

DES MINES DE
CENTRE de
MONTROUGE
FONTAINEBLEAU

C-76

LE KRIGEAGE DISJONCTIF ET LE
PARAMETRAGE LOCAL DES RESERVES

C-76
G. MATHERON

FONTAINEBLEAU SEPTEMBRE 1978

LE KRIGEAGE DISJONCTIF ET LE PARAMETRAGE LOCAL DES RESERVES

T A B L E

<u>LE KRIGEAGE DISJONCTIF</u>	1
0 - <u>INTRODUCTION</u>	1
1 - <u>MODELE ISOFATORIEL GAUSSIEN</u>	3
Les polynomes d'Hermite	3
Le K.D. isofactoriel	5
Estimation par le K.D. de la loi de Y_0	7
2 - <u>APPLICATION AUX F.A.</u>	8
Améliorations	11
<u>LE PARAMETRAGE LOCAL DES RESERVES</u>	13
1 - <u>K.D. DES FONCTIONS DE RECUPERATION DIRECTE</u>	14
2 - <u>K.D. DES FONCTIONS DE RECUPERATION INDIRECTE</u>	15
3 - <u>AUTRES METHODES</u>	17
a) Conditionnement uniforme sans permanence	18
b) Conditionnement uniforme avec permanence	20

LE KRIGEAGE DISJONCTIF
=====

0 - INTRODUCTION

Pourquoi un nouveau type d'estimateur ? Il s'agissait simplement au départ de combler une lacune existant entre les deux types les plus usuels. Considérons le problème qui consiste à estimer (ou prévoir) la valeur d'une V.A. Z_0 connaissant les valeurs prises par N autres V.A. Z_α ($\alpha = 1, 2, \dots, N$). La pratique usuelle ne retient que deux possibilités :

- ~ régression linéaire ou krigeage
- ~ espérance conditionnelle

Le premier type est simple, facile à mettre en oeuvre et nécessite peu de prérequis (la matrice de covariance). Mais, en contre partie, c'est un estimateur assez fruste, peu précis et assez mal adapté dans le cas de problèmes non linéaires.

En contraste, l'espérance conditionnelle est, en théorie, le meilleur de tous les estimateurs possibles. Mais il faudrait savoir la former, et cela nécessite la connaissance de la loi multi-variable des $N+1$ V.A. Z_0, Z_1, \dots, Z_N : d'où de difficiles problèmes d'inférence statistique, que l'on ne peut résoudre que moyennant des hypothèses très fortes et à peu près incontrôlables. On peut redouter un grave manque de robustesse.

Entre l'estimateur simple, robuste, nécessitant peu d'hypothèses et de prérequis, mais fruste et peu précis (le krigeage) et d'autre part l'estimateur ultra-sophistiqué, théoriquement parfait mais peu robuste et d'ailleurs le plus souvent inaccessible en pratique, (l'espérance conditionnelle) il y avait place pour un moyen terme.

Donnons quelques ordres de grandeur sur l'étendue de la lacune à combler.

Dans le cas du krigeage (avec condition $\sum \lambda^\alpha = 1$) l'espace de travail (l'ensemble des combinaisons linéaires vérifiant cette condition) est de dimension $N-1$. Pour l'espérance conditionnelle, c'est parmi l'ensemble de toutes les fonctions mesurables de N variables que nous souhaitons trouver la meilleure d'entre elles. L'espace de travail est de dimension infinie (théoriquement). En pratique, on est obligé de discrétiser. Si nous retenons seulement 10 classes de valeurs (ce qui est peu) pour chacune des variables, une fonction $f(z_1, z_2, \dots, z_N)$ dépend de 10^N paramètres : telle est la dimension de l'espace de travail.

Le contraste est écrasant. Pour $N = 10$ données, par exemple, la dimension est 9 dans le premier cas, dix milliards dans le second. Pour un physicien, le passage de 1 à 10^9 marque un changement d'échelle, et s'accompagne de l'apparition d'objets ou de phénomènes qualitativement différents. Pour nous, nous nous devons plus modestement mettre en doute notre capacité à appréhender et à optimiser réellement dix milliards de paramètres à partir de quelques centaines ou milliers de données expérimentales. La supériorité écrasante de l'espérance conditionnelle risque donc d'être en partie illusoire. Elle ne se maintient, en apparence, que parce que nous utilisons à fond les propriétés d'un modèle très riche (une loi à $N+1$ variables), mais échappant à toute possibilité de contrôle expérimental.

A titre d'intermédiaire, le Krigeage Disjonctif (K.D.) consiste à rechercher la meilleure approximation de Z_0 par une somme de N fonctions dépendant chacune d'une seule variable :

$$(1) \quad Z_0^* = f_1(Z_1) + \dots + f_N(Z_N)$$

L'espace de travail, dans notre exemple a donc $10 \times N = 100$ dimensions. C'est substantiellement plus riche que

le krigeage (dimension 9) mais nettement plus réaliste que l'espérance conditionnelle (10^{10}).

La mise en oeuvre du K.D. nécessite la connaissance préalable de toutes les lois à deux variables (Z_0, Z_α) et (Z_α, Z_β) , soit $N(N+1)/2$ lois à deux variables : ici encore ces prérequisites sont plus exigeants que pour le krigeage (une simple matrice de covariance) mais beaucoup moins que pour l'espérance conditionnelle (la loi à $N+1$ variables dans sa totalité). On voit, en effet, facilement (théorème des projections) que les fonctions $f_\alpha(Z_\alpha)$ figurant dans l'expression (1) du K.D. sont déterminées par le système d'équations :

$$E[Z_0^* | Z_\alpha] = E[Z_0 | Z_\alpha] \quad (\alpha = 1, 2, \dots, N)$$

Malgré la simplicité de l'écriture, il s'agit d'un système d'équations intégrales plutôt difficile à résoudre. On est donc conduit à rechercher des cas de simplification : tel est la raison d'être des modèles isofactoriels, que nous n'aborderons que dans le cas particulier gaussien.

1 - LE MODELE ISOFACTORIEL GAUSSIEN

Pourvu que sa fonction de répartition F soit connue, toute V.A. Z peut être mise sous la forme $Z = \varphi(Y)$ avec une gaussienne réduite Y et une transformation croissante ou anamorphe. φ convenable. Nos $N+1$ variables Z_0, Z_α dérivent ainsi de $N+1$ V.A. Y_0, Y_α dont chacune, individuellement, est une gaussienne réduite. Nous admettrons (ce qui ne va d'ailleurs nullement de soi) que ces variables prises deux à deux obéissent à des lois de Gauss à 2 variables, caractérisées par leurs coefficients de corrélation $\rho_{0\alpha}$ et $\rho_{\alpha\beta}$: c'est cette hypothèse (fondamentale) qui sert de définition à notre modèle "isofactoriel" gaussien.

