NOTE GEOSTATISTIQUE N° 108

COMPLEMENTS SUR LE KRIGEAGE UNIVERSEL

Par

G. MATHERON

Avril 1970

NOTE GEOSTATISTIQUE N° 108

COMPLEMENTS SUR LE KRIGEAGE UNIVERSEL

Table des Matières

NOTE GEOSTATISTIQUE N° 108

COMPLEMENTS SUR LE KRIGEAGE UNIVERSEL

Ces compléments sont inspirés par des problèmes qui se posent en géodésie, et font suite à plusieurs entretiens que j'ai eus avec J.M. Monget.

1 - COMPARAISON AVEC LE MAXIMUM DE VRAISEMBLANCE

Dans le fascicule sur le Krigeage Universel, Annexe 1, j'ai établi l'identité de l'estimateur optimal de la dérive et de l'estimateur du maximum de vraisemblance. Pour éclairer ce résultat, considérons la forme quadratique :

$$(1)\quad Q = B^{\alpha\beta}(Z_\alpha - m_\alpha)(Z_\beta - m_\beta) = B^{\alpha\beta}\ Z_\alpha Z_\beta - 2B^{\alpha\beta}\ Z_\alpha m_\beta + B^{\alpha\beta}\ m_\alpha m_\beta$$

(avec $m_\alpha = a_\ell f_\alpha^\ell$, et $B = \sigma^{-1}$). On obtient l'estimateur du maximum de vraisemblance en donnant aux paramètres a_i les valeurs A_i qui minimisent Q. Exprimons Q au moyen des Y^ℓ définis par :

$$< Y^\ell,\ Y_\alpha > = f_\alpha^\ell$$

c'est-à-dire :

$$Y^\ell = \lambda^{\beta\ell}\ Y_\beta \quad , \quad \lambda^{\beta\ell} = f_\alpha^\ell\ B^{\alpha\beta}$$

Comme $Z_\alpha - m_\alpha$ ne dépend pas des a_i, on peut supposer que les vraies valeurs des a_i sont nulles, et remplacer dans Q les Z_α par les Y_α. Le terme rectangle s'écrit :

$$B^{\alpha\beta} Z_\alpha A_\ell f_\beta^\ell = \lambda^{\alpha\ell} Z_\alpha A_\ell = A_\ell Y^\ell$$

De même, le troisième terme est :

$$B^{\alpha\beta} m_\alpha m_\beta = B^{\alpha\beta} A_\ell A_s f_\alpha^\ell f_\beta^s = A_\ell A_s \lambda^{\beta\ell} f_\beta^s =$$

$$= A_\ell A_s \lambda^{\beta\ell} < Y^s , Y_\beta > = A_\ell A_s < Y^s , Y^\ell >$$

En particulier, la matrice

$$f_\alpha^\ell B^{\alpha\beta} f_\beta^s = < Y^\ell , Y^s >$$

est l'inverse de la matrice des covariances $\mu_{\ell s} = < Y_\ell , Y_s >$.
La forme Q(A) s'écrit ainsi :

$$Q(A) = B^{\alpha\beta} Y_\alpha Y_\beta - 2 A_\ell Y^\ell + A_\ell A_s < Y^\ell , Y^s >$$

L'estimateur du maximum de vraisemblance est donc solution du système :

$$A_\ell < Y^\ell , Y^s > = Y^s$$

d'où

$$A_\ell = \mu_{\ell s} Y^s = Y_\ell$$

Ainsi, l'estimateur A_ℓ du maximum de vraisemblance est bien identique à l'estimateur optimal Y_ℓ.

2 - LE K.U. COMME INTERPOLATEUR

a - Le Krigeage simple.

Dans le cas du krigeage simple, l'estimateur $Y_K(x)$ de $Y(x)$, $x \notin S$ est :

$$Y_K(x) = \lambda^\beta(x)\ Y_\beta$$

avec des coefficients $\lambda^\beta(x)$ solution du système :

$$\lambda^\beta\ \sigma_{\alpha\beta} = \sigma_{\alpha x}$$

soit

$$\lambda^\beta(x) = B^{\alpha\beta}\ \sigma_{\alpha x}$$

On peut regrouper les termes, et mettre $Y_K(x)$ sous la forme :

$$(2) \qquad Y_k(x) = H^\alpha\ \sigma_{\alpha x}$$

avec $H^\alpha = B^{\alpha\beta}\ Y_\beta$). Cet estimateur est un interpolateur exact en ce sens que l'on a $Y_K(x) = Y(x)$ pour $x \in S$, soit $Y_K(x_\beta) = Y_\beta$. Inversement, $Y_K(x)$ est le seul interpolateur exact de la forme(2), de sorte que cette propriété est caractéristique du krigeage simple. En effet, compte tenu de (2), la condition $Y_K(x_\beta) = Y_\beta$ donne :

$$Y_\beta = H^\alpha\ \sigma_{\alpha\beta} \Leftrightarrow H^\alpha = B^{\alpha\beta}\ Y_\beta$$

Pour caractériser le K.U. par ses propriétés d'interpolateur, nous aurons recours à la notion de dérive compatible avec un interpolateur donné. Soit $Y^*(x) = T^\beta(x)\ Y_\beta$. Nous dirons que l'interpolateur $Y^*(x)$ est compatible avec une dérive $m(x)$ si $Y^*(x)$ se transforme en $Y^*(x) + m(x)$ lorsque l'on remplace les Y_β par les $Z_\beta = Y_\beta + m_\beta$. Il en est ainsi si et seulement si :

$$m(x) \equiv m_\beta \; T^\beta(x)$$

De même, on dira que l'interpolateur est compatible avec une derive de la forme $b_\ell \, f^\ell(x)$ s'il est compatible avec $m(x) = b_\ell \, f^\ell(x)$ quels que soient les coefficients b_ℓ . Pour qu'il en soit ainsi, il faut et il suffit que l'on ait :

$$\text{(U)} \qquad f^\ell(x) = T^\beta(x) \; f^\ell_\beta$$

Cette condition de compatibilité avec la dérive $b_\ell f^\ell$ est donc identique à la condition d'universalité qui intervient dans la théorie du K.U.

D'après la note 107, on sait que le krigeage simple doit être compatible avec toute dérive de la forme :

$$\text{(3)} \qquad m(x) = D^\alpha \; \sigma_{\alpha x}$$

Vérifions-le directement. On a ici $T^\beta(x) = \lambda^\beta_K(x) = B^{\alpha\beta} \; \sigma_{\alpha x}$. D'où :

$$m_\beta \; T^\beta(x) = D^\gamma \; \sigma_{\alpha\gamma} \; B^{\alpha\beta} \; \sigma_{\beta x} = D^\alpha \; \sigma_{\alpha x} = m(x)$$

Ainsi, le krigeage simple est un interpolateur compatible avec les dérives de la forme (3). Cette propriété est d'ailleurs caractéristique comme on le vérifie sans peine. Voyons s'il est possible de la généraliser pour obtenir une caractérisation du K.U. en tant qu'interpolateur.

b - Le Krigeage Universel comme interpolateur.

