

PHILIPPE BESSE

Approximation spline de l'analyse en composantes principales d'une variable aléatoire hilbertienne

Annales de la faculté des sciences de Toulouse 5^e série, tome 12, n° 3 (1991), p. 329-349

http://www.numdam.org/item?id=AFST_1991_5_12_3_329_0

© Université Paul Sabatier, 1991, tous droits réservés.

L'accès aux archives de la revue « Annales de la faculté des sciences de Toulouse » (<http://picard.ups-tlse.fr/~annaales/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

Approximation spline de l'analyse en composantes principales d'une variable aléatoire hilbertienne

PHILIPPE BESSE⁽¹⁾

RÉSUMÉ. — L'analyse en composantes principales d'une variable aléatoire à valeurs dans un espace hilbertien séparable est définie dans le cadre fonctionnel abstrait des fonctions splines d'interpolation ou d'ajustement. La description statistique de divers types de processus est alors obtenue en particulierisant les espaces hilbertiens considérés et le formalisme adopté permet de traiter les problèmes d'approximation par discrétisation et échantillonnage, et ainsi de convergence dans le cas général.

MOTS-CLÉS : Analyse en composantes principales, variable aléatoire hilbertienne, fonctions splines.

ABSTRACT. — Principal components analysis of an hilbertian random variable is defined in the abstract functional framework of interpolating and smoothing spline functions. The statistical description of different kinds of processes is then obtained by particularizing the Hilbert spaces to be considered. This theoretical approach allows to deal with the approximation by discretization and sampling and furthermore to solve the problems of convergence in the general case.

KEY-WORDS : Principal components analysis, hilbertian random variable, spline functions.

1. Introduction

L'objet de cet article est de présenter un cadre synthétique pour la description statistique par une analyse en composantes principales (a.c.p.) d'un processus ou, plus généralement, d'une variable aléatoire admettant un moment d'ordre 2 et prenant ses valeurs dans un espace hilbertien

⁽¹⁾ Laboratoire de Statistique et Probabilités, U.A. C.N.R.S. 745, Université Paul Sabatier, 31062 Toulouse cedex (France)

supposé séparable. Cette approche initiée par Besse [5] à la suite de l'étude synthétique des analyses factorielles proposée par Dauxois et Pousse [16] a suscité quelques développements et applications récemment parus (Saporta ([24] et [25]), Bouhaddou [13], Besse et Ramsay [9], Besse ([6], [7] et [8]), Winsberg [30], Ramsay et Dalzell [23]). De plus, Biritxinaga ([11] et [12]) s'intéresse à l'estimation de la moyenne d'un processus en utilisant le même type d'outils (fonctions splines).

L'a.c.p. d'une variable aléatoire hilbertienne (v.a.h) Y nécessite l'observation et donc la mesure de plusieurs réalisations $Y(\omega)$ sur un échantillon. Aussi, après avoir donné la définition de l'a.c.p. de Y dans un espace hilbertien (§ 2) il faut préciser celle-ci en fonction de la façon dont sont réalisées les mesures et en quoi elles permettent d'approcher les réalisations "vraies" de la variable aléatoire. Le formalisme adopté et rappelé au paragraphe 3 est celui introduit par Atteia [2] pour définir l'interpolation et l'ajustement (ou le lissage) par des fonctions splines dans le cadre abstrait d'espaces hilbertiens. Laurent [22] et Atteia [3] fournissent une description détaillée de ces outils.

Il est alors possible de prendre en compte diverses situations rencontrées dans l'étude descriptive d'un processus : les mesures sont des évaluations en un point ou un moyennage pondéré autour de ce point, elles sont ou non entachées d'erreur, conduisant ainsi à une interpolation ou un ajustement (lissage) des trajectoires de Y par des fonctions présentant certaines propriétés de régularité. Enfin, il faut choisir le produit scalaire utilisé pour l'a.c.p.. Il est à noter que cette approche descriptive et non prédictive ne requiert aucune hypothèse de stationnarité.

L'application de ces techniques permet donc de définir une approximation de l'a.c.p. de Y appelée approximation spline (§ 4) qui permet de retrouver les situations concrètes abordées dans la littérature en particulierisant les espaces considérés.

Enfin, un dernier paragraphe (§ 5) expose les conditions requises pour être assuré de la convergence et donc de la stabilité des ces approximations définies simultanément par une "discrétisation" abstraite et un échantillonnage.

Les abréviations et notations suivantes sont couramment utilisées :

- a.c.p. analyse en composantes principales,
- h.s. Hilbert-Schmidt (opérateur ou norme de),
- i.i.d. indépendantes identiquement distribuées,

- f.s.i. fonction spline d'interpolation,
- f.s.a. fonction spline d'ajustement ou de lissage,
- p.s. presque sûrement
- v.a.h.(r.) variable aléatoire hilbertienne (réelle);
- $\|\cdot\|_H, (\cdot | \cdot)_H, \langle \cdot, \cdot \rangle_{HH'}, \theta_H$ désignent respectivement la norme, le produit scalaire, la dualité et l'origine de l'espace hilbertien indiqué en indice;
- $\mathcal{L}(1, 2)$ est l'ensemble des applications linéaires continues de l'espace 1 dans l'espace 2;
- ${}^t\varphi, \varphi^*, \ker \varphi$ désignent respectivement l'application transposée, l'adjointe et le noyau de φ .

Tous les espaces hilbertiens considérés sont supposés séparables; l'espace bidual H'' est identifié à H .

2. Analyse en composantes principales d'une variable aléatoire hilbertienne

2.1. Notations

H est un espace hilbertien réel (le même travail pourrait être réalisé sur le corps des complexes), H' son dual topologique et \mathcal{B}_H la tribu des boréliens.

Soit Y une variable aléatoire définie sur l'espace probabilisé $(\Omega, \mathcal{A}, \mu)$ et à valeurs dans (H, \mathcal{B}_H) . On note $L^2(\mu)$ l'espace $L^2(\Omega, \mathcal{A}, \mu)$ des v.a.r. de carré intégrable muni du produit scalaire usuel. Cet espace est supposé séparable, ce qui est équivalent à dire que μ admet une base dénombrable (cf. [28]). L'isométrie canonique entre H et son dual H' (resp. $L^2(\mu)$ et $L^2(\mu)'$) est notée M (resp. D).

Dans tout l'article, Y est supposée être du second ordre, ceci s'écrit :

$$(H1) \ E\|Y\|_H^2 < \infty \text{ c'est-à-dire } Y \in L^2(\Omega, \mathcal{A}, \mu; H, \mathcal{B}_H) \text{ noté } L_H^2(\mu).$$

Dans ce cas, le théorème de Riesz assure l'existence d'un élément EY de H (l'espérance) et d'un opérateur V de $\mathcal{L}(H', H)$ (la covariance) définis de manière unique par les relations :

$$\begin{cases} \forall u \in H', & E\langle Y, u \rangle_{HH'} = \langle EY, u \rangle_{HH'}, \\ \forall u_1, u_2 \in H', & \\ E[\langle Y - EY, u_1 \rangle_{HH'} \langle Y - EY, u_2 \rangle_{HH'}] = \langle Vu_1, u_2 \rangle_{HH'}. & \end{cases} \quad (2.1)$$

On note par la suite $X = Y - EY$ la v.a.h. centrée associée à Y . L'opérateur de covariance est borné, positif et admet pour trace $E\|X\|_H^2$ qui est finie sous (H1) (cf. [20]); c'est un opérateur nucléaire.

