

**Les Cahiers du Centre de Morphologie Mathématique
DE FONTAINEBLEAU**

Fascicule 5

**LA THÉORIE DES
VARIABLES RÉGIONALISÉES,
ET SES APPLICATIONS**

par

G. MATHERON

ENSMP

1970

LA THEORIE DES VARIABLES REGIONALISEES, ET SES APPLICATIONS

Table des Matières

<u>PREFACE</u>	1
<u>Chapitre 0 - INTRODUCTION</u>	2
0-1 <u>Notations</u>	2
0-2 <u>Produit de convolution</u>	3
0-3 <u>Géostatistique et Théorie des Variables Régionalisées</u>	4
0-4 <u>Méthodes transitives et théorie intrinsèque</u>	6
<u>Chapitre 1 - LES METHODES TRANSITIVES</u>	9
1-1 <u>Exemple introductif</u>	9
1-2 <u>Le covariogramme transitif</u>	11
1-3 <u>Régularisation et montée</u>	14
1-3-1 Régularisation d'une Variable Régionalisée	14
1-3-2 La montée	15
1-3-3 La montée dans le cas isotrope	16
1-4 <u>L'estimation des Variables Régionalisées</u>	20
1-4-1 Variance d'estimation (maille régulière)	20
1-4-2 Cas d'une maille aléatoire stratifiée	25
1-4-3 Formules d'approximation pour l'espace à une dimension	27
1-4-4 Formules d'approximation dans \mathbb{R}^n	30
1-4-5 Application à l'estimation d'une surface	36
1-4-6 Passage à la version probabiliste de la théorie	39
1-5 <u>Exercices sur les méthodes transitives</u>	41
<u>Chapitre 2 - LES FONCTIONS ALEATOIRES INTRINSEQUES</u>	50
2-1 <u>Définitions générales</u>	50
2-2 <u>Propriétés de la covariance et du variogramme</u>	54
2-2-1 Les combinaisons linéaires autorisées	55
2-2-2 Continuité m.q. et autres propriétés	57
2-3 <u>Régularisation d'une F.A.I.</u>	60
2-3-1 Intégrale stochastique	60
2-3-2 Convolution stochastique	61
2-3-3 La montée droite sous puissance constante	62

2-4	<u>Variances d'extension et d'estimation</u>	64
2-4-1	Variance d'extension	64
2-4-2	Variance d'estimation	66
2-4-3	Variance de v dans V	67
2-4-4	Mailles aléatoires, et aléatoires stratifiées	70
2-5	<u>Méthodes d'approximation dans \mathbb{R}^1</u>	71
2-5-1	Le principe de correspondance	71
2-5-2	Le principe de composition des variances d'extension élémentaires	72
2-6	<u>Méthodes d'approximation dans \mathbb{R}^n</u>	74
2-7	<u>L'effet de pépité</u>	78
2-7-1	Genèse d'un effet de pépité	78
2-7-2	Influence macroscopique d'un effet de pépité	79
2-8	<u>Le schéma de De Wijs</u>	81
2-8-1	Les équivalents linéaires	81
2-8-2	Le schéma de De Wijs à deux dimensions	83
2-8-3	Le schéma de De Wijs dans \mathbb{R}^3	85
2-9	<u>Le schéma sphérique</u>	86
	Abaques du schéma de De Wijs	88
	Abaques du schéma sphérique	94
2-10	<u>Inférence statistique et quasi-stationnarité</u>	97
2-10-1	F.A. quasi stationnaire	97
2-10-2	Variance d'estimation	98
2-10-3	Possibilité de l'inférence statistique	100
2-11	<u>Exercices sur les F.A.I.</u>	102
2-11-1	Construction de F.A.I.	102
2-11-2	Exercices sur les variances d'estimation	103
2-11-3	Exercices sur les paragraphes 2-7 à 2-9	105
2-11-4	Exercices sur les grandes mailles	107
2-11-5	Inférence statistique pour les F.A.	112

<u>Chapitre 3 - LE KRIGEAGE</u>	117
3-1 <u>Les buts du krigeage</u>	117
3-2 <u>Notations</u>	121
3-3 <u>F.A. stationnaire d'espérance nulle ou connue</u>	122
3-4 <u>F.A. stationnaire d'espérance inconnue</u>	125
3-4-1 Les équations du krigeage	125

3-4-2	L'estimation optimale de m	127
3-4-3	Le théorème d'additivité	128
3-5	<u>Cas d'une F.A.I. sans covariance</u>	130
3-6	<u>Exercices sur le krigeage</u>	132
 <u>Chapitre 4 - LE KRIGEAGE UNIVERSEL</u>		141
4-1	<u>Introduction</u>	141
4-1-1	Critique des méthodes de moindre carré	141
4-1-2	Hypothèses générales	143
4-2	<u>L'estimation Optimale de la dérive</u>	147
4-2-1	Estimation de la dérive en un point	147
4-2-2	Estimation des coefficients de la dérive	151
4-2-3	L'invariance tensorielle	153
4-2-4	Le variogramme des résidus	155
4-2-5	Comparaison avec le maximum de vraisemblance	158
4-2-6	Cas d'une dérive donnée sous forme implicite	159
4-2-7	Comparaison avec les moindres carrés	162
4-3	<u>Le Krigeage</u>	167
4-3-1	Les équations du krigeage universel	167
4-3-2	Le théorème d'additivité	169
4-3-3	Le krigeage considéré comme un interpolateur	171
4-4	<u>Krigeage universel pour une dérive aléatoire</u>	176
4-4-1	Hypothèses	176
4-4-2	Estimation d'une dérive	179
4-4-3	Un exemple ambigu	181
4-4-4	Le problème du terme constant a_0	185
4-4-5	Le krigeage	186
4-5	<u>Le Cokrigeage</u>	187
4-5-1	Notations	187
4-5-2	Estimation de la dérive	189
4-5-3	Cokrigeage	190
4-6	<u>Les indéterminations du variogramme sous-jacent</u>	191
4-6-1	Position du problème	191
4-6-2	Les estimateurs quadratiques universels	194
4-6-3	Forme générale des matrices de covariance admissibles	196

4-6-4	Conséquences pour les estimateurs optimaux	198
4-7	<u>Exercices sur le Chapitre 4</u>	199
4-7-1	Exercices sur les dérivées	199
4-7-2	Exercices sur le krigeage et le cokrigeage	204
4-7-3	Les indéterminations de la covariance sous-jacente	209
	<u>BIBLIOGRAPHIE</u>	212

P R E F A C E

Ce texte est la rédaction du cours que j'ai donné en 1970 à l'Ecole d'Eté de Fontainebleau. Il est le fruit de plusieurs années de réflexion sur le caractère ambigu de l'opération qui consiste à interpréter en termes probabilistes un phénomène naturel unique mais partiellement inconnu, et sur le problème difficile que pose la recherche des conditions de possibilité de l'inférence statistique à partir d'une réalisation unique d'une fonction aléatoire non stationnaire. J'ai donc mis l'accent sur les problèmes méthodologiques plus que sur l'aspect mathématique de la théorie. Du point de vue des applications pratiques, les résultats utiles sont ceux-là même que j'avais déjà donnés dans le fascicule 2 de ces cahiers (à l'exception du krigeage universel), mais présentés sous un éclairage nouveau. La solution du problème méthodologique mentionné ci-dessus consiste à introduire une hypothèse de quasi-stationnarité, suffisamment faible pour être toujours physiquement plausible. Le rapprochement des chapitres 1 (méthodes transitives) et 2 (Théorie des F.A. intrinsèques) montre que le problème de l'inférence statistique est alors soluble (et donc la théorie opératoire), en ce qui concerne l'estimation globale d'une variable régionalisée. Le chapitre 4 consacré au krigeage universel aboutit à une conclusion analogue relativement au problème de l'estimation locale. Je n'ai pas présenté le krigeage universel en termes d'espaces de Hilbert, pour éviter toute difficulté mathématique et concentrer toute l'attention sur les problèmes méthodologiques. On devra donc se reporter au fascicule N° 1 de ces Cahiers pour la démonstration de certains résultats (notamment les théorèmes d'existence et d'unicité). Par contre, j'ai examiné la notion capitale de dérive à partir de points de vue très différents (dérive fonctionnelle, dérive aléatoire, maximum de vraisemblance, théorie des interpolateurs) qui en soulignent le caractère physiquement ambigu, ou du moins lié étroitement à des considérations d'échelle - et montré que ces points de vue apparemment opposés conduisaient à des conclusions convergentes : ce qui apparaîtra au moins comme rassurant sur le plan méthodologique.

Un mot encore sur les exercices proposés à la fin de chaque chapitre, et dont la plupart ont été effectivement traités lors de l'Ecole d'Eté de 1970. Certains d'entre eux sont de simples exercices d'application destinés à familiariser le lecteur avec le maniement réel d'une théorie un peu abstraite. D'autres sont de véritables compléments au cours lui-même, et apportent des résultats nouveaux dont certains ont une importance méthodologique fondamentale. Pour chacun d'eux j'indique la solution, la marche à suivre pour y parvenir, et les conclusions méthodologiques qu'il convient d'en tirer. J'espère en avoir assez dit pour convaincre le lecteur de la nécessité de faire effectivement ces exercices.

G. MATHERON

0 - INTRODUCTION

0-1 NOTATIONS

Nous désignerons par x, y, \dots des points de l'espace à n dimensions ($n = 1, 2$ ou 3), par dx, dy, \dots des éléments de longueur ($n = 1$), de surface ($n = 2$) ou de volumes ($n = 3$) centrés en ces points, par $f(x), g(y)$ etc.... des fonctions de ces points.

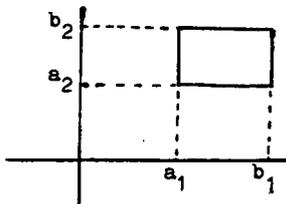
L'intégrale (étendue a tout l'espace) d'une fonction $f(x)$ se note :

$$\int f(x) dx$$

Par exemple, pour $n = 3$ dimensions, et en désignant par (x_1, x_2, x_3) les trois coordonnées du point x , cette notation s'explique ainsi :

$$\int f(x) dx = \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 \int_{-\infty}^{+\infty} f(x_1, x_2, x_3) dx_3$$

De même, nous noterons $\int_A f(x) dx$ l'intégrale de la fonction $f(x)$ étendue à un domaine A de l'espace à n dimensions.



Pour $n = 2$, par exemple, et si A est le rectangle représenté ci-joint, on a explicitement :

$$\int_A f(x) dx = \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} f(x_1, x_2) dx_2$$

Ces notations sont a la fois très condensées et très parlantes : en première lecture, on peut les interpréter dans le cas de l'espace a une seule dimension, où leur signification est, en général, très claire. La généralisation aux espaces à deux ou trois dimensions est ainsi rendue plus facile.

Exemple : Soit V un volume, et $g(h) = g(h_1, h_2, h_3)$ une fonction du vecteur h de coordonnées h_1, h_2, h_3 . Soient $x = (x_1, x_2, x_3)$ et $y = (y_1, y_2, y_3)$ l'origine et l'extrémité de ce vecteur h , c'est-à-dire $h = y - x$. La valeur moyenne de la fonction $g(h)$ lorsque les deux extrémités x et y du vecteur h décrivent (chacune pour son propre compte) le volume V se note :

$$\frac{1}{V} \int_V dx \int_V g(y-x) dy$$

En notation explicite, ceci s'écrirait :

$$\frac{1}{V} \iiint_V dx_1 dx_2 dx_3 \iiint_V g(y_1 - x_1 ; y_2 - x_2 ; y_3 - x_3) dy_1 dy_2 dy_3$$

La première notation est un symbole, qui désigne directement le concept de valeur moyenne ; la deuxième est un algorithme qui indique la marche à suivre pour calculer cette quantité, mais dont la signification n'apparaît plus au premier coup d'oeil.

0-2 PRODUIT DE CONVOLUTION (moyenne mobile)

Le produit de convolution de deux fonctions $f_1(x)$ et $f_2(x)$ est la fonction $g(x)$ définie par :

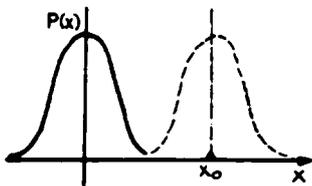
$$g(x) = \int f_1(y) f_2(x-y) dy = \int f_2(y) f_1(x-y) dy$$

On écrit en notation symbolique :

$$g = f_1 * f_2$$

Cette opération de convolution joue un rôle fondamental en géostatistique (comme aussi en probabilités et dans toute la physique théorique). On peut la rattacher à la notion intuitive de "moyenne mobile" :

Régularisée d'une fonction f (moyenne mobile pondérée de cette fonction)



Soit $p(y)$ une fonction de pondération. La valeur au point x_0 de la moyenne mobile de la fonction f pondérée par p est :

$$f_p(x_0) = \int p(y) f(x_0+y) dy = \int p(-y) f(x_0-y) dy$$

(on attribue le poids $p(y) dy$ à la valeur prise par f au point x_0+y , et on obtient la deuxième expression en changeant y en $-y$).

Soit \check{p} la transposée de la fonction p (définition : $\check{p}(x) = p(-x)$). $f_p(x_0)$ est la valeur en x_0 de $f * \check{p}$:

$$f_p = f * \check{p}$$

Cette moyenne mobile f_p de la fonction f (pondérée par p) s'appelle régularisée (de f par p).

Exemple 1 (prélèvement d'un échantillon v implanté au point x)

Soit v l'échantillon implanté à l'origine des coordonnées, et $k(x)$ sa fonction indicatrice ($k(x) = 1$ si $x \in v$ et $k(x) = 0$ si $x \notin v$). Prenons comme fonction de pondération $p(x) = \frac{1}{v} k(x)$. La moyenne mobile correspondante

$$f_v = \frac{1}{v} f * k$$

représente la teneur moyenne de l'échantillon v prélevé au point x . Explicitement, on a :

$$f_v(x) = \frac{1}{v} \int_v f(x+y) dy$$

Exemple 2 (radioactivités)

Si l'on place à l'origine O des coordonnées une masse unité de substance radioactive, on observe à la distance d une radioactivité $A e^{-\lambda d}/d$. (A est une constante, λ le coefficient d'absorption du milieu). Soit alors $f(x)$ la teneur au point x en cette substance radioactive. Désignons par $d(x-x_0)$ la distance entre deux points x et x_0 . La masse $f(x) dx$ placée en x entraîne en x_0 une radioactivité $A e^{-\lambda d(x_0-x)}/d(x_0-x) f(x) dx$. Au total, on observe en x la radioactivité :

$$A \int f(x) \frac{e^{-\lambda d(x_0-x)}}{d(x_0-x)} dx$$

Ce n'est pas autre chose que le produit de convolution $A f * \check{p}$ (avec une fonction de pondération $p(x) = e^{-\lambda d}/d$, d'ailleurs symétrique : $p = \check{p}$).

0-3 GEOSTATISTIQUE ET THEORIE DES VARIABLES REGIONALISEES

La Géostatistique est l'application de la théorie des variables régionalisées à l'esti-

mation des gisements miniers (avec tous les tours de main que cela implique). Plus généralement, nous dirons d'un phénomène qui se déploie dans l'espace et y manifeste une certaine structure, qu'il est régionalisé. Les sciences de la terre, entre autres, en fournissent de nombreux exemples. Si $f(x)$ désigne la valeur au point x d'une caractéristique f de ce phénomène, nous dirons que $f(x)$ est une variable régionalisée, en abrégé V.R. C'est là un terme neutre, purement descriptif, antérieur, en particulier, à toute interprétation probabiliste.

Du point de vue mathématique, une V.R. est donc simplement une fonction $f(x)$ du point x , mais c'est, en général, une fonction fort irrégulière : ex. : une teneur dans un gisement minier. Elle se présente sous deux aspects contradictoires (ou complémentaires) :

- un aspect aléatoire (haute irrégularité, et variations imprévisibles d'un point à l'autre).
- un aspect structuré (elle doit cependant refléter à sa manière les caractéristiques structurales du phénomène régionalisé).

La théorie des V.R. se propose donc deux objectifs principaux :

- sur le plan théorique, exprimer ces caractéristiques structurales sous une forme mathématique adéquate ;

sur le plan pratique, résoudre le problème de l'estimation d'une V.R. à partir d'un échantillonnage fragmentaire.

Ces deux objectifs sont liés : pour un même réseau de prélèvements, l'erreur d'estimation dépend des caractéristiques structurales ; elle est, par exemple, d'autant plus élevée que la V.R. est plus irrégulière et plus discontinue dans sa variation spatiale.

Champ et Support d'une V.R.

Le champ V d'une V.R. est le domaine où celle-ci est différente de 0. Un panneau est un sous-ensemble V' de V .

Support - Souvent, on ne connaît pas $f(x)$ elle-même, mais seulement sa valeur moyenne $f_v(x)$ dans l'échantillon v prélevé au point x . Cette régularisée $f_v(x)$ est effectivement plus régulière que la V.R. $f(x)$. Le volume v s'appelle le support de la V.R. f_v , régularisée de f . Une autre tâche importante de la théorie des V.R. consistera donc à déterminer les caractéristiques de f_v connaissant celles de f . Ex. : dans un gisement, prévoir les caractéristiques des panneaux V' (variable f_v), connaissant celles de f ou de f_v (échantillons).

Plus généralement, on cherchera à relier les caractéristiques de f à celles d'une régularisée f_p par une fonction p donnée (exemple des radioactivités).

O-4 METHODES TRANSITIVES ET THEORIE INTRINSEQUE

Pour atteindre ces objectifs, nous disposons de deux groupes de méthodes :

- méthodes transitives : absolument générales, et ne nécessitant en particulier aucune hypothèse de nature probabiliste, et a fortiori aucune hypothèse de stationnarité.

- théorie intrinsèque : c'est une application de la théorie des fonctions aléatoires ; on introduit donc des interprétations probabilistes, et même une certaine hypothèse de stationnarité (hypothèse intrinsèque).

Du point de vue théorique, ces deux groupes de méthodes conduisent à des résultats équivalents : c'est important pour la méthodologie, car cela montre que les résultats de la théorie intrinsèque ne sont pas réellement liés à l'hypothèse de la stationnarité (on peut d'ailleurs construire une théorie probabiliste affranchie de cette hypothèse et permettant de retrouver les principaux résultats de la géostatistique).

Nous rencontrons ici un problème méthodologique qui présente une importance capitale aussi bien pour la théorie elle-même que pour l'examen critique de la valeur des résultats auxquels elle conduit. Il est clair que le caractère ambigu, localement erratique, de nos V.R. appelle une interprétation probabiliste, et, de fait, notre deuxième groupe de méthodes fera explicitement appel à la théorie des fonctions aléatoires. Mais alors se posent deux questions fondamentales, dont la première est la suivante :

a/ Quelle est la signification épistémologique véritable de l'opération qui consiste à considérer un phénomène naturel unique (notre V.R.) comme une réalisation d'une fonction aléatoire (c'est-à-dire comme le résultat d'un tirage au sort effectué parmi une population infinie de V.R. considérées comme "possibles")?

A dire vrai, une interprétation probabiliste de ce genre est par elle-même une conceptualisation de la réalité (un modèle constitutif) plutôt qu'une hypothèse susceptible d'être infirmée ou confirmée expérimentalement. Elle se justifie seulement dans la mesure où elle permet de mieux saisir la réalité et de résoudre effectivement des problèmes pratiques insolubles autrement. Il semble donc prudent d'essayer, tout d'abord, de voir jusqu'où il est possible d'aller sans recourir

à cette interprétation - et nous verrons qu'en fait il est possible d'aller assez loin. De là ce premier groupe de méthodes que nous appelons transitives où n'intervient en apparence aucun concept probabiliste. La V.R. est ici caractérisée par son covariogramme transitif $g(h)$, non probabiliste, qui résume ses caractères structuraux essentiels et permet - si on le connaît - de résoudre entièrement certains problèmes pratiques comme celui de l'estimation.

On pourrait donc en rester là, et faire l'économie de toute interprétation probabiliste, s'il était effectivement possible de déterminer ce $g(h)$ à partir de données expérimentales fragmentaires. Mais il n'en est rien, et une analyse plus fine montrera qu'il est nécessaire ici (si l'on veut, par exemple, calculer effectivement une variance d'estimation) d'introduire un certain type d'hypothèses concernant le comportement analytique du $g(h)$ au voisinage de l'origine, hypothèses dont la signification épistémologique est exactement celle d'un passage camouflé à l'espérance mathématique. Il n'est donc pas réellement possible d'éviter l'interprétation probabiliste, et mieux vaut alors l'introduire explicitement. C'est ici que se pose la seconde question fondamentale:

b/ Une fois admise cette interprétation probabiliste, l'inférence statistique est-elle possible à partir d'une réalisation unique? Autrement dit, à partir du seul matériel expérimental disponible (qui est la V.R. elle-même, ou un échantillonnage fragmentaire de cette V.R. unique) est-il réellement possible de reconstituer, au moins en partie, la loi de probabilité de la fonction aléatoire hypothétique dont notre V.R. serait une réalisation?

En vue de répondre affirmativement à cette seconde question, on introduit souvent des hypothèses, du type stationnarité et ergodicité, beaucoup plus fortes qu'il n'est réellement nécessaire : dans beaucoup d'applications, ces hypothèses sont grossièrement fausses (exemples : un gisement où les teneurs décroissent plus ou moins régulièrement à partir d'un coeur riche ; la topographie sous-marine, où les profondeurs augmentent quand on s'éloigne des côtes, etc... : il s'agit là de phénomènes manifestement non stationnaires). Plus fréquemment encore, ces hypothèses apparaissent comme invérifiables. C'est pourquoi, nous chercherons constamment à les affaiblir, jusqu'à les réduire à un minimum à la fois indispensable et admissible. Plus précisément, nous remplacerons la question b/ ci-dessus par la suivante :

b'/ Quelle caractéristique probabiliste minimale avons-nous réellement besoin de connaître pour résoudre un problème pratique donné (par exemple, le calcul d'une variance d'estimation) , et quelle hypothèse minimale est-il nécessaire d'introduire en vue de rendre possible l'estima-

tion de cette caractéristique à partir d'une réalisation unique?

Le plus souvent, cette hypothèse minimale, suffisamment faible pour être à peu près toujours au moins physiquement plausible, consistera à supposer que la fonction aléatoire est quasi-intrinsèque (c'est-à-dire localement assimilable à une F.A. intrinsèque), tandis que la caractéristique minimale qu'il sera à la fois nécessaire et possible d'estimer sera liée essentiellement au comportement au voisinage de l'origine d'un variogramme ou d'une fonction de covariance quasi-stationnaire.

Cette réduction du problème à ses caractéristiques probabilistes minimales permet de comprendre le succès des méthodes transitives (qui, une fois explicitement probabilisées, ne nécessitent évidemment aucune hypothèse du type stationnaire) et aussi l'équivalence de leurs résultats avec ceux de la théorie intrinsèque. Elle souligne de plus l'intérêt capital que va présenter pour nous le comportement au voisinage de l'origine d'un variogramme ou d'un covariogramme transitif. Du point de vue pratique, enfin, elle montre que l'on a très souvent le droit d'appliquer à des phénomènes manifestement non stationnaires des procédés de calcul dont on pouvait croire, au départ, que la légitimité était liée à quelque hypothèse de stationnarité.

En pratique, la théorie intrinsèque est plus facile à mettre en oeuvre, et c'est donc presque toujours elle que l'on utilise, sauf toutefois dans le cas particulier très important de l'estimation d'une surface ou d'un volume (problème géométrique).

En ce qui concerne la bibliographie (voir références à la fin de ce fascicule), on trouvera dans [4] l'exposé complet de la théorie des variables régionalisées, et dans [5], un exposé abrégé de cette même théorie suivi d'un traité complet de Géostatistique appliquée, comprenant en particulier l'étude des problèmes d'optimisation économique (le texte français original peut être consulté à la Bibliothèque de l'Ecole des Mines de Paris). La thèse de Serra [7] donne d'une part un exposé général très concret (et dépourvu de toute difficulté mathématique), de l'autre une étude détaillée du schéma sphérique. L'ancien traité [3] est en grande partie périmé, sauf en ce qui concerne le schéma de De Wijs. La thèse de Carlier [1] étudie (uniquement en schéma De Wijsien) les problèmes spéciaux posés par les gisements de substances radioactives. Enfin, les problèmes d'optimisation économique sont traités dans deux articles des Annales des Mines [8], dont l'essentiel a été repris dans [5]. Pour le krigeage universel, enfin, on consultera [6].

CHAPITRE 1

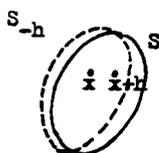
LES METHODES TRANSITIVES

1-1 EXEMPLE INTRODUCTIF

Considérons le phénomène de "transition" le plus simple que l'on puisse imaginer : présence ou absence d'un caractère. Soit, par exemple, une formation géologique S d'extension limitée. Un sondage foré au point x la rencontre, ou ne la rencontre pas. Désignons par $k(x)$ l'indicatrice de l'ensemble S, c'est-à-dire la fonction définie par :

$$k(x) = \begin{cases} 0 & \text{si } x \notin S \\ 1 & \text{si } x \in S \end{cases}$$

Il s'agit d'un phénomène unique, pour lequel aucune formulation probabiliste n'est possible : parler de la probabilité pour qu'un point x donné appartienne à S n'aurait pas grand sens. On peut noter aussi que tout l'intérêt se concentre ici sur la frontière de S. En effet, $k(x)$ est constante à l'intérieur comme à l'extérieur de S, et c'est au franchissement de cette frontière seulement que $k(x)$ varie, passant de 1 à 0 ou de 0 à 1. De là le nom de phénomène de transition, et de méthodes transitives.



L'aire S de notre formation est évidemment donnée par :

$$S = \int k(x) dx$$

La valeur de S constitue un renseignement fort intéressant du point de vue pratique : le plus souvent, c'est elle que l'on cherche à estimer à partir d'un réseau de sondages. Toutefois, ce paramètre scalaire ne nous apporte encore aucune information de nature réellement structurale. On peut, en effet, définir la structure d'un ensemble comme le système des relations existant entre les éléments ou les parties de cet ensemble. Nous n'aurons donc d'information de nature structurale sur notre surface S qu'à la condition de faire intervenir simultanément au moins deux points.

Soient donc x et x+h deux points (c'est-à-dire le plus petit ensemble structurant que nous

puissions imaginer). Considérons alors l'expression $k(x) k(x+h)$: elle vaut 1 si x et $x+h$ appartiennent tous les deux à S , et 0 autrement. Mais dire que $x+h$ appartient à S équivaut à dire que x appartient au translaté S_{-h} de S dans la translation $-h$. On a donc :

$$k(x) k(x+h) = \begin{cases} 1 & \text{si } x \in S \cap S_{-h} \\ 0 & \text{si } x \notin S \cap S_{-h} \end{cases}$$

Intégrons en x cette expression : nous obtenons une fonction de h :

$$K(h) = \int k(x) k(x+h) dx = \text{Mes} (S \cap S_{-h})$$

qui représente la mesure (l'aire) de l'intersection de S et de son translaté par $-h$. Cette fonction est symétrique, car les deux intersections $S \cap S_{-h}$ et $S \cap S_h$ se déduisent l'une de l'autre par translation. Cette fonction $K(h)$ est le covariogramme géométrique associé à S . Il donne une certaine image de la forme de l'ensemble S :

Propriétés du covariogramme géométrique $K(h)$

a/ Symétrie : $K(h) = K(-h)$; inégalités $0 \leq K(h) \leq K(0)$

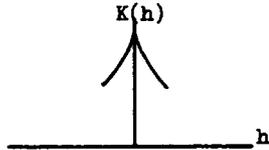
Relations :

$$S = K(0)$$

$$S^2 = \int K(h) dh$$

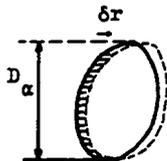
b/ Portées : la portée $a(\alpha)$ dans la direction α est la distance à partir de laquelle $K(h)$ s'annule dans cette direction. C'est donc la plus grande dimension de S dans cette direction.

c/ Pente à l'origine (dérivées à droite et à gauche)



Si le module δr de h est petit, on a :

$$K(h) = K(0) - \delta r D_\alpha$$



$\delta r D_\alpha$ est la moitié de la petite surface balayée par le vecteur δr dont l'origine décrit le contour de S . D_α est la variation diamétrale de S dans la direction α (si S est convexe c'est le diamètre apparent dans cette direction).

Bien qu'il existe une dérivée $K'_\alpha(0) = -D_\alpha$ dans chaque direction α , on notera soigneusement que la fonction $K(h)$, en tant que telle, n'est pas dérivable en $h = 0$. L'exercice 6 (à la fin de ce chapitre) montre comment la formule de Minkowski permet de relier les dérivées $K'_\alpha(0)$ dans les différentes directions α au périmètre de S (ou à sa surface si l'on se place dans l'espace à 3 dimensions). Ce sont là des relations très utiles en morphologie mathématique.

1-2 LE COVARIOGRAMME TRANSITIF

Soit $f(x)$ une V.R. nulle en dehors d'un champ V borné. Le covariogramme transitif de cette V.R. est la fonction $g(h)$ définie par :

$$(1-1) \quad g(h) = \int f(x) f(x+h) dx$$

Propriétés du covariogramme transitif $g(h)$

a/ Symétrie : $g(h) = g(-h)$, inégalité $|g(h)| \leq g(0) = \int [f(x)]^2 dx$

Soit $Q = \int f(x) dx$ la quantité de métal. On a :

$$(1-2) \quad Q^2 = \int g(h) dh$$

(pour démontrer (2), remplacer g par son expression (1), et intégrer d'abord en h).

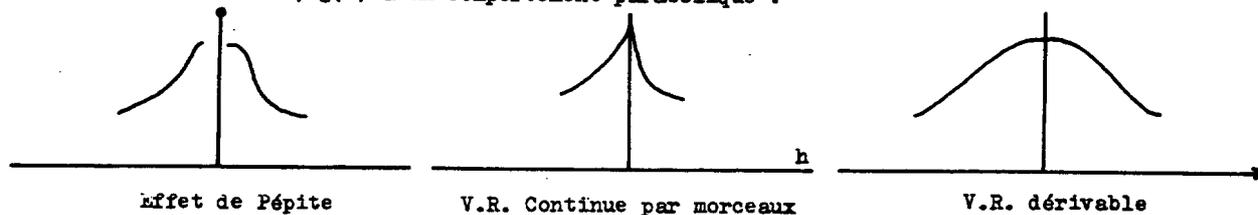
b/ Portée : $a(\alpha)$ est définie par la condition $g(h) = 0$ dès que $|h| > a(\alpha)$ pour le vecteur h de direction α : c'est une propriété du champ de la V.R.

c/ Comportement du $g(h)$ au voisinage de l'origine. La régularité de $g(h)$ au voisinage de l'origine reflète les propriétés de continuité de la V.R. dans sa variation spatiale. Cela résulte de :

$$g(0) - g(h) = \frac{1}{2} \int [f(x+h) - f(x)]^2 dx$$

Si $f(x)$ est continue par morceaux, $g(h)$ a un comportement linéaire au voisinage de l'origine.

Si $f(x)$ est dérivable, $g(h)$ a un comportement parabolique :



Parfois, le $g(h)$ n'est pas continu en $h = 0$: autrement dit, $g(0)$ est supérieur à la limite de $g(h)$ pour h tendant vers 0. On dit alors qu'il y a effet de pépité. Plutôt que d'une véritable discontinuité, il s'agit d'ailleurs le plus souvent d'une zone de transition très rapide, qui, expérimentalement, se présente comme une discontinuité. Nous reviendrons plus longuement, dans la version probabiliste de la théorie, sur cet effet de pépité et son interprétation en termes d'échelles. (voir aussi Exercice 7 de ce chapitre, et, pour une représentation de l'effet de pépité par une mesure de Dirac, les exercices 16 à 20).

Cas isotrope - Dans les applications, on cherche (moyennant, par exemple, une transformation linéaire) à se ramener au cas où le covariogramme g ne dépend que du module

$$r = |h| = \sqrt{h_1^2 + h_2^2 + \dots + h_n^2}$$

du vecteur h , et se met donc sous la forme $g(r)$ d'une fonction du seul argument r . En tant que fonctions des coordonnées h_1, h_2, \dots , seules les puissances paires r^{2k} de r sont indéfiniment dérivables en $h = 0$. Les puissances impaires r^{2k+1} , les puissances r^λ , λ non entier, et aussi les termes logarithmiques du type $r^{2n} \log r$ présentent tous une irrégularité en $h = 0$. Ainsi, dans le cas isotrope, on distinguera deux parties dans le développement limité du $g(r)$ au voisinage de l'origine : une partie régulière, qui ne comporte que des puissances paires, et une partie irrégulière regroupant les termes en r^λ , λ différent d'un entier pair, avec éventuellement aussi des termes logarithmiques :

$$g(r) = g(0) + a_2 r^2 + \dots + \sum_{\lambda} r^\lambda + \sum_k r^{2k} \log r$$

partie régulière
partie irrégulière

Le terme de plus bas degré de la partie irrégulière donne une mesure précise du degré d'irrégularité de la V.R. dans sa variation spatiale. Par exemple, l'indicatrice $k(x)$ d'un ensemble S est

caractérisée de ce point de vue par un terme irrégulier d'ordre 1 : $K(h) = S - D|h|$, comme on l'a montré ci-dessus.

d/ Le covariogramme $g(h)$ est une fonction de type positif

On dit qu'une fonction $g(h)$ est de type positif si pour tout entier $k > 0$, tout ensemble x_1, x_2, \dots, x_k de points de l'espace \mathbb{R}^n et tout système $\lambda_1, \dots, \lambda_k$ de nombres réels, on a :

$$(1-3) \quad \sum_{i,j} \lambda_i \lambda_j g(x_i - x_j) \geq 0$$

Les fonctions de type positif jouent un rôle important en physique (où elles ont en général une signification énergétique) et en théorie des probabilités (les fonctions caractéristiques des lois de probabilité sont de ce type). Nous verrons au chapitre 2 que les fonctions de covariance des F.A. sont également de type positif.

Montrons qu'un covariogramme transitif est de type positif. En effet, d'après la relation (1-1), on peut écrire :

$$\begin{aligned} \sum_{i,j} \lambda_i \lambda_j g(x_i - x_j) &= \sum_{i,j} \lambda_i \lambda_j \int f(y) f(y + x_i - x_j) dy = \\ &= \sum_{i,j} \lambda_i \lambda_j \int f(y + x_j) f(y + x_i) dy = \int \left[\sum_i \lambda_i f(y + x_i) \right]^2 dy \geq 0 \end{aligned}$$

Ainsi, dans le choix d'un modèle de covariogramme transitif on doit obligatoirement se limiter aux fonctions de type positif (nous verrons au paragraphe 1-4 que cette condition exprime que les variances d'estimation sont nécessairement positives).

D'après le théorème classique de BOCHNER, une fonction continue est de type positif si et seulement si elle est transformée de Fourier d'une mesure positive sommable. D'où une deuxième vérification de cette propriété capitale du covariogramme transitif. Désignons par Φ la transformée de Fourier de la V.R. $f(x)$, et par G celle de son covariogramme $g = f * \check{f}$. On sait que la transformée de Fourier d'un produit de convolution est le produit multiplicatif ordinaire des transformées des facteurs. Comme la transposée \check{f} de f a pour transformée l'imaginaire conjuguée $\bar{\Phi}$ de Φ , on trouve ici :

$$(1-4) \quad G = |\Phi|^2$$

et par suite on a bien $G \geq 0$.

Remarque - Du point de vue mathématique, le covariogramme g et sa transformée G sont des données parfaitement équivalentes. Comme la relation (1-4) est particulièrement simple, il est parfois plus commode, dans certaines démonstrations théoriques, d'utiliser G plutôt que le covariogramme g lui-même. Mais du point de vue expérimental, et dans les applications, le covariogramme g est en général un bien meilleur outil que G . En effet, le comportement du $g(h)$ au voisinage de l'origine est presque toujours le critère décisif, et, dans la transformation de Fourier, les propriétés de $g(h)$ au voisinage de $h = 0$ ont pour image des propriétés de G au voisinage de l'infini. Il n'est en général pas très facile d'apprécier le comportement à l'infini d'une courbe expérimentale. C'est pourquoi le point de vue de l'analyse harmonique n'est pas très intéressant dans les applications pratiques de la théorie des V.R.. La même remarque s'appliquera d'ailleurs à la version probabiliste de la théorie, où l'on travaillera directement sur la fonction de covariance ou sur le variogramme plutôt que sur leurs transformées de Fourier.

1-3 REGULARISATION D'UNE V.R., ET MONTEE

1-3-1 Régularisation d'une V.R.

Soit $f(x)$ une V.R., $p(x)$ une fonction de pondération, $f_p = f * \check{p}$ la régularisée de f par p . La relation (1-1) peut déjà s'écrire :

$$g = f * \check{f}$$

et montre que le covariogramme transitif est l'autorégularisée de la V.R. f . Le covariogramme $g_p = f * \check{p} * \check{f} * p = f * \check{f} * p * \check{p} = f_p * \check{f}_p$ se met aussi sous la forme :

$$g_p = g * P$$

avec $P = p * \check{p}$: on obtient le covariogramme de la régularisée en régularisant g par le covariogramme transitif P de la fonction de pondération p : c'est effectivement une fonction plus régulière que g , de même que f_p est plus régulière que la V.R. initiale f .

1-3-2 La Montée

Nous allons consacrer le reste de ce paragraphe à une opération appelée montée qui joue un rôle très important en théorie des V.R. (car elle permet, par exemple, de ramener le calcul d'une variance d'estimation dans un espace à n dimensions à une suite d'opérations beaucoup plus simples effectuées dans l'espace à une seule dimension). En termes miniers, la montée est simplement la transposition abstraite de l'opération qui consiste à implanter un sondage en un point (x_1, x_2) de la surface topographique.

Si $f_3(x) = f_3(x_1, x_2, x_3)$ est une V.R. dans l'espace à 3 dimensions (par exemple), on dit que la V.R. :

$$f_2(x_1, x_2) = \int f_3(x_1, x_2, x_3) dx_3$$

définie dans l'espace à 2 dimensions se déduit de f_3 par montée (d'ordre 1, parallèlement à l'axe des x_3). Par exemple, si $f_3(x)$ est une teneur ponctuelle, $f_2(x_1, x_2)$ est l'accumulation (quantité de métal au mètre carré) du sondage implanté au point (x_1, x_2) de la surface topographique.

Il n'y a aucune difficulté à définir des montées d'ordre supérieur. Par exemple la montée d'ordre 2 (parallèlement au plan des x_1, x_2) conduit à la V.R.

$$f_1(x_3) = \iint f_3(x_1, x_2, x_3) dx_1 dx_2$$

définie sur un espace à une seule dimension (l'axe des x_3) et qui représente, en termes miniers, la quantité de métal au mètre d'approfondissement portée par le niveau de cote x_3 . Il se trouve que le covariogramme $g_2(h_1, h_2)$ de la V.R. f_2 déduite de f_3 par montée d'ordre 1 s'obtient en effectuant directement cette même opération de montée sur le covariogramme $g_3(h_1, h_2, h_3)$ de la V.R. f_3 , autrement dit, en langage bref : le covariogramme monte en même temps que la V.R. à laquelle il est associé.

Donnons une démonstration de ce résultat à l'aide de la relation (1-4). Soit $f_n(x_1, \dots, x_n)$ une V.R. dans \mathbb{R}^n , $\Phi_n(u_1, \dots, u_n)$ sa transformée de Fourier. Par montée d'ordre 1 le long de l'axe des x_n , on obtient la V.R. $f_{n-1}(x_1, \dots, x_{n-1})$ dont la transformée de Fourier est :

$$(1-5) \quad \Phi_{n-1}(u_1, \dots, u_{n-1}) = \Phi_n(u_1, \dots, u_{n-1}, 0)$$

Autrement dit, conformément à un mécanisme de calcul évident (que l'on rencontre aussi, en théorie des probabilités, dans le passage à une loi marginale) on obtient Φ_{n-1} en faisant $u_n = 0$ dans l'expression de Φ_n . Désignons alors par ξ_n et ξ_{n-1} , les covariogrammes de f_n et f_{n-1} , par G_n et G_{n-1} leurs transformées de Fourier prises dans leurs espaces respectifs \mathbb{R}^n et \mathbb{R}^{n-1} . D'après (1-4), on a :

$$G_n = |\Phi_n|^2 \quad , \quad G_{n-1} = |\Phi_{n-1}|^2$$

D'après (1-5), on en déduit :

$$G_{n-1}(u_1, \dots, u_{n-1}) = |\Phi_n(u_1, \dots, u_{n-1}, 0)|^2 = G_n(u_1, \dots, u_{n-1}, 0)$$

Ainsi, on obtient G_{n-1} en annulant u_n dans l'expression de G_n . D'après la réciprocité de la transformation de Fourier, et l'expression (1-5) de la montée, il en résulte bien que ξ_{n-1} se déduit de ξ_n par montée d'ordre 1.

1-3-3 La montée dans le cas isotrope

Examinons maintenant le cas isotrope, c'est-à-dire le cas où la V.R. $f_n(x)$, $x \in \mathbb{R}^n$ admet un covariogramme $\xi_n(r)$ dépendant uniquement du rayon vecteur $r = |h|$. Il n'est plus nécessaire de spécifier dans quelle direction s'effectue la montée. Les montées d'ordre 1, 2, ... donnent des V.R. f_{n-1} , f_{n-2} , ... dont les covariogrammes $\xi_{n-1}(r)$, $\xi_{n-2}(r)$, ... dépendent uniquement du rayon vecteur r (de l'espace à $n-1$, $n-2$, ... dimensions). D'après ce qui précède, ces covariogrammes se déduisent de ξ_n par montée, et on trouve sans difficulté :

$$\begin{cases} \xi_{n-1}(r) = 2 \int_0^\infty \xi_n(\sqrt{h^2+r^2}) dh \\ \xi_{n-2}(r) = 2 \pi \int_0^\infty \xi_n(\sqrt{h^2+r^2}) h dh \end{cases}$$

Le changement de variable : $u^2 = h^2 + r^2$ donne :

$$(1-6) \quad \begin{cases} \xi_{n-1}(r) = 2 \int_r^\infty \xi_n(u) \frac{udu}{\sqrt{u^2-r^2}} \\ \xi_{n-2}(r) = 2 \pi \int_r^\infty \xi_n(u) u du \end{cases}$$

Il apparaît ainsi une différence remarquable entre les montées d'ordre 1 et 2. La montée d'ordre 2 (et, plus généralement, d'ordre pair) est une opération élémentaire. Elle s'inverse facilement par dérivation :

$$(1-7) \quad \xi_n(r) = -\frac{1}{2\pi r} \xi'_{n-2}(r)$$

Au contraire, la montée d'ordre 1 (et, plus généralement, d'ordre impair) est une opération plus difficile, à cause de la présence du radical sous le signe d'intégration, et ne s'inverse pas aisément.

Mais il est possible de montrer que les montées, et leurs inverses, ou descentes, considérées comme opérateurs agissant sur les fonctions isotropes du type $\xi_n(r)$ constituent un groupe (cf [4]). Pour trouver l'expression de la descente d'ordre 1 ($\xi_{n-1} - \xi_n$), on peut donc effectuer d'abord une descente d'ordre 2, qui conduit, d'après (1-7), à :

$$\xi_{n+1}(r) = -\frac{1}{2\pi r} \xi'_{n-1}(r)$$

puis une remontée d'ordre 1 qui donne, d'après (1-6) :

$$\xi_n(r) = 2 \int_r^\infty \xi_{n+1}(u) \frac{u \, du}{\sqrt{u^2 - r^2}} = -\frac{1}{\pi} \int_r^\infty \xi'_{n-1}(u) \frac{du}{\sqrt{u^2 - r^2}}$$

Autrement dit, nous avons ainsi établi les expressions réciproques de la montée et de la descente d'ordre 1 :

$$(1-8) \quad \left\{ \begin{array}{l} \xi_{n-1}(r) = 2 \int_r^\infty \xi_n(u) \frac{u \, du}{\sqrt{u^2 - r^2}} \\ \xi_n(r) = -\frac{1}{\pi} \int_r^\infty \xi'_{n-1}(u) \frac{du}{\sqrt{u^2 - r^2}} \end{array} \right.$$

On trouvera plus loin (Exercice 11 bis) une application intéressante de ces formules de réciprocity au problème qui consiste à reconstituer la granulométrie d'une population de sphères de l'espace à 3 dimensions à partir des granulométries des cercles ou des traversées (cordes) induits sur des plans ou sur des droites.

Adoptons maintenant le point de vue de la transformation de Fourier, et soit $G = \mathfrak{F}_n g_n$ la transformée du covariogramme g_n (prise dans son espace à n dimensions). Comme la fonction g_n est isotrope, sa transformée $\mathfrak{F}_n g_n$ ne dépend que du rayon vecteur

$$\rho = \sqrt{u_1^2 + \dots + u_n^2}$$

de l'espace à n dimensions où sont définies les transformées \mathfrak{F}_n . Ainsi G est elle-même une fonction isotrope du type $G(\rho)$. D'après (1-5), on obtient $\mathfrak{F}_{n-1} g_{n-1}$ en annulant u_n dans l'expression de $\mathfrak{F}_n g_n$. Or, annuler u_n dans l'expression du rayon vecteur conduit à l'expression $\sqrt{u_1^2 + \dots + u_{n-1}^2}$, c'est-à-dire au rayon vecteur de l'espace à $n-1$ dimensions, que nous pouvons encore désigner par ρ . Par conséquent, en tant que fonction $G(\rho)$ du rayon vecteur ρ de l'espace de Fourier, la transformée G du covariogramme isotrope g_n est invariante par montée (et, aussi bien, par descente) :

$$(1-9) \quad G(\rho) = \mathfrak{F}_n g_n = \mathfrak{F}_{n-1} g_{n-1} = \mathfrak{F}_{n+1} g_{n+1} = \dots$$

On dispose là d'un moyen analytique souvent commode pour expliciter une montée ou une descente.

Principe de correspondance.

Examinons, maintenant, comment se comporte à la montée ce qui nous intéresse le plus dans un covariogramme isotrope, c'est-à-dire la partie irrégulière de son développement limité au voisinage de l'origine (paragraphe 1-2). Le résultat essentiel (dont on trouvera la démonstration en [4]) est le suivant : on peut effectuer la montée terme à terme sur la partie irrégulière de $g_n(r)$, suivant la règle de correspondance :

$$(1-10) \quad r^\lambda \rightarrow A_\lambda r^{1+\lambda} \quad (\lambda \text{ non entier})$$

et ce principe de correspondance donne le développement limité de $g_{n-1}(r)$ à une série entière paire près. Il n'y a pas de règle analogue pour la partie régulière (mais cela ne sera pas gênant car, nous le verrons, les variances d'estimation dépendent seulement de la partie irrégulière). L'ordre du terme irrégulier de plus bas degré est ainsi augmenté d'une unité, et cette circonstance souligne bien l'effet régulateur de la montée.

En ce qui concerne les termes de degré entier impair et les termes logarithmiques de la partie irrégulière, le principe de correspondance terme à terme leur est également applicable sous la forme :

$$(1-11) \quad \begin{cases} r^{2k} \log r - A_{2k} r^{2k+1} \\ r^{2k+1} - A_{2k+1} r^{2k+2} \log r \end{cases}$$

On obtient ainsi, par montées successives, la série singulière où l'on voit alterner termes impairs et termes logarithmiques. Par exemple :

$$\begin{aligned} \log r - \pi r \\ r - - r^2 \log r \end{aligned}$$

Calcul des coefficients A_λ - Pour une démonstration rigoureuse, on se reportera à [4]. Nous donnons ici simplement une justification symbolique, dont le principe nous servira également pour le calcul des variances d'estimation. En tant que distribution (et non en tant que fonction, et c'est en cela que cette justification n'est que symbolique) r^λ admet, dans l'espace à n dimensions, la transformée de Fourier :

$$\mathfrak{F}_n r^\lambda = \frac{1}{\pi^{\lambda + \frac{n}{2}}} \frac{\Gamma(\frac{\lambda+n}{2})}{\Gamma(-\frac{\lambda}{2})} \frac{1}{\rho^{\lambda+n}}$$

(λ non entier). D'après (1-9), la montée d'ordre 1 effectuée sur r^λ conduit à $\mathfrak{F}_{n-1} \mathfrak{F}_n r^\lambda$. En calculant $\mathfrak{F}_{n-1} \rho^{-\lambda-n}$ d'après la formule même écrite ci-dessus, on obtient :

$$\mathfrak{F}_{n-1} \rho^{-\lambda-n} = \pi^{\lambda + \frac{n+1}{2}} \frac{\Gamma(-\frac{\lambda+1}{2})}{\Gamma(\frac{\lambda+n}{2})} r^{\lambda+1}$$

et on en déduit l'expression du coefficient A_λ de la règle (1-10) de correspondance, à l'aide des fonctions eulériennes Γ , soit :

$$(1-12) \quad A_\lambda = \sqrt{\pi} \frac{\Gamma(-\frac{1+\lambda}{2})}{\Gamma(-\frac{\lambda}{2})} = \sqrt{\pi} \operatorname{tg} \lambda \frac{\Gamma(1 + \frac{\lambda}{2})}{\Gamma(1 + \frac{1+\lambda}{2})}$$

Cette justification, répétons-le, n'est que symbolique, et, en particulier, ne montre pas que la règle de correspondance (1-10) ne donne le résultat de la montée qu'à une série entière près. Mais le résultat (1-12) est correct. Dans le cas de la série singulière (1-11), on obtient (par des passages à la limite sur lesquels il n'y a pas lieu d'insister ici) :

$$(1-13) \quad \left\{ \begin{array}{l} A_{2k} = \pi \frac{2^k k!}{1.3\dots(2k+1)} = \pi \frac{2^{2k} (k!)^2}{(2k+1)^2} \\ A_{2k+1} = - \frac{2^{-2k} (2k+1)!}{k! (k+1)!} \end{array} \right.$$

1-4 L'ESTIMATION DES V.R.

Dans ce paragraphe capital, nous définissons la variance d'estimation pour une maille régulière, et nous en donnons deux expressions équivalentes (rigoureuses) valables dans l'espace à n dimensions (par. 1-4-1). Dans le paragraphe (1-4-2), nous examinons le cas d'une maille aléatoire stratifiée. Nous cherchons ensuite des méthodes d'approximation, d'abord dans l'espace à 1 dimension (1-4-3), puis à deux ou plusieurs dimensions (1-4-4) et nous appliquons ces résultats au problème géométrique (1-4-5). L'examen critique des conditions moyennant lesquelles il est effectivement possible de mettre en oeuvre ces résultats nous permet enfin de préparer le passage à la version probabiliste de la théorie (1-4-6).

1-4-1 Expression rigoureuse de la variance d'estimation (maille régulière)

Pour abrégé, nous expliciterons les calculs dans l'espace à une seule dimension. Mais les résultats se généralisent sans peine à l'espace à n dimensions. Soit $f(x)$ une V.R. sur la droite réelle et :

$$Q = \int_{-\infty}^{+\infty} f(x) dx$$

la quantité de métal associée. Pour estimer Q , nous disposons d'un réseau de prélèvement à maille régulière a . Autrement dit, si nous désignons par x_0 l'abscisse de l'un (quelconque) des prélèvements du réseau, nous connaissons les valeurs numériques $f(x_0+pa)$ de la V.R. $f(x)$ pour p entier positif ou négatif. En réalité, $f(x)$ n'étant différente de 0 que sur un domaine borné, seul un nombre fini de ces valeurs $f(x_0+pa)$ sont différentes de zéro, et, dans ce qui suit, nous n'aurons pas à nous préoccuper de la convergence de séries (qui ne comporteront, en réalité, qu'un nombre fini de termes). Dans la pratique, il ne sera pas nécessaire que le réseau des prélèvements soit réellement infini ; il suffira qu'il déborde quelque peu le champ géométrique de la V.R.

Pour estimer la quantité de métal Q , on forme alors l'estimateur :

$$Q^*(x_0) = a \sum_{p=-\infty}^{+\infty} f(x_0+pa)$$

qui est une fonction périodique (de période a) de l'abscisse x_0 choisie arbitrairement comme origine du réseau. L'erreur $Q - Q^*(x_0)$ associée à cette estimation est également une fonction périodique de x_0 , et nous allons caractériser son amplitude possible par une variance d'estimation que nous noterons $\sigma^2(a)$.

Conformément à notre fil directeur méthodologique (paragraphe 0-4) nous n'introduisons encore aucune interprétation probabiliste relativement à la V.R. $f(x)$ elle-même : celle-ci représente une réalité physique unique, parfaitement déterminée (même si nous ne la connaissons pas effectivement), mais, justement parce que nous ne connaissons rien, au départ, à son sujet, nous implantons notre réseau de prélèvements "n'importe où" ou "au hasard" relativement à cette réalité physique inconnue. Il n'est pas difficile de donner un statut épistémologique plus précis à ce "n'importe où", en admettant que tout se passe comme si l'origine x_0 du réseau avait été implantée au hasard sur un segment de longueur a selon une loi de probabilité de densité constante. Cette hypothèse exprime simplement l'ignorance où nous sommes, au départ, quant à la localisation exacte d'un phénomène déterministe, et n'implique donc pas encore un engagement épistémologique bien grave.

Relativement à cette implantation aléatoire de l'origine x_0 sur le segment $(0, a)$, l'estimateur $Q^*(x_0)$ devient une variable aléatoire. Calculons son espérance mathématique, et sa variance : nous dirons, par définition, que cette variance est la variance d'estimation $\sigma^2(a)$. Pour l'espérance, on trouve :

$$E(Q^*) = \int_0^a \frac{dx_0}{a} a \sum_p f(x_0+pa) = \int_{-\infty}^{+\infty} f(x) dx = Q$$

Notre estimateur est donc sans biais. La variance d'estimation est alors :

$$\sigma^2(a) = E(Q^{*2}) - Q^2 = \frac{1}{a} \int_0^a [Q^*(x_0)]^2 dx_0 - Q^2$$

Evaluons l'intégrale de $[Q^*(x_0)]^2$. On a :

$$[Q^*(x_0)]^2 = a^2 \sum_{p=-\infty}^{+\infty} \sum_{q=-\infty}^{+\infty} f(x_0+pa) f(x_0+qa) = a^2 \sum_k \sum_p f(x_0+pa) f(x_0+pa+ka)$$

Intégrons de 0 à a. Il vient :

$$\begin{aligned} \int_0^a [Q^*(x_0)]^2 dx_0 &= a^2 \sum_k \int_0^a \sum_p f(x_0+pa) f(x_0+pa+ka) dx_0 = \\ &= a^2 \sum_k \int_{-\infty}^{+\infty} f(x_0) f(x_0+ka) dx_0 = a^2 \sum_k g(ka) \end{aligned}$$

Compte tenu aussi de la relation (1-2), on obtient donc la formule :

$$(1-14) \quad \sigma^2(a) = a \sum_{k=-\infty}^{+\infty} g(ka) - \int_{-\infty}^{+\infty} g(h) dh$$

Dans l'espace à n dimension, on a un résultat tout-à-fait analogue. Par exemple, pour $n = 3$, et une maille parallélépipédique a_1, a_2, a_3 , il vient :

$$\sigma^2(a_1, a_2, a_3) = (a_1 a_2 a_3) \sum_{k_1} \sum_{k_2} \sum_{k_3} g(k_1 a_1, k_2 a_2, k_3 a_3) - \int g(h) dh$$

Remarque - La variance d'estimation (1-14) est la différence entre une valeur approchée et la valeur exacte de l'intégrale $\int g(h) dh$. Elle est donc d'autant plus petite que :

- la maille a est plus petite,
- la fonction g, donc aussi la V.R. elle-même est plus régulière.

Si la maille a est petite vis-à-vis de la portée, la formule (1-14) comporte un grand nombre de termes. On est conduit à chercher des formules d'approximation. (Voir parag. 1-4-3 et 1-4-4). Donnons maintenant une seconde expression (équivalente) de la variance d'estimation, en utilisant cette fois la transformée de Fourier G du covariogramme g. L'estimateur $Q^*(x_0)$, en tant que fonction périodique, admet le développement en série de Fourier :

$$\begin{cases} Q^*(x_0) = \sum_{p=-\infty}^{+\infty} c_p e^{2i\pi p \frac{x_0}{a}} \\ c_p = \frac{1}{a} \int_0^a Q^*(x_0) e^{-2i\pi p \frac{x}{a}} dx \end{cases}$$

Si nous remplaçons $Q^*(x_0)$ par son expression explicite, on voit que le coefficient de Fourier c_p est :

$$c_p = \sum_k \int_0^a f(x+ka) e^{-2i\pi p \frac{x}{a}} dx = \int_{-\infty}^{+\infty} f(y) e^{-2i\pi p \frac{y}{a}} dy$$

Par conséquent, si $\Phi(u)$ désigne la transformée de Fourier de $f(x)$, il vient :

$$c_p = \Phi\left(\frac{p}{a}\right)$$

En particulier $c_0 = \Phi(0) = Q$. Les propriétés d'orthogonalités des fonctions trigonométriques dorment ensuite simplement :

$$\sigma^2(a) = \sum_p |c_p|^2 - Q^2$$

et, compte tenu de (1-4) :

$$(1-15) \quad \sigma^2(a) = \sum_p G\left(\frac{p}{a}\right) - G(0)$$

Cette formule, équivalente à (1-14), montre que $\sigma^2(a)$ est positive dès que G est une fonction positive (donc dès que g est de type positif). Elle se généralise au cas pluridimensionnel. Pour \mathbb{R}^2 , par exemple, avec la maille rectangulaire (a_1, a_2) on trouve

$$(1-16) \quad \sigma^2(a_1, a_2) = \sum_{p,q} G\left(\frac{p}{a_1}, \frac{q}{a_2}\right) - G(0,0)$$

Remarque - De ce qui précède, on retiendra surtout ce résultat capital : la variance d'estimation $\sigma^2(a)$ ne dépend que du covariogramme transitif $g(h)$ et de la maille a . La formule (1-14) montre même qu'elle dépend linéairement de $g(h)$. Si donc nous connaissions le vrai covariogramme g d'une V.R., nous serions effectivement à même de résoudre le problème de l'estimation d'une manière entièrement satisfaisante - et sans probabiliser le phénomène réel. En vue de préparer le renversement de point de vue qui aura lieu au paragraphe (1-4-6), montrons qu'il y a là quelque chose d'illusoire.

En effet, si les seules données expérimentales sont les $f(x_0+pa)$, pour un x_0 fixé, on ne connaîtra pas le vrai covariogramme $g(h)$. On pourra seulement estimer les valeurs numériques des $g(ka)$ à l'aide des estimateurs :

$$g^*(ka) = a \sum_p f(x_0+pa) f(x_0+pa+ka)$$

De même, pour estimer le terme $Q^2 = \int g(h)dh$, on ne pourra que prendre le carré $[Q^*(x_0)]^2$ de l'estimateur $Q^*(x_0)$ lui-même. Mais, si nous substituons ces estimateurs dans la formule (1-14), nous obtenons identiquement Q, comme le montre un calcul facile :

$$a \sum_k g^*(ka) - [Q^*(x_0)]^2 = 0$$

Ce résultat se comprend assez bien : en remplaçant dans (1-14) les vraies valeurs par leurs estimations, nous avons implicitement substitué à la vraie V.R. $f(x)$, définie sur la droite réelle, sa restriction $f(x_0+pa)$ à l'ensemble discret constitué par les points de prélèvements x_0+pa eux-mêmes. Le résultat obtenu signifie donc simplement que $Q^*(x_0)$ est un excellent estimateur de $Q^*(x_0)$ lui-même - comme on pouvait s'en douter.

En termes moins triviaux, cela signifie qu'il est illusoire d'espérer, par des procédés purement empiriques, extraire d'un même matériel expérimental (les $f(x_0+pa)$) à la fois une estimation de Q et la précision de cette estimation.

Pour sortir de cette impasse méthodologique, le recours à un modèle théorique est obligatoire. On fera choix d'une expression mathématique appropriée $g(h;\lambda,\mu\dots)$ dépendant d'un ou plusieurs paramètres $\lambda, \mu\dots$ que l'on ajustera au mieux aux $g^*(ka)$ expérimentaux (il est inutile, et illusoire, de chercher à faire passer exactement la courbe théorique par les points expérimentaux), et c'est à ce $g(h)$ théorique que l'on appliquera la formule (1-14). Le résultat que l'on obtiendra aura exactement autant de valeur que le choix du modèle de covariogramme.

Indiquons tout de suite que ce choix n'est nullement arbitraire. Tout d'abord, la fonction $g(h;\lambda,\mu)$ doit être de type positif (paragraphe 1-2). Ensuite, on devra attacher une importance extrême au comportement analytique de cette fonction au voisinage de l'origine, c'est-à-dire à sa partie irrégulière : l'ordre du terme irrégulier de plus bas degré, éventuellement la présence d'un effet de pépite, ont la signification de lois physiques et constituent des caractères objectifs du phénomène réel. L'expérience aidant, on sait vite (en profitant des occasions où l'on disposait de réseaux de prélèvements particulièrement denses) que tel ou tel type de phénomène se caractérise par tel ou tel type de comportement analytique de son covariogramme. Comme en définitive (par. 1-4-3) c'est également de ce comportement que dépendra pour l'essentiel la variance d'estimation,

le problème se ramènera, en gros, à déterminer le coefficient du terme irrégulier de plus bas degré et cela pourra, en général, se faire expérimentalement d'une manière satisfaisante.

1-4-2 Cas d'une maille aléatoire stratifiée.

Soit $f(x)$ une V.R. dans \mathbb{R}^n , $g(h)$ son covariogramme, et soit V le parallélépipède de côtés a_1, a_2, \dots, a_n centré à l'origine. Désignons par h_i les vecteurs dont les composantes sont des multiples entiers $p_1 a_1, \dots, p_n a_n$ des côtés du parallélépipède V , de manière à ce que les translats V_{h_i} de V par les h_i recouvrent l'espace entier. La maille aléatoire stratifiée définie par le parallélépipède V se construit comme suit :

~ on choisit une origine $x_0 \in V$

~ pour chaque indice i , on choisit au hasard un point $X_i \in V$ de telle manière que les X_i soient indépendants et admettent la même loi de densité uniforme dans V .

~ on effectue un prélèvement en chacun des points $x_0 + h_i + X_i$.

L'estimateur $Q^*(x_0)$ de la quantité de métal, à $x_0 \in V$ fixé, est alors :

$$Q^*(x_0) = \sum_i v f(x_0 + h_i + X_i)$$

A x_0 fixé, les $v f(x_0 + h_i + X_i)$ sont des variables aléatoires indépendantes. Exprimons leurs espérances et leurs variances à l'aide de la régularisée :

$$f_v(x) = \frac{1}{v} \int_V f(x+y) dy$$

(qui représente la moyenne de f dans le parallélépipède v_x centré en x). On trouve :

$$\begin{cases} E(f(x_0 + h_i + X_i) | x_0) = \frac{1}{v} \int_V f(x_0 + h_i + y) dy = f_v(x_0 + h_i) \\ D^2(f(x_0 + h_i + X_i) | x_0) = \frac{1}{v} \int_V f^2(x_0 + h_i + y) dy - f_v^2(x_0 + h_i) \end{cases}$$

Ces variables étant indépendantes à x_0 fixé, on en déduit l'espérance et la variance conditionnelle de $Q^*(x_0)$:

$$\left\{ \begin{array}{l} E(Q^* | x_0) = \sum_1 v f_v(x_0 + h_1) = Q \\ D^2(Q^* | x_0) = v^2 \sum_1 D^2(f(x_0 + h_1 + x_1) | x_0) = v g(o) - v^2 \sum_1 f_v^2(x_0 + h_1) \end{array} \right.$$

A x_0 fixé, $Q^*(x_0)$ est ainsi un estimateur sans biais. mais - exactement comme dans le paragraphe précédent - notre ignorance initiale relativement à la localisation du phénomène réel se formalise en assimilant x_0 à une variable aléatoire, indépendante des X_1 , admettant elle aussi une densité uniforme dans v . Comme $E(Q^* | x_0) = Q$ ne dépend pas de x_0 , la variance a priori de $Q^*(x_0)$ - c'est à-dire la variance d'estimation de la maille aléatoire stratifiée - s'obtient en prenant directement l'espérance en x_0 de la variance conditionnelle $D^2(Q^* | x_0)$, soit

$$\begin{aligned} D^2(Q^*) &= \frac{1}{v} \int_v D^2(Q^* | x_0) dx_0 = v g(o) - v \sum_1 \int_v f_v^2(x_0 + h_1) dx_0 = \\ &= v g(o) - v \int_v f_v^2(x) dx \end{aligned}$$

Désignons par g_v le covariogramme de la régularisée f_v . D'après le paragraphe (1-3-1), il se déduit du covariogramme géométrique $K(h)$ du parallélépipède v par la relation :

$$g_v = \frac{1}{v^2} g * K$$

En particulier :

$$g_v(o) = \frac{1}{v^2} \int g(h) K(h) dh = \int f_v^2(x) dx$$

D'après l'algorithme de Cauchy, $g_v(o)$ représente la valeur moyenne de $g(x-y)$ lorsque x et y décrivent séparément le parallélépipède v (cf. Exercice 2), et la variance d'estimation cherchée s'écrit :

$$(1-17) \quad D^2(Q^*) = v[g(o) - g_v(o)]$$

Elle dépend donc uniquement du comportement du $g(h)$ dans un voisinage de l'origine défini précisément par le parallélépipède v lui-même.