L'estimateur du K.U. (que nous noterons $Z_U(x)$, plutôt que Y_U, puisque maintenant il peut y avoir des dérives quelconques) est de la forme :

$$Z_U(x) = \lambda^\beta(x)\ Z_\beta$$

avec des coefficients λ^β que le théorème d'additivité permet de rattacher aux coefficients λ^β_K du krigeage simple. Si l'on désigne par

$$A_\ell = \lambda^\beta_\ell\ Z_\beta \quad , \quad \mu_{\ell s} = \mathrm{Cov}(A_\ell\ A_s)$$

les estimateurs optimaux de la dérive et la matrice de leur covariance, on a, en effet :

$$(4) \qquad Z_U(x) = \lambda^\beta_K(x)\ (Z_\beta - A_\ell\ f^\ell_\beta) + A_\ell\ f^\ell_x$$

Sous cette forme, on voit immédiatement que le K.U. est compatible avec une dérive du type $b_\ell f^\ell(x)$: si l'on remplace Z_β par Z_β + $+\ b_\ell\ f^\ell_\beta$, A_ℓ est majoré de b_ℓ , et Z_U de $b_\ell\ f^\ell_x$.

En explicitant (4), on trouve :

$$(5) \qquad \begin{cases} \lambda^\beta(x) = \lambda^\gamma_K\ (\delta^\beta_\gamma - \lambda^\beta_\ell\ f^\ell_\gamma) + \lambda^\beta_\ell\ f^\ell(x) \\ \\ (\lambda^\gamma_K = B^{\alpha\gamma}\ \sigma_{\alpha x}) \end{cases}$$

Ainsi, les coefficients du K.U. sont de la forme :

$$(5') \qquad \lambda^\beta(x) = H^{\alpha\beta}\ \sigma_{\alpha x} + \lambda^\beta_\ell\ f^\ell(x)$$

avec deux matrices H et λ indépendantes de x. Le K.U. possède donc les 3 propriétés suivantes :

a/ ses coefficients sont des <u>combinaisons linéaires</u> des $\sigma_{\alpha x}$ <u>et des $f^\ell(x)$</u>, conformément à (5').

b/ le K.U. est un <u>interpolateur exact</u>.

c/ le K.U. <u>est compatible avec les dérives</u> du type $b_\ell f^\ell(x)$.

Ces trois propriétés <u>ne suffisent pas</u> pour caractériser le K.U. Dans le cas du krigeage simple, $Y_K(x)$ était compatible avec les n fonctions $\sigma_{\alpha x}$ ($\alpha = 1,2,\ldots n$), et, comme on le vérifie sans peine, $Y_K(x)$ est le suel interpolateur compatible avec ces n fonctions. Mais ici la condition c/ ne fait plus intervenir que k fonctions f^ℓ ($\ell = 1,2,\ldots k$) avec lesquelles Z_U doit être compatible, et nous devons nous attendre à une indétermination d'ordre n-k. C'èst bien ce que nous allons voir.

Inversement, soit un interpolateur dont les coefficients sont de forme (5'). Cherchons à quelles conditions il vérifie les conditions b/ et c/. La condition d'exactitude b/ équivaut a :

$$H^{\alpha\beta}\ \sigma_{\alpha\gamma} + \lambda^\beta_\ell\ f^\ell_\gamma = \delta^\beta_\gamma \tag{6}$$

et la condition de compatibilité c/ à :

$$f^\ell(x) = H^{\alpha\beta}\ f^\ell_\beta\ \sigma_{\alpha x} + \lambda^\beta_s\ f^\ell_\beta\ f^s_x \tag{6'}$$

Cette condition (6') doit être vraie pour tout $x \in \mathbb{R}^n$. Nous supposons les f^ℓ_x et les $\sigma_{\alpha x}$ linéairement indépendants sur $\mathbb{R}^n$, autrement dit qu'il n'existe pas de combinaison linéaire à coefficients C_ℓ non tous nuls tels que $C_\ell f_x^\ell \equiv D^\alpha \sigma_{\alpha x}$ (le cas où il en existerait ne présente qu'un intérêt limité, comme on le verra ci-dessous ; en particulier, si toutes les fonctions f^ℓ sont des combinaisons linéaires des σ_α, on retombe sur le krigeage simple d'après une remarque faite ci-dessus). La condition (6') est donc équivalente à l'ensemble des deux conditions (7) et (7') suivantes :

(7) $$H^{\alpha\beta} f_\beta^\ell = 0$$

(7') $$\lambda_s^\beta f_\beta^\ell = \delta_s^\ell$$

(7') est la condition <u>d'universalité</u> habituelle. Les 3 conditions (6), (7) et (7'), nécessaires et suffisantes pour que l'interpolateur vérifie les conditions a/, b/ et c/, <u>ne sont pas indépendantes</u>. On peut les remplacer à volonté par le groupe (6)-(7), ou par le groupe (6)-(7').

Montrons, en effet, que <u>(6) et (7) entraînent (7')</u>. Pour celà, multiplions (6) par f_β^s. Compte tenu de (7), il vient :

$$f_\gamma^s = (f_\beta^s \lambda_\ell^\beta) f_\gamma^\ell$$

et (moyennant la condition habituelle d'indépendance linéaire des f^ℓ sur S) il en résulte

$$f_\beta^s \lambda_\ell^\beta = \delta_\ell^s$$

c'est-à-dire (7'). Inversement, <u>montrons que (6) et (7') entraînent (7)</u>. En multipliant (6) par f_β^s et en tenant compte de (7'), il vient en effet :

$$f^s_\beta \; H^{\alpha\beta} \; \sigma_{\alpha\gamma} = 0$$

et (7) en résulte, puisque σ est inversible.

Finalement, l'interpolateur est exact et compatible avec les f^ℓ si et seulement si on a :

$$(8) \qquad \begin{cases} H^{\alpha\beta} \; \sigma_{\alpha\gamma} + \lambda^\beta_\ell \; f^\ell_\gamma = \delta^\beta_\gamma \\ \\ \lambda^\beta_s \; f^\ell_\beta = \delta^\ell_s \end{cases}$$

Le système (8) comporte $n^2 + nk$ inconnues (les $H^{\alpha\beta}$ et les λ^β_s) et seulement $n^2 + k^2$ équations. La solution générale dépend effectivement de k(n-k) paramètres arbitraires, car le système $f^\ell_\cdot$ de k vecteurs linéairement indépendants admet dans $\mathbb{R}^n$ un complémentaire orthogonal de dimensions n-k.

