

Moyenne de Fréchet pour l'analyse statistique de formes et des images

Jérémie Bigot

Institut de Mathématiques de Toulouse
Université de Toulouse

**Workshop MASCOT NUM - Anestis' Fiesta
Mars 2011**

Collaboration avec Anestis Antoniadis, Sébastien Gadat, Fabrice Gamboa, Sophie Lambert-Lacroix, Frédérique Letué, Jean-Michel Loubes, Clément Marteau, Myriam Vimond

- 1 Motivations
- 2 Choix d'une distance
- 3 Moyenne de Fréchet
- 4 Mesures de dissimilarité et ACP géométrique
 - Action de groupe
 - Mesure de dissimilarité
 - ACP géométrique
- 5 Statistique non-paramétrique et modèles déformables

Problématique

Objectif : comparer des objets présentant des caractéristiques similaires et extraire des informations sur la loi de distribution de ces objets

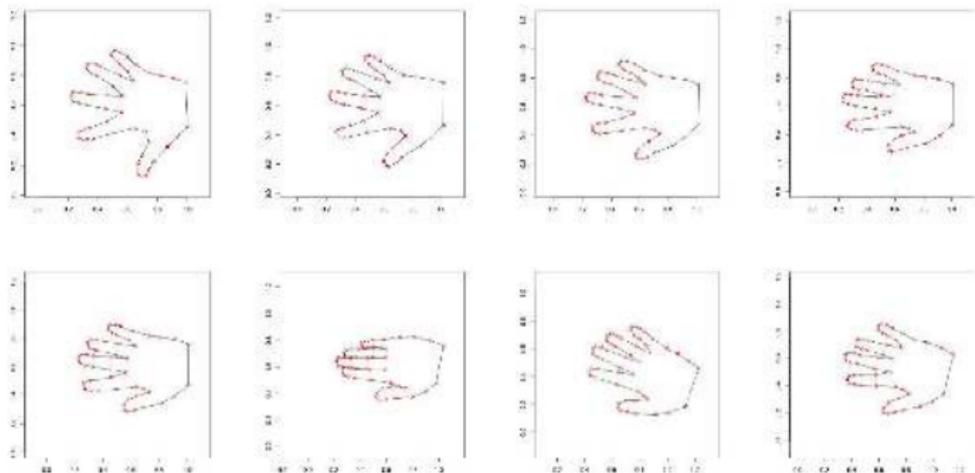
Problématique

Données : images 2D de taille $N_1 \times N_2$ pixels



Problématique

Données : formes dans le plan décrites par $k = 50$ landmarks (points rouges dans \mathbb{R}^2)



Problématique

Observations : collection de n “objets” $Y_i, i = 1, \dots, n$ qui peuvent être :

- des courbes ou des images i.e. $Y_i : \Omega \rightarrow \mathbb{R}$ avec $\Omega \subset \mathbb{R}^d$ pour $d = 1, 2, 3$
- des points dans le plan $Y_i \in \mathbb{R}^{2k}$ (ensemble de k landmarks dans \mathbb{R}^2)
- courbes paramétrées i.e $Y_i : [0, 1] \rightarrow \mathbb{R}^p$ pour $p = 2, 3$
-

Variabilité d'un ensemble de données

L'Analyse en Composante Principale (**ACP**) est très utilisée pour visualiser les principaux modes de variabilité d'un ensemble de données autour de leur moyenne

Soit $Y_i, i = 1, \dots, n$ des variables aléatoires iid à valeur dans un espace de Hilbert \mathcal{H} .

- Moyenne empirique $\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$
- Si $\mathcal{H} = \mathbb{R}^p$, ACP = diagonaliser la matrice de covariance

$$S = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y}_n)(Y_i - \bar{Y}_n)',$$

et on définit

- **le premier mode de variation par** $\bar{Y}_n \pm \sqrt{\lambda_1} \hat{w}_1$
- **le deuxième mode de variation par** $\bar{Y}_n \pm \sqrt{\lambda_2} \hat{w}_2$

où $\hat{w}_1, \hat{w}_2, \dots$ sont les premiers vecteurs propres de S associés aux plus grandes valeurs propres $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$

Variabilité d'un ensemble d'images

Données : images de taille $N_1 \times N_2$ pixels i.e. $\mathcal{H} = \mathbb{R}^{N_1 N_2}$



Moyenne empirique - \bar{Y}_n n'est pas un visage (effet de flou) !

