad , \quad \forall\ F$$

Notre estimateur doit donc vérifier :

$$\sum_\alpha f_\alpha(z) = \varphi(z) \qquad \forall\ z \in R$$

Nous devons ensuite, sous cette condition de non biais, minimiser Var E^*_φ dans le cadre d'un premier modèle.

Comme premier modèle, nous faisons choix d'une anamorphose gaussienne $Z(x) = \Phi(Y(x))$, obtenue à partir d'une première estimation de l'histogramme, par exemple l'estimation "naïve", et nous admettons que la F.A. $Y(x)$ est stationnaire et gaussienne (il suffit, en fait, que les lois à deux variables soient gaussiennes). Par hypothèse, dans ce modèle, $E(Y(x)) = 0$, Var $Y(x) = 1$, et les méthodes habituelles permettent d'estimer la covariance

$$\rho(h) = E[Y(x)\ Y(x+h)]$$

Dans l'anamorphose Φ, on a $\varphi(z) = \psi(y)$ avec $\psi = \varphi \circ \Phi$, et il s'agit d'estimer $E\ \psi(Y)$. Dans le modèle, évidemment, cette espérance est $\int \psi(y)\ G(dy)$, où G est la loi normale (0,1), mais justement nous voulons améliorer cette première estimation. Nous prendrons donc un estimateur disjonctif :

(1-3) $$E^*_\varphi = \sum_\alpha f_\alpha(Y_\alpha)$$

avec la condition $E(E^*_\varphi) = E(\psi(Y))$ quelle que soit la loi de Y : ce qui donne, comme ci-dessus :

(1-4) $$\sum_\alpha f_\alpha(y) = \psi(y)$$

Avec les développements hermitiens :

$$\psi(y) = \sum \frac{\psi_n}{n!} H_n(y)$$
$$f_\alpha(y) = \sum \frac{f_{n,\alpha}}{n!} H_n(y)$$

nous obtenons la condition équivalente :

(1-4') $$\psi_n = \sum_\alpha f_{n,\alpha} \quad , n = 0, 1, 2, \ldots$$

Il reste maintenant à introduire une condition d'optimalité. Alors que la condition (1-4), ou (1-4') est, en un sens, indépendante du choix du premier modèle (puisqu'elle exprime une condition de non biais quelle que soit la "vraie" loi inconnue de Y ou de Z), la seconde condition est liée à ce choix : elle exprime que la variance de l'estimateur (1-3) est minimale dans ce modèle, sous la contrainte (1-4). On tombe alors que le système de krigeage habituel :

$$\begin{cases} \sum_\alpha f_{n,\alpha}\, \rho^n_{\alpha\beta} = C_n \\ \sum_\alpha f_{n,\alpha} = \psi_n \end{cases}$$

En particulier, si $\psi(y) = H_n(y)$ est le polynome d'Hermite d'ordre n, on aura un estimateur H^*_n de $E(H_n(Y)]$ de la forme

$$H^*_n = \sum_\alpha \lambda^\alpha_n H_n(Y_\alpha)$$

c'est-à-dire dépendant uniquement des polynomes du même degré $H_n(Y_\alpha)$, puisqu'ici $\psi_m = 0$ et donc $f_{m,\alpha} = 0$ pour $m \neq n$. Les poids sont donnés par le système familier :

$$(1\text{-}5) \qquad \begin{cases} \sum\limits_{\alpha} \lambda_n^{\alpha}\, \rho_{\alpha\beta}^{n} = \mu_n \\ \sum\limits_{\alpha} \lambda_n^{\alpha} = 1 \end{cases}$$

Dans le cas général d'une fonction $\varphi(z) = \psi(y)$, l'estimateur sera :

$$E_{\varphi}^{*} = \sum \frac{\psi_n}{n!} \; H_n^{*}$$

Autrement dit, ici encore, on krige séparément chacune des composantes hermitiennes, mais avec une condition $\sum \lambda_n^{\alpha} = 1$. En particulier, si $\varphi(z) = 1_{z \le z_o} = 1_{y \le y_o}$ (avec $z_o = \Phi(y_o)$), on utilise le développement :

$$1_{y \le y_o} = G(y_o) + \sum_{n \ge 1} \frac{H_n(y)}{n!} \; H_{n-1}(y_o)\; g(y_o)$$

où g et G representent la densité et la fonction de répartition de la loi normale réduite, et il vient donc :

$$(1\text{-}6) \qquad F^*(z_o) = G(y_o) + \sum_{n \ge 1} \frac{H_n^{*}}{n!} \; H_{n-1}(y_o)\; g(y_o)$$

De cet estimateur, on déduit une nouvelle anamorphose gaussienne, et on peut réitérer la procédure jusqu'à stabilisation.

(Je n'aborde pas, pour l'instant, la difficulté très réelle que l'on risque de rencontrer si l'estimateur $F^*(z_o)$ ainsi obtenu n'est pas une fonction croissante de z_o : il faudra voir, dans les applications, si cette difficulté peut être négligée en première approximation, quitte à donner un léger coup de pouce, ou s'il est vraiment nécessaire d'introduire le formalisme beaucoup plus lourd de la projection sur les convexes fermés).

2 - K.D. INTRINSEQUE A L'ORDRE 0.

L'estimateur ci-dessus est l'équivalent disjonctif de l'estimateur linéaire optimal m^* de l'espérance $m = E(Z(x))$ d'une FAST.

La relation

$$Z^*(x) = m^* + \lambda_K^\alpha \, (Z_\alpha - m^*)$$

exprimant le krigeage $Z^*(x)$ avec condition $\sum \lambda^\alpha = 1$ en fonction de l'estimateur optiaml m^* de l'espérance et des poids λ_K^α du krigeage sans condition de non biais, va également avoir son équivalent disjonctif.