Pour caractériser le K.U., nous devons donc trouver n-k conditions supplémentaires. Comme le K.S. était compatible avec les fonctions $\sigma_{\alpha x}$, cherchons à quelle condition le K.U. est compatible avec une combinaison linéaire du type $D^\alpha \; \sigma_{\alpha x}$. Pour celà, revenons aux relations (5). Elles donnent :

$$\lambda^\beta(x) \; D^\alpha \; \sigma_{\alpha\beta} = D^\alpha \; \sigma_{\alpha x} + D^\alpha \; \sigma_{\alpha\beta} \; (\lambda^\beta \; f^\ell_x - \lambda^\gamma_K \; \lambda^\beta_\ell \;\; f^\ell_\gamma)$$

La condition de compatibilité du K.U. avec $D^\alpha \; \sigma_{\alpha x}$ est donc :

$$\sigma_{\alpha\beta} \; D^\alpha \; (\lambda^\beta_\ell \;\; f^\ell_x - \lambda^\gamma_K(x) \; \lambda^\beta_\ell \;\; f^\ell_\gamma) = 0$$

Explicitons, en remarquant $\sigma_{\alpha\beta}\ \lambda_\ell^\beta = \mu_{\ell s}\ f_\alpha^s$. La condition est :

$$0 = D^\alpha\ (\mu_{\ell s}\ f_\alpha^s\ f_x^\ell - \mu_{\ell s}\ \lambda_K^\gamma(x)\ f_\alpha^s\ f_\gamma^\ell) =$$

$$= D^\alpha\ \mu_{\ell s}\ f_\alpha^s\ (f_x^\ell - B^{\gamma\gamma'}\ \sigma_{\gamma' x}\ f_\gamma^\ell\ )$$

En l'absence de relation linéaire entre les f_x^ℓ et les $\sigma_{\gamma' x}$, cette condition est donc :

$$D^\alpha\ f_\alpha^s = 0 \tag{9}$$

Ainsi, outre les propriétés a/, b/ et c/, le K.U. vérifie :

d/ <u>le K.U. est un interpolateur compatible avec les dérivés de la forme $D^\alpha\ \sigma_{\alpha x}$ telles que $D^\alpha\ f_\alpha^s = 0$</u>.

Cette condition (5) s'écrit :

$$0 = D^\alpha\ f_\alpha^s = D^\alpha < Y_\alpha,\ Y^s >$$

Elle exprime par conséquent que l'image $D^\alpha\ Y_\alpha$ dans $\bar{H}$ de la fonction $D^\alpha\ \sigma_{\alpha\beta}$ sur S est orthogonale au sous-espace engendré par les Y^s. Ces propriétés a/, b/, c/ et d/ <u>sont caractéristiques du K.U</u>. En effet, soit $Z^*(x)$ un interpolateur de la forme :

$$Z^*(x) = \lambda^\beta(x)\ Z_\beta$$

vérifiant ces conditions, donc :

$$\lambda^\beta(x) = H^{\alpha\beta}\ \sigma_{\alpha x} + \lambda_\ell^\beta\ f^\ell(x)$$

avec des matrices H et λ vérifiant (8) (propriétés a/, b/ et c/).

Pour exprimer que d/ est vérifiée, il faut écrire :

$$D^\alpha f^s_\alpha = 0 \Rightarrow \lambda^\beta(x)\, D^\alpha \sigma_{\alpha\beta} = D^\alpha \sigma_{\alpha x}$$

autrement dit, il doit exister des fonctions $C_s(x)$ telles que l'on ait :

$$\sigma_{\alpha\beta}\, \lambda^\beta(x) - \sigma_{\alpha x} = C_s(x)\, f^s_\alpha$$

Mais ceci n'est pas autre chose que la première équation du système (K.U.). Jointe à la seconde relation (8), qui est la condition d'universalité, cette condition montre que <u>le K.U. est le seul interpolateur vérifiant a/, b/, c/ et d/</u>.

<u>Remarque</u> - On peut réduire le nombre de ces critères. On a déjà vu que le krigeage simple est le seul interpolateur compatible avec les dérives de la forme $D^\alpha \sigma_{\alpha x}$ (D^α coefficients quelconques). La propriété d'exactitude en découle - car les relations :

$$D^\alpha \lambda^\beta \sigma_{\alpha\beta} = D^\alpha \sigma_{\alpha x}$$

pour tout (D^α) entraînent bien $\lambda^\beta \sigma_{\alpha\beta} = \sigma_{\alpha x}$.

De même, le K.U. <u>est le seul interpolateur compatible avec les $f^\ell(x)$ et avec les fonctions D^α</u> $\sigma_{\alpha x}$ telles que $D^\alpha f^s_\alpha = 0$. Le fait que cet estimateur soit de la forme (5) et soit un interpolateur exact est alors une conséquence de ces propriétés de compatibilité. En effet, la compatibilité avec les $f^\ell(x)$ s'écrit :

$$\lambda^\beta f^\ell_\beta = f^\ell_x$$

(condition d'universalité), et la compatibilité avec les fonctions du second type, soit :

$$D^{\alpha} f_{\alpha}^{s} = 0 \Rightarrow D^{\alpha} \lambda^{\beta} \sigma_{\alpha\beta} = D^{\alpha} \sigma_{\alpha x}$$

se met sous la forme :

$$\lambda^{\beta}(x) \sigma_{\alpha\beta} - \sigma_{\alpha x} = \mu_s(x) f_{\alpha}^{s}$$

(condition d'optimalité), et l'ensemble des deux conditions caractérise le K.U.

Les conditions d'optimalité ont un contenu probabiliste, mais non les conditions d'universalité qui ne font pas intervenir la covariance. <u>Lorsque k varie de 0 à n</u>, <u>le contenu probabiliste du K.U. va donc en décroissant</u>. Pour n = k, en particulier, le K.U. se réduit à une pure interpolation :

$$\begin{cases} Z_U(x) = \Sigma \lambda^{\beta} Z_{\beta} \\ \\ \lambda^{\beta} f_{\beta}^{\ell} = f_x^{\ell} \end{cases}$$

Si φ_{ℓ}^{β} est la matrice inverse de f_{β}^{ℓ} (qui existe pour n = k, d'après l'indépendance linéaire) on a :

$$\lambda^{\beta}(x) = \varphi_{\ell}^{\beta} f^{\ell}(x)$$

et

$$Z_U(x) = (\varphi_{\ell}^{\beta} Z_{\beta}) f^{\ell}(x)$$

est déterminé par la condition d'être de la forme $b_{\ell} f^{\ell}(x)$ et de réaliser l'interpolation exacte.

