Table des matières

Notations et conventions

- $cste$ signifie constante.
- $v.a.r.$ signifie variable aléatoire réelle.
- $i.i.d.$ signifie indépendantes et identiquement distribuées.
- P_0 signifie loi de probabilité sous l'hypothèse **H0**.
- Soient X et Y deux $v.a.r.$: $X \overset{loi}{\sim} Y$ signifie que X a pour loi celle de Y.
- Soient (F_T) et (G_T) deux suites de fonctions de répartition sur $(\mathbb{R}, \mathcal{R})$, où $\mathcal{R}$ est la $\sigma-$algèbre borélienne de $\mathbb{R}$. On dit que (F_T) et (G_T) sont asymptotiquement équivalentes (noté : $F_T \underset{T \to +\infty}{\overset{loi}{\sim}} G_T$) si pour tout x dans $\mathbb{R}$:

$$\lim_{T \to +\infty} (F_T(x) - G_T(x)) = 0$$

- λ_n, $n \in \mathbb{N}^*$, désigne la mesure de Lebesgue sur $\mathbb{R}^n$.
- δ_x désigne la mesure masse de Dirac au point x.
- Soit μ une mesure positive et bornée sur $\mathbb{R}^n$, $\mu^{\otimes^2}$ désigne la mesure produit de μ par μ et ϕ_μ désigne la fonction caractéristique de μ.
- On dit qu'une suite de variables aléatoires $\{X_n, n \geq 1\}$ converge presque complètement ($p.c.$) vers une constante M si, pour tout $\epsilon > 0$:

$$\sum_{n=1}^{+\infty} P\left[|X_n - M| > \epsilon\right] < +\infty$$

- $p.s.$ signifie presque sûre.
- $i.s.$ signifie infiniment souvent.

- Sur $(\mathbb{R}^n, \mathcal{R}^n)$, où $\mathcal{R}^n$ est la $\sigma-$algèbre borélienne de $\mathbb{R}^n$, on dit que la mesure μ_n converge étroitement (noté : *c.e.*) vers μ si, pour toute fonction f continue bornée de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ (ou $\mathbb{C}$) :

$$\int f d\mu_n \underset{n\to+\infty}{\longrightarrow} \int f d\mu$$

- E signifie espérance.

- Var signifie variance.

- Cov signifie covariance.

- Cor signifie corrélation.

- $\chi^2(p)$ représente la loi du $khi - deux$ à p degrés de liberté.

- $\sum_{k=1}^{n} \lambda_k \chi_k^2(1)$, où $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ sont des réels strictement positifs et les $\chi_k^2(1)$, $k \in [1, \ldots, n]$, sont des $khi - deux$ indépendants à un degré de liberté, représente "une somme pondérée de $khi - deux$ indépendants" (*cf.* annexe B).

- $\mathcal{N}_p(m, \Sigma)$ représente la loi normale de dimension p, de moyenne m et de matrice de variances-covariances Σ.

- $\|.\|_n$, représente la norme euclidienne dans $\mathbb{R}^n$.

- $\mathbb{1}_n, n \in \mathbb{N}^*$ désigne le vecteur de dimension n, composé uniquement de 1.

- $I_n, n \in \mathbb{N}^*$ désigne la matrice identité de rang n.

- $A = [(A)_{i,j}]_{n,m}$ désigne une matrice à n lignes et m colonnes de terme général $(A)_{i,j}$ situé à la ligne i et à la colonne j.

- On munit l'espace vectoriel réel des matrices $A = [(A)_{i,j}]_{n,m}$ de la norme sup matricielle :

$$\|A\|_{sup} = \sup_{i,j} |(A)_{i,j}|$$

- A' désigne la transposée de la matrice A.

- $tr(A)$ désigne la trace de la matrice carrée A et $|A|$ son déterminant.

- Soit A une matrice symétrique, définie positive et sa décomposition $A = PDP'$, où P est orthogonale et D est diagonale ; on note $A^{1/2}$ la racine carrée $PD^{1/2}P'$.

- $\|.\|_\infty$ désigne la norme (de la convergence) uniforme.

- $C^k([0,1]^n)$, $k \in \mathbb{N}$ et $n \in \mathbb{N}^*$, désigne l'espace vectoriel de toutes les fonctions de classe C^k, ou k fois continuement différentiables, de $[0,1]^n$ à valeurs dans $\mathbb{R}$, muni de la norme uniforme ; $k = 0$ signifie qu'elle est seulement continue.

- $L^2([0,1])$ désigne l'espace vectoriel des fonctions à valeurs réelles, de carré intégrable sur $[0,1]$ par rapport à la mesure de Lebesgue.

- $L^2_{\mathbb{C}}(\Omega, \mu)$ désigne l'espace vectoriel des fonctions à valeurs complexes, de carré intégrable sur Ω par rapport à la mesure μ.

- f^{-1} représente la fonction réciproque, quand elle existe, de la fonction f.

- $f(., x')$ désigne l'application partielle $x \longmapsto f(x, x')$ de la fonction $f : x, x' \longmapsto f(x, x')$.

- Une fonction f, de $\mathbb{R}$ dans $\mathbb{R}$, est de classe C^k par morceaux dans $[0,1]$, $k \in \mathbb{N}$, s'il existe une subdivision $0 = x_0 < x_1 < \cdots < x_n = 1$, $n \in \mathbb{N}^*$, telle que la restriction de f à $]x_{i-1}, x_i[$ soit de classe C^k et admette un prolongement par continuité de classe C^k dans $[x_{i-1}, x_i]$, pour tout i dans $[1, \ldots, n]$.

- Pour une fonction f à valeurs dans $\mathbb{R}$, deux fois différentiables en $x \in \mathbb{R}^n$, on note $f^{(1)}(x)$ le vecteur des dérivées premières en x, $f^{(2)}(x)$ la matrice des dérivées secondes en x. Pour une fonction f à valeurs dans $\mathbb{R}^m$, $m \geq 2$, $f^{(1)}(x)$ représente la matrice jacobienne de f en x.

- Pour une fonction f de $\mathbb{R}$ dans $\mathbb{R}$, $f_d^{(1)}(x)$ désigne la dérivée à droite, quand elle existe, de f en x.

- La fonction :
$$x, x' \longmapsto g^{(m,m')}(x, x'), \text{ avec } (m, m') \in \mathbb{N}^2$$
représente la fonction dérivée partielle d'ordres m et m' par rapport respectivement à x et x' :
$$g^{(m,m')}(x, x') = \frac{\partial^{m+m'} g}{\partial x^m \partial x'^{m'}}(x, x')$$

- Nous définissons les quantités :
$$\begin{aligned} g^{(m,m')}(x, x^-) &= \lim_{x' \uparrow x} g^{(m,m')}(x, x') \\ g^{(m,m')}(x, x^+) &= \lim_{x' \downarrow x} g^{(m,m')}(x, x') \end{aligned}$$
quand ces limites existent, et où $x' \uparrow x$ (respectivement $x' \downarrow x$) signifie $x' \to x$ et $x' < x$ (respectivement $x' \to x$ et $x' > x$).

- $o(u)$ représente un infiniment petit par rapport à u au voisinage de 0.
- Soit $\gamma > 0$, la suite :
$$u_n = O(n^{-\gamma}),\ n \to +\infty$$
signifie qu'il existe une constante positive c telle que :
$$n^{\gamma}|u_n| \leq c,\ n \to +\infty$$
et
$$v_n = o(n^{-\gamma}),\ n \to +\infty$$
se traduit par :
$$\lim_{n\to+\infty} n^{\gamma} v_n = 0$$
- $u_n \sim v_n$ signifient que les deux suites u_n et v_n sont du même ordre, $n \to +\infty$.
- $\mathbb{1}_{[a,b]}(.)$, $a < b$ est la fonction définie par :
$$\mathbb{1}_{[a,b]}(x) = \begin{cases} 1 & \text{si } a \leq x \leq b \\ 0 & \text{sinon} \end{cases}$$
- Si E est un ensemble de $\mathbb{R}^n$, $\mathring{E}$ désigne l'intérieur de E, $\bar{E}$ désigne sa fermeture.
- $Card(E)$ désigne le cardinal d'un ensemble E.
- Si B est un ensemble inclus dans un ensemble Ω, B^c désigne le complémentaire de B dans Ω.
- Si B et E sont deux ensembles :
$$E \setminus B = \{x : x \in E, x \notin B\}$$
- $a \vee b$ représente le maximum de a et b, et $a \wedge b$ représente le minimum.
- Si x est un nombre réel, $\lfloor x \rfloor$ représente l'entier immédiatement inférieur à x, c'est-à-dire sa partie entière, $\lceil x \rceil$ l'entier immédiatement supérieur.
- On note $\sup\limits_{[0,1]}$ pour $\sup\limits_{x\in[0,1]}$ et $\sup\limits_{[0,1]^2}$ pour $\sup\limits_{(x,x')\in[0,1]^2}$.
- Sauf mention expresse, $x \leq x'$ signifie $0 \leq x \leq x' \leq 1$ dans le contexte.
- Sauf mention expresse, $i \leq j$ signifie $1 \leq i \leq j \leq n$ dans le contexte.
- Si W est un nombre complexe, $\bar{W}$ représente son conjugué et $|W|$ son module.
- L'énoncé d'une propriété, d'un lemme, d'un théorème ou d'un corollaire se termine par $\bigstar$, une démonstration par $\blacksquare$, une définition par $\spadesuit$ et une remarque par $\clubsuit$.

Introduction

Nous présentons d'abord la motivation du sujet et le contexte originel dans lequel elle est apparue. Dans un second temps, nous formalisons le sujet et nous explicitons l'objet de cette thèse. En dernier lieu, nous présentons les résultats obtenus, les méthodes utilisées et les exemples traités.

La motivation du sujet

Dans l'étude statistique des phénomènes spatio-temporels environnementaux (phénomènes qui varient dans le temps et dans l'espace géographique) on suppose très souvent que le phénomène est homogène (stationnaire) sur l'espace tout entier. Autrement dit :

1. le phénomène est constant en moyenne (en espérance) sur l'espace ;
2. les covariances du phénomène entre deux points du domaine ne dépendent que de la distance entre ces points.

De la sorte, on limite la complexité de la composante spatiale du phénomène analysé. En conséquence, cette "modélisation" n'est qu'une représentation simplifiée de la réalité du phénomène. Le recours à cette hypothèse de stationnarité est soit la conséquence d'une hypothèse *a priori* souvent difficile à tester, soit la reconnaissance de l'impossibilité de trouver le "vrai" mécanisme probabiliste engendrant les observations. En fait, les problèmes relevant des phénomènes environnementaux présentent souvent de très fortes hétérogénéités spatiales (une non-stationnarité) évidentes (Guttorp et Sampson). Par exemple, la covariance entre deux séries d'observations, portant sur des données pluviométriques, mesurées sur un couple de stations, va dépendre de la manière dont ce couple est orienté par rapport aux systèmes météorologiques dominants ; les covariances sont plus fortes si l'orientation des stations est dans la direction du passage des systèmes nuageux, et plus faibles dans la direction perpendiculaire (Monestiez, Sampson, Guttorp).

Pour ce type de structures "spatialement non-stationnaires", et si l'on veut faire des prédictions en des sites non observés, on ne peut plus utiliser les méthodes classiques. Ces dernières, tel le krigeage simple (Matheron), supposent en effet une structure de covariance stationnaire et une espérance constante connue. Diverses méthodes alternatives, le krigeage universel et les fonctions aléatoires intrinsèques d'ordre k (FAI-k) (Matheron, Journel et Huijbregts), ont alors été développées quand l'hypothèse de stationnarité n'a aucune raison d'être. Mais ces méthodes consistent à ne tenir compte de la non-stationnarité qu'au travers d'une moyenne (dérive) qu'il s'agit de filtrer afin de se ramener à un "résidu" supposé stationnaire.

Or, dans de nombreux cas, ce résidu n'est pas stationnaire et la non-stationnarité ne peut pas être modélisée au travers d'une moyenne. Si filtrer une dérive permet toujours d'obtenir une estimation relativement bonne, c'est-à-dire non biaisée, de la moyenne en un point, conditionnellement à l'information dont on dispose en d'autres points, l'estimation de la variance du prédicteur pourra être mauvaise si cette covariance est mal modélisée. Et dans le cadre de l'amélioration d'un réseau de stations de mesure, le choix de nouvelles stations améliorant le réseau risque d'être erroné si le modèle choisi *a priori* est inexact.

Précisons néanmoins, que ces méthodes géostatistiques (krigeage universel et FAI-k) ont été développées dans le contexte minier et géologique lorsqu'on dispose d'une seule réalisation du phénomène étudié (Matheron).

Par contre, dans l'étude des phénomènes environnementaux, nous avons à notre disposition plusieurs réalisations dans le temps qui en permettent une modélisation plus riche (Monestiez et Switzer).

Pour prendre en compte ces structures spatialement non-stationnaires, une alternative originale a été développée. Cette approche (Sampson (1986)) consiste à modéliser la non-stationnarité en réduisant la covariance non-stationnaire en une covariance stationnaire grâce à une anamorphose, c'est-à-dire une transformation de l'espace géographique. Depuis, cette méthode a été développée, formalisée et appliquée à l'analyse de nombreux phénomènes environnementaux (*cf.* par exemple Sampson et Guttorp, Mardia et Goodall, Monestiez et *al.*, Meiring).

Cette nouvelle approche est l'objet principal de cette thèse, à savoir, **l'étude de modèles de covariance non-stationnaires par déformation de l'espace et de leur statistique.**

Formalisation du sujet

La modélisation d'une covariance d'un processus non-stationnaire par déformation de l'espace peut être formalisée de la manière suivante :

- soit $\left\{Z_t(x), t \in \mathbb{N}, x \in G \subset \mathbb{R}^d, d \geq 1\right\}$, une suite temporelle telle que pour tout t, $\{Z_t(x), x \in G\}$ est un champ aléatoire non-stationnaire rendant compte du phénomène sous-jacent étudié. Dans tout notre travail, on suppose les répétitions temporelles *i.i.d.* ; autrement dit, pour tout $T \in \mathbb{N}^*$, $(Z_1, Z_2, \ldots, Z_T)$ constitue un échantillon *i.i.d.* du champ aléatoire parent Z. On suppose que Z possède des moments au moins jusqu'à l'ordre 2 et on note r_Z sa fonction de covariance ;

- soit $\{Y_t(u), t \in \mathbb{N}, u \in D \subset \mathbb{R}^d\}$, un processus tel que pour tout t, $\{Y_t(u), u \in D\}$ est un champ aléatoire stationnaire isotrope de fonction de covariance $r_Y(u, u') = R(\|u' - u\|_d)$. On suppose que $(Y_1, Y_2, \ldots, Y_T)$ constitue un échantillon *i.i.d.* du champ aléatoire parent Y ;
- soit f une transformation, ou déformation, déterministe bijective et bicontinue de G (espace géographique) sur D (espace déformé).

Le modèle est le suivant :

$$\{Z(x), x \in G\} = \{Y(f(x)), x \in G\}$$

qui peut aussi s'écrire de façon équivalente :

$$\{Z(f^{-1}(u)), u \in D\} = \{Y(u), u \in D\}$$

Pour ce modèle, on a pour tout (x, x') dans G^2 :

$$\begin{array}{rcccl} r_Z(x, x') & = & Cov\,(Z(x), Z(x')) & = & Cov\,(Y(f(x)), Y(f(x'))) \\ & = & r_Y\,(f(x'), f(x)) & = & R\,(\|f(x') - f(x)\|_d) \end{array}$$

Tout problème faisant intervenir le champ observé Z est donc transposé, par la déformation f, à Y, champ stationnaire pour lequel les méthodes usuelles, comme le krigeage simple, s'appliquent. Les résultats obtenus sur Y se transposent ensuite à Z, par la déformation inverse f^{-1}.

Dans ce travail, nous allons étudier la situation $d = 1$. La raison fondamentale est qu'il est beaucoup plus simple de décrire les transformations bijectives sur des parties de $\mathbb{R}$ que sur des parties de $\mathbb{R}^d$, $d \geq 2$. Sur $\mathbb{R}$, la classe des transformations bijectives et bicontinues est celle des fonctions strictement monotones que l'on sait caractériser complètement. Il est alors possible de définir des familles de modèles non-paramétriques de transformation et de proposer des estimateurs fonctionnels. En particulier, nos résultats pourront être appliqués aux séries chronologiques, où la déformation du support d'un processus peut s'avérer utile pour rendre compte de non-linéarités (Stock). Enfin, une étude complète sur $\mathbb{R}$ sera d'une grande utilité pour les dimensions supérieures. Comme on le verra, certaines analyses resteront inchangées sur $\mathbb{R}^2$. Mais de nouveaux problèmes, non résolus pour le moment, surgiront, comme la description paramétrique, ou non-paramétrique, de transformations bijectives de $\mathbb{R}^2$ dans $\mathbb{R}^2$.

Dans Meiring (1995), on trouve une synthèse détaillée de la méthode ainsi que des développements récents et originaux. Toutefois, aucune étude ne s'est intéressée de près à la caractérisation des fonctions de covariance réductibles à une fonction de covariance stationnaire *via* une transformation bijective de l'espace. Toute fonction de covariance non-stationnaire est-elle réductible à une covariance stationnaire ? D'autre part, si des méthodes d'estimation ont été proposées, les propriétés des estimateurs n'ont été analysées qu'au travers de simulations. C'est pourquoi, notre travail s'attache principalement à l'étude des qualités des estimateurs obtenus, ceci dans divers cadres asymptotiques, pour les différents types de convergence, *p.s*, en probabilité et en loi, et pour des modèles paramétriques ou non-paramétriques. On illustre ces résultats par des exemples.

Nous nous plaçons dans le cadre suivant :

1. on dispose d'un échantillon du processus aléatoire Z ; cette hypothèse simplificatrice permet d'alléger la présentation des résultats et les techniques de démonstration. Mais la plupart des propriétés que nous allons montrer s'étendent au cas où le processus Z est ergodique dans le temps et faiblement dépendant. Il suffit alors d'utiliser les résultats asymptotiques classiques (loi forte des grands nombres, théorème limite centrale) pour les processus faiblement dépendants, par exemple pour les processus mélangeants (Doukhan, Guyon) ;

2. nous supposons également que Z est de variance constante, c'est-à-dire, pour tout x dans $[0,1]$:
$$r_Z(x,x) = \sigma_0^2 > 0$$
Dans le cas d'un modèle paramétrique, nous montrons comment cette hypothèse peut être facilement relaxée ;

3. enfin, nous nous limitons à la situation $G = [0,1]$, $D = [0,1]$ et aux déformations bijectives, bicontinues et strictement croissantes de $[0,1]$ sur $[0,1]$ ($f(0) = 0$ et $f(1) = 1$). Ces restrictions ne font perdre aucune généralité. Si f est une transformation bijective de $[0,1]$ sur $[a,b]$ telle que $Y(u)$ soit stationnaire au second ordre, le processus $\{Y'(v)\} = \{Y(a + v(b-a))\}, v \in [0,1]$, est aussi stationnaire au second ordre. Si f est strictement décroissante, il suffira de mener l'étude avec $-f$.

Présentation des chapitres

Dans le **chapitre 1**, après quelques rappels sur les fonctions de covariance, nous montrons qu'une covariance non-stationnaire ne peut pas se réduire, en général, à une covariance stationnaire *via* une déformation bijective et bicontinue de l'espace. Les trois exemples sont les suivants :

1. $Z = \{Z(x)\}$ est un processus défini sur $[0,1]$ selon :
$$Z(x) = \epsilon_1 \mathbb{1}_{[0,\frac{1}{2}]}(x) + \epsilon_2 \mathbb{1}_{]\frac{1}{2},1]}(x)$$
où ϵ_1 et ϵ_2 sont deux *v.a.r.* indépendantes et de variance 1 ;

2. (Assouad (1980)) $Z = \{Z(x), x \in [0,1]\}$ est un processus continu en moyenne quadratique, dont la covariance vérifie, pour tout $1 \leq i, j \leq 6$:

$$r_Z\left(\frac{i-1}{5}, \frac{j-1}{5}\right) = < X_i, X_j >_3$$

où $X_1, X_2, \ldots, X_6$ sont 6 vecteurs unitaires de $\mathbb{R}^3$ (muni du produit scalaire canonique $< ., . >_3$) liés par les 3 relations linéaires suivantes :

$$\begin{cases} X_1 + X_2 &= 2X_6 \cos\alpha \\ X_2 + X_3 &= 2X_4 \cos\alpha \\ X_1 + X_3 &= 2X_5 \cos\alpha \end{cases}$$

avec $\alpha \in \left]\frac{\pi}{3}; \frac{\pi}{2}\right[$;

3. $Z = \{Z(s), s \in [0,1]\}$ est défini à partir de $P = \{P(x,y)\}$, champ stationnaire isotrope sur $\mathbb{R}^2$, comme suit :

$$Z(s) = P(x,y), \text{ où } \begin{cases} s = \frac{1}{2} - y, \text{ si } x = 0 \text{ et } y \in [0, \frac{1}{2}] \\ \\ s = \frac{1}{2} + x, \text{ si } y = 0 \text{ et } x \in]0, \frac{1}{2}] \end{cases}$$

Pour ces trois exemples, on montre qu'il n'existe pas d'homéomorphisme f de $[0,1]$ sur $[0,1]$, vérifiant $f(0) = 0$ et $f(1) = 1$, et de processus $Y = \{Y(u)\}$ stationnaire sur $[0,1]$, de fonction de covariance R, tels que, pour tout u dans $[0,1]$:

$$Y(u) = Z(f^{-1}(u))$$

Par conséquent, la covariance r_Z de Z ne peut pas se réduire à une covariance stationnaire *via* un homéomorphisme croissant de $[0,1]$ sur $[0,1]$, autrement dit, il n'existe pas de couple (R, f) où R est une covariance stationnaire et f un homéomorphisme croissant de $[0,1]$ sur $[0,1]$ tels que :

$$r_Z(x, x') = R(f(x') - f(x))$$

Ensuite, nous établissons une condition nécessaire et suffisante caractérisant les fonctions de covariance réductibles. Plus précisément :

- soit r une fonction de covariance continue dans $[0,1]^2$ et différentiable dans $\{(x,x') \in [0,1]^2, x \neq x'\}$
- soit f un difféomorphisme croissant de $[0,1]$ sur $[0,1]$ ($f(0) = 0$ et $f(1) = 1$)

alors, il existe une fonction R , telle que, pour tout (x, x') dans $[0,1]^2$:

$$r(x,x') = R\left(f(x') - f(x)\right)$$

si et seulement si, pour tout $x \neq x'$:

$$f^{(1)}(x')r^{(1,0)}(x,x') + f^{(1)}(x)r^{(0,1)}(x,x') = 0$$

Ce résultat nous permet de plus d'affirmer qu'il y a unicité de la déformation f et de la covariance stationnaire R, c'est-à-dire identifiabilité du modèle.

Dans les chapitres 2, 3 et 4, nous allons nous intéresser à l'estimation de la déformation bijective de $[0,1]$ sur $[0,1]$ f_0 et de la covariance stationnaire R_0 associées au modèle de covariance r_z :

$$r_Z(x,x') = R_0(f_0(x') - f_0(x)),\ (x,x') \in [0,1]^2$$

dans différentes situations : cadre paramétrique (chapitre 2) ; déformation non-paramétrique (chapitre 3) ; observation d'une seule réalisation et cas gaussien (chapitre 4).

Dans les chapitres 2 et 3, on suppose que l'on dispose de réalisations indépendantes du processus. Dans ce contexte expérimental, la seule hypothèse nécessaire sur le processus est l'existence d'un moment d'ordre 2ν, $\nu > 2$.

Dans le **chapitre 2**, la covariance stationnaire $R_0 = R_{\beta_0}$ appartient à une famille paramétrique de covariances stationnaires continues $\mathcal{R} = \{R_\beta\}$, $\beta \in \mathbb{R}^q$, $q \in \mathbb{N}^*$, et la déformation $f_0 = f(.;\alpha_0)$ appartient à une famille paramétrique d'homéomorphismes croissants de $[0,1]$ sur $[0,1]$ $\mathcal{F} = \{f(.;\alpha)\}$, $\alpha \in \mathbb{R}^p$, $p \in \mathbb{N}^*$.

On suppose que Z est de moyenne M_0 inconnue mais constante, de variance σ_0^2 inconnue mais constante et vérifie l'hypothèse **M1** ou **M3** selon le résultat étudié :

- **M1** : $\sup\limits_{[0,1]} E|Z(x)|^4 < +\infty$
- **M3** : $\exists \nu > 2, \sup\limits_{[0,1]} E|Z(x)|^{2\nu} < +\infty$

On dispose de T, $T \in \mathbb{N}^*$, réalisations *i.i.d.* du processus Z en un nombre fini et fixé n de stations positionnées en $x_1 < x_2 < \cdots < x_n$ sur $[0,1]$. Ces positions sont les mêmes quelque soit la réalisation $t \in [1,\ldots,T]$. L'asymptotique porte uniquement sur le nombre T de réalisations.

On pose $\theta = (\alpha, \beta)$ et on prend pour estimateur du vrai paramètre du modèle $\theta_0 = (\alpha_0, \beta_0)$, celui du minimum de contraste (*cf.* Dacunha-Castelle et Duflo (1993)) :

$$\hat{\theta}_T = \arg\min_\theta \sum_{i \leq j} \left[\hat{r}_T(x_i,x_j) - r(x_i,x_j;\theta)\right]^2$$

avec, pour tout $1 \leq i \leq j \leq n$:

$$\left\{ \begin{array}{rcl} \hat{r}_T(x_i, x_j) & = & \dfrac{1}{T-1} \displaystyle\sum_{t=1}^{T} \left(Z_t(x_i) - \bar{Z}_T(x_i)\right) \left(Z_t(x_j) - \bar{Z}_T(x_j)\right) \\ r(x_i, x_j; \theta) & = & R\left(f(x_j; \alpha) - f(x_i; \alpha); \beta\right) \end{array} \right.$$

où, pour tout $1 \leq i \leq n$:

$$\bar{Z}_T(x_i) = \frac{1}{T} \sum_{t=1}^{T} Z_t(x_i)$$

On étudie tout d'abord les propriétés asymptotiques de l'estimateur empirique de la covariance $\hat{r}_T(x_i, x_j)$: les convergences *p.s.* et en loi, et la loi du logarithme itéré.

Grâce à ces propriétés, on obtient ensuite sous l'hypothèse **M1** ou **M3** et des hypothèses de compacité de l'espace des paramètres, d'identifiabilité paramétrique du modèle et de régularité du modèle, les résultats suivants :

- les convergences *p.s.*, en probabilité et en loi de l'estimateur $\hat{\theta}_T$;
- des propriétés asymptotiques à la base de tests statistiques basés sur des sommes pondérées de $khi - deux$ indépendants :
    1. test de différence de contraste d'une hypothèse complètement spécifiée ;
    2. test de différence de contraste d'une sous-hypothèse régulière ;
    3. test d'adéquation du modèle.

On obtient aussi un résultat d'identification *p.s.* du modèle par contraste pénalisé (*cf.* Guyon (1993)).

On illustre ces résultats à l'aide d'un exemple, où on prend la famille de modèles de déformation puissance $\mathcal{F} = \{x^\alpha, \alpha > 0\}$ et la famille de modèles de covariance exponentielle $\mathcal{R} = \{e^{-\beta|u|}, \beta > 0\}$.

Enfin, on s'intéresse à la généralisation de toute cette étude, d'une part au cas où la variance de Z n'est pas constante, d'autre part au cas où Z est défini sur une partie G de $\mathbb{R}^d, d \geq 2$, et où la déformation, rendant ce processus stationnaire isotrope, est un homéomorphisme de G sur une partie D de $\mathbb{R}^d$.

Pour la généralisation au cas spatial, la difficulté et la spécificité de l'étude sur $\mathbb{R}^d$, $d \geq 2$, résident uniquement dans la description des bijections de $\mathbb{R}^d$ dans $\mathbb{R}^d$, problème qui est largement ouvert.

Dans le **chapitre 3**, le paramètre β_0 de la covariance stationnaire est toujours de dimension finie. Mais à la différence du chapitre 2, la déformation f_0 appartient à la famille non-paramétrique des homéomorphismes croissants f de $[0,1]$ sur $[0,1]$, continuement dérivables par morceaux et tels que leur dérivée à droite vérifie pour tout x dans $[0,1[$:

$$0 < f_d^{(1)}(x) \leq K$$

où K, indépendante de f, est une constante fixée (nécessairement $K \geq 1$).

On suppose que Z est centré, de variance σ_0^2 inconnue mais constante, continu en moyenne quadratique et vérifie l'hypothèse **M3**.

Pour estimer le paramètre fonctionnel f_0 sur tout $[0,1]$ à partir des observations sur Z, les points d'observation, ou stations, doivent être "denses" dans $[0,1]$. On dispose de T, $T \in \mathbb{N}^*$, réalisations indépendantes de Z. Mais au lieu d'être observées en un nombre fini et fixé de stations, ces T réalisations sont observées sur un nombre fini et croissant de stations qui tendent à densifier $[0,1]$. Dans ce contexte, nous proposons le schéma expérimental "triangulaire" suivant :

1. de nouvelles stations sont ajoutées, en supplément des précédentes, aux instants $T_k \in \mathbb{N}^*$, $k = 1, 2, \ldots, m, \ldots$:

$$T_1 < T_2 \cdots < T_m \leq T < T_{m+1} < \cdots$$

2. si on note $n_k, k \geq 1$, le nombre de stations présentes en T_k, et $\delta_k = n_k - n_{k-1} > 0$ le nombre de stations supplémentaires introduites en T_k, entre T_{k-1} et T_k il y a $(T_k - T_{k-1} + 1)$ répétitions *i.i.d.* des n_{k-1} observations de Z aux stations présentes en T_{k-1} ;

3. le nombre de stations n_m tend vers $+\infty$ avec m, de telle sorte que les positions des stations densifient $[0,1]$.

Dans toutes les études asymptotiques, c'est m qu'on fera tendre vers $+\infty$.

Dans ce contexte général où l'espace des paramètres $\Theta = \{\theta = (f, \beta)\}$, ici de dimension infinie, est relativement compact, on cherche un ϵ_m−estimateur du minimum de contraste (*cf.* Birgé et Massart (1993)) que l'on note $\tilde{\theta}_m = \left(\tilde{f}_m, \tilde{\beta}_m\right)$, c'est-à-dire un estimateur vérifiant pour tout θ dans Θ :

$$\int_{[0,1]^2} \left[\tilde{r}_m(x,x') - r_{\tilde{\theta}_m}(x,x')\right]^2 \mu_m^{\otimes^2}(dxdx') \leq \int_{[0,1]^2} \left[\tilde{r}_m(x,x') - r_\theta(x,x')\right]^2 \mu_m^{\otimes^2}(dxdx') + \epsilon_m$$

avec $r_\theta(x,x') = R(f(x') - f(x); \beta)$ et où la suite (ϵ_m) tend vers 0, $m \to +\infty$. $\int_{[0,1]^2} \left[\tilde{r}_m(x,x') - r_\theta(x,x')\right]^2 \mu_m^{\otimes^2}(dxdx')$ est le processus de contraste, $\mu_m = \dfrac{1}{n_{m-1}} \sum_x \delta_x$ est la mesure empirique associée aux stations x présentes en T_{m-1} et $\tilde{r}_m$ est un estimateur de la covariance r_z défini de manière empirique sur tout $[0,1]^2$ à l'aide de la fonction plus proche voisin.

Cette recherche d'un ϵ_m-estimateur se fait en construisant explicitement $\tilde{f}_m$ à partir d'une fonction affine croissante par intervalle, avec raccordements continus aux extrémités de l'intervalle (ligne brisée).

Après avoir montré les convergences, uniformes sur $[0,1]^2$, en probabilité et *p.s.* de l'estimateur empirique $\tilde{r}_m$ vers r_z, on obtient sous l'hypothèse **M3** et des hypothèses de densification des stations, de bonne vitesse d'échantillonnage, d'identifiabilité paramétrique du modèle, de régularité du modèle, les convergences en probabilité et *p.s.* de $\widetilde{\theta}_m$ vers $\theta_0 = (f_0, \beta_0)$.

Enfin, on étudie un cas particulier d'estimation lorsque la fonction de covariance stationnaire est connue et bijective sur $[0,1]$; dans ce cadre, nous traitons un exemple d'estimation de la déformation $f_0 : x \longmapsto x^2$ et de la covariance gaussienne $R(.;\beta_0) : u \longmapsto e^{-\beta_0 u^2}$.

La difficulté et la spécificité de la généralisation de l'étude sur $\mathbb{R}^d$, $d \geq 2$, résident encore une fois dans la description des bijections de $\mathbb{R}^d$ dans $\mathbb{R}^d$, mais aussi dans leur estimation non-paramétrique.

Dans le **chapitre 4**, on suppose que Z est gaussien, centré et réduit, et qu'il vérifie les deux hypothèses suivantes :

- **O1** : Z est continu en moyenne quadratique, c'est-à-dire r_Z est continue dans $[0,1]^2$
- **O2** : r_Z possède des dérivées secondes bornées uniformément dans $\{(x,x') \in [0,1]^2, x \neq x'\}$

La covariance $R_0 = R_{\beta_0}$ appartient à une famille paramétrique de fonctions de covariance stationnaire continues $\mathcal{R} = \{R_\beta\}$, paramétrée par $\beta \in \mathbb{R}^{*+}$, telles que, au voisinage de 0 :

$$R_\beta(u) = (1 - \beta|u| + o(|u|))$$

La déformation non-paramétrique f_0 est un difféomorphisme croissant de $[0,1]$ sur $[0,1]$.
On se propose d'estimer à la fois le paramètre d'échelle β_0 et la déformation f_0, à partir d'une seule trajectoire, ou réalisation, de Z sur $[0,1]$ observée finement.

Dans un premier temps, à partir des hypothèses **O1**-**O2**, nous donnons la définition de la fonction de singularité α_0 (*cf.* Istas (1996)) :

$$\alpha_0(x) = \lim_{h \uparrow 0} \frac{r_Z(x, x+h) - r_Z(x,x)}{h} - \lim_{h' \downarrow 0} \frac{r_Z(x, x+h') - r_Z(x,x)}{h'}$$

Cette fonction, positive et continue dans $[0,1]$, est liée à la déformation f_0 et au paramètre d'échelle β_0 de la manière suivante :

$$\alpha_0(x) = 2\beta_0 f_0^{(1)}(x)$$

On en déduit :

$$\beta_0 = \frac{1}{2}\int_0^1 \alpha_0(v)dv \tag{0.1}$$

$$f_0(x) = \frac{\displaystyle\int_0^x \alpha_0(v)dv}{\displaystyle\int_0^1 \alpha_0(v)dv} \tag{0.2}$$

Par conséquent, si on sait estimer $\int_0^x \alpha_0(v)dv$ pour tout x, on est capable d'estimer β_0 et f_0, c'est-à-dire d'identifier le modèle.

Dans un second temps, partant d'un résultat de Baxter (1956) sur la convergence *p.s.* de la variation quadratique d'un processus gaussien calculée sur la partition dyadique de $[0,1]$, on montre, grâce à un résultat de convergence presque-complète (*cf.* Stout (1974)), le résultat plus général suivant, pour tout x dans $[0,1]$:

$$V_n^2(x) = \sum_{k=1}^n \left[Z\left(\frac{k}{n}\right) - Z\left(\frac{k-1}{n}\right)\right]^2 \times \mathbb{1}_{[0,x]}\left(\frac{k}{n}\right) \underset{n\to+\infty}{\longrightarrow} \int_0^x \alpha_0(v)dv,\ p.s. \tag{0.3}$$

où $V_n^2(x)$ est la variation quadratique de Z le long de la partition générale $\Pi_n = \left\{0 < \frac{1}{n} < \cdots < \frac{n-1}{n} < 1\right\}$ sur $[0,x]$.

Compte tenu de la relation (0.1), le résultat (0.3) nous conduit à proposer pour estimateur de β_0 :

$$\hat{\beta}_n = \frac{1}{2}V_n^2(1)$$

estimateur pour lequel on obtient les convergences *p.s.*, en moyenne quadratique et en loi. De la même façon, compte tenu de la relation (0.2), le résultat (0.3) nous conduirait à proposer pour estimateur de $\int_0^x \alpha_0(v)dv$ pour tout x dans $[0,1]$: $V_n^2(x)$. Cependant cet estimateur n'est pas continu, car fonction en escalier. On propose donc une interpolation linéaire avec raccordements continus aux points de Π_n, qui nous conduit à l'estimateur $\hat{g}_n(x)$ de $\int_0^x \alpha_0(v)dv$, pour tout x dans $[0,1]$, défini par :

$$\begin{cases} \hat{g}_n(x) = V_n^2(x) + n\left[Z\left(\frac{\lfloor nx\rfloor + 1}{n}\right) - Z\left(\frac{\lfloor nx\rfloor}{n}\right)\right]^2\left(x - \frac{\lfloor nx\rfloor}{n}\right),\ \text{si } x \in [0,1[\\ \hat{g}_n(1) = V_n^2(1),\ \text{si } x = 1 \end{cases}$$

On propose alors, comme estimateur de f_0, la fonction $\hat{f}_n$:

$$\hat{f}_n(x) = \frac{\hat{g}_n(x)}{\hat{g}_n(1)},\ x \in [0,1]$$

estimateur pour lequel on obtient la convergence uniforme *p.s.*

Pour illustrer les résultats, on étudie deux exemples de processus dont la covariance est réductible à une covariance stationnaire :

1. un processus $Z = \{Z(x), 0 \leq x \leq 1\}$ qui après déformation de son support par $f_0 : x \longmapsto x^2$ est un processus d'Ornstein-Uhlenbeck $Y = \{Y(u), 0 \leq u \leq 1\}$, $u = f_0(x)$, de fonction de covariance exponentielle :

$$R_{\beta_0}(u) = e^{-\beta_0|u|}$$

2. un processus Z qui après déformation de son support par $f_0 : x \longmapsto \dfrac{\ln(1+x)}{\ln(2)}$ est un processus stationnaire de covariance sphérique :

$$R_{\beta_0}(u) = 1 - \beta_0|u| + \frac{4\beta_0^3|u|^3}{27}$$

Chapitre 1

Covariances réductibles par déformation à une covariance stationnaire

Nous allons rappeler, dans la section 1.1, quelques résultats sur les fonctions de covariance. Dans la section 1.2, nous montrons, en construisant trois exemples, que toutes les fonctions de covariance non-stationnaires ne sont pas réductibles à une covariance stationnaire via une déformation de l'espace. Pour celles qui sont réductibles, dans la section 1.3 nous donnons un résultat les caractérisant, sous certaines conditions de régularité. Ce résultat nous garantit de plus l'unicité de la transformation et de la covariance stationnaire.

1.1 Quelques rappels

1.1.1 Fonctions de covariance

Soit $Z = \{Z(x)\}$ un processus aléatoire défini sur $\mathbb{R}$ à valeurs réelles et possédant des moments d'ordre 2 (*cf.* Guikhman et Skorokhod (1980)). Cela signifie que pour tout x dans $\mathbb{R}$:

$$E|Z(x)|^2 < +\infty$$

On peut donc, pour tout (x, x') dans $\mathbb{R}^2$, définir une fonction de covariance :

$$r_Z(x, x') = Cov(Z(x), Z(x')) = Cov(Z(x'), Z(x)) = r_Z(x', x)$$

Quand $x = x'$, $r_Z(x, x)$ est la variance de Z en x. D'après l'inégalité de Cauchy-Schwarz :

$$|r_Z(x, x')| \leq \sqrt{r_Z(x, x)}\sqrt{r_Z(x', x')}$$

Une fonction de covariance est un noyau défini positif (Guikhman et Skorokhod (1980), p.18), c'est-à-dire que pour tout entier $n \geq 1$, tout $(\lambda_1, \ldots, \lambda_n)$ dans $\mathbb{R}^n$ et tout $(x_1, \ldots, x_n)$ dans

$\mathbb{R}^n$:

$$Var\left(\sum_{i=1}^{n} \lambda_i Z(x_i)\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j r_Z(x_i, x_j) \geq 0$$

On parlera indifféremment de covariance ou de fonction de covariance. Si $r_Z(x,x) > 0$ pour tout x dans $\mathbb{R}$, on peut également définir la corrélation de Z, pour tout (x, x') dans $\mathbb{R}^2$:

$$Cor(Z(x), Z(x')) = \frac{r_Z(x,x')}{\sqrt{r_Z(x,x)}\sqrt{r_Z(x',x')}}$$

D'après ce qui précède, la corrélation appartient à $[-1, 1]$.

1.1.2 Continuité en moyenne quadratique

***Définition* 1.1** *(Priestley (1981), p.151)*
On dit que Z est continu en moyenne quadratique en $x = x_0$ si et seulement si :

$$\lim_{x \to x_0} E\left[Z(x) - Z(x_0)\right]^2 = 0$$

♠

On a le résultat suivant reliant la continuité en moyenne quadratique de Z et la continuité de r_Z.

***Lemme* 1.1** *(Loève (1978), p.136)*
Z est continu en moyenne quadratique dans $\mathbb{R}$ si et seulement si r_Z est continue en tout point de la diagonale $x = x'$. De plus, r_Z est continue en tout point de la diagonale si et seulement si r_Z est continue partout dans $\mathbb{R}^2$.

★

1.1.3 Processus stationnaires

***Définition* 1.2** *Z est strictement stationnaire si toutes les lois de dimension finie sont invariantes par translation.*

♠

Mais le plus souvent, on ne considère que les processus faiblement stationnaires ; pour les processus gaussiens, les deux notions sont équivalentes.

***Définition* 1.3** *Z est un processus faiblement stationnaire, ou stationnaire à l'ordre 2, si Z possède des moments d'ordre 2 vérifiant pour tout (x, x') dans $\mathbb{R}^2$:*

$$\begin{aligned} EZ(x) &= \mu_Z \\ r_Z(x,x') &= R(x'-x) \end{aligned}$$

où μ_Z est une constante et R est une fonction définie dans $\mathbb{R}$ et définie positive. ♠

La variance d'un processus stationnaire est donc constante et égale à $R(0)$. Lorsque $R(0)$ est égale à 1, la fonction de covariance est identique à la fonction de corrélation. Dans la suite, stationnaire signifiera toujours stationnaire à l'ordre 2. Pour l'étude et des développements sur les processus stationnaires sur $\mathbb{R}$, on pourra se reporter à Yaglom (1987).

À toutes fins utiles, rappelons la notion de champ stationnaire isotrope (*cf.* Yadrenko (1983), chapitre 1) :

***Définition* 1.4** *Un champ aléatoire Z_d, processus aléatoire défini sur $\mathbb{R}^d$, $d \geq 1$, est stationnaire isotrope si pour tout (x, x') dans $\mathbb{R}^{2d}$:*

$$\begin{array}{rcl} EZ_d(x) & = & \mu_{Z_d} \\ r_{Z_d}(x, x') & = & R\left(\|x' - x\|_d\right) \end{array}$$

♠

Avec $d = 1$, la définition 1.4 est équivalente à la définition 1.3.

1.1.4 Fonctions définies positives

Pour une revue des principales propriétés des fonctions définies positives, on pourra se référer à Schœnberg (1938a), Stewart (1976) et Rudin (1990). Précisons néanmoins ici quelques propriétés fondamentales des fonctions définies positives à valeurs dans $\mathbb{R}$:

***Propriétés* 1.1**

1. *R est paire : $R(u) = R(-u)$*
2. *R est bornée : $|R(u)| \leq R(0)$*
3. *$(R(u') - R(u))^2 \leq 2R(0)[R(0) - R(u' - u)]$*

★

La troisième propriété signifie que R est uniformément continue dans $\mathbb{R}$ si et seulement si R est continue en 0. Cela est encore équivalent au fait que Z est continu en moyenne quadratique dans $\mathbb{R}$ (Priestley (1981), pp.151-152). La continuité de R nous conduit directement au théorème de Bochner :

***Théorème* 1.1** *(Rudin (1990), p.19)*
Une fonction $R : \mathbb{R} \longmapsto \mathbb{R}$ continue dans $\mathbb{R}$ est définie positive si et seulement si elle est la transformée de Fourier d'une mesure positive finie μ :

$$R(u) = \int_{-\infty}^{+\infty} e^{i\omega u} \mu(d\omega)$$

★

Enfin on introduit une classe particulière de fonctions :

***Définition* 1.5** *On définit $PD([0,1])$ comme la classe de toutes les fonctions $R : [-1,1] \longmapsto \mathbb{R}$ définies positives et continues.*

♠

Krein (1940) a montré que toute fonction de $PD([0,1])$ admet un prolongement continu et défini positif (pas forcément unique) sur $\mathbb{R}$. Ce résultat n'est plus nécessairement vrai sur $\mathbb{R}^2$.

1.2 Exemple de covariances non réductibles : les limites de la méthode

On montre à travers trois exemples que la covariance d'un processus aléatoire non-stationnaire n'est pas toujours réductible à une covariance stationnaire.

Le premier exemple, élémentaire, considère un processus non continu en moyenne quadratique. Le deuxième exemple reprend le début de la réponse d'Assouad (1980) à la question suivante de Krivine :

Comment distinguer, parmi les noyaux hermitiens r de type positif (vérifiant $r(x,x) =$ cste) ceux de la forme $r(x,x') = R(f(x') - f(x))$, où R est une fonction de type positif sur un groupe commutatif localement compact D et f est une application de G dans D ?

Dans ce second exemple, on utilise un processus continu en moyenne quadratique. Enfin dans le troisième, qui ne suppose aucune hypothèse de continuité, on construit un processus aléatoire sur une partie de $\mathbb{R}$ à partir d'un champ aléatoire stationnaire, isotrope défini sur $\mathbb{R}^2$. Par rapport au problème formulé par Krivine, nous nous restreignons, dans ce paragraphe, au cas d'applications f bijectives et bicontinues, c'est-à-dire aux homéomorphismes.

1.2.1 Exemple 1 : cas d'un processus non-continu en moyenne quadratique

Nous commençons par démontrer le lemme suivant :

***Lemme* 1.2** *Soit $Y = \{Y(u)\}$ un processus stationnaire sur $[0,1]$, de variance 1. Si sa fonction de covariance, R, définie sur $[-1,1]$, vaut 1 sur un intervalle $[-\gamma,\gamma], 0 < \gamma < 1$, alors elle vaut 1 sur $[-1,1]$.* ★

Démonstration :

Si on a, pour tout u dans $[0,\gamma]$:

$$R(u) = 1$$

Alors, pour tout h dans $[0, \gamma]$:

$$Cov(Y(\gamma), Y(\gamma + h)) = R(h) = 1$$

Comme :

$$Cov(Y(0), Y(\gamma)) = R(\gamma) = 1$$

On en déduit pour tout h dans $[0, \gamma]$:

$$Cov(Y(0), Y(\gamma + h)) = 1$$

Soit, pour tout u dans $[0, 2\gamma]$:

$$R(u) = 1$$

En répétant ce raisonnement $\left(\lceil \frac{1}{\gamma} \rceil - 2\right) \vee 2$ fois on obtient le résultat voulu.

■

Soient ϵ_1 et ϵ_2 deux $v.a.r.$ indépendantes et de variance 1. On définit un processus aléatoire réel $Z = \{Z(x)\}$ sur $[0, 1]$ de la façon suivante :

$$Z(x) = \epsilon_1 \mathbb{1}_{[0,\frac{1}{2}]}(x) + \epsilon_2 \mathbb{1}_{]\frac{1}{2},1]}(x)$$

Théorème 1.2 *Il n'existe pas d'homéomorphisme f de $[0, 1]$ sur $[0, 1]$, vérifiant $f(0) = 0$ et $f(1) = 1$, et de processus $Y = \{Y(u)\}$ stationnaire sur $[0, 1]$, de fonction de covariance R, tels que, pour tout u dans $[0, 1]$:*

$$Y(u) = Z(f^{-1}(u))$$

★

Démonstration :

La fonction de covariance de Z vaut, pour tout $x \leq x'$:

$$r_Z(x, x') = \begin{cases} 1 & \text{si } x' \leq \frac{1}{2} \text{ ou } x > \frac{1}{2} \\ 0 & \text{sinon} \end{cases}$$

Si un tel f et un tel processus Y existent, on a pour $Y(f(x)) = Z(x)$ et pour tout $x \leq x'$:

$$r_Z(x, x') = R(f(x') - f(x))$$

Soit $\gamma = f(\frac{1}{2}) > 0$. Alors pour tout $0 \leq u, u' \leq \gamma$, on a :

$$R(u' - u) = r_Z(f^{-1}(u), f^{-1}(u')) = Var(\epsilon_1) = 1$$

Autrement dit $R(u)$ vaut 1 sur $[-\gamma, \gamma]$. On en déduit donc, d'après le lemme 1.1, que $R(u)$ vaut 1 partout sur $[-1, 1]$. Or cela contredit le fait que pour tout $0 \leq u \leq \gamma < u' \leq 1$:

$$R(u' - u) = 0$$

■

1.2.2 Exemple 2 : cas d'un processus continu en moyenne quadratique

Voici un exemple plus sophistiqué donné par Guyon (communication personnelle, 1996), d'après une communication de Assouad (1980).

(i) Construisons tout d'abord un jeu de covariances $(r_{i,j}), 1 \leq i, j \leq 6$.

Considérons un angle α dans $\left]0, \frac{\pi}{3}\right[$ et 6 vecteurs unitaires de $\mathbb{R}^3$ (muni du produit scalaire canonique $< .,. >_3$) liés par les 3 relations linéaires suivantes :

$$\left\{\begin{array}{rcl} X_1 + X_2 & = & 2X_6 \cos\alpha \\ X_2 + X_3 & = & 2X_4 \cos\alpha \\ X_1 + X_3 & = & 2X_5 \cos\alpha \end{array}\right.$$

Soit encore:

$$\begin{aligned} \|X_6 \cos\alpha - \frac{1}{2}(X_1 + X_2)\|_3^2 + \|X_4 \cos\alpha - \frac{1}{2}(X_2 + X_3)\|_3^2 \\ + \|X_5 \cos\alpha - \frac{1}{2}(X_1 + X_3)\|_3^2 = 0 \end{aligned} \tag{1.1}$$

Un tel choix est toujours possible. On commence par choisir X_1 quelconque puis on prend X_2 et X_3 tels que :

$$\|X_i + X_j\|_3 = 2\cos\alpha,\ 1 \leq i \neq j \leq 3$$

Les intersections de X_1, X_2, X_3 avec la sphère unité sont les sommets d'un triangle équilatéral, et X_4, X_5, X_6, 3 vecteurs unitaires donnés par les relations :

$$\left\{\begin{array}{rcl} X_6 & = & \dfrac{X_1 + X_2}{2\cos\alpha} \\[2ex] X_4 & = & \dfrac{X_2 + X_3}{2\cos\alpha} \\[2ex] X_5 & = & \dfrac{X_1 + X_3}{2\cos\alpha} \end{array}\right.$$

On considère ensuite la matrice de Gram

$$(r_{i,j}) = (< X_i, X_j >_3), 1 \leq i, j \leq 6$$

associée à cet ensemble de 6 vecteurs $(X_1, X_2, \ldots, X_6)$.

La matrice $(r_{i,j})$ est symétrique et définie positive ; c'est une matrice de covariance.

(ii) Construisons ensuite un processus aléatoire $\{Z(x)\}$ sur $[0, 1]$, à partir de ce jeu de covariances $(r_{i,j})$.

***Lemme* 1.3** *Soit le* $6-$*uplet* $\left\{x_i = \frac{i-1}{5}, i = 1, 2, \ldots, 6\right\}$. *Il existe un processus aléatoire gaussien réduit et continu en moyenne quadratique* $\{Z(x)\}$ *défini sur* $[0, 1]$, *tel que sa fonction de covariance vérifie, pour tout* $1 \leq i, j \leq 6$ *:*

$$r_Z(x_i, x_j) = r_{i,j}$$

★

Démonstration :

Soit $\epsilon = (\epsilon_1, \ldots, \epsilon_6)$, un vecteur aléatoire gaussien centré de matrice de covariance $(r_{i,j})$. On définit d'abord sur $[x_1, x_6] = [0, 1]$ un processus $\epsilon = \{\epsilon(x)\}$ de la façon suivante :

$$\left\{\begin{array}{rcl} \epsilon(x) & = & \sum_{i=2}^{6} \left(\epsilon_{i-1} + \frac{(\epsilon_i - \epsilon_{i-1})}{x_i - x_{i-1}}(x - x_{i-1})\right) \mathbb{1}_{[x_{i-1}, x_i[}(x) \\ \epsilon(x_6) & = & \epsilon_6 \end{array}\right.$$

Ainsi défini, ϵ est un processus aléatoire (mesurable) gaussien, continu en moyenne quadratique, de variance :

$$\left\{\begin{array}{rcl} r_\epsilon(x, x) & = & \sum_{i=2}^{6} \left(1 + 2\frac{x - x_{i-1}}{x_i - x_{i-1}}(r_{i-1,i} - 1) + \left(\frac{x - x_{i-1}}{x_i - x_{i-1}}\right)^2 (2 - 2r_{i-1,i})\right) \mathbb{1}_{[x_{i-1}, x_i[}(x) \\ r_z(x_6, x_6) & = & 1 \end{array}\right.$$

On prend ensuite $Z = \{Z(x)\}$, défini par :

$$Z(x) = \frac{\epsilon(x)}{\sqrt{r_\epsilon(x, x)}}$$

■

***Théorème* 1.3** *Soit un processus* $Z = \{Z(x)\}$, *continu en moyenne quadratique, dont la covariance vérifie, pour tout* $1 \leq i, j \leq 6$ *:*

$$r_Z(x_i, x_j) = r_{i,j}$$

Il n'existe pas d'homéomorphisme f *de* $[0, 1]$ *sur* $[0, 1]$, *vérifiant* $f(0) = 0$ *et* $f(1) = 1$, *et de processus* $Y = \{Y(u)\}$ *stationnaire sur* $[0, 1]$, *de fonction de covariance* R, *tels que, pour tout* u *dans* $[0, 1]$ *:*

$$Y(u) = Z(f^{-1}(u))$$

★

Grâce au lemme 1.3, l'existence de Z tel que $r_Z(x_i, x_j) = r_{i,j}$ est assurée.

Nous commençons par démontrer le lemme suivant :

***Lemme* 1.4** *Soient $W_1, \dots, W_6$, six nombres complexes de module égal à 1. Alors, pour tout $\alpha \in \left]0, \frac{\pi}{3}\right[$, on a :*

$$|W_6 \cos\alpha - \frac{1}{2}(W_1 + W_2)|^2 + |W_4 \cos\alpha - \frac{1}{2}(W_2 + W_3)|^2 + |W_5 \cos\alpha - \frac{1}{2}(W_1 + W_3)|^2$$

$$\geq (\cos\alpha - \frac{1}{2})^2 > 0$$

★

Démonstration :

Il suffit de remarquer que, quelques soient W_1, W_2 et W_3 unitaires :

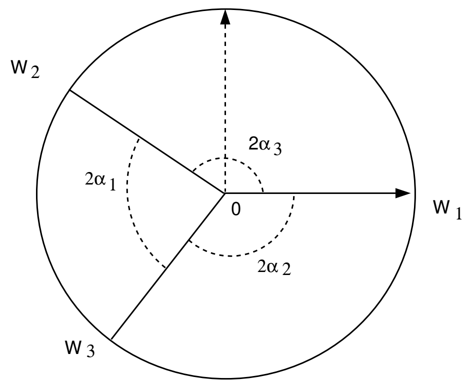
$$\inf\{|W_1 + W_2|, |W_2 + W_3|, |W_1 + W_3|\} \leq 1$$

Ce résultat étant équivalent au suivant : $\forall(\alpha_1, \alpha_2, \alpha_3), \alpha_1 + \alpha_2 + \alpha_3 = \pi$,

$$\inf\{|\cos\alpha_1|, |\cos\alpha_2|, |\cos\alpha_3|\} \leq \frac{1}{2}$$

Il suffit de prendre par exemple :

$$\begin{aligned} W_1 &= 1 = e^{2i(\alpha_1+\alpha_2+\alpha_3)} \\ W_2 &= e^{2i\alpha_3} \\ W_3 &= e^{2i(\alpha_1+\alpha_3)} \end{aligned}$$



Quitte à réordonner les W_1, W_2 et W_3, on peut prendre $|W_1 + W_2| \leq 1$.