Soit, en effet, à estimer l'intégrale

$$I = \int \varphi(Z(x)) \, \mu(dx) = \int \psi(Y_x) \, \mu(dx) \qquad \left(\int \mu(dx) = 1\right)$$

Nous cherchons un estimateur disjonctif :

$$I^* = \sum_\alpha f_\alpha \, (Y_\alpha)$$

avec la condition de non biais : $E(I) = E(I^*)$ quelle que soit la loi F de Y (ou de Z). Si $\int \mu(dx) = 1$, cette condition s'exprime comme ci-dessus par

$$\psi(y) = \sum_\alpha f_\alpha(y)$$

ou, avec les coefficients de développement hermitien :

$$\psi_n = \sum_\alpha f_{n,\alpha}$$

Il reste ensuite à minimiser (dans le modèle) la variance d'estimation Var $(I-I^*)$, compte tenu de cette condition. Ici encore on constate que l'on peut traiter séparément chacune des composantes hermitiennes, et on trouve :

(2-1) $$I^* = \sum \frac{\psi_n}{n!} \, \tilde{H}_n$$

où $\tilde{H}_n$ est l'estimateur de $\int H_n(Y_x) \, \mu(dx)$ soit

(2-2) $$\tilde{H}_n = \tilde{\lambda}_n^\alpha \, H_n(Y_\alpha)$$

avec des poids $\tilde{\lambda}_n^\alpha$ donnés par le système du krigeage ordinaire :

$$(2\text{-}3) \qquad \begin{cases} \tilde{\lambda}_n^\alpha \; \rho_{\alpha\beta}^n = \int \rho_{\beta x}^n \, \mu(dx) + \tilde{\mu}_n \\ \sum_\alpha \tilde{\lambda}_n^\alpha = 1 \end{cases}$$

Par exemple, pour estimer <u>l'histogramme régional</u>

$$F_\mu(z_o) = \int 1_{Z_x \leq z_o} \, \mu(dx)$$

on utilisera l'estimateur

$$(2\text{-}4) \qquad \tilde{F}_\mu(z_o) = G(y_o) + \sum_{n\geq 1} \frac{\tilde{H}_n}{n!} \; H_{n-1}(y_o) \; g(y_o)$$

qui n'est évidemment pas identique à (1-6). La relation entre ces deux estimateurs est donnée par le théorème habituel d'additivité. Soit

$$H_n = \lambda_{K,n}^\alpha \; H_n(Y_\alpha)$$

le krigeage (sans condition de non biais) de l'intégrale $\int H_n(Y_x) \; \mu(dx)$ donné par le système :

$$\lambda_{Kn}^\alpha \; \rho_{\alpha\beta}^n = \int \rho_{\beta x}^n \; \mu(dx)$$

Alors :

$$\tilde{H}_n = H_n^* + \lambda_{Kn}^\alpha \; (H_n(Y_\alpha) - H_n^*)$$

$$= (1 - \sum_\alpha \lambda_{Kn}^\alpha) \; H_n^* + H_n$$

La structure du terme correctif est exactement la même que dans la théorie linéaire, et le théorème d'additivité relatif aux variances s'applique de la même façon.

Du point de vue des applications, le K.D. intrinsèque d'ordre 0 (i.e. avec condition $\sum_\alpha \lambda_n^\alpha = 1$) doit sans doute présenter davantage de robustesse vis-à-vis du choix du modèle (de l'anamorphose) que la procédure que nous utilisons actuellement. De même, l'estimateur (1-6) de l'histogramme doit sans doute permettre d'améliorer le choix de l'anamorphose gaussienne.

3 - EXEMPLES SIMPLES.

Supposons $\rho_{\alpha\beta} = 0$ pour $\alpha \neq \beta$ (distance entre points expérimentaux supérieure à la portée), sans que cela implique $\rho_{\alpha x} = 0$ pour tout x. Si N est le nombre des points expérimentaux, (1-5) se réduit à $\lambda_n^\alpha = 1/N$ $(\forall \alpha, \forall n)$, donc

$$H_n^* = \frac{1}{N} \sum_\alpha H_n(Y_\alpha)$$

et par suite

$$F^*(z_o) = \frac{1}{N} \sum_\alpha 1_{z_\alpha \leq z}$$

L'estimateur naïf est ici optimal. Par contre, pour le krigeage sans condition, nous trouvons

$$\lambda_{Kn}^\alpha = \int \rho_{\alpha x}^n \mu(dx) \qquad (n \geq 1)$$

et, avec $H_n = \lambda_{Kn}^\alpha H_n(Y_\alpha)$, $n \geq 1$, et $H_0 = 1$

$$F_K(z_o) = G(y_o) + \sum_{n\geq 1} \frac{H_n}{n!} H_{n-1}(y_o)\, g(y_o)$$

$$= G(y_o) + \sum_{n\geq 1} \frac{H_{n-1}(y_o)\, g(y_o)}{n!} \sum_\alpha H_n(Y_\alpha) \int \rho_{\alpha x}^n \mu(dx)$$

Or (dans le modèle)

$$G(y_o) + \sum_{n\geq 1} \frac{H_{n-1}(y_o)\, g(y_o)}{n!} \rho_{\alpha x}^n H_n(Y_\alpha) = P(Y_x \leq y_o | Y_\alpha)$$

est la probabilité conditionnelle de $Y_x \leq y_o$ lorsque Y_α est donné. Il vient donc ainsi

$$F_K(z_o) = \sum_\alpha \int P(Y_x \leq y_o | Y_\alpha)\, \mu - (N-1)\, G(y_o)$$