Les polynomes d'Hermite

Désignons par g et G la densité et la fonction de répartition de la loi de Gauss réduite :

Orthogonality

$$g^{(1)}(x) = -x g(x) \quad (1)$$

On differentiating $n+1$ times

$$g^{(n+2)} = -x g^{(n+1)} - (n+1) g^{(n)}$$

$$\therefore H_{n+2} = -x H_{n+1} - (n+1) H_n$$

$$\therefore H_{n+2} + x H_{n+1} = -(n+1) H_n \quad (2)$$

Differentiating

$$H_{n+1} = \frac{1}{g} g^{(n+1)}$$

gives

$$H_{n+1}' = \frac{1}{g} g^{(n+2)} + \frac{g^{(1)}}{g^2} g^{(n+1)}$$

$$= \frac{1}{g} g^{(n+2)} + \frac{x}{g} g^{(n)}$$

$$= H_{n+2} + x H_{n+1} \quad (3)$$

Equating (2) & (3) gives

$$-(n+1) H_n = H_{n+1}'$$

Now

$$I = \int H_n H_m g \, dx$$

$$= \int H_n g^{(m)} \, dx$$

$$= \left[H_n g^{(m-1)} \right]_{-\infty}^{\infty} - \int g^{(m-1)} H_n' \, dx$$

$$= 0 + \int n \cdot H_{n-1} g^{(m-1)} \, dx$$

$$= \int_{n-(n-m)}^{n-1} H_{n-m} g \, dx$$

Now if $n \neq m$

$$I = (n-m) \dots \int_{-\infty}^{\infty} g^{(n-m)} \, dx$$

$$= 0$$

if $n = m$

$$I = \frac{n!}{n!} \int g^0 \, dx$$

$$= 1$$

$$g(y) = \frac{1}{\sqrt{2\pi}} \exp(-y^2/2) ; G(y) = \int_{-\infty}^y g(x)dx$$

A cette loi sont étroitement associés les polynomes d'Hermite $H_n(x)$ définis par

$$H_n(x) = \frac{1}{g(x)} \frac{d^n}{dx^n} g(x) \quad \left\{ \begin{array}{l} H_0 = 1 \\ H_1 = -x \\ H_2 = x^2 - 1 \\ \dots \dots \dots \end{array} \right.$$

qui vérifient la propriété d'orthogonalité :

$$\int H_n(x) H_m(x) g(x)dx = 0 \quad (m \neq n)$$

$$\int [H_n(x)]^2 g(x)dx = n!$$

Ceci entraîne pour toute fonction f (telle que $\int f^2(x)dx < \infty$) l'existence d'un développement en polynomes d'Hermite :

$$(1-1) \quad \left\{ \begin{array}{l} f(x) = \sum_{n=0}^{\infty} \frac{f_n}{n!} H_n(x) \\ f_n = \int H_n(x) f(x) g(x) dx \end{array} \right.$$

En ce qui nous concerne, la propriété capitale de ces polynomes est la suivante : si X et Y sont deux gaussiennes réduites et ρ leur coefficient de corrélation, l'espérance conditionnelle de $H_n(X)$ à Y fixé est :

see N 360
for details.

$$(1-2) \quad \boxed{E[H_n(X)|Y] = \rho^n H_n(Y)}$$

Conséquences immédiates : pour toute fonction f (telle que $E[f^2(X)] < \infty$), on déduit du développement (1-1) l'expression suivante de l'espérance conditionnelle :

$$E[f(X)|Y] = \sum \frac{f_n \rho^n}{n!} H_n(Y)$$

Pour $n \neq m$, $H_n(X)$ et $H_m(Y)$ sont orthogonaux. Pour $n = m$, $H_n(X)$ et $H_n(Y)$ ont un coefficient de corrélation égal à ρ^n :

$$(1-2) \quad E[H_n(X) H_m(Y)] = \rho^n n! \delta_{nm}$$

Si f et φ sont deux fonctions de carré intégrables, f_n et φ_n les coefficients de leurs développements en polynômes d'Hermite :

$$E[f^2(X)] = \sum_{n=0}^{\infty} \frac{f_n^2}{n!}$$

$$E[f(X)\varphi(Y)] = \sum_{n=0}^{\infty} \frac{f_n \varphi_n}{n!} \rho^n$$

Comme $f_0 = E[f(X)]$, on obtient variance et covariance en retranchant le terme en $n = 0$, soit :

$$(1-3) \quad \left\{ \begin{array}{l} \text{Var } f(X) = \sum_{n \geq 1} \frac{f_n^2}{n!} \\ \text{Cov}(f(X), \varphi(X)) = \sum_{n \geq 1} \frac{f_n \varphi_n}{n!} \rho^n \end{array} \right.$$

[N.B. : Dans la terminologie de l'analyse des données, les H_n constituent les "facteurs" associés à la loi bigaussienne. C'est pourquoi notre modèle est dit "isofactoriel"].

Le K.D. isofactoriel

Dans ce modèle, le K.D. est très facile à réaliser : tout revient, en effet, à kriger séparément chacun des facteurs H_n . On peut le voir indirectement à partir des équations générales posées plus haut, mais il vaut mieux repartir à 0. Comme nos variables Z initiales s'expriment à l'aide des gaussiennes Y , notre problème peut être reformulé comme suit : étant donné une fonction f de carré sommable, trouver pour $f(Y_0)$ le meilleur estimateur de la forme

$$f^* = \sum_{\alpha} f_{\alpha}(Y_{\alpha})$$

Comme les fonctions f_{α} admettent des développements du type (1-1), l'estimateur f^* est une combinaison linéaire des divers $H_n(Y_{\alpha})$ ($n = 0, 1, \dots$ et $\alpha = 1, 2, \dots, N$). Pour $n = 0$ les divers $H_0(Y_{\alpha})$ sont tous égaux à 1 et doivent être regroupés en

un terme unique (constant). Notons aussi que f admet elle-même un développement de type (1-1) : les projecteurs étant des opérateurs linéaires, il nous suffira de former le K.D. H_n^* de chaque facteur $H_n(Y_0)$ pour en déduire :

$$(1-4) \quad f^* = f_0 + \sum_{n \geq 1} \frac{f_n}{n!} H_n^*$$

(il est clair que $H_0^* = 1$, puisque $H_0(Y_0) = 1$, de sorte que le terme en $n = 0$ est nécessairement l'espérance $f_0 = E[f(X)]$).

