

Cours de Géostatistique

Multivariable

Hans WACKERNAGEL

Publication C-146
4e Edition: Mars 1993

Centre de Géostatistique, Ecole des Mines de Paris
35 rue Saint Honoré, 77305 Fontainebleau, France

L'Analyse des Données est "un outil pour dégager de la gangue des données le pur diamant de la véridique nature".

J.-P. Benzécri (d'après [96])

Préface

Nous avons rassemblé dans ce document des notes du cours de Géostatistique Multivariée dispensé dans le cadre du D.E.A. au Centre de Géostatistique de l'École des Mines de Paris.

Le premier Chapitre énumère quelques résultats du calcul matriciel (sans démonstrations), ce qui permet ensuite, dans les chapitres suivants, une présentation accélérée de trois méthodes de l'Analyse des Données, importantes pour la Géostatistique, à savoir: l'analyse en composantes principales, l'analyse canonique et l'analyse des correspondances.

Le cinquième Chapitre donne des définitions élémentaires du variogramme et de la fonction de covariance, précise la notion de variogramme gîgogne et présente trois modèles de régionalisation particuliers donnant lieu à des structures gîgognes. Le sixième Chapitre examine le krigeage des composantes de la variation spatiale définies par les modèles linéaires de régionalisation.

Le septième Chapitre présente une première méthode géostatistique multivariée, qui consiste à introduire dans le système de krigeage, sous la forme de dérives externes, des variables connues sur tout le domaine spatial. Le huitième Chapitre parle du filtrage d'une dérive temporelle dans le cas de données spatio-temporelles.

Le neuvième Chapitre étudie la fonction de covariance croisée d'une paire de variables, le variogramme croisé, ainsi que la relation entre ces deux fonctions structurales. Le dixième Chapitre expose la théorie du cokrigeage en liaison avec les notions d'isotopie/hétérotopie, de corrélation intrinsèque et d'autokrigeabilité.

Le onzième et dernier Chapitre est consacré au modèle linéaire de corégionalisation et à l'analyse krigeante multivariée, qui permettent de pratiquer l'Analyse des Données dans un contexte spatial.

Ce document est destiné à évoluer au fil du temps: toutes remarques, commentaires et critiques sont bienvenues!

Note

Ce document a servi de point de départ à la rédaction du livre *Multivariate Geostatistics* paru chez Springer-Verlag, Berlin, 256p, ISBN 3-540-60127-9.

Table des matières

Préface	ii
1 Rappel de Calcul Matriciel	1
Le tableau des données	1
Matrice, vecteur, scalaire	1
Addition	2
Multiplication par un scalaire	2
Multiplication de deux matrices	2
Transposition	2
Matrice carrée	3
Matrice diagonale	3
Matrice orthogonale	3
Matrice symétrique	3
Indépendance linéaire	4
Rang d'une matrice	4
Matrice inverse	5
Déterminant d'une matrice	5
Trace	5
Valeurs propres	6
Vecteurs propres	6
Matrice définie positive	6
Décomposition en valeurs et vecteurs propres	7
Décomposition en valeurs singulières	8
Inverse généralisée de Moore-Penrose	9
2 Analyse en Composantes Principales	11
Transformation en facteurs	11
Maximisation de la variance d'un facteur	12
Interprétation des variances des facteurs	13
Corrélation des variables avec les facteurs	14
3 Analyse Canonique	16
Etude des facteurs de deux groupes de variables	16
Intermède:décomposition en valeurs singulières	17

	Maximisation de la corrélation	17
4	Analyse des Correspondances	19
	Tableau disjonctif	19
	Tableau de contingence	19
	Analyse canonique de tableaux disjonctifs	20
	Codage d'une variable quantitative	20
	Contingences entre deux variables quantitatives	20
	Analyse des correspondances continue	21
5	Modèle Linéaire de Régionalisation	22
	Variogramme expérimental	22
	Variable régionalisée et fonction aléatoire	23
	Variogramme théorique	23
	Fonction de covariance	24
	Fonction de type positif	24
	Fonction de type négatif conditionnel	25
	Ajustement du variogramme par une fonction de covariance	26
	Variogramme gigogne	27
	Décomposition de la fonction aléatoire	28
	Régionalisation stationnaire d'ordre deux	29
	Régionalisation avec une composante intrinsèque	29
	Régionalisation localement stationnaire	29
6	Krigeage de Composantes Spatiales	30
	Krigeage ordinaire	30
	Krigeage de la moyenne locale	31
	Krigeage de la composante intrinsèque	32
	Krigeage d'une composante stationnaire	33
7	Dérive externe	38
	Deux variables mesurant la même chose	38
	Estimation avec une fonction de forme	38
	Dérive invariante par translation	39
	Covariance généralisée	40
	Krigeage avec dérive externe	41
8	Dérive temporelle	49
	Dérive spatiale et temporelle	49
	Filtrage de la variation temporelle	49
9	Covariance Croisée	51
	Fonction de covariance croisée	51
	Effet de retard	51

Variogramme croisé	52
Représentation spectrale	53
Densité spectrale	55
Déphasage	56
Corrélation intrinsèque	56
10 Cokrigeage	58
Isotopie et hétérotopie	58
Cokrigeage ordinaire	58
Cokrigeage simple	60
Autokrigeabilité	61
11 Analyse Krigeante	63
Modèle linéaire de corégionalisation	63
Ajustement bivarié de variogrammes expérimentaux	64
Ajustement multivarié	64
Analyse de corégionalisation	65
Cokrigeage de facteurs	68
Solution des Exercices	69
Références	75
Mots-clefs	75
Bibliographie	76

1 Rappel de Calcul Matriciel

Ce bref rappel consiste en l'énoncé des principaux concepts du calcul matriciel, parsemé de quelques exercices et exemples. Pour un exposé plus détaillé, on pourra consulter avec profit le livre de STRANG [94].

Le tableau des données

En Analyse des Données, on manipule des tableaux de nombres (matrices) et le tableau de départ est généralement le tableau des données \mathbf{Z}

$$\begin{array}{c} \text{Echantillons} \\ \text{(lignes)} \end{array} \begin{array}{c} \text{Variables} \\ \text{(colonnes)} \end{array} \begin{pmatrix} z_{1,1} & \cdots & z_{1,i} & \cdots & z_{1,N} \\ \vdots & & \vdots & & \vdots \\ z_{\alpha,1} & & z_{\alpha,i} & & z_{\alpha,N} \\ \vdots & & \vdots & & \vdots \\ z_{n,1} & \cdots & z_{n,i} & \cdots & z_{n,N} \end{pmatrix} = [z_{\alpha i}] = \mathbf{Z}$$

L'élément $z_{\alpha i}$ dénote la valeur numérique placée au croisement de la ligne numéro α (indice de l'échantillon) et de la colonne numéro i (indice de la variable).

Matrice, vecteur, scalaire

Une matrice est un tableau rectangulaire de nombres

$$\mathbf{A} = [a_{ij}]$$

avec des indices

$$i = 1, \dots, n \quad j = 1, \dots, m$$

On parle d'une matrice d'ordre $n \times m$ pour désigner une matrice de n lignes et m colonnes.

Un vecteur (par convention: un vecteur colonne) de dimension n est une matrice d'ordre $n \times 1$, c'est-à-dire une matrice ne possédant qu'une seule colonne.

Un scalaire est un nombre (si l'on veut: une matrice d'ordre 1×1).

Du point de vue des notations, les matrices seront désignées par des majuscules en caractères gras et les vecteurs par des minuscules en caractères gras.

Addition

Deux matrices de même ordre peuvent être additionnées

$$\mathbf{A} + \mathbf{B} = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}]$$

L'addition de matrices se fait en additionnant les éléments de même indice i et j .

Multiplication par un scalaire

Une matrice peut être multipliée à droite ou à gauche par un scalaire λ

$$\mathbf{A} \lambda = \lambda \mathbf{A} = [\lambda a_{ij}]$$

Cela revient à multiplier tous les éléments a_{ij} par le scalaire λ .

Multiplication de deux matrices

Le produit \mathbf{AB} de deux matrices \mathbf{A} et \mathbf{B} ne peut être réalisé que si \mathbf{B} a autant de lignes que \mathbf{A} a de colonnes. Soit \mathbf{A} d'ordre $n \times m$ et \mathbf{B} d'ordre $m \times l$. La multiplication de \mathbf{A} par \mathbf{B} donne une matrice \mathbf{C} d'ordre $n \times l$,

$$\underset{(n \times m)}{\mathbf{A}} \cdot \underset{(m \times l)}{\mathbf{B}} = \left[\sum_{j=1}^m a_{ij} b_{jk} \right] = [c_{ik}] = \underset{(n \times l)}{\mathbf{C}}$$

où $i = 1, \dots, n$ et $k = 1, \dots, l$.

Le produit \mathbf{BA} de ces mêmes matrices n'est possible que si n est égal à l et le résultat est une matrice d'ordre $m \times m$.

Transposition

La transposée \mathbf{A}^\top d'une matrice \mathbf{A} d'ordre $n \times m$ est obtenue en inversant la séquence des indices, de sorte que les lignes de \mathbf{A} deviennent les colonnes de \mathbf{A}^\top

$$\underset{(n \times m)}{\mathbf{A}} = [a_{ij}], \quad \underset{(m \times n)}{\mathbf{A}^\top} = [a_{ji}]$$

La transposée du produit de deux matrices est égal au produit des transposées en séquence inverse

$$(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$$

EXERCISE 1.1 Soit $\mathbf{1}$ le vecteur de dimension n dont les éléments sont tous égaux à 1. Effectuer les produits $\mathbf{1}^\top \mathbf{1}$ et $\mathbf{1}\mathbf{1}^\top$.

EXERCISE 1.2 Calculer les matrices résultant des produits

$$\frac{1}{n}\mathbf{Z}^\top \mathbf{1} \quad \text{et} \quad \frac{1}{n}\mathbf{1}\mathbf{1}^\top \mathbf{Z}$$

où \mathbf{Z} est la matrice $n \times N$ des données.

Matrice carrée

Une matrice carrée a autant de lignes que de colonnes.

Matrice diagonale

Une matrice diagonale \mathbf{D} est une matrice carrée dont les seuls éléments non nuls se trouvent sur la diagonale

$$\mathbf{D} = \begin{pmatrix} d_{11} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_{nn} \end{pmatrix}$$

En particulier, il y a la matrice identité

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

qui, multipliant à droite ou à gauche une matrice de même ordre, ne la modifie pas

$$\mathbf{A}\mathbf{I} = \mathbf{I}\mathbf{A} = \mathbf{A}$$

Matrice orthogonale

Une matrice \mathbf{A} carrée est orthogonale, si elle vérifie

$$\mathbf{A}^\top \mathbf{A} = \mathbf{A}\mathbf{A}^\top = \mathbf{I}$$

Matrice symétrique

Une matrice carrée \mathbf{A} est symétrique, si elle est égale à sa transposée

$$\mathbf{A} = \mathbf{A}^\top$$

EXEMPLE 1.3 Un exemple de matrice symétrique est la matrice de variance-covariance \mathbf{V} , contenant les variances sur la diagonale et les covariances en dehors de celle-ci

$$\mathbf{V} = [\sigma_{ij}] = \frac{1}{n} (\mathbf{Z} - \mathbf{M})^\top (\mathbf{Z} - \mathbf{M})$$

où \mathbf{M} est la matrice rectangulaire $n \times m$ des moyennes (solution de l'exercice 1.2), dont tous les éléments de chaque colonne sont égaux à la moyenne de la variable correspondant à cette colonne.

EXEMPLE 1.4 Un autre exemple de matrice symétrique est la matrice \mathbf{R} des corrélations

$$\mathbf{R} = [\rho_{ij}] = \mathbf{D}_{\sigma^{-1}} \mathbf{V} \mathbf{D}_{\sigma^{-1}}$$

où $\mathbf{D}_{\sigma^{-1}}$ est la matrice diagonale contenant les inverses des écarts-types des variables

$$\mathbf{D}_{\sigma^{-1}} = \begin{pmatrix} \frac{1}{\sqrt{\sigma_{11}}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{\sqrt{\sigma_{NN}}} \end{pmatrix}$$

Indépendance linéaire

Un ensemble de vecteurs $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$ est linéairement indépendant, si il n'existe pas d'ensemble non trivial de scalaires $\{x_1, \dots, x_m\}$, tel que

$$\sum_{j=1}^m \mathbf{a}_j x_j = \mathbf{0}$$

Autrement dit, l'indépendance linéaire des colonnes d'une matrice \mathbf{A} est acquise, si seul le vecteur nul $\mathbf{x} = \mathbf{0}$ satisfait à l'équation $\mathbf{A}\mathbf{x} = \mathbf{0}$.

Rang d'une matrice

Une matrice rectangulaire peut être subdivisée en l'ensemble des vecteurs colonnes qui la constituent. De même, on peut considérer les vecteurs "lignes" de cette matrice, qui sont en fait les vecteurs colonnes de sa transposée.

Le rang des colonnes d'une matrice est le nombre maximal de vecteurs colonnes linéairement indépendants. Le rang des lignes de la matrice est défini de manière analogue. On démontre que le rang des lignes est égal au rang des colonnes.

Le rang d'une matrice \mathbf{A} rectangulaire $n \times m$ est donc au plus égal à la plus petite de ses deux dimensions

$$\text{rang}(\mathbf{A}) \leq \min(n, m)$$

Le rang de la matrice \mathbf{A} indique la dimension¹ des espaces vectoriels $\mathcal{R}(\mathbf{A}^\top)$ et $\mathcal{R}(\mathbf{A})$ engendrés par les colonnes et les lignes de \mathbf{A}

$$\mathcal{R}(\mathbf{A}^\top) = \{\mathbf{y} : \mathbf{y} = \mathbf{A}\mathbf{x}\} \quad \text{et} \quad \mathcal{R}(\mathbf{A}) = \{\mathbf{x} : \mathbf{x} = \mathbf{y}^\top \mathbf{A}\}$$

où \mathbf{x} désigne un vecteur de dimension m et \mathbf{y} un vecteur de dimension n .

Matrice inverse

Une matrice \mathbf{A} carrée $n \times n$ est singulière, si $\text{rang}(\mathbf{A}) < n$, et non singulière, si $\text{rang}(\mathbf{A}) = n$.

Si \mathbf{A} est non singulière (invertible), il existe une matrice inverse \mathbf{A}^{-1} tel que

$$\mathbf{A} \mathbf{A}^{-1} = \mathbf{A}^{-1} \mathbf{A} = \mathbf{I}$$

L'inverse \mathbf{Q}^{-1} d'une matrice orthogonale \mathbf{Q} est sa transposée \mathbf{Q}^\top .

Déterminant d'une matrice

Le déterminant $|\mathbf{A}|$ d'une matrice \mathbf{A} carrée $n \times n$ est

$$|\mathbf{A}| = \det(\mathbf{A}) = \sum (-1)^{N(k_1, \dots, k_n)} \prod_{i=1}^n a_{ik_i}$$

où la somme porte sur l'ensemble des permutations (k_1, \dots, k_n) des entiers $(1, \dots, n)$ et où $N(k_1, \dots, k_n)$ est le nombre de transpositions de deux entiers nécessaires pour passer de l'ensemble de départ $(1, \dots, n)$ à la permutation (k_1, \dots, k_n) de cet ensemble.

Dans le cas d'une matrice 2×2 on a la formule bien connue

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \det(\mathbf{A}) = ad - bc$$

Un déterminant non nul indique que la matrice est invertible.

Trace

La trace d'une matrice carrée $n \times n$ est égale à la somme de ses éléments diagonaux

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii}$$

1. ne pas confondre: la dimension d'un vecteur et la dimension d'un espace vectoriel.

Valeurs propres

Soit \mathbf{A} une matrice carrée $n \times n$. L'équation caractéristique

$$|\lambda \mathbf{I} - \mathbf{A}| = 0$$

a n solutions λ en général complexes, appelées valeurs propres de \mathbf{A} .

La somme des valeurs propres est égale à la trace de la matrice

$$\text{tr}(\mathbf{A}) = \sum_{p=1}^n \lambda_p$$

Le produit des valeurs propres est égal au déterminant de la matrice

$$\det(\mathbf{A}) = \prod_{p=1}^n \lambda_p$$

Si \mathbf{A} est symétrique, toutes les valeurs propres sont réelles.

Vecteurs propres

Soit \mathbf{A} une matrice carrée et λ une valeur propre de \mathbf{A} . Il existe des vecteurs \mathbf{x} et \mathbf{y} , différents du vecteur nul $\mathbf{0}$, satisfaisant

$$(\lambda \mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}, \quad \mathbf{y}^\top (\lambda \mathbf{I} - \mathbf{A}) = \mathbf{0}$$

c'est-à-dire

$$\mathbf{A} \mathbf{x} = \lambda \mathbf{x}, \quad \mathbf{y}^\top \mathbf{A} = \lambda \mathbf{y}^\top$$

Les vecteurs \mathbf{x} sont les vecteurs propres des colonnes de \mathbf{A} et les \mathbf{y} sont les vecteurs propres des lignes de \mathbf{A} .

Lorsque \mathbf{A} est symétrique, il n'y a pas lieu de distinguer entre des vecteurs propres de colonnes et de lignes.

EXERCISE 1.5 Montrer que les valeurs propres du carré d'une matrice \mathbf{A} symétrique, $\mathbf{A}^2 = \mathbf{A}\mathbf{A}$, sont égales au carré des valeurs propres de \mathbf{A} . Et que tout vecteur propre de \mathbf{A} est un vecteur propre de \mathbf{A}^2 .

Matrice définie positive

DÉFINITION:

Une matrice \mathbf{A} symétrique $n \times n$ est définie positive, ssi (si et seulement si) pour tout vecteur \mathbf{x} non nul la forme quadratique

$$\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$$

De même, A est qualifiée de semi-définie positive (définie non négative), ssi $x^T A x \geq 0$ pour tout vecteur x . Ou encore, A est indéfinie, ssi $x^T A x > 0$ pour certains x et $x^T A x < 0$ pour d'autres x .

On remarque que cette définition est identique à celle d'une fonction de type positif.

Suivent trois critères très utiles, énoncés pour des matrices semi-définies positives.

PREMIER CRITÈRE:

A est semi-définie positive, ssi il existe une matrice W telle que $A = W^T W$.

Une telle matrice W s'écrit parfois \sqrt{A} .

DEUXIÈME CRITÈRE:

A est semi-définie positive, ssi toutes les valeurs propres $\lambda_p \geq 0$.

TROISIÈME CRITÈRE:

A est semi-définie positive, ssi tous les mineurs principaux sont non négatifs.

Un mineur principal est le déterminant d'une sous-matrice principale de A . Une sous-matrice principale est obtenue en biffant k colonnes ($k = 0, 1, \dots, n-1$) et les lignes correspondantes les croisant sur les éléments diagonaux. La combinatoire des mineurs principaux à vérifier rend ce critère peu intéressant dans les applications avec $n > 3$.

REMARQUE 1.6 *On vérifie aisément avec le troisième critère que la matrice $(n+1) \times (n+1)$ membre gauche du Krigeage Ordinaire, écrite en covariances, n'est pas semi-définie positive.*

La sous-matrice principale S obtenue en biffant toutes les colonnes et les lignes se croisant en des éléments diagonaux, sauf les deux dernières,

$$S = \begin{pmatrix} c_{nn} & 1 \\ 1 & 0 \end{pmatrix}$$

a un déterminant qui vaut -1 .

Le calcul des valeurs propres de la matrice membre gauche du KO donne une valeur propre négative (dûe à la condition d'universalité) et n valeurs propres positives (ou nulles) dans le cas de covariances. Constituée à l'aide de la fonction de covariance, cette matrice est indéfinie, alors qu'elle est (semi-)définie négative dans le cas du variogramme.

Décomposition en valeurs et vecteurs propres

Pour une matrice A symétrique, les valeurs propres λ_p et des vecteurs propres q_p normés à l'unité forment le système

$$A Q = Q \Lambda \quad \text{avec} \quad Q^T Q = I$$

où Λ est la matrice diagonale des valeurs propres et Q est la matrice orthogonale des vecteurs propres.

Etant donné que $Q^T = Q^{-1}$, on aboutit à une décomposition de la matrice symétrique A ,

$$A = Q \Lambda Q^T$$

Décomposition en valeurs singulières

L'Analyse des Données étant l'art de décomposer des tableaux, une décomposition s'appliquant à une matrice rectangulaire (dans le même esprit que la décomposition d'une matrice symétrique en valeurs et vecteurs propres) va évidemment jouer un rôle central.

La décomposition en valeurs singulières μ_p d'une matrice A rectangulaire $n \times m$ de rang r s'écrit:

$$\begin{array}{ccccc} A & = & Q_1 & \cdot & \Sigma & \cdot & Q_2^T \\ (n \times m) & & (n \times n) & & (n \times m) & & (m \times m) \end{array}$$

où Q_1 et Q_2 sont des matrices orthogonales et où Σ est une matrice rectangulaire avec r valeurs positives μ_p sur la diagonale (à savoir l'ensemble des éléments d'indices égaux) et des zéros ailleurs. Par exemple, dans le cas où $n > m$ et $r = m$, la matrice Σ a la structure suivante

$$\Sigma = \begin{pmatrix} \mu_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mu_r \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{pmatrix}$$

Une telle décomposition existe toujours et peut être obtenue en calculant les valeurs propres λ_p de AA^T et de $A^T A$, qui sont identiques et positives ou nulles. Les valeurs singulières sont les racines carrées des valeurs propres non nulles,

$$\mu_p = \sqrt{\lambda_p}$$

Dans cette décomposition, Q_1 est la matrice des vecteurs propres de AA^T , tandis que Q_2 est la matrice des vecteurs propres de $A^T A$.

EXERCISE 1.7 *Quelle est la décomposition en valeurs singulières d'une matrice symétrique?*

Inverse généralisée de Moore-Penrose

Une matrice inverse existe pour toute matrice carrée non singulière. Il est intéressant de généraliser le concept d'inverse à des matrices singulières et à des matrices rectangulaires.

Une matrice \mathbf{X} d'ordre $m \times n$ est inverse généralisée de Moore-Penrose d'une matrice \mathbf{A} rectangulaire $n \times m$, si elle vérifie les quatre conditions suivantes

$$\begin{aligned}\mathbf{A X A} &= \mathbf{A} \\ \mathbf{X A X} &= \mathbf{X} \\ (\mathbf{A X})^\top &= \mathbf{A X} \\ (\mathbf{X A})^\top &= \mathbf{X A}\end{aligned}$$

Une telle matrice \mathbf{X} est notée \mathbf{A}^+ .

EXERCISE 1.8 *Est-ce que l'inverse \mathbf{A}^{-1} d'une matrice carrée non singulière \mathbf{A} est une inverse généralisée de Moore-Penrose ?*

L'inverse de Moore-Penrose s'obtient à partir de la décomposition en valeurs singulières en intervertissant les deux matrices orthogonales, en transposant Σ et en inversant chacune des valeurs singulières:

$$\mathbf{A}^+ = \mathbf{Q}_2 \Sigma^+ \mathbf{Q}_1^\top$$

La matrice Σ^+ est d'ordre $m \times n$ et a, dans le cas où $n > m$ et $r = m$, la structure

$$\Sigma^+ = \begin{pmatrix} \mu_1^{-1} & 0 & 0 & 0 & \dots & 0 \\ 0 & \ddots & 0 & 0 & \dots & 0 \\ 0 & 0 & \mu_r^{-1} & 0 & \dots & 0 \end{pmatrix} = \begin{pmatrix} \lambda_1^{-1/2} & 0 & 0 & 0 & \dots & 0 \\ 0 & \ddots & 0 & 0 & \dots & 0 \\ 0 & 0 & \lambda_r^{-1/2} & 0 & \dots & 0 \end{pmatrix}$$

EXEMPLE 1.9 (KRIGEAGE SIMPLE AVEC UN DOUBLON) *On peut se demander si l'inverse de Moore-Penrose peut être utilisée pour résoudre un système de krigeage dont le membre gauche est singulier. Nous n'allons pas traiter le problème de manière générale, mais nous contenter d'un petit exercice très simple.*

On cherche à estimer une valeur en \mathbf{x}_0 à partir de deux valeurs situées au même point \mathbf{x}_1 . Le système de krigeage simple est

$$\begin{pmatrix} a & a \\ a & a \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} b \\ b \end{pmatrix}$$

où a est la variance des données et où b est la covariance entre les points \mathbf{x}_1 et \mathbf{x}_0 .

La matrice \mathbf{A} étant singulière on va recourir à l'inverse généralisée de MP.

Soit

$$\mathbf{A}\mathbf{A}^\top = \mathbf{A}^\top\mathbf{A} = \begin{pmatrix} 2a^2 & 2a^2 \\ 2a^2 & 2a^2 \end{pmatrix} = \begin{pmatrix} c & c \\ c & c \end{pmatrix} = \mathbf{C}$$

On a

$$\begin{aligned} \det(\mathbf{C}) = 0 &\Rightarrow \lambda_2 = 0 \\ \text{tr}(\mathbf{C}) = 2c &\Rightarrow \lambda_1 = 2c \end{aligned}$$

et une matrice de vecteurs propres normés

$$\mathbf{Q} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

La solution du système au sens de l'inverse généralisée est

$$\mathbf{w} = \mathbf{A}^+ \mathbf{b} = \mathbf{Q} \Sigma^+ \mathbf{Q}^\top \mathbf{b} = \begin{pmatrix} \frac{b}{\sqrt{2c}} \\ \frac{b}{\sqrt{2c}} \end{pmatrix} = \begin{pmatrix} \frac{b}{2a} \\ \frac{b}{2a} \end{pmatrix}$$

En considérant des données de moyenne nulle, on a en fin de compte l'estimation par krigeage simple

$$z^*(\mathbf{x}_0) = \frac{b}{a} \left(\frac{z_1(\mathbf{x}_1) + z_2(\mathbf{x}_1)}{2} \right)$$

Cette solution satisfait l'esprit, puisque l'estimateur prend la moyenne des deux valeurs z_1 et z_2 mesurées au point \mathbf{x}_1 et la multiplie par le pondérateur que livre le krigeage simple lorsqu'il n'y a qu'une seule information disponible.

2 Analyse en Composantes Principales

Transformation en facteurs

Le problème de base que résoud l'analyse en composantes principales est de transformer un jeu de variables corrélées en des variables non corrélées, qui, dans un contexte idéal (Gaussien) pourraient être interprétées comme des facteurs indépendants sous-jacents au phénomène. C'est pourquoi ces quantités orthogonales seront appelées "facteurs", bien que cette interprétation ne soit pas toujours parfaitement adéquate.

\mathbf{Z} est la matrice $n \times N$ des données centrées et \mathbf{V} est la matrice $N \times N$ de variance-covariance correspondante

$$\mathbf{V} = [\sigma_{ij}] = \frac{1}{n} \mathbf{Z}^\top \mathbf{Z}$$

Soit \mathbf{Y} une matrice $n \times N$ contenant dans ses lignes les n échantillons de facteurs Y_p ($p = 1, \dots, N$) non corrélés et d'espérances nulles.

La matrice de variance-covariance des facteurs est évidemment diagonale, puisque les covariances entre les facteurs sont nulles par définition

$$\mathbf{D} = \frac{1}{n} \mathbf{Y}^\top \mathbf{Y} = \begin{pmatrix} d_{11} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_{NN} \end{pmatrix}$$

et les éléments diagonaux d_{pp} sont les variances des facteurs.

On cherche une matrice \mathbf{A} , orthogonale $N \times N$, qui transforme linéairement les variables mesurées en facteurs synthétiques

$$\mathbf{Y} = \mathbf{Z} \mathbf{A} \quad \text{avec} \quad \mathbf{A}^\top \mathbf{A} = \mathbf{I}$$

En multipliant cette équation à gauche par $\frac{1}{n}$ et par \mathbf{Y}^\top on a

$$\frac{1}{n} \mathbf{Y}^\top \mathbf{Y} = \frac{1}{n} \mathbf{Y}^\top \mathbf{Z} \mathbf{A}$$

et en remplaçant \mathbf{Y} par $\mathbf{Z} \mathbf{A}$ à droite, il vient

$$\frac{1}{n} (\mathbf{Z} \mathbf{A})^\top (\mathbf{Z} \mathbf{A}) = \frac{1}{n} \mathbf{A}^\top \mathbf{Z}^\top \mathbf{Z} \mathbf{A} = \mathbf{A}^\top \frac{1}{n} (\mathbf{Z}^\top \mathbf{Z}) \mathbf{A}$$

En fin de compte,

$$\mathbf{D} = \mathbf{A}^\top \mathbf{V} \mathbf{A}$$

c'est-à-dire

$$\mathbf{V} \mathbf{A} = \mathbf{A} \mathbf{D}$$

On voit immédiatement que la matrice \mathbf{Q} de vecteurs propres orthonormés offre une solution au problème, et que les valeurs propres λ_p sont alors tout simplement les variances des facteurs Y_p . L'analyse en composantes principales n'est rien d'autre qu'une interprétation statistique du problème de valeur propre

$$\mathbf{V} \mathbf{Q} = \mathbf{Q} \mathbf{\Lambda} \quad \text{avec } \mathbf{Q}^\top \mathbf{Q} = \mathbf{I}$$

définissant des facteurs par

$$\mathbf{Y} = \mathbf{Z} \mathbf{Q}$$

Maximisation de la variance d'un facteur

Un autre aspect important de l'analyse en composantes principales est qu'elle permet de définir une suite de facteurs orthogonaux, qui absorbent successivement une part maximale de la variance des données.