En ce qui concerne l'estimateur disjonctif intrinsèque d'ordre 0 nous aurons, d'après (2-5) :

$$\tilde{H}_n = (1 - \sum_\alpha \lambda_{Kn}^\alpha)\, H_n^* + H_n$$

et donc :

$$\tilde{F}_\mu(z_o) = G(y_o) + \sum_{n\geq 1} \frac{H_{n-1}(y_o)\ g(y_o)}{n!} [(1 - \sum_\alpha \lambda^\alpha_{Kn})\ H^*_n + H_n]$$

$$= F^*(z_o) + F_K(z_o) - G(y_o) - \sum_{n\geq 1} \frac{H_{n-1}(y_o)\ g(y_o)}{n!}\ H^*_n \sum_\alpha \lambda^\alpha_{K,n}$$

Or, on a :

$$\sum_{n\geq 1} \rho^n_{\alpha x} \frac{H_{n-1}(y_o)\ g(y_o)}{n!}\ H^*_n =$$

$$\frac{1}{N} \sum_\gamma P(Y_x \leq y_o | Y_a = y_\gamma) - G(y_o)$$

et par suite

$$\sum_{n\geq 1} \frac{H_{n-1}(y_o)\ g(y_o)}{n!}\ H^*_n \sum_\alpha \lambda^\alpha_{Kn} = \frac{1}{N} \sum_{\alpha,\gamma} \int \mu(dx)\ P(Y_x \leq y_o | Y_\alpha = Y_\gamma)$$
$$- N\ G(y_o)$$

et finalement :

$$(3\text{-}1) \left\{ \begin{aligned} \tilde{F}_\mu(z_o) = F^*(z_o) &+ \sum_\alpha \int \mu(dx)\ P(Y_x \leq y_o | Y_\alpha = y_\alpha) \\ &- \frac{1}{N} \sum_{\alpha,\gamma} \int \mu(dx)\ P(Y_x \leq y_o | Y_\alpha = y_\gamma) \end{aligned} \right.$$

formule curieuse ou, à côté des probabilités conditionnées par $Y_\alpha = y_\alpha$, apparaissent les valeurs moyennes de ces mêmes probabilités prises conditionnellement lorsque la variable Y_α prend l'une ou l'autre des valeurs numériques y_γ affectées aux autres variables.:

Un examen plus attentif révèle des simplifications. Prenons, pour abréger, $\mu = \delta_x$, c'est-à-dire le cas de l'estimateur de l'histogramme au point x (conditionné par les z_α) :

$$\tilde{F}_x(z_o) = F^*(z_o) + \sum_\alpha P(Y_x \leq y_o | Y_\alpha = y_\alpha)$$
$$- \frac{1}{N} \sum_{\alpha,\gamma} P(Y_x \leq y_o | Y_\alpha = y_\gamma)$$

Comme F* est l'estimateur naïf, on a

$$F^*(dz) = \frac{1}{N} \sum_{\gamma} \delta_{z_\gamma}(dz)$$

et par suite, dans le modèle, pour tout α :

$$\frac{1}{N} \sum_{\gamma} P(Y_x \leq y_o | Y_\alpha = y_\gamma) = E[P(Y_x \leq y_o | Y_\alpha] = P(Y_x \leq y_o) = F^*(z_o)$$

Il vient donc :

$$\tilde{F}_x(z_o) = \sum_{\alpha} P(Y_x \leq y_o | Y_\alpha = y_\alpha) - (N-1) F^*(z_o)$$

De plus, dans le modèle, Y_x est indépendant de Y_α dès que la distance $|x-x_\alpha|$ dépasse la portée. On a donc $P(Y_x \leq y_o | Y_\alpha = y_\alpha) = P(Y_x \leq y_o) = F^*(z_o)$ pour tous les termes de la sommation en α, à la seule exception des indices α tels que $|x-x_\alpha| <$ portée.

Supposons, pour simplifier, qu'il n'y ait, pour chaque $x_o \in S$ qu'un seul point x_{α_o} expérimental ($\alpha_o = \alpha_o(x_o)$) vérifiant cette condition de proximité. Alors, il reste simplement :

$$\tilde{F}_{x_o}(z_o) = P(Y_{x_o} \leq y_o | Y_{\alpha_o} = y_{\alpha_o})$$

c'est-à-dire la loi de $Z(x_o)$ conditionnée par $Z(x_{\alpha_o})$

A titre de <u>2ème exemple</u>, prenons maintenant

$$\rho(h) = e^{-a|h|}$$

et plaçons-nous (à 1 dimension) dans le cas où $Z(x)$ et $Y(x)$ sont connues sur l'intervalle $(-R, R)$. Pour former l'estimateur optimal $F^*(z_o)$ de l'histogramme, calculons d'abord H_n^*. Comme la F.A. $H_n(Y_x)$ admet la covariance $e^{-na|h|}$, l'estimateur optimal H_n^* de son espérance est :

$$H_n^* = \frac{n\,a\,R}{1+n\,a\,R}\,\overline{H}_n + \frac{1}{1+n\,a\,R}\left(\frac{H_n(Y_{-R}) + H_n(Y_R)}{2}\right)$$

(avec évidemment $\overline{H}_n = (1/2\,R) \int_{-R}^{R} H_n(Y_x)dx$)

Il vient donc :

$$F*(z_o) = G(y_o) + \sum_{n\geq 1} \frac{H_{n-1}(y_o)\ g(y_o)}{n!}\ H_n^* \tag{3-2}$$

avec la valeur ci-dessus de H_n^*, que nous ré-écrivons :

$$H_n^* = \overline{H}_n + \frac{1}{1+n\ a\ R}\left(\frac{H_n(Y_{-R}) + H_n(Y_R)}{2} - \overline{H}_n\right) \tag{3-3}$$