Or $H_n(Y_0)$ est orthogonal à tous les $H_m(Y_\alpha)$ pour $m \neq n$, comme nous l'avons vu. Le meilleur estimateur que nous puissions en former sera donc une combinaison linéaire finie des N facteurs de même degré n , soit :

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

Il reste à exprimer que cet estimateur minimise la variance d'estimation : on obtient donc le système habituel du krigeage simple, écrit avec les coefficients de corrélation $\rho_{\alpha\beta}^n$ (cf. relation (1-2')), soit :

$$(1-5) \quad \left\{ \begin{array}{l} \sum_{\alpha} \lambda_n^\alpha \rho_{\alpha\beta}^n = \rho_{\alpha\beta}^n \\ \sigma_{K_n}^2 = n! [1 - \lambda_n^\alpha \rho_{0\alpha}^n] \end{array} \right.$$

Ainsi, le K.D. de H_n est un krigeage simple effectué sur les facteurs de même degré, et le K.D. de $f(Y_0)$ s'en déduit selon (1-4). En ce qui concerne la variance d'estimation sur $f(Y_0)$, on trouve

$$\sigma_K^2(f) = \sum_{n \geq 1} \frac{f_n^2}{(n!)^2} \sigma_{K_n}^2$$

puisque les facteurs de degré différent sont toujours orthogonaux.

Ainsi, en pratique, on doit kriger séparément chacun des facteurs H_n , opération simple et facilement réalisable. Le nombre de krigeages à effectuer dépend, évidemment, de la précision souhaitée et de la rapidité plus ou moins grande de la convergence de l'estimateur. Notons que la série (1-4) définissant cet estimateur converge toujours beaucoup plus vite que la série originelle (1-1), de sorte qu'en pratique on peut en général se limiter à un nombre très modeste de facteurs (5 ou 10 au plus).

En effet, dans la matrice $\rho_{\alpha\beta}^n$ du système (1-5), les termes non diagonaux deviennent négligeables lorsque n est grand, et il reste $\lambda_n^\alpha = \rho_{0\alpha}^n$, d'où la rapidité de la convergence. Remarquons de plus que pour des valeurs de n relativement peu élevées on a une approximation très suffisante en prenant :

$$\lambda_n^\alpha \approx \rho_{0\alpha}^n - \rho_{0\alpha}^n \sum_{\beta \neq \alpha} \rho_{\alpha\beta}^n$$

C'est donc seulement pour les toutes premières valeurs de n (4 ou 5 en pratique) que l'on a une résolution numérique effective à réaliser.

Estimation par K.D. de la loi de Y_0 aux Y_α fixés

D'après l'expression (1-1) des coefficients φ_n du développement en polynome d'Hermite d'une fonction φ , le K.D. φ^* de $\varphi(Y_0)$, qui est $\sum(\varphi_n/n!)H_n^*$ d'après (1-4), peut se mettre sous la forme :

(1-6)
$$\varphi^* = \int_{-\infty}^{+\infty} f_{KD}(y)\varphi(y)dy$$

avec :

(1-7)
$$f_{KD}(y) = \sum_{n=0}^{\infty} \frac{H_n^* H_n(y)}{n!} g(y)$$

(la remarque faite ci-dessus sur le comportement des λ_n^α pour $n \rightarrow \infty$ garantit la convergence de cette série).

Cette "densité" f_{KD} constitue l'estimation, par K.D., de la loi de Y_0 lorsque les Y_α sont connus. Elle permet, en effet, d'effectuer le calcul du K.D. de toute fonction $\varphi(Y_0)$ comme s'il s'agissait de l'espérance de $\varphi(Y)$ avec une V.A. Y admettant la densité f_{KD} .

Exercice proposé : Soit $F_{KD}(y)$ le krigeage disjonctif de $1_{Y_0 < y}$. Montrer que l'on a :

$$F_{KD}(y) = G(y) + \sum_{n \geq 1} \frac{H_n^*}{n!} H_{n-1}(y)g(y)$$

et retrouver (1-7) [former le développement de $1_{Y_0 < y}$].

Remarque : Cette "densité" $f_{KD}(y)$ n'est pas obligatoirement ≥ 0 pour tout y . Il faut voir là la triste conséquence d'un théorème général selon lequel, dans nos espaces de travail, les seuls projecteurs qui garantissent cette positivité sont les espérances conditionnelles. (Le krigeage usuel respecte encore moins cette positivité). En pratique, on constate que ces valeurs négatives sont suffisamment rares et d'amplitude assez faible pour qu'on puisse les tolérer. Elles n'apparaissent, le plus souvent, que pour des valeurs de y très grandes, en valeurs absolues, et n'altèrent donc légèrement que les queues de la distribution.

2 - APPLICATION AUX F.A.

Ce qui précède ne constitue pas, semble-t-il, un argument décisif en faveur du K.D., puisque, dans le cadre du même modèle, l'espérance conditionnelle ne serait pas beaucoup plus difficile à former. Mais les choses vont changer lorsqu'il va s'agir d'opérations répétitives : le K.D. manifeste une puissance de récapitulation qui lui confère (en temps de calcul) un avantage écrasant sur l'espérance conditionnelle.

Prenons un exemple - Soit un phénomène dont nous désirons étudier la zonéographie. Supposons que ce phénomène soit bien

représenté par une F.A. stationnaire $Z(x)$. La surface de l'ensemble $\{Z(x) \geq z\}$ limitée par l'isoteneur z est, en moyenne, proportionnelle à $1-F(z)$. Mais nous pouvons aussi nous intéresser aux aspects locaux du phénomène, par exemple chercher à estimer dans un domaine S donné la fraction :

$$\frac{1}{S} \int_S 1_{\{Z(x) \geq z\}} dx$$

occupée par les teneurs $\geq z$ (estimation d'un "histogramme local").

Admettons que, moyennant une anamorphose $Z(x) = \varphi(Y(x))$, notre F.A. se ramène à un modèle multi-gaussien stationnaire $Y(x)$. Il s'agit donc d'estimer des expressions du type :

$$\frac{1}{S} \int_S 1_{Y(x) \geq y} dx$$

ou, plus généralement :

$$\frac{1}{S} \int_S f(Y(x)) dx$$

Pour chaque x , nous pouvons former la loi conditionnelle de $Y(x)$ lorsque les $Y(x_\alpha) = Y_\alpha$ sont connus : c'est (dans ce modèle) une loi de Gauss de moyenne $m(x) = \lambda^\alpha(x) Y_\alpha$ dépendant de x , dont la variance $\sigma^2(x)$ dépend aussi de x : il faut donc former explicitement pour chaque x (ou au moins sur une grille assez fine dans S) l'expression de l'espérance conditionnelle de $f(Y(x))$ et sommer en x : procédure longue et ennuyeuse.