1-4-3 Formules d'approximation pour l'espace à une dimension.

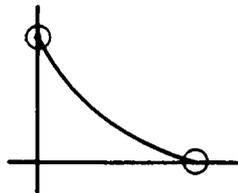
Soit maintenant $f(x)$ une V.R. de l'espace à une dimension, $g(h)$ son covariogramme, et a une maille de prélèvement. Lorsque la maille a est petite, la formule (1-14) comporte un assez grand nombre de termes, et il est avantageux d'obtenir des formules d'approximation permettant un calcul rapide. En réalité, l'intérêt des formules d'approximation que nous allons établir va bien au delà de simples raisons de commodité, et présente une signification épistémologique fondamentale pour le problème qui nous occupe. En effet, nous allons voir que (à un terme fluctuant près dont nous devons examiner le sens) la variance d'estimation $\sigma^2(a)$, pour a petit, dépend uniquement du comportement analytique du $g(h)$ au voisinage de $h = 0$.

La structure de la formule (1-14) montre que la variance d'estimation $\sigma^2(a)$ n'est autre que l'erreur numérique que l'on commettrait si l'on voulait évaluer l'intégrale $\int g(h) dh$ à partir des sommes discrètes $\sum_k g(ka)$. Si donc $g(h)$ était une fonction suffisamment régulière (c'est-à-dire partout dérivable un nombre suffisant de fois) on pourrait évaluer directement $\sigma^2(a)$ par les méthodes classiques fondées sur la formule d'Euler-Mac Laurin. Mais un $g(h)$ n'est pas, en général, suffisamment régulier en tout h . Les modèles théoriques de covariogrammes que l'on utilise en pratique localisent les irrégularités analytiques en deux endroits cruciaux :

~ au voisinage de l'origine, comme nous l'avons déjà vu

~ mais aussi au voisinage de $h = b$, b désignant la portée, car le raccordement en b du $g(h)$ avec l'axe des h peut être plus ou moins brutal.

Partout ailleurs, les modèles théoriques de $g(h)$ ne présentent jamais d'irrégularité analytique. Cela ne veut évidemment pas dire qu'il en soit encore ainsi pour les vrais covariogrammes. Mais nous ne connaissons jamais expérimentalement ces vrais covariogrammes, et, comme nous l'avons vu à la fin du paragraphe (1-4-1), nous sommes toujours obligés de leur substituer un modèle théorique. Nous examinerons dans le paragraphe 1-4-6 la portée épistémologique de cette opération ambiguë qui consiste à remplacer le vrai covariogramme (inconnu) par un modèle théorique analytiquement simple, dont les irrégularités se trouvent concentrées au voisinage de l'origine et de la portée :



Pour l'instant, nous partons de l'hypothèse que les irrégularités se trouvent effectivement localisées en ces deux points. Par des calculs sur lesquels nous n'insisterons pas ici (cf. [4]), il est alors possible de montrer que la variance d'estimation, pour a petit, est la somme de deux termes :

$$\sigma^2(a) = T(a) + Z(a)$$

Le premier, $T(a)$, ou terme d'extension, dépend du comportement de $g(h)$ en $h = 0$, le second, $Z(a)$, terme fluctuant ou Zitterbewegung, du comportement de $g(h)$ au voisinage de la portée b .

En ce qui concerne le terme fluctuant $Z(a)$, on peut montrer qu'il dépend essentiellement de la quantité $\varepsilon = b/a$ modulo 1 (précisément : on peut trouver un entier n tel que $na \leq b < (n+1)a$, et on pose $\varepsilon = \frac{b-na}{a}$). $Z(a)$ est une fonction périodique de ε , de période 1, et admet un développement en série de Fourier sans terme constant, donc de valeur moyenne nulle lorsque ε varie entre 0 et 1 :

$$\int_0^1 Z_\varepsilon(a) d\varepsilon = 0$$

Il est possible de déterminer l'expression théorique de ce terme fluctuant. Son ampleur n'est absolument pas négligeable, elle est souvent même considérable (voir Fig. 1 et Ex. 14). Mais, dans les applications, on ne peut jamais calculer la valeur exacte de ce terme, puisque précisément la valeur exacte de la portée b n'est jamais connue qu'à $\pm a$ près, donc ε , qui est b/a modulo 1, totalelement indéterminé. On s'autorise alors du fait que la valeur moyenne en $\varepsilon \in [0,1]$ de ce terme fluctuant est toujours nulle pour le négliger purement et simplement, et c'est ce que nous ferons dans ce qui suit. On notera le caractère probabiliste camouflé de cette approximation, sur lequel nous reviendrons dans le paragraphe 1-4-6.

Examinons maintenant le terme principal $T(a)$, ou terme d'extension, lié uniquement au comportement analytique du $g(h)$ en $h = 0$. Une étude analytique détaillée (cf [4]) montre qu'il ne dépend que de la partie irrégulière du développement limité de $g(h)$ au voisinage de l'origine. Plus précisément, on établit la validité d'un principe de correspondance terme à terme analogue à celui que nous avons rencontré lors de l'étude de la montée isotrope : chaque terme en $|h|^\lambda$ du développement limité de $g(h)$ apporte à $\sigma^2(a)$ une contribution en $T_\lambda a^{1+\lambda}$, avec un coefficient T_λ qui ne dépend que de λ :

$$(1-18) \quad |h|^\lambda = T_\lambda a^{1+\lambda}$$

Pour λ entier pair, on trouve $T_\lambda = 0$, et ceci confirme que la partie régulière du covariogramme n'apporte aucune contribution au développement limité de la variance d'estimation. Pour λ entier impair, la règle (1-18) subsiste sans modification. Pour un terme logarithmique en $h^{2n} \log h$, on trouve sans peine :

$$h^{2n} \log h = T_{2n}' a^{1+2n}$$

avec un coefficient T_{2n}' égal à la valeur en $\lambda = 2n$ de la dérivée $\frac{d}{d\lambda} T_\lambda$.

Calcul du coefficient T_λ . - On trouvera dans [4] un calcul rigoureux du coefficient T_λ qui figure dans la règle (1-18) de correspondance. Nous nous contenterons ici d'une justification symbolique (analogue à celle que nous avons déjà donnée pour le calcul du coefficient A_λ de la montée). Si $G(u)$ est la transformée de Fourier du covariogramme g , on a, d'après (1-15) :

$$\sigma^2(a) = 2 \sum_{p=1}^{\infty} G\left(\frac{p}{a}\right)$$

Si cela avait un sens de prendre comme covariogramme la fonction $|h|^\lambda$, la transformée G (au sens des distributions) serait :

$$G(u) = \frac{1}{\pi^{\lambda + \frac{1}{2}}} \frac{\Gamma(\frac{\lambda+1}{2})}{\Gamma(-\frac{\lambda}{2})} \frac{1}{|u|^{1+\lambda}}$$

et, si la formule (1-15) restait valable pour cette transformée prise au sens des distributions, on trouverait donc :

$$\sigma^2(a) = \frac{2 a^{1+\lambda}}{\pi^{\lambda + \frac{1}{2}}} \frac{\Gamma(\frac{\lambda+1}{2})}{\Gamma(-\frac{\lambda}{2})} \sum_{p=1}^{\infty} \frac{1}{p^{1+\lambda}}$$

D'où l'on déduirait l'expression de T_λ :

$$(1-19) \quad T_\lambda = \frac{2}{\pi^{\lambda + \frac{1}{2}}} \frac{\Gamma(\frac{\lambda+1}{2})}{\Gamma(-\frac{\lambda}{2})} \sum_{p=1}^{\infty} \frac{1}{p^{1+\lambda}}$$

Il se trouve que la valeur de T_λ obtenue par ce raisonnement approximatif est bien la valeur correcte. En particulier :

$$\begin{cases} T_1 = -\frac{1}{6} \\ T_2 = 0,0609\dots \end{cases}$$

Le principe de correspondance permet un calcul facile et rapide de la variance d'estimation. Mais on n'oubliera pas que l'on a négligé un Zitterbewegung, dont l'ampleur peut être énorme, ni que ce que l'on obtient est en réalité un développement limité de $\sigma^2(a)$ au voisinage de $a = 0$, donc un résultat utilisable seulement pour les petites mailles. Pour une maille a grande, d'ailleurs, la formule générale (1-14) ne comporte plus qu'un nombre modeste de termes et peut être utilisée directement. Lorsque la maille a devient supérieure à la portée, seul le terme correspondant à $k = 0$ subsiste dans (1-14), et on obtient alors la formule simple :

$$(1-20) \quad \sigma^2(a) = a g(0) - \int g(h) dh \quad (a \geq b)$$

dont le sens probabiliste est indiqué dans l'Exercice 15.

Remarque - Si $L = na$ est la longueur minéralisée, et n le nombre de prélèvements positifs, au covariogramme en r^λ correspond la variance d'estimation :

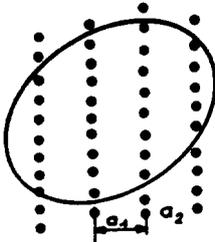
$$\sigma^2(a) = T_\lambda a^{1+\lambda} = (T_\lambda L^{1+\lambda}) \frac{1}{n^{1+\lambda}}$$

Elle est en raison inverse de $n^{1+\lambda}$ (et non en $1/n$ comme l'aurait suggéré une application maladroite de la statistique classique).

1-4-4 Formules d'approximation dans \mathbb{R}^n

Nous allons étudier en détail le cas de l'espace à deux dimensions, la généralisation à \mathbb{R}^n se faisant ensuite d'elle-même. Notre but est de dégager un second principe d'approximation, permettant de ramener le calcul d'une variance d'estimation $\sigma^2(a_1, a_2)$ à la somme de deux variances d'estimations à une seule dimension, justifiables séparément de la règle de correspon-

dance (1-18) - (1-19), donc très faciles à calculer. Nous désignerons ce principe sous le nom de principe de composition des termes de lignes et des termes de tranches. Examinons d'abord le sens intuitif de ce principe. Soit (a_1, a_2) , avec $a_1 > a_2$, une maille rectangulaire, $f(x, y)$ une V.R. définie dans le plan, $g(h_1, h_2)$ son covariogramme, et $g_1(h)$ le covariogramme que l'on déduit de g



par montée de long de l'axe des y . Notre principe de composition énonce que la variance d'estimation $\sigma^2(a_1, a_2)$ de la maille rectangulaire est la somme :

$$(1-20) \quad \sigma^2(a_1, a_2) = \sigma_1^2(a_1) + a_1 \sigma_2^2(a_2)$$

de deux termes : le premier, $\sigma_1^2(a_1)$ se calcule par la relation (1-14) appliquée au variogramme monté g_1 , et représente l'erreur que l'on commettrait en estimant la quantité de métal Q à partir des lignes de plus grande densité de prélèvement (parallèle à l'axe des y) ; c'est le terme de ligne. Le second, $a_1 \sigma_2^2(a_2)$ s'appelle terme de tranche. Il représente l'erreur que l'on commet en estimant l'accumulation des lignes à partir des prélèvements ponctuels. On calcule $\sigma_2^2(a_2)$ en appliquant la formule (1-14) à la maille a_2 et au covariogramme à 1 dimension $g(o, h)$. Le principe (1-20) exprime donc que, du point de vue du calcul des variances d'estimation, les deux erreurs commises en estimant l'accumulation des lignes à partir des prélèvements, puis la quantité de métal à partir de ces accumulations supposées connues, peuvent être regardées comme sans corrélation.

Précisons l'énoncé de ce principe, avant d'en donner une justification approchée. Pour abrégé les notations, nous désignerons par δ_a l'opérateur linéaire associé, selon (1-14), à la variance d'estimation pour une maille a :

$$\delta_a g = a \sum_{k=-\infty}^{+\infty} g(ka) - \int_{-\infty}^{+\infty} g(h) dh$$

Soit (x_0, y_0) les coordonnées de l'un des prélèvements. Désignons par

$$L_1(x_0) = \int f(x_0 + i a_1; y) dy$$

l'accumulation de la ligne i d'abscisse $x_0 + i a_1$, et par :

$$L_1^*(x_0, y_0) = a_2 \sum_{p=-\infty}^{+\infty} f(x_0 + i a_1; y_0 + p a_2)$$

l'estimateur que l'on peut en former. On a vu au paragraphe 1-4-1 que x_0 et y_0 doivent être considérés comme deux variables aléatoires indépendantes admettant des densités uniformes sur $(0, a_1)$ et $(0, a_2)$ respectivement. L'accumulation $L_1(x_0)$ de la ligne 1 est donc une variable aléatoire, et il est clair qu'à x_0 fixé on a :

$$E(L_1^* | x_0) = L_1(x_0)$$

Pour estimer $Q = \iint f(x, y) dx dy$, on forme l'estimateur :

$$Q^*(x_0, y_0) = a_1 \sum_i L_1^*(x_0, y_0)$$

et l'erreur d'estimation peut s'écrire :

$$(1-21) \quad Q - Q^* = (Q - a_1 \sum_i L_1) + a_1 \sum_i (L_1 - L_1^*)$$

Le premier terme est l'erreur que l'on commet en estimant Q à partir des accumulations. D'après la définition même de la montée, la variance de ce terme (terme de tranche) s'obtient en appliquant (1-14) au variogramme g_1 à 1 dimension déduit de $g_2(h_1, h_2)$ par montée le long de l'axe des y , soit, avec nos notations :

$$\sigma_1^2(a_1) = \delta_{a_1} g_1$$

Le second terme est la somme des erreurs $L_1 - L_1^*$ commises en estimant chacune des accumulations à partir des prélèvements ponctuels. Pour chaque ligne i , on vient de voir que l'espérance de cette erreur est nulle à x_0 fixé :

$$E(L_1 - L_1^* | x_0) = 0$$

A x_0 fixé, la V.R. $f(x_0 + i a_1; y)$ définie sur la ligne i admet le covariogramme g_1 à une dimension défini par :

$$(1-22) \quad g_1(h) = \int_{-\infty}^{+\infty} f(x_0 + i a_1; y) f(x_0 + i a_1; y+h) dy$$

et la variance conditionnelle de $L_1 - L_1^*$ est donc :

$$D^2(L_i - L_i^* | x_0) = \delta_{a_2} \varepsilon_i$$

Nous faisons alors l'hypothèse (que nous justifierons plus loin) selon laquelle les $L_i - L_i^*$ peuvent être considérées comme sans corrélation à x_0 fixé. La variance conditionnelle du terme de ligne est alors :

$$a_1^2 \sum_i D^2(L_i - L_i^* | x_0) = a_1^2 \sum_i \delta_{a_2} \varepsilon_i$$

Comme $E(L_i - L_i^*) = 0$, on en déduit la variance a priori du terme de ligne en prenant l'espérance en x_0 , soit

$$a_1 \int_0^{a_1} dx_0 \sum_i \delta_{a_2} \varepsilon_i$$

Mais il est clair que l'opérateur linéaire δ_{a_2} permute avec la sommation en i et l'intégrale en x_0 . En utilisant (1-22), on trouve donc pour le terme de ligne :

$$\begin{aligned} a_1 \delta_{a_2} \int_0^{a_1} dx_0 \sum_i \int_{-\infty}^{+\infty} f(x_0 + i a_1; y) f(x_0 + i a_1; y+h) dy = \\ = a_1 \delta_{a_2} \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} f(x, y) f(x, y+h) dy = a_1 \delta_{a_2} g(o, h) \end{aligned}$$

Le terme de ligne est donc bien :

$$a_1 \delta_{a_2} g(o, h) = a_1 \sigma_2^2(a_2)$$

conformément à (1-20). Pour obtenir la relation (1-20) elle-même, il reste encore à admettre l'absence de corrélation des deux erreurs $Q - a_1 \sum L_i$ et $a_1 \sum (L_i - L_i^*)$ qui figurent en (1-21).

Ce principe de composition des termes de ligne et des termes de tranches présente une importance très grande, pour des raisons pratiques, d'abord, puisqu'il permet de calculer très facilement une valeur approchée de la variance d'estimation, mais aussi pour des raisons méthodologiques : car, compte tenu du principe de correspondance, il montre que, dans l'espace à deux dimensions également, la variance d'estimation dépend essentiellement du comportement du $g(h)$ au voisinage de l'origine. Nous retrouverons ce principe de composition, sous une forme équivalente, dans la

version probabiliste de la théorie. Il n'est donc pas superflu d'en donner une justification au moins symbolique.

Pour cela nous appliquerons le principe de correspondance. A tout terme en r^λ de la partie irrégulière du covariogramme g correspond par montée le terme $A_\lambda r^{1+\lambda}$ dans le covariogramme monté g_1 (relations (1-10) et (1-12)), puis le terme $A_\lambda T_{1+\lambda} a_1^{2+\lambda}$ dans l'expression du terme de tranche (relations (1-18) et (1-19)). De même, r^λ contribue pour $T_\lambda a_1 a_2^{1+\lambda}$ au terme de ligne. Notre principe de composition se traduit donc par la règle de correspondance :

$$(1-23) \quad r^\lambda \rightarrow A_\lambda T_{1+\lambda} a_1^{2+\lambda} + T_\lambda a_1 a_2^{1+\lambda}$$

Inversement, la règle (1-23) permet de reconstituer le principe (1-20) (pourvu seulement que le covariogramme g soit isotrope ou se laisse ramener à la forme isotrope par une transformation linéaire). C'est cette règle que nous allons justifier.

Justification de la règle (1-23).

Partons de la formule générale (1-16), où $G(u,v)$ est la transformée de Fourier du covariogramme g . Cette formule s'écrit :

$$\sigma^2(a_1, a_2) = \sum_{p,q} G \left(\frac{p}{a_1}, \frac{q}{a_2} \right) - G(0,0)$$

Dans la somme double, les termes correspondant à $q = 0$ donnent le terme de tranche

$$\sigma_1^2(a_1) = \sum_p G \left(\frac{p}{a_1}; 0 \right) - G(0,0)$$

En effet, $G(u,0)$ est la transformée du covariogramme monté $g_1(h)$, et la formule (1-15) montre que le terme ci-dessus est justement la variance à 1 dimension $\sigma_1^2(a_1) = \sigma_{a_1}^2 g_1$ calculée sur le covariogramme monté g_1 . Le deuxième terme, ou terme de ligne, a donc pour expression exacte :

$$2 \sum_{q=1}^{\infty} \sum_{p=-\infty}^{+\infty} G \left(\frac{p}{a_1}, \frac{q}{a_2} \right)$$

Pour justifier (1-23), il faut montrer que la contribution à cette expression d'un terme en r^λ dans le développement de g est $T_\lambda a_1 a_2^{1+\lambda}$. Dans l'espace à deux dimensions, la transformée de r^λ est :

$$\int_2 r^\lambda = \frac{1}{\pi^{1+\lambda}} \frac{\Gamma(1 + \frac{\lambda}{2})}{\Gamma(-\frac{\lambda}{2})} \frac{1}{(u^2+v^2)^{1 + \frac{\lambda}{2}}}$$

Nous devons donc établir la relation :

$$(1-24) \quad \frac{2}{\pi^{1+\lambda}} \frac{\Gamma(1 + \frac{\lambda}{2})}{\Gamma(-\frac{\lambda}{2})} \sum_{q=1}^{\infty} \sum_{p=-\infty}^{+\infty} \frac{1}{\left(\frac{p^2}{a_1^2} + \frac{q^2}{a_2^2}\right)^{1 + \frac{\lambda}{2}}} = T_\lambda a_1 a_2^{1+\lambda}$$

pour un entier q fixé, on a :

$$(1-25) \quad \sum_{p=-\infty}^{+\infty} \frac{1}{\left(\frac{q^2}{a_2^2} + \frac{p^2}{a_1^2}\right)^{1 + \frac{\lambda}{2}}} = \left(\frac{a_2}{q}\right)^{\lambda+2} \sum_{p=-\infty}^{+\infty} \frac{1}{\left(1 + p^2 \frac{a_2^2}{q^2 a_1^2}\right)^{1 + \frac{\lambda}{2}}}$$

Or a_2 est supposé petit devant a_1 , et la quantité :

$$b = \frac{a_2}{q a_1}$$

est petite. La somme de Riemann $\sum_p \frac{b}{(1 + b^2 p^2)^{1 + \frac{\lambda}{2}}}$ donne donc une bonne approximation de l'intégrale

$$\int_{-\infty}^{+\infty} \frac{dx}{(1 + x^2)^{1 + \frac{\lambda}{2}}}$$

On peut d'ailleurs montrer que l'erreur commise en remplaçant cette intégrale par la somme discrète est d'un ordre très élevé en b . On en déduit, avec une approximation excellente :

$$\sum_p \frac{1}{(1 + b^2 p^2)^{1 + \frac{\lambda}{2}}} \approx \frac{1}{b} \int_{-\infty}^{+\infty} \frac{dx}{(1 + x^2)^{1 + \frac{\lambda}{2}}} = \frac{1}{b} \sqrt{\pi} \frac{\Gamma(\frac{1+\lambda}{2})}{\Gamma(1 + \frac{\lambda}{2})}$$

En remplaçant b par sa valeur $a_2/q a_1$ et en portant dans (1-25), on trouve ainsi :

$$\sum_p \frac{1}{\left(\frac{q^2}{a_2^2} + \frac{p^2}{a_1^2}\right)^{1 + \frac{\lambda}{2}}} = a_1 a_2^{1+\lambda} \sqrt{\pi} \frac{\Gamma(\frac{1+\lambda}{2})}{\Gamma(1 + \frac{\lambda}{2})} \frac{1}{q^{1+\lambda}}$$

Portons ce résultat dans le premier membre de (1-24). Il vient :

$$a_1 a_2^{1+\lambda} \frac{2}{\pi^{\lambda + \frac{1}{2}}} \frac{\Gamma(\frac{1+\lambda}{2})}{\Gamma(-\frac{\lambda}{2})} \sum_{q=1}^{\infty} \frac{1}{q^{1+\lambda}}$$

Mais, d'après (1-19) cette expression est justement égale à $T_{\lambda} a_1 a_2^{1+\lambda}$, de sorte que la relation (1-24) est vérifiée, donc aussi la règle de correspondance (1-23) et le principe de composition se trouve justifié.

1-4-5 Application à l'estimation d'une surface.

Appliquons le principe de composition du paragraphe précédent au cas de l'estimation d'une surface S. La V.R. à estimer est ici l'indicatrice k(x) de l'ensemble S. Nous avons vu que le covariogramme géométrique K(h) associé est linéaire au voisinage de l'origine (paragraphe 1), soit

$$K(h) = S - |h|D_{\alpha} + \dots$$

D_{α} désignant la variation diamétrale dans la direction α du vecteur h.

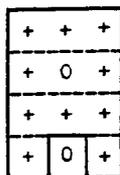
a/ Plaçons-nous d'abord dans le cas isotrope, c'est-à-dire dans le cas où la variation diamétrale $D_{\alpha} = D$ est à peu près indépendante de la direction α . Avec $T_1 = -\frac{1}{6}$ et $A_1 T_2 = -0,0609$, la règle (1-23) donne la variance d'estimation de S pour une maille rectangulaire (a_1, a_2) sous la forme :

$$(1-26) \quad \sigma_S^2 = D \left(\frac{1}{6} a_1 a_2^2 + 0,0609 a_1^3 \right) \quad (a_1 \leq a_2)$$

Cherchons plutôt la variance relative $\frac{\sigma_S^2}{S^2}$. A dire vrai, nous ne connaissons pas la vraie valeur de S, mais seulement son estimation $n a_1 a_2$, n désignant le nombre des points positifs du réseau. En divisant par $(n a_1 a_2)^{3/2} \sqrt{S}$, on obtient une expression intéressante :

$$(1-27) \quad \frac{\sigma_S^2}{S^2} = \frac{1}{n^{3/2}} \frac{D}{\sqrt{S}} \left(\frac{1}{6} \sqrt{\frac{a_1}{a_2}} + 0,0609 \left(\frac{a_2}{a_1} \right) \right)^{3/2}$$

En fait, on ne connaît pas la vraie variation diamétrale D. Pour l'estimer, on remplace l'ensemble



S inconnu par la réunion des zones d'influences (rectangles a_1 , a_2) des points positifs. On compte les nombres $2 N_1$ et $2 N_2$ d'éléments linéaires a_1 et a_2 constituant le périmètre du contour ainsi obtenu, et les variations diamétrales correspondantes sont $D_1 = N_1 a_1$ et $D_2 = N_2 a_2$. Comme nous sommes, pour l'instant, dans le cas isotrope, il vient :

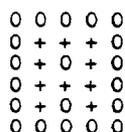
$$D = N_1 a_1 = N_2 a_2$$

Substituons dans (1-26), et divisons par $S^2 = (n a_1 a_2)^2$. Il vient :

$$(1-28) \quad \frac{\sigma_S^2}{S^2} = \frac{1}{N^2} \left[\frac{1}{6} N_2 + 0,061 \frac{(N_1)^2}{N_2} \right] \quad (N_2 \leq N_1)$$

b/ En général, cependant, le contour ne sera pas suffisamment isotrope pour que D_α puisse être regardée comme une constante D . Il présentera, par exemple, une direction principale d'allongement. Si l'un des côtés de la maille est parallèle à cette direction principale (ce qui sera souvent le cas), la formule (1-28) précédente reste applicable : en effet, prenant cette direction principale comme axe des x , et multipliant les ordonnées par un module convenable, nous obtenons une nouvelle figure, isotrope cette fois (au moins en première approximation) pour laquelle (1-28) est donc valable. Mais cette transformation linéaire n'a modifié ni N_1 , ni N_2 , ni la variance relative σ^2/S^2 , de sorte que (1-28) est valable aussi dans le cas de la figure anisotrope initiale.

Exemple - Sur la figure ci-après, l'aire minéralisée est estimée à 10 fois le rectangle de maille $a_1 a_2$. Elle comporte un trou (une lacune). Dans le décompte de N_1 et de N_2 doivent figurer aussi bien les éléments extérieurs que les éléments intérieurs. On lit donc sur la figure :



$$2 D_1 = 12 a_1 \quad \text{soit} \quad N_1 = 6$$

$$2 D_2 = 8 a_2 \quad \text{soit} \quad N_2 = 4$$

D'où, par conséquent :

$$\frac{\sigma_S^2}{S^2} = \frac{1}{100} \left[\frac{4}{6} + 0,061 \frac{36}{4} \right] = \frac{1,21}{100}$$

soit un écart type relatif $\sigma_S/S = 11/100$, et une fourchette d'erreur relative de $\pm 22\%$.

On se gardera toutefois d'oublier que ce calcul fait abstraction du terme fluctuant, ou Zitterbewegung, dont nous savons que l'amplitude peut être énorme.

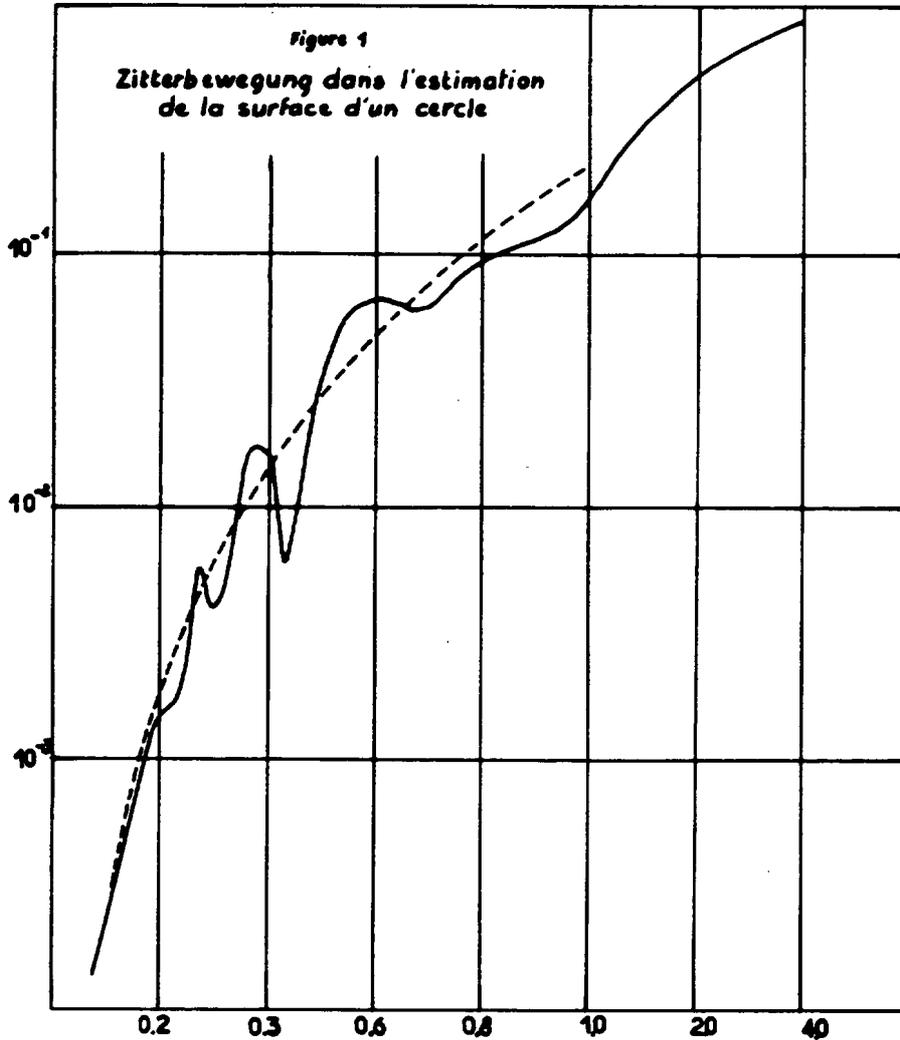


Fig 1. — Zitterbewegung dans l'estimation de la surface du cercle de diamètre unité à l'aide d'un réseau à maille carrée. En abscisse la maille a . En ordonnée la variance d'estimation correspondante $\sigma^2(a)$. la courbe en trait plein représente la valeur exacte, calculée à partir de la formule exacte (IV, 1, 12). la courbe en pointillé représente la formule $\sigma^2(a) = 0.2276 a^3 + 0.00477 a^5$ obtenue en négligeant le Zitterbewegung et en retenant les deux premiers termes du développement limité donné par le principe de correspondance.

1-4-6 Passage à la version probabiliste de la théorie.

Dans ce premier chapitre, nous avons adopté le point de vue spontané et, pourrait-on dire, naïf qui consiste à prendre tel qu'il se présente le phénomène que l'on veut étudier, sans faire à son sujet aucune hypothèse particulière. Mais, si les méthodes transitives prétendaient, au départ, construire la théorie des V.R. sans introduire d'interprétation probabiliste, cette prétention n'a pas tardé à se révéler injustifiée. Tout d'abord, un phénomène inattendu comme le Zitterbewegung (paragraphe 1-4-3) est très caractéristique d'une situation préaléatoire que les physiciens connaissent bien : celle où une légère modification des conditions initiales, indécidable expérimentalement (ici, le rapport portée sur maille modulo 1) provoque au bout d'un certain temps une modification radicale du phénomène observable. Plus profondément encore, le covariogramme $g(h)$ - si nous le connaissions parfaitement - présenterait de petites ondulations, des points singuliers, toute une structure de détail qui fournirait autant d'information, ou presque, que la donnée de la V.R. elle-même. Mais cette riche structure de détail n'est jamais accessible expérimentalement. A partir de points expérimentaux discontinus, nous ajustons un modèle théorique de covariogramme, qui est une courbe régulière et continue sauf au voisinage de l'origine et de la portée (remarque terminale du paragraphe 1-4-1). Plus précisément, nous sommes amenés à faire une hypothèse sur le comportement analytique (en h^λ par exemple) du $g(h)$ au voisinage de l'origine. En fait, il n'est nullement certain que le vrai $g(h)$ présente un comportement analytique d'un type aussi simple. D'un point de vue mathématique, les opérations de lissage ou de régularisation qu'entraîne une hypothèse de ce genre sont difficiles à analyser avec précision. Mais il est clair qu'elles oblitèrent nécessairement la structure fine du $g(h)$ et toute la richesse d'information potentielle qu'elle contient. Si le caractère mathématique de ces opérations est plutôt obscur, leur signification épistémologique est évidente : elles constituent un équivalent camouflé d'un passage à l'espérance mathématique. Ainsi, les méthodes transitives, qui se voulaient au départ purement géométriques, se révèlent, lorsque l'on analyse les conditions de leur mise en oeuvre effective, comme riches d'un contenu probabiliste implicite.

Nous allons donc faire apparaître au grand jour ces hypothèses probabilistes que dissimulaient les méthodes transitives, et ce sera notre théorie des fonctions aléatoires intrinsèques. Mais les phénomènes naturels auxquels nous nous intéressons doivent, en général, être regardés comme uniques, et l'inférence statistique ne sera (en principe) possible que moyennant une hypothèse spéciale, du type stationnaire ou intrinsèque. Cette hypothèse paraîtra parfois plausible, parfois franchement contraire aux données expérimentales et, dans tous les cas, comme arbitraire. Or, il peut sembler illusoire d'interpréter un phénomène de la nature comme une réalisation d'une fonction aléatoire si l'on n'est pas certain de pouvoir reconstituer sans ambiguïté la loi de probabilité de cette

F.A. à partir du phénomène réel. On voit ainsi apparaître une nouvelle contradiction. Tout-à-l'heure, les représentations transitives se sont révélées posséder un contenu probabiliste inavoué: mais elles n'impliquaient aucune hypothèse de stationnarité (et c'est en cela que réside leur intérêt méthodologique). Maintenant, nous allons mettre en pleine lumière ces hypothèses probabilistes auparavant cachées, mais il va nous sembler qu'elles n'ont de sens que moyennant cette hypothèse de stationnarité, dont justement les méthodes transitives n'avaient aucunement besoin.

Le rapprochement de ces points de vue opposés indique clairement la solution : l'interprétation probabiliste est inévitable, mais l'hypothèse stationnaire n'est pas réellement nécessaire. La V.R. devra être considérée, en général, comme une réalisation d'une F.A. non stationnaire, et l'outil de base sera une covariance non stationnaire $C(x,y)$ dépendant séparément des deux points d'appui x et y (et non pas seulement de la différence $x-y$). L'inférence statistique restera impossible, en ce sens que nous ne pourrons pas déterminer $C(x,y)$ à partir des données disponibles. Mais nous n'aurons pas besoin de connaître effectivement cette covariance. Pour le voir, il suffira de transposer en termes probabilistes les résultats des méthodes transitives. S'il s'agit, par exemple, d'estimer un domaine V , dont l'indicatrice est $k(x)$ et le covariogramme géométrique $K(h)$, les méthodes transitives montrent qu'il faut connaître la fonction :

$$g(h) = \int k(x) f(x) k(x+h) f(x+h) f(x) dx$$

Dans la version probabiliste, le covariogramme géométrique $g(h)$ devient lui-même une F.A., et c'est son espérance $E(g(h))$ qu'il est alors nécessaire et suffisant de connaître pour résoudre le problème de l'estimation. Or cette espérance se déduit de la covariance non stationnaire par la relation :

$$(1-29) \quad E[g(h)] = \int k(x) k(x+h) C(x,x+h) dx$$

C'est d'ailleurs déjà cette relation (1-29) que l'on appliquait implicitement lorsque l'on substituait un modèle théorique au vrai $g(h)$ inconnu. Cette espérance, à un facteur près, ne diffère pas de la covariance pseudo-stationnaire $\bar{C}(h)$, moyenne de $C(x,y)$ dans le domaine V à estimer, que l'on attribuerait à la fonction aléatoire si on la considérait comme stationnaire (alors qu'elle ne l'est peut-être pas). On a clairement, en effet :

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

Ainsi, c'est cette espérance $E[g(h)]$, ou cette covariance moyenne $\bar{C}(h)$ qu'il suffit de connaître pour résoudre le problème de l'estimation. De plus, les méthodes d'approximation que nous avons dégagées, montrent qu'il suffit même de connaître le comportement au voisinage de l'origine de cette fonction $E[g(h)]$ ou $\bar{C}(h)$. Nous montrerons dans ce qui suit que l'inférence statistique à partir d'une réalisation unique est possible pour ce qui nous intéresse (ce comportement au voisinage de l'origine) moyennant seulement une hypothèse de quasi-stationnarité que nous préciserons chemin faisant : mais cette hypothèse sera cette fois suffisamment faible pour être physiquement admissible dans tous les cas qui nous intéressent. Autrement dit, il sera possible de résoudre le problème de l'estimation dans un cadre explicitement probabiliste et en utilisant seulement les données expérimentales disponibles.

On mesure ainsi l'importance méthodologique de ce premier chapitre, qui a pu paraître un peu théorique au premier abord : une fois probabilisés, les résultats des méthodes transitives montrent qu'il est possible de s'affranchir de la fameuse hypothèse de stationnarité, et c'est pourquoi nous avons cru devoir en donner un exposé assez complet. Passons maintenant à la version probabiliste de la théorie.

1-5 EXERCICES SUR LES METHODES TRANSITIVES.

1-5-1 Exercices sur les paragraphes 1-1 et 1-2.

Exercice 1 (Fonction triangle) - Dans l'espace à une dimension, trouver le covariogramme géométrique $K(h)$ associé à l'indicatrice du segment de longueur b .

(Solution : $K(h) = b - |h|$ pour $|h| \leq b$, et 0 ailleurs. Noter le comportement linéaire au voisinage de $h = 0$ et le raccordement brutal en $h = b$).

Même question dans \mathbb{R}^2 pour le rectangle $a \times b$ (Sol. : $K(h_1, h_2) = (a - |h_1|)(b - |h_2|)$ pour $|h_1| \leq a$ et $|h_2| \leq b$) et dans \mathbb{R}^3 pour le parallélépipède $a \times b \times c$.

Exercice 2 (Algorithme de Cauchy) - Soit V un ensemble dans \mathbb{R}^n , $K(h)$ son covariogramme géométrique, et $g(h)$ une fonction quelconque d'un argument vectoriel h . Montrer que la valeur moyenne de $g(x-y)$ lorsque les extrémités x et y de l'argument de g décrivent séparément V se déduit de $K(h)$ par la formule :

$$\frac{1}{V^2} \int_V dx \int_V g(x-y) dy = \frac{1}{V^2} \int g(h) K(h) dh$$

(Solution : expliciter l'intégrale double à l'aide de l'indicatrice de V , faire un changement de variable et intervertir l'ordre des intégrations. Cet algorithme est d'un emploi constant en théorie des V.R.).

Application - (cf. Ex. 1) si $g = g(r)$ ne dépend que du module $r = |h|$, la valeur moyenne de g dans le rectangle $a \times b$ est :

$$\frac{4}{a^2 b^2} \int_0^a dx \int_0^b (a-x)(b-y) g(\sqrt{x^2+y^2}) dx dy$$

Exercice 3 - Soit dans \mathbb{R}^1 la fonction $f(x) = e^{-ax}$ pour $x \geq 0$ et $f(x) = 0$ pour $x < 0$. Montrez que le covariogramme associé est $g(h) = \frac{1}{2a} e^{-a|h|}$ (comportement linéaire à l'origine).

Exercice 4 - Soit $f(x)$ la fonction égale à x sur l'intervalle $(-b, b)$ et à 0 ailleurs. Montrez $g(h) = \frac{2}{3} b^3 - b^2|h| + \frac{1}{6} |h^3|$ (linéarité à l'origine). Vérifier $\int g(h) dh = 0$ et interpréter.

Exercice 5 - Soit dans \mathbb{R}^1 la fonction $f(x) = (b^2 - x^2)$ pour $-b \leq x \leq b$ et $f(x) = 0$ pour $x \in [-b, b]$. Montrez que le covariogramme associé est :

$$g(h) = b^5 \left(\frac{16}{15} - \frac{4}{3} \frac{h^2}{b^2} + \frac{2}{3} \frac{|h|^3}{b^3} - \frac{1}{30} \frac{|h|^5}{b^5} \right)$$

(Noter le terme irrégulier principal en h^3 : il s'agit d'une fonction continue. Plutôt que de faire un calcul direct, on montrera que si f est dérivable, son covariogramme $g = f * \check{f}$ a pour dérivée seconde $g'' = -f' * (\check{f}')$, et on utilisera l'exercice 3. On intégrera ensuite deux fois g'' , en notant $g'(0) = 0$ et en calculant directement $g(0)$).

Exercice 6 (Formule de Minkowski) - Soit S un ensemble suffisamment régulier dans \mathbb{R}^2 , $K(h)$ son covariogramme géométrique, $2\mathcal{L}$ son périmètre. Etablir la relation :

$$2\mathcal{L} = - \int_0^\pi K'_\alpha(0) d\alpha$$

(mettre $-K'_\alpha(0)$ sous la forme d'une intégrale curviligne $\frac{1}{2} \int |\cos(\theta-2)| ds$ étendue au contour de S , et intégrer d'abord en α).

De même si V est un ensemble de \mathbb{R}^3 , $K(h)$ son covariogramme géométrique, et S sa surface, montrez :

$$S = -\frac{1}{\pi} \int K'_\omega(o) d\omega$$

($d\omega$ angle solide élémentaire autour de la direction ω , intégration sur 4π stéradians)
(Ces formules sont très utiles en morphologie mathématique : elles permettent de reconstituer le périmètre ou la surface spécifique à partir de la dérivée de la fonction de covariance ou du covariogramme, donc d'observations purement linéaires).

Exercice 7 (Effet de pépité) - Dans \mathbb{R}^1 , on divise l'intervalle $(0, b)$ en n intervalles égaux de longueur $a = b/n$, et on considère n variables aléatoires indépendantes X_i admettant la même espérance m et la même variance σ^2 . On pose $f(x) = X_i$ pour $(i-1)a \leq x < ia$, $i = 1, 2, \dots, n$, et $f(x) = 0$ à l'extérieur de $(0, b)$. Trouver l'espérance $g(h)$ du covariogramme de $f(x)$.

(Solution : $\sigma^2(b-n|h|) + m^2(b-|h|)$ pour $|h| \leq a$, $m^2(b-|h|)$ pour $a < |h| \leq b$, 0 ailleurs. Pour a petit, ou n grand, ce schéma correspond à un effet de pépité).

Exercice 8 (Granulométrie d'un grain convexe) - a/ Soit S un ensemble convexe dans \mathbb{R}^2 , g son covariogramme géométrique qu'on exprimera sous la forme $g(r, \alpha)$, r et α désignant les coordonnées polaires. Pour α fixé, la dérivée en r , $-g'(r; \alpha)$ est la mesure de la projection de l'intersection $S \cap S_h$ sur une perpendiculaire à la direction α . En déduire que la granulométrie des traversées de S dans la direction α admet la fonction de répartition $F_\alpha(r)$ définie par :

$$1 - F_\alpha(r) = \frac{g'(r; \alpha)}{g'(0; \alpha)}$$

En déduire que, pour α fixé et $r \geq 0$, $g(r, \alpha)$ est une fonction convexe de r .

b/ On tire maintenant au sort la direction α selon la loi de densité

$$\frac{D}{2\pi} d\alpha = -\frac{g'(0; \alpha)}{2\pi} d\alpha \quad (0 \leq \alpha \leq \pi)$$

(cf. exercice 6) pour vérifier qu'il s'agit bien d'une loi de probabilité. Cette loi attribuée à chaque direction α une probabilité proportionnelle au contour apparent de S dans cette direction, et correspond exactement à la loi (conditionnelle) des inter-

sections de S avec un réseau préalable de droites aléatoires dont les directions sont uniformément distribuées sur $0, \pi$). Montrer que la granulométrie de la traversée correspondante est

$$1 - F(h) = - \frac{1}{2\mathcal{L}} \int_0^\pi g'(h; \alpha) d\alpha$$

avec la valeur moyenne

$$E(h) = \pi \frac{S}{2\mathcal{L}}$$

Exercice 9 (Covariogramme géométrique de la sphère) - Former le covariogramme $K(h)$ de la sphère de diamètre D dans \mathbb{R}^3 .

$$\begin{aligned} (K(h) = \frac{\pi D^3}{6} (1 - \frac{3}{2} \frac{|h|}{D} + \frac{1}{2} \frac{|h|^3}{D^3}) \text{ pour } |h| \leq D. \text{ Partir de la relation } K'(h) = \\ = \frac{\pi}{4} (D^2 - h^2) \text{ (cf ex. précédent) et intégrer en remarquant } g(0) = \frac{\pi D^3}{6}) \end{aligned}$$

1-5-2 Exercices sur la Montée.

Exercice 10 (montée d'ordre 2) - Effectuer directement la montée isotrope d'ordre 2 sur la fonction $g_n(r) = r^\lambda$ ($0 \leq r \leq 1$) $g_n(r) = 0$ ($r > 1$). En déduire que les coefficients A_λ de la relation (1-10) vérifient :

$$A_\lambda A_{\lambda+1} = \frac{2\pi}{\lambda+2}$$

(Solution : $g_{n-2}(r) = \frac{2\pi}{\lambda+2} (1 - r^{\lambda+2})$ pour $r \leq 1$, et 0 pour $r > 1$. La montée d'ordre 2 vérifie la règle $r^\lambda \rightarrow A_\lambda A_{\lambda+1} r^{\lambda+2}$ en ce qui concerne la partie irrégulière ; cette règle ne donne le résultat qu'à une série paire près, ici réduite à une constante).

Exercice 11 - Covariogramme de la fonction $f(x) = \frac{\pi}{4} (D^2 - x^2)$ pour $-D \leq x \leq D$, 0 ailleurs sur \mathbb{R}^1 (section de la sphère de diamètre D par un plan de cote x).

(Solution : $g(h) = \frac{\pi^2}{6} (\frac{1}{5} - \frac{h^2}{D^2} + \frac{|h|^3}{D^3} - \frac{1}{5} \frac{|h|^5}{D^5}) D^5$. Faire directement la montée d'ordre 2 sur le covariogramme géométrique de l'Ex. 9. Comparer avec Ex. 5).

Exercice 11 bis (Reconstitution de la granulométrie des sphères) - On se donne dans \mathbb{R}^3 une population de sphères. On désigne par θ_3 le nombre des sphères par unité de volume, et

par $F_3(D)$ la granulométrie de ces sphères. Autrement dit $\theta_3[1-F_3(D)]$ est le nombre des sphères de diamètre $\geq D$ présentes par unité de volume. Ces sphères induisent des cercles et des traversées (cordes) respectivement sur les plans et les droites de \mathbb{R}^3 . Désignons par θ_2 le nombre des cercles par unité de surface et par $1-F_2(D)$ leur granulométrie (qui est la granulométrie induite par F_3 sur les plans) ; et de même par θ_1 le nombre des traversées par unité de longueur et par $1-F_1(D)$ la granulométrie de ces traversées (induite sur les droites par F_3 ou, aussi bien, par F_2).

a/ Etablir les relations

$$\theta_1[1-F_1(h)] = \theta_2 \int_h^\infty \frac{D}{\sqrt{D^2-h^2}} (1-F_2(D)) dD = \frac{\pi}{4} \theta_3 \int_h^\infty (1-F_3(D)) 2 D dD$$

$$\theta_2[1-F_2(h)] = \theta_3 \int_h^\infty \frac{D}{\sqrt{D^2-h^2}} (1-F_3(D)) dD$$

(établir d'abord, par exemple, $\theta_1[1-F_1(h)] = \theta_2 \int_h^\infty \sqrt{D^2-h^2} F_2(dD)$, et faire une intégration par parties).

b/ Inverser ces relations : en utilisant (1-8), montrer que l'on a par exemple :

$$\begin{cases} \theta_1 F_1'(h) = \frac{\pi}{2} \theta_3 h[1-F_3(h)] \\ \theta_2[1-F_2(h)] = \frac{2}{\pi} \theta_1 \int_h^\infty F_1'(u) \frac{du}{\sqrt{u^2-h^2}} \end{cases}$$

En déduire $\theta_3 = \frac{2}{\pi} F_1''(0) \theta_1$, $\theta_2 = \frac{2}{\pi} \int_0^\infty \frac{1}{u} F_1'(du)$, et reconstituer les granulométries des cercles et des sphères à partir de celles des traversées.

Exercice 12 (Montée sur l'exponentielle de Gauss) - La montée d'ordre 1 transforme e^{-ar^2} en $\sqrt{\frac{\pi}{a}} e^{-ar^2}$.

1-5-3 Exercices sur l'estimation

Exercice 13 (Covariogramme exponentiel) - Soit, dans l'espace à une seule dimension, une V.R. dont le covariogramme transitif est $g(h) = e^{-\lambda|h|}$. Former l'expression exacte de la variance d'estimation $\sigma^2(a)$, trouver sa partie principale, et conclure à l'absence

de Zitterbewegung.

$$\text{(Solution : } \sigma^2(a) = a \frac{2}{1 - e^{-\lambda a}} - 1 - \frac{2}{\lambda a} \approx \frac{1}{6} \lambda a^2$$

(En poussant le développement plus loin, on identifierait ainsi les termes $T_1, T_3, T_5 \dots$ des relations (1-18). En particulier, $T_1 = -\frac{1}{6}$).

Exercice 14 (Covariogramme triangulaire) - Soit dans l'espace à une dimension la V.R. $f(x)$ égale à 1 sur l'intervalle $(0, b)$.

a/ Former le covariogramme de $f(x)$.

(Solution : $g(h) = b - |h|$ pour $|h| \leq b$ et 0 pour $|h| > b$; établir ce résultat par un raisonnement géométrique).

b/ Calculer la variance d'estimation pour une maille a petite, par la formule d'approximation. (Sol. : $1/6 a^2$).

c/ On dispose de $n+2$ sondages $S_0, S_1, \dots, S_n, S_{n+1}$ à maille régulière, dont le premier et le dernier sont négatifs et les autres positifs. On admet que chacune des deux extrémités de l'intervalle dont on veut estimer la longueur b peut tomber n'importe où sur le segment (S_0, S_1) et (S_n, S_{n+1}) respectivement, indépendamment l'une de l'autre et avec une probabilité uniforme. La longueur b devient ainsi une variable aléatoire. Montrer $E(b) = na$; $D^2(b) = a^2/6$: on retrouve bien ainsi la variance d'estimation telle que la donne la formule d'approximation, mais on n'atteint pas le terme fluctuant.

d/ Calculer la valeur exacte de la variance d'estimation $\sigma^2(a)$ du b/.

(Solution : poser $b = na + \varepsilon a$, avec $0 \leq \varepsilon < 1$. Si le premier prélèvement positif est en x ($0 \leq x < a$), on a l'estimation $(n+1)a$ pour $x \leq \varepsilon a$, et na pour $\varepsilon a < x < a$. Le point x étant implanté au hasard sur le segment $(0, a)$ avec une densité uniforme de probabilité, cette estimation est une variable aléatoire d'espérance b et de variance $(\varepsilon - \varepsilon^2)a^2$. Cette expression tient compte du Zitterbewegung. Lorsque ε (qui est toujours inconnu dans la pratique) est assimilé à une variable aléatoire uniformément distribuée sur $(0, 1)$, la valeur moyenne de cette expression coïncide avec l'expression approchée du b/ (Zitterbewegung exclu).

Exercice 15 (Variance d'estimation dans le cas aléatoire) - Soit, dans l'espace à deux dimensions, une V.R. $f(x)$ de covariogramme $g(h)$ et de support S . Si le rectangle de maille

(a_1, a_2) est grand devant S , un prélèvement au plus est positif, et tout se passe comme si ce prélèvement unique était implanté au hasard dans un rectangle de côtés a_1, a_2 contenant S . Si ce prélèvement tombe en un point $x \in S$, on prend l'estimation $Q^*(x) = a_1 a_2 f(x)$, et 0 si $x \notin S$. $Q^*(x)$ est ainsi une variable aléatoire. Montrer directement $E(Q^*) = \int f(x) dx = Q$, $E[(Q^*)^2] = \left[\int [f(x)]^2 dx \right] a_1 a_2$ et $D^2(Q^*) = a_1 a_2 g(0) - \int g(h) dh$.

1-5-4 Théorie transitive pour les mesures.

Exercice 16 (Généralisation de la théorie transitive pour les mesures) - Soit $f(x)$ une fonction positive sommable sur \mathbb{R}^n . On suppose $\int f(x) dx = 1$ (ce qui est loisible) et on considère $f(x)$ comme la densité de probabilité associée à une variable vectorielle $X = (X_1, X_2, \dots, X_n)$. Soient X et X' deux V.A. indépendantes admettant la même loi f . Montrer que le covariogramme $g = f * \check{f}$ est la densité de probabilité de la variable $X - X'$.

Exercice 17 (Suite) - Soit $X = (X_1, X_2, \dots, X_n)$ une variable vectorielle, μ sa loi de probabilité (mesure positive de somme unité définie sur \mathbb{R}^n) n'admettant plus nécessairement de densité. On définira le covariogramme $g = \mu * \check{\mu}$ comme la mesure donnant la loi de $X - X'$, X et X' désignant deux V.A. indépendantes de même loi μ .

a/ Examiner le cas où μ est une loi discrète concentrée en N points x_1, x_2, \dots, x_N avec $P(X = x_i) = \frac{1}{N}$. Montrer que g est la loi discrète comportant la masse $\frac{1}{N}$ à l'origine et la masse $\frac{1}{N^2}$ en chacun des points $x_i - x_j$, $i \neq j$.

b/ Les N points x_i du a/ ci-dessus sont maintenant considérés comme aléatoires, mutuellement indépendants, implantés dans \mathbb{R}^n selon une même loi de densité f . La loi g du paragraphe a est alors considérée comme une loi conditionnelle à x_i fixés. Montrer que la loi a priori (déconditionnée) correspondante est de la forme

$$g = \frac{1}{N} \delta + \left(1 - \frac{1}{N}\right) f * \check{f}$$

(δ mesure de Dirac : $\frac{1}{N} \delta$ désigne donc une masse $\frac{1}{N}$ placée en 0, et représente un effet de pépite : dans le paragraphe 1-2, l'effet de pépite a été représenté par une zone de variation très rapide et de dimension très petite entourant l'origine. Il est naturel, à la limite, d'idéaliser l'effet de pépite en le représentant par une

mesure de Dirac placée en 0).

Exercice 18 (Suite) - Soit p une fonction positive de somme unité. Montrer que la régularisée $\mu * \check{p}$ est la densité de la loi de $X-Y$, X et Y désignant deux variables vectorielles indépendantes de lois μ et p respectivement. Interpréter le covariogramme $\mu * \check{\mu} * p * \check{p}$ de la régularisée $\mu * \check{p}$ (loi de $X-X'-Y+Y'$). Examiner le cas où μ est la loi discrète de Ex. 17 a/. Lorsque les points x_i sont aléatoires, comme dans Ex. 17, b/, le covariogramme $\mu * \check{\mu} * p * \check{p}$ ci-dessus représente une loi conditionnelle. Montrer en déconditionnant relativement aux x_i , que la loi a priori correspondante est :

$$\frac{1}{N} p * \check{p} + (1 - \frac{1}{N}) p * \check{p} * f * \check{f}$$

On conclura de ces exercices qu'il est naturel de représenter un effet de pépité par une mesure de Dirac placée en 0. Un covariogramme avec effet de pépité sera de la forme $g = C\delta + g_1$, où g_1 est continue à l'origine. Si l'on régularise par une fonction p la mesure dont le covariogramme est g , on obtient une fonction dont le covariogramme est $g * P = C P + g_1 * P$ ($P = p * \check{p}$) de sorte que la règle du paragraphe 1-3-1 reste valable.

Exercice 19 (Suite) - Soit μ_n la loi de la variable vectorielle (X_1, \dots, X_n) à n composantes. La montée selon l'axe des x_n est l'opération qui fait passer de μ_n à la loi marginale μ_{n-1} des $n-1$ variables $(X_1, X_2, \dots, X_{n-1})$. Montrer que le covariogramme $\mu_{n-1} * \check{\mu}_{n-1}$ se déduit du covariogramme $\mu_n * \check{\mu}_n$ par la même montée.

Plus généralement, si p_n est une fonction (ou une mesure) sur \mathbb{R}^n , et p_{n-1} la fonction (ou la mesure) qui s'en déduit par montée, la régularisée $\mu_n * \check{p}_n$ devient $\mu_{n-1} * \check{p}_{n-1}$ par cette même montée. Comment se comporte à la montée un effet de pépité?

Exercice 20 (Suite) - (Estimation en présence d'un effet de pépité) - Soit f une mesure dont le covariogramme $g = C\delta + g_1$ comporte un effet de pépité et une composante continue g_1 . Soit $k(x)$ l'indicatrice d'un petit volume v , K son covariogramme. On régularise f par $1/v k(x)$, soit $f_v = \frac{1}{v} f * k$: $f_v(x)$ est la teneur d'un échantillon v implanté en x . On prélève des échantillons v à maille régulière $a = (a_1, a_2, \dots, a_n)$. Montrer que la variance d'estimation est du type $\sigma^2(a) = \sigma_p^2(a) + \sigma_1^2(a)$, $\sigma_1^2(a)$ se calculant à partir de g_1 par les formules du paragraphe 1-4-1, et le terme supplémentaire de

pépité est $\sigma_p^2(a) = C \left(\frac{a_1 a_2 \dots a_n}{V} - 1 \right)$ ou encore $C[V/V' - 1]$: V volume du champ, V' volume total des prélèvements (ainsi ce terme supplémentaire dû à l'effet de pépité est en raison inverse du volume total des prélèvements).

CHAPITRE 2LES FONCTIONS ALEATOIRES INTRINSEQUES
-----2-1 DEFINITIONS GENERALESNotion de Fonction Aléatoire.

En théorie des probabilités, on définit la notion de variable aléatoire (V.A.) vectorielle $Y = (Y_1, Y_2, \dots, Y_k)$ à k composantes : c'est une famille de k V.A. ordinaires Y_1, Y_2, \dots, Y_k (en général non indépendantes). Lorsque le nombre k de ces composantes devient infini, on obtient une famille infinie de variables aléatoires : c'est une fonction aléatoire. En particulier, si x est un point décrivant l'espace à n dimensions R^n , on peut définir une famille infinie $(Y_x)_{x \in R^n}$. A tout point x_0 de l'espace correspond ainsi une V.A. ordinaire Y_{x_0} , que l'on peut aussi bien noter $Y(x_0)$. $Y(x)$ est alors une fonction du point x , dont la "valeur" en x_0 n'est pas un nombre, mais une V.A. (à savoir $Y(x_0)$). On dit que $Y(x)$ est une fonction aléatoire (en abrégé, F.A.). On remarquera bien qu'en général les V.A. correspondant à deux points d'appui x_1 et x_2 , soient $Y(x_1)$ et $Y(x_2)$ ne sont pas indépendantes.

Si Y est une V.A. ordinaire, le résultat d'un tirage au sort effectué selon la loi de probabilité de Y est une valeur numérique particulière y . De même, si Y est une V.A. vectorielle (Y_1, Y_2, \dots, Y_k) , un tirage au sort selon la loi (à k variables) de Y donne un vecteur $y = (y_1, y_2, \dots, y_k)$, c'est-à-dire k valeurs numériques particulières. Enfin, si $Y(x)$ est une F.A. - c'est-à-dire une V.A. vectorielle à une infinité de composantes - un tirage au sort, effectué selon la loi (à une infinité de variables) de $Y(x)$ donne une fonction numérique particulière $y(x)$, en général extraordinairement irrégulière. On dit que $y(x)$ est une réalisation de la F.A. $Y(x)$.

On peut toujours considérer une réalisation $y(x)$ d'une F.A. $Y(x)$ comme une variable régionalisée. Inversement, on peut interpréter une V.R. donnée $y(x)$ comme une réalisation d'une certaine F.A. $Y(x)$: cette interprétation permet d'appliquer aux V.R. les résultats de la théorie probabiliste des F.A.

Remarques - 1/ On ne peut jamais dire qu'une V.R. donnée $y(x)$ est une F.A. Cela n'aurait pas plus de sens que de dire : le nombre 98 est une V.A. L'énoncé correct de l'hypothèse probabiliste de base que nous désirons introduire est : $y(x)$ est une réalisation d'une F.A. $Y(x)$.

2/ Pour que cette hypothèse probabiliste ait une signification réelle, il faut que l'on puisse reconstituer au moins en partie la loi de la F.A. $Y(x)$ dont la V.R. $y(x)$ est supposée être une réalisation, et cela suppose que l'inférence statistique soit possible. Or, l'inférence statistique n'est en général pas possible si l'on ne dispose que d'une seule réalisation $y(x)$ de $Y(x)$ (de même, on ne peut pas reconstituer la loi d'une V.A. Y à partir du résultat numérique $y = 98$ d'une épreuve unique). Pour que l'inférence statistique soit possible, il est nécessaire d'introduire des hypothèses supplémentaires sur la F.A. $Y(x)$, de manière à réduire le nombre des "paramètres" dont dépend sa loi. Tel est le but de l'hypothèse stationnaire que nous allons définir : une F.A. stationnaire se répète en quelque sorte elle-même dans l'espace, et cette répétition rend à nouveau possible l'inférence statistique à partir d'une réalisation unique. Précisons donc cette hypothèse (que nous affaiblirons considérablement par la suite).

F.A. stationnaire. - On dit qu'une F.A. $Y(x)$ est stationnaire si sa loi est invariante par translation : autrement dit, si x_1, x_2, \dots, x_k sont k points d'appui arbitraires (k entier quelconque) et si h est un vecteur quelconque, les k V.A. $Y(x_1), Y(x_2), \dots, Y(x_k)$ ont la même loi (à k variables) que les k V.A. $Y(x_1+h), Y(x_2+h), \dots, Y(x_k+h)$. Dans ce qui suit, $Y(x)$ désignera une F.A. stationnaire.

Espérance mathématique. - Considérons un point d'appui x_0 . Si la V.A. ordinaire $Y(x_0)$ admet une espérance mathématique, celle-ci est une fonction $m(x_0) = E[Y(x_0)]$ du point d'appui x_0 . Mais $Y(x)$ est stationnaire, et on a par suite $m(x_0+h) = m(x_0)$ pour tout vecteur h , et $m(x_0)$ est donc une constante m indépendante de x_0 :

$$m = E[Y(x)]$$

Quitte à remplacer $Y(x)$ par $Y(x) - m$, nous supposerons souvent $m = 0$, (pourvu toutefois que cette espérance existe).

La covariance $K(h)$. - Considérons maintenant deux points d'appui x_0 et x_0+h . Si les deux V.A. $Y(x_0)$ et $Y(x_0+h)$ admettent des variances finies (donc aussi une espérance m que nous supposerons nulle), elles admettent aussi une covariance $K(x_0 ; h)$, qui dépend en principe du point d'appui x_0 et du vecteur h . Mais, $Y(x)$ étant stationnaire, on a $K(x_0+a ; h) = K(x_0 ; h)$ pour tout vecteur a :

$K(x_0 ; h)$ ne dépend donc pas de x_0 , et nous écrirons simplement $K(h)$:

$$(2-1) \quad K(h) = E[Y(x) Y(x+h)]$$

On comparera cette définition de la covariance avec celle du covariogramme transitif : $K(h)$ est la transposition probabiliste de $g(h)$. Pour $h = 0$, on a $K(0) = E \left([Y(x)]^2 \right)$: c'est la variance de la V.A. $Y(x_0)$. Pour que la F.A. stationnaire $Y(x)$ admette une fonction de covariance $K(h)$, il faut et il suffit qu'elle admette une variance finie $K(0)$.

Hypothèse stationnaire d'ordre 2. - Nous dirons qu'une F.A. $Y(x)$ est stationnaire d'ordre 2 si la V.A. $Y(x_0)$ admet une espérance m indépendante du point d'appui x_0 , et si pour tout vecteur h la covariance :

$$K(h) = E[Y(x_0 + h) Y(x_0)] - m^2$$

existe et ne dépend pas de x_0 . Cette hypothèse (qui n'entraîne pas la stationnarité au sens strict, telle que nous l'avons définie ci-dessus) est suffisante pour la théorie des V.R. Mais elle suppose l'existence d'une variance a priori finie $K(0)$.

