République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Mouloud Mammeri de Tizi Ouzou
Faculté des Sciences
Département de Mathématiques

Mémoire de Master

Présentée par Mme. Aissiou Nassima

En vue de l'obtention de Master II en Mathématiques

Option : Probabilités et Statistique

Thème

Traitement de données sensorielles par une analyse en Composantes principales de Densités de Probabilité.

Soutenue le $30 juin 2024$ **devant le jury suivant :**

Président	Hocine Kourat	MCB
Rapporteur	Smail Yousfi	MCB
Examinateur	Youcef Ibaoune	MCB
Examinateur	Farid Graiche	MCB

Remerciements

Je remercie Mensieur Smail Yousfi, rapporteur de ce manuscrit, pour avoir accepté et assuré la direction de ce travail.
Je remercie Hocine Kourat, Youcef Ibaoune et Farid Graiche pour m'avoir fait l'honneur de juger ce travail de mémoire.
Je remercie ma mére, mon pére, mes soeurs, mon frére, mon marie. Tout ceci n'aurait jamais été possible sans vous.

Table des matières

Introduction générale

L'analyse statistique des courbes ou analyse des données fonctionnelles, est un thème de recherche en statistique qui est en plein expansion et dont les applications concernent de nombreux domaines scientifiques (climatologie, médecine, économie, chimie, ...). On poura ce raporter à [48] pour une revue de différentes méthodes d'analyse illustrées sur des exemples variées. Les outils statistiques mis on oeuvre sont issues de l'anaslyse fonctionnelle et généralement des procédures classiques multivariées .
Le statisticien cherche généralement, dans une première étape, à représenter au mieux ces données fonctionnelles dans un espace de dimention plus petite par l'intermédiaire d'une analyse en composantes principales adaptée au cadre fonctionnel. On peut ainsi déterminer le comportement moyen (courbe moyenne) et les principaux modes de variations autour de la moyenne grâce aux éléments propres de l'opérateur de covariance [22].
Un cas particulier de données fonctionnelle [7], [40] est une méthode bien adaptée à ce type de données, nous consacrons un chapitre dans ce mémoire pour exposer l'essentiele de cette méthode.
La mise en oeuvre de l'analyse en composantes principales de densités sur des données réelles et confrontée au problème des densités inconues, dans la littérature deux solutions ont été proposées, la première consiste en estimation paramétrique avec des hypothèses sur l'appartenance de la fonction de densité inconnue à une famille paramétrique connue (comme par exemple , les lois normales multidimentionnelles) [9]. La deuxième solution porte sur l'estimation non paramétrique sans l'hypothèses d'appartenance de la fonction de densité inconnue à une famille paramétrique connue. Parmi les méthodes non paramètrique on s'intéresse à l'estimation par la méthode noyau voir [55] , [38], [40],[63].
En appliquant dans ce travail cette méthode dans le cadre de l'analyse en composantes principales de densité, et on illustre avec des exemples de simulation, la convergence de l'ACP estimée vers l'ACP théorique, une application sur des données réelles (sensorielle et aussi exposée).
Comme dans toute les méthodes statistiques d'analyse des données multidimensionnelles la question d'intérprétation des résultats est impotante et rendue plus difficile par l'expression non linéaire des modes de variabilité et la perte des propriétés mathématiques des fonction propres obtenues par décomposition des densités. Cette question à été résolue en partie dans Boumaza et al (2014), elle consiste en interprétation de la position de la densité en fonction de la corrélation linéaire entre le vecteur des moyennes, le vecteur des variances, les vecteurs des covariances et les vecteurs des corrélations.
Ce mémoire est organisé comme suit :
Dans le chapitre 1, on donnera des rappels mathématiques nécaissairent à la méthode (ACP de densité) en particulier : la théorie des opérateurs et l'estimation de densité.
Le chapitre 2, est consacré à l'exposer de l'ACP théorique et estimê par noyau, et en illustre avec des exemples, la convergence de l'ACP estimé vers l'ACP théorique. Une

application sur les données réelles est aussi réalisée. Le chapitre 3 et à l'interprétation des sorties de l'analyse en composantes principales de densité.

Chapitre 1

Rappels

1.1 Estimateur de la densité de probabilité multivariée par la méthode du noyau

1.1.1 Estimateur à noyau

Soit $X = (X^{(1)}, ..., X^{(p)})$ un vecteur aléatoire à p-dimensions de fonction de densité $f : \mathbb{R}^p \longmapsto \mathbb{R}_+$, et soit $X_1, ..., X_i, ..., X_n$ un échantillon i.i.d de taille n.

Définition 1.1.1. On appelle estimateur à noyau de la densité de probabilité f la statistique suivante :

$$\hat{f}(x) = n^{-1} \sum_{i=1}^{n} K_H(x - X_i), \tag{1.1}$$

avec

$$K_H(x) = |H|^{-1/2} K(H^{-1/2} x), \tag{1.2}$$

où

- H une matrice carrée d'ordre p, symétrique et définie positive, appelée matrice des paramétres de lissage ou matrice des fenêtres (banduidth matrix) .
- $K(\cdot)$ appelée fonction noyau multivariée, qui est aussi une densité de probabilité.

1.2 Formes possibles du noyau multivariée

Il existe dans la littérature ([55], [59]) deux variantes possibles du noyau.

1.2.1 Noyau produit

Il est de la forme

$$K(x) = \prod_{i=1}^{p} w(x_i). \tag{1.3}$$

où w est une densité de probabilité réelle .

Quelques formes possibles de la fonction w :

Les différentes formes de la fonction w utilisées dans la littérature

Noyau	w(x).
Rectangulaire	$\frac{1}{2}I_{[-1,1]}(u)$.
Triangulaire	$(1-\|u\|)I_{[-1,1]}(u)$.
Gaussien	$\frac{1}{\sqrt{2\pi}}\exp^{\frac{-u^2}{2}}, u \in \mathbb{R}$.
Epanchnikov	$\frac{3}{4\sqrt{5}}(1-\frac{u^2}{5})I_{[-\sqrt{5},\sqrt{5}]}(u)$.
Biweight	$\frac{15}{16}(1-u^2)^2 I_{[-1,1]}(u)$.
Triweight	$\frac{35}{32}(1-u^2)^3 I_{[-1,1]}(u)$.
cosinus	$\frac{\pi}{4}\cos(\frac{\pi u}{2})I_{[-1,1]}(u)$.
Gamma	$w_{\lambda,r}(u) = \frac{\lambda^r}{\Gamma(r)}\exp^{-\lambda u} u^{r-1} I_{[0,\infty[}(u)$.
Beta	$\frac{u^{\alpha-1}(1-u)^{\beta-1}}{\beta(\alpha,\beta)} I_{]0,1[}(u)$.

1.2.2 Noyau sphérique

Il est de la forme

$$K(x) = C_{w.d} w\{(x^T x)^{\frac{1}{2}}\}, \tag{1.4}$$

avec

$$C_{w.d}^{-1} = \int w\{(x^T x)^{\frac{1}{2}}\}dx. \tag{1.5}$$

Remarque 1.2.1. Lorsque le support de la densité est $\mathbb{R}^p$, le noyau gaussien suivant

$$K_{N(\mu,\Sigma)}(z) = \frac{1}{(2\pi)^{\frac{p}{2}}(|\Sigma|)^{\frac{1}{2}}} exp^{\{-\frac{1}{2}(z-\mu)^t\Sigma^{-1}(z-\mu)\}} \mu \in \mathbb{R}^p \tag{1.6}$$

est le plus populaire. L'estimateur à noyau obtenu possède des bonnes propriétés asymptotiques .

1.3 Espace de Hilbert

Définition 1.3.1. Un produit scalaire sur un espace vectoriel réel H est une application de $H \times H$ dans $\mathbb{R}$, bilinéaire symétrique notée $< \cdot, \cdot >_H$ définé positif.

Définition 1.3.2. Un espace de Hilbert H est un espace vectoriel muni d'un produit scalaire qui est complet pour la norme induite, définie par :

$$\|x\|_H = (\langle x, x\rangle_H)^{1/2}. \tag{1.7}$$

La distance associée à cette norme est

$$\forall (x, x') \in H^2, \quad d_H(x, x') = \|x - x'\|_H. \tag{1.8}$$

L'orthogonalité est une notion très importante en particulier dans l'étude des opérateurs linéaires.

Définition 1.3.3. Deux vecteurs $x, y \in H$ sont orthogonaux si

$$\langle x, y\rangle_H = 0. \tag{1.9}$$

Définition 1.3.4. On appelle orthogonal d'une partie F d'un espace de Hilbert H noté par $F^{\perp}$, le sous espace vectoriel de H constitué des vecteurs orthogonaux à tous les vecteurs de F :

$$F^{\perp} = \{y \in H / \forall x \in F, \langle x, y\rangle_H = 0\}.$$

Définition 1.3.5. On note $L^2(\Omega)$ l'ensemble des fonction de carré intégrable sur Ω de carré intégrable sur Ω. Une fonction u définie sur Ωà valeur complexes est dite de carré intégrable si u est mesurable et $u \in L^2(\Omega)$. On définie alors la norme sur L^2

$$\|u\|_{L^2} = (\int_{\Omega} |u|^2)^{\frac{1}{2}}$$

L^2 est un espace vectoriel.

1.4 Projection sur un sous espace vectorielle

Le théorème suivant est une caractérisation de la notion de projection orthogonale sur un sous espace vectoriel.

Théorème 1.4.1. Soit H un espace de Hilbert et $F \subset H$ un convexe fermé, alors pour tout x dans H, il existe un unique $y \in F$ tel que :

$$\|x - y\| = d(x, F) = \inf\{\|x - z\|_H, z \in F\}$$

.

Notons alors par $y = p_F(x)$, ainsi

1. $p_F(x)$ est l'unique vecteur dans F qui satisfait $x - p_F(x) \perp F$.
2. $\forall x \in H; \|x\|^2 = \|p_F(x)\|^2 + \|p_{F^\perp}(x)\|^2$.
3. Le complémentaire orthogonal $F^\perp$ est un sous espace fermé de H, avec

$$H = F \oplus F^\perp.$$

$p_F(x)$ est appelé la projection orthogonale de x sur F .

Définition 1.4.1. Une famille de vecteur $(e_i)_{i\in I\subset\mathbb{N}}$ dans H est dite système orthonormal si

$$\langle e_i, e_j\rangle_H = \begin{cases} 0, & \text{si } \forall i \neq j, \\ 1, & \text{sinon.} \end{cases}$$

On écrit alors : $\|e_i\|_H^2 = \langle e_i, e_i\rangle_H$.

1.4.1 Décomposition sur un sous espace de projection

Soit $B = \{e_1, ..., e_n\}$ un système orthonormal et F le sous espace vectoriel de H engendré par B, alors

$$\forall x \in H, \quad p_F(x) = \sum_{i=1}^{n} \langle x, e_i\rangle_H e_i. \tag{1.10}$$

Définition 1.4.2. Un système orthonormal $(e_i)_{\{i\in I\}}$ est dit total dans H si

$$\{e_i, i \in I\}^\perp = \{0\}. \tag{1.11}$$

Autrement dit si $x \in H$ est tel que $\langle x, e_i\rangle_H = 0$ pour tout $i \in I$, alors $x = 0$.

Théorème 1.4.2. Si $(e_i)_{i\in I}$ est un système orthonormal total de H , alors

$$\forall x \in H, \quad x = \sum_{i\in I} \langle x, e_i\rangle_H e_i. \tag{1.12}$$

Remarque 1.4.1. Un système total H est appelé aussi base Hilbertienne de H .

Définition 1.4.3. Un espace de Hilbert H est dit séparable s'il possède une suite de points qui est dense dans H.

Théorème 1.4.3. Tout espace de Hilbert séparable possède une base orthonormale .

Théorème 1.4.4. Un espace de Hilbert est séparable si est seulement si possède une base hilbertienne .

1.5 Généralités sur les opérateurs

Définition 1.5.1. Soient H et $H^{'}$ deux espaces de hilbert, on appelle opérateur linéaire de H vers $H^{'}$ toute application linéaire de H dans $H^{'}$. Notons alors $L(H)$ l'ensemble des opérateurs linéaires de H vers H, on définit la norme d'un opérateur $T \in L(H)$ par

$$\|U\| = \sup\{\|Ux\|_H, x \in H, \|x\| \leq 1\}.$$

Définition 1.5.2. Soit $U \in L(H)$, alors il existe un unique opérateur $U^* \in L(H)$, tel que

$$\forall x, y \in H, \quad \langle Ux, y\rangle = \langle x, U^* y\rangle.$$

L'opérateur U^* est appelé **l'adjoint** dans H de l'opérateur U.

Remarque 1.5.1. L'opérateur U est dit

1. Auto-adjoint ou hermitien si $U^* = U$, dans ce cas :

$$\forall x, y \in H, \quad \langle Ux, y\rangle = \langle x, Uy\rangle.$$

2. Positif s'il est auto-adjoint et de plus $\langle Ux, x\rangle \geq 0, \ \forall x \in H$.

Propriétés. Soient U, U_1 et U_2, trois opérateurs linéaires dans H, on a alors les propriétés suivantes

1. Pour tout $U \in L(H)$, $(U^*)^* = U$ et $\|U^*\| = \|U\|$.
2. $(U_1 \circ U_2)^* = U_2^* \circ U_1^*$.
3. Si U est auto-adjoint , alors

$$\|U\| = sup\{|\langle Ux, x\rangle|, \|x\| = 1\}.$$

Définition 1.5.3. Un opérateur U de H dans H' est dit compact si l'image de la boule unité de H par U est relativement compact dans H'.

Définition 1.5.4. Soit U un opérateur de $L(H)$ et $(e_j)_{j\in I}$ une base hilbertienne de H, on appelle trace de l'opérateur U notée $tr(U)$ le réel

$$tr(U) = \sum_{i=1} \langle U(e_i), e_i\rangle.$$

Ce nombre ne dépend pas de la base hilbertienne choisie .

Définition 1.5.5. On dit qu'un opérateur $U \in L(H)$ admet une décomposition spectrale s'il existe une base $(e_i)_{i\in I}$ de H et des réels $(\lambda_i)_{i\in I}$ tel que

$$U = \sum_{i=1} \lambda_i e_i \otimes e_i.$$

$\otimes$ désigne le produit de kroneeker.
Ces réels sont appelés valeurs propres de U et les vecteurs $\{e_i\}$ sont les vecteurs propres associés .

Théorème 1.5.1. Tout opérateur $U \in L(H)$ auto-adjoint positif, admet une décomposition spectrale .

- L'opérateur étant positif, les valeurs propres (simples ou multiples) sont positives ou nulles .
- L'opérateur étant symétrique , les espaces propres associés à deux valeurs propres distinctes sont orthogonaux .