Avec le K.D., au contraire, nous pouvons effectuer l'opération en une seule fois. De fait, utilisons le développement :

$$f(Y(x)) = \sum \frac{f_n}{n!} H_n(Y(x))$$

Au lieu de former le K.D. de $H_n(Y(x))$ pour chaque x et de sommer ensuite, nous pouvons profiter de la linéarité du

second membre de (1-5) pour faire l'opération d'un seul coup :
le K.D. H_n^* de $1/S \int_S H_n(Y_x) dx$ est, en effet :

$$H_n^* = \lambda_n^\alpha H_n(Y_\alpha)$$

les λ_n^α étant déterminés par le système :

$$\sum_{\alpha} \lambda_n^\alpha \rho_{\alpha\beta}^n = \frac{1}{S} \int_S \rho_{x\beta}^n dx$$

L'estimateur cherché de $(1/S) \int_S f(Y_x) dx$ sera ensuite donné par :

$$\sum \frac{f_n}{n!} H_n^*$$

exactement comme au paragraphe précédent. On peut également associer à cette opération "une "densité"

$$f_{KD}(y) = \sum \frac{H_n^* H_n(y)}{n!} g(y)$$

qui, une fois connue, permet le K.D. de toutes les intégrales du type précédent. En particulier, l'estimateur de l'"histogramme local"

$$1-F_S(z) = \frac{1}{S} \int_S 1_{Z(x) \geq z} dx$$

sera

$$1-F_S(z) = \int_y^\infty f_{KD}(x) dx = 1-G(y) - \sum_{n \geq 1} \frac{H_n^* H_{n-1}(y)}{n!} g(y)$$

(avec un y tel que $z = \varphi(y)$, φ représentant l'anamorphose gaussienne $Z(x) = \varphi(Y_x)$).

Dans la pratique, le problème crucial est de choisir l'anamorphose gaussienne φ , c'est-à-dire, ce qui revient au même, d'estimer la loi $F(z) = P(Z(x) \leq z)$ à partir des valeurs expérimentales $Z_\alpha = Z(Y_\alpha)$. Nous ne développons pas ce point ici. Notons seulement

ceci : lorsque les données ne sont pas à maille régulière, il faut impérativement pondérer la fréquence de chacune d'elles par un poids tenant compte de l'étendue de sa "zone d'influence". Par exemple, on formera l'estimateur :

$$F^*(z) = \sum_{\alpha} \lambda^{\alpha} 1_{Z_{\alpha} \leq z}$$

en prenant pour λ^{α} les coefficients du krigeage de la moyenne générale (avec $\sum \lambda^{\alpha} = 1$).

Améliorations : La démarche qui précède repose sur une confiance peut être exagérée en notre modèle (stationnarité, anamorphose φ , binormalité). Il y a parfois intérêt à relâcher un peu cette dépendance excessive à l'égard du modèle et à augmenter la robustesse de la démarche, moyennant quelques sacrifices du côté de la précision.

Par exemple, on peut imposer aux coefficients λ_n^{α} la condition de non biais $\sum \lambda_n^{\alpha} = 1$: on retrouve le système du krigeage usuel

$$(1-5') \quad \begin{cases} \lambda_n^{\alpha} \rho_{\alpha\beta}^n = \rho_{0\beta}^n + \mu_n & (\text{ou } \frac{1}{S} \int_S \rho_{x\beta}^n dx + \mu_n) \\ \sum \lambda_n^{\alpha} = 1 \end{cases}$$

La variance d'estimation est légèrement augmentée. En contre partie, la condition de non biais $E[f^*] = E[f]$ est garantie, même si nous nous sommes lourdement trompés sur le choix de l'anamorphose φ (c'est à dire sur l'estimation de la loi F de $Z(x)$). En effet, si notre anamorphose φ est biaisée, les Y_x ne sont plus des gaussiennes réduites mais obéissent à une loi G' quelconque. Alors avec cette loi G' , on a :

$$E[H_n^*] = \sum_{\alpha} \lambda_n^{\alpha} \int H_n(y_{\alpha}) G'(dy_{\alpha}) = \int H_n(y) G'(dy) = E[H_n(Y_0)]$$

à cause de la condition $\sum \lambda_n^{\alpha} = 1$, et donc aussi $E[f^*] = E[f]$ pour toute fonction f , quelle que soit la loi réelle de nos variables.

(on reste, cependant, tributaire de la stationnarité).

On peut, sur cette base, définir une procédure itérative pour améliorer l'estimation de la loi F et de l'anamorphose φ . A partir d'une première estimation de F (obtenue, par exemple, en pondérant les fréquences des données par les poids du krigeage) on construit un premier modèle et une première anamorphose φ_1 . Dans le cadre de ce premier modèle, on procède alors selon le système (1-5') au K.D. des facteurs $\frac{1}{V} \int_V H_n(Y_x) dx$, V représentant le champ total : on obtient ainsi une estimation $f_{KD}(y)$ de la loi g théorique : ce qui permet d'améliorer l'anamorphose initiale φ_1 , en la remplaçant par :

$$\varphi_2(y) = \varphi_1[F_{KD}^{-1}(G(y))]$$

Cette procédure est, cependant, assez lourde.

LE PARAMETRAGE LOCAL DES RESERVES

Nous avons vu comment le modèle gaussien discrétisé permet de procéder au paramétrage global des réserves récupérables d'un gisement minier. A un stade plus avancé, on souhaite disposer d'un paramétrage analogue, mais au niveau local et conditionnellement à l'information disponible au moment de l'estimation (les $Z_\alpha = \varphi(Y_\alpha)$). Appelons blocs v les unités sur lesquelles portera la sélection ultime du minerai, $Z(v)$ la teneur d'un bloc v et $Z^*(v)$ son estimateur ultime qui servira de critère à cette sélection ($Z^*(v)$ sera formée sur la base d'une information beaucoup plus riche que l'information actuelle (les Z_α) et n'est donc pas numériquement connu au moment où l'on procède au paramétrage).

Les blocs v sont très nombreux (plusieurs milliers ou dizaines de milliers par exemple) et l'information actuelle (les Z_α) ne permet pas de différencier de manière significative deux blocs v voisins dans l'espace (par exemple : les blocs seront des cubes $5 \times 5 \times 5 \text{ m}^3$, et les Z_α des teneurs de sondages implantés à maille $100 \times 100 \text{ m}^2$). On renoncera donc à estimer les blocs v individuellement, et on les regroupera en panneaux V . (dans notre exemple : les panneaux V correspondront aux carrés 100×100 centrés sur les sondages, ou admettant ces sondages comme sommets).