Variance a priori infinie. - Or d'assez nombreux phénomènes présentent une capacité de dispersion illimitée, et ne peuvent être décrits correctement si on leur attribue une variance a priori finie : cette affirmation surprendra peut-être, mais il faut bien voir que la nature nous tend ici une sorte de piège. Lorsque l'on prélève des échantillons v dans un champ V , on obtient un histogramme à partir duquel on peut toujours calculer numériquement une variance, qui prend ainsi une valeur parfaitement définie. Mais cette variance expérimentale est en réalité une fonction $\sigma^2(v|V)$ du support v et du champ V . Elle augmente, en particulier, lorsque le champ V augmente. Si les échantillons de taille v possèdent une variance a priori finie, celle-ci doit apparaître comme la limite pour V infini de la variance expérimentale $\sigma^2(v|V)$.

C'est ainsi que les auteurs d'Afrique du Sud (D.G. Krige, etc...), à partir de centaines de milliers d'échantillons prélevés dans le grand gisement d'or du Rand, ont pu calculer la variance de ces échantillons dans des panneaux de plus en plus grands, puis dans une concession entière, puis dans le gisement du Rand dans son ensemble : ils ont ainsi obtenu une relation expérimentale de la forme :

$$\sigma^2(v|V) = \alpha \log (V/v)$$

La croissance de la variance se poursuit sans défaillance selon cette loi logarithmique (formule de De Wijs) jusqu'au dernier point expérimental, pour lequel V/v est de l'ordre de la dizaine de milliards. On peut conclure en toute certitude qu'il n'existe pas ici de variance a priori finie.

On est donc conduit à remplacer l'hypothèse stationnaire d'ordre 2 par une hypothèse plus faible mais de signification analogue :

Hypothèse intrinsèque. - Même lorsque la variance a priori $K(o)$ n'existe pas (est infinie), il peut arriver que les accroissements $Y(x_0+h) - Y(x_0)$ aient une variance finie. Nous dirons donc que la F.A. $Y(x)$ vérifie l'hypothèse intrinsèque si, pour tout vecteur h , l'accroissement $Y(x_0+h) - Y(x_0)$ admet une espérance et une variance indépendantes du point d'appui x (mais dépendant de h), soit :

$$E[Y(x+h) - Y(x)] = m(h)$$

$$D^2[(Y(x+h) - Y(x))^2] = 2 \gamma(h)$$

La fonction $m(h)$ est la dérive linéaire. Pour montrer qu'elle est linéaire en h , on part de la relation évidente :

$$Y(x+h'' + h') - Y(x) = [Y(x+h'' + h') - Y(x+h')] + [Y(x+h') - Y(x)]$$

et on passe aux espérances, d'où : $m(h'+h'') = m(h') + m(h'')$. On peut toujours supposer que cette dérive linéaire $m(h)$ est nulle, quitte à remplacer $Y(x)$ par $Y(x) - m(x)$.

La fonction $\gamma(h)$:

$$(2-2) \quad \gamma(h) = \frac{1}{2} D^2[(Y(x+h) - Y(x))^2]$$

s'appelle le demi-variogramme, ou fonction intrinsèque. Une F.A. vérifiant l'hypothèse intrinsèque constitue ce qu'on appelle un schéma intrinsèque, caractérisé par son demi-variogramme.

Remarque. Si $Y(x)$ vérifie l'hypothèse stationnaire d'ordre 2, elle vérifie aussi l'hypothèse intrinsèque, et l'on a dans ce cas :

$$(2-3) \quad \gamma(h) = K(o) - K(h)$$

Comme $|K(h)| \leq K(0)$, on a $\gamma(h) \leq 2 K(0)$, de sorte que le demi-variogramme d'une F.A. stationnaire d'ordre 2 est nécessairement borné. Il existe des schémas intrinsèques d'utilisation très courante dont le $\gamma(h)$ n'est pas borné, et qui par conséquent ne peuvent pas vérifier l'hypothèse stationnaire d'ordre 2 (ont une variance a priori infinie). EX. schéma de De Wijs ($\gamma(h) = 3 \alpha \log |h|$), schéma linéaire ($\gamma(h) = A |h|$).

2-2 PROPRIETES DE LA COVARIANCE ET DU VARIOGRAMME

Une covariance ou un demi-variogramme sont des fonctions symétriques :

$$K(h) = K(-h) \quad , \quad \gamma(h) = \gamma(-h)$$

Une covariance vérifie de plus l'inégalité de Schwarz :

$$|K(h)| \leq K(0)$$

Pour un variogramme, on trouve seulement $\gamma(h) \geq 0$ et $\gamma(0) = 0$. Cependant, il est possible de montrer que la croissance à l'infini d'un variogramme est nécessairement moins rapide que celle de $|h|^2$; soit :

$$\lim_{|h| \rightarrow \infty} \frac{1}{|h|^2} \gamma(h) = 0$$

Ces conditions sont nécessaires, mais il ne suffit pas qu'une fonction K ou γ les vérifie pour qu'il existe une F.A. stationnaire ou intrinsèque admettant cette covariance ou ce demi-variogramme. En fait, il faut et il suffit que K appartienne à la classe des fonctions "de type positif" et $-\gamma$ à celle des fonctions de type positif conditionnel. Par exemple, les fonctions $\log r$ et r^λ avec $\lambda < 2$ peuvent servir de demi-variogramme, mais non r^λ pour $\lambda \geq 2$. Ces conditions expriment entre autres, que les formules que nous établirons ci-dessous pour les variances d'extension ou d'estimation conduisent nécessairement à des valeurs positives (ce qui ne serait pas toujours le cas si l'on prenait une fonction quelconque comme demi-variogramme).

Ces conditions méritent d'être examinées de plus près, ne serait-ce que pour préciser la classe des combinaisons linéaires que l'on a le droit d'utiliser lorsqu'il n'existe pas de variance a priori finie.

2-2-1 Les combinaisons linéaires autorisées.

Plaçons-nous d'abord dans le cas stationnaire d'ordre 2, prenons $m = 0$ et donnons-nous N points x_1, x_2, \dots, x_N de \mathbb{R}^n et N coefficients arbitraires $\lambda_1, \dots, \lambda_N$. L'expression

$$(2-4) \quad Y = \sum_1 \lambda_1 Y(x_1)$$

est une variable aléatoire (en tant que combinaison linéaire finie de V.A.). Comme il existe, par hypothèse, une fonction de covariance $K(h)$, Y admet une variance finie, et un calcul élémentaire donne :

$$(2-5) \quad D^2(Y) = \sum_{i,j} \lambda_i \lambda_j K(x_i - x_j)$$

La condition évidente $D^2(Y) \geq 0$ exprime (par. 1-3) que toute fonction de covariance est de type positif. Inversement, si $K(h)$ est une fonction de type positif, on montre qu'il est possible de construire une F.A. dont la covariance est, précisément, $K(h)$. Cette formule (2-5) réapparaîtra, sous des formes légèrement différentes, à tous les étages de la théorie. Notons, pour l'instant, que dans le cas stationnaire d'ordre 2 toutes les combinaisons linéaires finies ont une variance finie, et que leur usage est donc toujours légitime.

Il n'en est pas de même lorsqu'il existe seulement un variogramme, mais pas de covariance. Dans ce cas, en effet, seules les combinaisons linéaires du type (2-4) vérifiant la condition :

$$(2-6) \quad \sum_1 \lambda_1 = 0$$

ont une variance finie. Ainsi, dans le cas d'une F.A.I. sans covariance, seul sera autorisé l'usage des combinaisons linéaires pour lesquelles la somme des coefficients sera égale à 0. Cette remarque joue un rôle capital dans toute la théorie intrinsèque. Montrons que cette condition (2-6) est suffisante. Si elle est vérifiée, on peut, en effet, écrire :

$$\sum_1 \lambda_1 Y(x_1) = \sum_1 \lambda_1 [Y(x_1) - Y(o)]$$

Or la F.A. (non stationnaire) définie par :

$$Z(x) = Y(x) - Y(o)$$

admet une fonction de covariance $C(x,y) = E(Z(x) Z(y))$, puisque $Z(x)$, en tant qu'accroissement de la F.A.I. $Y(x)$, admet une variance finie. Calculons cette fonction de covariance (non stationnaire) $C(x,y)$ dont l'expression nous sera utile ultérieurement. Le plus simple, pour cela, est de partir de la définition du variogramme :

$$2 \gamma(x-y) = E[Y_x - Y_0 - Y_y + Y_0]^2 = 2 \gamma(x) + 2 \gamma(y) - 2 C(x,y)$$

D'où :

$$(2-7) \quad C(x,y) = \gamma(x) + \gamma(y) - \gamma(x-y)$$

Lorsque (2-6) est vérifiée, la combinaison linéaire $\sum \lambda_i Y_{x_i} = \sum \lambda_i (Y_{x_i} - Y_0)$ admet la variance finie :

$$D^2(Y) = \sum_{i,j} \lambda_i \lambda_j C(x_i, x_j)$$

Soit, d'après (2-7) :

$$D^2(Y) = \sum_i \lambda_i \sum_j \lambda_j \gamma(x_j) + \sum_j \lambda_j \sum_i \lambda_i \gamma(x_i) - \sum_{i,j} \lambda_i \lambda_j \gamma(x_i - x_j)$$

et, compte tenu de (2-6) :

$$(2-8) \quad D^2(Y) = - \sum_{i,j} \lambda_i \lambda_j \gamma(x_i - x_j)$$

D'où un second résultat qui sera très utile dans ce qui suit : Dès que la somme de coefficients d'une combinaison linéaire est égale à 0, on a le droit de calculer sa variance selon la formule (2-5) comme s'il existait une covariance $K(h)$, à condition de remplacer cette covariance par $-\gamma$. C'est là un mécanisme de calcul très général, qui apporte souvent de grandes simplifications.

Inversement, s'il n'y a pas de variance a priori finie, une combinaison linéaire de variance finie est nécessairement une combinaison linéaire d'accroissements de la F.A. $Y(x)$, puisque seuls, par hypothèse, ces accroissements ont des variances finies. Par suite cette combinaison vérifie la condition (2-6) qui est bien nécessaire et suffisante, comme nous l'avions annoncé.

On peut alors introduire de manière naturelle la condition suivant laquelle $-\gamma$ doit être de type positif conditionnel. Par définition, nous dirons qu'une fonction $g(h)$ est de type positif conditionnel si pour tout entier N , tout système de N points x_1, x_2, \dots, x_N de l'espace \mathbb{R}^n et pour tout système de coefficients $\lambda_1, \lambda_2, \dots, \lambda_N$ vérifiant la condition $\sum \lambda_i = 0$, on a :

$$\sum_{i,j} \lambda_i \lambda_j g(x_i - x_j) \geq 0$$

Comme les combinaisons linéaires autorisées ont une variance positive ou nulle, la relation (2-8) signifie bien que $-\gamma$ est de type positif conditionnel. Inversement, si $-\gamma$ est de type positif conditionnel, on peut montrer qu'il est effectivement possible de construire une F.A.I. admettant γ comme demi-variogramme. Il s'agit donc d'une condition nécessaire et suffisante.

2-2-2 Continuité m.q. et autres propriétés.

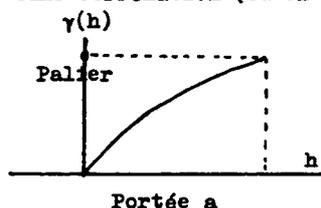
Voici quelques indications complémentaires sur les propriétés d'une covariance ou d'un variogramme. Comme dans le cas des méthodes transitives, l'intérêt va se concentrer sur le comportement au voisinage de l'origine.

a/ On peut remarquer que le variogramme donne un contenu précis à la notion traditionnelle de zone d'influence d'un échantillon : sa croissance plus ou moins rapide reflète, en effet, la manière plus ou moins rapide dont se détériore l'influence d'un échantillon sur des zones de plus en plus lointaines du gisement.

b/ Les anisotropies se manifestent par le comportement différent du variogramme dans les différentes directions de l'espace. En l'absence d'anisotropie, $\gamma(h) = \gamma(r)$ ne dépend que du module r de h , et non de la direction de ce vecteur. On dit qu'il y a anisotropie géométrique lorsqu'une simple transformation linéaire des coordonnées suffit à rétablir l'isotropie.

Il y a des types plus complexes d'anisotropies. Par exemple, dans l'espace à trois dimensions, il peut arriver que $Y(x)$ ne dépende que de la troisième coordonnée x , et reste par suite constante dans les plans parallèles aux deux premiers axes de coordonnées. Le demi-variogramme $\gamma(h) = \gamma(h_3)$ ne dépend que de la troisième composante de h . Le plus souvent, $Y(x)$ ne sera pas réellement constante dans les plans horizontaux, mais y variera moins vite ou plus régulièrement que dans la direction verticale. On prendra alors un demi-variogramme de la forme $\gamma(h) = \gamma_0(h_1, h_2, h_3) + \gamma_1(h_3)$ (anisotropie zonale).

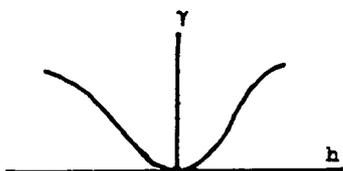
c/ Portée. Dans le cas stationnaire d'ordre 2, la portée $a(\alpha)$ dans une direction α est la valeur de la distance au-delà de laquelle, dans cette direction, $Y(x)$ et $Y(x+h)$ sont sans corrélation (ou en corrélation négligeable) : $K(h) = 0$ (ou $\neq 0$) pour $|h| \geq a(\alpha)$.



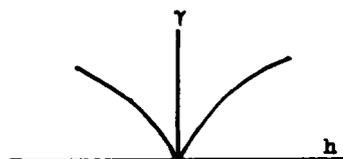
D'après c/ , $K(h) = 0$ équivaut à $\gamma(h) = K(0) = \gamma(\infty)$:

la portée est aussi la distance à partir de laquelle le demi-variogramme atteint sa valeur limite $\gamma(\infty)$ ou palier. Ainsi, une F.A. intrinsèque dont le variogramme n'est pas borné ne peut pas avoir de portée finie.

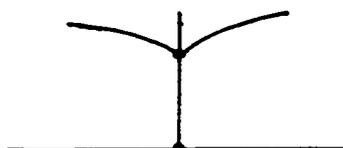
d/ Comportement au voisinage de l'origine. La continuité et la régularité dans l'espace de la F.A. $Y(x)$ s'expriment dans le comportement du $\gamma(h)$ au voisinage de l'origine. Par ordre de régularité décroissante, on peut distinguer quatre types :



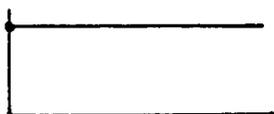
Comportement parabolique : $\gamma(h)$ est deux fois dérivable en $h = 0$, $Y(x)$ est alors elle-même dérivable (en moyenne quadratique), donc présente un haut degré de régularité dans l'espace.



Comportement linéaire (tangente oblique à l'origine) : $\gamma(h)$ est continu en $h = 0$, mais non dérivable, $Y(x)$ est continue en moyenne quadratique, mais non dérivable, donc déjà moins régulière.



Effet de pépité : $\gamma(h)$ ne tend pas vers 0 lorsque h tend vers 0 (discontinuité à l'origine). $Y(x)$ n'est même pas continue en moyenne quadratique, donc extraordinairement irrégulière.



Cas limite complètement aléatoire : $Y(x)$ et $Y(x')$ sont indépendantes pour deux points distincts quelconques, si rapprochés soient-ils (bruit blanc des physiciens, ou effet de pépité à l'état pur).

Partie irrégulière. Dans le cas isotrope ($\gamma(h) = \gamma(r)$) et en l'absence d'effet de pépité, on caractérisera le comportement du $\gamma(h)$ par un développement limité de la forme :

$$\gamma(r) = \sum a_{2n} r^{2n} + \sum c_{\lambda} r^{\lambda} + \sum c_{2n} r^{2n} \log r$$

Exactement comme pour les covariogrammes transitifs, on distinguera une partie régulière (termes

de degré entier pair), et une partie irrégulière (termes en r^λ avec λ différent d'un entier pair, et aussi termes logarithmiques du type $r^{2n} \log r$). En l'absence de partie irrégulière, la F.A. serait indéfiniment dérivable, donc parfaitement régulière. C'est donc la partie irrégulière seule qui représente le degré d'irrégularité de la F.A., et dans cette partie irrégulière, c'est le terme de plus bas degré qui joue le rôle principal : on peut définir le degré de régularité de la F.A. comme le degré λ du terme irrégulier principal. Donnons quelques indications mathématiques complémentaires (mais sans entrer dans les détails, puisqu'il s'agit là de résultats classiques).

e/ Continuité et dérivation en moyenne quadratique. On dit que la F.A. $Y(x)$ est continue en moyenne quadratique (continue m.q.) si l'on a :

$$E \left([Y(x+h) - Y(x)]^2 \right) \rightarrow 0 \quad \text{pour } |h| \rightarrow 0$$

Il en est ainsi (par définition) si et seulement si $\gamma(h)$ est continu en $h = 0$; c'est-à-dire en l'absence d'effet de pépite.

Dans l'espace à une dimension, on dit que la F.A. $Y'(x)$ est la dérivée en moyenne quadratique (dérivée m.q.) de la F.A. $Y(x)$ si :

$$E \left(\left[\frac{Y(x+h) - Y(x)}{h} - Y'(x) \right]^2 \right) \rightarrow 0 \quad \text{pour } |h| \rightarrow 0$$

On a des définitions analogues dans l'espace à $n \neq 1$ dimensions.

On démontre qu'une F.A. intrinsèque $Y(x)$ admet une dérivée m.q. $Y'(x)$ si et seulement si $\gamma(h)$ est deux fois dérivable en $h = 0$. La dérivée seconde $\gamma''(h)$ existe alors pour tout h , $Y'(x)$ est stationnaire d'ordre 2 (même si $Y(x)$ est seulement intrinsèque) et admet comme covariance la fonction $\gamma''(h)$. De même, $Y(x)$ est n fois dérivable m.q. si et seulement si la dérivée $\gamma^{(2n)}$ existe en $h = 0$ (elle existe alors pour tout h).

Si λ est le degré du terme irrégulier principal de $\gamma(h)$, $Y(x)$ admet donc une dérivée m.q. d'ordre n si et seulement si $\lambda > 2n$ (si ce terme irrégulier principal est $r^{2n} \log r$, $Y(x)$ est $n-1$ fois m.q. dérivable, mais non n fois).

2-3 REGULARISATION D'UNE F.A. INTRINSEQUE

Pour simplifier l'exposé, nous ferons les raisonnements dans le cas d'une F.A. stationnaire d'ordre 2 admettant une fonction de covariance $K(h)$, mais tous les résultats que nous exprimerons à l'aide de la fonction intrinsèque $\gamma(h) = K(0) - K(h)$ resteront valables dans le cas d'une F.A. intrinsèque (donc, même si $\gamma(h)$ n'est pas borné, et si par suite la covariance $K(h)$ n'existe pas) : cela résultera très simplement du mécanisme de calcul que nous avons vu dans le paragraphe 2-2-1. Ici encore, nous n'entrerons pas dans les détails, puisqu'il s'agit d'un simple rappel de résultats classiques.

$$\text{2-3-1 Intégrale Stochastique} \quad I = \int_{\mathcal{V}} Y(x) p(x) dx$$

On définit l'intégrale I comme la limite en moyenne quadratique (si elle existe) des sommes discrètes :

$$I_n = \sum_{i=1}^n Y(x_i) p(x_i) \Delta x_i$$

(les Δx_i sont des petits éléments de volume disjoints dont la réunion est \mathcal{V} , et x_i est un point de Δx_i). L'intégrale stochastique I est donc une variable aléatoire, comme les I_n .

On démontre que cette V.A. existe si et seulement si l'intégrale

$$(2-9) \quad D^2(I) = \int_{\mathcal{V}} p(x) dx \int_{\mathcal{V}} K(x-y) p(y) dy$$

est finie, et $D^2(I)$ est alors la variance (finie) de cette V.A. On retrouve facilement (2-9) par le calcul (non rigoureux) suivant :

$$\begin{aligned} I^2 &= \int_{\mathcal{V}} Y(x) p(x) dx \int_{\mathcal{V}} Y(y) p(y) dy \\ E(I^2) &= \int_{\mathcal{V}} p(x) dx \int_{\mathcal{V}} E[Y(x) Y(y)] p(y) dy = \\ &= \int_{\mathcal{V}} p(x) dx \int_{\mathcal{V}} K(x-y) p(y) dy \end{aligned}$$

(le calcul n'est pas rigoureux, car il n'est pas tout-à-fait évident au départ que l'on a le droit d'intervertir les symboles E et \int : en fait, on montre que cette interversion est bien légitime).

2-3-2 Convolution Stochastique.

Le produit de convolution $Y * f$ de la F.A. $Y(x)$ par une fonction ordinaire $f(x)$ est l'intégrale stochastique (si elle existe) :

$$\int Y(x-x') f(x') dx'$$

On peut ainsi définir la régularisée $Y_p = Y * \check{p}$ de la F.A. $Y(x)$ par une fonction de pondération $p(x)$. C'est la "moyenne mobile pondérée" :

$$Y_p(x) = \int Y(x+x') p(x') dx'$$

Attention : la régularisée $Y(x)$ est encore une fonction aléatoire (plus régulière que $Y(x)$, mais toujours aléatoire) : il ne suffit pas de lisser ou régulariser une F.A. par un procédé de moyenne mobile pour faire disparaître comme par magie le caractère aléatoire de cette fonction.

La variance de la régularisée Y_p est donnée par la formule (2-9) ci-dessus, et $Y_p(x)$ existe au sens de l'intégration m.q. si et seulement si $D^2(I) < \infty$. Calculons la covariance

$$E[Y(x_0) Y(x_0+h)] \text{ de } Y(x_0) \text{ et } Y(x_0+h)$$

On a :

$$Y_p(x_0) Y_p(x_0+h) = \iint p(x') p(x'') Y(x_0+x') Y(x_0+h+x'') dx' dx''$$

Passons aux espérances, en intervertissant E et \int . Il vient :

$$E[Y_p(x_0) Y_p(x_0+h)] = \iint K(h+x''-x') p(x') p(x'') dx' dx''$$

Cette covariance ne dépend pas de x_0 , mais seulement de h . Donc, la régularisée Y est stationnaire d'ordre 2, et admet la covariance :

$$(2-10) \quad K_p(h) = \int p(x) dx \int K(h+x-y) p(y) dy$$

Cette formule généralise (2-9). Pour la présenter sous forme plus synthétique, faisons le change-

ment de variable $x = y + z$. Il vient :

$$K_p(h) = \int K(x+z) dz \int p(y+z) p(y) dy$$

Soit P le covariogramme transitif ($P = p * \check{p}$) de la fonction de pondération p . On a par définition $P(z) = \int p(y+z) p(y) dy$, et par suite :

$$K_p(h) = \int K(h+z) P(z) dz$$

c'est-à-dire, sous forme de produit de convolution (car $P = \check{P}$) :

$$(2-11) \quad K = K * P$$

On obtient la covariance K_p de la régularisée Y_p en régularisant la covariance $K(h)$ de Y par le covariogramme transitif de la fonction de pondération. On comparera ce résultat avec celui que l'on a obtenu dans le cas des méthodes transitives.

Le demi-variogramme γ_p de la régularisée Y_p est $K_p(0) - K_p(h)$. En remplaçant $K(h)$ par $K(0) - \gamma(h)$ dans (2-9) et (2-10), on constate que $K(0)$ s'élimine, et il reste :

$$(2-12) \quad \gamma_p(h) = \int \gamma(h+z) P(z) dz - \int \gamma(z) P(z) dz$$

On peut encore écrire :

$$\gamma_p = \gamma * P - A$$

avec une constante A déterminée par la condition $\gamma_p(0) = 0$. Ces relations restent vraies pour toute F.A. intrinsèque (même si la covariance n'existe pas) (voir remarque ci-dessus).

2-3-3 La montée droite sous puissance constante.

La montée constitue un cas particulier de la régularisation. Dans les méthodes transitives, nous pouvions intégrer la V.R. de $-\infty$ à $+\infty$ par rapport à l'une des coordonnées, par exemple x_3 , parce que cette V.R. était nulle à l'extérieur d'un champ borné. Ici, nous ne pouvons

plus intégrer que sur une longueur finie ℓ . Nous appellerons donc montée sous puissance constante ℓ l'opération qui nous fait passer de la F.A. $Y_3(x_1, x_2, x_3)$ définie dans l'espace à trois dimensions à la F.A. $Y_2(x_1, x_2)$ définie dans l'espace à deux dimensions par :

$$Y_2(x_1, x_2) = \frac{1}{\ell} \int_0^{\ell} Y(x_1, x_2, x_3) dx_3$$

Si $Y_3(x)$ est la teneur ponctuelle dans une formation stratiforme de puissance ℓ , $Y_2(x_1, x_2)$ est la teneur moyenne du sondage implanté au point (x_1, x_2) de la surface topographique.

Il n'est pas très difficile d'établir la formule donnant le variogramme $\gamma_2(h_1, h_2)$ de la F.A. $Y_2(x_1, x_2)$ - qui est manifestement intrinsèque comme $Y_3(x)$ - en fonction du variogramme γ_3 de Y_3 (le lecteur pourra le faire à titre d'exercice). Dans le cas particulier où $\gamma_3 = \gamma_3(r)$ est isotrope, on obtient l'algorithme suivant (de la montée droite sous puissance constante dans le cas isotrope) :

$$\gamma_{n-1}(r) = \frac{2}{\ell^2} \int_0^{\ell} (\ell-x) \gamma_n(\sqrt{r^2+x^2}) dx - \frac{2}{\ell^2} \int_0^{\ell} (\ell-x) \gamma_n(x) dx$$

qui est du reste un simple cas particulier de (2-12).

Comme dans le cas des méthodes transitives, il existe ici encore une règle de correspondance terme à terme entre les parties irrégulières de $\gamma_n(r)$ et $\gamma_{n-1}(r)$. Ces règles, un peu moins simples que dans le cas transitif, s'écrivent :

$$(2-13) \quad \begin{cases} r^\lambda - A_\lambda \frac{r^{1+\lambda}}{\ell} + A'_\lambda \frac{r^{2+\lambda}}{\ell^2} \\ r^{2n} \log r - A_{2n} \frac{r^{2n+1}}{\ell} + A'_{2n} \frac{r^{2n+2}}{\ell^2} \log r \\ r^{2n+1} - A_{2n+1} \frac{r^{2n+2}}{\ell} \log r + A'_{2n+1} \frac{r^{2n+3}}{\ell^2} \end{cases}$$

Les coefficients A_λ , A_{2n} , A_{2n+1} sont les mêmes que ceux qui apparaissent dans les méthodes transitives (formules (1-10 à 1-13)), et c'est là une première manifestation de la parenté profonde des deux aspects (transitifs et probabilistes) de la théorie des V.R. Notons, toutefois, que les règles (2-13) sont moins simples que les règles (1-10). Au lieu de $r^{1+\lambda}$, nous trouvons $\frac{r^{1+\lambda}}{\ell}$, mais

ce coefficient $1/\ell$ s'introduit pour des raisons dimensionnelles évidentes (en intrinsèque, on raisonne sur des teneurs moyennes et non plus sur des accumulations, comme en transitif). Mais surtout, à côté du terme principal $A_\lambda \frac{r^{1+\lambda}}{\ell}$ - qui est l'exact équivalent du terme transitif - nous voyons apparaître un terme complémentaire $A'_\lambda \frac{r^{2+\lambda}}{\ell^2}$, avec un coefficient A'_λ dont il n'est pas utile ici de donner l'expression explicite (cf. [4]). Ce terme supplémentaire provient du caractère fini (montée sous puissance ℓ finie) du champ d'intégration, ou, si l'on préfère, de la présence du terme en $x \gamma(\sqrt{r^2+x^2})$ sous le signe d'intégration dans l'algorithme de la montée droite. Toutefois, ce terme supplémentaire est en $r^{2+\lambda}$, donc d'ordre plus élevé que le terme principal, et la partie principale du $\gamma_{n-1}(r)$ au voisinage de $r = 0$ est donc identique à celle que fournirait la règle transitive.

Notons aussi que le domaine de validité d'un développement de $\gamma_{n-1}(r)$ obtenu par la règle (2-13) est nécessairement intérieur à la boule de rayon ℓ , Aux grandes distances, on déduit facilement de l'algorithme de la montée droite :

$$\gamma_{n-1}(r) \doteq \gamma_n(r) - \frac{2}{\ell^2} \int_0^\ell (\ell-x) \gamma_n(x) \quad (r \gg \ell)$$

de sorte qu'à une constante près le variogramme monté est identique au variogramme initial.

2-4 VARIANCES D'EXTENSION OU D'ESTIMATION

2-4-1 Variance d'extension.

Définissons d'abord la notion capitale de variance d'extension. Soit $Y(x)$ une F.A. que nous supposons pour l'instant stationnaire d'ordre 2, et soit $K(h)$ sa covariance. Désignons par $Z(v)$ et $Z(v')$ les "tenseurs moyennes" de deux domaines v et v' de l'espace à n dimensions, c'est-à-dire les intégrales stochastiques :

$$Z(v) = \frac{1}{V} \int_v Y(x) dx \quad ; \quad Z(v') = \frac{1}{V'} \int_{v'} Y(x) dx$$

La formule (2-9) montre que, si v et v' sont bornés, $Z(v)$ et $Z(v')$ ont des variances finies. La première, par exemple, a pour expression :

$$\sigma^2(v) = \frac{1}{v^2} \int_v dx \int_v K(x-y) dy$$

Calculons de même la covariance $\sigma(v, v')$ de $Z(v)$ et $Z(v')$. De :

$$Z(v) Z(v') = \frac{1}{vv'} \int_v Y(x) dx \int_{v'} Y(y) dy$$

on déduit, en passant aux espérances mathématiques :

$$\sigma(v, v') = \frac{1}{vv'} \int_v dx \int_{v'} E[Y(x) Y(y)] dy = \frac{1}{vv'} \int_v dx \int_{v'} K(x-y) dy$$

Nous appellerons variance d'extension de v à v' (ou de v' à v) la variance de l'erreur $Z(v') - Z(v)$ que l'on commet en attribuant à v' la teneur moyenne $Z(v)$ de v . Cette variance d'extension σ_E^2 est égale à :

$$\sigma_E^2 = \sigma^2(v) + \sigma^2(v') - 2 \sigma(v, v')$$

Compte tenu des valeurs calculées plus haut, on trouve donc :

$$\begin{aligned} \sigma_E^2 &= \frac{1}{v^2} \int_v dx \int_v K(x-y) dy + \frac{1}{v'^2} \int_{v'} dx \int_{v'} K(x-y) dy \\ &\quad - \frac{2}{vv'} \int_v dx \int_{v'} K(x-y) dy \end{aligned}$$

Remplaçons $K(h)$ par $K(0) - \gamma(h)$: on constate que la constante $K(0)$ disparaît de l'expression de σ_E^2 , et il vient la formule fondamentale :

$$(2-14) \quad \sigma_E^2 = \frac{2}{vv'} \int_v dx \int_{v'} \gamma(x-y) dy - \frac{1}{v^2} \int_v dx \int_v \gamma(x-y) dy - \frac{1}{v'^2} \int_{v'} dx \int_{v'} \gamma(x-y) dy$$

On peut montrer que (2-14) reste valable pour toute F.A. intrinsèque, donc même si la covariance $K(h)$ n'existe pas (cela résulte du mécanisme de calcul du paragraphe 2-2-1).

2-4-2 Variance d'estimation.

Supposons maintenant qu'au lieu de connaître la "teneur moyenne" $Z(v')$ de $Y(x)$ dans un volume v' , nous connaissions la teneur moyenne :

$$Z' = \frac{1}{N} \sum_{i=1}^N Y(x_i)$$

de N prélèvements effectués aux N points x_i . Z' est une variable aléatoire, dont les caractéristiques se déduisent sans peine de $K(h)$ ou de $\gamma(h)$. Nous appellerons variance d'estimation σ_N^2 (de v par les N prélèvements effectués aux N points x_i) la variance de la différence $Z(v) - Z'$. Pour obtenir l'expression de σ_N^2 , on doit, à chacune des étapes du raisonnement qui nous a conduit à (2-14), remplacer les intégrales étendues à v' par des sommes discrètes étendues aux N points x_i . On trouve ainsi la seconde formule fondamentale :

$$(2-15) \quad \sigma_N^2 = \frac{2}{Nv} \sum_i \int_v \gamma(x_i - x) dx - \frac{1}{v^2} \int_v dx \int_v \gamma(x-y) dy - \frac{1}{N^2} \sum_i \sum_j \gamma(x_i - x_j)$$

Cette formule, où l'on voit alterner des expressions exactes et approchées des mêmes intégrales, présente une structure remarquable; analogue, quoique plus complexe, à celle de la formule (1-14) des méthodes transitives. On note, en particulier, que la variance d'estimation est d'autant plus faible :

- que le réseau des prélèvements x_i est plus serré et plus représentatif de la géométrie du volume v que l'on veut estimer,

- que la fonction $\gamma(h)$ est plus régulière, donc que la F.A. $Y(x)$ est elle-même plus continue dans sa variation spatiale.

En pratique, cependant, si N est grand, la formule (2-15) conduirait à des calculs assez longs. Nous donnerons plus loin des formules d'approximation beaucoup plus simples, qui permettront aussi une comparaison méthodologiquement très instructive avec les formules analogues des méthodes transitives.

Remarque - Il n'y a aucune différence conceptuelle entre les notions de variances d'extension et d'estimation : la formule (2-15) est un cas particulier de (2-14), v' étant remplacé par la réunion des n points x_i . L'usage a réservé le terme de variance d'extension à l'extension d'un échantillon unique dans sa "zone d'influence", et celui de variance d'estimation à l'extension d'un plus grand

nombre d'échantillons dans le gisement entier, ou dans un grand panneau.

2-4-3 Variance de v dans V .

La notion de variance $\sigma^2(v|V)$ d'un échantillon v dans un domaine V semble, de prime abord, expérimentalement évidente. Elle n'a cependant de sens précis que lorsque le domaine V apparaît comme la réunion $V = \cup v_i$ de volumes v_i disjoints tous égaux à v et translattés les uns des autres. Lorsque l'on a prélevé n échantillons v_i dans un champ V' quelconque et que l'on connaît leurs teneurs Y_i , ce que l'on calcule numériquement :

$$\bar{Y} = \frac{1}{n} \sum_1 Y_i$$

$$s^2 = \frac{1}{n} \sum_1 (Y_i - \bar{Y})^2$$

constitue une estimation de la variance de v dans $V = \cup v_i$ (réunion des n volumes v_i des prélèvements réels), et non dans le domaine V' comme on serait tenté de le croire. C'est cette démarche de l'expérimentateur qui inspire les définitions qui suivent.

a/ Examinons d'abord le cas où les échantillons v_i sont ponctuels (et où la relation $v = \cup v_i$ est évidemment vérifiée, puisque V est bien la réunion des ensembles ponctuels qu'il contient), et définissons la variance $\sigma^2(0|V)$ des échantillons ponctuels dans V . Désignons par :

$$Z(V) = \frac{1}{V} \int_V Y(x) dx$$

la teneur moyenne de V . Lorsque la réalisation $Y(x)$ est fixée, $Z(V)$ est une intégrale ordinaire, et la variance expérimentale des teneurs des échantillons ponctuels est :

$$(2-16) \quad s^2(0|V) = \frac{1}{V} \int_V [Y(x) - Z(V)]^2 dx$$

On peut aussi interpréter cette relation (2-16) en disant que $s^2(0|V)$ est la variance conditionnelle (à réalisation fixée) de la variable aléatoire $Y(x)$ obtenue en implantant un point x dans V au hasard avec une densité de probabilité uniforme. Nous appellerons variance $\sigma^2(v|V)$ de v dans V la variance de la V.A. obtenue en déconditionnant la variable aléatoire $Y(x)$ vis-a-vis de la réalisation, c'est-à-dire la variance de la V.A. $Y = Y(x)$, où $Y(x)$ est cette fois la F.A. (et non plus sa

réalisation) et x un point aléatoire dans V . A réalisation fixée, l'espérance conditionnelle de $Y(x)$ est évidemment $Z(V)$. On obtient donc :

$$\sigma^2(O|V) = E(s^2(O|V)) = \frac{1}{V} \int_V E[Y(x)-Z(V)]^2 dx$$

Or $E[Y(x)-Z(V)]^2$ est également la variance d'extension dans V de l'échantillon ponctuel implanté en x . Ainsi, la variance $\sigma^2(O|V)$ est la valeur moyenne pour $x \in V$ de la variance d'extension dans V de l'échantillon ponctuel implanté en x . On déduit alors facilement de (2-14) l'expression de cette variance :

$$(2-17) \quad \sigma^2(O|V) = \frac{1}{V^2} \int_V dx \int_V \gamma(x-y) dy$$

C'est la valeur moyenne de $\gamma(x-y)$ lorsque x et y décrivent séparément V .

b/ Lorsque V est réunion de N volumes v_i disjoints, de teneurs Z_i , la variance expérimentale :

$$s^2(v|V) = \frac{1}{N} \sum_{i=1}^N [Z_i - Z(V)]^2$$

est la variance de la population finie constituée des N Z_i (qui sont des nombres, à réalisation fixée) ou, si l'on veut, la variance de la V.A. Z_i obtenue en tirant au sort l'un des N indices i . En déconditionnant relativement à la réalisation, on obtient une V.A. Z_i , où l'indice i est tiré au sort comme ci-dessus, et où Z_i est maintenant une intégrale stochastique (et non plus numérique). La variance de cette variable déconditionnée est (par définition) la variance $\sigma^2(v|V)$ de v dans V . On constate, comme ci-dessus, qu'elle est égale à la valeur moyenne en i

$$\sigma^2(v|V) = \frac{1}{N} \sum_i E[Z_i - Z(V)]^2$$

de la variance d'extension de v_i dans V . A partir de la formule (2-14), un calcul facile montre que l'on a cette fois :

$$(2-18) \quad \sigma^2(v|V) = \frac{1}{V^2} \int_V dx \int_V \gamma(x-y) dy - \frac{1}{V^2} \int_V dx \int_V \gamma(x-y) dy$$

En particulier, on a :

$$(2-19) \quad \sigma^2(v|V) = \sigma^2(0|V) - \sigma^2(0|v)$$

c/ Dans le cas général où v et V sont des volumes quelconques, non nécessairement géométriquement compatibles, on définit la variance par cette même formule (2-18) : cette quantité ne représente alors qu'un simple artifice de calcul. Elle peut même prendre des valeurs négatives ; en particulier, on a toujours $\sigma^2(v|V) = -\sigma^2(V|v)$.

d/ Relation d'additivité. Soient v , V et V' trois volumes, avec par exemple $v \subset V \subset V'$. De (2-19), on tire :

$$\sigma^2(v|V) = \sigma^2(0|V) - \sigma^2(0|v)$$

$$\sigma^2(v|V') = \sigma^2(0|V') - \sigma^2(0|v)$$

et par différence :

$$\sigma^2(v|V') - \sigma^2(v|V) = \sigma^2(0|V') - \sigma^2(0|V) = \sigma^2(V|V')$$

c'est-à-dire :

$$(2-20) \quad \sigma^2(v|V') = \sigma^2(v|V) + \sigma^2(V|V')$$

La variance de l'échantillon v dans le champ V' est égale à la somme des variances de v dans le panneau V et du panneau V dans le champ V' .

e/ Covariance de v et v' dans V . On définit d'une manière analogue la covariance $\sigma(v, v'|V)$ de deux échantillons v et v' (dont la distance et la disposition mutuelles restent fixes) dans le champ V . On trouve :

$$\sigma(v, v'|V) = \frac{1}{V^2} \int_V dx \int_V \gamma(x-y) dy - \frac{1}{v v'} \int_v dx \int_{v'} \gamma(x-y) dy$$

f/ En particulier, il est souvent commode d'exprimer la variance d'estimation (2-15) ou la variance d'extension (2-14) par l'intermédiaire des variances et covariances des échantillons dans un champ V arbitraire. On trouve, par exemple, identiquement :

$$\sigma_E^2 = \sigma^2(v|V) + \sigma^2(v'|V) - 2 \sigma(v, v'|V)$$

comme on le voit en substituant les expressions de ces variances et covariances dans V : le terme $\frac{1}{V^2} \int_V \int_V$ disparaît, et les trois termes qui subsistent donnent le second membre de (2-14). On obtient ainsi une expression peut-être plus intuitive de la variance d'extension.

2-4-4 Application : mailles aléatoires et aléatoires stratifiées.

Nous traiterons plus loin le cas des mailles régulières, qui est de loin le plus difficile, et nous allons examiner ici le cas des deux types usuels de mailles aléatoires.

a/ maille aléatoire pure. Pour estimer la teneur $Z(V)$ d'un volume V , on dispose des valeurs $Y(x_i)$ de la F.A. en N points x_i implantés "n'importe où" dans V . Nous admettrons que tout se passe comme si chaque x_i avait été implanté au hasard dans V avec une densité de probabilité uniforme et indépendamment des autres prélèvements. On peut obtenir la variance d'estimation σ_N^2 en intégrant (2-15) dans V relativement à chacun des x_i . Ce calcul est facile, mais il est encore plus simple de remarquer que les erreurs partielles $Y(x_i) - Z(V)$ sont indépendantes les unes des autres à réalisation fixée et admettent la même variance $s^2(O|V)$: l'erreur résultante, qui est $\frac{1}{N} \sum [Y(x_i) - Z(V)]^2$ admet donc la variance :

$$\sigma_N^2 = \frac{1}{N} s^2(O|V)$$

Il suffit ensuite de déconditionner relativement à la réalisation pour voir que la variance d'estimation, dans le cas d'une maille aléatoire pure, est égale à la variance d'un prélèvement dans le champ V divisée par le nombre N des prélèvements, soit :

$$\sigma_N^2 = \frac{1}{N} \sigma^2(O|V)$$

b/ maille aléatoire stratifiée. Le volume V que l'on veut estimer est ici divisé en N zones d'influence égales et disjointes v_i , et dans chaque v_i on plante un prélèvement en un point x_i choisi au hasard dans la zone d'influence v_i avec une densité de probabilité uniforme, et indépendamment des autres prélèvements. On peut calculer σ_N^2 en intégrant (2-15) en x_i dans v_i pour chacune des v_i , mais il est plus simple de remarquer que l'erreur totale est $\frac{1}{N} \sum_i [Y(x_i) - Z(v_i)]$ et que chacune des erreurs partielles $Y(x_i) - Z(v_i)$ est indépendante des autres et admet la variance $\sigma^2(O|V)$. On en déduit, en raisonnant, comme ci-dessus, d'abord à réalisation fixée, puis en déconditionnant relativement à la réalisation :

$$\sigma_N^2 = \frac{1}{N} \sigma^2(O|v)$$

La variance d'estimation, dans le cas d'une maille aléatoire stratifiée, est donc égale à la variance d'un prélèvement dans sa zone d'influence divisée par le nombre N des prélèvements.

Remarque - Il est clair que la maille aléatoire stratifiée donne toujours de meilleurs résultats que la maille aléatoire pure. D'après (2-20), en effet, on a :

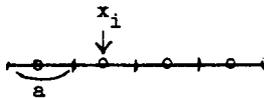
$$\frac{1}{N} [\sigma^2(O|v) - \sigma^2(O|v)] = \frac{1}{N} \sigma^2(v|v) \leq 0$$

2-5 METHODE D'APPROXIMATION DANS \mathbb{R}^1

La formule générale (2-15) est un peu lourde à calculer numériquement dès que le nombre N des prélèvements est un peu élevé, et il y a lieu de mettre au point des méthodes d'approximation permettant le calcul rapide d'une variance d'estimation. De plus, nous verrons à cette occasion réapparaître cette parenté profonde avec les formules analogues obtenues par les méthodes transitives, parenté dont nous avons déjà mentionné l'importance méthodologique à propos de la montée. Plaçons-nous, pour l'instant, dans l'espace à une seule dimension.

2-5-1 Le principe de correspondance.

Soit un segment de longueur $L = Na$ constitué de la réunion de n segments jointifs de même longueur a au centre x_i de chacun desquels on a implanté un échantillon ponctuel. Pour estimer la teneur



$$Z(L) = \frac{1}{L} \int_L Y(x) dx$$

on forme l'estimateur :

$$Z^*(L) = \frac{1}{N} \sum_i Y(x_i)$$

La variance d'estimation correspondante est donnée par la formule générale (2-15). En effectuant des calculs analogues à ceux que l'on a présentés à l'occasion des méthodes transitives, on met en évidence un principe de correspondance terme à terme entre la partie irrégulière du demi-vario-

gramme $\gamma(h)$ et le développement limité de la variance d'estimation au voisinage de $a = 0$. Ce principe se résume par les règles suivantes :

$$(2-21) \quad \begin{cases} |h|^\lambda \rightarrow -T_\lambda \frac{a^\lambda}{N} + T'_\lambda \frac{a^\lambda}{N^2} & (\lambda > 0 \text{ différent d'un entier pair}) \\ |h|^{2n} \log |h| \rightarrow -T_{2n} \frac{a^{2n}}{N} \end{cases}$$

Pour λ entier pair, T_λ est nul (mais non T'_λ). Les T_λ ont les mêmes valeurs que ceux qui figurent dans la règle analogue (1-18) des méthodes transitives.

Exactement comme dans le cas de la montée, il apparaît un terme supplémentaire, d'ordre plus élevé (en $1/N^2$), lié ici encore au caractère fini des opérations que l'on effectue en intrinsèque (moyenne d'un nombre N fini d'échantillons, et non plus, comme en transitif, somme étendue théoriquement de $-\infty$ à $+\infty$). En pratique, ce terme supplémentaire est négligeable dès que N n'est pas très petit.

Les règles (2-21) permettent d'obtenir une expression approchée de la variance d'estimation pour les mailles a très petites. Cette expression ne peut rester valable que si a est inférieure à la distance r_0 au-delà de laquelle le développement limité de $\gamma(h)$ lui-même cesse d'être utilisable. Pour les valeurs de a supérieures à r_0 , on est conduit à introduire un second principe d'approximation :

2-5-2 Principe de composition des variances d'extension élémentaires.

a/ Les fonctions intrinsèques auxiliaires. Outre $\gamma(h)$, on utilise constamment dans les applications les fonctions suivantes :

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

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

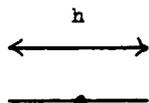
$F(h)$ est la valeur moyenne de $\gamma(h')$ lorsque les deux extrémités de h' décrivent le segment $(0, h)$. Ainsi, la variance du segment h dans le segment L est :

$$\sigma^2(h|L) = F(L) - F(h)$$

$\chi(h)$ permet de même le calcul des covariances : la covariance dans L du segment h avec l'une de ses extrémités (ponctuelle) est :

$$\sigma(0, h | L) = F(L) - \chi(h)$$

b/ Variance d'extension élémentaire. On appelle ainsi la variance d'extension au segment h de l'échantillon ponctuel implanté au centre de ce segment. Elle est donnée par :



$$(2-22) \quad \sigma_E^2 = 2 \chi\left(\frac{h}{2}\right) - F(h)$$

On déduira (2-22) de (2-15) et des deux relations établies en a/ .

c/ Principe de composition des variances d'extension élémentaires. Soit à estimer la teneur du segment $L = na$ divisé en n zones d'influences de longueur a, au centre de chacune



desquelles on a effectué un prélèvement ponctuel. Soit Y_i la teneur du prélèvement i, Z_i celle de sa zone d'influence, et $Z = \frac{1}{N} \sum Z_i$ celle de L. L'erreur totale est la moyenne :

$$\frac{1}{N} \sum_i (Y_i - Z_i)$$

des erreurs partielles $Y_i - Z_i$. Le principe d'approximation consiste à admettre que ces erreurs partielles sont indépendantes les unes des autres (ce principe est vérifié avec une approximation très raisonnable pour les $\gamma(h)$ de type usuel). Comme la variance de $Y_i - Z_i$ est justement la variance d'extension élémentaire calculée en (2-21), on trouve :

$$(2-23) \quad \sigma_n^2 = \frac{1}{N} \sigma_E^2 = \frac{1}{n} [2 \chi\left(\frac{a}{2}\right) - F(a)]$$

On obtient donc la variance d'estimation en divisant la variance d'extension élémentaire d'un échantillon dans sa zone d'influence par le nombre n des échantillons.

d/ Cas du dispositif fermé. A partir de n+1 prélèvements à maille régulière a, on



veut estimer le segment $L = na$ compris entre le premier et le dernier prélèvement. On montre que ce dispositif fermé est équivalent au dispositif centré examiné en c/ , et que la variance d'estimation est encore

donnée par (2-23) (pourvu que n soit supérieur à 1) : le dispositif fermé à $N+1$ échantillons est donc équivalent au dispositif centré à n échantillon.

Remarque - Pour $n = 1$, le dispositif fermé à deux échantillons donne la variance d'estimation :

$$\sigma_{\bar{F}}^2 = 2 \chi(h) - F(h) - \frac{1}{2} \gamma(h)$$

qui diffère de la variance d'extension élémentaire (2-23). On n'a pas le droit de diviser $\sigma_{\bar{F}}^2$ par n , car les erreurs partielles commises dans l'estimation de deux segments fermés consécutifs  ne sont pas du tout indépendantes (ces deux figures, en particulier, ont en commun l'échantillon central).

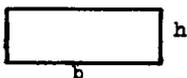
e/ Comparaison des deux principes d'approximation. Pour a petit, on déduit de (2-23) un principe de correspondance analogue à (2-22), avec des coefficients numériquement peu différents des T_λ (du moins pour λ pas trop grand), de sorte que les règles (2-23) et (2-22) sont pratiquement équivalentes. Dans un but de simplification, on utilise toujours, en pratique, la règle (2-23) de composition des variances d'extension élémentaires, puisqu'elle reste à peu près valable pour les grandes et les petites mailles.

2-6 METHODE D'APPROXIMATION DANS \mathbb{R}^n

Nous traiterons explicitement le cas de l'espace à 2 dimensions. Le cas général s'en déduit facilement par analogie. Le principe d'approximation que nous utiliserons est le principe de composition des terme de ligne et de tranche étudié en détail au paragraphe 1-4-4. On pourrait en donner, dans le cas intrinsèque, une justification analogue à celle que nous avons présentée à propos des méthodes transitives, mais les calculs seraient nettement plus compliqués.

Nous supposons que le demi-variogramme est isotrope (ne dépend que de r). En cas d'anisotropie géométrique, il est aisé de se ramener à ce cas.

a/ Les fonctions auxiliaires. Il est commode d'introduire les fonctions suivantes :



$\gamma_b(h)$: moyenne de $\gamma(x-y)$ lorsque x et y décrivent les deux côtés parallèles de longueur b du rectangle $b \times h$ (à une constante près, $\gamma_b(h)$)

se déduit de $\gamma(h)$ par montée d'ordre 1.

$\chi_b(h)$: moyenne de $\gamma(x-y)$, x décrivant l'un des côtés b du rectangle et y le rectangle lui-même. On

a :

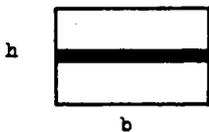
$$\chi_b(h) = \frac{1}{h} \int_0^h \gamma_b(x) dx$$

$F(b,h)$: moyenne de $\gamma(x-y)$ lorsque x et y décrivent le rectangle. Cette fonction est symétrique en b et h , et vérifie :

$$F(b,h) = \frac{2}{h^2} \int_0^h x \chi_b(x) dx = \frac{2}{h^2} \int_0^h (h-x) \gamma_b(x) dx$$

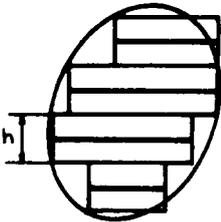
$Q(b,h)$: moyenne de $\gamma(x-y)$, x décrivant un côté b et y un côté h , ou encore, x décrivant le rectangle, y restant fixe en l'un des sommets.

b/ Extension d'un segment médian dans son rectangle d'influence.



$$(2-24) \quad \sigma_E^2 = 2 \chi_b\left(\frac{h}{2}\right) - F(h,b) - \gamma_b(0)$$

c/ Estimation de S par des traçages parallèles équidistants.



Soient b_1, b_2, \dots, b_n les longueurs des n traçages, h leur équidistance. On assimile la surface S que l'on veut estimer à la réunion des rectangles d'influence des n traçages. La formule (2-24) donne l'extension de b_i à son rectangle d'influence, soit $\sigma_{E_i}^2$. Si Y_i est la teneur de b_i , Z_i celle de sa zone d'influence, on admet que les $(Y_i - Z_i)$ sont indépendantes, et

l'erreur totale :

$$\frac{\sum b_i (Y_i - Z_i)}{b_i}$$

admet une variance que l'on peut calculer en pondérant par les carrés des b_i les variances d'extension $\sigma_{E_i}^2$. D'où la variance d'estimation :

$$(2-25) \quad \sigma_E^2 = \frac{\sum b_i^2 \sigma_{E_i}^2}{(\sum b_i)^2}$$

Remarque - Si les b sont égaux, il reste simplement $\sigma_E^2 = \frac{1}{N} \sigma_{E_1}^2$.

d/ Composition d'un terme de ligne et d'un terme de tranche.

Dans le cas c/ précédent, il arrive que l'on ne connaisse pas les teneurs réelles des traçages, mais seulement une estimation de ces teneurs à partir de prélèvements ponctuels à maille régulière a ($a < h$). On admet alors que les erreurs commises en estimant les lignes à partir des prélèvements et S elle-même à partir des lignes (supposées connues) peuvent être regardées comme indépendantes. La variance d'estimation est alors :

$$(2-26) \quad \sigma_E^2 = \frac{1}{N} \sigma^2(a) + \frac{\sum b_i^2 \sigma_{E_i}^2}{(\sum b_i)^2}$$

$\sigma^2(a)$ est la variance d'extension élémentaire (2-23) d'un prélèvement ponctuel dans son segment a d'influence, N est le nombre total des prélèvements. $\frac{1}{N} \sigma^2(a)$ est le terme de ligne, variance de l'erreur commise en estimant les lignes à partir des prélèvements.

Le deuxième terme, qui figure déjà en (2-25) est le terme de tranche : variance de l'erreur commise en étendant les teneurs des lignes à leurs tranches d'influence.

Ce principe (2-26) de composition est valable pourvu que a soit inférieur à h . Il s'applique, en particulier, au cas d'une reconnaissance à maille rectangulaire $a \times h$.

e/ Cas d'une maille carrée. La variance d'extension d'un sondage à son carré d'influence,

au centre duquel il est implanté, est :



(2-27)

$$\sigma_E^2 = 2 Q\left(\frac{a}{2}, \frac{a}{2}\right) - F(a, a)$$

Pour une maille carrée, on peut admettre que les erreurs commises en estimant chaque carré à partir de son sondage central sont indépendantes. La variance d'estimation par N sondages est alors :

$$\sigma_N^2 = \frac{1}{N} \sigma_E^2$$

avec un σ_E^2 donné par (2-27).

f/ Abaques. Pour chaque schéma isotrope, c'est-à-dire pour chaque fonction $\gamma(r)$, on doit présenter sous forme d'abaques :

1/ La fonction $F(a,b)$ définie en a/, qui sert au calcul de la variance d'un échantillon dans le rectangle a, b.

2/ La variance d'extension (2-24) d'un traçage médian de longueur b dans le rectangle b,h - avec le cas particulier $b = 0$, où l'on retrouve la variance d'extension élémentaire (2-22).

3/ La variance d'extension (2-27) d'un sondage dans son carré a,a d'influence.

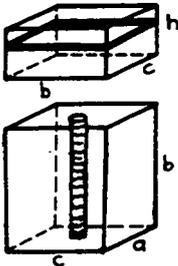
Ces trois documents permettent le calcul rapide de toutes les variances d'estimation.

Cas de l'espace à trois dimensions.- On procédera exactement comme dans le cas de l'espace à deux dimensions, en composant cette fois trois sortes de termes :

- terme de ligne : extension des prélèvements ponctuels dans les lignes.
- terme de section : extension des lignes dans les sections.
- terme de tranche : extension des sections dans les tranches.

On devra disposer des abaques suivants(en plus des précédents) :

- variance d'extension d'un plan b,c médian dans son parallélépipède rectangle d'influence.



- extension d'un sondage de longueur b dans son prisme droit a,a,b d'influence.

- moyenne de $\gamma(h)$ dans le parallélépipède rectangle a,b,c, : cette fonction $F(a,b,c,)$ est destinée au calcul de la variance d'un échantillon dans ce parallélépipède.

2-7 L'EFFET DE PÉPITE

2-7-1 Genèse d'un effet de pépité.

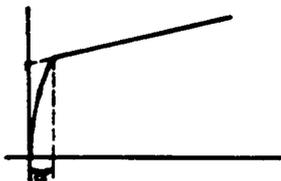
Un demi-variogramme de portée a finie caractérise ce que l'on appelle un phénomène de transition : au-delà de la distance a, l'indépendance est atteinte, et la portée a donne l'échelle des structures élémentaires du phénomène régionalisé correspondant. Il y a souvent, d'ailleurs, superposition de plusieurs structures d'échelles bien différentes, emboîtées les unes dans les autres. Le variogramme expérimental montre alors une succession de seuils et de paliers, dont l'analyse permet de reconstituer la hiérarchie de ces "structures gigogne".

La notion d'échelle joue ici un rôle primordial. A l'échelle de la dizaine ou de la centaine de mètres, un phénomène de transition dont la portée est, par exemple, centimétrique ne se manifeste plus, sur le $\gamma(h)$ expérimental, que comme une discontinuité à l'origine, c'est-à-dire comme un effet de pépité. D'une manière générale, tout effet de pépité est une réminiscence d'une structure de transition dont les dimensions sont depuis longtemps dépassées à l'échelle à laquelle on travaille : les détails et les caractéristiques qualitatives de cette structure antérieure ont depuis longtemps cessé d'être perceptibles, et l'échelle supérieure n'a guère conservé qu'un paramètre unique - la constante de pépité - qui donne une sorte de mesure globale indifférenciée de l'"intensité" de cette structure dépassée.

Pour analyser la genèse d'un effet de pépité, plaçons-nous d'abord au niveau ponctuel, et supposons qu'à une structure primaire de dimension a se superpose une macrorégionalisation, c'est-à-dire une structure secondaire de dimensions beaucoup plus grandes. Si la structure primaire existait seule, on pourrait décrire la V.R. correspondante comme une réalisation d'une P.A. admettant une covariance $C(h)$ de portée a, ou un demi-variogramme :

$$\gamma_1(h) = C - C(h)$$

de portée a et vérifiant $\gamma_1(\infty) = C = C(0)$. Pour tenir compte de la macrorégionalisation, on doit ajouter une deuxième composante $\gamma_2(h)$, qui représente la structure secondaire, et ne varie qu'avec une extrême lenteur à l'échelle a de la première structure :



$$\gamma(h) = C - C(h) + \gamma_2(h)$$

Ce $\gamma(h)$ présente donc au voisinage de l'origine une zone de crois-

sance très rapide, dont la dimension est de l'ordre de a . A l'échelle de la macrorégionalisation, ce $\gamma(h)$ présentera donc un effet de pépité d'amplitude C .

2-7-2 Influence macroscopique d'un effet de pépité.

Examinons les conséquences de cet effet de pépité. Tout d'abord, au niveau macroscopique, les prélèvements ne seront plus ponctuels, mais seront des volumes v déjà grands vis-à-vis de a . Déterminons donc le demi-variogramme $\gamma_v(h)$ de ces prélèvements v (le seul qui sera expérimentalement accessible). γ_v est la somme de la composante très continue γ_2 (qui n'a pas été sensiblement altérée dans cette régularisation) et de ce que l'on obtient en appliquant la formule (2-12) à la composante $C - C(h)$: cette composante pépitique est donc :

$$\gamma_p(h) = \frac{1}{v^2} \int_v dx \int_v C(x-y) dy - \frac{1}{v^2} \int_v dx \int_v C(h+x-y) dy$$

Etudions donc γ_p . On a $\gamma_p(0) = 0$, mais dès que h dépasse les dimensions du volume v , on a $|h+x-y| \geq a$ si $x \in v$ et $y \in v$, et $C(h+x-y) = 0$. Le premier terme subsiste donc seul dès que h n'est pas très petit, et l'on observe expérimentalement une discontinuité à l'origine, dont la valeur σ_p^2 (variance de pépité) est :

$$\sigma_p^2 = \frac{1}{v^2} \int_v dx \int_v C(x-y) dy$$

D'autre part, les dimensions de v sont grandes, par hypothèse, vis-à-vis de a , et on a $\int C(x-y) dy = \int C(h) dh$ (intégrale étendue à tout l'espace) sauf si x est situé à une distance inférieure à a de la frontière de v . Mais ces points occupent un volume négligeable, et l'on peut aussi négliger leur influence, qui est de l'ordre de a . Il reste donc :

$$\sigma_p^2 = \frac{1}{v^2} \int_v dx \int C(h) dh = \frac{1}{v} \int C(h) dh$$

Ainsi, la constante de pépité σ_p^2 que l'on observe expérimentalement est en raison inverse du volume des prélèvements :

$$\sigma_p^2 = \frac{A}{v}$$

et le coefficient $A = \int C(h) dh$, qui est l'intégrale de la covariance $C(h)$ des microstructures,

est le seul souvenir de celles-ci qui subsiste à l'échelle des volumes v .

L'effet de pépite majore de la même manière la variance de v dans V , les variances d'extension et les variances d'estimation. Pour $\sigma^2(v|V)$, par exemple, l'apport de l'effet de pépite sera :

$$\frac{1}{v^2} \int_V dx \int_V C(x-y) dy - \frac{1}{V^2} \int_V dx \int_V C(x-y) dy$$

c'est-à-dire, d'après le calcul précédent :

$$A \left(\frac{1}{v} - \frac{1}{V} \right)$$

Dans la variance d'estimation de V par N prélèvements de taille v , on trouverait de même une composante pépitique égale à :

$$A \left(\frac{1}{Nv} - \frac{1}{V} \right)$$

Dans tous les cas, la variance due à l'effet de pépite est en raison inverse du volume des prélèvements. Tout se passe donc bien comme si la V.R. elle-même admettait deux composantes indépendantes, l'une très régulière, correspondant au demi-variogramme γ_2 , l'autre complètement aléatoire et discontinue, prenant en charge l'effet de pépite.

Représentation d'un effet de pépite par une mesure de Dirac.

Les effets macroscopiquement observables d'un effet de pépite dépendent uniquement, on l'a vu, de la constante :

$$A = \int C(h) dh$$

et nullement de la forme exacte de la microcovariance $C(h)$. Au niveau macroscopique, donc, on peut aussi bien symboliser l'effet de pépite en remplaçant la fonction $C(h)$ par la mesure de Dirac $A\delta$ (masse A placée en 0). Les exercices 16 à 20 des méthodes transitives nous ont déjà familiarisés avec ce point de vue. Dans la version probabiliste, on remplace le variogramme ponctuel γ par $\gamma_1(h) - A\delta$ (avec le signe -, puisqu'il s'agit de variogramme). Les formules habituelles permettent de retrouver immédiatement le terme de pépite = A/v observable macroscopiquement.

En toute rigueur, pour justifier l'introduction de cette mesure de Dirac dans l'expression

du variogramme ou de la covariance ponctuelle, il serait nécessaire de recourir à la théorie des mesures aléatoires (et non plus des F.A.), ce que nous ne faisons pas ici. En fait, la mesure de Dirac est la covariance d'une mesure aléatoire correspondant, par exemple, à une implantation poissonnienne de points dans l'espace \mathbb{R}^n . On trouvera quelques détails à ce sujet dans les Exercices.

2-8 LE SCHEMA DE DE WIJS

Le schéma de De Wijs est défini par la fonction intrinsèque (isotrope)

$$\gamma(r) = 3 \alpha \log r$$

Le coefficient α est la dispersion absolue : elle caractérise la dispersion du phénomène à l'état pur, c'est-à-dire indépendamment de toute influence de la géométrie du champ ou du support. On notera que $\log r$ tend vers $-\infty$ pour $r \rightarrow 0$, de sorte que la fonction $\log r$ ne peut pas être, en toute rigueur, le variogramme d'une F.A.I. En fait, ce variogramme logarithmique (mesure $\log r$, et non plus fonction) caractérise une mesure aléatoire, et non plus une F.A. Mais les régularisées de cette mesure aléatoire sont à nouveau des F.A., et on peut montrer que les variogrammes de ces régularisées se déduisent de la fonction logarithmique conformément aux règles valables pour les F.A. En pratique, cela veut dire que l'on peut travailler sur le schéma de De Wijs comme s'il s'agissait d'une F.A.I., à condition de ne jamais raisonner au niveau ponctuel, mais toujours sur des régularisées.