Rn substituant cette expression dans (3-2), le terme $\overline{H}_n$ donnera sans problème l'histogramme "naïf" (qui est ici l'histogramme régional):

$$\overline{F}(z_o) = \frac{1}{2R}\int_{-R}^{R} 1_{Z_x\leq z_o}\ dx$$

Le second terme de (3-3) s'obtient à partir d'expressions de la forme

$$\frac{1}{1+n\ a\ R}\ H_n(Y_x)$$

Désignons par $G_\rho(y_o|y_x)$ la probabilité conditionnelle de $Y \leq y_o$ lorsque $Y_x = y_x$ est fixé, Y et Y_x étant gaussiennes réduites, avec le coefficient de corrélation ρ. Alors :

$$G_\rho(y_o|y_x) = G(y_o) + \sum_{n\geq 1} \rho^n \frac{H_{n-1}(y_o)\ g(y_o)}{n!}\ H_n(y_x)$$

Par suite :

$$\frac{1}{\alpha}\int_0^1 \rho^{1/\alpha-1}\ G_\rho(y_o|y_x)\ d\rho = G(y_o) + \sum_{n\geq 1} \frac{1}{1+\alpha n}\ \frac{H_{n-1}(y_o)\ g(y_o)}{n!}\ H_n(y_x)$$

La substitution dans (3-2) du second terme de (3-3) donnera donc :

$$\frac{1}{\alpha}\int_0^1 \rho^{1/\alpha-1}\ d\rho\left[\frac{G_\rho(y_o|y_{-R}) + G_\rho(y_o|y_{+R})}{2} - \frac{1}{2R}\int_{-R}^{R} G_\rho(y_o|y_x)\ dx\right]$$

avec $\alpha = a\ R$

D'où l'expression cherchée de l'estimateur optimal de l'histogramme :

$$
(3\text{-}4)\quad \begin{cases} F^*(z_o) = \dfrac{1}{2R} \displaystyle\int_{-R}^{R} 1_{Z_x \le z_o}\, dx \; + \\[2ex] + \dfrac{1}{\alpha} \displaystyle\int_o^1 \rho^{1/\alpha - 1}\, d\rho \left[\dfrac{G_\rho(y_o|y_{-R}) + G_\rho(y_o|y_R)}{2} - \dfrac{1}{2R} \displaystyle\int_{-R}^{R} G_\rho(y_o|y_x)\, dx \right] \end{cases}
$$

4 - VERS UN K.D. UNIVERSEL A L'ORDRE k.

Ce qui précède constitue la transposition disjonctive du krigeage universel à l'ordre 0, et de l'estimation optimale de l'espérance m. Nous remarquons que le rôle de la constante inconnue m est joué ici par une fonction de répartition inconnue $F(z)$. Est-il possible, de la même façon, d'obtenir une version disjonctive du krigeage universel d'ordre k ? A l'ordre 0, nous sommes arrivés à la conclusion que l'on pouvait kriger séparément chacune des composantes hermitiennes en imposant aux coefficients λ_n^α la condition habituelle d'ordre 0, soit $\sum \lambda_n^\alpha = 1$. A l'ordre k, nous pouvons espérer aboutir à la même conclusion, c'est-à-dire au système :

$$
\begin{cases} \lambda_n^\alpha\, \rho_{\alpha\beta}^n = \rho_{\beta x}^n + \mu_{n,\ell}\, f_\beta^\ell \\ \lambda_n^\alpha\, f_\alpha^\ell = f_x^\ell \end{cases}
$$

Il reste à dégager la signification de ce système.

Comme je l'ai indiqué en introduction, l'idée de départ consiste à supposer que la loi $F_x(z)$ de $Z(x)$ au point x est de la forme :

$$
F_x(z) = \varpi_x^\ell\, F_\ell(z)
$$

Les ϖ_x^ℓ sont des poids connus (choisis par nous), qui joueront le rôle des fonctions de base $f^\ell(x)$. Ils dépendent évidemment du point x. Les F_ℓ seront des lois inconnues, qui joueront le rôle des coefficients inconnus a_ℓ , et qu'il s'agira d'estimer. En ce qui concerne les ϖ_x^ℓ, ils devront vérifier, en principe, des conditions de positivité : $\varpi_x^\ell \ge 0$ (dont je ne m'occuperai pas pour l'instant) et aussi, évidemment, la condition

$$(4\text{-}1) \qquad \sum_{\ell} \varpi_x^{\ell} = 1$$

Cette condition (4-1) n'est d'ailleurs limitative qu'en apparence. Elle nous interdit d'identifier ϖ_x^{ℓ} et le monome f_x^{ℓ}, mais non de supposer que l'espace vectoriel engendré par les fonctions ϖ^{ℓ} est identique à l'espace des polynomes de degré $\leq k$.

Désignons par m_{ℓ} l'espérance (inconnue) associée à la loi F_{ℓ}. Alors, dans ce modèle

$$E[Z(x)] = m_{\ell}\, \varpi_x^{\ell}$$

Si nous voulons que la dérive soit un polynome du type $a_{\ell} f_x^{\ell}$, il faut donc bien que les espaces $\{m_{\ell}\, \varpi^{\ell}\}$ et $\{a_{\ell}\, f^{\ell}\}$ coïncident. Cela conduit donc à adopter pour ϖ_x une expression de la forme :

$$(4\text{-}2) \qquad \varpi_x^{\ell} = C_s^{\ell}\, f_x^{s}$$

avec une matrice C_s^{ℓ} assurant la condition (4-1), c'est-à-dire vérifiant

$$(4\text{-}3) \qquad \sum_{\ell} C_o^{\ell} = 1 \quad , \quad \sum_{\ell} C_s^{\ell} = 0 \qquad (s \neq 1)$$

(l'indice o est associé à la fonction constante $f^o = 1$). Le choix de la matrice C_s^{ℓ} n'a du reste pas d'importance. (sous réserve de vérifier (4-3)), pourvu qu'elle soit régulière. Si, en effet, nous effectuons une transformation linéaire sur les ϖ^{ℓ}, les F_{ℓ} (et leurs estimateurs) se modifieront de manière contravariante, de façon à assurer l'invariance tensorielle de l'expression $F_x = \varpi_x^{\ell}\, F_{\ell}$ (ou de son estimateur).