De plus, si \mathcal{H} désigne un autre espace hilbertien, pour toute application linéaire continue Π de H dans \mathcal{H} ($\Pi \in \mathcal{L}(H, \mathcal{H})$), ΠY est une v.a.h. élément de $L^2_{\mathcal{H}}(\mu)$. Alors ΠY admet pour espérance $\Pi[EY]$ et pour covariance l'opérateur nucléaire :

$$\mathcal{V} = \Pi V^t \Pi. \tag{2.2}$$

2.2. Définition

Il existe de nombreuses façons équivalentes de définir une a.c.p. qui se ramènent toutes, lorsque (H1) est vérifiée, à l'analyse spectrale d'un opérateur compact. La procédure adoptée ici est une démarche pas à pas classique ([1]) mais d'autres approches globales conduiraient aux mêmes résultats.

Comme tout élément u de H' est continu, $u(X(\cdot))$ noté aussi $\langle X(\cdot), u \rangle_{HH'}$ est une v.a.r. sur $(\Omega, \mathcal{A}, \mu)$ et intégrable si (H1) est vérifiée car :

$$|\langle X(\cdot), u \rangle_{HH'}| \leq \|u\|_{H'} \|X(\cdot)\|_H. \tag{2.3}$$

Notons U l'application linéaire de H' dans $L^2(\mu)$ définie par

$$Uu = \langle X(\cdot), u \rangle_{HH'}. \tag{2.4}$$

PROPOSITION 2.1. — U défini ci-dessus est un opérateur de h.s. ainsi que l'application transposée tU définie par :

$$\forall f \in L^2(\mu)', \quad {}^tUf = E(XD^{-1}f).$$

De plus, la covariance V de X (ou de Y) se met sous la forme tUDU .

Preuve. — Si $\{\beta_i \mid i \in I\}$ désigne une base orthonormée de H' ,

$$\sum_{i \in I} \|U\beta_i\|_{L^2(\mu)}^2 = \sum_{i \in I} E\langle X, \beta_i \rangle_{HH'}^2 = E\|X\|_H^2 \leq \infty.$$

Alors, d'après un théorème de Guelfand et Vilenkin ([21] p. 30), U est de h.s. ainsi que sa transposée qui vérifie, $\forall f \in L^2(\mu)', \forall u \in H'$,

$$\begin{aligned} \langle Uu, f \rangle_{L^2(\mu)L^2(\mu)'} &= (Uu \mid D^{-1}f)_{L^2(\mu)} = E\left(\langle X, u \rangle_{HH'} D^{-1}f\right) \\ &= E\langle XD^{-1}f, u \rangle_{HH'} = \langle E(XD^{-1}f), u \rangle_{HH'} \\ &= \langle {}^tUf, u \rangle_{HH'}. \end{aligned}$$

De plus, $\forall u_1, u_2 \in H'$,

$$\begin{aligned} \langle {}^tUDUu_1, u_2 \rangle_{HH'} &= \langle E(\langle X, u_1 \rangle_{HH'} X), u_2 \rangle_{HH'} \\ &= E[\langle X, u_1 \rangle_{HH'} \langle X, u_2 \rangle_{HH'}] \\ &= \langle Vu_1, u_2 \rangle_{HH'}. \end{aligned}$$

Ce qui montre également que V est nucléaire comme produit de deux opérateurs de h.s. \square

Ceci se résume dans le diagramme commutatif ou "schéma de dualité" suivant :

$$\begin{array}{ccc} H & \xleftarrow{{}^tU} & L^2(\mu)' \\ \begin{array}{c} \downarrow M \\ \uparrow V \end{array} & & \uparrow D \\ H' & \xrightarrow{U} & L^2(\mu) \end{array}$$

Remarques

M est un opérateur borné mais pas nécessairement compact.

U est appelée la fonctionnelle aléatoire linéaire associée à X (cf. [4]).

DÉFINITION 2.1. — *On appelle a.c.p. linéaire centrée de Y dans H la recherche dans H' d'un élément u de norme 1 tel que Uu soit de norme maximale dans $L^2(\mu)$ et itération sous contraintes d'orthonormalité dans H' .*

Une "itération sous contrainte d'orthogonalité" est définie de façon récurrente :

Si les k premiers u_1, \dots, u_k ont pu être obtenus, alors le $k + 1^{\text{ème}}$ est solution de :

$$\max \left\{ \|Uu\|_{L^2(\mu)} \mid u \in H', \|u\|_{H'} = 1 \text{ et } \forall i \leq k, (u_i \mid u)_{H'} = 0 \right\}. \quad (2.5)$$

PROPOSITION 2.2. — *Si (H1) est vérifiée, l'a.c.p. linéaire centrée de Y dans H conduit à l'analyse spectrale de l'opérateur nucléaire, auto-adjoint et positif MV .*

Preuve. — En remplaçant dans (2.5) la norme par son carré, le problème équivalent s'écrit pour la première étape :

$$\begin{aligned} \max \{ & \|Uu\|_{L^2(\mu)}^2 \mid u \in H', \|u\|_{H'} = 1 \} = \\ & = \max \{ \langle Vu, u \rangle_{HH'} \mid u \in H', \|u\|_{H'} = 1 \} \\ & = \max \{ (MVu \mid u)_{H'} \mid u \in H', \|u\|_{H'} = 1 \}. \end{aligned}$$

Comme V est positif, nucléaire et M positif, borné, le produit MV est aussi positif et nucléaire. La solution est un vecteur propre u_1 de MV associé à la plus grande valeur propre λ_1 . De plus, d'après les propriétés spectrales d'un opérateur compact positif, le maximum est atteint pour les autres étapes avec u_k vecteur propre de MV associé au $k^{\text{ième}}$ élément de la suite pleine décroissante $\{\lambda_i \mid i \in I\}$ des valeurs propres. \square

2.3. Propriétés

Un certain nombre de résultats admis sans démonstration découlent d'autres propriétés de l'analyse spectrale d'un opérateur compact, ils permettent de retrouver les éléments classiques de l'a.c.p. standard.

- i) La suite $\{u_i \mid i \in I\}$ est appelée suite des facteurs principaux, elle constitue une base orthonormée de H' .
- ii) La suite $\{v_i = M^{-1}u_i \mid i \in I\}$ est un système orthonormé total dans H . Il est obtenu par l'analyse spectrale de l'opérateur nucléaire, auto-adjoint et positif VM ; $\{v_i \mid i \in I\}$ est appelé la suite des vecteurs principaux de l'a.c.p..
- iii) V admet la décomposition orthogonale (de h.s.) :

$$V = \sum_{i \in I} \lambda_i u_i \otimes v_i \quad \text{dans } \mathcal{L}(H', H) \quad (2.6)$$

où la suite $\{\lambda_i \mid i \in I\}$ est un élément de ℓ^1 car V est nucléaire.