2-8-1 Equivalents linéaires.

Posons :

$$F(v) = \frac{1}{v^2} \int_v dx \int_v \log |x-y| dy$$

(valeur moyenne de $\log |h|$ dans le volume v : c'est une fonction qui dépend non seulement de la mesure du volume v , mais aussi de sa forme). D'après la relation (2-18), on a :

$$(2-28) \quad \sigma^2(v|V) = 3 \alpha [F(V) - F(v)]$$

Dans un même gisement De Wijsien, si deux échantillons v et v' vérifient $F(v) = F(v')$, ils ont même variance dans n'importe quel panneau. Nous diront qu'ils sont équivalents. Si nous prenons comme ensemble v un segment de droite de longueur ℓ , nous avons :

$$F(\ell) = \log \ell - \frac{3}{2}$$

(le démontrer à l'aide des fonctions auxiliaires de 2-5-2, a/). Nous appellerons équivalent linéaire d'un échantillon v la longueur du segment ℓ équivalent à v , c'est-à-dire la longueur définie par :

$$\log \ell - \frac{3}{2} = F(v)$$

Soit, de même, L l'équivalent linéaire du champ V :

$$\log L - \frac{3}{2} = F(V)$$

Par différence, il vient :

$$F(V) - F(v) = \log (L/\ell)$$

et (2-28) donne alors :

$$(2-29) \quad \sigma^2(v|V) = 3 \alpha \log \left(\frac{L}{\ell} \right)$$

Cette formule permet un calcul rapide de la variance. On l'utilise souvent en sens inverse, pour évaluer la dispersion absolue à partir de la variance expérimentale des échantillons dans leur champ : si v et V sont géométriquement semblables, on a $V/v = (L/\ell)^3$, et (2-28) se réduit à la formule de De Wijs : $\sigma^2(v|V) = \alpha \log (V/v)$.

Calcul approché des équivalents linéaires. - Avec une excellente précision, l'équivalent linéaire du rectangle de côtés a et b est $a + b$.

Pour le parallélépipède rectangle (a, b, c) , c désignant le plus petit côté, on a, avec une approximation cette fois assez grossière, un équivalent linéaire de $a + b + 0,7 c$. Pour plus de précision, on consultera l'abaque 1.

Application : comportement du demi-variogramme des échantillons v pour h grand devant les dimensions de v . (On n'utilise jamais le schéma de De Wijs au niveau ponctuel, car $\log r$ est infini en $r = 0$, mais toujours ses régularisées¹). Soit γ_v le demi-variogramme d'échantillons v distants de h . $2 \gamma_v(h)$ est la variance d'extension de ces deux échantillons, et la formule (2-14) nous donne :

$$\gamma_v(h) = \frac{3\alpha}{v^2} \int_v \int_v \log |h+x-y| dx dy - 3 \alpha F(v)$$

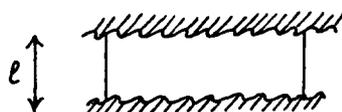
Dès que h dépasse les dimensions de v , le premier terme ne diffère plus beaucoup de $3 \alpha \log r$. Le deuxième terme est $3 \alpha (\log \ell - 3/2)$, ℓ désignant l'équivalent linéaire de v . Pour h assez grand devant v , on a donc :

$$(2-30) \quad \gamma_v(r) = 3 \alpha \log \frac{r}{\ell} + 9 \frac{\alpha}{2}$$

Cette formule est précieuse dans les applications expérimentales.

2-8-2 Le schéma de De Wijs à deux dimensions.

Dans l'application précédente, on peut prendre pour v un segment de longueur ℓ se déplaçant parallèlement à lui-même. D'après (2-30), on a :



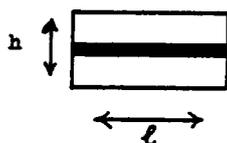
$$\gamma_\ell(h) = 3 \alpha \log \frac{r}{\ell} + \frac{9\alpha}{2} \quad \text{pour } h \gg \ell$$

Pour h petit, au contraire, la règle de montée $\log r \rightarrow \pi r/\ell$ donne

$$(2-31) \quad \gamma_\ell(h) = 3 \alpha \pi \frac{h}{\ell}$$

A partir de l'expression rigoureuse de $\gamma_\ell(h)$ (déduite de $\log r$ par montée) on calcule la variance d'extension d'un segment ℓ dans son rectangle ℓ, h d'influence : Abaque 2. Lorsque h est petit devant ℓ (en pratique, il suffit que h soit inférieur ou égal à ℓ), on peut utiliser (2-31) pour calculer directement cette variance d'extension, ce qui donne :

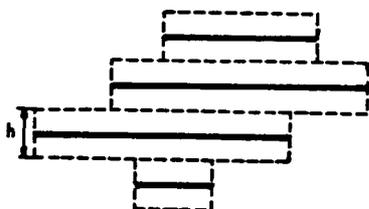
1 - Au niveau ponctuel, le schéma de De Wijs ne représente plus une fonction, mais une distribution aléatoire [4] - Mais les régularisées de ce schéma sont, à nouveau, des F.A.



(2-32)

$$\sigma_E^2 = \alpha \frac{\pi}{2} \frac{h}{l}$$

Application : Gisement stratiforme reconnu par traçages parallèles.



a/ Supposons d'abord que les teneurs réelles des traçages soient connues, et calculons le terme de tranche. Si h est l'équidistance des traçages, on lit sur l'abaque 2 la variance d'extension $\sigma_{E_1}^2$ du traçage de longueur l_1 dans sa tranche d'influence. Le terme de tranche est alors :

$$\sigma_E^2 = \frac{\sum l_1^2 \sigma_{E_1}^2}{(\sum l_1)^2}$$

Si chacun des traçages a une longueur l_1 supérieure à h, $\sigma_{E_1}^2$ est donnée par (2-32), et, en posant $L = \sum l_1$ (longueur tracée totale), on trouve :

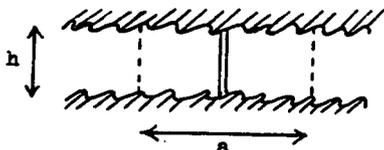
$$\sigma_E^2 = \alpha \frac{\pi}{2} \frac{\sum l_1^2 \frac{h}{l_1}}{L^2} = \alpha \frac{\pi}{2} \frac{h}{L}$$

ou encore, avec $S = L h$ (surface minéralisée totale) :

$$\sigma_E^2 = \alpha \frac{\pi}{2} \frac{S}{L^2}$$

Conséquence : un gisement 4 fois plus gros, toutes choses égales d'ailleurs, nécessite, à précision égale, deux fois moins de dépenses de reconnaissance à la tonne de minerai.

b/ Si les teneurs des traçages sont estimées à partir de rainures implantés à maille régulière a, on doit ajouter un terme de ligne. Pour le calculer, il n'est plus possible de négliger la puissance p de la formation (p est petit, par hypothèse, devant h ou les longueurs l_1 des traçages, mais non devant a). C'est à nouveau l'abaque 2 qui donne la variance d'extension $\sigma^2(a)$ de la rainure de longueur p dans son rectangle d'influence a, p. Le terme



de ligne est alors :

$$\frac{1}{h} \sigma^2(a) = \frac{a}{L} \sigma^2(a)$$

Remarque - Dans la plupart des gisements stratiformes ou filoniens, la troisième dimension, celle des puissances, n'est pas équivalente aux deux autres. Il y a anisotropie. La longueur p que l'on doit attribuer au rainurage peut donc différer de la puissance réelle. On détermine cette puissance équivalente p en remarquant, d'après (2-30), que le demi-variogramme expérimental des rainurages a pour équation :

$$\gamma(h) = 3 \alpha \left[\log \left(\frac{h}{p} \right) + \frac{3}{2} \right] \quad \text{pour } h > p$$

C'est la valeur de p ainsi déterminée qui doit servir au calcul de $\sigma^2(a)$.

Couche mince reconnue par sondages à maille carrée. - Soit p la puissance équivalente de la formation, définie comme dans la remarque qui précède. On doit calculer la variance d'extension d'un sondage dans son carré (a, a) d'influence. En assimilant ce carré au cercle de même surface, on trouve (cf. Exercice 11) :

$$\sigma_E^2 = 3 \alpha \left[\log \frac{2a}{p} - \frac{1}{2} \right]$$

La variance d'estimation s'en déduit en divisant par le nombre n des sondages :

$$\sigma_n^2 = \frac{3\alpha}{n} \left[\log \frac{2a}{p} - \frac{1}{2} \right]$$

2-8-3 Le schéma de De Wijs dans l'espace à trois dimensions.

L'abaque 1 donne l'équivalent linéaire du parallélépipède rectangle.

Estimation des amas. - On calculera le terme de ligne et le terme de section comme ci-dessus (à partir de l'abaque 2). Pour le terme de tranche on utilisera, selon le cas, l'abaque 3, 4 ou 5 (variance d'extension des grandes sections ab , des sections moyennes ac et des petites sections bc , ($a \geq b \geq c$)).

Amas reconnu par sondages à maille carrée. - On lit sur l'abaque 6 la variance d'extension $\sigma_{E_1}^2$ d'un sondage de puissance h_1 dans son prisme droit h_1, a, a d'influence. On a ensuite :

$$\sigma_E^2 = \frac{\sum h_i^2 \sigma_{E_i}^2}{(\sum h_i)^2}$$

(si les h_i sont égaux ou peu différents : $\frac{1}{n} \sigma_{E_i}^2$).

2-9 LE SCHEMA SPHERIQUE

Pour représenter un phénomène de transition, on peut utiliser des schémas de la forme :

$$\gamma(h) = A[K(o) - K(h)]$$

où $K(h)$ est le covariogramme géométrique d'un volume v (cf. Exercice 9, ch. 1). Si l'on prend pour v la sphère de diamètre a , on a (cf. ex. 9 des méthodes transitives) :

$$K(h) = \begin{cases} \frac{\pi}{6} a^3 \left(1 - \frac{3}{2} \frac{h}{a} + \frac{1}{2} \frac{h^3}{a^3}\right) & (|h| < a) \\ 0 & (|h| \geq a) \end{cases}$$

Le schéma sphérique sera donc défini par le demi-variogramme :

$$\gamma(r) = \begin{cases} C \left(\frac{3}{2} \frac{r}{a} - \frac{1}{2} \frac{r^3}{a^3}\right) & \text{pour } r < a \\ C & \text{pour } r \geq a \end{cases}$$

La portée est a , le palier est $C = \gamma(\infty)$, la pente à l'origine $\frac{3}{2} \frac{C}{a}$.

Dans les applications, on déterminera a et C par approximations successives : on part de valeurs a_1 et C_1 obtenues par simple interpolation sur le demi-variogramme expérimental ; on utilise l'abaque 7 pour calculer la variance théorique d'un échantillon dans son champ ; si celle-ci est, par exemple, plus forte que la variance expérimentale des échantillons dans leur champ, on peut soit diminuer C_1 , soit augmenter a_1 . On arrive assez vite à l'ajustement cherché.

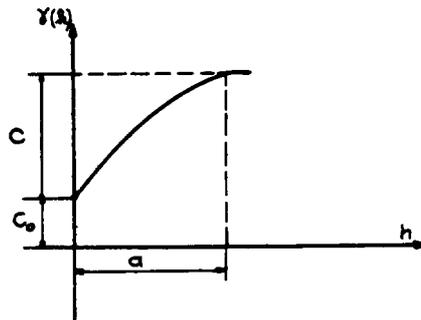
Nous ne disposons que des abaques relatifs aux problèmes à une et à deux dimensions :

Abaque 7 : variance d'un point dans un rectangle.

Abaque 8 : variance d'extension élémentaire d'un point dans son segment d'influence (pour le calcul du terme de ligne) ; variance d'extension d'un point dans son carré d'influence (pour les sondages à maille carrée).

Abaque 9 : variance d'extension d'un segment médian dans son rectangle d'influence (pour le calcul des termes de tranche).

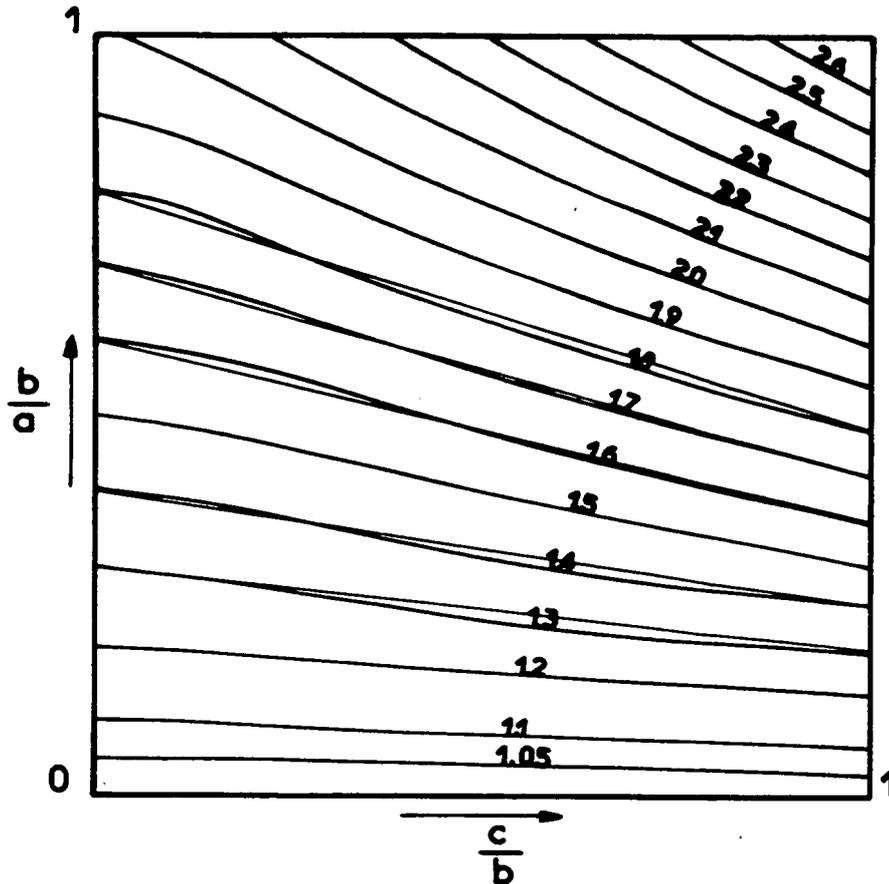
Remarque - Dans les applications, les phénomènes transitifs sont très souvent accompagnés d'un effet de pépité : on les représentera donc par un schéma sphérique à effet de pépité C_0 :

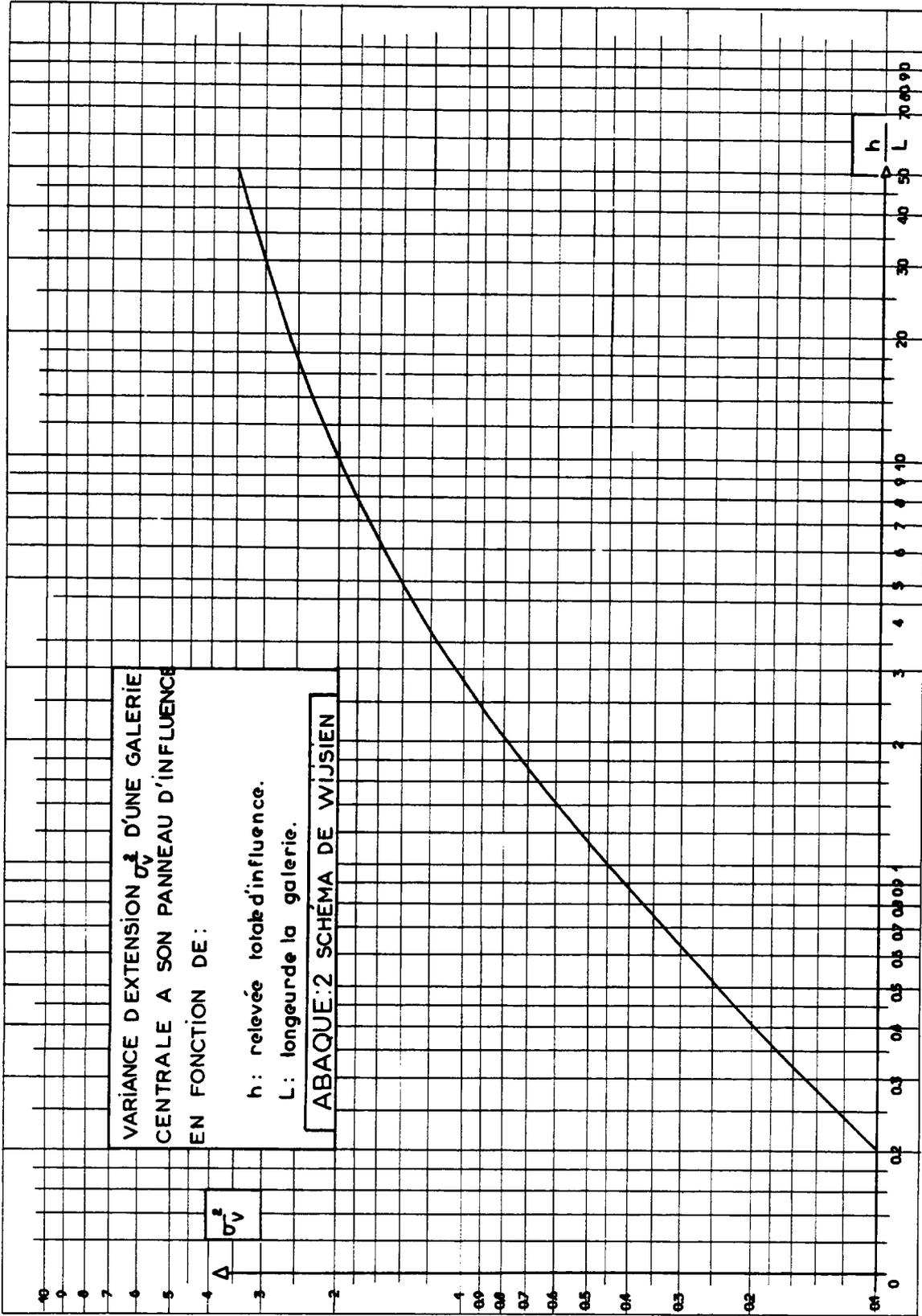


ABAQUE 1

equivalence lineaire en schéma de wijsien

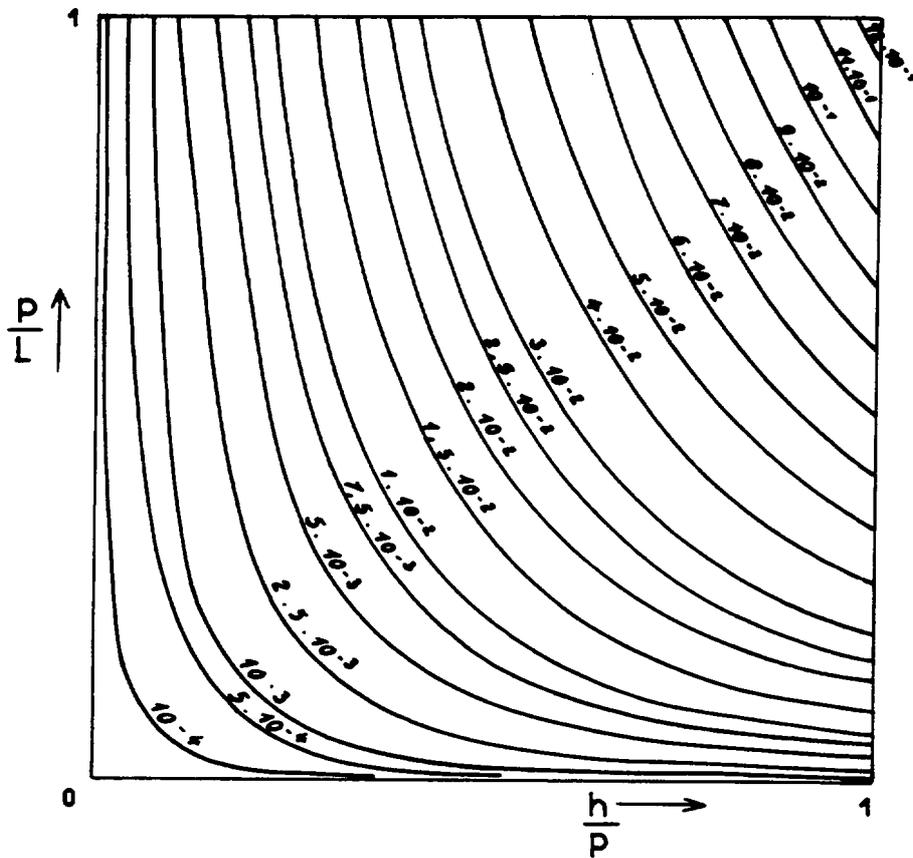
le parallélépipède $a \times b \times c$ (a b c) est equivalent
a l'échantillon linéaire λa
les courbes sont graduées en λ





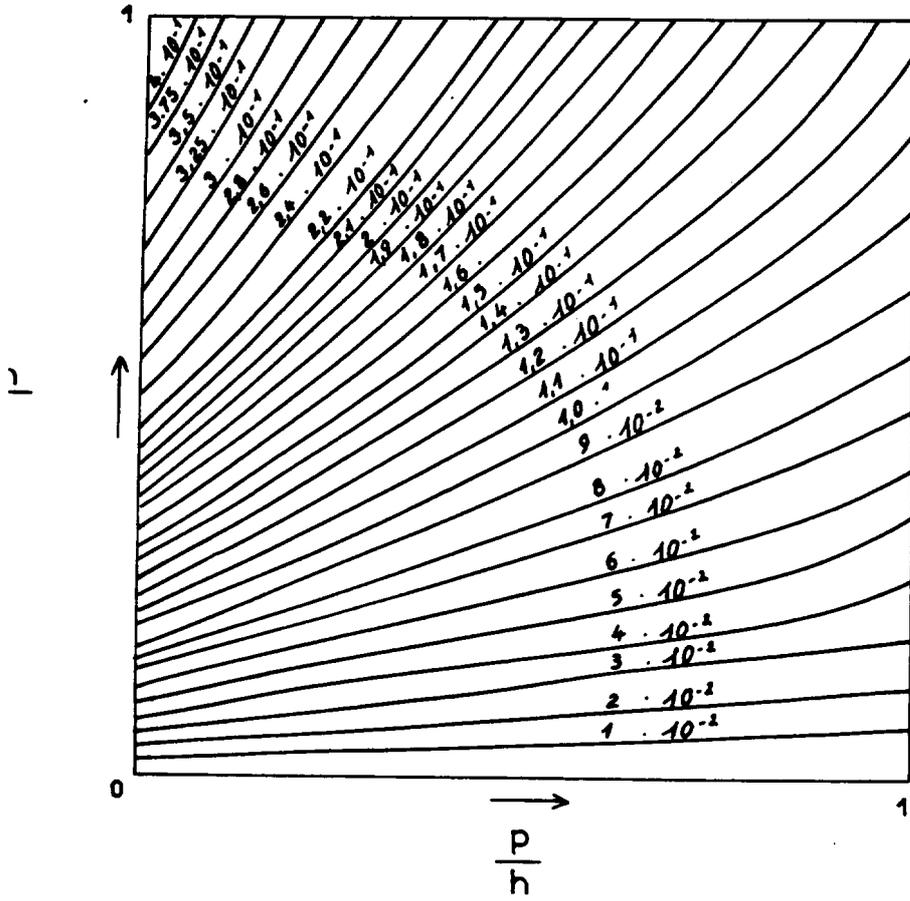
Abaque 3

schéma de wijsien

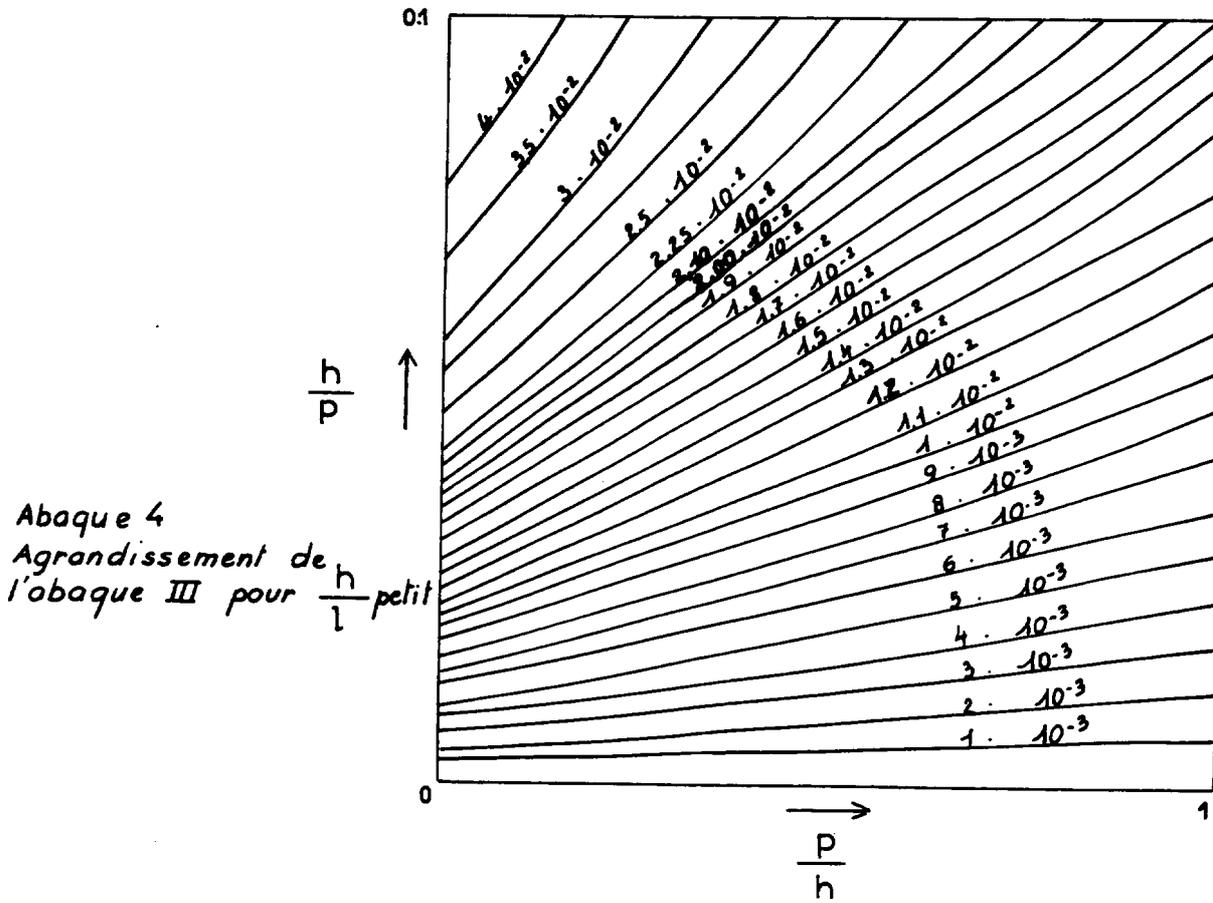


Abaque 3
 Variance d'extension
 d'une grande section
 d'influence de relevée
 $h: l \gg p \gg h$.
 les courbes sont
 graduées en $\frac{1}{36} \sigma_E^2$

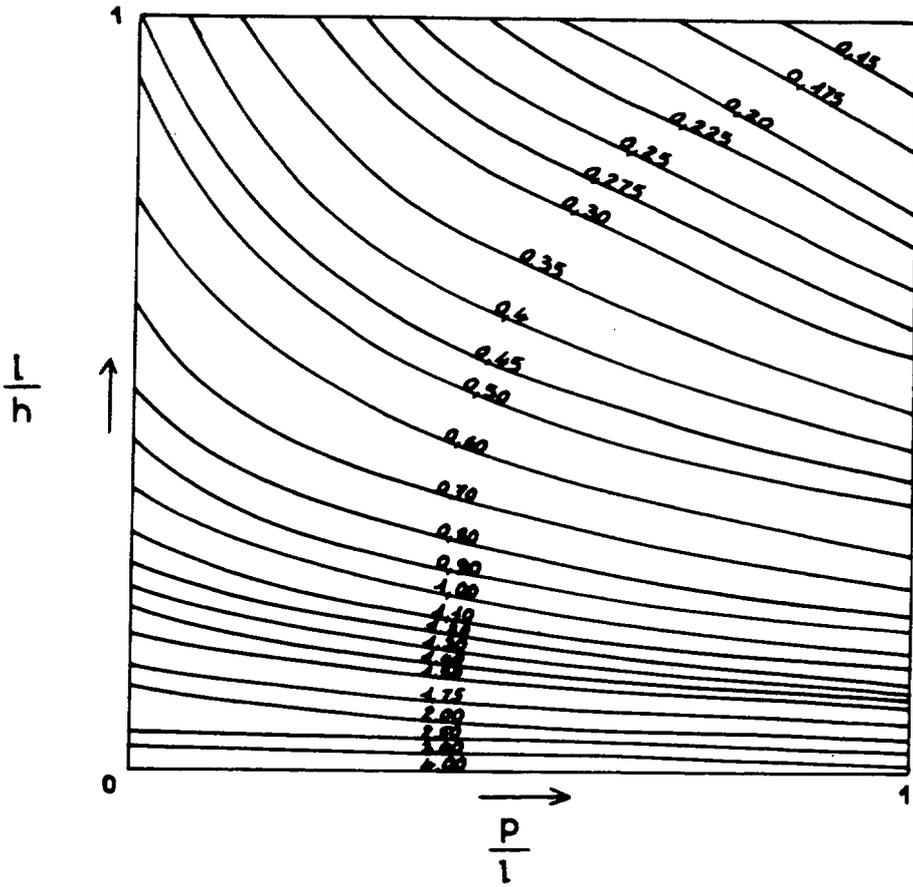
Abaque 4 schéma de wijnsien



Abaque 4 . Variance d'extension d'une section intermediaire $l \times p$ dans sa zone d'influence de relevé $h = l \geq h \geq p$ les courbes sont gradueés en $\frac{1}{3\pi} \sigma_E^2$.



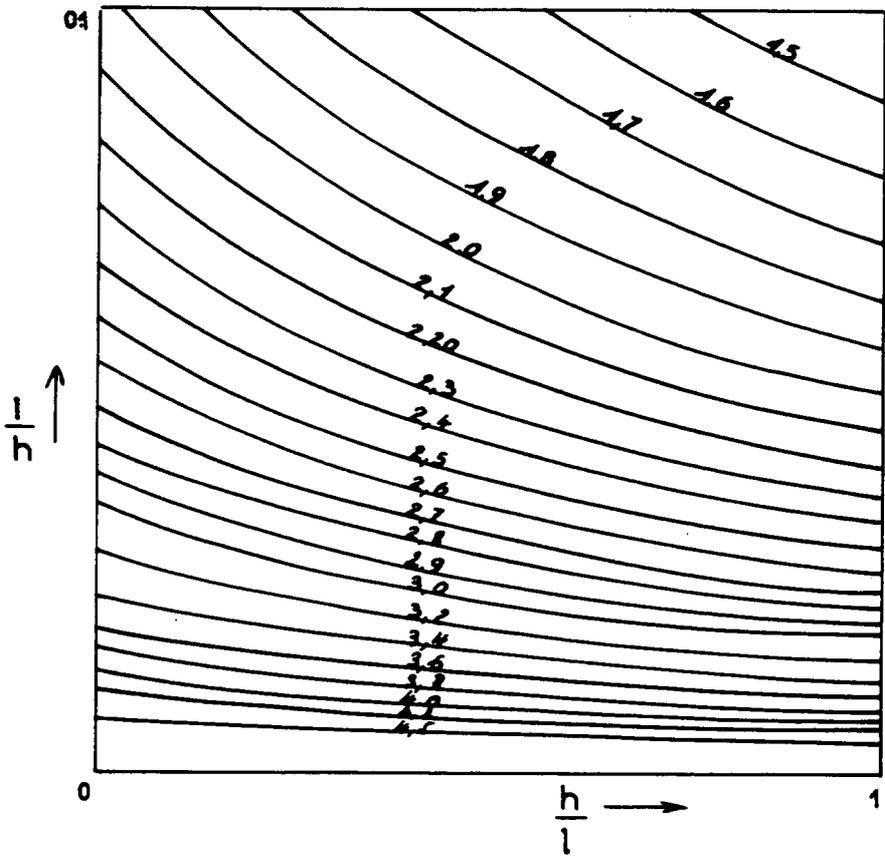
Abaque 4
Agrandissement de l'abaque III pour $\frac{l}{h}$ petit

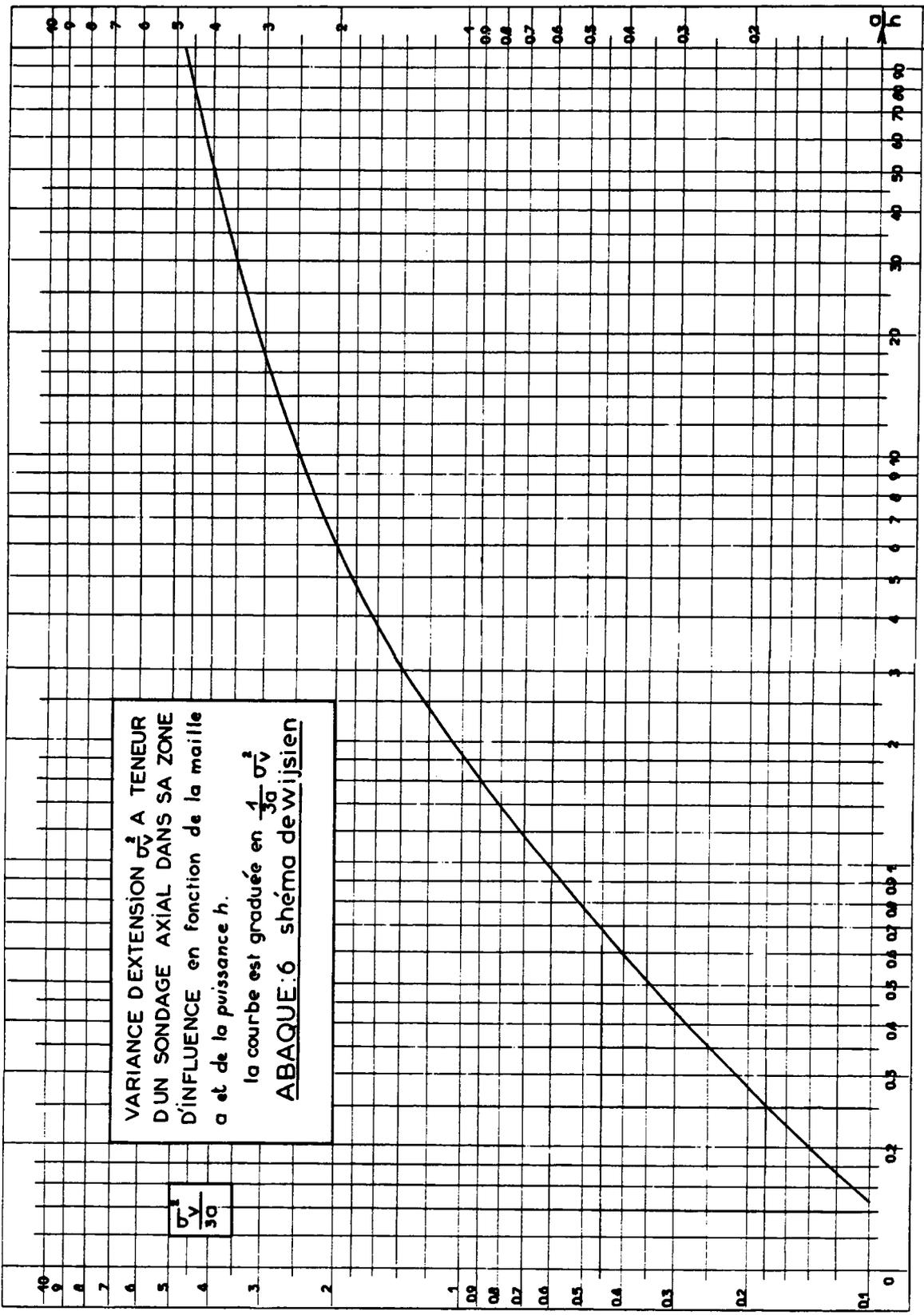


Abaque 5
schéma de wigsien

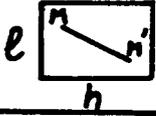
Abaque 5
Variance d'extension
d'une petite section
 $l \times p$ dans sa zone
d'influence de relevé
 $h = h \gg l \gg p$
les courbes sont
graduées en $\frac{1}{34} \sigma^2 E$

Abaque 5
Agrandissement de
l'abaque IV
pour $\frac{l}{h}$ petit

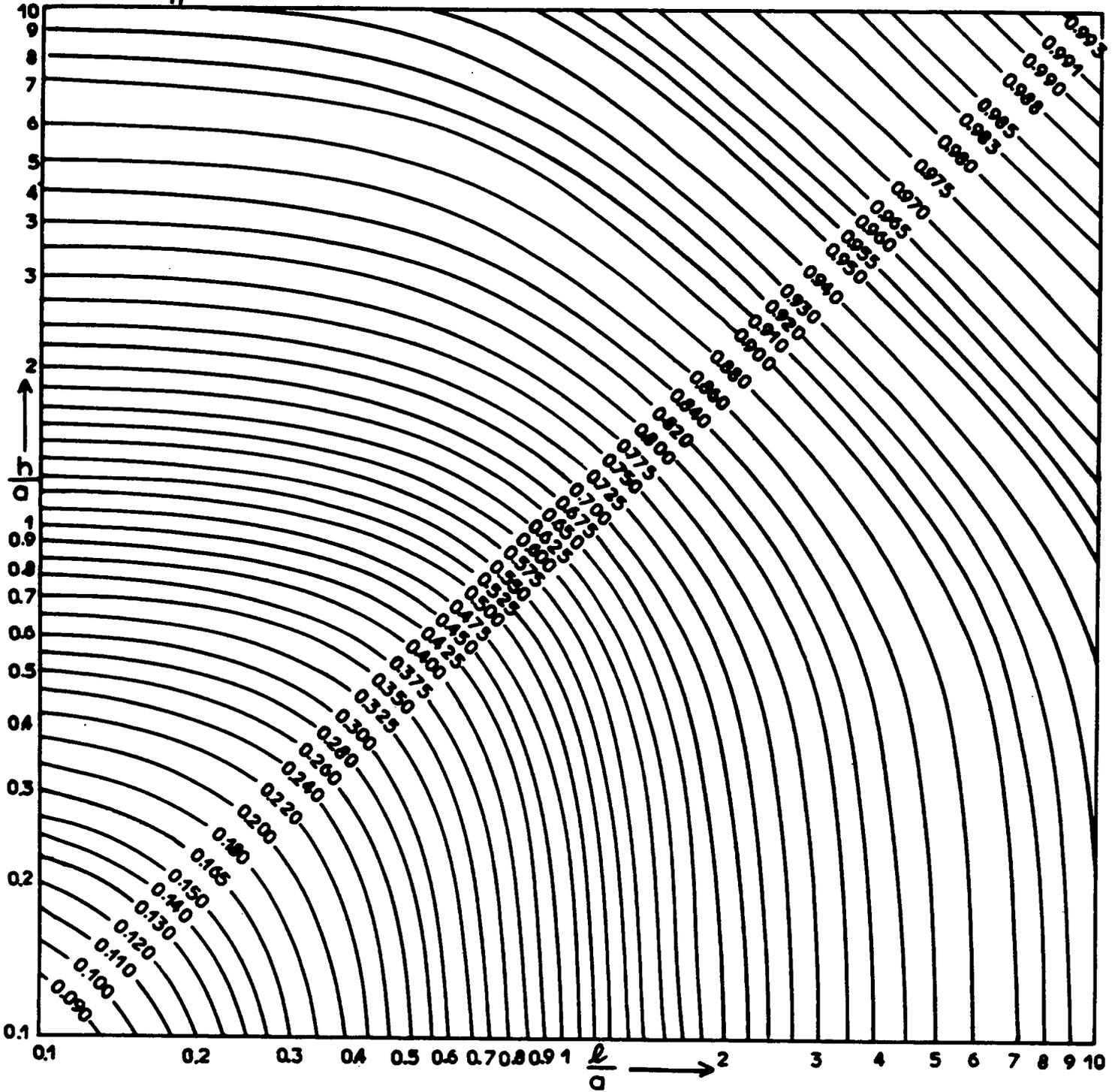


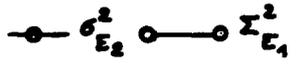


ABAUQUE : 7 — SCHEMA SPHÉRIQUE —

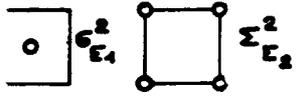


Fonction $\frac{1}{C} F\left(\frac{l}{a}, \frac{h}{a}\right)$ = Variance d'un point dans le rectangle $l \times h$

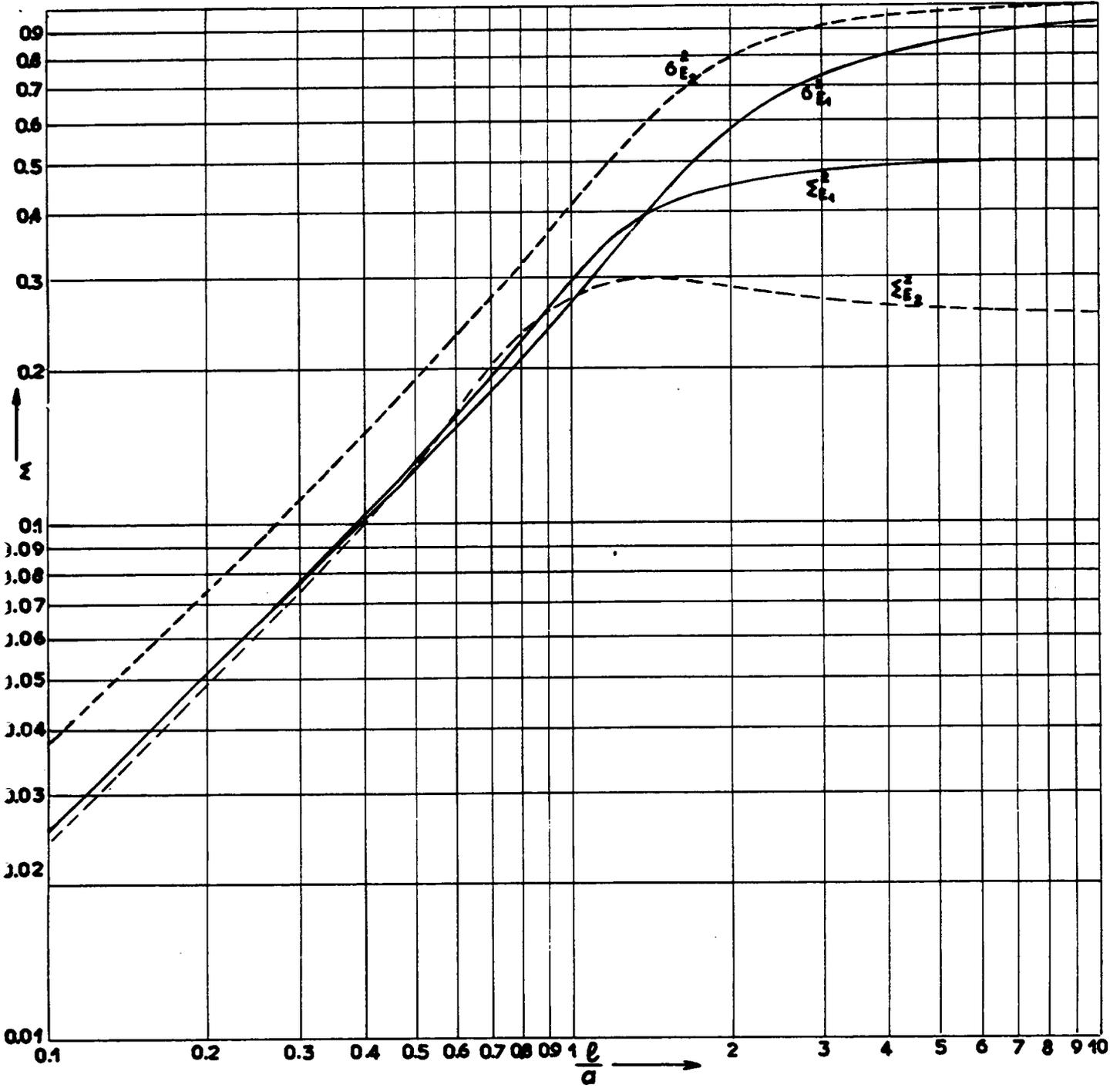




ABAQUE 8 schéma sphérique.

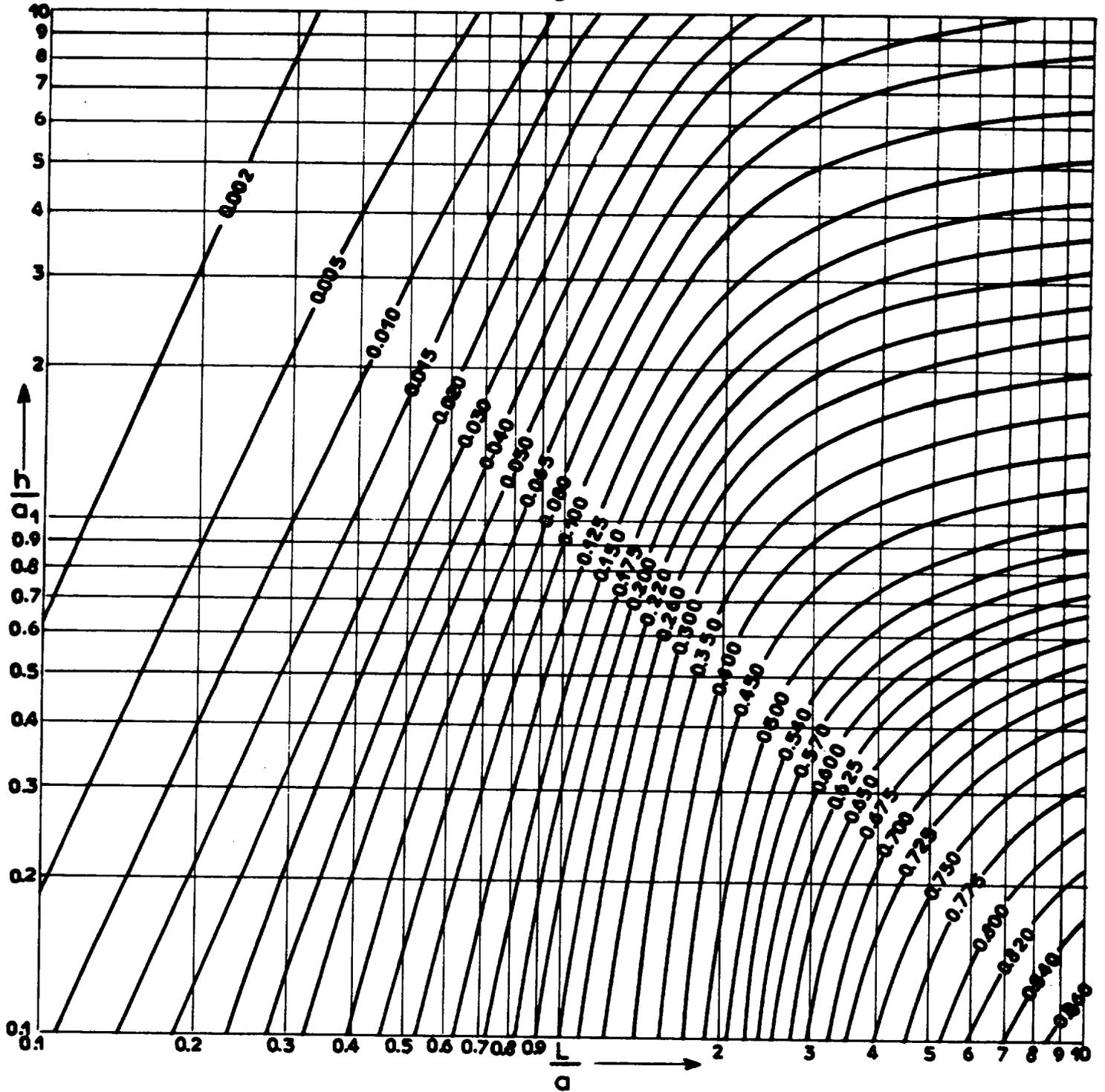


variance d'extention diverses.



ABAQUE : 9 SCHÉMA SPHÉRIQUE.

variance d'extension du tracé médian h
dans le rectangle $L \times h$



2-10 INFERENCE STATISTIQUE ET QUASI-STATIONNARITE.2-10-1 F.A. quasi-stationnaire.

Revenons maintenant au problème méthodologique évoqué au paragraphe 1-4-6. Il s'agit de montrer - moyennant une hypothèse assez faible du type quasi-stationnaire, que le problème de l'estimation globale d'un domaine V donné est soluble. Cela résultera des deux points suivants :

~ La variance d'estimation ne dépend (du moins pour les petites mailles) que du comportement au voisinage de l'origine d'un variogramme ou d'une covariance "moyenne".

~ Ce comportement est accessible expérimentalement, autrement dit l'inférence statistique de ce qu'il est utile de connaître est possible à partir d'une réalisation unique, pourvu que la F.A. ne soit pas dérivable en moyenne quadratique.

Il serait possible de poser comme hypothèse minimale le caractère quasi-intrinsèque, c'est-à-dire l'existence d'un variogramme γ localement stationnaire et ne se déformant que lentement dans l'espace. Nous nous limiterons ici à l'hypothèse légèrement plus forte de quasi-stationnarité que nous allons maintenant préciser : une fonction aléatoire $Z(x)$ sera quasi-stationnaire si elle possède une espérance $m(x)$ et une covariance centrée $C(x,y)$ telles que :

a/ $m(x)$ est une fonction très régulière et lentement variable dans l'espace (à l'échelle de la maille utilisée) ; plus précisément, $m(x)$ peut être considérée comme constante sur un domaine dont les dimensions sont celles de la maille.

b/ Il existe une fonction de trois arguments $K(h;x,y)$ telle que l'on ait $C(x,y) = K(x-y;x,y)$ et que, à h fixé, $K(h;x,y)$ soit une fonction régulière et lentement variable (dans le même sens que ci-dessus) des deux arguments x et y . Autrement dit, lorsque x et y appartiennent à un même domaine dont les dimensions sont celles de la maille, $K(h;x,y)$ ne dépend que de h , et tout se passe comme si la covariance C était localement stationnaire.

En fait, nous ferons même une hypothèse un peu plus restrictive. Nous admettrons que, sur le domaine V que l'on veut estimer, la fonction K admet un développement en série entière de la forme :

$$K(h;x,y) = \sum_n \omega_n(x) \omega_n(y) C_n(h)$$

les ω_n désignant des fonctions régulières et lentement variables à l'échelle de la maille, et les $C_n(h)$ des fonctions de covariance.

Il est alors possible de trouver des F.A. $Y_n(x)$ stationnaires d'ordre 2 d'espérance nulle, mutuellement indépendantes et telles que l'on ait :

$$Z(x) = m(x) + \sum_n \omega_n(x) Y_n(x)$$

Cette décomposition est suffisamment générale pour ce que nous avons en vue. Mais les opérations du type calcul d'une variance d'estimation, estimation d'un variogramme, variance de cette estimation etc... vont se présenter linéairement relativement aux composantes indépendantes $Y_n(x)$. Il va donc suffire de traiter séparément les composantes $Y_n(x)$, et d'additionner ensuite les résultats obtenus. Il en résulte une simplification importante, puisque nous sommes ramenés à l'étude du cas où $Z(x)$ est de la forme

$$(2-33) \quad Z(x) = m(x) + \omega(x) Y(x)$$

donc admet l'espérance $m(x)$ et une covariance centrée de la forme :

$$(2-34) \quad C(x,y) = \omega(x) \omega(y) C_0(x-y)$$

Examinons alors comment se présentent nos deux problèmes lorsque ces relations (2-33) et (2-34) sont vérifiées.

2-10-2 - Calcul de la variance d'estimation.

Pour simplifier, nous nous limiterons au cas de l'espace R^1 , et nous supposons que l'on veuille estimer l'intervalle $(0,L)$ à partir de n échantillons à maille régulière $a = L/n$ implantés selon le dispositif centré. On désignera par $x_i = (i - \frac{1}{2})a$ les points de prélèvements, et par

$$Z_i = \frac{1}{a} \int_{(i-1)a}^{ia} Z(x) dx$$

la teneur réelle de la zone d'influence numéro i . Si l'on prend l'estimateur $\frac{1}{n} \sum_n Z(x_i)$, l'erreur

d'estimation est :

$$\frac{1}{n} \sum_1 (Z(x_1) - Z_1)$$

Dans le terme $Z(x_1) - Z_1$, la contribution de $m(x)$ est négligeable, puisque $m(x)$ est considérée comme constante sur une zone d'influence : $m(x)$ ne joue aucun rôle dans ce problème. Il reste donc :

$$Z(x_1) - Z_1 = \omega(x_1) Y(x_1) - \frac{1}{a} \int_{(i-1)a}^{ia} \omega(x) Y(x) dx$$

et, $\omega(x)$ étant considérée comme constante sur une zone d'influence :

$$Z(x_1) - Z_1 = \omega(x_1) \left[Y(x_1) - \frac{1}{a} \int_{(i-1)a}^{ia} Y(x) dx \right] = \omega(x_1) [Y(x_1) - Y_1]$$

La variance de cette erreur individuelle est donc :

$$D^2[Z(x_1) - Z_1] = [\omega(x_1)]^2 \sigma_{E_0}^2$$

en désignant par $\sigma_{E_0}^2$ la variance d'extension de x_1 dans sa zone d'influence calculée sur la covariance C_0 de $Y(x)$. Or, les méthodes d'approximation que nous avons exposées au paragraphe 2-5 montrent que les différentes erreurs $Y(x_1) - Y_1$ peuvent être considérées comme sans corrélation. Il en résulte que notre variance d'estimation est :

$$\sigma_{Est}^2 = \sigma_{E_0}^2 \frac{1}{n^2} \sum_1 \omega^2(x_1) + \frac{1}{n^2} \sigma_{E_0}^2 \frac{1}{a} \int_0^L \omega^2(x) dx$$

soit :

$$\sigma_{Est}^2 = \frac{1}{n} \sigma_{E_0}^2 \frac{1}{L} \int_0^L \omega^2(x) dx$$

Ainsi, cette variance d'estimation est la même que si l'on avait la covariance stationnaire :

$$C(h) = C_0(h) \frac{1}{L} \int_0^L \omega^2(x) dx$$

Désignons alors par $\bar{C}(h)$ la covariance (pseudo-stationnaire) moyenne :

$$\bar{C}(h) = \frac{1}{L-h} \int_0^{L-h} C(x, x+h) dx = \frac{C_0(h)}{L-h} \int_0^{L-h} \omega(x) \omega(x+h) dx$$

Pour $h \leq a$, $\omega(x+h)$ ne diffère pas de $\omega(x)$ et on a :

$$\frac{1}{L-h} \int_0^{L-h} \omega(x) \omega(x+h) dx \neq \frac{1}{L} \int_0^L \omega^2(x) dx$$

Donc, pour $h \leq a$, $C(h) = \bar{C}(h)$, et la variance d'estimation est la même que pour une F.A. stationnaire dont la covariance serait précisément cette covariance moyenne $\bar{C}(h)$.

Pour a petit, il reste à montrer qu'il est possible d'estimer effectivement $\bar{C}(0) - \bar{C}(h)$.

2-10-3 Possibilité de l'inférence statistique.

Nous poserons $\bar{\gamma}(h) = \bar{C}(0) - \bar{C}(h)$, $\gamma_0(h) = C_0(0) - C_0(h)$ etc... Nous cherchons à estimer le comportement de $\bar{C}(h)$ au voisinage de l'origine, c'est-à-dire $\bar{\gamma}(h)$ pour h petit. Pour cela nous allons prendre l'estimateur :

$$\gamma^*(h) = \frac{1}{2(L-h)} \int_0^{L-h} (\omega(x+h) Y_{x+h} - \omega(x) Y_x)^2 dx \neq \frac{1}{2(L-h)} \int_0^{L-h} \omega^2(x) [Y_{x+h} - Y_x]^2 dx$$

Avec les approximations faites, cet estimateur est sans biais :

$$E(\gamma^*(h)) = \bar{\gamma}(h)$$

Il reste à évaluer sa variance. Il est nécessaire, ici, d'introduire une hypothèse sur la loi spatiale de la F.A., puisque les moments d'ordre 4 vont nécessairement intervenir. Le plus simple est de se limiter au cas gaussien, comme dans l'Exercice 16. Les calculs effectués dans les paragraphes a/ et b/ restent valables ici et donnent :

$$(2-35) \quad D^2(\gamma^*(h)) = \frac{1}{2(L-h)^2} \int_0^{L-h} \omega^2(x) dx \int_0^{L-h} \omega^2(y) [\gamma_0(x-y+h) + \gamma_0(x-y-h) - 2\gamma_0(x-y)]^2 dy$$

Evaluons cette intégrale lorsque h est petit. Il est clair qu'elle est équivalente à :

$$\frac{1}{2L^2} \int_0^L \int_0^L \omega^2(x) \omega^2(y) [\gamma_0(x-y+h) + \gamma_0(x-y-h) - 2\gamma_0(x-y)]^2 dx dy$$

Posons :

$$F(u) = [\gamma_0(u+h) + \gamma_0(u-h) - 2\gamma_0(u)]^2$$

et désignons par R le covariogramme transitif de la fonction égale à $\omega^2(x)$ sur $(0, L)$ et à 0 ailleurs. L'algorithme de Cauchy nous donne :

$$D^2(\gamma^*(h)) = \frac{1}{L^2} \int_0^L R(u) F(u) du$$

Pour $u > h$, $F(u)$ est une fonction régulière de u , équivalente à $h^4(\gamma''(u))^2$. Donc notre variance est de la forme :

$$D^2(\gamma^*(h)) = A h^4 + \frac{R(0)}{L^2} \int_0^h F(u) du$$

Tout va dépendre maintenant du comportement de $F(u)$ au voisinage de 0. Désignons par $b h^\lambda$ (éventuellement $b h^2 \log h$) la partie principale de γ_0 en $h = 0$. Celle de $F(u)$ est :

$$b^2 h^{2\lambda} \left[\left(1 + \frac{u}{h}\right)^\lambda + \left(1 - \frac{u}{h}\right)^\lambda - 2u^\lambda \right]^2$$

Par suite, la variance sera de la forme :

$$(2-36) \quad D^2(\gamma^*(h)) = A h^4 + \frac{B}{L^2} h^{1+2\lambda}$$

Il y alors deux cas à distinguer :

a/ $\lambda < 2$ (F.A. non dérivable m.q.) - La variance relative, pour h petit est du type :

$$\frac{D^2(\gamma^*(h))}{(\gamma(h))^2} = A' h^{4-2\lambda} + \frac{B'}{L^2} h$$

Elle est donc infiniment petite avec h , et l'inférence statistique est possible. Dans le cas limite où $\gamma(h)$ aurait un comportement en $h^2 \log h$, on trouverait une variance relative en $A'/\log h$: l'inférence serait encore théoriquement possible, mais ne pourrait se faire que très difficile-

ment en pratique, à cause de la décroissance très lente de $1/\log h$.

b/ Pour $\lambda = 2$ (F.A. dérivable m.q.), la variance relative $D^2(\gamma^*)/\gamma^2$ tend vers une constante A' non nulle lorsque h tend vers 0. Cela signifie (à moins que L ne soit très grand vis-à-vis de la portée de $C_0(h)$) que l'inférence statistique est impossible.

En résumé, l'inférence statistique est toujours possible lorsque la F.A. n'est pas dérivable m.q. - et c'est donc seulement dans ce cas que notre problème est théoriquement soluble en toute généralité. Cette circonstance un peu étrange qui pénalise les F.A. trop régulières et interdit l'inférence de leur covariance si l'on ne dispose pas d'une réalisation très étendue prendra tout son sens lorsque nous étudierons le krigeage universel, et, particulièrement, la notion de dérive aléatoire.

2-11 EXERCICES SUR LES FONCTIONS ALÉATOIRES INTRINSEQUES.

2-11-1 Construction de F.A.I.

Exercice 1 - Une origine c_0 étant choisie au hasard sur $(0, a)$, on divise la droite en segments de longueur a au moyen des points de subdivision $x_k + ka$ (k entier, positif ou négatif). Soit $Y(x)$ la F.A. prenant sur chacun de ces segments de longueur a une valeur aléatoire constante Y , tirée au sort indépendamment d'un segment à l'autre selon une même loi de moyenne m et de variance σ^2 . Montrer que la probabilité pour que deux points x et $x+h$ appartiennent au même segment a est $1 - |h|/a$ pour $|h| \leq a$ et 0 pour $|h| > a$. En déduire le demi-variogramme de $Y(x)$ ($\frac{|h|}{a} \sigma^2$ pour $|h| \leq a$ et σ^2 pour $|h| > a$).

Exercice 2 - même question que dans l'exercice précédent, mais les longueurs des segments successifs sont aléatoires, indépendantes les unes des autres, et obéissent à la même loi exponentielle $e^{-\lambda \ell}$ (autrement dit, les points de discontinuité constituent un processus de Poisson).

(Solution : la probabilité pour que x et $x+h$ soient dans le même segment est $e^{-\lambda h}$. On en déduit $\gamma(h) = (1 - e^{-\lambda h}) \sigma^2$.)

Exercice 3 - On se donne sur la droite réelle un processus de Poisson (points aléatoires séparés par des distances indépendantes admettant la même loi de densité $a e^{-ax}$). On définit (à une constante près) une F.A. $Y(x)$ en prenant $Y(x) = C^{\text{ste}}$ en dehors des points poissonniens x_i , et, en chaque x_i , $Y_+(x_i) - Y_-(x_i) = X_i$, avec des V.A. X_i indépendantes obéissant à une même loi d'espérance m et de variance σ^2 .

Montrer que la F.A. est intrinsèque (mais non stationnaire d'ordre 2), admet la dérivée $m a h$ et le demi-variogramme $\frac{1}{2} a(m^2 + \sigma^2) |h|$ (linéaire). (Ce Processus est un processus de Poisson composé, à accroissements indépendants et stationnaires. Pour calculer la dérivée, on pourra calculer les espérances $E(Y_{x+h} - Y_x)$ et $E(Y_{x+h} - Y_x)^2$ conditionnellement lorsque le nombre n de points poissonniens tombés entre x et $x+h$ est fixé, puis déconditionner en n).

Exercice 4 - a/ Soit $Y(x)$ une F.A. stationnaire d'ordre 2, m son espérance, $C(h)$ sa covariance. Montrer que la F.A. $Z(x) = \int_0^x Y(y) dy$ est intrinsèque (mais non stationnaire d'ordre 2) et admet la dérivée $m h$ et le demi-variogramme $\gamma(h) = \int_0^h (h-x) C(x) dx$.

(Solution : partir de $\gamma'' = C$).

b/ Application du processus à pentés aléatoires : on se donne des points de discontinuités poissonniens, comme dans Exercice 3, et la pente Π reste constante sur chacun des intervalles séparant les points de discontinuités. Ces pentes Π sont aléatoires et indépendantes, d'espérance nulle et de variance ω_2 sur chacun des intervalles. Montrer que le processus $Y(x)$ ainsi défini (à une constante près) est une F.A.I. (non stationnaire d'ordre 2) sans dérivée, de demi-variogramme $\omega_2 \frac{e^{-ah} - 1 + ah}{a^2}$. Noter le comportement parabolique en $h = 0$ de ce variogramme.

(Solution : le processus dérivé est du type de l'Exercice 2, et il suffit d'appliquer a/ à $Y'(x)$).

2-11-2 Exercices sur les variances d'estimation.