Pour simplifier l'écriture, nous prendrons $V = 1$ (unité de tonnage égale au tonnage total du panneau V). Nous souhaitons estimer, conditionnellement à l'information actuelle :

~ le pourcentage des blocs $v \subset V$ dont la teneur sera $\geq z$, et la teneur moyenne de ces blocs (effet de support). Bien que cette terminologie soit discutable, on parle dans ce cas des fonctions de transfert directes (fonction de récupération serait peut être meilleur)

~ en général, il faut tenir compte également de l'effet d'information, puisque l'on coupera sur les estimateurs ultimes $Z^*(v)$ et non sur les teneurs réelles. Il s'agit alors d'estimer, en pourcentage, le nombre des blocs $v \subset V$ solutionnés à $Z^*(v) \geq z$, ainsi que leur teneur moyenne (réelle). On parlera dans ce cas de fonctions de transfert (ou de récupération) indirectes.

1 - K.D. DES FONCTIONS DE RECUPERATION DIRECTE

Si on utilisait les lois conditionnelles, on devrait calculer une à une ces lois pour chacun des blocs $v \subset V$ pour ne retenir ensuite que la moyenne (arithmétique) de ces différentes lois : procédure longue et coûteuse, encore que parfaitement réalisable. Il est plus économique d'utiliser la capacité récapitulative du K.D. pour obtenir en une seule fois le résultat global qui seul nous intéresse.

Dans notre modèle discrétisé, nous connaissons l'anamorphose $Z(v) = \varphi_r(Y_v)$ des blocs, avec :

$$\varphi_r(y) = \sum \frac{r^n}{n!} H_n(y)$$

ainsi que le coefficient de corrélation $\rho_{v_i, \alpha}$ de chaque bloc $v_i \subset V$ avec la donnée Y_α . Formons le K.D. de l'expression :

$$\frac{1}{N} \sum_i H_n(Y_{v_i})$$

(sommation sur les N blocs $v_i \subset V$). Il est de la forme :

$$H_n^* = \sum_\alpha \lambda_n^\alpha H_n(Y_\alpha)$$

avec des coefficients λ_n^α fournis par le système :

$$\lambda_n^\alpha \rho_{\alpha\beta}^n = \frac{1}{N} \sum_i \rho_{\alpha v_i}^n$$

Soit y la valeur de la gaussienne réduite associée à la coupure z , c'est-à-dire telle que $z = \varphi_r(y)$. Les quantités à estimer

sont :

$$\frac{1}{N} \sum_i 1_{Y_{v_i} \geq y} \quad \text{et} \quad \frac{1}{N} \sum_i \varphi_r(Y_{v_i}) 1_{Y_{v_i} \geq y}$$

Pour la première, le K.D. fournit l'estimateur :

$$T_v(z) = 1 - G(y) - \sum_{n \geq 1} \frac{H_n^*}{n!} H_{n-1}(y) g(y)$$

Pour le second, il est plus commode d'introduire la "densité"

$$f_{KD}(y) = \sum_{n=0}^{\infty} \frac{H_n^*}{n!} H_n(y) g(y)$$

Il vient alors :

$$(1-1) \quad \left\{ \begin{array}{l} T_v(z) = \int_y^{\infty} f_{KD}(x) dx \\ Q_v(z) = \int_y^{\infty} \varphi_r(x) f_{KD}(x) dx \end{array} \right.$$

Telles sont les fonctions de récupération directe (ou du moins leur K.D.).

2 - K.D. DES FONCTIONS DE RECUPERATION INDIRECTE

Nous allons maintenant tenir compte de l'effet d'information, et désigner par $h(z)$ l'espérance conditionnelle de $Z(v)$ lorsque l'estimateur ultime $Z^*(v)$ est égal à z . Nous ferons également l'approximation suivante : lorsque l'on connaît $Z^*(v) = z$ et les Z_α , l'espérance conditionnelle de $Z(v)$ est encore $h(z)$ (n'est pas influencée par la donnée supplémentaire des Z_α). En toute rigueur, cette hypothèse n'est correcte que si l'estimateur ultime est déjà lui-même une espérance conditionnelle (celle de $Z(v)$ lorsque l'information ultime - qui contient les Z_α - est connue). En pratique, $Z^*(v)$ est plutôt une simple combinaison linéaire (un krigeage par exemple), mais approche $Z(v)$ avec une précision tellement plus grande qu'il n'est possible de le faire sur la base des données actuelles (les Z_α) que ces dernières ne peuvent guère modifier beaucoup l'espérance conditionnelle de $Z(v)$ à $Z^*(v)$ fixé,

de sorte qu'il est légitime de poser en lère approximation :

$$E[Z(v)/Z^*(v) = z, Z_\alpha] = E[Z(v)/Z^*(v) = z] = h(z)$$

Dans notre modèle discrétisé, nous avons associé aux estimateurs ultimes $Z^*(v)$ l'anamorphose φ_{r^*} et la gaussienne réduite Y_v^* , et nous avons appris à calculer les coefficients de corrélation :

$$\sim \rho = \rho_{v v^*} \text{ entre } Y_v \text{ et } Y_v^*$$

$\sim \rho_{\alpha i}^*$ entre Y_α (donnée ponctuelle) et $Y_{v_i}^*$ (estimateur ultime du bloc $v_i \subset V$), et établi la relation :

$$h(Z_v^*) = \varphi_{r\rho}(Y_v^*) = \sum \frac{C_n}{n!} (r\rho)^n H_n(Y_v^*)$$

Le reste en découle.

On formera d'abord le K.D. de $(1/N) \sum_i H_n(Y_{v_i}^*)$, soit

$$H_n^* = \sum_\alpha \lambda_n^\alpha H_n(Y_\alpha)$$

$$\sum_\alpha \lambda_n^\alpha \rho_{\alpha\beta}^n = \frac{1}{N} \sum_i (\rho_{\alpha i}^*)^n$$

D'où le tonnage récupérable à $Z^*(v) \geq z$:

$$(2-1) \quad \left\{ \begin{array}{l} z = \varphi_{r^*}(y) \\ F_{v^*}(z) = 1 - G(y) - \sum_{n \geq 1} \frac{H_n^*}{n!} H_{n-1}(y) g(y) \end{array} \right.$$

et, en introduisant la "densité"

$$f_{KD}(y) = \sum_{n=0}^{\infty} \frac{H_n^*}{n!} H_n(y) g(y)$$

la quantité de métal récupérable :

$$(2-2) \quad Q_{v^*}(z) = \int_y^{\infty} \varphi_{r\rho}(x) f_{KD}(x) dx$$

Remarque : Dans (1-1), la fonction φ_r à intégrer pour obtenir le métal Q_v est celle là même qui relie y et la coupure $z = \varphi_r(y)$.