Ainsi, le K.U. apparaît comme un compromis entre une technique de pure interpolation et un filtrage probabiliste. Le contenu probabiliste du K.U. est lié à la dimension n-k de l'espace des fonctions du type $D^{\alpha}\,\sigma_{\alpha x}$ avec lesquelles il doit rester compatible, et son contenu purement interpolatoire à la dimension k de l'espace des fonctions $a_{\ell} f^{\ell}$. Celà ne surprendra pas si l'on se souvient que le K.U. doit réaliser simultanément deux opérations : l'interpolation d'une fonction continue m(x) par une expression de la forme $a_{\ell} f^{\ell}(x)$, et le filtrage des résidus. Mais on voit aussi que si k n'est pas nettement plus petit que n l'aspect interpolatoire (donc, en un sens, arbitraire) du procédé risque de l'emporter sur son aspect probabiliste.

Abandonnant (momentanément) le point de vue probabiliste, montrons qu'il est possible de caractériser un interpolateur exact par ses dérives associées.

3 - CARACTERISATION D'UN INTERPOLATEUR PAR SES DERIVES ASSOCIEES

Soit $Z^*(x) = \lambda^{\beta}\, Z_{\beta}$ un interpolateur (Z_{β}, valeurs numériques aux points $x_{\beta} \in S$, $\beta = 1,2,n$, $\lambda^{\beta} = \lambda^{\beta}(x)$ est une fonction de $x \in \mathbb{R}^n$). Pour qu'une fonction g(x) soit une dérive compatible avec $Z^*(x)$, il faut et il suffit que l'on ait :

$$g(x) = \lambda^{\beta}(x)\, g_{\beta} \qquad \left(g_{\beta} = g(x_{\beta})\right)$$

en tout $x \in \mathbb{R}^n$. S'il existe n fonctions $g^{\ell}(x)$ linéairement indépendantes sur S et compatibles avec Z^*, on a :

$$g^{\ell}(x) = \lambda^{\beta}(x)\, g^{\ell}_{\beta}$$

et, en désignant par γ^{β}_{ℓ} la matrice inverse des g^{ℓ}_{β} :

$$(10) \qquad \lambda^{\beta}(x) = \gamma^{\beta}_{\ell}\; g^{\ell}(x)$$

de sorte que <u>l'interpolateur Z^* est déterminé par la donnée des n fonctions g^{ℓ}</u> compatibles avec lui. Cet interpolateur est <u>nécessairement exact</u>. Posant, en effet, $\lambda^{\beta}(x_{\gamma}) = \lambda^{\beta}_{\gamma}$ et $g^{\ell}(x_{\gamma}) = g^{\ell}_{\gamma}$, la relation (10) ci-dessus donne pour $x = x_{\gamma}$:

$$\lambda^{\beta}_{\gamma} = \gamma^{\beta}_{\ell}\, g^{\ell}_{\gamma} = \delta^{\beta}_{\gamma}$$

puisque les matrices g et γ sont inverses l'une de l'autre. Inversement, soit $Z^*(x) = \lambda^{\beta}(x)\; Z_{\beta}$ un interpolateur exact, c'est-à-dire tel que :

$$(11) \qquad \lambda^{\beta}_{\gamma} = \delta^{\beta}_{\gamma}$$

Il suffit de prendre n vecteurs $g^{\ell} = (g^{\ell}_{\alpha})$ linéairement indépendants dans $\mathbb{R}^n$ et de poser :

$$g^{\ell}(x) = \lambda^{\beta}(x)\; g^{\ell}_{\beta}$$

pour obtenir n fonctions linéairement indépendantes sur S (car (11) entraîne $g^{\ell}(x_{\gamma}) = g^{\ell}_{\gamma}$) compatibles avec l'interpolateur Z^* et vérifiant (10).

Il y a donc <u>correspondance biunivoque entre l'ensemble des</u> interpolateurs exacts et l'ensemble des sous-espaces de $\mathcal{F}(\mathbb{R}^n)$ qui sont de dimensions n ainsi que leurs restrictions à $\mathcal{F}(S)$.

Ces fonctions g^ℓ ne sont pas obligatoirement continues, de sorte qu'à l'extérieur de S tout interpolateur peut être assimilé à un interpolateur exact. Soit en effet $Z^*(x) = \lambda^\beta(x)\ Z_\beta$ un interpolateur quelconque. Posons

$$\mu^\beta(x) = \begin{cases} \lambda^\beta(x) & \text{si } x \notin S \\ \delta^\beta_\gamma & \text{si } x = x_\gamma \in S \end{cases}$$

L'interpolateur :

$$M^*(x) = \mu^\beta(x)\ Z_\beta \tag{12}$$

est exact, continu sur S^c si $Z^*(x)$ était continu dans $\mathbb{R}^n$, mais non continu dans $\mathbb{R}^n$ si la condition $\lambda^\beta_\gamma = \delta^\beta_\gamma$ n'était pas remplie. Il est compatible avec (et caractérisé par) n fonctions linéairement indépendantes sur S de la forme :

$$g^\ell(x) = \mu^\beta(x)\ g^\ell_\beta \tag{12'}$$

pour lesquelles les points de S sont, en général, des points de discontinuité. L'estimateur optimal de la dérive est un interpolateur de ce type. En effet, soient :

$$A_\ell = \lambda^\beta_\ell\ Z_\beta$$

les estimateurs optimaux de la dérive, obtenus en résolvant le système :

$$\text{(D)} \qquad \begin{cases} \lambda^\beta_\ell\ \sigma_{\alpha\beta} = \mu_{\ell s}\ f^s_\alpha \\ \lambda^\beta_\ell\ f^s_\beta = \delta^s_\ell \end{cases}$$

pour $k < n$ fonctions f^{ℓ} linéairement indépendantes sur S. L'interpolateur défini par :

$$M^*(x) = A_{\ell} f^{\ell}(x) = \left(\lambda_{\ell}^{\beta} f^{\ell}(x)\right) Z_{\beta}$$

pour $x \notin S$ (et $M^*(x_\beta) = Z_\beta$ pour $x = x_\beta \in S$) est compatible avec les k fonctions f^{ℓ} d'après la seconde relation (D) : plus généralement, tout estimateur vérifiant cette condition d'universalité constitue un interpolateur compatible avec les f^{ℓ}. Si les g^i ($i = 1,2,\ldots n$) sont n vecteurs linéairement indépendants dans $\mathbb{R}^n$, $M^*(x)$ est caractérisé par les n fonctions associées :

$$(13) \qquad \begin{cases} g^i(x) = \lambda_s^{\beta} f^s(x)\, g_{\beta}^i & (x \notin S) \\ \\ g^i(x_\gamma) = g_\gamma^i \end{cases}$$