Estimation optimale des F_{ℓ}^{*}.

Nous nous plaçons dans les conditions d'une procédure itérative : une première anamorphose Φ_x (dépendant ou non du point x) nous a donné un "premier modèle"

$$Z(x) = \Phi_x(Y_x)$$

où Y_x est (considéré comme) gaussienne et stationnaire. Pour passer au deuxième modèle, plus fin, nous devons former une estimation de la loi $F_x(y)$ de $Y(x)$, loi dépendant du point x selon une relation (supposée être) de la forme

$$F_x(y) = \varpi_x^\ell F(y)$$

avec des ϖ_x^ℓ connus et des $F_\ell(y)$ inconnus.

Plus généralement, si $\varphi(y)$ est une fonction donnée, nous désirons estimer $E[\varphi(Y_x)] = \int \varphi(y)\, F_x(dy)$. Pour cela, nous allons former un estimateur de type disjonctif :

$$E_x^* \varphi = \sum_\alpha f_\alpha(Y_\alpha) \tag{4-4}$$

et imposer la condition d'universalité : quelles que soient les lois (inconnues) F_ℓ, on doit avoir $E(E_x^* \varphi) = E[\varphi(Y_x)]$, c'est-à-dire

$$\sum_{\alpha,\ell} \varpi_\alpha^\ell \int f_\alpha(y)\, F_\ell(dy) = \varpi_x^\ell \int \varphi(y)\, F_\ell(dy)$$

Ceci devant avoir lieu quelles que soient les lois F_ℓ , notre condition d'universalité s'écrit :

$$\sum_\alpha \varpi_\alpha^\ell\, f_\alpha(y) = \varpi_x^\ell\, \varphi(y) \tag{4-5}$$

pour tout indice ℓ et tout $y \in \mathbb{R}$.

Une autre manière d'exprimer ceci consiste à introduire les développements

$$f_\alpha = \sum_n \frac{f_{n,\alpha}}{n!} H_n \quad , \quad \varphi = \sum \frac{\varphi_n}{n!} H_n$$

et a imposer les conditions :

(4-6) $$\sum_{\alpha} \varpi_{\alpha}^{\ell} f_{n,\alpha} = \varpi_x^{\ell} \varphi_n \qquad (\forall \ell, \forall n)$$

Ces conditions étant imposées, il reste à choisir les coefficients $f_{n,\alpha}$ qui minimisent (dans le premier modèle) la variance de cet estimateur. Ici encore, on constate que les différentes composantes hermitiennes se laissent traiter séparément.

Pour estimer l'espérance de $H_n(Y_x)$, on forme donc un estimateur du type

$$H_n^*(x) = \lambda_n^{\alpha} H_n(Y_{\alpha})$$

vérifiant la condition (4-6), soit

$$\lambda_n^{\alpha} \varpi_{\alpha}^{\ell} = \varpi_x^{\ell}$$

c'est-à-dire la condition d'universalité, et minimisant, sous cette condition, la variance de H_n^*. On obtient ainsi le système classique de l'estimation optimale de la dérive au point x :

(4-7) $$\begin{cases} \lambda_n^{\alpha} \rho_{\alpha\beta}^{n} = \mu_{n\ell} \varpi_{\beta}^{\ell} \\ \lambda_n^{\alpha} \varpi_{\alpha}^{\ell} = \varpi_x^{\ell} \end{cases}$$

Dans le cas d'une fonction $\varphi = \sum \frac{\varphi_n}{n!} H_n$, on trouve de même

(4-8) $$E_x^* \varphi = \sum_n \frac{\varphi_n}{n!} H_n^*(x)$$

D'après la structure du système (4-7), les λ_n^{α} dépendent linéairement des ϖ_x^{ℓ} qui figurent au second membre. On a donc, exactement comme en krigeage universel :

(4-9) $$H_n^*(x) = H_{n,\ell}^* \varpi_x^{\ell}$$

avec :

$$(4\text{-}10) \qquad \begin{cases} H^*_{n,\ell} = \lambda^{\alpha}_{n,\ell} \; H_n(Y_\alpha) \\ \lambda^{\alpha}_{n\ell} \; \rho^n_{\alpha\beta} = \mu_{n,\ell s} \; \varpi^s_\beta \\ \lambda^{\alpha}_{n,\ell} \; \varpi^s_\alpha = \delta^s_\ell \end{cases}$$

et aussi (dans le modèle) :

$$(4\text{-}11) \qquad \frac{1}{n!} < H^*_{n,\ell} \, , \, H^*_{ms} > = \delta_{nm} \, \mu_{n,\ell s}$$

En substituant (4-9) dans (4-8), nous mettons notre estimateur sous la forme :

$$E^*_x \, \varphi = \varpi^\ell_x \int \varphi(y) \; F^*_\ell(dy)$$

avec une (pseudo)-loi F^*_ℓ définie par :

$$\int \varphi(y) \; F^*_\ell(dy) = \sum_n \frac{\varphi_n}{n!} \; H^*_{n,\ell}$$

c'est-à-dire :

$$(4\text{-}12) \qquad F^*_\ell(dy) = \sum_n \frac{H^*_{n,\ell} \, H_n(y)}{n!} \quad g(y) \, dy$$

ou encore, sous forme de fonction de répartition :

$$(4\text{-}12') \qquad F^*_\ell(y) = G(y) + \sum_{n\geq 1} \frac{H^*_{n,\ell}}{n!} \; H_{n-1}(y) \, g(y)$$

Tel est l'estimateur optimal de la loi F_ℓ , que l'on peut aussi obtenir directement en prenant $\varphi = 1_{y \leq y_o}$ dans (4-8).