Si $(\lambda_i)_{i\in I}$ est la suite pleine décroissante des valeurs propre de U (c'est à dire répétée autant de fois que leur ordre de multiplicité l'indique) et $(u_i)_{i\in I}$ une suite de vecteurs propres unitaires associés formant une base orthonormale de H, l'opérateur U se décompose comme suit

$$T = \sum_{i\in I} \lambda_i u_i \otimes u_i. \tag{1.13}$$

La formule (1.13) appelée aussi décomposition de Hilbert-Schmidt.

1.5.1 Théorème d'identification de Riesz

Définition 1.1. On appelle dual d'un espace espace de Hilbert H noté H^*, l'espace des formes linéaire continues de H dans $K(K = \mathbb{R} \quad \text{où} \quad \mathbb{C})$, c'est-à-dire

$$H^* = L(H, K). \tag{1.14}$$

Le théorème de Riesz est d'une importance capitale en théorie des opérateurs, il permet l'identification d'un espace de Hilbert avec son dual topologique.

Théorème 1.5.2. (identification de Riesz) A tout élément x d'un espace de Hilbert H, on associe la forme linéaire f de H^* définie par :

$$\forall y \in H, \quad f(x, y) = \langle x, y\rangle_H.$$

1.6 Analyse en composantes principales d'un opérateur

Soient H et H' deux espaces de Hilbert séparables, l'espaces H (resp. H') est identifié à son dual. On pose $\langle \cdot, \cdot \rangle_H$ (resp. $\langle \cdot, \cdot \rangle_{H'}$) le produit scalaire dans H (resp. H')et $\| \cdot \|_H$ (resp. $\| \cdot \|_{H'}$) la norme associée.
Soit U un opérateur continu non nul de H dans H' et U^* son adjoint.

Définition 1.6.1. On appelle analyse en composantes principales "pas à pas"de U, tout couple $(\{\lambda_i\}_{i\in I}, \{u_i\}_{i\in I})$ où

1. I est une section commençante de N^*.
2. $\{\lambda_i\}_{i\in I}$ est une suite décroissante de réels positifs ou nuls.
3. $\{u_i\}_{i\in I}$ est une suite d'élèments de H' vérifiant les conditions suivantes

   (a) $\forall (i,j) \in I^2, \langle u_i, u_j\rangle_{H'} = \begin{cases} 1, & \text{si } \forall i = j \\ 0, & \text{sinon.} \end{cases}$

   (b) $\forall i \in I, \lambda_i = sup\{\frac{\|U^*u\|_H^2}{\|u\|_{H'}}, u \in H' \; et \; \forall j \leq i, \; \langle u, u_j\rangle'_H = 0\}$.

Proposition 1.6.1. [22] (Dauxois et Pausse , 1982)
Soit $(\{\lambda_i\}_{i\in I^*}, \{U_i\}_{i\in I^*})$, une ACP "pas à pas" de l'opérateur compact U, le couple $(\{\lambda_i\}_{i\in I^*}, \{\frac{U^*u}{\|U^*u\|_H}\}_{i\in I^*}$ une ACP "pas à pas " de U^* dite associée à la précédente .
où $\{\lambda_i\}_{i\in I^*}$ la suite des valeurs principales, $\{f_i\}_{i\in I^*} = \frac{U^*u}{\|U^*u\|_H}\}$ la suite des composantes principales normalisées et $\{u_i\}_{i\in I^*}$ la suite des facteurs principaux de l'opérateur U .
Par la proposition précédente on peut chercher l'ACP de U en faisant l'analyse spectrale de V ou bien celle de W. On peut écrire l'opérateur V comme suit

$$V = \sum_{i\in I^*} \lambda_i u_i \otimes u_i$$

$\forall \; (x,y) \in H^2, \; y \otimes x$ est l'opérateur définie par

$$y \otimes x f(x) = \langle y, f\rangle_H x.$$

1.7 Affinité L^2 entre deux densités de probabilité

La notion d'affinité est un concept et outils mathématiques trés important dans la suite de ce mémoire, il est à la base de la théorie de l'ACP de densités.

Définition 1.7.1. (Qannari, 1983) Soit f, g deux densités de probabilités de carrées intégrables sur $(\mathbb{R}^p, B_{\mathbb{R}^p})$, on appelle mesure d'affinité L^2 entre f et g la quantité suivante :

$$\langle f, g\rangle = \int_{\mathbb{R}^p} f(x)g(x)dx$$

Cette affinité n'est que le produit scalaire classique sur H.

1.7.1 Mesure d'affinité de quelques lois usuelles

Cas de deux densités gaussiennes

Proposition 1.1. [9] Soit f et g deux densités gaussiennes dans H, de paramétres respectifs (μ_f, Σ_f) et (μ_g, Σ_g). La mesure d'affinité L^2 de f et g est égale à

$$\langle f, g\rangle = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{|\Sigma_f + \Sigma_g|^{\frac{1}{2}}} \exp[-\frac{1}{2}(\mu_f - \mu_g)^t(\Sigma_f + \Sigma_g)^{-1}(\mu_f - \mu_g)].$$

Cas de deux densités de loi uniformes

La mesure d'affinité L^2 entre deux densités de lois uniformes sur deux domaines ouverts et bornés de $\mathbb{R}^p, D_f, D_g$

$$f = \frac{1}{vol(D_f)} I_{D_f}, \quad g = \frac{1}{vol(D_g)} I_{D_g}.$$

est définie par

$$\langle f, g\rangle = \frac{vol(D_f \cap D_g)}{vol(D_f)vol(D_g)}.$$

I_D est l'indicatrice du domaine D et $vol(D)$ son volume.

Lois gamma unidimensionnelle

La mesure d'affinité L^2 entre deux densités f et g de lois gamma $G(r_f, a_f)$ et $G(r_g, a_g)$ respectivement est définie par

$$\langle f, g\rangle = \frac{a_f^{r_f+1} a_g^{r_g+1}}{(a_f + a_g)^{r_f+r_g+1}} \frac{\Gamma(r_f + r_g + 1)}{\Gamma(r_f + 1)\Gamma(r_g + 1)}.$$

avec $\Gamma(a) = \int_0^{+\infty} xe^{-ax}dx$.

Lois de poisson

Soient θ_f et θ_g deux paramètres de deux lois de poisson. La mesure d'affinité L^2 entre les deux lois vaut

$$\frac{e^{\theta_f} e^{\theta_g}}{e^{\theta_f + \theta_g}}.$$

Lois binomiale

Soient $B(n, p_f)$ et $B(n, p_d)$ deux lois binomiale, alors l'affinité L^2 est égale à

$$[(1 - p_f)(1 - p_g) + p_f p_g]^n.$$

Chapitre 2

Analyse en composantes principales de densités de probabilité

2.1 Introduction, position du problème

En théorie des probabilités et en statistique, une densité de probabilité est une fonction qui permet de représenter une loi de probabilité sous forme d'une integrale.
Formellent, si une loi de probabilité posséde une densité f, intégrable, positive sur $\mathbb{R}^p$, et si A est un Borélien de $\mathbb{R}$ alors

$$P(A) = \int_A f(x)dx.$$

En statistique une densité de probabilité peut être utiliser pour estimer des caractéristque d'une population de données. En analyse des données fonctionnelles une densité de probabilité joue le rôle d'une observation relative à un échantillon de données. Considérons alors le cas où on observe plusieurs échantillons de données, on s'intéresse alors à une discription globale de ces échantillons au moyen de leurs densités associés. Un exemple pratique de ce type de données sont les données sensorielles, où l'objectif est de construire des cartographies afin d'apprécier les ressemblances et les défférences entre les produits. L'exposer théorique pour réaliser cet objectifs est donné dans la partie suivante.

Soit $f_1, ..., f_L$, L densités de probabilités dans $L^2(\mathbb{R}^p)$, l'objectifs est de trouver une représentation reprochés des L densités dans un sous-espace de faible dimention.

Notons pour cela P_g le projecteur orthogonal sur le sous espace engendré par le vecteur g de H.

1. Cherchons $g_1 = \sum_{t=1}^{L} \alpha_t^{(1)} f_t$ de norme 1 dans H, qui minimise la quantite

$$I_{g_1} = \sum_{t=1}^{L} \|P_{g_1}(f_t) - f_t\|^2$$

2. Cherchons $g_2 = \sum_{t=1}^{L} \alpha_t^{(2)} f_t$ de norme 1 dans H, orthogonale à g_1 qui minimise la quantite suivante :

$$I_{g_2} = \sum_{t=1}^{L} \|P_{g_2}(f_t) - f_t\|^2.$$

Ainsi de suite .
Les fonction $g_1, g_2, ...$ ainsi obtenues, ne sont pas des densités de probabilités, mais constituent un système orthonormal dans H.
Soit U l'opérateur compact défini sur $\mathbb{R}^L$ par :

$$\forall v \in \mathbb{R}^L, \quad U_v = \sum_{t=1}^{L} v_t f_t, \qquad v = (v_1, ..., v_L).$$

sont adjoint U^* est définit par

$$\forall g \in H,\ U_g^* = (\langle g, f_1\rangle, ..., \langle g, f_L\rangle)^t,$$

car

$$\langle v, U_g^*\rangle_{\mathbb{R}^L} = \langle g, U_v\rangle_H = \sum_{t=1}^{L} \langle f_t, g\rangle v_t.$$

On a la définition suivante .

Définition 2.1. On appelle ACP de densités de probabilités, l'ACP "pas à pas " de l'opérateur U.

Par application de théorème de Phytagore la minimisation de $I_{g_1} = \sum_{t=1}^{L} \|P_{g_1}(f_t) - f_t\|^2$ est équivalente à la maximisation de la quantitie suivante

$$I'_{g_1} = \sum_{t=1}^{L} \|P_{g_1}(f_t)\|^2 = \sum_{t=1}^{L} \langle f_t, g_1\rangle^2 = \|U^* g_1\|^2_{\mathbb{R}^L}$$

avec $\|.\|_{\mathbb{R}^L}$ est la norme usuelle de $\mathbb{R}^L$.

D'autre part

$$\|U^* g_1\|^2_{\mathbb{R}^L} = \langle U^* g_1, U^* g_1\rangle_{\mathbb{R}^L} = \langle g_1, U \circ U^*\ g_1\rangle$$

$$\max_{\|g_1\|=1} (I' g_1) = \max_{\|g_1\|=1} \langle g_1, V g_1\rangle$$

De cette dernière définition on peut déduire que l'ACP de ces densités équivalente à l'analyse spectrale de l'opérateur autoadjoint $W = U^* \circ U$. Dans la base canonique $e_1, ..., e_L$ de $\mathbb{R}^L$, la matrice W s'écrit

$$W_{t,s} = \begin{pmatrix} \langle f_1, f_1\rangle & ... & \langle f_1, f_s\rangle & ... & \langle f_1, f_L\rangle \\ \langle f_t, f_1\rangle & ... & \langle f_t, f_s\rangle & ... & \langle f_t, f_L\rangle \\ \langle f_L, f_1\rangle & ... & \langle f_L, f_s\rangle & ... & \langle f_L, f_L\rangle \end{pmatrix} \tag{2.1}$$

Remarque 2.1.1.

1. Si u de $\mathbb{R}^L$ est un vecteur propre de l'opérateur W associé à la valeur propre non nulle λ , alors $g = \frac{Uu}{\sqrt{\lambda}}$ est un vecteur propre de V associé à la même valeur propre non nulle λ.
2. Les vecteurs propres de l'opérateur V sont les facteurs principaux, et leurs images par U^* sont les composantes principales .

2.1.1 Formule de reconstitution des densités de probabilité

Soit $g_1, g_2, ..., g_T$ le système de vecteur propre de l'opérateur V, où T désigne le nombre de valeurs propre non nulles de V (resp W), alors chaque densité f_t s'écrit comme suit

$$f_t = \sum_{j=1}^{T} \langle f_t, g_j \rangle_H g_j.$$

La coordonnée $\langle f_t, g_j \rangle$ de f_t suivant g_j, est égale à la t-ième composante du vecteur $U^* g_j$. De la relation

$$U^* g_j = \frac{1}{\sqrt{\lambda_j}} (U^* \circ U) u_j = \frac{1}{\sqrt{\lambda_j}} W u_j = \sqrt{\lambda_j} u_j.$$

On déduit

$$\langle f_t, g_j \rangle_H = \sqrt{\lambda_j} u_{j,t}. \tag{2.2}$$

$$f_t = \sum_{j}^{L} \sqrt{\lambda_j} u_{j,t} g_j. \tag{2.3}$$

$$g_j = \frac{1}{\sqrt{\lambda_j}} \sum_{t=1}^{T} u_{j,t} f_t. \tag{2.4}$$

Pour obtenir une représentation approchée du nuage initial il suffit de tronquer la relation (2.4) .

2.2 Qualité globale de l'ACP

On mesure la qualité globale de l'ACP par la somme des proportions d'inerties expliquées par les axe retenus, l'axe j expliquant une quantité d'inertie égale à

$$\frac{\lambda_j}{\sum_{r=1}^{T} \lambda_r} .$$

2.2.1 Qualité de représentation de f_t suivant g_j

Elle est égale à

$$\frac{\|P g_j(f_t)\|}{\|f_t\|}.$$

2.2.2 ACP normée

L'ACP normée des densités de propabilité consiste à diviser chaque densité par sa norme associée dans H, cette ACP conduit alors à diagonaliser la matrice de terme generale

$$W_{s,t}^{(N)} = \frac{1}{\|f_t} \frac{1}{\|f_s\|} \langle f_t, f_s \rangle_H.$$

Remarque 2.2.1. Cette normalisation conserve dans H les angles entre les densités mais déforme leur distances .

2.2.3 ACP centrée

Pour obtenir une représentation approchée qui restitue les distances entre les densités, on definit un autre nuage dont le centre de gravité est lui même l'origine de l'espace vectoriel où les densités sont representées. Théoriquement elle consiste à prendre comme nuage l'ensemble des fonctions $f_t^{(c)} = f_t - f_u$ dans $L^2(\mathbb{R}^p)$, où $f_u = \frac{1}{L}\sum_{s=1}^{L} f_s$. L'ACP de ce nuage conduit à diagonaliser la matrice $W_{t,s}^{(c)}$, de terme général

$$W_{t,s}^{(c)} = \langle f_t - f_u, f_s - f_u \rangle.$$

L'opérateur de covariance associé s'ecrit

$$w^{(c)} = \sum_{j=1}^{T} (f_t - f_u) \otimes (f_t - f_u).$$

Les densites $f_t, t \in \{1, ..., L\}$ s'écrivent dans la base des fonctions propres $g_1^{(c)}, g_2^{(c)}, ...$ de l'opérateur $V^{(c)}$ comme suit

$$f_t = f_u + \sum_{j=1}^{T} \sqrt{\lambda_j^{(c)}} u_{t,j}^{(c)} g_j^{(c)}$$

avec $\lambda_j^{(c)}$ et la j-iéme plus grande valeur propre et de $u_j^{(c)}$ est le vecteur propre associé à $W^{(c)}$.