Prenons le vecteur \mathbf{y}_1 correspondant à un premier facteur obtenu par transformation de la matrice des données \mathbf{Z} avec un vecteur \mathbf{a}_1 calibré à l'unité

$$\mathbf{y}_1 = \mathbf{Z} \mathbf{a}_1 \quad \text{avec } \mathbf{a}_1^\top \mathbf{a}_1 = 1$$

La variance de \mathbf{y}_1 vaut

$$\text{var}(\mathbf{y}_1) = \frac{1}{n} \mathbf{y}_1^\top \mathbf{y}_1 = \frac{1}{n} \mathbf{a}_1^\top \mathbf{Z}^\top \mathbf{Z} \mathbf{a}_1 = \mathbf{a}_1^\top \mathbf{V} \mathbf{a}_1$$

Pour attribuer une part maximale de la variance des données à \mathbf{y}_1 , on définit une fonction objectif ϕ_1 avec un paramètre de Lagrange λ_1 , qui multiplie la condition de norme unité du vecteur \mathbf{a}_1

$$\phi_1 = \mathbf{a}_1^\top \mathbf{V} \mathbf{a}_1 - \lambda_1 (\mathbf{a}_1^\top \mathbf{a}_1 - 1)$$

En annulant la dérivée selon \mathbf{a}_1

$$\frac{\partial \phi_1}{\partial \mathbf{a}_1} = 0 \quad \iff \quad 2\mathbf{V} \mathbf{a}_1 - 2\lambda_1 \mathbf{a}_1 = 0$$

on voit que λ_1 est une valeur propre de la matrice de variance-covariance et que \mathbf{a}_1 est égal au vecteur propre \mathbf{q}_1 associé à cette valeur propre

$$\mathbf{V} \mathbf{q}_1 = \lambda_1 \mathbf{q}_1$$

On s'intéresse à un deuxième vecteur \mathbf{y}_2 orthogonal au premier

$$\text{cov}(\mathbf{y}_2, \mathbf{y}_1) = \text{cov}(\mathbf{Z}\mathbf{a}_2, \mathbf{Z}\mathbf{a}_1) = \mathbf{a}_2^\top \mathbf{V}\mathbf{a}_1 = \mathbf{a}_2^\top \lambda_1 \mathbf{a}_1 = 0$$

La fonction ϕ_2 à maximiser contient deux contraintes: le fait que l'on choisisse \mathbf{a}_2 normé à l'unité, et l'orthogonalité de \mathbf{a}_2 avec \mathbf{a}_1 . Ces contraintes nécessitent l'introduction de deux nouveaux pondérateurs de Lagrange λ_2 et μ

$$\phi_2 = \mathbf{a}_2^\top \mathbf{V}\mathbf{a}_2 - \lambda_2 (\mathbf{a}_2^\top \mathbf{a}_2 - 1) + \mu \mathbf{a}_2^\top \mathbf{a}_1$$

On annule la dérivée selon \mathbf{a}_2

$$\frac{\partial \phi_2}{\partial \mathbf{a}_2} = 0 \quad \iff \quad 2\mathbf{V}\mathbf{a}_2 - 2\lambda_2 \mathbf{a}_2 + \mu \mathbf{a}_1 = 0$$

Que vaut μ ? En multipliant l'équation par \mathbf{a}_1^\top à gauche

$$2 \underbrace{\mathbf{a}_1^\top \mathbf{V}\mathbf{a}_2}_0 - 2\lambda_2 \underbrace{\mathbf{a}_1^\top \mathbf{a}_2}_0 + \mu \underbrace{\mathbf{a}_1^\top \mathbf{a}_1}_1 = 0$$

on voit que μ est égal à zéro (la contrainte n'est pas active) et ainsi

$$\mathbf{V}\mathbf{a}_2 = \lambda_2 \mathbf{a}_2$$

A nouveau λ_2 s'avère être une valeur propre de la matrice de variance-covariance et \mathbf{a}_2 un vecteur propre \mathbf{q}_2 correspondant.

En numérotant les valeurs propres de \mathbf{V} , de la plus forte à la plus faible, on obtient une suite de N facteurs non corrélés, qui assurent un redécoupage optimal (au sens des moindres carrés) de la variance totale, puisque

$$\text{tr}(\mathbf{V}) = \sum_{i=1}^N \sigma_{ii} = \sum_{p=1}^N \lambda_p$$

Interprétation des variances des facteurs

Les valeurs propres indiquent la part de la variance totale associée à chaque facteur et le rapport

$$\frac{\text{variance du facteur}}{\text{variance totale}} = \frac{\lambda_p}{\text{tr}(\mathbf{V})}$$

donne une indication numérique, habituellement indiquée en %, de l'importance du facteur.

Il est généralement préférable de standardiser les variables (en soustrayant d'emblée les moyennes et en divisant par les écarts-types), de sorte que l'analyse en composantes principales est basée sur la matrice de corrélation \mathbf{R} . Dans ce cadre, lorsqu'une valeur propre est inférieure à 1, on considère parfois que le facteur associé a moins de valeur explicative que chacune des variables mesurées, car sa variance est inférieure à la variance unité des variables.

Corrélation des variables avec les facteurs

En général on travaille avec des données standardisées $\tilde{z}_{\alpha i}$ pour ramener toutes les variables à une échelle unique et les rendre comparables

$$\tilde{z}_{\alpha i} = \frac{z_{\alpha i} - m_i}{\sqrt{\sigma_{ii}}}$$

où m_i et $\sqrt{\sigma_{ii}}$ sont la moyenne et l'écart-type de la variable z_i .

La matrice de variance-covariance associée aux données standardisées est la matrice de corrélation

$$\mathbf{R} = \frac{1}{n} \tilde{\mathbf{Z}}^\top \tilde{\mathbf{Z}}$$

dont la décomposition en valeurs propres donne

$$\mathbf{R} = \tilde{\mathbf{Q}} \tilde{\Lambda} \tilde{\mathbf{Q}}^\top = \tilde{\mathbf{Q}} \sqrt{\tilde{\Lambda}} \left(\tilde{\mathbf{Q}} \sqrt{\tilde{\Lambda}} \right)^\top = \mathbf{A} \mathbf{A}^\top$$

Les vecteurs \mathbf{a}_i , colonnes de \mathbf{A}^\top , ont ceci de remarquable qu'ils indiquent la corrélation entre une variable \tilde{z}_i et les différents facteurs y_p , car

$$\text{corr}(\tilde{z}_i, y_p) = \sqrt{\tilde{\lambda}} \tilde{q}_{ip} = a_{ip}$$

La longueur des vecteurs \mathbf{a}_i est égale à l'unité, et leur produit scalaire est égal au coefficient de corrélation

$$\mathbf{a}_i^\top \mathbf{a}_j = \rho_{ij}$$

Du point de vue géométrique les vecteurs \mathbf{a}_i peuvent être utilisés pour représenter la position des variables sur la surface de l'hypersphère unité centrée à l'origine. Les coefficients de corrélation ρ_{ij} sont les cosinus des angles entre deux vecteurs se rapportant à des variables différentes.

La projection de la position des variables depuis la surface de l'hypersphère dans des plans définis par des axes factoriels donne une représentation graphique, appelée "cercle des corrélations", qui montre leur proximité à l'intérieur d'un cercle unité centré à l'origine. Ces représentations sont utiles pour juger de l'affinité ou de l'antagonisme existant entre des variables. Ces graphiques sont fiables, si les variables auxquelles on s'intéresse sont situées près de la circonférence du cercle. Sinon, il faut examiner plusieurs plans factoriels pour s'assurer qu'une proximité constatée sur un cercle correspond effectivement à une proximité sur l'hypersphère.

Dans le cas plus général de données non standardisées, on peut également fabriquer des cercles de corrélation entre les variables et des paires de facteurs. On multiplie la matrice \mathbf{V} à gauche et à droite par la matrice diagonale des inverses des écarts-types

$$\mathbf{D}_{\sigma^{-1}} \mathbf{V} \mathbf{D}_{\sigma^{-1}} = \mathbf{D}_{\sigma^{-1}} \mathbf{Q} \Lambda \mathbf{Q}^\top \mathbf{D}_{\sigma^{-1}} = \mathbf{D}_{\sigma^{-1}} \mathbf{Q} \sqrt{\Lambda} \left(\mathbf{D}_{\sigma^{-1}} \mathbf{Q} \sqrt{\Lambda} \right)^\top$$

d'où l'on déduit la formule générale des corrélations entre variables et facteurs

$$\text{corr}(\mathbf{z}_i, \mathbf{y}_p) = \sqrt{\frac{\lambda_p}{\sigma_{ii}}} q_{ip}$$

EXERCISE 2.1 *Déduire de*

$$\frac{1}{n} \tilde{\mathbf{Z}}^\top \mathbf{Y} = \mathbf{R} \tilde{\mathbf{Q}} = \tilde{\mathbf{Q}} \tilde{\mathbf{\Lambda}} \quad \text{avec } \tilde{\mathbf{Q}}^\top \tilde{\mathbf{Q}} = \mathbf{I}$$

que le coefficient de corrélation entre deux vecteurs $\tilde{\mathbf{z}}_i$ et \mathbf{y}_p vaut

$$\rho_{ip} = \sqrt{\tilde{\lambda}_p} \tilde{q}_{ip}$$

Il s'ensuit que, lorsqu'une variable standardisée est orthogonale à toutes les autres, elle sera identifiée à un facteur de variance unité dans l'analyse en composantes principales.

EXERCISE 2.2 *Soit $\tilde{\mathbf{Z}}$ une matrice de données standardisées et \mathbf{Y} la matrice de facteurs correspondants obtenue par analyse en composantes principales, $\mathbf{Y} = \tilde{\mathbf{Z}} \tilde{\mathbf{Q}}$, où $\tilde{\mathbf{Q}}$ est la matrice des vecteurs propres orthonormés de la matrice de corrélation \mathbf{R} .*

Montrer que:

$$\text{corr}(\tilde{\mathbf{Z}}, \tilde{\mathbf{Z}}) = \mathbf{R} = \text{corr}(\tilde{\mathbf{Z}}, \mathbf{Y}) \left[\text{corr}(\tilde{\mathbf{Z}}, \mathbf{Y}) \right]^\top$$

3 Analyse Canonique

Etude des facteurs de deux groupes de variables

L'analyse canonique propose d'étudier simultanément deux groupes de variables mesurant des propriétés différentes sur le même jeu d'échantillons. Ceci afin de mettre en évidence, successivement, des couples formés d'un facteur du premier groupe ayant un lien maximal avec un facteur du deuxième groupe. Le but étant, au moment de l'interprétation, de voir pour chaque facteur sous-jacent à un groupe de propriétés, s'il peut être mis en rapport avec un facteur sous-jacent à un autre ensemble de mesures de nature différente, réalisées sur les mêmes échantillons.

Le tableau de données centré \mathbf{Z} d'ordre $n \times N$ est scindé en deux groupes de variables

$$\mathbf{Z} = [\mathbf{Z}_1 | \mathbf{Z}_2]$$

c'est-à-dire, la matrice \mathbf{Z} est formée par juxtaposition des colonnes d'une matrice \mathbf{Z}_1 d'ordre $n \times M$ et d'une matrice \mathbf{Z}_2 d'ordre $n \times L$, avec $N = M + L$.

La matrice de variance-covariance \mathbf{V} associée à \mathbf{Z} est subdivisée en quatre blocs

$$\mathbf{V} = \frac{1}{n} [\mathbf{Z}_1 | \mathbf{Z}_2]^\top [\mathbf{Z}_1 | \mathbf{Z}_2] = \begin{pmatrix} \mathbf{V}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{V}_{22} \end{pmatrix}$$

où les matrices \mathbf{V}_{11} d'ordre $M \times M$ et \mathbf{V}_{22} d'ordre $L \times L$ sont des matrices de variance-covariance associées à \mathbf{Z}_1 et \mathbf{Z}_2 . Les matrices de covariances \mathbf{C}_{12} d'ordre $M \times L$ et \mathbf{C}_{21} d'ordre $L \times M$ contiennent les covariances entre variables de groupes différents

$$\mathbf{C}_{12}^\top = \mathbf{C}_{21}$$

On cherche des couples de facteurs $\{\mathbf{u}_p, \mathbf{v}_p\}$

$$\mathbf{u}_p = \mathbf{Z}_1 \mathbf{a}_p \quad \text{et} \quad \mathbf{v}_p = \mathbf{Z}_2 \mathbf{b}_p$$

qui soient non corrélés à l'intérieur de leurs groupes respectifs

$$\mathbf{u}_p \perp \mathbf{u}_k \quad \text{et} \quad \mathbf{v}_p \perp \mathbf{v}_k \quad \text{pour } p \neq k$$

et pour lesquels la corrélation $\text{corr}(\mathbf{u}_p, \mathbf{v}_p)$ est maximale. Cette corrélation (appelée corrélation canonique) entre deux facteurs \mathbf{u} et \mathbf{v} est

$$\text{corr}(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{a}^\top \mathbf{Z}_1^\top \mathbf{Z}_2 \mathbf{b}}{\sqrt{\mathbf{a}^\top \mathbf{Z}_1^\top \mathbf{Z}_1 \mathbf{a}} \cdot \sqrt{\mathbf{b}^\top \mathbf{Z}_2^\top \mathbf{Z}_2 \mathbf{b}}}$$

En décidant de normer \mathbf{a} et \mathbf{b} selon la métrique donnée par les matrices de variance-covariance (supposées définies positives) de leurs groupes respectifs

$$\mathbf{a}^\top \mathbf{V}_{11} \mathbf{a} = \mathbf{b}^\top \mathbf{V}_{22} \mathbf{b} = 1$$

la corrélation s'écrit plus simplement

$$\text{corr}(\mathbf{u}, \mathbf{v}) = \mathbf{a}^\top \mathbf{C}_{12} \mathbf{b}$$

Intermède: décomposition en valeurs singulières

Il est instructif de se livrer à un petit détour algébrique. On pose

$$\mathbf{x} = \sqrt{\mathbf{V}_{11}} \mathbf{a} \quad \text{et} \quad \mathbf{y} = \sqrt{\mathbf{V}_{22}} \mathbf{b} \quad \text{avec} \quad \mathbf{x}^\top \mathbf{x} = \mathbf{y}^\top \mathbf{y} = 1$$

de sorte que

$$\text{corr}(\mathbf{u}, \mathbf{v}) = \mathbf{x}^\top \underbrace{\left[\left(\sqrt{\mathbf{V}_{11}} \right)^{-1} \mathbf{C}_{12} \left(\sqrt{\mathbf{V}_{22}} \right)^{-1} \right]}_{\mathbf{G}} \mathbf{y}$$

Pour le système complet des vecteurs \mathbf{x} et \mathbf{y} on a

$$\mathbf{X}^\top \mathbf{G} \mathbf{Y} = \mathbf{\Sigma}$$

c'est-à-dire la décomposition en valeurs singulières

$$\mathbf{G} = \mathbf{X} \mathbf{\Sigma} \mathbf{Y}^\top$$

où les corrélations canoniques μ , seuls éléments non nuls de la matrice rectangulaire $\mathbf{\Sigma}$, se révèlent être les valeurs singulières du tableau rectangulaire \mathbf{G} .

Maximisation de la corrélation

On cherche une paire de vecteurs de transformation $\{\mathbf{a}, \mathbf{b}\}$ rendant maximale la corrélation entre les facteurs $\{\mathbf{u}, \mathbf{v}\}$

$$\text{corr}(\mathbf{u}, \mathbf{v}) = \mathbf{a}^\top \mathbf{C}_{12} \mathbf{b} \quad \text{avec} \quad \mathbf{a}^\top \mathbf{V}_{11} \mathbf{a} = \mathbf{b}^\top \mathbf{V}_{22} \mathbf{b} = 1$$

La fonction objectif ϕ est

$$\phi = \mathbf{a}^\top \mathbf{C}_{12} \mathbf{b} - \frac{1}{2} \mu (\mathbf{a}^\top \mathbf{V}_{11} \mathbf{a} - 1) - \frac{1}{2} \mu' (\mathbf{b}^\top \mathbf{V}_{22} \mathbf{b} - 1)$$

où μ et μ' sont des paramètres de Lagrange.

L'annulation des dérivées partielles selon \mathbf{a} et \mathbf{b} donne

$$\frac{\partial \phi}{\partial \mathbf{a}} = 0 \quad \Longleftrightarrow \quad \mathbf{C}_{12} \mathbf{b} - \mu \mathbf{V}_{11} \mathbf{a} = 0$$

$$\frac{\partial \phi}{\partial \mathbf{b}} = 0 \quad \Longleftrightarrow \quad \mathbf{C}_{21} \mathbf{a} - \mu' \mathbf{V}_{22} \mathbf{b} = 0$$

Quelle est la relation entre μ et μ' ? En prémultipliant la première équation par \mathbf{a}^\top et la seconde par \mathbf{b}^\top

$$\mathbf{a}^\top \mathbf{C}_{12} \mathbf{b} - \mu \mathbf{a}^\top \mathbf{V}_{11} \mathbf{a} = 0 \quad \Longleftrightarrow \quad \mu = \mathbf{a}^\top \mathbf{C}_{12} \mathbf{b}$$

$$\mathbf{b}^\top \mathbf{C}_{21} \mathbf{a} - \mu' \mathbf{b}^\top \mathbf{V}_{22} \mathbf{b} = 0 \quad \Longleftrightarrow \quad \mu' = \mathbf{b}^\top \mathbf{C}_{21} \mathbf{a}$$

on voit que

$$\mu = \mu' = \text{corr}(\mathbf{u}, \mathbf{v})$$

Ayant trouvé plus haut que

$$\mathbf{C}_{12} \mathbf{b} = \mu \mathbf{V}_{11} \mathbf{a} \quad \text{et} \quad \mathbf{C}_{21} \mathbf{a} = \mu' \mathbf{V}_{22} \mathbf{b}$$

on prémultiplie la première équation par $\mathbf{C}_{21} \mathbf{V}_{11}^{-1}$

$$\mathbf{C}_{21} \mathbf{V}_{11}^{-1} \mathbf{C}_{12} \mathbf{b} = \mu \mathbf{C}_{21} \mathbf{a}$$

et en tenant compte de la deuxième équation

$$\mathbf{C}_{21} \mathbf{V}_{11}^{-1} \mathbf{C}_{12} \mathbf{b} = \mu^2 \mathbf{V}_{22} \mathbf{b}$$

Cela donne, au bout du compte, deux problèmes de valeur propre, en posant $\lambda = \mu^2$

$$\mathbf{V}_{22}^{-1} \mathbf{C}_{21} \mathbf{V}_{11}^{-1} \mathbf{C}_{12} \mathbf{b} = \lambda \mathbf{b}$$

$$\mathbf{V}_{11}^{-1} \mathbf{C}_{12} \mathbf{V}_{22}^{-1} \mathbf{C}_{21} \mathbf{a} = \lambda \mathbf{a}$$

Seul le plus petit de ces deux systèmes a besoin d'être résolu. La solution de l'autre peut alors être obtenue par les relations de transition entre les vecteurs de transformation

$$\mathbf{a} = \frac{1}{\sqrt{\lambda}} \mathbf{V}_{11}^{-1} \mathbf{C}_{12} \mathbf{b} \quad \text{et} \quad \mathbf{b} = \frac{1}{\sqrt{\lambda}} \mathbf{V}_{22}^{-1} \mathbf{C}_{21} \mathbf{a}$$

EXERCISE 3.1 *Montrer que* $\text{cov}(\mathbf{u}_p, \mathbf{u}_k) = \delta_{pk}$.

EXERCISE 3.2 *Montrer que* $\text{cov}(\mathbf{u}_p, \mathbf{v}_k) = \mu \delta_{pk}$.

EXERCISE 3.3 *Montrer que le vecteur propre \mathbf{a} dans l'équation*

$$\mathbf{V}_{11} \mathbf{a} = \lambda \mathbf{C}_{12} \mathbf{V}_{22}^{-1} \mathbf{C}_{21} \mathbf{a}$$

résoud le problème de l'analyse canonique.

4 Analyse des Correspondances

On a vu que l'analyse canonique étudie, en quelque sorte, les correspondances entre des facteurs de deux groupes de variables quantitatives. La même approche, appliquée à deux variables qualitatives, qui représentent, chacune d'elles, un groupe de catégories mutuellement exclusives, est connue sous le nom évocateur d'analyse des correspondances.

Tableau disjonctif

Une variable qualitative z est un système de catégories (classes) mutuellement exclusives. L'appartenance d'un échantillon z_α à une catégorie \mathcal{C}_i peut être représentée numériquement à l'aide d'une fonction indicatrice

$$1_{\mathcal{C}_i}(z_\alpha) = \begin{cases} 1 & \text{si } z_\alpha \in \mathcal{C}_i \\ 0 & \text{sinon} \end{cases}$$

La matrice indiquant les appartenances des n échantillons aux N catégories de la variable qualitative est le tableau disjonctif \mathbf{H} d'ordre $n \times N$

$$\mathbf{H} = \left[1_{\mathcal{C}_i}(z_\alpha) \right]$$

Chaque ligne du tableau disjonctif ne contient qu'un seul élément égal à 1 et a des zéros ailleurs, étant donné qu'il est exclu qu'un échantillon appartienne à deux catégories à la fois.

Le produit de \mathbf{H}^\top avec \mathbf{H} donne une matrice diagonale, dont les éléments diagonaux n_{ii} indiquent le nombre d'échantillons se trouvant dans la catégorie numéro i . La division des n_{ii} par n donne la proportion F_{ii} d'échantillons contenus dans une certaine catégorie et l'on obtient

$$\mathbf{V} = \frac{1}{n} \mathbf{H}^\top \mathbf{H} = \begin{pmatrix} F_{11} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & F_{NN} \end{pmatrix}$$

Tableau de contingence

Deux variables qualitatives mesurées sur le même ensemble d'échantillons sont représentées, respectivement, par un tableau \mathbf{H}_1 d'ordre $n \times M$ et un tableau \mathbf{H}_2 d'ordre $n \times L$.

Le produit des deux tableaux disjonctifs a pour éléments les effectifs n_{ij} d'échantillons appartenant simultanément à la catégorie i de la première variable qualitative et à la catégorie j de la deuxième variable. Les éléments de ce tableau, qui est le tableau de contingence d'ordre $M \times L$, peuvent être divisés par le nombre total n d'échantillons, mesurant ainsi la proportion d'échantillons se trouvant au croisement des catégories des deux variables

$$\mathbf{C}_{12} = \frac{1}{n} \mathbf{H}_1^\top \mathbf{H}_2 = [F_{ij}]$$

Analyse canonique de tableaux disjonctifs

L'analyse des correspondances consiste à appliquer le formalisme de l'analyse canonique aux tableaux disjonctifs \mathbf{H}_1 et \mathbf{H}_2 en formant les matrices (non centrées)

$$\mathbf{C}_{21}^\top = \mathbf{C}_{12}, \quad \mathbf{V}_{11} = \frac{1}{n} \mathbf{H}_1^\top \mathbf{H}_1, \quad \mathbf{V}_{22} = \frac{1}{n} \mathbf{H}_2^\top \mathbf{H}_2$$

Vu la structure diagonale des matrices \mathbf{V}_{11} et \mathbf{V}_{22} , qui n'interviennent que dans le calibrage des vecteurs de transformation, on peut dire que l'analyse des correspondances revient à étudier les liens entre les lignes et les colonnes du tableau de contingence \mathbf{C}_{12} .

Codage d'une variable quantitative

Une variable quantitative z peut être transformée en une variable qualitative en la découpant en un ensemble d'intervalles disjoints \mathcal{C}_i . Des fonctions indicatrices de ces intervalles permettent de constituer le tableau disjonctif \mathbf{H} associé à z

$$1_{\mathcal{C}_i}(z_\alpha) = \begin{cases} 1 & \text{si } z_\alpha \text{ appartient à l'intervalle } \mathcal{C}_i \\ 0 & \text{sinon} \end{cases}$$

Une subdivision est parfois donnée a priori, comme c'est par exemple le cas pour des fractions granulométriques d'un échantillon de sol.

Le codage d'une variable quantitative permet d'effectuer une analyse des correspondances croisant les intervalles de cette variable avec les catégories d'une variable qualitative.

La matrice diagonale $\mathbf{V} = (1/n) \mathbf{H}^\top \mathbf{H}$ contient les valeurs de l'histogramme des échantillons z_α , puisque ses éléments diagonaux exhibent les fréquences d'apparition des échantillons dans les classes de valeurs de z .

Contingences entre deux variables quantitatives

Le tableau de contingence $\mathbf{C}_{12} = (1/n) \mathbf{H}_1^\top \mathbf{H}_2$ de deux variables quantitatives z_1 et z_2 renferme l'information sur un histogramme bivariable, car on y trouve les fréquences des échantillons au croisement des classes de z_1 et de z_2 .

Analyse des correspondances continue

L'histogramme bivariable, c'est-à-dire le tableau de contingence de deux variables quantitatives, peut être modélisé par une loi bivariable $F(dZ_1, dZ_2)$. Si cette loi admet une densité $\varphi(Z_1, Z_2)$, elle s'écrit

$$F(dZ_1, dZ_2) = \varphi(Z_1, Z_2) F_1(dZ_1) F_2(dZ_2)$$

où $F_1(dZ_1)$, $F_2(dZ_2)$ sont les lois marginales de Z_1 et Z_2 , modélisant les histogrammes respectifs. En supposant, en outre, que la densité φ est de carré intégrable, on peut la décomposer en un système de valeurs propres $\lambda_p = \mu_p^2$ et de fonctions propres $f_p(Z_1)$, $g_p(Z_2)$

$$\varphi(Z_1, Z_2) = \sum_{p=0}^{\infty} \mu_p f_p(Z_1) g_p(Z_2)$$

Cette décomposition est l'analogie, en continu, de la décomposition en valeurs singulières.

EXEMPLE 4.1 Lorsque la loi bivariable est Gaussienne, la décomposition en valeurs propres donne des coefficients μ_p égaux à un coefficient ρ mis à la puissance p

$$\mu_p = \rho^p \quad \text{avec } |\rho| < 1$$

ainsi que des fonctions propres égales à des polynômes d'Hermite η_p normalisés

$$\eta_p = \frac{H_p}{\sqrt{p!}}$$

La loi bivariable Gaussienne $G(dZ_1, dZ_2)$ s'écrit

$$G(dZ_1, dZ_2) = \sum_{p=0}^{\infty} \rho^p \eta_p(Z_1) \eta_p(Z_2) G(dZ_1) G(dZ_2)$$

REMARQUE 4.2 En géostatistique non linéaire, on s'intéresse à la loi bivariable de deux variables aléatoires $Z(\mathbf{x}_\alpha)$ et $Z(\mathbf{x}_\beta)$ situées en deux points \mathbf{x}_α et \mathbf{x}_β d'un domaine spatial. La décomposition de cette loi bivariable donne lieu, sous certaines hypothèses, à un modèle isofactoriel. Les modèles isofactoriels sont d'une grande utilité pour la mise en oeuvre de ce qui s'appelle, non par hasard, le krigeage disjonctif.

En particulier, dans le cas du modèle isofactoriel Gaussien, le coefficient ρ sera tout simplement égal à la valeur $\rho(\mathbf{x}_\alpha - \mathbf{x}_\beta)$ de la fonction de corrélation pour la paire de points concernée.

EXERCISE 4.3 Qu'est-ce qu'un modèle isofactoriel qualifié de 'discret'?

5 Modèle Linéaire de Régionalisation

Variogramme expérimental

Nous allons mesurer la variabilité à différentes échelles d'une variable régionalisée $z(\mathbf{x})$ tout simplement en calculant une mesure de dissemblance entre deux données z_1 et z_2 situées en deux points \mathbf{x}_1 et \mathbf{x}_2 d'un domaine spatial. Cette dissemblance entre deux valeurs, désignée par γ^* , vaudra

$$\gamma^* = \frac{(z_2 - z_1)^2}{2}$$

c'est-à-dire la moitié du carré de la différence entre deux valeurs.