Soit $\alpha \in \left]0, \frac{\pi}{3}\right[$, $\frac{1}{2} < \cos\alpha < 1$, d'où, en utilisant l'inégalité $|a - b| \geq ||a| - |b||$:

$$|W_6 \cos\alpha - \frac{1}{2}(W_1 + W_2)| \geq |\cos\alpha - \frac{1}{2}|W_1 + W_2||$$

Soit encore :

$$|W_6 \cos\alpha - \frac{1}{2}(W_1 + W_2)| \geq \cos\alpha - \frac{1}{2} > 0$$

■

Démonstration du théorème 1.3 :

Supposons qu'il existe un homéomorphisme f et un processus $Y = \{Y(u)\}$ stationnaire sur $[0, 1]$ tels que, pour tout u dans $[0, 1]$:

$$Y(u) = Z(f^{-1}(u))$$

Comme Z est continu en moyenne quadratique et f^{-1} est continue, Y est aussi continu en moyenne quadratique ; sa fonction de covariance R est par conséquent continue dans $[-1, 1]$. D'après un résultat de Krein (1940), R admet un prolongement continu et défini positif sur $\mathbb{R}$. D'après le théorème de Bochner, on peut donc écrire :

$$R(u) = \int_{-\infty}^{\infty} e^{i\omega u} \mu(d\omega)$$

où μ est une mesure positive finie $(0 < \mu(\mathbb{R}) = m < +\infty)$.

De plus, comme le processus est réel : $R(u) = R(-u)$. On en déduit alors, pour tout $1 \leq i \leq j \leq 6$:

$$r_{i,j} = \int_{-\infty}^{\infty} e^{i\omega(f(x_j) - f(x_i))} \mu(d\omega)$$

Posant $W_i(\omega) = e^{i\omega f(x_i)}$ pour $1 \leq i \leq 6$, on a :

$$r_{i,j} = \int_{-\infty}^{\infty} W_j(\omega) \bar{W}_i(\omega) \mu(d\omega) = < W_i, W_j >_{L^2_{\mathbb{C}}(\Omega,\mu)} = < X_i, X_j >_3$$

Il s'ensuit donc que, pour α dans $\left]0, \frac{\pi}{3}\right[$:

$$\begin{aligned}\|W_6 \cos\alpha - \frac{1}{2}(W_1 + W_2)\|^2_{L^2_{\mathbb{C}}(\Omega,\mu)} &= \int_{-\infty}^{\infty} \left| W_6(\omega)\cos\alpha - \frac{1}{2}(W_1(\omega) + W_2(\omega)) \right|^2 \mu(d\omega) \\ &= \|X_6 \cos\alpha - \frac{1}{2}(X_1 + X_2)\|^2_3\end{aligned}$$

Posons :

$$\begin{aligned} D(w) &= |W_6(\omega)\cos\alpha - \frac{1}{2}(W_1(\omega)+W_2(\omega))|^2 \\ &+ |W_4(\omega)\cos\alpha - \frac{1}{2}(W_2(\omega)+W_3(\omega))|^2 \\ &+ |W_5(\omega)\cos\alpha - \frac{1}{2}(W_1(\omega)+W_3(\omega))|^2 \end{aligned}$$

Du lemme 1.4 précédent, on déduit pour tout ω :

$$D(\omega) \geq \left(\cos\alpha - \frac{1}{2}\right)^2 > 0$$

Par intégration par rapport à μ, on obtient donc :

$$\begin{aligned} \|X_6\cos\alpha - \frac{1}{2}(X_1+X_2)\|_3^2 + \|X_5\cos\alpha - \frac{1}{2}(X_1+X_3)\|_3^2 \\ +\|X_4\cos\alpha - \frac{1}{2}(X_2+X_3)\|_3^2 = 0 \\ \geq m(\cos\alpha - \frac{1}{2})^2 > 0 \end{aligned}$$

La première égalité provenant de la définition des 6 point $X_1, X_2, \ldots, X_6$ (*cf.* l'égalité (1.1)). D'où la contradiction.

■

1.2.3 Exemple 3 : cas d'un processus défini à partir d'un champ stationnaire, isotrope

Dans l'exemple précédent nous avons construit un processus aléatoire à partir de covariances s'exprimant à partir de produits scalaires sur $\mathbb{R}^3$. L'idée sous-jacente ici est encore de partir d'une dimension strictement supérieure à 1 pour se ramener à la dimension 1. Plus précisément, soit $P = \{P(x,y)\}$ un champ stationnaire isotrope, défini sur $\mathbb{R}^2$, de fonction de covariance :

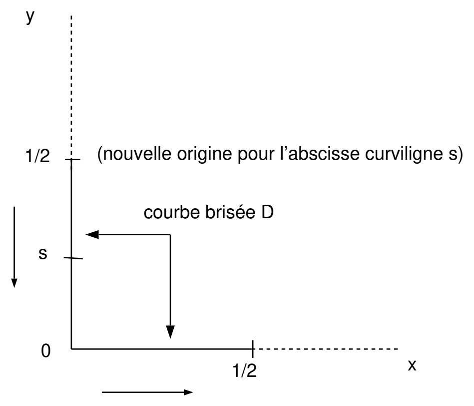
$$r_P(x,y,x',y') = k\left(\sqrt{(x'-x)^2+(y'-y)^2}\right)$$

où k est une fonction de covariance, stationnaire isotrope, définie sur $\mathbb{R}^+$ et injective.

Observons ce processus sur la courbe brisée :

$$D = \left\{x = 0, y \in [0, \frac{1}{2}]\right\} \cup \left\{x \in [0, \frac{1}{2}], y = 0\right\}$$

En reparamétrant $P(x,y), (x,y) \in D^2$, à l'aide de l'abscisse curviligne s, calculée à partir du point $\left(0, \frac{1}{2}\right)$, on obtient finalement un processus $\{Z(s)\}$ défini sur $[0,1]$ comme suit :



$$Z(s) = P(x,y), \text{ où } \begin{cases} s = \frac{1}{2} - y, \text{ si } x = 0 \text{ et } y \in [0, \frac{1}{2}] \\ \\ s = \frac{1}{2} + x, \text{ si } y = 0 \text{ et } x \in]0, \frac{1}{2}] \end{cases}$$

Sa fonction de covariance $r_Z(s,s')$ est alors définie pour tout $s \leq s'$, par :

$$r_Z(s,s') = \begin{cases} k(s'-s), \text{ si } 0 \leq s' \leq \frac{1}{2} \\ k(s'-s), \text{ si } \frac{1}{2} < s \leq 1 \\ k\left(\sqrt{(s'-\frac{1}{2})^2 + (s-\frac{1}{2})^2}\right) \text{ sinon} \end{cases}$$

***Théorème* 1.4** *Il n'existe pas d'homéomorphisme f de $[0,1]$ sur $[0,1]$, vérifiant $f(0) = 0$ et $f(1) = 1$, et de processus $Y = \{Y(u)\}$ stationnaire sur $[0,1]$, de fonction de covariance R, tels que, pour tout u dans $[0,1]$:*

$$Y(u) = Z(f^{-1}(u))$$

★

Démonstration :

Si un tel processus existe, on a, pour tout $0 \leq s \leq \frac{1}{2} < s' \leq 1$:

$$r_Z(s,s') = R(f(s') - f(s)) = k\left(\sqrt{(s'-\frac{1}{2})^2 + (s-\frac{1}{2})^2}\right)$$

Pour $s = 0$, on a :

$$R(f(s')) = k\left(\sqrt{(s' - \frac{1}{2})^2 + \frac{1}{4}}\right)$$

Posant $s' = f^{-1}(u') = g(u')$, on a :

$$R(u') = k\left(\sqrt{\left(g(u') - \frac{1}{2}\right)^2 + \frac{1}{4}}\right)$$

Soit :

$$R(u' - u) = k\left(\sqrt{\left(g(u' - u) - \frac{1}{2}\right)^2 + \frac{1}{4}}\right)$$

Grâce à l'injectivité de k, l'égalité :

$$k\left(\sqrt{\left(g(u') - \frac{1}{2}\right)^2 + \left(g(u) - \frac{1}{2}\right)^2}\right) = k\left(\sqrt{(g(u' - u) - \frac{1}{2})^2 + \frac{1}{4}}\right)$$

conduit à :

$$g^2(u' - u) - g(u' - u) = g^2(u') + g^2(u) - g(u') - g(u)$$

Posant $h(u) = g^2(u) - g(u)$, on aboutit à :

$$h(u' - u) = h(u) + h(u')$$

ce qui entraîne $h(0) = 0 = 2h(u)$, pour tout u. f^{-1} serait une application constante, d'où la contradiction.

■

Remarque :

Ce dernier exemple apporte un élément de réflexion sur la classe des processus aléatoires qu'il serait possible ou non de rendre stationnaire par l'intermédiaire d'une transformation de l'espace. En effet, ne pas trouver une telle transformation de $[0, 1]$ dans $[0, 1]$ peut s'expliquer par le fait que le processus dépende d'une ou de plusieurs autres dimensions supplémentaires, non prises en compte *a priori* dans les données. Ce point est d'ailleurs soulevé dans Sampson et Guttorp (1992). Par exemple, l'altitude peut être un facteur déterminant dans l'étude de certains phénomènes spatiaux, la pluviométrie par exemple (Monestiez, Sampson et Guttorp (1993)). Sa prise en compte pourrait conduire à l'utilisation de transformations de $\mathbb{R}^2$ dans $\mathbb{R}^3$, quand il apparaît évident que les transformations de $\mathbb{R}^2$ dans $\mathbb{R}^2$ ne permettent pas d'obtenir une représentation "stationnaire, isotrope" satisfaisante, au sens où la bijection est respectée. D'une façon générale, on pourrait se demander quelle serait la dimension minimale de l'espace sur lequel le processus transformé soit stationnaire, isotrope au second ordre. On rejoint alors les méthodes de "multidimensional scaling", méthodes d'ailleurs à l'origine du travail de Sampson ((1986), p.8).

♣

1.3 Caractérisation des covariances réductibles et identifiabilité du modèle

Nous présentons un résultat caractérisant les fonctions de covariance réductibles à une fonction de covariance stationnaire. En corollaire, nous obtenons l'unicité du couple (f, R) pour lequel une covariance r peut s'écrire, pour tout (x, x') dans $[0,1]^2$:

$$r(x, x') = R(f(x') - f(x))$$

L'indice Z dans r_Z est omis dans les énoncés qui suivent. Définissons les ensembles :

- $\mathcal{A}_1$: ensemble des fonctions h symétriques, continues dans $[0,1]^2$, différentiables dans $\Delta = \left\{(x, x') \in [0,1]^2, x \neq x'\right\}$
- $\mathcal{A}_2$: ensemble des fonctions k paires, continues dans $[-1,1]$, différentiables dans $[-1,1] - \{0\}$
- $\mathcal{A}_3$: ensemble des difféomorphismes croissants f de $[0,1]$ sur $[0,1]$ (nécessairement $f(0) = 0$ et $f(1) = 1$)

On a le lemme suivant :

***Lemme* 1.5** *Soient h appartenant à $\mathcal{A}_1$, k appartenant à $\mathcal{A}_2$ et f appartenant à $\mathcal{A}_3$, telles que, pour tout (x, x') dans $[0,1]^2$:*

$$h(x, x') = k\left(f(x') - f(x)\right)$$

Alors h est une fonction de covariance si et seulement si k appartient à $PD([0,1])$. De plus, h admet sur $\mathbb{R}^2$ tout entier un prolongement continu défini positif.

★

Démonstration :

h est une fonction de covariance si et seulement si pour tout $n \geq 2$, pour tout $(x_1, \ldots, x_n)$ dans $[0,1]^n$ et tout $(\lambda_1, \ldots, \lambda_n)$ dans $\mathbb{R}^n$:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j h(x_i, x_j) \geq 0$$

$$\Longleftrightarrow$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j k\left(f(x_j)-f(x_i)\right)\geq 0$$

Comme f est une fonction bijective de $[0,1]$ dans $[0,1]$, cette dernière inégalité est équivalente à :

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j k\left(u_j-u_i\right)\geq 0$$

où $u_i=f(x_i)$, pour tout i dans $[1,\ldots,n]$.

k, appartenant à $PD([0,1])$, admet alors un prolongement continu défini positif, pas forcément unique, sur $\mathbb{R}$ (Krein (1940)). On en déduit que h admet sur $\mathbb{R}^2$ un prolongement continu défini positif.

■

***Théorème* 1.5** *Soient r une fonction de covariance appartenant à $\mathcal{A}_1$ et f une déformation appartenant à $\mathcal{A}_3$. Alors, il existe une fonction R appartenant à $\mathcal{A}_2$, telle que, pour tout (x,x') dans $[0,1]^2$:*

$$r(x,x')=R\left(f(x')-f(x)\right)$$

si et seulement si, pour tout (x,x') dans Δ :

$$f^{(1)}(x')r^{(1,0)}(x,x')+f^{(1)}(x)r^{(0,1)}(x,x')=0$$

★

Démonstration :

Le lemme 1.5 précédent nous assure déjà que R appartient à $PD([0,1])$.

(i) La condition est évidemment nécessaire car on vérifie aisément que, pour tout (x,x') dans Δ :

$$\begin{aligned} r^{(1,0)}(x,x') &= -f^{(1)}(x)R^{(1)}\left(f(x')-f(x)\right) \\ r^{(0,1)}(x,x') &= f^{(1)}(x')R^{(1)}\left(f(x')-f(x)\right) \end{aligned}$$

Par conséquent, pour tout (x,x') dans Δ :

$$f^{(1)}(x')r^{(1,0)}(x,x')+f^{(1)}(x)r^{(0,1)}(x,x')=0$$

(ii) Pour ce qui est de la condition suffisante, faisons le changement de variable bijectif, pour tout (x,x') dans Δ :

$$\begin{aligned} u &= f(x')-f(x) \\ u' &= f(x')+f(x) \end{aligned}$$

et posons $H(u,u') = r\left(x(u,u'), x'(u,u')\right)$. L'équation différentielle :

$$f^{(1)}(x')r^{(1,0)}(x,x') + f^{(1)}(x)r^{(0,1)}(x,x') = 0$$

conduit à la nouvelle équation :

$$H^{(0,1)}(u,u') = 0$$

Par conséquent, on en déduit l'existence d'une fonction R paire et différentiable dans $[-1,0[\cup]0,1]$, telle que :

$$R(u) = H(u,u')$$

et donc, pour tout (x,x') dans Δ :

$$R(f(x') - f(x)) = r(x,x')$$

Soit, pour tout $u \in [-1,0[\cup]0,1]$:

$$R(u) = r\left(0, f^{-1}(|u|)\right)$$

Comme r est continue pour $x = x'$ et f^{-1} est continue en 0, un prolongement par continuité de R en 0 entraîne R continue sur $[-1,1]$. Par conséquent R appartient à $\mathcal{A}_2$.

■

Remarques : généralisations du théorème 1.5

1. Nous pensons qu'il est possible d'affaiblir les conditions de régularité faites sur les ensembles $\mathcal{A}_1$ et $\mathcal{A}_3$. Plus précisément, nous conjecturons que le théorème 1.5 est encore vrai quand $\mathcal{A}_1$ et $\mathcal{A}_3$ sont remplacés respectivement par :

   - ${\mathcal{A}'}_1$: l'ensemble des fonctions h symétriques, continues dans $[0,1]^2$, différentiables $\lambda_2 - p.s.$
   - ${\mathcal{A}'}_3$: l'ensemble des homéomorphismes croissants f de $[0,1]$ sur $[0,1]$, dérivables $\lambda_1 - p.s.$, tels qu'en tout point x où la dérivée existe :

   $$f^{(1)}(x) > 0$$

2. C'est par commodité que nous avons considéré les covariances h définies dans $[0,1]^2$ et les déformations f définies dans $[0,1]$ à valeurs dans $[0,1]$. Le théorème 1.5 reste vrai quand $\mathcal{A}_1 - \mathcal{A}_3$ sont remplacés respectivement par :

   - ${\mathcal{A}''}_1$: l'ensemble des fonctions h symétriques, continues dans $[a,b]^2$, $a < b$, différentiables dans $\Delta' = \left\{(x,x') \in [a,b]^2, x \neq x'\right\}$
   - ${\mathcal{A}''}_2$: l'ensemble des fonctions k paires, continues dans $[-1,1]$, différentiables dans $[-1,1] - \{0\}$

- $\mathcal{A}''_3$: l'ensemble des difféomorphismes croissants f de $[a,b]$ sur $[0,1]$ (nécessairement $f(a)=0$ et $f(b)=1$, c'est-à-dire qu'on fixe les translations et les changements d'échelle)

♣

Un exemple de processus non-stationnaire, dont la covariance peut se réduire à une covariance stationnaire, est donné par le mouvement brownien fractionnaire non dégénéré $Z=\{Z(x), x\in\mathbb{R}\}$ d'indice $H\in]0,1[$. Sa covariance s'écrit, pour tout (x,x') dans $\mathbb{R}^2$:

$$r_Z(x,x')=\frac{\sigma^2}{2}\left[x^{2H}+x'^{2H}-|x'-x|^{2H}\right]$$

où $\sigma^2=Var(Z(1))$. Dans le cas où $\sigma=1$, on obtient le mouvement brownien fractionnaire non dégénéré standard, et si en plus $H=\frac{1}{2}$, on retrouve le mouvement brownien standard ou processus de Wiener. Pour appliquer notre résultat qui nécessite $r(x,x)=cste$, nous travaillons avec la corrélation :

$$c_Z(x,x')=\frac{1}{2}\frac{\left[x^{2H}+x'^{2H}-|x'-x|^{2H}\right]}{x^H x'^H}$$

Cette corrélation n'étant pas définie pour x, ou x', nul, nous travaillons avec Z défini sur un intervalle $[a,b]$ avec $0<a<b$ et considérons les ensembles $\mathcal{A}''_1-\mathcal{A}''_3$ du point 2. de la remarque précédente. On vérifie que c_Z appartient à $\mathcal{A}''_1$.

Existe-t-il un difféomorphisme croissant f, de $[a,b]$ sur $[0,1]$ appartenant à $\mathcal{A}''_3$, et un processus $Y=\{Y(u)\}$ stationnaire sur $[0,1]$, de fonction de corrélation $c_Y(u,u')=C_Y(u'-u)$ appartenant à $\mathcal{A}''_2$, tels que :

$$c_Z(x,x')=C_Y(f(x')-f(x))\,?$$

On a, pour tout (x,x') dans Δ' :

$$\frac{1}{x'}c_Z^{(1,0)}(x,x')+\frac{1}{x}c_Z^{(0,1)}(x,x')=0$$

et on en déduit la transformation :

$$f(x)=\frac{\ln(x)-\ln(a)}{\ln(b)-\ln(a)}$$

qui permet d'avoir :

$$c_Z(x,x')=C_Y\left(f(x')-f(x)\right)$$

avec :

$$C_Y(u) = \frac{1}{2}\frac{\left[1 + \exp\left(2H(\ln(b) - \ln(a))u\right) - |\exp\left((\ln(b) - \ln(a))u\right) - 1|^{2H}\right]}{\exp\left(H(\ln(b) - \ln(a))u\right)}$$

En conclusion, si $\{Z(x), x \in [a,b], 0 < a < b\}$ est un mouvement brownien fractionnaire d'indice $H \in]0,1[$, alors, pour tout x dans $[a,b]$:

$$Z(x) = \sigma x^H Y\left(\frac{\ln(x) - \ln(a)}{\ln(b) - \ln(a)}\right)$$

où $\{Y(u), u \in [0,1]\}$ est un processus stationnaire centré et réduit, de covariance stationnaire $C_Y(u)$.

Grâce au théorème précédent 1.5, nous avons le corollaire suivant donnant l'unicité du couple (f, R) :

***Corollaire* 1.1** *Si r vérifie, pour tout (x, x') dans $[0,1]^2$:*

$$r(x,x') = R(f(x') - f(x))$$

alors f et R sont uniques.

Démonstration :

Supposons qu'il existe deux couples distincts (f_1, R_1) et (f_2, R_2) tels que :

$$r(x,x') = R_1(f_1(x') - f_1(x)) = R_2(f_2(x') - f_2(x))$$

D'après le théorème 1.5, on en déduit, pour tout (x, x') dans Δ :

$$\frac{r^{(1,0)}(x,x')}{r^{(0,1)}(x,x')} = -\frac{f_1^{(1)}(x)}{f_1^{(1)}(x')} = -\frac{f_2^{(1)}(x)}{f_2^{(1)}(x')}$$

Prenant $x' = 0$, on a, pour tout x dans $]0,1]$:

$$f_1^{(1)}(x) = f_2^{(1)}(x)\frac{f_1^{(1)}(0)}{f_2^{(1)}(0)} \tag{1.2}$$

Comme :

$$\int_0^1 f_1^{(1)}(x)dx = \int_0^1 f_2^{(1)}(x)dx = 1$$

en intégrant l'égalité (1.2) sur $[0,1]$, on obtient $f_1^{(1)}(0) = f_2^{(1)}(0)$. Par conséquent, pour tout x dans $[0,1]$:

$$f_1(x) = f_2(x) + c$$

Ayant $f_1(0) = f_2(0) = 0$, alors $c = 0$, et finalement :

$$f_1 = f_2$$

De plus, pour tout u dans $[-1, 1]$:

$$R_1(u) = r\left(0, f_1^{-1}(|u|)\right) = r\left(0, f_2^{-1}(|u|)\right) = R_2(u)$$

■

Nous avons un autre théorème pour caractériser les covariances r réductibles à une covariance stationnaire R, mais d'une utilisation peu commode.

***Théorème* 1.6** *Soient r, une fonction de covariance continue dans $[0,1]^2$, f, un homéomorphisme croissant de $[0,1]$ sur $[0,1]$. Alors, il existe R dans $PD([0,1])$, telle que, pour tout (x, x') dans $[0,1]^2$:*

$$r(x, x') = R\left(f(x') - f(x)\right)$$

si et seulement si, pour tout (x, x') dans $[0,1]^2$:

$$r(x, x') = r\left(0, f^{-1}\left(|f(x') - f(x)|\right)\right)$$

★

Démonstration :

(i) La condition nécessaire est évidente :

$$r\left(0, f^{-1}\left(|f(x') - f(x)|\right)\right) = R\left(|f(x') - f(x)| - f(0)\right) = R\left(f(x') - f(x)\right) = r(x, x')$$

(ii) Pour démontrer la condition réciproque, posons, pour tout u dans $[-1, 1]$:

$$R(u) = r(0, f^{-1}(|u|))$$

On en déduit, pour tout (x, x') dans $[0,1]^2$:

$$r(x, x') = r\left(0, f^{-1}\left(|f(x') - f(x)|\right)\right) = R(f(x') - f(x))$$

■

Au vu de ce dernier théorème, il est suffisant de connaître $r(0, y)$, pour tout y dans $[0,1]$, pour connaître $r(x, x')$ partout. Ainsi, d'un point de vue pratique, connaissant la transformation f, il suffira d'estimer $r(x, x')$, pour tout $x' \geq x$, aux points $(0, y)$, $(y = f^{-1}(f(x') - f(x)))$.

Mettons en parallèle le théorème 1.5 et le deuxième contre-exemple exhibé au paragraphe précédent. On obtiendrait, pour tout $(x, x') \in [0, \dfrac{1}{5}[\times[\dfrac{1}{5}, \dfrac{2}{5}[$:

$$\frac{x'(10a + 2 - 5x) + x(5a + 2) - 3a - \frac{3}{5}}{a + 2(x' - \frac{1}{5}) - 10(x' - \frac{1}{5})^2} r_Z^{(1,0)}(x, x') + \frac{x(-10a - 5x') - 5ax' + 3a + \frac{1}{5}}{a + 2x - 10x^2} r_Z^{(0,1)}(x, x') = 0$$

où $a = \dfrac{1}{5(r_{1,2}-1)}$.

Par conséquent, on vérifie bien que r_Z n'est pas réductible à une covariance stationnaire.

De même, mettons en parallèle le théorème 1.5 et le troisième contre-exemple exhibé au paragraphe précédent. On suppose en outre, pour être dans les conditions d'application du théorème 1.5, que k appartient à $\mathcal{A}_2$. On obtiendrait, pour tout $(s, s') \in [0, \frac{1}{2}] \times]\frac{1}{2}, 1]$:

$$(s' - \frac{1}{2}) r_Z^{(1,0)}(s, s') + (\frac{1}{2} - s) r_Z^{(0,1)}(s, s') = 0$$

Par conséquent, on vérifie encore une fois que r_Z n'est pas réductible à une covariance stationnaire.

On ne peut pas appliquer le théorème 1.5 au premier contre-exemple car la continuité de la fonction de covariance n'est même pas assurée. On voit donc les limites de ce théorème de caractérisation, en terme de régularité des fonctions de covariance r_Z. Enfin, dans la pratique, on ne dispose, le plus souvent, que d'une estimation de r_Z calculée seulement en un nombre fini de points ...

Chapitre 2

Estimation du modèle : cas paramétrique

On s'intéresse à l'estimation des paramètres de la fonction de covariance d'un processus aléatoire Z, réductible à une fonction de covariance stationnaire paramétrique, lorsque la déformation de l'espace appartient à une famille paramétrique de modèles. Dans la section 2.1, on introduit le modèle et on présente le schéma expérimental. Dans la section 2.2, on définit l'estimateur empirique de la covariance et on présente les propriétés asymptotiques de cet estimateur. Dans la section 2.3, on s'intéresse à l'estimation des paramètres de la déformation et de la fonction de covariance stationnaire, puis aux propriétés asymptotiques, la convergence p.s. et la convergence en loi, des estimateurs obtenus. Différents tests statistiques asymptotiques basés sur des sommes pondérées de khi − deux à un degré de liberté indépendants, test d'hypothèse simple et test d'adéquation du modèle, sont proposés dans la section 2.4. Enfin une identification du modèle par contraste pénalisé est étudiée à la section 2.5. L'identification du modèle par contraste pénalisé est utile lorsque la dimension du modèle est inconnue et qu'on souhaite se prémunir contre une mauvaise paramétrisation. Dans la section 2.6, on illustre les résultats théoriques obtenus par un exemple. Dans la section 2.7, on voit comment cette étude peut se généraliser :

- *au cas où la variance de Z n'est pas constante ;*
- *au cas où Z est défini sur une partie G de $\mathbb{R}^d, d \geq 2$, et où la déformation, rendant ce champ stationnaire isotrope, est un homéomorphisme de G sur une partie D de $\mathbb{R}^d$. Un exemple est traité dans ce cas.*

2.1 Présentation du modèle et schéma expérimental

Le modèle est défini dans le paragraphe 2.1.1. Ensuite, on présente le schéma expérimental et on précise dans quel cadre asymptotique on se place dans le paragraphe 2.1.2.

2.1.1 Le modèle

Soit $Z = \{Z(x)\}$, un processus aléatoire défini dans $[0,1]$, continu en moyenne quadratique, à valeurs réelles, de moyenne M_0 inconnue mais constante, de variance σ_0^2 inconnue mais constante et de fonction de covariance r_Z. On suppose en outre :

- **M1** : $\sup\limits_{[0,1]} E|Z(x)|^4 < +\infty$

Soit $\mathcal{R} = \{R(.;\beta)\}$ une famille de fonctions de covariance stationnaire continues dans $[-1,1]$, paramétrée par $\beta \in \mathbb{R}^q$, $q \in \mathbb{N}^*$. $\sqrt{R(0;\beta)} = \sigma$, le paramètre courant de l'écart-type de Z, est inclus dans β.

On considère par ailleurs $\mathcal{F} = \{f(.;\alpha)\}$, $\alpha \in \mathbb{R}^p$, $p \in \mathbb{N}^*$, une famille paramétrique d'homéomorphismes croissants de $[0,1]$ sur $[0,1]$.

On note $\theta = (\alpha, \beta)$ et on considère l'hypothèse :

- **H0** : il existe une unique valeur α_0 et un processus stationnaire $Y = \{Y(u)\}$, défini dans $[0,1]$, de moyenne M_0, de fonction de covariance $r_Y(u,u') = R(u'-u;\beta_0)$ appartenant à $\mathcal{R}$, tels que, pour tout x dans $[0,1]$:

$$Y\left(f(x;\alpha_0)\right) = Z(x)$$

Sous l'hypothèse **H0**, le processus Z a pour fonction de covariance, pour tout (x,x') dans $[0,1]^2$:

$$\begin{aligned} r_Z(x,x') &= r_Y\left(f(x;\alpha_0), f(x';\alpha_0)\right) \\ &= R\left(f(x';\alpha_0) - f(x;\alpha_0);\beta_0\right) \\ &= r(x,x',\theta_0) \end{aligned}$$

avec $\theta_0 = (\alpha_0, \beta_0)$.

L'objet de ce chapitre est d'estimer le paramètre de la déformation α_0 et le paramètre de la covariance β_0, puis d'étudier les propriétés asymptotiques des estimateurs de ces deux paramètres.

2.1.2 Le schéma expérimental

Pour $T \in \mathbb{N}^*$, soient T réalisations *i.i.d.* du processus Z en un nombre fini et fixé n de stations positionnées en $x_1 < x_2 < \cdots < x_n$ sur $[0,1]$. Ces positions sont les mêmes quelque soit la réalisation $t \in [1, \ldots, T]$. L'étude asymptotique porte uniquement sur le nombre T de réalisations, supposées indépendantes et issues de la même loi. En résumé, en notant $Z_t(x_i)$

l'observation en x_i, $i \in [1, \ldots, n]$, pour la réalisation t, $t \in [1, \ldots, T]$, du processus Z, on dispose du schéma rectangulaire suivant :

	$\forall 1 \leq i \leq n, 0 \leq x_i \leq 1$			
	x_1	$< x_2$	$\ldots$	$< x_n$
1	$Z_1(x_1)$	$Z_1(x_2)$	$\ldots$	$Z_1(x_n)$
2	$Z_2(x_1)$	$Z_2(x_2)$	$\ldots$	$Z_2(x_n)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
T	$Z_T(x_1)$	$Z_T(x_2)$	$\ldots$	$Z_T(x_n)$

2.2 Estimateur empirique de la fonction de covariance

On définit, dans le paragraphe 2.2.1, un estimateur empirique de la fonction de covariance r_Z aux points de mesure (x_i, x_j), $i \leq j$, du processus Z. On présente au paragraphe 2.2.2 les propriétés asymptotiques de cet estimateur : la convergence en moyenne quadratique, la convergence *p.s.*, la convergence en loi et la loi du logarithme itéré. Ces propriétés seront utilisées dans l'étude des estimateurs des paramètres α_0 et β_0 du modèle.

2.2.1 Définitions

À l'instant $T \geq 2$, on estime la covariance $r_Z(x_i, x_j)$, pour tout $i \leq j$, de façon empirique par :

$$\hat{r}_T(x_i, x_j) = \frac{1}{T-1} \sum_{t=1}^{T} \left(Z_t(x_i) - \bar{Z}_T(x_i) \right) \left(Z_t(x_j) - \bar{Z}_T(x_j) \right)$$

avec, pour tout $1 \leq i \leq n$:

$$\bar{Z}_T(x_i) = \frac{1}{T} \sum_{t=1}^{T} Z_t(x_i)$$

Notons $\hat{r}_T$, le vecteur constitué des $\dfrac{n(n+1)}{2}$ estimateurs empiriques $\hat{r}_T(x_i, x_j)$:

$$\hat{r}_T = \begin{pmatrix} \hat{r}_T(x_1, x_1) \\ \hat{r}_T(x_1, x_2) \\ \vdots \\ \hat{r}_T(x_n, x_n) \end{pmatrix}$$

$\hat{r}_T$ estime :

$$r_0 = \begin{pmatrix} r_Z(x_1, x_1) \\ r_Z(x_1, x_2) \\ \vdots \\ r_Z(x_n, x_n) \end{pmatrix}$$

Posons, pour tout $i \leq j$ et $i' \leq j'$:

$$C\left((i,j),(i',j')\right) = Cov\left((Z(x_i) - M_0)(Z(x_j) - M_0), (Z(x_{i'}) - M_0)(Z(x_{j'}) - M_0)\right)$$

et notons Σ_0 la matrice carrée $\left(\frac{n(n+1)}{2}\right) \times \left(\frac{n(n+1)}{2}\right)$ constituée des $C\left((i,j),(i',j')\right)$:

$$\Sigma_0 = \begin{pmatrix} C\left((1,1),(1,1)\right) & C\left((1,1),(1,2)\right) & \cdots & C\left((1,1),(1,n)\right) \\ C\left((1,2),(1,1)\right) & C\left((1,2),(1,2)\right) & \cdots & C\left((1,2),(1,n)\right) \\ \vdots & \vdots & \ddots & \vdots \\ C\left((n,n),(1,1)\right) & C\left((n,n),(1,2)\right) & \cdots & C\left((n,n),(n,n)\right) \end{pmatrix}$$

vérifiant l'hypothèse :

- **M2** : Σ_0 est inversible

L'estimateur empirique $\hat{\Sigma}_T$ de Σ_0 a pour terme général :

$$\begin{aligned} &\hat{C}_T\left((i,j),(i',j')\right) \\ &= \frac{1}{T}\sum_{t=1}^{T}(Z_t(x_i) - \bar{Z}_T(x_i))(Z_t(x_j) - \bar{Z}_T(x_j))(Z_t(x_{i'}) - \bar{Z}_T(x_{i'}))(Z_t(x_{j'}) - \bar{Z}_T(x_{j'})) \\ &-\hat{r}_T(x_i, x_j)\hat{r}_T(x_{i'}, x_{j'}) \end{aligned}$$

Sous l'hypothèse **M1**, la loi forte des grands nombres nous assure de la convergence *p.s.* de $\hat{C}_T\left((i,j),(i',j')\right)$ vers $C\left((i,j),(i',j')\right)$, $T \to +\infty$, pour tout $i \leq j$ et $i' \leq j'$. $\hat{\Sigma}_T$ converge donc *p.s.* vers Σ_0, $T \to +\infty$.

2.2.2 Comportements asymptotiques

Les propriétés que nous présentons dans ce paragraphe sont des résultats classiques et vont servir de résultats préliminaires pour l'étude des propriétés des estimateurs des paramètres α_0 et β_0 du modèle. Nous étudions successivement la convergence en moyenne quadratique, la convergence *p.s.* et la convergence en loi. Tous les résultats de ce paragraphe sont valables indépendamment de l'hypothèse **H0**.

Convergence en moyenne quadratique

Pour tout $i \leq j$, l'estimateur empirique de la covariance :

$$\hat{r}_T(x_i, x_j) = \frac{1}{T-1} \sum_{t=1}^{T} \left(Z_t(x_i) - \bar{Z}_T(x_i) \right) \left(Z_t(x_j) - \bar{Z}_T(x_j) \right)$$

est sans biais et de variance (*cf.* annexe A) :

$$\frac{1}{T(T-1)} \left((T-1) E \left((Z_t(x_i) - M_0)^2 (Z_t(x_j) - M_0)^2 \right) + \sigma_0^4 + (T-2) r_Z^2(x_i, x_j) \right)$$

Il converge donc en moyenne quadratique vers $r_Z(x_i, x_j)$, $T \to +\infty$, et, par suite, en probabilité.

Convergence *p.s.*

La loi forte des grands nombres nous assure de la convergence *p.s.* de $\hat{r}_T(x_i, x_j)$ vers $r_Z(x_i, x_j)$, $T \to +\infty$, pour tout $i \leq j$.

Convergence en loi

Rappelons le théorème de Slutsky sous sa forme univariée :

***Théorème* 2.1** *Si :*

$$a_T \underset{T \to +\infty}{\longrightarrow} a \text{ en probabilité}$$

avec a cste réelle, et si :

$$x_T \xrightarrow[T \to +\infty]{loi} x$$

avec x v.a.r., alors :

$$a_T x_T \xrightarrow[T \to +\infty]{loi} a x$$

★

***Lemme* 2.1** *Sous l'hypothèse **M1** :*

$$\sqrt{T} \left(\hat{r}_T - r_0 \right) \xrightarrow[T \to +\infty]{loi} \mathcal{N}_{\frac{n(n+1)}{2}} \left(0, \Sigma_0 \right)$$

★

Démonstration :

D'après le théorème de Cramer-Wold (Monfort (1980), p.162), une condition nécessaire et suffisante pour que :

$$\sqrt{T}\left(\hat{r}_T - r_0\right) \underset{T\to+\infty}{\xrightarrow{loi}} \mathcal{N}_{\frac{n(n+1)}{2}}\left(0, \Sigma_0\right)$$

est que, pour tout vecteur u de $\mathbb{R}^{\frac{n(n+1)}{2}}$:

$$u'\sqrt{T}\left(\hat{r}_T - r_0\right) \underset{T\to+\infty}{\xrightarrow{loi}} u'\mathcal{N}_{\frac{n(n+1)}{2}}\left(0, \Sigma_0\right)$$

Pour $u = (u_{1,1}, u_{1,2}, \cdots, u_{n,n})$, un vecteur quelconque de $\mathbb{R}^{\frac{n(n+1)}{2}}$, en utilisant la décomposition, pour tout $i \leq j$:

$$\begin{aligned}\hat{r}_T(x_i, x_j) &= \frac{1}{T-1}\sum_{t=1}^{T}\left(Z_t(x_i) - M_0\right)\left(Z_t(x_j) - M_0\right) \\ &- \frac{T}{T-1}\left(\bar{Z}_T(x_i) - M_0\right)\left(\bar{Z}_T(x_j) - M_0\right)\end{aligned}$$

on a la relation :

$$\begin{aligned}&u'\sqrt{T}\left(\hat{r}_T - r_0\right) \\ &= \frac{T}{T-1}\sqrt{T}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{i\leq j} u_{i,j}\left((Z_t(x_i) - M_0)(Z_t(x_j) - M_0) - r_Z(x_i, x_j)\right)\right] \\ &- \frac{T}{T-1}\sum_{i\leq j} u_{i,j}\sqrt{T}(\bar{Z}_T(x_i) - M_0)(\bar{Z}_T(x_j) - M_0) \\ &+ \frac{\sqrt{T}}{T-1}\sum_{i\leq j} u_{i,j} r_Z(x_i, x_j)\end{aligned}$$

Ayant supposé que $(Z_t(x_1), \ldots, Z_t(x_n))'_{t\geq 1}$ constitue une suite de vecteurs aléatoires indépendants et de même loi, possédant des moments d'ordre 4 (hypothèse **M1**), on en déduit que $((Z_t(x_i) - M_0)(Z_t(x_j) - M_0) - r_Z(x_i, x_j), i \leq j)'_{t\geq 1}$ constitue une suite de vecteurs aléatoires indépendants et centrés, possédant des moments d'ordre 2. Grâce au théorème limite centrale multidimensionnel sur la somme de vecteurs aléatoires réels *i.i.d.* centrés et de carré intégrable (Monfort (1980), p.144), on déduit :

$$\sqrt{T}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{i\leq j} u_{i,j}\left((Z_t(x_i) - M_0)(Z_t(x_j) - M_0) - r_Z(x_i, x_j)\right)\right] \underset{T\to+\infty}{\xrightarrow{loi}} \mathcal{N}_1\left(0, u'\Sigma_0 u\right)$$

On a aussi, pour tout $1 \leq i \leq n$:

$$\sqrt{T}(\bar{Z}_T(x_i) - M_0) \underset{T\to+\infty}{\xrightarrow{loi}} \mathcal{N}_1\left(0, \sigma_0^2\right)$$

et, d'après la loi faible des grands nombres de Kintchine, pour tout $1 \leq j \leq n$:

$$\bar{Z}_T(x_j) \underset{T\to+\infty}{\longrightarrow} M_0, \text{ en probabilité}$$

Une application répétée du théorème de Slutsky (*cf.* théorème 2.1) à $\left(\bar{Z}_T(x_j) - M_0\right)$ et $\sqrt{T}(\bar{Z}_T(x_i) - M_0)$ pour tout $i \leq j$ donne :

$$\sum_{i\leq j} u_{i,j}\sqrt{T}(\bar{Z}_T(x_i) - M_0)(\bar{Z}_T(x_j) - M_0) \underset{T\to+\infty}{\overset{loi}{\longrightarrow}} 0$$

Bien entendu :

$$\frac{\sqrt{T}}{T-1}\sum_{i\leq j} u_{i,j} r_Z(x_i, x_j) \underset{T\to+\infty}{\longrightarrow} 0$$

On obtient finalement :

$$u'\sqrt{T}\left(\hat{r}_T - r_0\right) \underset{T\to+\infty}{\overset{loi}{\longrightarrow}} \mathcal{N}_1\left(0, u'\Sigma_0 u\right)$$

■

Loi du logarithme itéré

***Lemme* 2.2** *Sous les hypothèses **M1-M2**, l'ensemble :*

$$\left\{\frac{T\Sigma_0^{-1/2}\left(\hat{r}_T - r_0\right)}{\sqrt{2T\ln\ln(T)}}\right\}$$

est p.s. relativement compact dans la boule unité de $\mathbb{R}^{n(n+1)/2}$.

★

Démonstration :

En utilisant la décomposition, pour tout $i \leq j$:

$$\begin{aligned}
\hat{r}_T(x_i, x_j) - r_Z(x_i, x_j) &= \frac{1}{T-1}\left[\sum_{t=1}^{T}(Z_t(x_i) - M_0)(Z_t(x_j) - M_0) - r_Z(x_i, x_j)\right] \\
&+ \frac{1}{T-1} r_Z(x_i, x_j) \\
&- \frac{T}{T-1}(\bar{Z}_T(x_i) - M_0)(\bar{Z}_T(x_j) - M_0)
\end{aligned}$$

on obtient la relation entre vecteurs :

$$\hat{r}_T - r_0 = \frac{1}{T-1}\Omega\mathbb{1}_T + \frac{1}{T-1}r_0 - \frac{T}{T-1}\left(\left(\bar{Z}_T(x_i) - M_0\right)\left(\bar{Z}_T(x_j) - M_0\right)\right)_{i\leq j}$$

avec $\Omega_T = [(Z_t(x_i) - M_0)(Z_t(x_j) - M_0) - r_Z(x_i, x_j)]_{(i \leq j),(1 \leq t \leq T)}$ une matrice $\left(\frac{n(n+1)}{2}, T\right)$.

On a donc :

$$\begin{aligned}\frac{\Sigma_0^{-1/2}(T-1)(\hat{r}_T - r_0)}{\sqrt{2T \ln\ln(T)}} &= \frac{\Sigma_0^{-1/2}(\Omega_T \mathbb{1}_T)}{\sqrt{2T \ln\ln(T)}} \\ &+ \frac{\Sigma_0^{-1/2} r_0}{\sqrt{2T \ln\ln(T)}} \\ &+ \frac{T\Sigma_0^{-1/2}\left((\bar{Z}_T(x_i) - M_0)(\bar{Z}_T(x_j) - M_0)\right)_{i \leq j}}{\sqrt{2T \ln\ln(T)}}\end{aligned}$$

(i) On a :

$$\frac{\Sigma_0^{-1/2} r_0}{\sqrt{2T \ln\ln(T)}} \underset{T \to +\infty}{\longrightarrow} 0$$

(ii) De plus :

$$\frac{T\Sigma_0^{-1/2}\left((\bar{Z}_T(x_i) - M_0)(\bar{Z}_T(x_j) - M_0)\right)_{i \leq j}}{\sqrt{2T \ln\ln(T)}} = \Sigma_0^{-1/2}\left(\frac{\sum_{t=1}^{T}(Z_t(x_i) - M_0)}{\sqrt{2T \ln\ln(T)}} \times (\bar{Z}_T(x_j) - M_0)\right)_{i \leq j}$$

Pour tout $1 \leq i \leq n$, $(Z_t(x_i) - M_0)_{t \geq 1}$ est une suite de *v.a.r. i.i.d.*, centrées et de variance σ_0^2. D'après une version de la loi du logarithme itéré (Hartman et Wintner (1941)), on en déduit que l'ensemble :

$$\left\{\frac{\sum_{t=1}^{T}(Z_t(x_i) - M_0)}{\sqrt{2T \ln\ln(T)}}\right\}$$

est *p.s.* relativement compact dans $[-\sigma_0, \sigma_0]$. Comme, pour tout $1 \leq j \leq n$:

$$\bar{Z}_T(x_j) \underset{T \to +\infty}{\longrightarrow} M_0, \; p.s.$$

on a finalement :

$$\frac{T\Sigma_0^{-1/2}\left((\bar{Z}_T(x_i) - M_0)(\bar{Z}_T(x_j) - M_0)\right)_{i \leq j}}{\sqrt{2T \ln\ln(T)}} \underset{T \to +\infty}{\longrightarrow} 0, \; p.s.$$

(iii) Enfin, $\Omega_T \mathbb{1}_T$ constitue une somme de vecteurs aléatoires indépendants centrés et de matrice de variances-covariances Σ_0 inversible. D'après un résultat de Senoussi (1990), pp.42-44), on en déduit alors que l'ensemble :

$$\left\{ \frac{\Sigma_0^{-1/2}\left(\Omega_T \mathbb{1}_T\right)}{\sqrt{2T\ln\ln(T)}} \right\}$$

est $p.s.$ relativement compact dans la boule unité de $\mathbb{R}^{n(n+1)/2}$.

On a aussi :

$$\frac{\Sigma_0^{-1/2}(T-1)\left(\hat{r}_T - r_0\right)}{\sqrt{2T\ln\ln(T)}} = \left(\frac{T-1}{T}\right)\frac{T\Sigma_0^{-1/2}\left(\hat{r}_T - r_0\right)}{\sqrt{2T\ln\ln(T)}}$$

avec $\dfrac{T-1}{T} \underset{T\to+\infty}{\longrightarrow} 1$

Par conséquent, l'ensemble :

$$\left\{ \frac{T\Sigma_0^{-1/2}\left(\hat{r}_T - r_0\right)}{\sqrt{2T\ln\ln(T)}} \right\}$$

est $p.s.$ relativement compact dans la boule unité de $\mathbb{R}^{n(n+1)/2}$.

■

Remarque : le cas gaussien

Dans le cas où le processus Z est gaussien, on peut exprimer chacun des termes de la matrice de variances-covariances Σ_0 simplement en fonction des covariances :

– pour tout $i \leq j$:

$$Var\left((Z(x_i) - M_0)(Z(x_j) - M_0)\right) = r_Z(x_i, x_j)^2 + \sigma_0^4$$

– pour tout $i \leq j$ et $i' \leq j'$:

$$Cov\left((Z(x_i) - M_0)(Z(x_j) - M_0), (Z(x_{i'}) - M_0)(Z(x_{j'}) - M_0)\right)$$
$$= r_Z(x_i, x_{j'})r_Z(x_{i'}, x_j) + r_Z(x_i, x_{i'})r_Z(x_j, x_{j'})$$

Cela vient du fait que si $(X_1, X_2, X_3, X_4)'$ suit une loi normale d'espérance nulle :

$$E(X_1X_2X_3X_4) = E(X_1X_2)E(X_3X_4) + E(X_1X_3)E(X_2X_4) + E(X_1X_4)E(X_2X_3)$$

Cette relation nous permet donc, dans le cas gaussien, d'estimer Σ_0 simplement à partir des covariances empiriques $\hat{r}(x_i, x_j)$, $i \leq j$.

♣

2.3 Estimation des paramètres et propriétés des estimateurs

Dans le paragraphe 2.3.1, nous définissons le processus de contraste, appelé encore contraste par abus de langage, dont la minimisation fournit les estimateurs des paramètres α_0 et β_0 relatifs à la transformation $f(.;\alpha_0)$ et à la fonction de covariance $R(.;\beta_0)$. Le théorème 2.2 fournit un résultat de convergence $P_0-p.s.$ de ce contraste vers une quantité positive ou nulle appelée fonction de contraste, selon la terminologie de Dacunha-Castelle et Duflo ((1993), p.93) (*cf.* également Hardouin (1992), p.6, Guyon (1993), p.106, ou encore Bayomog et *al* (1996), pp.115-116). Cette fonction est notée :

$$\theta \longmapsto K(\theta_0, \theta)$$

Dans le paragraphe 2.3.2, on montre l'existence des estimateurs du minimum de contraste que l'on note $\hat{\alpha}_T$ et $\hat{\beta}_T$. La convergence $P_0 - p.s.$ et la convergence en loi de ces estimateurs sont étudiées au paragraphe 2.3.3.

Précisons que cette méthode d'estimation, par minimum de contraste, permet de se placer dans un cadre très général avec des hypothèses *a priori* sur la loi du processus étudié très faibles.

2.3.1 Définition et convergence $P_0 - p.s.$ du processus de contraste

Définition

On définit le contraste $U_T : \theta \longmapsto U_T(\theta)$ comme la somme des écarts au carré entre les covariances empiriques $\hat{r}_T(x_i, x_j)$ et les covariances théoriques $r(x_i, x_j; \theta)$ sur les $i \leq j$, c'est-à-dire, pour tout θ :

$$U_T(\theta) = \sum_{i \leq j} \left[\hat{r}_T(x_i, x_j) - r(x_i, x_j; \theta)\right]^2$$

où l'on rappelle que :

$$\left\{ \begin{array}{rcl} \hat{r}_T(x_i, x_j) & = & \dfrac{1}{T-1} \displaystyle\sum_{t=1}^{T} \left(Z_t(x_i) - \bar{Z}_T(x_i)\right) \left(Z_t(x_j) - \bar{Z}_T(x_j)\right) \\ r(x_i, x_j; \theta) & = & R\left(f(x_j; \alpha) - f(x_i; \alpha); \beta\right) \end{array} \right.$$

avec $\theta = (\alpha, \beta)$.

On rappelle aussi que sous l'hypothèse **H0**, pour tout (x, x') dans $[0, 1]^2$:

$$r_Z(x, x) = r(x, x'; \theta_0)$$

Convergence $P_0 - p.s.$ du contraste

On a vu au paragraphe 2.2.2 que les estimateurs empiriques $\hat{r}_T(x_i, x_j)$, $i \leq j$, convergent $p.s.$ vers $r(x_i, x_j; \theta_0)$, $T \to +\infty$, ce qui entraîne *ipso facto* le résultat :

***Théorème* 2.2** *Pour tout θ :*

$$U_T(\theta) \underset{T \to +\infty}{\longrightarrow} K(\theta_0, \theta), \; P_0 - p.s.$$

avec :

$$K(\theta_0, \theta) = \sum_{i \leq j} [r(x_i, x_j; \theta_0) - r(x_i, x_j; \theta)]^2 \geq 0$$

★

$K(\theta_0, \theta)$ est nulle pour $\theta_0 = \theta$. On donne ci-après une condition d'identifiabilité paramétrique du modèle (hypothèse **B3**), c'est-à-dire une condition assurant $K(\theta_0, \theta)$ nulle si et seulement si $\theta_0 = \theta$.

2.3.2 Existence de $\hat{\theta}_T = (\hat{\alpha}_T, \hat{\beta}_T)$

Considérons les hypothèses suivantes :

- **B1** : $\theta = (\alpha, \beta) \in \Theta$ compact de $\mathbb{R}^{p+q}$, $(p, q) \in \mathbb{N}^{*2}$, et $\theta_0 \in \mathring{\Theta}$
- **B2** : $\theta \longmapsto r(x, x'; \theta)$ est continue dans Θ, pour tout (x, x') dans $[0, 1]^2$
- **B3** : l'application :

$$\theta \longmapsto r(\theta) = \begin{pmatrix} r(x_1, x_1; \theta) \\ r(x_1, x_2; \theta) \\ \vdots \\ r(x_n, x_n; \theta) \end{pmatrix}$$

  est injective

La condition de compacité de l'espace des paramètres est classique pour ce type de problème. La condition **B3** assure l'identifiabilité paramétrique du modèle ; sous cette condition $K(\theta_0, .)$ admet un minimum unique en $\theta = \theta_0$, c'est-à-dire que cette fonction sépare θ et θ_0. L'hypothèse **B3** signifie que, pour le réseau de points de mesure $x_1, x_2, \ldots, x_n$ dont on dispose, la paramétrisation en θ du modèle de covariance r_Z est propre. D'une manière générale, la condition **B3** est vérifié pour un choix "correct" des positions $x_i, 1 \leq i \leq n$. Mais dans la pratique, où la valeur du paramètre θ_0 est inconnue, il n'est pas toujours simple de faire un tel choix fondé sur des méthodes analytiques. On suppose donc *a priori* cette condition vérifiée. On verra dans la section 2.6, sur un exemple, comment s'assurer de l'hypothèse **B3**.

Pour prouver l'existence d'un estimateur du minimum de contraste $\hat{\theta}_T = (\hat{\alpha}_T, \hat{\beta}_T)$ qui soit une fonction mesurable des observations $Z_t(x_i), 1 \leq i \leq n, 1 \leq t \leq T$, on utilise un lemme de Jennrich ((1969), p.637).

***Théorème* 2.3** *Sous les hypothèses **B1-B2**, il existe une fonction mesurable $\hat{\theta}_T$ des covariances empiriques $\hat{r}_T(x_i, x_j)$, $i \leq j$, et a posteriori des observations $Z_t(x_i)$, $1 \leq i \leq n$ et $1 \leq t \leq T$ telle que :*

$$U_T(\hat{\theta}_T) = \inf_{\theta \in \Theta} U_T(\theta)$$

★

Démonstration :

Elle suit exactement la démarche de Jennrich (1969).

■

2.3.3 Propriétés asymptotiques de $\hat{\theta}_T = (\hat{\alpha}_T, \hat{\beta}_T)$

Dans un premier temps, on s'intéresse à la convergence $P_0 - p.s.$, puis on étudie la convergence en loi.