($i = 1,2,\ldots n$, $\beta,\gamma = 1,2,\ldots n$, $s = 1,2,\ldots k$). Nous pouvons prendre $g_\beta^i = f_\beta^i$ pour $i = 1,2,\ldots k$, avec, d'après la relation d'universalité sur $\mathbb{R}^n$ entier (et pas seulement sur $\mathbb{R}^n \setminus S$) :

$$g^i(x) = f^i(x) \qquad (x \in \mathbb{R}^n,\ i = 1,2,\ldots k)$$

Il reste à déterminer n-k autres fonctions g. Munissons $\mathbb{R}^n$ de la norme $\|f\|^2 = B^{\alpha\beta} f_\alpha f_\beta$, B désignant la matrice inverse des $\sigma_{\alpha\beta}$. Nous pouvons trouver n-k vecteurs g^i ($i= k+1,\ldots n$) linéairement indépendants et orthogonaux aux vecteurs f^{ℓ}($\ell = 1,2,\ldots k$), c'est-à-dire vérifiant :

$$(14) \qquad B^{\alpha\beta} f_\alpha^{\ell} g_\beta^i = 0 \qquad (\ell = 1,2,\ldots k,\ i = k+1,\ldots n)$$

A ces n-k vecteurs g^i, les relations (13) associent les n-k fonctions compatibles avec M^* dont nous avions besoin pour achever de caractériser cet interpolateur. Mais la première relation (D) donne :

$$\lambda_\ell^\beta = \mu_{\ell s} f_\alpha^s B^{\alpha\beta}$$

Portons dans la première relation (13) :

$$g^i(x) = \mu_{\ell s} f_\alpha^s B^{\alpha\beta} f^\ell(x) g_\beta^i \qquad (x \notin S)$$

Pour $i > k$ et $x \notin S$, on a donc, d'après (14) :

$$g^i(x) = 0 \qquad (x \notin S)$$

Ainsi, l'estimateur optimal de la dérive, considéré comme un interpolateur, est défini par la donnée des n fonctions compatibles suivantes :

$$\begin{cases} g^i(x) = f^i(x) & (i = 1,2,\ldots k,\ x \in \mathbb{R}^n) \\ g^i(x) \begin{cases} 0 & \text{si } x \notin S \\ g_\gamma^i & \text{si } x = x_\gamma \in S \end{cases} & (i = k+1,\ \ldots n) \end{cases}$$

les g_γ^i désignant les composantes de n-k vecteurs g^i linéairement indépendants et vérifiant la relation d'orthogonalité (14) :

$$B^{\alpha\beta} f_\alpha^\ell g_\beta^i = 0$$

(on note que cette relation équivaut à $\lambda_\ell^\beta g_\beta^i = 0$). Les n-k fonctions complémentaires g^i ($i= k+1,\ldots n$) sont <u>identiquement nulles hors de S</u>.

Dans le cas de <u>l'estimateur des moindres carrés</u>, on peut prendre $B^{\alpha\alpha} = 1$, $B^{\alpha\beta} = 0$ pour $\alpha \neq \beta$, et la condition (14) d'orthogonalité se réduit à :

$$\sum_{\alpha=1}^{n} f_\alpha^\ell \, g_\alpha^i = 0$$

4 - GENERALISATION AU CAS CONTINU

Plaçons-nous maintenant dans le cas où l'ensemble S des données expérimentales est infini (en général, on supposera S compact). On pourrait définir un interpolateur (ou peut-être un "extrapolateur") à partir d'un noyau $\lambda(x;dy)$, qui serait une mesure dépendant du point d'appui x et à support dans S permettant d'associer à toute fonction Z(y) définie pour $y \in S$ la fonction :

$$Z^*(x) = \int \lambda(x,dy) \; Z(y)$$

définie sur $\mathbb{R}^n$ (ou, à volonté, seulement sur $\mathbb{R}^n \setminus S$). Ce point de vue, toutefois, n'est pas exactement celui qui convient si nous désirons retrouver le K.U. comme cas particulier. Nous allons donc adopter une formulation mieux adaptée à notre point de vue probabiliste. Nous désignerons - comme d'habitude - par Y(x) une F.A. d'espérance nulle et de covariance $\sigma(x,y)$, par $\overline{H}$ et $\overline{H}_f$ les espaces de Hilbert engendrés par les Y(x) et les Y(x)+f(x) pour $x \in S$, et par $\mathfrak{F}(S)$ l'ensemble des fonctions $\mathfrak{F}$ telles qu'il existe $Y_f \in \overline{H}$ avec $f(x) = \langle Y_f, Y(x) \rangle$ pour $x \in S$. Nous munirons l'espace $\mathfrak{F}(S)$ de la norme $\|f\| = \|Y_f\|$ pour laquelle il est un espace de Hilbert homéomorphe à $\overline{H}$ (cf. Note 107, Annexe).

Nous appellerons <u>interpolateur</u> une application $Y^* : \mathbb{R}^n \setminus S \to \overline{H}$ associant à tout $x \in \mathbb{R}^n \setminus S$ l'élément $Y^*(x) \in \overline{H}$, et nous dirons que <u>l'interpolateur Y^* est compatible avec une fonction f</u> définie sur $\mathbb{R}^n$ dont la restriction à S (que nous désignerons encore par f) est dans $\mathcal{F}(S)$ si l'on a :

$$(15) \qquad f(x) = \langle Y^*(x), Y_f \rangle$$

pour tout $x \in \mathbb{R}^n \setminus S$. (La relation (15) est vérifiée, par définition, pour $x \in S$ si l'on pose $Y^*(x) = Y(x)$ pour $x \in S$, ce que l'on peut toujours faire).

Toute fonction $f \in \mathcal{F}(S)$ (donc définie sur S) peut être prolongée sur $\mathbb{R}^n$ en posant

$$f^*(x) = \langle Y^*(x), Y_f \rangle \quad , \quad x \notin S$$

et une fonction f sur $\mathbb{R}^n$ est compatible avec l'interpolateur Y^* si et seulement si $f = f^*$ sur $\mathbb{R}^n \setminus S$.

<u>Exemple 1</u> - Soit Π le projecteur de $L^2(\Omega,\mathcal{A},P)$ dans $\overline{H}$. Si l'on pose $Y_K(x) = \Pi Y(x)$ $(x \in \mathbb{R}^n)$, on obtient l'interpolateur associé au <u>krigeage simple</u>. Il est compatible avec toutes les fonctions f sur $\mathbb{R}^n$ de la forme $f(x) = \langle Y(x), Y \rangle$ pour un $Y \in \overline{H}$ (cf. Note 107), et d'ailleurs avec celles-là seulement, puisque $\langle \Pi\, Y(x), Y_f \rangle =$ $= \langle Y(x), Y_f \rangle$ dès que $Y_f \in \overline{H}$. Inversement, si un interpolateur Y^* est compatible avec toutes les fonctions $\langle Y(.), Y \rangle$, $Y \in \overline{H}$, on a d'après (15)

$$\langle Y^*(x), Y \rangle = \langle Y(x), Y \rangle$$

pour tout $Y \in \overline{H}$ et tout $x \notin S$, donc(puisque $Y^*(x) \in \overline{H}$ par définition), $Y^*(x) = \Pi\, Y(x) = Y_K(x)$. La propriété ci-dessus est donc caractéristique du krigeage simple.