On fait dépendre la dissemblance γ^* de la distance et de l'orientation d'une paire de points, décrits par le vecteur $\mathbf{h} = \mathbf{x}_2 - \mathbf{x}_1$, indifféremment de la position de la paire dans le domaine étudié

$$\gamma^*(\mathbf{h}) = \frac{1}{2} \left(z(\mathbf{x}_1 + \mathbf{h}) - z(\mathbf{x}_1) \right)^2$$

En formant la moyenne des dissemblances γ^* entre valeurs pour toutes les $n_{\mathbf{h}}$ paires de points reliées par un vecteur \mathbf{h} donné pour une maille donnée (avec, le cas échéant, une certaine tolérance sur la longueur et l'angle du vecteur), on obtient la notion de variogramme expérimental

$$\gamma^*(\mathbf{h}) = \frac{1}{2n_{\mathbf{h}}} \sum_{\alpha=1}^{n_{\mathbf{h}}} \left(z(\mathbf{x}_{\alpha} + \mathbf{h}) - z(\mathbf{x}_{\alpha}) \right)^2$$

Habituellement on observe que la dissemblance des valeurs augmente en moyenne en fonction de l'éloignement spatial des points de mesure et atteint fréquemment un palier de variation aux grandes distances. Lorsque la pente du variogramme change abruptement, on peut penser à des paliers intermédiaires.

Le comportement aux très petites échelles, près de l'origine du variogramme, est d'une importance capitale, car il est un indicateur du degré de continuité de la variable régionalisée, à savoir: différentiable, continue mais non différentiable, ou carrément discontinue. Dans ce dernier cas on aura affaire à une variable régionalisée donnant lieu à un effet de pépité, symptôme

de valeurs changeant abruptement à très petite échelle par rapport à la maille, comme les teneurs de l'or lorsqu'il y a des pépites.

Lorsque la dissemblance moyenne des valeurs est constante pour toutes les distances $|\mathbf{h}|$, il y a une absence complète de structuration spatiale des valeurs. A l'inverse, une pente non nulle du variogramme près de l'origine indique une structuration des données. Un changement soudain de la pente du variogramme indique le passage à une structuration des valeurs de nature différente.

Ces transitions seront d'abord modélisées par des variogrammes gigognes, puis les différents types de structuration spatiale des valeurs pourront être visualisés séparément sous forme de cartes en effectuant le krigeage de composantes spatiales.

Variable régionalisée et fonction aléatoire

La variable régionalisée $z(\mathbf{x})$ est considérée comme une réalisation d'une fonction aléatoire $Z(\mathbf{x})$ (la famille infinie des variables aléatoires implantées en tout point \mathbf{x} d'un domaine spatial \mathcal{D}). Les données $z(\mathbf{x}_\alpha)$ sont des informations fragmentaires que l'on possède sur une réalisation particulière $z(\mathbf{x})$ de $Z(\mathbf{x})$.

Cette approche est radicalement différente de celle de la statistique, où les échantillons z_α sont des réalisations d'une variable aléatoire Z .

Les implications épistémologiques de l'étude d'un phénomène unique par des méthodes probabilistes sont débattues dans MATHERON [66].

Variogramme théorique

Le variogramme théorique $\gamma(\mathbf{h})$ est défini par l'hypothèse intrinsèque. L'hypothèse intrinsèque est formée de deux conditions sur les accroissements $Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x})$ de la fonction aléatoire:

- La moyenne des accroissements est invariante pour toute translation du vecteur \mathbf{h} dans le domaine. Plus spécifiquement, la moyenne des accroissements est supposée nulle, quelle que soit la position de \mathbf{h} dans le domaine.
- La variance des accroissements admet une valeur finie en fonction de \mathbf{h} et indépendante de la position de \mathbf{h} dans le domaine.

C'est-à-dire,

$$\begin{cases} \mathbb{E}[Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x})] = 0 & \text{pour tout } \mathbf{x}, \mathbf{x}+\mathbf{h} \in \mathcal{D} \\ \text{var}[Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x})] = 2\gamma(\mathbf{h}) & \text{pour tout } \mathbf{x}, \mathbf{x}+\mathbf{h} \in \mathcal{D} \end{cases}$$

ce qui donne le variogramme théorique

$$\gamma(\mathbf{h}) = \frac{1}{2} \mathbb{E} [(Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x}))^2] \quad \text{pour tout } \mathbf{x}, \mathbf{x}+\mathbf{h} \in \mathcal{D}$$

L'existence de l'espérance des accroissements d'une fonction aléatoire intrinsèque n'implique pas celle de l'espérance de la fonction aléatoire. Une fonction aléatoire intrinsèque peut avoir une variance infinie, tout en ayant une variance des accroissements finie pour tout vecteur \mathbf{h} .

Fonction de covariance

La fonction de covariance $C(\mathbf{h})$ est définie sur la base de l'hypothèse de stationnarité des deux premiers moments, moyenne et covariance, de la fonction aléatoire

$$\begin{cases} \mathbb{E} [Z(\mathbf{x})] = m & \text{pour tout } \mathbf{x} \in \mathcal{D} \\ \mathbb{E} [Z(\mathbf{x}) \cdot Z(\mathbf{x}+\mathbf{h})] - m^2 = C(\mathbf{h}) & \text{pour tout } \mathbf{x} \in \mathcal{D} \end{cases}$$

La fonction de covariance est bornée et n'excède pas la variance

$$|C(\mathbf{h})| \leq C(\mathbf{0}) = \text{var}(Z(\mathbf{x}))$$

C'est une fonction paire: $C(-\mathbf{h}) = C(+\mathbf{h})$.

On peut déduire un variogramme d'une fonction de covariance par la formule

$$\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h})$$

Etant donné que l'hypothèse de stationnarité d'ordre deux est moins générale que l'hypothèse intrinsèque, tout modèle de variogramme ne peut pas être déduit d'un modèle de covariance.

EXEMPLE 5.1 *Par exemple, le variogramme puissance*

$$\gamma_\theta(\mathbf{h}) = b |\mathbf{h}|^\theta \quad \text{avec } 0 < \theta < 2$$

déborde manifestement du cadre fixé par l'hypothèse de stationnarité d'ordre deux.

Fonction de type positif

Du point de vue algébrique, l'emploi d'une fonction de covariance $C(\mathbf{h})$ dans le calcul de la variance d'une combinaison linéaire de $n+1$ variables aléatoires $Z(\mathbf{x}_\alpha)$ (sous-ensemble d'une fonction aléatoire stationnaire d'ordre

deux) avec $n+1$ pondérateurs λ_α doit nécessairement conduire à une matrice semi-définie positive

$$\text{var} \left(\sum_{\alpha=0}^n \lambda_\alpha Z(\mathbf{x}_\alpha) \right) = \sum_{\alpha=0}^n \sum_{\beta=0}^n \lambda_\alpha \lambda_\beta C(\mathbf{x}_\alpha - \mathbf{x}_\beta) = \mathbf{I}^\top \mathbf{C} \mathbf{I} \geq 0$$

quels que soient les points \mathbf{x}_α et les pondérateurs λ_α . Une fonction de covariance est donc une fonction de type positif.

Une fonction de covariance continue est définie par le théorème de Bochner comme étant la transformée de Fourier d'une mesure positive.

Fonction de type négatif conditionnel

Le variogramme, quant-à-lui, doit être une fonction de type négatif conditionnel pour garantir la positivité de la variance d'une combinaison linéaire de $n+1$ variables aléatoires, sous-ensemble d'une fonction aléatoire intrinsèque, avec $n+1$ pondérateurs de somme nulle

$$\text{var} \left(\sum_{\alpha=0}^n \lambda_\alpha Z(\mathbf{x}_\alpha) \right) = - \sum_{\alpha=0}^n \sum_{\beta=0}^n \lambda_\alpha \lambda_\beta \gamma(\mathbf{x}_\alpha - \mathbf{x}_\beta) \geq 0 \quad \text{si} \quad \sum_{\alpha=0}^n \lambda_\alpha = 0$$

c'est-à-dire

$$[\lambda_\alpha]^\top [\gamma(\mathbf{x}_\alpha - \mathbf{x}_\beta)] [\lambda_\alpha] = \mathbf{I}^\top \mathbf{\Gamma} \mathbf{I} \leq 0 \quad \text{pour} \quad \sum_{\alpha=0}^n \lambda_\alpha = 0$$

Pour comprendre l'utilité (le caractère suffisant) de la condition sur la somme des pondérateurs, il faut se livrer à un calcul explicite de la variance de la combinaison linéaire.

D'abord, la fonction aléatoire $Z(\mathbf{x})$ étant intrinsèque, seule l'espérance d'accroissements a un sens. Pour pouvoir former des accroissements, nous introduisons dans l'expression de la variance une variable aléatoire supplémentaire $Z(\mathbf{0})$, placée arbitrairement à l'origine $\mathbf{0}$ (on suppose que $\mathbf{0} \in \mathcal{D}$), multipliée par les pondérateurs de somme nulle, ce qui ne modifie pas la variance

$$\begin{aligned} \text{var} \left(\sum_{\alpha=0}^n \lambda_\alpha Z(\mathbf{x}_\alpha) \right) &= \text{var} \left(-Z(\mathbf{0}) \cdot \underbrace{\sum_{\alpha=0}^n \lambda_\alpha}_0 + \sum_{\alpha=0}^n \lambda_\alpha Z(\mathbf{x}_\alpha) \right) \\ &= \text{var} \left[\sum_{\alpha=0}^n \lambda_\alpha \left(Z(\mathbf{x}_\alpha) - Z(\mathbf{0}) \right) \right] \\ &= \sum_{\alpha=0}^n \sum_{\beta=0}^n \lambda_\alpha \lambda_\beta \text{E} \left[\left(Z(\mathbf{x}_\alpha) - Z(\mathbf{0}) \right) \cdot \left(Z(\mathbf{x}_\beta) - Z(\mathbf{0}) \right) \right] \end{aligned}$$

$$= \sum_{\alpha=0}^n \sum_{\beta=0}^n \lambda_{\alpha} \lambda_{\beta} C_I(\mathbf{x}_{\alpha}, \mathbf{x}_{\beta})$$

où $C_I(\mathbf{x}_{\alpha}, \mathbf{x}_{\beta})$ est la covariance des accroissements avec la variable supplémentaire.

Dans un calcul intermédiaire, on introduit le point supplémentaire $Z(\mathbf{0})$ dans l'expression du variogramme

$$\begin{aligned} \gamma(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}) &= \frac{1}{2} \text{E} \left[\left(Z(\mathbf{x}_{\alpha}) - Z(\mathbf{0}) - Z(\mathbf{x}_{\beta}) + Z(\mathbf{0}) \right)^2 \right] \\ &= \frac{1}{2} \left(2\gamma(\mathbf{x}_{\alpha}) + 2\gamma(\mathbf{x}_{\beta}) - 2C_I(\mathbf{x}_{\alpha}, \mathbf{x}_{\beta}) \right) \end{aligned}$$

de sorte que

$$C_I(\mathbf{x}_{\alpha}, \mathbf{x}_{\beta}) = \gamma(\mathbf{x}_{\alpha}) + \gamma(\mathbf{x}_{\beta}) - \gamma(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta})$$

avec deux termes non stationnaires $\gamma(\mathbf{x}_{\alpha})$ et $\gamma(\mathbf{x}_{\beta})$.

En reprenant le calcul de la variance de la combinaison linéaire, on constate que la condition sur la somme des pondérateurs annule les termes non stationnaires

$$\begin{aligned} \text{var} \left(\sum_{\alpha=0}^n \lambda_{\alpha} Z(\mathbf{x}_{\alpha}) \right) &= \sum_{\alpha=0}^n \sum_{\beta=0}^n \lambda_{\alpha} \lambda_{\beta} C_I(\mathbf{x}_{\alpha}, \mathbf{x}_{\beta}) \\ &= \underbrace{\sum_{\beta=0}^n \lambda_{\beta}}_0 \sum_{\alpha=0}^n \lambda_{\alpha} \gamma(\mathbf{x}_{\alpha}) + \underbrace{\sum_{\alpha=0}^n \lambda_{\alpha}}_0 \sum_{\beta=0}^n \lambda_{\beta} \gamma(\mathbf{x}_{\beta}) - \sum_{\alpha=0}^n \sum_{\beta=0}^n \lambda_{\alpha} \lambda_{\beta} \gamma(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}) \end{aligned}$$

Ajustement du variogramme par une fonction de covariance

Quand le variogramme expérimental présente un palier aux grandes distances $|\mathbf{h}|$, on supposera souvent que la fonction aléatoire $Z(\mathbf{x})$ a une variance finie et l'on ajustera le variogramme par une fonction de covariance $C(\mathbf{h})$ en se basant sur la formule

$$\gamma(\mathbf{h}) = b - C(\mathbf{h})$$

où b caractérise un palier de variation.

Dans la suite, le modèle $C(\mathbf{h})$ sera soit un effet de pépité

$$C_{\text{pep}}(\mathbf{h}) = \begin{cases} b & \text{lorsque } |\mathbf{h}| = 0 \\ 0 & \text{lorsque } |\mathbf{h}| > 0 \end{cases}$$

soit un modèle sphérique

$$C_{sph}(\mathbf{h}) = \begin{cases} b \left(1 - \frac{3|\mathbf{h}|}{2a} + \frac{1}{2} \frac{|\mathbf{h}|^3}{a^3} \right) & \text{pour } 0 \leq |\mathbf{h}| \leq a \\ 0 & \text{pour } |\mathbf{h}| > a \end{cases}$$

Le paramètre a indique la portée de la covariance: la covariance s'éteint lorsque la portée est atteinte. Le paramètre b indique la valeur maximale de la covariance: la covariance sphérique diminue de manière continue à partir de la valeur maximale b , située à l'origine, jusqu'à disparaître complètement lorsque la portée est atteinte. Le modèle de l'effet de pépite pourrait être considéré comme un cas particulier d'un sphérique avec une portée a infiniment petite. Il y a cependant une différence notable entre les deux modèles: $C_{pep}(\mathbf{h})$ décrit un phénomène discontinu, pour lequel les valeurs changent brutalement d'un point à un autre, tandis que $C_{sph}(\mathbf{h})$ décrit un phénomène continu, mais non différentiable, qui serait en quelque sorte rugueux au toucher.

Variogramme gigogne

On peut souvent distinguer plusieurs paliers de variation qui sont en rapport étroit avec la morphologie de la variable régionalisée¹.

Définissons d'abord une fonction de corrélation $\rho(\mathbf{h})$, obtenue en normant la fonction de covariance à un palier unitaire

$$\rho(\mathbf{h}) = \frac{C(\mathbf{h})}{b} \quad \text{c'est-à-dire} \quad C(\mathbf{h}) = b \rho(\mathbf{h})$$

On numérotera les paliers de variation b_u observés sur le variogramme avec un index $u = 0, \dots, S$. On peut ainsi construire un variogramme gigogne

$$\gamma(\mathbf{h}) = \sum_{u=0}^S \gamma_u(\mathbf{h}) = \sum_{u=0}^S b_u g_u(\mathbf{h})$$

où les $g_u(\mathbf{h})$ sont des variogrammes normés.

Si les variogrammes $\gamma_u(\mathbf{h})$ peuvent être déduits de fonctions de covariance $C_u(\mathbf{h})$, le modèle gigogne devient

$$\gamma(\mathbf{h}) = b - \sum_{u=0}^S C_u(\mathbf{h}) = b - \sum_{u=0}^S b_u \rho_u(\mathbf{h}) \quad \text{où} \quad \sum_{u=0}^S b_u = b$$

avec, typiquement, une covariance pépitique $C_0(\mathbf{h})$ et plusieurs covariances sphériques $C_u(\mathbf{h})$ de portées a_u différentes.

EXEMPLE 5.2 *La Figure 5.1 montre le variogramme du deuxième canal d'une*

1. SERRA [90] a mené une étude remarquable du gisement de fer Lorrain, dans laquelle il a pu distinguer jusqu'à sept paliers de variation ayant une signification géologique, dans les multiples transitions entre l'échelle micrométrique et l'échelle kilométrique.

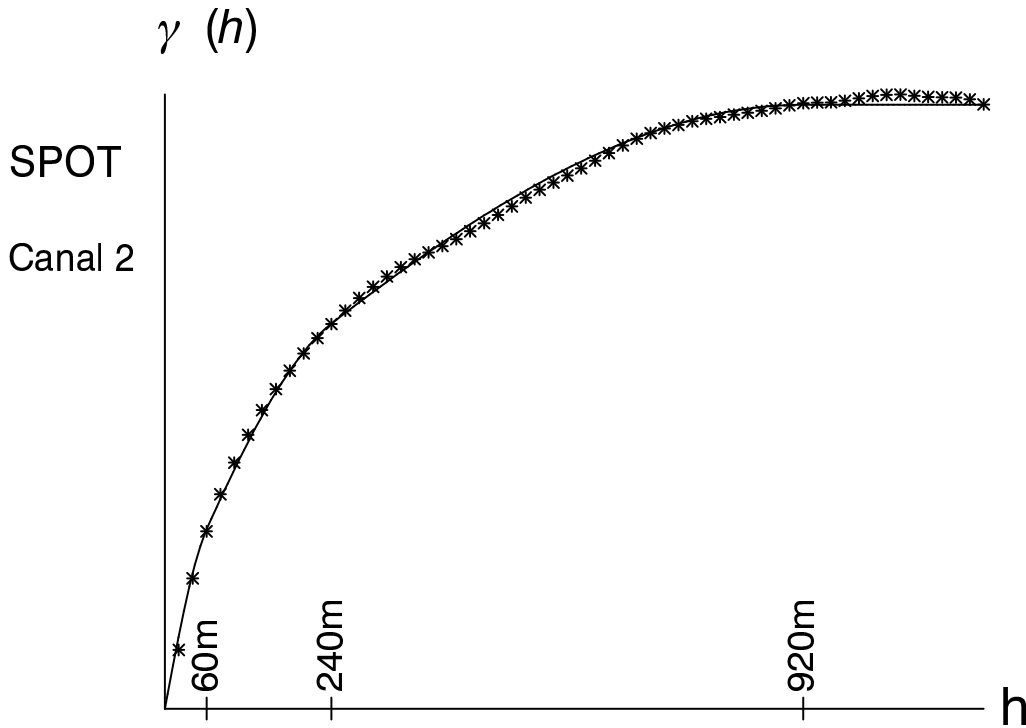


FIG. 5.1 – Variogramme du deuxième canal d'une image SPOT.

image SPOT sur une zone de $5 \times 4 \text{ km}^2$, ajusté avec un modèle gigogne composé de trois structures sphériques de portées 60m, 240m et 920m (d'après [50]).

Décomposition de la fonction aléatoire

En association avec le variogramme gigogne, la fonction aléatoire $Z(x)$ est décomposée en une somme de composantes spatiales indépendantes agissant à différentes échelles spatiales, i.e. atteignant différents paliers de variation b_u à des échelles différentes, mis à part éventuellement un coefficient b_S , qui ne représente pas forcément un palier.

Il est clair qu'une composante à petite échelle ne pourra être identifiée que si la maille d'échantillonnage est suffisamment fine. De même, une composante à grande échelle n'apparaîtra sur le variogramme que si le diamètre du domaine échantillonné est assez important.

Une composante déterministe (moyenne m ou dérive $m(x)$) pourra, le cas échéant, également faire partie de la décomposition.

Nous présentons trois modèles linéaires de régionalisation particuliers, obtenus en mélangeant des composantes spatiales de types différents.

Régionalisation stationnaire d'ordre deux

Une fonction aléatoire stationnaire d'ordre deux $Z(\mathbf{x})$ peut être constituée à partir d'une somme de fonctions aléatoires d'ordre deux $Z_u(\mathbf{x})$, non corrélées et de moyennes nulles, plus une constante m représentant l'espérance de $Z(\mathbf{x})$

$$Z(\mathbf{x}) = Z_0(\mathbf{x}) + \dots + Z_u(\mathbf{x}) + \dots + Z_S(\mathbf{x}) + m$$

Un calcul simple montre que le modèle de covariance correspondant est

$$C(\mathbf{h}) = C_0(\mathbf{h}) + \dots + C_u(\mathbf{h}) + \dots + C_S(\mathbf{h})$$

Régionalisation avec une composante intrinsèque

Un modèle de régionalisation intrinsèque particulier, qui nous sera utile plus loin, peut être construit en additionnant une fonction aléatoire intrinsèque $Z_S(\mathbf{x})$ et S fonctions aléatoires stationnaires d'ordre deux $Z_u(\mathbf{x})$ ($u = 0, \dots, S - 1$), de moyennes nulles, non corrélées entre elles et non corrélées avec les accroissements de $Z_S(\mathbf{x})$

$$Z(\mathbf{x}) = Z_0(\mathbf{x}) + \dots + Z_u(\mathbf{x}) + \dots + Z_S(\mathbf{x})$$

Le modèle de variogramme associé est un modèle gigogne composé de S structures déduites de fonctions de covariance et d'une structure purement intrinsèque.

Régionalisation localement stationnaire

Une fonction aléatoire non stationnaire $Z(\mathbf{x})$ peut être obtenue en superposant $S+1$ fonctions aléatoires d'ordre deux non corrélées de moyennes nulles $Z_u(\mathbf{x})$ et une dérive $m(\mathbf{x})$ (l'espérance de $Z(\mathbf{x})$ au point \mathbf{x})

$$Z(\mathbf{x}) = Z_0(\mathbf{x}) + \dots + Z_u(\mathbf{x}) + \dots + Z_S(\mathbf{x}) + m(\mathbf{x})$$

En choisissant une fonction $m_l(\mathbf{x})$ très continue, qui ne varie que très lentement d'un bout à l'autre du domaine, on peut lui donner le sens d'une moyenne locale et qualifier

$$Z(\mathbf{x}) = Z_0(\mathbf{x}) + \dots + Z_u(\mathbf{x}) + \dots + Z_S(\mathbf{x}) + m_l(\mathbf{x})$$

de localement stationnaire, l'espérance de $Z(\mathbf{x})$ étant approximativement constante dans tout voisinage local du domaine.

6 Krigeage de Composantes Spatiales

Krigeage ordinaire

Le krigeage ordinaire est une opération qui est répétée en chaque noeud \mathbf{x}_0 d'une grille régulière recouvrant le domaine étudié. Pour un ensemble de n points de données \mathbf{x}_α d'un voisinage centré autour d'un point \mathbf{x}_0 de la maille d'estimation, on peut construire, en minimisant la variance d'estimation, le système de krigeage ordinaire suivant

$$\begin{pmatrix} C(\mathbf{x}_1 - \mathbf{x}_1) & \dots & C(\mathbf{x}_1 - \mathbf{x}_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C(\mathbf{x}_n - \mathbf{x}_1) & \dots & C(\mathbf{x}_n - \mathbf{x}_n) & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ -\mu \end{pmatrix} = \begin{pmatrix} C(\mathbf{x}_1 - \mathbf{x}_0) \\ \vdots \\ C(\mathbf{x}_n - \mathbf{x}_0) \\ 1 \end{pmatrix}$$

où les λ_α sont des pondérateurs à affecter aux points de données et où μ est un paramètre de Lagrange qui intervient pour des raisons algébriques. Le membre gauche du système contient les covariances entre les points de données, tandis que le membre droit contient les covariances entre chaque point de donnée et le point d'estimation \mathbf{x}_0 .

Ce système, une fois résolu, permet de transférer au point \mathbf{x}_0 de l'information en provenance des points de données voisins, par le calcul d'une moyenne pondérée

$$z^*(\mathbf{x}_0) = \sum_{\alpha=1}^n \lambda_\alpha z(\mathbf{x}_\alpha)$$

En effectuant la multiplication matricielle, on peut réécrire le système sous la forme

$$\begin{cases} \sum_{\beta=1}^n \lambda_\beta C(\mathbf{x}_\alpha - \mathbf{x}_\beta) - \mu = C(\mathbf{x}_\alpha - \mathbf{x}_0) & \text{pour } \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_\beta = 1 \end{cases}$$

La condition que les poids soient de somme unité exprime le fait que dans le cas extrême où les données sont égales à une constante dans un voisinage donné, la valeur estimée reproduira cette constante.

Krigage de la moyenne locale

Dans le cadre d'un modèle de fonction aléatoire localement stationnaire, on est intéressé par une estimation de la moyenne locale $m_l(\mathbf{x}_0)$.

L'estimateur basé sur une combinaison linéaire de variables aléatoires Z_α implantées en n points de données \mathbf{x}_α à proximité d'un point arbitraire \mathbf{x}_0

$$m_l^*(\mathbf{x}_0) = \sum_{\alpha=1}^n \lambda_m^\alpha Z(\mathbf{x}_\alpha)$$

L'estimateur doit être non biaisé au sens que l'on veut que l'erreur d'estimation, c'est-à-dire la différence entre la valeur estimée $m_l^*(\mathbf{x}_0)$ et la valeur effective $m_l(\mathbf{x}_0)$, soit nulle en moyenne. L'espérance de l'erreur est égale à

$$\mathbb{E} \left[m_l^*(\mathbf{x}_0) - m_l(\mathbf{x}_0) \right] = \mathbb{E} \left[\sum_{\alpha=1}^n \lambda_m^\alpha Z(\mathbf{x}_\alpha) - m_l(\mathbf{x}_0) \right] = m_l(\mathbf{x}_0) \sum_{\alpha=1}^n \lambda_m^\alpha - m_l(\mathbf{x}_0)$$

en identifiant $m_l(\mathbf{x}_\alpha)$ à $m_l(\mathbf{x}_0)$ sur la base de l'hypothèse de stationnarité locale. L'espérance de l'erreur s'annule avec des pondérateurs de somme unité, ce qui donne la condition de non biais

$$\sum_{\alpha=1}^n \lambda_m^\alpha = 1$$

La variance σ_E^2 de l'erreur d'estimation (variance d'estimation) vaut, en utilisant la condition de non biais

$$\begin{aligned} \sigma_E^2 &= \text{var} \left(m_l^*(\mathbf{x}_0) - m_l(\mathbf{x}_0) \right) \\ &= \mathbb{E} \left[\left(\sum_{\alpha=1}^n \lambda_m^\alpha Z(\mathbf{x}_\alpha) - m_l(\mathbf{x}_0) \right)^2 \right] \\ &= \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_m^\alpha \lambda_m^\beta C(\mathbf{x}_\alpha - \mathbf{x}_\beta) \end{aligned}$$

Dans le but d'optimiser la variance d'estimation, on construit une fonction objectif ϕ en introduisant la contrainte sur les pondérateurs avec un paramètre de Lagrange μ_m

$$\phi = \sigma_E^2 - 2 \mu_m \left(\sum_{\alpha=1}^n \lambda_m^\alpha - 1 \right)$$

En annulant les dérivées partielles selon λ_m^α

$$\frac{\partial \phi}{\partial \lambda_m^\alpha} = 0 \quad \iff \quad 2 \sum_{\beta=1}^n \lambda_m^\beta C(\mathbf{x}_\alpha - \mathbf{x}_\beta) - 2 \mu_m = 0$$

et selon μ_m

$$\frac{\partial \phi}{\partial \mu_m} = 0 \quad \Longleftrightarrow \quad 2 \left(\sum_{\beta=1}^n \lambda_m^\beta - 1 \right) = 0$$

on obtient le système de krigeage de la moyenne locale

$$\begin{cases} \sum_{\beta=1}^n \lambda_m^\beta C(\mathbf{x}_\alpha - \mathbf{x}_\beta) - \mu_m = 0 & \text{pour } \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_m^\beta = 1 \end{cases}$$

La variance de krigeage σ_K^2 , c'est à dire la variance d'estimation à l'optimum, est égale à

$$\sigma_K^2 = \sum_{\alpha=1}^n \lambda_m^\alpha \underbrace{\sum_{\beta=1}^n \lambda_m^\beta C(\mathbf{x}_\alpha - \mathbf{x}_\beta)}_{\mu_m} = \mu_m$$

On notera bien la différence entre un krigeage de la moyenne locale et le calcul de la moyenne arithmétique des données dans le voisinage de krigeage: cette dernière ne tient pas compte de la disposition spatiale des échantillons.