2.2.4 ACP de 20 densités gaussiennes

Considérons une famille de 20 densités gaussienne à deux dimentions

$$f_t \equiv N(\mu_t, \Sigma_t),\ \mu_t = (t, t)^t,\ \Sigma_t = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix}, t = 1...20. \tag{2.5}$$

La mesure d'affinité entre f_t et f_r est égale à :

$$\langle f_t, f_r \rangle = \frac{1}{(2\pi)^{10}} \frac{1}{|\Sigma_t + \Sigma_r|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mu_t - \mu_r)^t (\Sigma_t + \Sigma_r)^{-1}(\mu_t - \mu_r)\right],$$

$$= \frac{1}{(2\pi)^{10}}(t + r) \exp(-\frac{(t-r)^2}{t+r}).$$

En effectuant une ACP sur les 20 densités on obtient les résultats suivants :

La liste des valeurs propres non nulles et les pourcentages d'inertie
Les quatres plus grandes valeurs propres ainsi que les pourcentage d'inerties associés sont données dans le tableau (2.1) :

λ	valeurs propres	pourcentages d'inertie
λ_1	0.117	40.9
λ_2	0.065	22.7
λ_3	0.044	15.4
λ_4	0.029	10.1

TABLE 2.1 – Les 4 premieres valeurs propres et les pourcentage d'inerties associés

L'inertie globale est égale à

$$I_g = \sum_{r=1}^{20} \lambda_r = 0.285 \tag{2.6}$$

et montre que 3 axes sont suffissants pour reproduire 80% d'informations

Contribution des densités à l'inertie
Le tableau suivant nous montre les densités qui contribuent le plus à l'inertie sur les quatres premiers axes.

densités	axe 1	axe 2	axe 3	axe 4
1	51.4	17.0	10.2	8.7
2	27.8	0.0	1.7	6.8
3	12.0	6.0	9.6	10.4
4	5.0	13.0	7.9	1.6

TABLE 2.2 – Contribution des densités à l'inertie

Les coordonnées des densités sur les deux premiers axes
Le tableaux suivant nous donne les coordonnés des 20 densités sur l'axe1 et l'axe2.

Densité	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}
axe.1	0.245	0.180	0.119	0.077	0.050	0.032	0.021	0.014	0.009	0.006
axe.2	0.105	0.004	-0.063	-0.092	-0.099	-0.094	-0.084	-0.072	-0.061	-0.050
Densité	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_{16}	f_{17}	f_{18}	f_{19}	f_{20}
axe1	0.004	0.003	0.002	0.001	0.001	0.001	0.000	0.000	0.000	0.000
axe.2	-0.041	-0.033	-0.027	-0.021	-0.017	-0.013	-0.010	-0.008	-0.006	-0.005

TABLE 2.3 – les coordonnées des densités sur les deux premiers axes

La projection des densités sur les quatres premiesrs plans principaux est donnée par la figure suivante.

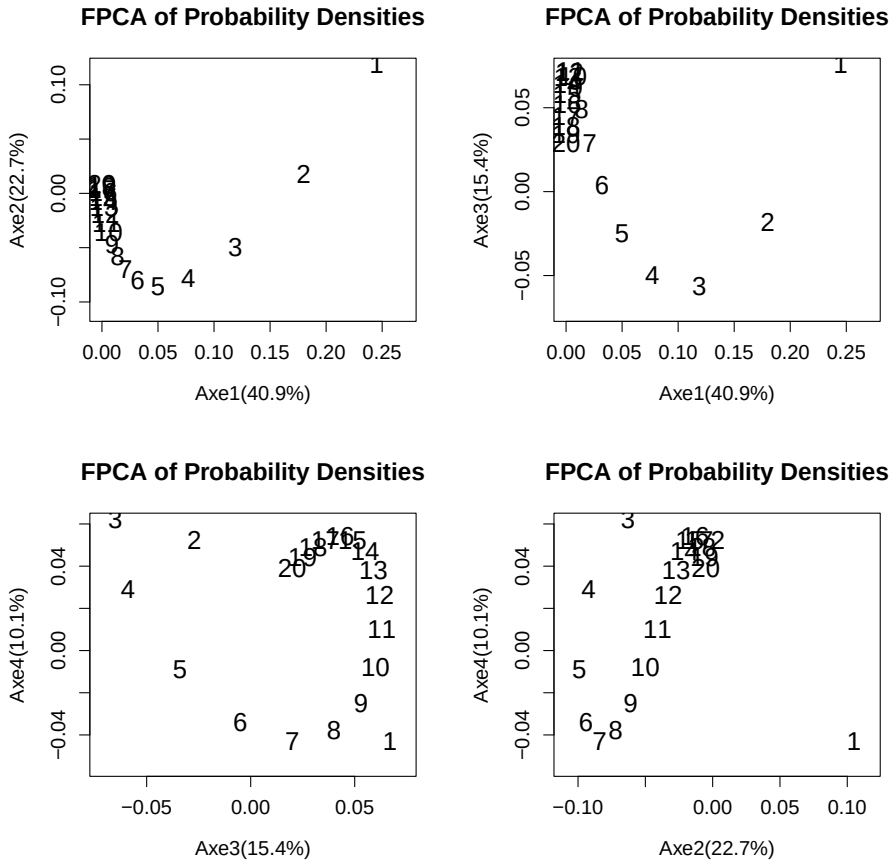

FIGURE 2.1 – projection des densites sur les quatres premiers plans

2.2.5 ACP de 20 densités issues d'un mélange de deux gaussiennes

Considérons une famille 20 de densité $f_1, ..., f_{20}$ où

$$f_t = \alpha f_t^{(1)} + (1-\alpha) f_t^{(2)},\ \alpha \in [0,1] \tag{2.7}$$

avec

$$f_t^{(1)} \equiv N(\mu_t^{(1)}, \Sigma_t^{(1)}) \tag{2.8}$$

$$f_t^{(2)} \equiv N(\mu_t^{(2)}, \Sigma_t^{(2)}) \tag{2.9}$$

On réalise une ACP sur ces 20 densités avec :

$$\mu_t^{(1)} = (t,t)^t,\ \mu_t^{(2)} = (2t,2t)^t, \alpha = 0.5 \tag{2.10}$$

$$\Sigma_t^{(1)} = \Sigma_t^{(2)} = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} \tag{2.11}$$

Ce qui conduit à diagonaliser la matrice W de terme général

$$\begin{aligned} W_{tr} &= \langle f_t, f_r \rangle & (2.12)\\ &= \langle \alpha f_t^{(1)} + (1-\alpha) f_t^{(2)}, \alpha f_r^{(1)} + (1-\alpha) f_r^{(2)} \rangle & (2.13)\\ &= \alpha^2 \langle f_t^{(1)}, f_r^{(1)} \rangle + \alpha(1-\alpha)\langle f_t^{(1)}, f_r^{(2)} \rangle + \alpha(1-\alpha)\langle f_t^{(2)}, f_r^{(1)} \rangle + (1-\alpha)^2 \langle f_t^{(2)}, f_r^{(2)} \rangle \end{aligned}$$

Les résultats obtenus sont donnés dans la partie suivante.

Les quatres premieres valeurs propres et pourcentages d'inerties sont donnés dans le tablaux suivante

λ	valeurs propres	pourcentages d'inertie
λ_1	0.087	48.3
λ_2	0.043	23.9
λ_3	0.023	12.8
λ_4	0.009	5.0

TABLE 2.4 – Valeurs propres et pourcentages d'inertie

La contribution des quatres premiers densités sur les quatres premiers axes est donnée par le tableau (2.5) :

Densités	axe 1	axe 2	axe 3	axe 4
1	59.6	19.8	12.1	6.0
2	23.9	1.0	11.9	25.0
3	9.3	9.7	12.6	1.6
4	3.8	13.4	4.4	2.2

TABLE 2.5 – La contribution des densités sur les premiers plans principaux

La présentation graphique sur les plans (1,2), (1,3), (3,4) et (2,4)
Les 4 graphiques de la figure (2.2.6) nous montrent l'allure des densités sur les 4 premiers plans.

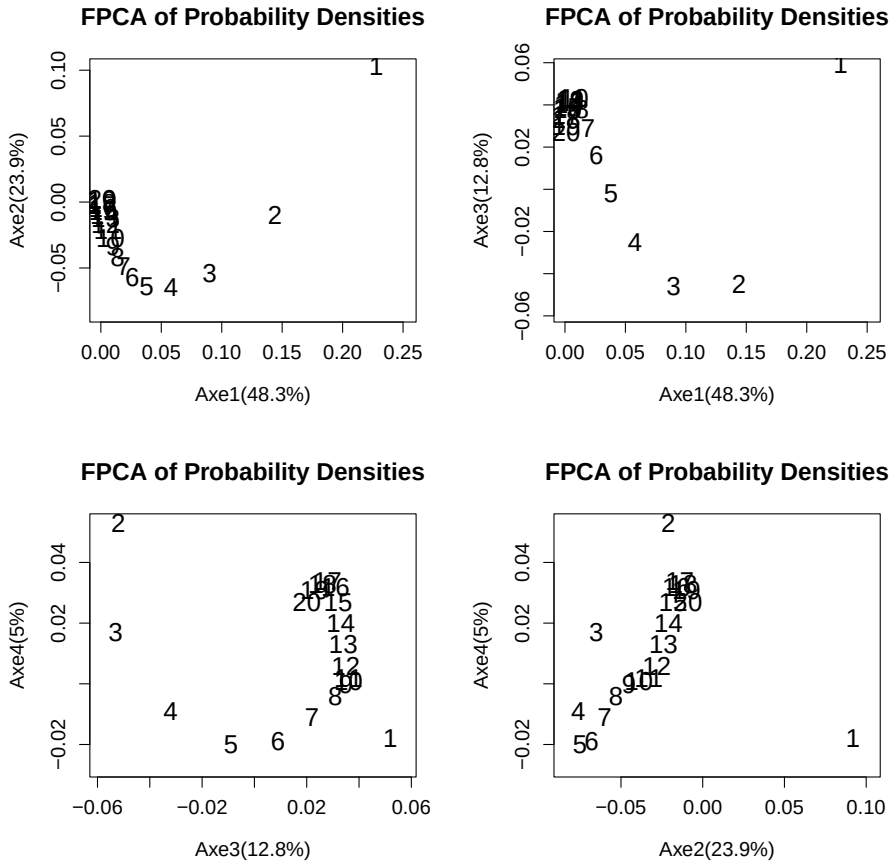

FIGURE 2.2 – Projection des densités sur les premiers plans principaux

2.3 ACP de densités estimées par la méthode du noyau multidimenionnel

L'estimation des densités peut être faite par une méthode non paramétrique ou paramétrique. L'estimation paramétrique repose sur l'hypothèse de l'appartenance de la fonction de densité inconnue à une famille paramétrique (par exemple les lois normales multidimentionnelles) connue et vise ensuite à estimer les paramètres par les techniques d'estimation classique voir[9]. Les méthodes non paramétriques ne repose sur aucune hypothèse sur l'appartenance de la fonction de densité inconnue à une famille connue. Parmi ces méthodes on s'interesse dans ce travail à l'estimation par la méthode de noyau ([55]) .En appliquant cette méthode dans le cadre de l'analyse en composantes principales ([38], [40],[63], [64], [67]), on obtient grâce à la bilinéarité de la mesure d'affinité une estimation de produit scalaire $\langle f, g\rangle$. La suite de cette section est concacrée à l'étude assymptotique des sorties de l'ACP de densités lorsque elles sont estimées par noyau.

Soit $X_{t1}, ..., X_{t_{nt}}$ un échantillon (iid) de densité inconnue f_t , avec $t \in \{1, ..., T\}$. $\hat{f}_1, ..., \hat{f}_T$ sont resprctivement les estimateurs à noyau des densités $f_1, ..., f_T$, alors le terme général de l'estimateur $\hat{W}$ de la matrice W déduit de la bilinéarité du produit scalaire est donnée par

$$\hat{W}_{tr} = \frac{1}{n_t n_r} \sum_{i=1}^{n_t} \sum_{j=1}^{n_r} \int_{\mathbb{R}^l} K_{H_{n_t}}(z - X_{ti}) K_{H_{nr}}(z - X_{rj}) dz.$$

Par définition , l'estimation non-parametrique de l'ACP de densités est l'analyse spectrale de la matrice $\hat{W}$.

2.3.1 Propriétés asymptotiques

Sous la condition de la convergence en moyenne quadratique integrée des estimateurs à noyau $\hat{f}_t$ vers f_t $(t = 1, ..., L)$ nous avons les résultats suivantes.

Théorème 2.3.1. (Yousfi et al)

$$\|\hat{W}\|_2 \xrightarrow{p} \|W\|_2$$

$$\|\hat{W} - W\| \xrightarrow{p} 0$$

où $\|\cdot\|_2$ eat la norme matricielle de frobenious

Soient $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq ... \geq \hat{\lambda}_L > 0$, les valeurs propres non nulles de $\hat{W}$ rangées dans l'ordre décroissant et répétées autant de fois que leurs ordres de multiplicités, et $\hat{u}_1, ..., \hat{u}_L$ la suite des vecteurs propres orthonormés associés .

Corollaire 2.3.1. (Yousfi et al)

$\hat{\lambda}_r$ converge en probabilités vers λ_r . De plus, si l'ordre de multiplicités des valeurs propres égale à 1, alors le signe de $\hat{u}_r$ peut être choisis tel que :

$$\hat{u}_t \xrightarrow{p} u_r.$$

2.4 Influence de la matrice de lissage sur la qualité de l'estimation de l'ACP

Supposons dans tous ce qui suit que la taille des données est fixée.
Nous avons alors les résultat de convergence suivant

Théorème 2.4.1. (Yousfi et al)

Si $H_t = hI_p$, on a

(a) Si $h \to 0_+$

$$\langle \hat{f}_t, \hat{f}_r \rangle \to 0$$

$$\|\hat{f}_t\| et \|\hat{f}_r\| \to +\infty$$

$$\frac{\langle \hat{f}_t,\ \hat{f}_r \rangle}{\|\hat{f}_t\| \|\hat{f}_r\|} \to 0.$$

(b) Si $h \to \infty$

$$\langle \hat{f}_t \hat{f}_r \rangle \to 0$$

$$\frac{\langle \hat{f}_t \hat{f}_r \rangle}{\|\hat{f}_t\| \|\hat{f}_r\|} \to 1.$$

2.4.1 Produit de noyau gaussien

Si K est le noyau gaussien réel, le noyau gaussien produit est définie par

$$K_p(z) = \prod_{k=1}^{p} K(z_k), \forall z \in (z_1, ..., z_p). \tag{2.14}$$

L'estimateur de la mesure d'affinité L^2 dans ce cas vaut

$$\langle \hat{f}_t, \hat{f}_r \rangle = \frac{1}{n_t n_r} \sum_{i}^{n_t} \sum_{j}^{n_r} \prod_{k=1}^{p} \frac{1}{\sqrt{2\pi}} \frac{1}{h(S_{tk}^2 + S_{rk}^2)^{\frac{1}{2}}} \exp\left(-\frac{(X_{ti}^{(k)} - X_{rj}^{(k)})^2}{2h^2(S_{tk}^2 + S_{rk}^2)}\right)$$

avec $H_t = h\ diag(S_{t1}, ..., S_{tp})\ t = 1, ..., p$ où S_{tj} est la variance de j-éme variable . On a alors le théorème suivant.