Convergence $P_0 - p.s.$

Considérons l'hypothèse suivante qui renforce **M1**:

- **M3** : $\exists \nu > 2, \sup_{[0,1]} E|Z(x)|^{2\nu} = M_3 < +\infty$

et rappelons une inégalité de grande déviation :

***Théorème* 2.4** *(Brillinger (1962))*
Soient $X_1, X_2, \ldots, X_n$ un échantillon i.i.d. de la variable aléatoire parente X centrée. S'il existe $\zeta \geq 2$ tel que :

$$E|X|^\zeta < +\infty$$

alors, pour tout $\epsilon > 0$:

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i\right| \geq \epsilon\right) \leq \frac{K_\zeta(E|X|^\zeta + 1)}{\epsilon^\zeta n^{\zeta/2}}$$

où K_ζ est une constante positive qui dépend uniquement de ζ. ★

Théorème* 2.5** *Sous les hypothèses* ***B1-B3 *et* ***M3****, les estimateurs du minimum de contraste,* $\hat{\alpha}_T$ *et* $\hat{\beta}_T$*, convergent* $P_0 - p.s.$ *vers* α_0 *et* β_0*, c'est-à-dire :*

$$\hat{\theta}_T \underset{T \to +\infty}{\longrightarrow} \theta_0, \; P_0 - p.s.$$

★

Démonstration :

Notre démarche suit celle de Hardouin ((1992), pp.6-8) qui donne la version *p.s.* du résultat de Dacunha-Castelle et Duflo ((1993), p.94). Les conditions d'application sont les suivantes :

- *(i)* Θ est compact et $\theta_0 \in \mathring{\Theta}$
- *(ii)* $\theta \longmapsto U_T(\theta)$ est continue $P_0 - p.s.$ et $\theta \longmapsto K(\theta_0, \theta)$ est continue, dans Θ
- *(iii)* $K(\theta_0, \theta) = 0 \Longleftrightarrow \theta = \theta_0$
- *(iv)* $\liminf_{T \to +\infty} U_T(\theta) \geq K(\theta_0, \theta), \; P_0 - p.s.$
- *(v)* Notant W_T le module de continuité de U_T :

$$W_T(\eta) = \sup \left\{ |U_T(\theta) - U_T(\theta')|, \|\theta - \theta'\|_{p+q} \leq \eta, (\theta, \theta') \in \Theta^2 \right\}$$

  il existe une suite (ϵ_k) qui tend vers zéro, $k \to +\infty$, telle que, pour tout $k \geq 1$:

$$P_0 \left(\limsup_{T \to +\infty} \left\{ W_T \left(\frac{1}{k} \right) \geq \epsilon_k \right\} \right) = 0$$

Il est immédiat que les hypothèses *(i)*, *(ii)* et *(iii)* sont assurées sous les conditions **B1**-**B3**. L'hypothèse *(iv)*, quant à elle, est vérifiée grâce au théorème 2.2.

Reste à montrer que *(v)* est aussi vérifiée. Une décomposition de $[U_T(\theta') - U_T(\theta)]$ fournit l'inégalité :

$$\begin{aligned}
&|U_T(\theta') - U_T(\theta)| \\
&\leq \left| \sum_{i \leq j} (r(x_i, x_j; \theta') + r(x_i, x_j; \theta)) \times (r(x_i, x_j; \theta') - r(x_i, x_j; \theta)) \right| \\
&+ 2 \left| \sum_{i \leq j} (\hat{r}_T(x_i, x_j) - r(x_i, x_j; \theta_0)) \times (r(x_i, x_j; \theta') - r(x_i, x_j; \theta)) \right| \\
&+ 2 \left| \sum_{i \leq j} r(x_i, x_j; \theta_0) \, (r(x_i, x_j; \theta') - r(x_i, x_j; \theta)) \right|
\end{aligned}$$

Puisque $|r(x_i, x_j; \theta)| \leq \sqrt{r(x_i, x_i; \theta)}$ et que σ^2 appartient à un compact de $\mathbb{R}$, on a, pour tout $i \leq j$:

$$\sup_{\theta \in \Theta} |r(x_i, x_j; \theta)| \leq M < +\infty$$

On obtient alors :

$$\begin{aligned} |U_T(\theta') - U_T(\theta)| &\leq 2Mn(n+1) \sup_{i \leq j} |r(x_i, x_j; \theta') - r(x_i, x_j; \theta)| \\ &+ 4M \sum_{i \leq j} |\hat{r}_T(x_i, x_j) - r(x_i, x_j; \theta_0)| \end{aligned}$$

Par conséquent, pour tout k dans $\mathbb{N}^*$:

$$\begin{aligned} W_T\left(\frac{1}{k}\right) &\leq 2Mn(n+1) \sup_{\|\theta - \theta'\|_{p+q} \leq \frac{1}{k}} \sup_{i \leq j} |r(x_i, x_j; \theta') - r(x_i, x_j; \theta)| \\ &+ 4M \sum_{i \leq j} |\hat{r}_T(x_i, x_j) - r(x_i, x_j; \theta_0)| \end{aligned}$$

Posons, pour tout k dans $\mathbb{N}^*$:

$$u_k = 2Mn(n+1) \sup_{\|\theta - \theta'\|_{p+q} \leq \frac{1}{k}} \sup_{i \leq j} |r(x_i, x_j; \theta') - r(x_i, x_j; \theta)|$$

et considérons l'ensemble, pour une suite ϵ_k strictement positive :

$$B_{k,T} = \left\{ u_k + 4M \sum_{i \leq j} |\hat{r}_T(x_i, x_j) - r(x_i, x_j; \theta_0)| \geq \epsilon_k \right\}$$

On a alors pour tout k :

$$0 \leq P_0\left(\limsup_{T \to +\infty} \left\{ W_T\left(\frac{1}{k}\right) \geq \epsilon_k \right\} \right) \leq P_0\left(\limsup_{T \to +\infty} B_{k,T} \right)$$

Prenons désormais :

$$\epsilon_k = u_k + \frac{1}{k}$$

Comme Θ est compact (hypothèse **B1**), on déduit de l'hypothèse **B2** que $\theta \longmapsto r(x, x'; \theta)$ est uniformément continue dans Θ. On a donc :

$$\lim_{k \to +\infty} \epsilon_k = 0$$

Pour vérifier *(v)*, il reste à montrer que, pour tout k dans $\mathbb{N}^*$:

$$P_0\left(\limsup_{T \to +\infty} B_{k,T} \right) = 0$$

D'après le théorème de Borel-Cantelli, il est suffisant de vérifier que :

$$\sum_{T=1}^{+\infty} P_0 \left(4M \sum_{i \leq j} |\hat{r}_T(x_i, x_j) - r(x_i, x_j; \theta_0)| \geq \frac{1}{k} \right) < +\infty \tag{2.1}$$

Les inégalités de Bonferroni (Durret (1995), p.22) nous donnent la majoration :

$$\begin{aligned} & P_0 \left(4M \sum_{i \leq j} |\hat{r}_T(x_i, x_j) - r(x_i, x_j; \theta_0)| \geq \frac{1}{k} \right) \\ & \quad \leq \sum_{i \leq j} P_0 \left(|\hat{r}_T(x_i, x_j) - r(x_i, x_j; \theta_0)| \geq \frac{1}{2Mn(n+1)k} \right) \end{aligned}$$

En utilisant la décomposition, pour tout $i \leq j$:

$$\begin{aligned} \hat{r}_T(x_i, x_j) &= \frac{1}{T-1} \sum_{t=1}^{T} (Z_t(x_i) - M_0)(Z_t(x_j) - M_0) \\ &- \frac{T}{T-1} \left(\bar{Z}_T(x_i) - M_0 \right) \left(\bar{Z}_T(x_j) - M_0 \right) \end{aligned}$$

on obtient, en posant $\rho = \dfrac{1}{2Mn(n+1)k}$:

$$\begin{aligned} & P_0 \left(|\hat{r}_T(x_i, x_j) - r(x_i, x_j; \theta_0)| \geq \rho \right) \\ & \leq P_0 \left(\frac{1}{T-1} \left| \sum_{t=1}^{T} [(Z_t(x_i) - M_0)(Z_t(x_j) - M_0) - r(x_i, x_j; \theta_0)] \right| \geq \frac{\rho}{2} \right) \\ & + P_0 \left(\frac{1}{T-1} \left| r(x_i, x_j; \theta_0) - T(\bar{Z}_T(x_i) - M_0)(\bar{Z}_T(x_j) - M_0) \right| \geq \frac{\rho}{2} \right) \end{aligned}$$

Choisissons $T > 2 \vee \left(\dfrac{4M + \rho}{\rho} \right)$:

$$\begin{aligned} & P_0 \left(|\hat{r}_T(x_i, x_j) - r(x_i, x_j; \theta_0)| \geq \rho \right) \\ & \leq P_0 \left(\frac{1}{T} \left| \sum_{t=1}^{T} [(Z_t(x_i) - M_0)(Z_t(x_j) - M_0) - r(x_i, x_j; \theta_0)] \right| \geq \frac{\rho}{4} \right) \\ & + P_0 \left(|\bar{Z}_T(x_i) - M_0||\bar{Z}_T(x_j) - M_0| \geq \frac{\rho}{8} \right) \\ & \leq P_0 \left(\frac{1}{T} \left| \sum_{t=1}^{T} [(Z_t(x_i) - M_0)(Z_t(x_j) - M_0) - r(x_i, x_j; \theta_0)] \right| \geq \frac{\rho}{4} \right) \\ & + P_0 \left(|\bar{Z}(x_i) - M_0| \geq \frac{\sqrt{\rho}}{2\sqrt{2}} \right) \\ & + P_0 \left(|\bar{Z}(x_j) - M_0| \geq \frac{\sqrt{\rho}}{2\sqrt{2}} \right) \end{aligned}$$

Pour s'assurer de (2.1), il suffit de majorer chacune des 3 dernières probabilités ci-dessus par le terme d'une série convergente. Pour cela, on s'appuie sur l'inégalité de grande déviation fournie par le théorème 2.4. D'après ce résultat, on a grâce à l'hypothèse **M3**, pour tout $i \leq j$:

$$P_0\left(|\hat{r}_T(x_i, x_j) - r(x_i, x_j; \theta_0)| \geq \rho\right) = O\left(\frac{1}{T^{\nu/2}}\right)$$

et on obtient (2.1). ∎

Remarque : la convergence en probabilité

Pour montrer la convergence en P_0-probabilité nous n'avons besoin que des hypothèses **B1**-**B3**. Les hypothèses *(i)*-*(iii)* restent inchangées, et les hypothèses *(iv)* et *(v)* sont affaiblies en :

- *(iv)'* $\liminf\limits_{T\to+\infty} U_T(\theta) \geq K(\theta_0, \theta)$, en P_0 − probabilité
- *(v)'* $\lim\limits_{T\to+\infty} P_0\left(W_T(\frac{1}{k}) \geq \epsilon_k\right) = 0$, pour tout $k \geq 1$

Cette fois, pour s'assurer que l'hypothèse *(v)'* est vérifiée, on vérifie facilement que $\lim\limits_{T\to+\infty} P_0\left(B_{k,T}\right) = 0$, pour tout $k \geq 1$, où $B_{k,T}$ a été défini dans la démonstration précédente. Cette vérification s'effectue grâce à la propriété de convergence en P_0-probabilité de $\hat{r}_T(x_i, x_j)$ vers $r(x_i, x_j; \theta_0)$ vue au paragraphe 2.2.2.

♣

Convergence $P_0-p.s.$ dans le cas de contrastes convexes avec déformations splines et covariances stationnaires strictement décroissantes sur $[0, 1]$

Nous supposons ici que β_0 est connu, et nous notons R_0 la covariance stationnaire $R(.;\beta_0)$. Nous supposons de plus que R_0 est strictement décroissante sur $[0, 1]$, ce qui est le cas, par exemple, pour :

- la famille des modèles sphériques (*cf.* figure 4.8, chapitre 4) :

$$R_0(u) = \sigma_0^2\left(1 - a_0|u| + \frac{4a_0^3|u|^3}{27}\right),\ 0 < a_0 \leq \frac{3}{2},\ \sigma_0 > 0$$

- la famille des modèles exponentiels (*cf.* figure 2.3) :

$$R_0(u) = \sigma_0^2 \exp\left(-a_0|u|\right),\ a_0 > 0,\ \sigma_0 > 0$$

- la famille des modèles à covariance gaussienne (*cf.* figure 3.4, chapitre 3) :

$$R_0(u) = \sigma_0^2 \exp\left(-a_0 u^2\right),\ a_0 > 0,\ \sigma_0 > 0$$

Nous prenons comme famille de modèles de déformation $\mathcal{F}$ l'ensemble des fonctions $f(.;\alpha)$, avec $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p)$, définies pour tout x dans $[0,1]$ par :

$$f(x;\alpha) = \sum_{k=1}^{p} \alpha_k I_k^l(x) \tag{2.2}$$

où les fonctions I_k^l, $k = 1, \ldots, p$, sont les I-splines de degré l, $l \in \mathbb{N}^*$, associées aux $m = p - l > 0$ nœuds $0 < a_1 < a_2 < \cdots < a_m < 1$. Ces fonctions I-splines, appelées splines monotones, sont définies dans l'annexe C. Disons simplement ici que les I-splines de degré l, $l \geq 1$, sont des polynômes croissants de degré l par intervalle, avec raccordements continus des dérivées jusqu'à l'ordre $l-1$ aux extrémités de l'intervalle. Dans le cas des I-splines de degré 1, on obtient les fonctions affines croissantes par intervalle, avec raccordements continus aux extrémités de l'intervalle, appelées encore lignes brisées (*cf.* figure 2.1).

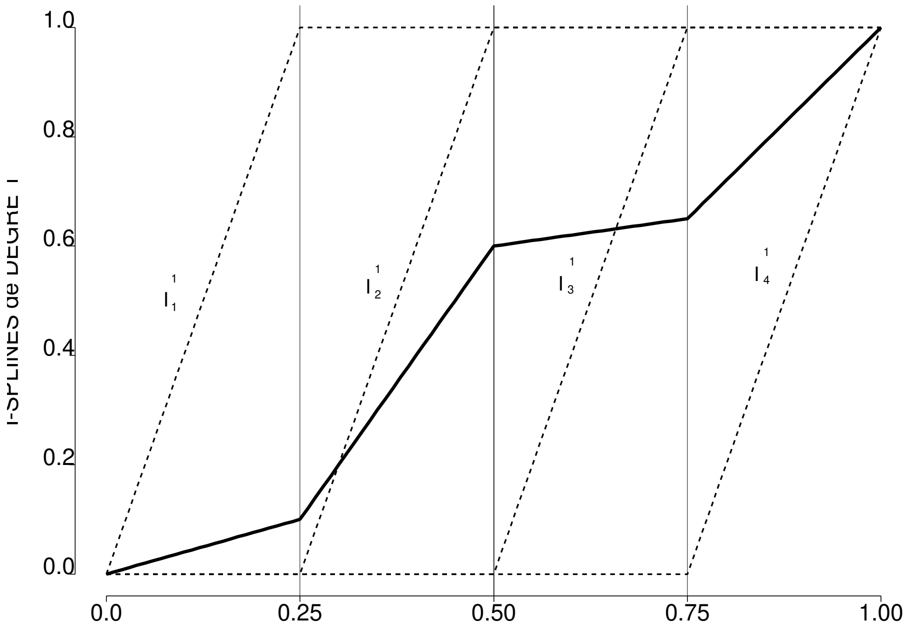


FIG. 2.1 – *Les quatre fonctions I-splines (en pointillé) de degré 1 définies sur* $[0,1]$ *associées aux 3 nœuds intérieurs* 0.25, 0.50 *et* 0.75 *et leur combinaison convexe, fournissant* $0.10 I_1^1 + 0.50 I_2^1 + 0.05 I_3^1 + 0.35 I_4^1$, *représentée par la courbe en trait fort*

De plus, pour tout $k = 1, \ldots, p$, I_k^l dans (2.2) vérifie, en vertu des propriétés des I-splines (*cf.* Ramsay (1988), pp.425-429, ou l'annexe C) :

$$\begin{cases} I_k^l \text{ est continue et croissante dans } [0,1] \\ I_k^l(0) = 0 \\ I_k^l(1) = 1 \end{cases}$$

Pour assurer la contrainte de stricte croissance imposée aux déformations et la contrainte $f(1;\alpha) = 1$, il suffit donc que les p paramètres vérifient :

$$(S_1) : \begin{cases} \alpha_k > 0, \text{ pour tout } k \text{ dans } [1, \ldots, p] \\ \displaystyle\sum_{k=1}^{p} \alpha_k = 1 \end{cases}$$

α appartient donc à un ouvert convexe de $\mathbb{R}^p$. Nous notons $\alpha_0 = (\alpha_{1,0}, \ldots, \alpha_{p,0})$ la vraie valeur du paramètre α.

Le contraste s'écrit :

$$U_T(\alpha) = \sum_{i \leq j} \left[\hat{r}_T(x_i, x_j) - r(x_i, x_j; \alpha)\right]^2$$

où l'on rappelle que :

$$\begin{cases} \hat{r}_T(x_i, x_j) &= \displaystyle\frac{1}{T-1} \sum_{t=1}^{T} \left(Z_t(x_i) - \bar{Z}_T(x_i)\right)\left(Z_t(x_j) - \bar{Z}_T(x_j)\right) \\ r(x_i, x_j; \alpha) &= R_0\left(f(x_j;\alpha) - f(x_i;\alpha)\right) \end{cases}$$

R_0 étant strictement décroissante sur $[0,1]$ nous avons, pour tout $i < j$:

$$R_0\left(f(x_j;\alpha) - f(x_i;\alpha)\right) = r(x_i, x_j; \alpha_0) \Longleftrightarrow f(x_j;\alpha) - f(x_i;\alpha) = R_0^{-1}\left(r(x_i, x_j; \alpha_0)\right)$$

Le fait de connaître R_0, et en particulier la variance σ_0^2, nous permet de ne considérer que les couples de stations (x_i, x_j) avec $i < j$ et de définir un nouveau contraste par :

$$\begin{aligned} U_T^*(\alpha) &= \sum_{i<j} \left[R_0^{-1}\left(\hat{r}_T(x_i, x_j)\right) - \left(f(x_j;\alpha) - f(x_i;\alpha)\right)\right]^2 \\ &= \sum_{i<j} \left[R_0^{-1}\left(\hat{r}_T(x_i, x_j)\right) - \sum_{k=1}^{p} \alpha_k \left(I_k^l(x_j) - I_k^l(x_i)\right)\right]^2 \end{aligned}$$

U_T^* est un polynôme convexe de degré 2 par rapport aux α_k pour tout $T \geq 2$, ce qui est satisfaisant d'un point de vue numérique car on est assuré de l'unicité de l'estimateur $\hat{\alpha}_T$

du minimum de contraste U_T^*. Pour tout α, $U_T^*(\alpha)$ converge $P_0 - p.s$ vers la fonction de contraste :

$$\begin{aligned} K(\alpha_0, \alpha) &= \sum_{i<j} \left[R_0^{-1}\left(r(x_i, x_j; \alpha_0)\right) - \sum_{k=1}^{p} \alpha_k \left(I_k^l(x_j) - I_k^l(x_i)\right)\right]^2 \\ &= \sum_{i<j} \left[\sum_{k=1}^{p} (\alpha_{k,0} - \alpha_k)\left(I_k^l(x_j) - I_k^l(x_i)\right)\right]^2 \end{aligned}$$

Sous l'hypothèse d'identifiabilité :

- **B3'** : le rang de la matrice

$$\begin{pmatrix} I_1^l(x_2) - I_1^l(x_1) & \cdots & I_p^l(x_2) - I_p^l(x_1) \\ I_1^l(x_3) - I_1^l(x_1) & \cdots & I_p^l(x_3) - I_p^l(x_1) \\ \vdots & \vdots & \vdots \\ I_1^l(x_n) - I_1^l(x_{n-1}) & \cdots & I_p^l(x_n) - I_p^l(x_{n-1}) \end{pmatrix}_{(n(n-1)/2),p}$$

  est égal à p

  on déduit :

$$K(\alpha_0, \alpha) = 0 \Longleftrightarrow \alpha = \alpha_0$$

Pour conclure sur la convergence $P_0 - p.s.$ de l'estimateur $\hat{\alpha}_T$ du minimum de contraste U_T^* vers α_0, on applique directement le résultat de Senoussi (1990) :

***Théorème* 2.6** *Soit Θ un ouvert convexe de $\mathbb{R}^p$, $g_n : \Theta \longmapsto \mathbb{R}$ une suite de fonctions convexes convergeant simplement vers g, telle que α_0 est l'unique point réalisant le minimum de g.*

Alors, si α_n réalise le minimum de g_n :

$$\lim_{n \to +\infty} \alpha_n = \alpha_0$$

★

D'un point de vue numérique, la minimisation du contraste U_T^* peut être abordée de trois façons :

- soit on minimise U_T^* sous les contraintes (S_1) ;
- soit T est suffisamment grand pour penser que ces contraintes (S_1) seront automatiquement vérifiées pour l'estimateur obtenu ;

– soit on adopte la paramétrisation :

$$\alpha_k = \frac{e^{\theta_k}}{1+\sum_{i=2}^{p} e^{\theta_i}}, \text{ avec } \theta_1 = 0, k = 1, \ldots, p$$

et le problème de minimisation du contraste re-paramétré devient non contraint.

Propriété asymptotique du biais

Puisque la suite $(\hat{\theta}_T)$ est bornée $P_0 - p.s.$ ($\hat{\theta}_T \in \Theta$ compact), donc équi-intégrable, et qu'elle converge $P_0 - p.s.$ vers θ_0, elle converge en moyenne vers θ_0 et le biais tend vers 0, $T \to +\infty$.

Convergence en loi

Pour la démonstration de la convergence en loi, les réalisations dans le temps étant *i.i.d.*, nous allons nous inspirer de la démonstration de Dacunha-Castelle et Duflo ((1993), pp.102-103). On peut trouver la généralisation de cette démonstration au cas non-ergodique dans Hardouin ((1992), pp.34-36), Guyon ((1993), pp.110-111) ou encore dans Bayomog et *al* ((1996), pp.118-119). Tout d'abord, nous introduisons les hypothèses :

– **D1** : $\theta \longmapsto r(x, x'; \theta)$ est de classe C^2 dans $\mathring{\Theta}$, pour tout (x, x') dans $[0,1]^2$

– **D2** : la matrice jacobienne de l'application :

$$\theta \longmapsto r(\theta) = \begin{pmatrix} r(x_1, x_1; \theta) \\ r(x_1, x_2; \theta) \\ \vdots \\ r(x_n, x_n; \theta) \end{pmatrix}$$

est de rang $(p+q)$ en θ_0

puis nous considérons les matrices :

$$\begin{aligned}
R_{p+q}(\theta) &= 2\left[r^{(1)}(x_1, x_1; \theta), r^{(1)}(x_1, x_2; \theta), \ldots, r^{(1)}(x_n, x_n; \theta)\right]_{(p+q), \frac{n(n+1)}{2}} \\
J_{p+q}(\theta_0) &= R_{p+q}(\theta_0)\Sigma_0 R'_{p+q}(\theta_0) \\
J_{p+q}(\hat{\theta}_T) &= R_{p+q}(\hat{\theta}_T)\hat{\Sigma}_T R'_{p+q}(\hat{\theta}_T) \\
L_{p+q}(\theta) &= \frac{1}{2} R_{p+q}(\theta) R'_{p+q}(\theta) = \left[2\sum_{i \le j} \frac{\partial r(x_i, x_j; \theta)}{\partial \theta_u}(\theta) \times \frac{\partial r(x_i, x_j; \theta)}{\partial \theta_v}(\theta)\right]_{(p+q),(p+q)}
\end{aligned}$$

On remarque que $\frac{1}{2}R_{p+q}(\theta)$ est la transposée de la matrice jacobienne de $\theta \longmapsto r(\theta)$ et, d'après l'hypothèse **D2**, $R_{p+q}(\theta)$ est de rang $(p+q)$ dans un voisinage de θ_0. On en déduit alors que $L_{p+q}(\theta)$ est inversible dans un voisinage de θ_0, ce qui nous permet de définir la matrice carrée :

$$\Delta_{p+q}(\theta_0) = L_{p+q}^{-1}(\theta_0) J_{p+q}(\theta_0) L_{p+q}^{-1}(\theta_0)$$

Si de plus **M2** est vérifiée, $J_{p+q}(\theta_0)$ est inversible et par conséquent $\Delta_{p+q}(\theta_0)$ l'est aussi.

***Théorème* 2.7** *Sous les hypothèses **B1**, **B3**, **D1-D2** et **M1-M2**, l'estimateur du minimum de contraste $\hat{\theta}_T$ vérifie :*

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow[T\to+\infty]{loi} \mathcal{N}_{p+q}\left(0, \Delta_{p+q}(\theta_0)\right), \ \textit{sous } \boldsymbol{H0}$$

★

Démonstration :

Les hypothèses **B1**, **B3** et **D1** nous assurent de la convergence en $P_0 -$ probabilité de $\hat{\theta}_T$ vers θ_0. On en déduit que $\hat{\theta}_T \in \mathring{\Theta}$ avec une probabilité tendant vers 1. Utilisant le développement de Taylor-Lagrange à l'ordre 0, on obtient pour un τ dans $]0,1[$:

$$U_T^{(1)}(\hat{\theta}_T) = 0 = U_T^{(1)}(\theta_0) + U_T^{(2)}\left(\theta_0 + \tau(\hat{\theta}_T - \theta_0)\right)(\hat{\theta}_T - \theta_0)$$

Le terme général de $U_T^{(2)}(\theta)$ s'écrit, pour tout $1 \leq u, v \leq (p+q)$ et tout $\theta \in \mathring{\Theta}$:

$$\begin{aligned}\left(U_T^{(2)}(\theta)\right)_{u,v} &= 2\sum_{i\leq j}\left[\frac{\partial r(x_i,x_j;\theta)}{\partial\theta_u}(\theta) \times \frac{\partial r(x_i,x_j;\theta)}{\partial\theta_v}(\theta)\right.\\ &\qquad \left. - \left(\hat{r}_T(x_i,x_j) - r(x_i,x_j;\theta)\right) \times \frac{\partial^2 r(x_i,x_j;\theta)}{\partial\theta_u\partial\theta_v}(\theta)\right]\end{aligned}$$

D'après l'hypothèse **D1**, pour tout $i \leq j$, $r(x_i,x_j;.)$, $r^{(1)}(x_i,x_j;.)$ et $r^{(2)}(x_i,x_j;.)$ sont continues dans $\mathring{\Theta}$. Comme $\hat{r}_T(x_i,x_j)$ converge en $P_0 -$ probabilité vers $r(x_i,x_j;\theta_0)$ et $\hat{\theta}_T$ converge en $P_0 -$ probabilité vers θ_0, on déduit :

$$U_T^{(2)}\left(\theta_0 + \tau(\hat{\theta}_T - \theta_0)\right) \xrightarrow[T\to+\infty]{} L_{p+q}(\theta_0), \text{ en } P_0 - \text{probabilité}$$

Par conséquent, $U_T^{(2)}(\theta_0 + \tau(\hat{\theta}_T - \theta_0))$ est inversible sur un ensemble dont la probabilité tend vers 1. Sur cet ensemble, on peut donc définir :

$$A_T = L_{p+q}(\theta_0)\left[U_T^{(2)}\left(\theta_0 + \tau(\hat{\theta}_T - \theta_0)\right)\right]^{-1}$$

Posons :

$$Y_T = \sqrt{T} L_{p+q}(\theta_0)(\hat{\theta}_T - \theta_0) = -A_T \sqrt{T} U_T^{(1)}(\theta_0)$$

Un calcul direct donne :

$$U_T^{(1)}(\theta_0) = -R_{p+q}(\theta_0)\,(\hat{r}_T - r(\theta_0))$$

En rappelant que (lemme 2.1) :

$$\sqrt{T}\,(\hat{r}_T - r(\theta_0)) \underset{T\to+\infty}{\xrightarrow{loi}} \mathcal{N}_{\frac{n(n+1)}{2}}(0, \Sigma_0), \text{ sous } \mathbf{H0}$$

on en déduit :

$$\sqrt{T} U_T^{(1)}(\theta_0) \underset{T\to+\infty}{\xrightarrow{loi}} \mathcal{N}_{p+q}\left(0, J_{p+q}(\theta_0)\right), \text{ sous } \mathbf{H0}$$

Comme de plus :

$$A_T \underset{T\to+\infty}{\longrightarrow} I_{p+q}, \text{ en } P_0 - \text{probabilité}$$

on a le résultat voulu d'après l'application du théorème de Slutsky, sous sa forme multivariée, à A_T et $\sqrt{T} U_T^{(1)}(\theta_0)$, c'est-à-dire :

$$Y_T = \sqrt{T} L_{p+q}(\theta_0)(\hat{\theta}_T - \theta_0) \underset{T\to+\infty}{\xrightarrow{loi}} \mathcal{N}_{p+q}\left(0, J_{p+q}(\theta_0)\right)$$

■

***Corollaire* 2.1** *Sous les hypothèses **B1**, **B3**, **D1-D2** et **M1-M2**, l'estimateur du minimum de contraste $\hat{\theta}_T$ vérifie :*

$$\sqrt{T} \Delta_{p+q}^{-\frac{1}{2}}(\hat{\theta}_T)(\hat{\theta}_T - \theta_0) \underset{T\to+\infty}{\xrightarrow{loi}} \mathcal{N}_{p+q}\left(0, I_{p+q}\right), \text{ sous } \boldsymbol{H0}$$

★

Démonstration :

D'après l'hypothèse **D1**, $r^{(1)}(x_i, x_j; .)$ est continue dans $\overset{\circ}{\Theta}$. Comme $\hat{\Sigma}_T$ converge en P_0 − probabilité vers Σ_0 (paragraphe 2.2.1) et $\hat{\theta}_T$ converge en P_0 − probabilité vers θ_0 (paragraphe 2.3.3), on déduit :

$$\Delta_{p+q}(\hat{\theta}_T) = L_{p+q}^{-1}(\hat{\theta}_T) J_{p+q}(\hat{\theta}_T) L_{p+q}^{-1}(\hat{\theta}_T) \underset{T\to+\infty}{\longrightarrow} \Delta_{p+q}(\theta_0), \text{ en } P_0 - \text{probabilité}$$

avec $\Delta_{p+q}(\theta_0)$ inversible (hypothèses **D2** et **M2**). Par conséquent, $\Delta_{p+q}(\hat{\theta}_T)$ est inversible sur un ensemble dont la probabilité tend vers 1. On conclut, en appliquant encore une fois le théorème de Slutsky à $\Delta_{p+q}^{-\frac{1}{2}}(\hat{\theta}_T)$ et $\sqrt{T}(\hat{\theta}_T - \theta_0)$.

■

2.4 Tests d'hypothèse simple et d'adéquation du modèle

Nous développons dans cette section des propriétés asymptotiques à la base de tests statistiques asymptotiques. On s'intéresse tout d'abord, dans le paragraphe 2.4.1., au test paramétrique de l'hypothèse nulle complètement spécifiée :

$$\mathbf{H0} : \theta = \theta_0$$

contre l'hypothèse alternative

$$\mathbf{H1} : \theta \neq \theta_0$$

On parlera dans ce cas du test de contraste qui est l'équivalent du test du rapport de vraisemblances. Au paragraphe 2.4.2, on propose un test d'adéquation du modèle. Ce dernier s'avérera être un cas particulier du test de différence de contraste d'une sous-hypothèse régulière générale.

2.4.1 Test de différence de contraste d'une hypothèse complètement spécifiée

Le test de l'hypothèse simple complètement spécifiée $\theta = \theta_0$ contre l'hypothèse alternative $\theta \neq \theta_0$ est basé sur le résultat de la convergence en loi, sous **H0**, de la statistique :

$$T\left(U_T(\theta_0) - U_T(\hat{\theta}_T)\right)$$

***Théorème* 2.8** *Sous les hypothèses **B1**, **B3**, **D1-D2** et **M1-M2** :*

$$2T\left(U_T(\theta_0) - U_T(\hat{\theta}_T)\right) \xrightarrow[T\to+\infty]{loi} \sum_{k=1}^{p+q} \lambda_k \chi_k^2(1), \ \textit{sous } \boldsymbol{H0}$$

où les $\{\lambda_k, 1 \leq k \leq (p+q)\}$ sont les $(p+q)$ valeurs propres strictement positives de $L_{p+q}^{-1}(\theta_0) J_{p+q}(\theta_0)$.

★

Démonstration :

Les hypothèses **B1**, **B3** et **D1** nous assurent que $\hat{\theta}_T$ est un estimateur du minimum de contraste convergeant en P_0-probabilité vers θ_0. Utilisant le développement de Taylor-Lagrange à l'ordre 1, on obtient pour un τ dans $]0, 1[$:

$$\begin{aligned} U_T(\theta_0) &= U_T(\hat{\theta}_T) + (\theta_0 - \hat{\theta}_T)' U_T^{(1)}(\hat{\theta}_T) \\ &+ \frac{1}{2}(\theta_0 - \hat{\theta}_T)' U_T^{(2)}(\hat{\theta}_T + \tau(\theta_0 - \hat{\theta}_T))(\theta_0 - \hat{\theta}_T) \end{aligned}$$

soit encore :

$$U_T(\theta_0) - U_T(\hat{\theta}_T) = \frac{1}{2}(\hat{\theta}_T - \theta_0)' U_T^{(2)}(\hat{\theta}_T + \tau(\theta_0 - \hat{\theta}_T))(\hat{\theta}_T - \theta_0)$$

On a :

$$\begin{cases} \sqrt{T}(\theta_0 - \hat{\theta}_T) \underset{T\to+\infty}{\overset{loi}{\longrightarrow}} \mathcal{N}_{p+q}\left(0, L_{p+q}^{-1}(\theta_0) J_{p+q}(\theta_0) L_{p+q}^{-1}(\theta_0)\right), \text{ sous } \mathbf{H0} \\ U_T^{(2)}(\hat{\theta}_T + \tau(\theta_0 - \hat{\theta}_T)) \underset{T\to+\infty}{\longrightarrow} L_{p+q}(\theta_0), \text{ en } P_0 - \text{probabilité} \end{cases}$$

Sur un ensemble dont la probabilité tend vers 1, $U_T^{(2)}(\hat{\theta}_T + \tau(\theta_0 - \hat{\theta}_T))$ est définie positive. Une application du théorème de Slutsky sous sa forme multivariée donne alors :

$$\sqrt{T}\left(U_T^{(2)}(\hat{\theta}_T + \tau(\theta_0 - \hat{\theta}_T))\right)^{1/2} (\theta_0 - \hat{\theta}_T) \underset{T\to+\infty}{\overset{loi}{\longrightarrow}} \mathcal{N}_{p+q}\left(0, L_{p+q}^{-1/2}(\theta_0) J_{p+q}(\theta_0) L_{p+q}^{-1/2}(\theta_0)\right), \text{ sous } \mathbf{H0}$$

Or, si Y est un vecteur aléatoire gaussien de loi $\mathcal{N}_n(0, \Delta)$, Δ inversible, $Y'Y$ a la loi $\sum_{k=1}^{n} \lambda_k \chi_k^2(1)$, où les λ_k sont les valeurs propres strictement positives de Δ (Johnson et Kotz (1970), pp. 149-151). Par conséquent :

$$2T\left(U_T(\theta_0) - U_T(\hat{\theta}_T)\right) \underset{T\to+\infty}{\overset{loi}{\longrightarrow}} \sum_{k=1}^{p+q} \lambda_k \chi_k^2(1), \text{ sous } \mathbf{H0}$$

où les $\{\lambda_k, 1 \leq k \leq (p+q)\}$ sont les valeurs propres strictement positives de $L_{p+q}^{-1/2}(\theta_0) J_{p+q}(\theta_0) L_{p+q}^{-1/2}(\theta_0)$. Ces valeurs propres sont encore celles de $L_{p+q}^{-1}(\theta_0) J_{p+q}(\theta_0)$.

■

Remarque : distribution d'une somme pondérée de $khi - deux$

Les lois des sommes pondérées de $\chi_k^2(1)$ indépendants sont tabulées pour différents jeux de taille $(p+q)$ et de poids strictement positifs $\{\lambda_k\}$ (*cf.* annexe B). Elles peuvent également être calculées numériquement.

♣

Régions de confiance

Le test de contraste introduit précédemment est basé sur la convergence en loi de la statistique $2T\left(U_T(\theta_0) - U_T(\hat{\theta}_T)\right)$ vers une somme pondérée de $khi - deux$ à un degré de liberté indépendants. Or, la loi asymptotique de $2T\left(U_T(\theta_0) - U_T(\hat{\theta}_T)\right)$ dépend du paramètre inconnu θ_0 qu'on cherche justement à estimer, ce qui pose problème pour la détermination des régions de rejet ou de confiance asymptotiques. C'est pourquoi il est intéressant, pour palier cet inconvénient, d'utiliser le test de Wald. Ce test repose sur le résultat de convergence en loi suivant, conséquence directe du corollaire 2.1 :

Théorème* 2.9** *Sous les hypothèses* ***B1, ***B3***, ***D1-D2*** *et* ***M1-M2*** *:*

$$T(\hat{\theta}_T - \theta_0)'\Delta_{p+q}^{-1}(\hat{\theta}_T)(\hat{\theta}_T - \theta_0) \xrightarrow[T\to+\infty]{loi} \chi^2(p+q), \ \textit{sous } \boldsymbol{H0}$$

★

Il est ainsi possible de construire une région de confiance asymptotique (ellipsoïde de confiance (centré en $\hat{\theta}_T$)) pour θ_0 ou de tester l'hypothèse :

$$\mathbf{H0} : \theta = \theta_0$$

contre l'hypothèse alternative :

$$\mathbf{H1} : \theta \neq \theta_0$$

La région de rejet de ce test au niveau asymptotique a est :

$$\{T(\hat{\theta}_T - \theta_0)'\Delta_{p+q}^{-1}(\hat{\theta}_T)(\hat{\theta}_T - \theta_0) \geq Q_{1-a}(p+q)\}$$

où $Q_{1-a}(p+q)$ est le quantile d'ordre $(1-a)$ de $\chi^2(p+q)$.

Il est également possible de construire un intervalle de confiance ponctuel asymptotique pour le modèle de covariance $r(x, x'; \theta_0)$. Cette construction s'obtient à partir du résultat suivant connu sous le nom de méthode *delta* :

***Lemme* 2.3** *(Monfort (1980), p.167))*
Soit $(X_T)_{T\geq 1}$ *une suite de vecteurs aléatoires à valeurs dans* $(\mathbb{R}^n, \mathcal{R}^n)$ *et une fonction* $h : \mathbb{N} \longrightarrow \mathbb{R}^+$ *telles que :*

$$\begin{cases} h(T) \xrightarrow[T\to+\infty]{} +\infty \\ h(T)(X_T - \theta) \xrightarrow[T\to+\infty]{loi} \mathcal{N}_n(0, \Delta) \end{cases}$$

Soit g *une fonction borélienne :* $\mathbb{R}^n \longrightarrow \mathbb{R}^p$, *différentiable en* θ, *alors :*

$$h(T)\left[g(X_T) - g(\theta)\right] \xrightarrow[T\to+\infty]{loi} \mathcal{N}_p\left(0, g^{(1)}(\theta)\Delta g^{(1)'}(\theta)\right)$$

★

Théorème* 2.10** *Sous les hypothèses* ***B1, ***B3***, ***D1-D2*** *et* ***M1-M2*** *:*

$$2\sqrt{T}\left(r(\hat{\theta}_T) - r(\theta_0)\right) \xrightarrow[T\to+\infty]{loi} \mathcal{N}_{\frac{n(n+1)}{2}}\left(0, R'_{p+q}(\theta_0)\Delta_{p+q}(\theta_0)R_{p+q}(\theta_0)\right), \ \textit{sous } \boldsymbol{H0}$$

★

Démonstration :

On rappelle :

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow[T\to+\infty]{loi} \mathcal{N}_{p+q}\left(0, \Delta_{p+q}(\theta_0)\right), \text{ sous } \mathbf{H0}$$

L'application du lemme 2.3 donne le résultat voulu.

■

Comme précédemment, on peut construire un test de Wald basé sur cette convergence en loi. On déduit aussi de ce résultat, pour tout (x, x') dans $[0, 1]^2$:

$$\sqrt{T}\left(r(x,x';\hat{\theta}_T) - r(x,x';\theta_0)\right) \xrightarrow[T\to+\infty]{loi} \mathcal{N}_1\left(0, (r^{(1)}(x,x';\theta_0))'\Delta_{p+q}(\theta_0)(r^{(1)}(x,x';\theta_0))\right), \text{ sous } \mathbf{H0}$$

Posons :

$$\Lambda_1(\hat{\theta}_T) = (r^{(1)}(x,x';\hat{\theta}_T))'\Delta_{p+q}(\hat{\theta}_T)(r^{(1)}(x,x';\hat{\theta}_T))$$

En tout point (x, x') de $[0, 1]^2$, tel que $r^{(1)}(x, x'; \theta_0)$ soit non-nulle, un intervalle de confiance pour $r(x, x'; \theta_0)$ de niveau de confiance asymptotique $(1 - a)$ est donné par :

$$\left\{T\Lambda_1^{-1}(\hat{\theta}_T)\left(r(x,x';\hat{\theta}_T) - r(x,x';\theta_0)\right)^2 \leq Q_{1-a}(1)\right\}$$

où $Q_{1-a}(1)$ est le quantile d'ordre $(1 - a)$ de $\chi^2(1)$.

2.4.2 Test d'adéquation du modèle

On a utilisé comme critère d'ajustement du modèle, celui des moindres carrés. Un moyen d'étudier la validité du modèle, compte tenu des observations, est d'exhiber un test d'adéquation du modèle $r(x, x'; \theta_0)$ aux données, ou plus exactement aux covariances empiriques $\hat{r}_T(x_i, x_j), i \leq j$. Ce test, basé sur la convergence en loi de $TU_T(\hat{\theta}_T)$, se déduit d'un résultat général relatif au test d'une sous-hypothèse régulière.

Test de différence de contraste d'une sous-hypothèse : le test de la somme pondérée de $khi - deux$ indépendants

Soit **Hm**, l'hypothèse du modèle majorant :

- **Hm** : θ appartient à Θ compact de $\mathbb{R}^m$ avec $\theta_0 \in \overset{\circ}{\Theta}$ et toujours l'hypothèse d'identifiabilité : $\theta \longmapsto r(\theta)$ est une injection continue

Nous voulons tester la sous-hypothèse de dimension r **Hr**, $r < m$:

- **Hr** : $\theta = s(\gamma), \gamma \in \Omega$ ouvert borné de $\mathbb{R}^r$,

  $s(\Omega) \subset \Theta$, s est injective, de classe C^2 dans Ω et :

  $$\theta_0 = s(\gamma_0), rang(s^{(1)}(\gamma_0)) = r$$

  (γ_0 désigne la vraie valeur du paramètre γ)

Nous gardons les notations $\hat{\theta}_T$ et U_T pour l'estimateur $\hat{\theta}_T$ (de θ_0) du minimum de contraste U_T sous **Hm**, et introduisons l'estimateur contraint $\bar{\theta}_T$:

soit $\bar{U}_T(\gamma) = U_T(s(\gamma))$ le contraste sous **Hr**, et soit $\hat{\gamma}_T$ un estimateur du minimum de contraste sous **Hr** ; on note $\bar{\theta}_T = s(\hat{\gamma}_T)$ et ainsi :

$$\bar{U}_T(\hat{\gamma}_T) = U_T(\bar{\theta}_T)$$

Considérons les matrice carrées :

$$\begin{aligned} J_m(\theta_0) &= R_m(\theta_0)\Sigma_0 R'_m(\theta_0) \\ \Psi_m &= L_m^{-1}(\theta_0) - s^{(1)}(\gamma_0)\left[s^{(1)'}(\gamma_0)L_m(\theta_0)s^{(1)}(\gamma_0)\right]^{-1} s^{(1)'}(\gamma_0) \end{aligned}$$

où l'on rappelle que :

$$L_m(\theta_0) = \frac{1}{2}R_m(\theta_0)R'_m(\theta_0)$$

avec :

$$R_m(\theta) = 2\left[r^{(1)}(x_1, x_1; \theta), r^{(1)}(x_1, x_2; \theta), \ldots, r^{(1)}(x_n, x_n; \theta)\right]_{m, \frac{n(n+1)}{2}}$$

Lemme* 2.4** *Sous* ***Hr *:*

$$\hat{\gamma}_T \underset{T\to+\infty}{\longrightarrow} \gamma_0, \ \text{en } P_0 - \text{probabilité}$$

★

Démonstration :

Il suffit de voir que les hypothèses faites sur **Hr** entraînent les conditions assurant la convergence en P_0-probabilité de l'estimateur du minimum de contraste contraint $\bar{U}_T$.

■

Théorème* 2.11** *On suppose que les hypothèses* ***D1-D2 *sont vérifiées sous* ***Hm*** *comme sous* ***Hr****. Si de plus, les hypothèses* ***M1-M2*** *sont vérifiées, alors :*

$$2T\left(U_T(\bar{\theta}_T) - U_T(\hat{\theta}_T)\right) \underset{T\to+\infty}{\overset{loi}{\longrightarrow}} \sum_{k=1}^{m-r} \lambda_k \chi_k^2(1), \ \text{sous } \boldsymbol{Hr}$$

où les $\{\lambda_k, 1 \leq k \leq (m-r)\}$ *sont les* $(m-r)$ *valeurs propres strictement positives de* $J_m(\theta_0)\Psi_m$.

★

Démonstration :

Elle suit une démarche classique ; on examine les développements de Taylor-Lagrange à l'ordre 1 de U_T en $\hat{\theta}_T$ et $\bar{U}_T$ en $\hat{\gamma}_T$, puis leur différence. Puisque

$$U_T^{(1)}(\hat{\theta}_T) = \bar{U}_T^{(1)}(\hat{\gamma}_T) = 0$$

on a, pour un couple (τ_1, τ_2) dans $]0,1[^2$:

$$\left\{\begin{array}{l} U_T(\theta_0) - U_T(\hat{\theta}_T) = \dfrac{1}{2}(\hat{\theta}_T - \theta_0)' U_T^{(2)}(\hat{\theta}_T + \tau_1(\theta_0 - \hat{\theta}_T))(\hat{\theta}_T - \theta_0) \\ \\ \bar{U}_T(\gamma_0) - \bar{U}_T(\hat{\gamma}_T) = \dfrac{1}{2}(\hat{\gamma}_T - \gamma_0)' \bar{U}_T^{(2)}(\hat{\gamma}_T + \tau_2(\gamma_0 - \hat{\gamma}_T))(\hat{\gamma}_T - \gamma_0) \end{array}\right.$$

D'autre part, les développements de Taylor-Lagrange à l'ordre 0 de $U_T^{(1)}$ en θ_0 et $\bar{U}_T^{(1)}$ en γ_0 donnent, pour un couple (τ_3, τ_4) dans $]0,1[^2$:

$$\left\{\begin{array}{l} \hat{\theta}_T - \theta_0 = -\left[U_T^{(2)}(\theta_0 + \tau_3(\hat{\theta}_T - \theta_0))\right]^{-1} U_T^{(1)}(\theta_0) \\ \hat{\gamma}_T - \gamma_0 = -\left[\bar{U}_T^{(2)}(\gamma_0 + \tau_4(\hat{\gamma}_T - \gamma_0))\right]^{-1} \bar{U}_T^{(1)}(\gamma_0) \end{array}\right.$$

Enfin, nous avons sous **Hr** les relations :

$$\left\{\begin{array}{lcl} U_T(\theta_0) & = & \bar{U}_T(\gamma_0) \\ \\ \bar{U}_T(\hat{\gamma}_T) & = & U_T(\bar{\theta}_T) \\ \\ \bar{U}_T^{(1)}(\gamma_0) & = & s^{(1)'}(\gamma_0) U_T^{(1)}(\theta_0) = -s^{(1)'}(\gamma_0) R_m(\theta_0)(\hat{r}_T - r(\theta_0)) \end{array}\right.$$

On obtient alors :

$$2T\left[U_T(\bar{\theta}_T) - U_T(\hat{\theta}_T)\right] = T\left(\hat{r}_T - r(\theta_0)\right)' R_m'(\theta_0)\left[B_T - s^{(1)}(\gamma_0) C_T s^{(1)'}(\gamma_0)\right] R_m(\theta_0)\left(\hat{r}_T - r(\theta_0)\right)$$

avec :

$$\left\{\begin{array}{lcl} B_T & = & \left[U_T^{(2)}(\theta_0 + \tau_3(\hat{\theta}_T - \theta_0))\right]^{-1} U_T^{(2)}(\theta_0 + \tau_1(\hat{\theta}_T - \theta_0)) \left[U_T^{(2)}(\theta_0 + \tau_3(\hat{\theta}_T - \theta_0))\right]^{-1} \\ C_T & = & \left[\bar{U}_T^{(2)}(\gamma_0 + \tau_4(\hat{\gamma}_T - \gamma_0))\right]^{-1} \bar{U}_T^{(2)}(\gamma_0 + \tau_2(\hat{\gamma}_T - \gamma_0)) \left[\bar{U}_T^{(2)}(\gamma_0 + \tau_4(\hat{\gamma}_T - \gamma_0))\right]^{-1} \end{array}\right.$$

On vérifie alors que :

$$\left\{\begin{array}{lcl} B_T & \underset{T\to+\infty}{\longrightarrow} & L_m^{-1}(\theta_0), \text{ en } P_0 - \text{probabilité} \\ C_T & \underset{T\to+\infty}{\longrightarrow} & \left[s^{(1)'}(\gamma_0) L_m(\theta_0) s^{(1)}(\gamma_0)\right]^{-1}, \text{ en } P_0 - \text{probabilité} \end{array}\right.$$

On vérifie que la matrice $\left[I_m - L_m^{1/2}(\theta_0) s^{(1)}(\gamma_0)\left[s^{(1)'}(\gamma_0) L_m(\theta_0) s^{(1)}(\gamma_0)\right]^{-1} s^{(1)'}(\gamma_0) L_m^{1/2}(\theta_0)\right]$ est idempotente de rang $(m - r)$. Comme $L_m(\theta_0)$ est régulière,

$\left[L_m^{-1}(\theta_0) - s^{(1)}(\gamma_0)\left[s^{(1)'}(\gamma_0)L_m(\theta_0)s^{(1)}(\gamma_0)\right]^{-1} s^{(1)'}(\gamma_0)\right]$ est aussi de rang $(m-r)$ et ses valeurs propres sont positives ou nulles.

Soit

$$D_T = \left[B_T - s^{(1)}(\gamma_0)C_T s^{(1)'}(\gamma_0)\right]^{1/2} R_m(\theta_0)$$

On a :

$$\sqrt{T}\left(\hat{r}_T - r(\theta_0)\right) \xrightarrow[T\to+\infty]{loi} \mathcal{N}_{\frac{n(n+1)}{2}}(0,\Sigma_0), \text{ sous } \mathbf{Hr}$$

et :

$$\begin{array}{rcl} D_T & \underset{T\to+\infty}{\longrightarrow} & \left[L_m^{-1}(\theta_0) - s^{(1)}(\gamma_0)\left[s^{(1)'}(\gamma_0)L_m(\theta_0)s^{(1)}(\gamma_0)\right]^{-1} s(1)'(\gamma_0)\right]^{1/2} R_m(\theta_0) \\ & = & \Psi_m^{1/2} R_m(\theta_0), \text{ en } P_0-\text{probabilité.} \end{array}$$

Par conséquent :

$$\sqrt{T} D_T \left(\hat{r}_T - r(\theta_0)\right) \xrightarrow[T\to+\infty]{loi} \mathcal{N}_{\frac{n(n+1)}{2}}\left(0, \Psi_m^{1/2} J_m(\theta_0)\Psi_m^{1/2}\right), \text{ sous } \mathbf{Hr}$$

On en déduit finalement :

$$2T\left(U_T(\bar{\theta}_T) - U_T(\hat{\theta}_T)\right) \xrightarrow[T\to+\infty]{loi} \sum_{k=1}^{m-r} \lambda_k \chi_k^2(1), \text{ sous } \mathbf{Hr}$$

où les $\{\lambda_k, 1 \leq k \leq (m-r)\}$ sont les valeurs propres strictement positives de $\Psi_m^{1/2} J_m(\theta_0)\Psi_m^{1/2}$, valeurs propres qui sont encore celles de $J_m(\theta_0)\Psi_m$ (Searle (1982), p.303). ■

Remarque : une autre traduction de la sous-hypothèse **Hr**

Une autre façon de procéder au test de la sous-hypothèse **Hr** est la suivante :
traduire **Hr** sous la forme d'une contrainte $\phi_0 = C(\theta_0) = 0$ où $C : \theta \longrightarrow \mathbb{R}^{m-r}$ est de classe C^2 et de matrice jacobienne J_C de plein rang $(m-r)$ en θ_0. Comme :

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow[T\to+\infty]{loi} \mathcal{N}_m\left(0, \Delta_m(\theta_0)\right), \text{ sous } \mathbf{Hr}$$

alors, d'après le lemme 2.3, pour $\hat{\phi}_T = C(\hat{\theta}_T)$:

$$\sqrt{T}(\hat{\phi}_T - \phi_0) \xrightarrow[T\to+\infty]{loi} \mathcal{N}_{m-r}\left(0, J_C(\theta_0)\Delta_m(\theta_0)J_C'(\theta_0)\right), \text{ sous } \mathbf{Hr}$$

Par conséquent :

$$T\hat{\phi}_T' \left[J_C(\theta_0)\Delta_m(\theta_0)J_C'(\theta_0)\right]^{-1} \hat{\phi}_T \xrightarrow[T\to+\infty]{loi} \chi^2(m-r), \text{ sous } \mathbf{Hr}$$

♣

La situation limite où $m = \dfrac{n(n+1)}{2}$ (c'est-à-dire où le nombre de paramètres est égal au nombre de covariances empiriques $\hat{r}_T(x_i, x_j)$, $1 \leq i \leq j \leq n$, dont on dispose pour construire notre contraste) conduit au cas particulier, décrit ci-après, qui va fournir un critère pour tester la validité d'un sous-modèle.

Test d'adéquation

On prend ici comme modèle majorant le modèle saturé, ou maximal, sans contrainte à $m = \dfrac{n(n+1)}{2}$ paramètres et on teste la validité du sous-modèle à r paramètres correspondant à l'hypothèse **Hr**. On considère l'hypothèse **D2'** qui renforce **D2** :

- **D2'** : la matrice jacobienne de l'application $\theta \longmapsto r(\theta)$ est de rang $(p+q)$ pour tout θ dans $\overset{\circ}{\Theta}$

Corollaire* 2.2** *Dans ce contexte particulier et sous les hypothèses du théorème 2.11 avec* ***D2 *renforcé par* ***D2'****:*

$$2TU_T(\bar{\theta}_T) \xrightarrow[T\to+\infty]{loi} \sum_{k=1}^{\frac{n(n+1)}{2}-r} \lambda_k \chi_k^2(1), \ \ sous\ \boldsymbol{Hr}$$

où les $\left\{\lambda_k, 1 \leq k \leq \left(\dfrac{n(n+1)}{2} - r\right)\right\}$ *sont les valeurs propres strictement positives de* $J_{\frac{n(n+1)}{2}}(\theta_0)\Psi_{\frac{n(n+1)}{2}}$.

★

Démonstration :

Si on note $\hat{\theta}_T^s$ un estimateur du minimum de contraste sous l'hypothèse du modèle saturé, on a :

$$U_T(\hat{\theta}_T^s) = 0$$

En effet, l'hypothèse **D2'** signifie que $R_{\frac{n(n+1)}{2}}(\theta)$, matrice carrée, est inversible dans $\overset{\circ}{\Theta}$. On a vu, au cours de la démonstration du théorème 2.7, la relation :

$$U_T^{(1)}(\theta) = -R_{\frac{n(n+1)}{2}}(\theta)\left(\hat{r}_T - r(\theta)\right)$$

On a donc :

$$-R^{-1}_{\frac{n(n+1)}{2}}(\theta)U_T^{(1)}(\theta) = \left(\hat{r}_T - r(\theta)\right)$$

Comme, par définition du minimum de contraste, $U_T^{(1)}(\hat{\theta}_T^s) = 0$, on a :

$$\hat{r}_T - r(\hat{\theta}_T^s) = 0$$

D'où :

$$U_T(\hat{\theta}_T^s) = \left\| \hat{r}_T - r(\hat{\theta}_T^s) \right\|^2_{\frac{n(n+1)}{2}} = 0$$

■

2.5 Identification *p.s.* du modèle par contraste pénalisé

S'il est raisonnable de supposer que le vrai modèle appartient à une famille de modèles de dimensions variables, il nous faudra alors identifier le vrai modèle puis estimer ses paramètres. Le choix du modèle et du nombre de paramètres pour un bon ajustement aux données est un problème crucial. Une stratégie consiste à utiliser un nombre de paramètres important afin de minimiser au mieux le critère d'ajustement. Cela conduit en général à une sur-paramétrisation du modèle ; aux points observés l'écart avec le modèle est minime, par contre, pour une prévision en dehors des points observés, l'écart est fort. Pour éviter ce type de problème, l'idée est d'introduire un compromis entre le critère d'ajustement, qu'on veut le plus faible possible, et la dimension du modèle, qu'on souhaite modérée pour la raison soulignée précédemment mais aussi pour des raisons d'interprétation du modèle. Notre critère d'ajustement étant ici la minimisation d'un contraste, cela va donner lieu à une identification du modèle par contraste pénalisé. Le théorème 2.12 ci-après permet de se prémunir contre une sur-paramétrisation du modèle, mais également contre une sous-paramétrisation éventuelle.