Exemple 2 - Soit maintenant f^ℓ, $\ell = 1,2,\dots k$, k fonctions sur $\mathbb{R}^n$ dont les restrictions à S sont dans $\mathfrak{F}(S)$, et $Y^\ell \in \overline{H}$ les éléments de $\overline{H}$ vérifiant $< Y^\ell, Y(x) > = f^\ell(x)$, $x \in S$, et soit Y^* l'interpolateur associé au K.U. relatif à ces k fonctions, vérifiant donc :

$$(K.U.) \begin{cases} < Y^*(x), Y(y) > = \sigma(x,y) + \mu_\ell(x)\, f^\ell(y) & (x \notin S,\ y \in S) \\ < Y^*(x), Y^\ell > = f^\ell(x) & (x \notin S) \end{cases}$$

La seconde relation exprime que Y^* est compatible avec les f^ℓ. Soit g une fonction sur $\mathbb{R}^n$ dont la restriction à S est dans $\mathfrak{F}(S)$. Y^* est compatible avec g si et seulement si

$$(16) \qquad g(x) = < Y^*(x), Y_g > \quad , \quad g \in \mathbb{R}^n$$

Pour déterminer l'ensemble $\mathcal{G}^*$ des fonctions compatibles avec Y^*, il suffit (puisque les f^ℓ sont dans $\mathcal{G}^*$) de déterminer l'orthogonal dans $\mathfrak{F}^*(S)$ de la restriction à S des $\{f^\ell\}$ et de prolonger sur $\mathbb{R}^n$ les fonctions g correspondantes au moyen de la relation (16). Sur S, ces fonctions g sont de la forme :

$$g(x) = < Y(x), Y > \quad , \ x \in S$$

pour des $Y \in \overline{H}$ vérifiant $< Y, Y^\ell > = 0$. Pour $x \notin S$, le prolongement de g, d'après (16) et (K.U.), se met sous la forme :

$$g(x) = < Y^*(x), Y > = < Y(x), Y > + \mu_\ell(x) < Y^\ell, Y >$$

donc :

$$g(x) = < Y^*(x), Y > \quad , \quad x \notin S$$

puisque $< Y^\ell, Y > = 0$ par définition. Inversement, à tout Y vérifiant

(17) $$< Y, Y^\ell > = 0$$

correspond la fonction g sur $\mathbb{R}^n$ définie par :

(17') $$g(x) = \begin{cases} < Y(x), Y > & \text{si } x \in S \\ < Y^*(x), Y > & \text{si } x \in \mathbb{R}^n \setminus S \end{cases}$$

qui est manifestement compatible avec Y^*. Ainsi, les fonctions compatibles avec le K.U. sont les combinaisons linéaires des f^ℓ et des fonctions g vérifiant (17) et (17').

Cette propriété est caractéristique du K.U., comme on le verra ci-dessous.

<u>Exemple 3</u> - Avec les mêmes notations que dans l'exemple 2, désignons par

$$M^*(x) = Y_\ell \, f^\ell(x) \qquad (x \in \mathbb{R}^n \setminus S)$$

l'interpolateur associé à <u>l'estimateur optimal de la dérive</u>. Les fonctions associées sont les f^ℓ d'une part, de l'autre les fonctions g définies par :

$$g(x) = \begin{cases} < Y(x), Y > & , \quad x \in S \\ 0 & , \quad x \in \mathbb{R}^n \setminus S \end{cases}$$

pour $Y \in \bar{H}$ vérifiant $< Y, Y^\ell > = 0$.

Caractérisation d'un interpolateur par ses dérives associées.

Soit Y^* un interpolateur et $\mathcal{G}^*$ l'ensemble des fonctions g compatibles avec Y^*. On a $g \in \mathcal{G}^*$ si et seulement si il existe $Y_g \in \overline{H}$ avec :

$$g(x) = \begin{cases} < Y(x), Y_g > & \text{si } x \in S \\ < Y^*(x), Y_g > & \text{si } x \in \mathbb{R}^n \setminus S \end{cases}$$

de sorte que l'application $g \to Y_g$ est une bijection de $\mathcal{G}^*$ sur $\overline{H}$. Il est clair que l'interpolateur Y^* est déterminé sans ambigüité par la donnée de l'espace $\mathcal{G}^*$ des fonctions compatibles avec lui (c'est-à-dire par la donnée d'un prolongement sur $\mathbb{R}^n$ de chacune des fonctions de $\mathfrak{F}(S)$). En effet, pour chaque $x \notin \mathbb{R}^n$ on connaît alors la valeur $g(x)$ du produit scalaire $< Y^*(x), Y_g >$ pour chaque $Y_g \in \overline{H}$.

Pour chaque $g \in \mathcal{G}^*$, considérons l'espace $\overline{H}_g$ engendré par les $Y(x) + g(x)$, $x \in S$, et l'homéomorphisme $Y \to Z_g = Y + <Y, Y_g>$ de $\overline{H}$ sur $\overline{H}_g$. L'image de $Y^*(x)$, $x \notin \mathbb{R}^n$, dans cet homéomorphisme est :

$$Z_g^*(x) = Y^*(x) + < Y^*(x), Y_g > \qquad (x \in \mathbb{R}^n \setminus S)$$

c'est-à-dire par définition de $\mathcal{G}^*$:

$$Z_g^* = Y^*(x) + g(x)$$

Autrement dit, la définition donnée en (15) d'une fonction compatible avec Y^* constitue bien la généralisation de la définition utilisée dans le paragraphe précédent.

Soit Y^* un interpolateur, et $\mathcal{G}^*$ l'ensemble des fonctions compatibles avec Y. $\mathcal{G}^*$ est un espace vectoriel isomorphe à $\overline{\overline{H}}$ ou à $\mathfrak{F}(S)$ et caractérise l'interpolateur Y^*. Désignons par $\mathcal{G}_p^*$ le sous-espace de $\mathcal{G}^*$ constitué des fonctions g vérifiant :

$$g(x) = \langle Y^*(x), Y_g \rangle \quad , \; x \in \mathbb{R}^n \setminus S$$

et par $\mathcal{G}_i^*$ son supplémentaire orthogonal. L'espace $\mathcal{G}^*$ se présente comme la somme directe :

$$\mathcal{G}^* = \mathcal{G}_p^* \oplus \mathcal{G}_i^*$$

des espaces $\mathcal{G}_p^*$ et $\mathcal{G}_i^*$ qui prennent en charge, respectivement, <u>le contenu probabiliste et le contenu purement interpolatoire</u> de l'estimateur utilisé.

5 - DERIVE DEFINIE SOUS FORME IMPLICITE

Considérons le problème d'inférence statistique suivant, qui se rencontre en géodésie (par exemple lorsqu'on a mesuré les trois côtés et les trois angles d'un triangle, avec des résultats géométriquement incompatibles, et que l'on veut reconstituer le triangle "le plus probable").