- iv) Par U , la suite $\{u_i \mid i \in I\}$ est transformée en une suite $\{f_i \mid i \in I\}$ (les composantes principales) qui est une suite de vecteurs propres de l'opérateur UM^tUD associé aux mêmes valeurs propres; $\{f_i \mid i \in I\}$ constitue une famille de v.a.r. non corrélées et de variance λ_i .
- v) La décomposition orthogonale de tU s'écrit :

$${}^tU = \sum_{i \in I} D^{-1} f_i \otimes v_i \quad \text{dans } \mathcal{L}(L^2(\mu)', H). \quad (2.7)$$

Approximation spline de l'analyse en composantes principales d'une v.a.h.

De plus, tUD s'exprime comme un opérateur intégral de noyau X et comme il est de h.s., son noyau admet la décomposition (cf. [26]) :

$$X = \sum_{i \in I} f_i v_i \quad \text{et} \quad Y = X + EY \quad (2.8)$$

qui est la formule de reconstitution des données ou encore la décomposition de Karhunen-Loève dans le cas d'un processus.

Remarque. — L'a.p.c. linéaire d'une v.a.h. du second ordre se généralise à celle d'une fonctionnelle aléatoire linéaire U admettant un opérateur de covariance V . C'est une généralisation au sens où il n'existe pas toujours de v.a. associée à une fonction aléatoire par la relation (2.4). Pour que cette v.a. existe, il suffit que V soit nucléaire (cf. [4]).

2.4. Autre définition

Une définition globale de l'a.c.p. pourrait être la suivante.

DÉFINITION 2.2. — On appelle a.c.p. linéaire centrée de Y dans H la recherche du sous-espace H_q de H de dimension finie q qui maximise $\|P_{H_q} {}^tU\|_{\text{h.s.}}^2$ où P_{H_q} est le projecteur orthogonal sur H_q et la norme celle de l'espace $\sigma_2(L^2(\mu)', H)$ des opérateurs de h.s..

Remarques

L'explicitation de la norme de h.s. conduit au même type de maximisation exprimée cette fois dans l'espace H . La solution est un espace engendré par les vecteurs propres de VM (les vecteurs principaux $\{v_i \mid i = 1, \dots, q\}$), orthonormés dans H et associés aux q plus grandes valeurs propres.

Cette définition est équivalente à la précédente car les solutions sont emboîtées : $H_{q+1} = H_q \oplus \{v_{q+1}\}$ pour $q = 1, 2, \dots$

En outre, Techené [29] (ou Dauxois [15]) obtient les mêmes résultats en remplaçant la norme de h.s. par l'une quelconque de celles des espaces $\sigma_p(H, L^2(\mu))$ ($p \in [1, \infty[$).

2.5. Exemple : analyse en composantes principales d'une fonction aléatoire

Soit (T, ξ) un espace mesurable et ν une mesure σ -finie sur cet espace. On suppose que ν admet une base dénombrable de sorte que l'espace $L^2(T, \xi, \nu)$

est séparable. On considère une fonction aléatoire $\{X_t \mid t \in T\}$ définie sur $(\Omega, \mathcal{A}, \mu)$ et à valeurs dans un espace hilbertien H ; Xt est supposée du second ordre et centrée ($X \in L^2(\Omega \times T, \mathcal{A} \otimes \xi, \mu \otimes \nu; H, \mathcal{B}_H)$).

Il est alors équivalent de considérer X comme une fonction aléatoire X_t à valeurs dans H ou comme une v.a.h., encore notée X , à valeurs dans l'espace hilbertien séparable $L^2(T, \xi, \nu; H, \mathcal{B}_H)$, désigné par \mathcal{H} , car les espaces $L^2(\Omega \times T, \mathcal{A} \otimes \xi, \mu \otimes \nu; H, \mathcal{B}_H)$ et $L^2(\Omega, \mathcal{A}, \mu; \mathcal{H}, \mathcal{B}_{\mathcal{H}})$ sont isométriques.

L'a.c.p. de la fonction aléatoire $\{X_t \mid t \in T\}$ étant définie comme celle de la v.a.h. X dans \mathcal{H} , on retrouve, avec des hypothèses plus faibles, les résultats de Dauxois et Pousse [16] ou ceux de Deville [18].

3. "Fonctions" splines

Les développements et preuves concernant ce paragraphe sont à rechercher dans Laurent [22] ou Atteia [3].

3.1. Notations, hypothèses

Soient F, G et H trois espaces hilbertiens, φ (resp. ψ) une application linéaire continue de H dans F (resp. G) et φ^* (resp. ψ^*) l'application adjointe; φ^* (resp. ψ^*) est bijective de F (resp. G) sur $(\ker \varphi)^\perp$ (resp. $(\ker \psi)^\perp$) et $\varphi^* \varphi$ (resp. $\psi^* \psi$) est bijective de $(\ker \varphi)^\perp$ (resp. $(\ker \psi)^\perp$) sur lui-même. On suppose :

(H2) $\ker \varphi + \ker \psi$ est fermé dans H ,

(H3) $\ker \varphi \cap \ker \psi = \theta_H$.

3.2. Fonctions splines d'interpolation

Si (H2) et (H3) sont vérifiées, pour tout élément g de G il existe un seul σ dans H tel que :

$$\|\varphi \sigma\|_F = \min \{ \|\varphi h\|_F \mid h \in H \text{ et } \psi h = g \}. \quad (3.1)$$

σ est appelée la fonction spline d'interpolation (f.s.i.) relativement au triplet (φ, ψ, g) . On note :

$$S = (\varphi^* \varphi)^{-1} [(\ker \psi)^\perp] = \varphi^{-1} [(\varphi(\ker \psi))^\perp]. \quad (3.2)$$

C'est un sous-espace fermé de H appelé espace des fonctions splines associé au couple (φ, ψ) ; S contient $\ker \varphi$. De plus, l'application définie de H dans S par :

$$\forall h \in H, \sigma(h) \text{ est la f.s.i. associée au triplet } (\varphi, \psi, \psi(h)) \quad (3.3)$$

est linéaire.

Enfin, si $\{\zeta_i \mid i \in I\}$ est une base de G (I fini ou égal à \mathbb{N}) et si σ_i désigne la f.s.i. associée au triplet (φ, ψ, ζ_i) alors,

i) $\{\sigma_i \mid i \in I\}$ est une base de S muni de la topologie induite par celle de H et

ii) si σ est une f.s.i. associée à $\left(\varphi, \psi, \sum_{i \in I} \alpha_i \zeta_i\right)$ (où les α_i sont des réels)

alors σ se met sous la forme $\sum_{i \in I} \alpha_i \sigma_i$.

3.3. Fonctions splines d'ajustement

Avec les mêmes hypothèses et notations que précédemment, on désigne par E l'espace produit $F \times G$ muni du produit scalaire :

$$\begin{aligned} (e_1 \mid e_2)_E &= ((f_1, g_1) \mid (f_2, g_2))_E \\ &= (f_1 \mid f_2)_F + \rho(g_1 \mid g_2)_G \quad \text{avec } \rho \in \mathbb{R}^{+*}. \end{aligned} \quad (3.4)$$

On note ℓh le couple $(\varphi h, \psi h)$ pour tout élément h de H et L l'image de H par ℓ muni du produit scalaire et de la norme induits par E . Alors, L est un sous-espace vectoriel fermé de E , ℓ est linéaire, bijective, continue et d'inverse continue de H sur L enfin, $\ell^* \ell = \varphi^* \varphi + \rho \psi^* \psi$.