Krigeage de la composante intrinsèque

Si $Z(\mathbf{x})$ est définie dans le cadre du modèle de régionalisation intrinsèque décrit plus haut, on peut chercher à estimer la composante intrinsèque $Z_S(\mathbf{x})$ à partir de données dans le voisinage en formant l'estimateur

$$Z_S^*(\mathbf{x}_0) = \sum_{\alpha=1}^n \lambda_S^\alpha Z(\mathbf{x}_\alpha)$$

On obtient le non biais de l'erreur d'estimation en utilisant des pondérateurs de somme unité

$$\begin{aligned} \mathbb{E} \left[Z_S^*(\mathbf{x}_0) - Z_S(\mathbf{x}_0) \right] &= \mathbb{E} \left[\sum_{\alpha=1}^n \lambda_S^\alpha Z(\mathbf{x}_\alpha) - Z_S(\mathbf{x}_0) \cdot \underbrace{\sum_{\alpha=1}^n \lambda_S^\alpha}_1 \right] \\ &= \sum_{\alpha=1}^n \lambda_S^\alpha \mathbb{E} \left[Z(\mathbf{x}_\alpha) - Z_S(\mathbf{x}_0) \right] \\ &= \sum_{\alpha=1}^n \lambda_S^\alpha \left(\underbrace{\sum_{u=0}^{S-1} \mathbb{E} \left[Z_u(\mathbf{x}_\alpha) \right]}_0 + \underbrace{\mathbb{E} \left[Z_S(\mathbf{x}_\alpha) - Z_S(\mathbf{x}_0) \right]}_0 \right) \end{aligned}$$

$$= 0$$

La variance d'estimation σ_E^2 s'écrit

$$\begin{aligned}\sigma_E^2 &= \text{var} (Z_S^*(\mathbf{x}_0) - Z_S(\mathbf{x}_0)) \\ &= \text{E} \left[\left(\sum_{u=0}^{S-1} \sum_{\alpha=1}^n \lambda_S^\alpha Z_u(\mathbf{x}_\alpha) + \sum_{\alpha=1}^n \lambda_S^\alpha (Z_S(\mathbf{x}_\alpha) - Z_S(\mathbf{x}_0)) \right)^2 \right]\end{aligned}$$

En tenant compte de la non corrélation

$$\begin{aligned}\sigma_E^2 &= \sum_{u=0}^{S-1} \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_S^\alpha \lambda_S^\beta C^u(\mathbf{x}_\alpha - \mathbf{x}_\beta) \\ &\quad - \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_S^\alpha \lambda_S^\beta \gamma^S(\mathbf{x}_\alpha - \mathbf{x}_\beta) - \gamma^S(\mathbf{x}_0 - \mathbf{x}_0) + 2 \sum_{\alpha=1}^n \lambda_S^\alpha \gamma^S(\mathbf{x}_\alpha - \mathbf{x}_0) \\ &= \sum_{u=0}^{S-1} C^u(\mathbf{x}_0 - \mathbf{x}_0) \\ &\quad - \gamma^S(\mathbf{x}_0 - \mathbf{x}_0) - \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_S^\alpha \lambda_S^\beta \gamma(\mathbf{x}_\alpha - \mathbf{x}_\beta) + 2 \sum_{\alpha=1}^n \lambda_S^\alpha \gamma^S(\mathbf{x}_\alpha - \mathbf{x}_0)\end{aligned}$$

En optimisant sous contrainte, on a le système

$$\begin{cases} \sum_{\beta=1}^n \lambda_S^\beta \gamma(\mathbf{x}_\alpha - \mathbf{x}_\beta) + \mu_S = \gamma^S(\mathbf{x}_\alpha - \mathbf{x}_0) & \text{pour } \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_S^\beta = 1 \end{cases}$$

Ce système diffère de celui du krigeage ordinaire par l'apparition dans le membre droit de termes $\gamma^S(\mathbf{x}_\alpha - \mathbf{x}_0)$ spécifiques à la composante $Z_S(\mathbf{x})$.

Krigeage d'une composante stationnaire

Pour kriger une composante stationnaire d'ordre deux particulière Z_{u_0} , dans le modèle de régionalisation intrinsèque, on part de la combinaison linéaire

$$Z_{u_0}^*(\mathbf{x}_0) = \sum_{\alpha=1}^n \lambda_{u_0}^\alpha Z(\mathbf{x}_\alpha)$$

Le non biais est obtenu sur la base d'une somme des pondérateurs nulle

$$\text{E} \left[Z_{u_0}^*(\mathbf{x}_0) - Z_{u_0}(\mathbf{x}_0) \right] = \text{E} \left[\sum_{u=0}^S \sum_{\alpha=1}^n \lambda_{u_0}^\alpha Z_u(\mathbf{x}_\alpha) - Z_{u_0}(\mathbf{x}_0) \right]$$

on introduit un point fictif pour former des accroissements

$$\begin{aligned}
&= \mathbb{E} \left[\sum_{u=0}^{S-1} \sum_{\alpha=1}^n \lambda_{u_0}^\alpha Z_u(\mathbf{x}_\alpha) - Z_{u_0}(\mathbf{x}_0) \right. \\
&\quad \left. + \sum_{\alpha=1}^n \lambda_{u_0}^\alpha Z_S(\mathbf{x}_\alpha) - Z_S(\mathbf{0}) \cdot \underbrace{\sum_{\alpha=1}^n \lambda_{u_0}^\alpha}_0 \right] \\
&= \sum_{u=0}^{S-1} \sum_{\alpha=1}^n \lambda_{u_0}^\alpha \underbrace{\mathbb{E} [Z_u(\mathbf{x}_\alpha)]}_0 - \underbrace{\mathbb{E} [Z_{u_0}(\mathbf{x}_0)]}_0 \\
&\quad + \sum_{\alpha=1}^n \lambda_{u_0}^\alpha \underbrace{\mathbb{E} [Z_S(\mathbf{x}_\alpha) - Z_S(\mathbf{0})]}_0 \\
&= 0
\end{aligned}$$

La variance d'estimation est

$$\begin{aligned}
\sigma_E^2 &= \text{var} \left(Z_{u_0}^*(\mathbf{x}_0) - Z_{u_0}(\mathbf{x}_0) \right) \\
&= C^{u_0}(\mathbf{x}_0 - \mathbf{x}_0) - \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{u_0}^\alpha \lambda_{u_0}^\beta \gamma(\mathbf{x}_\alpha - \mathbf{x}_\beta) + 2 \sum_{\alpha=1}^n \lambda_{u_0}^\alpha \gamma^{u_0}(\mathbf{x}_\alpha - \mathbf{x}_0)
\end{aligned}$$

Le système de krigeage de la composante $Z_{u_0}(\mathbf{x}_0)$ est alors

$$\begin{cases} \sum_{\beta=1}^n \lambda_{u_0}^\beta \gamma(\mathbf{x}_\alpha - \mathbf{x}_\beta) + \mu_{u_0} = \gamma^{u_0}(\mathbf{x}_\alpha - \mathbf{x}_0) & \text{pour } \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_{u_0}^\beta = 0 \end{cases}$$

On est ainsi capable d'extraire des données des composantes de variation spatiale, en partant d'une interprétation du variogramme expérimental par un variogramme gigogne.

Ces techniques sont proches des méthodes d'analyse spectrale, très populaires en prospection géophysique.

REMARQUE 6.1 (FILTRAGE) Soit $Z(\mathbf{x})$ une fonction aléatoire définie dans un modèle de régionalisation localement stationnaire. On filtre une composante spatiale en retirant les covariances correspondantes $C_u(\mathbf{x}_\alpha - \mathbf{x}_0)$ du membre droit de l'équation du krigeage ordinaire. Ainsi peut-on par exemple éliminer l'effet de pépite en retirant les termes correspondant à $C_0(\mathbf{h})$ du membre droit et l'on obtient alors une estimation amputée de la composante pépitique $z_0(\mathbf{x})$.

REMARQUE 6.2 (MAILLE D'ÉCHANTILLONNAGE IRRÉGULIÈRE ET KO) Prenons une variable régionalisée pour laquelle on aurait identifié trois composantes de variation non corrélées, que l'on interprète comme des réalisations de fonctions aléatoires stationnaires d'ordre deux, plus une moyenne locale

$$z(\mathbf{x}) = z_0(\mathbf{x}) + z_1(\mathbf{x}) + z_2(\mathbf{x}) + m_l(\mathbf{x})$$

A cette situation correspond le modèle de covariance

$$C(\mathbf{h}) = C_0(\mathbf{h}) + C_1(\mathbf{h}) + C_2(\mathbf{h})$$

avec $C_0(\mathbf{h})$ un effet de pépite et $C_1(\mathbf{h}), C_2(\mathbf{h})$ des modèles sphériques de portées a_1 et a_2 , numérotés de sorte que $a_1 < a_2$.

Supposons que la maille d'échantillonnage soit très irrégulière, entraînant une répartition des données extrêmement inégale dans l'espace. Afin d'établir une carte, on va construire un réseau de points, aussi fin que l'on voudra, et l'on exécutera en chaque point l'opération de krigeage ordinaire.

Vu la grande irrégularité dans la disposition des échantillons, différents cas de figure peuvent se présenter, que nous allons examiner séparément dans la suite.

1. \mathbf{x}_0 est éloigné $> a_2$ des données

Ce cas de figure peut se produire si \mathbf{x}_0 se trouve en plein milieu d'un grand trou dans la maille d'échantillonnage. Le krigeage ordinaire est alors équivalent à un krigeage de la moyenne.

Le membre droit est nul puisque toutes les composantes s'annulent pour un écartement supérieur à a_2 . L'information transférée au point \mathbf{x}_0 sera une estimation de la moyenne locale dans le voisinage.

2. \mathbf{x}_0 est écarté $< a_2$ et $> a_1$ du point \mathbf{x}_α le plus proche

Avec cette disposition spatiale le système de krigeage ordinaire est équivalent à un filtrage des composantes $z_0(\mathbf{x})$ et $z_1(\mathbf{x})$,

$$\begin{cases} \sum_{\beta=1}^n \lambda_\beta C(\mathbf{x}_\alpha - \mathbf{x}_\beta) - \mu = \boxed{C_2(\mathbf{x}_\alpha - \mathbf{x}_0)} & \text{pour } \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda^\beta = 1 \end{cases}$$

3. x_0 est situé $< a_1$ et > 0 du point x_α le plus proche

Dans cette situation le krigeage ordinaire transférera de l'information non seulement sur la composante de longue portée $z_2(\mathbf{x})$, mais aussi sur la composante $z_1(\mathbf{x})$, qui varie plus rapidement dans l'espace et qui suscitera une description plus détaillée de la variable régionalisée dans les zones de la carte où les données abondent. On devine que le krigeage ordinaire est maintenant équivalent au filtrage,

$$\left\{ \begin{array}{l} \sum_{\beta=1}^n \lambda_\beta C(\mathbf{x}_\alpha - \mathbf{x}_\beta) - \mu = \boxed{C_1(\mathbf{x}_\alpha - \mathbf{x}_0) + C_2(\mathbf{x}_\alpha - \mathbf{x}_0)} \quad \text{pour } \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_\beta = 1 \end{array} \right.$$

4. x_0 coïncide avec un point de donnée

Dans ce dernier cas, le krigeage ordinaire ne filtrera rien du tout et restituera fidèlement la donnée mesurée au point $\mathbf{x}_\alpha = \mathbf{x}_0$: c'est un interpolateur exact!

Cette fidélité peut cependant être indésirable lorsqu'il n'y a que quelques points de coïncidence, purement fortuits, entre une maille d'échantillonnage irrégulière et une maille d'estimation régulière. Pour éviter des sauts abrupts des valeurs estimées en ces quelques points, certains programmes de krigeage à vocation cartographique (tel Bluepack) proposeront de filtrer systématiquement la composante pépitique $z_0(\mathbf{x})$.

EXERCISE 6.3 On dit qu'il y a interpolation exacte, si l'estimation en un point de donnée restitue la valeur de la donnée. Montrer que le krigeage ordinaire (ponctuel) est un interpolateur exact.

EXERCISE 6.4 $Z(\mathbf{x})$ est une FA stationnaire d'ordre 2 d'espérance nulle, scindée en deux composantes non corrélées d'espérance nulle, selon le modèle de régionalisation

$$Z(\mathbf{x}) = Y^P(\mathbf{x}) + Y^G(\mathbf{x})$$

où $Y^P(\mathbf{x})$ est la composante de petite portée de covariance $C^P(\mathbf{h})$, et où $Y^G(\mathbf{x})$ est la composante de grande portée de covariance $C^G(\mathbf{h})$, avec

$$C(\mathbf{h}) = C^P(\mathbf{h}) + C^G(\mathbf{h})$$

Les données sont disposées sur une grille régulière, et l'on fait du krigeage simple (sans condition d'universalité) aux noeuds de cette même grille pour

estimer les composantes de courte et de longue portée, en utilisant un voisinage constitué de toutes les données.

Quelle relation peut-on établir entre les pondérateurs λ_α^P et λ_α^G des deux composantes en chaque point du voisinage de krigeage?

EXERCISE 6.5 $Z(\mathbf{x})$ est une FA localement stationnaire d'ordre 2, constituée de composantes $Y^P(\mathbf{x})$ et $Y^G(\mathbf{x})$ non corrélées d'espérances nulles, ainsi que d'une dérive assimilable à une constante dans tout voisinage local du domaine.

Montrer que la somme des krigeages des composantes et du krigeage de la moyenne est égale au krigeage ordinaire de $Z(\mathbf{x})$

$$Y_P^*(\mathbf{x}_0) + Y_G^*(\mathbf{x}_0) + m_0^* = Z^*(\mathbf{x}_0)$$

7 Dérive externe

Deux variables mesurant la même chose

Il arrive que deux variables mesurées par des procédés différents traduisent le même phénomène et que la première variable est précise, mais connue en peu de points seulement, tandis que la seconde variable est connue à peu près partout dans le domaine spatial, mais avec une précision plus approximative, donnant uniquement l'allure générale de la morphologie du phénomène régionalisé.

L'exemple classique provient de l'exploration pétrolière, où l'on cherche à cartographier un niveau géologique, qui est en général continu (mis à part des failles), parce que le pétrole se trouve dans des formations sédimentaires. On dispose de deux sources d'information:

- de mesures précises de la profondeur d'une couche faites sur des forages pétroliers. Ces mesures sont cependant peu nombreuses à cause du coût très élevé des forages. Elles sont modélisées par une fonction aléatoire $Z(\mathbf{x})$.
- de mesures imprécises, mais très abondantes spatialement, de la profondeur de la couche déduite des temps sismiques. Cette deuxième variable est représentée par la variable régionalisée $s(\mathbf{x})$.

Puisque $Z(\mathbf{x})$ et $s(\mathbf{x})$ sont deux expressions du même phénomène "profondeur de la couche", on suppose que $Z(\mathbf{x})$ est en moyenne proportionnelle à $s(\mathbf{x})$ à une constante a_0 près

$$E[Z(\mathbf{x})] = a_0 + b_1 s(\mathbf{x})$$

$s(\mathbf{x})$ est une fonction qui décrit l'allure générale ("shape" en Anglais) de la carte de la couche, tandis que les données de $Z(\mathbf{x})$ donnent en quelques endroits seulement une information précise sur la profondeur de la couche.

Estimation avec une fonction de forme

Dans un premier temps on va examiner le cas d'une fonction aléatoire $Z(\mathbf{x})$ dont on voudrait améliorer l'estimation en introduisant la fonction de forme $s(\mathbf{x})$.

On impose que la valeur estimée avec des pondérateurs de somme unité soit en moyenne cohérente avec une interpolation exacte de $s(\mathbf{x})$

$$E \left[Z^*(\mathbf{x}_0) \right] = \sum_{\alpha=1}^n \lambda_{\alpha} E \left[Z(\mathbf{x}_{\alpha}) \right] = a_0 + b_1 \sum_{\alpha=1}^n \lambda_{\alpha} s(\mathbf{x}_{\alpha}) = a_0 + b_1 s(\mathbf{x}_0)$$

et on obtient la condition suivante sur les pondérateurs

$$s(\mathbf{x}_0) = \sum_{\alpha=1}^n \lambda_{\alpha} s(\mathbf{x}_{\alpha})$$

Cela donne une condition supplémentaire intégrée dans la fonction objectif

$$\phi = \sigma_E^2 - \mu_1 \left(\sum_{\alpha=1}^n \lambda_{\alpha} - 1 \right) - \mu_2 \left(\sum_{\alpha=1}^n \lambda_{\alpha} s(\mathbf{x}_{\alpha}) - s(\mathbf{x}_0) \right)$$

Le système de krigeage ordinaire devient alors un système de krigeage universel

$$\begin{cases} \sum_{\beta=1}^n \lambda_{\beta} C(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}) - \mu_1 - \mu_2 s(\mathbf{x}_{\alpha}) = C(\mathbf{x}_{\alpha} - \mathbf{x}_0) & \text{pour } \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_{\beta} = 1 \\ \sum_{\beta=1}^n \lambda_{\beta} s(\mathbf{x}_{\beta}) = s(\mathbf{x}_0) \end{cases}$$

Le cadre fixé jusqu'ici paraît évidemment trop étroit pour la plupart des applications et il convient de passer à des modèles non stationnaires.

Dérive invariante par translation

Considérons une fonction aléatoire $Z(\mathbf{x})$ non stationnaire (fonction aléatoire intrinsèque d'ordre k)

$$Z(\mathbf{x}) = Z_k(\mathbf{x}) + m_k(\mathbf{x})$$

constituée d'une partie déterministe $m_k(\mathbf{x})$, modélisant la dérive par un polynôme, et d'une partie aléatoire $Z_k(\mathbf{x})$ de covariance généralisée $K(\mathbf{h})$. La dérive $m_k(\mathbf{x})$ est un polynôme de degré k de fonctions des coordonnées $f_l(\mathbf{x})$ avec des coefficients a_l

$$m_k(\mathbf{x}) = \sum_{l=0}^L a_l f_l(\mathbf{x})$$

Les fonctions $f_l(\mathbf{x})$, au nombre de $L + 1$, doivent engendrer un espace vectoriel invariant par translation. On peut montrer que les seuls ensembles de fonctions satisfaisant à cette condition sont les exponentielles-polynômes.

REMARQUE 7.1 *A deux dimensions spatiales, de vecteur de coordonnées $\mathbf{x} = (x_1, x_2)$, on prend généralement les monômes suivants:*

$$f_0(\mathbf{x}) = 1, f_1(\mathbf{x}) = x_1, f_2(\mathbf{x}) = x_2, f_3(\mathbf{x}) = (x_1)^2, f_4(\mathbf{x}) = x_1 \cdot x_2, f_5(\mathbf{x}) = (x_2)^2$$

qui engendrent un espace vectoriel invariant par translation. Le nombre de fonctions de base, $L+1$, dépend alors du degré k de la dérive de la manière suivante:

$$k = 0 \quad \Rightarrow \quad L = 0 \quad (\text{une fonction de base})$$

$$k = 1 \quad \Rightarrow \quad L = 2 \quad (\text{trois fonctions de base})$$

$$k = 2 \quad \Rightarrow \quad L = 5 \quad (\text{six fonctions de base})$$

Ayant en vue des problèmes d'estimation, on regarde les pondérateurs λ_α interpolant exactement les termes $a_l f_l(\mathbf{x})$ de la dérive

$$\sum_{\alpha=1}^n \lambda_\alpha a_l f_l(\mathbf{x}_\alpha) = a_l f_l(\mathbf{x}_0)$$

Les coefficients a_l sont manifestement redondants et il suffit de considérer l'interpolation des fonctions de base $f_l(\mathbf{x})$.

En se limitant à des pondérateurs ainsi définis, on constate que l'erreur d'estimation s'annule

$$\mathbb{E} \left[Z^*(\mathbf{x}_0) - Z(\mathbf{x}_0) \right] = \mathbb{E} \left[Z_k^*(\mathbf{x}_0) - Z_k(\mathbf{x}_0) \right] = 0$$

pour $\sum_{\alpha=1}^n \lambda_\alpha f_l(\mathbf{x}_\alpha) = f_l(\mathbf{x}_0)$ et $l = 0, \dots, L$.

REMARQUE 7.2 *En introduisant un pondérateur $\lambda_0 = -1$, on peut aussi écrire*

$$\mathbb{E} \left[Z^*(\mathbf{x}_0) - Z(\mathbf{x}_0) \right] = \mathbb{E} \left[\sum_{\alpha=0}^n \lambda_\alpha Z_k(\mathbf{x}_\alpha) \right] = 0 \quad \text{pour} \quad \sum_{\alpha=0}^n \lambda_\alpha f_l(\mathbf{x}_\alpha) = 0, \quad \forall l$$

où des pondérateurs de somme nulle filtrent une dérive invariante par translation.

Covariance généralisée

Une fonction symétrique $K(\mathbf{h})$ est la covariance généralisée d'une fonction aléatoire intrinsèquement stationnaire à l'ordre k , si

$$\text{var} \left(\sum_{\alpha=0}^n \lambda_\alpha Z_k(\mathbf{x}_\alpha) \right) = \mathbb{E} \left[\left(\sum_{\alpha=0}^n \lambda_\alpha Z_k(\mathbf{x}_\alpha) \right)^2 \right] = \sum_{\alpha=0}^n \sum_{\beta=0}^n \lambda_\alpha \lambda_\beta K(\mathbf{x}_\alpha - \mathbf{x}_\beta)$$

pour tout ensemble de pondérateurs satisfaisant $\sum_{\alpha=0}^n \lambda_{\alpha} f_l(\mathbf{x}_{\alpha}) = 0$ pour chaque $l = 0, \dots, L$.

Une covariance généralisée est une fonction de type positif conditionnel d'ordre k

$$\sum_{\alpha=0}^n \sum_{\beta=0}^n \lambda_{\alpha} \lambda_{\beta} K(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}) \geq 0 \quad \text{pour} \quad \sum_{\alpha=0}^n \lambda_{\alpha} f_l(\mathbf{x}_{\alpha}) = 0, \quad l = 0, \dots, L$$

L'inférence du degré maximal du polynôme à filtrer et de la covariance généralisée se fait habituellement via des algorithmes automatisés.

REMARQUE 7.3 *Un modèle possible de covariance généralisée d'ordre k est*

$$K_{\theta}(\mathbf{h}) = \Gamma\left(-\frac{\theta}{2}\right) |\mathbf{h}|^{\theta} \quad \text{avec} \quad 0 < \theta < 2k+2$$

où $\Gamma(\cdot)$ est la fonction gamma.

REMARQUE 7.4 *On peut construire à partir du modèle précédent une autre covariance généralisée d'ordre k , appelée "polynomiale"*

$$K_{pol}(\mathbf{h}) = \sum_{u=0}^k b_u (-1)^{u+1} |\mathbf{h}|^{2u+1} \quad \text{avec}^1 \quad b_u \geq 0$$

Cette covariance généralisée gigogne a la particularité de renfermer des structures avec des comportements très différents à l'origine. Pour $k=1$ (dérive linéaire) le terme en $u=0$ est linéaire à l'origine et est adapté pour des variables régionalisées continues mais non dérivables, tandis que le terme en $u=1$ est approprié pour des variables régionalisées dérivables.

En rajoutant une covariance pépitique à ce modèle, on couvre également des phénomènes plus erratiques (discontinus). Le modèle prend alors un caractère polyvalent en ce qui concerne le comportement à l'origine, et est particulièrement bien adapté pour des ajustements automatiques, quitte à ce que les structures inutiles dans la modélisation d'une application aient des coefficients b_u nuls.

Les covariances généralisées polynomiales sont importantes, parce que l'ont peut les utiliser en voisinage glissant.

Krigeage avec dérive externe

La méthode de la dérive externe consiste à intégrer dans le système de krigeage des conditions d'universalité supplémentaires, relatives à une ou plusieurs variables externes $s_i(\mathbf{x})$, $i = 1, \dots, N$ mesurées de manière exhaustive dans le domaine. En fait, on doit connaître les $s_i(\mathbf{x})$ en tous les points de données \mathbf{x}_{α} des échantillons de $Z(\mathbf{x})$, ainsi qu'aux noeuds de la grille d'estimation.

Il faut bien voir que les conditions

$$\sum_{\alpha=1}^n \lambda_{\alpha} s_i(\mathbf{x}_{\alpha}) = s_i(\mathbf{x}_0) \quad i = 1, \dots, N$$

sont rajoutées au système de krigeage indépendamment du choix de la classe de covariances $K(\mathbf{h})$, d'où la dénomination tout-à-fait pertinente de dérivées "externes". La classe de covariances généralisées ne pourra en effet être fonction que de l'ensemble de fonctions de base $f_l(\mathbf{x})$, qu'on peut qualifier dans ce contexte de dérivée "interne".

REMARQUE 7.5 *Si on adopte une définition aussi restrictive pour la dérivée externe, il apparaît que la théorie classique du Krigeage Universel, dans laquelle on se limite d'emblée à la classe des variogrammes $\gamma(\mathbf{h})$ ou à la classe des fonctions de covariances $C(\mathbf{h})$, ne fait usage que de dérivées externes ! On pourrait même dire que les problèmes d'inférence du KU proviennent du fait que les fonctions de dérivée $f_l(\mathbf{x})$ ($l > 0$) ou $s_i(\mathbf{x})$ restent toujours externes (hormis le cas où les dérivées sont constantes dans une direction donnée, qui permet l'inférence du variogramme sous-jacent moyennant une hypothèse d'isotropie).*

Le système de krigeage avec dérivée interne et dérivées externes multiples s'écrit finalement

$$\left\{ \begin{array}{l} \sum_{\beta=1}^n \lambda_{\beta} K(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}) - \sum_{l=0}^L \mu_l f_l(\mathbf{x}_{\alpha}) - \sum_{i=1}^N \mu_i s_i(\mathbf{x}_{\alpha}) = K(\mathbf{x}_{\alpha} - \mathbf{x}_0) \\ \sum_{\beta=1}^n \lambda_{\beta} f_l(\mathbf{x}_{\beta}) = f_l(\mathbf{x}_0) \\ \sum_{\beta=1}^n \lambda_{\beta} s_i(\mathbf{x}_{\beta}) = s_i(\mathbf{x}_0) \end{array} \right. \quad \begin{array}{l} \text{pour } \alpha = 1, \dots, n \\ \text{pour } l = 0, \dots, L \\ \text{pour } i = 1, \dots, N \end{array}$$

Quand on travaille en voisinage glissant, il est utile de cartographier également les coefficients b_i des dérivées externes pour se rendre compte de l'influence exercée par les différentes variables externes au niveau de chaque noeud de la grille d'estimation.

EXEMPLE 7.6 (d'après [41]) *La Figure 7.1 montre la carte des températures moyennes au mois de Janvier réalisée par krigeage (en voisinage glissant) avec une dérivée linéaire, à partir de mesures en 146 stations météorologiques toutes situées en dessous de 400m d'altitude (alors que le plus haut point de l'Ecosse, le Ben Nevis, est situé à 1344m). On peut raisonnablement faire l'hypothèse que la température est en relation linéaire avec l'altitude et utiliser une carte topographique digitalisée pour introduire l'altitude comme dérivée*

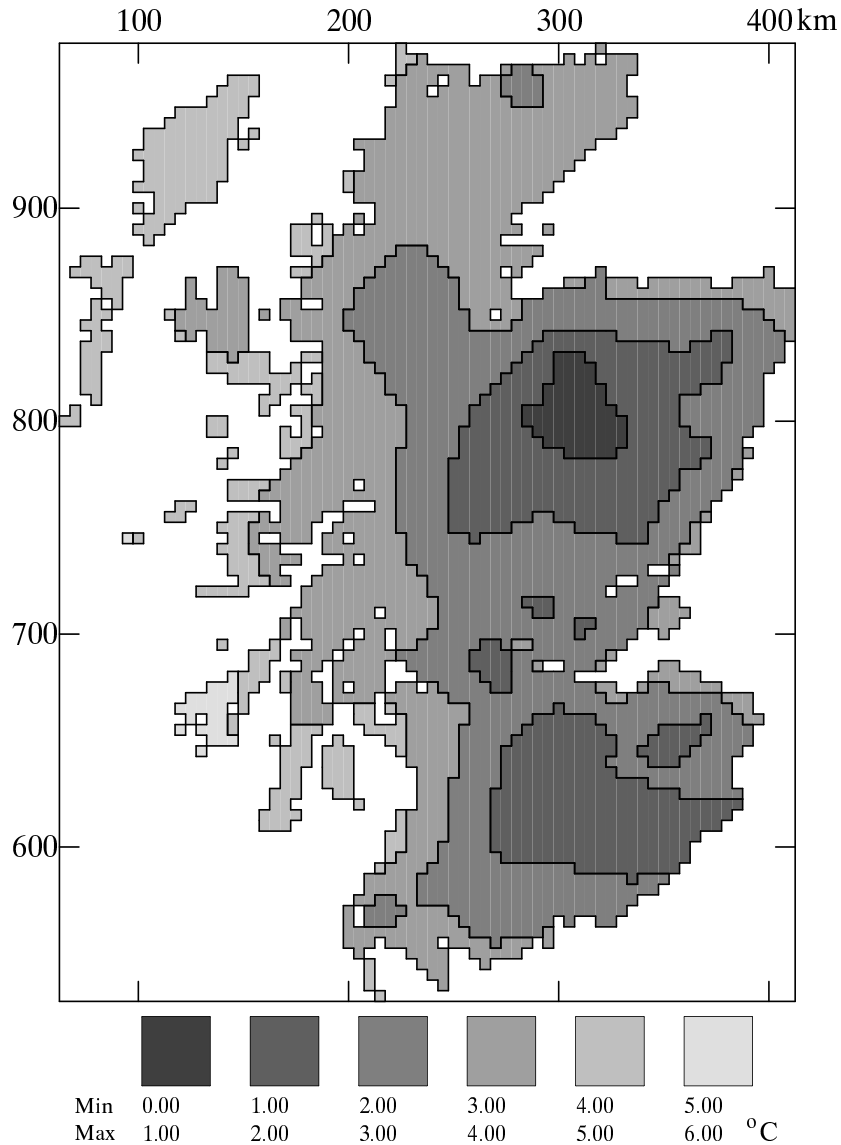


FIG. 7.1 – Krigeage avec une dérive linéaire des températures en Janvier en Ecosse.

externe $s(\mathbf{x})$ dans le krigeage de des températures. La carte correspondante est reproduite sur la Figure 7.2, tandis qu'une carte du coefficient b de la dérive externe est exhibée sur la Figure 7.3.