Cette circonstance ne subsiste plus dans le cas de la récupération indirecte : la fonction que l'on doit intégrer pour obtenir le métal Q_{v^*} , soit φ_{pr} n'est plus la même que celle qui relie y et la coupure $z = \varphi_{r^*}(y)$. Cela résulte, évidemment, de ce que les blocs vérifiant $Z^*(v) = z$ ont la teneur moyenne conditionnelle $h(z) = \varphi_{pr}(z)$ et non $z = \varphi_{r^*}(y)$.

On peut éviter ce désagrément en remplaçant l'estimateur ultime $Z^*(v)$ par $H = E[Z(v)/Z^*(v)] = h(Z^*(v))$. En coupant cette fois à $H \geq z$ (et non plus à $Z^*(v) \geq z$), on retrouve des formules du type (1-1) :

$$(2-3) \quad \left| \begin{array}{l} z = \varphi_{pr}(y) \\ T_h(z) = \int_y^\infty f_{KD}(x) dx \\ Q_h(z) = \int_y^\infty \varphi_{pr}(x) f_{KD}(x) dx \end{array} \right.$$

Cette présentation est plus réaliste, puisque c'est la coupure sur H (et non sur $Z^*(v)$) qui se prête le mieux à la recherche de l'optimum économique.

3 - AUTRES METHODES

Nous avons déjà mentionné la méthode qui consiste à former chacune des lois conditionnelles de chacun des blocs $v_i \subset V$. Théoriquement meilleure que le K.D. (en pratique, vraisemblablement, à peu près équivalente), cette méthode présente surtout le désavantage d'être beaucoup plus onéreuse en temps calcul. On peut, en sens inverse, chercher plutôt des méthodes moins onéreuses. Nous en mentionnerons deux seulement, de valeur inégale, qui ont été effectivement utilisées surtout dans le cas de gisements lognormaux.

a) Conditionnement uniforme (sans hypothèse de "permanence" pour les panneaux).

La longueur de la méthode basée sur les lois conditionnelles des blocs provient surtout du fait que l'espérance conditionnelle de chaque Y_{v_i} , $v_i \subset V$ est de la forme $\sum \lambda^\alpha(v_i) Y_\alpha$, avec des poids $\lambda^\alpha(v_i)$ qui dépendent de la position du bloc v_i dans le panneau V , et nécessite donc la résolution d'autant de systèmes de krigeage qu'il y a de blocs v_i dans V .

On réaliserait donc une économie substantielle, dans la mise en oeuvre de cette méthode, en remplaçant ces combinaisons linéaires variables par une seule combinaison $\sum \lambda^\alpha Y_\alpha$ fixée une fois pour toutes, (la même pour tous les blocs v_i). On renonce ainsi à l'information différentielle que nous apporte chacun des sondages Y_α , compte tenu de son implantation relativement à chacun des blocs v_i , et l'on ne retient qu'un résumé global, indifférencié, de toute cette information, à savoir l'unique combinaison linéaire $\sum \lambda^\alpha Y_\alpha$. Cette perte d'information risque d'être assez considérable lorsque certains des sondages Y_α tombent à l'intérieur du panneau V , et donc rencontrent certains de nos blocs v_i . On peut penser qu'elle reste, au contraire, relativement faible si tous les sondages sont extérieurs à V .

Désignons alors par Y la gaussienne réduite associée à la combinaison linéaire choisie, soit :

$$Y = \frac{\sum \lambda^\alpha Y_\alpha}{\sqrt{\sum \lambda^\alpha \lambda^\beta \rho_{\alpha\beta}}}$$

et par ρ_i le coefficient de corrélation de Y et Y_{v_i} (que l'on calcule sans difficulté dans le modèle discrétisé). On a alors :

$$E[H_n(Y_{v_i}) | Y] = \rho_i^n H_n(Y)$$

de sorte que l'on peut former sans difficulté l'espérance de n'importe quelle expression du type $\frac{1}{N} \sum_i f(Y_{v_i})$ conditionnée par l'unique variable Y . Posant :

$$H_n^* = \left(\frac{1}{N} \sum_i \rho_i^n \right) H_n(Y)$$

et formant la "densité" :

$$\tilde{f}(y) = \sum_{n \geq 0} \frac{H_n^*}{n!} H_n(y) g(y)$$

on obtient les estimations :

$$\left\{ \begin{array}{l} T_V(z) = \int_y^\infty \tilde{f}(x) dx \\ Q_V(z) = \int_y^\infty \varphi_r(x) \tilde{f}(x) dx \\ (z = \varphi_r(y)) \end{array} \right.$$

pour les fonctions de transfert direct.

Pour les fonctions de récupération indirecte, on a un schéma analogue. Désignons par ρ_i^* le coefficient de corrélation (que l'on sait calculer) entre Y et Y_{vi}^* , et par ρ celui de Y_v et Y_v^* : on prendra cette fois

$$H_n^* = \left(\frac{1}{N} \sum_i (\rho_i^*)^n \right) H_n(Y)$$

et on formera la densité \tilde{f} comme ci-dessus. Alors, en coupant sur $Z^*(v) \geq z$:

$$\left\{ \begin{array}{l} T_{V^*}(z) = \int_y^\infty \tilde{f}(x) dx \\ Q_{V^*}(z) = \int_y^\infty \varphi_{r^*}(x) \tilde{f}(x) dx \\ z = \varphi_{r^*}(y) \end{array} \right.$$

Si, au contraire, on remplace $Z^*(v)$ par $H = h(Z^*(v))$, et si on coupe en $H \geq z$:

$$\left\{ \begin{array}{l} T_h(z) = \int_y^{\infty} \tilde{f}(x) dx \\ Q_h(z) = \int_y^{\infty} \varphi_{pr}(x) \tilde{f}(x) dx \\ z = \varphi_{pr}(y) \end{array} \right.$$

Il reste à choisir les coefficients λ^α de la combinaison linéaire Y. Une méthode simple (suffisante en pratique) consiste à prendre les poids λ^α du krigeage (ordinaire) du panneau V.

Une méthode plus élaborée (mais peu en accord avec le but poursuivi, qui est la recherche de simplifications) consisterait à choisir les λ^α qui minimisent la variance d'estimation de telle ou telle expression retenue par nous pour son importance économique : par exemple, si l'on sait que la coupure réelle sera (à peu près) $z = z_0$, minimiser la variance d'estimation de $Q_h(z_0)$, ou (mieux) de la grandeur $Q_h(z_0) - z_0 T_h(z_0)$ qui représente la valeur du panneau V (dans l'hypothèse d'une formule de valorisation linéaire).

[Exercice : former les expressions correspondantes].