<u>Itération</u>. N'oublions pas qu'il s'agit d'une méthode itérative. Partis d'une première anamorphose gaussienne $Z_x = \Phi_x(Y_x)$, attribuant par construction la loi G à Y_x, nous obtenons maintenant l'estimation plus fine $F^*_x = \varpi^\ell_x \; F^*_\ell$ de cette même loi, a partir de laquelle nous allons définir le deuxième modèle en prenant :

$$P(Z_x \leq z) = P(Y_x \leq \Phi^{-1}_x(z)) = (F^*_x \circ \Phi^{-1}_x)(z)$$

Par conséquent, notre deuxième anamorphose gaussienne sera $Z_x = \tilde{\Phi}_x$, avec

$$\tilde{\Phi}_x = \Phi_x \circ (F_x^*)^{-1} \circ G$$

On peut aussi passer directement des Y_x, tels que $Z_x = \Phi_x(Y_x)$ à leurs nouvelles versions, i.e. les $\tilde{Y}_x$ tels que $Z_x = \tilde{\Phi}_x(\tilde{Y}_x)$, en posant directement

$$\tilde{Y}_x = (G^{-1} \circ F_x^*)(Y_x)$$

Il reste un point délicat : du fait que rien ne nous garantit que l'expression (4-12) est toujours ≥ 0 , il se peut que les $\varpi_x^\ell F_\ell^*$ ne soient pas tous des fonctions croissantes, et donc que $\tilde{\Phi}_x$ ne soit pas une vraie anamorphose. Seule la pratique nous dira si les écarts éventuels peuvent être négligés, ou rectifiés par quelque artifice à imaginer.

Le choix de la première anamorphose.

Dans le cas d'ordre 0, on peut définir la première anamorphose à partir de l'estimateur "naïf" de l'histogramme (ou, éventuellement, d'un estimateur plus élaboré déjà, par exemple un estimateur attribuant à chaque indicatrice $1_{Z_\alpha \leq z_0}$ le poids λ^α de Z_α dans le krigeage de la moyenne spatiale). Dans le cas présent $k > 0$, on peut encore obtenir un estimateur "naïf" de l'histogramme F_x en appliquant une procédure de moindres carrés aux indicatrices $1_{Z_\alpha \leq z_0}$.

Pour légitimer cette démarche, considérons l'exemple simple ou, dans le premier modèle, on aurait :

$$\rho_{\alpha\beta} = 0 \quad \text{pour} \quad \alpha \neq \beta$$

Dans ce cas, les équations (4-10) donnent la même valeur $\lambda_{n\ell}^\alpha = \lambda_\ell^\alpha$, où les λ_ℓ^α sont les coefficients des moindres carrés : explicitement, la matrice $\mu_{\ell s}$ est l'inverse de la matrice

$$\nu^{\ell s} = \sum_\alpha \varpi_\alpha^\ell \varpi_\alpha^s$$

et $$\lambda_\ell^\alpha = \mu_{\ell s} \; \varpi_\alpha^s$$

On a donc ici

$$H_n^*(x) = \varpi_x^\ell \; \lambda_\ell^\alpha \; H_n(Y_\alpha)$$

et, plus généralement, pour toute fonction φ :

$$E_x^* \varphi = \varpi_x^\ell \; \lambda_\ell^\alpha \; \varphi(Y_\alpha)$$

En particulier :

$$F_x^*(y_o) = \varpi_x^\ell \; \lambda_\ell^\alpha \; 1_{y_\alpha \leq y_o}$$

Il s'agit bien d'une procédure de moindres carrés effectués sur les indicatrices des $1_{y_\alpha \leq y_o}$, et son résultat, en ce qui concerne la loi de Z_x lui-même, ne dépend pas de l'anamorphose Φ_x initiale : on trouve, en effet, comme estimateur de $P(Z_x \leq z_o)$:

$$P^*(Z_x \leq z_o) = \varpi_x^\ell \; \lambda_\ell^\alpha \; 1_{z_\alpha \leq z_o} \tag{4-13}$$

c'est-à-dire le résultat de la procédure des moindres carrés effectuée directement sur les indicatrices $1_{z_\alpha \leq z_o}$ de la V.R. $z(\underline{x})$ elle-même. Il est donc assez indiqué de choisir cet estimateur (4-13) de la loi de Z_x pour définir l'anamorphose initiale $\tilde{\Phi}_x$.

Notons aussi qu'en raison de l'invariance tensorielle, cet estimateur naïf (comme d'ailleurs tous les autres estimateurs) ne dépend pas du choix des fonctions ϖ_x^ℓ individuelles, mais seulement de l'espace vectoriel qu'elles engendrent. En particulier, si cet espace vectoriel est celui des polynomes de degré $\leq k$, $P^*(Z_x \leq z_o)$ est le polynome de degré k ajusté par moindres carrés aux indicatrices $1_{z_\alpha \leq z_o}$.