EXERCISE 7.7 (d'après [4]) On s'intéresse à des couples de valeurs (z_α, s_α) d'une variable d'intérêt $Z(\mathbf{x})$ de fonction de covariance $C(\mathbf{h})$ et d'une dérive externe $s(\mathbf{x})$ en des points $\mathbf{x}_\alpha, \alpha = 1, \dots, n$ d'un domaine \mathcal{D} . Ces valeurs sont rangées respectivement dans des vecteurs \mathbf{z} et \mathbf{s} de dimension n .

Pour fixer les notations, nous rappelons des systèmes de krigeage relatifs

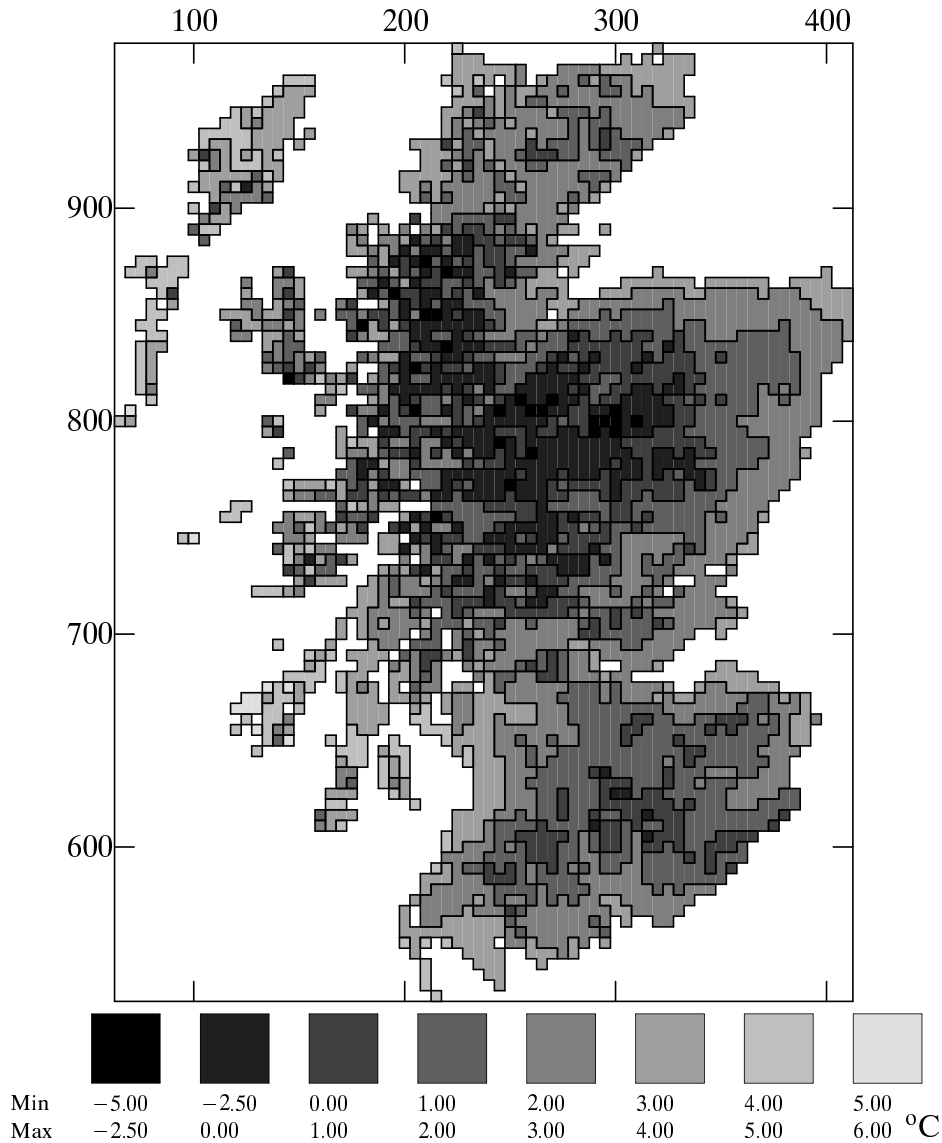


FIG. 7.2 – Krigeage avec une dérive linéaire des températures en Janvier en Ecosse en utilisant une carte topographique digitalisée comme dérive externe.

à l'estimation de la variable d'intérêt, ainsi que le krigeage de la moyenne,

Krigeage Simple (KS):

$$C \mathbf{l}_{KS} = \mathbf{c}_0 \quad \text{et} \quad \mathbf{l}_{KS} = \mathbf{R} \mathbf{c}_0 \quad \text{où:} \quad \mathbf{R} = \mathbf{C}^{-1}$$

Krigeage Ordinaire (KO):

$$\begin{pmatrix} \mathbf{C} & \mathbf{1} \\ \mathbf{1}^\top & 0 \end{pmatrix} \begin{pmatrix} \mathbf{l}_{KO} \\ -\mu_{KO} \end{pmatrix} = \begin{pmatrix} \mathbf{c}_0 \\ 1 \end{pmatrix} \quad \text{c'est-à-dire:} \quad \mathbf{K} \begin{pmatrix} \mathbf{l}_{KO} \\ -\mu_{KO} \end{pmatrix} = \mathbf{k}_0$$

et

$$\begin{pmatrix} \mathbf{l}_{KO} \\ -\mu_{KO} \end{pmatrix} = \mathbf{K}^{-1} \mathbf{k}_0 \quad \text{avec:} \quad \mathbf{K}^{-1} = \begin{pmatrix} \mathbf{A} & \mathbf{v} \\ \mathbf{v}^\top & w \end{pmatrix}$$

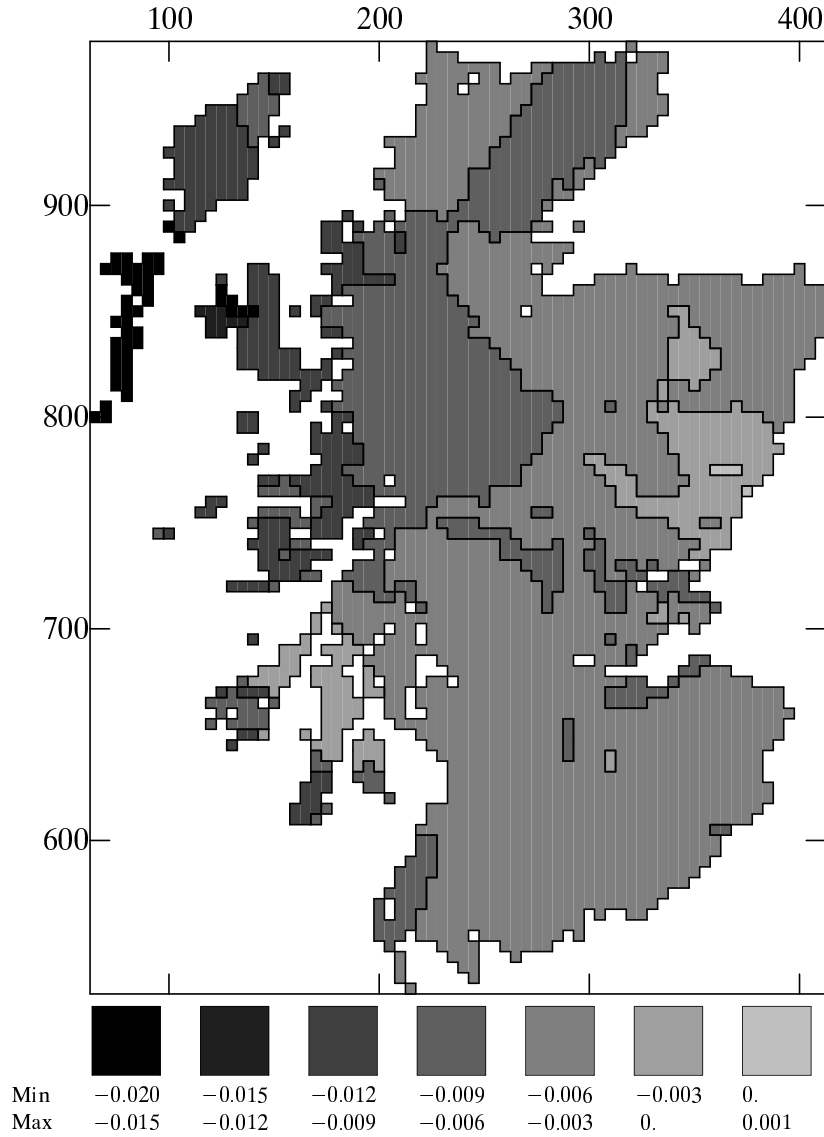


FIG. 7.3 – Carte du coefficient b de la dérive externe dans le krigage des températures en voisinage glissant.

Krigeage avec dérive Externe (KE):

$$\begin{pmatrix} \mathbf{C} & \mathbf{1} & \mathbf{s} \\ \mathbf{1}^\top & 0 & 0 \\ \mathbf{s}^\top & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{1}_{KE} \\ -\mu_{KE} \\ -\mu'_{KE} \end{pmatrix} = \begin{pmatrix} \mathbf{c}_0 \\ 1 \\ s_0 \end{pmatrix} \quad \text{c'est-à-dire:} \quad \mathbf{F} \begin{pmatrix} \mathbf{1}_{KE} \\ -\mu_{KE} \\ -\mu'_{KE} \end{pmatrix} = \mathbf{f}_0$$

Krigeage de la Moyenne (KM):

$$\mathbf{K} \begin{pmatrix} \mathbf{1}_{KM} \\ -\mu_{KM} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \quad \text{et} \quad m^* = \sum_{\alpha=1}^n \lambda_{KM}^\alpha z_\alpha = (\mathbf{z}^\top, 0) \mathbf{K}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$$

a) Montrer que la variance du KO s'écrit:

$$\sigma_{KO}^2 = c_{00} - \mathbf{k}_0^\top \mathbf{K}^{-1} \mathbf{k}_0$$

b) En se basant sur les relations suivantes entre les matrices et vecteurs du KO:

$$\begin{aligned} \mathbf{C} \mathbf{A} + \mathbf{1} \mathbf{v}^\top &= \mathbf{I} \\ \mathbf{C} \mathbf{v} + \mathbf{1} w &= \mathbf{o} \\ \mathbf{1}^\top \mathbf{A} &= \mathbf{o}^\top \\ \mathbf{1}^\top \mathbf{v} &= 1 \end{aligned}$$

montrer que le krigeage de la moyenne peut s'exprimer comme:

$$m^* = \frac{\mathbf{z}^\top \mathbf{R} \mathbf{1}}{\mathbf{1}^\top \mathbf{R} \mathbf{1}}$$

c) On définit un produit scalaire (de Castelier):

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{\mathbf{x}^\top \mathbf{R} \mathbf{y}}{\mathbf{1}^\top \mathbf{R} \mathbf{1}}$$

et, au moyen de celui-ci, une moyenne et une covariance:

$$\begin{aligned} E_n[\mathbf{z}] &= \langle \mathbf{z}, \mathbf{1} \rangle \\ E_n[\mathbf{z}, \mathbf{s}] &= \langle \mathbf{z}, \mathbf{s} \rangle \\ \text{cov}_n(\mathbf{z}, \mathbf{s}) &= E_n[\mathbf{z}, \mathbf{s}] - E_n[\mathbf{z}] E_n[\mathbf{s}] \end{aligned}$$

Sur la base d'un pseudo-produit scalaire de Castelier défini au moyen de l'inverse \mathbf{K}^{-1} du KO (au lieu de l'inverse \mathbf{R} du KS):

$$\mathcal{K}_n(\mathbf{z}, \mathbf{s}) = (\mathbf{z}^\top, 0) \mathbf{K}^{-1} \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix}$$

montrer que:

$$\mathcal{K}_n(\mathbf{z}, \mathbf{s}) = \mathbf{z}^\top \mathbf{A} \mathbf{s} = \mathbf{1}^\top \mathbf{R} \mathbf{1} \text{cov}_n(\mathbf{z}, \mathbf{s})$$

d) Sachant que la dérive externe est liée à l'espérance de la variable d'intérêt par une relation linéaire:

$$E[Z(\mathbf{x})] = a + b s(\mathbf{x})$$

et que la constante a et le coefficient b de la dérive externe sont estimés par:

$$b^* = (\mathbf{z}^\top, 0, 0) \mathbf{F}^{-1} \begin{pmatrix} \mathbf{o} \\ 0 \\ 1 \end{pmatrix} \quad \text{et} \quad a^* = (\mathbf{z}^\top, 0, 0) \mathbf{F}^{-1} \begin{pmatrix} \mathbf{o} \\ 1 \\ 0 \end{pmatrix}$$

montrer que b^* et a^* sont les coefficients d'une régression linéaire au sens du produit scalaire de Castelier:

$$b^* = \frac{\text{cov}_n(\mathbf{z}, \mathbf{s})}{\text{cov}_n(\mathbf{s}, \mathbf{s})} \quad \text{et} \quad a^* = E_n[\mathbf{z}] - b^* E_n[\mathbf{s}]$$

EXERCISE 7.8 (d'après [4]) *Considérons, dans un premier temps, que nous possédons une donnée z_0 en un point supplémentaire \mathbf{x}_0 , et écrivons la matrice membre gauche du KO avec $n + 1$ informations:*

$$\mathbf{K}_0 = \begin{pmatrix} c_{00} & \mathbf{k}_0^\top \\ \mathbf{k}_0 & \mathbf{K} \end{pmatrix} \quad \text{et} \quad \mathbf{K}_0^{-1} = \begin{pmatrix} w_{00} & \mathbf{v}_0^\top \\ \mathbf{v}_0 & \mathbf{A}_0 \end{pmatrix}$$

où $c_{00} = C(\mathbf{x}_0 - \mathbf{x}_0)$, \mathbf{K} , \mathbf{k}_0 sont définis comme dans l'exercice précédent.

Etant donné que $\mathbf{K}_0 \mathbf{K}_0^{-1} = \mathbf{I}$, on peut poser les relations:

$$\begin{aligned} c_{00} w_{00} + \mathbf{k}_0^\top \mathbf{v}_0 &= 1 \\ \mathbf{k}_0 w_{00} + \mathbf{K} \mathbf{v}_0 &= \mathbf{o} \\ c_{00} \mathbf{v}_0^\top + \mathbf{k}_0^\top \mathbf{A}_0 &= \mathbf{o}^\top \\ \mathbf{k}_0 \mathbf{v}_0^\top + \mathbf{K} \mathbf{A}_0 &= \mathbf{I} \end{aligned}$$

a) Montrer que:

$$w_{00} = \frac{1}{\sigma_{KO}^2}$$

b) Soient $\mathbf{z}_0^\top = (z_0, \mathbf{z}^\top)$ et $\mathbf{s}_0^\top = (s_0, \mathbf{s}^\top)$ des vecteurs de dimension $n + 1$ et:

$$\mathcal{K}_{n+1}(\mathbf{z}_0, \mathbf{s}_0) = (\mathbf{z}_0^\top, 0) \mathbf{K}_0^{-1} \begin{pmatrix} \mathbf{s}_0 \\ 0 \end{pmatrix}$$

Montrer que:

$$\mathcal{K}_{n+1}(\mathbf{z}_0, \mathbf{s}_0) = \mathcal{K}_n(\mathbf{z}, \mathbf{s}) + w_{00} \left(z_0 - (\mathbf{z}^\top, 0) \mathbf{K}^{-1} \mathbf{k}_0 \right) \cdot \left(s_0 - (\mathbf{s}^\top, 0) \mathbf{K}^{-1} \mathbf{k}_0 \right)$$

c) Soit $\Delta(z_0 | \mathbf{z})$ l'erreur du KO:

$$\Delta(z_0 | \mathbf{z}) = z_0 - z_0^* = z_0 - (\mathbf{z}^\top, 0) \mathbf{K}^{-1} \mathbf{k}_0$$

Montrer que:

$$\mathcal{K}_{n+1}(\mathbf{z}_0, \mathbf{s}_0) = z_0 \frac{\Delta(s_0 | \mathbf{s})}{\sigma_{KO}^2} - (\mathbf{z}^\top, 0) \mathbf{K}^{-1} \mathbf{k}_0 \frac{\Delta(s_0 | \mathbf{s})}{\sigma_{KO}^2} + \mathcal{K}_n(\mathbf{z}, \mathbf{s})$$

et que:

$$\frac{\Delta(s_0 | \mathbf{s})}{\sigma_{KO}^2} = (w_{00}, \mathbf{v}_0^\top) \cdot \begin{pmatrix} \mathbf{s}_0 \\ 0 \end{pmatrix}$$

EXERCISE 7.9 (d'après [4]) *Dans un deuxième temps, on s'intéresse à un KO avec $n - 1$ points et l'on note $\Delta(z_\alpha | \mathbf{z}_{-\alpha})$ l'erreur en un point \mathbf{x}_α d'un krigeage n'utilisant que les $n - 1$ données contenues dans le vecteur $\mathbf{z}_{-\alpha}^\top = (z_1, \dots, z_{\alpha-1}, z_{\alpha+1}, \dots, z_n)$ et l'on écrit σ_α^2 la variance de krigeage correspondante.*

a) Dédurre de ce qui précède la très élégante relation:

$$\mathbf{K}^{-1} \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix} = \begin{pmatrix} \vdots \\ \frac{\Delta(s_\alpha | \mathbf{s}_{-\alpha})}{\sigma_\alpha^2} \\ \vdots \\ m^* \end{pmatrix}$$

b) Sachant que l'on a pour le coefficient de la dérive externe:

$$b^* = \sum_{\alpha=1}^n \lambda_b^\alpha z_\alpha = \frac{\mathcal{K}_n(\mathbf{z}, \mathbf{s})}{\mathcal{K}_n(\mathbf{s}, \mathbf{s})}$$

montrer que l'on a:

$$\lambda_b^\alpha = \frac{\Delta(s_\alpha | \mathbf{s}_{-\alpha})}{\sigma_\alpha^2 \mathcal{K}_n(\mathbf{s}, \mathbf{s})}$$

On saisit l'intérêt, dans la pratique, du calcul des erreurs $\Delta(s_\alpha | \mathbf{s}_{-\alpha})$. Elles peuvent, en effet, servir à montrer l'influence des différentes valeurs s_α sur un KE dans un voisinage donné.

8 Dérive temporelle

Dérive spatiale et temporelle

Dans l'étude du magnétisme terrestre à proximité de l'équateur, on constate une forte variation temporelle avec une période de 24 h, dûe à la rotation de la terre. Les mesures prises par un bateau zigzaguant pendant plusieurs jours dans un domaine d'étude en région équatoriale doivent être corrigées de l'effet de cette variation diurne, si l'on veut établir une carte, en quelque sorte "intemporelle", du magnétisme (SÉGURET & HUCHON [89]).

Le magnétisme est modélisé par une fonction aléatoire spatio-temporelle $Z(\mathbf{x}, t)$ composée de deux fonctions aléatoires intrinsèques d'ordre k $Z_{k,\omega}(\mathbf{x})$, $Z_{k,\omega}(t)$ (supposées non corrélées), d'une dérive spatiale polynomiale $m_k(\mathbf{x})$ et d'une dérive temporelle périodique $m_\omega(t)$

$$Z(\mathbf{x}, t) = Z_{k,\omega}(\mathbf{x}) + m_k(\mathbf{x}) + Z_{k,\omega}(t) + m_\omega(t)$$

où \mathbf{x} est un point appartenant au domaine spatial \mathcal{D} et t est une coordonnée du temps \mathcal{T} .

La variation temporelle du magnétisme est connue à partir de stations fixes. A proximité de l'équateur, la composante prépondérante est une dérive périodique

$$m_{\omega,\varphi}(t) = \sin(\omega t + \varphi)$$

de période $\omega=24$ heures et de phase φ . En utilisant une relation bien connue, on peut aussi écrire

$$m_\omega(t) = a \sin(\omega t) + b \cos(\omega t)$$

où $a = \cos \varphi$ et $b = \sin \varphi$ sont des constantes multiplicatives qui seront prises en charge, dans les calculs, par les divers pondérateurs affectés à la dérive, de sorte que le paramètre de phase φ devient implicite.

Filtrage de la variation temporelle

En zone équatoriale, la composante temporelle aléatoire $Z_{k,\omega}(t)$ du magnétisme se traduisant par un bruit de faible ampleur par rapport à $m_\omega(t)$, on peut adopter dans un premier temps le modèle

$$Z(\mathbf{x}, t) \cong Z_{k,\omega}(\mathbf{x}) + m_k(\mathbf{x}) + m_\omega(t)$$

Pour cartographier un magnétisme intemporel $Z(\mathbf{x}) = Z_{k,\omega}(\mathbf{x}) + m_k(\mathbf{x})$, on fait d'abord l'inférence de la covariance généralisée en tenant compte de la dérive spatio-temporelle, puis on filtre la dérive temporelle en l'annulant dans le membre droit du système de krigeage

$$\left\{ \begin{array}{l} \sum_{\beta=1}^n \lambda_{\beta} K(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}) - \sum_{l=0}^L \mu_l f_l(\mathbf{x}_{\alpha}) \quad -\nu_1 \sin(\omega t_{\alpha}) - \nu_2 \cos(\omega t_{\alpha}) \\ \hspace{15em} = K(\mathbf{x}_{\alpha} - \mathbf{x}_0) \quad \text{pour } \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_{\beta} f_l(\mathbf{x}_{\beta}) = f_l(\mathbf{x}_0) \quad \text{pour } l = 0, \dots, L \\ \sum_{\beta=1}^n \lambda_{\beta} \sin(\omega t_{\beta}) = \boxed{0} \\ \sum_{\beta=1}^n \lambda_{\beta} \cos(\omega t_{\beta}) = \boxed{0} \end{array} \right.$$

avec $\alpha = 1, \dots, n$ et $l = 0, \dots, L$.

Si toutefois la composante temporelle aléatoire $Z_{k,\omega}(t)$ est non négligeable, il faut d'abord faire l'inférence de la covariance généralisée $K(t)$ correspondante, puis filtrer $Z(t) = Z_{k,\omega}(t) + m_{\omega}(t)$ par le système (avec $\alpha = 1, \dots, n$ et $l = 0, \dots, L$)

$$\left\{ \begin{array}{l} \sum_{\beta=1}^n \lambda_{\beta} \boxed{K(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}, t_{\alpha} - t_{\beta})} - \sum_{l=0}^L \mu_l f_l(\mathbf{x}_{\alpha}) \quad -\nu_1 \sin(\omega t_{\alpha}) - \nu_2 \cos(\omega t_{\alpha}) \\ \hspace{15em} = K(\mathbf{x}_{\alpha} - \mathbf{x}_0) \quad \text{pour } \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_{\beta} f_l(\mathbf{x}_{\beta}) = f_l(\mathbf{x}_0) \quad \text{pour } l = 0, \dots, L \\ \sum_{\beta=1}^n \lambda_{\beta} \sin(\omega t_{\beta}) = 0 \\ \sum_{\beta=1}^n \lambda_{\beta} \cos(\omega t_{\beta}) = 0 \end{array} \right.$$

où les termes $K(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}, t_{\alpha} - t_{\beta})$ sont la somme des covariances $K(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta})$ et $K(t_{\alpha} - t_{\beta})$.

9 Covariance Croisée

Fonction de covariance croisée

Les fonctions de covariance $C_{ij}(\mathbf{h})$ d'un ensemble de N fonctions aléatoires $Z_i(\mathbf{x})$ sont définies par l'hypothèse de stationnarité d'ordre deux conjointe

$$\begin{cases} \mathbb{E}[Z_i(\mathbf{x})] = m_i & \text{pour tout } \mathbf{x} \in \mathcal{D}; i = 1, \dots, N \\ \mathbb{E}\left[\left(Z_i(\mathbf{x}) - m_i\right)\left(Z_j(\mathbf{x} + \mathbf{h}) - m_j\right)\right] = C_{ij}(\mathbf{h}) & \text{pour tout } \mathbf{x} \in \mathcal{D}; i, j = 1, \dots, N \end{cases}$$

La moyenne de chaque variable $Z_i(\mathbf{x})$ en tout point du domaine est égale à une constante m_i . La covariance d'une paire de variables ne dépend que du vecteur \mathbf{h} reliant une paire de points et elle est invariante pour toute translation de la paire de points dans le domaine.

La matrice de fonctions de covariance simples et croisées est de type positif, la variance de toute combinaison linéaire de N variables en $n+1$ points avec un ensemble de poids λ_α^i devant être positive

$$\sum_{i=1}^N \sum_{j=1}^N \sum_{\alpha=0}^n \sum_{\beta=0}^n \lambda_\alpha^i C_{ij}(\mathbf{x}_\alpha - \mathbf{x}_\beta) \lambda_\beta^j \geq 0 \quad \text{pour tout } \mathbf{x}_\alpha, \mathbf{x}_\beta \in \mathcal{D}; \lambda_\alpha^i, \lambda_\beta^j \in \mathbb{R}$$

Effet de retard

La fonction de covariance croisée n'est pas a priori une fonction paire ou impaire. Généralement on a pour $i \neq j$

$$C_{ij}(\mathbf{h}) \neq C_{ji}(\mathbf{h}) \quad \text{et} \quad C_{ij}(-\mathbf{h}) \neq C_{ij}(\mathbf{h})$$

mais toujours

$$C_{ij}(\mathbf{h}) = C_{ji}(-\mathbf{h})$$

En particulier, il arrive que le maximum de la fonction de covariance soit décalé de l'origine $\mathbf{h} = \mathbf{0}$ vers un point $\mathbf{r} \neq \mathbf{0}$. Ce décalage de la corrélation maximale est fréquent pour des séries temporelles, où une variable peut avoir une influence sur une deuxième variable. Le temps de réaction de la deuxième

variable à des fluctuations de la première provoque alors un “retard” dans la corrélation entre les séries temporelles.

EXEMPLE 9.1 Soit $Z_1(\mathbf{x})$ une fonction aléatoire obtenue en multipliant par une constante a une fonction aléatoire $Z_2(\mathbf{x})$ décalée de \mathbf{r}

$$Z_1(\mathbf{x}) = a Z_2(\mathbf{x} - \mathbf{r}) + E(\mathbf{x})$$

où $E(\mathbf{x})$ est une erreur de mesure sans corrélation spatiale, de moyenne nulle et de fonction de covariance $C_{pep}(\mathbf{h})$.

La fonction de covariance simple de $Z_1(\mathbf{x})$ est proportionnelle à celle de $Z_2(\mathbf{x})$

$$C_{11}(\mathbf{h}) = a^2 C_{22}(\mathbf{h}) + C_{pep}(\mathbf{h})$$

et la fonction de covariance croisée est obtenue à partir de la fonction paire $C_{22}(\mathbf{h})$ translatée de \mathbf{r}

$$C_{12}(\mathbf{h}) = a C_{22}(\mathbf{h} - \mathbf{r})$$

EXERCISE 9.2 Calculer la fonction de covariance croisée entre $Z_1(\mathbf{x})$ et $Z_2(\mathbf{x})$ pour

$$Z_1(\mathbf{x}) = a_1 Z_2(\mathbf{x} - \mathbf{r}_1) + a_2 Z_2(\mathbf{x} - \mathbf{r}_2)$$

dans laquelle interviennent deux décalages \mathbf{r}_1 et \mathbf{r}_2 .