Exercice 5 - On considère le demi-variogramme h^λ ($0 < \lambda < 2$) dans l'espace à une dimension. Calculer les fonctions auxiliaires χ et F , la variance d'extension élémentaire d'un point central à son segment d'influence, la variance d'extension à ce segment de ses deux extrémités (dispositif fermé). Comparer et discuter.

$$X(h) = \frac{h^\lambda}{\lambda+1} ; F(h) = \frac{2 h^\lambda}{(\lambda+1)(\lambda+2)} ; \frac{2a^\lambda}{\lambda+1} \left[\frac{1}{2^\lambda} - \frac{1}{\lambda+2} \right] ; \left(\frac{2}{\lambda+2} - \frac{1}{2} \right) a^\lambda$$

On a $\sigma_E^2 > \sigma_E^2$, pour $\lambda > 1$ et inversement, avec égalité pour $\lambda = 1$: pour $\lambda > 1$, haute continuité, et l'échantillon unique mais bien placé est supérieur aux deux échantillons mal placés; pour $\lambda < 1$, on est proche du cas aléatoire pur, et les deux échantillons, même mal placés, donnent toujours un meilleur résultat que l'échantillon unique ; pour $\lambda = 1$, les deux dispositifs sont équivalents).

Exercice 6 - Avec le demi-variogramme h^λ dans l'espace à une dimension, calculer la variance d'estimation du segment $L = n a$ par n échantillons ponctuels, dans les trois cas suivants :

- maille régulière (dispositif centré) $\left(\frac{1}{n} \frac{2}{\lambda+1} \left[\frac{1}{2^\lambda} - \frac{1}{\lambda+2} \right] a^\lambda \right)$
- maille aléatoire stratifiée $\left(\frac{1}{n} \frac{2 a^\lambda}{(\lambda+1)(\lambda+2)} \right)$
- maille aléatoire pure $\left(\frac{1}{n} \frac{2}{(\lambda+1)(\lambda+2)} L^\lambda \right)$

Exercice 7 - Soit, dans l'espace à deux dimensions, le demi-variogramme r^λ . Soit un gisement reconnu par traçages de longueurs supérieures à leur équidistance h .

a/ Calculer la variance d'extension d'un traçage de longueur ℓ dans son rectangle ℓ, h d'influence.

(la montée sous puissance ℓ transforme r^λ en $A r^{\lambda+1}/\ell$ (ou $\log r$ en $\pi r/\ell$) ; d'où :

$$\sigma_E^2 = A \frac{2}{\lambda+2} \left[\frac{1}{2^{1+\lambda}} - \frac{1}{\lambda+3} \right] \frac{h^{1+\lambda}}{\ell} = B \frac{h^{1+\lambda}}{\ell}$$

b/ En déduire la variance d'estimation (pondérer par les carrés des longueurs des traçages, poser $L = \sum \ell_i$ et $S = L h$, et mettre la variance d'estimation sous la forme $B \frac{S^{1+\lambda}}{L^{2+\lambda}}$. Cas de Wijsien : $A = \pi$, $B = \frac{\pi}{6}$, $\frac{\pi}{6} \frac{S}{L^2}$.

Exercice 8 - Même problème qu'en 7, mais on suppose de plus que les traçages sont estimés à partir de rainurages à maille a . Calculer le terme de ligne $\left(\frac{C}{L} a^{1+\lambda} \right)$ avec $C = \frac{2}{\lambda+1} \left[\frac{1}{2^\lambda} - \frac{1}{\lambda+2} \right]$. Optimisation économique : soit ω le prix du mètre de galerie, p_0 le prix d'un prélèvement, et $\ell_0 = p_0/\omega$. Optimiser la précision à prix de revient donné, ou, ce qui revient au même, le prix de revient pour une précision donnée. (Tout revient à minimiser $C h a^{1+\lambda} + B h^{2+\lambda}$ sous la condition $\frac{1}{h \ell_0} + \frac{1}{h a} = C$. On obtient la relation suivante :

$$B(2+\lambda) h^{1+\lambda} = \lambda C a^{1+\lambda} + C(1+\lambda) \frac{a^{2+\lambda}}{p_0}$$

qui donne l'équidistance h des niveaux en fonction de la maille a des rainurages.

2-11-3 Exercices sur l'effet de pépite, et les schémas De Wijsiens et sphériques.

Exercice 9 : Effet de pépite à l'état pur - On se donne des pépites ponctuelles réparties dans l'espace selon un schéma de Poisson (définition : le nombre $N(v)$ des pépites dans un volume v est une variable de Poisson de moyenne λv ; si v et v' sont disjoints, $N(v)$ et $N(v')$ sont indépendants).

a/ On prélève des volumes v , et on pose $Y(x) = N(v_x)$, v_x désignant le translaté de v implanté au point x . Montrer que la covariance de $Y(x)$ et $Y(x+h)$ est la variance de $N(v_x \cap v_{x+h})$, soit $\lambda K(h)$, $K(h)$ désignant le covariogramme géométrique du volume v .

b/ On suppose maintenant que les pépites ont des poids aléatoires indépendants (moyenne p_0 , variance σ_p^2), et on prend pour $Y(x)$ la somme $P_1 + \dots + P_N$ des poids des $N = N(v_x)$ pépites contenues dans v . Calculer la moyenne et la variance de $Y(x)$ ($E(Y) = \lambda v p_0$, $D^2(Y) = \lambda v(p_0^2 + \sigma_p^2)$). En déduire la covariance de $Y(x)$ et $Y(x+h)$ (remplacer v par $K(h)$).

Exercice 10 : Schéma de dilution. - On a des germes poissoniens comme dans l'Exercice 7. Soit $f(x)$ une fonction. On pose $Y(x) = \sum_1 f(x-x_i)$, x désignant les implantations des différents germes. Il s'agit donc d'une dilution de ces germes par la fonction f . Former la covariance $K(h)$ de $Y(x)$. ($\lambda g(h)$, avec $g = f * \check{f}$).

(on pourra se limiter au cas où f est à support compact, considérer deux points x_0 et x_0+h et un domaine borné V contenant les translats par x_0 et x_0+h du support de V . On raisonnera alors conditionnellement lorsque le nombre n des points poissoniens tombés dans V est fixé, puis on déconditionnera en n . On notera aussi combien l'interprétation à l'aide de la mesure-covariance $\lambda \delta$ des points poissoniens rend ce résultat intuitif).

Exercice 11 - Soit dans le plan un schéma de De Wijs $\gamma(r) = \log r$. On rappelle que si g est une fonction harmonique à l'intérieur d'un cercle, sa valeur moyenne dans le cercle (ou

sur sa circonférence) est égale à sa valeur au centre : $\log r$ est une fonction harmonique sauf en $r = 0$.

a/ On désigne par $C(R)$, $\chi(R)$, $F(R)$, $C_0(R)$ la valeur moyenne de $\log r$ sur la circonférence du cercle de rayon R , entre ce cercle et sa circonférence, dans ce cercle, entre ce cercle et son centre. Montrer $C(R) = \chi(R) = \log R$; $F(R) = \log R - 1/4$; $C_0(R) = \log R - 1/2$ (les deux premières relations résultent du théorème sur les fonctions harmoniques. Pour calculer $F(R)$, faire varier $R^4 F(R)$ et montrer $\frac{d}{dR} R^4 F(R) = 4 R^3 \chi(R)$. Pour $C_0(R)$, intégrer $2 \pi r \chi(r)$).

b/ Calculer la variance d'estimation d'un panneau circulaire à partir de sa circonférence ($\sigma_E^2 = 2 \chi(R) - F(R) - C(R) = \frac{1}{4}$).

c/ Variance d'extension dans le cercle d'un petit élément central dont l'équivalent linéaire est ℓ ($\log R/\ell + 3/4$). Si L est l'équivalent linéaire du cercle, cette variance d'estimation est encore $\log(L/\ell) - 1/2$. En déduire que la variance d'extension d'un échantillon équivalent à ℓ dans le carré (a, a) au centre duquel il est implanté est approximativement $\log(\frac{2a}{\ell}) - 1/2$.

Exercice 12 : schéma sphérique à une dimension - Calculer les fonctions auxiliaires du schéma sphérique :

$$\gamma(\ell) = \frac{3}{2} \frac{\ell}{a} - \frac{1}{2} \frac{\ell^3}{a^3} \quad \text{pour } \ell \leq a, \quad 1 \text{ pour } \ell \geq a$$

$$\chi(\ell) = \frac{3}{4} \frac{\ell}{a} - \frac{1}{8} \frac{\ell^3}{a^3} \quad " \quad 1 - \frac{3}{8} \frac{a}{\ell}$$

$$F(\ell) = \frac{1}{2} \frac{\ell}{a} - \frac{1}{20} \frac{\ell^3}{a^3} \quad " \quad 1 - \frac{3}{4} \frac{a}{\ell} + \frac{1}{5} \frac{a^2}{\ell^2}$$

b/ Variance d'extension élémentaire d'un échantillon dans son segment b d'influence pour $b \leq a$ et $b \geq 2a$. Interpréter.

$$\left(\frac{1}{4} \frac{b}{a} + \frac{3}{160} \frac{b^3}{a^3} \right) \quad \text{et} \quad \left(1 - \frac{3}{4} \frac{a}{b} - \frac{1}{5} \frac{a^2}{b^2} \right)$$

2-11-4 Exercices sur les grandes mailles.

Les trois exercices qui suivent constituent un test. Si vous êtes capables de comprendre la démarche des deux premiers, et d'effectuer les calculs correspondants (qui sont faciles par eux-mêmes), c'est que vous avez bien assimilé les mécanismes mentaux qui doivent être ceux du géostatisticien. Si de plus vous pouvez accepter sans être troublés les conclusions critiques du troisième exercice (non validité aux grandes mailles du principe de composition des termes de ligne et de tranche, et validité seulement très approximative du principe de composition des variances d'extension élémentaire), cela prouve que votre compréhension s'étend en profondeur, et vous permet de jauger le sens et les limites de chacune des hypothèses d'approximation qui rendent la géostatistique opératoire.

D'un point de vue épistémologique, on notera aussi qu'aux grandes mailles (c'est-à-dire supérieures à la portée) la forme exacte de la covariance ou du $\gamma(h)$ perd toute importance: le $K(h)$ n'intervient plus que par sa valeur C à l'origine, et ses premiers moments (A_0 et A_1 dans R^1 , A_0 , A_1 , A_2 et A_3 dans R^2 ; dans R^3 , il faudrait aller jusqu'à A_5). Ainsi, dans R^2 par exemple, tous les schémas se ramènent, aux grandes mailles, à un type unique dépendant seulement des 4 paramètres essentiels A_0 , A_1 , A_2 et A_3 et du paramètre C (qui n'intervient que comme un facteur). Aux grandes mailles, on est relativement proche des conditions de validité de la statistique classique : à la variance d'estimation C/n à laquelle conduirait cette dernière, le calcul géostatistique ajoute des termes correctifs, dont les coefficients sont précisément les A_i .

Exercice 13 (1 dimension) - On désignera par $\gamma(h) = \gamma(r)$ un demi-variogramme (isotrope) de type transitif, c'est-à-dire de la forme :

$$\gamma(r) = C - K(r) \quad \text{avec } K(r) = 0 \text{ pour } r > \text{portée.}$$

1/ Dans l'espace à 1 dimension, pour $\ell > \text{portée}$, on a :

$$C - F(\ell) = \frac{2}{\ell^2} \int_0^{\ell} (\ell-x) K(x) dx = \frac{2}{\ell^2} \int_0^{\infty} (\ell-x) K(x) dx$$

En déduire :

$$F(\ell) = C - \frac{A_0}{\ell} + \frac{A_1}{\ell^2}$$

$$K(\ell) = C - \frac{A_0}{2\ell}$$

avec :

$$A_0 = 2 \int_0^{\infty} K(x) dx, \quad A_1 = 2 \int_0^{\infty} x K(x) dx$$

2/ Variance d'extension pour une maille $a >$ portée

$$\sigma_E^2 = 2 \chi\left(\frac{a}{2}\right) - F(a) = C - \frac{A_0}{a} - \frac{A_1}{a^2}$$

Pour une galerie de longueur $\ell = n a$, la variance d'estimation est :

$$\frac{1}{n} \sigma_E^2 = \frac{a}{\ell} C - \frac{A_0}{\ell} - \frac{A_1}{a\ell}$$

Exercice 14 (2 dimensions) - 1/ Fonction $F(a,b)$. La valeur moyenne de $\gamma(h) = \gamma(r)$ dans le rectangle (a,b) est :

$$F(a,b) = \frac{4}{a^2 b^2} \int_0^a \int_0^b (a-x)(b-y) \gamma(\sqrt{x^2+y^2}) dx dy$$

2/ Soit $\gamma(h) = C - K(h)$ un schéma de type transitif (i.e. $K(h) = 0$ pour $r = |h| > \epsilon$, ϵ désignant la portée). Aux grandes mailles (a et $b >$ portée) 1/ donne :

$$C - F(a,b) = \frac{4}{a^2 b^2} \int_0^{\infty} \int_0^{\infty} (a-x)(b-y) K(\sqrt{x^2+y^2}) dx dy$$

En déduire :

$$F(a,b) = C - \frac{\pi A_1}{ab} + \frac{2 A_2}{ab} \left(\frac{1}{a} + \frac{1}{b}\right) - A_3 \frac{1}{a^2 b^2}$$

avec :

$$\begin{cases} A_1 = \frac{1}{\pi} \int_{\mathbb{R}^2} K(h) dh = 2 \int_0^{\infty} r K(r) dr \\ A_2 = 2 \int_0^{\infty} r^2 K(r) dr \\ A_3 = 2 \int_0^{\infty} r^3 K(r) dr \end{cases}$$

3/ En déduire $\chi(a;b)$, $\gamma(a;b)$, $Q(a;b)$ (toujours pour $a, b >$ portée).

$$\text{Appliquer } \chi(a;b) = \frac{1}{a} \frac{\partial}{\partial a} \frac{a^2 F}{2}, \quad \gamma(a,b) = \frac{\partial^2}{\partial a^2} \left(\frac{a^2}{2} F \right)$$

$$Q(a,b) = \frac{1}{ab} \frac{\partial^2}{\partial a \partial b} \frac{a^2 b^2 F}{4}$$

$$\text{Ceci doit donner : } \chi = C - \frac{\pi A_1}{2ab} + \frac{A_2}{a b^2}, \quad \gamma = C$$

$$Q = C - \frac{\pi A_1}{4 ab}$$

4/ Variance d'extension d'une galerie dans $\ell \times h$  $h, \frac{h}{2} >$ portée.

$$\sigma_E^2 = 2 \chi \left(\frac{h}{2}, \ell \right) - F(h, \ell) - F(\ell)$$

$$\sigma_E^2 = \frac{A_0}{\ell} - A_1 \left(\frac{\pi}{2h} + \frac{1}{\ell^2} \right) + 2 A_2 \left(\frac{1}{h\ell^2} - \frac{1}{h^2\ell} \right) + A_3 \frac{1}{h^2\ell^2}$$

5/ Variance d'estimation de S par galeries de longueurs $\ell_1 >$ portée (équidistance $h >$ portée). Poser $\sum \ell_1 = L$.- N nombre de galeries

$$\begin{aligned} \frac{\sum \ell_1 \sigma_{E_i}^2}{(\sum \ell_1)^2} &= \frac{A_0}{L} - A_1 \left(\frac{\pi}{Lh} + \frac{N}{L^2} \right) + 2 A_2 \left(\frac{N}{hL^2} - \frac{1}{h^2L} \right) + A_3 \frac{N}{h^2L^2} \\ &= A_0 \frac{h}{S} - A_1 \left(\frac{\pi}{S} + \frac{h^2 N}{S^2} \right) + 2 A_2 \left(\frac{hN}{S^2} - \frac{1}{hS} \right) + \frac{A_3 N}{S^2} \end{aligned}$$

6/ Estimation de S par galeries (équidistance $h >$ portée) échantillonnées à maille a supérieure à portée :

$$\sigma_{Est}^2 = \frac{ah}{S} C - A_1 \left(\frac{h}{aS} + \frac{\pi}{S} + \frac{h^2 N}{S^2} \right) + 2 A_2 \left(\frac{hN}{S^2} - \frac{1}{hS} \right) + \frac{A_3 N}{S^2}$$

7/ Variance d'extension d'un sondage 

$$\sigma_E^2 = 2 Q \left(\frac{a}{2}, \frac{a}{2} \right) - F(a, a) = C - \frac{\pi A_1}{a^2} - 4 \frac{A_2}{a^2} + \frac{A_3}{a^4}$$

d'où variance d'estimation de S :

$$\frac{1}{n} \sigma_E^2 = C \frac{a^2}{S} - \pi \frac{A_1}{S} - 4 \frac{A_2}{aS} + \frac{A_3}{a^2 S}$$

Exercice 15 (Critique) - 1/ Jeter un regard critique : pour $h = a$, la formule de 6/ n'est pas compatible avec le résultat trouvé en 7/ (à part le premier terme $\frac{C}{n}$, qui est l'approximation de la statistique classique). Des deux principes d'approximation :

~ composition des termes de tranche et des termes de ligne (utilisé en 6/)

~ composition des variances d'extensions élémentaires (utilisé en 7/)

l'un au moins est inutilisable dans le cas des grandes mailles.

A la réflexion, il apparaît que la variance d'estimation associée à une maille rectangulaire (a,b), dès que a et b sont supérieurs à la portée, ne doit dépendre que de la surface ab du rectangle de maille, et non de son allongement (a/b). Or la formule 6/ fait intervenir explicitement ce rapport a/h. On peut conclure que le principe de composition des termes de tranche n'est pas applicable aux grandes mailles, et que la formule 7/ doit être préférée.

2/ Pour le vérifier de manière rigoureuse, partir de la formule générale :

$$\sigma_{Est}^2 = \frac{1}{S^2} \int_S \int_S K(x-y) dx dy - \frac{2}{NS} \sum_i \int_S K(x_i-y) dy + \frac{1}{N^2} \sum_{ij} K(x_i-x_j)$$

Dès que les distances $|x_i-x_j|$ entre sondages, ainsi que les distances des sondages à la frontière de S sont supérieures à la portée, ceci se réduit à la formule (rigoureuse)

$$\sigma_{Est}^2 = C - F(S) + \frac{C}{N} - \frac{2\pi A_1}{S}$$

formule valable indépendamment de la forme de la maille (et applicable, en particulier, même à des mailles irrégulières).

$C - F(S)$ ne dépend que de la surface S (et non de la maille) et a comme partie principale (pour S grand) : $\pi A_1/S$, d'où la formule d'approximation :

$$\sigma_{Est}^2 = \frac{C}{N} - \frac{\pi A_1}{S}$$

3/ Examinons si le principe de composition des variances d'extension élémentaires donne de meilleurs résultats. On trouve :

$$\frac{1}{N} [2 Q(\frac{a}{2}, \frac{b}{2}) - F(a,b)] = \frac{C}{N} - \pi \frac{A_1}{S} - \frac{2A_2}{S} (\frac{1}{a} + \frac{1}{b}) + \frac{A_3}{Sab}$$

Les deux premiers termes sont bons. Les termes d'ordre supérieur (en A_2 et A_3) ne sont pas acceptables, puisqu'ils dépendent encore de a et b , alors que la formule exacte ne fait intervenir que la géométrie de S . Néanmoins, cette formule est meilleure que σ (dont le premier terme seul est bon).

4/ Dans le même esprit qu'au 2/ ci-dessus, montrer que la variance d'estimation d'un segment de longueur L par une maille $a >$ portée est (rigoureusement) $C - F(L) + \frac{C}{N} - 2 \frac{A_0}{L}$.

En déduire que la variance d'estimation à une dimension (Ex. 13, paragraphe 2/) est en réalité :

$$\sigma_{\text{Est}}^2 = \frac{C}{N} - \frac{A_0}{L} - \frac{A_1}{L^2}$$

Ainsi, à une dimension, le principe de composition des variances d'extension donne la valeur exacte des deux premiers termes, mais non celle du troisième.

5/ Conclusion du paragraphe précédent que seuls les deux premiers termes de la formule du paragraphe 5/ Ex. 14 sont valables (variance d'estimation de S par des galeries parfaitement échantillonnées).

Pour reconstituer la formule exacte, on supposera que S est le rectangle $\ell \times H$, et on raisonnera directement sur le variogramme γ' déduit de γ par montée sous puissance ℓ . Pour h supérieur à la portée, la fonction $F'(h)$ de ce schéma transitif à 1 dimension est de la forme :

$$F'(h) = C' - \frac{A'_0}{h} + \frac{A'_1}{h^2}$$

(cf. Ex. 13). Montrer, par un raisonnement direct, que la nouvelle constante C' est :

$$C' = C - F(\ell, 0) = \frac{A_0}{\ell} - \frac{A_1}{\ell^2}$$

Pour déterminer les nouveaux coefficients A'_0 et A'_1 , procéder par identification avec $F(\ell, h)$ (Ex. 14, par. 2/), ce qui donne

$$A'_0 = \pi \frac{A_1}{\ell} - 2 \frac{A_2}{\ell^2}$$

$$A'_1 = 2 \frac{A_2}{\ell} - \frac{A_3}{\ell^2}$$

Montrer, en appliquant le résultat du 4/ ci-dessus, que la variance d'estimation pour des galeries de longueur ℓ équidistantes de $h = \frac{H}{N}$ est :

$$\begin{aligned} \sigma_{\text{Est}}^2 &= \frac{C'}{N} - \frac{A'_0}{H} - \frac{A'_1}{H^2} \\ &= \frac{A_0}{N\ell} - A_1 \left(\frac{1}{N\ell^2} + \frac{\pi}{S} \right) + 2 \frac{A_2}{S} \left(\frac{1}{\ell} - \frac{1}{H} \right) + \frac{A_3}{S^2} \end{aligned}$$

Comparer avec Ex. 14, par. 5/.

2-11-5 Inférence statistique pour les F.A.

Exercice 16 (Estimation d'un variogramme) - Soit $Y(x)$ une F.A.I. sur la droite \mathbb{R}^1 , $\gamma(h)$ son demi-variogramme. On suppose connue une réalisation de $Y(x)$ sur l'intervalle $(0, L)$, et, pour estimer $\gamma(h)$, on forme l'estimateur :

$$\gamma^*(h) = \frac{1}{2(L-h)} \int_0^{L-h} [Y(x+h) - Y(x)]^2 dx$$

a/ Montrer que γ^* est un estimateur sans biais; soit $E(\gamma^*(h)) = \gamma(h)$.

b/ (Lemme) Soient X_1, X_2, X_3 et X_4 quatre V.A. gaussiennes d'espérances nulles et σ_{ij} la matrice de leurs covariances. Montrer :

$$E(X_1 X_2 X_3 X_4) = \sigma_{12} \sigma_{34} + \sigma_{13} \sigma_{24} + \sigma_{14} \sigma_{23}$$

(identifier le terme en $u_1 u_2 u_3 u_4$ de la fonction caractéristique à 4 variables. On

peut aussi bien admettre ce lemme sans démonstration et passer à la suite de cet exercice).

c/ On suppose que les accroissements de la F.A.I. $Y(x)$ ont des lois gaussiennes (N.B. : cette hypothèse a simplement pour but de permettre le calcul des moments d'ordre 4). En utilisant le lemme b/, établir la relation :

$$E[(Y_{x+h} - Y_x)^2 (Y_{y+h} - Y_y)^2] = 4 \gamma^2(h) + 2 [\gamma(x-y+h) + \gamma(x-y-h) - 2\gamma(x-y)]^2$$

et en déduire la variance $D^2(\gamma^*(h))$ de l'estimateur $\gamma^*(h)$:

$$D^2(\gamma^*(h)) = \frac{\omega^2}{(L-h)^2} \int_0^{L-h} dx \int_0^x [\gamma(x-y+h) + \gamma(x-y-h) - 2\gamma(x-y)]^2 dy$$

(utiliser la symétrie en x et y de la fonction à intégrer).

d/ Effectuer les calculs explicites pour $\gamma(h) = \omega|h|$ (le processus $Y(y)$ est alors le mouvement brownien). Pour cela, montrer d'abord (avec $h \geq 0$) :

$$D^2(\gamma^*(h)) = \frac{4\omega^2}{(L-h)^2} \int_0^{L-h} dx \int_{\text{Sup}(0, (x-h))}^x [h-x+y]^2 dy$$

et distinguer deux cas : pour $h \leq L/2$, il vient :

$$D^2(\gamma^*(h)) = \left(\frac{4}{3} \frac{h^3}{L-h} - \frac{1}{3} \frac{h^4}{(L-h)^2} \right) \omega^2$$

et pour $h \geq L/2$, $D^2(\gamma^*(h)) = \left(2h^2 + \frac{1}{3}(L-h)^2 - \frac{4}{3}h(L-h) \right) \omega^2$

En particulier, pour h très petit, la variance relative est :

$$\frac{D^2(\gamma^*(h))}{(\gamma(h))^2} \approx \frac{4}{3} \frac{h}{L}$$

Elle tend vers 0 avec h , et par conséquent l'inférence statistique est possible dans des conditions acceptables en ce qui concerne le comportement de $\gamma(h)$ au voisinage de l'origine, et cela bien que l'on ne dispose que d'une réalisation unique.

Par contre, dès que h n'est plus très petit devant L , la variance relative devient

énorme et l'inférence statistique n'est plus possible : le variogramme expérimental γ^* doit "normalement" différer considérablement de son espérance $\gamma(h)$. Par exemple, pour $h = L/2$, la variance relative est égale à 1 (c'est-à-dire énorme).

Exercice 17 (Estimation de la variance $\sigma^2(0|L)$). - Dans les mêmes conditions que dans l'exercice précédent, on estime la variance dans L des échantillons ponctuels par l'estimateur :

$$S^2 = \frac{1}{L} \int_0^L [Y(x) - \bar{Y}]^2 dx$$

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

a/ mettre S^2 sous la forme $\frac{1}{L^3} \int_0^L \int_0^L \int_0^L (Y_x - Y_y)(Y_x - Y_{y'}) dx dy dy'$ et montrer $E(S^2) = \sigma^2(0|L)$.

b/ Calculer la variance $D^2(S^2)$ en supposant gaussiens les accroissements de $Y(x)$. Pour cela, on formera l'expression de S^4 , qui est une intégrale sextuple, on utilisera le lemme b/ de l'exercice 16 pour expliciter l'argument de cette intégrale, et on montrera que le résultat s'exprime à l'aide des fonctions auxiliaires χ et F sous la forme :

$$D^2(S^2) = 2(F(L))^2 + \frac{4}{L^2} \int_0^L (L-x) \gamma^2(x) dx - \frac{8}{L^3} \int_0^L x^2 \chi^2(x) dx - \frac{8}{L^3} \int_0^L x(L-x) \chi(x) \chi(L-x) dx$$

c/ Appliquer b/ au cas de $\gamma(h) = \omega|h|$, et montrer que la variance relative sur S^2 est égale à 4/5 (c'est-à-dire énorme) : les fluctuations de la variance expérimentale autour de son espérance ont toujours une ampleur énorme.

Exercice 18 (pseudo-covariance) - On a souvent l'habitude d'admettre, sans examen approfondi que les V.R. que l'on étudie peuvent être considérées comme des réalisations de F.A. stationnaires d'ordre 2 (alors que bien souvent une étude plus fine montre que seuls les accroissements de cette F.A. ont des moments d'ordre 2 - autrement dit, bien souvent, il n'existe pas de covariance $C(h)$, mais seulement un variogramme $\gamma(h)$). Or il se trouve que les procédés d'inférence statistique que l'on utilise ont systématiquement pour effet, par suite des biais qu'ils entraînent, de rendre cette hypothèse plausible, même

lorsqu'elle est tout-à-fait fausse.

Nous nous contenterons de montrer ce phénomène sur un exemple : celui du mouvement brownien, de demi-variogramme $\gamma(h) = |h|$, dont on connaît une réalisation sur l'intervalle $(0, L)$. On suppose donc :

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

Dans la pratique usuelle, pour estimer l'hypothétique covariance $C(h)$, qui en fait n'existe pas ici, on calcule successivement les valeurs (expérimentales) des expressions :

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

$$C^*(x, y) = (Y(x) - \bar{Y})(Y(y) - \bar{Y})$$

et on en déduit la "covariance expérimentale" :

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

Or, avec le demi-variogramme $\gamma(h) = |h|$, l'espérance de $C^*(x, y)$ est :

$$E[C^*(x, y)] = \frac{2}{3} L + \frac{x^2 + y^2}{L} - 2 \text{Sup}(x, y)$$

pour $h \geq 0$, on a donc

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

et, en intégrant en x , on voit que la "covariance expérimentale" admet l'espérance :

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

C'est une parabole, de pente $-\frac{4}{3}$ à l'origine. On trouvera donc une variance apparente $C(0) = \frac{1}{3} L$, liée uniquement à la longueur L de l'intervalle sur lequel on travaille, et qui constitue un pur artefact (puisque la variance réelle est infinie). Bien

qu'il n'existe pas de covariance stationnaire, mais seulement un variogramme linéaire, les biais qu'introduit ce procédé d'estimation ont pour effet de donner une confirmation apparente de l'existence de cette covariance. On note que la structure du phénomène est profondément déformée: non seulement la droite est remplacée par une parabole, mais même la pente à l'origine se trouve affectée ($\frac{4}{3}$ au lieu de 1). Le $C^*(h)$ est un pur artefact, et ne conserve à peu près plus rien de la structure réelle du processus.

On notera que le variogramme expérimental :

$$2 \gamma^*(h) = \frac{1}{L-h} \int_0^{L-h} [Y(x+h) - Y(x)]^2 dx$$

a comme espérance

$$E[\gamma^*(h)] = \gamma(h) = |h|$$

il n'est pas affecté par le biais précédent, et constitue par conséquent pour l'expérimentateur un outil plus sûr que la covariance.

CHAPITRE 3

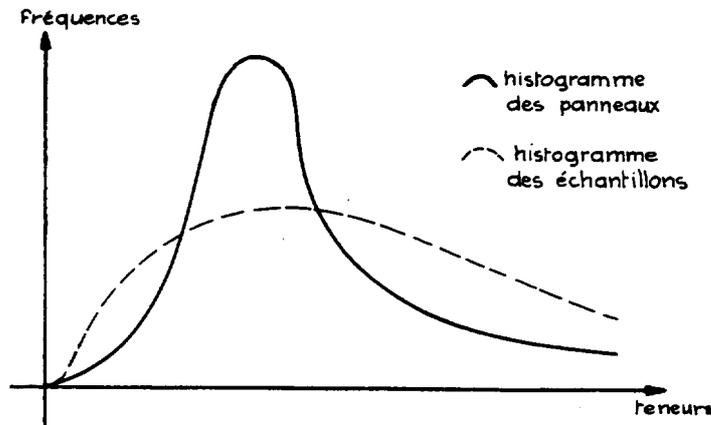
LE KRIGEAGE

3-1 LES BUTS DU KRIGEAGE

En termes miniers, le problème du krigeage consiste à trouver la meilleure estimation linéaire possible de la teneur d'un panneau, compte tenu de l'information disponible, c'est-à-dire des teneurs des différents échantillons qui ont été prélevés, soit à l'intérieur, soit à l'extérieur du panneau que l'on veut estimer. Le krigeage revient à effectuer une pondération, c'est-à-dire à attribuer un poids à la teneur de chaque échantillon, ces poids étant calculés de manière à rendre minimale la variance d'estimation résultante, compte tenu des caractéristiques géométriques du problème (formes, dimensions et implantation relative du panneau et des échantillons). En gros, comme cela est naturel, le krigeage attribuera des poids faibles aux échantillons éloignés, et inversement. Mais cette règle intuitive peut être partiellement mise en défaut lorsqu'apparaissent des phénomènes plus complexes d'effet d'écran et de transfert d'influence. Il n'est naturellement possible de résoudre un problème de krigeage, c'est-à-dire de calculer effectivement le poids optimal qu'il convient d'attribuer à chaque échantillon, qu'à la condition de faire certaines hypothèses sur les caractéristiques géostatistiques du gisement que l'on étudie, c'est-à-dire, essentiellement, de se donner la fonction de covariance ou le variogramme de la F.A. dont les teneurs ponctuelles sont supposées constituer une réalisation. Il n'est pas nécessaire, en principe, d'introduire une hypothèse stationnaire ou intrinsèque, et les équations du krigeage ont une portée générale. (en pratique, naturellement, il faut commencer par estimer la covariance ou le variogramme à partir des données expérimentales, et c'est ici que l'hypothèse en question retrouve son importance). Nous aborderons, dans le chapitre suivant, le cas général des F.A. non stationnaires, et nous nous limiterons pour l'instant aux F.A. intrinsèques ou stationnaires d'ordre 2.

Le premier intérêt du krigeage découle de sa définition même. En minimisant la variance d'estimation, on est sûr de tirer le meilleur parti possible des informations disponibles ou, si l'on préfère, d'obtenir l'estimation la plus précise possible du panneau en cause. Cet avantage est souvent notable, mais ne justifierait cependant pas toujours les complications supplémentaires qu'introduit nécessairement une pondération. L'intérêt pratique le plus important - et de loin - du krigeage provient non pas de ce qu'il assure la meilleure précision possible, mais de ce qu'il permet d'éviter une erreur systématique. Dans la plupart des gisements métalliques, on est conduit à sélectionner pour l'exploitation un certain nombre de panneaux jugés rentables et à en abandonner

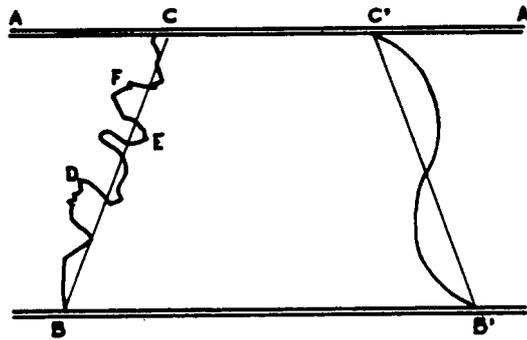
d'autres jugés inexploitable. D.G. KRIGE a montré que, si cette sélection était faite au seul vu des échantillons intérieurs à chaque panneau, il en résultait nécessairement - en moyenne - une surestimation de la teneur des panneaux sélectionnés. La raison de ce phénomène très général est que la variance des teneurs réelles des panneaux est toujours plus faible que la variance des résultats d'un échantillonnage intérieur. Autrement dit, l'histogramme des teneurs réelles des panneaux comporte toujours moins de teneurs extrêmes (riches ou pauvres) et davantage de teneurs intermédiaires que l'histogramme déduit des échantillonnages intérieurs, et, si l'on calcule l'effet d'une sélection sur ce dernier histogramme, les panneaux éliminés seront en réalité moins pauvres qu'on ne l'avait prévu, et les panneaux conservés moins riches.



Notre notion de krigeage permet d'interpréter facilement ce phénomène, et d'en corriger les effets. Du fait même que l'on sélectionne un panneau riche, l'auréole des échantillons extérieurs est en général à teneur plus faible que les échantillons intérieurs, et cependant son influence sur le panneau à estimer n'est pas négligeable, puisque le krigeage lui attribue un poids non nul. En ne tenant pas compte de cette auréole extérieure, on introduit nécessairement une cause d'erreur systématique par excès. Pour illustrer cette notion, imaginons un gisement filonien reconnu par deux traçages AA' et BB'.

Les teneurs de BB' ont été jugées exploitables, tandis que celles de AA' n'ont été jugées intéressantes que sur un segment CC'. Si l'on se contente de faire la moyenne des teneurs de BB' et de CC', on est sûr de commettre une erreur par excès : car les tronçons AC et C'A' - pauvres, par construction - ont une influence non négligeable sur le trapèze BB', CC' que l'on se propose d'exploiter. S'il était possible de tracer une frontière précise de teneur entre minerais exploitables et inexploitable, la frontière réelle, en général, ne coïnciderait pas avec les droites BC ou B'C', mais serait une courbe quelconque, le plus souvent très irrégulière, telle que la ligne brisée BDEF... De plus, des enclaves pauvres subsisteraient dans le panneau riche, et réciproque-

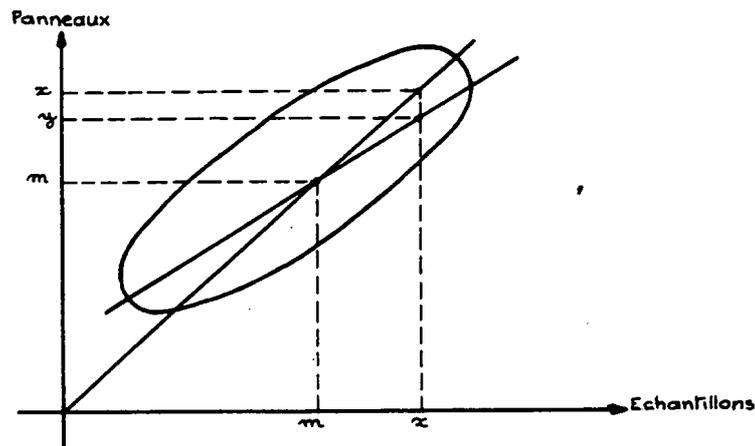
ment. A l'exploitation, on serait contraint d'abandonner certaines portions riches, et de prendre certaines parties pauvres, et ce salissage traduit de façon concrète l'influence des tronçons pauvres AC et C'A' sur le panneau riche (comme il s'agit d'un gisement homogène, et que l'existence même d'une frontière précise de teneur est douteuse, nous parlons de krigeage, et non de salissage. Nous employons le mot salissage dans le cas où les minerais riches et pauvres sont géologiquement hétérogènes, et séparés par une frontière continue représentable par une courbe relativement régulière, telle que B', D'E'...., qui n'a d'ailleurs pas non plus de raison de coïncider avec la droite B'C'). En pareil cas, le krigeage conduit à pondérer les teneurs de BB', CC' et des tronçons



AC, C'A' par des coefficients convenables, qui seraient - par exemple - 60% pour BB', 27% pour CC' et 13% pour les deux tronçons.

L'effet essentiel d'une telle pondération est d'éliminer - en moyenne - une erreur systématique par excès, donc particulièrement redoutable. Vis-à-vis de cet objectif primordial, l'amélioration de la précision proprement dite apparaît comme relativement secondaire.

Donnons, maintenant, quelques indications sur la manière dont D.G. KRIGE a formulé le problème au début des années 50. Depuis longtemps, les mineurs d'Afrique du Sud connaissaient cet effet de surestimation des panneaux riches, et appliquaient des coefficients correcteurs empiriques. Pour retrouver ces coefficients, KRIGE part de l'hypothèse que l'échantillonnage est bien fait, autrement dit que l'espérance mathématique de la teneur des échantillons que l'on prélève dans un panneau est égale à la teneur moyenne réelle de ce panneau. Si ces variables sont gaussiennes (dans le cas de l'Afrique du Sud, elles étaient en réalité lognormales, mais cela ne change rien d'essentiel), la droite de régression donnant l'espérance conditionnelle de l'échantillon en fonction de la teneur du panneau est donc identique à la première bissectrice.



L'autre droite de régression, celle qui donne l'espérance du panneau en fonction de l'échantillon (et c'est celle-là qui est intéressante) a donc une pente inférieure à l'unité. Si la teneur x de l'échantillon est supérieure à la moyenne générale m , l'espérance y du panneau est alors inférieure à x , et inversement. KRIGE corrigeait donc cette erreur systématique en utilisant l'équation

$$(3-1) \quad y = m + \beta(x-m)$$

de cette seconde droite de régression, avec un coefficient de régression $\beta < 1$. Plus précisément, si ρ est le coefficient de corrélation, σ_x et σ_y les écarts type des échantillons et des panneaux, les deux coefficients de régression (1 et β respectivement) sont

$$1 = \rho \frac{\sigma_x}{\sigma_y}, \quad \beta = \rho \frac{\sigma_y}{\sigma_x}$$

et par suite :

$$\beta = \frac{\sigma_y^2}{\sigma_x^2} < 1$$

Le coefficient β est égal au rapport des variances (dans le gisement entier) des panneaux et des échantillons.

Un examen plus attentif de (3-1) suggère d'autres possibilités. La teneur moyenne m n'est pas connue, en général, et on l'estime en prenant la moyenne arithmétique $m^* = \frac{1}{n} \sum X_i$ des échantillons disponibles dans le gisement. Mais si l'on remplace m par m^* , (3-1) prend l'aspect :

$$Y^* = \sum_1 a_1 X_1$$

d'une combinaison linéaire des teneurs X_1 des échantillons disponibles, avec d'ailleurs la condition :

$$\sum a_1 = 1$$

sur laquelle nous reviendrons longuement. Mais alors, au lieu d'attribuer aux échantillons extérieurs le même poids $a_1 = \frac{1-\beta}{n}$ indifférencié, qu'il s'agisse d'échantillons proches ou lointains, il est naturel de chercher à attribuer à chacun d'eux un poids a_1 approprié, tenant compte de sa localisation par rapport au panneau à estimer. On arrive ainsi à la recherche des poids optimaux a_1 , c'est-à-dire à la notion du krigeage tel que nous l'avons présentée en premier lieu.

3-2 Notations

Dans tout ce chapitre, et dans le chapitre suivant, nous utiliserons le même système de notations :

~ $Z(x)$ désignera une F.A. (stationnaire ou non), en général d'espérance non nulle, et $Y(x)$ une F.A. (pas nécessairement stationnaire ni intrinsèque) d'espérance nulle : par exemple, on prendra souvent

$$Y(x) = Z(x) - E(Z(x))$$

~ On désignera par S l'ensemble de \mathbb{R}^n sur lequel la réalisation de la F.A. est connue expérimentalement. Dans le cas où S est un ensemble fini, on désignera par des indices grecs les points expérimentaux $x_\alpha, x_\beta \dots$ de S , soit :

$$S = \{x_\alpha, \alpha = 1, 2, \dots, N\}$$

~ Suivant les cas, on cherchera à estimer soit la valeur $Z(x_0)$ de la F.A. en un point $x_0 \notin S$, soit la moyenne $\frac{1}{V} \int_V Z(x) dx$ dans un domaine V différent de S , soit plus généralement (et pour réunir tous les cas possibles en une écriture unique) une moyenne

$$Z_0 = \int p(dx) Z(x) \quad \left(\int p(dx) = 1 \right)$$

$Z(x)$ pondérée par une mesure p de somme unité et de support disjoint de S .

~ pour cela on formera des estimateurs linéaires à partir des données disponibles, qui seront les $Z(x)$, $x \in S$. Dans le cas discret, on posera toujours

$$Z_{\alpha} = Z(x_{\alpha}) \quad (x_{\alpha} \in S)$$

et de même, pour des fonctions $f(x)$, $C(x,y)$ etc... :

$$f_{\alpha} = f(x_{\alpha}) \quad , \quad C(x_{\alpha}, y_{\beta}) = C_{\alpha\beta} \quad \text{etc...}$$

Dans le cas discret, on appliquera toujours la convention de sommation suivant laquelle on doit sommer sur tout indice apparaissant deux fois (en général, une fois en position inférieure, ou covariante, et une fois en position supérieure, ou contravariante). Ainsi, on écrira $\lambda^{\alpha} Z_{\alpha}$ au lieu de $\sum_{\alpha} \lambda^{\alpha} Z_{\alpha}$. Les estimateurs que nous utiliserons s'écriront donc toujours :

$$Z^* = \lambda^{\alpha} Z_{\alpha}$$

Dans le cas où S est infini, on utilisera pour former les estimateurs des mesures λ à support dans \underline{S} , soit

$$Z^* = \int_S \lambda(dx) Z(x)$$

on écrira même, en général, simplement $\int \lambda(dx) Z(x)$, puisque la mesure λ est à support dans S par définition.

Posons maintenant les équations du krigeage, en distinguant 3 cas (F.A. stationnaire d'espérance nulle ou connue ; F.A. stationnaire d'espérance inconnue ; F.A. intrinsèque sans dérive). Dans tous les cas nous supposons connus la covariance ou le variogramme.

3-3 F.A. STATIONNAIRE D'ESPERANCE NULLE OU CONNUE A PRIORI.

Soit $Y(x)$ une F.A. d'espérance nulle, $\sigma(x;y)$ sa covariance, $S = \{x_{\alpha}\}$ l'ensemble des points expérimentaux, que nous supposons d'abord fini, $Y_0 = \int p(dx) Z(x)$ la variable à estimer.

Comme estimateur, on va utiliser une combinaison linéaire :

$$Y_K = \lambda_K^\alpha Y_\alpha$$

et déterminer les coefficients λ_K^α par la condition de minimiser l'espérance de $(Y_0 - Y_K)^2$, c'est-à-dire (puisque l'espérance de la F.A. est nulle) la variance $D^2(Y_0 - Y_K)$. Or cette variance est :

$$E(Y_0 - Y_K)^2 = D^2(Y_0) - 2 \lambda_K^\alpha \sigma_{\alpha Y_0} + \lambda_K^\alpha \lambda_K^\beta \sigma_{\alpha\beta}$$

avec :

$$\begin{cases} \sigma_{\alpha Y_0} = \int p(dx) \sigma(x_\alpha, x) \\ D^2(Y_0) = \iint p(dx) \sigma(x, y) p(dy) \end{cases}$$

En annulant les dérivées partielles en λ_K^α de cette forme quadratique, on obtient le système :

$$(3-2) \quad \lambda_K^\beta \sigma_{\alpha\beta} = \sigma_{\alpha Y_0}$$

Ce système de N équations à N inconnues est régulier, et admet une solution unique, si et seulement si la matrice des covariances $\sigma_{\alpha\beta}$ est strictement définie positive (donc de déterminant > 0), ce que nous supposons toujours. La variance σ_K^2 (ou variance de krigeage) de cette estimation optimale est égale à la valeur de la forme quadratique $E(Y_0 - Y_K)^2$ lorsque l'on prend comme coefficients λ_K^α la solution de (3-2). Or (3-2) entraîne :

$$\lambda_K^\alpha \lambda_K^\beta \sigma_{\alpha\beta} = \lambda_K^\alpha \sigma_{\alpha Y_0}$$

Il y a donc, à l'optimum, égalité du terme rectangle et du terme quadratique, et on trouve :

$$(3-3) \quad \sigma_K^2 = D^2(Y_0) - \lambda_K^\alpha \sigma_{\alpha Y_0}$$

Dans le cas du krigeage ponctuel (estimation de $Y(x_0)$ pour un point $x_0 \notin S$) le système se réduit à :

$$(3-4) \quad \begin{cases} \lambda_K^\beta \sigma_{\alpha\beta} = \sigma_{\alpha, x_0} \\ \sigma_K^2 = \sigma_{x_0, x_0} - \lambda_K^\alpha \sigma_{\alpha, x_0} \end{cases}$$

La solution $\lambda_K^\alpha(x_0)$ dépend évidemment de x_0 , et l'estimateur correspondant :

$$Y_K(x_0) = \lambda_K^\alpha(x_0) Y_\alpha$$

est un interpolateur exact en ce sens que $Y_K(x_{\alpha_0}) = Y_{\alpha_0}$ si x_0 vient à coïncider avec un point expérimental $x_{\alpha_0} \in S$: on peut le vérifier directement à l'aide du système (3-4), mais il est évident a priori que l'estimateur Y_{α_0} est optimal, puisqu'il est de variance nulle.

Lorsque l'ensemble S est infini, on cherche à estimer Y_0 au moyen d'un estimateur de la forme :

$$(3-5) \quad Y_K = \int_S \lambda_K(dx) Y(x)$$

et le système (3-2) devient :

$$(3-6) \quad \begin{cases} \forall y \in S & \int_S \lambda_K(dx) \sigma(x, y) = \sigma_{y, Y_0} \\ \sigma_K^2 = D^2(Y_0) - \int_S \lambda_K(dx) \sigma_{x, Y_0} \end{cases}$$

Mais ici, quelques réserves s'imposent. S'il existe une mesure λ_K à support sur S vérifiant (3-6), cette solution est unique, et (3-5) est l'estimateur optimal. Mais il n'existe pas nécessairement de mesure λ_K (du moins lorsque la covariance est très régulière, cf. exercice 8) vérifiant ce système. On montre (cf. [6]) par contre qu'il existe toujours un estimateur optimal unique Y_K appartenant à l'espace de Hilbert $H(S)$ engendré par les $Y(x)$, $x \in S$ (c'est-à-dire : limite en moyenne quadratique de combinaison linéaires finies des $Y(x)$, $x \in S$), mais Y_K n'admet pas nécessairement une représentation de la forme (3-5) avec une mesure λ_K .

On peut éclairer ce résultat en remarquant que les systèmes (3-4) ou (3-6) peuvent s'écrire :

$$(3-7) \quad \begin{cases} \text{Cov}(Y_K - Y_0, Y(y)) = 0 & (\forall y \in S) \\ \sigma_K^2 = D^2(Y_0) - \text{Cov}(Y_K, Y_0) \end{cases}$$

Ainsi Y_K est l'unique élément de l'espace de Hilbert $H(S)$ tel que $Y_K - Y_0$ soit orthogonal à tous les éléments $Y \in H(S)$ (c'est-à-dire en covariance nulle avec tout $Y \in H$). Cette propriété caractérise Y_K comme la projection orthogonale de Y_0 dans $H(S)$, d'où l'existence et l'unicité, mais cela n'implique évidemment pas (sauf dans le cas fini où H est un espace euclidien) l'existence d'une représentation intégrale de la forme (3-5).

Soit maintenant $Z(x)$ une F.A. admettant une covariance centrée $\sigma(x,y)$ et une espérance constante $m = E(Z(x))$, non nulle, mais connue. On se ramène immédiatement au cas précédent en raisonnant sur la F.A. $Y(x) = Z(x) - m$. L'estimateur optimal est donc :

$$(3-8) \quad Z_K = m + \lambda_K^\alpha (Z_\alpha - m)$$

ou :

$$(3-8') \quad Z_K = m + \int_S \lambda_K(dx) (Z(x) - m)$$

avec des coefficients λ_K^α ou une mesure λ_K vérifiant les mêmes systèmes (3-2) ou (3-6), et la variance σ_K^2 conserve la même expression.

3-4 FONCTION ALEATOIRE STATIONNAIRE D'ESPÉRANCE INCONNUE.

3-4-1 Les équations du Krigeage

Soit maintenant $Z(x)$ une F.A. d'espérance m constante mais inconnue, et $\sigma(x,y)$ sa covariance centrée. On veut estimer $Z_0 = \int p(dx) Z(x)$ à partir des Z_α , valeurs de la réalisation sur un ensemble fini $S = \{x_\alpha, \alpha = 1, 2, \dots, N\}$ au moyen d'une combinaison linéaire de la forme :

$$Z^* = \lambda^\alpha Z_\alpha$$

Du fait que l'espérance m n'est pas connue, il est nécessaire d'imposer aux coefficients λ^α la condition dite d'universalité :

$$(3-9) \quad \sum_\alpha \lambda^\alpha = 1$$

En effet, la meilleure combinaison linéaire possible est celle qui minimise l'espérance de $(Z_0 - Z^*)^2$.

Or on a :

$$\begin{aligned} E(Z_0 - Z^*)^2 &= E(Z_0^2) - 2 E(Z_0 Z^*) + E(Z^*)^2 = \\ &= m^2(1 - \sum_{\alpha} \lambda^{\alpha})^2 + D^2(Z_0) - 2 \lambda^{\alpha} \sigma_{\alpha Z_0} + \lambda^{\alpha} \lambda^{\beta} \sigma_{\alpha\beta} \end{aligned}$$

Comme m est inconnue, on ne peut minimiser cette expression que si elle ne dépend pas de m , c'est-à-dire si (3-9) est vérifiée.

On peut encore justifier cette condition (3-9) en imposant à l'estimateur Z^* utilisé d'être sans biais quelle que soit la valeur (inconnue) de m (estimateur universel). Or

$$E(Z_0 - Z^*) = m - m \sum_{\alpha} \lambda^{\alpha}$$

et, à nouveau, cette espérance est nulle quelle que soit m si et seulement si (3-9) est vérifiée.

Lorsque la condition d'universalité (3-9) est vérifiée, $E(Z_0 - Z^*)$ est nulle, et par suite

$$E(Z_0 - Z^*)^2 = D^2(Z_0 - Z^*) = D^2(Z_0) - 2 \lambda^{\alpha} \sigma_{\alpha, Z_0} + \lambda^{\alpha} \lambda^{\beta} \sigma_{\alpha\beta}$$

En exprimant que cette forme quadratique est minimale compte tenu de la condition (3-9), on obtient le système suivant où figure un paramètre μ de Lagrange :

$$(3-10) \quad \begin{cases} \lambda^{\beta} \sigma_{\alpha\beta} = \sigma_{\alpha, Z_0} + \mu \\ \sum_{\alpha} \lambda^{\alpha} = 1 \end{cases}$$

En multipliant la première équation (3-9) et en tenant compte de la seconde, on trouve ensuite :

$$\lambda^{\alpha} \lambda^{\beta} \sigma_{\alpha\beta} = \lambda^{\alpha} \sigma_{\alpha, Z_0} + \mu \sum_{\alpha} \lambda^{\alpha} = \lambda^{\alpha} \sigma_{\alpha, Z_0} + \mu$$

D'où l'expression de la variance du krigeage, qui fait intervenir le paramètre de Lagrange :

$$(3-10') \quad D^2(Z^* - Z_0) = D^2(Z_0) + \mu - \lambda^\alpha \sigma_{\alpha, Z_0}$$

Dans le cas du krigeage ponctuel d'un point x_0 , ce système devient :

$$(3-11) \quad \begin{cases} \lambda^\beta \sigma_{\alpha\beta} = \sigma_{\alpha, x_0} + \mu \\ \sum_{\alpha} \lambda^\alpha = 1 \\ D^2(Z(x_0) - Z^*) = \sigma_{x_0 x_0} + \mu - \lambda^\alpha \sigma_{\alpha, Z_0} \end{cases}$$

Ici encore on vérifie directement que le krigeage ponctuel est un interpolateur exact.

On démontre que le système (3-10) est toujours régulier. Dans le cas continu, il existe toujours un élément unique $Z^* \in H(S)$ limite en moyenne quadratique de combinaisons linéaires finies vérifiant la condition (3-9) et tel que l'on ait :

$$\text{Cov}(Z^* - Z_0, Z(y)) = 0 \quad \forall y \in S$$

Cet élément unique est la projection de Z_0 dans la variété linéaire fermée définie dans $H(S)$ par la condition d'universalité, d'où l'existence et l'unicité de l'estimateur optimal Z^* . Ici encore, Z^* n'admet pas forcément une représentation de la forme $\int_S \lambda(dx) Z(x)$ avec une mesure λ vérifiant :

$$(3-12) \quad \begin{cases} \int_S \lambda(dx) \sigma(x, y) = \sigma_{y, Z_0} + \mu \quad \forall y \in S \\ \int_S \lambda(dx) = 1 \end{cases}$$

mais, si l'on peut trouver une mesure λ vérifiant (3-12), alors $Z^* = \int_S \lambda(dx) Z(x)$ est l'unique solution du problème.

3-4-2 L'estimation optimale de m.

Au lieu d'estimer une moyenne spatiale du type $\int p(dx) Z(x)$, on peut aussi chercher à estimer l'espérance $m = E(Z(x))$ elle-même. Nous allons former l'estimateur optimal m^* de m ,

et, dans le paragraphe suivant, nous interroger sur le rapport existant entre m^* et Z^* . Nous nous limiterons au cas où S est fini, l'extension au cas infini se faisant ensuite d'elle-même, avec les réserves d'usage quant à l'existence d'une représentation de nos estimateurs au moyen de mesures à support dans S .

Pour estimer m , on forme une combinaison linéaire

$$m^* = \lambda_0^\alpha Z_\alpha$$

On impose aux coefficients λ_0^α la condition d'universalité

$$\sum \lambda_0^\alpha = 1$$

qui exprime que m^* est sans biais quelle que soit la valeur inconnue de m , et on choisit les coefficients λ_0^α qui minimisent $E(m-m^*)^2 = D^2(m^*)$ compte tenu de cette condition. De :

$$D^2(m-m^*) = \lambda_0^\alpha \lambda_0^\beta \sigma_{\alpha\beta}$$

on déduit que les λ_0^α constituent l'unique solution du système suivant, où figure un paramètre de Lagrange μ_0 :

$$(3-13) \quad \begin{cases} \lambda_0^\beta \sigma_{\alpha\beta} = \mu_0 \\ \sum_\alpha \lambda_0^\alpha = 1 \end{cases}$$

A l'optimum, on trouve $\lambda^\alpha \lambda^\beta \sigma_{\alpha\beta} = \mu_0 \sum_\alpha \lambda^\alpha = \mu_0$, de sorte que le paramètre de Lagrange μ_0 est égal à la variance de l'estimateur optimal

$$(3-13') \quad D^2(m^*) = \mu_0$$

3-4-3 Le théorème d'additivité.

A la fin du paragraphe 3-3, nous avons indiqué comment former l'estimateur optimal lorsque l'espérance m n'est pas nulle mais se trouve connue : on calcule les coefficients λ_K^α

optimaux du cas $m = 0$, et on effectue la correction très simple qui consiste à remplacer $Y(x)$ par $Z(x) - m$, soit :

$$(3-8) \quad Z_K = m + \lambda_K^\alpha (Z_\alpha - m)$$

Montrons que l'estimateur optimal Z^* du cas où m n'est pas connue (par. 3-4-1) et l'estimateur optimal m^* de m elle-même sont liés par la relation suivante obtenue en démarquant (3-8) :

$$(3-14) \quad Z^* = m^* + \lambda_K^\alpha (Z_\alpha - m^*)$$

avec les mêmes coefficients λ_K^α , solution du système (3-2). Ce résultat signifie qu'on a le droit de kriger comme si m était connue, à condition de remplacer la valeur inconnue de m par son estimation optimale m^* .

En effet, considérons le second membre de (3-14) :

$$m^* + \lambda_K^\alpha (Z_\alpha - m^*) = [\lambda_K^\alpha + \lambda_0^\alpha (1 - \sum_\beta \lambda_K^\beta)] Z_\alpha$$

Les quantités :

$$\lambda'^\alpha = \lambda_K^\alpha + \lambda_0^\alpha (1 - \sum_\beta \lambda_K^\beta)$$

vérifient la condition d'universalité, car :

$$\sum_\alpha \lambda'^\alpha = \sum_\alpha \lambda_K^\alpha + \sum_\alpha \lambda_0^\alpha (1 - \sum_\beta \lambda_K^\beta) = 1$$

compte tenu de $\sum_\alpha \lambda_0^\alpha = 1$ (système 3-13)). Formons ensuite l'expression :

$$\lambda'^\beta \sigma_{\alpha\beta} = \lambda_K^\beta \sigma_{\alpha\beta} + (1 - \sum_\gamma \lambda_K^\gamma) \lambda_0^\beta \sigma_{\alpha\beta}$$

D'après (3-2) et (3-13), il vient :

$$\lambda'^\beta \sigma_{\alpha\beta} = \sigma_{\alpha, Z_0} + (1 - \sum_\gamma \lambda_K^\gamma) \mu_0$$

Donc, les λ'^β vérifient bien la première relation (3-11), avec le paramètre de Lagrange

$$(3-15) \quad \mu = (1 - \sum_{\gamma} \lambda_K^{\gamma}) \mu_0$$

D'après l'unicité de la solution, on a donc bien $\lambda^{\alpha} = \lambda^{\alpha}$, et la relation (3-14) en résulte.

En ce qui concerne les variances, on a également un théorème d'additivité. En effet, d'après (3-14), on a :

$$Z^* - Z_0 = (\lambda_K^{\alpha} Z_{\alpha} - Z_0) + (1 - \sum_{\beta} \lambda_K^{\beta}) m^*$$

Or le système (3-2) exprime précisément que $\lambda_K^{\alpha} Z_{\alpha} - Z_0$ est en covariance nulle avec tous les Z_{β} , donc aussi avec leurs combinaisons linéaires, et, en particulier, avec m^* . On a donc :

$$D^2(Z^* - Z_0) = D^2(\lambda_K^{\alpha} Z_{\alpha} - Z_0) + (1 - \sum_{\beta} \lambda_K^{\beta})^2 D^2(m^*)$$

c'est-à-dire

$$(3-16) \quad D^2(Z^* - Z_0) = \sigma_K^2 + (1 - \sum_{\beta} \lambda_K^{\beta})^2 D^2(m^*)$$

Le premier terme est la variance σ_K^2 du krigeage lorsque m est connue. Le second donne une mesure exacte de la perte d'information que l'on subit relativement à Z_0 faute de connaître la vraie valeur de m .

3-5 CAS D'UNE F.A.I. SANS COVARIANCE.

Soit maintenant $Z(x)$ une F.A.I. sans dérive, admettant un variogramme γ mais pas de covariance. Pour estimer $Z_0 = \int p(dx) Z(x)$ avec $\int p(dx) = 1$ nous cherchons une combinaison linéaire

$$Z^* = \lambda^{\alpha} Z_{\alpha}$$

telle que :

a/ l'erreur $Z^* - Z_0$ soit une combinaison linéaire autorisée (paragraphe 2-2-1), c'est-à-

dire admettant une variance finie.

b/ telle que cette variance d'estimation soit minimale.

La condition a/ donne $\sum_{\alpha} \lambda^{\alpha} - \int p(dx) = 0$, c'est-à-dire la même condition d'universalité que dans le paragraphe précédent :

$$\sum_{\alpha} \lambda^{\alpha} = 1$$

Lorsque cette condition est vérifiée, on sait, d'après le mécanisme de calcul du paragraphe (2-2-1) que l'on peut calculer la variance de $Z^* - Z_0$ comme s'il existait une covariance égale à $-\gamma$. On peut donc transposer directement le système (3-11) : les coefficients λ^{α} de l'estimateur optimal constituent l'unique solution du système :

$$(3-17) \quad \begin{cases} \lambda^{\beta} \gamma_{\alpha\beta} = \int p(dx) \gamma(x, x_{\alpha}) - \mu \\ \sum_{\alpha} \lambda^{\alpha} = 1 \end{cases}$$

et la variance σ_K^2 correspondante est :

$$(3-17') \quad \sigma_K^2 = - \iint p(dx) \gamma(x, y) p(dy) + \mu + \lambda^{\alpha} \int p(dx) \gamma(x, x_{\alpha})$$

Dans le cas continu (avec les restrictions d'usage concernant l'existence d'une représentation de Z^* à l'aide d'une mesure) on a de même :

$$(3-18) \quad \begin{cases} \int_S \lambda(dx) \gamma(x, y) = \int p(dx) \gamma(x, y) - \mu & (\forall y \in S) \\ \int_S \lambda(dx) = 1 \\ \sigma_K^2 = - \iint p \gamma p + \mu + \iint p \gamma \lambda \end{cases}$$

Dans le cas ponctuel, enfin, on obtient encore un interpolateur exact.

Remarque - Soit $\lambda^{\alpha}(x_0)$ la solution du krigeage du point x_0 :

$$(3-19) \quad \begin{cases} \lambda^\beta(x_0) \sigma_{\alpha\beta} = \gamma(x_0, x_\alpha) - \mu(x_0) \\ \sum_\alpha \lambda^\alpha(x_0) = 1 \\ \sigma_K^2(x_0) = \mu(x_0) + \lambda^\alpha(x_0) \gamma(x_0, x_\alpha) \end{cases}$$

D'après le caractère linéaire des seconds membres de (3-17), la solution du krigeage de $\int p(dx)Z(x)$ est :

$$\lambda^\alpha = \int p(dx) \lambda^\alpha(x)$$

Il y a superposition, ou combinaison linéaire des krigeages ponctuels : cette relation s'étend d'ailleurs au paramètre de Lagrange :

$$\mu = \int p dx \mu(x)$$

Par contre la variance (3-17') ne s'obtient pas par combinaison linéaire des variances $\sigma_K^2(x_0)$ des krigeages ponctuels.

3-6 EXERCICES SUR LE KRIGEAGE.

Exercice 1 (Krigage d'un segment de longueur ℓ) a/ On se donne sur la droite une F.A.I. sans dérive dont le variogramme est $\gamma(h)$ et quatre points de prélèvements $x_1, x_2 = x_1 + \ell, x_3 = x_2 + \ell, \text{ et } x_4 = x_3 + \ell$. On veut kriger le segment (x_2, x_3) de longueur ℓ à partir des valeurs Z_1, Z_2, Z_3, Z_4 de la réalisation en x_1, x_2, x_3, x_4 . Expliciter le système (3-17) en utilisant les fonctions auxiliaires χ et F du paragraphe 2-5-2. [montrer $\lambda_1 = \lambda_4 = \frac{\lambda}{2}, \lambda_2 = \lambda_3 = \frac{1-\lambda}{2}$, ce qui est évident par symétrie, avec λ solution de :

$$\frac{\lambda}{2} \gamma(\ell) + \frac{1-\lambda}{2} \gamma(\ell) + \frac{\lambda}{2} \gamma(2\ell) = \chi(\ell) + \mu$$

$$\frac{1-\lambda}{2} \gamma(\ell) + \frac{1-\lambda}{2} \gamma(2\ell) + \frac{\lambda}{2} \gamma(3\ell) = 2\chi(2\ell) - \chi(\ell) + \mu$$

et éliminer le paramètre de Lagrange μ , d'où :

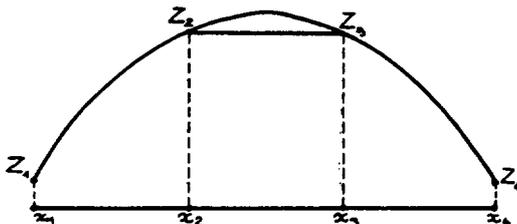
$$\lambda \left[\frac{1}{2} \gamma(3\ell) - \gamma(2\ell) - \frac{1}{2} \gamma(\ell) \right] = 2 \chi(2\ell) - 2 \chi(\ell) - \frac{1}{2} \gamma(2\ell)$$

b/ appliquera au cas $\gamma(h) = |h|^\alpha$ ($0 \leq \alpha < 2$) et interpréter.