Théorème 2.4.2. (Yousfi et al)

Presque sûrement, nous avons

$$\lim_{h \to \infty} \frac{\langle \hat{f}_t, \hat{f}_r \rangle}{\|\hat{f}_t\| \|\hat{f}_r\|} = \prod_{k=1}^{p} \left(\frac{2S_{tk}S_{rk}}{S_{tk}^2 + S_{rk}^2}\right)^{\frac{1}{2}}.$$

2.4.2 Noyau gaussienne sphérique

On définit le noyau gaussienne sphérique par

$$K_p(z) = (\frac{1}{2\pi})^{\frac{p}{2}} \exp(\frac{-z'z}{2}).$$

Donc l'estimateur de la mesure d'affinité L^2 entre f_t et f_r est égale à

$$\langle \hat{f}_t, \hat{f}_r \rangle = \frac{1}{n_t n_r} \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{|\hat{V}_t + \hat{V}_r|^{\frac{1}{2}}} \frac{1}{h^p} \sum_i^{n_t} \sum_j^{n_r} \exp(-\frac{1}{2h^2}(X_{ti} - X_{rj})'(\hat{V}_t + \hat{V}_r)^{-1}(X_{ti} - X_{rj})). \tag{2.15}$$

Où $H_t = h\hat{V}_t^{\frac{1}{2}}$ est la matrice de lissage. Et $\hat{V}_t^{\frac{1}{2}}$ la racine carrée de la matrice variance covariance empirique.

Théorème 2.4.3. (Yousfi et al)

On a presque sûrement

$$\lim_{h \to +\infty} \frac{\langle \hat{f}_t, \hat{f}_r \rangle}{\|\hat{f}_t\| \|\hat{f}_r\|} = 2^{\frac{p}{2}} \frac{|\hat{V}|^{\frac{1}{4}} |\hat{V}_r|^{\frac{1}{4}}}{|\hat{V}_t + \hat{V}_r|^{\frac{1}{2}}}.$$

2.4.3 Exemple de 20 densités gaussiennes estimées par la méthode du noyau gaussien

Considérons l'exemple des 20 densités gaussienne $f_t \equiv N(\mu_t, \Sigma_t)$, $t = 1, ..., 20$ de paramétres

$$\mu_t = (t, t), \Sigma_t = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix}. \tag{2.16}$$

Pour illustrer la convergence de l'ACP estimé vers l'ACP théorique on réalise pour differentes tailles de $(n = 10, 30, 100)$ une ACP sur les 20 densités gaussienne estimées par la méthode du noyau avec un noyau gaussien bivarié. Les résultats obtenus sont données dans la partie suivante :

Le tableau suivant nous montre la convergance des valeurs propres estimées vers les valeurs réelles correspandantes.

λ	n=10	n=30	n=100	val réel
λ_1	0.081	0.095	0.092	0.117
λ_2	0.061	0.056	0.053	0.065
λ_3	0.040	0.040	0.036	0.044
λ_4	0.030	0.024	0.022	0.029

TABLE 2.6 – Les valeurs propres

Le tableau (2.7) suivant nous montre la convergance des inerties expliqués par les axes retunus vers les valeures réelles sont

axe	n=10	n=30	n=100	val réel
axe 1	27.2	36.6	39.5	40.9
axe 2	20.5	21.6	22.7	22.7
axe 3	13.5	15.4	15.5	15.4
axe 4	10.1	9.2	9.4	10.1

TABLE 2.7 – le pourcentage d'inertie

Les contributions des densités à l'inertie pour les différentes valeurs de n est donnée dans les tableaux (2.8, 2.9, 2.10), et montre une bonne aproximations des contribution réelles (le tableau 2.2).

densités	axe 1	axe 2	axe 3	axe 4
1	48.0	14.5	2.4	2.1
2	21.4	3.2	0.1	0.0
3	10.8	0.0	1.4	2.1
4	7.7	2.1	4.5	5.5

TABLE 2.8 – Contribution des densités à l'inertie pour n=10

densités	axe 1	axe 2	axe 3	axe 4
1	31.9	17.5	15.3	7.3
2	26.8	3.6	0.4	0.4
3	17.9	0.4	4.7	6.0
4	11.7	3.7	9.0	4.5

TABLE 2.9 – Contribution des densités à l'inertie pour n=30

densités	axe 1	axe 2	axe 3	axe 4
1	39.3	14.6	13.1	11.3
2	25.5	2.0	0.0	1.8
3	16.6	0.3	3.6	7.6
4	9.2	5.5	10.3	1.3

TABLE 2.10 – Contribution des densités à l'inertie pour n=100

La présentation graphique sur le plans (1,2)
Les graphiques de la figure (2.3) montrent la convergence de l'ACP estimé par noyau vers l'ACP réelle sur le plans (1,2).

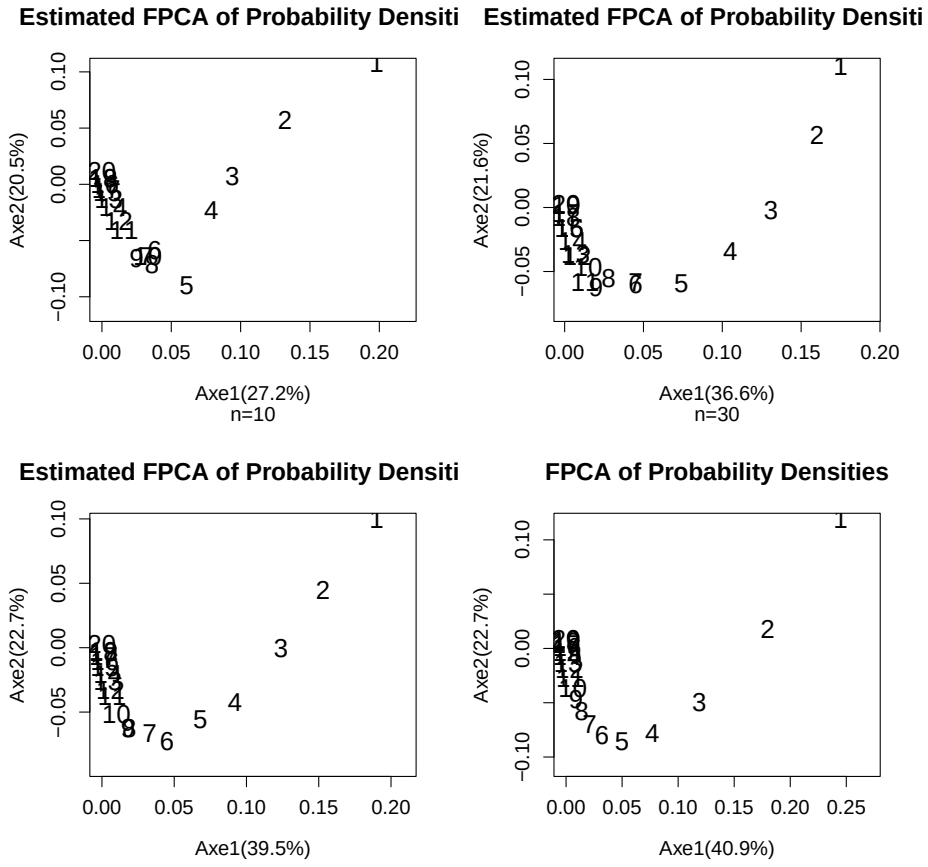

FIGURE 2.3 – la projection sur le plans (1,2)

Les graphiques de la figure (2.4) montrent la convergence de l'ACP estimé par noyau vers l'ACP réelle sur le plans (1,3).

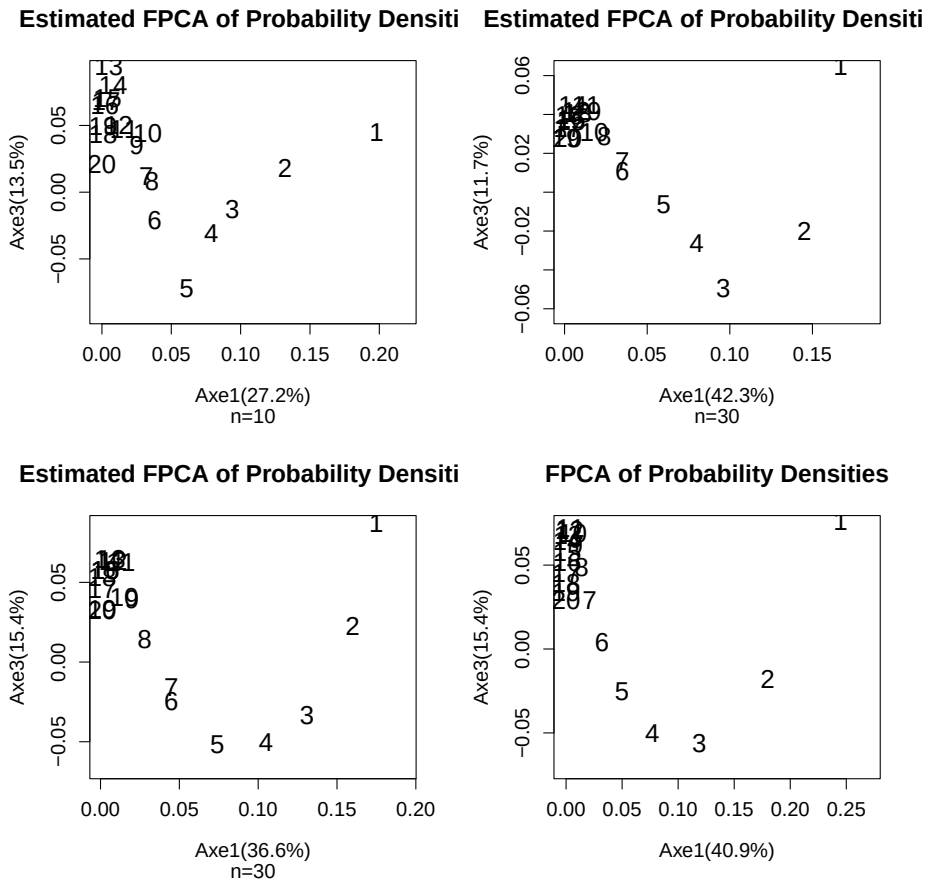

FIGURE 2.4 – la projection sur le plan (1,3)

La présentation graphique sur le plans(2,4)

Les graphiques de la figure (2.5) montrent la convergence de l'ACP estimé par noyau vers l'ACP réelle sur le plans (2,4).

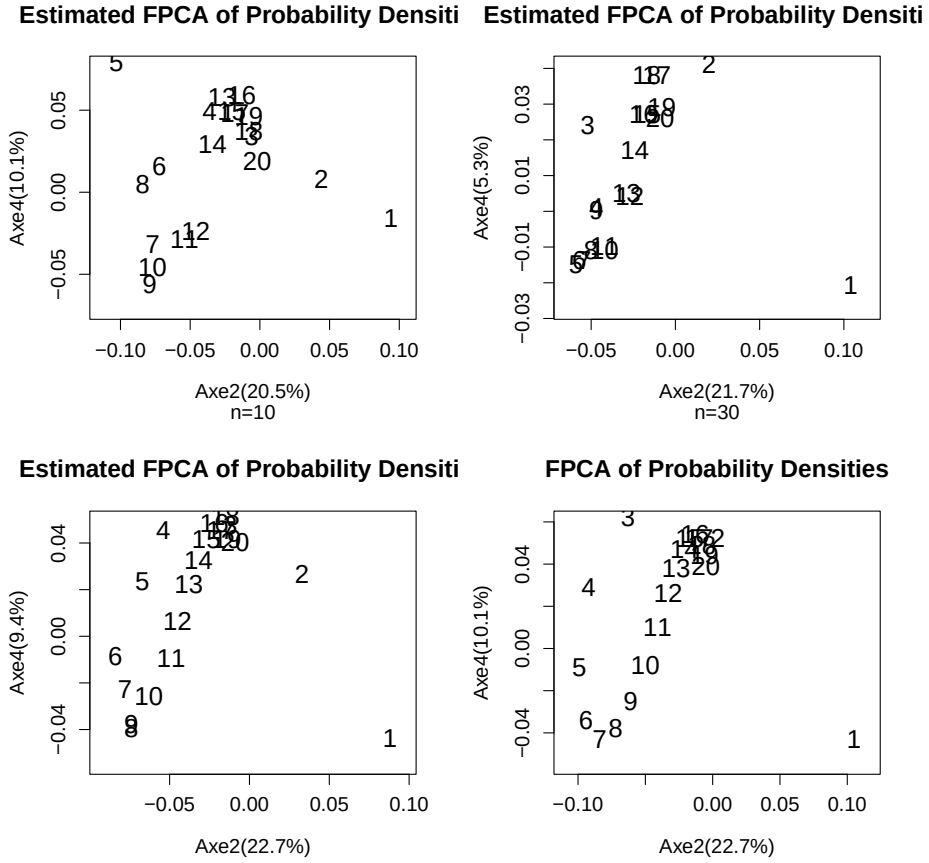

FIGURE 2.5 – la projection sur le plan (2,4)

La présentation graphique sur le plans (3,4)

Les graphiques de la figure (2.6) montrent la convergence de l'ACP estimé par noyau vers l'ACP réelle sur le plans (3,4).