Variogramme croisé

Le variogramme croisé $\gamma_{ij}(\mathbf{h})$ fait son apparition dans le cadre d’une hypothèse de stationnarité intrinsèque conjointe de N fonctions aléatoires

$$\begin{cases} \mathbb{E} \left[Z_i(\mathbf{x} + \mathbf{h}) - Z_i(\mathbf{x}) \right] = 0 & \forall \mathbf{x} \in \mathcal{D}; i = 1, \dots, N \\ \text{cov} \left[\left(Z_i(\mathbf{x} + \mathbf{h}) - Z_i(\mathbf{x}) \right), \left(Z_j(\mathbf{x} + \mathbf{h}) - Z_j(\mathbf{x}) \right) \right] = 2 \gamma_{ij}(\mathbf{h}) & \forall \mathbf{x} \in \mathcal{D}; i, j = 1, \dots, N \end{cases}$$

Il s’ensuit qu’il est défini par

$$\gamma_{ij}(\mathbf{h}) = \frac{1}{2} \mathbb{E} \left[\left(Z_i(\mathbf{x} + \mathbf{h}) - Z_i(\mathbf{x}) \right) \cdot \left(Z_j(\mathbf{x} + \mathbf{h}) - Z_j(\mathbf{x}) \right) \right]$$

Le variogramme croisé étant de toute évidence une fonction paire, il est intéressant d’étudier sa relation avec la covariance croisée dans un contexte de stationnarité d’ordre deux. On obtient facilement la formule

$$\gamma_{ij}(\mathbf{h}) = C_{ij}(0) - \frac{1}{2} \left(C_{ij}(-\mathbf{h}) + C_{ij}(+\mathbf{h}) \right)$$

qui montre que le variogramme croisé prend la moyenne de la valeur en $-\mathbf{h}$ et en $+\mathbf{h}$ de la fonction de covariance correspondante. En décomposant la fonction de covariance croisée en une fonction paire et une fonction impaire

$$C_{ij}(\mathbf{h}) = \underbrace{\frac{1}{2} \left(C_{ij}(+\mathbf{h}) + C_{ij}(-\mathbf{h}) \right)}_{\text{terme pair}} + \underbrace{\frac{1}{2} \left(C_{ij}(+\mathbf{h}) - C_{ij}(-\mathbf{h}) \right)}_{\text{terme impair}}$$

on voit que le variogramme croisé n'incorpore que le terme pair de la fonction de covariance croisée.

EXEMPLE 9.3 (GAS FURNACE DATA) On trouve dans BOX & JENKINS [2] (p. 371, 532) l'exemple de deux séries temporelles correspondant, d'une part, à la fluctuation du volume d'un gaz introduit dans un fourneau, d'autre part, à la fluctuation du taux de CO_2 en résultant. La réaction chimique prenant quelques dizaines de secondes, on observe un effet de retard sur des mesures prises toutes les 9 secondes.

La Figure 9.1 montre la covariance croisée expérimentale

$$C_{ij}^*(\mathbf{h}) = \frac{1}{n_{\mathbf{h}}} \sum_{\alpha=1}^{n_{\mathbf{h}}} \left(z_i(\mathbf{x}_{\alpha}) - m_i \right) \cdot \left(z_j(\mathbf{x}_{\alpha} + \mathbf{h}) - m_j \right)$$

qui permet de constater un retard de 45 secondes entre les fluctuations de l'apport de gaz et leur effet sur le taux de gaz carbonique mesuré en sortie du système. La même Figure 9.1 exhibe le variogramme croisé expérimental

$$\gamma_{ij}^*(\mathbf{h}) = \frac{1}{2n_{\mathbf{h}}} \sum_{\alpha=1}^{n_{\mathbf{h}}} \left(z_i(\mathbf{x}_{\alpha} + \mathbf{h}) - z_i(\mathbf{x}_{\alpha}) \right) \cdot \left(z_j(\mathbf{x}_{\alpha} + \mathbf{h}) - z_j(\mathbf{x}_{\alpha}) \right)$$

qui est symétrique par rapport à l'ordonnée.

La Figure 9.2 reproduit les valeurs du terme pair et du terme impair de la covariance croisée expérimentale. On y retrouve, dans le terme pair, la structure suggérée par le variogramme croisé. Le terme impair mesure le degré d'asymétrie de la covariance croisée par rapport à l'ordonnée (en cas de symétrie il serait identiquement nul).

Représentation spectrale

La matrice $\mathbf{C}(\mathbf{h})$ de fonctions de covariance d'un vecteur de fonctions aléatoires complexes conjointement stationnaires $\mathbf{z}(\mathbf{x})$ (de moyennes nulles, sans perte de généralité)

$$\mathbf{C}(\mathbf{h}) = \mathbb{E} \left[\mathbf{z}(\mathbf{x}) \overline{\mathbf{z}(\mathbf{x} + \mathbf{h})}^{\top} \right]$$

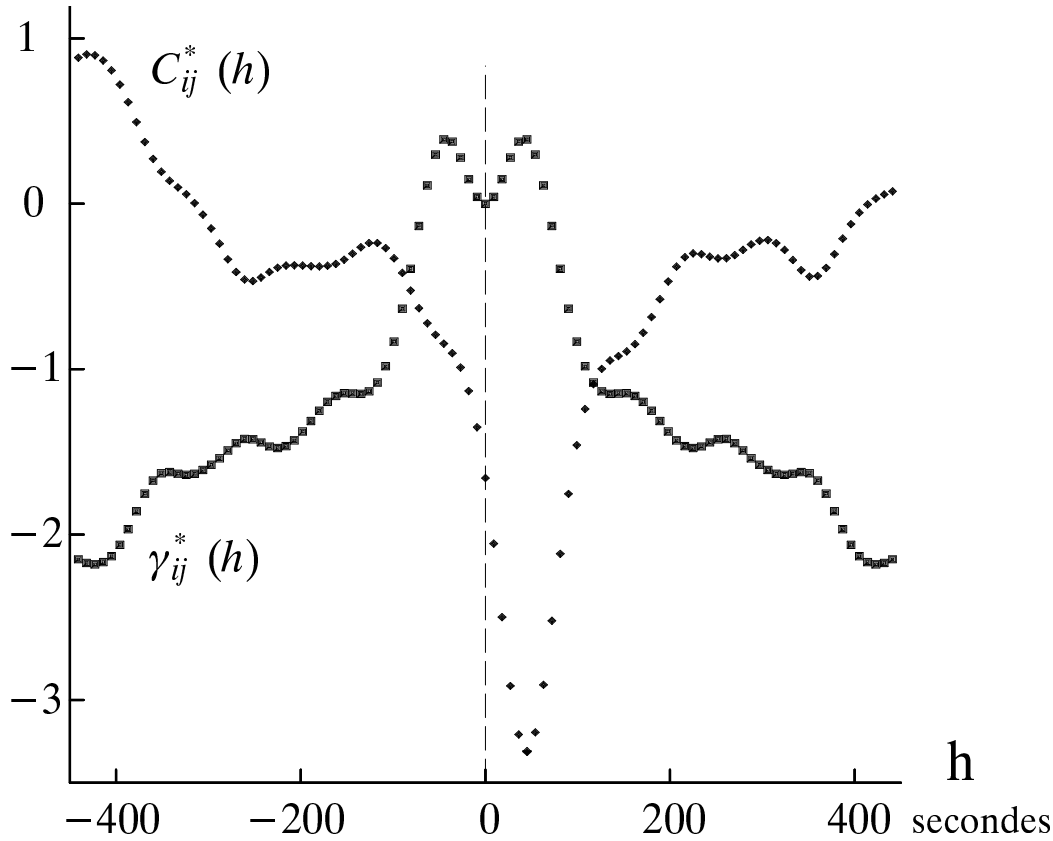


FIG. 9.1 – Covariance croisée expérimentale $C_{ij}^*(\mathbf{h})$ et variogramme croisé expérimental $\gamma_{ij}^*(\mathbf{h})$.

est semi-définie positive hermitienne¹, c'est-à-dire

$$\sum_{i=1}^N \sum_{j=1}^N \sum_{\alpha=0}^n \sum_{\beta=0}^n \lambda_{\alpha}^i C_{ij}(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}) \overline{\lambda_{\beta}^j} \geq 0 \quad \text{pour tout } \mathbf{x}_{\alpha}, \mathbf{x}_{\beta} \in \mathcal{D}; \lambda_{\alpha}^i, \lambda_{\beta}^j \in \mathbb{C}$$

Selon la généralisation multivariable et multidimensionnelle du théorème de Bochner, une matrice de fonctions continues $\mathbf{C}(\mathbf{h})$, $\mathbf{h} \in \mathbb{R}^d$, est une matrice de fonctions de covariances, ssi elle admet une représentation spectrale

$$\mathbf{C}(\mathbf{h}) = \int_{\mathbb{R}^d} \mathbf{e}^{i \mathbf{k}^{\top} \mathbf{h}} \mathbf{F}(d\mathbf{k})$$

avec une matrice de fonctions hermitienne $\mathbf{F}(\cdot)$ semi-définie positive, quel que soit l'argument de $\mathbf{F}(\cdot)$.

1. Une matrice hermitienne est la généralisation dans le domaine complexe d'une matrice symétrique réelle. Les éléments diagonaux d'une matrice hermitienne \mathbf{A} d'ordre $N \times N$ sont réels et les éléments non diagonaux sont égaux à la conjuguée complexe de leur transposée: $a_{ij} = \overline{a_{ji}}$.

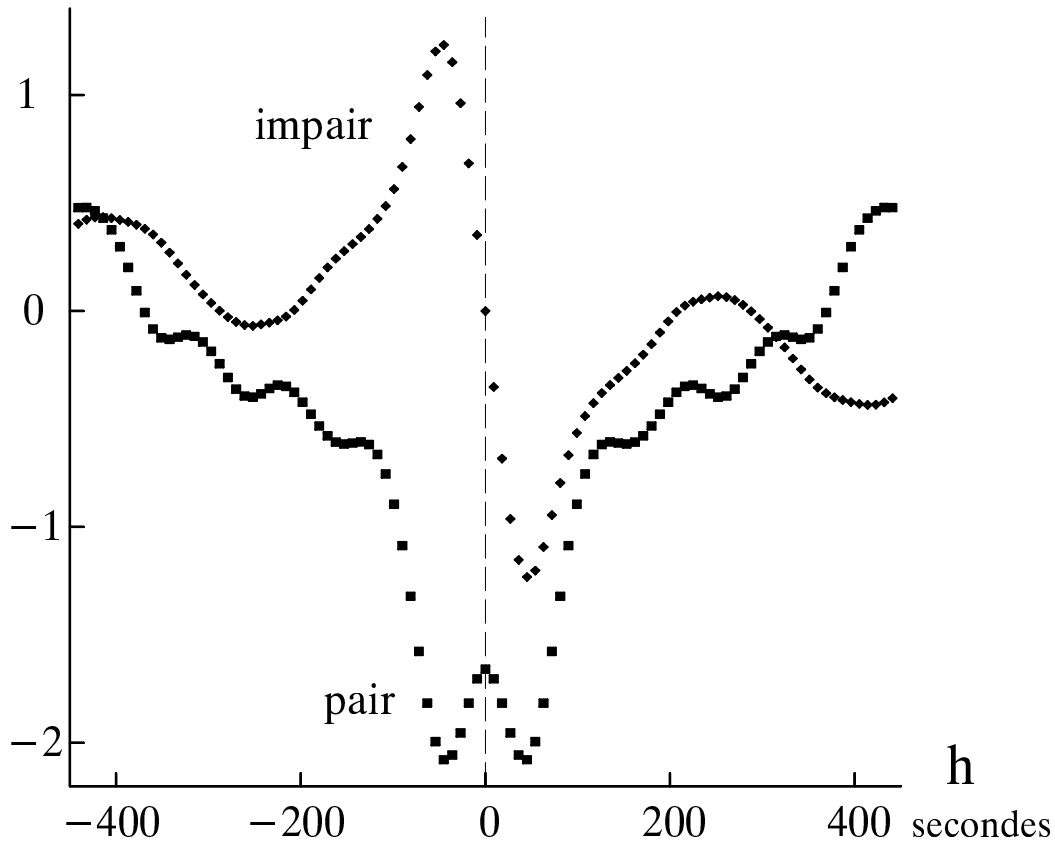


FIG. 9.2 – Termes pair et impair de la covariance croisée expérimentale.

Densité spectrale

Si l'on se limite à des fonctions de covariance intégrables en valeur absolue, une densité spectrale $f_{ij}(\mathbf{k})$ existe, tel que

$$C_{ij}(\mathbf{h}) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \mathbf{e}^{i\mathbf{k}^T \mathbf{h}} f_{ij}(\mathbf{k}) d\mathbf{k}$$

et, en inversant la transformée de Fourier

$$f_{ij}(\mathbf{k}) = \frac{1}{(2\pi)^n} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \mathbf{e}^{-i\mathbf{k}^T \mathbf{h}} C_{ij}(\mathbf{h}) d\mathbf{h}$$

La matrice des densités spectrales d'un ensemble de fonctions de covariances continues doit être semi-définie positive pour toute valeur de \mathbf{k} . Cela implique la relation de Schwarz

$$|f_{ij}(\mathbf{k})|^2 \leq f_{ii}(\mathbf{k}) f_{jj}(\mathbf{k})$$

Déphasage

Dans l'espace à une dimension, le long de l'axe temporel par exemple, on peut aisément interpréter des décalages spectraux. Si l'on considère les transformées de Fourier inverses (en admettant leur existence) du terme pair et du terme impair de la fonction de covariance croisée, qui sont appelées traditionnellement le "cospectre" $c_{ij}(k)$ et le "spectre de quadrature" $q_{ij}(k)$, on a

$$f_{ij}(k) = c_{ij}(k) - i q_{ij}(k)$$

En écriture polaire, la fonction complexe $f_{ij}(k)$ s'exprime sous la forme

$$f_{ij}(k) = |f_{ij}(k)| e^{i\varphi(k)}$$

où $\varphi(k)$ est le déphasage moyen de la proportion de $Z_i(x)$ de fréquence k sur la proportion de $Z_j(x)$ à cette fréquence, et

$$\tan \varphi(k) = \frac{\text{Im}(f_{ij}(k))}{\text{Re}(f_{ij}(k))} = \frac{-q_{ij}(k)}{c_{ij}(k)}$$

Un déphasage nul à une fréquence k implique une partie imaginaire $\text{Im}(f_{ij}(k))$ nulle à cette fréquence. Lorsqu'il n'y a pas de déphasage à aucune fréquence, la densité spectrale croisée $f_{ij}(k)$ est réelle et sa transformée de Fourier est une fonction paire.

Corrélation intrinsèque

Le modèle multivariable de fonctions de covariance réelles le plus simple que l'on puisse adopter consiste à décrire les relations entre variables par la matrice de variance-covariance \mathbf{V} et les relations entre les points d'échantillonnage par une fonction de corrélation $\rho(\mathbf{h})$ qui est la même pour toutes les variables

$$\mathbf{C}(\mathbf{h}) = \mathbf{V} \rho(\mathbf{h})$$

Ce modèle, dit "de corrélation intrinsèque", revient à choisir des covariances simples et croisées qui sont toutes proportionnelles à une même fonction de corrélation de base

$$C_{ij}(\mathbf{h}) = \sigma_{ij} \rho(\mathbf{h})$$

EXERCISE 9.4 *Montrer que le modèle de corrélation intrinsèque est de type positif.*

EXERCISE 9.5 *La covariance croisée réelle stationnaire $C_{ij}(\mathbf{h})$ vérifie-t-elle les inégalités suivantes:*

- a) $C_{ii}(\mathbf{0}) C_{jj}(\mathbf{0}) \geq |C_{ij}(\mathbf{h})|^2$
- b) $C_{ij}(\mathbf{0}) \geq |C_{ij}(\mathbf{h})|$
- c) $C_{ii}(\mathbf{h}) C_{jj}(\mathbf{h}) \geq |C_{ij}(\mathbf{h})|^2$

EXERCISE 9.6 Soit $C(\mathbf{h}) = \int e^{i\mathbf{k}^\top \mathbf{h}} \mathbf{F}(\mathbf{k}) d\mathbf{k}$ la représentation spectrale d'une matrice de fonctions de covariance réelles, continues et intégrables en valeur absolue. Montrer qu'il peut arriver que la matrice de distributions spectrales $\mathbf{F}(\mathbf{k}) = [f_{ij}(\mathbf{k})]$ soit hermitienne, bien que $C(\mathbf{h})$ soit réelle.

EXERCISE 9.7 Récemment, une définition alternative du variogramme croisé a fait son apparition dans la littérature sous le nom de "pseudo-variogramme croisé" (PVC), noté $\pi_{ij}(\mathbf{h})$. Un auteur le considère comme une généralisation plus naturelle du variogramme simple et augure même qu'il remplacera à terme le variogramme croisé habituel $\gamma_{ij}(\mathbf{h})$ (VCH).

Faites preuve de sens critique en examinant le PVC!

DÉFINITION (PVC):

$$\pi_{ij}(\mathbf{h}) = \frac{1}{2} \text{var} \left(Z_i(\mathbf{x} + \mathbf{h}) - Z_j(\mathbf{x}) \right)$$

- a) S'agit-il d'une généralisation du variogramme simple $\gamma_{ii}(\mathbf{h})$?
- b) Dans le contexte d'une hypothèse de stationnarité d'ordre 2, quelle relation peut-on établir entre le PVC et la fonction de covariance croisée?
- c) Dans le contexte d'une hypothèse intrinsèque, quelle hypothèse supplémentaire est-on obligé de faire pour le PVC, que l'on n'a pas besoin de faire pour le VCH?
- d) Discuter l(es) avantage(s) et inconvénient(s) que peut présenter le PVC par rapport au VCH (les examiner, en particulier, à la lumière des concepts d'hétérotopie et de corrélation négative).

10 Cokrigage

Isotopie et hétérotopie

Les mesures disponibles sur deux variables $Z_1(\mathbf{x})$ et $Z_2(\mathbf{x})$ dans un domaine donné peuvent être situées, soit aux mêmes points de mesure, soit en des points différents. On distingue les cas suivants:

- *isotopie*: en tous les points d'échantillonnage on dispose d'informations sur la paire de variables.
- *hétérotopie*: les deux variables ont été mesurées sur deux ensembles de points disjoints.
- *hétérotopie partielle*: une partie des points de mesure ne sont pas communs aux deux variables. Un cas particulier d'hétérotopie partielle est celui où l'ensemble des points de mesure d'une variable est inclu dans l'ensemble des points de mesure de l'autre variable.

L'estimateur du cokrigage ordinaire est une combinaison linéaire de pondérateurs λ_α^i avec des variables situées en des points d'un voisinage du domaine

$$Z_{i_0}^*(\mathbf{x}_0) = \sum_{i=1}^N \sum_{\alpha=1}^{n_i} \lambda_\alpha^i Z_i(\mathbf{x}_\alpha)$$

où l'index i_0 désigne une variable particulière de l'ensemble de N variables. On remarque que le nombre d'échantillons n_i dépend de l'indice i des variables, afin d'inclure dans la notation le cas de l'hétérotopie, partielle ou totale.

Cokrigage ordinaire

Dans le cadre d'une hypothèse de stationnarité intrinsèque conjointe, on désire estimer une variable particulière d'un ensemble de N variables, en se basant sur une erreur d'estimation nulle en moyenne. Cette condition est remplie en choisissant des pondérateurs de somme unité pour la variable d'intérêt, et de somme nulle pour les variables auxiliaires

$$\sum_{\alpha=1}^{n_i} \lambda_\alpha^i = \delta_{ii_0} = \begin{cases} 1 & \text{si } i = i_0, \\ 0 & \text{sinon} \end{cases}$$

On développe l'erreur d'estimation

$$\begin{aligned}
& \mathbb{E} \left[Z_{i_0}^*(\mathbf{x}_0) - Z_{i_0}(\mathbf{x}_0) \right] \\
&= \mathbb{E} \left[\sum_{i=1}^N \sum_{\alpha=1}^{n_i} \lambda_{\alpha}^i Z_i(\mathbf{x}_{\alpha}) - \underbrace{\sum_{\alpha=1}^{n_{i_0}} \lambda_{\alpha}^{i_0}}_1 Z_{i_0}(\mathbf{x}_0) - \sum_{\substack{i=0 \\ i \neq i_0}}^N \underbrace{\sum_{\alpha=1}^{n_i} \lambda_{\alpha}^i}_0 Z_i(\mathbf{x}_0) \right] \\
&= \sum_{i=1}^N \sum_{\alpha=1}^{n_i} \lambda_{\alpha}^i \underbrace{\mathbb{E} \left[Z_i(\mathbf{x}_{\alpha}) - Z_i(\mathbf{x}_0) \right]}_0 \\
&= 0
\end{aligned}$$

Pour la variance de l'erreur d'estimation il vient alors

$$\sigma_{\mathbb{E}}^2 = \mathbb{E} \left[\left(\sum_{i=1}^N \sum_{\alpha=1}^{n_i} \lambda_{\alpha}^i Z_i(\mathbf{x}_{\alpha}) - Z_{i_0}(\mathbf{x}_0) \right)^2 \right]$$

En introduisant des pondérateurs λ_0^i définis par

$$\lambda_0^i = -\delta_{ii_0} = \begin{cases} -1 & \text{si } i = i_0, \\ 0 & \text{si } i \neq i_0 \end{cases}$$

ajoutés aux sommations, on condense l'expression de la variance d'estimation

$$\sigma_{\mathbb{E}}^2 = \mathbb{E} \left[\left(\sum_{i=1}^N \sum_{\alpha=0}^{n_i} \lambda_{\alpha}^i Z_i(\mathbf{x}_{\alpha}) \right)^2 \right]$$

puis, en introduisant des variables aléatoires fictives situées arbitrairement à l'origine, on peut former des incréments

$$\begin{aligned}
\sigma_{\mathbb{E}}^2 &= \mathbb{E} \left[\left(\sum_{i=1}^N \left(\sum_{\alpha=0}^{n_i} \lambda_{\alpha}^i Z_i(\mathbf{x}_{\alpha}) - Z_i(\mathbf{0}) \underbrace{\sum_{\alpha=0}^{n_i} \lambda_{\alpha}^i}_0 \right) \right)^2 \right] \\
&= \mathbb{E} \left[\left(\sum_{i=1}^N \sum_{\alpha=0}^{n_i} \lambda_{\alpha}^i \underbrace{\left(Z_i(\mathbf{x}_{\alpha}) - Z_i(\mathbf{0}) \right)}_{\text{incrément}} \right)^2 \right]
\end{aligned}$$

et une covariance croisée d'incrément $C_{ij}^I(\mathbf{x}_{\alpha}, \mathbf{x}_{\beta})$

$$\sigma_{\mathbb{E}}^2 = \sum_{i=1}^N \sum_{j=1}^N \sum_{\alpha=0}^{n_i} \sum_{\beta=0}^{n_j} \lambda_{\alpha}^i \lambda_{\beta}^j C_{ij}^I(\mathbf{x}_{\alpha}, \mathbf{x}_{\beta})$$

Il est important de voir que l'on est obligé d'introduire, du point de vue physique, l'hypothèse que les covariances croisées d'incrémentes sont symétriques, pour pouvoir les remplacer par des variogrammes croisés. A cette condition, on a finalement

$$\begin{aligned} \sigma_E^2 = & 2 \sum_{i=1}^N \sum_{\alpha=1}^n \lambda_\alpha^i \gamma_{ii_0}(\mathbf{x}_\alpha - \mathbf{x}_0) - \gamma_{i_0 i_0}(\mathbf{x}_0 - \mathbf{x}_0) \\ & - \sum_{i=1}^N \sum_{j=1}^N \sum_{\alpha=1}^{n_i} \sum_{\beta=1}^{n_j} \lambda_\alpha^i \lambda_\beta^j \gamma_{ij}(\mathbf{x}_\alpha - \mathbf{x}_\beta) \end{aligned}$$

Après minimisation, où les contraintes sur les pondérateurs suscitent N paramètres de Lagrange μ_i , vient le cokrigage ordinaire

$$\begin{cases} \sum_{j=1}^N \sum_{\beta=1}^{n_j} \lambda_\beta^j \gamma_{ij}(\mathbf{x}_\alpha - \mathbf{x}_\beta) + \mu_i = \gamma_{ii_0}(\mathbf{x}_\alpha - \mathbf{x}_0) & \text{pour } i = 1, \dots, N; \alpha = 1, \dots, n_i \\ \sum_{\beta=1}^{n_i} \lambda_\beta^i = \delta_{ii_0} & \text{pour } i = 1, \dots, N \end{cases}$$

et la variance de cokrigage

$$\sigma_{CK}^2 = \sum_{i=1}^N \sum_{\alpha=1}^{n_i} \lambda_\alpha^i \gamma_{ii_0}(\mathbf{x}_\alpha - \mathbf{x}_0) + \mu_{i_0} - \gamma_{i_0 i_0}(\mathbf{x}_0 - \mathbf{x}_0)$$

Cokrigage simple

Le cokrigage ordinaire n'a en général pas de sens lorsqu'aucune information n'est disponible sur la variable d'intérêt dans le voisinage où l'on désire le pratiquer. Par contre, le cokrigage simple s'appuie sur une connaissance des moyennes des variables, ce qui lui permet de calibrer l'estimation d'une variable sans posséder aucune donnée sur cette variable dans le voisinage de cokrigage.

L'estimateur du cokrigage simple est constitué de la moyenne de la variable d'intérêt et d'une combinaison linéaire de pondérateurs λ_α^i avec les résidus de toutes les variables par rapport à leurs moyennes

$$Z_{i_0}^*(\mathbf{x}_0) = m_{i_0} + \sum_{i=1}^N \sum_{\alpha=1}^{n_i} \lambda_\alpha^i \left(Z_i(\mathbf{x}_\alpha) - m_i \right)$$

On lui associe le système de cokrigage simple, écrit sous forme matricielle

$$\begin{pmatrix} \mathbf{W}_{11} & \dots & \mathbf{C}_{1j} & \dots & \mathbf{C}_{1N} \\ \vdots & \ddots & & & \vdots \\ \mathbf{C}_{i1} & & \mathbf{W}_{ii} & & \mathbf{C}_{iN} \\ \vdots & & & \ddots & \vdots \\ \mathbf{C}_{N1} & \dots & \mathbf{C}_{Nj} & \dots & \mathbf{W}_{NN} \end{pmatrix} \begin{pmatrix} \mathbf{l}_1 \\ \vdots \\ \mathbf{l}_i \\ \vdots \\ \mathbf{l}_N \end{pmatrix} = \begin{pmatrix} \mathbf{c}_{1i_0} \\ \vdots \\ \mathbf{c}_{ii_0} \\ \vdots \\ \mathbf{c}_{Ni_0} \end{pmatrix}$$

où la matrice du membre gauche est construite avec des blocs symétriques \mathbf{W}_{ii} d'ordre $n_i \times n_i$ sur la diagonale, et des blocs rectangulaires \mathbf{C}_{ij} d'ordre $n_i \times n_j$ en dehors de la diagonale avec

$$\mathbf{C}_{ij} = \mathbf{C}_{ji}^\top$$

Dans un contexte d'isotopie, on peut réécrire le vecteur des pondérateurs sous la forme de l'opérateur de vectorisation $\text{vec}(\cdot)$ appliqué à une matrice \mathbf{L} d'ordre $n \times N$ contenant les vecteurs \mathbf{l}_i de pondérateurs des variables dans ses colonnes

$$\begin{pmatrix} \mathbf{l}_1 \\ \vdots \\ \mathbf{l}_i \\ \vdots \\ \mathbf{l}_N \end{pmatrix} = \text{vec}(\mathbf{L})$$

De même, le membre droit peut être représenté comme vectorisation d'une matrice \mathbf{C}_0 d'ordre $n \times N$ contenant les vecteurs de covariances \mathbf{c}_{ii_0} de chaque variable avec la variable d'intérêt. En désignant par \mathbf{W} la matrice membre gauche d'ordre $nN \times nN$, le cokrigage simple peut être posé comme

$$\mathbf{W} \text{vec}(\mathbf{L}) = \text{vec}(\mathbf{C}_0)$$

Autokrigeabilité

Une variable est dite "autokrigeable" par rapport à un ensemble de variables, si son krigeage propre coïncide avec son cokrigage. Le cas trivial est celui où toutes les variables sont non-corrélées avec la variable d'intérêt, et l'on voit facilement que tous les pondérateurs du cokrigage sont nuls, mis-à-part ceux affectés aux échantillons de la variable d'intérêt.