On se propose de déterminer les valeurs inconnues de n paramètres m_α ($\alpha = 1,2,\ldots n$) à partir des résultats

$$Z_\alpha = Y_\alpha + m_\alpha$$

de mesures entachées d'erreurs Y_α vérifiant :

$$E\,Y_\alpha = 0 \quad , \quad E\,Y_\alpha\,Y_\beta = \sigma_{\alpha\beta}$$

avec une matrice de covariance $\sigma_{\alpha\beta}$ connue, sachant que les vraies valeurs (inconnues) des m_α sont liées par n-k relations physiques du type :

$$M_u^\alpha\, m_\alpha = g_u \qquad (u = 1,2,\ldots n-k)$$

où les g_u et la matrice M sont connues. Quitte à remplacer Z_α par $Z_\alpha' - \mu_\alpha$ et m_α par $m_\alpha - \mu_\alpha$ pour un μ_α particulier vérifiant $M\,\mu = g$, on peut supposer $g = 0$ sans perdre de généralité, soit :

$$(18) \qquad M_u^\alpha\, m_\alpha = 0 \qquad (u = 1,2,\ldots n-k)$$

Cette relation (18) exprime que l'on a :

$$m_\alpha = a_i\, f_\alpha^i \qquad (i = 1,2,\ldots k)$$

avec des f_α^i linéairement indépendants en α et k paramètres inconnus a_i à déterminer. L'estimation des m_α vérifiant (18) se laisse donc traiter comme une estimation optimale de dérive. Mais, en fait, la résolution explicite n'est pas obligatoire, et l'on peut former l'estimateur optimal de m_α sans déterminer effectivement les f_α^i.

En effet, on se propose de déterminer un estimateur de la forme :

$$m_\alpha^* = \lambda_\alpha^\beta\, Z_\beta$$

vérifiant les deux conditions habituelles d'universalité et d'optimalité. La condition d'universalité dit que l'on doit avoir

$E\, m^*_\alpha = m_\alpha$ pour tout vecteur $m = (m_\alpha) \in \mathbb{R}^n$ vérifiant les relations (18), soit :

$$M^\alpha_u\, m_\alpha = 0 \;\Rightarrow\; \lambda^\beta_\alpha\, m_\beta = m_\alpha$$

Or cette condition équivaut à l'existence d'une matrice S^u_α (u = = 1, ...n-k) telle que l'on ait identiquement :

$$(U.) \qquad \lambda^\beta_\alpha = \delta^\beta_\alpha - S^u_\alpha\, M^\beta_u$$

Cette condition étant remplie, l'estimateur m^*_α se met sous la forme :

$$(19) \qquad m^*_\alpha = Z_\alpha - S^u_\alpha\, M^\beta_u\, Z_\beta$$

et il reste, parmi les estimateurs de cette forme, à choisir celui dont la variance est minimale pour chaque α. Pour chaque α, on a :

$$D^2(m^*_\alpha) = \sigma^2_\alpha - 2\, S^u_{(\alpha)}\, M^\beta_u\, \sigma_{(\alpha)\beta} + S^u_{(\alpha)}\, S^v_{(\alpha)}\, M^\beta_u\, \sigma_{\beta\gamma}\, M^\gamma_v$$

(sans sommation sur l'indice (α) muet). Le minimum est atteint pour les S^u_α vérifiant le système :

$$(19') \qquad S^v_\alpha\, M^\beta_u\, \sigma_{\beta\gamma}\, M^\gamma_u = M^\beta_u\, \sigma_{\alpha\beta}$$

qui admet une solution et une seule si la matrice M est de rang k. On vérifie sans peine que la variance optimale est alors :

$$(19'') \qquad D^2(m^*_\alpha) = \sigma^2_\alpha - S^u_{(\alpha)}\, M^\beta_u\, \sigma_{(\alpha)\beta}$$

L'estimateur défini par (19) et (19') est identique à celui que l'on aurait trouvé sous forme explicite $A_i\ f^i_\alpha$ par la technique du K.U. - Il est donc identique également à l'estimateur du maximum de vraisemblance.

Généralisation.

Il arrive que les conditions de liaison imposées aux paramètres m_α à estimer soient données, d'une manière moins directe qu'en (18), sous la forme :

$$(20) \qquad M^\alpha_u\ m_\alpha = A^\ell_u\ c_\ell$$

$$(u = 1,2,\ldots K\ ;\ \ell = 1,2,\ldots k\ ;\ 0 < K-k < n)$$

où M et A sont des matrices connues et les c_ℓ des coefficients pouvant prendre des valeurs quelconques dans $\mathbb{R}^k$. En langage algébrique, donc, (20) exprime que Mm ne doit pas être un point quelconque de $M(\mathbb{R}^n) \subset \mathbb{R}^K$ mais doit appartenir à l'image $A(\mathbb{R}^k) \subset \mathbb{R}^K$ de $\mathbb{R}^k$ dans $\mathbb{R}^K$ par la matrice A. Nous supposons (ce qui est loisible) que A est de rang k, donc $A(\mathbb{R}^k)$ de dimension k dans $\mathbb{R}^K$. De même nous supposons que M est de rang $\mu = \text{Inf}(n,K)$. Si $K \leq n$, le rang μ est K, on a $M(\mathbb{R}^n) = \mathbb{R}^K$ et la condition $M\ m \in A(\mathbb{R}^k)$ peut s'exprimer au moyen de K-k relations linéaires indépendantes imposées à m. Si $K > n$, le rang μ de M est $\mu = n$, les choses sont un peu moins simples. Si l'on désigne par p la dimension de $M(\mathbb{R}^n) \cap A(\mathbb{R}^k) \subset \mathbb{R}^k$ (donc $p \leq \text{Inf}\ (n,k)$) la condition :

$$M\ m \in M(\mathbb{R}^n) \cap A(\mathbb{R}^k)$$

s'exprime à l'aide de n-p relations indépendantes imposées à m.

Nous devons chercher un estimateur du type

$$m^*_\alpha = \lambda^\beta_\alpha \, m_\beta$$

vérifiant les conditions habituelles :

Conditions d'universalité. - Désignons par $\Lambda : \mathbb{R}^n \to \mathbb{R}^n$ l'application de matrice λ. Pour tout $m \in \mathbb{R}^n$ tel que $M\,m \in A(\mathbb{R}^k)$, on doit avoir $\lambda^\beta_\alpha \, m_\beta = m_\alpha$, soit :

(21) $$m \in \mathbb{R}^n \ , \ M\,m \in A(\mathbb{R}^k) \Rightarrow (I - \Lambda)\, m = 0$$

Or on a toujours $0 \in A(\mathbb{R}^k)$. La relation (21) entraîne donc que le noyau de M est contenu dans le noyau de $(I - \Lambda)$, donc que $(I - \Lambda)$ se laisse factoriser par M : si la condition d'universalité (21) est remplie, il existe une application linéaire S de $M(\mathbb{R}^n) \subset \mathbb{R}^K$ dans $\mathbb{R}^n$ telle que l'on ait :

(22) $$I - \Lambda = S\,M$$

Il faut alors envisager séparément les cas $K \leq n$ et $K > n$.