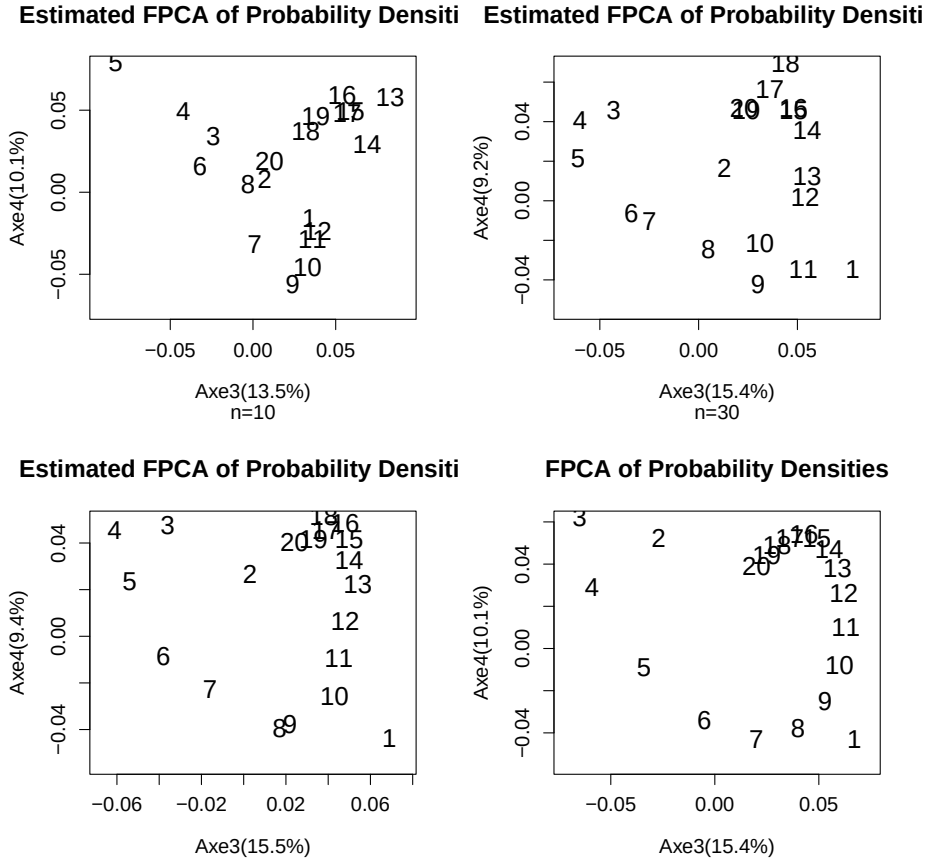

FIGURE 2.6 – la projection sur plan (3,4)

2.4.4 Exemple de 20 densités issues d'un mélange de deux gaussiennes estimées par la méthode du noyau gaussien

Considérons une famille 20 de densité $f_1, ..., f_{20}$ où

$$f_t = \alpha f_t^{(1)} + (1-\alpha) f_t^{(2)},\ \alpha \in [0,1] \tag{2.17}$$

avec

$$f_t^{(1)} \equiv N(\mu_t^{(1)}, \Sigma_t^{(1)}) \tag{2.18}$$

$$f_t^{(2)} \equiv N(\mu_t^{(2)}, \Sigma_t^{(2)}) \tag{2.19}$$

On réalise une ACP sur ces 20 densités avec :

$$\mu_t^{(1)} = (t,t)^t,\ \mu_t^{(2)} = (2t,2t)^t, \alpha = 0.5 \tag{2.20}$$

$$\Sigma_t^{(1)} = \Sigma_t^{(2)} = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} \tag{2.21}$$

On a alors les résultats suivantes .
Le tableau (2.11) suivant nous montre la convergance pour diffèrentes taille de n des 4 premieres valeurs propres estimé vers les valeurs propres réelles.

valeur propre	n=10	n=30	n=100	val propre réelles
λ_1	0.061	0.072	0.066	0.087
λ_2	0.038	0.037	0.034	0.043
λ_3	0.021	0.020	0.018	0.023
λ_4	0.012	0.009	0.007	0.009

TABLE 2.11 – Les quatres premiers valeurs propres

Le tableau (2.12) suivant nous illustre l'évolution en fonction de n des porcentage d'inertie expliquées, et montre ce passage leurs convergences versl'inertie réelle correspondante.

axe	n=10	n=30	n=100	inertie réelle
axe1	31.5	42.3	45.3	48.3
axe2	19.6	21.7	23.3	23.9
axe3	10.8	11.7	12.3	12.8
axe4	6.2	5.3	4.8	5.0

TABLE 2.12 – Pourcentage d'inertie pour les quatres premiers axes

Les contributions des densités à l'inertie des axes, en fonction de n sont données par les tableaux (2.13, 2.14, 2.15)

Densité	axe 1	axe 2	axe 3	axe 4
1	54.8	20.0	8.7	2.5
2	20.3	0.3	2.3	2.5
3	8.6	4.2	8.2	3.7
4	7.5	4.5	8.8	1.7

TABLE 2.13 – La contribution des quatres premiers plans principaux pour n=10

Densités	axe 1	axe 2	axe 3	axe 3
1	38.8	29.0	16.2	7.3
2	29.0	1.1	3.9	15.6
3	12.8	7.2	16.1	4.6
4	8.9	6.0	5.7	0.1

TABLE 2.14 – La contribution des quatres premiers plans principaux pour n=30

Densités	axe 1	axe 2	axe 3	axe 3
1	46.6	23.2	19.0	7.5
2	24.6	0.2	6.3	22.9
3	12.8	3.4	13.0	2.1
4	7.7	8.0	9.4	0.1

TABLE 2.15 – La contribution des quatres premiers plans principaux pour n=100

La présentation graphique sur le plans (1,2)

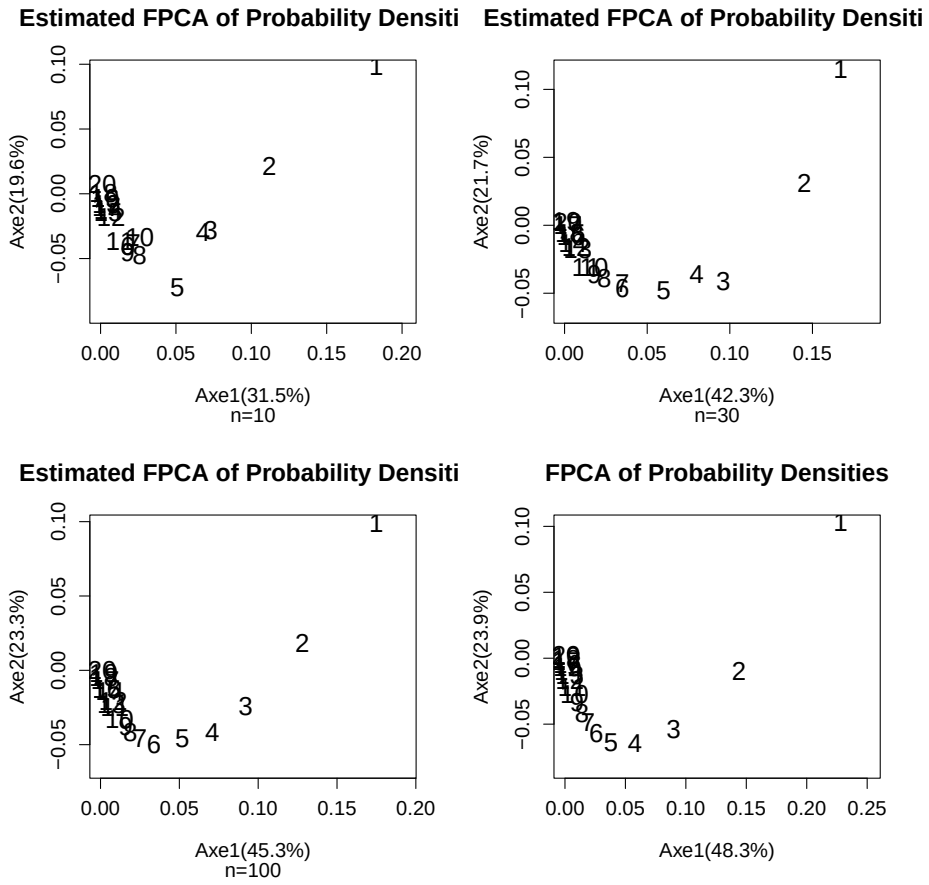

FIGURE 2.7 – La projection des densités sur le plans(1,2)

La présentation graphique sur le plans (1,3)

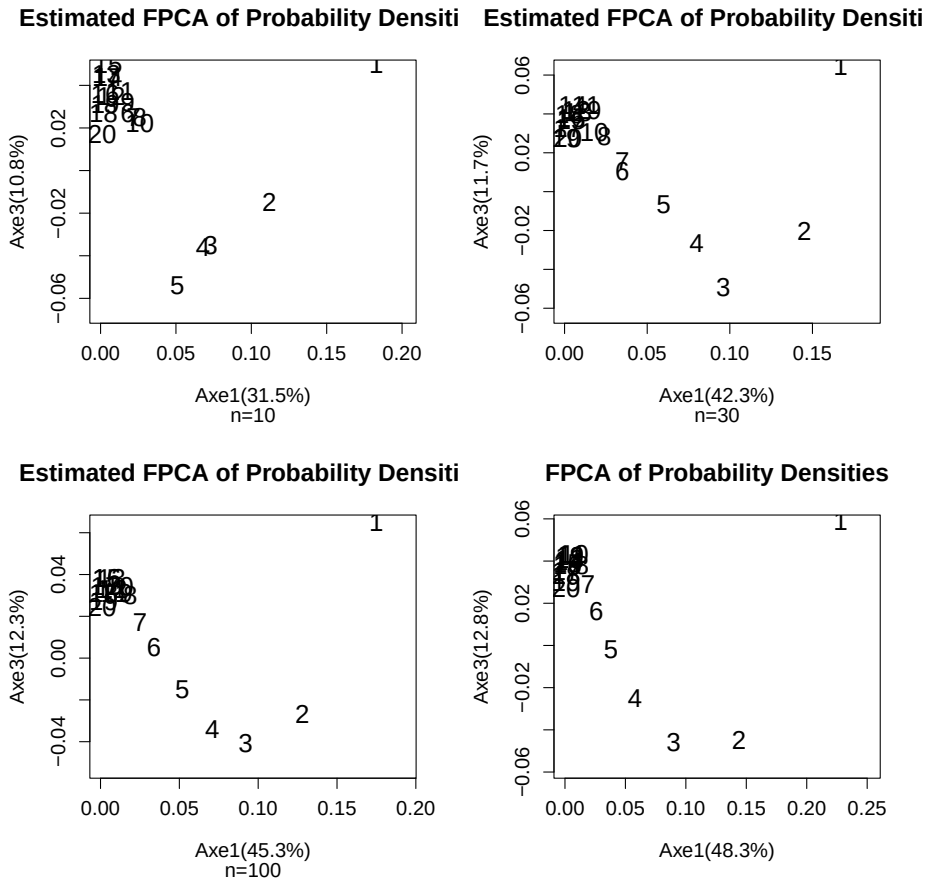

FIGURE 2.8 – La projection des densités sur le plans(1,3)

La présentation graphique sur le plans (3,4)

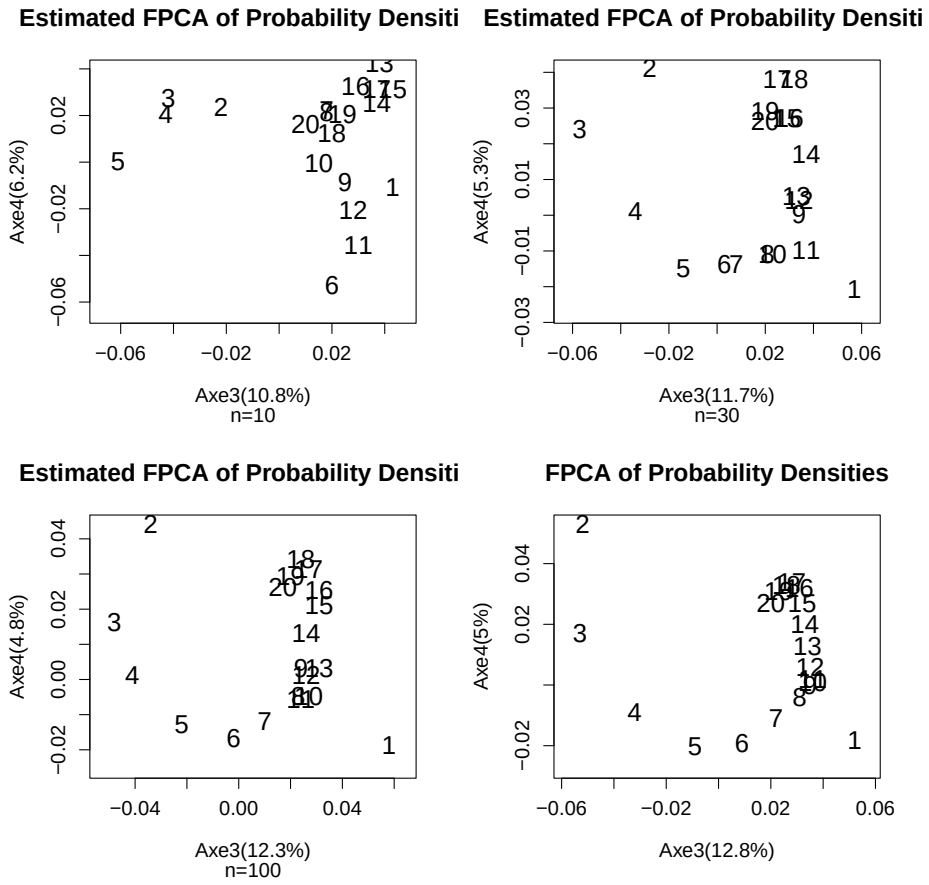

FIGURE 2.9 – La projection des densités sur le plans (3,4)

La présentation graphique sur le plans (2,4)

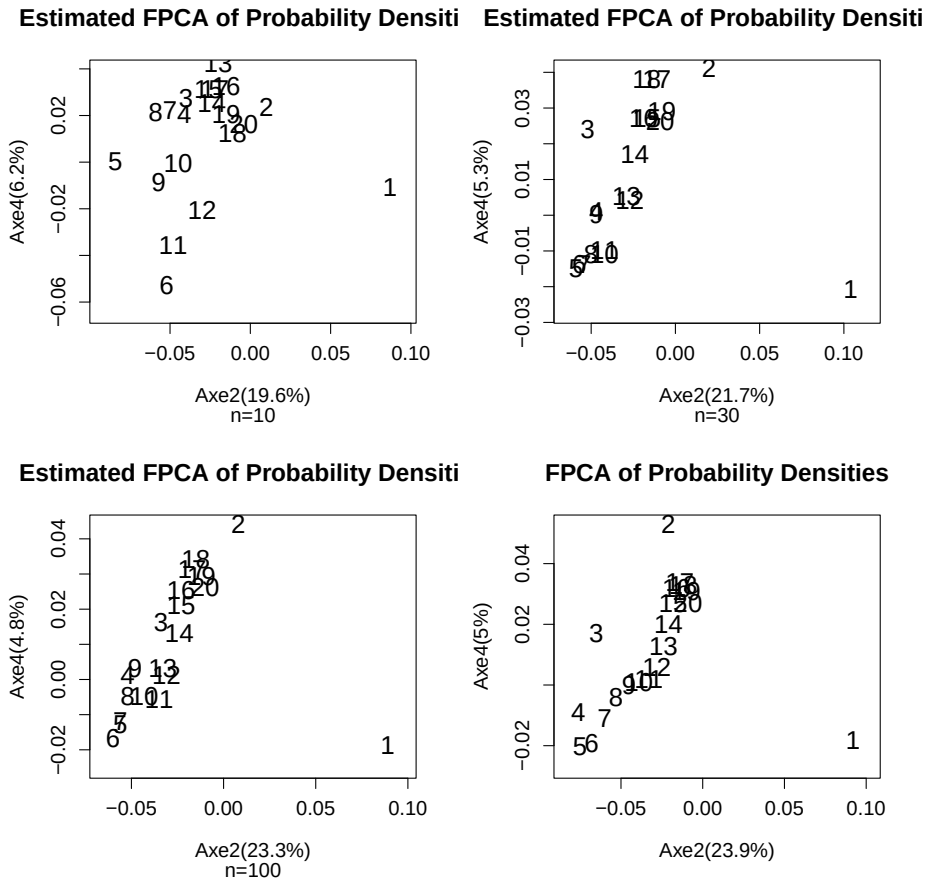

FIGURE 2.10 – La projection des densités sur le plans (2,4)

2.4.5 Analyse des données sonsorielles par une ACP de densité

On considère les données issues d'une application de l'analyse sensorielle en horticulture ornementale (données roses du R-Package dad) : 14 juges ont évalués à trois reprises, 10 rosiers (sur photo) selon entre autre 2 descripteurs sensoriels visuels (symetrie et forme global et hauteur) sur une échelle structurée à 9 niveaux Ainsi, au rosier $t(t = 1, ..., 10)$ est associé un tableau à 16 colonnes et 42 lignes considérés comme 42 observations indépendantes d'un vecteur aléatoire X_t à 16 dimensions, et de densité de probabilité ${f_t}^{(k)}$pour chacune des dimensions. On realise une ACP sur les 10 densités, estimies à la base des données par noyau gaussien, les résultats obtenus sont données dans la partie suivante.