(Solution : utiliser l'exercice 5 du chapitre 2. On trouve

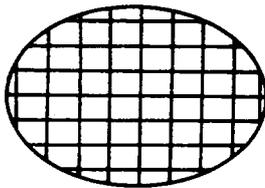
$$\lambda = \frac{2^\alpha - \frac{4}{\alpha+1} (2^\alpha - 1)}{1 + 2^{\alpha+1} - 3^\alpha}$$

Pour $\alpha = 0$, $\lambda = \frac{1}{2}$ (effet de pépite pur, l'estimateur optimal est la moyenne arithmétique des 4 échantillons). Lorsque α augmente, λ décroît, s'annule pour $\alpha = 1$ (propriété markovienne du variogramme linéaire) et devient négatif pour $1 < \alpha < 2$: pour $\alpha > 1$, la réalisation présente de bonnes propriétés de continuité et le dessin ci-dessous fait comprendre que, pour $\frac{Z_1 + Z_4}{2} < \frac{Z_2 + Z_3}{2}$ (par exemple), l'estimateur doit être supérieur à $\frac{Z_2 + Z_3}{2}$, donc $\lambda < 0$:



(Comparer avec l'exercice 12 du Ch. 4)

Exercice 2 (Gisement reconnu par traçages et montages) - Soit un gisement reconnu par deux réseaux réguliers de galeries (traçages et montages) respectivement parallèles à deux direc-



tions perpendiculaires. On désigne par Z_M et Z_T les teneurs moyennes des montages et des traçages, par Z la teneur inconnue du gisement et par $\sigma_M^2 = D^2(Z - Z_M)$ et $\sigma_T^2 = D^2(Z - Z_T)$ les variances d'estimation du gisement par les montages seuls et les traçages seuls respectivement. On admet (ce qui est vrai avec une bonne approximation) $E(Z - Z_M)(Z - Z_T) = 0$.

a/ Montrer que l'estimateur optimal (krigeage) de Z est $Z^* = \lambda Z_M + (1-\lambda) Z_T$ avec

$$\lambda = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_M^2} \quad (\text{pondération par l'inverse des variances d'estimation respectives}) \text{ et}$$

que la variance correspondante est $\frac{\sigma_M^2 \sigma_T^2}{\sigma_T^2 + \sigma_M^2}$.

(les coefficients sont de la forme λ et $1-\lambda$ à cause de la condition d'universalité. Partir ensuite de $D^2(Z-Z^*) = D^2[\lambda(Z-Z_M) + (1-\lambda)(Z-Z_T)] = \lambda^2 \sigma_M^2 + (1-\lambda)^2 \sigma_T^2$, et minimiser en λ).

b/ Application au cas De Wijsien, lorsque les longueurs des traçages et des montages sont supérieures à leur équidistance.

[Désigner par L_M et L_T les longueurs (totales) des montages et des traçages, et utiliser la relation $\sigma_E^2 = \alpha \frac{\pi S}{L^2}$ de 2-8-2. On trouve $\lambda = \frac{L_M^2}{L_M^2 + L_T^2}$ (pondération par les carrés des longueurs lontées et tracées, et $\sigma_K^2 = \alpha \frac{\pi S}{L_M^2 + L_T^2}$]

Exercice 3 (Propriété markovienne de la covariance exponentielle et du variogramme linéaire).

Cet exercice préparatoire sera utile pour la compréhension des exercices suivants.

a/ Une F.A. $Z(x)$ sur la droite réelle est markovienne si, pour tout x_0 , les $Z(x)$, $x > x_0$ et les $Z(x')$, $x' < x_0$ sont conditionnellement indépendants à $Z(x_0)$ fixé. Montrer qu'une F.A. stationnaire d'ordre 2 à loi gaussienne est markovienne si et seulement si sa covariance centrée est de la forme $C e^{-a|h|}$.

[Deux gaussiennes sont indépendantes dès que leur coefficient de corrélation est nul. Si X_1, X_2, X_3 sont gaussiennes, le coefficient de corrélation de X_1 et X_3 à X_2 fixé est :

$$\frac{\rho_{13} - \rho_{12} \rho_{23}}{\sqrt{(1 - \rho_{12}^2)(1 - \rho_{23}^2)}}$$

En déduire que la F.A. est markovienne si et seulement si $C(h+h') = C(h) C(h')$ (h, h' positifs)]

b/ Soit $Z(x)$ une F.A.I. sur la droite réelle. On dit que $Z(x)$ est à accroissements indépendants si les $Z(x_i) - Z(y_i)$ sont indépendants dès que les intervalles (y_i, x_i) sont disjoints ou possèdent un seul point commun. Montrer qu'une F.A.I. dont les accroissements sont gaussiens est à accroissements indépendants si et seulement si son variogramme est linéaire (la F.A.I. constitue alors un processus de Wiener-Lévy, ou mouvement brownien).

[un calcul simple montre que les accroissements sont sans corrélation si et seulement si le variogramme est linéaire].

Exercice 4 (covariance exponentielle, cas continu) - a/ Soit, sur la droite, une F.A. $Y(x)$ d'espérance nulle et $C(h) = e^{-a|h|}$ sa covariance. On connaît la réalisation de $Y(x)$ sur

l'intervalle $(-R, +R)$. Krieger le point $x_0 = R + h$ ($h > 0$).

[Si $Y(x)$ est gaussienne, on déduit de la propriété markovienne de Ex. 3 que l'estimateur optimal est $e^{-a|h|} Y(R)$ avec la variance $1 - e^{-2a|h|}$. Cette propriété, liée à la seule covariance, subsiste si $Y(x)$ n'est pas gaussienne. On vérifiera directement que la mesure $\lambda_R = e^{-a|h|} \delta_R$ (Dirac placé en R) est effectivement solution de (3-6) pour tout $y \leq R$, et on calculera directement σ_Y^2].

b/ Soit sur la droite une F.A. $Z(x)$ d'espérance m inconnue et $e^{-a|x|}$ sa covariance centrée. Montrer que l'estimateur optimal m^* de m (lorsque la réalisation est connue sur $(-R, R)$) vérifie :

$$\begin{cases} m^* = \frac{aR}{1+aR} \bar{Z} + \frac{1}{1+aR} \frac{Z_R + Z_{-R}}{2} & (\bar{Z} = \frac{1}{2R} \int_{-R}^R Z(x) dx) \\ D^2(m^*) = \frac{1}{1+aR} \end{cases}$$

[établir les relations $\frac{1}{2} \int (\delta_R + \delta_{-R}) e^{-a|x-y|} = e^{-aR}$ et $\int_{-R}^R e^{-a|x-y|} dx =$

$\frac{2}{a} - \frac{2}{a} e^{-aR}$ pour $-R \leq y \leq R$. En déduire que la mesure $\nu_0 = \frac{a}{2} dx + \frac{1}{2} (\delta_R + \delta_{-R})$ vérifie $\int_{-R}^R \nu_0(dx) e^{-a|x-y|} = 1$, et que la mesure $\lambda_0 = \frac{1}{1+aR} \nu_0$ est la solution cherchée avec le paramètre de Lagrange $\mu_0 = D^2(m^*) = \frac{1}{1+aR}$]

c/ Dans les mêmes conditions qu'en b/ , krieger le point $x_0 = R + h$ ($h > 0$) en appliquant le théorème d'additivité

$$[Z^* = e^{-a|h|} Z_R + (1 - e^{-a|h|}) m^* ; D^2(Z_{x_0} - Z^*) = 1 - e^{-2ah} + \frac{(1 - e^{-ah})^2}{1+aR}]$$

Exercice 5 (Covariance exponentielle, cas discret) a/ soit, sur la droite, $Y(x)$ une F.A. d'espérance nulle de covariance, stationnaire $e^{-a|h|}$. On connaît la réalisation sur les $n+1$ points d'abscisse $0, 1, \dots, n$. Former le krigeage ponctuel du point $x_0 = i + \varepsilon$ ($0 < \varepsilon < 1$) ($i < n$).

[D'après la propriété markovienne, on doit chercher une solution de la forme $Y_K = \lambda Y_i + \lambda' Y_{i+1}$. Montrer que le système (3-4) est vérifié pour $\lambda_j = 0$ ($j \neq i, j \neq i+1$)

$$\lambda_i = \lambda = \frac{Sh(1-\varepsilon)a}{Sha} , \lambda_{i+1} = \lambda' = \frac{Sh \varepsilon a}{Sha}]$$

b/ soit, dans les mêmes conditions, $Z(x)$ une F.A. de covariance $e^{-a|h|}$ et d'espérance m inconnue. Montrer que l'estimateur optimal m^* est de la forme

$$m^* = (1-b) \bar{Z} + b \frac{Z_0 + Z_n}{2} \quad \left(\bar{Z} = \frac{1}{n+1} \sum_{i=0}^n Z_i \right)$$

$$\text{avec } b = \frac{2}{(n+1) e^{a-(n-1)}} \quad , \quad \mu_0 = D^2(m^*) = \frac{1 + e^a}{(n+1) e^{a-(n-1)}}$$

[établir d'abord la relation $\sum_{j=0}^n e^{-a|i-j|} = \frac{e^{a+1}}{e^a-1} - \frac{e^{-a1} + e^{-(n-1)a}}{e^a-1}$;

chercher pour le système (3-13) une solution de la forme $\lambda^\beta = A + B(\delta_0^\beta + \delta_n^\beta)$ et identifier]

c/ Dans les mêmes conditions, kriger un point x_0 compris entre i et $i+1$ lorsque m est inconnue.

[appliquer la relation d'additivité]

Exercice 6 (Variogramme linéaire) a/ Soit sur la droite une F.A.I. sans dérive et $\gamma(h) = |h|$ son variogramme. On connaît la réalisation en $x = R$, et en un nombre fini ou infini de points $\llcorner R$. Montrer que le krigeage de $x_0 = R + h$ ($h > 0$) est $Z(R)$ lui-même, avec $\sigma_K^2 = 2|h|$.

[cette solution est suggérée par la propriété markovienne de $|h|$. Vérifier que la mesure $\lambda(dx) = \delta_R$ est bien solution du système (3-18) pour tout $y \leq R$]

b/ On suppose maintenant la réalisation connue en deux points x_1 et x_2 et en un nombre quelconque de points extérieurs à l'intervalle (x_1, x_2) . Montrer que le krigeage d'un point x_0 compris entre x_1 et x_2 est :

$$Y^*(x_0) = \frac{x_2 - x_0}{x_2 - x_1} Y(x_1) + \frac{x_0 - x_1}{x_2 - x_1} Y(x_2) \quad (x_1 \leq x_0 \leq x_2)$$

$$\sigma_K^2 = 2 \frac{(x_2 - x_0)(x_0 - x_1)}{x_2 - x_1}$$

[la propriété markovienne suggère que la solution est du type $\lambda Y_1 + (1-\lambda) Y_2$. Montrer qu'avec $\lambda = \frac{x_2 - x_0}{x_2 - x_1}$ (interpolation linéaire entre les deux points connus les plus proches) la mesure $\nu = \lambda \delta_{x_1} + (1-\lambda) \delta_{x_2}$ vérifie $\int \nu(dx) |x-y| = |x_0 - y|$ pour

$y < x_1$ ou $y \geq x_2$. Le paramètre de Lagrange est nul, et le calcul de σ_K^2 est immédiat).

Exercice 7 (Variogramme linéaire ou effet de pépité) - On connaît, sur le segment $(-R, +R)$, une réalisation d'une F.A.I. dont le variogramme comporte un terme linéaire et un terme de pépité. Comme on est dans le cas continu, on prendra $\gamma = |h| - C\delta$ (effet de pépité représenté par une mesure de Dirac). Kriger le point $x_0 = R + h$ ($h > 0$). Pour cela :

a/ Etablir les relations $\int_{-R}^R |x-y| \text{ chax } dx = \frac{2R}{a} \text{ sha } R + \frac{2}{a^2} \text{ chay}$, $\int_{-R}^R |x-y| \text{ shax } dx = -2 \frac{\text{cha } R}{a} y + \frac{2}{a^2} \text{ shay}$. En déduire, avec $a = \sqrt{\frac{2}{C}}$:

$$\int_{-R}^R (|x-y| dx - C\delta(dx)) \text{ chax} = \frac{2R}{a} \text{ sha } R, \int_{-R}^R (|x-y| dx - C\delta(dx)) \text{ shax} = -2 \frac{\text{Cha } R}{a} y$$

b/ Montrer que la fonction $\lambda(x) = \frac{a \text{ shax}}{2 \text{ cha} R} + \frac{a \text{ chax}}{2 \text{ sha} R}$ vérifie le système :

$$\int_{-R}^R \lambda(x) dx = 1, \int_{-R}^R (|x-y| dx - C\delta(dx)) \lambda(x) = -y + R. \text{ En déduire que l'estimateur optimal est } Z^*(x_0) = \int_{-R}^R \lambda(x) Z(x) dx, \text{ avec le paramètre de Lagrange } \mu = x_0 - R = h \text{ et } \sigma_K^2 = 2h + \frac{1}{a} \text{ cha } R.$$

[remarquer que la présence d'un effet de pépité a fait disparaître la propriété de Markov. La densité $\lambda(x)$ attribue des poids positifs à tous les échantillons. On dit que l'effet de pépité lève les écrans].

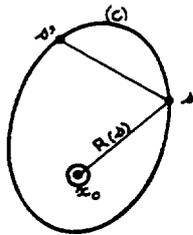
Exercice 8 (Exponentielle de Gauss) - Lorsque la covariance est très régulière, il arrive que la solution du krigeage ne soit pas représentable par une mesure. Par exemple, considérons sur la droite une F.A. stationnaire admettant la covariance $C(h) = e^{-\frac{h^2}{2}}$ et une espérance nulle. Soit à kriger $x_0 = R + h$, $h > 0$ connaissant la réalisation sur $(-R, +R)$. Si l'estimateur optimal a une représentation du type $Y_K = \int_{-R}^R \lambda(dx) Y(x)$, la mesure λ vérifie

$$\int_{-R}^R \lambda(dx) e^{-\frac{(x-y)^2}{2}} = e^{-\frac{(x_0-y)^2}{2}} \quad \forall y \in (-R, +R)$$

Montrer qu'il ne peut pas exister de mesure λ vérifiant cette relation.

[la mesure $\nu(dx) = e^{-\frac{x^2}{2}} \lambda(dx)$ devrait vérifier $\int_{-R}^R e^{xy} \nu(dx) = e^{-\frac{x_0^2}{2} + x_0 y}$. D'après les propriétés de la transformation de Laplace, l'unique mesure vérifiant cette relation serait $e^{-\frac{x_0^2}{2}} \delta_{x_0}$, mais son support n'est pas dans $(-R, +R)$].

Exercice 9 (Krigage De Wijsien dans \mathbb{R}^2) - Dans \mathbb{R}^2 , le schéma De Wijsien possède une propriété à caractère markovien : lorsque l'on connaît la réalisation sur un contour fermé C, il y a absence de corrélation entre l'intérieur et l'extérieur de C : dans le krigage d'un élément intérieur à C, la connaissance d'informations extérieures à C ne change pas la solution. Il y a un effet d'écran total, que l'analogie électrique permet de comprendre ($\log r$ est le potentiel harmonique du plan, et joue dans \mathbb{R}^2 le même rôle que le potentiel newtonien $1/r$ dans \mathbb{R}^3).



a/ Soit $Z(x)$ une F.A.I. De Wijsienne dans \mathbb{R}^2 , (C) un contour fermé sur lequel on connaît la réalisation, s l'abscisse curviligne d'un point de C, $r(s, s')$ la distance de deux points d'abscisses s et s' sur C. On veut kriger un petit cercle centré en un point x_0 intérieur. On désigne par $R(s)$ la distance de x_0 et du point courant de C d'abscisse s . Montrez que l'estimateur optimal est $Z^* = \int_C \lambda(s) Z(s) ds$ avec une fonction $\lambda(s)$ vérifiant :

$$\begin{cases} \int_C \lambda(s) \log r(s, s') ds = \log R(s') + \mu \\ \int_C \lambda(s) ds = 1 \end{cases}$$

En déduire que $\lambda(s)$ est la densité induite par une masse -1 placée en x_0 (lorsque le potentiel est $\log r$), autrement dit que (C) est une équipotentielle lorsque l'on place la masse -1 en x_0 et la densité $\lambda(s)$ sur C.

On sait que le potentiel est alors constant non seulement sur C mais également à l'extérieur de C. En déduire que si y est un point extérieur quelconque, on a encore :

$$\int_C \lambda(s) \log r(s, y) ds = \log r(x_0, y) + \mu$$

En déduire que la densité $\lambda(s)$ donne encore la solution du krigage lorsque la réali-

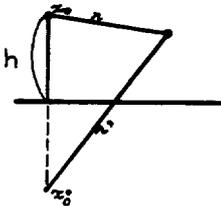
sation est connue sur C et sur un ensemble S' quelconque disjoint de l'intérieur de C (effet d'écran total).

b/ La fonction de Green $G(x)$ relative au point x_0 est caractérisée par les deux propriétés suivantes : 1/ $G(x)$ est constante sur (C)

2/ $G(x) - \log r(x_0, x)$ est une fonction régulière et harmonique en tout point intérieur x , y compris le point x_0 . On sait que la densité induite $\lambda(s)$ est donnée par $\frac{1}{2\pi} \frac{dG}{dn}$ (dérivée suivant la normale positive).

En déduire le krigeage d'un point x_0 connaissant la réalisation sur une droite.

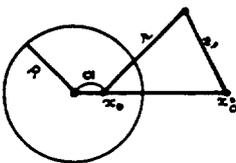
[la fonction de Green est $\log \frac{r}{r'}$, cf figure ci-jointe, et on trouve $\lambda(x) = \frac{1}{\pi} \frac{h}{x^2+h^2}$].



c/ Même question pour un point x_0 intérieur connaissant la réalisation sur le cercle de rayon R .

[ici $G = \log \frac{r}{r'}$, avec comme point x_0' l'inverse de x_0 dans l'inversion de module R^2 . Utiliser les coordonnées polaires :

$$G = \frac{1}{2} \log \frac{\rho^2 + a^2 - 2 a \rho \cos \theta}{\rho^2 + \frac{R^4}{a^2} - 2 \frac{R^2}{a} \rho \cos \theta}$$



On trouve la densité

$$\lambda(\theta) = \frac{R^2 - a^2}{2 \pi R} \frac{1}{r^2} = \frac{R^2 - a^2}{2 \pi R} \frac{1}{\rho^2 + a^2 - 2 a \rho \cos \theta}$$

Exercice 10 (Krigage aux grandes mailles) (cf. Ex. 13 à 15 du Chapitre 2) - On veut kriger une surface S à l'aide de $n+k$ échantillons, dont k sont prélevés à l'intérieur de S et n à l'extérieur, en supposant que les distances entre deux échantillons distincts, et la distance de chaque échantillon à la frontière de S sont toutes supérieures à la portée. Les notations sont celles des Exercices 13 à 15 du Chapitre 2.

a/ Poser les équations du krigeage. Montrer qu'elles se résument à

$$\forall i \quad \lambda_i = \sigma_{si} + \mu$$

$$\sum \lambda_i = 1$$

avec $\sigma_{si} = 0$ pour les échantillons i extérieurs et $\sigma_{si} = \pi \frac{A_1}{S}$ pour les échantillons intérieurs.

En déduire que les k échantillons intérieurs ont le même poids λ et les n échantillons extérieurs le même poids λ' avec :

$$\lambda = \frac{1}{n+k} + \frac{n}{n+k} \pi \frac{A_1}{CS}$$

$$\lambda' = \frac{1}{n+k} - \frac{k}{n+k} \pi \frac{A_1}{CS}$$

b/ Calculer la variance de ce krigeage :

$$(\sigma_k^2 = C - F(S) - \frac{2k}{n+k} \pi \frac{A_1}{S} - \frac{nk}{n+k} (\pi \frac{A_1}{S})^2 \frac{1}{C} + \frac{C}{n+k})$$

En particulier, si S est grand, on a la formule approchée :

$$\sigma_k^2 \approx \frac{C}{n+k} + \frac{n-k}{n+k} \pi \frac{A_1}{S}$$

c/ Comparer ces résultats avec la formule (3-1) du paragraphe 3-1, et montrer que, dans le cas des grandes mailles, l'estimateur optimal est identique à celui que KRIGE a obtenu originellement par sa méthode de régression linéaire.

[Le coefficient de régression est $\beta = \frac{\sigma_y^2}{\sigma_x^2} = \frac{C-F(S)}{C/k} = \frac{k}{C} \pi \frac{A_1}{S}$. Lorsque m est connu, l'estimateur optimal est identique à celui que donne la régression linéaire :

$Z_K = m + \beta \left(\frac{1}{k} \sum_1 X_i - m \right)$, X_i teneurs des échantillons intérieurs. D'après (3-4-3) on doit prendre :

$$Z^* = m^* + \beta \left(\frac{1}{k} \sum_1 X_i - m^* \right)$$

avec $m^* = \frac{1}{n+k} \left(\sum_1 X_i + \sum_j Y_j \right)$ (Y_j : échantillons extérieurs). On retrouve ainsi : λ et λ'].

Calculer la variance à partir du théorème d'additivité.

[à m connue, la variance est $\sigma_y^2(1-\rho^2)$. Montrer $\rho^2 = \beta = \frac{k}{C} \pi \frac{A_1}{S}$. On en déduira :

$$\sigma_k^2 = \sigma_y^2(1-\beta) + (1-\beta)^2 D^2(m^*) = (1-\beta) \pi \frac{A_1}{S} + (1-\beta)^2 \frac{C}{n+k}$$

et on vérifiera que ce résultat est identique à celui du parag. b/.]

CHAPITRE 4

LE KRIGEAGE UNIVERSEL

4-1 INTRODUCTION

4-1-1 Critique des méthodes de moindres carrés.

Nous nous proposons, dans ce qui suit, de formuler en termes de fonctions aléatoires non stationnaires le problème de l'estimation des dérivées¹ (ou tendances), et de montrer que ce problème admet une solution optimale, bien différente de celles auxquelles conduisent les méthodes dites de "trend surface analysis" : ces dernières, qui consistent en général à ajuster un polynôme par une méthode de moindres carrés, donnent un sentiment de malaise; on a l'impression, en effet, de faire violence à la nature, en lui imposant de force une expression polynomiale, qui n'a aucune raison a priori de présenter le moindre rapport avec la structure réelle du phénomène que l'on veut représenter. D'une manière plus précise, nous adresserons trois reproches principaux à ces méthodes de moindres carrés :

Tout d'abord, ces méthodes entraînent souvent une confusion du mode opératoire et du concept. Peu d'auteurs se donnent la peine de définir la signification de ce "trend" qu'ils estiment par une méthode de moindres carrés. On a souvent l'impression que ce fameux trend n'est rien de plus que le résultat numérique auquel conduit un mode opératoire - c'est-à-dire, peut-être, un pur et simple artefact - A l'analyse, il semble que le terme peu clair de "trend" se réfère tour à tour, et parfois même simultanément dans un même contexte, à trois notions bien distinctes au moins : si $Z(x)$ est la variable régionalisée à laquelle on s'intéresse, et $P(x)$ le polynôme ajusté par moindres carrés à partir des valeurs expérimentalement connues aux points x_1, x_2, \dots , on peut attribuer à la valeur en x de $P(x)$ l'une ou l'autre des significations incompatibles suivantes :

a/ $P(x)$ est (une estimation de) l'espérance a priori $E(Z(x))$: c'est toujours ce sens a/ que nous attribuerons au terme "dérive".

1 - Pour éviter tout anthropomorphisme, nous utiliserons toujours le terme "dérive" au lieu de "tendance", ou "trend".

b/ $P(x)$ est une estimation de la vraie valeur (inconnue) $Z(x)$ prise par la variable régionalisée au point x : c'est à ce sens b/ que se rattache notre terme "krigeage ponctuel". Plus précisément, le krigeage ponctuel sera le meilleur estimateur linéaire de $Z(x)$ construit à partir des données disponibles $Z(x_1), Z(x_2), \dots$

c/ On attribue parfois, enfin, à $P(x)$ le sens d'une "moyenne mobile". $P(x_0)$ serait alors (une estimation de) la valeur moyenne de $Z(x)$ dans une zone plus ou moins grande (à préciser) entourant un point donné x_0 . Dans notre terminologie, le krigeage désignera le meilleur estimateur linéaire de cette moyenne mobile.

Ces distinctions ont une grande importance théorique et pratique. Il est essentiel de définir avec précision l'objectif que l'on vise (a/, b/ ou c/). En cartographie sous-marine, ce sont les vraies valeurs $Z(x)$ que l'on doit estimer et cartographier (cas b/). En exploitation minière, on s'intéresse à la teneur moyenne d'un panneau de taille donnée (cas c/). Dans certaines études à caractère plus fondamental, enfin, où l'on cherche à reconstituer les mécanismes généraux qui ont donné naissance au phénomène que l'on étudie, c'est plutôt dans la dérive elle-même (cas a/) que l'on peut espérer trouver un reflet de la structure de ces mécanismes. En géophysique, par exemple, la notion d'anomalie régionale correspond bien à notre concept de dérive.

Nous en arrivons ainsi au deuxième grief que l'on peut formuler à l'encontre des méthodes de moindres carrés. Il n'est pas possible que le même polynôme $P(x)$ résolve à la fois les trois problèmes a/, b/ et c/. Nous verrons que $P(x)$ constitue en fait une solution du problème a/. Mais, en général, ce n'est pas la meilleure solution possible : ces méthodes passe-partout qui utilisent toujours les mêmes polynômes, quelles que soient les caractéristiques structurales du phénomène que l'on étudie, n'ont aucune chance en général, de conduire à un optimum.

En troisième lieu, enfin, les méthodes de moindre carré ne permettent en aucune façon d'évaluer l'erreur que l'on commet en estimant la dérive à l'aide du polynôme $P(x)$. La variance des résidus, contrairement à ce que l'on croit parfois, n'est pas une variance d'estimation : la variance des écarts $Z(x_1) - P(x_1)$ aux points x_1 où les données sont connues est, par construction, systématiquement plus faible (et même beaucoup plus faible) que celle de l'écart $Z(x) - P(x)$ en un point x différent des x_1 . Ainsi, la variance des résidus n'est pas la variance d'estimation de $Z(x)$ (problème b/). Elle n'est pas davantage la variance d'estimation de la dérive (problème a/), car il n'y a cette fois plus aucun lien conceptuel entre un écart $Z(x_1) - P(x_1)$ et la qualité de $P(x_1)$ considéré comme un estimateur de l'espérance a priori $E[Z(x_1)]$.

Cependant, si les méthodes de moindres carrés se révèlent peu satisfaisantes, l'objectif qu'elles visaient, et qu'elles n'atteignent pas ou atteignent mal, correspond, lui, à un problème très réel et très important. Il existe réellement des phénomènes que l'on ne peut absolument pas assimiler à des (réalisations de) fonctions aléatoires stationnaires. Pour reprendre l'exemple de la cartographie des fonds sous-marins, il est certain que la profondeur va en augmentant lorsqu'on s'éloigne des côtes. Nous allons essayer de formuler et de résoudre dans le cadre de la théorie des fonctions aléatoires non stationnaires les problèmes importants que posent ces dérives.

4-1-2 Position du problème et hypothèses générales.

La variable régionalisée $z(x)$ que l'on étudie sera interprétée comme une réalisation d'une fonction aléatoire $Z(x)$ non stationnaire en général. Dans un premier groupe d'hypothèses, nous supposons que $Z(x)$ admet des moments d'ordre 1 et 2 :

$$(I) \quad \begin{cases} E[Z(x)] = m(x) \\ E[Z(x) Z(y)] = m(x) m(y) + C(x,y) \end{cases}$$

Nous dirons que la fonction $m(x)$ est la dérive de la F.A. $Z(x)$. Il nous semble, en effet, que la seule définition conceptuellement claire de la notion de dérive est bien celle-ci : moment d'ordre 1 d'une F.A. non stationnaire. La fonction $C(x,y)$ qui dépend séparément des deux points x et y est la covariance (non stationnaire) habituelle.

Dans bien des cas (comme le montre la pratique de la Géostatistique), ces hypothèses (I) seront encore trop restrictives, et devront être remplacées par les hypothèses (I') suivantes qui expriment que les accroissements de la fonction aléatoire $Z(x)$ (et non $Z(x)$ elle-même) admettent des moments d'ordre 1 et 2 :

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

La dérive $m(x)$ n'est plus ici déterminée qu'à une constante près, car, en général, l'espérance $E[Z(x)]$ n'existe plus : le point de vue des hypothèses (I') revient donc à étudier la fonction

aléatoire $Z(x)$ à une constante additive près. La fonction $\gamma(x,y)$ est le demi-variogramme habituel. (Dans l'espace à une seule dimension, des processus aussi usuels que le mouvement brownien ou les processus de Poisson donnent des exemples simples de F.A. sans espérance, vérifiant les hypothèses (I') mais non les hypothèses (I)).

Nous avons à résoudre les deux problèmes (liés, mais distincts) suivants : connaissant les valeurs numériques prises sur un ensemble S (l'ensemble des points où l'on dispose de données expérimentales) par une réalisation $z(x)$ de la F.A. $Z(x)$ nous devons :

1°/ estimer la fonction $m(x)$ (sur S et à l'extérieur de S) : c'est le problème a/ ci-dessus.

2°/ estimer $z(x)$ en un point $x \notin S$ (problème b/), ou, plus généralement estimer une "moyenne mobile" $\int \mu(dx) z(x)$ où μ est une mesure donnée dont le support est disjoint de S (problème c/).

De plus, nous devons être capables de représenter par des variances d'estimation les erreurs commises dans ces opérations, et nous efforcer de choisir nos estimateurs de manière à minimiser (si faire se peut) cette variance d'estimation.

En ce qui concerne ce dernier point, indiquons que nous nous limiterons en fait à rechercher le meilleur estimateur linéaire de $m(x)$ ou de $\int \mu(dx) z(x)$ que l'on puisse former à partir des valeurs numériques des $z(y)$, $y \in S$: les estimateurs non linéaires sont beaucoup trop compliqués pour qu'il soit possible de les mettre en oeuvre dans les applications, et d'autre part leurs propriétés ne sont plus liées aux seuls moments d'ordre 1 et 2, qui figurent dans les hypothèses (I) ou (I'), mais font intervenir la totalité de la loi spatiale de la F.A. $Z(x)$. (Dans le cas où cette loi spatiale est gaussienne, on sait d'ailleurs que le meilleur estimateur possible coïncide avec le meilleur estimateur linéaire).

Ainsi formulé en toute généralité (les fonctions $m(x)$ et $C(x,y)$ étant complètement inconnues) notre problème est manifestement insoluble, et d'ailleurs probablement dépourvu de sens. mais, dans les problèmes concrets, la notion de dérive ne peut présenter une signification réelle que si la fonction $m(x)$ correspondante varie d'une manière continue et régulière relativement à l'échelle à laquelle on travaille (et aux données expérimentalement disponibles) : si la fonction $m(x)$ était irrégulière et chaotique à cette échelle, on devrait la considérer elle-même comme une réalisation d'une nouvelle fonction aléatoire. Du point de vue de l'interprétation physique (et non mathématique) la notion de dérive est ainsi manifestement liée à celle d'échelle. Si l'on étudie des cotes topographiques, à l'échelle de la dizaine de mètres, la notion de montagne se traduit par une dérive ; à l'échelle de la dizaine de kilomètres, il ne lui correspond plus qu'une fonction aléatoire,

et la dérive, à cette échelle, exprimera plutôt la notion de chaîne de montagnes. (cf. aussi la notion de structures gigogne, [7]).

Mais cette condition de régularité imposée a priori à la fonction $m(x)$ pour que la notion de dérive ait un contenu physique réel, entraîne aussi qu'une estimation de $m(x)$ doit toujours être plus ou moins possible localement. Cette condition exprime en effet que - sur un certain voisinage V d'un point x_0 donné - la fonction $m(x)$ peut être approchée avec une excellente précision par une fonction de la forme :

$$(II) \quad m(x) \approx f(x) = \sum_{\ell=0}^k a_{\ell} f^{\ell}(x) = a_{\ell} f^{\ell}(x)$$

où les $f^{\ell}(x)$ sont des fonctions connues, choisies une fois pour toutes (par exemple des polynômes), et les a_{ℓ} des coefficients inconnus : quant au voisinage V de x_0 sur lequel l'approximation (II) est acceptable, sans être très grand, il doit - si le problème a un sens - contenir un nombre suffisant de points expérimentaux pour qu'il soit possible d'estimer les $k+1$ coefficients inconnus a_0, a_1, \dots, a_k .

Il reste le problème de la fonction $C(x,y)$, ou $\gamma(x,y)$ sur laquelle on ne sait rien, non plus, a priori. Mais, ici encore, et pour les mêmes raisons, on peut supposer que C ou γ sont localement assimilables à des fonctions de type connu, et ne se déforment qu'assez lentement dans l'espace à l'échelle à laquelle on travaille.

Le cas le plus favorable sera celui d'une fonction γ (ou C) de la forme $\gamma(x,y) = \omega \gamma_0(x-y)$ où γ_0 est une fonction connue, et ω un facteur lentement variable, que l'on pourra regarder comme constant sur le voisinage V du point x_0 : les équations qui déterminent les estimateurs optimaux étant linéaires et homogènes, ces estimateurs ne dépendront que de γ_0 et non du facteur ω ; ce dernier ne se répercutera donc que sur les variances d'estimation, et non sur les estimateurs eux-mêmes. Dans certains cas, on pourra prendre l'expression très simple :

$$\gamma(x,y) = \omega r \quad (r = |x-y|)$$

(variogramme linéaire). Il suffit, pour cela, en effet, que le vrai variogramme ou la vraie covariance ait un comportement linéaire au voisinage de $x = y$ et jusqu'à des distances comparables aux dimensions du voisinage V ci-dessus : cette circonstance se rencontre plus souvent qu'on ne croit.

En général, pourtant, en plus du facteur ω , il conviendra d'introduire un ou plusieurs paramètres supplémentaires. Par exemple, on prendra

$$\gamma(x,y) = \omega r^{\alpha-1} \quad (0 < \alpha < 2)$$

ou encore

$$C(x,y) = \omega e^{-ar}$$

Dans le premier cas, le paramètre α est en relation avec le degré de continuité de la variable régionalisée. Dans le deuxième (covariance exponentielle), le paramètre a , ou plutôt son inverse, donne la mesure de la portée du phénomène (distance au-delà de laquelle les corrélations s'éteignent).

Naturellement, le contrôle expérimental d'une hypothèse de ce genre, et l'estimation des paramètres correspondants (ω et surtout α ou a) posera des problèmes assez délicats de statistique mathématique : malgré leur importance cruciale pour les applications, il ne nous sera pas possible de traiter ici complètement ces problèmes, et nous ne ferons guère que les évoquer.

En résumé, le problème que nous devons traiter se trouve schématisé comme suit : on a une F.A. $Z(x)$ vérifiant les hypothèses (I) ou (I'). Sur un voisinage V d'un point x_0 , la dérivée $m(x)$ est de la forme (II) avec des coefficients a_ρ inconnus. On connaît a priori (éventuellement à un facteur près) la covariance $C(x,y)$ ou le variogramme $\gamma(x,y)$. Enfin, on connaît les valeurs numériques de (la réalisation de) $Z(x)$ pour les points x appartenant à un ensemble $S \subset V$. On veut former les meilleurs estimateurs linéaires :

1°/ des coefficients a_ρ inconnus de la dérivée.

2°/ de la valeur numérique de (la réalisation de) $Z(x)$ en $x_0 \notin S$ (mais $x_0 \in V$) ou de $\int \mu(dx) Z(x)$ pour une mesure μ dont le support est disjoint de S .

3°/ nous voulons de plus contrôler la validité de l'hypothèse que nous avons faite en choisissant pour la covariance ou le variogramme une expression mathématique particulière, et estimer les paramètres dont cette expression dépend.

Pour simplifier l'exposé, nous ne formulerons pas la théorie en termes d'espaces de Hilbert. On devra donc se reporter à [6] pour certaines démonstrations (notamment pour établir les théorèmes d'existence et d'unicité). Nous utiliserons les notations définies au paragraphe 3-2. Il n'y

aura d'ailleurs de difficultés réelles que dans le cas où l'ensemble S des données expérimentales est infini : l'existence (au sens des espaces de Hilbert) de nos estimateurs optimaux est toujours assurée, mais il n'en résulte pas nécessairement qu'ils admettent des représentations du type $\int \lambda(dx) Z(x)$ avec des mesures λ à support dans S. Nous écrirons en général, dans ce qui suit, les équations intégrales relatives au cas continu sans répéter explicitement cette réserve essentielle, que le lecteur devra toujours garder présente à l'esprit.

4-2 L'ESTIMATION OPTIMALE DE LA DERIVE.

4-2-1 Estimation de la dérive en un point.

Examinons d'abord le cas où les hypothèses (I) et (II) du paragraphe 4-1-2 sont vérifiées : existence d'une dérive admettant la représentation :

$$m(x) = a_{\ell} f^{\ell}(x) \quad (\ell = 0, 1, \dots, k)$$

sur un certain voisinage V du domaine S des données expérimentales, et existence d'une covariance centrée $\sigma(x, y)$. Nous ne connaissons pas les valeurs numériques des k+1 coefficients a_{ℓ} , et nous voulons estimer la valeur $m(x_0) = a_{\ell} f^{\ell}(x_0)$ de la dérive en un point x_0 (qui peut, indifféremment, être ou non dans S, mais doit appartenir à V).

Nous allons pour cela former un estimateur linéaire de la forme :

$$m^*(x_0) = \lambda^{\alpha} Z_{\alpha} \quad \text{ou} \quad \int_S \lambda(dx) Z(x)$$

et imposer deux conditions aux coefficients λ^{α} ou à la mesure λ :

a/ L'estimateur doit être sans biais quelles que soient les valeurs (inconnues) des coefficients a_{ℓ} (condition d'universalité).

b/ Compte tenu de la condition d'universalité, la variance de cette estimation doit être minimale (condition d'optimalité).

En général, nous ferons les raisonnements dans le cas où l'ensemble S est fini, et nous nous

contenterons de transposer au cas continu les résultats ainsi obtenus (avec les réserves rappelées au paragraphe précédent quant à la représentabilité des estimateurs au moyen de mesures).

Examinons d'abord la condition d'universalité. Elle s'écrit $E[m^*(x_0)] = m(x_0)$, soit :

$$a_\ell \lambda^\alpha f_\alpha^\ell = a_\ell f^\ell(x_0)$$

quels que soient les a_ℓ , donc :

$$(4-2-1) \quad \lambda^\alpha f_\alpha^\ell = f^\ell(x_0)$$

ou, dans le cas continu :

$$\int_S \lambda(dx) f^\ell(x) = f^\ell(x_0)$$

Lorsque cette condition est vérifiée, $E(m-m^*)^2$ est égal à la variance $D^2(m-m^*)$ et ne dépend donc plus des coefficients inconnus a^ℓ . Cette variance est :

$$D^2(m^*) = \lambda^\alpha \lambda^\beta \sigma_{\alpha\beta}$$

Il reste à déterminer les coefficients λ^α en minimisant cette forme quadratique compte tenu de la condition d'universalité (4-2-1). Il apparaît $k+1$ paramètres de Lagrange μ_ℓ , et on obtient le système suivant :

$$(4-2-2) \quad \begin{cases} \lambda^\beta \sigma_{\alpha\beta} = \mu_\ell f_\alpha^\ell \\ \lambda^\alpha f_\alpha^\ell = f^\ell(x_0) \end{cases} \quad (\ell = 0, 1, \dots, k)$$

On suppose toujours que la matrice des covariances $\sigma_{\alpha\beta}$ est strictement définie positive. On démontre alors (cf [6]) que le système (4-2-2) est régulier, et admet donc une solution et une seule, si et seulement si les $k+1$ fonctions f^ℓ sont linéairement indépendantes sur S (au sens algébrique), c'est-à-dire si :

$$c_\ell f_\alpha^\ell = 0 \Rightarrow c_\ell = 0$$

Nous supposerons toujours, dans ce qui suit, que cette condition est vérifiée : mais, dans les applications, il faudra toujours s'assurer qu'il en est effectivement ainsi.

Dans le cas continu, on cherchera, avec les réserves d'usage, un estimateur de la forme $m^*(x_0) = \int_S \lambda(dx) Z(x)$, avec une mesure λ à support dans S vérifiant le système suivant :

$$(4-2-3) \quad \begin{cases} \int \lambda(dx) \sigma(x,y) = \mu_\ell f^\ell(y) & , \quad \forall y \in S \\ \int \lambda(dx) f^\ell(x) = f^\ell(x_0) \end{cases}$$

Cherchons maintenant la variance de notre estimateur. A l'optimum, le système (4-2-2) est vérifié. En multipliant par λ^α la première relation de ce système, et en utilisant la seconde relation, on trouve :

$$\lambda^\alpha \lambda^\beta \sigma_{\alpha\beta} = \mu_\ell \lambda^\alpha f_\alpha^\ell = \mu_\ell f^\ell(x_0)$$

Par suite, dans le cas discret comme dans le cas continu, la variance de notre estimateur est liée aux paramètres de Lagrange par la formule :

$$(4-2-4) \quad D^2(m^*(x_0)) = \mu_\ell f^\ell(x_0)$$

Cas où il existe un variogramme, mais pas de covariance.

Dans tout ce qui suit, la fonction f^0 correspondant à $\ell = 0$ sera toujours la fonction identiquement égale à 1, $f^0(x) = 1$, et le coefficient a_0 jouera un rôle assez particulier. En effet, la dérivée, comme la F.A. elle-même, n'est réellement définissable qu'à une constante près, et a_0 est en réalité indéterminé lorsqu'il n'existe pas de covariance. Il n'est, en principe, pas possible d'estimer $m(x_0)$ lui-même. On le voit aussi en remarquant que, pour $\ell = 0$, la condition d'universalité (4-2-1) se réduit à $\sum_\alpha \lambda^\alpha = 1$: il ne s'agit donc pas d'une combinaison linéaire autorisée (paragraphe 2-2-1). Par contre, il est possible d'estimer $m(x_0) - m(y_0)$ pour deux points x_0 et y_0 quelconques. En effet, fixons y_0 et considérons la F.A. :

$$Z'(x) = Z(x) - Z(y_0)$$

Elle admet la dérivée $m(x) - m(y_0) = m'(x)$ et la covariance :

$$C'(x,y) = -\gamma(x,y) + \gamma(x,y_0) + \gamma(y,y_0)$$

(Voir paragraphe 2-2-1). On a cette fois :

$$m'(x) = a_\ell \left(f^\ell(x) - f^\ell(y_0) \right) \quad (\ell = 1, 2, \dots, k)$$

la constante a_0 s'éliminant. On peut alors reprendre les raisonnements et estimer $m'(x_0)$ à l'aide d'un estimateur linéaire autorisé :

$$\begin{cases} m'^*(x_0) = \lambda^\alpha Z'_\alpha = \lambda^\alpha Z_\alpha \\ \sum_\alpha \lambda^\alpha = 0 \end{cases}$$

Les conditions d'universalité s'écrivent alors :

$$\lambda^\alpha \varphi_\alpha^\ell = \varphi^\ell(x_0) \quad \ell = 1, 2, \dots, k$$

(l'équation en $\ell = 0$ disparaît, puisque $\varphi^0 = 1 - 1 = 0$). La variance $D^2(m'^*)$ est :

$$D^2(m'^*) = \lambda^\alpha \lambda^\beta (-\gamma_{\alpha\beta} + \gamma_{\alpha x_0} + \gamma_{\beta y_0}) = -\lambda^\alpha \lambda^\beta \gamma_{\alpha\beta}$$

à cause de la condition $\sum \lambda^\alpha = 0$, et on obtient le système :

$$\begin{cases} -\lambda^\beta \gamma_{\alpha\beta} = \mu_\ell \varphi_\alpha^\ell + \mu'_0 \\ \lambda^\alpha \varphi_\alpha^\ell = \varphi^\ell(x_0) \\ \sum \lambda^\alpha = 0 \end{cases}$$

D'ailleurs $\mu_\ell \varphi_\alpha^\ell = \mu_\ell f_\alpha^\ell - \mu_\ell f^\ell(y_0)$ et la constante $\mu_\ell f^\ell(y_0)$ peut être incorporée au paramètre de Lagrange μ'_0 . De même $\lambda^\alpha \varphi_\alpha^\ell = \lambda^\alpha f_\alpha^\ell$ à cause de $\sum \lambda^\alpha = 0$. Finalement, on obtient le système suivant, qui se déduit de (4-2-2) en remplaçant $\sigma_{\alpha\beta}$ par $-\gamma_{\alpha\beta}$, en limitant l'indice ℓ aux valeurs $1, 2, \dots, k$, et en introduisant en contrepartie la condition d'universalité supplémentaire $\sum \lambda^\alpha = 0$ et le paramètre μ_0 correspondant :

$$(4-2-5) \quad \begin{cases} \lambda^\beta \gamma_{\alpha\beta} = -\mu_\ell f_\alpha^\ell - \mu_0 \\ \lambda^\alpha f_\alpha^\ell = f^\ell(x_0) - f^\ell(y_0) \quad (\ell = 1, 2, \dots, k) \\ \sum_\alpha \lambda^\alpha = 0 \end{cases}$$

L'estimateur $\lambda^\alpha Z_\alpha$ correspondant (qui estime l'accroissement $m(x_0) - m(y_0)$) a pour variance :

$$(4-2-6) \quad D^2(\lambda^\alpha Z_\alpha) = \mu_\ell [f^\ell(x_0) - f^\ell(y_0)]$$

4-2-2 Estimation des coefficients de la dérive.

Plaçons-nous d'abord dans le cas où il existe une covariance. La solution de (4-2-2) $\lambda^\alpha = \lambda^\alpha(x_0)$ dépend évidemment de x_0 , et l'examen des seconds membres montre qu'elle dépend linéairement de $f^\ell(x_0)$. Lorsque x décrit V , ou même simplement S , les $f^\ell(x)$ sont linéairement indépendants. Il existe donc une matrice unique λ_ℓ^α et une matrice $\mu_{\ell S}$ telles que l'on ait :

$$\lambda^\alpha(x_0) = \lambda_\ell^\alpha f^\ell(x_0), \quad \mu_\ell(x_0) = \mu_{\ell S} f^S(x_0)$$

et cette matrice unique est la solution du système déduit de (4-2-2) en annulant le coefficient de $f^\ell(x_0)$, soit

$$(4-2-6) \quad \begin{cases} \lambda_\ell^\beta \sigma_{\alpha\beta} = \mu_{\ell S} r_\alpha^S \\ \lambda^\alpha r_\alpha^S = \delta_\ell^S \end{cases} \quad (\ell = 0, 1, \dots, k)$$

L'estimateur $m^*(x_0) = \lambda_\ell^\alpha Z_\alpha f^\ell(x_0)$ est donc de la forme $A_\ell f^\ell(x_0)$ avec

$$(4-2-7) \quad \begin{cases} A_\ell = \lambda_\ell^\alpha Z_\alpha \\ E(A_\ell) = a \end{cases}$$

(quelles que soient les valeurs inconnues des a_ℓ). Ainsi A_ℓ est un estimateur universel de a_ℓ , caractérisé par le fait que pour tout vecteur b^ℓ , la variance de $A_\ell b^\ell$ est minimale (compte tenu de cette condition d'universalité). A_ℓ est donc l'estimateur universel optimal du coefficient a_ℓ .

D'après (4-2-4), on a pour tout vecteur b^ℓ :

$$D^2(A_\ell b^\ell) = \mu_{\ell S} b^\ell b^S$$

Par suite la matrice $\mu_{\ell\beta}$ des paramètres de Lagrange est identique à la matrice des covariances des estimateurs optimaux A_ℓ (en particulier, elle est toujours symétrique):

$$(4-2-8) \quad \mu_{\ell\beta} = \text{Cov}(A_\ell, A_\beta) \quad (\ell = 0, 1, \dots, k)$$

Cas où il n'y a pas de covariance, mais seulement un variogramme.

Le système (4-2-5) dépend, lui aussi, linéairement de $f^\ell(x_0) - f^\ell(y_0)$. Il existe donc une matrice unique λ_ℓ^α ($\ell = 1, 2, \dots, k$) vérifiant :

$$(4-2-9) \quad \left\{ \begin{array}{l} \lambda_\ell^\beta \gamma_{\alpha\beta} = -\mu_{\ell\beta} f_\alpha^\beta - \mu_{0\ell} \\ \lambda_\ell^\alpha f_\alpha^\beta = \delta_\ell^\beta \\ \sum_\alpha \lambda_\ell^\alpha = 0 \end{array} \right. \quad (\ell = 1, 2, \dots, k)$$

Les $\lambda_\ell^\alpha Z_\alpha$, $\ell = 1, 2, \dots, k$, sont les estimateurs optimaux A_1, A_2, A_k des coefficients a_1, a_2, \dots, a_k (à l'exclusion de a_0), et la matrice $\mu_{\ell\beta}$ est égale à la matrice des covariances de ces estimateurs.

$$(4-2-10) \quad \mu_{\ell\beta} = \text{Cov}(A_\ell, A_\beta) \quad (\ell = 1, 2, \dots, k)$$

Remarque - On peut, pour compléter le système (4-2-9), introduire un vecteur λ_0^α et la fonction $f^0(x) = 1$, et déterminer λ_0^α par les conditions $\sum_\alpha \lambda_0^\alpha = 1$, $\lambda_0^\beta \gamma_{\alpha\beta} = -\mu_{0\beta} f_\alpha^\beta - \mu_{00}$. Le système ainsi complété est identique à (4-2-6).

$$\left\{ \begin{array}{l} \lambda_\ell^\beta \gamma_{\alpha\beta} = -\mu_{\ell\beta} f_\alpha^\beta \\ \lambda_\ell^\alpha f_\alpha^\beta = \delta_\ell^\beta \end{array} \right. \quad (\ell = 0, 1, \dots, k)$$

Nous reviendrons au paragraphe 4-6 sur l'interprétation que l'on peut donner des λ_0^α et de la combinaison linéaire

$$A_0 = \lambda_0^\alpha Z_\alpha$$

qui n'est pas une combinaison autorisée, donc n'admet pas, en général, de variance finie (la

relation (4-2-10) ne reste valable que pour ℓ et $s \neq 0$.

Cas Continu. - Il n'y a aucune difficulté à transposer les systèmes (4-2-6) ou (4-2-8) au cas où l'ensemble S des données est infini. Avec les réserves d'usage sur la possibilité de représenter les estimateurs A_ℓ (qui existent toujours au sens des espaces de Hilbert) au moyen des mesures λ_ℓ :

$$A_\ell = \int_S \lambda_\ell(dx) Z(x)$$

le système (4-2-9) par exemple, devient :

$$(4-2-11) \quad \left\{ \begin{array}{l} \int_S \lambda_\ell(dx) \gamma(x,y) = -\mu_{\ell s} f^s(y) - \mu_{0\ell} \quad (y \in S) \\ \int_S \lambda_\ell(dx) f^s(x) = \delta_\ell^s \\ \int_S \lambda_\ell(dx) = 0 \\ (\mu_{\ell s} = \text{Cov}(A_\ell, A_s)) \end{array} \right. \quad (\ell = 1, 2, \dots, k)$$

4-2-3 L'Invariance tensorielle

Montrons que l'estimateur optimal de la dérive n'est pas modifié si l'on effectue une transformation linéaire sur les fonctions f^ℓ . Plaçons-nous, par exemple, dans le cas où il existe une covariance σ (la démonstration serait analogue dans le cas du variogramme). Soit $B = (B_s^\ell)$ une matrice régulière, $B' = (B'^\ell_s)$ la matrice inverse. Posons

$$\varphi^\ell(x) = B_s^\ell f^s(x)$$

La dérive $m(x)$ se met sous la forme

$$\left\{ \begin{array}{l} m(x) = a'_\ell \varphi^\ell(x) \\ a'_\ell = B'^\ell_s a_s \end{array} \right.$$

Avec ces nouvelles fonctions φ^ℓ , l'estimateur optimal de $m^*(x_0)$ est

$$m^{*\ell}(x_0) = \lambda'^\alpha Z_\alpha$$

avec des coefficients λ'^{α} solution du système :

$$4-2-2') \quad \begin{cases} \lambda'^{\beta} \sigma_{\alpha\beta} = \mu'_{\ell} f_{\alpha}^{\ell} \\ \lambda'^{\alpha} \varphi_{\alpha}^{\ell} = \varphi^{\ell}(x_0) \end{cases}$$

mais $\varphi_{\alpha}^{\ell} = B_S^{\ell} f_{\alpha}^S$; $\varphi^{\ell}(x_0) = B_S^{\ell} f^S(x_0)$ et la matrice B est inversible. Par suite, la seconde relation (4-2-2') est équivalente à :

$$\lambda'^{\alpha} f_{\alpha}^{\ell} = f^{\ell}(x_0)$$

Le système (4-2-2') lui-même est donc équivalent au système (4-2-2). On a donc $\lambda'^{\alpha} = \lambda^{\alpha}$, puisque la solution est unique, et par suite :

$$m'^{*}(x_0) = m^{*}(x_0)$$

L'estimateur optimal est invariant pour les transformations linéaires des fonctions f^{ℓ} . En ce qui concerne les estimateurs optimaux A'_{ℓ} des nouveaux coefficients a'_{ℓ} , ils vérifient

$$A'_{\ell} \varphi^{\ell}(x_0) = A_{\ell} f^{\ell}(x_0)$$

pour tout x_0 , donc :

$$A'_S B_S^{\ell} = A_{\ell} \quad \text{ou} \quad A'_S = B_S^{\ell} A_{\ell}$$

Les estimateurs optimaux se transforment donc selon la même loi que les vrais coefficients, c'est-à-dire de façon contravariante vis-à-vis des f^{ℓ} .

Remarque - Dans les applications, il arrive souvent que l'on ait à étudier une zone trop étendue pour que la représentation de $m(x)$ sous la forme $a_{\ell} f^{\ell}(x)$ y soit admissible. On découpe alors cette zone en plusieurs domaines V_1, V_2, \dots sur chacun desquels cette approximation est légitime, mais avec des coefficients a_{ℓ} différents sur V_1, V_2 etc...

Pour étudier le domaine V_1 , il est plus commode de placer l'origine des coordonnées en un point $x_1 \in V_1$, et cela revient à effectuer le changement $\varphi^{\ell}(x) = f^{\ell}(x-x_1)$ sur les fonctions de base. Si les fonctions f^{ℓ} sont des polynomes, des fonctions trigonométriques, ou plus généralement

des exponentielles polynomes, les nouvelles fonctions φ^ℓ s'expriment linéairement au moyen des anciennes f^ℓ , et par suite, d'après l'invariance tensorielle, ce changement d'origine n'affecte pas l'estimateur optimal de la dérive.

On peut d'ailleurs montrer que les exponentielles polynomes sont les seules fonctions possédant cette propriété (invariance par translation de l'espace vectoriel de dimension finie engendrée par les f^ℓ). C'est pourquoi, dans les applications, on n'utilise jamais que des fonctions de ce type (à moins que les conditions physiques du problème ne fassent jouer un rôle privilégié à un point donné de l'espace que l'on doit alors choisir comme origine absolue des coordonnées).

4-2-4 Le variogramme des résidus.

Nous appellerons résidu au point expérimental $x \in S$ la différence entre la valeur observée et la dérive estimée (optimalement) en ce point, soit

$$R(x) = Z(x) - A_\ell f^\ell(x)$$

Lorsqu'il existe une covariance et que S est finie, les deux matrices :

$$A_\alpha^\beta = \lambda_\ell^\beta f_\alpha^\ell, \quad U_\alpha^\beta = \delta_\alpha^\beta - A_\alpha^\beta$$

donnent, respectivement, l'estimateur optimal m_α^* de la dérive en x_α et le résidu en ce point :

$$(4-2-13) \quad m_\alpha^* = A_\alpha^\beta Z_\beta, \quad R_\alpha = U_\alpha^\beta Z_\beta$$

Nous reviendrons ultérieurement sur les propriétés algébriques de ces deux matrices qui jouent un rôle important dans le problème de l'indétermination du variogramme sous-jacent (paragraphe 4-6).

S'il n'existe qu'un variogramme, mais pas de covariance, les résidus eux-mêmes ne sont plus définissables (puisque A_0 n'existe plus), mais leurs accroissements :

$$R(x) - R(y) = Z(x) - Z(y) - m^*(x) + m^*(y)$$

sont bien définis, puisqu'ils ne dépendent pas de A_0 . On peut donc parler du variogramme des

résidus : ce sera, par définition, la fonction

$$\gamma_R(x,y) = \frac{1}{2} E[R(x)-R(y)]^2 \quad (x,y \in S)$$

Calculons cette espérance mathématique (dans l'hypothèse où les $m^*(x)$ sont effectivement les estimateurs optimaux de la dérive). On a :

$$\gamma_R(x,y) = \frac{1}{2} D^2(Z_x - Z_y) - \text{Cov}(Z_x - Z_y)(m_x^* - m_y^*) + \frac{1}{2} D^2(m_x^* - m_y^*)$$

Par définition, $\frac{1}{2} D^2(Z_x - Z_y) = \gamma(x,y)$. D'après (4-2-6) et (4-2-9), on a aussi :

$$\frac{1}{2} D^2(m_x^* - m_y^*) = \frac{1}{2} \mu_{\ell_S} [f_x^\ell - f_y^\ell] [f_x^S - f_y^S]$$

Evaluons le terme rectangle. En premier lieu, on trouve :

$$(Z_x - Z_y)(m_x^* - m_y^*) = [f_x^\ell - f_y^\ell] [Z_x - Z_y] \int_S \lambda_\ell(dz) Z(z)$$

puis :

$$E[(Z_x - Z_y) \int_S \lambda_\ell(dz) Z(z)] = \int_S \lambda_\ell(dz) [-\gamma(x,z) + \gamma(y,z)]$$

puisqu'il s'agit de combinaisons linéaires autorisées. Comme x et y appartiennent à S , la première relation (4-2-11) donne ensuite :

$$\int_S \lambda_\ell(dz) [-\gamma(x,z) + \gamma(y,z)] = \mu_{\ell_S} (f_x^\ell - f_y^\ell)$$

d'où :

$$E(Z_x - Z_y)(m_x^* - m_y^*) = \mu_{\ell_S} (f_x^\ell - f_y^\ell)(f_x^S - f_y^S)$$

Du fait que les estimateurs sont optimaux, la simplification habituelle se produit, le terme rectangle est égal au terme quadratique, et le variogramme des résidus est :

$$(4-2-14) \quad \gamma_R(x,y) = \gamma(x,y) - \frac{1}{2} \mu_{\ell_S} [f_x^\ell - f_y^\ell] [f_x^S - f_y^S]$$

Comme μ_{ℓ_S} est définie positive, le variogramme des résidus est toujours inférieur (et même nota-

blement inférieur) au variogramme vrai (ou, comme nous dirons, variogramme sous-jacent).

Désignons alors par $k(x)$ l'indicatrice de l'ensemble S des données expérimentales et par $K(h)$ son covariogramme géométrique (dans le cas continu. Si S est discret, on aurait, des résultats analogues). Posons :

$$\gamma^*(h) = \frac{1}{K(h)} \int k(x) k(x+h) [z_{x+h} - z_x - m_{x+h}^* + m_x^*]^2 dx$$

$\gamma^*(h)$ est un estimateur sans biais du variogramme moyen des résidus:

$$E(\gamma^*(h)) = \bar{\gamma}_R(h) = \frac{1}{K(h)} \int k(x) k(x+h) \gamma_R(x, x+h) dx$$

Mais $\bar{\gamma}_R(h)$ diffère profondément du vrai variogramme moyen :

$$\bar{\gamma}(h) = \frac{1}{K(h)} \int k(x) k(x+h) \gamma(x, x+h) dx$$

De la relation (4-2-14) résulte, en effet :

$$(4-2-15) \quad \begin{cases} \bar{\gamma}_R(h) = \bar{\gamma}(h) - \frac{1}{2} \mu_{\ell s} T^{\ell s} \\ T^{\ell s} = \frac{1}{K(h)} \int k(x) k(x+h) [f_{x+h}^{\ell} - f_x^{\ell}][f_{x+h}^s - f_x^s] \end{cases}$$

et par suite $\bar{\gamma}_R(h) < \bar{\gamma}(h)$: on voit que $\gamma^*(h)$ donne une estimation profondément biaisée du vrai variogramme moyen.

On verra, dans l'Exercice 1, paragraphe c/, que ce biais peut être considérable. Ce biais peut être si important qu'il peut inciter (à tort) l'expérimentateur à conclure à l'indépendance des résidus. Cet effet est évidemment général, et se manifeste encore lorsque l'on estime la dérive par un procédé quelconque (non optimal), par exemple par une méthode de moindres carrés : il donne à l'expérimentateur l'impression (erronée) que le phénomène qu'il étudie se réduit à la simple somme d'une dérive et d'un bruit de fond.

4-2-5 Comparaison avec le maximum de vraisemblance.

Dans le cas où l'ensemble S est fini et où la F.A. $Z(x)$ est gaussienne, l'estimateur optimal de la dérive formé au paragraphe 4-2-1 est identique à l'estimateur classique du maximum de vraisemblance. En effet, soient Z_α N variables gaussiennes, $\sigma_{\alpha\beta}$ leur matrice de covariance, $B^{\alpha\beta}$ la matrice inverse, et :

$$E(Z_\alpha) = m_\alpha = a_\ell r_\alpha^\ell$$

les espérances des Z_α , dépendant de $k+1$ paramètres a_ℓ à estimer. On obtient l'estimateur du maximum de vraisemblance en choisissant pour les a_ℓ les valeurs qui rendent maximale l'expression de la densité de probabilité des V.A. Z_α au point z_α correspondant aux valeurs numériques de la réalisation des Z_α . Autrement dit, on minimise la forme quadratique :

$$B^{\alpha\beta} (z_\alpha - m_\alpha)(z_\beta - m_\beta) = B^{\alpha\beta} (z_\alpha - a_\ell r_\alpha^\ell)(z_\beta - a_s r_\beta^s)$$

relativement aux a_ℓ . Désignons par A'_ℓ les valeurs des a_ℓ réalisant ce minimum. On les obtient en résolvant le système :

$$(4-2-16) \quad A'_\ell r_\beta^\ell r_\alpha^s B^{\alpha\beta} = r_\alpha^s B^{\alpha\beta} z_\beta$$

Or le système (4-2-6) donne (en inversant la première relation):

$$(4-2-16') \quad \lambda_\ell^\beta = B^{\alpha\beta} r_\alpha^s \mu_{s\ell}$$

puis (en portant dans la seconde relation) :

$$r_\beta^\ell r_\alpha^s B_{\alpha\beta} \mu_{s\ell} = \delta_{\ell i}$$

Donc, la matrice $\mu_{\ell s}$ des paramètres de Lagrange de (4-2-6) est l'inverse de la matrice $f B f$ du système (4-2-16). La solution de ce système est donc :

$$A'_\ell = \mu_{\ell s} r_\alpha^s B^{\alpha\beta} z_\beta$$

Mais l'estimateur optimal A_ℓ se met, d'après (4-2-16'), sous la forme :

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

On a donc bien l'égalité numérique $A'_{\ell} = A_{\ell}$ à $z_{\beta} = z_{\beta}$ fixé, et donc l'identité des deux estimateurs, comme nous l'avions annoncé.

Remarque - Notre estimateur optimal est plus général que l'estimateur du maximum de vraisemblance, puisqu'il n'est pas lié à une hypothèse gaussienne et se généralise au cas où S est infini. Dans les applications, le système (4-2-6) est plus facile à résoudre, car il ne nécessite qu'une seule inversion de matrice.

4-2-6 Cas d'une dérivée donnée sous forme implicite.