Le contexte mathématique s'inspire directement du travail de Guyon ((1993), pp.114-120).

Supposons que l'espace des paramètres Θ soit inclus dans $\mathbb{R}^m$, où $\mathbb{R}^m$ correspond à la dimension du modèle majorant ; soit $\mathcal{E}$ une famille finie de sous-espaces de $\mathbb{R}^m$, $\delta \in \mathcal{E}$ l'élément générique de $\mathcal{E}$, $|\delta|$ la dimension de δ, $\Theta_\delta = \Theta \bigcap \delta$ le sous-espace (sous-modèle) paramétrique associé. On suppose que la vraie valeur θ_0 est dans Θ_{δ_0}, où δ_0, élément de $\mathcal{E}$, est le sous-espace minimal associé à θ_0 : si $\delta_0 \nsubseteq \delta$ et $\delta \in \mathcal{E}$ alors $\theta_0 \notin \Theta_\delta$.

Le problème est d'estimer δ_0, c'est-à-dire, au vu des observations $Z_t(x_i)$, $1 \leq i \leq n$ et $1 \leq t \leq T$, déterminer un support $\hat{\delta}_T \in \mathcal{E}$. On définit pour cela un nouveau contraste V_T égal au contraste U_T pénalisé, à la vitesse $\dfrac{c(T)}{T}$, par la dimension du modèle :

$$V_T(\theta) = U_T(\theta) + \frac{c(T)}{T}|\delta(\theta)|$$

Ce nouveau contraste répond au principe de parcimonie d'Akaïke (1969).

On estime δ_0 par :

$$\hat{\delta}_T = \arg\min_{\delta \in \mathcal{E}} \left(U_T(\hat{\theta}_{T,\delta}) + \frac{c(T)}{T}|\delta| \right)$$

où :

$$\hat{\theta}_{T,\delta} = \arg\min_{\theta \in \Theta_\delta} U_T(\theta)$$

On introduit les hypothèses suivantes :

- **B'** : les hypothèses **B1**-**B3** adaptées aux différents sous-modèles sont vérifiées
- **D'** : les hypothèses **D1**-**D2** adaptées aux différents sous-modèles sont vérifiées

et concernant la vitesse de pénalisation, on fait le choix :

- **P** : $c(T) = T^\alpha, 0 < \alpha < 1$

***Théorème* 2.12** *Sous les hypothèses **B'**, **D'**, **M3** et **P**, on a :*

$$(\hat{\delta}_T, \hat{\theta}_{T,\hat{\delta}_T}) \underset{T\to+\infty}{\longrightarrow} (\delta_0, \theta_0), \; P_0 - p.s.$$

★

Démonstration :

On utilise le résultat de Guyon ((1993), propriété 3.9). Pour appliquer ce résultat directement à notre problème, on doit montrer que les hypothèses **I1**-**I3** et **L1**-**L2** de la propriété 3.9 sur l'identification *p.s.* d'un modèle sont bien vérifiées ici :

- **I1** :
  1. $\liminf_{T\to+\infty} U_T(\theta) \geq K(\theta_0, \theta), \; P_0 - p.s.$
  2. Si $\delta \in \mathcal{E}, \theta \longmapsto K(\theta_0, \theta)$ de Θ_δ dans $\mathbb{R}^+$ admet un minimum unique θ_δ sur Θ_δ
  3. $\theta \longmapsto U_T(\theta)$ est de classe C^2 dans $\overset{\circ}{\Theta}$ $P_0 - p.s.$ et $\theta \longmapsto K(\theta_0, \theta)$ est continue dans Θ
  4. $P_0 \left(\limsup_{T\to+\infty} \left\{ W_T \left(\frac{1}{k} \geq \epsilon_k \right) \right\} \right) = 0$, pour une suite $\epsilon_k \underset{T\to+\infty}{\longrightarrow} 0$

- **I2** : il existe une suite (l_T) tendant vers $+\infty$, vérifiant $\frac{T}{l_T} \underset{T\to+\infty}{\longrightarrow} +\infty$, et $\alpha \geq 0$ *cste* finie telles que :

$$\limsup_{T\to+\infty} \sqrt{\frac{T}{l_T}} \|U_T^{(1)}(\theta_0\| \leq \alpha < +\infty, \; P_0 - p.s.$$

– **I3** : il existe L symétrique et définie positive telle que :

$$\liminf_{T\to+\infty}\left(U_T^{(2)}(\theta_0)\right)\geq L,\ P_0-p.s.$$

– **L1** : $\lim_{T\to+\infty}\frac{c(T)}{T}=0$

– **L2** : $\liminf_{T\to+\infty}\frac{c(T)}{l_T}>\frac{\alpha^2}{2\lambda}$, où λ est la plus petite valeur propre de L

(i) Grâce aux hypothèses **B'**, **D'** et **M3**, l'hypothèse **I1** est bien vérifiée.

(ii) On a :

$$U_T^{(1)}(\theta_0)=-R_m(\theta_0)\left(\hat{r}_T-r(\theta_0)\right)$$

$R_m(\theta_0)$ définissant une application linéaire de $\mathbb{R}^{n(n+1)/2}$ dans $\mathbb{R}^m$, on a :

$$\|U_T^{(1)}(\theta_0)\|_m\leq\|R_m(\theta_0)\|_{sup}\|\hat{r}_T-r(\theta_0)\|_{n(n+1)/2}$$

Posant $K_1=\|R_m(\theta_0)\|_{sup}$, on obtient :

$$\|U_T^{(1)}(\theta_0)\|_m\leq K_1\|\hat{r}_T-r(\theta_0)\|_{n(n+1)/2}$$

On sait, d'après le lemme 2.2, que l'ensemble :

$$\left\{\frac{T\Sigma_0^{-1/2}\left(\hat{r}_T-r(\theta_0)\right)}{\sqrt{2T\ln\ln(T)}}\right\}$$

est $p.s.$ relativement compact dans la boule unité de $\mathbb{R}^{n(n+1)/2}$. Par conséquent, il existe $K_2<+\infty$ telle que :

$$\limsup_{T\to+\infty}\frac{T\|\hat{r}_T-r(\theta_0)\|_{n(n+1)/2}}{\sqrt{2T\ln\ln(T)}}\leq K_2$$

On a finalement, avec $K=K_1\times K_2<+\infty$:

$$\limsup_{T\to+\infty}\frac{\sqrt{T}}{\sqrt{2\ln\ln(T)}}\|U_T^{(1)}(\theta_0)\|_m\leq K<+\infty$$

L'hypothèse **I2** est donc bien vérifiée avec $l_T=2\ln\ln(T)$.

(iii) Quant à l'hypothèse **I3**, elle est vérifiée, puisqu'on a vu dans la démonstration du théorème 2.7 que :

$$U_T^{(2)}(\theta_0)\underset{T\to+\infty}{\longrightarrow}L_{p+q}(\theta_0),\ P_0-p.s.$$

où $L_{p+q}(\theta_0)$ est symétrique définie positive.

(iv) Enfin, avec $l_T=2\ln\ln(T)$ et $c(T)=T^\alpha, 0<\alpha<1$, les hypothèses **L1** et **L2** sont elles aussi vérifiées.

■

2.6 Exemple et une étude expérimentale

Nous considérons un processus Z gaussien défini sur $[0,1]$, non-stationnaire, centré et réduit, et continu en moyenne quadratique.

Nous prenons comme famille de modèles de déformation, les fonctions puissance :

$$\mathcal{F} = \{x^\alpha, \alpha > 0\}$$

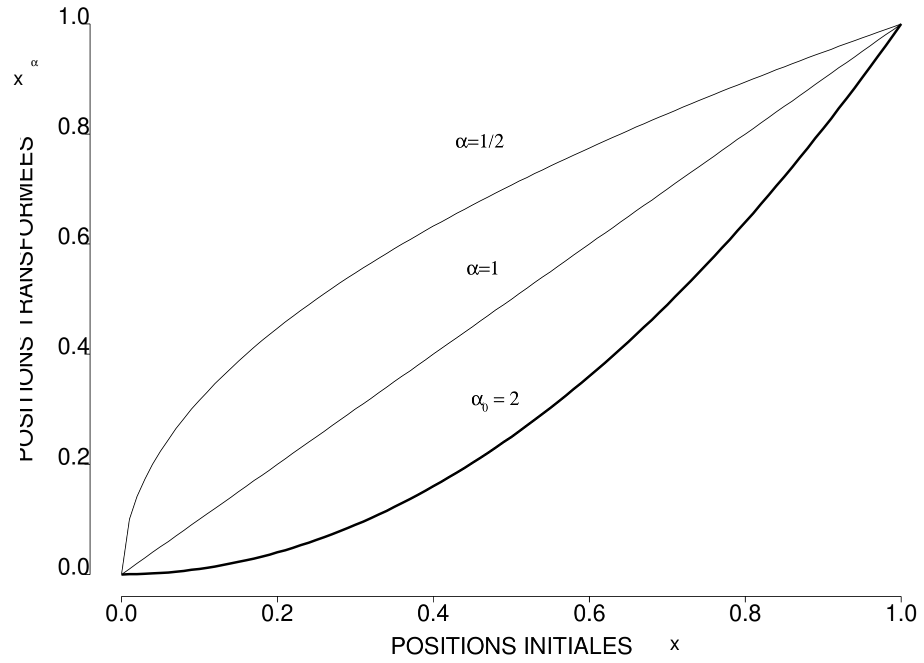


FIG. 2.2 – *Trois fonctions puissance, α_0 est le paramètre de la vraie déformation*

et comme famille de modèles de covariance stationnaire, les covariances exponentielles :

$$\mathcal{R} = \{e^{-\beta|u|}, \beta > 0\}$$

Ces covariances exponentielles, souvent utilisée en géostatistique, sont définies positives quelque soit la dimension du support du processus étudié, c'est-à-dire, pour la norme euclidienne $\|u\|_p$:

$$u \in \mathbb{R}^p \longmapsto e^{-\beta\|u\|_p} \text{ est définie positive pour tout } p \in \mathbb{N}^*$$

Nous prenons comme vraies valeurs des paramètres :

$$\begin{cases} \alpha_0 = 2 \\ \beta_0 = 1 \end{cases}$$

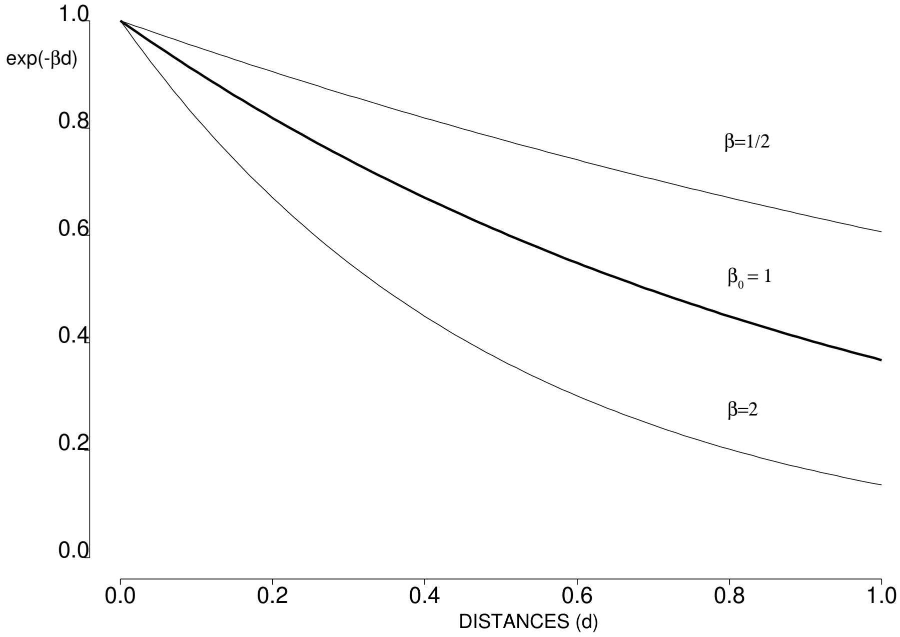


FIG. 2.3 – *Trois covariances exponentielles, β_0 est le paramètre de la vraie covariance stationnaire*

L'étude numérique qui suit a pour but d'illustrer expérimentalement le résultat de normalité asymptotique des estimateurs du minimum de contraste de α_0 et β_0 (théorème 2.7) et le résultat de convergence en loi de la différence de contraste vers une somme pondérée de $khi - deux$ indépendants (théorème 2.8).

2.6.1 Identifiabilité paramétrique du modèle

Soit $\mathcal{P}_n = \{x_1 < x_2 < \cdots < x_n\}$ le plan d'expérience constitué de n stations fixes positionnées en $x_1, x_2, \ldots, x_n$ et en lesquelles on dispose des valeurs du processus Z aux instants $t \in [1, \ldots, T]$. On prend dans la suite $n = 26$ et $x_i = \dfrac{i-1}{25}$, $i = 1, \ldots, 26$. On suppose que (α, β) est dans un compact Θ de $\left(\mathbb{R}^{*+}\right)^2$ et que $(\alpha_0, \beta_0) \in \mathring{\Theta}$; pour notre modèle, l'hypothèse **B1** est vérifiée.

Sous l'hypothèse **H0**, le modèle de covariance s'écrit, pour tout (x, x') dans $[0,1]^2$:

$$r(x, x'; \alpha_0, \beta_0) = \exp\left(-\beta_0 |x'^{\alpha_0} - x^{\alpha_0}|\right)$$

L'application :

$$(\alpha, \beta) \longmapsto \exp\left(-\beta|x'^{\alpha} - x^{\alpha}|\right)$$

est continue pour tout (x, x') dans $[0,1]^2$; l'hypothèse **B2** est bien vérifiée.

Le fait de savoir que Z est réduit nous permet de ne considérer que les couples de stations (x_i, x_j) avec $i < j$. Pour les couples (x_1, x_2) et (x_1, x_3), on a :

$$\begin{pmatrix} \exp\left(-\beta x_2^{\alpha}\right) \\ \exp\left(-\beta x_3^{\alpha}\right) \end{pmatrix} = \begin{pmatrix} \exp\left(-\beta' x_2^{\alpha'}\right) \\ \exp\left(-\beta' x_3^{\alpha'}\right) \end{pmatrix} \Longleftrightarrow (\alpha, \beta) = (\alpha', \beta')$$

autrement dit, l'hypothèse **B3** est assurée et l'identifiabilité paramétrique du modèle aussi.

2.6.2 Estimation des paramètres

$\hat{\alpha}_T$ et $\hat{\beta}_T$ minimisent le contraste :

$$U_T(\alpha, \beta) = \sum_{i<j} \left[\hat{r}_T(x_i, x_j) - r(x_i, x_j; \alpha, \beta)\right]^2$$

avec :

$$\hat{r}_T(x_i, x_j) = \frac{1}{T} \sum_{t=1}^{T} Z_t(x_i) Z_t(x_j)$$

Pour calculer les covariances empiriques $\hat{r}_T(x_i, x_j)$, $i < j$, nous simulons, sous l'hypothèse **H0**, $T = 100$ réalisations *i.i.d.* du vecteur gaussien $(Z(x_1), Z(x_2), \ldots, Z(x_n))$ centré de matrice de variances-covariances $\left[r_Z(x_i, x_j)\right]_{n,n}$, avec pour tout $1 \leq i, j \leq n$:

$$r_Z(x_i, x_j) = \exp\left(-\beta_0 \left|x_j^{\alpha_0} - x_i^{\alpha_0}\right|\right)$$

Pour relaxer les contraintes de positivité :

$$(S_2) : \begin{cases} \alpha > 0 \\ \beta > 0 \end{cases}$$

on fait le changement de paramètres :

$$\begin{cases} \alpha = e^{\alpha'} \\ \beta = e^{\beta'} \end{cases}$$

avec (α', β') dans $\mathbb{R}^2$, et on minimise U_T par rapport aux nouveaux paramètres α' et β' sans contrainte, c'est-à-dire dans $\mathbb{R}^2$.

Les estimateurs du minimum de contraste ont été obtenus à l'aide de la fonction ***nlregb*** du logiciel de Statistiques "S-Plus", version 3.4 (*cf.* Spector (1994)).

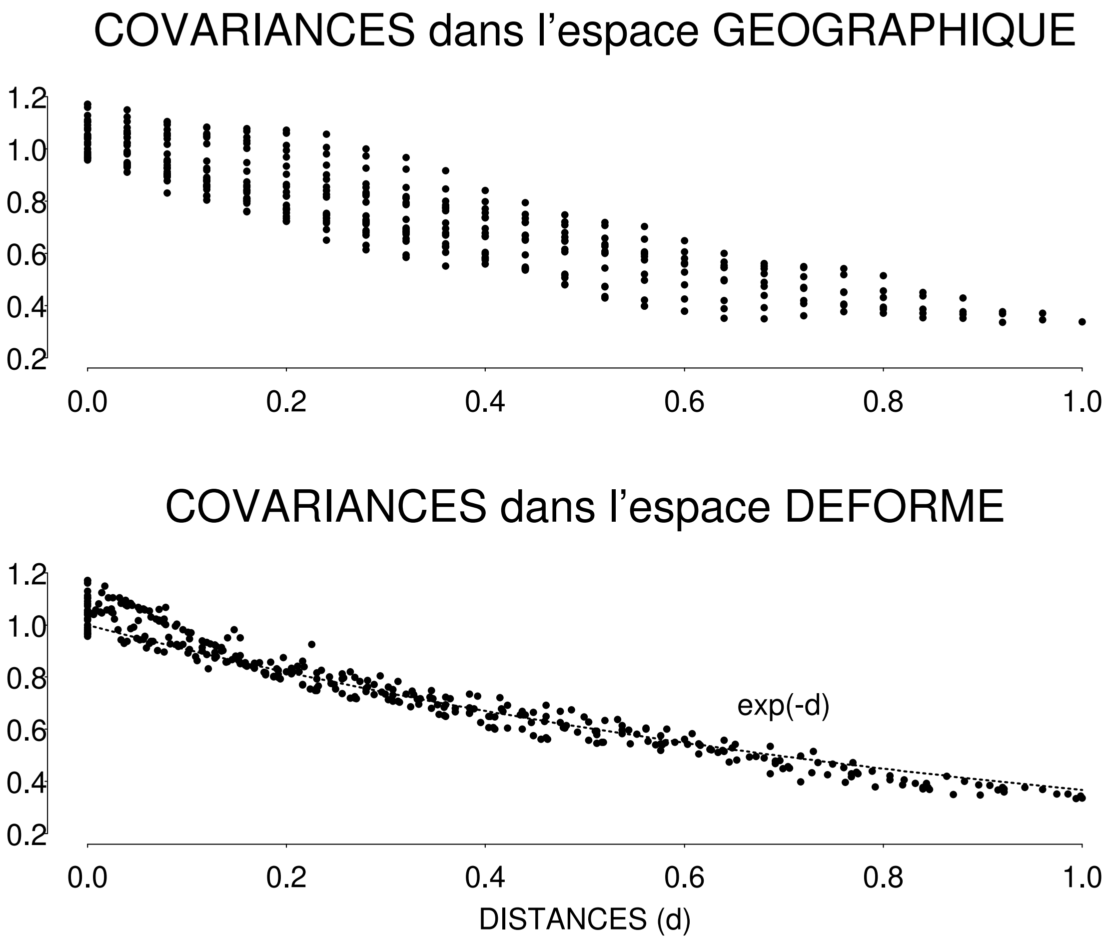


FIG. 2.4 – *Les covariances empiriques en fonction des distances dans l'espace géographique (figure de dessus) et en fonction des distances dans l'espace déformé (figure de dessous)*

2.6.3 Normalité asymptotique des estimateurs

L'hypothèse **D1** est bien assurée ainsi que **M2** (le calcul numérique donne $|\Sigma_0| \neq 0$). Pour vérifier l'hypothèse **D2**, il suffit de montrer que la matrice :

$$R_2(\alpha_0, \beta_0) = -2 \left[\begin{array}{c} \beta_0(x_j^{\alpha_0} \ln(x_j) - x_i^{\alpha_0} \ln(x_i)) \exp\left(-\beta_0(x_j^{\alpha_0} - x_i^{\alpha_0})\right) \\ (x_j^{\alpha_0} - x_i^{\alpha_0}) \exp\left(-\beta_0(x_j^{\alpha_0} - x_i^{\alpha_0})\right) \end{array} \right]_{2,(i<j)}$$

(avec, par prolongement par continuité en $x_1 = 0$, $x_1^{\alpha_0} \ln(x_1) = 0$) est de rang 2. Cela est vraie dès qu'il existe deux couples (x_{i_1}, x_{j_1}) et (x_{i_2}, x_{j_2}) tels que :

$$\left| \begin{array}{cc} \left(x_{j_1}^{\alpha_0} \ln(x_{j_1}) - x_{i_1}^{\alpha_0} \ln(x_{i_1})\right) & \left(x_{j_2}^{\alpha_0} \ln(x_{j_2}) - x_{i_2}^{\alpha_0} \ln(x_{i_2})\right) \\ \left(x_{j_1}^{\alpha_0} - x_{i_1}^{\alpha_0}\right) & \left(x_{j_2}^{\alpha_0} - x_{i_2}^{\alpha_0}\right) \end{array} \right| \neq 0$$

Pour les couples (x_1, x_2) et (x_1, x_3), ce déterminant est égal à $x_2^{\alpha_0} x_3^{\alpha_0} (\ln(x_2) - \ln(x_3)) \neq 0$.

On a donc :

$$\sqrt{T}\left((\hat{\alpha}_T, \hat{\beta}_T) - (\alpha_0, \beta_0)\right) \underset{T\to+\infty}{\xrightarrow{loi}} \mathcal{N}_2\left(0, \Delta_2(\alpha_0, \beta_0)\right), \text{ sous } \mathbf{H0}$$

où la matrice de variances-covariances $\Delta_2(\alpha_0, \beta_0)$, définie p.61, vaut :

$$\Delta_2(\alpha_0, \beta_0) = \begin{bmatrix} 55.56 & -9.79 \\ -9.79 & 12.52 \end{bmatrix}$$

Pour avoir une idée de la distribution des estimateurs, on simule, sous l'hypothèse **H0**, 100, 200 et 500 fois 100 réalisations *i.i.d.* du vecteur gaussien $(Z(x_1), Z(x_2), \ldots, Z(x_n))$. On construit ensuite les histogrammes de chacun des estimateurs $\hat{\alpha}_T$ et $\hat{\beta}_T$, pour chaque série de simulations, afin d'avoir une idée, à distance finie, de la distribution des estimateurs et comparer cette distribution empirique à la distribution gaussienne.

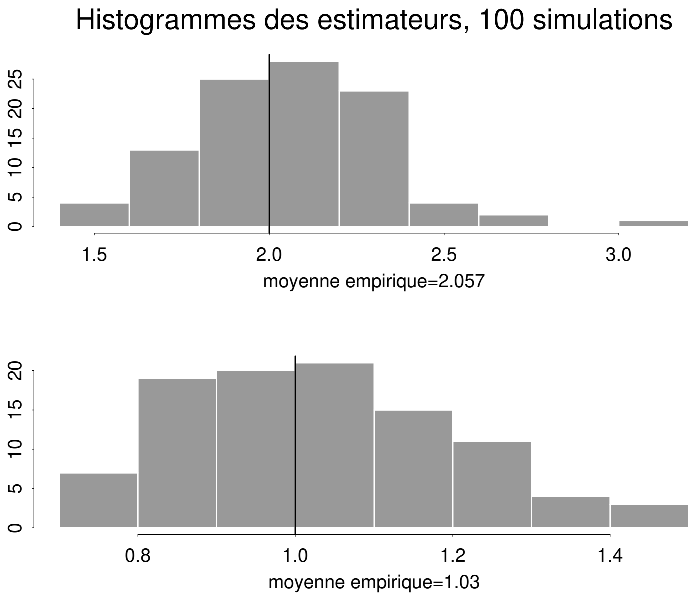


FIG. 2.5 – *Distribution empirique de $\hat{\alpha}_T$ (figure de dessus) et de $\hat{\beta}_T$ (figure de dessous)*

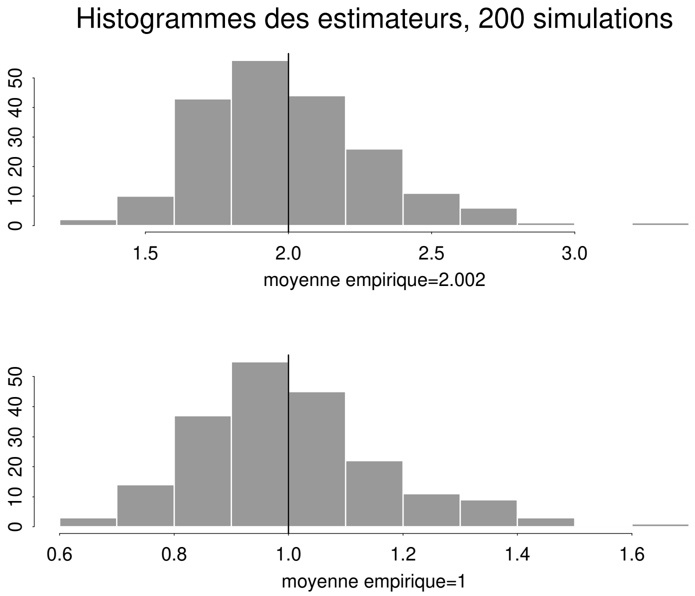


FIG. 2.6 – *Distribution empirique de $\hat{\alpha}_T$ (figure de dessus) et de $\hat{\beta}_T$ (figure de dessous)*

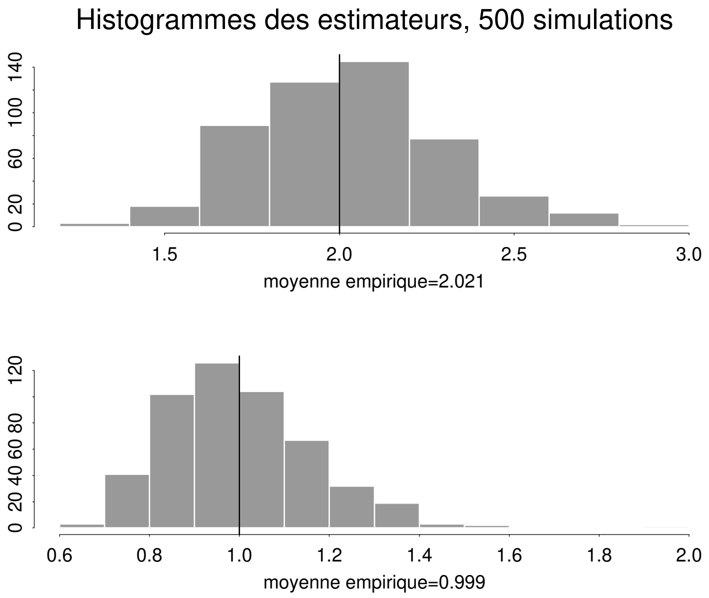


FIG. 2.7 – *Distribution empirique de $\hat{\alpha}_T$ (figure de dessus) et de $\hat{\beta}_T$ (figure de dessous)*

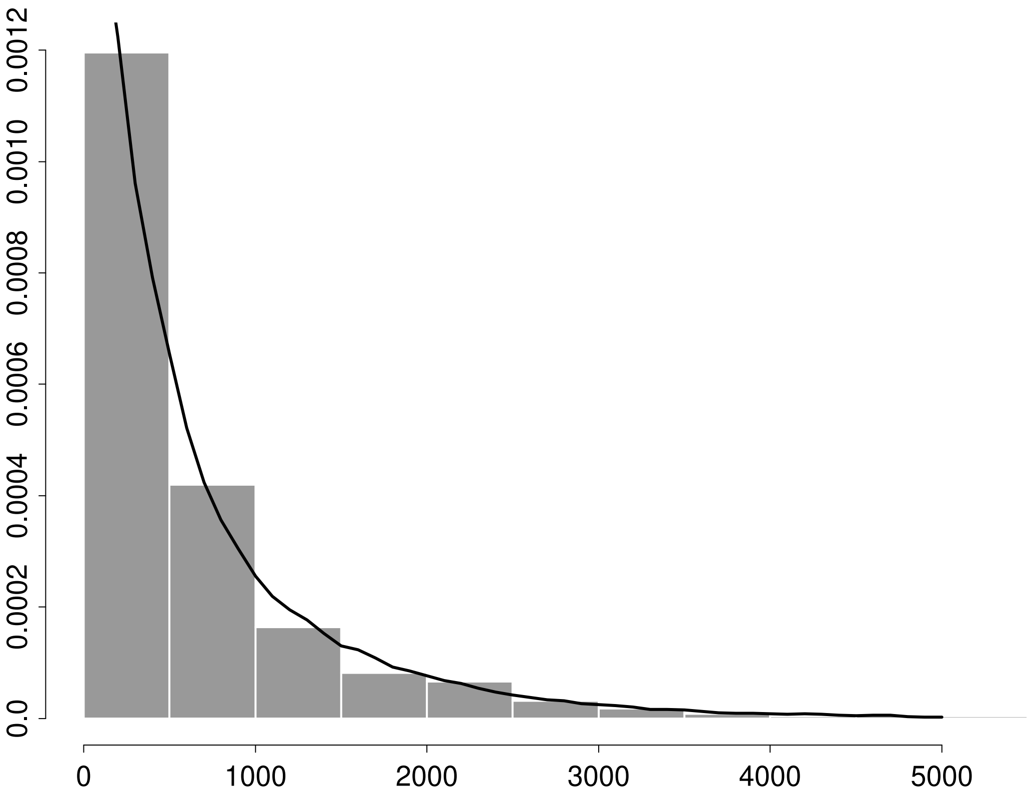


FIG. 2.8 – *Histogramme de* $2T\Big(U_T(\alpha_0,\beta_0) - U_T(\hat{\alpha}_T,\hat{\beta}_T)\Big)$, *la courbe représente la densité de* $\sum_{k=1}^{2} \lambda_k \chi_k^2(1)$

On constate que si le nombre N de simulations pour valablement approcher les lois de $\hat{\alpha}_T$ et $\hat{\beta}_T$ est suffisant ($N \geq 100$, *cf.* figures 2.5 à 2.7), la moyenne empirique de chaque estimateur est proche de la valeur réelle et les histogrammes sont approximativement gaussiens.

2.6.4 Test de différence de contraste d'une hypothèse complètement spécifiée

On rappelle le résultat du théorème 2.8 :

$$2T\left(U_T(\alpha_0,\beta_0) - U_T(\hat{\alpha}_T,\hat{\beta}_T)\right) \xrightarrow[T\to+\infty]{loi} \sum_{k=1}^{2} \lambda_k \chi_k^2(1), \text{ sous } \mathbf{H0}$$

où λ_1 et λ_2 sont les valeurs propres de $L_2^{-1}(\alpha_0,\beta_0)J_2(\alpha_0,\beta_0)$, $L_2(\alpha_0,\beta_0)$ et $J_2(\alpha_0,\beta_0)$ matrices définies p.60 :

$$\begin{cases} \lambda_1 = 557.20 \\ \lambda_2 = 99.35 \end{cases}$$

Sous l'hypothèse **H0**, on simule 1000 fois 1000 réalisations *i.i.d.* de $(Z(x_1), Z(x_2), \ldots, Z(x_n))$. On construit ensuite l'histogramme correspondant de la statistique $2T\left(U_T(\alpha_0,\beta_0) - U_T(\hat{\alpha}_T,\hat{\beta}_T)\right)$.

La figure 2.8 nous montre que l'histogramme de $2T\left(U_T(\alpha_0,\beta_0) - U_T(\hat{\alpha}_T,\hat{\beta}_T)\right)$ sur ces 1000 simulations dessine assez bien la loi de $\sum_{k=1}^{2} \lambda_k \chi_k^2(1)$, loi pour laquelle on trace la densité de la manière suivante :

1. on a simulé de manière indépendante deux fois un échantillon *i.i.d.* de 30000 points selon la loi du $khi - deux$ à un degré de liberté. On multiplie la première série de 30000 points par λ_1 et la deuxième série par λ_2 ;

2. à l'aide de ces 60000 points pondérés par λ_1 et λ_2, on estime la densité de $\sum_{k=1}^{2} \lambda_k \chi_k^2(1)$ par une méthode non-paramétrique à noyau gaussien (fonction ***density*** de "S-plus").

2.7 Généralisations

2.7.1 Cas où la variance de Z n'est pas constante

Nous allons considérer le cas, plus réaliste, où l'écart-type de Z n'est plus constant mais appartient à une famille paramétrique de modèles $\{\sigma(x;\kappa), \kappa \in \mathbb{R}^r, r \in \mathbb{N}^*\}$.

Tous les résultats obtenus avec une variance constante peuvent s'étendre ici, à des modifications mineures près qui tiennent à l'écriture du modèle.

On suppose toujours que Z est défini sur $[0,1]$, à valeurs réelles, continu en moyenne quadratique, de moyenne M_0 et que l'hypothèse **M2** est vérifiée ainsi que l'hypothèse **M1** ou **M3** selon le résultat cherché.

Soit $\mathcal{C} = \{C(.;\beta)\}$, une famille de fonctions de corrélation stationnaire continues dans $[-1,1]$, paramétrée par $\beta \in \mathbb{R}^q, q \in \mathbb{N}^*$, avec $\sqrt{C(0;\beta)} = 1$, pour tout β.

La famille de modèles de déformation est toujours $\mathcal{F}$.

On note $\varpi = (\alpha, \beta, \kappa)$ et on considère l'hypothèse :

- **h0** : il existe une valeur α_0, une valeur κ_0 et un processus stationnaire $Y = \{Y(u)\}$, défini dans $[0,1]$, de fonction de corrélation $c_Y(u,u') = C(u'-u;\beta_0)$ appartenant à $\mathcal{C}$, tels que, pour tout x dans $[0,1]$:
$$Y\left(f(x;\alpha_0)\right) = \frac{Z(x)}{\sigma(x;\kappa_0)}$$

Sous l'hypothèse **h0**, Z a pour fonction de covariance :

$$\begin{aligned} r_Z(x,x') &= c_Y\left(f(x;\alpha_0), f(x';\alpha_0)\right) \times \sigma(x;\kappa_0)\sigma(x';\kappa_0) \\ &= C\left(f(x';\alpha_0) - f(x;\alpha_0);\beta_0\right) \times \sigma(x;\kappa_0)\sigma(x';\kappa_0) \\ &= r(x,x',\varpi_0) \end{aligned}$$

avec $\varpi_0 = (\alpha_0, \beta_0, \kappa_0)$.

Si nous gardons le même schéma expérimental et le même estimateur empirique de la covariance que précédemment, les résultats du paragraphe 2.2.2, concernant les propriétés de cet estimateur, restent inchangés.

Quant à l'estimation des paramètres du modèle proprement dite :

(i) le contraste devient :

$$\begin{aligned} u_T(\varpi) &= \sum_{i \leq j} \left[\hat{r}_T(x_i,x_j) - r(x_i,x_j;\varpi)\right]^2 \\ &= \sum_{i \leq j} \left[\hat{r}_T(x_i,x_j) - C\left(f(x';\alpha) - f(x;\alpha);\beta\right) \times \sigma(x;\kappa)\sigma(x';\kappa)\right]^2 \end{aligned}$$

(ii) la fonction de contraste s'écrit :

$$\begin{aligned} k(\varpi_0, \varpi) &= \sum_{i \leq j} [r(x_i, x_j; \varpi_0) - r(x_i, x_j; \varpi)]^2 \\ &= \sum_{i \leq j} [C\left(f(x'; \alpha_0) - f(x; \alpha_0); \beta_0\right) \times \sigma(x; \kappa_0)\sigma(x'; \kappa_0) \\ &- C\left(f(x'; \alpha) - f(x; \alpha); \beta\right) \times \sigma(x; \kappa)\sigma(x'; \kappa)]^2 \end{aligned}$$

Finalement, on peut sans peine reprendre tous les résultats obtenus dans les sections 2.3, 2.4 et 2.5, et montrer qu'ils restent valables sous la nouvelle hypothèse h0. Toutefois, il faut adapter correctement les hypothèses. En particulier, les hypothèses **B1**-**B3** et **D1**-**D2** et deviennent :

- **b1** : $\varpi = (\alpha, \beta, \kappa) \in \Upsilon$ compact de $\mathbb{R}^{p+q+r}$, $(p, q, r) \in \mathbb{N}^{*3}$, et $\varpi_0 \in \mathring{\Upsilon}$
- **b2** : $\varpi \longmapsto r(x, x'; \varpi)$ est continue dans Υ, pour tout (x, x') dans $[0, 1]^2$
- **b3** : l'application :

$$\varpi \longmapsto r(\varpi) = \begin{pmatrix} r(x_1, x_1; \varpi) \\ r(x_1, x_2; \varpi) \\ \vdots \\ r(x_n, x_n; \varpi) \end{pmatrix}$$

  est injective

- **d1** : $\varpi \longmapsto r(x, x'; \varpi)$ est de classe C^2 dans $\mathring{\Upsilon}$, pour tout (x, x') dans $[0, 1]^2$
- **d2** : $\varpi \longmapsto r(\varpi)$ est de rang $(p + q + r)$ en ϖ_0

2.7.2 Cas où Z est défini sur une partie de $\mathbb{R}^d$, $d \geq 2$

Dans toute l'étude qu'on a menée jusqu'ici, la seule hypothèse faite sur la famille de modèles de déformation paramétrique de $[0, 1]$ dans $[0, 1]$ concernait la bijection et la continuité.

En dehors des exemples spécifiques étudiés (les fonctions I-splines (*cf.* paragraphe 2.3.3) et l'étude expérimentale de la section 2.6), la structure totalement ordonnée de $\mathbb{R}$ n'a jamais été utilisée que ce soit pour l'estimation des paramètres ou pour l'étude des propriétés asymptotiques des estimateurs. Cette structure n'est intervenue que pour décrire explicitement des familles de modèles de transformation strictement croissante f de $\mathbb{R}$ dans $\mathbb{R}$:

$$x < x' \Longleftrightarrow f(x) < f(x')$$

condition assurée pour f dérivable si, pour tout x dans $\mathbb{R}$:

$$f^{(1)}(x) \geq 0 \text{ et } f^{(1)}(x) \neq 0, \lambda_1 - p.s.$$

Considérons maintenant Z, un champ défini sur une partie G de $\mathbb{R}^d$, $d \geq 2$, continu en moyenne quadratique, de moyenne M_0 et supposons que l'hypothèse **M2** est vérifiée ainsi que l'une ou l'autre de ces hypothèses :

- **M1'** : $\sup_{x\in G} E|Z(x)|^4 < +\infty$
- **M3'** : $\exists \nu > 2, \sup_{x\in G} E|Z(x)|^{2\nu} < +\infty$

selon le résultat cherché.

Soit $\mathcal{R}_d = \{R(.;\beta)\}$, $\beta \in \mathbb{R}^q$, une famille de fonctions de covariance continues, stationnaires isotropes et définies sur $\mathbb{R}^+$ (*cf.* définition 1.4).

Soit $\mathcal{F}_d = \{f(.;\alpha)\}$, $\alpha \in \mathbb{R}^p$, une famille d'homéomorphismes définis sur G.

On considère l'hypothèse :

- $\mathbf{H0}_d$: il existe une valeur α_0 et un champ stationnaire isotrope $Y = \{Y(u)\}$, de moyenne M_0 et de fonction de covariance $r_Y(u, u') = R(\|u' - u\|_d; \beta_0)$ appartenant à $\mathcal{R}_d$, tels que, pour tout x dans G :
$$Y(f(x;\alpha_0)) = Z(x)$$

Sous l'hypothèse $\mathbf{H0}_d$, le champ Z a pour fonction de covariance, pour tout (x, x') dans G^2 :

$$\begin{aligned} r_Z(x,x') &= r_Y(f(x;\alpha_0), f(x';\alpha_0)) \\ &= R(\|f(x';\alpha_0) - f(x;\alpha_0)\|_d; \beta_0) \\ &= r(x,x',\theta_0) \end{aligned}$$

avec $\theta_0 = (\alpha_0, \beta_0)$.

On choisit encore comme estimateur de θ_0, l'estimateur du minimum de contraste :

$$U_{T,d}(\theta) = \sum_{i\leq j} \left[\hat{r}_T(x_i,x_j) - r(x_i,x_j;\theta)\right]^2$$

avec :

$$\begin{cases} \hat{r}_T(x_i,x_j) = \dfrac{1}{T-1}\displaystyle\sum_{t=1}^{T} \left(Z_t(x_i) - \bar{Z}_T(x_i)\right)\left(Z_t(x_j) - \bar{Z}_T(x_j)\right) \\ r(x_i,x_j;\theta) = R(\|f(x_j;\alpha) - f(x_i;\alpha)\|_d; \beta) \end{cases}$$

où $\theta = (\alpha, \beta)$ et $x_1, x_2, \ldots, x_n$ sont n stations de G en lesquelles on observe T réalisations *i.i.d.* de Z.

Que deviendrait l'étude précédente (de la section 2.3 à la section 2.5) sur les propriétés des estimateurs et les tests asymptotiques ?

Sous les conditions de compacité de l'espace des paramètres, d'identifiabilité paramétrique du modèle et de régularité :

- $\mathbf{B1}_d$: $\theta \in \Theta$ compact de $\mathbb{R}^{p+q}$, $(p+q) \in \mathbb{N}^{*2}$, et $\theta_0 \in \overset{\circ}{\Theta}$
- $\mathbf{B2}_d$: $\theta \longmapsto r(x, x'; \theta)$ est continue dans Θ, pour tout (x, x') dans G^2
- $\mathbf{B3}_d$: l'application :
$$\theta \longmapsto r(\theta) = \begin{pmatrix} r(x_1, x_1; \theta) \\ r(x_1, x_2; \theta) \\ \vdots \\ r(x_n, x_n; \theta) \end{pmatrix}$$
est injective
- $\mathbf{D1}_d$: $\theta \longmapsto r(x, x'; \theta)$ est de classe C^2 dans $\overset{\circ}{\Theta}$, pour tout (x, x') dans G^2
- $\mathbf{D2}_d$: $\theta \longmapsto r(\theta)$ est de rang $(p+q)$ en θ_0

toute l'étude statistique se maintiendrait.

En fait, la difficulté et la spécificité de l'étude sur $\mathbb{R}^d$ résident dans la description des bijections de $\mathbb{R}^d$ dans $\mathbb{R}^d$, $d \geq 2$. On dispose toutefois du résultat général suivant :

***Théorème* 2.13** *(Lelong-Ferrand et Arnaudiès (1977), p.245)*
Si G est un ouvert de $\mathbb{R}^d$, et si $f : G \longrightarrow \mathbb{R}^d$ est une application injective de classe C^k, $k \geq 1$, dont le jacobien ne s'annule pas, alors $f(G)$ est un ouvert de $\mathbb{R}^d$, et f définit un difféomorphisme de classe C^k de G sur $f(G)$.

★

et du résultat particulier pour $d = 2$, avec $G = [0, 1]^2$: si on fait l'hypothèse que la déformation selon l'axe des abscisses x est indépendante de la déformation selon l'axe des ordonnées y, on peut utiliser les déformations $f(.; \alpha)$, définies pour tout (x, y) dans $[0, 1]^2$ par :

$$f(x, y; \alpha) = \begin{pmatrix} f_1(x; \alpha_1) \\ f_2(y; \alpha_2) \end{pmatrix}$$

avec $\alpha = (\alpha_1, \alpha_2)$ et où $f_1(.; \alpha_1)$ et $f_2(.; \alpha_2)$ sont deux homéomorphismes de $[0, 1]$ sur $[0, 1]$.

Exemple

On considère un champ Z gaussien défini sur $[0,1]^2$, non-stationnaire, centré et réduit, et continu en moyenne quadratique. La déformation bijective et bicontinue de $[0,1]^2$ sur $[0,1]^2$ rendant Z stationnaire est la fonction :

$$f(x,y;\alpha_{1,0},\alpha_{2,0}) = \begin{pmatrix} x^{\alpha_{1,0}} \\ y^{\alpha_{2,0}} \end{pmatrix}, \text{ avec } \begin{cases} \alpha_{1,0} = 1/2 \\ \alpha_{2,0} = 2 \end{cases}$$

On prend comme modèle de covariance stationnaire isotrope, la covariance exponentielle (*cf.* figure 2.3) :

$$u \longmapsto \exp\left(-\beta_0 \|u\|_2\right), \quad \text{avec } \beta_0 = 1$$

Le plan d'expérience est constitué de $(n \times n)$ stations $\{(x_i, y_j),\ 1 \le i,j \le n \text{ et } n = 6\}$ régulièrement réparties sur $[0,1]^2$ (graphique en haut à gauche de la figure 2.9) :

$$\begin{cases} x_i = \dfrac{i-1}{n-1},\ i = 1,\ldots,n \\[2ex] y_j = \dfrac{j-1}{n-1},\ j = 1,\ldots,n \end{cases}$$

Le graphique en haut à droite de la figure 2.9 représente la déformation par $f(.,.;\alpha_{1,0},\alpha_{2,0})$ de ce plan.

Pour calculer les covariances empirique, pour tout $(i,j) \le (k,l)$ (au sens de l'ordre lexicographique sur $\mathbb{N}^2$) :

$$\hat{r}_T(x_i,y_j,x_k,y_l) = \frac{1}{T}\sum_{t=1}^{T} Z_t(x_i,y_j)Z_t(x_k,y_l)$$

on simule, sous l'hypothèse $\mathbf{H0}_2$, $T = 1000$ réalisations *i.i.d.* du vecteur gaussien $(Z(x_i,y_j))_{1\le i,j\le n}$ centré de matrice de variances-covariances $[r_Z(x_i,y_j,x_k,y_l)]_{n^2,n^2}$, avec pour tout $1 \le i,j,k,l \le n$:

$$r_Z\left(x_i,y_j,x_k,y_l\right) = \exp\left(-\sqrt{(x_k^{1/2} - x_i^{1/2})^2 + (y_l^2 - y_j^2)^2}\right)$$

Le graphique en bas à gauche de la figure 2.9 représente les covariances empiriques en fonction des distances $\|(x_i,y_j) - (x_k,y_l)\|_2$ dans l'espace géographique alors que le graphique en bas à droite représente les covariances empiriques en fonction des distances $\|(x_i^{1/2},y_j^2) - (x_k^{1/2},y_l^2)\|_2$ dans l'espace déformé.

L'étude numérique qui suit a pour but d'illustrer expérimentalement le théorème 2.11 à la base du test de différence de contraste d'une sous-hypothèse.

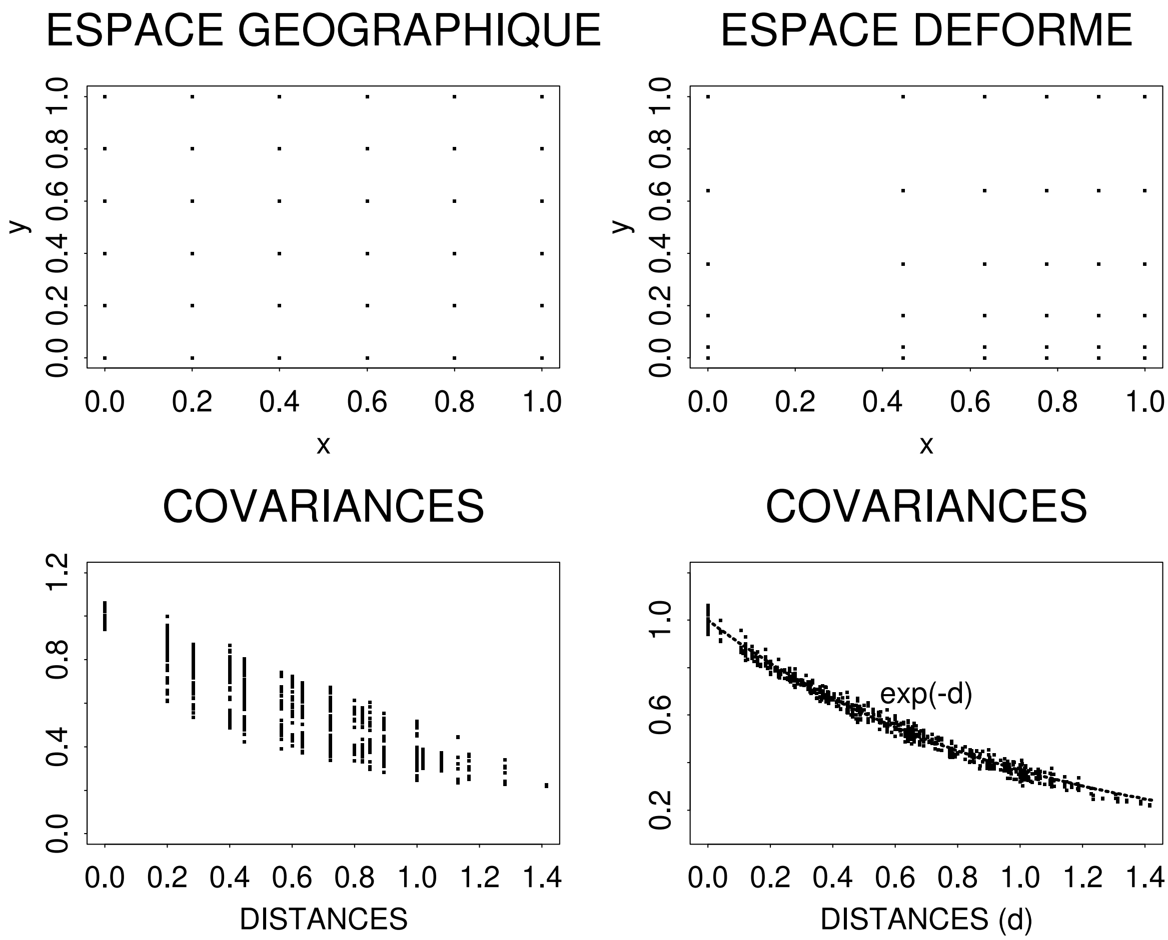


FIG. 2.9 – *Les covariances empiriques en fonction des distances dans l'espace géographique (à gauche) et en fonction des distances dans l'espace déformé (à droite)*

Identifiabilité paramétrique du modèle

Sous l'hypothèse $\mathbf{H0}_2$, le modèle de covariance s'écrit, pour tout (x, y, x', y') dans $[0,1]^4$:

$$r(x, y, x', y'; \alpha_{1,0}, \alpha_{2,0}, \beta_0) = \exp\left(-\beta_0\sqrt{\left(x'^{\alpha_{1,0}} - x^{\alpha_{1,0}}\right)^2 + \left(y'^{\alpha_{2,0}} - y^{\alpha_{2,0}}\right)^2}\right)$$

L'application :

$$(\alpha_1, \alpha_2, \beta) \longmapsto \exp\left(-\beta\sqrt{\left(x'^{\alpha_1} - x^{\alpha_1}\right)^2 + \left(y'^{\alpha_2} - y^{\alpha_2}\right)^2}\right)$$

est continue pour tout (x, y, x', y') dans $[0,1]^4$; l'hypothèse $\mathbf{B2}_2$ est bien vérifiée.

Le fait de savoir que Z est réduit nous permet de ne considérer que les couples de stations $((x_i, y_j), (x_k, y_l))$ avec $(i,j) < (k,l)$. Pour les couples $((x_1, y_1), (x_2, y_1))$, $((x_1, y_1), (x_3, y_1))$ et $((x_1, y_1), (x_1, y_2))$, on a :

$$\begin{pmatrix} \exp\left(-\beta x_2^{\alpha 1}\right) \\ \exp\left(-\beta x_3^{\alpha_1}\right) \\ \exp\left(-\beta y_2^{\alpha_2}\right) \end{pmatrix} = \begin{pmatrix} \exp\left(-\beta' x_2^{\alpha_1'}\right) \\ \exp\left(-\beta' x_3^{\alpha_1'}\right) \\ \exp\left(-\beta' y_2^{\alpha_2'}\right) \end{pmatrix} \Longleftrightarrow (\alpha_1, \alpha_2, \beta) = (\alpha_1', \alpha_2', \beta')$$

autrement dit, l'hypothèse $\mathbf{B3}_2$ est assurée et l'identifiabilité paramétrique du modèle aussi.

On suppose que $(\alpha_1, \alpha_2, \beta)$ est dans un compact Θ de $\left(\mathbb{R}^{*+}\right)^3$ et que $(\alpha_{1,0}, \alpha_{2,0}, \beta_0) \in \overset{\circ}{\Theta}$; pour notre modèle, l'hypothèse $\mathbf{B1}_2$ est vérifiée.

Estimation des paramètres

$\hat{\alpha}_{1,T}$, $\hat{\alpha}_{2,T}$ et $\hat{\beta}_T$ minimisent le contraste :

$$U_{T,2}(\alpha_1, \alpha_2, \beta) = \sum_{(i,j)<(k,l)} \left[\hat{r}_T(x_i, y_j, x_k, y_l) - r(x_i, y_j, x_k, y_l; \alpha_1, \alpha_2, \beta)\right]^2$$

Pour relaxer les contraintes de positivité :

$$(S_2') : \begin{cases} \alpha_1 > 0 \\ \alpha_2 > 0 \\ \beta > 0 \end{cases}$$

on fait le changement de paramètres :

$$\begin{cases} \alpha_1 = e^{\alpha_1'} \\ \alpha_2 = e^{\alpha_2'} \\ \beta = e^{\beta'} \end{cases}$$

avec $(\alpha_1', \alpha_2', \beta')$ dans $\mathbb{R}^3$, et on minimise $U_{T,2}$ par rapport aux nouveaux paramètres α_1', α_2' et β' sans contrainte, c'est-à-dire dans $\mathbb{R}^3$.

Test de différence de contraste d'une sous-hypothèse

Nous allons tester l'absence de déformation suivant l'axe des x : $\alpha_1 = 1$.

L'hypothèse $\mathbf{D1}_2$ est bien assurée ainsi que $\mathbf{M2}$ (le calcul numérique donne $|\Sigma_0| \neq 0$).