Cas 1. - $K \leq n$ - La matrice M est de rang K, donc $M(\mathbb{R}^n) = \mathbb{R}^K$ et l'application S de la relation (22) est définie sur $\mathbb{R}^K$ entier. La condition d'universalité (21) s'exprime alors simplement par $SA = 0$. Dans ce cas simple, donc, on doit avoir :

(23-1) $$\begin{cases} \Lambda = I - SM \\ SA = 0 \end{cases}$$

et la correspondance entre Λ et S est biunivoque.

Cas 2. - $n < K$ - La matrice M est de rang n, $M(\mathbb{R}^n)$, de dimension n, est strictement contenu dans $\mathbb{R}^K$. Les conditions (21) expriment que l'application S de $M(\mathbb{R}^n)$ dans $\mathbb{R}^n$ s'annule sur l'espace :

$$P = M(\mathbb{R}^n) \cap A(\mathbb{R}^k)$$

de dimension p. On peut prolonger S par une application $\tilde{S}$ de $\mathbb{R}^K$ dans $\mathbb{R}^n$. Pour S donnée, ce prolongement $\tilde{S}$ dépend de $n(K-n)$ paramètres arbitraires. La condition (21) s'écrit $S(P) = \{0\}$. On peut profiter de l'arbitraire qui préside au choix du prolongement $\tilde{S}$ pour imposer à $\tilde{S}$ de s'annuler non seulement sur P mais sur $A(\mathbb{R}^k)$ entier. Pour un S donné vérifiant $S(P) = 0$, le choix de $\tilde{S}$ dépend alors encore de $n(K-n-k+p)$ paramètres arbitraires. Λ et S sont en correspondance biunivoque, mais les $\tilde{S}$ associés à un même Λ dépendent de $n(K+p-n-k)$ paramètres. Les conditions (21) s'expriment alors ainsi : il existe une application $\tilde{S}$ (qui n'est plus déterminée univoquement) telle que l'on ait :

$$(23\text{-}2) \qquad \begin{cases} \Lambda = I - \tilde{S}\, M \\ \tilde{S}\, A = 0 \end{cases}$$

Les matrices $\tilde{S}$ vérifiant $\tilde{S}\, A = 0$ dépendent de $n(K-k)$ paramètres. Mais l'espace des $\tilde{S}$ telles que $\tilde{S}\, A = \tilde{S}\, M = 0$ possède $n(K+p-n-k)$ dimensions, de sorte que l'espace des Λ vérifiant (23-2) est de dimension :

$$n(K-k) - n(K-k+p-n) = n(n-p)$$

C'est bien à ce résultat qu'on devait aboutir, car l'espace des m vérifiant $M\, m \in P$ est de dimension p, et $I - \Lambda$ doit s'annuler sur cet espace (n p conditions).

Condition d'optimalité. - Choisissons maintenant, parmi tous les estimateurs vérifiant la condition d'universalité, celui qui minimise la variance.

Cas 1.- D'après (23-1), on doit déterminer la matrice S^u_α vérifiant :

$$S^u_\alpha A^\ell_u = 0$$

et minimisant pour chaque α la variance de l'estimateur :

$$m^*_\alpha = Z_\alpha - S^u_\alpha M^\beta_u Z_\beta$$

Cette variance est :

$$D^2(m^*_\alpha) = \sigma^2_\alpha - 2\, S^u_{(\alpha)} M^\beta_u \sigma_{(\alpha)\beta} + S^u_{(\alpha)} S^v_{(\alpha)} M^\beta_u \sigma_{\beta\gamma} M^\beta_v$$

(sans sommation sur l'indice muet α). La condition de minimum lié conduit au système :

$$(24) \qquad \begin{cases} S^v_\alpha M^\beta_u \sigma_{\beta\gamma} M^\gamma_v = M^\beta_u \sigma_{\alpha\beta} + \mu_{\ell\alpha} A^\ell_u \\ S^u_\alpha A^\ell_u = 0 \end{cases}$$

Ce système est nécessairement régulier. En effet, l'estimateur Λ est déterminé univoquement par les conditions d'universalité et d'optimalité, et, dans le cas 1 ($K \leq n$), S et Λ se correspondent biunivoquement. On peut le vérifier directement : M étant de rang $\leq n$, la matrice $M\, \sigma\, M$ est inversible, et la première relation

(24) se met sous la forme :

$$S_{\alpha}^{v} = T_{\alpha}^{v} + \mu_{\ell\alpha} R^{uv} A_{u}^{\ell}$$

avec les matrices $T = (M \sigma M)^{-1} M \sigma$ et $R = (M \sigma M)^{-1}$

En portant dans la seconde relation, on trouve :

$$\mu_{\ell\alpha} A_{u}^{\ell} R^{uv} A_{v}^{s} + T_{\alpha}^{v} A_{v}^{s} = 0$$

La matrice A étant de rang $k < K$, A R A est inversible, et les $\mu_{\ell\alpha}$ sont univoquement déterminés.

. La variance de cet estimateur optimal se met sous la forme :

$$(24') \qquad D^2(m_{\alpha}^{*}) = \sigma_{\alpha}^{2} - S_{(\alpha)}^{u} M_{u}^{\beta} \sigma_{(\alpha)\beta}$$

comme on le voit en saturant la première relation (24) par $S_{(\alpha)}^{u}$. On note qu'en raison de la forme particulière de la condition d'universalité $S A = 0$, cette variance optimale ne dépend pas des paramètres de Lagrange $\mu_{\ell\alpha}$.

Cas 2.- Examinons maintenant le cas $n < K$. Les conditions (23-2) expriment que l'estimateur m_{α}^{*} peut encore se mettre sous la forme :

$$m_{\alpha}^{*} = Z_{\alpha} - S_{\alpha}^{u} M_{u}^{\beta} Z_{\beta}$$

avec une matrice S vérifiant $S A = 0$ déterminée à $n(K+p-n-k)$ paramètres arbitraires près. Las conditions d'optimalité conduisent ensuite au même système (24) que ci-dessus. Mais cette fois, M étant de rang n, la matrice $M \sigma M$ n'est plus inversible, et la solution générale S du système (24) dépend de n(K+p-n-k) paramètres arbitraires. Par contre, les $S_\alpha^u M_u^\beta$ et donc l'estimateur lui-même sont déterminés de manière unique. Il suffit évidemment de trouver une solution particulière de (24) pour en déduire m_α^* .