Pour tout élément g de G désignons par $\eta(g)$ le couple (θ_F, g) et notons $\eta_L(g)$ la projection de $\eta(g)$ sur L ; il existe un et un seul élément σ_ρ de H appelé f.s.a. de paramètre ρ associée au triplet $(\varphi, \psi, \eta(g))$ et vérifiant :

$$\|\ell \sigma_\rho - \eta(g)\|_E = \min \{ \|\ell h - \eta(g)\|_E \mid h \in H \}; \quad (3.5)$$

σ_ρ est la f.s.i. associée au triplet $(\varphi, \psi, \psi(\ell^{-1} \eta_L(g)))$ et ainsi, les propriétés des f.s.a. se déduisent de celles des f.s.i.

La f.s.a. σ_ρ est un élément de l'espace S des fonctions splines (3.2) et l'application de H dans S définie par :

$$\forall h \in H, \sigma_\rho(h) \text{ est la f.s.a. associée au triplet } (\varphi, \psi, \eta(\psi(h))) \quad (3.6)$$

est linéaire. Soit $\{\zeta_i \mid i \in I\}$ une base de G et σ_{ρ_i} la f.s.a. associée à $(\varphi, \psi, \eta(\zeta_i))$. Si $\{\psi(\sigma_{\rho_i}) \mid i \in I\}$ est une base de G alors,

- i) $\{\sigma_{\rho_i} \mid i \in I\}$ est une base de S muni de la topologie induite de celle de H et
- ii) si $\sigma_\rho(g)$ est la f.s.a. associée à $(\varphi, \psi, \eta(g))$ où g , élément de G , se décompose en $\sum_{i \in I} \alpha_i \zeta_i$ alors $\sigma_\rho(g)$ s'écrit $\sum_{i \in I} \alpha_i \sigma_{\rho_i}$.

Le paramètre ρ contrôle le compromis entre une interpolation et une régression sur $\ker \varphi$ car avec :

- σ , f.s.i. associée au triplet (φ, ψ, g) ,
- σ_ρ , f.s.a. associée au triplet $(\varphi, \psi, \eta(g))$,
- τ , l'élément unique de $\ker \varphi$ qui rend $\|\psi\tau - g\|_G$ minimum,

on a :

$$\lim_{\rho \rightarrow \infty} \|\sigma_\rho - \sigma\|_H = 0 \quad \text{et} \quad \lim_{\rho \rightarrow 0} \|\sigma_\rho - \tau\|_H = 0.$$

3.4. Partition de H

En pratique, le produit scalaire de H est particularisé de la façon suivante :

$$\forall h_1, h_2 \in H, \quad (h_1 \mid h_2)_H = (\beta h_1 \mid \beta h_2)_\Gamma + (\varphi h_1 \mid \varphi h_2)_F \quad (3.7)$$

où β désigne une application linéaire de H dans un espace hilbertien Γ de sorte que l'intersection de $\ker \beta$ et $\ker \varphi$ soit réduite à θ_H et que leur somme soit fermée dans H . Ainsi, H se décompose en une somme directe hilbertienne

$$H = \ker \varphi \oplus (\ker \varphi)^\perp$$

dans laquelle $\ker \varphi$ est muni du produit scalaire $(\beta \cdot \mid \beta \cdot)_\Gamma$ et $(\ker \varphi)^\perp$ du produit scalaire $(\varphi \cdot \mid \varphi \cdot)_F$.

Plus précisément, si $\ker \varphi$ est de dimension finie m alors (H2) et la fermeture de $\ker \beta + \ker \varphi$ sont vérifiées et, de plus, β peut être choisie de la forme :

$$\beta h = [\beta_k(h) \mid \beta_k \in H', k = 1, \dots, m] \quad (3.8)$$

à valeurs dans \mathbb{R}^m muni du produit scalaire usuel. L'expression (3.8) définit un ensemble de m conditions "aux bornes" qui font de (3.7) un produit scalaire sur H .

D'autre part, l'espace quotient $H/\ker \varphi$ est hilbertien lorsqu'il est muni du produit scalaire $(\varphi \cdot | \varphi \cdot)_F$.

Exemples

- Avec $T = [0, 1]$ et $\varphi = D^2$, $\ker \varphi = \text{vect}\{1, t\}$ et si H est l'espace de Sobolev $W^{(2)}$ muni du produit scalaire (3.7), β peut prendre différentes formes :

$$\begin{aligned} \beta h &= [h(0), h'(0)] , \\ \beta h &= [h(0), h(1)] , \\ \beta h &= \left[\int_0^1 h(t) dt , \int_0^1 h'(t) dt \right] , \\ &\vdots \end{aligned}$$

- Avec $T = [0, 1]$ et $\varphi = \theta^2 D + D^3$, $\ker \varphi = \text{vect}\{1, \sin \theta t, \cos \theta t\}$ et $H = W^{(3)}$ est muni du produit scalaire (3.7) où β peut être :

$$\beta h = \left[\int_0^1 h(t) dt , \int_0^1 \sin(\theta t) h(t) dt , \int_0^1 \cos(\theta t) h(t) dt \right] .$$

4. Approximation spline de l'analyse en composantes principales

L'a.c.p. théorique définie au paragraphe 2 ne peut être mise en œuvre qu'à la condition de pouvoir observer plusieurs réalisations de la v.a.h. Y et d'en construire des approximations. Ceci est décrit en utilisant le formalisme abstrait du paragraphe 3 introduit pour l'étude des fonctions splines d'interpolation ou d'ajustement; l'observation d'une réalisation de $Y(\omega)$ permet de l'approcher par un élément d'un sous-espace fermé S de H . Enfin, l'a.c.p. est toujours définie par rapport à une métrique mesurant une distance entre les individus. Le choix de cette métrique est fondamental même s'il est occulté dans la pratique courante des logiciels commerciaux qui utilise implicitement un produit scalaire défini par la matrice identité dans l'analyse des données brutes ou réduites. Il conduit dans le cas général, à déterminer l'espace de référence dans lequel plonger les interpolations ou approximations.

4.1. Définition

Soit \mathcal{H} un espace hilbertien, \mathcal{M} l'isométrie canonique entre \mathcal{H} et son dual et Π une application linéaire continue de H dans \mathcal{H} ; Σ désigne indifféremment l'opérateur d'interpolation spline noté σ ou celui d'ajustement noté σ_ρ opérant de H dans S et qui à tout élément h de H associe la f.s.i. relativement au triplet $(\varphi, \psi, \psi(h))$, ou la f.s.a. associée au triplet $((\varphi, \psi, \eta(\psi(h)))$.

DÉFINITION 4.1. — *Avec les notations des paragraphes précédents ($E, F, G, H, \mathcal{H}, \varphi, \psi, \Pi$), on appelle approximation spline, relativement au triplet (φ, ψ, Π) , de l'a.c.p. linéaire centrée de la v.a.h. Y à valeurs dans H , l'a.c.p. dans \mathcal{H} de la v.a.h. $\Pi\Sigma Y$ d'espérance $\Pi\Sigma EY$ et de covariance $\Pi\Sigma V^t\Sigma^t\Pi$ notée \mathcal{V} .*

Cette définition a un sens car $\Pi\Sigma$ est bien une application linéaire continue de H dans \mathcal{H} dont la covariance, si celle, notée V , de Y existe, est donnée par l'expression (2.2). On parlera d'interpolation (resp. d'ajustement) spline de l'a.c.p. lorsque Σ est σ (resp. σ_ρ).