Pour vérifier l'hypothèse $\mathbf{D2}_2$, il suffit de montrer que la matrice $R_2(1/2,2,1)$

$$= -2 \begin{bmatrix} \left((x_k^{1/2}-x_i^{1/2})^2+(y_l^2-y_j^2)^2\right)^{-1/2}(x_k^{1/2}\ln(x_k)-x_i^{1/2}\ln(x_i))(x_k^{1/2}-x_i^{1/2})r_Z\,(x_i,y_j,x_k,y_l) \\ \left((x_k^{1/2}-x_i^{1/2})^2+(y_l^2-y_j^2)^2\right)^{-1/2}(y_l^2\ln(y_l)-y_j^2\ln(y_j))(y_l^2-y_j^2)r_Z\,(x_i,y_j,x_k,y_l) \\ \left((x_k^{1/2}-x_i^{1/2})^2+(y_l^2-y_j^2)^2\right)^{1/2} r_Z\,(x_i,y_j,x_k,y_l) \end{bmatrix}_{3,\,((i,j)<(k,l))}$$

(avec, par prolongement par continuité en $x_1=0$ et $y_1=0$, $x_1^{1/2}\ln(x_1)=y_1^2\ln(y_1)=0$) est de rang 3. Cela est vraie puisque :

$$\begin{vmatrix} \sqrt{x_2}\ln(x_2)r_z(x_1,y_1,x_2,y_1) & \sqrt{x_3}\ln(x_3)r_z(x_1,y_1,x_3,y_1) & 0 \\ 0 & 0 & y_2^2\ln(y_2)r_z(x_1,y_1,x_1,y_2) \\ \sqrt{x_2}r_z(x_1,y_1,x_2,y_1) & \sqrt{x_3}r_z(x_1,y_1,x_3,y_1) & y_2^2r_z(x_1,y_1,x_1,y_2) \end{vmatrix}$$

$$= y_2^2r_z(x_1,y_1,x_1,y_2)r_z(x_1,y_1,x_2,y_1)r_z(x_1,y_1,x_3,y_1)\sqrt{x_2x_3}\,(\ln(x_3)-\ln(x_2)) \neq 0$$

On rappelle le résultat du théorème 2.11 :

$$\Gamma_T = 2T\left(U_{T,2}(1,\bar{\alpha}_{1,T},\bar{\beta}_T)-U_{T,2}(\hat{\alpha}_{1,T},\hat{\alpha}_{2,T},\hat{\beta}_T)\right) \xrightarrow[T\to+\infty]{loi} \lambda\chi^2(1), \text{ sous } \alpha_1=1$$

où $(1,\bar{\alpha}_{1,T},\bar{\beta}_{2,T})$ est l'estimateur contraint et λ est la valeur propre de $J_3(\alpha_{1,0},\alpha_{2,0},\beta_0)\Psi_3$, $J_3(\alpha_{1,0},\alpha_{2,0},\beta_0)$ et Ψ_3 matrices définies p.67 :

$$\lambda = 59.42$$

Sous l'hypothèse $\mathbf{H0}_2$, on simule 1 fois $T=1000$ réalisations *i.i.d.* de $(Z(x_i,y_j))_{1\leq i,j\leq n}$. Sur cette simulation, on a :

$$\Gamma_{1000} = 3292.85 > \lambda Q_{0.95}(1) = 228.26$$

où $Q_{0.95}(1)=3.84$ est le quantile d'ordre 0.95 de $\chi^2(1)$. Par conséquent, au niveau asymptotique 0.05, on rejette l'hypothèse du sous-modèle $\alpha_1=1$.

Pour construire l'histogramme de la statistique Γ_{1000} sous l'hypothèse du sous-modèle $\alpha_1=1$, on simule sous cette hypothèse 1000 fois 1000 réalisations *i.i.d.* de $(Z(x_i,y_j))_{1\leq i,j\leq n}$. La figure 2.10 nous montre que l'histogramme de Γ_{1000} dessine assez bien la loi de $\lambda\chi^2(1)$.

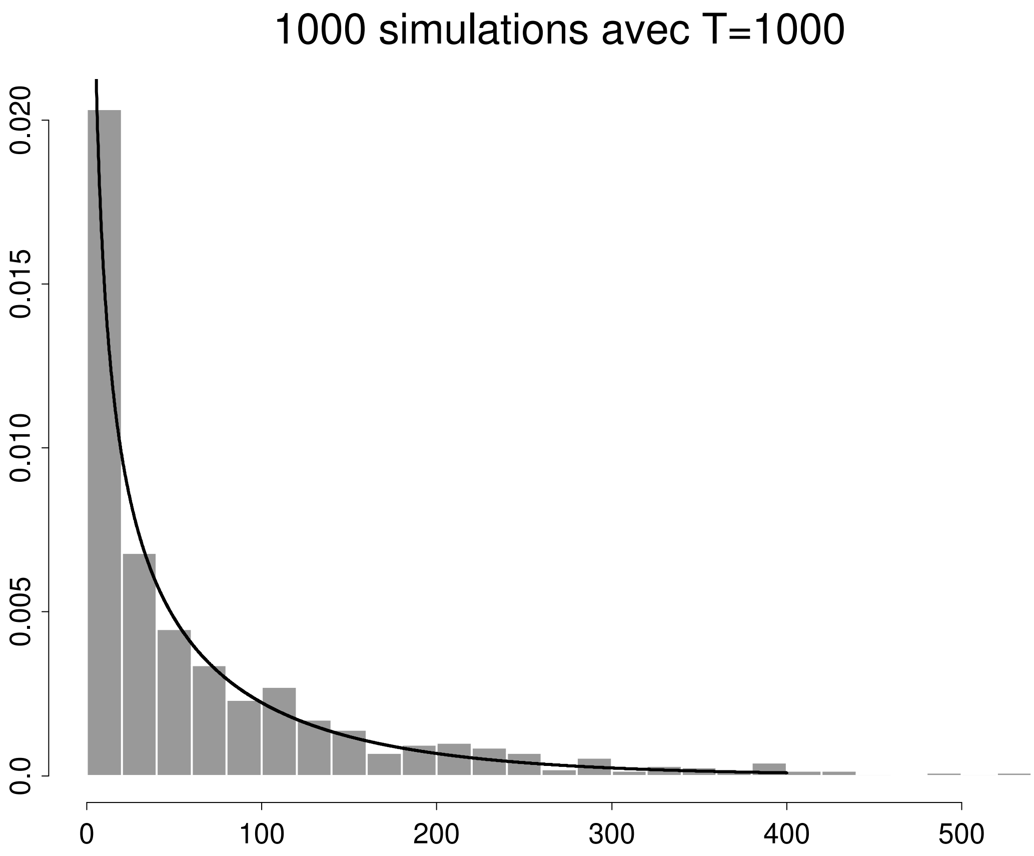


FIG. 2.10 – *Histogramme de* $2T\left(U_T(1,\bar{\alpha}_{2,T},\bar{\beta}_T)-U_T(\hat{\alpha}_{1,T},\hat{\alpha}_{2,T},\hat{\beta}_T)\right)$ *sous l'hypothèse du sous-modèle* $\alpha_1=1$*, la courbe représente la densité de* $\lambda\chi^2(1)$

Chapitre 3

Le cas d'une déformation non-paramétrique

Dans ce chapitre, on s'intéresse à l'estimation des paramètres de la fonction de covariance d'un processus Z, réductible à une covariance stationnaire, lorsque la déformation appartient à une famille non-paramétrique de modèles. Le modèle de covariance stationnaire appartient quant à lui à une famille paramétrique, comme dans le chapitre 2. Dans la section 3.1, on introduit le modèle et on présente un nouveau schéma expérimental dans lequel les stations "densifient" notre domaine d'observation pour permettre l'estimation de la déformation partout. Dans la section 3.2, on définit l'estimateur empirique de la covariance et on étudie ses convergences uniformes en probabilité et p.s. Dans la section 3.3, on s'intéresse à l'estimation non-paramétrique de la déformation et du paramètre de la covariance stationnaire. La méthode d'estimation utilisée est une méthode du minimum de contraste dans un cadre général où on suppose seulement que l'espace des paramètres, ici de dimension infinie, est relativement compact. Les propriétés asymptotiques des estimateurs, convergence en probabilité et convergence p.s., sont étudiées dans la section 3.4. Dans la section 3.5, on étudie un cas particulier d'estimation lorsque la fonction de covariance stationnaire est connue et bijective sur $[0,1]$ *; dans ce cadre, nous traitons un exemple.*

Dans ce chapitre, notre intérêt majeur porte sur la déformation et son estimation non-paramétrique.

3.1 Présentation du modèle et schéma expérimental

Le modèle est défini dans le paragraphe 3.1.1. Ensuite, on présente le schéma expérimental et on précise dans quel cadre asymptotique on se place dans le paragraphe 3.1.2.

3.1.1 Le modèle

Soit $Z = \{Z(x)\}$ un processus aléatoire défini sur $[0,1]$, à valeurs réelles, de moyenne $M_0 = 0$, de variance σ_0^2 inconnue mais constante et de fonction de covariance r_Z. On suppose en outre :

- **N** : Z est continu en moyenne quadratique

L'hypothèse **N** est équivalente à la continuité de r_Z dans $[0,1]^2$, compact ; r_Z est donc uniformément continue dans $[0,1]^2$.

Soit $\mathcal{R} = \{R(.;\beta)\}$, une famille de fonctions de covariance stationnaire continues dans $[-1,1]$, paramétrée par $\beta \in \mathbb{R}^q, q \in \mathbb{N}^*$. $\sqrt{R(0;\beta)} = \sigma$, le paramètre courant de l'écart-type de Z, est inclus dans β.

On considère par ailleurs la famille non-paramétrique $\mathcal{F}_K$ des homéomorphismes croissants f de $[0,1]$ sur $[0,1]$, continuement dérivables par morceaux et tels que leur dérivée à droite vérifie pour tout x dans $[0,1[$:

$$0 < f_d^{(1)}(x) \leq K$$

où K, indépendante de f, est une constante connue (nécessairement $K \geq 1$).

On munit $\mathcal{F}_K \subset C([0,1])$ de la topologie de la convergence uniforme. Les fonctions f de $\mathcal{F}_K$ sont lipschitziennes de rapport K sur $[0,1]$, c'est-à-dire, pour tout f dans $\mathcal{F}_K$ et tout (x,x') dans $[0,1]^2$:

$$|f(x') - f(x)| \leq K|x' - x|$$

De plus, pour tout f :

$$\|f\|_\infty = 1$$

D'après le théorème d'Arzelà-Ascoli, l'ensemble $\mathcal{F}_K$ est relativement compact dans $C([0,1])$, autrement dit $\bar{\mathcal{F}}_K$ est compact.

L'ensemble $\mathcal{F}_K$ est une classe assez riche de modèles de déformation. En particulier, les I-splines (*cf.* annexe C) de pentes inférieures ou égales à K appartiennent à $\mathcal{F}_K$, et l'ensemble des combinaisons convexes des I-splines de degré 1, ou lignes brisées, de pentes inférieures ou égales à K, est dense dans $\mathcal{F}_K$.

On note $\theta = (f,\beta)$ et on considère l'hypothèse :

- **H0** : il existe une fonction f_0 de $\mathcal{F}_K$ et un processus stationnaire $Y = \{Y(u)\}$, défini dans $[0,1]$, de fonction de covariance $r_Y(u,u') = R(u'-u;\beta_0)$ appartenant à $\mathcal{R}$, tels que, pour tout x dans $[0,1]$:

$$Y(f_0(x)) = Z(x)$$

Sous l'hypothèse **H0**, le processus Z a pour fonction de covariance, pour tout (x, x') dans $[0,1]^2$:

$$\begin{aligned} r_Z(x, x') &= r_Y\left(f_0(x), f_0(x')\right) \\ &= R\left(f_0(x') - f_0(x); \beta_0\right) \\ &= r_{\theta_0}(x, x') \end{aligned}$$

avec $\theta_0 = (f_0, \beta_0)$.

L'objet de ce chapitre est d'estimer le paramètre fonctionnel de la déformation f_0 et le paramètre de la covariance β_0, puis d'étudier les convergences en probabilité et *p.s.* des estimateurs de ces deux paramètres.

3.1.2 Schéma expérimental

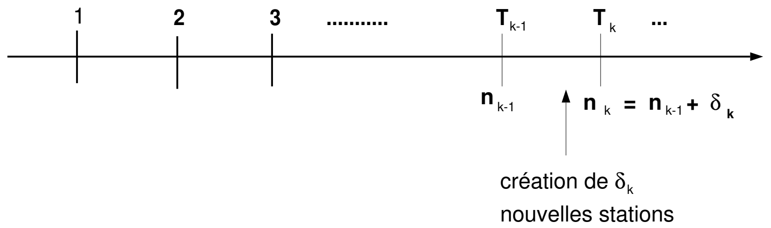
La déformation f_0 appartient à une famille non-paramétrique de modèles. Pour l'identifier, il faut donc l'estimer sur tout $[0, 1]$ à partir des observations sur Z : les points d'observation, ou stations, doivent donc être "denses" dans $[0, 1]$. On dispose de T, $T \in \mathbb{N}^*$, réalisations indépendantes de Z. Mais au lieu d'être observées en un nombre fini et fixé de stations, ces T réalisations sont observées sur un nombre fini et croissant de stations qui tendent à "densifier" $[0, 1]$. Dans ce contexte, nous proposons le schéma expérimental "triangulaire" suivant :

1. de nouvelles stations sont ajoutées, en supplément des précédentes, aux instants $T_k \in \mathbb{N}^*$, $k = 1, 2, \dots, m, \dots$:

$$T_1 < T_2 \cdots < T_m \leq T < T_{m+1} < \cdots$$

2. si on note $n_k, k \geq 1$, le nombre de stations présentes en T_k, et $\delta_k = n_k - n_{k-1} > 0$ le nombre de stations supplémentaires introduites en T_k, entre T_{k-1} et T_k il y a $(T_k - T_{k-1} + 1)$ réalisations *i.i.d.* des n_{k-1} observations de Z aux stations présentes en T_{k-1} ;

3. le nombre de stations n_m tend vers $+\infty$ avec m, de telle sorte que les positions des stations densifient $[0, 1]$.



On prend $T_1 = 1$, on pose $n_0 = 0$ et $T_0 = 0$, et on suppose que le nombre de stations initiales, $n_1 = \delta_1$, est supérieur ou égal à 2.

Dans toutes les études asymptotiques qui vont suivre, c'est m qu'on fera tendre vers $+\infty$.

En résumé, on dispose à l'instant $T \in [T_m, T_{m+1}[$ du schéma triangulaire suivant (*cf.* figure 3.2) :

	$\forall 1 \leq k \leq m, \forall 1 \leq i \leq \delta_k, 0 < x_i^k < 1$										
	x_1^1	$\dots$	$x_{\delta_1}^1$	x_1^2	$\dots$	$x_{\delta_2}^2$	$\dots$	x_1^m	$\dots$	$x_{\delta_m}^m$
$T_1 = 1$	$Z_{T_1}(x_1^1)$	$\dots$	$Z_{T_1}(x_{\delta_1}^1)$							
$\vdots$	$\vdots$	$\dots$	$\vdots$							
T_2	$Z_{T_2}(x_1^1)$	$\dots$	$Z_{T_2}(x_{\delta_1}^1)$	$Z_{T_2}(x_1^2)$	$\dots$	$Z_{T_2}(x_{\delta_2}^2)$				
$\vdots$	$\vdots$	$\dots$	$\vdots$	$\vdots$	$\dots$	$\vdots$				
T_m	$Z_{T_m}(x_1^1)$	$\dots$	$Z_{T_m}(x_{\delta_1}^1)$	$Z_{T_m}(x_1^2)$	$\dots$	$Z_{T_m}(x_{\delta_2}^2)$	$\dots$	$Z_{T_m}(x_1^m)$	$\dots$	$Z_{T_m}(x_{\delta_m}^m)$
$\vdots$	$\vdots$	$\dots$	$\vdots$	$\vdots$	$\dots$	$\vdots$	$\dots$	$\vdots$	$\dots$	$\vdots$
$T(< T_{m+1})$	$Z_T(x_1^1)$	$\dots$	$Z_T(x_{\delta_1}^1)$	$Z_T(x_1^2)$	$\dots$	$Z_T(x_{\delta_2}^2)$	$\dots$	$Z_T(x_1^m)$	$\dots$	$Z_T(x_{\delta_m}^m)$

Pour tout $k \geq 1$, notons $\mathcal{D}_k$, l'ensemble des stations supplémentaires à l'instant T_k :

$$\mathcal{D}_k = \{x_1^k, x_2^k, \dots, x_{\delta_k}^k\}$$

et Δ_k, l'ensemble de toutes les stations présentes à l'instant T_k :

$$\Delta_k = \bigcup_{l=1}^{k} \mathcal{D}_l$$

Quitte à réordonner les positions de toutes les stations présentes aux instants T_k, l'ensemble Δ_k peut aussi s'écrire :

$$\Delta_k = \{x_1 \leq x_2 \cdots \leq x_{n_k}\}$$

avec $x_1 > 0$ et $x_{n_k} < 1$.

Posons $\mu_m = \dfrac{1}{n_{m-1}} \displaystyle\sum_{x \in \Delta_{m-1}} \delta_x$. Ainsi définie, μ_m, la mesure empirique associée aux stations présentes en T_{m-1}, est une mesure ponctuelle positive et finie. On suppose que la procédure d'échantillonnage vérifie :

- **E1** : $\mu_m \xrightarrow[m \to +\infty]{c.e.} \mu$, avec μ mesure positive finie équivalente à la mesure de Lebesgue sur $[0, 1]$.

L'hypothèse **E1** implique que l'ensemble des stations devient dense dans $[0, 1]$ (*cf.* point 3. précédent), c'est-à-dire :

$$\bigcup_{k=1}^{+\infty} \mathcal{D}_k = \Delta_\infty \text{ est dense dans } [0, 1]$$

D'autres types de schémas auraient pu être pris en compte. Par exemple, on aurait pu imaginer un schéma expérimental où le nombre de stations serait fini mais où les positions des stations auraient été différentes à chaque instant (stations mobiles).

3.2 Estimateur empirique de la fonction de covariance

On définit dans le paragraphe 3.2.1 un estimateur empirique de la fonction de covariance r_Z sur $[0, 1]^2$. On étudie au paragraphe 3.2.2 les convergences uniformes en probabilité et *p.s.* de cet estimateur vers r_Z. Ces résultats de convergence seront directement utilisés dans l'étude des estimateurs des paramètres f_0 et β_0 du modèle.

3.2.1 Définition

À l'instant T_m, $m \geq 2$, on définit un estimateur empirique $\tilde{r}_m$ de la covariance r_Z dans $[0, 1]^2$ en deux étapes :

1. dans un premier temps, on estime de façon empirique r_Z dans Δ_{m-1} par :

$$\hat{r}_m(x, x') = \frac{1}{T_m - T_x \vee T_{x'} + 1} \sum_{t = T_x \vee T_{x'}}^{T_m} Z_t(x) Z_t(x')$$

   où T_x représente l'instant d'apparition de la station positionnée en x (*cf.* figure 3.1).

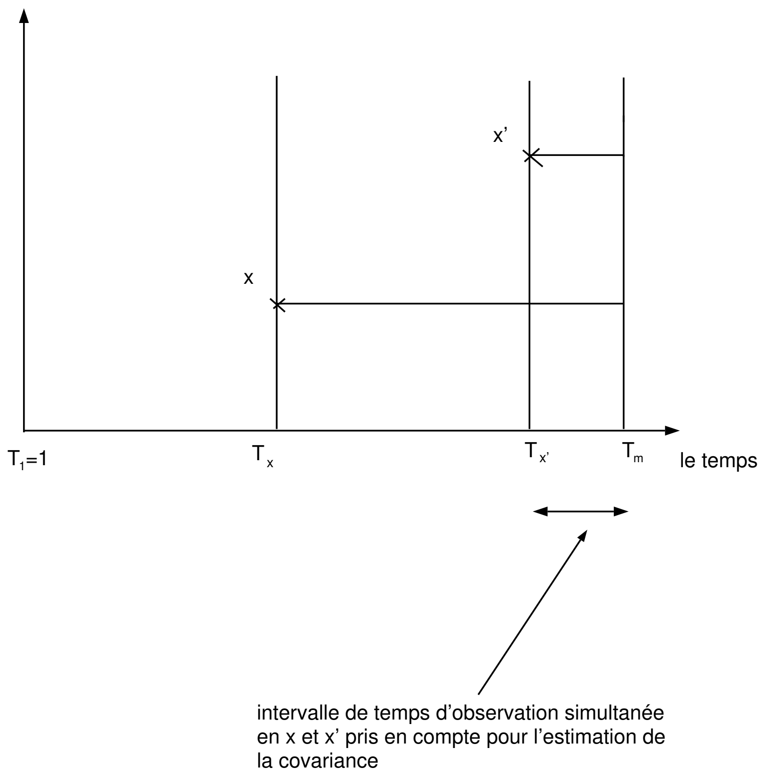


FIG. 3.1 –

L'estimateur $\hat{r}_m$ est défini seulement sur Δ^2_{m-1} et constitue de fait une version discrétisée de l'estimation de la fonction de covariance r_Z ;

2. dans un deuxième temps, pour avoir un estimateur sur $[0,1]^2$ tout entier, nous définissons, pour tout (x,x') dans $[0,1]^2$:

$$\tilde{r}_m(x,x') = \hat{r}_m\left(v_{m-1}(x,x')\right)$$

où

$$(x,x') \longmapsto v_{m-1}(x,x') = \arg \inf_{(y,y')\in\Delta^2_{m-1}} \|(x,x')-(y,y')\|_2 \tag{3.1}$$

est la fonction plus proche voisin de (x,x') dans Δ^2_{m-1}. $v_{m-1}(x,x')$ n'est pas forcément unique ; si tel est le cas, on choisit $v_{m-1}(x,x')$ selon l'ordre lexicographique sur $\mathbb{R}^2$ (si plusieurs couples de points vérifient (3.1), on prend celui qui a la plus grande abscisse et, si cela ne suffit pas (c'est-à-dire si on a l'égalité des abscisses), celui qui a la plus grande ordonnée).

Remarque : la moyenne M_0 est non-nulle

On utilise dans ce cas l'estimateur, pour tout (x, x') dans Δ^2_{m-1} :

$$\hat{r}_m(x, x') = \frac{1}{T_m - T_x \vee T_{x'} + 1} \sum_{t=T_x \vee T_{x'}}^{T_m} Z_t(x) Z_t(x') - \hat{M}_m^2$$

avec :

$$\hat{M}_m = \frac{1}{n_{m-1}} \sum_{x \in \Delta_{m-1}} \left(\frac{1}{T_m - T_x + 1} \sum_{t=T_x}^{T_m} Z_t(x) \right)$$

Par exemple si :

$$\begin{cases} \delta_1 = n_1 = 3 \\ \delta_2 = 4 \Longrightarrow n_2 = n_1 + \delta_2 = 7 \\ \delta_3 = 2 \Longrightarrow n_3 = n_2 + \delta_3 = 9 \end{cases}$$

alors l'estimateur de M_0 au temps T_3 vaut (*cf.* figure 3.2) :

$$\hat{M}_3 = \frac{1}{7} \left(\sum_{i=1}^{3} \left(\frac{1}{T_3} \sum_{t=1}^{T_3} Z_t(x_i^1) \right) + \sum_{i=1}^{4} \left(\frac{1}{T_3 - T_2 + 1} \sum_{t=T_2}^{T_3} Z_t(x_i^2) \right) \right)$$

♣

3.2.2 Convergences uniformes en probabilité et *p.s.*

Notons tout d'abord le résultat préliminaire suivant, lié à la densification des stations dans $[0, 1]$.

***Lemme* 3.1** *Sous l'hypothèse **E1** :*

$$\sup_{[0,1]^2} \|(x, x') - v_m(x, x')\|_2 \underset{m \to +\infty}{\longrightarrow} 0$$

★

Démonstration :

Posons, pour tout (x, x') dans $[0, 1]^2$:

$$d_m(x, x') = \|(x, x') - v_m(x, x')\|_2$$

Puisque $Card(\Delta^2_m) < +\infty$, d_m est continue dans $[0, 1]^2$. On a même plus :

la suite (d_m) est uniformément équicontinue dans $C([0, 1]^2)$.

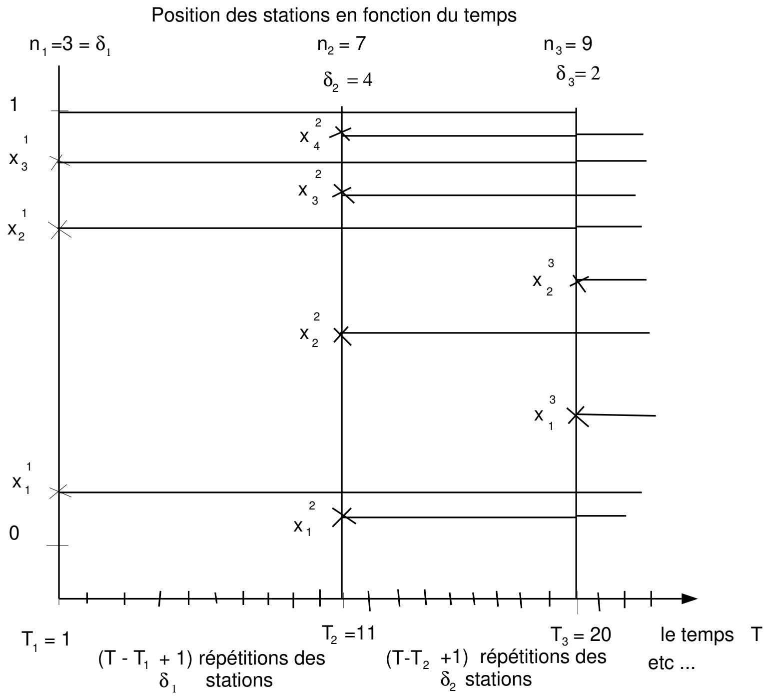


FIG. 3.2 –

En effet, on montre facilement que d_m est une application lipschitzienne de rapport 1, uniformément en m, c'est-à-dire que pour tout $((x,x'),(y,y'))$ dans $[0,1]^2 \times [0,1]^2$:

$$|d_m(x,x') - d_m(y,y')| \leq \|(x,x') - (y,y')\|_2$$

Comme Δ_∞ est dense dans $[0,1]$, Δ_∞^2 est dense dans $[0,1]^2$. Cela implique déjà que (d_m) converge simplement vers l'application nulle dans $[0,1]^2$, compact de $\mathbb{R}^2$. Pour conclure, on applique le lemme suivant :

***Lemme* 3.2** *(Dieudonné (1972), p.143)*
On suppose que E est un espace métrique compact, (f_m) une suite équicontinue dans l'espace vectoriel de toutes les applications continues de E dans l'espace normé F. Si (f_m) converge simplement vers f dans E, elle converge uniformément vers f dans E. ★

■

Pour établir les résultats de convergence qui vont suivre, nous avons encore besoin de l'hypothèse **M3**, que nous rappelons, et d'hypothèses supplémentaires sur les instants d'apparition T_k et les nombres n_k, $k \geq 1$, de stations :

– **M3** : $\exists \nu > 2, \sup\limits_{[0,1]} E|Z(x)|^{2\nu} = M_3 < +\infty$

avec l'une ou l'autre des hypothèses suivantes sur le plan d'échantillonnage :

- **E2** : $\lim_{m\to+\infty} \sum_{k=1}^{m-1} \frac{n_k^2 - n_{k-1}^2}{(T_m - T_k)^{\nu/2}} = 0$

- **E2'** : $\sum_{m=2}^{+\infty} \sum_{k=1}^{m-1} \frac{n_k^2 - n_{k-1}^2}{(T_m - T_k)^{\nu/2}} < +\infty$

L'hypothèse **E2'** implique l'hypothèse **E2** ; l'hypothèse **E2** suffira pour montrer la convergence uniforme en probabilité de $\tilde{r}_m$ vers r_Z alors que **E2'** permettra d'obtenir la convergence uniforme *p.s.*

Une condition nécessaire pour assurer l'une ou l'autre de ces deux hypothèses est :

$$\lim_{m\to+\infty}(T_m - T_{m-1}) = +\infty \text{ et } n_{m-1}^2 - n_{m-2}^2 = o\left((T_m - T_{m-1})^{\nu/2}\right)$$

L'hypothèse **E2** (ou **E2'**) signifie que la densification des stations est très lente, au sens où le nombre de stations augmente faiblement par rapport au nombre de répétitions des observations de Z en ces stations entre deux instants d'apparition, nombre de répétitions qui tend vers $+\infty$. Toutefois l'hypothèse **E2** est moins draconienne, comme on peut le constater avec les conditions suffisantes suivantes :

1. une condition suffisante pour assurer l'hypothèse **E2'** est :

$$\sum_{m=2}^{+\infty} \frac{n_{m-1}^2}{(T_m - T_{m-1})^{\nu/2}} < +\infty$$

   Par exemple, **E2'** est vérifiée si $n_k = k$ (à chaque instant T_k on ajoute une nouvelle station) et $T_k \sim k^4$ ($\Longrightarrow T_k - T_{k-1} \sim k^3$) ;

2. une condition suffisante pour assurer l'hypothèse **E2** est :

$$\lim_{m\to+\infty} \frac{n_{m-1}^2}{(T_m - T_{m-1})^{\nu/2}} = 0$$

   Si, comme au point 1. précédent, on ajoute une station à chaque étape, il "suffit" de prendre $T_k \sim k^3$ pour assurer **E2**.

Convergences uniformes en probabilité et *p.s.*

Le théorème 3.1 ci-après fournit un résultat de convergence uniforme en probabilité et un résultat de convergence uniforme *p.s.* sur $[0,1]^2$ de l'estimateur empirique $\tilde{r}_m$ de la fonction de covariance r_Z.

Pour démontrer le théorème 3.1, nous avons besoin du résultat préliminaire suivant :

***Lemme* 3.3**

1. *Sous les hypothèses **E2** et **M3** :*

$$\sup_{(x,x')\in\Delta^2_{m-1}} |\tilde{r}_m(x,x') - r_Z(x,x')| \underset{m\to+\infty}{\longrightarrow} 0, \text{ en probabilité}$$

2. *Si **E2'** remplace **E2** :*

$$\sup_{(x,x')\in\Delta^2_{m-1}} |\tilde{r}_m(x,x') - r_Z(x,x')| \underset{m\to+\infty}{\longrightarrow} 0, \text{ p.s.}$$

★

Démonstration :

Posons, pour tout (x,x') dans Δ^2_{m-1} :

$$\tilde{r}_m(x,x') = \hat{E}_m Z(x)Z(x')$$

Par définition, on a pour tout (x,x') dans Δ^2_{m-1} :

$$\tilde{r}_m(x,x') - r_Z(x,x') = \hat{E}_m Z(x)Z(x') - EZ(x)Z(x')$$

Soit $\epsilon > 0$ et définissons pour tout $m \geq 2$ l'ensemble :

$$\Phi^\epsilon_m = \left\{\omega : \sup_{(x,x')\in\Delta^2_{m-1}} \left|\hat{E}_m Z(x)Z(x') - EZ(x)Z(x')\right| \geq \epsilon\right\}$$

1. Pour obtenir la convergence uniforme en probabilité, il faut montrer que :

$$\lim_{m\to+\infty} P(\Phi^\epsilon_m) = 0 \tag{3.2}$$

Les quantités que nous manipulons étant symétriques en x et x', on a :

$$\Phi^\epsilon_m = \left\{\omega : \sup_{\substack{(x,x')\in\Delta^2_{m-1}\\ x\leq x'}} \left|\hat{E}_m Z(x)Z(x') - EZ(x)Z(x')\right| \geq \epsilon\right\}$$

et :

$$P(\Phi_m^\epsilon) \leq \sum_{\substack{(x,x')\in\Delta^2_{m-1}\\ x\leq x'}} P\left(\left\{\omega : \left|\hat{E}_m Z(x)Z(x') - EZ(x)Z(x')\right| \geq \epsilon\right\}\right)$$

On peut décomposer la somme dans le membre de droite de l'inégalité précédente selon :

$$\begin{aligned}
P(\Phi_m^\epsilon) \ \leq\ & \sum_{1\leq i\leq j\leq \delta_1} P\left(\left\{\omega : \left|\hat{E}_m Z(x_i^1)Z(x_j^1) - EZ(x_i^1)Z(x_j^1)\right| \geq \epsilon\right\}\right)\\
&+ \sum_{1\leq i\leq j\leq \delta_2} P\left(\left\{\omega : \left|\hat{E}_m Z(x_i^2)Z(x_j^2) - EZ(x_i^2)Z(x_j^2)\right| \geq \epsilon\right\}\right)\\
&+ \sum_{(i,j)\in[1,\ldots,\delta_1]\times[1,\ldots,\delta_2]} P\left(\left\{\omega : \left|\hat{E}_m Z(x_i^1)Z(x_j^2) - EZ(x_i^1)Z(x_j^2)\right| \geq \epsilon\right\}\right)\\
&\ \vdots\\
&+ \sum_{1\leq i\leq j\leq \delta_{m-1}} P\left(\left\{\omega : \left|\hat{E}_m Z(x_i^{m-1})Z(x_j^{m-1}) - EZ(x_i^{m-1})Z(x_j^{m-1})\right| \geq \epsilon\right\}\right)\\
&+ \sum_{k=1}^{m-2} \sum_{(i,j)\in[1,\ldots,\delta_k]\times[1,\ldots,\delta_{m-1}]} P\left(\left\{\omega : \left|\hat{E}_m Z(x_i^k)Z(x_j^{m-1}) - EZ(x_i^k)Z(x_j^{m-1})\right| \geq \epsilon\right\}\right)
\end{aligned}$$

On rappelle que pour tout $1 \leq k' \leq k \leq (m-1)$ et tout (i,j) dans $[1,\ldots,\delta_{k'}]\times[1,\ldots,\delta_k]$:

$$\left|\hat{E}_m Z(x_i^{k'})Z(x_j^k) - EZ(x_i^{k'})Z(x_j^k)\right| = \frac{1}{T_m - T_k + 1}\left|\sum_{t=T_k}^{T_m}\left(Z_t(x_i^{k'})Z_t(x_j^k) - EZ(x_i^{k'})Z(x_j^k)\right)\right|$$

Or,

$$\left(Z_{T_k}(x_i^{k'})Z_{T_k}(x_j^k) - EZ(x_i^{k'})Z(x_j^k)\right), \ldots, \left(Z_{T_m}(x_i^{k'})Z_{T_m}(x_j^k) - EZ(x_i^{k'})Z(x_j^k)\right)$$

sont $(T_m - T_k + 1)$ variables *i.i.d.* et centrées. De plus, d'après l'hypothèse **M3** :

$$\sup_{[0,1]^2} E|Z(x)Z(x')|^\nu \leq \sqrt{\sup_{[0,1]} E^{2\nu}Z(x)\sup_{[0,1]} E^{2\nu}Z(x')} = M_3 < +\infty \tag{3.3}$$

L'inégalité de grande déviation (Brillinger (1962), *cf.* théorème 2.4, chapitre 2) donne donc, pour tout $1 \leq k' \leq k \leq (m-1)$ et tout (i,j) dans $[1,\ldots,\delta_{k'}]\times[1,\ldots,\delta_k]$:

$$\begin{aligned}
&P\left(\left\{\omega : \left|\hat{E}_m Z(x_i^{k'})Z(x_j^k) - EZ(x_i^{k'})Z(x_j^k)\right| \geq \epsilon\right\}\right)\\
&\qquad\qquad \leq \frac{K_\nu\left(E|Z_t(x_i^{k'})Z_t(x_j^k) - EZ(x_i^{k'})Z(x_j^k)|^\nu + 1\right)}{\epsilon^\nu (T_m - T_k + 1)^{\nu/2}}
\end{aligned}$$

où K_ν est une constante positive qui dépend uniquement de ν. On déduit d'après (3.3), pour tout $1 \leq k' \leq k \leq (m-1)$:

$$\sup_{(i,j)\in[1,\dots,\delta_{k'}]\times[1,\dots,\delta_k]} P\left(\left\{\omega : \left|\hat{E}_m Z(x_i^{k'})Z(x_j^k) - EZ(x_i^{k'})Z(x_j^k)\right| \geq \epsilon\right\}\right) \leq \frac{K_\nu\,(2^\nu M_3+1)}{\epsilon^\nu (T_m - T_k + 1)^{\nu/2}}$$

Par conséquent :

$$\begin{aligned}
P(\Phi_m^\epsilon) &= \frac{n_1(n_1+1)}{2} \times O\left(\frac{1}{(T_m - T_1 + 1)^{\nu/2}}\right) \\
&+ \frac{(n_2-n_1)(n_2-n_1+1)}{2} \times O\left(\frac{1}{(T_m - T_2 + 1)^{\nu/2}}\right) \\
&+ n_1(n_2-n_1) \times O\left(\frac{1}{(T_m - T_2 + 1)^{\nu/2}}\right) \\
&\vdots \\
&+ \frac{(n_{m-1}-n_{m-2})(n_{m-1}-n_{m-2}+1)}{2} \times O\left(\frac{1}{(T_m - T_{m-1} + 1)^{\nu/2}}\right) \\
&+ n_{m-2}(n_{m-1}-n_{m-2}) \times O\left(\frac{1}{(T_m - T_{m-1} + 1)^{\nu/2}}\right)
\end{aligned}$$

Soit :

$$P(\Phi_m^\epsilon) = \frac{1}{2}\sum_{k=1}^{m-1}(n_k - n_{k-1})(n_k + n_{k-1} + 1) \times O\left(\frac{1}{(T_m - T_k + 1)^{\nu/2}}\right)$$

Soit encore :

$$P(\Phi_m^\epsilon) = \sum_{k=1}^{m-1}(n_k^2 - n_{k-1}^2) \times O\left(\frac{1}{(T_m - T_k)^{\nu/2}}\right) \tag{3.4}$$

Grâce à l'hypothèse **E2**, on obtient bien (3.2).

2. Pour obtenir la convergence uniforme *p.s.*, il suffit de montrer, d'après le théorème de Borel-Cantelli, que pour tout $\epsilon > 0$:

$$\sum_{m=2}^{+\infty} P(\Phi_m^\epsilon) < +\infty$$

Grâce à la relation (3.4) et l'hypothèse **E2'**, ce résultat est bien vérifié.

■

***Théorème* 3.1**

1. *Sous les hypothèses **E1-E2**, **M3** et **N**:*

$$\sup_{[0,1]^2} |\tilde{r}_m(x,x') - r_Z(x,x')| \underset{m\to+\infty}{\longrightarrow} 0, \text{ en probabilité}$$

2. *Si **E2'** remplace **E2**:*

$$\sup_{[0,1]^2} |\tilde{r}_m(x,x') - r_Z(x,x')| \underset{m\to+\infty}{\longrightarrow} 0, \text{ p.s.}$$

★

Démonstration:

On a:

$$\begin{aligned}\sup_{[0,1]^2} |\tilde{r}_m(x,x') - r_Z(x,x')| &\leq \sup_{[0,1]^2} |\tilde{r}_m(x,x') - r_Z(v_{m-1}(x,x'))| \\ &+ \sup_{[0,1]^2} |r_Z(v_{m-1}(x,x')) - r_Z(x,x')|\end{aligned}$$

(i) Pour la première partie du membre de droite de l'inégalité, par construction:

$$\begin{aligned}\sup_{[0,1]^2} |\tilde{r}_m(x,x') - r_Z(v_{m-1}(x,x'))| &= \sup_{[0,1]^2} |\tilde{r}_m(v_{m-1}(x,x')) - r_Z(v_{m-1}(x,x'))| \\ &= \sup_{(x,x')\in\Delta^2_{m-1}} |\tilde{r}_m(x,x') - r_Z(x,x')|\end{aligned}$$

Or, d'après le point 1. du lemme 3.3:

$$\lim_{m\to+\infty}\left(\sup_{(x,x')\in\Delta^2_{m-1}} |\tilde{r}_m(x,x') - r_Z(x,x')|\right) = 0, \text{ en probabilité}$$

(ii) Pour ce qui est de la deuxième partie, r_Z étant uniformément continue dans $[0,1]^2$:

$$\forall\epsilon>0, \exists\delta>0, \|(x,x')-(y,y')\|_2<\delta \Longrightarrow |r_Z(x,x') - r_Z(y,y')| < \epsilon$$

Or, d'après le lemme 3.1:

$$\exists m_0(\delta), \forall m > m_0(\delta), \sup_{[0,1]^2} \|(x,x') - v_{m-1}(x,x')\|_2 < \delta$$

On en déduit alors :

$$\forall \epsilon > 0, \exists m_0, \forall m > m_0, \sup_{[0,1]^2} |r_Z(x,x') - r_Z(v_m(x,x'))| < \epsilon$$

■

Remarque : le cas où Z est borné

Sous les conditions **E1** sur la densification des stations dans $[0,1]$, **M3** sur l'existence d'un moment d'ordre 2ν, $\nu > 2$, de Z et **N** sur la continuité en moyenne quadratique de Z, l'hypothèse **E2** (respectivement **E2'**) nous a permis d'obtenir la convergence uniforme en probabilité (respectivement *p.s.*) sur $[0,1]^2$ de $\tilde{r}_m$ vers r_Z. Comme nous l'avons vu, les hypothèses **E2** et **E2'** sont très contraignantes et limitent de fait de nombreux plans d'échantillonnage. Nous allons voir que si on renforce **M3** par :

- **M3'** : $\sup_{x\in[0,1]} |Z(x)| \leq M < +\infty$, *p.s.*

c'est-à-dire que si on suppose que Z est borné uniformément sur $[0,1]$ *p.s.*, les convergences uniformes en probabilité et *p.s.* sont encore vraies avec des conditions plus souples sur les nombres des stations n_k et leurs instants d'apparition T_k, $k \geq 1$. On utilise pour cela une inégalité de grande déviation de type exponentiel :

***Théorème* 3.2** *(Hœffding (1963))*
Soient $X_1, X_2, \ldots, X_n$ un échantillon i.i.d. de la variable aléatoire parente X centrée. Si $|X| \leq C < +\infty$, p.s., alors pour tout $\epsilon > 0$:

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i\right| \geq \epsilon\right) \leq 2\exp\left(-\frac{n\epsilon^2}{2C^2}\right)$$

★

En utilisant cette inégalité dans la démonstration du lemme 3.3, au lieu de l'inégalité de Brillinger, on aboutit à :

$$P(\Phi_m^\epsilon) \leq 2\sum_{k=1}^{m-1}(n_k^2 - n_{k-1}^2) \times \exp\left(-\frac{(T_m - T_k)\epsilon^2}{2M'}\right)$$

au lieu de (3.4) ($M' > 0$).

Si on remplace **E2** et **E2'** par :

- **E3** : $\begin{cases} \exists \nu_1 > 1, \forall m \geq 1, T_m = m^{\nu_1}, \text{ avec } \nu_1 \text{ aussi proche de 1 que l'on souhaite} \\ \exists \nu_2 > 1, \forall m \geq 1, n_m = m^{\nu_2}, \text{ avec } \nu_2 \text{ aussi grand que l'on souhaite} \end{cases}$

alors on obtient en même temps :

$$\begin{cases} \lim\limits_{m \to +\infty} P(\Phi_m^\epsilon) = 0 \\ \sum\limits_{m=2}^{+\infty} P(\Phi_m^\epsilon) < +\infty \end{cases}$$

et on en déduit donc simultanément les convergences uniformes en probabilité et *p.s.* dans $[0,1]^2$ de $\tilde{r}_m$ vers r_Z.

♣

3.3 Estimation des paramètres du modèle

Dans le paragraphe 3.3.1, nous définissons le processus de contraste, dont la minimisation fournit les estimations des paramètres f_0 et β_0. Le théorème 3.3 est un résultat de convergences en P_0-probabilité et $P_0 - p.s.$ du contraste vers la fonction de contraste, pour laquelle l'identifiabilité du modèle sous **H0** est assurée. Dans le paragraphe 3.3.2, on donne deux résultats, l'un sur l'uniforme continuité $P_0 - p.s.$ du contraste, l'autre sur l'uniforme continuité de la fonction de contraste. Dans le paragraphe 3.3.3, on construit un estimateur mesurable du minimum de contraste dans un sens que l'on précisera et que l'on note $\tilde{\theta}_m = \left(\tilde{f}_m, \tilde{\beta}_m\right)$.

3.3.1 Convergences du contraste et identifiabilité du modèle

Présentation du contraste

Comme au chapitre 2, on utilise une méthode d'estimation par minimum de contraste, mais ici la nature de l'espace des paramètres est différente. On définit ce contraste comme la moyenne des écarts au carré entre les covariances empiriques $\tilde{r}_m(x,x')$ et le modèle de covariance $r_\theta(x,x')$ calculée sur les (x,x') appartenant à Δ_{m-1}^2 :

$$\begin{aligned} U_m(\theta) &= \frac{1}{n_{m-1}^2} \sum_{(x,x') \in \Delta_{m-1}^2} \left[\tilde{r}_m(x,x') - r_\theta(x,x')\right]^2 \\ &= \int_{[0,1]^2} \left[\tilde{r}_m(x,x') - r_\theta(x,x')\right]^2 d\mu_m^{\otimes 2}(x,x') \end{aligned}$$

en rappelant que $\mu_m = \dfrac{1}{n_{m-1}} \sum\limits_{x \in \Delta_{m-1}} \delta_x$.

On rappelle aussi que pour tout (x,x') dans Δ_{m-1}^2 et $\theta = (f,\beta)$:

$$\begin{cases} \tilde{r}_m(x,x') &= \dfrac{1}{T_m - T_x \vee T_{x'} + 1} \sum\limits_{t = T_x \vee T_{x'}}^{T_m} Z_t(x) Z_t(x') \\ r_\theta(x,x') &= R\left(f(x') - f(x); \beta\right) \end{cases}$$

On rappelle aussi que sous l'hypothèse **H0**, pour tout (x,x') dans $[0,1]^2$:

$$r_Z(x,x) = r_{\theta_0}(x,x')$$

Considérons de plus les deux hypothèses :

- **J1** : $\beta \in \Psi$ compact de $\mathbb{R}^q$, $q \in \mathbb{N}^*$, et $\beta_0 \in \overset{\circ}{\Psi}$
- **J2** : $R : (u,\beta) \longmapsto R(u;\beta)$ est continue dans $[-1,1] \times \Psi$

Donc, sous **J1**-**J2**, l'application R est uniformément continue dans $[-1,1] \times \Psi$. On note $\Theta = \mathcal{F}_K \times \Psi$ l'ensemble des paramètres du modèle. On a le lemme suivant :

***Lemme* 3.4** *Sous l'hypothèse **E1** :*

$$\mu_m^{\otimes 2} \underset{m \to +\infty}{\overset{c.e.}{\longrightarrow}} \mu^{\otimes 2}$$

★

Démonstration :

On utilise le théorème de Paul Lévy (Monfort (1980), p.140). Par hypothèse, la suite (μ_m) est une suite de mesures positives finies qui converge étroitement vers la mesure positive finie μ. On en déduit alors que la suite des fonctions caractéristiques (ϕ_{μ_m}) converge simplement vers ϕ_μ dans $\mathbb{R}$. On a :

$$\phi_{\mu_m^{\otimes 2}}(y,y') = \phi_{\mu_m}(y) \times \phi_{\mu_m}(y')$$

Par conséquent, pour tout (y,y') dans $\mathbb{R}^2$:

$$\phi_{\mu_m^{\otimes 2}}(y,y') \underset{m \to +\infty}{\longrightarrow} \phi_\mu(y) \times \phi_\mu(y') = \phi_{\mu^{\otimes 2}}(y,y')$$

En tant que fonction caractéristique, ϕ_μ est continue dans $\mathbb{R}$ (Métivier (1979), p.152), et donc $\phi_{\mu^{\otimes 2}}$ l'est également. On en déduit que $\phi_{\mu^{\otimes 2}}$ est la fonction caractéristique d'une mesure finie vers laquelle converge la mesure empirique $\mu_m^{\otimes 2}$. Cette mesure limite n'est autre que $\mu^{\otimes 2}$.

■

Convergence en probabilité du contraste

Avant d'énoncer le théorème 3.3 sur les convergences en P_0-probabilité et $P_0 - p.s.$ du contraste vers une quantité positive, la fonction de contraste, on a :

***Corollaire* 3.1**

1. *Sous les hypothèses **E1-E2**, **M3** et **N** :*

$$\mu_m^{\otimes^2}(\tilde{r}_m) = \int_{[0,1]^2} \tilde{r}_m(x,x')d\mu_m^{\otimes^2}(x,x')$$

$$\underset{m\to+\infty}{\longrightarrow} \int_{[0,1]^2} r_Z(x,x')d\mu^{\otimes^2}(x,x') = \mu^{\otimes^2}(r_Z), \text{ en probabilité}$$

2. *Si **E2'** remplace **E2** :*

$$\mu_m^{\otimes^2}(\tilde{r}_m) \underset{m\to+\infty}{\longrightarrow} \mu^{\otimes^2}(r_Z), \ p.s.$$

★

Démonstration :

1. On a :

$$\mu_m^{\otimes^2}(\tilde{r}_m) - \mu^{\otimes^2}(r_Z) = \mu_m^{\otimes^2}(\tilde{r}_m - r_Z) + (\mu_m^{\otimes^2} - \mu^{\otimes^2})(r_Z)$$

(i) D'un côté, et d'après l'hypothèse **N**, on sait que r_Z est une fonction continue et bornée dans $[0,1]^2$. On en déduit, par définition de la convergence étroite :

$$(\mu_m^{\otimes^2} - \mu^{\otimes^2})(r_Z) \underset{m\to+\infty}{\longrightarrow} 0$$

(ii) De l'autre :

$$\begin{aligned}
\left|\mu_m^{\otimes^2}(\tilde{r}_m - r_Z)\right| &\leq \mu_m^{\otimes^2}(|\tilde{r}_m - r_Z|) \\
&\leq \sup_{[0,1]^2} |\tilde{r}_m(x,x') - r_Z(x,x')| \, \mu_m^{\otimes^2}\left([0,1]^2\right) \\
&\leq \sup_{[0,1]^2} |\tilde{r}_m(x,x') - r_Z(x,x')|
\end{aligned}$$

Or, d'après le point 1. du théorème 3.1 :

$$\lim_{m\to+\infty} \left(\sup_{[0,1]^2} |\tilde{r}_m(x,x') - r_Z(x,x')| \right) = 0, \text{ en probabilité}$$

On en déduit :

$$\lim_{m\to+\infty} \left|\mu_m^{\otimes^2}(\tilde{r}_m - r_Z)\right| = 0, \text{ en probabilité}$$

2. Pour la convergence *p.s.*, la démonstration se fait de la même façon, mais on utilise le point 2. du théorème 3.1.

$\blacksquare$

Théorème 3.3

1. *Sous les hypothèses **E1-E2**, **J1-J2** et **M3**, pour tout θ dans Θ :*

$$U_m(\theta) \underset{m\to+\infty}{\longrightarrow} \int_{[0,1]^2} \left[r_{\theta_0}(x,x') - r_\theta(x,x')\right]^2 d\mu^{\otimes^2}(x,x'), \text{ en } P_0 - \textit{probabilité}$$

2. *Si **E2'** remplace **E2**, pour tout θ dans Θ :*

$$U_m(\theta) \underset{m\to+\infty}{\longrightarrow} \int_{[0,1]^2} \left[r_{\theta_0}(x,x') - r_\theta(x,x')\right]^2 d\mu^{\otimes^2}(x,x'), P_0 - p.s.$$

$\bigstar$

Démonstration :

1.

$$\begin{aligned}
& U_m(\theta) - \int_{[0,1]^2} \left[r_{\theta_0}(x,x') - r_\theta(x,x')\right]^2 d\mu^{\otimes^2}(x,x') \\
= & \int_{[0,1]^2} \tilde{r}_m^2(x,x') d\mu_m^{\otimes^2}(x,x') - \int_{[0,1]^2} r_{\theta_0}^2(x,x')^2 d\mu^{\otimes^2}(x,x') && (3.5)\\
- & 2\int_{[0,1]^2} r_\theta(x,x')\tilde{r}_m(x,x') d\mu_m^{\otimes^2}(x,x') + 2\int_{[0,1]^2} r_\theta(x,x') r_{\theta_0}(x,x') d\mu^{\otimes^2}(x,x') && (3.6)\\
+ & \int_{[0,1]^2} r_\theta^2(x,x') d\mu_m^{\otimes^2}(x,x') - \int_{[0,1]^2} r_\theta^2(x,x') d\mu^{\otimes^2}(x,x') && (3.7)
\end{aligned}$$

De la convergence étroite de μ_m vers μ, on déduit la convergence en P_0-probabilité de chacune des quantités (3.5), (3.6) et (3.7) vers 0, grâce au point 1. du corollaire 3.1 précédent, à des modifications mineures près.

2. Pour la convergence $P_0-p.s.$, la démonstration se fait de la même façon, mais on utilise le point 2. du corollaire 3.1.

$\blacksquare$

Posons :

$$\begin{aligned}
K : \bar{\Theta} & \longrightarrow \mathbb{R}^+ \\
\theta & \longmapsto K(\theta_0,\theta) = \int_{[0,1]^2} \left[r_{\theta_0}(x,x') - r_\theta(x,x')\right]^2 d\mu^{\otimes^2}(x,x')
\end{aligned}$$

K est la fonction de contraste.

Identifiabilité du modèle

Faisons tout d'abord les hypothèses :

- **Q1** : la famille des modèles de covariance stationnaire $\mathcal{R}$ est identifiable, c'est-à-dire :

$$\beta = \beta' \Longleftrightarrow \|R(.;\beta) - R(.;\beta')\|_\infty = 0$$

- **Q2** : $\forall \beta \in \Psi, u \longmapsto R(u;\beta)$ est différentiable dans $[-1,1] - \{0\}$

On note que ces hypothèses sont vérifiées pour les familles usuelles de modèles de covariance stationnaire, par exemple :

- la famille des modèles sphériques (*cf.* figure 4.8, chapitre 4) :

$$R(u;\sigma_0,a_0) = \begin{cases} \sigma_0^2 \left(1 - a_0|u| + \dfrac{4a_0^3|u|^3}{27}\right) & \text{si } 0 \leq |u| \leq \dfrac{3}{2a_0} \\ 0 & \text{sinon} \end{cases}$$

- la famille des modèles exponentiels (*cf.* figure 2.3, chapitre 2) :

$$R(u;\sigma_0,a_0) = \sigma_0^2 \exp\left(-a_0|u|\right)$$

- la famille des modèles à covariance gaussienne (*cf.* figure 3.4) :

$$R(u;\sigma_0,a_0) = \sigma_0^2 \exp\left(-a_0 u^2\right)$$

Dans les trois exemples, $\beta_0 = (\sigma_0, a_0)$ avec $\sigma_0 > 0$ et $a_0 > 0$.

Par définition de $\mathcal{F}_K$ et d'après l'hypothèse **J2**, r_θ est continue dans $[0,1]^2$, pour tout θ dans Θ. Par conséquent :

$$K(\theta_0,\theta) = 0 \Longleftrightarrow \|r_{\theta_0} - r_\theta\|_\infty = 0$$

Sous les hypothèses **J2** et **Q1**-**Q2**, et d'après le corollaire 1.1 (*cf.* chapitre 1) sur l'unicité du couple $(f_0, R(.;\beta_0))$:

$$\|r_{\theta_0} - r_\theta\|_\infty = 0 \Longleftrightarrow \begin{cases} \|f - f_0\|_\infty = 0 \\ \|R(.;\beta) - R(.\beta_0)\|_\infty = 0 \end{cases} \Longleftrightarrow \begin{cases} \|f - f_0\|_\infty = 0 \\ \|\beta - \beta_0\|_q = 0 \end{cases} \Longleftrightarrow \theta = \theta_0$$

L'identifiabilité paramétrique du modèle sous **H0** est donc bien assurée.