En résumé, et sous réserve d'accepter la perte d'information qu'elle implique comme rançon des simplifications qu'elle apporte, cette méthode paraît absolument correcte : son mérite principal est de ne nécessiter aucune hypothèse supplémentaire (les formules de "permanence", c'est-à-dire la loi φ_r des anamorphoses ne sont utilisées qu'au niveau des blocs v, et non des panneaux V).

b) Conditionnement uniforme avec hypothèse de permanence pour les panneaux :

Dans ce deuxième groupe de méthode, on introduit (implicitement) des hypothèses beaucoup plus fortes. Le point de départ intuitif est le suivant : dans le cas lognormal, il semble "naturel" d'admettre que la loi conditionnelle des teneurs des blocs $v \subset V$ extraits d'un panneau V est encore lognormal lorsque la teneur du panneau V, ou même simplement l'estimation $Z^*(v)$ que l'on peut en

en former est fixée : intuition simpliste, qui, en fait, revient à admettre la validité de la permanence de la lognormalité non seulement pour les supports v (blocs) mais aussi pour les supports V (panneaux).

Formulons cette hypothèse de manière plus explicite.

Si \underline{v} désigne l'un quelconque des N blocs v dont la réunion constitue le panneau V , supposé choisi au hasard avec une loi de probabilité uniforme parmi les N blocs $v_i \subset V$, nous admettons que la gaussienne réduite correspondante $Y_{\underline{v}}$ et les N gaussiennes conditionnantes Y_{α} obéissent encore à une loi $(N+1)$ -gaussienne (une telle hypothèse équivaut, évidemment, à étendre à des domaines de la taille des panneaux V l'hypothèse des "permanences", c'est-à-dire la loi φ_r des anamorphoses). Pour achever de déterminer cette loi, il faut se donner les coefficients de corrélation $\rho_{\underline{v},\alpha}$ entre $Y_{\underline{v}}$ et Y_{α} : comme $S_{\underline{v},\alpha} = S_{V\alpha}$, covariance de $Z(V)$ et de Z_{α} , ces coefficients sont déterminés par la relation :

$$S_{V\alpha} = \sum_{n \geq 1} \frac{C_n^2}{n!} r^n (\rho_{\underline{v},\alpha})^n$$

Moyennant cette hypothèse, le reste est affaire de calculs simples : les Y_{α} étant fixés, on évalue la loi conditionnelle de $Y_{\underline{v}}$ à Y_{α} donné, compte tenu des valeurs trouvées pour les $\rho_{\underline{v},\alpha}$, et on en déduit sans difficulté les expressions $T_{\underline{v}}(z)$ et $Q_{\underline{v}}(z)$ de l'espérance conditionnelle du nombre et de la quantité de métal des blocs sélectionnés dans V à $Z(\underline{v}) \geq z$. On a une formulation analogue en ce qui concerne les fonction de récupération indirectes.

Je ne développe pas davantage, parce que (me semble-t-il) ces méthodes du type b) ne sont pas, réellement, tellement plus simples que les précédentes (type a)), mais présentent par contre le désavantage considérable de supposer valable jusqu'au niveau des panneaux V l'hypothèse de "permanence", c'est-à-dire la validité de la loi φ_r pour les anamorphoses. Or, lorsque les teneurs expérimentales (quasi ponctuelles) sont, par exemple, lognormales, on

peut sans grand problème admettre que la lognormalité reste valable pour des blocs de $5 \times 5 \times 5 \text{ m}^3$, mais il est vraisemblable que cette lognormalité sera déjà beaucoup moins bien vérifiée au niveau des panneaux V de (par exemple) $100 \times 100 \text{ m}^2$. Autrement dit : autant les méthodes de type a) ci-dessus me semblent admissibles (à titre de simplification d'une procédure plus lourde, avec, en contre partie, l'acceptation d'une certaine perte d'information), autant les procédures de type b) (auxquelles on se réfère parfois sous la terminologie ambiguë de "krigeage lognormal") me paraissent discutables, dans la mesure où elles supposent que l'on étende jusqu'à la taille de grands panneaux V des hypothèses de "permanence", admissibles pour les petits blocs v mais beaucoup plus difficilement pour les panneaux V .

Polynômes orthogonaux (Cramér: Mathematical methods of statistics §12.6)

$F(x)$ F. Repart. ayant des moments finis de tous ordres α , x_0 pt de croissance de $F(x) \Leftrightarrow F(x_0+h) > F(x_0-h) \quad \forall h > 0$

Thm $\{x_0\}$ infini \Rightarrow Th: \exists suite de polynômes $p_0(x), p_1(x), \dots, p_n(x)$ déterminés par les 2 conditions suivantes:
 a) $p_n(x)$ est de $D^{\circ} n$ et le coeff de x^n dans p_n est > 0
 b) On a: $\int_{-\infty}^{+\infty} p_m(x) p_n(x) dF(x) = \begin{cases} 1 & \text{pour } m=n \\ 0 & \text{'' } m \neq n \end{cases}$

On considère la forme quadratique:

$$Q = \int_{-\infty}^{+\infty} (u_0 + u_1 x + u_2 x^2 + \dots + u_n x^n) dF(x) = \sum_{i,j=0}^n \alpha_{i+j} u_i u_j$$

Q est définie positive: en effet Q semi-définie positive (ou ident par définition)

$F(x)$ a au moins $n+1$ pts de croissance, donc au moins un de ces pts n'est pas zéro de $u_0 + u_1 x + u_2 x^2 + \dots + u_n x^n$
 $\Rightarrow Q > 0$ dès que $(u_i \neq 0 \forall i)$

$$\Rightarrow D_n = \begin{vmatrix} \alpha_0 & \alpha_2 & \dots & \alpha_n \\ \alpha_2 & \alpha_4 & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_n & \alpha_{2n} & \dots & \alpha_{2n} \end{vmatrix} > 0$$

Nécess $p_0(x) = 1$

$$p_n(x) = u_0 + u_1 x + u_2 x^2 + \dots + u_n x^n$$

Calculer u_i à partir de a) et b)

$p_i(x)$ doit avoir exact^l le $\delta^{\circ} i \Rightarrow x^i$ peut se représenter comme comb. linéaire des $p_i(x) \quad i=0, \dots, i$

$$\Rightarrow \int_{-\infty}^{+\infty} x^i p_n(x) dF(x) = 0 \quad \text{pour } i=0, \dots, n-1$$

On intègre $\int_{-\infty}^{+\infty}$ n'équations linéaires et homogènes à $n+1$ inconnues. Donc $p_n(x)$ devra avoir la forme:

$$p_n(x) = K \begin{vmatrix} \alpha_0 & \dots & \alpha_n \\ \vdots & \ddots & \vdots \\ \alpha_{n-1} & \dots & \alpha_{2n} \\ 1 & x & x^n \end{vmatrix}$$