PROPOSITION 4.1. — *Si les hypothèses (H1), (H2) et (H3) sont vérifiées, l'approximation spline, relativement au triplet (φ, ψ, Π) , de l'a.c.p. linéaire centrée de la v.a.h. Y est obtenue par l'analyse spectrale de l'un des opérateurs auto-adjoints, positifs et nucléaires $\mathcal{M}\mathcal{V}$ dans \mathcal{H} , ${}^t\Pi|_S\mathcal{M}\Pi|_S\Sigma V^t\Sigma$ dans S' (définition 2.1), $\mathcal{V}\mathcal{M}$ dans \mathcal{H} ou $\Sigma V^t\Sigma^t\Pi|_S\mathcal{M}\Pi|_S$ dans S (définition 2.2).*

Remarques

$\Pi|_S$ désigne la restriction de Π à S ; c'est la composée de Π par l'injection canonique i_S de S dans H . L'application transposée ${}^t\Pi|_S$ est alors la composée de ti_S par ${}^t\Pi$.

Si S est de dimension finie p , ti_S est le projecteur orthogonal de H' sur S' . De plus, le dernier opérateur $\Sigma V^t\Sigma^t\Pi|_S\mathcal{M}\Pi|_S$, qui admet alors une représentation matricielle dans une base de S , est facile à manipuler concrètement.

La preuve découle directement de la proposition 2.2.

4.2. Rôle des paramètres

Chacun des paramètres cités dans la définition 4.1 a un rôle spécifique dans la chaîne d'analyse "mesure, approximation, a.c.p." que l'on va préciser ci-dessous. Implicitement et pour illustrer notre propos, on s'intéresse

maintenant à des processus c'est-à-dire à des variables aléatoires à valeurs dans un espace fonctionnel, par exemple un sous-espace hilbertien de \mathbb{R}^T où T est un ouvert de \mathbb{R} . C'est le cas de la plupart des exemples traités dans les références citées en introduction qui détaillent par ailleurs les explicitations matricielles nécessaires à une mise en œuvre pratique (voir Ramsay et Dalzell [23] pour un exposé complet). Ces dernières utilisent une base de S construite en utilisant les noyaux reproduisants K , K_φ , K_{φ^\perp} , des espaces H , $\ker \varphi$ et $(\ker \varphi)^\perp$ ($K = K_\varphi + K_{\varphi^\perp}$).

a. Type de mesure

Le couple (ψ, G) décrit l'appareil de mesure. Dans le cas élémentaire, G est \mathbb{R}^p , il s'agit d'une simple évaluation :

$$\psi = [\delta_{t_j}(h) \mid \delta_{t_j} \in H', j = 1, \dots, p]$$

est la fonctionnelle vectorielle associant des mesures de Dirac aux points (les nœuds) de discrétisation. Cette opération a un sens si la fonctionnelle δ_t est continue de H dans \mathbb{R} , c'est-à-dire si H est un sous-espace hilbertien auto-reproduisant de \mathbb{R}^T . L'hypothèse (H2) est vérifiée car G est de dimension finie tandis que (H3) dépend de la position des nœuds, les $\{t_j \mid j = 1, \dots, p\}$ doivent être $\ker \varphi$ -unisolvants.

Tant que les trajectoires de Y possèdent des propriétés de régularité suffisantes, cette démarche est admissible. Dans le cas contraire, par exemple si Y est une mesure aléatoire ou même simplement à valeurs dans un espace $L^2(T)$, la simple évaluation n'est plus valide et les mesures de Dirac de ψ doivent être remplacées par des fonctionnelles continues sur l'espace H . Besse [6] donne un exemple de cette approche en utilisant un opérateur intégral à noyau.

b. Type d'approximation

Σ permet de choisir entre σ et σ_ρ . Les f.s.i. sont adaptées à des trajectoires très régulières tandis que les f.s.a. permettent un lissage de trajectoires bruitées dont on peut supposer que ce sont les mesures entachées d'erreur d'un phénomène sous-jacent lisse. Le choix de la valeur du paramètre de lissage ρ est important. Il peut être déterminé en fonction des mesures par une technique de validation croisée (cf. [14] ou [10] pour une revue) associée au lissage de chaque trajectoire (cf. [23]) ou encore par l'utilisation d'un critère heuristique visant à "optimiser" les représentations de l'a.c.p. (cf. [7], [8]).

Un autre type de splines dites “sur un convexe” a été développé (cf. [22]) pour obtenir des ajustements passant à travers des intervalles d’erreur de taille fixée autour des points de mesure et minimisant une norme. Moins utilisé, il ne sera pas décrit ici mais il peut s’insérer dans ce cadre général. Biritxinaga [11] le compare aux autres types de splines pour l’estimation de la moyenne d’un processus.

Comme on s’intéresse essentiellement à des a.c.p. centrées, cette estimation de la moyenne est ici implicite grâce aux propriétés de linéarité de l’espérance et de Σ . Calculer l’approximation spline de l’a.c.p. à partir d’un tableau de données centrées revient à considérer que l’estimation de EY interpole, ou ajuste selon le cas, les moyennes empiriques :

$$\left\{ \frac{1}{n} \sum_{i=1}^n Y_{t_j}(\omega_i) \mid j = 1, \dots, p \right\}.$$

calculées sur les données brutes.

c. Type d’interpolation

Le couple (φ, F) conditionne la forme, optimale au sens de la norme d’un élément φh dans F , des fonctions de l’espace S entre les nœuds de discrétisation. Celle-ci peuvent être linéaires, polynômiales, exponentielles, ... par morceaux. En général, F est $L^2(T)$ et φ un opérateur différentiel plus ou moins sophistiqué. H est alors un espace de fonctions régulières de type espace de Sobolev, d’où l’importance de l’hypothèse (H1).

d. Type de métrique

Le couple (Π, \mathcal{H}) fixe le produit scalaire et donc la métrique qui est utilisée par l’a.c.p. pour mesurer les distances entre les approximations des trajectoires. Les situations rencontrées en pratique sont les suivantes :

- i) $\mathcal{H} = L^2(T)$ et Π est l’injection continue à image dense d’un espace de Sobolev H dans \mathcal{H} ; Besse [8] utilise ce contexte pour définir une a.c.p. lissante ($\Sigma = \sigma_\rho$).
- ii) $\mathcal{H} = H$ et Π est l’application identité.

Cette situation est reprise par tous les auteurs et Besse ([7], [8]) a proposé d’introduire un paramètre dans le produit scalaire afin de pondérer l’influence respective des composantes des éléments de H dans leur décomposition associée à la somme directe $H = \ker \varphi \oplus (\ker \varphi)^\perp$. La valeur de ce paramètre est alors optimisée

avec le même critère heuristique que celui proposé pour déterminer la valeur du paramètre de lissage.

iii) $\mathcal{H} = H/\ker \varphi$ ou $H/\ker \beta$ et Π est l'injection canonique.