Cette remarque s'applique aussi bien aux $H_n^*(x)$ et aux $E_x^* \varphi$ des relations (4-8) et (4-9) : autrement dit, dans les équations (4-10), on peut remplacer les ϖ_x^ℓ par les f_x^ℓ sans modifier le moins du monde les estimateurs relatifs à un point donné. Ainsi, il n'est pas réellement nécessaire de se donner explicitement les ϖ_x^ℓ . Par contre, évidemment, il faut vérifier que les F_x^* sont des fonctions croissantes pour chaque x.

Or, c'est là une condition très sévère. Le seul cas où nous soyons sûrs qu'elle est automatiquement vérifiée est celui où les fonctions ϖ_x^ℓ sont les indicatrices d'une partition (S_ℓ) du champ total S : avec ce choix des fonctions ϖ_x^ℓ , il est facile de voir que l'on est conduit à traiter séparément, et indépendamment des autres, chacune des composantes S_ℓ comme si le phénomène y était stationnaire. En particulier, pour $x \in S_\ell$, l'estimateur $F_x^*(z_o)$ des moindres carrés est identique à l'estimateur naïf F_ℓ^* dans S_ℓ . En dehors de ce cas trivial, pour tout autre choix des fonctions ϖ^ℓ,on peut déduire du théorème classique caractérisant les projecteurs positifs, que la condition de croissance de $F_x^*(z)$ ne sera jamais vérifiée. Naturellement, si les $F_x^*(z)$ diffèrent relativement peu de fonctions croissantes, on pourra s'en accomoder en pratique (moyennant un léger coup de pouce). Mais on peut craindre qu'il n'en soit pas toujours, ni même souvent, ainsi. Le contre-exemple suivant permet de comprendre pourquoi.

Un contre-exemple.

Soit, sur l'intervalle $(-R, R)$, la variable régionalisée $z(x) = x$ (ou, plus généralement, une dérive linéaire pure). Prenons $\varpi_x^o = 1$, $\varpi_x^1 = x$. Sur $(-R, R)$, la droite des moindres carrés ajustée à une fonction $\varphi(x)$ a pour expression

$$\varphi^*(x) = \frac{1}{2R}\int_{-R}^{R}\varphi(y)dy + \frac{3}{2R^3}\; x\int_{-R}^{R} y\,\varphi(y)dy$$

En appliquant ceci à $\varphi(x) = 1_{x \leq z_o}$, nous obtenons l'estimateur (4-13) de l'histogramme au point x :

$$F_x^*(z_o) = \frac{z_o + R}{2R} + \frac{3x}{4\,R^3}\,(z_o^2 - R^2) \qquad (-R \leq z_o \leq R)$$

La densité associée est donc

$$f_x^*(z) = \frac{1}{2R} + \frac{3\;xz}{2\;R^3}$$

Elle n'est positive que pour $3\,xz > -R^2$. Par exemple, au point $x = -R$, elle est $(1/2R)(1 - 3\,z/R)$, donc < 0 pour $z > R/3$. Seuls les points x de l'intervalle $(-R/3,\ R/3)$ ont une densité ≥ 0.

L'estimateur en question n'est donc pas utilisable. De plus, on ne voit vraiment pas le rapport existant entre cet estimateur bizarre et la réalité physique très simple (une dérive linéaire pure). Ceci jette un doute sérieux sur l'intérêt de la méthode.

K.D. universel.

Dans le cadre précédent, on est conduit à kriger une expression du type $\int \varphi(Y_x)\ p(dx)$, avec $\int p(dx) = 1$ et $\varphi = \sum \varphi_n \dfrac{H_n}{n!}$, en prenant :

$$\left\{\begin{array}{l} \varphi_p^* = \sum \dfrac{\varphi_n}{n!}\ H_n^* \\ H_n^* = \lambda_n^\alpha\ H_n(Y_\alpha) \\ \lambda_n^\alpha\ \rho_{\alpha\beta}^n = \int \rho_{\beta x}^n\ p(dx) + \mu_{n,\ell}\ f_\beta^\ell \\ \lambda_n^\alpha\ f_\alpha^\ell = \int f_x^\ell\ p(dx) \end{array}\right.$$

Ici, H_n^* krige l'expression $\int H_n(Y_x)\ p(dx)$ sous les conditions d'universalité habituelles. Ces conditions expriment que l'on a :

$$E\ \varphi_p^* = \int p(dx) \int \varphi(y)\ F_x(dy)$$

quelle que soit la loi $F_x(dy)$, pourvu qu'elle soit de la forme $f_x^\ell\ F_\ell(dy)$, avec des F_ℓ arbitraires.

Il n'est donc pas nécessaire d'estimer les F_ℓ eux-mêmes. En pratique, on pourra souvent se contenter d'une anamorphose obtenue comme dans le cas stationnaire, et kriger avec des conditions d'universalité d'ordre peu élevé (1 ou 2), pour conférer à nos estimateurs davantage de robustesse vis-à-vis d'un défaut éventuel de stationnarité. C'est surtout lorsque l'on travaille en voisinage glissant que ces techniques risquent de présenter le plus d'intérêt. Car, dans ce cas, on n'a besoin que d'une hypothèse de stationnarité locale approximative, qui peut rester compatible avec une non-stationnarité manifeste au niveau global. Pour le voir, le plus simple est d'utiliser le point de vue des représentations glissantes.

5 - POINT DE VUE DES REPRESENTATIONS GLISSANTES.