Valeurs propres et inertie

Densités	valeurs propres	pourcentages d'inerties
A	0.206	46.4
B	0.104	23.4
C	0.047	10.6
D	0.037	8.3
E	0.034	7.7
F	0.011	2.5
G	0.003	0.7
H	0.001	0.2
I	0.000	0.0
J	-0.002	0.5

TABLE 2.16 – Les valeurs propres et pourcentages d'inerties

La contribution des rosiers aux axes factoriels

rosier	axe 1	axe 2	axe 3	axe 4
A	14.3	7.0	0.0	5.6
B	13.0	9.7	0.1	8.4
C	0.0	0.0	97.8	1.3
D	16.3	0.3	0.9	43.2
E	13.3	7.4	0.0	5.0
F	11.9	5.7	0.1	3.9
G	16.3	16.6	0.0	2.1
H	5.7	2.1	1.1	8.5
I	5.7	26.5	0.0	7.4
J	3.6	24.6	0.0	14.5

TABLE 2.17 – Convtribution des 10 rosiers aux quatres premiers axes factoriel

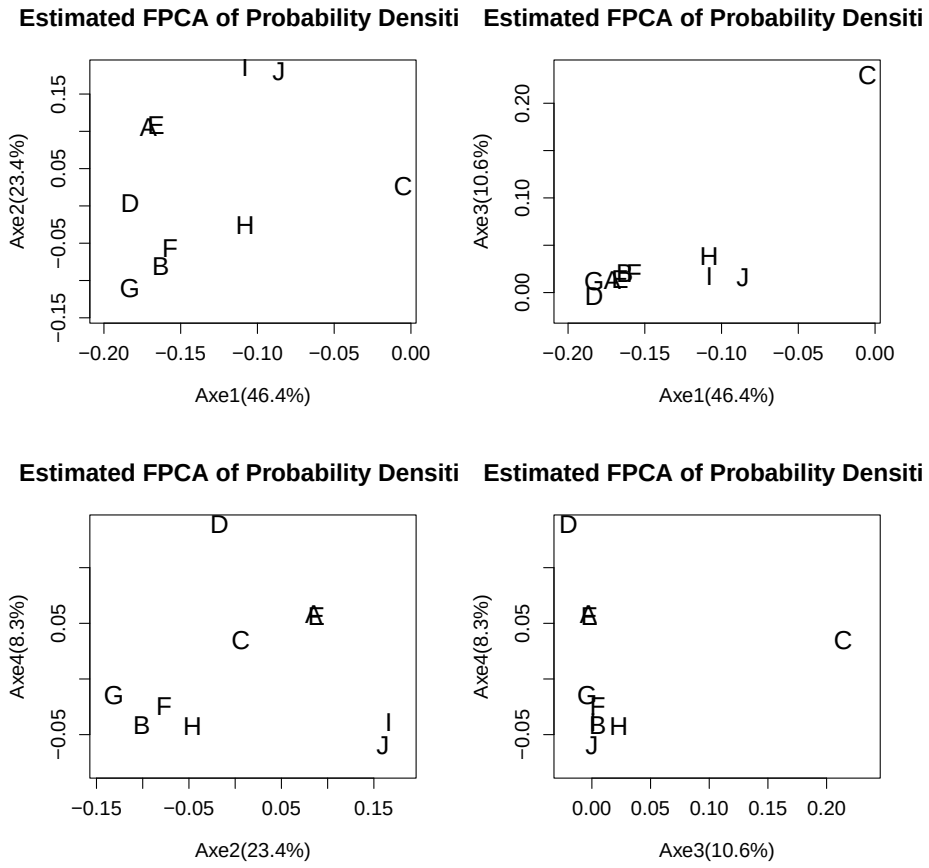

FIGURE 2.11 – La projection sur les quatres premiers axes

La projection des 10 rosiers sur les 4 premiers plans principaux montrent à chaque fois l'existance de groupes de rosiers similaire, en particulier sur le plan (1,3) qui sépare bien les rosiers, (A,B,D,E,F,G) (H,I,J) et C.

Chapitre 3

Interprétation des résultats de l'ACP de densité

3.1 Introduction, position de problème

Comme dans toutes méthodes statistique d'analyse des données mulidimentionelles, la question d'interprétation des résultats est importatnte et rendue plus difficile par l'expression non linéaire des modes de variabilité et la perte des propriétés mathématiques des fonctions propres obtenues par décomposition à des densités. Cette question est résolue en partie dans Boumaza et al 2014, et elle consistée l'interprétation de la position de la densité en fonction de la corrélation linéaire entre le vecteur de $\mathbb{R}^T$ des scores principaux des densités sur l'axe factoriel l et le vecteur des T moyennes de la j-iéme variables ($moy_j \in \mathbb{R}^T$), le vecteur des T variances de l j-iéme variable ($var_j \in \mathbb{R}^T$), les vecteurs des T covariances ($cov_{ij} \in \mathbb{R}^T,\ i = 1, p,\ j = 1, p$) entre la i-ième et la j-iéme variable marginale, et en fin, les vecteurs des T corrélations ($cor_{ij} \in \mathbb{R}^T,\ i = 1, p,\ j = 1, p$) entre la i-ième et la j-iéme variable marginale .
On peut interpréter la position d'une densité en fonction de la corrélation linéaire entre le vecteur de $\mathbb{R}^T$ ($CP.l$) des axes prinsipaux des densités sur l'axe factoriel l et le vecteur $\mathbf{moy}_j$ des T moyennes de la j-ième variable marginale $\mathbf{moy}_j = (moy_{j1}, ..., moy_{jT}) \in \mathbb{R}^T;\ 1 \leq j \leq p$, le vecteur $\mathbf{var}_j$ des T variances de la j-iéme variable marginale $\mathbf{var}_j = (var_{j1}, ..., varjT) \in \mathbb{R}^T;\ 1 \leq j \leq p$, le vecteur $\mathbf{cov}_{ij}$ des T covariances de la i-iéme et la j-iéme variable marginale $\mathbf{cov}_{ij} = (cov_{ij1}, ..., cov_{ijT}) \in \mathbb{R}^T,\ 1 \leq i \leq p,\ 1 \leq j \leq p$, et le vecteur $\mathbf{cor}_{ij}$ des T corrélation de la i-iéme et la j-iéme variable marginale $\mathbf{cor}_{ij} = (cor_{ij1}, ..., cor_{ijT}) \in \mathbb{R}^T,\ 1 \leq i \leq p,\ 1 \leq j \leq p$.
Ainsi, l'axe s'interprete en fonction de la valeur de la contribution obtenu. Par exemple :

- Si $\mathbf{cor}$(axel,moy(X^j)), $j = 1, 2$ est proche de 1, l'axe s'intérprete plus comme une moyenne de X^j, ainsi une densité qui a une forte coordonnée sur l'axe l a donc une forte moyenne par rapport à la variable j.
- Si $\mathbf{cor}$(axel,var(X^j)), $j = 1, 2$ est proche de -1 l'axe l s'intérprete comme l'anti-variance de la variable X^j.

Exemple 3.1. (Interprétation de la position des 20 densités gaussienne (exemple 2.2.4)). Du fait que l'axe1 et l'axe2 représentent à eux seul 80% d'inertie, on se fixe alors à l'intérprétation de ces deux axes. De plus comme les covariances (resp. coorélations) sont nulles, l'intérpretation des densités concerne uniquement les moyennnes et les variances. Le tableau (3.1) nous donne les corrélations entre les vecteurs des scores des densités sur les deux premiers axes et le vecteur des 20 moyennes et des 20 variances.

axe	moy.X1	moy.X2	var.X1	var.X2
axe1	-0.76	-0.76	-0.76	-0.76
axe2	0.19	0.19	0.19	0.19

TABLE 3.1 – Corrélation scores principaux, moyennes, variances

On remarque que l'axe1 est fortement corolé négativement avec les moyennes et les variances des deux variables.Ainsi, plus une densité possède une forte coordonée sur l'axe1, plus ses deux moyennes et ses deux variances marginales sont petites.

3.1.1 Interprétation des proximités inter densités sur les espaces de projection

En pratique les densités sont estimées par noyau, la question d'intérprétation des proximités inter densités passe essentielment par l'intérprétation de la mesure de similarité L^2, entre les même densité. C'est l'objet de paragraphe suivant.

3.2 Mesure de similarité L^2 entre deux populationns de données réelles

Soient $(x_1, ..., x_n)$ et $(y_1, ..., y_n)$ deux échantillons correspondant aux réalisations de deux variables aléatoires X et Y, de même famille de lois et de densités de probabilités f et g, de carrées intégrable par rapport à la mesure de Lebesgue sur $\mathbf{R}$. La mesure de similarité L^2 entre f et g est définie par

$$A = \int_{\mathbb{R}} f(z)g(z)dz. \tag{3.1}$$

En pratique cette quantité est estimée au moyen de l'estimation des densités. Par une méthode paramétrique si les densités appartiennent à une famille connue de lois de probabilité (voir Boumaza(1999) pour plus de détails sur les propriétés asymptotiques des estimateurs obtenus), soit par une approche non paramétrique (voir Yousfi et al, 2014)

pour une famille plus générale de densités. Si on note par $\hat{\mu}_f$, $\hat{\mu}_g$, $\hat{\sigma}_f$, $\hat{\sigma}_g$ les moyennes et les écart-types observés de X et Y respectivement, alors sous l'hypothèse de normalité l'estimation de A vaut :

$$\hat{A} = \frac{1}{\sqrt{2\pi}} \frac{1}{(\hat{\sigma}_f + \hat{\sigma}_g)(\frac{1}{2})} \exp[-\frac{1}{2} \frac{(\hat{\mu}_f - \hat{\mu}_g)^2}{(\hat{\sigma}_f + \hat{\sigma}_g)}]. \tag{3.2}$$

Dans le cas contraire, l'estimation par noyau gaussien de A est égale à (h un paramètre de lissage positif proportionnel à $n^{\frac{-1}{5}}$) :

$$\hat{A} = \frac{1}{n^2} \sum_{i}^{n} \sum_{j}^{n} \frac{1}{\sqrt{2\pi}} \frac{1}{h(\hat{\sigma}_f^2 + \hat{\sigma}_g^2)\frac{1}{2}} \exp[-\frac{1}{2} \frac{(x_i - y_i)^2}{h^2(\hat{\sigma}_f^2 + \hat{\sigma}_g^2)}]. \tag{3.3}$$

La lecture de la formule (3.1) peut s'effectuer de deux points de vue :

1. A est la moyenne de la variable aléatoire $f(Y)$.
2. A est la moyenne de la variable aléatoire $g(X)$.

Les deux points de vue jouent un rôle symétrique. Dans la suite de ce chapitre nous montrons en quoi cette mesure de similarité est porteuse d'information sur la similarité entre les deux échantillons, de X et de Y.

3.2.1 Démarche

L'intérprétation statistique de la mesure de similarité en relation avec la structure des données est une étape importante, peu de travaux dans la littérature traite de cette question, l'idée est donc de vérifier sur des exemples que la mesure de similarité est en quelque sorte une moyenne un peu particulière et, qu'elle est d'autant plus grande, si en moyenne les observations d'une population sont proches des zones de concentration des observations de l'autre population ce qu'on peut expliciter par le calcul de la moyenne arithmétique des images des observations d'une population par la densité de l'autre population. Ce calcul nous l'avons effectué sur des données gaussiennes(3.5), la variation des tailles d'échantillons permet de visualiser en comparant les graphique (c) des figures (2) et (3) où le fait de déplacer 25% des points noirs de la zone de concentration des points rouges a fait chuter la mesure de similatités de 40% de sa valeur.

n	20	100	1000
moyennes observées de $g(X)$	0.498	0.686	0.951
moyennes observées$f(Y)$	0.268	2.123	2.086
valeurs observées de $\hat{A}$	0.398	0.398	0.398

TABLE 3.2 – Valeurs observées en fonction de la taille des données n des moyennes arithmétiques de $g(X)$ et $f(Y)$, de la similarité $\hat{A}$, calculées pour deux populations de données simulées à partir de deux densités gaussiennes $f \equiv N(1,2)$ *et* $g \equiv N(2,3)$

n	50	100	1000
moyennes observées de $g(X)$	0.096	0.100	0.104
moyennes observées$f(Y)$	0.101	0.104	0.106
valeurs observées de $\hat{A}$	0.037	0.036	0.024

TABLE 3.3 – Valeurs observées en fonction de la taille des données n des moyennes arithmétiques de $g(X)$ et $f(Y)$, de la similarité $\hat{A}$, calculées pour deux populations de données simulées à partir d'un mélange de gaussiennes ($f \equiv \frac{1}{2}N(1,2) + \frac{1}{2}N(3,2)$ *et* $g \equiv \frac{1}{2}N(3,2) + \frac{1}{2}N(2,4)$)

Cas de données sensorielle : 10 rosiers de cultivar :