3.3.2 Uniforme continuité du contraste U_m et de la fonction de contraste K

On rappelle que Θ est l'ensemble produit $\mathcal{F}_K \times \Psi \subset C([0,1]) \times \mathbb{R}^q$. Soit d la distance sur $C([0,1]) \times \mathbb{R}^q$ définie par :

$$d((f,\beta),(f',\beta')) = \|f'-f\|_\infty + \|\beta'-\beta\|_q$$

$\blacksquare$

Uniforme continuité de U_m

Comme $\bar{\mathcal{F}_K}$ et Ψ sont compacts, $\bar{\Theta} = \bar{\mathcal{F}_K} \times \Psi$ est aussi compact.

Lemme 3.5 *Sous les hypothèses **J1-J2**, à m fixé, $m \geq 2$, la fonctionnelle :*

$$\begin{aligned} U_m : \bar{\Theta} &\longrightarrow \mathbb{R}^+ \\ \theta &\longmapsto U_m(\theta) \end{aligned}$$

est uniformément continue $P_0 - p.s.$ dans $\bar{\Theta}$.

$\bigstar$

Démonstration :

Pour tout $m \geq 2$:

$$U_m = F_{2,m} \circ F_1$$

où F_1, déterministe, et $F_{2,m}$, aléatoire, sont définis comme suit :

$$\begin{array}{cccc} \bar{\Theta} & \xrightarrow{F_1} & \{\text{fonctions de covariance}\} & \xrightarrow{F_{2,m}} & \mathbb{R}^+ \\ \theta & \longmapsto & r_\theta & \longmapsto & U_m(\theta) = F_{2,m} \circ F_1(\theta) \end{array}$$

(i) Montrons d'abord que l'opérateur F_1 est uniformément continue dans $\bar{\Theta}$. D'après les hypothèses **J1-J2**, l'application :

$$R : (u,\beta) \longmapsto R(u;\beta)$$

est uniformément continue dans $[-1,1] \times \Psi$, c'est-à-dire :

$$\forall \epsilon > 0, \exists (\eta,\eta') \in \mathbb{R}^{*2}, |u'-u| < \eta \text{ et } \|\beta'-\beta\|_q < \eta' \Longrightarrow |R(u';\beta') - R(u;\beta)| < \epsilon$$

Par conséquent :

$$\begin{aligned} \forall \epsilon > 0, \exists (\eta,\eta') \in \mathbb{R}^{*2}, |((f'(x')-f'(x)) - (f(x')-f(x))| < \eta \text{ et } \|\beta'-\beta\|_q < \eta' \\ \Longrightarrow |R(f'(x')-f'(x);\beta') - R(f(x')-f(x);\beta)| < \epsilon \end{aligned}$$

Or, pour tout (f, f') dans $\bar{\mathcal{F}_K}^2$:

$$\sup_{[0,1]^2} |(f'(x) - f'(x')) - (f(x) - f(x'))| \leq 2\|f' - f\|_\infty$$

et, pour tout (x, x') dans $[0,1]^2$:

$$r_\theta(x, x') = R(f(x') - f(x); \beta)$$

On a finalement :

$$\forall \epsilon > 0, \exists \eta'' = \frac{\eta}{2} + \eta', d(\theta, \theta') < \eta'' \Longrightarrow \sup_{[0,1]^2} |r_{\theta'}(x, x') - r_\theta(x, x')| < \epsilon$$

D'où l'uniforme continuité de F_1 dans $\bar{\Theta}$.

(ii) De plus, par construction, la fonctionnelle $F_{2,m}$ est continue $P_0 - p.s.$

On en conclut que $U_m = F_{2,m} \circ F_1$ est uniformément continue $P_0 - p.s.$ dans $\bar{\Theta}$, pour tout $m \geq 2$. En particulier, $U_m(\bar{\Theta})$ est un compact $P_0 - p.s.$ dans $\mathbb{R}^+$.

■

Uniforme continuité de K

***Lemme* 3.6** *Sous les hypothèses **J1-J2** la fonctionnelle K est uniformément continue dans $\bar{\Theta}$.*

★

Démonstration :

Elle est identique à celle du lemme 3.5, mais au lieu de considérer la fonctionnelle $F_{2,m}$, on fait intervenir :

$$\begin{array}{rcl} F_{2,\infty} : \{\text{fonctions de covariance}\} & \longrightarrow & \mathbb{R}^+ \\ r_\theta & \longmapsto & K(\theta_0, \theta) \end{array}$$

■

3.3.3 Existence de $\tilde{\theta}_m = (\tilde{f}_m, \tilde{\beta}_m)$

Θ étant un espace relativement compact, il n'existe pas *a priori* de solution au problème de minimisation de U_m dans Θ. C'est pourquoi nous chercherons une suite $\left(\tilde{\theta}_m\right)$ qui minimisera U_m à une constante strictement positive ϵ_m près, c'est-à-dire une suite telle que pour tout $\theta \in \Theta$:

$$U_m(\tilde{\theta}_m) \leq U_m(\theta) + \epsilon_m$$

où il est naturel de choisir la suite (ϵ_m) tendant vers 0, $m \to +\infty$.

Cette notion de ϵ_m-estimateur est classique dans les problèmes d'estimation dans un cadre non-paramétrique. On citera par exemple l'article de Wong et Severini ((1991), pp.603-606) pour les $\epsilon-$estimateurs du maximum de vraisemblance et l'article de Birgé et Massart ((1993), p.117) pour les $\epsilon-$estimateurs du minimum de contraste.

***Théorème* 3.4** *Sous les hypothèses **J1-J2**, pour toute suite (ϵ_m) strictement positive et tendant vers 0, $m \to +\infty$, il existe une suite $\left(\tilde{\theta}_m\right)$ de ϵ_m-estimateurs mesurables vérifiant pour tout θ dans Θ et tout $m \geq 2$:*

$$U_m(\tilde{\theta}_m) \leq U_m(\theta) + \epsilon_m$$

★

Démonstration :

Elle va consister à construire explicitement, pour la déformation f_0, un ϵ_m-estimateur mesurable $\tilde{f}_m$ du minimum de contraste, sous la forme d'une ligne brisée ou combinaison convexe de I-splines de degré 1 (*cf.* paragraphe 2.3.3 et annexe C) associée aux n_{m-1} nœuds intérieurs à $[0, 1]$:

$$0 < x_1 < x_2 < \cdots < x_{n_{m-1}} < 1$$

chaque I-spline de la combinaison étant de pente inférieure où égale à K.

Plus précisément, posons $x_0 = y_0 = 0$, $x_{n_{m-1}+1} = y_{n_{m-1}+1} = 1$, et considérons la suite d'ensembles de points :

$$\mathcal{Y}_{K,m} = \left\{(y_1, y_2, \ldots, y_{n_{m-1}}); 0 < y_1 < y_2 < \cdots < y_{n_{m-1}} < 1, \frac{y_i - y_{i-1}}{x_i - x_{i-1}} \leq K, i \in [1, \ldots, n_{m-1} + 1]\right\}$$

et définissons l'ensemble $\mathcal{F}_{K,m}$, associé à $\mathcal{Y}_{K,m}$, des lignes brisées $f_{K,m}$ de pentes inférieures où égales à K, associées aux nœuds $x_1, \ldots, x_{n_{m-1}}$:

$$\left\{\begin{array}{lcl} f_{K,m}(x) & = & \displaystyle\sum_{i=1}^{n_{m-1}+1} \left(y_{i-1} + \frac{y_i - y_{i-1}}{x_i - x_{i-1}}(x - x_{i-1})\right) \mathbb{1}_{[x_{i-1}, x_i[}(x) \\ f_{K,m}(1) & = & 1 \end{array}\right.$$

avec $\left(y_1, y_2, \ldots, y_{n_{m-1}}\right) \in \mathcal{Y}_{K,m}$.

$\mathcal{F}_{K,m}$ est un sous-ensemble de $\mathcal{F}_K$, vérifiant sous l'hypothèse **E1** :

$$\bigcup_{m=2}^{+\infty} \mathcal{F}_{K,m} \text{ est dense dans } \mathcal{F}_K$$

On définit ensuite un nouveau contraste sur l'espace $\mathcal{Y}_{K,m} \times \Psi$, par :

$$V_m\left(y_1, y_2, \ldots, y_{n_{m-1}}, \beta\right) = \frac{1}{n_{m-1}^2} \sum_{(i,j)\in[1,\ldots,n_{m-1}]^2} \left[\tilde{r}_m(x_i, x_j) - R(y_j - y_i; \beta)\right]^2$$

$\mathcal{Y}_{K,m}$ n'est pas un compact dans $\mathbb{R}^{n_{m-1}}$. Cependant, $\bar{\mathcal{Y}}_{K,m} \times \Psi$ est compact dans $\mathbb{R}^{n_{m-1}+q}$. De plus, grâce à **J1**-**J2**, V_m est uniformément continu $P_0 - p.s.$ dans $\bar{\mathcal{Y}}_{K,m} \times \Psi$. On est donc assuré de l'existence, dans $\mathcal{Y}_{K,m} \times \Psi$, d'un ϵ_m−estimateur mesurable du minimum de contraste V_m :

$$\left(\tilde{y}_{1,m}, \tilde{y}_{2,m}, \ldots, \tilde{y}_{n_{m-1},m}, \tilde{\beta}_m\right)$$

Donc, pour toute suite (ϵ_m) strictement positive et tendant vers 0, $m \to +\infty$, pour tout $\left(y_1, \ldots, y_{n_{m-1}}, \beta\right)$ dans $\mathcal{Y}_{K,m} \times \Psi$ et tout $m \geq 2$:

$$V_m\left(\tilde{y}_{1,m}, \tilde{y}_{2,m}, \ldots, \tilde{y}_{n_{m-1},m}, \tilde{\beta}_m\right) \leq V_m\left(y_1, \ldots, y_{n_{m-1}}, \beta\right) + \epsilon_m$$

De plus, par définition de $\mathcal{F}_K$, on vérifie que, pour tout f dans $\mathcal{F}_K$ et tout $m \geq 2$:

$$\left(f(x_1), f(x_2), \ldots, f(x_{n_{m-1}})\right) \in \mathcal{Y}_{K,m}$$

En effet, on a supposé que :

- *(i)* $0 < x_1 < x_2 < \cdots < x_{n_{m-1}} < 1$. Comme les fonctions de $\mathcal{F}_K$ sont des homéomorphismes croissants de $[0,1]$ dans $[0,1]$, on a pour tout f dans $\mathcal{F}_K$:

$$0 = f(0) < f(x_1) < f(x_2) < \cdots < f(x_{n_{m-1}}) < f(1) = 1$$

- *(ii)* les fonctions f de $\mathcal{F}_K$ sont lipschitziennes de rapport K sur $[0,1]$, et donc, pour tout i dans $[1, \ldots, n_{m-1}+1]$, $\dfrac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} \leq K$.

Par conséquent, on a bien, pour tout $m \geq 2$:

$$\left\{(f(x_1), f(x_2), \ldots, f(x_{n_{m-1}})); f \in \mathcal{F}_K\right\} \subset \mathcal{Y}_{K,m}$$

D'autre part, il est facile de voir que pour tout $\theta = (f, \beta)$ dans $\mathcal{F}_K \times \Psi$ et tout $m \geq 2$:

$$U_m(\theta) = U_m(f, \beta) = V_m\left(f(x_1), f(x_2), \ldots, f(x_{n_{m-1}}), \beta\right)$$

On en déduit que, pour tout $\theta = (f, \beta)$ dans $\mathcal{F}_K \times \Psi$ et tout $m \geq 2$:

$$V_m\left(\tilde{y}_{1,m}, \tilde{y}_{2,m}, \ldots, \tilde{y}_{n_{m-1},m}, \tilde{\beta}_m\right) \leq U_m(f, \beta) + \epsilon_m \tag{3.8}$$

Enfin, posant $\tilde{y}_{0,m} = 0$ et $\tilde{y}_{n_{m-1}+1,m} = 1$, la fonction $\tilde{f}_m$ ainsi définie (*cf.* figure 3.3) :

$$\begin{cases} \tilde{f}_m(x) & = & \displaystyle\sum_{i=1}^{n_{m-1}+1} \left(\tilde{y}_{i-1,m} + \frac{\tilde{y}_{i,m} - \tilde{y}_{i-1,m}}{x_i - x_{i-1}} (x - x_{i-1}) \right) \mathbb{1}_{[x_{i-1}, x_i[}(x) \\ \tilde{f}_m(1) & = & 1 \end{cases}$$

est continue, affine par morceaux dans $[0,1]$ et vérifie $\tilde{f}_m(x_i) = \tilde{y}_{i,m}$, $i \in [0, \ldots, n_{m-1}+1]$. Par définition, $\tilde{f}_m$ est dans $\mathcal{F}_{K,m}$ et donc dans $\mathcal{F}_K$.

En termes de combinaison convexe de I-splines de degré 1, associée aux nœuds $x_1, x_2, \ldots, x_{n_{m-1}}$, $\tilde{f}_m$ peut encore s'écrire, pour tout x dans $[0,1]$:

$$\tilde{f}_m(x) = \sum_{i=1}^{n_{m-1}+1} \tilde{\alpha}_i I_i^1(x)$$

avec :

$$\begin{cases} \tilde{\alpha}_1 = \tilde{y}_{1,m} - \tilde{y}_{0,m} \\ \tilde{\alpha}_2 = \tilde{y}_{2,m} - \tilde{y}_{1,m} \\ \vdots \\ \tilde{\alpha}_{n_{m-1}+1} = \tilde{y}_{n_{m-1}+1,m} - \tilde{y}_{n_{m-1},m} \end{cases}$$

Puisque :

$$U_m(\tilde{f}_m, \tilde{\beta}_m) = V_m \left(\tilde{y}_{1,m}, \tilde{y}_{2,m}, \ldots, \tilde{y}_{n_{m-1},m}, \tilde{\beta}_m \right) \tag{3.9}$$

on conclut, grâce aux relations (3.8) et (3.9), à ce que $\tilde{\theta}_m = (\tilde{f}_m, \tilde{\beta}_m)$, ainsi défini, constitue un $(\epsilon_m)-$estimateur mesurable du minimum de contraste $U_m(\theta)$.

■

3.4 Convergence en probabilité et $p.s.$ des ϵ_m-estimateurs

***Théorème* 3.5**

1. *Sous les hypothèses* ***E1-E2***, ***J1-J2***, ***M3*** *et* ***Q1-Q2***, *les ϵ_m-estimateurs du minimum de contraste, $\tilde{f}_m$ et $\tilde{\beta}_m$, convergent en P_0-probabilité vers f_0 et β_0, c'est-à-dire :*

$$d\left(\tilde{\theta}_m, \theta_0\right) \underset{m \to +\infty}{\longrightarrow} 0, \text{ en } P_0 - \textit{probabilité}$$

2. *Si* ***E2'*** *remplace* ***E2*** *:*

$$d\left(\tilde{\theta}_m, \theta_0\right) \underset{m \to +\infty}{\longrightarrow} 0, \ P_0 - p.s.$$

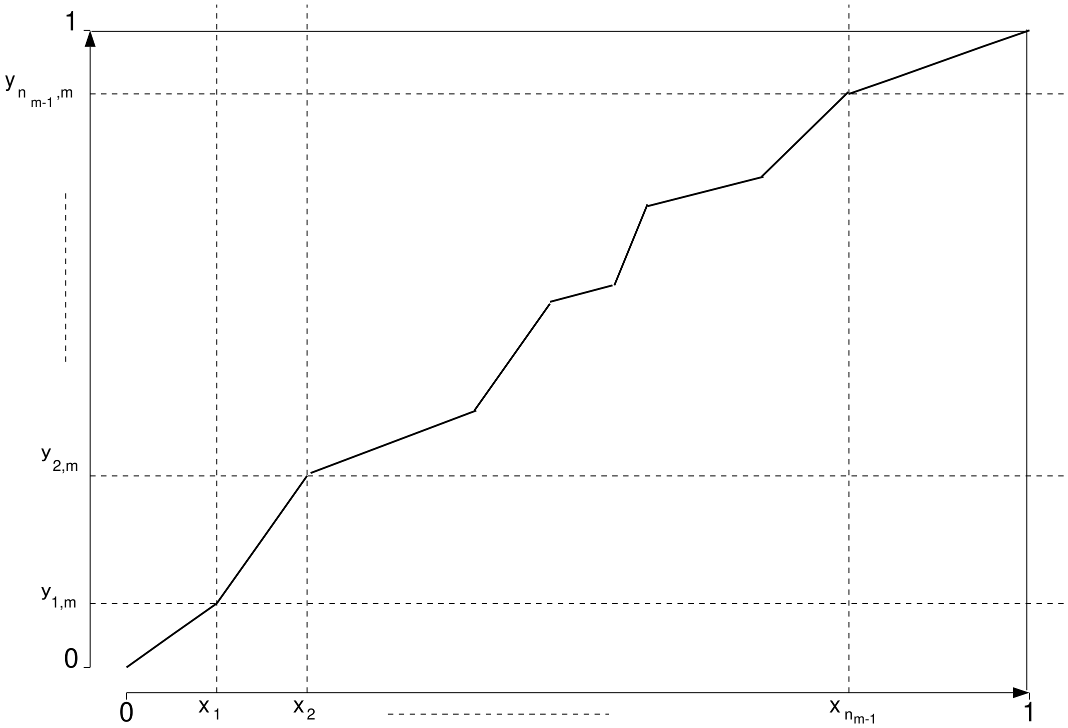


FIG. 3.3 – *Forme générale de* $\tilde{f}_m$

★

Démonstration :

1. Nous appliquons le théorème D.2 (*cf.* annexe D) qui est une généralisation du résultat de Hardouin ((1992), pp.6-8), au cas où l'espace des paramètres Θ n'est plus compact. Les conditions d'application sont similaires à celles de Hardouin (la seule différence concerne l'hypothèse *(i)* ci-dessous) :

   - *(i)* Θ est relativement compact (dans Hardouin, Θ est compact)
   - *(ii)* $\theta \longmapsto U_m(\theta)$ est continue $P_0 - p.s.$ et $\theta \longmapsto K(\theta_0, \theta)$ est continue, dans $\bar{\Theta}$
   - *(iii)* $K(\theta_0, \theta) = 0 \Longleftrightarrow \theta = \theta_0$
   - *(iv)* $\liminf_{m \to +\infty} U_m(\theta) \geq K(\theta_0, \theta)$, en $P_0 -$ probabilité
   - *(v)* Notant W_m le module de continuité de U_m :

$$W_m(\eta) = \sup \left\{ |U_m(\theta) - U_m(\theta')|, d(\theta, \theta') \leq \eta, (\theta, \theta') \in \bar{\Theta}^2 \right\}$$

il existe une suite (η_k) qui tend vers zéro, $k \to +\infty$, telle que, pour tout $k \geq 1$:

$$\lim_{m \to +\infty} P_0 \left(W_m \left(\frac{1}{k} \right) + \epsilon_m \geq \eta_k \right) = 0$$

L'hypothèse *(i)* est vérifiée par définition de l'ensemble $\mathcal{F}_K$ et grâce à l'hypothèse **J1**. L'hypothèse *(ii)* est vérifiée grâce aux lemmes 3.5 et 3.6. L'identifiabilité paramétrique du modèle sous **H0**, c'est-à-dire l'hypothèse *(iii)*, est assurée (*cf.* fin du paragraphe 3.3.1). Grâce au point 1. du théorème 3.3, l'hypothèse *(iv)* est aussi vérifiée.

Reste à montrer que *(v)* est également vérifiée sous nos conditions. Une décomposition de $(U_m(\theta') - U_m(\theta))$ fournit l'inégalité :

$$\begin{aligned} |U_m(\theta') - U_m(\theta)| &\leq \left| \int_{[0,1]^2} (r_{\theta'}(x,x') + r_\theta(x,x')) \times (r_{\theta'}(x,x') - r_\theta(x,x'))\, d\mu_m^{\otimes^2}(x,x') \right| \\ &+ 2 \left| \int_{[0,1]^2} (\tilde{r}_m(x,x') - r_{\theta_0}(x,x')) \times (r_{\theta'}(x,x') - r_\theta(x,x'))\, d\mu_m^{\otimes^2}(x,x') \right| \\ &+ 2 \left| \int_{[0,1]^2} r_{\theta_0}(x,x') (r_{\theta'}(x,x') - r_\theta(x,x'))\, d\mu_m^{\otimes^2}(x,x') \right| \end{aligned}$$

Puisque $|r_\theta(x,x')| \leq \sigma^2$ et que σ^2 appartient à un compact de $\mathbb{R}$, on a, pour tout (x,x') dans $[0,1]^2$:

$$\sup_{\theta \in \Theta} |r_\theta(x,x')| \leq N < +\infty$$

On obtient alors :

$$\begin{aligned} |U_m(\theta') - U_m(\theta)| &\leq 4N \sup_{[0,1]^2} |r_{\theta'}(x,x') - r_\theta(x,x')| \\ &+ 4N \int_{[0,1]^2} |\tilde{r}_m(x,x') - r_{\theta_0}(x,x')|\, d\mu_m^{\otimes^2}(x,x') \end{aligned}$$

Par conséquent, pour tout k dans $\mathbb{N}^*$:

$$\begin{aligned} W_m(\frac{1}{k}) &\leq 4N \sup_{d(\theta,\theta') \leq \frac{1}{k}} \sup_{[0,1]^2} |r_{\theta'}(x,x') - r_\theta(x,x')| \\ &+ 4N \int_{[0,1]^2} |\tilde{r}_m(x,x') - r_{\theta_0}(x,x')|\, d\mu_m^{\otimes^2}(x,x') \end{aligned}$$

Posons, pour tout k dans $\mathbb{N}^*$:

$$u_k = 4N \sup_{d(\theta,\theta') \leq \frac{1}{k}} \sup_{[0,1]^2} |r_{\theta'}(x,x') - r_\theta(x,x')|$$

et considérons l'ensemble, pour tout k dans $\mathbb{N}^*$:

$$B_{k,m} = \left\{ \epsilon_m + u_k + 4N \int_{[0,1]^2} |\tilde{r}_m(x,x') - r_{\theta_0}(x,x')| \, d\mu_m^{\otimes^2}(x,x') \geq \eta_k \right\}$$

On a alors pour tout k :

$$0 \leq P_0 \left(W_m \left(\frac{1}{k} \right) + \epsilon_m \geq \eta_k \right) \leq P_0 \left(B_{k,m} \right)$$

Prenons désormais :

$$\eta_k = u_k + \frac{1}{k}$$

On a vu au cours de la démonstration du lemme 3.5 que :

$$\forall \epsilon > 0, \exists \eta'' > 0, d(\theta, \theta') < \eta'' \Longrightarrow \sup_{[0,1]^2} |r_{\theta'}(x,x') - r_\theta(x,x')| < \epsilon$$

En conséquence :

$$\lim_{k \to +\infty} \eta_k = 0$$

Pour vérifier *(v)*, il reste à montrer que, pour tout k :

$$\lim_{m \to +\infty} P_0 \left(B_{k,m} \right) = 0$$

soit encore :

$$\lim_{m \to +\infty} P_0 \left(\epsilon_m + 4N \int_{[0,1]^2} |\tilde{r}_m(x,x') - r_{\theta_0}(x,x')| \, d\mu_m^{\otimes^2}(x,x') \geq \frac{1}{k} \right) = 0$$

Par définition :

$$\lim_{m \to +\infty} \epsilon_m = 0$$

et l'inégalité de Cauchy-Schwarz nous donne :

$$\int_{[0,1]^2} |\tilde{r}_m(x,x') - r_{\theta_0}(x,x')| \, d\mu_m^{\otimes^2}(x,x') \leq \sqrt{\int_{[0,1]^2} \left[\tilde{r}_m(x,x') - r_{\theta_0}(x,x') \right]^2 d\mu_m^{\otimes^2}(x,x')}$$

Or, d'après le point 1. du théorème 3.3 :

$$\lim_{m \to +\infty} \int_{[0,1]^2} \left[\tilde{r}_m(x,x') - r_{\theta_0}(x,x') \right]^2 d\mu_m^{\otimes^2}(x,x') = 0, \text{ en } P_0 - \text{probabilité}$$

2. Nous appliquons le théorème D.1 (*cf.* annexe D), version *p.s.* du théorème D.2, pour laquelle les hypothèses *(i)-(iii)* restent inchangées, et les hypothèses *(iv)* et *(v)* deviennent :

   - *(iv)'* $\liminf_{m \to +\infty} U_m(\theta) \geq K(\theta_0, \theta), \ P_0 - p.s.$

– *(v)'* $P_0\left(\limsup_{m\to+\infty}\left\{W_m\left(\frac{1}{k}\right)+\epsilon_m\geq\eta_k\right\}\right)=0$, pour tout $k\geq 1$

L'hypothèse *(iv)'* est vérifiée grâce au point 2. du théorème 3.3 Pour vérifier *(v)'*, quand **E2'** remplace **E2** on peut montrer, en utilisant le théorème de Borel-Cantelli dans la démonstration précédente, que, pour tout $k\geq 1$:

$$P_0\left(\limsup_{m\to+\infty} B_{k,m}\right)=0$$

■

3.5 La covariance stationnaire est bijective sur $[0,1]$

Comme nous l'avons précisé au début du chapitre, notre intérêt majeur porte sur l'estimation non-paramétrique de la déformation. C'est pourquoi nous supposons ici que β_0 est connu, et nous notons R_0 la covariance stationnaire $R(.;\beta_0)$. Nous allons voir que si R_0 est une fonction strictement décroissante sur $[0,1]$ (comme par exemple les covariances exponentielles ou gaussiennes), alors l'estimation de la déformation peut se ramener à un problème de minimisation d'un contraste convexe. Cela est satisfaisant d'un point de vue numérique car on est assuré de l'unicité de la solution et de la convergence des algorithmes d'optimisation.

3.5.1 Présentation de la méthode

On considère tout d'abord que $x_0=0$ et $x_{n_{m-1}+1}=1$ appartiennent à Δ_{m-1}, pour tout $m\geq 2$:

$$\Delta_{m-1}=\left\{0=x_0<x_1<\cdots<x_{n_{m-1}}<x_{n_{m-1}+1}=1\right\}$$

Nous avons vu au paragraphe 3.3.3 qu'on peut utiliser comme ϵ_m−estimateur de la déformation f_0, la fonction $\tilde{f}_m$ définie sur $[0,1]$ par :

$$\begin{cases}\tilde{f}_m(x) &= \displaystyle\sum_{i=1}^{n_{m-1}+1}\left(\tilde{y}_{i-1,m}+\frac{\tilde{y}_{i,m}-\tilde{y}_{i-1,m}}{x_i-x_{i-1}}(x-x_{i-1})\right)\mathbb{1}_{[x_{i-1},x_i[}(x)\\ \tilde{f}_m(1) &= 1\end{cases}$$

où $\tilde{y}_{1,m},\tilde{y}_{2,m},\ldots,\tilde{y}_{n_{m-1},m}$ sont les ϵ_m−estimateurs, dans $\mathcal{Y}_{K,m}$, du minimum de contraste :

$$V_m\left(y_1,y_2,\ldots,y_{n_{m-1}}\right)=\frac{1}{n_{m-1}^2}\sum_{(i,j)\in[1,\ldots,n_{m-1}]^2}\left[\tilde{r}_m(x_i,x_j)-R_0(y_j-y_i)\right]^2$$

Le fait de connaître R_0, donc en particulier la variance σ_0^2, nous permet de ne considérer que les couples de stations (x_i, x_j) avec $i < j$. Comme de plus $\tilde{r}_m$ est une fonction symétrique, $\tilde{y}_{1,m}, \tilde{y}_{2,m}, \ldots, \tilde{y}_{n_{m-1},m}$ sont aussi les estimateurs du minimum de contraste :

$$\frac{1}{(n_{m-1}-1)n_{m-1}} \sum_{1\le i<j\le n_{m-1}} \left[\tilde{r}_m(x_i,x_j) - R_0(y_j - y_i)\right]^2$$

R_0 étant strictement décroissante sur $[0,1]$, nous avons, pour tout $i<j$:

$$R_0\left(f(x_j) - f(x_i)\right) = r_{\theta_0}(x_i,x_j) \Longleftrightarrow f(x_j) - f(x_i) = R_0^{-1}\left(r_{\theta_0}(x_i,x_j)\right)$$

Finalement, nous estimons $y_1, y_2, \ldots, y_{n_{m-1}}$ par $\tilde{\tilde{y}}_{1,m}, \tilde{\tilde{y}}_{2,m}, \ldots, \tilde{\tilde{y}}_{n_{m-1},m}$ minimisant dans $\mathcal{Y}_{K,m}$ le contraste :

$$V_m^*\left(y_1,y_2,\ldots,y_{n_{m-1}}\right) = \sum_{1\le i<j\le n_{m-1}} \left[R_0^{-1}\left(\tilde{r}_m(x_i,x_j)\right) - (y_j - y_i)\right]^2$$

Pratiquement, nous proposons la méthode suivante pour estimer les déformations des positions des stations :

$$y_1 = f_0(x_1) < y_2 = f_0(x_2) < \cdots < y_{n_{m-1}} = f_0(x_{m-1})$$

À T_m, nous souhaitons minimiser V_m^* sous les contraintes :

$$(S_3) : \begin{cases} 0 < y_1 < y_2 < \cdots < y_{n_{m-1}} < 1 \\ \dfrac{y_i - y_{i-1}}{x_i - x_{i-1}} \le K,\ i \in [1,\ldots,n_{m-1}+1] \end{cases}$$

1. on effectue tout d'abord le changement de paramètres :

$$z_k = y_k - y_{k-1},\ k \in [1,\ldots,n_{m-1}+1]$$

2. puis on minimise par rapport aux nouveaux paramètres z_k, $k \in [1,\ldots,n_{m-1}+1]$:

$$\sum_{0\le i<j\le n_{m-1}+1} \left[R_0^{-1}\left(\tilde{r}_m(x_i,x_j)\right) - \sum_{k=i+1}^{j} z_k\right]^2 \tag{3.10}$$

sous les contraintes :

$$(S_3') : \begin{cases} \displaystyle\sum_{k=1}^{n_{m-1}+1} z_k = 1 \\ 0 < z_k \le K(x_k - x_{k-1}),\ k \in [1,\ldots,n_{m-1}+1] \end{cases}$$

Minimiser (3.10) sous (S_3') est satisfaisant d'un point de vue numérique car (3.10) est convexe par rapport aux z_k, $k \in [2,\ldots,n_{m-1}+1]$, et les contraintes (S_3') délimitent un domaine convexe ;

3. soient $\tilde{\tilde{z}}_k$, $k \in [1, \dots, n_{m-1}+1]$, les ϵ_m−estimateurs du minimum de (3.10). On a finalement :
$$\begin{cases} \tilde{\tilde{y}}_1 = \tilde{\tilde{z}}_1 \\ \tilde{\tilde{y}}_2 = \tilde{\tilde{z}}_2 + \tilde{\tilde{z}}_1 \\ \vdots \\ \tilde{\tilde{y}}_{n_{m-1}} = \displaystyle\sum_{k=1}^{n_{m-1}} \tilde{\tilde{z}}_k = 1 - \tilde{\tilde{z}}_{n_{m-1}+1} \end{cases}$$

4. l'estimateur $\tilde{\tilde{f}}_m$ de la déformation f_0 est défini sur $[0,1]$ par :
$$\begin{cases} \tilde{\tilde{f}}_m(x) & = & \displaystyle\sum_{i=1}^{n_{m-1}+1} \left(\tilde{\tilde{y}}_{i-1,m} + \frac{\tilde{\tilde{y}}_{i,m} - \tilde{\tilde{y}}_{i-1,m}}{x_i - x_{i-1}}(x - x_{i-1}) \right) \mathbb{1}_{[x_{i-1}, x_i[}(x) \\ \tilde{\tilde{f}}_m(1) & = & 1 \end{cases}$$

3.5.2 Application numérique

Dans l'exemple qui suit, Z est un processus gaussien défini sur $[0,1]$, non-stationnaire, centré et réduit, et vérifiant l'hypothèse **N**.

Nous prenons comme famille de modèles de déformation l'ensemble $\mathcal{F}_K$, avec $K = 3$. Dans $\mathcal{F}_3$, nous considérons la déformation $f_0(x) = x^2$ (*cf.* figure 2.2, chapitre 2). On prend comme fonction de covariance stationnaire R_0, la covariance gaussienne, définie pour tout u dans $[-1,1]$ par :
$$R_0(u) = e^{-u^2}$$
Comme la covariance exponentielle, la covariance gaussienne est définie positive dans $\mathbb{R}^n$ quelque soit n dans $\mathbb{N}^*$. En outre, elle permet de caractériser l'ensemble des fonctions définies positives dans $\mathbb{R}^n$ pour tout n, en vertu du théorème de Schœnberg (1938b, pp.816-821) :

***Théorème* 3.6** *Une fonction :*
$$\begin{array}{rcl} R : \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ x & \longmapsto & R(x) \end{array}$$
qui ne dépend que de la norme euclidienne $\|x\|_n$ est une fonction définie positive pour tout n dans $\mathbb{N}^*$ si et seulement si elle peut s'écrire sous la forme :
$$R(x) = \int_0^{+\infty} e^{-u\|x\|_n^2} \varphi(du)$$
où φ est une mesure positive et finie.

★

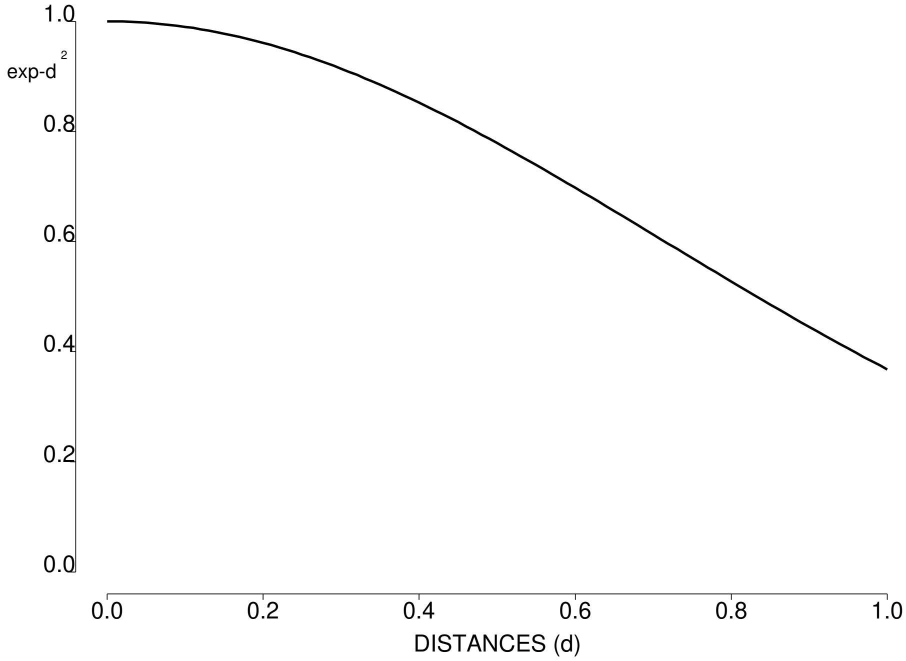


FIG. 3.4 – *Covariance gaussienne*

Nous allons estimer la déformation f_0 en trois dates :

- en $T_2 = 500$. En $T_1 = 1$, on prend $(n_1+1) = 25$ stations régulièrement réparties sur $[0,1]$, $x_i^1 = \dfrac{i-1}{n_1}, i = 1, \ldots, (n_1+1)$. Pour calculer les covariances empiriques $\tilde{r}_2(x_i^1, x_j^1)$, $i < j$, on simule, sous **H0**, T_2 réalisation *i.i.d.* du vecteur gaussien $\left(Z(0), Z(\dfrac{1}{n_1}), \ldots, Z(1)\right)$ centré de matrice de variances-covariances $\left[r_Z\left(\dfrac{i-1}{n_1}, \dfrac{j-1}{n_1}\right)\right]_{(n_1+1),(n_1+1)}$, avec pour tout $1 \leq i, j \leq (n_1+1)$:

$$r_Z\left(\frac{i-1}{n_1}, \frac{j-1}{n_1}\right) = \exp\left(-\frac{\left((j-1)^2 - (i-1)^2\right)^2}{n_1^4}\right)$$

Puis, on estime la déformation par $\tilde{\tilde{f}}_2$ selon la méthode proposée au paragraphe 3.5.1 précédent, mais sans tenir compte des contraintes, c'est-à-dire qu'on cherche les

$$\tilde{\tilde{z}}_k = \tilde{\tilde{f}}_2(x_k^1) - \tilde{\tilde{f}}_2(x_{k-1}^1) = \tilde{\tilde{y}}_k - \tilde{\tilde{y}}_{k-1},\ k \in [1, \ldots, n_1]$$

minimisant :

$$Q_{n_1}(z_1,\dots,z_{n_1}) = \sum_{0\le i<j\le n_1}\left[R_0^{-1}\left(\tilde r_2(x_i^1,x_j^1)\right) - \sum_{k=i+1}^{j} z_k\right]^2$$

sans tenir compte des contraintes :

$$(S_3') : \begin{cases} \displaystyle\sum_{k=1}^{n_1} z_k = 1 \\ 0 < z_k \le 3(x_k^1 - x_{k-1}^1),\ k\in[1,\dots,n_1] \end{cases}$$

On vérifiera *a posteriori* que ces contraintes sont bien respectées.

Q_{n_1} étant un polynôme convexe de degré 2 par rapport aux $z_1,\dots,z_{n_1}$, son minimum global est atteint pour les $\tilde{\tilde z}_k$, $k\in[1,\dots,n_1]$, solutions des n_1 équations normales :

$$\frac{\partial Q_{n_1}}{\partial z_k}(z_1,\dots,z_{n_1}) = 0,\ k\in[1,\dots,n_1]$$

Posons :

$$A_{n_1} = \begin{pmatrix} n_1 & n_1-1 & n_1-2 & \cdots & 1 \\ n_1-1 & 2(n_1-1) & 2(n_1-2) & \cdots & 2 \\ n_1-2 & 2(n_1-2) & 3(n_1-2) & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n_1 \end{pmatrix}$$

Les n_1 équations normales peuvent alors s'écrire sous la forme matricielle suivante :

$$A_{n_1} \times \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n_1} \end{pmatrix} = \begin{pmatrix} \displaystyle\sum_{0\le i<j\le n_1} R_0^{-1}\left(\tilde r_2(x_i^1,x_j^1)\right) \mathbb{1}_{[(i+1),j]}(1) \\ \displaystyle\sum_{0\le i<j\le n_1} R_0^{-1}\left(\tilde r_2(x_i^1,x_j^1)\right) \mathbb{1}_{[(i+1),j]}(2) \\ \vdots \\ \displaystyle\sum_{0\le i<j\le n_1} R_0^{-1}\left(\tilde r_2(x_i^1,x_j^1)\right) \mathbb{1}_{[(i+1),j]}(n_1) \end{pmatrix}$$

A_{n_1} est de plein rang. On montre en effet que son déterminant est différent de 0. On calcule ce déterminant de la manière suivante :

1. on divise la deuxième ligne de A_{n_1} par 2, on retranche le résultat à la première ligne et on obtient :

$$|A_{n_1}| = (n_1+1)\times \begin{vmatrix} n_1-1 & n_1-2 & \cdots & 1 \\ 2(n_1-2) & 3(n_1-2) & \cdots & 3 \\ \vdots & \vdots & \ddots & \vdots \\ 2 & 3 & \cdots & n_1 \end{vmatrix} = (n_1+1)\times|B_{n_1}|$$

2. on divise la deuxième ligne de B_{n_1} par 3, on retranche le résultat à la première ligne et on obtient :
$$|A_{n_1}| = (n_1+1)^2 \times \begin{vmatrix} n_1-2 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 3 & \cdots & n_1 \end{vmatrix}$$
3. on réitère ce calcul $(n_1 - 3)$ fois et on obtient :
$$|A_{n_1}| = (n_1+1)^{n_1-1} > 0$$

Finalement, il y a une solution unique au problème de minimisation de Q_{n_1}, solution donnée par :
$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n_1} \end{pmatrix} = A_{n_1}^{-1} \begin{pmatrix} \sum_{0\le i<j\le n_1} R_0^{-1}\left(\tilde{r}_2(x_i^1, x_j^1)\right) \mathbb{1}_{[(i+1),j]}(1) \\ \sum_{0\le i<j\le n_1} R_0^{-1}\left(\tilde{r}_2(x_i^1, x_j^1)\right) \mathbb{1}_{[(i+1),j]}(2) \\ \vdots \\ \sum_{0\le i<j\le n_1} R_0^{-1}\left(\tilde{r}_2(x_i^1, x_j^1)\right) \mathbb{1}_{[(i+1),j]}(n_1) \end{pmatrix}$$

Le graphique à gauche de la figure 3.5 représente $\tilde{\tilde{f}}_2(x)$ en fonction de $f_0(x)$, pour tout x dans $[0, 1]$.

Les contraintes (S_3') sont-elles respectées ? :
$$\begin{cases} \sum_{k=1}^{n_1} \tilde{\tilde{z}}_k = 0.89 \\ 1 \text{ estimation négative et } \sup_{1\le k\le n_1} \dfrac{\tilde{\tilde{z}}_k}{x_k - x_{k-1}} = 1.58 < K = 3 \end{cases}$$

Les contraintes ne sont donc pas respectées.

- en $T_3 = 1000$. En T_2, $\delta_2 = 12$ nouvelles stations régulièrement réparties, $x_i^2 = \dfrac{4(i-1)+1}{2n_1}$, $i = 1, \dots, \delta_2$, sont ajoutées aux 25 stations précédentes. On simule sur ces 37 stations $(T_3 - T_2 + 1)$ réalisations *i.i.d.* de Z. Pour estimer f_0 en T_3, on adopte la même méthode que précédemment. Le graphique au centre de la figure 3.5 représente l'estimation obtenue $\tilde{\tilde{f}}_3(x)$ en fonction de $f_0(x)$, pour tout x. Quant à la vérification des contraintes :
$$\begin{cases} \sum_{k=1}^{36} \tilde{\tilde{z}}_k = 0.99 \\ 2 \text{ estimations négatives et } \sup_{1\le k\le 36} \dfrac{\tilde{\tilde{z}}_k}{x_k - x_{k-1}} = 2.42 < K = 3 \end{cases}$$

Les contraintes ne sont donc pas respectées.

- en $T_4 = 10000$. En T_3, $\delta_3 = 12$ nouvelles stations régulièrement réparties, $x_i^3 = \dfrac{4(i-1)+3}{2n_1}$, $i = 1, \ldots, \delta_3$, sont ajoutées aux 37 stations précédentes. On simule sur ces 49 stations $(T_4 - T_3 + 1)$ réalisations *i.i.d.* de Z. Le graphique à droite de la figure 3.5 représente $\tilde{\tilde{f}}_4(x)$ en fonction de $f_0(x)$, pour tout x. Quant à la vérification des contraintes :

$$\begin{cases} \displaystyle\sum_{k=1}^{48} \tilde{\tilde{z}}_k = 0.99 \\ \text{aucune estimation négative et } \displaystyle\sup_{1 \leq k \leq 48} \frac{\tilde{\tilde{z}}_k}{x_k - x_{k-1}} = 2.18 < K = 3 \end{cases}$$

Les contraintes sont donc respectées.

Grâce à cet exemple, on constate, au vu de la figure 3.5, l'amélioration de la qualité de l'estimation de la déformation quand le nombre de stations croît avec l'augmentation du nombre de réalisations de Z. Toutefois, les contraintes ne sont pas systématiquement respectées. C'est pourquoi, il serait souhaitable d'en tenir compte directement dans la procédure d'estimation ; ce problème est ouvert.

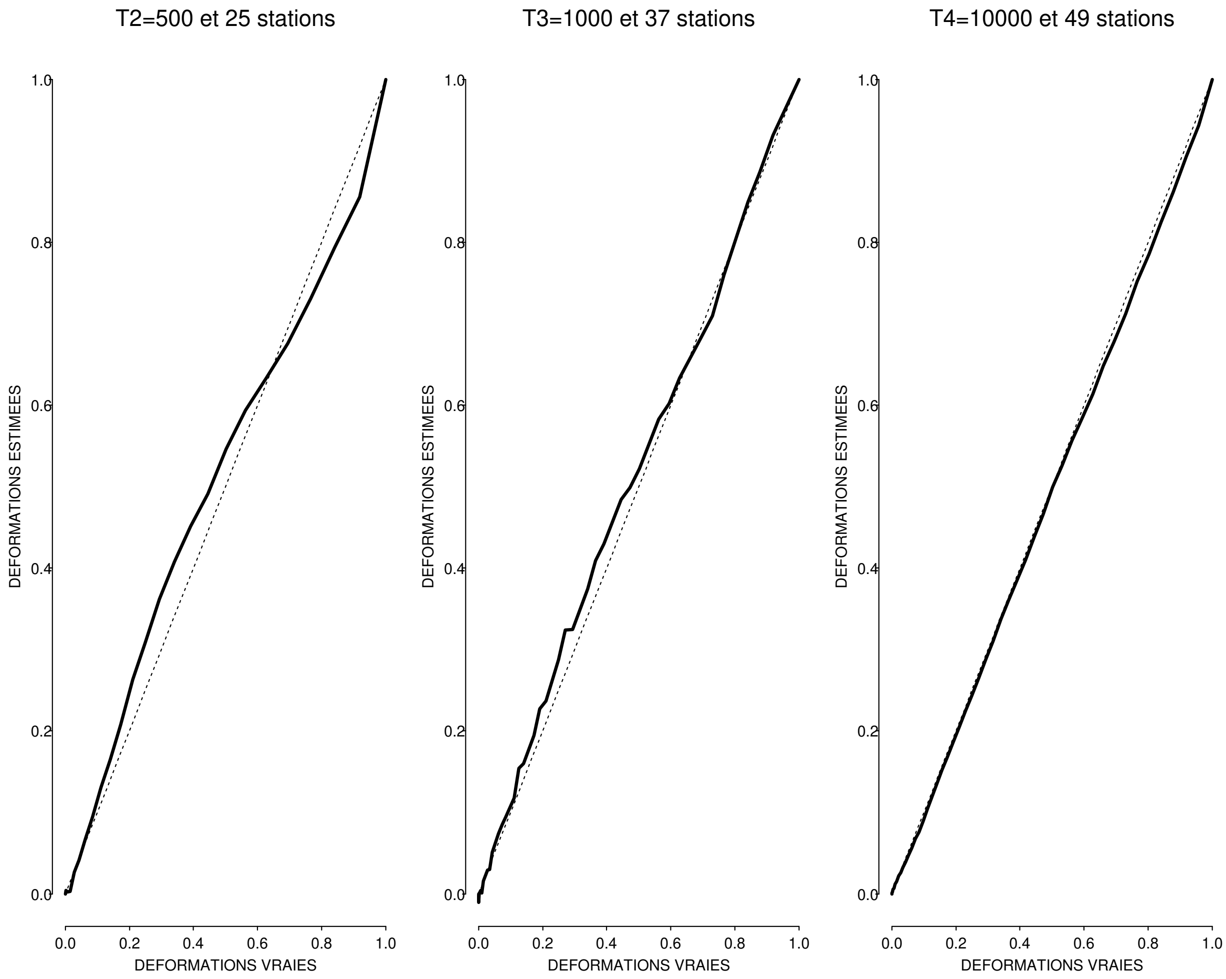


FIG. 3.5 – *Estimation de la déformation x^2 en trois dates différentes*

Chapitre 4

Processus gaussien avec observation fine sur une seule trajectoire

*Jusqu'à présent, la seule hypothèse que nous avons faite sur la loi du processus aléatoire Z portait sur l'existence d'un moment d'ordre 2ν, $\nu > 2$ (hypothèse **M3**). Dans ce chapitre, nous allons montrer que si le processus est gaussien et possède une fonction de covariance assez régulière, il suffit d'une seule trajectoire, ou réalisation, observée partout, pour identifier la déformation qui le rend stationnaire. La section 4.1 introduit le modèle. La section 4.2 présente la méthode utilisée et la technique que cette méthode utilise dans d'autres contextes. La section 4.3 s'intéresse à l'estimation du paramètre d'échelle de la fonction de covariance du processus stationnaire à l'ordre 2 associé à Z et aux propriétés asymptotiques, convergences p.s., en moyenne quadratique et en loi, de l'estimateur obtenu. L'estimation de la déformation et l'étude de sa convergence uniforme p.s. font l'objet de la section 4.4. La section 4.5 illustre l'étude menée dans ce chapitre par deux exemples.*

4.1 Le modèle

Soit $Z = \{Z(x)\}$, un processus aléatoire gaussien défini sur $[0,1]$, centré et réduit, de fonction de covariance r_Z. Posons $\Delta = \left\{(x,x') \in [0,1]^2, x \neq x'\right\}$ et supposons :

- **O1** : Z est continu en moyenne quadratique, c'est-à-dire r_Z est continue dans $[0,1]^2$
- **O2** : r_Z possède des dérivées secondes bornées uniformément dans Δ

De ces hypothèses, on déduit que pour tout x fixé dans $[0,1]$, $r_Z^{(0,1)}(x,.) = \dfrac{\partial r_Z}{\partial x'}(x,.)$ est continue dans $[0,1]$, sauf au point de discontinuité de première espèce x où elle admet une limite à gauche et une limite à droite. On déduit aussi que pour tout x' fixé dans $[0,1]$, $r_Z^{(0,1)}(.,x') = \dfrac{\partial r_Z}{\partial x'}(.,x')$ est continue dans $[0,1]$, sauf au point de discontinuité de première espèce x' où elle admet une limite à gauche et une limite à droite. Notons :

$$r_Z^{(0,1)}(x,x^-) = \lim_{x'\uparrow x} r_Z^{(0,1)}(x,x')$$
$$r_Z^{(0,1)}(x,x^+) = \lim_{x'\downarrow x} r_Z^{(0,1)}(x,x')$$

et définissons les deux fonctions :

$$\begin{array}{rcl} D^- : x \in]0,1] \longmapsto D^-(x) & = & r_Z^{(0,1)}(x,x^-) \\ D^+ : x \in [0,1[\longmapsto D^+(x) & = & r_Z^{(0,1)}(x,x^+) \end{array}$$

***Lemme* 4.1** *Sous les hypothèses **O1-O2**, D^- et D^+ sont continues dans* $[0,1]$.

★

Démonstration :

1. On montre tout d'abord que D^- est continue dans $]0,1]$ et que D^+ est continue dans $[0,1[$.

   (i) Pour D^-, il suffit de remarquer que pour tout x dans $]0,1]$ et tout $h>0$:

$$\begin{array}{rcl} |D^-(x+h) - D^-(x)| & \leq & |r_Z^{(0,1)}(x+h,(x+h)^-) - r_Z^{(0,1)}(x+h,x^-)| \\ & + & |r_Z^{(0,1)}(x+h,x^-) - r_Z^{(0,1)}(x,x^-)| \end{array}$$

   Sous l'hypothèse **O2**, le théorème des accroissements finis donne :

$$\begin{array}{rcl} |D^-(x+h) - D^-(x)| & \leq & h \sup\limits_{(x,x')\in\Delta} |r_Z^{(0,2)}(x,x')| \\ & + & h \sup\limits_{(x,x')\in\Delta} |r_Z^{(1,1)}(x,x')| \end{array}$$

   On en déduit alors la continuité de D^- dans $]0,1]$.

   (ii) Par le même type de raisonnement, on aboutirait à la continuité de D^+ dans $[0,1[$.

2. On fait ensuite un prolongement par continuité pour D^- et D^+ sur $[0,1]$.

   (i) Pour D^-, il suffit de remarquer que :

$$\begin{array}{rcl} \lim\limits_{x\downarrow 0} D^-(x) & = & \lim\limits_{x\downarrow 0}\lim\limits_{h\downarrow 0} r_Z^{(0,1)}(x,x-h) \\ & = & \lim\limits_{x\downarrow 0}\lim\limits_{h\downarrow 0} \left(r_Z^{(0,1)}(x,x-h) - r_Z^{(0,1)}(x,0) + r_Z^{(0,1)}(x,0) \right) \end{array}$$

D'une part, sous l'hypothèse **O2**, le théorème des accroissements finis donne :

$$|r_Z^{(0,1)}(x, x-h) - r_Z^{(0,1)}(x, 0)| \leq |x-h| \sup_{(x,x')\in\Delta} |r_Z^{(0,2)}(x, x')|$$

D'autre part, pour tout x' dans $[0,1]$, $r_Z^{(0,1)}(., x')$ étant continue par morceaux dans $[0,1]$:

$$\lim_{x\downarrow 0} r_Z^{(0,1)}(x, 0) = r_Z^{(0,1)}(0^+, 0)$$

On obtient finalement :

$$\lim_{x\downarrow 0} D^-(x) = r_Z^{(0,1)}(0^+, 0)$$

(ii) Pour D^+, on obtiendrait par le même type de raisonnement :

$$\lim_{x\uparrow 1} D^+(x) = r_Z^{(0,1)}(1^-, 1)$$

Pour conclure, on définit :

$$\begin{aligned} D^-(0) &= r_Z^{(0,1)}(0^+, 0) \\ D^+(1) &= r_Z^{(0,1)}(1^-, 1) \end{aligned}$$

■

Pour résumer, D^- et D^+ sont continues dans $[0,1]$, compact, donc uniformément continues et bornées dans $[0,1]$.

Remarque : le cas où le support de Z est $[a,b]$, $a < b$

C'est par commodité que nous avons pris Z défini sur $[0,1]$. Toute l'étude que nous allons menée dans ce chapitre resterait valable avec Z défini sur un intervalle $[a,b]$ et les hypothèses :

- **O1'** : r_Z est continue dans $[a,b]^2$
- **O2'** : r_Z possède des dérivées secondes bornées uniformément dans $\{(x,x') \in [a,b]^2, x \neq x'\}$

♣

Les conditions **O1**-**O2**, ou **O1'**-**O2'**, sont satisfaites pour les exemples suivants :

1. les processus à accroissements indépendants réduits, de fonction de covariance, pour tout (x,x') dans $[a,b]^2$, $0 < a < b$:

$$r_Z(x,x') = \frac{\gamma(x \wedge x')}{\sqrt{\gamma(x)\gamma(x')}}$$

   où γ est une fonction strictement positive, croissante et de classe C^2 dans $[a,b]$. Si $\gamma(x) = x$, on retrouve le processus de Wiener réduit ;

2. les processus de Wiener m fois intégré, $m \geq 1$, réduits, de fonction de covariance, pour tout (x, x') dans $[0,1]^2$:
$$r(x,x') = (2m+1)\frac{\displaystyle\int_0^1 ((x-y)\vee 0)^m((x'-y)\vee 0)^m dy}{\sqrt{x^{(2m+1)}x'^{(2m+1)}}}$$
3. les processus d'Ornstein-Uhlenbeck, de fonction de covariance, pour tout (x, x') dans $[0,1]^2$:
$$r(x,x') = e^{-\beta|x'-x|},\ \beta > 0$$
4. les mouvements browniens fractionnaires non dégénérés réduits d'indice $H \in [1/2, 1[$, de fonction de covariance, pour (x, x') dans $[a,b]^2$, $0 < a < b$:
$$r_Z(x,x') = \frac{1}{2}\frac{\left[x^{2H} + x'^{2H} - |x'-x|^{2H}\right]}{x^H x'^H}$$
Pour $H \in]0, 1/2[$, l'hypothèse **O2'** n'est pas vérifiée et pour $H = 1/2$, on retrouve le processus de Wiener réduit.