Soit z(x) la V.R., S le champ utile, V le voisinage de travail. En munissant S de la probabilité

$$\varpi(dx) = \frac{1}{S} \; 1_S(x)dx$$

on obtient une F.A. Z(h), définie pour $h \in V$, en posant :

$$Z(h) = z(\underline{x}+h)$$

où $\underline{x}$ est le point aléatoire dans S muni de $\varpi(dx)$. La loi $F_h(z)$ de Z(h), pour $h \in S$, est donc, d'après cette définition :

$$F_h(z) = \int 1_{z(x+h)\leq z} \; \varpi(dx)$$

Si le voisinage de travail V est petit vis-à-vis de S, $F_h(z)$ ne se modifie que relativement peu lorsque h parcourt V, et cela même si, globalement, z(x) n'apparait pas du tout comme stationnaire dans S. Car $F_h(z)$ peut s'écrire :

$$F_h(z) = \frac{1}{S} \int_{S_h} 1_{z_x \leq z} \; dx$$

et les translatés S_h, $h \in V$ ont une grosse intersection commune. On peut donc, dans bien des cas, supposer $F_h(z) = F(z)$ indépendant de $h \in V$ (et utiliser, pour l'estimer, un estimateur global, naïf ou non, comme ci-dessus).

On peut aussi chercher à représenter la légère variabilité de F_h par une expression de la forme :

$$\tilde{F}_h(z) = \varpi_h^\ell \, F_\ell(z)$$

Les $F_\ell(z)$ doivent être évalués par la méthode des moindres carrés, qui conduit à :

$$(5\text{-}1) \qquad F_\ell(z) = \frac{1}{V} \int_V \varpi_\ell(h') \, F_{h'}(z) \, dh'$$

(les $\varpi_\ell(h')$ sont les duales des ϖ_h^ℓ pour la métrique des moindres carrés sur V).

L'expression (5-1) ne représente pas un estimateur, mais constitue une <u>définition</u> (à savoir : les $F_\ell(z)$ sont définis par la technique des moindres carrés appliquée à $F_h(z)$ sur V). Explicitement :

$$F_\ell(z) = \frac{1}{VS} \int_{V\times S} \varpi_\ell(h') \, 1_{z_{x+h'} \leq z} \, dx \, dh'$$

soit

$$F_\ell(z) = \frac{1}{VS} \int K_{SV}^\ell(y) \, 1_{z_y \leq z} \, dy$$

avec

$$K_{SV}^\ell(y) = \int 1_S(x) \, 1_V(y-x) \, \varpi_\ell(y-x) \, dx$$

Pour <u>estimer</u> $F_\ell(z)$ à partir d'une maille, régulière ou non, x_α , il faudrait évidemment utiliser des expressions de la forme :

$$F_\ell^*(z) = \sum_\alpha W_\alpha^\ell \, 1_{z_0 \leq z}$$

avec, par exemple, W_α^ℓ proportionnel à $K^\ell(x_\alpha)$.

<u>Exemple</u> - A titre d'exemple, reprenons le cas de la V.R. $z(x) = x$ sur la droite, avec comme champ S l'intervalle $(-R, R)$ et comme voisinage mobile V l'intervalle $(-r, r)$, avec r petit devant R. $F_h(dz)$ est ici la loi de densité uniforme sur l'intervalle $(-R+h, R+h)$, soit :

$$f_h(z) = \frac{1}{2R} \, 1_{-R+h \leq z < R+h}$$

En ajustant la droite de moindres carrés sur (-r, r) on obtient :

$$\tilde{f}_h(z) = \frac{1}{2R}\left[\int_{-r}^{r} f_{h'}(z)dh' + \frac{3h}{2\,r^3}\int_{-r}^{r} h'\, f_{h'}(z)dh'\right]$$

On trouve :

$$\tilde{f}_h(z) = \frac{1}{2R} \qquad \text{pour} \quad -R + r \leq z \leq R \times r$$

$$\frac{1}{4\,Rr}(R + r - z_o)\left[1 + \frac{3h}{4\,r^2}(z_o + r - R)\right]$$

$$\text{pour } R - r < z \leq R + r$$

(et l'expression symétrique pour $-R + r > z \geq -R - r$).

Il apparait encore des valeurs négatives. Mais elles sont localisées aux extrémités de l'intervalle de variation de z, et leur importance reste faible. Il n'en reste pas moins que le modèle auquel on aboutit semble un peu bizarre pour représenter une pure dérive linéaire.

Comme conclusion provisoire, il ne semble pas que la technique précédente puisse conduire à une estimation raisonnable d'un histogramme dérivant, du moins dans le cas où le phénomène lui-même est affecté d'une dérive bien marquée. Mais cela n'exclut pas l'emploi du K.D. avec conditions d'universalité. La marche à suivre pourrait être la suivante : si on a une V.R. présentant une dérive apparente, on peut commencer par ajuster une dérive du type $a_\ell f_x^\ell$ et travailler sur les résidus $y_x = z_x - a_\ell f_x^\ell$ (cela sera légitime, en pratique, pour des estimations locales et en voisinage glissant). Cette dérive $a_\ell f_x^\ell$ peut être un ajustement de moindres carrés effectué soit sur les z_α soit, ce qui serait mieux, sur le krigeage z_x^* obtenu par les techniques habituelles (FAI-k). Notons qu'en présence de contraintes, (par exemple $z_x \geq 0$) il peut y avoir intérêt à effectuer une anamorphose préalable (par exemple passer aux logarithmes) avant d'appliquer les moindres carrés. C'est aux résidus y(x) que l'on appliquera ensuite une anamorphose gaussienne $y(x) = \Phi(U_x)$, comme si y(x) était stationnaire. En réalité, y(x) ne peut pas être réellement stationnaire, car, en vertu de la manière même dont la dérive a été estimée, son comportement doit dépendre de la plus ou moins grande proximité des points expérimentaux x_α. C'est pourquoi il sera peut-être judicieux, lors d'une estimation locale par K.D., d'ajouter des conditions d'universalité.