Dans une situation d'isotopie, une variable est autokrigeable si l'ensemble des variables est en corrélation intrinsèque, avec une matrice \mathbf{V} définie positive. Le cokrigage simple s'écrit alors

$$(\mathbf{V} \otimes \mathbf{R}) \text{vec}(\mathbf{L}) = \text{vec}(\mathbf{c}_0 \mathbf{v}_{i_0}^\top) = \mathbf{v}_{i_0} \otimes \mathbf{r}_0$$

La matrice membre gauche du cokrigage s'exprime ici comme le produit de Kronecker entre la matrice de variance-covariance \mathbf{V} et la matrice membre

gauche du krigeage \mathbf{R} (à une constante près). Le membre droit correspond à la vectorisation de la matrice résultant du produit entre le membre droit \mathbf{r}_0 du krigeage et la transposée d'un vecteur \mathbf{v}_{i_0} , lequel désigne la colonne de \mathbf{V} contenant la variance $\sigma_{i_0 i_0}$ de la variable d'intérêt. Or, comme on retrouve \mathbf{v}_{i_0} dans le membre gauche, il existe une solution du système pour laquelle les pondérateurs du vecteur \mathbf{l}_{i_0} sont ceux du krigeage direct et tous les autres sont nuls. C'est l'unique solution du système de cokrigeage, vu que \mathbf{V} et \mathbf{R} sont supposées définies positives.

En examinant les sous-blocs d'un système de cokrigeage quelconque établi avec un modèle de corégionalisation de type positif dans le cas isotopique, on s'aperçoit qu'il suffit, pour qu'une variable d'indice i_0 soit autokrigeable, que la fonction de covariance simple de cette variable soit proportionnelle à chacune des fonctions de covariances croisées

$$C_{i_0 j}(\mathbf{h}) = \sigma_{i_0 j} \rho(\mathbf{h})$$

EXERCISE 10.1 *Montrer que le concept d'autokrigeabilité s'applique sous des conditions identiques au cokrigeage ordinaire et au cokrigeage universel.*

11 Analyse Krigeante

L'étude des corégionalisations peut être subdivisée en deux parties:

- l'analyse de la corégionalisation d'un ensemble de variables, basée sur une interprétation par un modèle linéaire multivariable;
- un cokrigeage des facteurs définis par l'analyse des corégionalisations. Lorsqu'on a affaire qu'à une seule variable, l'opération se résume au krigage de composantes spatiales vu précédemment.

Cet ensemble de techniques a reçu le nom d'"analyse krigeante".

Modèle linéaire de corégionalisation

Un ensemble de fonctions aléatoires $\{Z_i(\mathbf{x}); i = 1, \dots, N\}$ peut être décomposé en des ensembles $\{Z_u^i(\mathbf{x}); u = 0, \dots, S\}$ de composantes spatialement non corrélées

$$Z_i(\mathbf{x}) = \sum_{u=0}^S Z_u^i(\mathbf{x})$$

Aux composantes spatiales sont associées des covariances $C_{ij}^u(\mathbf{h})$ de palier b_{ij}^u et de fonction de corrélation spatiale $\rho_u(\mathbf{h})$

$$C_{ij}(\mathbf{h}) = \sum_{u=0}^S C_{ij}^u(\mathbf{h}) = \sum_{u=0}^S b_{ij}^u \rho_u(\mathbf{h})$$

En regroupant les paliers b_{ij}^u en $S+1$ matrices de corégionalisation \mathbf{B}_u d'ordre $N \times N$, nécessairement semi-définies positives, on obtient le modèle de covariance

$$\mathbf{C}(\mathbf{h}) = \sum_{u=0}^S \mathbf{B}_u \rho_u(\mathbf{h})$$

EXERCISE 11.1 Dans quel cas ce modèle de covariance est-il équivalent au modèle de corrélation intrinsèque?

EXERCISE 11.2 Montrer qu'une fonction de corrélation $\rho_u(\mathbf{h})$ apparaissant avec un palier non nul sur une covariance croisée obtient nécessairement un palier non nul sur les deux covariances simples correspondantes.

Chacune des composantes spatiales $Z_u^i(\mathbf{x})$ peut elle-même être décomposée à l'aide de coefficients de transformation a_{up}^i en un ensemble de facteurs $Y_p^u(\mathbf{x})$ spatialement non corrélés et orthogonaux au sens d'une métrique donnée

$$Z_u^i(\mathbf{x}) = \sum_{p=1}^N a_{up}^i Y_p^u(\mathbf{x})$$

En combinant la décomposition spatiale avec la décomposition multivariable, on obtient finalement le modèle linéaire de corégionalisation

$$Z_i(\mathbf{x}) = \sum_{u=0}^S \sum_{p=1}^N a_{up}^i Y_p^u(\mathbf{x})$$

Ajustement bivariable de variogrammes expérimentaux

Le modèle de variogramme déduit d'un modèle linéaire de corégionalisation de fonctions aléatoires intrinsèques s'écrit

$$\Gamma(\mathbf{h}) = \sum_{u=0}^S \mathbf{B}_u g_u(\mathbf{h})$$

où les $g_u(\mathbf{h})$ sont des variogrammes normés.

Dans le cas bivariable, il est très simple de définir une procédure d'ajustement du modèle de variogramme aux variogrammes expérimentaux. On commence par ajuster les variogrammes simples en utilisant au moins une structure commune à ces deux variogrammes. Puis on peut ajuster les piliers b_{ij}^u du variogramme croisé en utilisant ceux des simples pour définir des bornes garantissant un modèle autorisé

$$|b_{ij}^u| \leq \sqrt{b_{ii}^u b_{jj}^u}$$

Il existe des algorithmes de moindres carrés sous contrainte permettant d'automatiser l'ajustement. L'extension de cette procédure à plus de deux variables ne garantit pas d'emblée un modèle autorisé, car les mineurs principaux d'ordre supérieur à deux des matrices \mathbf{B}_u ne sont pas forcément positifs.

Ajustement multivariable

L'ajustement multivariable de matrices de variogrammes expérimentaux $\Gamma^*(\mathbf{h}_k)$ calculées pour n_c classes de vecteurs \mathbf{h}_k par un algorithme itératif dû à GOULARD [35] est basé sur le critère de moindres carrés

$$\sum_{k=1}^{n_c} \text{tr} \left[\left(\Gamma^*(\mathbf{h}_k) - \Gamma(\mathbf{h}_k) \right)^2 \right]$$

que l'on cherche à minimiser sous la contrainte que les matrices de corégionalisation \mathbf{B}_u du modèle $\Gamma(\mathbf{h})$ soient semi-définies positives.

En se donnant S matrices \mathbf{B}_u semi-définies positives, on cherche dans un premier temps une $S+1$ -ième matrice \mathbf{B}_v qui comble au mieux les écarts $d\Gamma_{vk}^*$ entre un modèle à S matrices et les matrices expérimentales $\Gamma^*(\mathbf{h}_k)$

$$d\Gamma_{vk}^* = \Gamma^*(\mathbf{h}_k) - \sum_{\substack{u=0 \\ u \neq v}}^S \mathbf{B}_u g_u(\mathbf{h}_k)$$

La somme des écarts pondérés par $g_v(\mathbf{h}_k)$ est une matrice symétrique $d\Gamma_v^*$ qui n'est en général pas semi-définie positive. En la décomposant en valeurs et vecteurs propres

$$\sum_{k=1}^{n_c} d\Gamma_{vk}^* g_v(\mathbf{h}_k) = d\Gamma_v^* = \mathbf{Q}_v \Lambda_v \mathbf{Q}_v^\top \quad \text{avec } \mathbf{Q}_v^\top \mathbf{Q}_v = \mathbf{I}$$

on peut construire avec les valeurs propres positives une nouvelle matrice $\overline{d\Gamma}_v^*$, qui est la matrice semi-définie positive la plus proche de $d\Gamma_v^*$ au sens des moindres carrés

$$\overline{d\Gamma}_v^* = \mathbf{Q}_v \Lambda_v^o \mathbf{Q}_v^\top$$

où Λ_v^o correspond à Λ_v , si ce n'est que des valeurs propres nulles remplacent les valeurs propres négatives. En divisant par le carré du variogramme $g_v(\mathbf{h})$ de la $S + 1$ -ième structure, on obtient une matrice semi-définie positive \mathbf{B}_v^* minimisant le critère d'ajustement

$$\mathbf{B}_v^* = \frac{\overline{d\Gamma}_v^*}{\sum_{k=1}^{n_c} \left(g_v(\mathbf{h}_k) \right)^2}$$

Cette procédure est appliquée à tour de rôle à chacune des matrices \mathbf{B}_u , puis itérée, et l'on constate en pratique une assez bonne convergence du critère, bien que celle-ci ne soit pas assurée du point de vue théorique.

Analyse de corégionalisation

Lorsque l'on fait tendre \mathbf{h} vers l'infini dans un contexte de stationnarité d'ordre deux avec des fonctions $g_u(\mathbf{h})$ de palier unitaire, on voit que le modèle de variogramme tend vers la matrice de variance-covariance

$$\Gamma(\mathbf{h}) \rightarrow \mathbf{V} \quad \text{pour } \mathbf{h} \rightarrow \infty$$

Dans ce contexte, la matrice de variance-covariance est un mélange de matrices de corégionalisation

$$\mathbf{V} = \sum_{u=0}^S \mathbf{B}_u$$

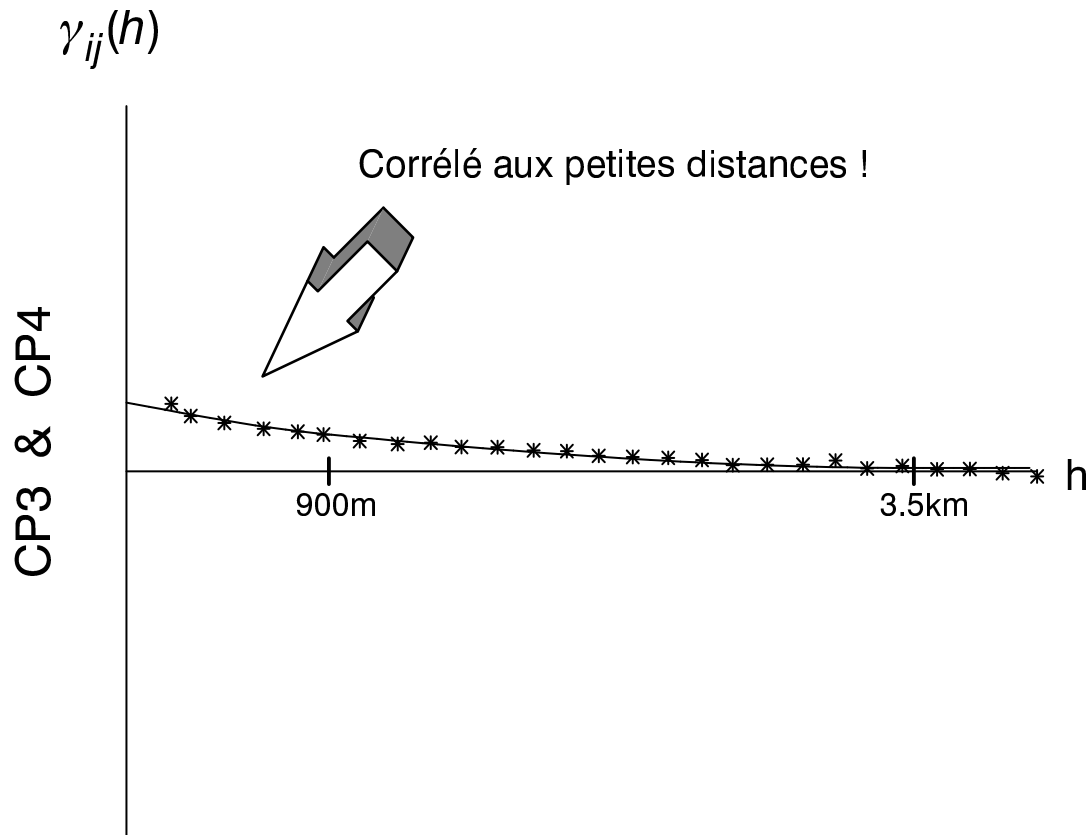


FIG. 11.1 – Variogramme croisé de deux composantes principales.

On voit que lorsqu'on ne se trouve pas dans le cas de la corrélation intrinsèque, il est du plus grand intérêt d'analyser séparément chacune des matrices de corégionalisation \mathbf{B}_u , considérées comme des matrices de variance-covariance d'un ensemble de variables $\{Z_u^i(\mathbf{x}); i = 1, \dots, N\}$. En outre, les matrices de corégionalisation \mathbf{B}_u peuvent être construites sous n'importe quelle hypothèse de stationnarité, alors que la matrice de variance-covariance \mathbf{V} n'a de sens que dans un contexte de stationnarité d'ordre deux.

EXEMPLE 11.3 (VARIOGRAMME CROISÉ DE DEUX COMPOSANTES PRINCIPALES)

La Figure 11.1 montre le variogramme croisé entre les troisième et quatrième composantes principales calculées sur la base de la matrice de variance-covariance de données pour lesquelles l'hypothèse de corrélation intrinsèque est manifestement fautive. En effet, près de l'origine les composantes principales sont corrélées de manière significative, symptôme du fait que la structure de corrélation des données est différente à petite échelle, ce dont une analyse en composantes principales non spatialisée ne peut tenir compte (d'après [102]).

L'analyse en composantes principales d'une matrice de corégionalisation

consiste à décomposer chaque matrice B_u en valeurs et vecteurs propres

$$B_u = A_u A_u^\top \quad \text{avec } A_u = Q_u \sqrt{\Lambda_u} \text{ et } Q_u Q_u^\top = I$$

On peut généraliser cette analyse en choisissant des vecteurs propres orthogonaux par rapport à une matrice symétrique M_u , représentant une métrique

$$A_u = Q_u M_u \sqrt{\Lambda_u} \quad \text{avec } Q_u M_u Q_u^\top = I$$

en utilisant, par exemple, la métrique

$$M_u = \sum_{u=0}^S B_u$$

Une autre possibilité, lorsqu'on s'intéresse à l'analyse conjointe de deux groupes de variables, est de partitionner les matrices de corégionalisation

$$B_u = \begin{pmatrix} B_u^{11} & B_u^{12} \\ B_u^{21} & B_u^{22} \end{pmatrix}$$

et de pratiquer une analyse canonique de la corégionalisation, c'est-à-dire, en prenant pour illustration le premier groupe

$$B_u^{11} = A_u^{11} \left(A_u^{11} \right)^\top$$

où

$$A_u^{11} = Q_u^{11} M_u^{11} \sqrt{\Lambda_u^{11}} \quad \text{avec } Q_u^{11} M_u^{11} \left(Q_u^{11} \right)^\top = I$$

et

$$M_u^{11} = B_u^{12} \left(B_u^{22} \right)^{-1} B_u^{21}$$

avec des matrices B_u^{22} évidemment non singulières.

L'analyse canonique donne souvent des résultats décevants, car difficilement interprétables. L'analyse en composantes principales avec des variables instrumentales (aussi appelée 'analyse de redondance') peut s'avérer une alternative plus performante. On passe à l'ACPVI simplement en remplaçant B_u^{22} par la matrice identité dans l'expression précédente, c'est-à-dire

$$M_u^{11} = B_u^{12} B_u^{21}$$

EXERCISE 11.4 *Montrer que*

$$A_u = Q_u M_u \sqrt{\Lambda_u} \quad \text{avec } Q_u M_u Q_u^\top = I$$

dans la factorisation $B_u = A_u A_u^\top$ est solution du problème de valeur propre généralisé

$$B_u Q_u = Q_u M_u \Lambda_u$$

Cokrigage de facteurs

Le modèle linéaire de corégionalisation permet d'estimer localement les facteurs régionalisés $Y_p^u(\mathbf{x})$ par cokrigage. L'estimateur utilisé pour évaluer un facteur $Y_{p_0}^{u_0}(\mathbf{x})$ donné est une combinaison linéaire de pondérateurs λ_α^i avec des variables dans le voisinage d'un point quelconque \mathbf{x}_0 du domaine

$$Y_{u_0 p_0}^*(\mathbf{x}) = \sum_{i=1}^N \sum_{\alpha=1}^{n_i} \lambda_\alpha^i Z_i(\mathbf{x}_\alpha)$$

En se plaçant dans le cadre d'une hypothèse de stationnarité locale, donnant un sens à des moyennes locales $m_i(\mathbf{x}_0)$, on construit un estimateur sans biais de facteurs synthétiques (auxquels on a attribué des moyennes nulles)

$$E \left[Y_{u_0 p_0}^*(\mathbf{x}) - Y_{p_0}^{u_0}(\mathbf{x}) \right] = \sum_{i=1}^N m_i(\mathbf{x}_0) \sum_{\alpha=1}^n \lambda_\alpha^i = 0 \quad \text{pour} \quad \sum_{\alpha=1}^n \lambda_\alpha^i = 0$$

La variance d'estimation σ_E^2 vaut

$$E \left[\left(Y_{u_0 p_0}^*(\mathbf{x}) - Y_{p_0}^{u_0}(\mathbf{x}) \right)^2 \right] = 1 + \sum_{i=1}^N \sum_{j=1}^N \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_\alpha^i \lambda_\beta^j C_{ij}(\mathbf{x}_\alpha - \mathbf{x}_\beta) - 2 \sum_{i=1}^N \sum_{\alpha=1}^n \lambda_\alpha^i a_{u_0 p_0}^i \rho_{u_0}(\mathbf{x}_\alpha - \mathbf{x}_0)$$

De sorte que l'on a le système de cokrigage

$$\left\{ \begin{array}{l} \sum_{j=1}^N \sum_{\beta=1}^{n_j} \lambda_\beta^j \mathbf{C}_{ij}(\mathbf{x}_\alpha - \mathbf{x}_\beta) - \mu_i = a_{u_0 p_0}^i \rho_{u_0}(\mathbf{x}_\alpha - \mathbf{x}_0) \\ \hspace{10em} \text{pour } i = 1, \dots, N; \alpha = 1, \dots, n \\ \sum_{\beta=1}^{n_i} \lambda_\beta^i = 0 \\ \hspace{10em} \text{pour } i = 1, \dots, N \end{array} \right.$$

Dans le membre droit de ce système apparaissent les coefficients de transformation a_{up}^i , résultats de la décomposition des matrices de corégionalisation \mathbf{B}_u . Dans le cas d'une analyse en composantes principales de ces matrices, les a_{up}^i indiquent les covariances entre les variables Z_i et les facteurs Y_p^u .

Le cokrigage de facteurs permet de réaliser des cartes de facteurs en n'utilisant que les données multivariées contenues dans un voisinage local autour de chaque point de la carte.

Solution des Exercices

EXERCICE 1.1 $\mathbf{1}^\top \mathbf{1} = n$;

$$\mathbf{1}\mathbf{1}^\top = \underbrace{\begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}}_{(n \times n)}$$

EXERCICE 1.2 $\frac{1}{n} \mathbf{Z}^\top \mathbf{1} = \mathbf{m}$ est le vecteur des moyennes m_i des variables.

$\frac{1}{n} \mathbf{1}\mathbf{1}^\top \mathbf{Z} = \mathbf{M}$ est une matrice $n \times N$ contenant dans chaque ligne la transposée du vecteur des moyennes; chaque colonne \mathbf{m}_i de \mathbf{M} a N éléments égaux à la moyenne m_i de la variable numéro i .

EXERCICE 1.5 $\mathbf{A}^2 \mathbf{x} = \mathbf{A} \lambda \mathbf{x} = \lambda \mathbf{A} \mathbf{x} = \lambda^2 \mathbf{x}$

EXERCICE 1.7 La décomposition classique en valeurs et vecteurs propres.

EXERCICE 1.8 Oui.

EXERCICE 2.2 $\text{corr}(\tilde{\mathbf{Z}}, \tilde{\mathbf{Z}}) = \tilde{\mathbf{Q}} \sqrt{\tilde{\mathbf{\Lambda}}} \sqrt{\tilde{\mathbf{\Lambda}}} \tilde{\mathbf{Q}}^\top = \text{corr}(\tilde{\mathbf{Z}}, \mathbf{Y}) \left[\text{corr}(\tilde{\mathbf{Z}}, \mathbf{Y}) \right]^\top$

EXERCICE 3.1 Il suffit de montrer que les contraintes d'orthogonalité correspondantes sont redondantes (de manière analogue à l'ACP).

EXERCICE 3.2 Idem.

EXERCICE 3.3 Multiplier l'équation par l'inverse de λ .

EXERCICE 4.3 Voir MATHERON [65], LAJAUNIE & LANTUÉJOL [48].

EXERCICE 6.3 Si un point de donnée coïncide avec le point d'estimation, le système d'équations correspondant aura une colonne du membre gauche (mis-à-part le dernier élément) égale au membre droit. Une solution du système est de prendre un ensemble de pondérateurs nuls, sauf le pondérateur de la colonne correspondant à la donnée coïncidente, valant un. Cette solution est la seule solution, puisque le membre gauche est supposé inversible, et elle restitue la valeur mesurée au point d'estimation.

EXERCICE 6.4 Etant donné qu'il y a interpolation exacte, les poids de $Z^*(\mathbf{x}_0)$ valent

$$\lambda_\alpha = \delta_{\mathbf{x}_\alpha, \mathbf{x}_0}$$

Vu la linéarité des estimateurs

$$\lambda_\alpha^P = \delta_{\mathbf{x}_\alpha, \mathbf{x}_0} - \lambda_\alpha^G$$

EXERCICE 6.5 Le krigeage ordinaire peut s'écrire

$$\underbrace{\begin{pmatrix} C_{1,1} & \dots & C_{1,N} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C_{N,1} & \dots & C_{N,N} & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix}}_{\mathbf{A}} \cdot \left[\underbrace{\begin{pmatrix} \lambda_1^P \\ \vdots \\ \lambda_N^P \\ \mu_P \end{pmatrix}}_{\mathbf{x}_P} + \underbrace{\begin{pmatrix} \lambda_1^G \\ \vdots \\ \lambda_N^G \\ \mu_G \end{pmatrix}}_{\mathbf{x}_G} + \underbrace{\begin{pmatrix} \lambda_1^{m_0} \\ \vdots \\ \lambda_N^{m_0} \\ \mu_{m_0} \end{pmatrix}}_{\mathbf{x}_{m_0}} \right]$$

$$= \underbrace{\begin{pmatrix} C_{0,1}^P \\ \vdots \\ C_{0,N}^P \\ 0 \end{pmatrix}}_{\mathbf{b}_P} + \underbrace{\begin{pmatrix} C_{0,1}^G \\ \vdots \\ C_{0,N}^G \\ 0 \end{pmatrix}}_{\mathbf{b}_G} + \underbrace{\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}}_{\mathbf{b}_{m_0}}$$

avec $\mathbf{A} \mathbf{x}_P = \mathbf{b}_P$, $\mathbf{A} \mathbf{x}_G = \mathbf{b}_G$, $\mathbf{A} \mathbf{x}_{m_0} = \mathbf{b}_{m_0}$.

On a alors

$$z^*(\mathbf{x}_0) = \mathbf{z}^\top \mathbf{l}_P + \mathbf{z}^\top \mathbf{l}_G + \mathbf{z}^\top \mathbf{l}_{m_0} = \mathbf{z}^\top (\mathbf{l}_P + \mathbf{l}_G + \mathbf{l}_{m_0}) = \mathbf{z}^\top \mathbf{l}$$

où \mathbf{z} est le vecteur des n données et $\mathbf{l}_P, \mathbf{l}_G, \mathbf{l}_{m_0}, \mathbf{l}$ sont des vecteurs de n pondérateurs correspondants.

EXERCICE 7.7 a)

$$\begin{aligned} \sigma_{\text{KO}}^2 &= c_{00} - \mathbf{k}_0^\top \mathbf{K}^{-1} \mathbf{k}_0 = c_{00} - \mathbf{k}_0^\top \begin{pmatrix} \mathbf{l}_{\text{KO}} \\ -\mu_{\text{KO}} \end{pmatrix} \\ &= C(\mathbf{x}_0 - \mathbf{x}_0) - \sum_{\alpha=1}^n \lambda_{\text{KO}}^\alpha C(\mathbf{x}_\alpha - \mathbf{x}_0) + \mu_{\text{KO}} \end{aligned}$$

b)

$$w = \frac{-1}{\mathbf{1}^\top \mathbf{R} \mathbf{1}} \quad \mathbf{v} = \frac{\mathbf{R} \mathbf{1}}{\mathbf{1}^\top \mathbf{R} \mathbf{1}} \quad \mathbf{A} = \mathbf{R} - \frac{\mathbf{R} \mathbf{1} (\mathbf{R} \mathbf{1})^\top}{\mathbf{1}^\top \mathbf{R} \mathbf{1}}$$

$$m^* = (\mathbf{z}^\top, 0) \mathbf{K}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} = (\mathbf{z}^\top, 0) \begin{pmatrix} \mathbf{v} \\ w \end{pmatrix} = \mathbf{z}^\top \mathbf{v} = \frac{\mathbf{z}^\top \mathbf{R} \mathbf{1}}{\mathbf{1}^\top \mathbf{R} \mathbf{1}}$$

c)

$$\begin{aligned} \mathcal{K}_n(\mathbf{z}, \mathbf{s}) &= \mathbf{z}^\top \mathbf{A} \mathbf{s} = \mathbf{1}^\top \mathbf{R} \mathbf{1} \left(\frac{\mathbf{z}^\top \mathbf{R} \mathbf{s}}{\mathbf{1}^\top \mathbf{R} \mathbf{1}} - \frac{\mathbf{z}^\top \mathbf{R} \mathbf{1}}{\mathbf{1}^\top \mathbf{R} \mathbf{1}} \cdot \frac{\mathbf{1}^\top \mathbf{R} \mathbf{s}}{\mathbf{1}^\top \mathbf{R} \mathbf{1}} \right) \\ &= \mathbf{1}^\top \mathbf{R} \mathbf{1} \left(E_n[\mathbf{z}, \mathbf{s}] - E_n[\mathbf{z}] E_n[\mathbf{s}] \right) = \mathbf{1}^\top \mathbf{R} \mathbf{1} \text{cov}_n(\mathbf{z}, \mathbf{s}) \end{aligned}$$

d)

Soient $\mathbf{z}' = \begin{pmatrix} \mathbf{z} \\ 0 \end{pmatrix}$, $\mathbf{s}' = \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix}$ et $\mathbf{F}^{-1} = \begin{pmatrix} \mathbf{K} & \mathbf{s}' \\ (\mathbf{s}')^\top & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{A}_F & \mathbf{v}_F \\ \mathbf{v}_F^\top & w_F \end{pmatrix}$

avec: $\mathbf{v}_F = \frac{\mathbf{K}^{-1} \mathbf{s}'}{(\mathbf{s}')^\top \mathbf{K}^{-1} \mathbf{s}'}$ et $\mathbf{A}_F = \mathbf{K}^{-1} - \frac{\mathbf{K}^{-1} \mathbf{s}' (\mathbf{K}^{-1} \mathbf{s}')^\top}{(\mathbf{s}')^\top \mathbf{K}^{-1} \mathbf{s}'}$

$$\begin{aligned} b^* &= (\mathbf{z}^\top, 0, 0) \mathbf{F}^{-1} \begin{pmatrix} \mathbf{0} \\ 0 \\ 1 \end{pmatrix} = (\mathbf{z}^\top, 0, 0) \begin{pmatrix} \mathbf{v}_F \\ w_F \end{pmatrix} = (\mathbf{z}')^\top \mathbf{v}_F \\ &= \frac{(\mathbf{z}')^\top \mathbf{K}^{-1} \mathbf{s}'}{(\mathbf{s}')^\top \mathbf{K}^{-1} \mathbf{s}'} = \frac{\mathcal{K}_n(\mathbf{z}, \mathbf{s})}{\mathcal{K}_n(\mathbf{s}, \mathbf{s})} = \frac{\text{cov}_n(\mathbf{z}, \mathbf{s})}{\text{cov}_n(\mathbf{s}, \mathbf{s})} \end{aligned}$$

$$\begin{aligned} a^* &= (\mathbf{z}^\top, 0, 0) \mathbf{F}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \\ 0 \end{pmatrix} = (\mathbf{z}^\top, 0) \mathbf{A}_F \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \\ &= \mathbf{z}' \left(\mathbf{K}^{-1} - \frac{\mathbf{K}^{-1} \mathbf{s}' (\mathbf{K}^{-1} \mathbf{s}')^\top}{(\mathbf{s}')^\top \mathbf{K}^{-1} \mathbf{s}'} \right) \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \\ &= \mathbf{z}' \mathbf{K}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} - \frac{\mathbf{z}' \mathbf{K}^{-1} \mathbf{s}'}{\mathbf{s}' \mathbf{K}^{-1} \mathbf{s}'} \cdot \mathbf{s}' \mathbf{K}^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \\ &= E_n[\mathbf{z}] - b^* E_n[\mathbf{s}] \end{aligned}$$