C'est sans doute l'une des applications les plus intéressantes de la théorie exposée; elle a été proposée par Besse et Ramsay [9] puis développée de façon générale par Ramsay et Dalzell [23]. calculer l'a.c.p. relativement à la métrique induite par celle de $H/\ker \varphi$ revient à négliger la composante "structurelle" des trajectoires pour s'intéresser aux "résidus". Ceci permet, sur les exemples présentés d'éliminer des composantes triviales (effet taille, phénomène périodique) pour permettre à l'a.c.p. de révéler des structures plus fines dans les données. Cette approche est indissociable du dernier choix.

e. Type de contraintes

Si m est la dimension de $\ker \varphi$ et si l'on souhaite utiliser le produit scalaire associé à la partition de H , il faut définir β à l'aide de m fonctionnelles continues sur H . Les solutions les plus simples utilisent les valeurs des fonctions ou de leurs dérivées aux bornes de T mais pour des applications de type statistique, Ramsay et Dalzell [23] trouvent plus efficace l'emploi de fonctionnelles intégrales.

5. Convergence

L'étude de la convergence de l'approximation définie précédemment est fondamentale pour s'assurer de la stabilité des résultats d'une analyse vis-à-vis des fluctuations pouvant intervenir sur le nombre de points de discrétisation ou la taille de l'échantillon. C'est une des raisons essentielles qui ont motivé les développements théoriques précédents permettant d'exhiber les analyses limites.

5.1. Convergence d'opérateurs compacts

Un premier lemme, dont on peut trouver une preuve dans Dunford et Schwartz [19] montre que la convergence uniforme des opérateurs, c'est-à-dire la convergence au sens de la norme de l'espace de Banach $\mathcal{L}(H_1, H_2)$, induit de bonnes propriétés sur la convergence des éléments propres.

LEMME 5.1. — Soit $\{T_p \mid p \in \mathbb{N}\}$ une suite d'opérateurs compacts de $\mathcal{L}(H)$ qui convergent uniformément vers T (T est donc compact). Soit $\lambda_k(T)$ une énumération des valeurs propres non nulles de T , chacune étant répétée suivant son ordre de multiplicité. Alors, il existe des énumérations $\lambda_k(T_p)$ des valeurs propres non nulles de T_p , chacune répétée suivant son ordre de multiplicité, telles que :

$$\lim_{p \rightarrow \infty} \lambda_k(T_p) = \lambda_k(T), \quad \text{pour } k \geq 1.$$

De plus, les sous-espaces propres (resp. les projecteurs propres) associés aux valeurs propres $\lambda_k(T_p)$ convergent uniformément vers les sous-espaces propres (resp. les projecteurs propres) associés aux $\lambda_k(T)$.

Remarques

L'expression "énumération $\lambda_k(T_p)$ des valeurs propres non nulles de T_p " signifie que celles-ci sont non nulles à partir d'un certain rang fini p_0 . En effet, si $\{T_p \mid p \in \mathbb{N}\}$ est une suite d'opérateurs de rang fini p convergeant uniformément vers un opérateur compact T , le nombre de valeurs propres non nulles croît alors avec le rang de p .

Si λ est de multiplicité 1, pour tout vecteur propre normé u associé à λ , il existe une suite $\{u_p \mid p \in \mathbb{N}\}$ convergeant vers u telle que :

$$\forall p \in \mathbb{N}, \quad T_p u_p = \lambda_p u_p \quad \text{et} \quad \lim_{p \rightarrow \infty} \lambda_p = \lambda.$$

Un autre lemme est utilisé.

LEMME 5.2. — Soit E_1, E_2, E_3 trois espaces de Banach et $\{T_p \mid p \in \mathbb{N}\}$ une suite d'éléments de $\mathcal{L}(E_2, E_3)$ qui converge fortement vers T de $\mathcal{L}(E_2, E_3)$. Si V est un opérateur compact de $\mathcal{L}(E_1, E_2)$ alors la suite $\{T_p V \mid p \in \mathbb{N}\}$ converge vers TV dans $\mathcal{L}(E_1, E_3)$.

Preuve. — En effet, si on note \mathcal{S} la sphère de rayon unité de E_1 alors pour tout n de \mathbb{N} :

$$\begin{aligned} \|(T_p - T)V\|_{\mathcal{L}(E_1, E_3)} &= \sup_{x \in \mathcal{S}} \|(T_p - T)Vx\|_{E_3} \leq \\ &\leq \sup_{x \in \mathcal{S}} \left\{ \frac{\|(T_p - T)Vx\|_{E_3}}{\|Vx\|_{E_2}} \right\} \|V\|_{\mathcal{L}(E_1, E_2)} \leq \\ &\leq \sup_{y \in V(\mathcal{S})} \left\{ \frac{\|(T_p - T)y\|_{E_3}}{\|y\|_{E_2}} \right\} \|V\|_{\mathcal{L}(E_1, E_2)}. \end{aligned}$$

D'après un corollaire du théorème de Banach-Steinhaus (cf. [27] p. 142), comme T est limite simple de la suite $\{T_p \mid p \in \mathbb{N}\}$, alors T est linéaire et continue, les normes T_p sont bornées et la suite $\{T_p \mid p \in \mathbb{N}\}$ converge uniformément sur tout compact de E_2 . Or V est compact donc $V(\mathcal{S})$ est relativement compacte et $\{T_p V \mid p \in \mathbb{N}\}$ converge uniformément vers TV sur \mathcal{S} donc dans $\mathcal{L}(E_1, E_3)$ d'après les inégalités ci-dessus. \square

5.2. Convergence de l'approximation

Soit $\{G_p \mid p \in \mathbb{N}\}$ une suite d'espace hilbertien et $\{\psi_p \mid p \in \mathbb{N}\}$ une suite d'applications linéaires bornées de H sur G_p ; Σ_p désigne l'opérateur d'interpolation (resp. d'ajustement) spline σ_p (resp. σ_{ρ_p}) relativement au couple (φ, ψ_p) . Ceci suppose que les hypothèses suivantes sont vérifiées.

$$(H2_p) \quad \forall p \in \mathbb{N}, \ker \varphi + \ker \psi_p \text{ est fermé dans } H.$$

$$(H3_p) \quad \forall p \in \mathbb{N}, \ker \varphi \cap \ker \psi_p = \theta_H.$$

DÉFINITION 5.1. — *On dit que l'approximation spline, relative au triplet (φ, ψ_p, Π) de l'a.c.p. d'une v.a.h. Y converge vers l'a.c.p. de ΠY dans H si l'opérateur $\mathcal{V}_p (= \Pi|_{\mathcal{S}} \Sigma_p V^t \Sigma_p^t \Pi|_{\mathcal{S}})$ converge vers l'opérateur \mathcal{V} de covariance de ΠY dans $\mathcal{L}(\mathcal{H}', \mathcal{H})$.*

Cette définition est motivée par le lemme 5.1 car si \mathcal{V}_p converge uniformément vers \mathcal{V} alors, comme \mathcal{M} est borné, $\mathcal{M}\mathcal{V}_p$ converge uniformément vers $\mathcal{M}\mathcal{V}$ et donc les éléments propres de $\mathcal{M}\mathcal{V}_p$ convergent uniformément vers ceux de $\mathcal{M}\mathcal{V}$. On est ainsi assuré de la convergence des éléments principaux et des représentations de l'a.c.p.