On considère les données issue d'une application de l'analyse sensorielle en horticulture ornementale (données rose du R-Pakage dad) : 14 juges ont évalué à trois reprises, 10 rosiers (sur photo, figure reference) selon 16 descripteurs sensoriels visuels sur une échelle structurée à 9 niveaux. Ainsi, au rosier $t = (t = 1, ..., 10)$ est associé un tableau à 16 colonnes et 42 lignes considérés comme 42 observation indépendantes d'un vecteur aléatoire X_t à 16 dimentions, et de densité de probabilité $f_t^{(k)}$ pour chacune des dimensions. Il est important de noter que dans ce cas les densités sont estimées par la formule (3.2). En effectue pour le discripteur Symétrique de la plante (**Sha**) le calcul des 03 grandeurs(3.5), les résultats obtenus montrent l'effet symétrique joué par les deux premières grandeurs. Les valeurs obtenues sont très proches de celles de la mesure de similarité, cela ait dû en partie à la raison suivante :
L'estimation à noyau gaussien d'une densité $f_t^{(k)}$ associé au rosier t pour un descripteur$k$ est donnée par :

n	50	100	1000
moyennes observées de $g(X)$	0.093	0.091	0.088
moyennes observées $f(Y)$	0.076	0.080	0.078
valeurs observées de $\hat{A}$	0.038	0.035	0.022

TABLE 3.4 – Valeurs observées en fonction de la taille des données n des moyennes arithmétiques de $g(X)$ et $f(Y)$, de la similarité $\hat{A}$, calculées pour deux populations de données simulées à partir d'un mélange de gaussiennes ($f \equiv \frac{1}{4}N(1;2) + \frac{3}{4}N(3;2)$ *et* $g \equiv \frac{1}{6}N(3;2) + \frac{5}{6}N(2;4)$)

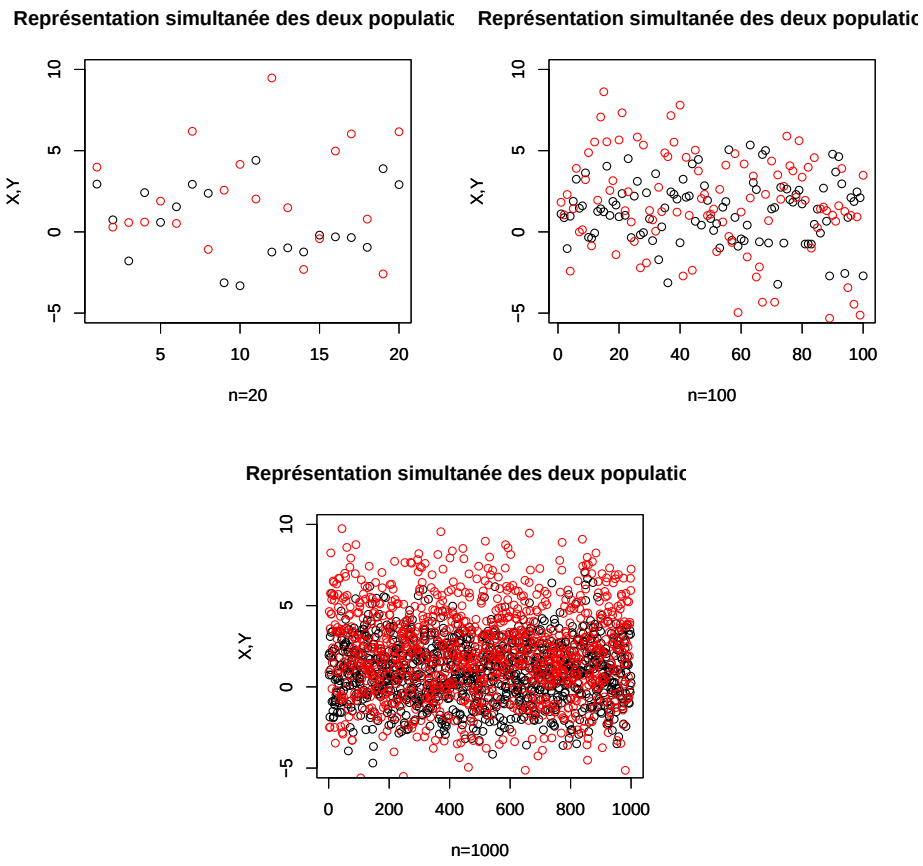

FIGURE 3.1 – Allure du melange des deux populations en fonction de la concentration des d'observations ´ ($f \equiv N(1,2) etg \equiv N(2,3)$)

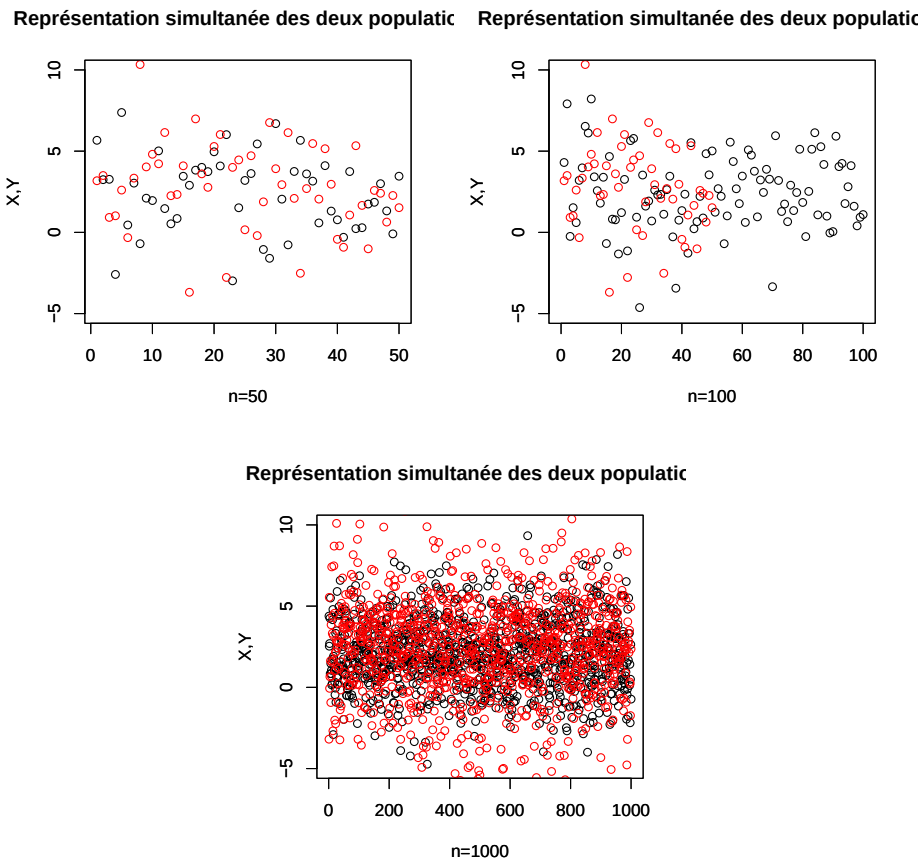

FIGURE 3.2 – Allure du mélange des données des deux populations en fonction de la concentration des d'observations ($f \equiv \frac{1}{2}N(1,2) + \frac{1}{2}N(3,2)$ *et* $g \equiv \frac{1}{2}N(3,2) + \frac{1}{2}N(2,4)$)

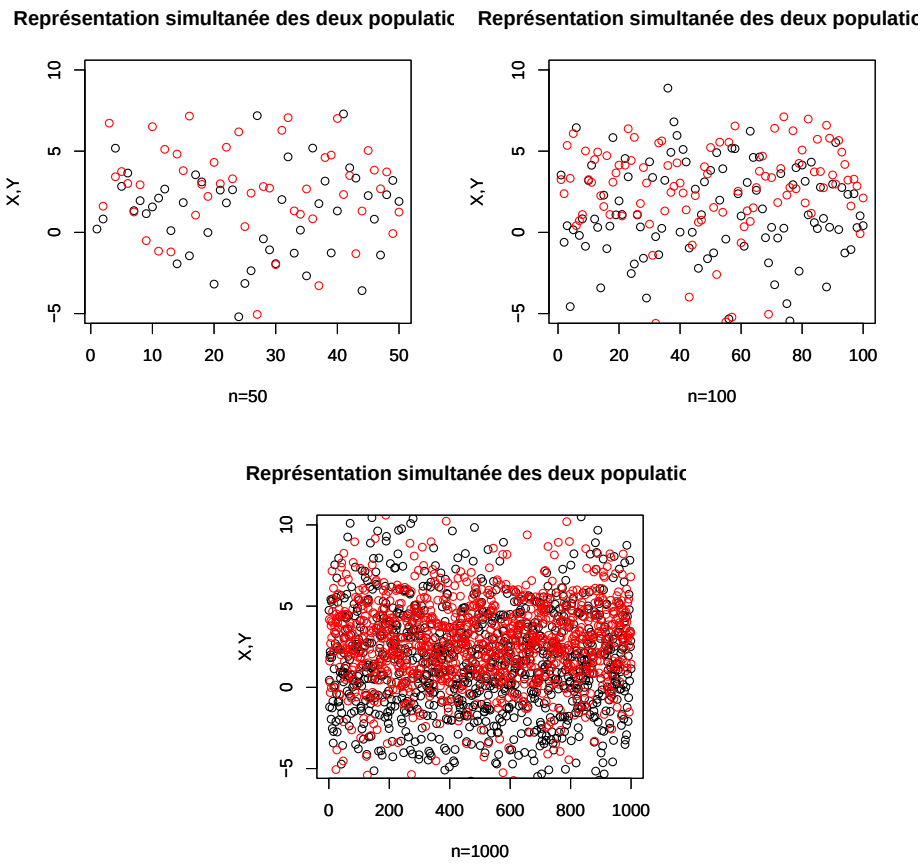

FIGURE 3.3 – Allure du mélange des données des deux populations en fonction de la concentration des d'observations ($f \equiv \frac{1}{4}N(1;2) + \frac{3}{4}N(3;2)$ *et* $g \equiv \frac{1}{4}N(3;2) + \frac{3}{4}N(2;4)$)

Cas de données sensorielle

On considère les données issues d'une application de l'analyse sensorielle en horticulture ornementale (données roses du R-Package dad) : 14 juges ont évalués à trois reprises, 10 rosiers (sur photo,) selon 16 descripteurs sensoriels visuels sur une échelle structurée à 9 niveaux. Ainsi, au rosier $t(t = \overline{1, ..., 10})$ est associé un tableau à 16 colonnes et 42 lignes considérés comme 42 observations indépendantes d'un vecteur aléatoireX_t à 16 dimensions, et de densité de probabilité ${f_t}^{(k)}$pour chacune des dimensions. Il est important de noter que dans ce cas les densités sont estimées par la formule (3.2). En effectue pour le discripteur Symétrie de la plante (**Sha**) le calcul des 03 grandeurs (3.5), les résultats obtenus montrent l'effet symétrique joué par les deux premières grandeurs. Les valeurs obtenues sont très proches de celles de la mesure de similarité, cela ait dû en partie à la raison suivante . L'estimation à noyau gaussien d'une densité${f_t}^{(k)}$ associé au rosier t pour un descripteur k est donnée par

$$\hat{f}_t^{(k)}(x) = \frac{1}{nh\sigma_k^{(t)}} \sum_i^{42} \exp \frac{-1}{2\pi} \frac{(x - x_{il}^{(k)})^2}{(h\sigma_k^{(t)})^2} \tag{3.4}$$

$\sigma_t^{(k)}$ est l'écart-type observé du descripteur k calculer sur le rosier t et $x_{it}^{(k)},\ i = 1, 42$ les 42 notes attribués par les 14 juges au rosier t suivant le descripteur k. Ainsi en remplaçant dans la formule (3.4) x par $x_{il}^{(k)}$ et en calculant la moyenne arithmétique on obtient

$$\frac{1}{n^2} \sum_i^n \sum_j^n \frac{1}{\sqrt{2\pi}} \frac{1}{h\sigma_t^{(k)}} \exp \frac{-1}{2} \frac{(x_{il}^{(k)} - x_{it}^k)^2}{(h\sigma_t^{(k)})^2} \tag{3.5}$$

En effectuant les même calculs sur la densité du rosier l, on obtient

$$\frac{1}{n^2} \sum_i^n \sum_j^n \frac{1}{\sqrt{2\pi}} \frac{1}{h\sigma_l^{(k)}} \exp \frac{-1}{2} \frac{(x_{il}^{(k)} - x_{it}^k)^2}{(} h\sigma_l^{(k)})^2 \tag{3.6}$$

En identifiant les formules (3.2), (3.5) et (3.6) on peut annoncer que l'idée d'estimer la mesure de similarité en utilisant la méthode du noyau est plus proche de l'intuition que la similarité mesure le taux d'observations d'une population donnée qui sont proches des zones de concentrations des observations de l'autre population. En analysant de plus près les résultats du tableau (3.5) par exemple les valeurs de la similarité entre les rosiers A et I on s'attend à une valeur inférieur à celle calculée entre A et lui même, ce qui n'est exactement pas le cas (même chose pour les rosiers E et I, A et B , . . .). Ce constat peut être expliquer par le fait que les zones de concentrations des observations des deux populations sont très proches et que l'une d'elle est absorbée par l'autre, comme on peut le voir sur les graphiques de la figure (3.2.1) où sont superposées les notes des rosiers A et I ensuite E et I. Ce constat a été observé aussi sur les deux premières grandeurs (rosiers F et E, G et E). Une solution naturelle consiste à diviser chaque densité par sa norme. Il est aussi intéressant de notés que l'expression analytique de la mesure de similarité estimée par noyau est moins évidente pour une large famille de noyaux, d'où l'intérêt d'utiliser une des deux premières grandeurs comme une alternative.