Si $K \neq 0$, $p_n(x)$ a $D^{\circ} n$ et le coeff de x^n est $D_{n-1} > 0$
 K se détermine par la condition $\int p_n^2 dF = 1$

Si $F(x)$ a un nbre fini N de x_0 , $p_n(x)$ existent et sont déterminés de manière unique $\forall n: 0, 1, \dots, N-1$ (voir démonst.)
 Les déterminants D_n sont > 0 , mais pour $n \geq N$, $D_n = 0$

polynômes d'Hermite

Définis par les relations:

$$\left(\frac{d}{dx}\right)^n e^{-x^2/2} = (-1)^n H_n(x) e^{-x^2/2}$$

$H_n(x)$ est un polynôme de degré n .

$$H_0(x) = 1$$

$$H_3(x) = x^3 - 3x$$

$$H_1(x) = x$$

$$H_4(x) = x^4 - 6x^2 + 3$$

$$H_2(x) = x^2 - 1$$

$$H_5(x) = x^5 - 10x^3 + 15x$$

$$\begin{aligned} \langle m | n \rangle &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} H_m(x) H_n(x) e^{-x^2/2} dx = \frac{(-1)^m}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} H_m(x) \left(\frac{d}{dx}\right)^n e^{-x^2/2} dx \\ &= \frac{(-1)^m}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} H_m(x) \left(\frac{d}{dx}\right)^{n-1} e^{-x^2/2} dx - \frac{(-1)^m}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} H'_m(x) e^{-x^2/2} H_{n-1}(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} H'_m(x) H_{n-1}(x) e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} H_{m-1}^{(m)}(x) H_{n-1}(x) e^{-x^2/2} dx \\ &= R \frac{(-1)^{n-m}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left(\frac{d}{dx}\right)^{n-m} e^{-x^2/2} dx \\ &= \frac{R (-1)^{n-m}}{\sqrt{2\pi}} \left[\left(\frac{d}{dx}\right)^{n-(m+1)} e^{-x^2/2} \right]_{-\infty}^{+\infty} = 0 \end{aligned}$$

Donc $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} H_m(x) H_n(x) e^{-x^2/2} dx = 0$ si $m \neq n$

Si $m = n$: $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} H_n^2(x) e^{-x^2/2} dx = \frac{R}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-x^2/2} dx = R$

R s'obtient en dérivant n fois $H_n(x)$

Coeff. de x^n : a_n

$$\left(\frac{d}{dx}\right)^n e^{-x^2/2} = (-1)^n H_n(x) e^{-x^2/2} = (-1)^n (a_n x^n + \dots) e^{-x^2/2}$$

$H_{n+1}(x) \rightarrow$ coeff de x^{n+1} : $a_{n+1} = a_n = a_{n-1} = \dots = a_0 = 1$

$$\Rightarrow R = n!$$

Donc $\left\{ \frac{1}{\sqrt{n!}} H_n(x) \right\}$ est la suite de polynômes orthogonaux associés à la distribution normale.

Relation de récurrence:

$H_n(x) - x H_{n-1}(x)$ est de degré $< n$

Donc

$$\langle H_n - x H_{n-1}, H_n \rangle = 0$$

$$\Rightarrow \langle H_n, H_n \rangle = \langle x H_{n-1}, H_n \rangle \quad (B)$$

On a: $H_n - xH_{n-1} = \sum_{i \leq n-2} c_i H_i(x)$

$$\Rightarrow (H_n - xH_{n-1} | H_i) = (\sum_{i \leq n-2} c_i H_i | H_i) \quad \forall i \leq n$$

$$\Rightarrow \underbrace{(H_n | H_i)}_0 = (xH_{n-1} | H_i) = c_i (H_i | H_i)$$

$$\Rightarrow - (xH_{n-1} | H_i) = c_i (H_i | H_i)$$

$$\Rightarrow - (H_{n-2} | xH_i) = c_i (H_i | H_i) = c_i i!$$

$$\Rightarrow \forall i = 0, 1, \dots, n-3 \quad c_i = 0$$

Pour $i = n-2$, on a:

$$- (H_{n-2} | xH_{n-2}) = - (H_{n-1} | H_{n-1}) = c_{n-2} (n-2)!$$

$$\Rightarrow c_{n-2} = - \frac{(n-1)!}{(n-2)!} = \underline{\underline{-(n-1)}}$$

Pour $i = n-1$

$$- (H_{n-1} | xH_{n-1}) = c_{n-1} (n-1)!$$

$$(H_{n-1} | xH_{n-1}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x H_{n-1}^2(x) e^{-x^2/2} dx$$

$$= \frac{(-1)^{n-1}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x H_{n-1}(x) \left(\frac{d}{dx}\right)^{n-1} e^{-x^2/2} dx$$

$$= \frac{(-1)^{n-1}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} [x H_{n-1}]' H_{n-2}(x) (-1)^{n-2} e^{-x^2/2} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} [x H_{n-1}]' H_{n-2}(x) e^{-x^2/2} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} [x H_{n-1}]^{(n-1)} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} [x H_{n-1} + (n-1) H_{n-1}] e^{-x^2/2} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} [(n-1)x + (n-1)x] e^{-x^2/2} dx = \underline{\underline{0}}$$

Donc: $\boxed{H_n(x) - xH_{n-1}(x) + (n-1)H_{n-2}(x) = 0} \quad (C)$

Relation entre dérivées

$$\left(\frac{d}{dx}\right)^{n-2} e^{-x^2/2} = (-1)^{n-2} H_{n-2}(x) e^{-x^2/2}$$

$$\left(\frac{d}{dx}\right)^{n-1} e^{-x^2/2} = (-1)^{n-1} H_{n-1}(x) e^{-x^2/2} = (-1)^{n-2} [H_{n-2}'(x) - xH_{n-2}(x)] e^{-x^2/2}$$

$$\left(\frac{d}{dx}\right)^n e^{-x^2/2} = (-1)^n H_n(x) e^{-x^2/2} = (-1)^{n-2} [H_{n-2}''(x) - 2xH_{n-2}'(x) - H_{n-2}(x) + x^2H_{n-2}(x)] e^{-x^2/2}$$

En portant dans (C), il vient:

$$H_{n-2}''(x) - 2xH_{n-2}'(x) - (1-x^2)H_{n-2}(x) + xH_{n-2}'(x) - x^2H_{n-2}(x) + (n-1)H_{n-2}(x) = 0$$

$$\boxed{H_{n-2}''(x) - xH_{n-2}'(x) + (n-3)H_{n-2}(x) = 0}$$