THÉORÈME 5.1. — *Convergence de l'interpolation spline de l'a.c.p.*

Si les hypothèses (H1), (H2_p), (H3_p) ainsi que (H4_p) et (H5_p) ci-dessous sont vérifiées :

$$(H4_p) \quad \forall p \in \mathbb{N}, \ker \psi_{p+1} \subset \ker \psi_p;$$

$$(H5_p) \quad \bigcap_{p \in \mathbb{N}} \ker \psi_p = \theta_H;$$

alors, l'interpolation spline de l'a.c.p. d'une v.a.h. converge uniformément.

Preuve. — C'est une conséquence du corollaire 2.1 de Atteia ([3], chap. 5) qui montre la convergence forte de σ_p vers l'application identité de H sous les hypothèses (H2_p), (H3_p), (H4_p) et (H5_p) :

$$\forall h \in H, \quad \lim_{p \rightarrow \infty} \|\sigma_p(h) - h\|_H = 0.$$

Ceci entraîne la convergence forte de $\Pi \upharpoonright_S \sigma_p$ (resp. ${}^t \sigma_p {}^t \Pi \upharpoonright_S$) vers Π (resp. ${}^t \Pi$) et comme ${}^t U$ est compact et D borné, le lemme 5.2 montre la convergence uniforme de $\Pi \upharpoonright_S \sigma_p {}^t U D$ vers $\Pi {}^t U D$ dans $\mathcal{L}(L^2(\mu)', \mathcal{H})$ ainsi que celle de ${}^t(\Pi \upharpoonright_S \sigma_p {}^t U)$ vers $U {}^t \Pi$ dans $\mathcal{L}(\mathcal{H}, L^2(\mu))$. Le résultat est obtenu en composant les deux suites. \square

THÉORÈME 5.2. — Convergence de l'ajustement spline de l'a.c.p.

Si les hypothèses (H1), (H2_p), (H3_p) ainsi que (H6_p), (H7_p) et (H8_p) ci-dessous sont vérifiées.

$$(H6_p) \forall h \in H, \lim_{p \rightarrow \infty} \|\psi_p h\|_{G_p} = +\infty.$$

(H7_p) Pour tout p de \mathbb{N} , il existe un espace hilbertien $G_{0,p}$ et une application linéaire continue $\psi_{0,p}$ de H dans $G_{0,p}$ tels que :

$$i) \forall h \in H, \|\psi_{0,p} h\|_{G_{0,p}} \leq \|\psi_p h\|_{G_p},$$

$$ii) \exists C \in \mathbb{R}^{+*} \forall p \in \mathbb{N}, \|\psi_{0,p}\|_{\mathcal{L}(H, G_{0,p})} \leq C,$$

$$iii) \exists c \in \mathbb{R}^{+*} \forall p \in \mathbb{N}, \forall h \in \ker \varphi, c \|h\|_H \leq \|\psi_{0,p} h\|_{G_{0,p}}.$$

(H8_p) Pour tout p de \mathbb{N} , il existe un endomorphisme ζ_p de G_p tel que :

$$i) \exists K \in \mathbb{R}^{+*}, \forall p \in \mathbb{N}, \|\zeta_p \psi_p\|_{\mathcal{L}(H, G_p)} \leq K,$$

ii) il existe un endomorphisme ξ de H qui admet un inverse à gauche continu tel que :

$$\lim_{p \rightarrow \infty} \|(\zeta_p \psi_p)^* \zeta_p \psi_p - \xi\|_{G_p} = +\infty.$$

Alors, l'ajustement spline de l'a.c.p. d'une v.a.h. converge uniformément.

La preuve est identique à la précédente en utilisant la proposition 5.3 de ([3], chap. 5).

5.3. Échantillonnage statistique

Le comportement asymptotique de l'a.c.p., obtenue à partir d'un échantillon statistique, a été étudié par Dauxois et Pousse [16] puis Dauxois et al. [17] dans le cadre du modèle suivant.

Y est une v.a.h. définie sur l'espace probabilité $(\Omega, \mathcal{A}, \mu)$ et à valeurs dans (H, \mathcal{B}_H) . On suppose que la tribu \mathcal{A} contient tous les singletons de Ω ; γ_i désigne la $i^{\text{ème}}$ projection canonique de $(\Omega, \mathcal{A}, \mu)^{\otimes \mathbb{N}^*}$ sur $(\Omega, \mathcal{A}, \mu)$ telle que :

$$\forall i \in \mathbb{N}^*, \forall n \in \mathbb{N}^*, \gamma_i([\omega_n \mid n \in \mathbb{N}^*]) = \omega_i. \quad (5.1)$$

Approximation spline de l'analyse en composantes principales d'une v.a.h.

Les v.a. $\{\gamma_i \mid i \in \mathbb{N}^*\}$ (resp. $\{Y_i = Y\gamma_i \mid i \in \mathbb{N}^*\}$) sont i.i.d. de loi μ (resp. μ_Y de Y). Les $\{Y_i \mid i \in \mathbb{N}^*\}$ constituent alors un échantillonnage statistique permettant de construire l'opérateur de covariance empirique usuel dont on montre (cf. [16]) qu'il converge uniformément p.s. vers V lorsque la taille de l'échantillonnage croît indéfiniment dans $\mathcal{L}(H', H)$.

THÉORÈME 5.3. — *Avec les notations et hypothèses du théorème 5.1 (resp. 5.2), l'interpolation (resp. l'ajustement) spline de l'a.c.p. d'une v.a.h. obtenue à partir d'un n -échantillon converge uniformément p.s.*

Preuve. — On s'intéresse à la convergence uniforme p.s. dans $\mathcal{L}(H', H)$ de l'opérateur d'échantillonnage obtenu par approximation spline :

$$V_{n_p} = \Pi \upharpoonright_S \Sigma_p V_n {}^t \Sigma_p {}^t \Pi \upharpoonright_S.$$

En décomposant la norme dans $\mathcal{L}(H', H)$:

$$\begin{aligned} \|\Pi \upharpoonright_S \Sigma_p V_n {}^t \Sigma_p {}^t \Pi \upharpoonright_S - V\| &\leq \|\Pi \upharpoonright_S \Sigma_p V_n {}^t \Sigma_p {}^t \Pi \upharpoonright_S - \Pi \upharpoonright_S \Sigma_p V {}^t \Sigma_p {}^t \Pi \upharpoonright_S\| + \\ &+ \|\Pi \upharpoonright_S \Sigma_p V {}^t \Sigma_p {}^t \Pi \upharpoonright_S - V\|. \end{aligned}$$

or

$$\begin{aligned} \|\Pi \upharpoonright_S \Sigma_p V_n {}^t \Sigma_p {}^t \Pi \upharpoonright_S - \Pi \upharpoonright_S \Sigma_p V {}^t \Sigma_p {}^t \Pi \upharpoonright_S\| &\leq \\ &\leq \|{}^t \Sigma_p {}^t \Pi \upharpoonright_S\|_{\mathcal{L}(\mathcal{H}, H')} \|\Pi \upharpoonright_S \Sigma_p\|_{\mathcal{L}(H, \mathcal{H})} \|V_n - V\| \end{aligned}$$

et comme Π (resp. ${}^t \Pi$) est limite simple de la suite $\left\{ \Pi \upharpoonright_S \Sigma_p \mid p \in \mathbb{N} \right\}$ (resp. $\left\{ {}^t \Sigma_p {}^t \Pi \upharpoonright_S \mid p \in \mathbb{N} \right\}$), l'application du corollaire du théorème de Banach-Steinhaus déjà utilisé montre l'existence d'une constante C indépendante de n et p qui majore les normes et telle que :

$$\|\Pi \upharpoonright_S \Sigma_p V_n {}^t \Sigma_p {}^t \Pi \upharpoonright_S - \Pi \upharpoonright_S \Sigma_p V {}^t \Sigma_p {}^t \Pi \upharpoonright_S\| \leq C \|V_n - V\|.$$