Soit $\mathcal{R} = \{R_\beta\}$, une famille de fonctions de covariance stationnaire continues dans $[-1,1]$, paramétrée par $\beta \in \mathbb{R}^{*+}$, telles que :

- **R1** : R_β admet une dérivée seconde bornée sur $[-1,1] - \{0\}$, pour tout β dans $\mathbb{R}^{*+}$

Pour tout β dans $\mathbb{R}^{*+}$, $R_\beta^{(1)}$ est donc continue dans $[-1,1]$, sauf au point de discontinuité de première espèce 0. On note $R_\beta^{(1)}(0^-)$ et $R_\beta^{(1)}(0^+)$ les dérivées premières à gauche et à droite en 0. À cause de la parité de R_β, on a :
$$R_\beta^{(1)}(0^-) = -R_\beta^{(1)}(0^+)$$
On suppose en outre que :

- **R2** : $R_\beta^{(1)}(0^-) = \beta\ (> 0)$

On en déduit que R_β vérifie au voisinage de 0 :
$$R_\beta(u) = 1 - \beta|u| + o(|u|)$$
Pour cette raison, on interprète β comme un paramètre d'échelle.

Cette famille de modèles de covariance, $\mathcal{R}$, est assez restreinte car la connaissance de la covariance au voisinage de 0 va impliquer sa connaissance partout. Pour d'autres familles à plusieurs paramètres, on ne pourra identifier que les paramètres exprimant le comportement local en 0 de la covariance. L'information récupérée est donc limitée à ce comportement local. Il faudrait donc proposer d'autres alternatives pour identifier l'ensemble des paramètres, c'est-à-dire pour récupérer toute l'information sur la covariance. On n'étudie pas ces alternatives dans ce travail. Toutefois, notons que sur $\mathbb{R}^2$ on disposera d'outils d'identification plus riches, fournis par un choix de variations plus important que sur $\mathbb{R}$ (Guyon (1987)).

Voici deux exemples de familles de covariances qui satisfont aux conditions **R1**-**R2** :

1. La famille des covariances exponentielles (*cf.* figure 2.3, chapitre 2) :

$$R_\beta(u) = \exp\left(-\beta|u|\right), \beta > 0$$

2. La famille des covariances sphériques (*cf.* figure 4.8) :

$$R_\beta(u) = 1 - \beta|u| + \frac{4\beta^3|u|^3}{27}, \text{ avec } 0 < \beta \leq \frac{3}{2}$$

Par contre, il n'en est pas ainsi pour la famille des modèles de covariance gaussienne (*cf.* figure 3.4, chapitre 3) :

$$R_\beta(u) = \exp\left(-\beta u^2\right), \beta > 0$$

pour laquelle l'hypothèse **R2** n'est pas vérifiée car :

$$R_\beta^{(1)}(0^-) = R_\beta^{(1)}(0^+) = 0$$

On considère par ailleurs la famille non-paramétrique $\mathcal{T}$ de tous les changements de variable f de $[0,1]$ dans $[0,1]$ suivante :

$$\mathcal{T} = \left\{\text{difféomorphismes croissants de classe } C^1 \text{ de } [0,1] \text{ sur } [0,1]\right\}$$

On considère alors l'hypothèse suivante :

- **H0** : il existe une fonction f_0 de $\mathcal{T}$ et un processus gaussien stationnaire $Y = \{Y(u)\}$ défini sur $[0,1]$, de fonction de covariance $r_Y(u,u') = R_{\beta_0}(u'-u)$ appartenant à $\mathcal{R}$, tels que, pour tout x dans $[0,1]$:

$$Y\left(f_0(x)\right) = Z(x)$$

Sous l'hypothèse **H0**, le processus Z a pour fonction de covariance, pour tout (x,x') dans $[0,1]^2$:

$$\begin{aligned} r_Z(x,x') &= r_Y\left(f_0(x), f_0(x')\right) \\ &= R_{\beta_0}\left(f_0(x') - f_0(x)\right) \end{aligned}$$

L'objet de ce chapitre est d'estimer, à partir d'une seule trajectoire de Z sur $[0,1]$, le paramètre fonctionnel f_0 et le paramètre d'échelle de la covariance β_0, puis d'étudier les propriétés asymptotiques des estimateurs de ces deux paramètres.

4.2 Principe de la méthode : utilisation des variations quadratiques

Dans le paragraphe 4.2.1, nous montrons que l'estimation de la transformation f_0 et de la covariance R_{β_0}, à travers son paramètre β_0, se ramène au problème de l'estimation de $\int_0^x \alpha_0(v)dv$, $x \in [0,1]$, où α_0 est la fonction de singularité du processus gaussien Z, fonction que nous allons définir ci-après. Dans le paragraphe 4.2.2, nous nous intéressons aux variations d'un processus gaussien et présentons certaines de leurs propriétés asymptotiques. Cela nous permettra de lier entre elles les deux notions, fonction de singularité et variation quadratique, et de justifier *a posteriori* la méthode d'estimation utilisée dans les sections 4.3 et 4.4, dont l'idée vient initialement de Istas (communication personnelle, 1996).

4.2.1 Principe de la méthode

Tout d'abord, on donne la définition de la fonction de singularité α_0 du processus gaussien Z sous les hypothèses **O1**-**O2** (*cf.* Sacks et Ylvisaker (1966)) :

***Définition* 4.1** *La fonction de singularité α_0 est définie, pour tout x dans $[0,1]$, par :*

$$\alpha_0(x) = D^-(x) - D^+(x)$$

♠

***Propriétés* 4.1**

1. *α_0 est positive.*
2. *$\int_0^x \alpha_0(v)dv$ existe, au sens de Riemann, pour tout x dans $[0,1]$.*

★

Démonstration :

1. Par définition de α_0, pour tout x dans $[0,1]$:

$$\begin{aligned}\alpha_0(x) &= r_Z^{(0,1)}(x,x^-) - r_Z^{(0,1)}(x,x^+) \\ &= \lim_{h\uparrow 0} \frac{r_Z(x,x+h) - r_Z(x,x)}{h} - \lim_{h'\downarrow 0} \frac{r_Z(x,x+h') - r_Z(x,x)}{h'}\end{aligned}$$

D'après les propriétés des fonctions de covariance, pour tout x dans $[0,1]$, tout $h < 0$ et tout $h' > 0$:

$$\begin{cases} r_Z(x,x+h) - r_Z(x,x) \leq 0 \\ r_Z(x,x+h') - r_Z(x,x) \leq 0 \end{cases}$$

Alors :

$$\lim_{h\uparrow 0} \frac{r_Z(x,x+h) - r_Z(x,x)}{h} - \lim_{h'\downarrow 0} \frac{r_Z(x,x+h') - r_Z(x,x)}{h'} \geq 0$$

Par conséquent, pour tout x dans $[0,1]$:

$$\alpha_0(x) \geq 0$$

2. C'est immédiat, car on a déduit des hypothèses **O1**-**O2** la continuité de D^- et D^+ dans $[0,1]$, compact (*cf.* lemme 4.1).

■

Dans les exemples précédents de processus satisfaisant les conditions **O1**-**O2**, ou **O1'**-**O2'** (*cf.* section 4.1), la fonction de singularité vaut :

1. pour les processus à accroissements indépendants réduits :

$$\alpha_0(x) = \frac{\gamma^{(1)}(x)}{\gamma(x)}$$

soit $\alpha_0(x) = \dfrac{1}{x}$ pour le processus de Wiener réduit ;

2. pour les processus de Wiener m fois intégré :

$$\alpha_0(x) = 0$$

3. pour les processus d'Ornstein-Uhlenbeck :

$$\alpha_0(x) = 2\beta$$

4. pour les mouvements browniens fractionnaires non dégénérés réduits d'indice $H \in]1/2, 1[$:

$$\alpha_0(x) = 0$$

Sous l'hypothèse **H0** on a, pour tout (x,x') dans $[0,1]^2$:

$$r_Z(x,x') = R_{\beta_0}(f_0(x') - f_0(x))$$

On en déduit alors, sous les hypothèses **O1**-**O2** et **R1**-**R2**, pour tout x dans $[0,1]$:

$$\begin{aligned} D^-(x) &= \beta_0 f_0^{(1)}(x) \\ D^+(x) &= -\beta_0 f_0^{(1)}(x) \end{aligned}$$

Par conséquent, sous l'hypothèse **H0**, la fonction de singularité de Z est définie dans $[0,1]$ par :

$$\alpha_0(x) = 2\beta_0 f_0^{(1)}(x) \tag{4.1}$$

On déduit de (4.1) :

1. $\beta_0 = \frac{1}{2}\int_0^1 \alpha_0(v)dv$

2. $f_0(x) = \int_0^x f_0^{(1)}(v)dv = \frac{\int_0^x \alpha_0(v)dv}{\int_0^1 \alpha_0(v)dv}$

L'estimation de $\int_0^x \alpha_0(v)dv$, pour tout x dans $[0,1]$, fournit le moyen d'estimer $f_0(x)$ et β_0 ; si $\hat{g}_n(x)$ est un estimateur de $\int_0^x \alpha_0(v)dv$, on peut donc estimer :

1. β_0 par $\hat{\beta}_n = \frac{1}{2}\hat{g}_n(1)$

2. $f_0(x)$ par $\hat{f}_n(x) = \frac{\hat{g}_n(x)}{\hat{g}_n(1)}$

4.2.2 Variations d'un processus gaussien

Ce paragraphe fournit quelques résultats sur différents types de variations de processus gaussiens. On verra *a posteriori* le lien qui existe entre les variations quadratiques et la fonction de singularité, et on en déduira, dans les sections 4.3 et 4.4, une méthode d'estimation.

Soient n dans $\mathbb{N}^*$ et $a(n)$ un entier strictement positif, et soit :

$$\Pi_{a(n)} = \left\{0 < \frac{1}{a(n)} < \cdots < \frac{a(n)-1}{a(n)} < 1\right\}$$

une partition de $[0,1]$ de pas régulier $\frac{1}{a(n)}$.

***Définition* 4.2** *Soit $\{Z(x), 0 \le x \le 1\}$ un processus gaussien et $p \in \mathbb{N}^*$:*

$$V_{a(n)}^{2p}(x) = \sum_{k=1}^{a(n)} \left[Z\left(\frac{k}{a(n)}\right) - Z\left(\frac{k-1}{a(n)}\right)\right]^{2p} \times \mathbb{1}_{[0,x]}\left(\frac{k}{a(n)}\right)$$

s'appelle la variation d'ordre $2p$ de Z le long de $\Pi_{a(n)}$ sur $[0,x]$.

♠

Baxter (1956) a démontré un résultat de convergence *p.s.* sur les variations quadratiques ($p = 1$) de processus gaussiens sur $[0,1]$, généralisant celui de convergence *p.s.* de la variation quadratique d'un processus de Wiener.

***Théorème* 4.1** *(Baxter (1956))*
Si $\{Z(x), 0 \leq x \leq 1\}$ est un processus gaussien satisfaisant les hypothèses ***O1-O2****, alors :*

$$V_{2^n}^2(1) \underset{n \to +\infty}{\longrightarrow} \int_0^1 \alpha_0(v) dv, \ p.s.$$

★

Le résultat de Baxter, (*cf.* Grenander (1981), pp.285-288), est utilisé pour fournir une condition suffisante d'orthogonalité entre deux mesures de probabilité gaussienne (Yaglom (1963), pp.338-342, et Grenander (1981), pp.288-290).

Guyon (1987), quant à lui, s'est intéressé, aux convergences dans L^2 des variations d'ordre $2p$, $p \geq 1$, de champs gaussiens stationnaires et non-stationnaires. Dans le cas d'un processus Z non-stationnaire de covariance r_Z, si $p > \dfrac{1}{2}$, nous avons sous les hypothèses :

- **T1** : $1 - \rho_Z(x, x+h) = \beta_0(x)|h|^{1/p} + o_x(h^{1/p}), h \longrightarrow 0$, avec β_0 continue, le petit o étant uniforme en x sur tout compact, ρ_Z étant la corrélation entre $Z(x)$ et $Z(x+h)$
- **T2** : $\sigma_0(x) = (r_Z(x,x))^{1/2}$ est de classe C^1
- **T3** : $r_Z^{(1,1)}$ existe et est continue dans Δ, $|r_Z^{(1,1)}(x,x')| \leq C(x,x')$ pour une fonction C symétrique, mesurable, avec $x' \longmapsto C(x,x')$ décroissante sur $x' > x$

***Théorème* 4.2** *(Guyon (1987), pp.186-187)*
Soit Z un processus gaussien sur $\mathbb{R}$, de covariance r_Z vérifiant ***T1-T3****. La variance v_n de la variation d'ordre $2p$ de Z le long de $[0,1]$, pour une partition Π_n de pas $\dfrac{1}{n}$, vérifie :*

$$v_n = O\left(\frac{1}{n} + \frac{1}{(n^{4-2/p})} \int_{0 \leq u,v \leq 1; u + \frac{2}{n} \leq v} C^2(u,v) du dv\right)$$

En particulier, dès que v_n tend vers 0, $n \longrightarrow +\infty$, la variation d'ordre $2p$ tend vers :

$$2^p a_{2p} \int_0^1 \left[\beta_0(v)\sigma_0^2(v)\right]^p dv$$

où a_{2p} est le moment centré d'ordre $2p$ de la loi normale réduite.

★

On rappelle que sous l'hypothèse **H0** :

$$r_Z(x, x') = R_{\beta_0}\left(f_0(x') - f_0(x)\right)$$

Dans le théorème 4.2, si on prend $p = 1$ (alors $a_{2p} = 1$) et si on suppose $\sigma_0(x) = 1$ pour tout x dans $[0, 1]$, on a la correspondance sous les hypothèses **H0** et **R1**-**R2**, pour tout x dans $[0, 1]$:

$$\beta_0(x) = \beta_0 f^{(1)}(x)$$

et le résultat de Guyon (1987) donne le résultat de Baxter dans sa version en moyenne quadratique.

Ces variations d'ordre $2p$ permettraient surtout de récupérer des comportements locaux de fonctions de covariance n'appartenant pas à $\mathcal{R}$, comme les covariances gaussiennes.

On peut également citer Guyon (1989) pour l'étude de la convergence faible des H-variations (basées sur $H\left(Z\left(\frac{k}{n}\right) - Z\left(\frac{k-1}{n}\right)\right)$, où H est une fonction réelle) d'un processus gaussien stationnaire sur $\mathbb{R}$ et Istas et Lang (1996) pour l'étude des convergences *p.s.* et en loi des a-variations quadratiques (généralisant les variations quadratiques) d'un processus gaussien à accroissements stationnaires sur $[0, +\infty[$ et l'utilisation de ces variations pour estimer l'exposant de Hölder des trajectoires.

Enfin, Istas (1995) s'est intéressé directement à l'estimation fonctionnelle de la fonction de singularité sur $[0, 1]$, en utilisant des noyaux de projection. Il a également établi les propriétés asymptotiques de cet estimateur, erreur quadratique moyenne intégrée et convergence en loi normalisée.

4.3 Estimation du paramètre d'échelle

D'après ce qu'on a vu dans la section 4.2, si on dispose d'une partition régulière Π_{2^n}, sous les hypothèses **O1**-**O2** et **R1**-**R2** on a, en appliquant le théorème de Baxter (1956) :

$$V_{2^n}^2(1) \underset{n \to +\infty}{\longrightarrow} \int_0^1 \alpha_0(v)dv = 2\beta_0, \; P_0 - p.s.$$

Dans le paragraphe 4.3.1, nous allons étudier la convergence *p.s.* de $V_n^2(1)$. Autrement dit, nous allons commencer par redémontrer le résultat de Baxter avec la partition Π_n. Ainsi $\hat{\beta}_n = \frac{1}{2}V_n^2(1)$ constituera un estimateur de β_0, convergeant $P_0 - p.s.$ vers β_0. Dans le paragraphe 4.3.2, nous étudions d'autres propriétés asymptotiques de $V_n^2(1)$, la convergence en moyenne quadratique et la convergence en loi. De ces propriétés, on en déduit celles de $\hat{\beta}_n$ dans le paragraphe 4.3.3.

4.3.1 Convergence *p.s.* de $V_n^2(1)$

Dans sa démonstration, Baxter utilise l'inégalité de Bienaymé-Cebyshev :

$$P\left\{\left|V_{2^n}^2(1) - E\left(V_{2^n}^2(1)\right)\right| > n/(2^n)^{1/2}\right\} < 2^n b_{2^n}/n^2$$

avec $b_{2^n} = Var\left(V_{2^n}^2(1)\right)$.

puis le lemme de Borel-Cantelli pour montrer que :

$$\left|V_{2^n}^2(1) - E\left(V_{2^n}^2(1)\right)\right| \underset{n\to+\infty}{\longrightarrow} 0, \ p.s.$$

L'utilisation de l'inégalité de Bienaymé-Cebyshev justifie *a posteriori* le fait de prendre un nombre d'observation de l'ordre de 2^n. Pour démontrer le théorème de Baxter avec une partition Π_n, c'est-à-dire avec $(n+1)$ observations, nous allons tout d'abord montrer, à l'instar de Baxter, que, pour tout $\epsilon > 0$:

$$\sum_{n=1}^{+\infty} P\left\{\left|V_n^2(1) - E\left(V_n^2(1)\right)\right| > \epsilon\right\} < +\infty$$

mais sans utiliser l'inégalité de Bienaymé-Cebyshev. Nous allons procéder en explicitant directement les propriétés de la variation quadratique

$$V_n^2(1) = \sum_{k=1}^{n}\left[Z\left(\frac{k}{n}\right) - Z\left(\frac{k-1}{n}\right)\right]^2$$

et en se servant d'une version simplifiée d'un théorème sur la convergence *p.c.* de sommes pondérées de variables aléatoires :

***Théorème* 4.3** *(Stout (1974), pp.226-229)*
Soit la somme pondérée :

$$S_n = \sum_{k=1}^{n} a_{k,n} X_k$$

où $\{X_n, n \geq 1\}$ est une suite de v.a.r. i.i.d., centrées, de carré intégrable et où les coefficients de pondération $a_{k,n}, k \in [1, \ldots, n]$, vérifient :

$$\begin{aligned} |a_{k,n}| &= O(1/n) \\ \sum_{k=1}^{n} a_{k,n}^2 &= o(1/\ln(n)) \end{aligned}$$

Alors S_n converge p.c. vers 0.

★

Soit $\Sigma_{0,n}$, la matrice de variances-covariances du vecteur gaussien :

$$W_n = \left(Z\left(\frac{0}{n}\right), Z\left(\frac{1}{n}\right), \ldots, Z\left(\frac{n}{n}\right) \right)'$$

$$\Sigma_{0,n} = \begin{pmatrix} r_Z(0/n,0/n) & r_Z(0/n,1/n) & \cdots & r_Z(0/n,n/n) \\ & r_Z(1/n,1/n) & \cdots & r_Z(1/n,n/n) \\ & & \ddots & \vdots \\ & & & r_Z(n/n,n/n) \end{pmatrix}_{(n+1),(n+1)}$$

Et soit L_n, la matrice des coefficients de variation :

$$L_n = \begin{pmatrix} -1 & +1 & 0 & \cdots & 0 \\ 0 & -1 & +1 & \cdots & 0 \\ & & \cdots & & \\ 0 & 0 & \cdots & -1 & +1 \end{pmatrix}_{n,(n+1)}$$

On a la relation :

$$V_n^2(1) = W_n' L_n' L_n W_n$$

On en déduit (Johnson et Kotz (1970), pp.149-151) :

***Lemme* 4.2**

$$V_n^2(1) \overset{\mathcal{L}}{\sim} \sum_{k=1}^{n} \lambda_{k,n} \chi^2_{k,n}(1)$$

où $\lambda_{1,n} \geq \lambda_{2,n} \geq \cdots \geq \lambda_{n,n}$ sont les n valeurs propres positives de la matrice de variances-covariances $L_n \Sigma_{0,n} L_n'$ du vecteur gaussien $L_n W_n$, et les $\chi^2_{k,n}(1)$, $k \in [1, \ldots, n]$, sont des $khi - deux$ indépendants à un degré de liberté.

★

Le lemme suivant donne l'ordre de grandeur des valeurs propres $\lambda_{k,n}$, $k \in [1, \ldots, n]$, contrôle qui va permettre de conclure sur la convergence *p.c.* vers 0 de $\big(V_n^2(1) - EV_n^2(1)\big)$.

***Lemme* 4.3** *Sous les hypothèses **O1-O2**, les n valeurs propres $\lambda_{1,n}, \lambda_{2,n}, \ldots, \lambda_{n,n}$ de $L_n \Sigma_{0,n} L_n'$ vérifient, uniformément en k :*

$$\lambda_{k,n} = 0(1/n)$$

★

Démonstration :

(i) Nous montrons dans un premier temps :

$$\sum_{k=1}^{n} \lambda_{k,n} = O(1)$$

On a :

$$\sum_{k=1}^{n} \lambda_{k,n} = tr\left(L_n \Sigma_{0,n} L'_n\right) = \sum_{k=1}^{n} \left[r_Z\left(\frac{k}{n}, \frac{k}{n}\right) - 2r_Z\left(\frac{k-1}{n}, \frac{k}{n}\right) + r_Z\left(\frac{k-1}{n}, \frac{k-1}{n}\right) \right]$$

Sous les hypothèses **O1**-**O2**, un développement de Taylor-Young à l'ordre 1 donne, pour tout $1 \leq k \leq n$:

$$r_Z\left(\frac{k}{n}, \frac{k}{n}\right) - 2r_Z\left(\frac{k-1}{n}, \frac{k}{n}\right) + r_Z\left(\frac{k-1}{n}, \frac{k-1}{n}\right) = \frac{1}{n}\left[D^-\left(\frac{k}{n}\right) - D^+\left(\frac{k-1}{n}\right)\right] + o(1/n)$$

où le $o(1/n)$ est uniforme en k, grâce aux hypothèses **O1**-**O2**.

Comme D^- et D^+ sont continues dans $[0,1]$, compact, donc bornées dans $[0,1]$, on obtient, pour tout $1 \leq k \leq n$:

$$r_Z\left(\frac{k}{n}, \frac{k}{n}\right) - 2r_Z\left(\frac{k-1}{n}, \frac{k}{n}\right) + r_Z\left(\frac{k-1}{n}, \frac{k-1}{n}\right) = O(1/n)$$

où le $O(1/n)$ est uniforme en k.

On en déduit donc:

$$\sum_{k=1}^{n} \lambda_{k,n} = O(1)$$

(ii) Montrons dans un second temps :

$$\sum_{k=1}^{n} \lambda_{k,n}^2 = O(1/n)$$

On a, d'après le lemme 4.2 :

$$Var\left(V_n^2(1)\right) = 2\sum_{k=1}^{n} \lambda_{k,n}^2 = 2tr(L_n \Sigma_{0,n} L'_n)^2$$

Notons, pour tout $1 \leq j, k \leq n$:

$$a_{j,k} = E\left[\left(Z\left(\frac{j}{n}\right) - Z\left(\frac{j-1}{n}\right)\right) \left(Z\left(\frac{k}{n}\right) - Z\left(\frac{k-1}{n}\right)\right)\right]$$

Rappelons que si le vecteur $(X_1, X_2, X_3, X_4)'$ suit une loi normale d'espérance nulle :

$$E(X_1X_2X_3X_4) = E(X_1X_2)E(X_3X_4) + E(X_1X_3)E(X_2X_4) + E(X_1X_4)E(X_2X_3)$$

On a alors :

$$
\begin{aligned}
E\left(V_n^2(1)\right) &= \sum_{k=1}^{n} a_{k,k} \\
E\left(V_n^2(1)\right)^2 &= 3\sum_{k=1}^{n} a_{k,k}^2 + 2\sum_{k=1}^{n}\sum_{j>k}\left(a_{k,k}a_{j,j} + 2a_{j,k}^2\right)
\end{aligned}
$$

On en déduit :

$$
Var\left(V_n^2(1)\right) = E\left(V_n^2(1)\right)^2 - E^2\left(V_n^2(1)\right) = 2\sum_{k=1}^{n}\sum_{j=1}^{n} a_{j,k}^2
$$

Sous les hypothèses **O1**-**O2**, un développement de Taylor-Young à l'ordre 2 donne, pour tout $1 \leq j \neq k \leq n$:

$$
|a_{j,k}| = \left| r_Z\left(\frac{j}{n}, \frac{k}{n}\right) + r_Z\left(\frac{j-1}{n}, \frac{k-1}{n}\right) - r_Z\left(\frac{j}{n}, \frac{k-1}{n}\right) - r_Z\left(\frac{j-1}{n}, \frac{k}{n}\right) \right| = O(1/n^2)
$$

où le $O(1/n^2)$ est indépendant de $j \neq k$, grâce aux hypothèses **O1**-**O2**.

Comme nous l'avons vu au (i), on a aussi, pour tout $1 \leq k \leq n$:

$$
a_{k,k} = O(1/n)
$$

où le $O(1/n)$ est uniforme en k.

Par conséquent :

$$
\begin{aligned}
Var\left(V_n^2(1)\right) &= 2\sum_{k=1}^{n} a_{k,k}^2 + 4\sum_{k=1}^{n}\sum_{j>k} a_{j,k}^2 \\
&= O(1/n)
\end{aligned}
$$

D'où :

$$
\sum_{k=1}^{n} \lambda_{k,n}^2 = O(1/n)
$$

Pour conclure, on majore la plus grande valeur propre $\lambda_{1,n}$. Pour cela, on utilise un résultat d'algèbre linéaire (Luenberger (1979), p.194) :

$$
\lambda_{1,n} \leq \sup_{1 \leq j \leq n} \sum_{k=1}^{n} |a_{j,k}|
$$

On a vu au point (ii) :

$$\begin{aligned} a_{k,k} &= O(1/n), \text{ pour tout } 1 \leq k \leq n \\ a_{j,k} &= O(1/n^2), \text{ pour tout } 1 \leq j \neq k \leq n \end{aligned}$$

où le $O(1/n)$ est uniforme en k et où le $O(1/n^2)$ est indépendant de $j \neq k$.

On en déduit donc :

$$\lambda_{1,n} = O(1/n)$$

Par conséquent, uniformément en k :

$$\lambda_{k,n} = O(1/n)$$

■

Du lemme 4.3 précédent, on en déduit la convergence *p.c.* de $\left(V_n^2(1) - E(V_n^2(1))\right)$ vers 0.

Lemme 4.4 *Sous les hypothèses **O1-O2**, pour tout $\epsilon > 0$:*

$$\sum_{n=1}^{+\infty} P\left\{\left|V_n^2(1) - E(V_n^2(1))\right| > \epsilon\right\} < +\infty$$

c'est-à-dire $\left(V_n^2(1) - E(V_n^2(1))\right)$ converge p.c. vers 0. ★

Démonstration :

Soit :

$$T_n = V_n^2(1) - E(V_n^2(1)) = \sum_{k=1}^{n} \lambda_{k,n} \left(\chi_{k,n}^2(1) - 1\right)$$

où, pour tout $n \geq 1$, $\left(\chi_{1,n}^2(1) - 1\right), \left(\chi_{2,n}^2(1) - 1\right), \ldots, \left(\chi_{n,n}^2(1) - 1\right)$ sont n variables *i.i.d.* et centrées issues de $\left(\chi^2(1) - 1\right)$. On a :

$$E\left|\chi^2(1) - 1\right|^2 < +\infty$$

et d'après le lemme 4.3, uniformément en k, pour n suffisamment grand et un $M < +\infty$:

$$\lambda_{k,n} \leq \frac{M}{n}$$

D'après le point *(ii)* de la démonstration du lemme 4.3 :

$$Var(T_n) = 2\sum_{k=1}^{n} \lambda_{k,n}^2 = O(1/n) = o(1/\ln n)$$

En suivant le même schéma de démonstration que celui du théorème de Stout (*cf.* théorème 4.1.3, p.226), nous pouvons conclure.

■

A partir des lemmes précédents, nous sommes en mesure de démontrer le résultat de Baxter avec la partition régulière Π_n générale, en identifiant la limite $E\left(V_n^2(1)\right)$:

Théorème 4.4 *Sous les hypothèses* ***O1-O2*** *:*

$$V_n^2(1) \underset{n\to+\infty}{\longrightarrow} \int_0^1 \alpha_0(v)dv, \ p.s.$$

★

Démonstration :

D'après le point *(i)* de la démonstration du lemme 4.3, pour tout $1 \le k \le n$:

$$a_{k,k} = \frac{1}{n}\left[D^-\left(\frac{k}{n}\right) - D^+\left(\frac{k-1}{n}\right)\right] + o(1/n)$$

où le $o(1/n)$ est uniforme en k.

D'après les hypothèses **O1**-**O2**, D^+ est continue et bornée dans $[0,1]$. On a donc, pour tout $1 \le k \le n-1$:

$$D^+\left(\frac{k-1}{n}\right) = D^+\left(\frac{k}{n}\right) + o(1)$$

où le $o(1)$ est uniforme en k, grâce aux hypothèses **O1**-**O2**.

On a aussi :

$$a_{n,n} = O(1/n) = o(1)$$

Par conséquent :

$$E(V_n^2(1)) = tr(L_n\Sigma_{0,n}L_n') = \frac{1}{n}\sum_{k=1}^{n-1}\left[D^-\left(\frac{k}{n}\right) - D^+\left(\frac{k}{n}\right)\right] + o(1)$$

Mais puisque :

$$\alpha_0\left(\frac{k}{n}\right) = D^-\left(\frac{k}{n}\right) - D^+\left(\frac{k}{n}\right)$$

on a :

$$E(V_n^2(1)) = \frac{1}{n}\sum_{k=1}^{n-1}\alpha_0\left(\frac{k}{n}\right) + o(1)$$

Comme α_0 est Riemann-intégrable sur $[0,1]$, on a par définition :

$$E(V_n^2(1)) \underset{n\to+\infty}{\longrightarrow} \int_0^1 \alpha_0(v)dv$$

Pour conclure sur la convergence *p.s.* de $V_n^2(1)$ vers $\int_0^1 \alpha_0(v)dv$, on utilise l'égalité :

$$V_n^2(1) = E(V_n^2(1)) + \left[V_n^2(1) - E(V_n^2(1))\right]$$

Le lemme 4.4 nous assure de la convergence presque-complète de $\left[V_n^2(1) - E(V_n^2(1))\right]$ vers 0. Par une application du théorème de Borel-Cantelli, on obtient la convergence *p.s.* de $\left[V_n^2(1) - E(V_n^2(1))\right]$ vers 0. On a finalement :

$$V_n^2(1) \underset{n\to+\infty}{\longrightarrow} \int_0^1 \alpha_0(v)dv, \ p.s.$$

■

4.3.2 Autres propriétés asymptotiques

Convergence en moyenne quadratique

***Théorème* 4.5** *Sous les hypothèses **O1-O2** :*

$$E\left[V_n^2(1) - \int_0^1 \alpha_0(v)dv\right]^2 \underset{n\to+\infty}{\longrightarrow} 0$$

★

Démonstration :

On a la décomposition :

$$E\left[V_n^2(1) - \int_0^1 \alpha_0(v)dv\right]^2 = Var(V_n^2(1)) + \left[E(V_n^2(1)) - \int_0^1 \alpha_0(v)dv\right]^2$$

La démonstration du théorème 4.4 prouve que $V_n^2(1)$ est un estimateur asymptotiquement sans biais de $\int_0^1 \alpha_0(v)dv$:

$$\left[E(V_n^2(1)) - \int_0^1 \alpha_0(v)dv\right] \underset{n\to+\infty}{\longrightarrow} 0$$

De plus, d'après le point *(ii)* de la démonstration du lemme 4.3 :

$$Var(V_n^2(1)) = O(1/n)$$

Par conséquent, on en déduit la convergence en moyenne quadratique de $V_n^2(1)$ vers $\int_0^1 \alpha_0(v)dv$.

■

Convergence en loi

Nous allons tout d'abord montrer le résultat suivant :

Lemme* 4.5** *Sous les hypothèses* ***O1-O2 *:*

$$n \sum_{k=1}^{n} \lambda_{k,n}^2 \underset{n\to+\infty}{\longrightarrow} \int_0^1 \alpha_0^2(v)dv$$

où on rappelle que $\lambda_{1,n} \geq \lambda_{2,n} \geq \cdots \geq \lambda_{n,n}$ sont les n valeurs propres positives de la matrice de variances-covariances $L_n\Sigma_{0,n}L'_n$.

★

Démonstration :

D'après le point *(ii)* de la démonstration du lemme 4.3 :

$$\sum_{k=1}^{n} \lambda_{k,n}^2 = tr(L_n\Sigma_{0,n}L'_n)^2 = \sum_{k=1}^{n} a_{k,k}^2 + 2\sum_{k=1}^{n}\sum_{j>k} a_{j,k}^2$$

Pour le deuxième terme du membre de droite de l'égalité précédente :

$$\sum_{k=1}^{n}\sum_{j>k} a_{j,k}^2 = O(1/n^2)$$

Pour ce qui est du premier terme, d'après la démonstration du théorème 4.4 :

$$\begin{cases} a_{k,k} & = & \dfrac{1}{n}\left[D^-\left(\dfrac{k}{n}\right) - D^+\left(\dfrac{k}{n}\right)\right] + o(1/n), \text{ pour tout } k \text{ dans } [1,\ldots,n-1] \\ a_{n,n} & = & O(1/n) \end{cases}$$

où le $o(1/n)$ est uniforme en k.

On en déduit alors :

$$\begin{cases} a_{k,k}^2 & = & \dfrac{1}{n^2}\left[D^-\left(\dfrac{k}{n}\right) - D^+\left(\dfrac{k}{n}\right)\right]^2 + o(1/n^2), \text{ pour tout } k \text{ dans } [1,\ldots,n-1] \\ a_{n,n}^2 & = & O(1/n^2) = o(1/n) \end{cases}$$

Par conséquent, on obtient :

$$n\sum_{k=1}^{n} a_{k,k}^2 = \frac{1}{n}\sum_{k=1}^{n-1}\left[D^-\left(\frac{k}{n}\right) - D^+\left(\frac{k}{n}\right)\right]^2 + o(1)$$

Soit encore :

$$n \sum_{k=1}^{n} a_{k,k}^2 = \frac{1}{n} \sum_{k=1}^{n-1} \alpha_0^2 \left(\frac{k}{n}\right) + o(1)$$

Comme α_0 est Riemann-intégrable sur $[0,1]$, on a le résultat cherché.

■

On en déduit le résultat de convergence en loi :

Théorème* 4.6** *Sous les hypothèses* ***O1-O2 *:*

$$\sqrt{n}\left(V_n^2(1) - E(V_n^2(1))\right) \underset{n\to+\infty}{\xrightarrow{loi}} \mathcal{N}_1\left(0, \int_0^1 \alpha_0^2(v)dv\right)$$

★

Démonstration :

D'après le lemme 4.2, on a :

$$V_n^2(1) - E(V_n^2(1)) \overset{\mathcal{L}}{\sim} \sum_{k=1}^{n} \lambda_{k,n}(\chi_{k,n}^2(1) - 1)$$

où, pour tout $n \geq 1$, les $(\chi_{k,n}^2(1) - 1)$, $k \in [1, \ldots, n]$, sont *i.i.d.* et centrées.

On applique alors le théorème limite centrale de Lyapunov (Durret (1996), p.121), corollaire du théorème limite centrale pour des tableaux triangulaires de Lindeberg-Feller (Durret (1996), p.116), aux variables aléatoires indépendantes et centrées :

$$X_{k,n} = \sqrt{n}\lambda_{k,n}(\chi_{k,n}^2(1) - 1)$$

Il suffit de vérifier la condition, pour un $\delta > 0$:

$$\sum_{k=1}^{n} E|X_{k,n}|^{2+\delta} \underset{n\to+\infty}{\longrightarrow} 0$$

Si on prend $\delta = 1$, on a, pour tout k dans $[1, \ldots, n]$:

$$E|X_{k,n}|^3 \leq 15 n^{3/2} \lambda_{k,n}^3$$

Or, le lemme 4.3 donne :

$$\lambda_{k,n} = O(1/n)$$

où le $O(1/n)$ est uniforme k.

On en déduit que, uniformément en k, pour n suffisamment grand et un $M' < +\infty$:

$$E|X_{k,n}|^3 \leq M' \frac{n^{3/2}}{n^3}$$

Soit :

$$\sum_{k=1}^{n} E|X_{k,n}|^3 \leq M' \frac{1}{n^{1/2}}$$

D'où le résultat.

■

4.3.3 Propriétés asymptotiques de $\hat{\beta}_n = \frac{1}{2} V_n^2(1)$

On déduit directement du paragraphe précédent, les résultats suivant :

Convergence *p.s.*

Corollaire* 4.1** *Sous les hypothèses* ***O1-O2 *et* ***R1-R2*** *:*

$$\hat{\beta}_n = \frac{1}{2} V_n^2(1) \underset{n \to +\infty}{\longrightarrow} \beta_0, \; P_0 - p.s.$$

★

Convergence en moyenne quadratique

Corollaire* 4.2** *Sous les hypothèses* ***O1-O2 *et* ***R1-R2*** *:*

$$E\left[\hat{\beta}_n - \frac{1}{2}\int_0^1 \alpha_0(v)dv\right]^2 \underset{n \to +\infty}{\longrightarrow} 0$$

★

Convergence en loi

Corollaire* 4.3** *Sous les hypothèses* ***O1-O2 *et* ***R1-R2*** *:*

$$\sqrt{n}\left(\hat{\beta}_n - E(\hat{\beta}_n)\right) \underset{n \to +\infty}{\overset{loi}{\longrightarrow}} \mathcal{N}_1\left(0, \frac{1}{4}\int_0^1 \alpha_0^2(v)dv\right), \textit{sous } \boldsymbol{H0}$$

★

4.4 Estimation de la déformation

On a vu au paragraphe 4.2.1 que sous l'hypothèse **H0** :

$$f_0(x) = \frac{\displaystyle\int_0^x \alpha_0(v)dv}{\displaystyle\int_0^1 \alpha_0(v)dv}$$

où α_0 est la fonction de singularité de Z. On dispose, comme estimateur de $\int_0^1 \alpha_0(v)dv$, de la variation quadratique sur $[0,1]$:

$$V_n^2(1) = \sum_{k=1}^{n} \left[Z\left(\frac{k}{n}\right) - Z\left(\frac{k-1}{n}\right) \right]^2$$

dont on connaît certaines propriétés asymptotiques (*cf.* section 4.3). Il reste donc à proposer un estimateur pour $\int_0^x \alpha_0(v)dv$. Dans le paragraphe 4.4.1, nous allons construire un estimateur continu et croissant *p.s.* de $\int_0^x \alpha_0(v)dv$. Dans le paragraphe 4.4.2, nous en déduisons un estimateur $\hat{f}_n$, continu et croissant $P_0 - p.s.$, de f_0, pour lequel on obtient un résultat de convergence uniforme $P_0 - p.s.$

4.4.1 Estimation de $\int_0^x \alpha_0(v)dv$

On pourrait utiliser comme estimateur de $\int_0^x \alpha_0(v)dv$, la variation quadratique $V_n^2(x)$ le long de Π_n calculée sur $[0,x]$:

$$V_n^2(x) = \sum_{k=1}^{n} \left[Z\left(\frac{k}{n}\right) - Z\left(\frac{k-1}{n}\right) \right]^2 \times \mathbb{1}_{[0,x]}\left(\frac{k}{n}\right)$$

On généralise ainsi la simple estimation de $\int_0^1 \alpha_0(v)dv$ par $V_n^2(1)$. Remarquons au passage que toutes les propriétés asymptotiques démontrées pour $V_n^2(1)$, la convergence *p.s.*, la convergence en moyenne quadratique et la convergence en loi, s'obtiennent de manière analogue pour $V_n^2(x)$, avec x quelconque dans $[0,1[$:

***Corollaire* 4.4** *Sous les hypothèses **O1-O2**, pour tout x dans $[0,1[$:*

$$\begin{array}{ll} (i) & V_n^2(x) \underset{n\to+\infty}{\longrightarrow} \int_0^x \alpha_0(v)dv, \ p.s. \\ (ii) & E\left[V_n^2(x) - \int_0^x \alpha_0(v)dv\right]^2 \underset{n\to+\infty}{\longrightarrow} 0 \\ (iii) & \sqrt{n}\left(V_n^2(x) - E(V_n^2(x))\right) \underset{n\to+\infty}{\xrightarrow{loi}} \mathcal{N}_1\left(0, \int_0^x \alpha_0^2(v)dv\right) \end{array}$$

★

Cependant, cet estimateur n'est pas continu, c'est une fonction en escalier avec des sauts aux points $\frac{k}{n}$, $k \in [0, \dots, n-1]$, de la partition Π_n, et $\frac{V_n^2(.)}{V_n^2(1)}$ ne constituerait pas un estimateur continu $P_0 - p.s.$ de f_0. On propose donc une autre alternative par interpolation linéaire avec raccordements continus aux points de Π_n.

Estimateur continu et croissant $p.s.$ de $\int_0^x \alpha_0(v)dv$

On définit un estimateur, $\hat{g}_n(x)$ de $\int_0^x \alpha_0(v)dv$ pour tout x dans $[0,1]$, de la façon suivante :

$$\hat{g}_n(x) = \sum_{k=1}^{\lfloor nx \rfloor} \left[Z\left(\frac{k}{n}\right) - Z\left(\frac{k-1}{n}\right)\right]^2 + n\left[Z\left(\frac{\lfloor nx \rfloor + 1}{n}\right) - Z\left(\frac{\lfloor nx \rfloor}{n}\right)\right]^2 \left(x - \frac{\lfloor nx \rfloor}{n}\right), \text{ si } x \in [0,1[$$

$$\hat{g}_n(1) = V_n^2(1), \text{ si } x = 1$$

Ainsi définie, $\hat{g}_n$ est une version continue $p.s.$ obtenue par interpolation linéaire de $V_n^2(.)$, puisqu'on peut encore écrire :

$$\begin{cases} \hat{g}_n(x) = V_n^2(x) + n\left[Z\left(\frac{\lfloor nx \rfloor + 1}{n}\right) - Z\left(\frac{\lfloor nx \rfloor}{n}\right)\right]^2 \left(x - \frac{\lfloor nx \rfloor}{n}\right), \text{ si } x \in [0,1[\\ \hat{g}_n(1) = V_n^2(1), \text{ si } x = 1 \end{cases}$$

Propriétés asymptotiques

Avant d'étudier la convergence uniforme $p.s.$ de $\hat{g}_n$ vers $\int_0^x \alpha_0(v)dv$, notons tout d'abord le résultat suivant sur le module de continuité des trajectoires de Z :

***Lemme* 4.6** *Sous les hypothèses **O1-O2**, pour toute constante* $0 < v < \frac{1}{2}$ *:*

$$\sup_{|x'-x| \leq h} |Z(x') - Z(x)| = o(h^v), \ p.s.$$

On en déduit que Z est p.s. à trajectoires continues.

★

Démonstration :

Il suffit de montrer, uniformément en (x, x') dans $[0,1]^2$, que :

$$E\left[Z(x') - Z(x)\right]^2 = O(|x' - x|) \tag{4.2}$$

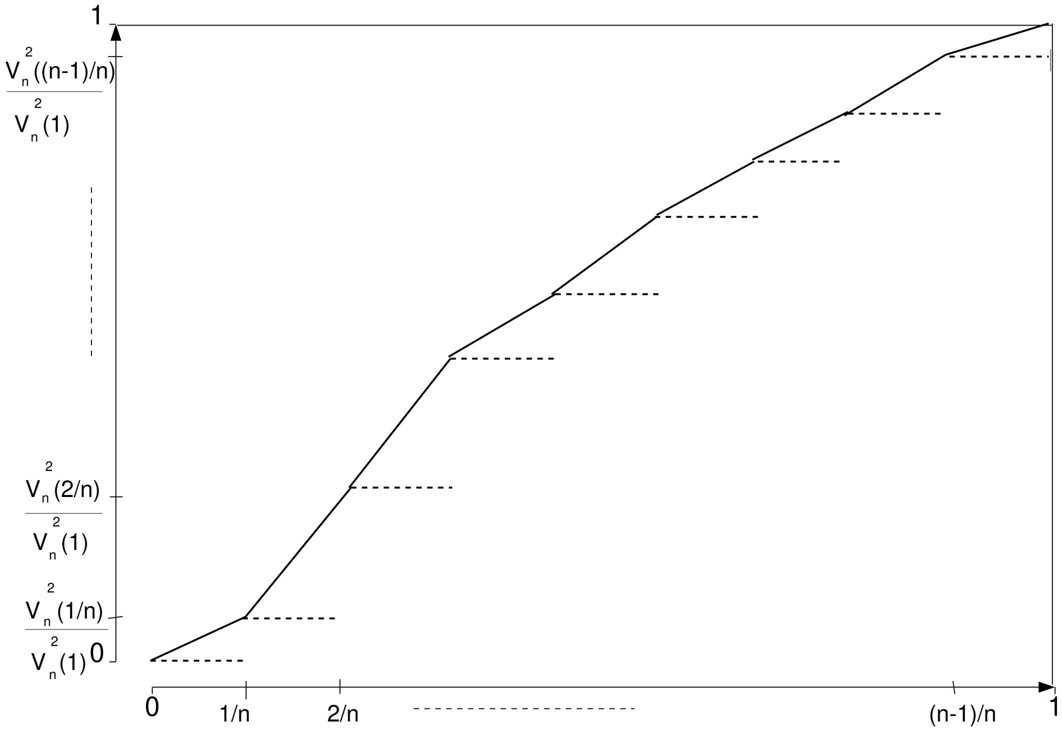


FIG. 4.1 – *Représentation de* $\hat{f}_n$ *(en trait plein) et de* $\dfrac{V_n^2(.)}{V_n^2(1)}$ *(en pointillé)*

puis d'utiliser un résultat de régularité des trajectoires de processus gaussiens (Neveu (1980), p.93).

(i) On a, pour tout (x, x') dans $[0,1]^2$, avec $x = x' - h, h > 0$:

$$E\left[Z(x') - Z(x)\right]^2 = E\left[Z(x') - Z(x'-h)\right]^2 = r_Z(x', x') - 2r_Z(x'-h, x') + r_Z(x'-h, x'-h)$$

Grâce aux hypothèses **O1**-**O2**, un développement de Taylor-Young à l'ordre 1 donne, pour tout $0 < x' \leq 1$:

$$r_Z(x', x') - 2r_Z(x'-h, x') + r_Z(x'-h, x'-h) = h\left[D^-(x') - D^+(x'-h)\right] + o(h)$$

où le $o(h)$ est indépendant de x'.

Comme D^- et D^+ sont bornées dans $[0,1]$, on obtient, pour tout $0 < x' \leq 1$:

$$E\left[Z(x') - Z(x'-h)\right]^2 = O(h)$$

où le $O(h)$ est indépendant de x'.

(ii) Le même raisonnement s'applique pour $x' < x$, et l'on obtient (4.2).

■

Théorème* 4.7** *Sous les hypothèses* ***O1-O2 *:*

$$\sup_{x\in[0,1]}\left|\hat{g}_n(x)-\int_0^x \alpha_0(v)dv\right| \underset{n\to+\infty}{\longrightarrow} 0,\ p.s.$$

★

Démonstration :

Comme $(\hat{g}_n)$ est une suite de fonctions croissantes dans $[0,1]$ compact, il est suffisant de montrer que $\hat{g}_n$ converge simplement $p.s.$ vers l'application continue $x \longmapsto \int_0^x \alpha_0(v)dv$, dans $[0,1]$ (Choquet (1992), p.115).

(i) D'une part, le théorème 4.4 nous donne la convergence de $\hat{g}_n(1) = V_n^2(1)$ vers $\int_0^1 \alpha_0(v)dv$.

(ii) D'autre part, pour tout x dans $[0,1[$, nous avons :

$$\hat{g}_n(x)=\sum_{k=1}^{\lfloor nx\rfloor}\left[Z\left(\frac{k}{n}\right)-Z\left(\frac{k-1}{n}\right)\right]^2+n\left[Z\left(\frac{\lfloor nx\rfloor+1}{n}\right)-Z\left(\frac{\lfloor nx\rfloor}{n}\right)\right]^2\left(x-\frac{\lfloor nx\rfloor}{n}\right)$$

Grâce au corollaire 4.4, nous avons pour le premier terme du membre de droite de l'égalité précédente et pour tout x dans $[0,1[$:

$$\sum_{k=1}^{\lfloor nx\rfloor}\left[Z\left(\frac{k}{n}\right)-Z\left(\frac{k-1}{n}\right)\right]^2 = V_n^2(x) \underset{n\to+\infty}{\longrightarrow} \int_0^x \alpha_0(v)dv,\ p.s.$$

Pour le second terme, pour tout x dans $[0,1[$ et tout $n\geq 1$:

$$0\leq n\left(x-\frac{\lfloor nx\rfloor}{n}\right)<1$$

et d'après le lemme 4.6, Z est $p.s.$ à trajectoires continues. On a donc, pour tout x dans $[0,1[$:

$$\left|Z\left(\frac{\lfloor nx\rfloor+1}{n}\right)-Z\left(\frac{\lfloor nx\rfloor}{n}\right)\right| \underset{n\to+\infty}{\longrightarrow} 0,\ p.s.$$

Soit :

$$n\left[Z\left(\frac{\lfloor nx\rfloor+1}{n}\right)-Z\left(\frac{\lfloor nx\rfloor}{n}\right)\right]^2\left(x-\frac{\lfloor nx\rfloor}{n}\right) \underset{n\to+\infty}{\longrightarrow} 0,\ p.s.$$

Finalement, pour tout x dans $[0,1[$:

$$\hat{g}_n(x) \underset{n\to+\infty}{\longrightarrow} \int_0^x \alpha_0(v)dv, \ p.s.$$

Par conséquent, on conclut à la convergence simple *p.s.* de $\hat{g}_n$ vers l'application continue $x \longmapsto \int_0^x \alpha_0(v)dv$, dans $[0,1]$. ■

Nous avons aussi un résultat de convergence en loi ponctuelle :

***Théorème* 4.8** *Sous les hypothèses **O1-O2**, pour tout x dans $[0,1]$:*

$$\sqrt{n}\left(\hat{g}_n(x) - E\left(\hat{g}_n(x)\right)\right) \xrightarrow[n\to+\infty]{loi} \mathcal{N}_1\left(0, \int_0^x \alpha_0^2(v)dv\right), sous\ \textbf{\textit{H0}}$$

★

Démonstration :

(i) Pour $x = 1$, le résultat est donné par le théorème 4.6.

(ii) Pour tout $x \in [0,1[$, on a la décomposition :

$$\begin{aligned}\sqrt{n}\left(\hat{g}_n(x) - E\left(\hat{g}_n(x)\right)\right) &= \sqrt{n}\left(V_n^2(x) - E\left(V_n^2(x)\right)\right)\\ &+ \sqrt{n}\left(nx - \lfloor nx\rfloor\right)\left[Z\left(\frac{\lfloor nx\rfloor+1}{n}\right) - Z\left(\frac{\lfloor nx\rfloor}{n}\right)\right]^2\\ &- \sqrt{n}\left(nx - \lfloor nx\rfloor\right)E\left[Z\left(\frac{\lfloor nx\rfloor+1}{n}\right) - Z\left(\frac{\lfloor nx\rfloor}{n}\right)\right]^2\end{aligned}$$

D'après le corollaire 4.4, pour tout x dans $[0,1[$:

$$\sqrt{n}\left(V_n^2(x) - E(V_n^2(x))\right) \xrightarrow[n\to+\infty]{loi} \mathcal{N}_1\left(0, \int_0^x \alpha_0^2(v)dv\right)$$

De plus, pour tout x dans $[0,1[$ et tout $n \geq 1$:

$$0 \leq n\left(x - \frac{\lfloor nx\rfloor}{n}\right) < 1$$

D'après le lemme 4.6 :

$$\sqrt{n}\left(nx - \lfloor nx\rfloor\right)\left[Z\left(\frac{\lfloor nx\rfloor+1}{n}\right) - Z\left(\frac{\lfloor nx\rfloor}{n}\right)\right]^2 \underset{n\to+\infty}{\longrightarrow} 0, \ p.s.$$

et, d'après le théorème de la convergence dominée de Lebesgue :

$$\sqrt{n}\left(nx - \lfloor nx\rfloor\right)E\left[Z\left(\frac{\lfloor nx\rfloor+1}{n}\right) - Z\left(\frac{\lfloor nx\rfloor}{n}\right)\right]^2 \underset{n\to+\infty}{\longrightarrow} 0$$

D'où le résultat.

■

4.4.2 Estimation de f_0

Définition de l'estimateur

Sous l'hypothèse **H0** :

$$f_0(x) = \frac{\displaystyle\int_0^x \alpha_0(v)dv}{\displaystyle\int_0^1 \alpha_0(v)dv}$$

On propose alors, comme estimateur de f_0, la fonction $\hat{f}_n$ définie, pour tout x dans $[0,1]$, par :

$$\hat{f}_n(x) = \frac{\hat{g}_n(x)}{\hat{g}_n(1)}$$

Convergence uniforme $P_0 - p.s.$

Elle est déduite directement de celle de $\hat{g}_n$:

***Corollaire* 4.5** *Sous les hypothèses **O1-O2** et **R1-R2** :*

$$\sup_{x\in[0,1]} |\hat{f}_n(x) - f_0(x)| \underset{n\to+\infty}{\longrightarrow} 0, \ P_0 - p.s.$$

★

Démonstration :

Comme $(\hat{f}_n)$ est une suite de fonctions croissantes définies dans $[0,1]$ compact, il est encore suffisant de montrer que $\hat{f}_n$ converge simplement $P_0 - p.s.$ vers l'application continue f_0. Cette dernière convergence est assurée grâce à la convergence simple $p.s.$ de $\hat{g}_n$ vers l'application continue $x \longmapsto \displaystyle\int_0^x \alpha_0(v)dv$ (*cf.* démonstration du théorème 4.7) et grâce à la convergence $p.s.$ de $V_n^2(1)$ vers $\displaystyle\int_0^1 \alpha_0(v)dv$ vue au théorème 4.4.

■

4.5 Exemples

Dans les deux exemples qui suivent, Z est un processus aléatoire gaussien défini sur $[0,1]$, non-stationnaire, centré et réduit, et vérifiant les hypothèses **O1-O2**.

4.5.1 Exemple 1 : processus d'Ornstein-Uhlenbeck

Dans cet exemple, nous prenons pour $Z = \{Z(x), 0 \leq x \leq 1\}$ un processus qui après déformation f_0 est un processus d'Ornstein-Uhlenbeck $Y = \{Y(u), 0 \leq u \leq 1\}$, $u = f_0(x)$, de fonction de covariance exponentielle (*cf.* figure 2.3, chapitre 2) :

$$R_{\beta_0}(u) = e^{-\beta_0 |u|},\ \beta_0 > 0$$

Les hypothèses **R1**-**R2** sont bien vérifiées. La covariance r_Z de Z vérifie donc, sous l'hypothèse **H0** :

$$r_Z(x, x') = e^{-\beta_0 |f_0(x') - f_0(x)|}$$

Nous prenons $\beta_0 = 1$ et $f_0(x) = x^2$ comme déformation appartenant à $\mathcal{T}$ (*cf.* figure 2.2, chapitre 2).

L'étude numérique qui suit a pour but d'illustrer expérimentalement le résultat de normalité asymptotique de l'estimateur $\hat{\beta}_n$ de β_0 (corollaire 4.3) et de fournir une estimation de f_0.