EXERCICE 7.8 a) $\mathbf{v}_0 = -\mathbf{K}^{-1} \mathbf{k}_0 w_{00}$ et

$$c_{00} w_{00} - \mathbf{k}_0^\top \mathbf{K}^{-1} \mathbf{k}_0 w_{00} = 1$$

$$w_{00} \left(c_{00} - \mathbf{k}_0^\top \mathbf{K}^{-1} \mathbf{k}_0 \right) = 1$$

$$w_{00} = \frac{1}{\sigma_{\text{KO}}^2}$$

b)

$$\mathcal{K}_{n+1}(\mathbf{z}_0, \mathbf{s}_0) = (z_0, \mathbf{z}^\top, 0) \begin{pmatrix} w_{00} & \mathbf{v}_0^\top \\ \mathbf{v}_0 & \mathbf{A}_0 \end{pmatrix} (s_0, \mathbf{s}^\top, 0)$$

avec: $\mathbf{A}_0 = \mathbf{K}^{-1} + w_{00} \mathbf{K}^{-1} \mathbf{k}_0 (\mathbf{K}^{-1} \mathbf{k}_0)^\top$

$$\begin{aligned} \mathcal{K}_{n+1}(\mathbf{z}_0, \mathbf{s}_0) &= z_0 w_{00} s_0 + (\mathbf{z}^\top, 0) \mathbf{v}_0 s_0 + z_0 \mathbf{v}_0^\top \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix} + (\mathbf{z}^\top, 0) \mathbf{A}_0 \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix} \\ &= w_{00} z_0 s_0 - w_{00} (\mathbf{z}^\top, 0) \mathbf{K}^{-1} \mathbf{k}_0 s_0 - w_{00} z_0 \mathbf{K}^{-1} \mathbf{k}_0 \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix} \\ &\quad + \mathcal{K}_n(\mathbf{z}, \mathbf{s}) + w_{00} (\mathbf{z}^\top, 0) \mathbf{K}^{-1} \mathbf{k}_0 (\mathbf{K}^{-1} \mathbf{k}_0)^\top \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix} \\ &= w_{00} \left(z_0 - (\mathbf{z}^\top, 0) \mathbf{K}^{-1} \mathbf{k}_0 \right) \cdot \left(s_0 - (\mathbf{s}^\top, 0) \mathbf{K}^{-1} \mathbf{k}_0 \right) + \mathcal{K}_n(\mathbf{z}, \mathbf{s}) \end{aligned}$$

c)

$$\mathcal{K}_{n+1}(\mathbf{z}_0, \mathbf{s}_0) = \left(z_0 - (\mathbf{z}^\top, 0) \mathbf{K}^{-1} \mathbf{k}_0 \right) \cdot \frac{\Delta(s_0 | \mathbf{s})}{\sigma_{\text{KO}}^2} + \mathcal{K}_n(\mathbf{z}, \mathbf{s})$$

et

$$\begin{aligned} (w_{00}, \mathbf{v}_0^\top) \cdot \begin{pmatrix} \mathbf{s}_0 \\ 0 \end{pmatrix} &= w_{00} s_0 + \mathbf{v}_0^\top \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix} = w_{00} s_0 - w_{00} \mathbf{K}^{-1} \mathbf{k}_0 \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix} \\ &= w_{00} \Delta(s_0 | \mathbf{s}) = \frac{\Delta(s_0 | \mathbf{s})}{\sigma_{\text{KO}}^2} \end{aligned}$$

EXERCICE 7.9 a) Pour $\alpha = 1$ on a:

$$(w_{11}, \mathbf{v}_1^\top) \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix} = \frac{\Delta(s_1 | \mathbf{s}_{-1})}{\sigma_1^2}$$

et pour $\alpha = n$:

$$(v_{1,n}, \dots, v_{n-1,n}, w_{nn}, v_{n+1,n}) \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix} = \frac{\Delta(s_n | \mathbf{s}_{-n})}{\sigma_n^2}$$

ainsi que pour $\alpha \neq 1$ et $\alpha \neq n$:

$$(v_{1,\alpha}, \dots, v_{\alpha-1,\alpha}, w_{\alpha\alpha}, v_{\alpha+1,\alpha}, \dots, v_{n+1,\alpha}) \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix} = \frac{\Delta(s_\alpha | \mathbf{s}_{-\alpha})}{\sigma_\alpha^2}$$

b)

$$b^* = \frac{\mathcal{K}_n(\mathbf{z}, \mathbf{s})}{\mathcal{K}_n(\mathbf{s}, \mathbf{s})} = (\mathbf{z}^\top, 0) \frac{\mathbf{K}^{-1} \begin{pmatrix} \mathbf{s} \\ 0 \end{pmatrix}}{\mathcal{K}_n(\mathbf{s}, \mathbf{s})} = \sum_{\alpha=1}^n z_\alpha \frac{\Delta(s_\alpha | \mathbf{s}_{-\alpha})}{\sigma_\alpha^2 \mathcal{K}_n(\mathbf{s}, \mathbf{s})}$$

EXERCICE 9.2 $C_{12}(\mathbf{h}) = a_1 C_{22}(\mathbf{h} - \mathbf{r}_1) + a_2 C_{22}(\mathbf{h} - \mathbf{r}_2)$

EXERCICE 9.4 La matrice \mathbf{V} est une matrice semi-définie positive en tant que matrice de variance-covariance et $\rho(\mathbf{h})$ est une fonction de type positif en tant que fonction de covariance, de sorte que l'on vérifie aisément que le produit des deux vérifie le théorème de Bochner multivariable.

EXERCICE 9.5 a) Oui. b) $C_{ij}(\mathbf{0})$ peut être négatif! c) $C_{ii}(\mathbf{h})$ peut être négatif!

EXERCICE 9.6 Lorsque $C_{ij}(\mathbf{h})$ n'est pas une fonction paire, le spectre de quadrature $q_{ij}(\mathbf{h})$ n'est pas nul et $f_{ij}(\mathbf{h})$ est complexe.

EXERCICE 9.7 a) $\pi_{ii}(\mathbf{h}) = \gamma_{ii}(\mathbf{h})$.

b) $\pi_{ij}(\mathbf{h}) = \frac{1}{2} (C_{ii}(\mathbf{0}) + C_{jj}(\mathbf{0})) - C_{ij}(\mathbf{h})$ et peut donc servir à détecter des effets de retard au même titre que la fonction de covariance croisée.

c) On est obligé de supposer pour $\pi_{ij}(\mathbf{h})$ que les accroissements croisés sont intrinsèques avec $E[Z_i(\mathbf{x} + \mathbf{h}) - Z_j(\mathbf{x})] = 0$. C'est aventureux, lorsque les variables sont de nature différente.

d) $\pi_{ij}(\mathbf{h})$ peut être calculé dans un contexte d'hétérotomie totale. Son inconvénient majeur est qu'il ne peut traduire une corrélation négative éventuelle entre deux variables.

EXERCICE 11.1 Lorsque toutes les matrices de corégionalisation sont proportionnelles à une même matrice \mathbf{B}

$$\mathbf{C}(\mathbf{h}) = \sum_{u=0}^S a_u \mathbf{B} \rho_u(\mathbf{h}) = \mathbf{B} \sum_{u=0}^S a_u \rho_u(\mathbf{h})$$

où les a_u sont des coefficients de proportionnalité.

EXERCICE 11.2 Etant donné que \mathbf{B}_u est semi-définie positive, la positivité de ses mineurs principaux d'ordre 2 implique la relation de Cauchy-Schwarz entre les paliers simples et croisés

$$|b_{ij}^u| \leq \sqrt{b_{ii}^u b_{jj}^u}$$

EXERCICE 11.4 $\mathbf{B}_u = \mathbf{Q}_u \mathbf{M}_u \mathbf{\Lambda}_u \mathbf{M}_u \mathbf{Q}_u^\top$ avec $\mathbf{Q}_u = (\mathbf{M}_u \mathbf{Q}_u^\top)^{-1}$, donc:

$$\mathbf{B}_u \mathbf{Q}_u = \mathbf{Q}_u \mathbf{M}_u \mathbf{\Lambda}_u \underbrace{(\mathbf{M}_u \mathbf{Q}_u^\top) \cdot (\mathbf{M}_u \mathbf{Q}_u^\top)^{-1}}_{\mathbf{I}}$$

Références

Mots-clefs

Théorie

- Rappel de calcul matriciel: [94]; [53]; [75].
- ACP, Analyse Canonique, Analyse des Correspondances: [75]; [55]; [85]; [32]; [38]; [31]; [71]; [104].
- Analyse de redondance (ACPVI): [74]; [95].
- Modèles isofactoriels, krigeage disjonctif: [62], [65]; [48]; [78]; [49]; [18].
- Modèle de régionalisation: [59].
- Krigeage de composantes spatiales: [59], [64]; [82]; [29]; [14]; [99]; [18].
- Déconvolution: [7]; [88]; [43].
- Dérive externe: [23]; [21]; [30]; [57]; [76]; [41]; [4].
- Dérive spatio-temporelle: [89]; [87].
- Covariance généralisée: [61]; [22]; [12]; [26].
- Covariance croisée: [15]; [105], [106]; [73]; [67]; [58].
- Pseudo-variogramme croisé: [69]; [16].
- Cokrigeage: [59], [63]; [5], [6].
- Equations aux dérivées partielles: [60]; [25]; [3].
- Ajustement multivariable de variogrammes expérimentaux: [35][36]; [46]; [37].
- Analyse de corégionalisation: [64]; [97]; [101]; [35], [36]; [33].
- Cokrigeage de facteurs: [64]; [83]; [98]; [19], [20]; [18].
- Classification: [91]; [70]; [103].
- Géostatistique exploratoire: [8]; [40].
- Fonction aléatoire complexe: [47].

Applications

- Analyse d'images: [52]; [19][20]; [43]; [18].
- Géodésie: [67].
- Hydrologie: [93]; [1]; [25]; [79]; [17]; [13].
- Exploration géochimique: [82]; [83]; [98]; [84]; [80]; [100]; [40]; [102].
- Exploration géophysique: [14]; [76]; [86]; [89]; [87].
- Exploration minière: [44]; [91].
- Exploration pétrolière et gazière: [24]; [21]; [29]; [57]; [30]; [42];

- Métallurgie: [19][20].
- Météorologie: [10]; [72]; [45]; [39]; [51]; [81]; [68]; [41].
- Science des sols: [103]; [36]; [70]; [104]; [37]; [92]; [33].
- Sylviculture: [54]; [27].
- Télédétection: [27]; [50].

Bibliographie

- [1] AHMED S & DE MARSILY (1987) *Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity*. Water Resources Research, 23, 1717–1737.
- [2] BOX GEP & JENKINS GM (1970) *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.
- [3] CASTELIER E (1992) *Le krigeage propagatif*. Mémoire de DEA, Publication S-291, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau, 22p.
- [4] CASTELIER E (1992) *La dérive externe vue comme une régression linéaire*. Publication N-34/92/G, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau, 22p.
- [5] CHAUVET P (1975) *Formulaire du cokrigeage en FAI-k. Application au cas Gaussien*. Publication N-425, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [6] CHAUVET P (1977) *Exemple d'application de la géostatistique non stationnaire: le cokrigeage*. Publication N-502, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [7] CHAUVET P (1977) *Exemple simplifié de déconvolution de courbe lecteur par cokrigeage*. Publication N-509, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [8] CHAUVET P (1982) *The variogram cloud*. In: Johnson TB & Barnes RJ (eds) 17th APCOM, 757–764, Society of Mining Engineers, New York.
- [9] CHAUVET P (1991) *Aide Mémoire de Géostatistique Linéaire*. Cahiers de Géostatistique, Fascicule 2, Ecole des Mines de Paris, Fontainebleau.
- [10] CHAUVET P, PAILLEUX J & CHILÈS JP (1976) *Analyse objective des champs météorologiques par cokrigeage*. La Météorologie, 6e série, No 4, 37–54.
- [11] CHAUVET P & GALLI A (1982) *Universal Kriging*. Publication C-96, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [12] CHILÈS JP (1977) *Géostatistique des Phénomènes Non Stationnaires (dans le plan)*. Thèse de Doctorat, Université de Nancy I, Nancy, 152p.
- [13] CHILÈS JP (1991) *Application du krigeage avec dérive externe à l'implantation d'un réseau de mesures piézométriques*. Sciences de la Terre, Série Informatique, 30, 131–147.
- [14] CHILÈS JP & GUILLEN A (1984) *Variogrammes et Krigeages pour la Gravimétrie et le Magnétisme*. Sciences de la Terre, Série Informatique, 20, 455–468.
- [15] CRAMER H & LEADBETTER MR (1967) *Stationary and Related Stochastic Processes*. Wiley, New York.
- [16] CRESSIE N (1991) *Statistics for Spatial Data*. Wiley, New York.
- [17] DAGAN G (1989) *Flow and Transport in Porous Formations*. Springer-Verlag, Berlin.
- [18] DALY C (1991) *Application de la Géostatistique à quelques Problèmes de Filtrage*. Thèse de Doctorat, Ecole des Mines de Paris, Fontainebleau.
- [19] DALY C, LAJAUNIE C & JEULIN D (1989) *Application of multivariate kriging to the processing of noisy images*. In: Armstrong M (ed) *Geostatistics*, 749–760, Kluwer Academic Publisher, Amsterdam, Holland.

- [20] DALY C, JEULIN D, BENOIT D & AUCLAIR G (1990) *Application of Multivariate Geostatistics to Macroprobe Mappings in Steels*. ISIJ International, 30, 529–534.
- [21] DELFINER P, DELHOMME JP & PELISSIER-COMBESCURE J (1983) *Application of geostatistical analysis to the evaluation of petroleum reservoirs with well logs*. 24th Annual Logging Symposium of the SPWLA, Calgary, June 27–30 1983.
- [22] DELFINER P & MATHERON G (1980) *Les fonctions aléatoires intrinsèques d'ordre k*. Publication C-84, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [23] DELHOMME JP (1979) *Réflexions sur la prise en compte simultanée des données de forages et des données sismiques*. Publication LHM/RC/79/41, Centre d'Informatique Géologique, Ecole des Mines de Paris, Fontainebleau.
- [24] DELHOMME JP, BOUCHER M, MEUNIER G & JENSEN F (1981) *Apport de la géostatistique à la description des stockages de gaz en aquifère*. Revue de l'Institut Français du Pétrole, 36, 309–327.
- [25] DONG A (1990) *Estimation Géostatistique des Phénomènes régis par des Equations aux Dérivées Partielles*. Thèse de Doctorat, Ecole des Mines de Paris, Fontainebleau.
- [26] DOWD PA (1989) *Generalized cross-covariances*. In: Armstrong M (ed) *Geostatistics*, 151–162, Kluwer Academic Publisher, Amsterdam, Holland.
- [27] FOUQUET C DE & MANDALLAZ D (1992) *Using geostatistics for forest inventory with air cover: an example*. Publication N-20/92/G, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau, 12p.
- [28] FRANÇOIS-BONGARÇON D (1981) *Les corégionalisations, le cokrigeage*. Publication C-86, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [29] GALLI A, GERDIL-NEUILLET F & DADOU C (1984) *Factorial kriging analysis: a substitute to spectral analysis of magnetic data*. In: Verly G et al (eds) *Geostatistics for Natural Resources Characterization*, 543–557, NATO ASI Series C-122, Reidel, Dordrecht, Hollande.
- [30] GALLI A & MEUNIER G (1987) *Study of a gas reservoir using the external drift method*. In: G Matheron et M Armstrong (eds) *Geostatistical Case Studies*, 105–119, Reidel, Dordrecht, Holland.
- [31] GIFI A (1990) *Nonlinear Multivariate Analysis*. Wiley, New York.
- [32] GITTINS R (1985) *Canonical Analysis: a Review with Applications in Ecology*. Biomathematics, Vol. 12, Springer-Verlag, Berlin, 351p.
- [33] GOOVAERTS P (1992) *Multivariate Geostatistical Tools for Studying Scale-Dependent Correlation Structures and Describing Space-Time Variations*. Doctoral Thesis, Université Catholique de Louvain, Louvain-la-Neuve, 233p.
- [34] GOOVAERTS P (1992) *Factorial kriging analysis: a useful tool for exploring the structure of multivariate spatial information*. Journal of Soil Science, 43, 597–619.
- [35] GOULARD M (1988) *Champs Spatiaux et Statistique Multidimensionnelle*. Thèse de Doctorat, Université des Sciences et Techniques du Languedoc, Montpellier.
- [36] GOULARD M (1989) *Inference in a coregionalization model*. In: Armstrong M (Ed.) *Geostatistics*, Kluwer Academic Publisher, Amsterdam, Holland, 397–408.
- [37] GOULARD M & VOLTZ M (1992) *Linear coregionalization model: tools for estimation and choice of cross-variogram matrix*. Mathematical Geology, 24, 269–286.
- [38] GREENACRE MJ (1984) *Theory and Applications of Correspondence Analysis*. Academic Press, London.
- [39] HASLETT J (1989) *Space time modelling in meteorology: a review*. Proceedings of 47th Session, International Statistical Institute, Paris, 229–246.
- [40] HASLETT J, BRADLEY R, CRAIG PS, WILLS G & UNWIN AR (1991) *Dynamic graphics for exploring spatial data, with application to locating global and local anomalies*. The American Statistician, 45, 234–242.

- [41] HUDSON G & WACKERNAGEL H (1992) *Kriging temperature in Scotland using the external drift method*. Publication N-2/92/G, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [42] JAQUET O (1989) *Factorial Kriging Analysis Applied to Geological Data from Petroleum Exploration*. *Mathematical Geology*, 21, 683–691.
- [43] JEULIN D & RENARD D (1991) *Practical limits of the deconvolution of images by kriging*. Publication N-13/91/G. Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [44] JOURNAL AG & HUIJBREGTS CJ (1978) *Mining Geostatistics*. Academic Press, London, 600p.
- [45] LAJAUNIE C (1984) *A geostatistical approach to air pollution modelling*. In: Verly G et al (eds) *Geostatistics for Natural Resources Characterization*, 877–891, NATO ASI Series C-122, Reidel, Dordrecht, Hollande.
- [46] LAJAUNIE C (1988) *Ajustement automatique de corégionalisation*. Note non publiée.
- [47] LAJAUNIE C & BÉJAOUI (1991) *Sur le krigeage des fonctions complexes*. Publication N-23/91/G, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [48] LAJAUNIE C & LANTUÉJOL C (1989) *Setting up the general methodology for discrete isofactorial models*. In: Armstrong M (Ed.) *Geostatistics*, Kluwer Academic Publisher, Amsterdam, Holland, 323–334.
- [49] LANTUÉJOL C (1990) *Cours de Sélectivité*. Publication C-140, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [50] LINDNER S & WACKERNAGEL H (1993) *Statistische Definition eines Lateritpanzer-Index für SPOT/Landsat-Bilder durch Redundanzanalyse mit bodengeochemischen Daten*. Publication N-8/93/G, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau, 11p.
- [51] LUMLEY JL (1970) *Stochastic Tools in Turbulence*. Academic Press, London.
- [52] MA YZ & ROYER JJ (1988) *Local geostatistical filtering: application to remote sensing*. *Sciences de la Terre, Série Informatique*, 27, 17–36.
- [53] MAGNUS JR & NEUDECKER H (1988) *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley, New York.
- [54] MARBEAU JP (1976) *Géostatistique Forestière*. Thèse de Doctorat, Université de Nancy, Nancy.
- [55] MARDIA KV, KENT JT & BIBBY JM (1979) *Multivariate Analysis*. Academic Press, London, 521p.
- [56] MARÉCHAL A (1970) *Cokrigeage et régression en corrélation intrinsèque*. Publication N-205, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [57] MARÉCHAL A (1984) *Kriging seismic data in presence of faults*. In: Verly G et al.(eds) *Geostatistics for Natural Resources Characterization*, 271–294, NATO ASI Series C-122, Reidel, Dordrecht, Hollande.
- [58] MATHERON G (1965) *Les Variables Régionalisées et leur Estimation*. Masson, Paris, 305p.
- [59] MATHERON G (1970) *La Théorie des Variables Régionalisées et ses Applications*. Fascicule 5, Les Cahiers du Centre de Morphologie Mathématique, Ecole des Mines de Paris, Fontainebleau, 211p.
- [60] MATHERON G (1971) *La Théorie des Fonctions Aléatoires Intrinsèques Généralisées*. Publication N-252, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [61] MATHERON G (1973) *The intrinsic random functions and their applications*. *Advances in Applied Probability*, 5, 439–468.
- [62] MATHERON G (1976) *A simple substitute for conditional expectation: the disjunctive kriging*. In: Guarascio M et al.(eds) *Advanced Geostatistics in the Mining Industry*, 221–236, NATO ASI Series C 24, Reidel, Dordrecht, Hollande.

- [63] MATHERON G (1979) *Recherche de simplification dans un problème de cokrigeage*. Publication N-628, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau, 19p.
- [64] MATHERON G (1982) *Pour une Analyse Krigeante des Données Régionalisées*. Publication N-732, Centre de Géostatistique, Fontainebleau, France, 22p.
- [65] MATHERON G (1989) *Two classes of isofactorial models*. In: Armstrong M (ed) *Geostatistics*, 309–322, Kluwer Academic Publisher, Amsterdam, Holland.
- [66] MATHERON G (1989) *Estimating and Choosing*. Springer-Verlag, Berlin, 141p.
- [67] MEIER S & KELLER W (1990) *Geostatistik*. Springer-Verlag, Wien.
- [68] MONESTIEZ P & SWITZER P (1991) *Semiparametric estimation of nonstationary spatial covariance models by multidimensional scaling*. Technical report No. 165, Stanford University, Stanford.
- [69] MYERS DE (1991) *Pseudo-cross variograms, positive-definiteness, and cokriging*. *Mathematical Geology*, 23, 805–816.
- [70] OLIVER MA & WEBSTER R (1989) *A geostatistical basis for spatial weighting in multivariate classification*. *Mathematical Geology*, 21, 15–35.
- [71] ORFEUIL JP (1973) *L'Analyse des Données*. Publication N-353, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [72] ORFEUIL JP (1977) *Etude, mise en oeuvre et test d'un modèle de prédiction à court terme de pollution atmosphérique*. Rapport CECA, Publication N-498, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [73] PRIESTLEY MB (1981) *Spectral Analysis and Time Series*. Academic Press, London.
- [74] RAO CR (1964) *The use and interpretation of principal component analysis in applied research*. *Sankhya, Series A*, 329–358.
- [75] RAO CR (1973) *Linear Statistical Inference and its Applications*. 2nd edition, Wiley, New York, 625p.
- [76] RENARD D & NAI-HSIEN M (1988) *Utilisation de dérivées externes multiples*. *Sciences de la Terre, Série Informatique*, 28, 281–301.
- [77] RIVOIRARD J (1988) *Modèles à résidus d'indicatrices autokrigeables*. *Sciences de la Terre, Série Informatique*, 28, 303–326.
- [78] RIVOIRARD J (1990) *Introduction to Disjunctive Kriging and Nonlinear Geostatistics*. Publication C-143, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [79] ROUHANI S & WACKERNAGEL H (1990) *Multivariate geostatistical approach to space-time data analysis*. *Water Resources Research*, 26, 585–591.
- [80] ROYER JJ (1988) *Analyse Multivariable et Filtrage des Données Régionalisées*. Thèse de Doctorat, Institut National Polytechnique de Lorraine, Nancy.
- [81] SAMPSON P & GUTTORP (1992) *Nonparametric estimation of nonstationary spatial structure*. *Journal of the American Statistical Association*, 87, 108–119.
- [82] SANDJIVY L (1983) *Analyse krigeante de données géochimiques*. *Sciences de la Terre, Série Informatique*, 18, 141–172.
- [83] SANDJIVY L (1984) *The factorial kriging analysis of regionalized data*. In: Verly G et al.(eds) *Geostatistics for Natural Resources Characterization*, 559–571, NATO ASI Series C-122, Reidel, Dordrecht, Hollande.
- [84] SANDJIVY L (1987) *Analyse krigeante des données de prospection géochimique*. Thèse de Doctorat, Ecole des Mines de Paris, Fontainebleau.
- [85] SAPORTA G (1990) *Probabilités, Analyse des Données et Statistique*. Technip, Paris, 493p.
- [86] SCHULZ-OHLBERG (1989) *Die Anwendung geostatistischer Verfahren zur Interpretation von gravimetrischen und magnetischen Felddaten*. *Wissenschaftlich-Technische Berichte*, 1989–6, Deutsches Hydrographisches Institut, Hamburg.
- [87] SÉGURET S (1991) *Géostatistique des Phénomènes à Tendance Périodique (dans l'espace-temps)*. Thèse de Doctorat, Ecole des Mines de Paris, Fontainebleau.

- [88] SÉGURET S & CHAUVET P (1988) *Pour une méthodologie de déconvolution de variogrammes*. Publication N-51/88/G, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.
- [89] SÉGURET S & HUCHON P (1990) *Trigonometric kriging: a new method for removing the diurnal variation from geomagnetic data*. Journal of Geophysical Research, Vol. 32, No. B13, 21.383–21.397.
- [90] SERRA J (1968) *Les structures gigognes: morphologie mathématique et interprétation métallogénique*. Mineral. Deposita, 3, 135–154.
- [91] SOUSA AJ (1989) *Geostatistical data analysis: an application to ore typology*. In: Armstrong M (ed) Geostatistics, 851–860, Kluwer Academic Publisher, Amsterdam, Holland.
- [92] STEIN A (1991) *Spatial Interpolation*. Doctoral thesis, Agricultural University, Wageningen, Netherlands.
- [93] STEIN M (1986) *A simple model for spatial-temporal processes*. Water Resources Research, Vol. 22, 2107–2110.
- [94] STRANG G (1984) *Linear Algebra and its Applications*. Academic Press, 2e édition.
- [95] TYLER DE (1982) *On the optimality of the simultaneous redundancy transformations*. Psychometrika, 47, 77–86.
- [96] VOLLE M (1985) *Analyse des Données*. 3e édition, Economica, Paris.
- [97] WACKERNAGEL H (1985) *L'inférence d'un Modèle Linéaire en Géostatistique Multivariable*. Thèse de Doctorat, Ecole des Mines de Paris, Fontainebleau.
- [98] WACKERNAGEL H (1988) *Geostatistical techniques for interpreting multivariate spatial information*. In: Chung CF et al.(eds) Quantitative Analysis of Mineral and Energy Resources, 393–409, NATO ASI Series C 223, Reidel, Dordrecht, Hollande.
- [99] WACKERNAGEL H (1991) *Changement de maille et changement de support*. In: Mullon C (ed) SEMINFOR 4: Le Transfert d'Echelle, 159–171, Editions de l'Orstom, Paris.
- [100] WACKERNAGEL H & BUTENUTH C (1989) *Caractérisation d'anomalies géochimiques par la géostatistique multivariable*. J. Geochem. Explor., 32, 437–444.
- [101] WACKERNAGEL H, PETITGAS P & TOUFFAIT Y (1989) *Overview of methods for coregionalization analysis*. In: Armstrong M (ed) Geostatistics, 409–420, Kluwer Academic Publisher, Amsterdam, Holland.
- [102] WACKERNAGEL H & SANGUINETTI H (1992) *Gold prospecting with factorial cokriging in the Limousin, France*. In: Davis JC & Herzfeld UC (eds) Computers in Geology: 25 years of progres. Studies in Mathematical Geology, Vol. 5, 33–43, Oxford University Press.
- [103] WACKERNAGEL H, WEBSTER R & OLIVER MA (1988) *A geostatistical method for segmenting multivariate sequences of soil data*. In: Bock HH (ed) Classification and Related Methods of Data Analysis, 641–650, Elsevier (North-Holland), Amsterdam.
- [104] WEBSTER R & OLIVER MA (1990) *Statistical Methods in Soil and Land Resource Survey*. Oxford University Press, Oxford.
- [105] YAGLOM AM (1962) *An Introduction to the Theory of Stationary Random Functions*. Dover, New York.
- [106] YAGLOM AM (1986) *Correlation Theory of Stationary and Related Random Functions*. Springer-Verlag, Berlin.