Rosiers	les 3 grandeurs	A	B	C	D	E	F	G	H	I	J
	moy obs de f(Y)	0.161	0.119	0.048	0.133	0.159	0.129	0.102	0.149	0.123	0.078
A	moy obs de g(X)	0.161	0.122	0.052	0.132	0.149	0.131	0.800	0.138	0.109	0.148
	val obs de $\hat{A}$	0.126	0.113	0.061	0.111	0.115	0.112	0.102	0.108	0.085	0.074
	moy obs de f(Y)		0.198	0.005	0.185	0.121	0.189	0.203	0.163	0.043	0.052
B	moy obs de g(X)		0.198	0.013	0.159	0.114	0.167	0.273	0.136	0.062	0.047
	val obs de $\hat{A}$		0.181	0.021	0.140	0.105	0.150	0.226	0.117	0.064	0.056
	moy obs de f(Y)			0.147	0.021	0.054	0.018	0.008	0.033	0.106	0.075
C	moy obs de g(x)			0.147	0.018	0.054	0.014	0.000	0.035	0.100	0.093
	val obs de $\hat{A}$			0.106	0.035	0.063	0.031	0.011	0.049	0.079	0.076
	moy obs de f(Y)				0.035	0.063	0.031	0.011	0.049	0.079	0.076
D	moy obs de g(X)				0.157	0.126	0.163	0.224	0.140	0.074	0.083
	val obs de $\hat{A}$				0.121	0.103	0.126	0.153	0.108	0.069	0.060
	moy obs de f(Y)					0.148	0.123	0.100	0.139	0.118	0.108
E	moy obs de g(X)					0.148	0.129	0.076	0.136	0.109	0.140
	val obs de $\hat{A}$					0.107	0.104	0.099	0.101	0.082	0.072
	moy obs def(Y)						0.164	0.166	0.153	0.064	0.048
F	moy obsde g(X)						0.164	0.236	0.139	0.070	0.075
	val obs de $\hat{A}$						0.132	0.167	0.110	0.068	0.059
	moy obs def(Y)							0.305	0.162	0.008	0.041
G	moy obs de g(x)							0.305	0.128	0.051	0.003
	val obs de $\hat{A}$							0.370	0.120	0.057	0.051
	moy obs de f(Y)								0.143	0.092	0.090
H	moy obs de g(X)								0.143	0.090	0.109
	val obs de $\hat{A}$								0.101	0.073	0.065
	moy obs de f(Y)									0.123	0.119
I	moy obs de g(X)									0.123	0.133
	val obs de $\hat{A}$									0.077	0.071
	moy obs de f(Y)										0.119
J	moy obs de g(X)										0.119
	val obs de $\hat{A}$										0.067

TABLE 3.5 – Valeurs des moyennes arithmétiques de $f_t^{(k)}(x_{il}^{(k)}), (t, l = A, ...J)$et de la similarité $\hat{A}$

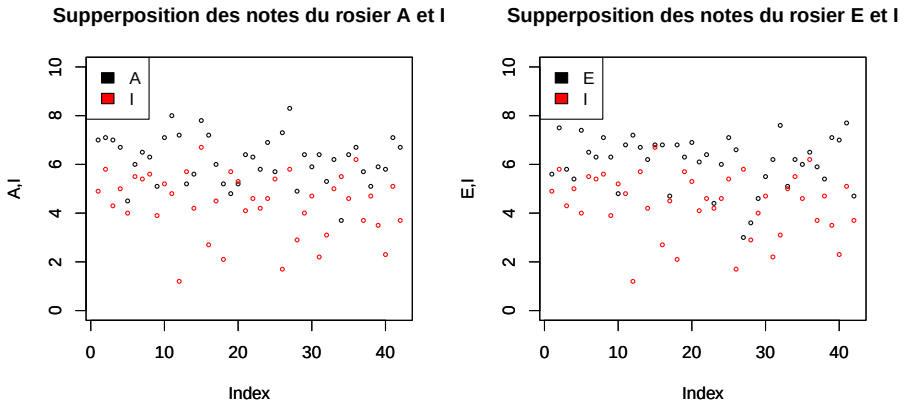

FIGURE 3.4 – Zone de concentration des notes des rosiers A, E et I (Descripteur Symétrie de la plante)

Conclusion générale

Dans ce travail nous avons exposé la méthode de l'analyse en composantes principales de densités de probabilité comme un cas particulier de l'analyse en composante principale "pas à pas" d'un opérateur compact auto-adjoint. Ainsi, nous avons appliqué la méthode sur une famille de 20 densités gaussiennes et non gaussiennes, les résultats obtenue montre que l'ACP reproduit une partie de la variabilité des densités visibles sur les graphiques des plans principaux de projection.

Pour une mise en oeuvre sur des données réelles nous avons effectué une estimation de l'ACP en utilisant une méthode non paramétrique à savoir la méthode du noyau et nous avons montrer sur des données simulées, la convergence de l'analyse en composantes principales estimée vers l'analyse en composantes principales réelle. Une application sur les donnés sensorielle à été effictuée.
En fin, nous avons proposé une approche d'interpretation des résultats de l'ACP basée sur l'interprétation statistique de la mesure d'affinité. En effet, nous avons montré en quoi la mesure d'affinité permet une comparaison de population de données et cette comparaison est en partie reproduite par une application de l'ACP sur les densités de ces populations.

Bibliographie

[1] Anderson, T.W., *Asymptotic theory for principal component analysis*, Annals of Mathematical Statistics 34 (1963), 122-148.

[2] Benko, M., Hardle, W., Kneip, A., *Common Functional Principal Compenents*, The Annals of Statistics. 37 (2009) 1-34.

[3] Benzecri, et al., *L'analyse des données*, tome I, 3^e éd., Dunod, Paris, 1979.

[4] Berrendéro, J.R., Justel, A., Suarc, M., *Principal Components for multivariate functional data*, Computational Statistics & Data Analysis. 55 (2011) 2619-2634.

[5] Bilodeau, M., Brenner, D., *Theory of multivariate statistics*, Springer-Verlag, New York, 1999.

[6] Bosq, D et Lecoutre, J-P., *Théorie de l'estimation fonctionnelle*. Collection économie et statistique avancées, 1987.

[7] Boumaza, R., *Analyses factorielles des distributions marginales de processus.* Thèse Doctorat, Université Joseph Fourier, 1999.

[8] Boumaza, R., *Contribution à l'étude descriptive d'une fonction aléatoire qualitative*,Thèse Doctorat de Spécialité, Université de Toulouse, 1980.

[9] Boumaza, R., *Analyse en composantes principales de distributions gaussiennes multidimennelles*, Revue de Statistique Appliquée XLVI. 2 (1998) 5-20.

[10] Boumaza, R., *Discriminant analysis with independetly repeated multivariate measurements : and L^2 approach*, Computational Statistics & Data Analysis. 47 (2004) 823-843.

[11] Boumaza, R. et al., *Sensory profiles and preference analysis in ornomental horticulture : The case of the rosebush*, Food Quality and Preference. 21 (2010) 987-997.

[12] Boumaza, R., Yousfi, S., Sabine, D.M., *Interpreting the principal component analysis of multivariate density functions*, Accepted for publication in Communications in Statistics-Theory and Methods, (2013).

[13] Brieuc, M.C, *Modélisation supervisée de données fonctionnelles par perceptron multicouches*, Thèse doctorat, Université Paris IX Dauphine, 2002.

[14] Bruce, Hasen, H., *Bandwith selection for nonparametric distribution estimation*, University of Winsconsin, 2004.

[15] Buchwalter, H., *Le calcul intégral*, ellipses, E.d, 1991.

[16] Cailliez, F., Pagés, J-P., *Introduction à l'analyse des données*, Smash, Paris 1976.

[17] Cardot, H., *Nonparametric estimation of smoothed principal components analysis of sampled noisy functions*, Journal of Nonparametric Statistics, 2000, vol.12, 503-538.

[18] Cardot, H., *Contribution à la modélisation statistique fonctionnelle*, Mémoire d'Habilitation à diriger des recherches, Laboratoire de Statistique et de probabilité, Université Paul, Sabatier, Toulouse, 2004.

Journal of Multivariate Analysis. 92 (2005) 24-41.

[19] Coppi, R., Bolasco, S., *Multivariate data analysis* North-Holland Amsterdam. Proceedings of the International Meeting on the Analysis of Multiway Data Matrices, Rome, Marsch 1988, 28-30.

[20] Cornillon, P-A. et al., *Statistiques avec R*, Société Française de statistique, Presses universitaires de Rennes edition, 2010.

[21] Dauxois, J., Pousse, A., *Les Analyses factorielles en calcul des probabilités et en statistique*, Essai d'étude synthétique, Thèse, Université Paul Sabatier, Toulouse, 1976.

[22] Dauxois, J., Pousse, A., Romain, Y., *Asymptotic theory for the principal component analysis of a vector random function*, Some Application to Statistical inference, Laboratoire de statistique et probabilités, Université Paul Sabatie, Toulouse, 1982, page 136-141.

[23] Dauxois, J. Nkiet, G. M. Romain, Y., *Projecteurs orthogonaux et opérateurs associes utiles en statistique multidimensionnelle*, Laboratoire de statistique et probabilités, UMR CNRS C55830, Université Paul Sabatier Toulouse, 2006.

[24] Deheuvels, P., *Estimation non pararamétrique de la densité par histogrammes generalisés*, Publication Inst. Stats. Univ. Paris, 1977 XXII-1-23

[25] Delicado, P., *Dimensionality reduction when data are density functions*, Computational Statistics & Data Analysis. 55 (2011) 401-420.

[26] Duong, T., Hazelton, M.L., *Plug-in bandwidth matrices for bivariate kernel density estimation*, Journal of Nonparametric Statistics, (2003) 15, 17-30.

[27] Duong, T., Hazelton, M.L., *Convergence rates for unconstrained bandwidth matrix selectors in multivariate kernel density estimation*, Journal of Multivariate Analysis, 93 (2005), 417-433.

[28] Duong, T., Hazelton, M.L., *Cross-validation bandwidth matrices for multivariate kernel density estimation*, Scand. J. Stat, (2005) 32, 485-506.

[29] Dunford, N., Schwartz, J., *Linear operators*, Intersience, New York, 1963.

[30] Everitt, B., Hothorn, T., *The analysis of repeated measures data. In : An introduction to applied multivariate analysis with R*, New York : Springer ; 2011 :225e57. Available from : http ://www.springerlink. com.myaccess.library.utoronto.ca/content/l4q1342u00075x06/. Accessed August 24, 2011

[31] Ferraty, F., Vieu, P., *Nonparametric Functional Data Analysis*, Springer, 2006.

[32] Gérard, B. *Estimation à noyau itérés : Synthèse bibliographique* Département des Mathémathiques, Laboratoire de Probabilité et Statistique, Université Montpelier II, 2005.

[33] Gnanadesikan, R., *Methods for Statistical Data Analysis of Multivariate Observations*, New York : Wiley, 1977.

[34] Healy, M., *Matrices for Statistics*, Oxford : Clarendon Press, (2000).

[35] Hildenbrand, W., Kneip, A., and Utikal, K. J., *Une analyse nonparamétrique des distributions du revenu et des caractéristiques des ménages*, Revue de Statistique Appliquée. (1999) 47, 39–56.

[36] L'Hermier des Plantes, H., *Structuration des tableaux à trois indices de la statistique*, Thèse de 3ème cycle, Université Montpellier II, Montpellier, France (1976).

[37] Joliffe, I.T, *Principal component analysis*, Springer, New-York, 2002.

[38] Jones, M.C., Rice, J.A., *Displaying the important features of large collection of similar curves*, The American Statistician. 46 (1992) 140-145.

[39] Jones, M.C., Marron, J.S., Sheather, S.J., *A brief survey of bandwidth selection for density estimation*, Journal of The American Statistical Association. 91 (1996) 401-407.

[40] Kneip, A., Utikal, K.J., *Inference for density families using functional principal component analysis*, Journal of the American Statistical Association. 96 (2001) 519-542.

[41] Kim, P et al. *Functional principal component analysis of density families with categorical and continuous data on Canadian Entrant Manufacturing Firms*, Journal of the American Statistical Association. 106 (2011) 858-878.

[42] Laming,D.,Laming, DRJ., *Sensory analysis*, Cambridge Univ Press, 1986

[43] Lancaster, P., *Theory of Matrices*, Academic Press, New York, 1969.

[44] Leurgans, S., Moyeed, R., Silverman, B., *Canonical correlation analysis when the data are curves*, Journal of The Royal Statistical Society. 55 (1993) 725-740.

[45] Pezzuli, S., Silverman, B.W., *Some propreties of smoothed principal components analysis for functional data*, Computational Analysis, 1993, page 1-16.

[46] Phillipe, B., Ramsay, J.O., *Principal components analysis of sampled function*, Psychometrika, june 1986, vol 51, page 285-311.

[47] Qannari, E.M., *Analyses factorielles de mesures et applications*, Thèse de 3ème cycle, Université Paul Sabatier, Toulouse, 1983.

[48] Ramsay, J.O., Silverman, B.W., *Functional data analysis*, Springer, second ed, New-York, 2005.

[49] Rachedi, M., Vieu, P. *Nonparametric regression for functional data : automatic smoothing parameter selection*, Journal of statistical Planing and Inference. 137 (2007) 2784-2801.

[50] Rossi, F., Conan-Guez, B., *Consistent estimation of parameters in a nonlinear model for functional data with randomly chosen evaluation points*, Comptes Rendus Mathématique. 340 (2005) 167-170.

[51] Romain, Y., *Contribution à la Statistique Multidimensionnelles Opèratorielle*, Thèse Laboratoire de Statistique et Probabilité, Université Paul Sabatier, Toulouse 2000.

[52] Saporta, G., *Probabilités, analyses des données et statistique*, Technip, Paris 1990.

[53] Scott, D.W., *Multivariate Density Estimation : Theory, Practice and Visualisation*, Wiley, New York, 1992.

[54] Scott, D.W., Tapia, R.A., Tompson, J.R., *Kernel Density Estimation revisited, Nonlinear Analysis*, Theory, Methods and Applications, Vol 1, pp 339-372.

[55] Silverman, B.W., *Density estimation for Statistical and Data analysis*, School of Mathematics, University of Bath, UK 1986.

[56] Silverman, B.W., Pezzuli, S., *Some Properties of Smoothed Principal Component Analysis for Functional Data*, Computational Statistics, 1993, vol 8 pp 1-16.

[57] Simonoff, J.S., *Smoothing methods in statistics*, Springer-Verlag, New-York, 1996.

[58] Timm, N., *Applied multivariate analysis*, Berlin Heidelberg New York : Springer, 2002.

[59] Wand, M.P., Jones, M.C., *Kernel Smoothing* Chapman and Hall Ltd, London, 1995.

[60] Wang, G. Lin, N. Zhang, B., *Functional contour regression*, Journal of Multivariate Analysis. 116 (2013) 1-13.

[61] Yao, F., Muller, Wang, J-L., *Functional linear regression analysis for longitudinal data*, Annals of Statistics. 33 (2005) 2873-2903.

[62] Ycart, B., Genon-Catalot, V., *Vecteurs et suites aléatoires. Martingales discrètes* Cahiers de Mathématiques Appliquées N°9-10, Tunisie, 2004.

[63] Yousfi, S., Boumaza, R., Aissani, D. *ACP de densités de probabilité et STATIS dual. Estimation par noyau et choix de la fenêtre.* Acte des 43 ième Journées de la Société Française de la Statistique, Gamamrth, Tunisie (2011).

[64] Yousfi, S., Aissani, D., *Contrôle de la performance en analyse sensorielle par l'ACP de densités de probabilités*, RAMA'2012. http ://www.usthb.dz/RAMA8.

[65] Yousfi, S., Boumaza, R., *ACP de densités de probabilité estimée par la méthode du noyau*, Acte du Séminaire International sur la Statistique et ses Applications, Université de Tizi-Ouzou, Mai 2006.

[66] Yousfi, S., Boumaza, R., *Analyses en Composantes Principales de Densités de Probabilité Estimée par la Méthode du noyau*, Mémoire de Magister, Université Mouloud Mammeri de Tizi ouzou, 2007.

[67] Yousfi, S., Boumaza, R., *Analyses en Composantes Principales de Densités de Probabilité Estimée par la Méthode du noyau*, Mémoire de Magister, Université Mouloud Mammeri de Tizi ouzou, 2007.