De la convergence uniforme du théorème 5.1 (resp. 5.2) et de celle de V_n vers V , on déduit que pour tout ϵ positif,

$$\exists p_0 \in \mathbb{N} \mid p > p_0 \Rightarrow \|\Pi \upharpoonright_S \Sigma_p V {}^t \Sigma_p {}^t \Pi \upharpoonright_S - V\| < \frac{\epsilon}{2},$$

$$\exists n_0 \in \mathbb{N} \mid n > n_0 \Rightarrow C \|V_n - V\| < \frac{\epsilon}{2} \text{ p.s.,}$$

et donc, d'après (5.2),

$$\exists p_0, n_0 \in \mathbb{N} \mid p > p_0 \text{ et } n > n_0 \Rightarrow \|\Pi \upharpoonright_S \Sigma_p V_n {}^t \Sigma_p {}^t \Pi \upharpoonright_S - V\| < \epsilon \text{ p.s. } \square$$

Remarque. — Malgré leur caractère très général, ces résultats de convergence n'englobent pas tous les cas déjà cités. Biritxinaga [11] considère une situation où les hypothèses sur les trajectoires de Y sont plus faibles que celles sur son espérance EY supposée plus régulière. Les techniques de démonstration de la convergence de l'estimation spline (interpolation, ajustement ou sur un convexe) de la moyenne d'un processus sont alors différentes.

Références

- [1] ANDERSON (T.W.) . — *Introduction to multivariate statistical analysis*, Wiley, New-York (1958).
- [2] ATTEIA (M.) . — *Généralisation de la définition et des propriétés des "spline-fonctions"*, C.R. Acad. Sci. Paris, **260** (1985) pp. 3550-3553.
- [3] ATTEIA (M.) . — *Hilbertian kernels and spline functions*, (1989) livre à paraître.
- [4] BENSOUSSAN (A.) . — *Filtrage optimal des systèmes linéaires*, Dunod, Paris (1971).
- [5] BESSE (P.) . — *Étude descriptive d'un processus; approximation, interpolation*, Thèse de spécialité, Toulouse (1979).
- [6] BESSE (P.) . — *Choix de la métrique pour l'a.c.p. d'événements discrets*, Statistique et Analyse des Données, **12**, n° 3 (1987) pp. 1-16.
- [7] BESSE (P.) . — *Spline functions and optimal metric in linear principal components analysis*, in *Components and Correspondence analysis* (Eds. J.L.A. van Rijkevorsel and J. de Leeuw), Wiley, Londres (1988).
- [8] BESSE (P.) . — *Optimal metric in principal components analysis of longitudinal data*, in *Data analysis and informatics V*, (Eds E. Diday et al.), Nord Holland, Amsterdam (1988).
- [9] BESSE (P.) and RAMSAY (J.O.) . — *Principal components analysis of sampled functions*, Psychometrika, **51**, n° 2 (1986) pp. 285-311.
- [10] BESSE (P.) et THOMAS-AGNAN (C.) . — *Le lissage par fonctions splines en statistique : revue bibliographique*, Statistique et Analyse des données, **14**, n° 1 (1989) pp. 55-83.
- [11] BIRITXINAGA (E.) . — *Sur la convergence de deux estimateurs spline de la moyenne d'un processus stochastique*, publ. n° 89/3, UA 1204 CNRS (1989).
- [12] BIRITXINAGA (E.) . — *Intervalles de confiance et estimation spline de l'espérance mathématique d'une variable aléatoire à valeurs dans $D[0, 1]$* publ. n° 89/4, UA 1204 CNRS (1989).

- [13] BOUHADDOU (O.) .— *Analyse en composantes principales et interpolation de processus*,
Thèse de spécialité, Grenoble (1984).
- [14] CRAVEN (P.) and WAHBA (G.) .— *Smoothing noisy data with spline functions*,
Numerische Mathematik 31 (1984) pp. 377-403.
- [15] DAUXOIS (J.) Cours de DEA, Université Toulouse III (1986).
- [16] DAUXOIS (J.) et POUSSE (A.) .— *Les analyses factorielles en calcul des probabilités et en statistique : essai d'étude synthétique*,
Thèse, Toulouse (1976).
- [17] DAUXOIS (J.), POUSSE (A.) and ROMAIN (Y.) .— *Asymptotic theory for the principal component analysis of a vector random function : some applications to statistical inference*,
Journal of Multivariate Analysis, 12 (1982) pp. 136-154.
- [18] DEVILLE (J.C.) .— *Méthodes statistiques et numériques de l'analyse harmonique*,
Annales de l'INSEE, n° 15 (1974) pp. 3-101.
- [19] DUNFORD (N.) and SCHWARTZ (J.T.) .— *Linear operators, part II*,
Interscience Publishers, 2^{ème} édition, New-York (1964).
- [20] GRENANDER (U.) .— *Probabilities on algebraic structures*,
Almquist et Wiksell, Stockholm (1963).
- [21] GUELFAND (I.M.) et VILENKIN (N.Y.) .— *Les distributions*,
Dunod, Paris, t. 4 (1967).
- [22] LAURENT (P.J.) .— *Approximation et optimisation*,
Hermann, Paris (1972).
- [23] RAMSAY (J.O.) and DALZELL (C.) .— *Some tools for functional data analysis*,
J. M. Statist. Soc. B 53 n° 3 (1991) pp. 539-561.
- [24] SAPORTA (G.) .— *Méthodes exploratoires d'analyse des données temporelles*,
Cahiers du BURO, Paris, n° 37-38 (1981).
- [25] SAPORTA (G.) .— *Data analysis for numerical and categorical individual time-series*,
Applied Stochastic Models and Data Analysis, 1 (1985) pp. 109-119.
- [26] SCHAEFER (H.H.) .— *Banach lattices and positive operators*,
Springer-Verlag, Berlin (1974).
- [27] SCHWARTZ (L.) .— *Analyse hilbertienne*
Hermann, Paris (1979).
- [28] TAYLOR (S.J.) .— *Introduction to measure and integration*,
Cambridge University Press (1966).
- [29] TECHENÉ (J.J.) .— *Réduction optimale d'opérateurs, application aux analyses factorielles*,
Thèse de spécialité, Pau (1980).
- [30] WINSBERG (S.) .— *Two techniques: monotone spline transformations for dimension reduction in p.c.a. and easy to generate metrics for p.c.a. of sampled functions*,
in *Components and Correspondance analysis* (Eds. J.L.A. van Rijkevorsel and J. de Leeuw), Wiley, Londres (1988).