Dans certaines applications, notamment en géodésie, il arrive que l'on ait à estimer une dérivée donnée non pas sous forme explicite $m_{\alpha} = a_{\ell} f_{\alpha}^{\ell}$, mais, sous forme implicite, par un système de relations linéaires que doivent vérifier les m_{α} . Le problème typique est le suivant: on veut estimer N paramètres m_{α} liés par $N-k < N$ relations physiques ou géométriques (par exemple, les 3 côtés et les 3 angles d'un triangle euclidien, grandeurs vérifiant trois relations géométriques). On dispose, pour ce faire, de mesures Z_{α} entachées d'erreurs Y_{α} , soit :

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

On suppose $E Y_{\alpha} = 0$ et on connaît la matrice $\sigma_{\alpha\beta}$ des covariances des erreurs Y_{α} , et on cherche à former à l'aide des Z_{α} le meilleur estimateur m^* possible des paramètres inconnus m_{α} (ce qui implique, en particulier, que les m_{α}^* vérifient eux-mêmes les contraintes imposées aux m_{α}). En général, on connaît au départ une solution suffisamment approchée pour qu'il soit possible de linéariser les contraintes. Autrement dit, on peut trouver une matrice M telle que les contraintes soient de la forme :

$$(4-2-17) \quad M_{\alpha}^u m_{\alpha} = 0 \quad (u = 1, 2, \dots, N-k)$$

et s'arranger pour que la matrice M soit de rang $N-k$. Il est clair que (4-2-17) exprime que le vecteur m appartient à un sous-espace à k dimensions de \mathbb{R}^N . Si l'on fait choix d'une base f^{ℓ} ($\ell = 1, 2, \dots, k$) de ce sous-espace, les conditions (4-2-17) sont équivalentes à l'existence de k coefficients a_{ℓ} tels que :

(4-2-17')

$$m_\alpha = a_\ell r_\alpha^\ell$$

On peut donc ramener ce problème à celui de l'estimation optimale de la dérivée m_α^* . Mais il n'est pas nécessaire d'explicitier les r_α^ℓ , et on peut traiter directement le problème à partir des conditions (4-2-17) données sous forme implicite.

Avant d'établir ce résultat, notons tout de suite que l'estimateur optimal sera (implicitement ou explicitement) de la forme :

$$m_\alpha^* = A_\ell r_\alpha^\ell$$

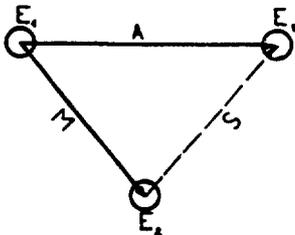
et par conséquent, d'après l'équivalence des relations (4-2-17) et (4-2-17') vérifiera de lui-même les conditions imposées, soit

$$M_u^\alpha m_\alpha^* = 0$$

sans qu'il soit nécessaire d'introduire de contraintes supplémentaires.

Pour effectuer les calculs sous forme implicite, nous aurons besoin d'un lemme algébrique simple :

Lemme algébrique : Soient E_1 , E_2 et E_3 trois espaces vectoriels de dimensions n_1 , n_2 et n_3 respectivement, avec $n_1 \geq n_2$; soient M une application linéaire de E_1 sur E_2 , A une application linéaire de E_1 dans E_3 . Pour que A se laisse factoriser par M (c'est-à-dire pour qu'il existe une application linéaire S de E_2 dans E_3 telle que $A = S \circ M$), il faut et il suffit que le noyau de A contienne le noyau de M (donc que $Mx_1 = 0$, $x_1 \in E_1$ entraîne $Ax_1 = 0$).



Cette condition est évidemment nécessaire. Supposons, inversement, qu'elle soit vérifiée, et soit x_2 un point de E_2 , $x_1 \in M^{-1}(x_2)$ un point de l'image inverse de x_2 dans A_1 . Tout autre point de cette image inverse est de la forme $x_1 + y_1$ avec $My_1 = 0$. On a alors $A(x_1 + y_1) = Ax_1$, puisque $My_1 = 0$ entraîne $Ay_1 = 0$. L'application S de E_2 dans E_3 définie par $Sx_2 = Ax_1$ pour un $x_1 \in M^{-1}(x_2)$ est donc définie sans ambiguïté, est manifestement linéaire, et vérifie par construction $Ax_1 = Sx_2 = SMx_1$ pour tout $x_1 \in E_1$, donc $A = SM$, ce qui achève d'établir le lemme.

Revenons à notre dérive implicite. Pour estimer les m_α , nous allons former des estimateurs linéaires :

$$m_\alpha^* = \Lambda_\alpha^\beta z_\beta$$

vérifiant les deux conditions habituelles d'universalité et d'optimalité.

Condition d'universalité : on doit avoir $E(m_\alpha^*) = m_\alpha$, c'est-à-dire

$$\Lambda_\alpha^\beta m_\beta = m_\alpha$$

quels que soient les paramètres inconnus m_α vérifiant (4-2-17). Autrement dit $M_\alpha^\alpha m_\alpha = 0$ doit entraîner $(\delta_\alpha^\beta - \Lambda_\alpha^\beta) m_\beta = 0$. D'après le lemme, il en est ainsi si et seulement si l'application linéaire $I - A$ de \mathbb{R}^N dans \mathbb{R}^N se laisse factoriser par l'application M de \mathbb{R}^N sur \mathbb{R}^{N-k} , donc si et seulement si il existe une matrice S_α^u telle que :

$$(4-2-18) \quad \delta_\alpha^\beta - \Lambda_\alpha^\beta = S_\alpha^u M_\alpha^\beta$$

Pour exprimer la condition d'optimalité, il reste, pour chaque α fixé, à choisir les $N-k$ coefficients S_α^u qui minimisent la variance $D^2(m_\alpha^*)$. D'après (4-2-18), on a :

$$m_\alpha^* = z_\alpha - S_\alpha^u M_\alpha^\beta z_\beta$$

d'où :

$$D^2(m_\alpha^*) = \sigma_{\alpha\alpha} - 2 S_\alpha^u M_\alpha^\beta \sigma_{(\alpha)\beta} + S_\alpha^u S_\alpha^\gamma M_\alpha^\beta \sigma_{\beta\gamma} M_\alpha^\gamma$$

(sans sommation sur α). En annulant les dérivées partielles en S_α^u , on obtient le système

$$(4-2-19) \quad S_\alpha^\gamma M_\alpha^\beta \sigma_{\beta\gamma} M_\alpha^\gamma = M_\alpha^\beta \sigma_{\alpha\beta}$$

Ce système est nécessairement régulier (puisque le problème algébriquement équivalent formulé avec les f_α^ℓ admet une solution unique). On peut d'ailleurs le vérifier directement (puisque $\sigma_{\beta\gamma}$ est régulière et M_α^β de rang $N-k$). La variance d'estimation correspondante est :

$$(4-2-20) \quad D^2(m_\alpha^*) = \sigma_{\alpha\alpha} - S_\alpha^u M_\alpha^\beta \sigma_{(\alpha)\beta}$$

4-2-7 Comparaison avec les moindres carrés.

Pour estimer les coefficients a_ℓ d'une dérivée $a_\ell f^\ell(x)$, où les fonctions connues $f^\ell(x)$ sont, (par exemple) des polynômes, on a parfois recours à la méthode des moindres carrés. Nous exposerons cette méthode et nous examinerons l'efficacité des estimateurs (non optimaux en général) auxquels elle conduit, en explicitant les calculs dans un cas simple. Enfin, nous examinerons dans quels cas la méthode des moindres carrés conduit au même estimateur optimal que le krigeage universel.

a/ Théorie des moindres carrés. - Soit $Z(x)$ une F.A. admettant une dérivée :

$$m(x) = a_0 + \sum_{\ell=1}^k a_\ell f^\ell(x)$$

avec des coefficients a_ℓ inconnus. On connaît la réalisation de $Z(x)$ sur un ensemble S dont l'indicatrice est $k(x)$ ($=1$ si $x \in S$, $=0$ si $x \notin S$). A partir de cette réalisation, on détermine les $k+1$ nombres B_0, B_1, \dots, B_k réalisant le minimum de l'intégrale :

$$I = \int k(x) [Z(x) - \sum_{\ell=0}^k B_\ell f^\ell(x)]^2 dx$$

Ces nombres vérifient le système obtenu en annulant des dérivées partielles $\frac{\partial I}{\partial B_\ell}$, soit

$$B_j \int k(x) f^j(x) f^\ell(x) dx = \int k(x) Z(x) f^\ell(x) dx$$

Posons

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

$S = \int k(x) dx$ désignant la mesure de S . Le système précédent s'écrit (avec sommation de $j = 0$ à $j = k$) :

$$(4-2-21) \quad B_j T^{j\ell} = \frac{1}{S} \int k(x) Z(x) f^\ell(x) dx$$

Dans les applications, on choisit toujours les fonctions $f^\ell(x)$ de manière à ce que la matrice des $T^{j\ell}$ admette une inverse $S_{j\ell}$. Les B_j sont donc donnés (avec sommation de $\ell = 0$ à $\ell = k$) par :

$$(4-2-22) \quad B_j = \frac{1}{S} S_{j\ell} \int k(x) Z(x) f^\ell(x) dx$$

On voit que ceci peut s'écrire :

$$B_j = \int \lambda_j(dx) Z(x)$$

avec des mesures λ_j dont le support est S et qui admettent les densités :

$$\frac{\lambda_j(dx)}{dx} = \frac{1}{S} k(x) S_{j\ell} f^\ell(x)$$

Ces mesures vérifient les conditions d'universalité (celles qui expriment que l'on a un estimateur sans biais de la dérive). En effet, on trouve :

$$\int \lambda_j(dx) f^s(x) = \frac{1}{S} S_{j\ell} \int k(x) f^\ell(x) f^s(x) dx = S_{j\ell} T^{\ell s}$$

Comme les matrices S et T sont inverses l'une de l'autre, on a bien :

$$(4-2-23) \quad \int \lambda_j(dx) f^s(x) = \delta_j^s \quad (j, s = 0, 1, \dots, k)$$

b/ Espérances et covariances des B_j . - Plaçons-nous d'abord dans le cas le plus simple, où $Z(x)$ admet une fonction de covariance $C(x, y)$. Les B_j , définies par (4-2-22), qui représentaient des nombres lorsque $Z(x)$ désignait la réalisation de la F.A. deviennent des variables aléatoires lorsque $Z(x)$ est considérée comme la F.A. elle-même (les intégrales stochastiques qui figurent au second membre de (4-2-22) ont un sens, puisqu'il existe une covariance). La relation (4-2-23) exprime, comme nous l'avons déjà remarqué, que les B_j sont des estimateurs sans biais des vrais coefficients inconnus a_j : en ce sens, on peut dire que les estimateurs des moindres carrés sont des estimateurs universels (bien qu'ils ne soient pas optimaux en général). On a ainsi :

$$E(B_j) = a_j \quad (j = 0, 1, \dots, k)$$

Calculons maintenant la matrice des covariances

$$\beta_{ij} = E[B_i B_j] - a_i a_j$$

Comme les B_j sont des estimateurs sans biais, on peut faire le calcul comme si on avait $a_j = 0$ (ce qui revient à remplacer $Z(x)$ par $Z(x) - a_j f^\ell(x)$). On déduit de (4-2-22) :

$$B_i B_j = \frac{1}{S^2} S_{i\ell} S_{js} \iint k(x) k(y) Z(x) Z(y) f^\ell(x) f^s(y) dx dy$$

d'où, en passant aux espérances :

$$(4-2-24) \quad \beta_{ij} = \frac{1}{S^2} S_{i\ell} S_{js} \iint k(x) k(y) f^\ell(x) f^s(y) C(x,y) dx dy \quad (i, j = 0, 1, \dots, k)$$

Examinons maintenant le cas où il n'existe qu'un demi-variogramme $\gamma(x,y)$, $Z(x)$ elle-même n'ayant plus de moments d'ordre 2, ni même, en général, d'espérance. Pour $\ell = 0$, l'intégrale stochastique qui figure au second membre de (4-2-22) n'existe plus (au sens de la convergence en moyenne quadratique), car, d'après (4-2-23), on a $\int \lambda_0(dx) = 1$. Cela équivaut à dire que la variance $D^2(B_0)$ est infinie, et que l'espérance $E(B_0)$ n'existe pas non plus en général : on retrouve ici une circonstance que nous avons déjà rencontrée en plusieurs occasions. Par contre, pour $j \neq 0$, la relation (4-2-23), écrite avec $s = 0$, donne $\int \lambda_j(dx) = 0$ ($j = 1, 2, \dots, k$). Cette condition nous garantit que l'intégrale stochastique (4-2-22) existe (au sens de la convergence en moyenne quadratique) et qu'elle admet une variance finie. La matrice des covariances β_{ij} , restreinte aux indices $i, j \neq 0$, est alors donnée par la relation suivante, analogue à (4-2-24) :

$$(4-2-25) \quad \beta_{ij} = -\frac{1}{S^2} S_{i\ell} S_{js} \iint k(x) k(y) f^\ell(x) f^s(y) \gamma(x,y) dx dy \quad (i, j = 1, 2, \dots, k)$$

c/ Exemple du variogramme linéaire. - Prenons $\gamma(x,y) = |x-y|$, plaçons-nous sur la droite réelle, et cherchons à estimer une dérive quadratique $a_0 + a_1x + a_2x^2$ à partir d'une réalisation de la F.A. $Z(x)$ donnée sur l'intervalle $(-R, +R)$.

Il n'y a aucune difficulté à calculer la matrice des T^{ij} et son inverse S_{ij} ($i, j = 0, 1, 2$). On trouve :

$$T = \begin{bmatrix} 1 & 0 & \frac{R^2}{3} \\ 0 & \frac{R^2}{3} & 0 \\ \frac{R^2}{3} & 0 & \frac{R^4}{5} \end{bmatrix} \quad S = \begin{bmatrix} \frac{9}{4} & 0 & -\frac{15}{4R^2} \\ 0 & \frac{3}{R^2} & 0 \\ -\frac{15}{4R^2} & 0 & \frac{45}{4R^4} \end{bmatrix}$$

Les estimateurs des moindres carrés sont donc :

$$\left\{ \begin{aligned} B_0 &= \frac{9}{4} \frac{1}{2R} \int_{-R}^R Z(x) dx - \frac{15}{4R^2} \frac{1}{2R} \int_{-R}^R x^2 Z(x) dx \\ B_1 &= \frac{3}{R^2} \frac{1}{2R} \int_{-R}^R x Z(x) dx \\ B_2 &= -\frac{15}{4R^2} \frac{1}{2R} \int_{-R}^R Z(x) dx + \frac{45}{4R^4} \frac{1}{2R} \int_{-R}^R x^2 Z(x) dx \end{aligned} \right.$$

En ce qui concerne la matrice des covariances restreintes aux indices $i, j = 1, 2$ (β_{00} serait ici infinie), on trouve

$$\beta_{11} = \frac{6}{5R}, \quad \beta_{12} = \beta_{21} = 0, \quad \beta_{22} = \frac{15}{14R^3}$$

On comparera avec l'Exercice 1 ($\frac{6}{5R}$ contre $\frac{1}{R}$ pour l'estimateur optimal, et $\frac{15}{14R^3}$ contre $\frac{3}{4R^3}$) : les estimateurs des moindres carrés sont nettement plus variants que les estimateurs optimaux.

Condition pour que l'estimateur des moindres carrés soit optimal. -

Parmi les estimateurs de la forme $B_j = \int \lambda_j(dx) Z(x)$ (support de λ_j inclus dans S), les estimateurs des moindres carrés sont caractérisés par les deux conditions suivantes :

~ ils sont sans biais ($E(B_j) = a_j, \forall a_j$)

~ les mesures λ_j ont des densités qui sont des combinaisons linéaires des fonctions $k(x) f^\ell(x)$.

On a déjà vu que l'estimateur des moindres carrés vérifie ces conditions. Inversement, supposons que des mesures λ_j vérifient les conditions d'universalité (4-2-23). Si les λ_j ont des densités de la forme $\frac{k(x)}{S} S'_{j\ell} f^\ell(x)$, ces conditions (4-2-23) expriment que la matrice des $S'_{j\ell}$ coïncide avec l'inverse $S_{j\ell}$ de la matrice des T^{1j} , de sorte que les estimateurs B_j sont bien donnés par (4-2-22).

Un estimateur vérifiant ces deux propriétés vérifie par définition les conditions d'universalité. Pour qu'il soit optimal, il faut et il suffit qu'il satisfasse également aux premières relations (4-2-6), c'est-à-dire que l'on puisse trouver des constantes $\mu_{j\ell}$ telles que :

$$(4-2-26) \quad \frac{1}{S} S_{j\ell} \int k(x) f^\ell(x) C(x,y) dx = \mu_{j\ell} f^\ell(y), \quad \forall y \in S$$

Il en résulte que l'estimateur des moindres carrés coïncide avec l'estimateur optimal de la dérive si et seulement si les $k+1$ fonctions $f^\ell(x)$ sont des combinaisons linéaires de $k+1$ fonctions propres distinctes du noyau $k(x,y) C(x,y) k(y)$, considéré comme un opérateur agissant sur les fonctions à support dans S .

En particulier, si les $f^\ell(x)$ sont elles-mêmes des fonctions propres ($\int k(x) f^\ell(x) C(x,y) dx = \mu^\ell f^\ell(y)$ pour $y \in S$), la relation (4-2-26) se réduit à :

$$(4-2-27) \quad \mu_{js} = \frac{1}{S} S_{js} \mu^s \quad (\text{sans sommation en } s)$$

[Dans le cas où il n'existe pas de covariance $C(x,y)$, mais seulement un variogramme $\gamma(x,y)$, la condition est un tout petit peu moins simple. Si l'on désigne par $\bar{Z} = \frac{1}{S} \int k(x) Z(x) dx$ la moyenne (spatiale) de $Z(x)$ sur S , $Z(x) - \bar{Z}$ est une F.A. admettant des moments d'ordre 1 et 2. Sa covariance $\bar{C}(x,y)$ est donnée par :

$$\begin{aligned} \bar{C}(x,y) = & -\gamma(x,y) + \frac{1}{S} \int k(y) \gamma(x,y) dx + \frac{1}{S} \int k(x) \gamma(x,y) dx \\ & - \frac{1}{S^2} \iint k(x) k(y) \gamma(x,y) dx dy \end{aligned}$$

On peut alors montrer que les B_j des moindres carrés ($j = 1, 2, \dots, k$) coïncident avec l'estimateur optimal si et seulement si les k fonctions $f^\ell(x)$ ($\ell = 1, 2, \dots, k$) sont des combinaisons linéaires de k fonctions propres du noyau $\bar{C}(x,y)$ correspondant à des valeurs propres non nulles].

Cette condition est de nature assez complexe, puisqu'elle fait intervenir à la fois les fonctions $f^\ell(x)$, le noyau $C(x,y)$ et la géométrie de l'ensemble S où les données sont disponibles. Mais il y a un cas où cette condition est automatiquement vérifiée. C'est le cas où la partie aléatoire de $Z(x)$ se réduit à un simple bruit de fond, (ou, en terminologie géostatistique, à un effet de pépité pur). Dans ce cas, la covariance $C(x,y)$ se réduit à $\omega \delta(x-y)$, δ désignant la mesure de Dirac et ω un facteur constant : toute fonction $f^\ell(x)$ continue sur S est alors fonction propre (pour la valeur propre ω) et cela quel que soit l'ensemble S . Dans ce cas, l'estimateur des moindres carrés est toujours optimal, et les covariances μ_{js} , d'après (4-2-27) sont données par :

$$(4-2-28) \quad \mu_{js} = \frac{\omega}{S} S_{js}$$

En fait, une covariance $\omega \delta(x-y)$ caractérise une mesure, et non plus une fonction aléatoire. Ce

n'est d'ailleurs pas gênant si l'ensemble S est continu, puisqu'en fait on ne travaille alors que sur des régularisées de cette mesure aléatoire (qui sont à nouveau des F.A.). En pratique, il suffit pour que l'estimateur des moindres carrés soit (presque) optimal que la covariance soit de la forme $C(x-y)$ avec une portée très faible (par exemple, $C(h) = e^{-ah}$ avec $\frac{1}{a}$ petit relativement aux dimensions de S). En particulier, si l'ensemble S est fini, il suffira que cette portée soit inférieure à la plus petite des distances séparant les points de S .

Dans ce cas (covariance $C(h)$ de portée très petite), l'estimateur des moindres carrés est (presque) optimal, et les covariances μ_{ij} sont encore données par la relation (4-2-28), à condition de prendre comme facteur ω l'intégrale de $C(h)$ étendue à l'espace entier, soit :

$$\mu_{ij} = \frac{S_{ij}}{S} \int C(h) dh$$

S_{ij} désignant toujours l'inverse de

$$T^{jS} = \frac{1}{S} \int k(x) r^j(x) r^S(x) dx$$

On trouvera dans l'exercice 6 des indications sur le cas où l'ensemble S est fini.

4-3 LE KRIGEAGE

4-3-1 Les équations du krigeage universel.

Dans le problème du krigeage proprement dit, on se propose d'estimer une moyenne pondérée du type $Z_0 = \int p(dx) Z(x)$ (avec $\int p(dx) = 1$) à partir des valeurs expérimentales connues sur un ensemble S . Nous écrirons les équations dans le cas où l'ensemble S est fini, et, moyennant les réserves habituelles, nous les transposerons ensuite au cas où S est infini.

Supposons d'abord qu'il existe une covariance centrée (connue) $\sigma(x,y)$ et une dérive $a_\ell r^\ell(x)$ ($\ell = 0, 1, \dots, k$, a_ℓ inconnues). Nous allons former un estimateur linéaire de Z_0 , soit

$$Z_U = Z_U^\alpha Z_\alpha \quad (\text{ou } Z_U = \int_S \lambda_U(dx) Z(x))$$

et lui imposer les conditions habituelles :

Condition d'universalité : on veut avoir $E(Z_U - Z_0) = 0$ quelles que soient les valeurs inconnues des coefficients a_ℓ , ce qui donne :

$$\begin{cases} \lambda^\alpha f_\alpha^\ell = b^\ell \\ (b^\ell = \int f^\ell(x) p(dx)) \end{cases}$$

Condition d'optimalité : compte tenu de la condition précédente, on choisit les coefficients $\lambda_{\alpha\beta}^\ell$ qui minimisent la variance d'estimation

$$\begin{cases} D^2(Z_0 - Z_U) = \sigma_{Z_0}^2 - 2 \lambda_U^\alpha \sigma_{Z_0, \alpha} + \lambda^\alpha \lambda^\beta \sigma_{\alpha\beta} \\ \sigma_{Z_0, \alpha} = \int p(dx) \sigma(x, x_\alpha) \\ \sigma_{Z_0}^2 = \int \int p(dx) \sigma(x, y) p(dy) \end{cases}$$

On obtient ainsi le système suivant, où figurent des paramètres de Lagrange μ_ℓ associés aux conditions d'universalité, et où la variance σ_U^2 est la valeur optimale de la forme quadratique $D^2(Z_0 - Z_U)$, calculée en tenant compte des deux premières relations :

$$(4-3-1) \quad \begin{cases} \lambda_U^\beta \sigma_{\alpha\beta} = \sigma_{Z_0, \alpha} + \mu_\ell f_\alpha^\ell \\ \lambda_U^\alpha f_\alpha^\ell = b^\ell \\ \sigma_U^2 = \sigma_{Z_0}^2 - \lambda^\alpha \sigma_{Z_0, \alpha} + \mu_\ell b^\ell \end{cases}$$

Ce système est toujours régulier (pourvu que $\sigma_{\alpha\beta}$ soit strictement définie positive et que les f_α^ℓ soient linéairement indépendants sur S).

Ce système se transpose sans difficulté au cas où il existe seulement un variogramme γ . En effet, pour $\ell = 0$, on a, comme d'habitude, $f^0 = 1$. La condition d'universalité relative $\ell = 0$ se réduit donc à :

$$\sum_\alpha \lambda^\alpha = 1$$

Or (exactement comme dans le cas du krigeage simple, paragraphe 3-5) c'est là précisément la condition pour que l'erreur :

$$Z_0 - Z_U = \int p(dx) Z(x) - \lambda^\alpha Z_\alpha$$

soit une combinaison linéaire autorisée (c'est-à-dire admette une variance finie). Comme $\int p(dx) = 1$, en effet, la somme $1 - \sum \lambda^\alpha$ des coefficients de la combinaison linéaire $Z_0 - Z_U$ doit, comme d'habitude, être égale à 0. Le terme constant a_0 de la dérive, sur lequel se concentrerait la difficulté lors de l'estimation optimale de la dérive elle-même, est éliminé automatiquement de l'expression $Z_0 - Z_U$ et n'intervient plus dans le problème du krigeage : c'est exactement là le sens de la condition $\sum \lambda^\alpha = 1$.

Transposons donc le système (4-3-1) en termes de variogramme, en adoptant, pour changer, les notations du cas continu.

$$(4-3-2) \quad \begin{cases} \int_S \lambda(dx) \gamma(x,y) = \int p(dx) \gamma(x,y) - \mu_\ell r^\ell(y) & (\forall y \in S) \\ \int_S \lambda(dx) r^\ell(x) = \int p(dx) r^\ell(x) & (\ell = 0, 1, \dots, k) \\ \sigma_u^2 = - \iint p(dx) \gamma(x,y) p(dy) + \iint \lambda(dx) p(dy) \gamma(x,y) + \mu_\ell \int p(dx) r^\ell(x) \end{cases}$$

Krigeage ponctuel. - Si l'on veut estimer la valeur $Z(x_0)$ de la réalisation en un point x_0 , le système précédent se simplifie, et devient (avec les notations du cas discret) :

$$(4-3-3) \quad \begin{cases} \lambda^\beta \gamma_{\alpha\beta} = \gamma(x_0, x_\alpha) - \mu_\ell r_\alpha^\ell \\ \lambda^\alpha r_\alpha^\ell = r^\ell(x_0) \\ \sigma_u^2 = \lambda^\alpha \gamma(x_\alpha, x_0) + \mu_\ell r^\ell(x_0) \end{cases}$$

On note que le krigeage universel ponctuel est un interpolateur exact. On peut le vérifier directement en faisant $x_0 = x_{\alpha_0} \in S$ dans (4-3-3). Il est plus simple de remarquer que Z_{α_0} vérifie $E(Z_{\alpha_0} - Z_{\alpha_0}) = 0$ et $D^2(Z_{\alpha_0} - Z_{\alpha_0}) = 0$ quels que soient les a_ℓ , donc satisfait aux conditions d'universalité et d'optimalité.

4-3-2 Le théorème d'additivité.

Si la dérive $m(x)$ était connue, l'estimateur optimal de $Z_0 = \int p(dx) Z(x)$ serait :

$$(4-3-4) \quad \lambda_K^\alpha (Z_\alpha - m_\alpha) + \int m(x) p(dx)$$

avec les coefficients λ_K^α du krigeage simple (paragraphe 3-4-1), soit :

$$(4-3-5) \quad \begin{cases} \lambda_K^\beta \sigma_{\alpha\beta} = \sigma_{\alpha, Z_0} \\ \sigma_K^2 = \sigma_{Z_0}^2 - \lambda_K^\alpha \sigma_{\alpha, Z_0} \end{cases}$$

Montrons que la circonstance déjà observée au paragraphe (3-4-2) est générale : autrement dit, l'estimateur optimal universel se met sous la forme :

$$(4-3-6) \quad Z_U = \lambda_K^\alpha (Z_\alpha - m_\alpha^*) + \int m^*(x) p(dx)$$

obtenue en remplaçant dans (4-3-4) l'expression inconnue de la dérive par son estimation optimale $m^*(x)$. Cela signifie, si l'on veut, que l'on a le droit de kriger les résidus estimés $Z(x) - m^*(x)$ comme s'il s'agissait des résidus vrais.

Pour vérifier (4-3-6), posons

$$Z_U = Z_K + Z_D, \quad Z_K = \lambda_K^\alpha Z^\alpha$$

Le terme Z_D représente la correction de dérive. Il est de la forme $Z_D = \lambda_D^\alpha Z_\alpha$, avec $\lambda_D^\alpha = \lambda_U^\alpha - \lambda_K^\alpha$. Cherchons les conditions que doivent vérifier les λ_D^α . Si nous remplaçons λ_U^α par $\lambda_K^\alpha + \lambda_D^\alpha$ dans la première relation (4-3-1), il résulte de (4-3-5) que les λ_K^α et σ_{α, Z_0} disparaissent. Il reste finalement le système suivant (équivalent à (4-3-1)) :

$$\begin{cases} \lambda_D^\beta \sigma_{\alpha\beta} = \mu_\ell f_\alpha^\ell \\ \lambda_D^\alpha f_\alpha^\ell = \int p(dx) f^\ell(x) - \lambda_K^\alpha f_\alpha^\ell \end{cases}$$

Or ce système est identique à (4-2-2), à ceci près que $f^\ell(x_0)$ a été remplacé par $\int [p(dx) - \lambda_K(dx)] f^\ell(x)$. Par conséquent, si nous posons :

$$(4-3-7) \quad b^\ell = \int p(dx) f^\ell(x) - \lambda_K^\alpha f_\alpha^\ell$$

les λ_D^α se déduisent des λ_ℓ^α du système délinéarisé (4-2-6) par les relations :

$$\lambda_D^\alpha = \lambda_\ell^\alpha b^\ell$$

Le terme Z_D représentant la correction de dérive est donc $\lambda_\ell^\alpha Z_\alpha b^\ell = A_\ell b^\ell$, en désignant par A_ℓ l'estimateur optimal de la dérive, et, en remplaçant b^ℓ par son expression (4-3-7), ceci donne exactement :

$$Z_D = \int p(dx) m^*(x) - \lambda_K^\alpha m_\alpha^*$$

Par suite, Z_U lui-même est bien de la forme (4-5-6) :

$$Z_U = Z_K + Z_D = \lambda_K^\alpha (Z_\alpha - m_\alpha^*) + \int m^*(x) p(dx)$$

Toujours comme dans le cas du krigeage simple, ce théorème d'additivité s'étend aux variances, soit :

$$(4-3-7) \quad D^2(Z_0 - Z_U) = \sigma_K^2 + \sigma_D^2$$

avec $\sigma_D^2 = D^2(Z_D)$ et en désignant par σ_K^2 la variance du krigeage simple. En effet, le système (4-3-5) exprime exactement que $Z_0 - Z_K$ est en covariance nulle avec les combinaisons linéaires des Z_α , donc, en particulier, avec Z_D .

On vérifie sans peine que ce théorème d'additivité subsiste lorsqu'il n'y a pas de covariance, mais seulement un variogramme (en effet, pour $\ell = 0$, la condition d'universalité relative à λ_D^α s'écrit :

$$\sum_\alpha \lambda_D^\alpha = \int p(dx) - \int \lambda_K(dx) = 0$$

et Z_D est une combinaison linéaire autorisée).

4-3-3 Le krigeage considéré comme interpolateur.

Nous avons déjà remarqué que le krigeage (simple ou universel) constitue un interpolateur exact. C'est là une propriété importante dans certaines applications telles que la cartographie. Du point de vue méthodologique, il n'est peut-être pas inutile d'oublier pour un instant le contenu probabiliste de la théorie, et d'analyser d'un point de vue purement pragmatique les manipulations auxquelles on soumet effectivement les données numériques expérimentales lorsque l'on

krige un point x .

a/ Définition des interpolateurs. - Soit $S \subset \mathbb{R}^N$ l'ensemble des points expérimentaux, et V un domaine contenant S et pouvant coïncider avec \mathbb{R}^N lui-même. On appellera interpolateur un opérateur linéaire T appliquant un espace $\mathcal{F}(S)$ de fonctions définies sur S dans un espace $\mathcal{F}(V)$ de fonctions définies sur V : $f \in \mathcal{F}(S) \rightarrow T f \in \mathcal{F}(V)$. Nous dirons que T est un interpolateur exact si la restriction à S de $T f$ est identique à f .

Dans ce qui suit, nous nous limiterons au cas où l'ensemble $S = \{x_\alpha, \alpha = 1, \dots, N\}$ est fini. Une fonction f sur S est alors un vecteur (f_α) , $\alpha = 1, \dots, N$, et l'opérateur T est défini par la donnée des $T e^\alpha$, e^α constituant une base de l'espace à N dimensions constitué de ces vecteurs. Par exemple prenons pour e^α le vecteur de composantes $e_\beta^\alpha = \delta_\beta^\alpha$, et soit $T^\alpha(x)$ l'image de ce vecteur. Pour un vecteur f quelconque, on a alors :

$$(4-3-8) \quad T f(x) = f_\alpha T^\alpha(x)$$

Pour que l'interpolateur T soit exact, il faut et il suffit que la matrice $T_\beta^\alpha = T^\alpha(x_\beta)$ vérifie $f_\alpha T_\beta^\alpha = f_\beta$ pour tout $f \in \mathbb{R}^N$, donc vérifie

$$(4-3-9) \quad T_\beta^\alpha = \delta_\beta^\alpha$$

Exemple : Le krigeage simple est l'interpolateur λ_K défini par les $\lambda_K^\alpha(x)$ obtenus en résolvant le système :

$$(4-3-10) \quad \lambda_K^\beta(x) \sigma_{\alpha\beta} = \sigma_{\alpha,x}$$

Nous désignerons par $B^{\alpha\beta}$ la matrice inverse de $\sigma_{\alpha\beta}$. On a donc :

$$(4-3-11) \quad \lambda_K^\beta(x) = B^{\alpha\beta} \sigma_{\alpha,x}$$

Pour $x = x_\gamma$, on trouve $\lambda_K^\beta(x_\gamma) = B^{\alpha\beta} \sigma_{\alpha\gamma} = \delta_\gamma^\beta$, et le krigeage simple est un interpolateur exact. De plus, les composantes $\lambda_K^\beta(x)$ sont, d'après (4-3-11), des combinaisons linéaires des N fonctions $\sigma_{\alpha,x} = \sigma(x_\alpha, x)$.

Ces deux propriétés sont caractéristiques : le krigeage simple est le seul interpolateur exact dont les composantes sont des combinaisons linéaires des fonctions $\sigma_{\alpha,x}$.

En effet, soit T un interpolateur de la forme :

$$T^\beta(x) = H^{\beta\alpha} \sigma_{\alpha,x}$$

Pour qu'il soit exact, il faut et il suffit que l'on ait $T^\beta_\alpha = \delta^\beta_\alpha$, soit :

$$\delta^\beta_\alpha = H^{\beta\gamma} \sigma_{\alpha\gamma}$$

Donc $H^{\beta\gamma} = \delta^{\beta\gamma}$, et $T^\beta(x) = \lambda^\beta_K(x)$.

Pour caractériser le krigeage universel, nous aurons besoin d'une notion nouvelle :

b/ Dérives compatibles avec un interpolateur. - Nous dirons qu'une fonction $f(x)$, $x \in V$ est une dérive compatible avec un interpolateur T si l'interpolation de f par T est exacte, autrement dit si l'on a :

$$(4-3-12) \quad f_\beta T^\beta(x) = f(x)$$

On reconnaît en (4-3-12) la condition familière d'universalité. Cette condition entraîne, pour $x = x_\alpha$:

$$(4-3-13) \quad T^\beta_\alpha f_\beta = f_\alpha$$

Autrement dit f_α doit être vecteur propre pour la valeur propre 1. Inversement, soit f_α un vecteur vérifiant (4-3-13). La fonction $f(x)$ définie sur V par :

$$f(x) = f_\beta T^\beta(x)$$

est, par construction, une dérive compatible avec T. Ainsi (4-3-12) établit une correspondance bijective entre les dérivées compatibles et les vecteurs propres associés à la valeur propre 1.

Pour que T soit un interpolateur exact, il faut et il suffit que l'on ait $T^\alpha_\beta = \delta^\alpha_\beta$, et il en est ainsi si et seulement si tous les vecteurs f_α sont vecteurs propres pour la valeur propre 1. Par conséquent, pour que T soit un interpolateur exact il faut et il suffit qu'il existe N dérivées compatibles avec T et linéairement indépendantes sur S.

c/ Caractérisation du Krigeage Universel. - Dans le cas du krigeage simple, ces N dérivées sont les fonctions $\sigma_{\alpha, x}$: cela résulte du système (4-3-9) lui-même. Considérons le krigeage universel dans le cas où la dérivée est de la forme $a_{\ell} f^{\ell}(x)$, avec $\ell = 1, 2, \dots, k$ ($k \leq N$). Il définit un interpolateur λ dont les composantes $\lambda^{\alpha}(x)$ constituent la solution du système :

$$(4-3-11) \quad \begin{cases} \lambda^{\beta}(x) \sigma_{\alpha\beta} = \sigma_{\alpha, x} + \mu_{\ell} f_{\alpha}^{\ell} \\ \lambda^{\alpha}(x) f_{\alpha}^{\ell} = f^{\ell}(x) \quad (\ell = 1, 2, \dots, k) \end{cases}$$

Les conditions d'universalité nous indiquent que les k fonctions f^{ℓ} sont des dérivées compatibles avec cet interpolateur. Etant entendu que ces k fonctions f^{ℓ} sont linéairement indépendantes sur S , nous connaissons donc un espace vectoriel de dimensions k de dérivées compatibles avec λ . Il reste, pour caractériser le krigeage universel, à déterminer un espace supplémentaire de dimension $N-k$.

Désignons par f^{ν} , $\nu = k+1, \dots, N$ une base de cet espace supplémentaire. Les $N-k$ dérivées manquantes seront donc les $f_{\beta}^{\nu} \lambda^{\beta}(x)$. Pour déterminer les f^{ν} , nous allons utiliser le théorème d'additivité du paragraphe 4-3-2. D'après ce théorème, l'interpolation des données Z_{α} conduit à la fonction :

$$Z_{\nu}(x) = \lambda_{\kappa}^{\alpha}(x) (Z_{\alpha} - A_{\ell} f_{\alpha}^{\ell}) + A_{\ell} f^{\ell}(x)$$

avec des coefficients A_{ℓ} de la forme :

$$A_{\ell} = \lambda_{\ell}^{\beta} Z_{\beta}$$

où la matrice λ_{ℓ}^{β} est la solution du système (4-2-6). En explicitant l'expression de $Z_{\nu}(x)$, on trouve :

$$Z_{\nu}(x) = \lambda_{\kappa}^{\beta}(x) (\delta_{\beta}^{\alpha} - \lambda_{\ell}^{\alpha} f_{\beta}^{\ell}) Z_{\alpha} + \lambda_{\ell}^{\alpha} f^{\ell}(x) Z_{\alpha}$$

Ainsi, les coefficients $\lambda^{\beta}(x)$ du système (4-3-11) admettent les expressions suivantes :

$$\lambda^{\alpha}(x) = \lambda_{\kappa}^{\beta}(x) (\delta_{\beta}^{\alpha} - \lambda_{\ell}^{\alpha} f_{\beta}^{\ell}) + \lambda_{\ell}^{\alpha} f^{\ell}(x)$$

On reconnaît l'expression de la matrice U :

$$U_{\beta}^{\alpha} = \delta_{\beta}^{\alpha} - \lambda_{\ell}^{\alpha} f_{\beta}^{\ell}$$

introduite au paragraphe 4-2-4 (le résidu $Z_{\alpha} - m_{\alpha}^*$ est $U_{\alpha}^{\beta} Z_{\beta}$). Ainsi :

$$(4-3-12) \quad \lambda^{\alpha}(x) = \lambda_{\ell}^{\beta}(x) U_{\beta}^{\alpha} + \lambda_{\ell}^{\alpha} f^{\ell}(x)$$

D'après la forme des conditions d'universalité ($\lambda_{\ell}^{\alpha} f_{\alpha} = \delta_{\ell}^{\beta}$) il est possible de choisir la base f^{ν} des dérivées manquantes dans l'orthogonal des λ_{ℓ}^{α} ($\ell = 1, 2, \dots, k$), c'est-à-dire de s'imposer les relations :

$$(4-3-13) \quad \lambda_{\ell}^{\alpha} f_{\alpha}^{\nu} = 0$$

Moyennant cette condition, il résulte de (4-3-13) que les $N-k$ dérivées manquantes sont les fonctions

$$f^{\nu}(x) = \lambda^{\beta}(x) f_{\beta}^{\nu} = \lambda_{\ell}^{\beta}(x) f_{\beta}^{\nu}$$

soit :

$$f^{\nu}(x) = f_{\beta}^{\nu} B^{\alpha\beta} \sigma_{\alpha, x}$$

Elles s'expriment linéairement à l'aide des fonctions $\sigma_{\alpha, x}$, soit :

$$(4-3-14) \quad \begin{cases} f^{\nu}(x) = D^{\alpha} \sigma_{\alpha, x} \\ D^{\alpha} = f_{\beta}^{\nu} B^{\alpha\beta} \end{cases}$$

Il reste à caractériser l'espace à $(n-k)$ dimensions des vecteurs D^{α} tels que les $D^{\alpha} \sigma_{\alpha, x}$ soient des dérivées compatibles. Pour cela, considérons les conditions (4-3-13), en remarquant que les λ_{ℓ}^{α} du système (4-2-6) sont données par :

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

La relation $f_{\alpha}^{\nu} \lambda_{\ell}^{\alpha} = 0$ équivaut donc (puisque la matrice $\mu_{\ell s}$ des paramètres de Lagrange est de type positif strict) à la relation :

$$f_{\alpha}^{\nu} B^{\alpha\beta} f_{\beta}^s = 0$$

D'après (4-3-14), ceci équivaut encore à :

$$(4-3-15) \quad D^\alpha f_\alpha^\ell = 0$$

Par conséquent, l'espace supplémentaire à N-k dimensions est constitué des fonctions de la forme $D^\alpha \sigma_{\alpha, x}$ pour des vecteurs D^α orthogonaux aux k vecteurs f^ℓ , c'est-à-dire vérifiant (4-3-15).

d/ Contenu probabiliste du krigeage universel. - Comme tout interpolateur exact, le krigeage universel est caractérisé par l'espace vectoriel (de dimension N) des dérivées compatibles avec lui. Cet espace comporte

1 - Un espace imposé, de dimension k (celui des fonctions $a_\ell f^\ell$) dont la signification est purement interpolatoire (non probabiliste)

2 - Un espace supplémentaire de dimension N-k (celui des fonctions $D^\alpha \sigma_{\alpha, x}$, avec la condition d'orthogonalité $D^\alpha f_\alpha^\ell = 0$ exprimant que le vecteur D décrit l'orthogonal du premier espace) de signification purement probabiliste.

Autrement dit, lorsque le nombre k des dérivées imposées augmente, le contenu probabiliste du krigeage universel décroît tandis que son contenu purement interpolatoire augmente.

En particulier, pour $k = 0$, le krigeage simple est purement probabiliste. Pour $k = N$, au contraire, le contenu probabiliste est nul, et le krigeage universel se réduit au procédé purement interpolatoire qui consiste à faire passer par N points expérimentaux une combinaison linéaire de N fonctions choisies d'avance.

4-4 KRIGEAGE UNIVERSEL POUR UNE DERIVE ALEATOIRE.

4-4-1 Hypothèses.

Au lieu de considérer la dérive comme une fonction $m(x)$ inconnue mais non aléatoire, on peut aussi envisager le cas où $m(x)$ est une fonction aléatoire $m(x, \omega)$, autrement dit, étudier une F.A. $Z(x, \omega)$ de la forme :

$$(4-4-1) \quad Z(x, \omega) = Y(x, \omega) + m(x, \omega)$$

avec $E Y(x) = 0$. Posons :

$$\begin{cases} E Y(x) Y(y) = \sigma(x,y) \\ E m(x) m(y) = K(x,y) \\ E Y(x) m(y) = R(x,y) \end{cases}$$

(la seconde covariance est non centrée).

Il n'est pas nécessaire que les F.A. Y et m soient indépendantes. Mais, pour que la dichotomie écrite en (4-3-1) présente une signification physique réelle, et non un pur artefact, il faut que chacune des deux composantes Y et m prennent en charge des traits structuraux se rapportant à des échelles nettement différenciées (structures gigogne de J. SERRA). La F.A. m doit avoir une portée (infinie ou) très grande vis-à-vis de celle de Y , et, de plus, à l'échelle de la portée de Y , la variabilité de m doit être très réduite. A l'échelle de Y , la réalisation $m(x,\omega)$ doit ainsi pouvoir être assimilée à une fonction très régulière et continue, de sorte que - sur un certain domaine V - on doit pouvoir écrire :

$$(4-4-2) \quad m(x,\omega) = a_\ell(\omega) f^\ell(x) \quad (x \in V)$$

les $a_\ell(\omega)$ désignant des V.A. (qui dépendent du domaine V), et les f^ℓ les fonctions de base habituelles (avec $f^0 = 1$), par exemple des polynomes, ou, plus généralement, des fonctions continues et suffisamment régulières, linéairement indépendantes sur V .

Examinons la signification de cette hypothèse (4-4-2). Nous poserons :

$$(4-4-3) \quad K_{\ell s} = E a_\ell a_s$$

de sorte que la covariance $K(x,y)$ de la F.A. $m(x)$ admet, dans le domaine V où l'approximation (4-4-2) est valable, la représentation :

$$(4-4-4) \quad K(x,y) = K_{\ell s} f^\ell(x) f^s(y) \quad (x,y \in V)$$

Pour la covariance $R(x,y)$, nous trouvons

$$R(x,y) = E Y(x) a_\ell f^\ell(y) = R_\ell(x) f^\ell(y) \quad (x,y \in V)$$

Si la F.A. Y peut être considérée comme stationnaire sur V , $R_\ell(x)$ est une constante R_ℓ . Plus généralement, $R_\ell(x)$ admettra sur V une bonne représentation approchée au moyen des mêmes fonctions $f^s(x)$, soit

$$R(x) \approx R_\ell f^s(x) \quad (x \in V)$$

(cette hypothèse n'est pas absolument nécessaire, mais elle simplifie beaucoup l'exposé) d'où finalement pour $R(x,y)$ un résultat analogue à (4-4-4):

$$(4-4-4') \quad R(x,y) \approx R_\ell f^\ell(x) f^s(y) \quad (x,y \in V)$$

Il apparaît d'ailleurs que les relations (4-4-4) et (4-4-4') sont les seules hypothèses réellement requises pour ce qui suit, et que la relation (4-4-2) n'en est qu'une traduction imagée. D'un point de vue purement pragmatique, elles signifient seulement que les fonctions $K(x,y)$ et $R(x,y)$ varient assez lentement dans l'espace à l'échelle du domaine V de référence pour qu'on puisse les remplacer sur V par les premiers termes de leurs développements limités en $f^\ell(x) f^s(y)$. L'exemple suivant montrera le sens de cette hypothèse.

Exemple d'une exponentielle de Gauss. - Pour alléger les écritures, plaçons-nous dans \mathbb{R}^1 (on aurait évidemment des résultats analogues dans \mathbb{R}^n) et supposons que la covariance K soit de la forme :

$$K(x,y) = B e^{-\frac{(x-y)^2}{2b^2}}$$

avec une portée b plus grande que la longueur L de l'intervalle de référence V . Développons cette exponentielle de Gauss :

$$K(x,y) = B \left[1 - \frac{(x-y)^2}{2b^2} + \frac{1}{8} \frac{(x-y)^4}{b^4} - \frac{1}{48} \frac{(x-y)^6}{b^6} + \dots \right]$$

chacun des termes $(x-y)^{2n}$ se développe à son tour en monomes $x^\ell y^s$, et l'on voit que $K(x,y)$ admet effectivement un développement limité de type (4-4-4) avec $f^\ell(x) = x^\ell$.

Dans le cas d'une "dérive quadratique", c'est-à-dire $\ell = 0, 1, 2$, on arrête ce développement au terme d'ordre 4, et l'erreur est majorée par

$$\frac{B}{48} \frac{(x-y)^6}{b^6} \leq \frac{4}{3} B \frac{L^6}{b^6}$$

Pour apprécier cette majoration, remarquons que B peut être très grand (voire, à la limite, infini, si m n'a pas de covariance, mais seulement un variogramme), mais B/b^2 est obligatoirement fini, puisqu'il représente le coefficient du terme en h^2 du variogramme de m. Posant $B/b^2 = B_2$, on voit que la majoration est en $B_2 L^2 (\frac{L}{b})^4$, donc d'ordre au moins 4 en L/b . Si donc la portée b est effectivement grande vis-à-vis de L, l'approximation faite sera tout-à-fait légitime.

Remarque - D'après (1), la covariance de Z serait $\sigma + 2R + K$. On peut se demander si cette dichotomie est vraiment utile, et s'il ne serait pas plus simple de faire du krigeage ou du cokrigeage ordinaire directement sur cette covariance. La réponse est que (dans les conditions de nos hypothèses) l'estimation de cette covariance n'est pas possible dans des conditions acceptables. La composante $K(x,y)$ n'est, en général, pas une covariance stationnaire. Même si, localement, (c'est-à-dire à l'échelle de V) on peut la mettre sous la forme $K(x-y)$, l'inférence statistique à partir de données couvrant un domaine de l'ordre de grandeur de V ne sera pas réalisable en général pour une covariance K trop régulière (voir par. 2-10).

4-4-2 Estimation d'une dérive.

Désignons par Z_α les valeurs de $Z(x)$ aux points expérimentaux $x_\alpha \in S \subset V$. Nous supposons S fini, mais ce qui suit se transpose sans difficulté au cas où S est infini.

Nous voulons estimer la dérive $m(x)$ en un point $x \in V$, et pour cela nous formons un estimateur de la forme

$$M^*(x) = \lambda^\alpha Z_\alpha = \lambda^\alpha Y_\alpha + \lambda^\alpha m_\alpha$$

L'espérance du carré de l'erreur est donc :

$$\begin{aligned} E(M_x^* - m_x)^2 &= E \left(\lambda^\alpha Y_\alpha + (\lambda^\beta m_\beta - m_x) \right)^2 = \\ &= \lambda^\alpha \lambda^\beta \sigma_{\alpha\beta} + 2 \lambda^\alpha (\lambda^\beta R_{\alpha\beta} - R_{\alpha x}) + \lambda^\alpha \lambda^\beta K_{\alpha\beta} - 2 \lambda^\alpha K_{\alpha x} + K_{xx} \end{aligned}$$

Les hypothèses (4-4-4) et (4-4-4') donnent :

$$\lambda^\beta R_{\alpha\beta} - R_{\alpha\alpha} = R_{\ell_S} (\lambda^\beta r_\beta^S - r_X^S) r_\alpha^\ell$$

$$\lambda^\alpha \lambda^\beta K_{\alpha\beta} - 2 \lambda^\alpha K_{\alpha\alpha} + K_{XX} = K_{\ell_S} (\lambda^\alpha \lambda^\beta r_\alpha^\ell r_\beta^S - \lambda^\alpha r_\alpha^\ell r_X^S -$$

$$- \lambda^\beta r_X^\ell r_\beta^S + K_{XX}) = K_{\ell_S} (\lambda^\alpha r_\alpha^\ell - r_X^\ell) (\lambda^\beta r_\beta^S - r_X^S)$$

Ainsi :

$$(4-4-5) \quad \left\{ \begin{aligned} E(m_X^* - m_X)^2 &= \lambda^\alpha \lambda^\beta \sigma_{\alpha\beta} + 2 \lambda^\alpha r_\alpha^\ell R_{\ell_S} (\lambda^\beta r_\beta^S - r_X^S) \\ &+ K_{\ell_S} (\lambda^\alpha r_\alpha^\ell - r_X^\ell) (\lambda^\beta r_\beta^S - r_X^S) \end{aligned} \right.$$

On voit apparaître le sens des conditions d'universalité. Nous ne pouvons pas espérer minimiser la forme quadratique (4-4-5) pour des λ^α quelconques, puisque nous ne connaissons pas les K_{ℓ_S} et les R_{ℓ_S} , et nous sommes ainsi conduits à imposer aux λ^α des conditions supplémentaires moyennant lesquelles (4-4-5) ne dépendra plus des R_{ℓ_S} et K_{ℓ_S} . La matrice K_{ℓ_S} étant toujours de type positif, ces conditions sont :

$$(4-4-6) \quad \lambda^\alpha r_\alpha^\ell = r_X^\ell$$

c'est-à-dire les conditions d'universalité habituelles. Moyennant ces conditions, il reste :

$$(4-4-5') \quad E(m_X^* - m_X)^2 = \lambda^\alpha \lambda^\beta \sigma_{\alpha\beta}$$

Il suffit alors de minimiser (4-4-5'), compte tenu de (4-4-6) pour retrouver le même système que dans la théorie habituelle - et la même variance d'estimation (on vérifie sans peine $E(m_X^* - m_X) = E(\lambda^\alpha m_\alpha - m_X) = 0$ dès que $m(x, \omega)$ - sans être obligatoirement stationnaire - admet elle-même une dérive $E m(x)$ représentable sur V par une expression de la forme $a_\ell f^\ell(x)$). Moyennant les hypothèses (4-4-4) et (4-4-4'), la dérive aléatoire relève du même traitement que la dérive fonctionnelle.

En délinéarisant le système relativement aux seconds membres de (4-4-6), on obtient, comme dans le cas classique, les estimateurs

$$A_\ell = \lambda_\ell^\alpha Z_\alpha$$

des coefficients $a_\rho = a_\rho(\omega)$ de la dérive $m(x) = a_\rho f_\rho(x)$ ($x \in V$), et les paramètres de Lagrange $\mu_{\rho s}$ conservent leur signification. Plus précisément - les a_ρ étant des V.A., on a :

$$\begin{cases} E(A_\rho - a_\rho) = 0 \\ E(A_\rho - a_\rho)(A_s - a_s) = \mu_{\rho s} \end{cases}$$

4-4-3 Un exemple ambigu.

Soit $Z(x, \omega)$ une F.A.I. sans dérive, c'est-à-dire :

$$\begin{cases} E(Z(x) - Z(y)) = 0 \\ \frac{1}{2} E(Z(x+h) - Z(x))^2 = \gamma(h) \end{cases}$$

et soit p une fonction de pondération très régulière et de somme égale à 1 (par exemple une exponentielle de Gauss). Posons :

$$\begin{cases} m(x) = Z * \check{p} = \int Z(x+y) p(y) dy \\ Y(x) = Z - Z * \check{p} \end{cases}$$

Dans la décomposition

$$Z = Y + m$$

m est une moyenne mobile, Y le "résidu" de Z autour de cette moyenne mobile. Le variogramme de m est $\gamma * p * \check{p} - \langle \gamma, p * \check{p} \rangle$; en vertu de la régularité de p , il admet un développement du type (4-4-4) sur un domaine V de dimension petite vis-à-vis de la portée de p . On est donc dans les conditions requises pour estimer la moyenne mobile $m(x)$, considérée comme une dérive aléatoire, par les méthodes du Krigeage Universel.

Y est une F.A. stationnaire d'ordre 2. Sa covariance σ se déduit du variogramme γ par la relation :

$$\sigma = -\gamma * (\delta - p - \check{p} + p * \check{p})$$

qui s'explique ainsi (avec $P = p * \check{p}$) :

$$(4-4-7) \quad \sigma(h) = -\gamma(h) + \int \gamma(y)[p(y+h) + p(y-h)] dy - \int \gamma(y) P(y+h) dy$$

En particulier :

$$(4-4-7') \quad \sigma^2 = \sigma(0) = 2 \int \gamma(x) p(x) dx - \int \gamma(x) P(x) dx$$

C'est cette covariance (4-4-7) qui doit figurer dans les premiers membres des conditions d'optimalité du système K.U. Mais les conditions imposées à p (régularité, et portée grande vis-à-vis du voisinage V) entraînent que $\sigma(x-y)$ admet elle aussi un développement de type (4-4-4), au terme $\gamma(h)$ près, soit :

$$(4-4-8) \quad \sigma(x-y) = -\gamma(x-y) + P_{\ell_S} r^{\ell}(x) r^S(y) \quad (x, y \in V)$$

La quantité à minimiser pour estimer $m(x)$ se met donc sous la forme :

$$\lambda^{\alpha} \lambda^{\beta} \sigma_{\alpha\beta} = -\lambda^{\alpha} \lambda^{\beta} \gamma_{\alpha\beta} + P_{\ell_S} \lambda^{\alpha} r_{\alpha}^{\ell} \lambda^{\beta} r_{\beta}^S$$

Mais les conditions d'universalité (4-4-6) entraînent alors :

$$\lambda^{\beta} \lambda^{\alpha} \sigma_{\alpha\beta} = -\lambda^{\alpha} \lambda^{\beta} \gamma_{\alpha\beta} + P_{\ell_S} r_x^{\ell} r_x^S$$

Or, d'après (4-4-8), on a aussi :

$$\sigma(0) = P_{\ell_S} r_x^{\ell} r_x^S$$

Donc, finalement - les conditions (4-4-6) étant supposées vérifiées - tout se ramène à minimiser l'expression :

$$\lambda^{\alpha} \lambda^{\beta} \sigma_{\alpha\beta} = \sigma(0) - \lambda^{\alpha} \lambda^{\beta} \gamma_{\alpha\beta}$$

D'où le système :

$$(4-4-9) \quad \begin{cases} -\lambda^\beta \gamma_{\alpha\beta} = \mu_\ell r_\alpha^\ell \\ \lambda^\alpha r_\alpha^\ell = r_x^\ell \end{cases}$$

et la variance correspondante est :

$$(4-4-9') \quad D^2(M_x^* - m_x) = \sigma(0) + \mu_\ell r_x^\ell$$

on note que le système (4-4-9) ne dépend pas du choix de la fonction de pondération p (pourvu qu'elle soit régulière et de grande portée) : le même estimateur $\lambda^\alpha Z_\alpha$ s'applique à toutes les moyennes mobiles, et c'est seulement sur la variance (4-4-9') que le choix de la fonction p se répercute, par l'intermédiaire de la variance $\sigma(0)$ définie en (4-4-7') - qui n'est autre que la variance a priori du résidu $Y(x) - m(x)$.

Ce résultat ambigu mérite quelques commentaires.

Remarque 1 - Le variogramme γ étant connu, l'estimation de la moyenne mobile $m(x)$ relève en réalité du krigeage ordinaire. Au lieu du système (4-4-9), on devrait donc utiliser le système (4-4-10) ci-après :

$$(4-4-10) \quad \begin{cases} \lambda^\beta \gamma_{\alpha\beta} = \int \gamma(y) p(y-x+x_\alpha) dy - \mu_0 \\ \sum_\alpha \lambda^\alpha = 1 \end{cases}$$

qui ne comporte qu'une seule condition d'universalité ($\sum \lambda^\alpha = 1$), mais où, en contrepartie, le second membre de la relation d'optimalité contient le variogramme γ , et fait intervenir explicitement les valeurs prises par $\gamma(h)$ pour des distances $|h|$ notablement supérieures aux dimensions de l'ensemble S des points expérimentaux x_α . Ce système (4-4-10) se rapporte donc à une situation dans laquelle nous connaissons bien le comportement réel de $\gamma(h)$ jusqu'à des distances assez grandes. Au contraire, dans le système (4-4-9), le variogramme n'intervient que par les termes $\gamma_{\alpha\beta}$, et ne suppose donc la connaissance du $\gamma(h)$ que pour les courtes distances.

On peut alors envisager la question d'une manière différente. Supposons (et la théorie de l'inférence statistique montre que c'est là une circonstance assez générale) que nous ne soyons

sûrs de notre $\gamma(h)$ qu'aux courtes distances (de l'ordre des dimensions de S), de sorte que nous connaissons bien les $\gamma_{\alpha\beta}$, mais non le second membre de (4-4-10) : dans ce cas, le système (4-4-9) nous donne la meilleure estimation possible d'une grande moyenne mobile $m(x)$, compte tenu de l'information dont nous disposons réellement.

Remarque 2 - En l'absence de dérive réelle, l'estimateur optimal de la dérive n'est donc pas dépourvu de sens : il estime une grande moyenne mobile, sans qu'il soit d'ailleurs possible de préciser laquelle - car il s'applique à toutes. La variance (4-4-9') de cet estimateur est d'autant plus élevée qu'il est censé représenter une moyenne mobile plus étendue.

Si nous posons :

$$m(x) = a_\ell f^\ell(x) \quad , \quad M^*(x) = A_\ell f^\ell(x)$$

nous obtenons les $A_\ell = \lambda_\ell^\alpha Z_\alpha$ en délinéarisant le système (4-4-9) :

$$(4-4-11) \quad \left\{ \begin{array}{l} -\lambda_\ell^\beta \gamma_{\alpha\beta} = \mu_{\ell s} f_\alpha^\ell \\ \lambda_\ell^\alpha f_\alpha^s = \delta_\ell^s \end{array} \right. \quad (\ell, s = 0, 1, \dots, n)$$

D'après (4-4-9'), la variance est alors :

$$D^2(M_x^* - m_x) = \mu_{\ell s} f_x^\ell f_x^s + \sigma(0)$$

d'où

$$(4-4-11') \quad \left\{ \begin{array}{l} E(A_0 - a_0)^2 = \sigma(0) + \mu_{00} \\ E(A_\ell - a_\ell)(A_s - a_s) = \mu_{\ell s} \end{array} \right. \quad (\ell, s \neq 0)$$

Ainsi les paramètres de Lagrange - sauf μ_{00} - représentent encore les covariances des $(A_\ell - a_\ell)$. Le terme μ_{00} - relatif aux coefficients λ_0^α de somme égale à 1 - par contre, ne donne pas la variance de $A_0 - a_0$: il doit être majoré du terme $\sigma(0)$, qui est le seul à dépendre de la fonction de pondération p . Cette remarque suggère donc la solution du problème apparemment insoluble que pose le terme constant A_0 d'une dérive lorsqu'il n'y a pas de covariance.

4-4-4 Le problème du terme constant a_0 .

Soit $Z(x) = Y_1(x) + m_1(x)$ une F.A. comportant une dérive $m_1(x)$ (fonctionnelle ou aléatoire) et une composante $Y_1(x)$ sans covariance, dont le variogramme est $\gamma(h)$. Nous pouvons poser :

$$\begin{cases} Y(x) = Y_1(x) - \int Y_1(x+y) p(y) dy \\ m(x) = m_1(x) + \int Y_1(x+y) p(y) dy \end{cases} \quad \begin{cases} Y = Y_1 - Y_1 * \check{p} \\ m = m_1 + Y_1 * \check{p} \end{cases}$$

Dans cette nouvelle décomposition, $Z = Y + m$, la composante Y admet une covariance $\sigma(x,y)$, et la dérive (aléatoire) $m(x)$ vérifie les hypothèses (4-4-4) et (4-4-4'), pourvu que la fonction de pondération $p(x)$ soit suffisamment régulière et de portée assez grande. L'estimation optimale $M^*(x) = A_\ell f^\ell(x)$, $A_\ell = \lambda_\ell^\alpha Z_\alpha$ de la dérive $m(x)$ s'obtient en résolvant le même système (4-4-11) que ci-dessus, et conduit en particulier à un estimateur A_0 : la variance $E(A_0 - a_0)^2 = \sigma(0) + \mu_{00}$ correspondante dépend explicitement du $\sigma(0)$ défini en (4-4-7'), c'est-à-dire du choix de la fonction de pondération p qui nous a servi à donner un sens à cette expression (soustraction d'une moyenne mobile). La valeur numérique de A_0 n'en dépend pas.

Or le système (4-4-11) n'est pas autre chose que le système (4-2-5) complété par les équations relatives à $\ell = 0$. Il n'est donc pas absurde d'utiliser le système ainsi complété, et d'en déduire une estimation A_0 du terme constant a_0 (qui, à strictement parler, n'existait pas dans le cadre des hypothèses du fascicule). L'expression $A_0 + \sum_{\ell=1}^n A_\ell f^\ell(x)$ donne une estimation de la vraie dérive corrigée d'une moyenne mobile, et sa variance est la variance de l'estimation de la dérive ainsi corrigée. Cette variance n'a qu'une signification relative (puisqu'elle dépend du choix de la moyenne mobile). Mais, du fait que l'estimateur lui-même

$$M^*(x) = A_0 + \sum_{\ell=1}^n A_\ell f^\ell(x)$$

ne dépend pas de ce choix, il peut être utilisé pour cartographier la dérive, sans que se posent de problèmes de raccordement entre les différents domaines $V_1, V_2 \dots$ de travail. Plus précisément, à chaque point x on peut affecter un voisinage V_x (translaté de V centré en x), puis l'ensemble $S_x \subset V_x$ des points expérimentaux contenus dans V_x . Il suffit alors de résoudre le système (4-4-9) à l'aide des $x_\alpha \in S_x$ pour former l'estimateur $M^*(x)$ de la dérive en x . En choisissant l'origine des coordonnées en ce point x , et avec $f^\ell(0) = 0$ ($\ell = 1, 2, \dots, n$) (cas des polynômes), c'est en fait $M^*(0) = A_0$ que l'on estime pour chaque V_x , et cette méthode du voisinage glissant permet de

cartographier la dérive.