Notons que cette exemple nous permet de simuler rapidement Z, sous **H0**, sur une partition fine de $[0, 1]$, régulière et de pas $\frac{1}{n}$, avec n pouvant être très grand ($n = 1000$ ou 10000 par exemple) :

1. on prend $Z(0) = 0$;
2. on simule n variables *i.i.d.* ϵ_k, $k = 1 \ldots n$, de loi normale centrée réduite ;
3. on calcule de manière récursive pour k allant de 1 à n :

$$Z\left(\frac{k}{n}\right) = \exp\left(-\frac{\left(k^2 - (k-1)^2\right)}{n^2}\right) Z\left(\frac{k-1}{n}\right) + \epsilon_k \sqrt{1 - \exp\left(-\frac{2\left(k^2 - (k-1)^2\right)}{n^2}\right)}$$

Pour simuler Z sur une partition de pas $\frac{1}{n}$, il n'est donc pas besoin de calculer une matrice de variances-covariances de taille $(n+1) \times (n+1)$, ce qui est satisfaisant d'un point de vue numérique et informatique.

À partir de l'observation fine d'une seule trajectoire de Z, nous sommes maintenant capables d'estimer β_0 et f_0 grâce aux variations quadratiques, et donc d'identifier r_Z sous l'hypothèse **H0**.

Estimation du paramètre d'échelle β_0 et normalité asymptotique de $\left(\hat{\beta}_n - E(\hat{\beta}_n)\right)$

On estime β_0 par :

$$\hat{\beta}_n = \frac{1}{2}V_n^2(1) = \frac{1}{2}\sum_{k=1}^{n}\left[Z\left(\frac{k}{n}\right) - Z\left(\frac{k-1}{n}\right)\right]^2$$

(i) Calculons d'abord analytiquement le biais de $\hat{\beta}_n$; d'après le point *(ii)* de la démonstration du lemme 4.3, on a :

$$\begin{aligned} E(\hat{\beta}_n) &= \frac{1}{2}E(V_n^2(1)) = \sum_{k=1}^{n}\left(1 - \exp\left(-\frac{k^2-(k-1)^2}{n^2}\right)\right) \\ &= n - \exp\left(-\frac{1}{n^2}\right)\left(\frac{1-\exp\left(-\frac{2}{n}\right)}{1-\exp\left(-\frac{2}{n^2}\right)}\right) \end{aligned}$$

Soit pour $n = 1000$, le biais vaut $3,34 \times 10^{-4}$ (donc est proche de 0).

(ii) On a ensuite :

$$\sqrt{n}\left(\hat{\beta}_n - E(\hat{\beta}_n)\right) \xrightarrow[n\to+\infty]{loi} \mathcal{N}_1\left(0, \frac{1}{4}\int_0^1 \alpha_0^2(v)dv\right), \text{sous } \mathbf{H0}$$

où la fonction de singularité α_0 est définie sur $[0,1]$ par :

$$\alpha_0(x) = 2\beta_0 f_0^{(1)}(x) = 4x$$

Par conséquent, on a pour la variance asymptotique :

$$\frac{1}{4}\int_0^1 \alpha_0^2(v)dv = \frac{4}{3}$$

Pour avoir une idée de la distribution de $\left(\hat{\beta}_n - E(\hat{\beta}_n)\right)$, on simule, sous l'hypothèse **H0**, 100, 200 et 500 fois une trajectoire de Z sur la partition Π_n, $n = 1000$. On construit ensuite l'histogramme de $\left(\hat{\beta}_n - E(\hat{\beta}_n)\right)$ pour chaque série de simulations, afin d'avoir une idée, à distance finie, de la distribution des estimateurs et comparer cette distribution empirique à une distribution gaussienne.

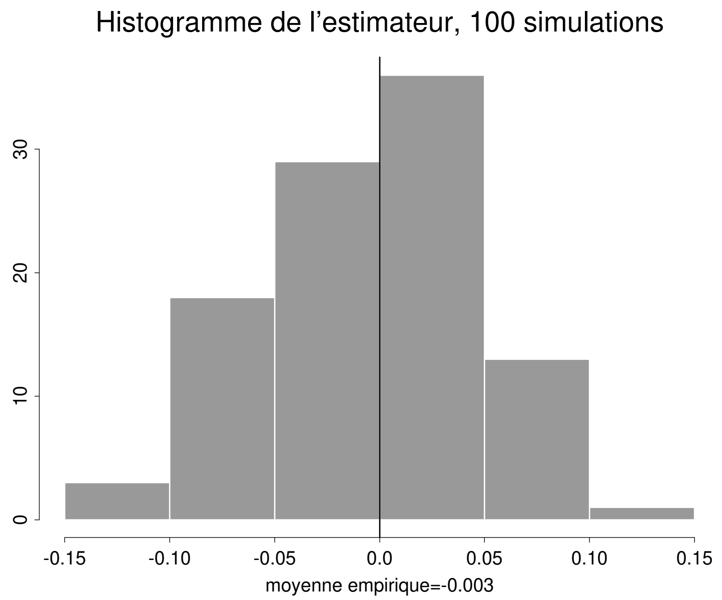


FIG. 4.2 – *Distribution empirique de* $\left(\hat{\beta}_n - E(\hat{\beta}_n)\right)$ *pour une covariance exponentielle*

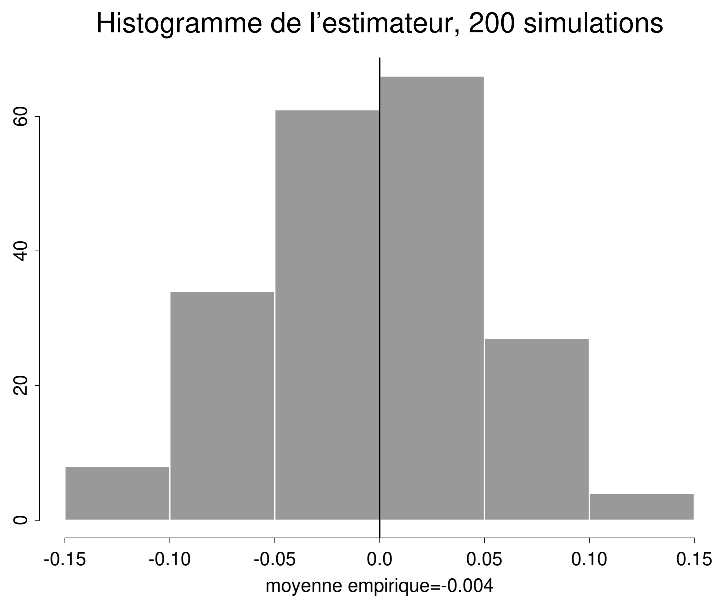


FIG. 4.3 – *Distribution empirique de* $\left(\hat{\beta}_n - E(\hat{\beta}_n)\right)$ *pour une covariance exponentielle*

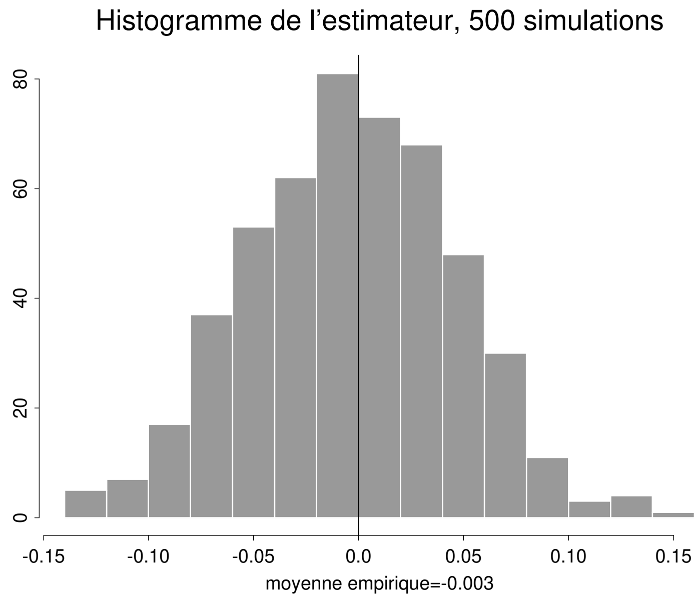


FIG. 4.4 – *Distribution empirique de* $\left(\hat{\beta}_n - E(\hat{\beta}_n)\right)$ *pour une covariance exponentielle*

On constate que si le nombre de simulations pour valablement approcher la loi de $\left(\hat{\beta}_n - E(\hat{\beta}_n)\right)$ est suffisant ($N \geq 100$, *cf.* figures 4.2 à 4.4), la moyenne empirique de $\left(\hat{\beta}_n - E(\hat{\beta}_n)\right)$est proche de 0 et les histogrammes sont approximativement gaussiens.

Estimation de la déformation

On estime uniformément $f_0 : x \longmapsto x^2$ sur $[0,1]$ par :

$$\hat{f}_n(x) = \frac{1}{V_n^2(1)} \left(\sum_{k=1}^{\lfloor nx \rfloor} \left[Z\left(\frac{k}{n}\right) - Z\left(\frac{k-1}{n}\right)\right]^2 + n \left[Z\left(\frac{\lfloor nx \rfloor + 1}{n}\right) - Z\left(\frac{\lfloor nx \rfloor}{n}\right)\right]^2 \left(x - \frac{\lfloor nx \rfloor}{n}\right)\right) \text{ si } x \in [0,1[$$

$$\hat{f}_n(1) = 1 \text{ si } x = 1$$

Nous nous donnons deux partitions régulières de $[0,1]$, la première de pas $\frac{1}{1000}$, la seconde de pas $\frac{1}{10000}$.

L'observation d'une seule trajectoire de Z sur la première partition nous donne un bon ajustement, comme le montre la figure 4.5 représentant les déformations estimées $\hat{f}_n(x)$ en fonction des vraies valeurs $f_0(x) = x^2$, pour tout x dans $[0,1]$.

Une observation plus fine grâce à la seconde partition nous donne un ajustement presque parfait (*cf.* figure 4.6).

4.5.2 Exemple 2 : cas d'une covariance sphérique

Nous prenons comme déformation appartenant à $\mathcal{T}$, la fonction :

$$f_0 : x \longmapsto \frac{\ln(1+x)}{\ln(2)}$$

et comme famille de modèles de covariance stationnaire $\mathcal{R}$, les covariances sphériques :

$$R_\beta(u) = 1 - \beta|u| + \frac{4\beta^3|u|^3}{27}, \text{ avec } 0 < \beta \leq \frac{3}{2}$$

Nous prenons $\beta_0 = 1$ comme vraie valeur du paramètre. La fonction de singularité α_0 est définie sur $[0,1]$ par :

$$\alpha_0(x) = 2\beta_0 f_0^{(1)}(x) = \frac{2}{\ln(2)x}$$

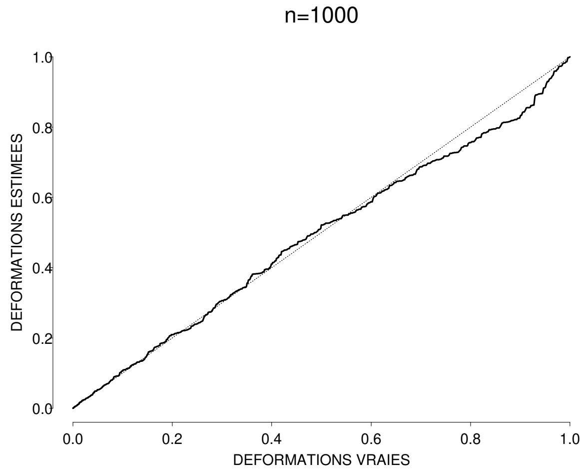


FIG. 4.5 – *Estimation de la déformation x^2 avec une partition régulière de pas* $\frac{1}{1000}$

Sans hypothèse supplémentaire sur Z, on ne dispose pas, comme dans l'exemple précédent, d'algorithme permettant de simuler rapidement des trajectoires de Z sur une partition fine et régulière de $[0,1]$. Toutefois, on peut toujours simuler, sous **H0**, p réalisations *i.i.d.* du vecteur gaussien $\left(Z(0), Z\left(\frac{1}{n}\right), \ldots, Z\left(\frac{n-1}{n}\right), Z(1)\right)$ centré et de matrice de variances-covariances $\left[r_Z\left(\frac{i-1}{n}, \frac{j-1}{n}\right)\right]_{(n+1),(n+1)}$, avec, pour tout $1 \leq i,j \leq n+1$:

$$\begin{aligned} r_Z\left(\frac{i-1}{n}, \frac{j-1}{n}\right) &= 1 - \frac{1}{\ln(2)} \left|\ln\left(1 + ((i-1)/n)\right) - \ln\left(1 + ((j-1)/n)\right)\right| \\ &+ \frac{4\left|\ln\left(1 + ((i-1)/n)\right) - \ln\left(1 + ((j-1)/n)\right)\right|^3}{27(\ln(2))^3} \end{aligned}$$

Avec $n = 1000$ et $p = 100$, nous obtenons d'une part une distribution empirique pour $\left(\hat{\beta}_n - E(\hat{\beta}_n)\right)$ proche d'une gaussienne (*cf.* figure 4.9), et d'autre part une bonne estimation de la déformation (*cf.* figure 4.10).

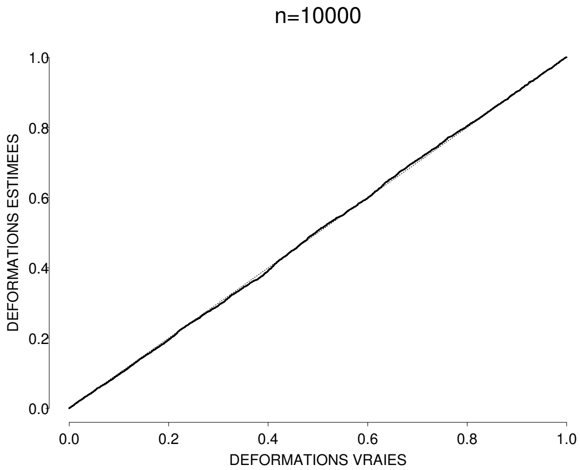


FIG. 4.6 – *Estimation de la déformation* x^2 *avec une partition régulière de pas* $\dfrac{1}{10000}$

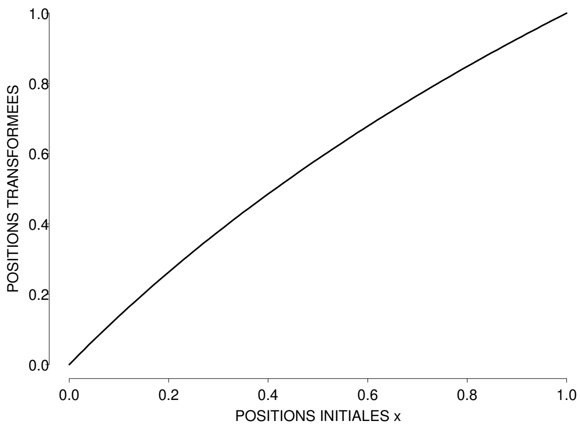


FIG. 4.7 – *La déformation* $\dfrac{\ln(1+x)}{\ln(2)}$

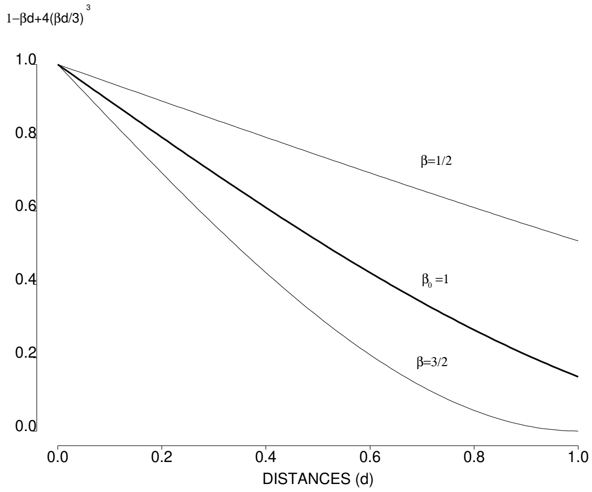


FIG. 4.8 – *Trois covariances sphériques avec* $0 < \beta \leq \frac{3}{2}$, β_0 *est le paramètre de la vraie covariance stationnaire*

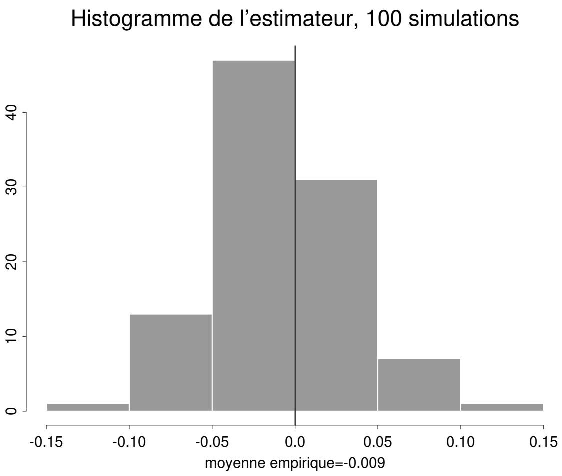


FIG. 4.9 – *Distribution empirique de* $\left(\hat{\beta}_n - E(\hat{\beta}_n)\right)$ *pour une covariance sphérique*

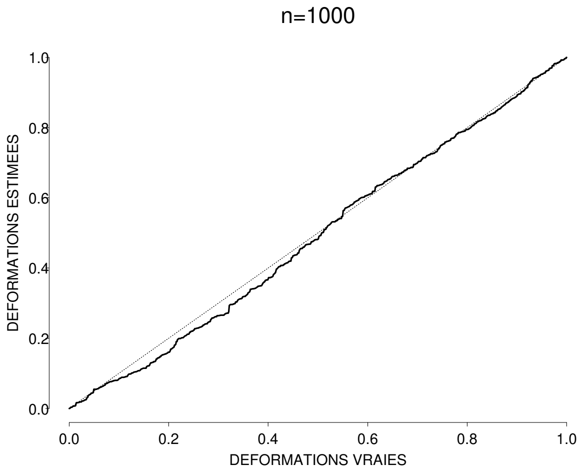


FIG. 4.10 – *Estimation de la déformation* $\dfrac{\ln(1+x)}{\ln(2)}$ *avec une partition régulière de pas* $\dfrac{1}{1000}$

Conclusion et perspectives

Nous n'allons pas reprendre ici l'ensemble des résultats obtenus dans les chapitres précédents. Rappelons simplement les objectifs fixés et en grande partie réalisés :

- caractériser les fonctions de covariance de processus non-stationnaires réductibles à une covariance stationnaire *via* une déformation de l'espace (chapitre 1) ;
- proposer différentes méthodes d'estimation de la déformation et de la covariance stationnaire d'un modèle de covariance non-stationnaire réductible, et obtenir des propriétés asymptotiques sur les estimateurs dans différentes situations :

  1. dans le chapitre 2, les paramètres de la déformation et de la covariance stationnaire sont de dimension finie. La méthode d'estimation est celle du minimum de contraste. On obtient les convergences $p.s.$ et en loi, ainsi que des propriétés asymptotiques à la base de tests statistiques de différence de contraste reposant sur des sommes pondérées de $khi - deux$ indépendants. On établit également un résultat d'identification $p.s.$ du modèle par contraste pénalisé ;
  2. dans le chapitre 3, le paramètre de la covariance stationnaire reste de dimension finie, mais la déformation appartient à une famille non-paramétrique d'homéomorphismes croissants. La méthode d'estimation est toujours celle du minimum de contraste, mais dans un cadre où l'espace des paramètres, de dimension infinie, est relativement compact. Ce type d'estimation non-paramétrique nécessite des hypothèses contraignantes sur la vitesse d'échantillonnage, plus précisément sur le nombre de points d'observation du processus et leurs instants d'apparition dans le plan d'expérience. Pour l'instant, seules les convergences en probabilité et $p.s.$ sont montrées ;
  3. dans le chapitre 4, le processus est gaussien. Cette hypothèse nous permet d'identifier la déformation non-paramétrique et le paramètre d'échelle de la covariance stationnaire à partir d'une seule réalisation du processus. La méthode d'estimation fait appel aux variations quadratiques et trouve sa justification dans la relation existant entre la fonction de singularité, le paramètre d'échelle et la déformation. On obtient les convergences $p.s$ et en loi pour le paramètre d'échelle. La convergence uniforme $p.s.$ est établie pour la déformation.

Dans les chapitre 2, 3 et 4, des études par simulation corroborent les résultats théoriques et donnent une idée de la qualité des estimateurs.

Outre la mise en œuvre des méthodes proposées à des cas concrets et leur application à des problèmes, tels que la prédiction d'un phénomène physique et l'amélioration d'un réseau de stations de mesure, ce travail ouvre des pistes de recherche intéressantes sur l'étude de processus spatiaux réductibles, par déformation de l'espace, à des processus stationnaires, et plus précisément sur :

- la caractérisation des covariances non-stationnaires réductibles à une covariance stationnaire *via* une déformation de l'espace dans $\mathbb{R}^d$, $d \geq 2$. On pourrait envisager la même démarche que celle adoptée dans le cas unidimensionnel du chapitre 1. Cela nous conduirait à une caractérisation sous forme d'un système d'équations différentielles qu'il faudrait ensuite exploiter pour des familles particulières de déformations ;
- la description paramétrique, ou non-paramétrique, des "bonnes familles" de déformations spatiales bijectives dans un premier temps. Dans le cas d'une description non-paramétrique d'une déformation bijective de $\mathbb{R}^2$ dans $\mathbb{R}^2$, on pourrait prendre, par exemple, la famille des déformations de réseaux triangulaires avec contrainte de non-repliement. Cette famille serait la "généralisation" des déformations lignes brisées avec contrainte de monotonie utilisées au chapitre 3. La mise en œuvre d'algorithmes d'estimation sous contrainte de telles déformations serait à étudier dans un deuxième temps ;
- la définition et l'utilisation de gammes plus riches de variations "quadratiques" sur le plan afin de tirer le maximum d'informations sur la covariance non-stationnaire et de mieux identifier le modèle. On sait en effet, d'après Guyon (1987), que les variations dans le plan représentent un outil d'identification, en moyenne quadratique, plus riche que sur la droite. Comment ces variations nous permettraient d'identifier, au sens *p.s.*, la déformation et les paramètres de la covariance stationnaire de notre modèle de covariance non-stationnaire ?

Annexes

Annexe A

Calcul de la variance de l'estimateur empirique de la covariance

On rappelle que pour tout $T \geq 2$ et tout $i \leq j$:

$$\hat{r}_T(x_i, x_j) = \frac{1}{T-1} \sum_{t=1}^{T} \left(Z_t(x_i) - \bar{Z}_T(x_i)\right) \left(Z_t(x_j) - \bar{Z}_T(x_j)\right) \tag{A.1}$$

avec, pour tout i dans $[1, \ldots, n]$:

$$\bar{Z}_T(x_i) = \frac{1}{T} \sum_{t=1}^{T} Z_t(x_i)$$

(i) On vérifie tout d'abord que, pour tout $i \leq j$:

$$E\hat{r}_T(x_i, x_j) = r_Z(x_i, x_j)$$

c'est-à-dire que l'estimateur empirique de la covariance est sans biais.

Posons, pour tout $1 \leq t \leq T$ et tout x dans $[0, 1]$:

$$Y_t(x) = Z_t(x) - M_0$$

Pour tout x, $(Y_1(x), Y_2(x), \ldots, Y_T(x))$ constitue une suite *i.i.d.* et centrée de la *v.a.r.* parente Y. L'égalité (A.1) devient, pour tout $i \leq j$:

$$\hat{r}_T(x_i, x_j) = \frac{1}{T-1} \sum_{t=1}^{T} Y_t(x_i) Y_t(x_j) - \frac{T}{T-1} \bar{Y}_T(x_i) \bar{Y}_T(x_j) \tag{A.2}$$

On vérifie :

$$\left\{ \begin{array}{rcl} EY(x_i)Y(x_j) & = & r_Z(x_i, x_j) \\ E\bar{Y}_T(x_i)\bar{Y}_T(x_j) & = & \frac{1}{T} r_Z(x_i, x_j) \end{array} \right.$$

Soit, pour tout $i \leq j$:

$$E\hat{r}_T(x_i, x_j) = r_Z(x_i, x_j)$$

(ii) Revenons à l'étude de la variance de notre estimateur. On a la décomposition classique :

$$Var\left(\hat{r}_T(x_i, x_j)\right) = E\hat{r}_T^2(x_i, x_j) - \left(E\hat{r}_T(x_i, x_j)\right)^2$$

A partir de (A.2), on obtient :

$$\begin{aligned} \hat{r}_T^2(x_i, x_j) &= \frac{1}{(T-1)^2} \left\{ \left[\sum_{t=1}^T Y_t(x_i) Y_t(x_j) \right]^2 \right. \\ &- \frac{2}{T} \sum_{t=1}^T Y_t(x_i) \sum_{t'=1}^T Y_{t'}(x_j) \sum_{t''=1}^T Y_{t''}(x_i) Y_{t''}(x_j) \\ &+ \left. \frac{1}{T^2} \left[\sum_{t=1}^T Y_t(x_i) \right]^2 \left[\sum_{t'=1}^T Y_{t'}(x_j) \right]^2 \right\} \end{aligned}$$

Analysons l'espérance de chacun des trois termes du membre de droite de l'égalité précédente :

1.
$$E\left[\sum_{t=1}^T Y_t(x_i)Y_t(x_j)\right]^2 = TEY^2(x_i)Y^2(x_j) + E\sum_{t\neq t'} Y_t(x_i)Y_t(x_j)Y_{t'}(x_i)Y_{t'}(x_j)$$

avec :

$$\begin{aligned} EY^2(x_i)Y^2(x_j) &= Cov(Y^2(x_i), Y^2(x_j)) + EY^2(x_i)EY^2(x_j) \\ &= Cov(Y^2(x_i), Y^2(x_j)) + \sigma_0^4 \end{aligned}$$

et :

$$\begin{aligned} E\sum_{t\neq t'} Y_t(x_i)Y_t(x_j)Y_{t'}(x_i)Y_{t'}(x_j) &= \sum_{t\neq t'} Cov\left(Y_t(x_i)Y_t(x_j), Y_{t'}(x_i)Y_{t'}(x_j)\right) \\ &+ \sum_{t\neq t'} E\left[Y_t(x_i)Y_t(x_j)\right] E\left[Y_{t'}(x_i)Y_{t'}(x_j)\right] \\ &= T(T-1)r_Z^2(x_i, x_j) \end{aligned}$$

Donc :

$$E\left[\sum_{t=1}^T Y_t(x_i)Y_t(x_j)\right]^2 = T\left(Cov(Y^2(x_i), Y^2(x_j)) + \sigma_0^4\right) + T(T-1)r_Z^2(x_i, x_j)$$

2.

$$E\sum_{t=1}^{T}Y_t(x_i)\sum_{t'=1}^{T}Y_{t'}(x_j)\sum_{t''=1}^{T}Y_{t''}(x_i)Y_{t''}(x_j)=\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{t''=1}^{T}EY_t(x_i)Y_{t'}(x_j)Y_{t''}(x_i)Y_{t''}(x_j)$$

avec :

$$EY_t(x_i)Y_{t'}(x_j)Y_{t''}(x_i)Y_{t''}(x_j)=\begin{cases} Cov(Y^2(x_i),Y^2(x_j))+\sigma_0^4 & \text{si } t=t'=t'' \\ r_Z^2(x_i,x_j) & \text{si } t=t'\neq t'' \\ 0 & \text{sinon}\end{cases}$$

Donc :

$$\begin{aligned} E\sum_{t=1}^{T}Y_t(x_i)\sum_{t'=1}^{T}Y_{t'}(x_j)\sum_{t''=1}^{T}Y_{t''}(x_i)Y_{t''}(x_j) &= T\left(Cov(Y^2(x_i),Y^2(x_j))+\sigma_0^4\right) \\ &+ T(T-1)r_Z^2(x_i,x_j)\end{aligned}$$

3.

$$\begin{aligned} E\left[\sum_{t=1}^{T}Y_t(x_i)\right]^2\times\left[\sum_{t=1}^{T}Y_t(x_i)\right]^2 &= E\left\{\left(\sum_{t=1}^{T}Y_t^2(x_i)+\sum_{t\neq t'}Y_t(x_i)Y_{t'}(x_i)\right)\right. \\ &\times \left.\left(\sum_{l=1}^{T}Y_l^2(x_j)+\sum_{l\neq l'}Y_l(x_j)Y_{l'}(x_j)\right)\right\}\end{aligned}$$

avec :

$$\begin{cases} E\sum_{t=1}^{T}Y_t^2(x_i)\sum_{l=1}^{T}Y_l^2(x_j) &= T\left(Cov(Y^2(x_i),Y^2(x_j))+\sigma_0^4\right)+T(T-1)\sigma_0^4 \\ E\sum_{t=1}^{T}Y_t^2(x_i)\sum_{l\neq l'}Y_l(x_j)Y_{l'}(x_j) &= 0 \\ E\sum_{l=1}^{T}Y_l^2(x_j)\sum_{t\neq t'}Y_t(x_i)Y_{t'}(x_i) &= 0 \\ E\sum_{t\neq t'}Y_t(x_i)Y_{t'}(x_i)\sum_{l\neq l'}Y_l(x_j)Y_{l'}(x_j) &= 2T(T-1)r_Z^2(x_i,x_j)\end{cases}$$

Donc :

$$\begin{aligned} E\left[\sum_{t=1}^{T}Y_t(x_i)\right]^2\left[\sum_{t=1}^{T}Y_t(x_i)\right]^2 &= T\left(Cov(Y^2(x_i),Y^2(x_j))+\sigma_0^4\right) \\ &+ T(T-1)(\sigma_0^4+2r_Z^2(x_i,x_j)\end{aligned}$$

De 1., 2. et 3., on déduit :

$$\begin{aligned}\hat{r}_T^2(x_i,x_j) &= \frac{1}{T(T-1)}\left\{(T-1)\left(Cov(Y^2(x_i),Y^2(x_j))+\sigma_0^4\right)\right.\\ &+ \left.(T^2-2T+2)r_Z^2(x_i,x_j)+\sigma_0^4\right\}\end{aligned}$$

Comme :

$$\begin{aligned}E(\hat{r}_T(x_i,x_j)) &= r_Z(x_i,x_j)\\ Cov(Y^2(x_i),Y^2(x_j))+\sigma_0^4 &= EY^2(x_i)Y^2(x_j)\end{aligned}$$

on a finalement :

$$Var(\hat{r}_T(x_i,x_j)) = \frac{1}{T(T-1)}\left\{(T-1)E(Y_t^2(x_i)Y_t^2(x_j))+\sigma_0^4+(2-T)r_Z^2(x_i,x_j)\right\}$$

■

Annexe B

Distribution d'une somme pondérée de $khi - deux$ indépendants

Si $\{W_j\}_{1 \leq j \leq n}$ est une suite de n variables normales $\mathcal{N}_1(0,1)$ indépendantes, nous appelons "une somme pondérée de $khi - deux$ indépendants", la forme quadratique suivante :

$$Q(W) = \sum_{j=1}^{n} \lambda_j W_j^2$$

où $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$ sont des réels strictement positifs. Sa fonction caractéristique est :

$$\phi_n(t) = E\left(e^{itQ(W)}\right) = \prod_{j=1}^{n} (1 - 2it\lambda_j)^{-\frac{1}{2}}$$

La densité $p_n(y)$ et la fonction de répartition $F_n(y)$ de $Q(W)$ ne posèdent pas une forme explicite. Toutefois, $p_n(y)$ et $F_n(y)$ peuvent être données formellement sous la forme d'un développement en polynômes de Laguerre. Des possiblités d'approximation numérique en résultent (Jonhson et Kotz (1970), pp.162-163)). Signalons également qu'il existe des tables donnant les fractiles usuels de $F_n(y)$ pour $2 \leq n \leq 10$ et une grille de valeurs des facteurs $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ (Shah (1986)).

Annexe C

Les fonctions I-splines : construction pratique

Présentation générale

L'ensemble des I-splines I_k^l, $k = 1, \ldots, p$, $p \in \mathbb{N}^*$, de degré l, $l \in \mathbb{N}^*$, définies sur l'intervalle $[0, 1]$ et associées aux $m = p - l > 0$ nœuds :

$$0 < a_1 < a_2 < \ldots < a_m < 1$$

est une base de l'espace des splines polynomiales sur $[0, 1]$, sous-espace hilbertien de dimension p de l'espace de Sobolev $\mathcal{H}^l([0, 1])$:

$$\mathcal{H}^l([0,1]) = \left\{ f \in C^{l-1}([0,1]), f^l \in L^2([0,1]) \right\}$$

Les I-splines sont, entre autres, utilisées dans des problèmes d'estimation de courbes de croissance (Champely et Deleuze (1995)), ou bien dans des problèmes de régression dans lesquels les données possèdent des structures de monotonie évidentes (Ramsay (1988)).

Construction

Il y a plusieurs façons de construire les p fonctions I-splines I_k^l. Celle que nous présentons ci-après est décrite de manière plus détaillée dans Ramsay (1988, pp.425-429).

1. On se donne tout d'abord la séquence de points suivante :

$$0 = x_1 = \cdots = x_l < x_{l+1} = a_1 < \cdots < x_p = a_m < x_{p+1} = \cdots = x_{p+l} = 1$$

2. On construit ensuite, de manière récurrente, les fonctions M-splines M_k^l, splines polynomiales de degré $l-1$:

   – à l'étape 1, pour tout $1 \leq k \leq (p+l-1)$ on définit les M-splines M_k^1 de degré 0 de la manière suivante :

$$\begin{cases} \text{si } x_{k+1} > x_k, \text{ alors } M_k^1(x) = \begin{cases} \dfrac{1}{x_{k+1}-x_k} & \text{si } x_k \leq x < x_{k+1} \\ 0 & \text{sinon} \end{cases} \\ \text{si } x_{k+1} = x_k, \text{ alors } M_k^1(x) = 0 \end{cases}$$

   – à l'étape l, pour tout $1 \leq k \leq p$ les M-splines M_k^l sont données par :

$$M_k^l(x) = \frac{l\left[(x-x_k)M_k^{l-1}(x) + (x_{k+l}-x)M_{k+1}^{l-1}(x)\right]}{(l-1)(x_{k+l}-x_k)}$$

   Les M-splines M_k^l sont positives et d'intégrale 1 sur $[x_k, x_{k+l}]$, elles se comportent donc comme des densités de probabilité.

3. À partir de ces M-splines, les fonctions I-splines de degré l sont définies pour tout x dans $[0,1]$ par :

$$I_k^l(x) = \int_0^x M_k^l(v)dv$$

Propriétés

Par construction, une fonction I-spline I_k^l sur un intervalle $[0,1]$ se comporte comme une fonction de répartition, plus précisément :

- elle est continue ;
- elle est croissante ;
- elle vaut 0 en 0 et 1 en 1.

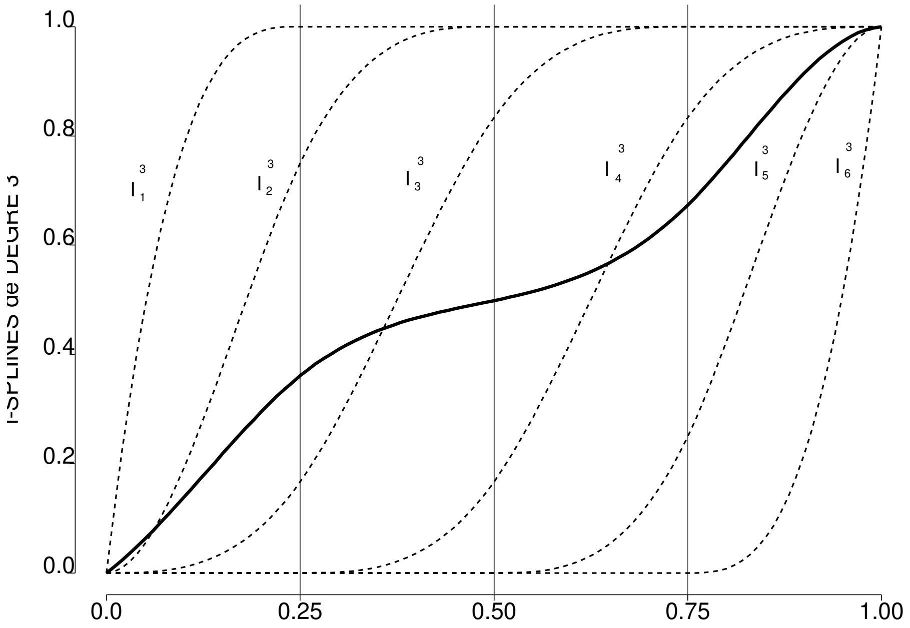


FIG. C.1 – *Les six fonctions I-splines (en pointillé) de degré 3 définies sur* $[0, 1]$ *associées aux nœuds intérieurs* 0.25, 0.50 *et* 0.75 *et leur combinaison convexe, fournissant* $0.09I_1^3 + 0.35I_2^3 + 0.05I_3^3 + 0.10I_4^3 + 0.40I_5^3 + 0.01I_6^3$, *représentée par la courbe en trait fort*

Annexe D

Convergence *p.s.* des estimateurs du minimum de contraste

Le théorème suivant est une généralisation du théorème portant sur la convergence *p.s.* des estimateurs du minimum de contraste de Hardouin ((1992), pp.6-8). La modification porte sur le fait que l'espace des paramètres considéré ici est un ensemble relativement compact dans un espace normé. On ne fait aucune hypothèse sur la dimension de cet espace qui peut être finie ou infinie. On note néanmoins que les espaces de paramètres de dimension infinie contenant des boules ouvertes non vides ne rentrent pas dans notre cadre d'étude. En effet, d'après Brézis ((1983), p.92), ces boules ouvertes ne sont jamais relativement compactes.

Soit $\mathcal{X}$ un espace normé et soit d la distance associée à la norme. Soit Y un processus défini sur un ensemble infini dénombrable S, observé sur une suite (D_n) de domaines finis, croissante vers S, et $Y(n) = \{Y_i, i \in D_n\}$. Considérons le modèle statistique, paramétrique ou non, $(\Omega, \mathcal{A}, P_f)$ où $f \in \mathcal{F} \subset \mathcal{X}$. Nous travaillerons par la suite dans $\bar{\mathcal{F}}$ muni de la topologie trace, induite par celle de $\mathcal{X}$.

Précisons que pour l'estimation par la méthode du maximum de vraisemblance dans des ensembles de dimension infinie, il existe de nombreux résultats. Par exemple, dans Bahadur ((1967), p.320), on trouve un théorème sur la convergence en probabilité d'un estimateur du maximum de vraisemblance dans un espace de dimension infinie. En fait, ce résultat est en filigrane dans Wald ((1949), p.601). Dans tous les cas, le point essentiel est que l'ensemble des paramètres doit vérifier une hypothèse de relative compacité.

$\star$ Soit f_0 dans $\mathcal{F}$, la vraie valeur du paramètre inconnu.

$\star\star$ Une **fonctionnelle de contraste** pour f_0 est une application :

$$\begin{array}{rcl} K : \bar{\mathcal{F}} & \longrightarrow & \mathbb{R}^+ \\ f & \longmapsto & K(f_0, f) \end{array}$$

ayant un minimum unique dans $\bar{\mathcal{F}}$ (minimum que l'on peut supposer nul) en $f = f_0$.

$\star\star\star$ Un **processus de contraste** pour K est une suite de fonctionnelles aléatoires :

$$\begin{array}{rcl} U_n : \bar{\mathcal{F}} & \longrightarrow & \mathbb{R}^+ \\ f & \longmapsto & U_n(f) \end{array}$$

adaptées à $Y(n)$ et vérifiant :

$$\liminf_{n \longrightarrow \infty} (U_n(f) - U_n(f_0)) \geq K(f_0, f), \; P_{f_0} - p.s. \tag{D.1}$$

On appelle ϵ_n-**estimateur du minimum de contraste**, s'il existe, toute suite aléatoire de fonctions $\hat{f}_n$ appartenant à $\mathcal{F}$ et vérifiant :

$$U_n(\hat{f}_n) \leq \inf_{f \in \mathcal{F}} U_n(f) + \epsilon_n$$

avec (ϵ_n) une suite tendant vers 0, $n \to +\infty$.

Le module de continuité d'une fonction f d'un espace métrique (A, d_A) à valeurs dans un autre espace métrique (B, d_B), est noté :

$$W_f(\gamma) = \sup \left\{ d_B(f(x), f(y)), d_A(x, y) \leq \gamma, (x, y) \in A^2 \right\}$$

Soient les hypothèses :

- **V1** $\mathcal{F}$ est relativement compact
- **V2** $f \longmapsto K(f_0, f)$ est continue dans $\bar{\mathcal{F}}$
- **V3** $f \longmapsto U_n(f)$ est continue $P_{f_0} - p.s.$ dans $\bar{\mathcal{F}}$
- **V4** il existe une suite (η_k) qui tend vers 0, $k \to +\infty$, telle que pour tout $k \geq 1$:

$$P_{f_0} \left(\limsup_{n \longrightarrow +\infty} \left\{ W_{U_n} \left(\frac{1}{k} \right) + \epsilon_n \geq \eta_k \right\} \right) = 0$$

Sous les hypothèses **V1** et **V3**, l'existence d'un ϵ_n-estimateur du minimum de contraste est assurée, car $\bar{\mathcal{F}}$ compact entraîne U_n uniformément continue.

***Théorème* D.1** *Soit $\hat{f}_n$ un ϵ_n-estimateur du minimum de contraste, alors :*

$$d(\hat{f}_n, f_0) \underset{n \to +\infty}{\longrightarrow} 0, \ P_{f_0} - p.s.$$

★

Démonstration :

- Prenons un ensemble F contenu dans $\mathcal{F}$ dénombrable et dense dans $\mathcal{F}$, c'est-à-dire $\bar{F} = \bar{\mathcal{F}}$. L'existence d'un tel ensemble est garantie car tout ensemble relativement compact dans un espace métrique est séparable (Trénoguine (1985), p.221). On déduit alors de la continuité *p.s.* des U_n que les variables aléatoires :

  $$\inf_{f \in \mathcal{F}} U_n(f) = \inf_{f \in F} U_n(f)$$

  et

  $$W_{U_n}(\gamma) = \sup \left\{ |U_n(f) - U_n(g)|, d(f,g) \leq \gamma, (f,g) \in F^2 \right\}$$

  sont mesurables pour tout n.

- Puisque $K(f_0, .)$ est continue dans $\bar{\mathcal{F}}$ et admet un minimum unique en f_0, alors :

  > $\forall r > 0, \exists \rho_r > 0$ tel que si B_r est une boule ouverte dans $\bar{\mathcal{F}}$ (pour la topologie trace), centrée en f_0 et de rayon r :

  $$\inf_{\mathcal{F} \setminus B_r} K(f_0, f) \geq \inf_{\bar{\mathcal{F}} \setminus B_r} K(f_0, f) \geq 2\rho_r$$

  $\bar{\mathcal{F}} \setminus B_r$ est fermé dans $\bar{\mathcal{F}}$ compact, donc compact. Prenons alors un recouvrement de $\mathcal{F} \setminus B_r$ par des boules $B\left(f, \dfrac{1}{k}\right), f \in F \setminus B_r$ et k tel que $\eta_k < \rho_r$. Il recouvre aussi $\bar{\mathcal{F}} \setminus B_r$, donc on peut en tirer un sous-recouvrement fini de $N = N(r,k)$ boules $B\left(f_i, \dfrac{1}{k}\right), 1 \leq i \leq N$.

  On a :

  $$\left\{ \hat{f}_n \notin B_r \right\} \subset \left\{ \inf_{f \in \mathcal{F} \setminus B_r} U_n(f) \leq U_n(f_0) + \epsilon_n \right\}$$

En utilisant le sous-recouvrement fini de $\mathcal{F} \setminus B_r$ et le module de continuité de U_n, on a :

$$\begin{aligned}\inf_{f\in\mathcal{F}\setminus B_r}(U_n(f)-U_n(f_0)) &= \inf_{1\le i\le N}\left(\inf_{f\in B(f_i,\frac{1}{k})}(U_n(f)-U_n(f_0))\right)\\ &\ge \inf_{1\le i\le N}\left(U_n(f_i)-U_n(f_0)-W_{U_n}\left(\frac{1}{k}\right)\right)\end{aligned}$$

Soit :

$$\left\{\hat{f}_n\notin B_r\right\}\subset\left\{\inf_{1\le i\le N}(U_n(f_i)-U_n(f_0))\le W_{U_n}\left(\frac{1}{k}\right)+\epsilon_n\right\}$$

– De (D.1) on déduit la minoration suivante :

$$\begin{aligned}\liminf_{n\to+\infty}\left(\inf_{1\le i\le N}(U_n(f_i)-U_n(f_0))\right) &= \inf_{1\le i\le N}\left(\liminf_{n\to+\infty}(U_n(f_i)-U_n(f_0))\right)\\ &\ge \inf_{1\le i\le N}K(f_0,f_i)\ge 2\rho_r\end{aligned}$$

Par conséquent, pour tout $r>0$:

$$P_{f_0}(C_r)=1,\ \text{où}\ C_r=\left\{\omega:\liminf_{n\to+\infty}\left(\inf_{1\le i\le N}(U_n(f_i)-U_n(f_0))\right)\ge 2\rho_r\right\}$$

– L'hypothèse **V4** se traduit, pour tout $k\ge 1$, par :

$$P_{f_0}(D_k)=0,\ \text{où}\ D_k=\left\{\omega:W_{U_n}\left(\frac{1}{k}\right)+\epsilon_n\ge\eta_k\ (i.s.)\ \text{en n}\right\}$$

– En résumé :

$\forall r>0$, k tel que $\eta_k<\rho_r$ et $\omega\in C_r\cap D_k^c$, alors il existe $N'=N(\omega,r,k)$ tel que, pour $n\ge N'$:

$$\inf_{1\le i\le N}(U_n(f_i)-U_n(f_0))\ge\rho_r>\eta_k>W_{U_n}\left(\frac{1}{k}\right)+\epsilon_n$$

Ce qui signifie :

$$\begin{aligned}\left\{\hat{f}_n\notin B_r\ (i.s.)\right\} &\subseteq \left\{\inf_{1\le i\le N}(U_n(f_i)-U_n(f_0)\le W(U_n,\frac{1}{k})+\epsilon_n\ (i.s.)\right\}\\ &\subseteq (C_r\cap D_k^c)^c\end{aligned}$$

c'est-à-dire :

$$P_{f_0}\left(\hat{f}_n\notin B_r\ (i.s.)\right)=0$$

Ceci étant vrai pour toute boule B_r centrée en f_0 et de rayon $r>0$, on en déduit :

$$P_{f_0}(\lim_{n\to+\infty}d(\hat{f}_n,f_0)=0)=1$$

■

Remarque : la convergence en probabilité

Si on remplace (D.1) et **V4** par leur version en probabilité :

$$\liminf_{n \longrightarrow \infty} (U_n(f) - U_n(f_0)) \geq K(f_0, f), \text{ en } P_{f_0} - \text{probabilité} \tag{D.2}$$

– **V4'**: $\lim_{n \to +\infty} P_{f_0} \left(W_{U_n} \left(\frac{1}{k} \right) + \epsilon_n \geq \eta_k \right) = 0$, pour tout $k \geq 1$

les hypothèses **V1**-**V3** restant identiques, on a le résultat :

***Théorème* D.2** *Sois $\hat{f}_n$ un ϵ_n-estimateur du minimum de contraste, alors :*

$$d(\hat{f}_n, f_0) \underset{T \to +\infty}{\longrightarrow} 0, \ \textit{en } P_{f_0} - \textit{probabilité}$$

★

Références bibliographiques :

Akaïke H. (1969). Statistical predictor identification. *Annals of the Institute of Statistical Mathematics*, **22**, 203-217.

Assouad P. (1980). Plongements isométriques dans L^1 : aspect analytique. Séminaire d'Initiation à l'Analyse, **14**, 1-23.

Bahadur R.R. (1967). Rates of convergence of estimates and test statistics. *Annals of Mathematical Statistics*, **38**, 303-324.

Baxter G. (1956). A strong limit theorem for Gaussian processes. *Proceedings of the American Mathematical Society*, **7**, 522-527.

Bayomog S., Guyon X., Hardouin C., Yao J.F. (1996). Test de différence de contraste et somme pondérée de $khi - deux$. *The Canadian Journal of Statistics*, **24**, 115-130.

Birge L., Massart P. (1993). Rates of convergence for minimum contrast estimators. *Probability Theory and Related Fields*, **97**, 113-150.

Brezis H. (1987). *Analyse fonctionnelle : théorie et applications.* Masson.

Brillinger D.R. (1962). A note on the rate of convergence of a mean. *Biometrika*, **49**, 574-576.

Champely S., Deleuze Ch. -1995. Co-inertie de deux variables aléatoires hilbertiennes et aproximation B-spline. Application en foresterie. *Revue de Statistique Appliquée*, **XLIII**, 91-107.

Choquet G. (1992). *Cours de topologie.* Masson.

Dacunha-Castelle D., Duflo M. (1993). *Probabilités et statistiques 2 : problèmes à temps mobile.* Masson, 2nd éd.

Dieudonné J. (1972). *Éléments d'analyse, tome 1 : fondements de l'analyse moderne.* Gauthier-Villars.

Doukhan P. (1994). *Mixing, properties and examples.* Lecture Notes in Statistics, **85**, Springer.

Durret R. (1996). *Probability : theory and examples.* Duxbury Press, 2nd ed.

Grenander, U. (1981). *Abstract Inference.* Wiley.

Guikhman I., Skorokhod A. (1980). *Introduction à la théorie des processus aléatoires.* Éds Mir.

Guttorp P., Sampson P.D. (1994). Methods for estimating heterogeneous spatial covariance functions with environmental applications. *In* Patil G.P. and C.R. Rao eds, *Handbook of statistics XII : environmental statistics*, Elsevier/North Holland, 663-690.

Guyon X., Leon J. (1989). Convergence en loi des H-variations d'un processus gaussien stationnaire sur $\mathbb{R}$. *Annales de l'Institut Henri Poincaré*, **25**, 265-282.

Guyon X. (1987). Variations de champs gaussiens stationnaires : application à l'identification. *Probability Theory and Related Fields*, **75**, 179-193.

Guyon X. (1993). *Champs aléatoires sur un réseau : modélisations, statistiques et applications.* Masson.

Hardouin-Ceccantini C. (1992). Quelques résultats nouveaux en statistique des processus : contraste fort, régression à résidus à longue portée et estimation par log-périodogramme. Thèse, Université Paris VII.

Hartman P., Wintner A. (1941). On the law of the iterated logarithm. *American Journal of Mathematics*, **63**, 169-176.

Hœffding W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, **58**, 13-30.

Istas J. (1996). Estimating the singularity fonction of a Gaussian process with applications. *Scandinavian Journal of Statistics*, **23**, pp.581-596.

Istas J., Lang G. (1997). Quadratic variations and estimation of the local Hölder index of a Gaussian process. A paraître dans *Annales de l'Institut Henri Poincaré.*

Jennrich R. I. (1969). Asymptotic properties of non-linear least squares estimators. *The Annals of Mathematical Statistics*, **40**, 633-643.

Johnson N.L., Kotz S. (1970). *Continuous univariate distributions-2.* Houghton Mifflin Company.

Journel A.G., Huijbregts Ch.J. (1978). *Mining Geostatistics.*

Krein M. (1940). Sur le problème du prolongement des fonctions hermitiennes positives et continues. *Comptes Rendus (Doklady) de l'Académie des Sciences de l'U.R.S.S.*, **XXVI**.

Lelong-Ferrand J., Arnaudiès J. M. (1977). *Cours de mathématiques, tome 2 : analyse.* Dunod Université, 4ème édition.

Loève M. (1978). *Tome 2 : probability theory.* Springer-Verlag, 4th ed.

Luenberger L. (1979). *Introduction to dynamical systems : theory and applications.* Wiley.

Mardia K.V., Goodall C. R. (1993). *Spatial-temporal analysis of multivariate environmental monitoring data. In* G.P. Patil and C.R. eds, *Multivariate environmental statistics*, Elsevier Science Publishers, 347-386.

Matheron, G. (1965). *Les variables régionalisées et leur estimation.* Masson.

Matheron, G. (1969). Le krigeage universel. *Cahiers du Centre de Morphologie Mathématique*, **1**, Fontainebleau, France.

Matheron, G. (1973). The intrinsic random functions and their applications. *Advances in Applied Probability*, **5**, 439-468.

Meiring W. (1995). Estimation of Heterogeneous Space-Time Covariance. Thèse, Université du Washington.

Monestiez P., Sampson P.D., Guttorp P. (1993). Modelling of heterogeneous spatial correlation structure by spatial deformation. *Cahiers de Géostatistique*, Fasicule **3**, Compte Rendu des Journées de Géostatistique, Fontainebleau. Publié par l'Ecole Nationale Supérieure des Mines de Paris, 1-12.

Monfort A. (1980). *Cours de probabilités.* Economica.

Neveu J. (1980). *Bases mathématiques du calcul des probabilités.* Masson, 2nd éd.

Priestley M.B. (1981). *Spectral analysis and time series.* Academic Press.

Ramsay J.O. (1988). Monotone regression splines. *Statistical Science*, **3**, 425-461.

Rudin W. (1962). The extension problem for positive-definite functions. *Illinois Journal of Mathematics*, **7**, 532-539.

Rudin W. (1990). *Fourier Analysis on Groups.* Wiley Classics Library.

Sacks J., Ylvisaker D. (1966). Designs for regression problems with correlated errors. *The Annals of Mathematical Statistics*, **37**, 66-89.

Sampson P.D. (1986). Spatial covariance estimation by scaled-metric scaling and biorthogonal grids. *Technical Report*, Department of Statistics of the University of Washington, **91**, 1-40.

Sampson P.D., Guttorp P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association*, **87**, 108-119.

Schœnberg I.J. (1938a). Metric spaces and positive definite functions. *Transactions of the American Mathematical Society*, **44**, 522-536.

Schœnberg I.J. (1938b). Metric spaces and completely monotone functions. *Annals of Mathematics*, **39**, 811-841.

Searle S. R. (1982). *Matrix algebra useful for statistics.* Wiley.

Senoussi R. (1990). Statistique asymptotique presque sûre de modèles statistiques convexes. *Annales de l'Institut Henri Poincaré*, **26**, 19-44.

Shah B.K. (1986). The distribution of positive quadric forms. *Selected Tables in Mathematical Statistics*, **10**, IMS-AMS Providence ed.

Spector P. (1994). *An introduction to S and S-plus.* Duxbury Press.

Stock J. H. (1988). Estimating continuous-time processes subject to time deformation. *Journal of the American Statistical Association*, **83**, 77-85.

Stout W.F. (1974). *Almost sure convergence.* Academic Press.

Stewart J. (1976). Positive definite functions and generalizations, an historical survey. *Rocky Mountain, Journal of Mathematics*, **6**, 409-434.

Wald A. (1949). Note on the consistency of the maximum likelihood estimate. *Annals of Mathematical Statistics*, **20**, 595-601.

Wong H.W., Severini T.A. (1991). On maximum likelihood estimation in infinite dimensional parameter spaces. *The Annals of Statistics*, **19**, 603-632.

Yadrenko M.I. (1983). *Spectral theory of random fields.* Optimization Software, Inc., Publication Division.

Yaglom, A.M. (1987). *Correlation theory of stationary and related random functions, tome 1 : basics results.* Springer-Verlag.

Yaglom, A.M. (1963). On the equivalence or perpendicularity of two Gaussian probability measures in function space. M. Rosenblatt, Ed., *Proceedings of the Symposium on Time Series Analysis*, 327-346.