4-4-5 Le Krigeage.

Le cas du krigeage va se présenter d'une manière très semblable. Pour abrégier les écritures, nous nous limitons au cas du krigeage d'un point $x \in V$ (bien que le cas général puisse se traiter de la même manière). Nous supposons toujours vérifiées les hypothèses du paragraphe 1. Il s'agit de former un estimateur $Z_x^* = \lambda^\alpha Z_\alpha$ de $Z_x = Z(x)$. Considérons l'expression

$$Z_x^* - Z_x = \lambda^\alpha Z_\alpha - Z_x = \lambda^\alpha Y_\alpha - Y_x + \lambda^\alpha m_\alpha - m_x$$

et prenons l'espérance de son carré :

$$E(Z_x^* - Z_x)^2 = \lambda^\alpha \lambda^\beta \sigma_{\alpha\beta} - 2 \lambda^\alpha \sigma_{\alpha x} + \sigma_{xx} + 2 E(\lambda^\alpha Y_\alpha - Y_x)(\lambda^\beta m_\beta - m_x) + E(\lambda^\alpha m_\alpha - m_x)^2$$

Compte tenu des hypothèses (4-4-4) et (4-4-4'), nous trouvons

$$\begin{cases} E(\lambda^\alpha Y_\alpha - Y_x)(\lambda^\beta m_\beta - m_x) = R_{\rho_s} (\lambda^\alpha f_\alpha^\ell - f_x^\ell)(\lambda^\beta f_\beta^s - f_x^s) \\ E(\lambda^\alpha m_\alpha - m_x)^2 = K_{\rho_s} (\lambda^\alpha f_\alpha^\ell - f_x^\ell)(\lambda^\beta f_\beta^s - f_x^s) \end{cases}$$

Donc, ici encore, $E(Z_x^* - Z_x)^2$ est indépendant de K_{ρ_s} (inconnus) si et seulement si les conditions d'universalité (4-4-6) sont vérifiées. Moyennant ces conditions, on a à nouveau $E(Z_x^* - Z_x) = 0$ et :

$$D^2(Z_x^* - Z_x) = \lambda^\alpha \lambda^\beta \sigma_{\alpha\beta} - 2 \lambda^\alpha \sigma_{\alpha x} + \sigma_{xx}$$

en minimisant cette forme quadratique, compte tenu des conditions d'universalité (4-4-6), on retrouve le système habituel du K.U. pour les dérivées fonctionnelles.

4-5 LE COKRIGEAGE.

Dans certains problèmes, on doit étudier et estimer simultanément plusieurs variables régionalisées. On trouvera dans [4] ou [5] une étude systématique des corégionalisations, avec les deux aspects habituels (méthodes transitives, c'est-à-dire non probabilistes, et théorie intrinsèque). Nous nous contentons ici de voir comment se pose le problème du krigeage universel lorsque l'on a simultanément affaire à plusieurs F.A. non stationnaires. Du point de vue conceptuel, il n'apparaît aucun élément nouveau, comme on le verra. Toutes les difficultés (qui sont réelles) proviennent des notations, nécessairement plus complexes. Pour éviter d'avoir à manipuler trois systèmes d'indices, nous utiliserons les notations du cas continu, où interviennent des mesures plutôt que des coefficients. Ce sont d'ailleurs les notations les plus générales, valables aussi bien pour le cas infini que pour le cas fini.

4-5-1 Les Notations.

Soient $Z_i(x)$, $i = 1, 2, \dots, d$, d fonctions aléatoires (non stationnaires en général) définies sur l'espace \mathbb{R}^n . Chacune d'elles possède une dérive $m_i(x)$, représentable sur un certain domaine V_i , par une combinaison linéaire (à coefficients inconnus) de fonctions données $f_i^\ell(x)$, $\ell = 0, 1, \dots, k_i$. Nous nous limiterons au cas où ces F.A. admettent des fonctions de covariance, et nous laisserons au lecteur le soin de procéder à l'extension au cas où certaines d'entre elles n'admettraient que des covariogrammes. Outre les covariances (centrées) des d F.A. Z_i , il convient d'introduire les covariances rectangles. Nous poserons

$$(4-5-1) \quad \sigma_{ij}(x, y) = E(Z_i(x) Z_j(y)) - m_i(x) m_j(y)$$

Du point de vue théorique, le problème de l'estimation de ces dérivées et du krigeage de ces F.A. lorsque l'on connaît leurs réalisations sur des ensembles S_1, S_2, \dots, S_d (en général distincts) ne diffère pas des problèmes analogues que nous avons déjà résolus dans le cas d'une seule F.A. En effet, à aucun moment nous n'avons utilisé le fait que l'espace sur lequel était définie cette F.A. unique était l'espace euclidien \mathbb{R}^n , et tous les résultats que nous avons obtenus restent donc valables pour une F.A. définie sur un espace abstrait E quelconque. En particulier, nous pouvons prendre comme espace de définition l'espace produit

$$E = \mathbb{R}^n \times D$$

de \mathbb{R}^n par l'ensemble $D = \{1, 2, \dots, d\}$ des indices. Un élément de E sera un couple (x, i) constitué d'un point $x \in \mathbb{R}^n$ et d'un indice $i \in D$. Nos d fonctions aléatoires $Z_i(x)$ constituent alors une F.A. unique sur cet espace produit, et nous la noterons $Z(x, i)$. De même, l'ensemble des d dérivées $m_i(x)$ constitue une dérivée unique $m(x, i)$ sur l'espace produit, et cette dérivée admet - sur un certain domaine V de $\mathbb{R}^n \times D$, défini par la donnée de D domaines V_i - une représentation de la forme :

$$(4-5-2) \quad m(x, i) = a_{\ell} f^{\ell}(x, i) \quad (x, i) \in V$$

La fonction $f^{\ell}(x, i)$ sur $\mathbb{R}^n \times D$ s'explicité en d fonctions $f_i^{\ell}(x)$ sur \mathbb{R}^n et $(x, i) \in V$ désigne l'ensemble des couples (x, i) tels que $x \in V_i$. De même, au lieu de $\sigma_{ij}(x, y)$ nous devrions écrire :

$$\sigma((x, i)(y, j)) = E[Z(x, i) Z(y, j)] - m(x, i) m(y, j)$$

Pour des raisons de commodité, nous conserverons cependant la notation $\sigma_{ij}(x, y)$.

Comme tous les résultats antérieurs restent valables pour la F.A. $Z(x, i)$, il n'y a pas lieu de donner de nouvelles démonstrations, mais seulement de retranscrire ces résultats dans le nouveau système de notations. Il y a lieu, cependant, de prendre garde à la condition d'indépendance linéaire des f^{ℓ} sur S , condition moyennant laquelle seulement sont garanties l'existence et l'unicité des solutions de nos systèmes d'équations.

Désignons par S_i , $i \in D$ l'ensemble des points expérimentaux sur lequel est connu la réalisation de la F.A. $Z_i(x)$, avec en général $S_i \neq S_j$ pour $i \neq j$. Certains de ces ensembles peuvent être vides si l'on n'a aucune information expérimentale sur les F.A. correspondantes. En notation d'espace produit, l'ensemble S des points expérimentaux (x, i) où l'on connaît la réalisation de $Z(x, i)$ est l'ensemble :

$$S = \{(x, i) : x \in \mathbb{R}^n, i \in D, x \in S_i\}$$

des couples (x, i) tels que le point x appartienne à l'ensemble S_i .

La condition d'indépendance linéaire sur S des fonctions $f^{\ell}(x, i)$ se formule donc ainsi :

$$c_{\ell} f^{\ell}(x, i) = 0 \text{ pour tout } i \in D \text{ et tout } x \in \mathbb{R}^n \text{ tel que } x \in S_i$$

entraîne $c_{\ell} = 0$.

L'exemple simple suivant montrera combien cette condition est naturelle.

Exemple : Supposons $D = \{1, 2\}$, autrement dit nous avons affaire à 2 fonctions aléatoires seulement. Prenons comme dérivées $f^{\ell(x,i)}$ $k_1 + k_2$ fonctions vérifiant

$$(4-5-3) \quad \begin{cases} f^1(x,2) = \dots = f^{k_1}(x,2) = 0 \\ f^{k_1+1}(x,1) = \dots = f^{k_1+k_2}(x,1) = 0 \end{cases}$$

Ces relations décrivent une situation où les deux dérivées $m_1(x)$ et $m_2(x)$ sont algébriquement indépendantes. Les k_1 premières fonctions, nulles sur V_2 , servent uniquement à représenter sur V_1 la dérivée $m_1(x)$ de la première F.A., et les k_2 fonctions suivantes, nulles sur V_1 , représentent sur V_2 la dérivée $m_2(x)$ de la seconde F.A.

Supposons que S_2 soit vide, donc que nous ne disposions d'aucune information expérimentale sur $Z_2(x)$. Comme les dérivées $m_1(x)$ et $m_2(x)$ sont, par hypothèses, algébriquement indépendantes, il est clair que la connaissance de $Z_1(x)$ sur S_1 ne nous apportera aucune espèce de renseignement sur la dérivée $m_2(x)$ de la seconde F.A. De fait, les fonctions $f^{\ell(x,i)}$ ne sont pas linéairement indépendantes sur S . En effet, comme S_2 est vide, S se réduit à $\{S_1\} \times \{1\}$, c'est-à-dire à l'ensemble (x,i) des couples tels que $i = 1$ et $x \in S_1$. Or, précisément, on a $f^{\ell(x,1)}$ identiquement nulle pour $\ell > k_1$, et par suite ces fonctions ne sont pas linéairement indépendantes sur S_1 .

Passons maintenant à l'écriture des systèmes eux-mêmes. On prendra bien garde que la notation :

$$\forall (y,j) \in S$$

signifie précisément : quel que soit l'indice $j \in D$ et quel que soit le point $y \in \mathbb{R}^n$ tel que $y \in S_j$.

4-5-2 Estimation optimale de la dérive.

Proposons-nous de former l'estimateur optimal $\hat{m}^*(x_0, i_0)$ de la dérive de la F.A. $Z_{i_0}(x)$ au point x_0 , c'est-à-dire, en notation d'espace produit, l'estimateur optimal de la dérive $m(x_0, i_0)$ au point $(x_0, i_0) \in \mathbb{R}^n \times D$. Cet estimateur doit être linéaire relativement aux $Z(x,i)$,

$(x, i) \in S$. Il sera donc de la forme :

$$m^*(x_0, i_0) = \sum_1 \int_{S_1} \lambda^1(dx) z_1(x)$$

avec d mesures λ^1 à supports respectivement contenus dans les S_1 . La transposition du système (4-2-3) donne alors :

$$(4-5-4) \quad \begin{cases} \sum_1 \int_{S_1} \lambda^1(dx) \sigma_{1j}(x, y) = \mu_\ell f^\ell(y, j) & \forall (y, j) \in S \\ \sum_1 \int_{S_1} \lambda^1(dx) f^\ell(x, i) = f^\ell(x_0, i_0) \\ D^2(m^*(x_0, i_0)) = \mu_\ell f^\ell(x_0, i_0) \end{cases}$$

De même, en délinéarisant le système (4-5-4), on obtient les estimateurs A_ℓ optimaux des coefficients a_ℓ de la dérive à l'aide d'une famille de mesures $\lambda_\ell^1(dx)$ à supports, respectivement, dans S_1 , soit

$$A_\ell = \sum_1 \int_{S_1} \lambda_\ell^1(dx) z_1(x)$$

en résolvant le système que l'on obtient en transposant (4-2-6), soit :

$$(4-5-5) \quad \begin{cases} \sum_1 \int_{S_1} \lambda_\ell^1(dx) \sigma_{1j}(x, y) = \mu_{\ell S} f^S(y, j) & \forall (y, j) \in S \\ \sum_1 \int_{S_1} \lambda_\ell^1(dx) f^S(x, i) = \delta_\ell^S \\ \mu_{\ell S} = \text{Cov}(A_\ell, A_S) \end{cases}$$

4-5-3 Le Cokrigeage ponctuel.

En toute généralité, on pourrait former les équations relatives à l'estimation d'une grandeur de la forme :

$$z_0 = \sum_1 \int p^1(dx) z_1(x)$$

avec des mesures p^i quelconques (dont les supports seraient contenus respectivement dans les V_i). Mais nous avons déjà observé que les seconds membres de nos systèmes d'équations dépendent linéairement de ces mesures $p^i(dx)$. Il suffit donc de savoir former l'estimateur $Z^*(x,i)$ du krigeage relatif au point $(x,i) \in \mathbb{R}^n \times D$ pour en déduire l'estimateur Z_0^* du krigeage de Z_0 selon la relation :

$$Z_0^* = \sum_i \int p^i(dx) Z^*(x,i)$$

Nous nous limiterons donc à écrire les systèmes d'équation pour le krigeage ponctuel.

Soit donc à estimer $Z_{i_0}(x_0)$, c'est-à-dire la valeur de la F.A. au point (x_0, i_0) de l'espace produit. La transposition du système (4-3-1) montre alors que l'estimateur optimal universel est de la forme :

$$Z_{i_0}^*(x_0) = \sum_i \int_{S_i} \lambda^i(dx) Z_i(x)$$

avec des mesures λ^i à support dans S_i vérifiant :

$$(4-5-6) \quad \left\{ \begin{array}{l} \sum_i \int_{S_i} \lambda^i(dx) \sigma_{ij}(x,y) = c_{ji_0}(y, i_0) + \mu_\rho f^\rho(y,j) \quad \forall (y,j) \in S \\ \sum_i \int_{S_i} \lambda^i(dx) f^\rho(x,i) = f^\rho(x_0, i_0) \\ \sigma_v^2 = \sigma_{i_0 i_0}(x_0, x_0) + \mu_\rho f^\rho(x_0, i_0) - \sum_i \int_{S_i} \lambda^i(dx) \sigma_{ii_0}(x, x_0) \end{array} \right.$$

On notera que le théorème d'additivité s'applique sans changement au cas du cokrigeage, mais il n'est pas utile ici d'explicitier les relations correspondantes.

4-6 LES INDETERMINATIONS DU VARIOGRAMME SOUS-JACENT.

4-6-1 Position du problème.

Nous allons maintenant aborder le problème difficile (mais évidemment capital

pour les applications) que pose l'estimation du variogramme $\gamma(x,y)$ d'une F.A. non stationnaire $Z(x)$ admettant une dérive $m(x)$. Ce variogramme est le variogramme des résidus vrais $Z(x)-m(x)$, et la présence de la dérive $m(x)$ ne permet pas de l'estimer directement à partir du variogramme expérimental brut ou du variogramme des résidus estimés $Z(x)-m^*(x)$ (voir paragraphe 4-2-4). Pour exprimer le fait que ce "vrai" variogramme est masqué (de manière partiellement irrémédiable, comme nous le verrons) nous dirons qu'il s'agit du variogramme sous-jacent, et le problème crucial qui se pose ici est celui de l'identification du variogramme sous-jacent à partir d'une réalisation unique. Nous ne pouvons pas aborder ici dans toute leur ampleur les problèmes très délicats de statistique mathématique concernant la possibilité d'une telle opération. Il conviendrait, en particulier, d'examiner avec précision le type d'hypothèse de quasi-stationnarité qu'il est nécessaire de faire au sujet des résidus vrais $Z(x)-m(x)$ pour qu'il soit effectivement possible de reconstituer ce qu'il est théoriquement possible d'estimer concernant ce variogramme sous-jacent.

C'est qu'en effet, nous allons voir que le problème est théoriquement indéterminé : en mettant les choses au mieux, c'est-à-dire en supposant que l'on connaisse les espérances des estimateurs qu'on a le droit d'utiliser (les estimateurs universels) et non pas seulement les valeurs expérimentales correspondantes calculées sur plusieurs morceaux d'une réalisation unique, il apparaît en effet que la reconstitution du variogramme sous-jacent reste entachée d'une indétermination fondamentale, de nature algébrique, et c'est essentiellement à l'étude algébrique précise de cette indétermination que nous allons consacrer ce dernier paragraphe.

Notons d'ailleurs qu'il s'agit là d'un problème absolument général, que l'on rencontre dès les applications les plus banales de la statistique classique (voir Ex. 16). Les statisticiens éludent le problème de cette indétermination fondamentale en introduisant des hypothèses supplémentaires en général extrêmement fortes (par exemple, l'indépendance des différentes V.A. dont on connaît une réalisation). Ici encore, il serait utile d'examiner jusqu'à quel point il est possible d'affaiblir ces conditions trop fortes sans que le problème cesse d'être soluble. Pour des raisons d'ordre physique, on peut penser qu'une hypothèse de quasi stationnarité devrait convenir. Nous nous contenterons de vérifier qu'il est, en effet, possible de réduire considérablement cette indétermination en choisissant parmi tous les variogrammes sous-jacents compatibles avec les données, celui qui est "le plus stationnaire" possible (au sens, par exemple, des moindres carrés). (cf. Ex. 17).

Il apparaîtra, d'autre part, une circonstance remarquable, et très rassurante sur le plan méthodologique : cette indétermination fondamentale affectera uniquement les variances de nos es-

estimateurs optimaux, mais non ces estimateurs eux-mêmes. Autrement dit, tous les variogrammes sous-jacents théoriquement possibles conduisent aux mêmes estimateurs optimaux pour la dérive comme pour le krigeage, et ces estimateurs ne dépendent en aucune façon des hypothèses supplémentaires de nature physique que l'on peut introduire pour lever ces indéterminations (par contre, la variance que l'on attribuera à ces estimateurs en dépendra).

On rapprochera ce résultat de la conclusion analogue concernant le terme constant a_0 de la dérive, conclusion à laquelle l'interprétation de la dérive par une F.A. très régulière nous a permis d'aboutir dans le paragraphe 4-4-4. Ici encore, la théorie des dérivées aléatoires nous permet de prévoir à l'avance le type d'indétermination que nous allons rencontrer. C'est ce que nous allons voir maintenant, avant de passer dans les paragraphes suivants, à l'étude proprement algébrique du problème.

Soit, en effet, $Z(x)$ une F.A. admettant une dérive fonctionnelle du type $a_\ell f^\ell(x)$ et une covariance $\sigma(x,y)$, et soit aussi $m(x)$ une fonction aléatoire très régulière dont la covariance $K(x,y)$ admet, sur le domaine V de travail, une bonne représentation approchée de la forme

$$K(x,y) = K_{\ell_S} f^\ell(x) f^\ell(y)$$

La fonction aléatoire

$$Z'(x) = Z(x) + m(x)$$

admet la covariance sous-jacente

$$(4-6-1) \quad \sigma'(x,y) = \sigma(x,y) + K_{\ell_S} f^\ell(x) f^\ell(y)$$

(en supposant $m(x)$ et $Y(x)$ indépendantes) et la dérive :

$$a_\ell f^\ell(x) + m(x) = (a_\ell + b_\ell) f^\ell(x)$$

avec a_ℓ déterminé mais inconnu, tandis que les b_ℓ sont des V.A. (dont la matrice de covariance est, justement, K_{ℓ_S}).

Or, lorsque l'on dispose d'une réalisation unique de la fonction aléatoire, seules ont un sens expérimental les réalisations correspondantes des V.A. b_ℓ , qui sont de simples nombres.

mais, du point de vue épistémologique, il serait illusoire de croire que l'on gagne réellement de l'information en décidant d'interpréter une valeur numérique donnée comme une réalisation d'une V.A. dont on ignore la loi. Dès lors que les K_{ρ_s} sont inaccessibles expérimentalement (et nous avons vu que c'est le cas général) les deux paramètres a_{ρ} et b_{ρ} ne sont pas réellement séparables. Nous pouvons aussi bien décider d'interpréter les b_{ρ} comme des nombres : au lieu de $\sigma'(x,y)$, notre F.A. $Z'(x)$ admet alors la même covariance $\sigma(x,y)$ que $Z(x)$. Rien donc, à partir d'une réalisation unique de $Z'(x)$, ne doit permettre de choisir entre ces deux covariances. L'indétermination algébrique, en ce qui concerne la reconstitution de la covariance sous-jacente, sera donc au moins (et nous verrons qu'elle est en fait exactement) celle qu'exprime la relation (4-6-1).

Mais en contre-partie - circonstance rassurante - la théorie des dérivées aléatoires exposées au paragraphe 4-4 montre que les équations des estimateurs optimaux sont les mêmes pour les dérivées aléatoires et les dérivées fonctionnelles. Par conséquent, nous devons nous attendre, et nous allons le vérifier, à ce que l'indétermination fondamentale (4-6-1) ne se répercute pas sur les estimateurs optimaux eux-mêmes (mais seulement sur les variances que nous leur attribuerons).

Passons maintenant à l'étude algébrique du problème, en nous limitant pour simplifier, au cas où il existe une covariance $\sigma(x,y)$ et où l'ensemble S des données expérimentales est fini.

4-6-2 Les estimateurs quadratiques universels.

Dans le cadre d'une théorie d'ordre 2 (c'est-à-dire où l'on cherche à estimer la covariance, sans rien connaître a priori sur la loi spatiale de la F.A.) les seuls estimateurs que l'on puisse utiliser pour estimer cette covariance sont des estimateurs quadratiques, du type

$$Q^* = Q^{\alpha\beta} Z_{\alpha} Z_{\beta}$$

Nous dirons qu'un estimateur quadratique Q^* est universel si son espérance mathématique ne dépend pas des coefficients a_{ρ} de la dérive $m(x) = a_{\rho} f^{\rho}(x)$. De :

$$E(Q^*) = Q^{\alpha\beta} \sigma_{\alpha\beta} + a_{\rho} a_{\sigma} Q^{\alpha\beta} f_{\alpha}^{\rho} f_{\beta}^{\sigma}$$

résulte alors qu'un estimateur quadratique Q^* est universel si et seulement si il vérifie la con-

dition :

$$(4-6-2) \quad Q^{\alpha\beta} f_{\alpha}^{\ell} f_{\beta}^s = 0$$

Lorsque les coefficients a_{ℓ} de la dérive sont inconnus, ces estimateurs universels sont les seuls utilisables pour l'inférence statistique de la fonction de covariance. Il résulte de cette remarque que le maximum d'information qu'il soit théoriquement possible d'obtenir concernant la matrice $\sigma_{\alpha\beta}$ des covariances sous-jacentes est contenu dans l'ensemble des valeurs numériques $E(Q^*)$, Q^* décrivant la classe des estimateurs quadratiques universels, donc dans l'ensemble des $Q^{\alpha\beta}$ $\sigma_{\alpha\beta}$ pour les matrices $Q^{\alpha\beta}$ vérifiant (4-6-2).

Pour caractériser la classe des estimateurs quadratiques universels, quelques considérations algébriques préliminaires sont nécessaires. Soit N le nombre des points expérimentaux $x_{\alpha} \in S$, k celui des fonctions f^{ℓ} ($\ell = 1, 2, \dots, k$) supposées linéairement indépendantes sur S . Nous pouvons trouver $N-k$ autres fonctions f^u ($u = k+1, \dots, N$) telles que les N vecteurs $f^i = (f_{\alpha}^i)$, $i = 1, 2, \dots, N$ constituent une base de l'espace \mathbb{R}^N , et nous pouvons associer à cette base la base duale constituée de N co-vecteurs $\lambda_i = (\lambda_i^{\alpha})$ vérifiant les conditions :

$$(4-6-3) \quad \lambda_i^{\alpha} f_{\alpha}^j = \delta_i^j, \quad \lambda_i^{\alpha} f_{\beta}^i = \delta_{\beta}^{\alpha}$$

(autrement dit, les matrices λ_i^{α} et f_{α}^i sont inverses l'une de l'autre). Parmi ces co-vecteurs λ_i , $i = 1, 2, N$, les k premiers, soit λ_{ℓ} , $\ell = 1, 2, \dots, k$ vérifient la condition d'universalité :

$$(4-6-4) \quad \lambda_{\ell}^{\alpha} f_{\alpha}^s = \delta_{\ell}^s$$

Autrement dit, quelle que soit la matrice $\sigma_{\alpha\beta}$, l'estimateur

$$(4-6-4') \quad A_{\ell} = \lambda_{\ell}^{\alpha} z_{\alpha}$$

vérifie $E(A_{\ell}) = a_{\ell}$ quels que soient les a_{ℓ} inconnus : c'est un estimateur universel (mais non optimal en général).

On posera :

$$m_{\alpha}^* = A_{\ell} f_{\alpha}^{\ell}$$

et m_α^* sera un estimateur universel (non optimal) de la valeur de la dérive en x_α . Les $N-k$ autres vecteurs λ_u , soient λ_u , $u = k+1, \dots, N$ constituent une base de l'orthogonal des f^ℓ . Ils vérifient :

$$\lambda_u^\alpha f_\alpha^\ell = 0$$

Le sous-espace engendré par les λ_u est donc indépendant du choix des fonctions f^u , puisqu'il est identique à l'orthogonal des f^ℓ . Posons alors :

$$A_u = \lambda_u^\alpha Z_\alpha$$

Ces variables aléatoires vérifient $E(A_u) = 0$ quels que soient les coefficients a_ℓ . Comme la matrice λ_u^α est régulière, nous pouvons effectuer un changement de variables, et remplacer les Z_α par les :

$$A_i = \lambda_i^\alpha Z_\alpha \quad (i = 1, 2, \dots, N)$$

Tout estimateur quadratique Q^* est alors de la forme :

$$Q^* = Q^{ij} A_i A_j$$

Calculons $E(Q^*)$. Comme $E(A_u) = 0$ ($u = k+1, \dots, N$) quels que soient les a_ℓ , il vient :

$$E(Q^*) = Q^{ij} \text{Cov}(A_i A_j) + Q^{\ell s} a_\ell a_s$$

Par conséquent, Q^* est un estimateur quadratique universel si et seulement si on a $Q^{\ell s} = 0$ ($\ell, s = 1, 2, \dots, k$). Ainsi, toute l'information qu'il est théoriquement possible de reconstituer en ce qui concerne la matrice des covariances se trouve contenue dans la matrice $(N-k) \times N$ suivante :

$$(4-6-10) \quad S_{ui} = \text{Cov}(A_u A_i) \quad (u = k+1, \dots, N ; i = 1, 2, \dots, N)$$

4-6-3 Forme générale des matrices de covariance admissible.

Posons-nous maintenant le problème inverse : supposons connue toute l'information

qu'il est théoriquement possible de reconstituer, c'est-à-dire les $N(N-k)$ quantités S_{ij} de la relation (4-6-10). La matrice :

$$S_{ij} = \text{Cov}(A_i, A_j) = \lambda_i^\alpha \lambda_j^\beta \sigma_{\alpha\beta}$$

est inversible, et on a donc théoriquement :

$$(4-6-11) \quad \sigma_{\alpha\beta} = S_{ij} f_\alpha^i f_\beta^j$$

Dans cette écriture, les $S_{\ell s} = \text{Cov}(A_\ell, A_s)$ ($\ell, s = 1, 2, \dots, k$) représentent la partie inaccessible de la covariance. Nous pouvons donc écrire :

$$(4-6-12) \quad \sigma_{\alpha\beta} = C_{\alpha\beta} + S_{\ell s} f_\alpha^\ell f_\beta^s$$

Le premier terme représente la partie connaissable de la matrice de covariance $\sigma_{\alpha\beta}$, tandis que la forme quadratique $S_{\ell s} f_\alpha^\ell f_\beta^s$ représente la partie indéterminée. Or, on a :

$$S_{\ell s} f_\alpha^\ell f_\beta^s = \lambda_\ell^{\alpha'} \lambda_s^{\beta'} f_\alpha^\ell f_\beta^s \text{Cov}(Z_\alpha, Z_\beta) = E(m_\alpha^* m_\beta^*) - m_\alpha m_\beta$$

et par suite aussi, d'après (4-6-12) :

$$C_{\alpha\beta} = \sigma_{\alpha\beta} + m_\alpha m_\beta - E(m_\alpha^* m_\beta^*)$$

c'est-à-dire :

$$(4-6-13) \quad C_{\alpha\beta} = E(Z_\alpha Z_\beta - m_\alpha^* m_\beta^*)$$

Telle est la partie connaissable de la covariance. Toute matrice de la forme :

$$(4-6-14) \quad \tau_{\alpha\beta} = C_{\alpha\beta} + D_{\ell s} f_\alpha^\ell f_\beta^s$$

où $D_{\ell s}$ est une matrice symétrique arbitraire est tout aussi admissible (pourvu qu'elle soit de type positif) que la "vraie" covariance $\sigma_{\alpha\beta}$. Ces covariances admissibles se déduisent les unes

des autres selon la relation suivante, où T_{ℓ_s} est une matrice symétrique arbitraire :

$$(4-6-15) \quad \tau_{\alpha\beta} = \sigma_{\alpha\beta} + T_{\ell_s} f_{\alpha}^{\ell} f_{\beta}^s$$

4-6-4 Conséquences pour les estimateurs optimaux.

Considérons, par exemple, le système d'équations de l'estimation optimale de la dérive écrit en prenant comme covariance $\tau_{\alpha\beta}$ la solution générale (4-6-15) :

$$\begin{cases} \tilde{\lambda}_{\ell}^{\beta} \tau_{\alpha\beta} = \tilde{\mu}_{\ell_s} f_{\alpha}^s \\ \tilde{\lambda}_{\ell}^{\beta} f_{\beta}^s = \delta_{\ell}^s \end{cases}$$

La condition d'universalité ne fait pas intervenir $\tau_{\alpha\beta}$ et ne dépend donc pas de la matrice indéterminée. La première relation se met sous la forme équivalente :

$$\tilde{\lambda}_{\ell}^{\beta} \sigma_{\alpha\beta} = \tilde{\mu}_{\ell_s} f_{\alpha}^s - T_{\ell_s} f_{\alpha}^{\ell} f_{\beta}^s \tilde{\lambda}_{\ell}^{\beta}$$

soit, en tenant compte de la condition d'universalité :

$$(4-6-19) \quad \begin{cases} \tilde{\lambda}_{\ell}^{\beta} \sigma_{\alpha\beta} = (\tilde{\mu}_{\ell_s} - T_{\ell_s}) f_{\alpha}^s \\ \tilde{\lambda}_{\ell}^{\beta} f_{\beta}^s = \delta_{\ell}^s \end{cases}$$

Or ce système est identique au système (4-2-6) écrit avec la "vraie" covariance sous-jacente inconnue, à condition de prendre :

$$(4-6-20) \quad \mu_{\ell_s} = \tilde{\mu}_{\ell_s} - T_{\ell_s}$$

Par conséquent les indéterminations de la covariance sous-jacente ne se répercutent pas sur l'estimation optimale de la dérive. Par contre, d'après (4-6-20) elles affectent la matrice des covariances des estimateurs optimaux, qui reste théoriquement entièrement indéterminée.

Lorsqu'il n'existe pas de covariance $\sigma_{\alpha\beta}$ mais seulement un variogramme, l'espérance $E(Z_\alpha Z_\beta - m_\alpha^* m_\beta^*)$ n'existe pas en général. On devra remplacer l'expression (4-6-17) par l'expression analogue construite à partir des accroissements de la F.A. et de sa dérivée estimée, soit

$$\gamma_{\alpha\beta} = \frac{1}{2} E[(Z_\alpha - Z_\beta)^2 - (m_\alpha^* - m_\beta^*)^2] + \frac{1}{2} D^2 (m_\alpha^* - m_\beta^*)$$

dont le premier terme représente la partie accessible du variogramme sous-jacent, et le second la partie inaccessible.

On trouvera dans l'Exercice 17 quelques indications sur la manière dont une hypothèse physiquement plausible comme la stationnarité locale permet de lever pour l'essentiel ces indéterminations (on notera que le terme $\sigma(o) = \sigma^2$ représentant la variance a priori reste, en principe, inaccessible).

4-7 EXERCICES SUR LE CHAPITRE 4.

4-7-1 Exercices sur les Dérivées.

Exercice 1 (Variogramme linéaire, cas continu) -

a/ Soit, sur la droite, $f(x)$ une fonction admettant une dérivée seconde continue. Montrer (en effectuant une intégration par partie) que l'on a pour $-R \leq y \leq R$:

$$\int_{-R}^R |x-y| f''(x) dx = 2 f(y) - f(R) - f(-R) + R(f'(R) - f'(-R)) - y[f'(R) + f'(-R)]$$

En désignant par δ_R et δ_{-R} les mesures de Dirac placées en R et en $-R$ respectivement, montrer, toujours pour $-R \leq y \leq R$

$$\int \frac{\delta_R + \delta_{-R}}{2} |x-y| = R, \quad \int \frac{\delta_R - \delta_{-R}}{2} |x-y| = -y, \quad \int_{-R}^R |x-y| dx = y^2 + R^2$$

b/ Soit, sur la droite, une F.A. admettant la variogramme $\gamma(x,y) = |x-y|$ et une dé-

rive quadratique $a_1 x + a_2 x^2$ (définie à une constante près a_0) à coefficients inconnus. On suppose la réalisation connue sur $(-R, R)$. Montrer que la solution du système (4-2-11) complété pour la valeur 0 de l'indice ℓ conduit aux estimateurs :

$$A_0 = \frac{3}{2} \bar{Z} - \frac{1}{4}(Z_R + Z_{-R}), \quad A_1 = \frac{1}{2R} (Z_R - Z_{-R}), \quad A_2 = \frac{3}{4R^2} (Z_R + Z_{-R}) - \frac{3}{2R^2} \bar{Z}$$

$$(\bar{Z} = \frac{1}{2R} \int_{-R}^R Z(x) dx) \text{ avec } D^2(A_1) = \frac{1}{R}, \quad D^2(A_2) = \frac{3}{4R^3}, \quad \text{Cov}(A_1, A_2) = 0$$

[Remarquer, d'après a/, que $A_\ell = \int_{-R}^R \lambda_\ell(dx) Z(x)$ avec une mesure λ_ℓ qui est une combinaison linéaire des mesures $\frac{\delta_R + \delta_{-R}}{2}$, $\frac{\delta_R - \delta_{-R}}{2}$ et dx , et utiliser les relations d'universalité pour déterminer les coefficients de ces combinaisons. Pour déterminer les variances, identifier les paramètres de Lagrange dans la première relation (4-2-11).

c/ (Variogramme des résidus) - Montrer que $\gamma^*(h) = \frac{1}{1-h} \int_{-R}^{R-h} \frac{1}{2} [Z(x+h) - m^*(x+h) - Z(x) + m^*(x)]^2$

a, dans les conditions ci-dessus, l'espérance

$$E(\gamma^*(h)) = h - \frac{h^2}{R} + \frac{1}{2} \frac{h^3}{R^3} - \frac{1}{8} \frac{h^4}{R^3}$$

[Utiliser la relation (4-2-15). Le graphe est une courbe assez aplatie. La pente à l'origine n'est pas altérée, mais dès que h n'est pas très petit le biais est énorme. Pour $h = 2R$, cette espérance s'annule. Remarquer qu'un expérimentateur risque de conclure à l'indépendance des résidus].

Exercice 2 - Dans les mêmes conditions que dans Ex. 1, on suppose que la F.A. admet seulement une dérivée linéaire. Montrer que l'estimation $A_0 + A_1 x$ représente la droite joignant les deux points extrêmes Z_R et Z_{-R} , et que le variogramme des résidus a pour espérance

$$E(\gamma^*(h)) = h - \frac{h^2}{2R}$$

Exercice 3 - (Influence d'un effet de pépité) - Soit, sur la droite, une F.A.I. admettant une dérivée linéaire et le variogramme $\gamma(h) = |h| - C \delta$ (effet de pépité représenté par un dirac). Montrer que la solution du système (4-2-11) complété pour $\ell = 0$ conduit aux

estimateurs :

$$A_0 = b_0 \int_{-R}^R \text{chax } Z(x) dx \quad , \quad A_1 = b_1 \int_{-R}^R \text{shax } Z(x) dx$$

$$(b_0 = \frac{a}{2 \text{sha } R} \quad , \quad b_1 = \frac{a^2}{2(a R \text{cha } R - \text{sha } R)}) \quad D^2(A_1) = 2 b_1 \text{cha } R \quad a = \sqrt{\frac{2}{C}}$$

(utiliser Ex. 7, Ch. 3 pour montrer que les mesures λ_0 et λ_1 sont du type $b_0 \text{chax } dx$ et $b_1 \text{shax } dx$, identifier b_0 et b_1 par les conditions d'universalité et μ_{11} par la première relation (4-2-11)).

Exercice 4 (Variogramme linéaire, cas discret) - Soit $Z(x)$ la même F.A. que dans l'exercice 1 (dérive quadratique et variogramme linéaire) dont la réalisation est connue aux $n+1$ points $x_i = i h$, $i = 0, 1, \dots, n$. On posera $L = nh$.

a/ Soient v_1 la mesure $\delta_L - \delta_0$ et v_2 la mesure $\sum_{i=0}^n \delta_{ih} - \frac{n+1}{2} (\delta_L - \delta_0)$. Montrer que ces mesures vérifient pour $j = 0, 1, \dots, n$:

$$\int v_1(dx) |x-jh| = (n-2j) h \quad , \quad \int v_2(dx) |x-jh| = -j(n-j) h$$

b/ En déduire les estimateurs optimaux de a_1 et a_2 :

$$A_1 = \frac{Z_n - Z_0}{Ah} + \frac{6}{(n-1)h} \left(\bar{Z} - \frac{Z_0 + Z_n}{2} \right) \quad (\bar{Z} = \frac{1}{n+1} \sum Z_i)$$

$$A_2 = -\frac{6}{h^2 n(n-1)} \left(\bar{Z} - \frac{Z_0 + Z_n}{2} \right)$$

et la matrice de covariance :

$$\mu_{11} = \frac{8n^2 - 2}{n(n^2 - 1)} \frac{1}{h} \quad , \quad \mu_{12} = \mu_{21} = -\frac{6}{n^2 - 1} \frac{1}{h^2} \quad , \quad \mu_{22} = \frac{6}{n(n^2 - 1)} \frac{1}{h^3}$$

(déduire de a/ que les mesures λ_1 et λ_2 sont des combinaisons linéaires de v_1 et v_2 ; déterminer les coefficients par les conditions d'universalité ; identifier les $\mu_{\beta\beta}$ sur le second membre de l'équation d'optimalité).

c/ Comparer avec Exercice 1, b/ - Pour cela, déterminer les estimateurs A_0' , A_1' et A_2'

du cas continu lorsque la réalisation est connue sur $(0, L)$.

[Poser $L = 2R$, placer l'origine des coordonnées en $-R$ et utiliser l'invariance tensorielle, d'où $A'_0 = A_0 - A_1 R + A_2 R^2$; $A'_1 = A_1 - 2 A_2 R$, $A'_2 = A_2$].

Exercice 5 (Covariance exponentielle, cas continu) -

a/ Montrer que, sur la droite, les mesures :

$$v_0(dx) = \frac{a}{2} dx + \frac{1}{2} (\delta_R + \delta_{-R}) ; \quad v_1(dx) = \frac{a}{2} x dx + \frac{1+aR}{2a} (\delta_R - \delta_{-R})$$

vérifient les équations intégrales : $(-R \leq y \leq R)$

$$\int_{-R}^R v_0(dx) e^{-a|x-y|} = 1, \quad \int_{-R}^R v_1(dx) e^{-a|x-y|} = y$$

b/ Soit sur la droite une F.A. admettant une dérive linéaire et la covariance $e^{-a|x-y|}$ dont on connaît la réalisation sur l'intervalle $(-R, R)$. Montrer que les estimateurs optimaux vérifient :

$$A_0 = \frac{aR}{1+aR} \bar{z} + \frac{1}{1+aR} \frac{z_R + z_{-R}}{2}$$

$$A_1 = \frac{a}{R(1+aR + \frac{1}{3} a^2 R^2)} \left[\frac{a}{2} \int_{-R}^R x z(x) dx + \frac{1+aR}{2a} (z_R - z_{-R}) \right]$$

$$\mu_{00} = \frac{1}{1+aR}, \quad \mu_{10} = \mu_{01} = 0, \quad \mu_{11} = \frac{a}{R(1+aR + \frac{1}{3} a^2 R^2)}$$

Exercice 6 (Cas des grandes mailles) -

a/ Les notations sont celles de l'Ex. 10 du Ch. 3 et des Ex. 13 à 15 du Ch. 2. On se donne, dans \mathbb{R}^2 , une F.A. dont la covariance $K(x-y)$ admet une portée petite, et N points expérimentaux x_α , $\alpha = 1, 2, \dots, N$ dont les distances mutuelles sont supérieures à la portée. On suppose qu'il existe une dérive de la forme $a_\ell f^\ell(x)$, $\ell = 1, \dots, k$. Montrer que les estimateurs optimaux A_ℓ sont identiques à ceux que donne la méthode des moindres carrés. Former leurs expressions et calculer la matrice des covariances $\mu_{\ell s}$ en fonction de la matrice

$$r^{\ell s} = \sum_{\alpha=1}^n f_{\alpha}^{\ell} f_{\alpha}^s \quad \text{et de son inverse } S_{\ell s}$$

$$(\mu_{\ell s} = C S_{\ell s} \quad , \quad \lambda_{\ell}^{\alpha} = S_{\ell s} f_{\alpha}^s)$$

b/ Calculer l'espérance mathématique de l'expression

$$I = \frac{1}{n} \sum_{\alpha} (Z_{\alpha} - A_{\ell} f_{\alpha}^{\ell})^2$$

[Profiter de la relation, à l'optimum, entre terme rectangle et terme quadratique. Il vient $I = (1 - \frac{k}{N}) C$].

Exercice 7 (Dérive linéaire sur la circonférence d'un cercle) -

a/ Soit, dans \mathbb{R}^2 , une F.A. dont la covariance $C(r)$ ne dépend que de $r = |x-y|$. On suppose connue la réalisation sur la circonférence du cercle de centre 0 et de rayon R. Montrer que s'il y a une dérive quadratique le système donnant les estimations optimales n'est pas régulier [les f^{ℓ} ne sont pas linéairement indépendantes sur S, à cause de $x^2 + y^2 = R^2$].

b/ On suppose que la F.A. précédente admet une dérive linéaire $a_0 + a_1 x + a_2 y$. Former les estimateurs optimaux A_0 , A_1 et A_2 , montrer qu'ils ne dépendent pas de la covariance $C(r)$ et coïncident avec ceux de la méthode des moindres carrés.

[Sur la circonférence, la dérive est $a_0 + a_1 R \cos \theta + a_2 R \sin \theta$; la covariance se développe en série de Fourier $C(2R |\sin \frac{\theta-\theta'}{2}|) = \sum b_p \cos(\theta-\theta')$ avec des b_p tous différents de 0 si $C(r)$ est de type positif strict. Les fonctions 1, $\sin \theta$ et $\cos \theta$ sont alors fonctions propres pour le noyau $C(\theta-\theta')$. On en déduira

$$A_0 = \frac{1}{2\pi} \int_0^{2\pi} Z(\theta) d\theta \quad , \quad A_1 = \frac{1}{\pi R} \int_0^{2\pi} Z(\theta) \cos \theta d\theta \quad , \quad A_2 = \frac{1}{\pi R} \int_0^{2\pi} Z(\theta) \sin \theta d\theta$$

La matrice des covariances dépend, seule, de $C(r) = \mu_{00} = b_0$, $\mu_{11} = \mu_{22} = \frac{b_1}{R^2}$, termes rectangles nuls].

4-7-2 Krigeage et cokrigeage.

Exercice 8 - Dans les conditions de l'Ex. 1, kriger le point $x_0 = R + h$ (appliquer le théorème d'additivité et exercice 6, a/ du Ch. 3 ; l'estimateur est $Z_U(x_0) = Z_R + A_1 h + A_2 h(2R + h)$ avec la variance

$$4 \frac{h^2}{R} + 3 \frac{h^3}{R^2} + 3 \frac{h^4}{4R^3} + 2h$$

rapidement croissante avec h : l'extrapolation n'est pas possible très loin).

Exercice 9 - Dans les conditions de l'Exercice 4, kriger le point $x_0 = ih + \varepsilon h$ ($0 \leq \varepsilon \leq 1$), $0 < i < n$.

(appliquer le théorème d'additivité, et l'Exercice 6, b/ du Ch. 3. On trouve

$$Z^*(x_0) = \varepsilon Z_{i+1} + (1-\varepsilon)Z_i - A_2 h^2 \varepsilon(1-\varepsilon), \quad \sigma_U^2 = 2 \varepsilon(1-\varepsilon)h + \frac{6 \varepsilon^2(1-\varepsilon)^2 h}{n(n^2-1)}$$

Exercice 10 - Dans les conditions de l'Exercice 5, kriger $x_0 = R + h$. (appliquer théorème d'additivité et Ex. 4 du Ch. 3).

Exercice 11 (Cas des grandes mailles) - Dans les conditions de l'Exercice 6, on veut estimer un domaine V grand vis-à-vis de la portée. On suppose que parmi les $N = n + n'$ points x_α , n sont intérieurs à V et n' extérieurs, et qu'ils sont tous situés à une distance de la frontière de V supérieure à la portée. Former l'estimateur optimal de V en présence d'une dérive $a_\ell f^\ell(x)$ et calculer la variance associée.

(appliquer le théorème d'additivité et l'exercice 10 du Ch. 3. Pour éviter les confusions de notation on désignera par B le terme A_1 de cet exercice. En posant

$$\lambda = \frac{1}{n+n'} + \frac{n'}{n+n'} \pi \frac{B}{CV}, \quad \lambda' = \frac{1}{n+n'} - \frac{n}{n+n'} \pi \frac{B}{CV}$$

en désignant par i et j' les échantillons intérieurs et extérieurs respectivement, l'estimateur est :

$$Z_U = \lambda \sum_{i=1}^n (Z_i - A_\ell f_i^\ell) + \lambda' \sum_{j'=1}^{n'} (Z_{j'} - f_{j'}^\ell, A_\ell)$$

avec la variance

$$\sigma_K^2 + \mu_{\ell_S} H^{\ell_S}, H^{\ell_S} = (\lambda \sum_i f_i^{\ell} + \lambda' \sum_j f_j^{\ell}), (\lambda \sum_i f_i^S + \lambda' \sum_j f_j^S)$$

Exercice 12 (Interpolateur polynomial) - Soit n points x_1, x_2, \dots, x_n sur la droite. Le polynome de degré $n-1$ tel que $P(x_i) = Z_i$ est :

$$P(x) = \sum_i \frac{(x-x_1)(x-x_2)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_1)(x_i-x_2)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)} Z_i$$

Former l'expression de l'interpolateur cubique associé aux quatre points

$$x_1 = -\frac{3}{2}a, x_2 = -\frac{a}{2}, x_3 = \frac{a}{2}, x_4 = \frac{3}{2}a, \text{ soit } b_0 + b_1x + b_2x^2 + b_3x^3 = P(x)$$

Intégrer de $-\frac{a}{2}$ à $\frac{a}{2}$. Montrer :

$$\frac{1}{a} \int_{-\frac{a}{2}}^{\frac{a}{2}} P(x) dx = b_0 + \frac{1}{12} b_2 a^2 = \frac{13}{12} \frac{Z_2+Z_3}{2} - \frac{1}{12} \frac{Z_1+Z_4}{2}$$

Comparer avec l'Exercice 1 du Ch. 3 pour le poids négatif des données externes.

Exercice 13 (Cokrigage en présence d'une erreur non systématique) - Soit $Z_0(x)$ une F.A., $C(x,y)$ sa covariance, $a_{\ell} f^{\ell}(x)$ sa dérivée avec des a_{ℓ} inconnus. Soit $\varepsilon(x)$ une autre F.A. d'espérance nulle, de covariance $C_{\varepsilon}(x,y)$, indépendante de Z_0 . On pose $Z_1(x) = Z_0(x) + \varepsilon(x)$.

a/ Montrer $C_{00} = C_{01} = C_{10} = C$ et $C_{11} = C + C_{\varepsilon}$.

b/ On suppose connue la réalisation de Z_1 sur S_1 (S_0 est vide). Montrer que les estimateurs optimaux $A_{\ell} = \int_{S_1} \lambda_{\ell}(dx) Z_1(x)$ s'obtiennent en résolvant le système :

$$\int_{S_1} \lambda_{\ell}(dx) [C(x,y) + C_{\varepsilon}(x,y)] = \mu_{\ell_S} f^S(y) \quad (y \in S_1)$$

$$\int_{S_1} \lambda_{\ell}(dx) f^S(x) = \delta_{\ell}^S$$

(identique à celui que l'on formerait pour estimer la dérive de Z_1).

c/ Kriger $Z_0(x_0)$ en un point x_0 (appartenant, ou non, à S_1) : montrer que $Z_0^*(x_0) =$

$$\int_{S_1} \lambda(dx) Z_1(x) \text{ avec}$$

$$\int_{S_1} \lambda(dx) [C(x,y) + \epsilon_\epsilon(x,y)] = C(x_0,y) + \mu_\epsilon f^\ell(y) \quad (y \in S_1)$$

$$\int_{S_1} \lambda(dx) f^\ell(x) = f^\ell(x_0)$$

(Ce système n'est pas identique à celui du krigeage de $Z_1(x_0)$; en particulier, $Z^*(x_0)$ n'est pas un interpolateur exact pour $Z_1(x)$).

Ecrire de même le système du krigeage de $Z_1(x_0)$, ($x_0 \notin S_1$) et comparer les deux variances dans l'hypothèse où C_ϵ est un effet de pépité (l'estimation de $Z_0(x_0)$ est meilleure que celle de $Z_1(x_0)$).

Exercice 14 (Cokrigeage en présence d'une erreur systématique) - Soit $Z_0(x)$ une F.A. stationnaire,

$C(h)$ sa covariance (connue), m_0 son espérance inconnue. Soit $\epsilon(x)$ une autre F.A. stationnaire, $C_\epsilon(h)$ sa covariance et ϵ_0 son espérance (inconnue). $\epsilon(x)$ est indépendante de $Z_0(x)$. On pose $Z_1(x) = Z_0(x) + \epsilon(x)$.

a/ Montrer $C_{00} = C_{01} = C_{10} = C$; $C_{11} = C + C_\epsilon$. Pour représenter les dérivées, introduire les deux fonctions $f(x,i)$ définies par :

$$f^0(x,0) = 1, f^0(x,1) = 0, f^1(x,0) = 0, f^1(x,1) = 1$$

La dérive $m(x,i)$ de $Z(x,i)$ est $a_0 f^0(x,i) + a_1 f^1(x,i)$, avec $a_0 = m_0$, $a_1 = m_1 = m_0 + \epsilon_0$. Montrer que la condition d'indépendance linéaire n'est réalisée que si les ensembles S_0 et S_1 sont tous deux non vides (si S_0 est vide, il n'y a aucun espoir de séparer, dans m_1 , la part de l'erreur systématique ϵ_0 et de la vraie espérance m_0).

b/ Former les équations des estimateurs optimaux $A_0 = \int_{S_0} \lambda_0^0(dx) Z_0(x) + \int_{S_1} \lambda_0^1(dx) Z_1(x)$ et $A_1 = \int_{S_0} \lambda_1^0(dx) Z_0(x) + \int_{S_1} \lambda_1^1(dx) Z_1(dx)$.

Les conditions d'universalité sont $\int_{S_1} \lambda_j^i(dx) = \delta_j^i$, celles d'optimalité :

$$\int_{S_0} \lambda_{\ell}^0 C(x,y) + \int_{S_1} \lambda_{\ell}^1 C(x,y) = \mu_{\ell 0} \quad (y \in S_0)$$

$$\int_{S_0} \lambda_{\ell}^0 C(x,y) + \int_{S_1} \lambda_{\ell}^1 [C(x,y) + C_{\varepsilon}(x,y)] = \mu_{\ell 1} \quad (y \in S_1)$$

c/ Kriger $Z_0(x_0)$ en un point $x_0 \notin S_0$.

(Universalité : $\int_{S_0} \lambda_0(dx) = 1$, $\int_{S_1} \lambda_1(dx) = 0$; optimalité :

$$\int_{S_0} \lambda_0(dx) C(x,y) + \int_{S_1} \lambda_1(dx) C(x,y) = C(x_0,y) + \mu_0 \quad (y \in S_0)$$

$$\int_{S_0} \lambda_0(dx) C(x,y) + \int_{S_1} \lambda_1(dx) [C(x,y) + C_{\varepsilon}(x,y)] = C(x_0,y) + \mu_1 \quad (y \in S_1)$$

Exercice 15 (Cas de simplification du cokrigage) -

a/ Si $S_1 = S_2 = \dots = S_d = S$ (les réalisations des $Z_i(x)$ étant connues sur le même ensemble S) et si les d dérivées $m_i(x)$ s'expriment à l'aide des mêmes fonctions $f^{\ell}(x)$ et sont linéairement indépendantes, il y a intérêt à changer les notations et à poser $m_i(x) = a_{i\ell} f^{\ell}(x)$ avec une matrice $a_{i\ell}$ inconnue. En désignant par $\sigma_{ij}(x,y)$ les covariances rectangles, par $A_{i\ell} = \lambda_{i\ell}^{j\alpha} Z_{j\alpha}$ (notation du cas fini $S = \{x_{\alpha}, \alpha = 1, 2, \dots, N\}$, $i, j = 1, 2, \dots, d$, $\ell = 0, 1, \dots, k$) l'estimateur optimal de $a_{i\ell}$, par $Z_i^*(x_0) = \lambda_{i\ell}^{j\alpha} Z_{j\alpha}$ le krigage de $x_0 \notin S$, former les systèmes d'équations des $\lambda_{i\ell}^{j\alpha}$ et des $\lambda_{i\ell}^{j\alpha}$.

(On trouve respectivement :

$$\left\{ \begin{array}{l} \lambda_{i\ell}^{j\alpha} r_{\alpha}^s = \delta_i^j \delta_{\ell}^s \\ \lambda_{i\ell}^{j\beta} \sigma_{jj';\beta\alpha} = \mu_{ij';\ell s} r_{\alpha}^s \end{array} \right. \quad (\mu_{ij';\ell s} = \text{Cov } A_{i\ell} A_{j's})$$

et :

$$\left\{ \begin{array}{l} \lambda_{i\ell}^{j\alpha} r_{\alpha}^s = \delta_i^j r^s(x_0) \\ \lambda_{i\ell}^{j\beta} \sigma_{jj';\beta\alpha} = \mu_{ij';\ell s} r_{\alpha}^s \end{array} \right.$$

b/ Soit B_j^i une matrice $d \times d$ régulière, et $B_j^{i'}$ son inverse. On considère les fonctions aléatoires $Z_i' = B_1^j Z_j$ déduites des précédentes par la matrice B_1^j . Montrer

que les coefficients $\lambda_{i\ell}^{j\alpha}$ et $\lambda_1^{j\alpha}$ des estimateurs optimaux précédents varient suivant une loi une fois covariante et une fois contravariante, tandis que les paramètres de Lagrange se modifient de manière covariante (c'est-à-dire $\mu_{1j}^i, - = B_1^{i'} B_j^{j'} \mu_{1,j}^i, - ; \lambda_{j'}^{i'} = B_1^{i'} B_j^{j'} \lambda_j^i$).

c/ Soient $Z_1(x)$ d F.A. comme ci-dessus, indépendantes, admettant toutes la même fonction de covariance $\sigma(x,y)$. Montrer que $\sigma_{jj',\beta\alpha} = 0$ ou $\sigma_{\beta\alpha}$ selon que $j \neq j'$ ou $j = j'$. En déduire que les estimateurs optimaux de la dérive et du krigeage pour la F.A. $Z_1(x)$ ne font pas intervenir les autres F.A. $Z_j, j \neq 1$, soit :

$$A_{i\ell} = \lambda_{\ell}^{\alpha} Z_{i\alpha} \quad Z_i^*(x_0) = \lambda^{\alpha} Z_{i\alpha}$$

avec des coefficients (indépendants de i) vérifiant

$$\left\{ \begin{array}{l} \lambda_{\ell}^{\alpha} r_{\alpha}^s = \delta_{\ell}^s \\ \lambda_{\ell}^{\beta} \sigma_{\alpha\beta} = \mu_{\ell s} r_{\alpha}^s \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{l} \lambda^{\alpha} r_{\alpha}^s = r^s(x_0) \\ \lambda^{\beta} \sigma_{\alpha\beta} = \mu_{\ell} r_{\alpha}^{\ell} \end{array} \right.$$

d/ Soient $Z_1^i(x)$ des F.A. définies comme en a/ telles que $\sigma_{ij}(x,y) = K_{ij} \sigma(x,y)$ avec une matrice strictement positive K_{ij} et une fonction de covariance $\sigma(x,y)$. Montrer que l'on peut trouver i fonctions aléatoires $Z_i(x)$ indépendantes comme en c/ et admettant la même fonction de covariance $\sigma(x,y)$, telles que $Z_1^i(x) = B_1^j Z_j(x)$ avec une matrice B_1^j vérifiant :

$$K_{ii'} = \sum_j B_1^j B_1^{j'}$$

(diagonaliser la matrice K_{ij}).

En déduire que les coefficients des estimateurs optimaux sont les mêmes qu'au paragraphe c/, donc que l'on estime la dérive ou le krigeage de l'une de ces F.A. comme si elle était seule. ($B_1^{i'} B_j^{j'}, \delta_j^i = \delta_{j'}^{i'}$)

(appliquer la loi de variation des coefficients démontrée en b/ : dans les applications, ce cas où toutes les covariances rectangles sont proportionnelles à une même covariance fixe est particulièrement intéressant, en raison des simplifications énormes qui se produisent alors).

4-7-3 Les indéterminations de la covariance sous-jacente.

Exercice 16 (n variables gaussiennes corrélées) - Soient X_1, X_2, \dots, X_n n variables gaussiennes admettant la même espérance m , la même variance σ^2 et telles que les X_i, X_j aient, deux à deux, le même coefficient de corrélation $\rho_{ij} = \rho$ ($i \neq j$).

a/ Montrer que la matrice des covariances est $\sigma_{ij} = [(1-\rho) \delta_{ij} + \rho] \sigma^2$, et admet comme inverse :

$$B_{ij} = \frac{1}{(1-\rho)\sigma^2} \left(\delta_{ij} - \frac{\rho}{1 + (n-1)\rho} \right)$$

Calculer le déterminant de σ_{ij} ($\sigma^{2n} (1-\rho)^{n-1} (1+(n-1)\rho)$). En déduire que la densité $f(x_1, \dots, x_n)$ de ces variables vérifie :

$$\begin{aligned} -2 \log f &= n \log 2\pi + n \log \sigma^2 + (n-1) \log (1-\rho) + [1 + (n-1)\rho] \\ &+ \sum_{i,j} B_{ij} (x_i - m)(x_j - m) \end{aligned}$$

(On remarquera que les valeurs propres de σ_{ij} sont $(1-\rho) \sigma^2$ avec l'ordre $n-1$ (vecteurs propres : a_i tels que $\sum a_i = 0$) et $1 + (n-1)\rho$ avec l'ordre 1 (vecteurs propres : $a_i = 1$). Noter la condition $1 \geq \rho \geq -\frac{1}{n-1}$)

b/ On suppose connue une réalisation (x_1, x_2, \dots, x_n) de (X_1, X_2, \dots, X_n) . Former les estimateurs du maximum de vraisemblance m^* , s^2 et r de m , σ^2 et ρ . On posera pour cela :

$$S_1 = \sum_1 (x_i - m) \quad S_2 = \sum_1 (x_i - m)^2$$

et on notera $\frac{\partial S_1}{\partial m} = -n$, $\frac{\partial S_2}{\partial m} = -2 S_1$, et :

$$\sum_{i,j} B_{ij} (x_i - m)(x_j - m) = \frac{1}{(1-\rho)\sigma^2} \left(S_2 - \frac{\rho}{1 + (n-1)\rho} S_1^2 \right)$$

Montrer que l'équation en m donne $S_1 = 0$, soit $m^* = \frac{1}{n} \sum x_i$, que l'équation en s^2 conduit à :

$$(1-r) s^2 = \frac{1}{n} S_2$$

et que l'équation en r conduit à un résultat absurde, où ne figurent plus les données expérimentales.

Conclure que l'on peut (à la rigueur) estimer $(1-\rho) \sigma^2$, mais non séparément ρ et σ^2 .

c/ Expliquer ce résultat à l'aide de la théorie des estimateurs quadratiques universels.

($Q^{ij} X_i X_j$ est universel si $\sum_{ij} Q^{ij} = 0$. Mais $E(Q^{ij} X_i X_j)$ est alors égal à $\sum_{ij} Q^{ij} \sigma_{ij} = (1-\rho) \sigma^2 \sum_1 Q^{ii}$: $(1-\rho) \sigma^2$ représente la partie accessible de la covariance, et on ne peut pas séparer ρ et σ^2).

d/ Montrer que ces conclusions subsistent pour k variables vectorielles indépendantes (X_1^k, \dots, X_n^k) dont chacune admet la même covariance σ_{ij} que ci-dessus, mais une espérance $m_k = E(X_1^k)$ différente.

Exercice 17 - Soit, sur la droite, une F.A. admettant une covariance $\sigma(x,y)$ et une dérivée linéaire $a_0 + a_1 x$. On ne connaît ni la covariance, ni les coefficients a_0 et a_1 . On suppose la réalisation connue sur l'intervalle $(-R, R)$.

a/ Pour estimer a_0 et a_1 on utilise les estimateurs universels (mais évidemment non optimaux en général) $A_0 = \frac{Z_R + Z_{-R}}{2}$, $A_1 = \frac{Z_R - Z_{-R}}{2R}$ (cf exercice 2). Montrer

$$D^2(A_0) = \frac{1}{4} (\sigma_{RR} + \sigma_{-R,-R} + 2 \sigma_{R,-R}) \quad , \quad D^2(A_1) = \frac{1}{4R^2} (\sigma_{RR} + \sigma_{-R,-R} - 2 \sigma_{R,-R})$$

et $\text{Cov}(A_0, A_1) = \text{Cov}(A_1, A_0) = 0$. (Ces deux variances représentent la partie inaccessible de la covariance).

En déduire la forme générale des covariances $\tau(x,y)$ admissibles, en fonction de l'une d'elles $\sigma(x,y)$:

$$\tau(x,y) = \sigma(x,y) + D_0 + D_1 x y \quad (x,y \in (-R, R))$$

b/ On suppose que, parmi toutes les covariances possibles, la covariance "vraie" inconnue est la plus stationnaire possible. Pour l'estimer, on part d'une solution particulière $\tau(x,y)$, égale, par exemple, à $E(Z_x Z_y - m_x^* m_y^*)$, obtenue expérimentalement. On pose :

$$\sigma(x, x+h) = \tau(x, x+h) - D_0 - D_1 x(x+h) \quad (0 \leq h \leq 2R)$$

$$\bar{\sigma}(h) = \frac{1}{2R-h} \int_{-R}^{R-h} \sigma(x, x+h) dx, \quad \bar{\tau}(h) = \frac{1}{2R-h} \int_{-R}^{R-h} \tau(x, x+h)$$

d'où

$$\bar{\sigma}(h) = \bar{\tau}(h) - D_0 + D_1 \left[\frac{1}{6} (2R-h)^2 - R(R-h) \right]$$

Déterminer D_1 de manière à minimiser l'intégrale :

$$\int_0^{2R} dh \int_{-R}^{R-h} [\sigma(x, x+h) - \bar{\sigma}(h)]^2 dx$$

(noter que D_0 s'élimine et reste indéterminé).

BIBLIOGRAPHIE

- [1] CABLIER, A. - Contribution aux méthodes d'estimation des gisements d'uranium.
(Thèses) - Fontenay-aux-Roses, 1964.
- [2] FORMERY, Ph. - Cours de Géostatistique - Ecole des Mines de Nancy (non publié).
- [3] MATHERON, G. - Traité de Géostatistique Appliquée. Tome 1 (1962), Tome 2 (1963).
Editions Technip, Paris.
- [4] MATHERON, G. - Les Variables Régionalisées et leur estimation. (Thèse) - 1965.
Masson et Cie, Paris.
- [5] MATHERON, G. - Osnovy Prikladnoï geostatistiki. 1968, Editions Mir, Moscou.
- [6] MATHERON, G. - Le Krigeage Universel - Les Cahiers du Centre de Morphologie Mathématique,
Fascicule 1, 1969 - Fontainebleau.
- [7] SERRA, J. - Echantillonnage et Estimation locale des phénomènes de transition miniers.
(Thèse) - Nancy, 1967.
- [8] FORMERY, Ph. et MATHERON, G. - Recherche d'optima dans la reconnaissance et la mise en
exploitation des gisements miniers. Annales des Mines, Mai et Juin 1963.