Variabilité d'un ensemble d'images

Données : images de taille $N_1 \times N_2$ pixels i.e. $\mathcal{H} = \mathbb{R}^{N_1 N_2}$



Premier mode de variation - $\bar{Y}_n + \sqrt{\lambda_1} \hat{w}_1$

Variabilité d'un ensemble d'images

Données : images de taille $N_1 \times N_2$ pixels i.e. $\mathcal{H} = \mathbb{R}^{N_1 N_2}$



Premier mode de variation - \bar{Y}_n

Variabilité d'un ensemble d'images

Données : images de taille $N_1 \times N_2$ pixels i.e. $\mathcal{H} = \mathbb{R}^{N_1 N_2}$



Premier mode de variation - $\bar{Y}_n - \sqrt{\lambda_1} \hat{w}_1$

Variabilité d'un ensemble d'images

Données : images de taille $N_1 \times N_2$ pixels i.e. $\mathcal{H} = \mathbb{R}^{N_1 N_2}$



Moyenne empirique - \bar{Y}_n

Variabilité d'un ensemble d'images

Données : images de taille $N_1 \times N_2$ pixels i.e. $\mathcal{H} = \mathbb{R}^{N_1 N_2}$



Deuxième mode de variation - $\bar{Y}_n + \sqrt{\lambda_2} \hat{w}_2$

Variabilité d'un ensemble d'images

Données : images de taille $N_1 \times N_2$ pixels i.e. $\mathcal{H} = \mathbb{R}^{N_1 N_2}$



Deuxième mode de variation - \bar{Y}_n

Variabilité d'un ensemble d'images

Données : images de taille $N_1 \times N_2$ pixels i.e. $\mathcal{H} = \mathbb{R}^{N_1 N_2}$



Deuxième mode de variation - $\bar{Y}_n - \sqrt{\lambda_2} \hat{w}_2$

- 1 Motivations
- 2 Choix d'une distance
- 3 Moyenne de Fréchet
- 4 Mesures de dissimilarité et ACP géométrique
 - Action de groupe
 - Mesure de dissimilarité
 - ACP géométrique
- 5 Statistique non-paramétrique et modèles déformables

Le problème du choix d'une distance

Idée : la comparaison d'objets repose sur le choix d'une bonne distance

Modèle utilisé classiquement : choix d'une **distance euclidienne** d sur \mathcal{H} associée à un produit scalaire

$$d(y, y') = \|y - y'\|_{\mathcal{H}} = \sqrt{\langle y - y', y - y' \rangle_{\mathcal{H}}} \text{ pour } y, y' \in \mathcal{H}$$

Le choix d'une distance induit des statistiques spécifiques

Distance euclidienne \Rightarrow moyenne usuelle dans \mathcal{H}

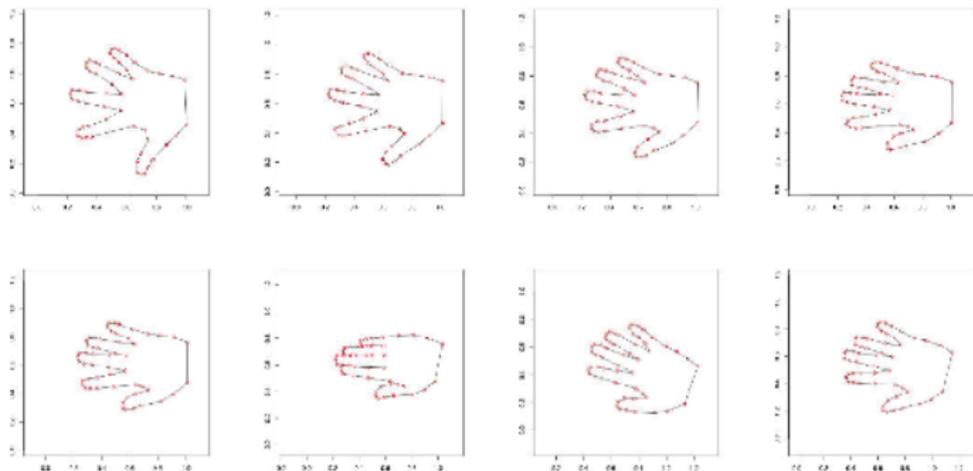
$$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i = \arg \min_{y \in \mathcal{H}} \sum_{i=1}^n \|Y_i - y\|_{\mathcal{H}}^2$$

Questions :

- **le choix de la distance euclidienne** sur \mathcal{H} est-il bien adaptée à la comparaison d'images ou de formes ?
- \bar{Y}_n est-il toujours dans "**le même espace**" que les Y_i ?

Cas de formes dans le plan

Données : formes dans le plan décrites par $k = 50$ landmarks (points rouges dans \mathbb{R}^2)



Problème : l'espace euclidien $\mathcal{H} = \mathbb{R}^{k \times 2}$ est-il bien adapté pour modéliser des formes dans le plan ?

Espace des formes de Kendall

Soit x_1, \dots, x_k des points dans \mathbb{R}^2 (vecteur dans \mathbb{R}^{2k} ou \mathbb{C}^k)

Définition d'une forme (Kendall (1984)) : "Shape is what remains when location, size, and rotational effects are filtered out"

Standardisation (échelle + translation)

$$\tau(x_1, \dots, x_k) = \left(\frac{x_1 - \bar{x}}{\sqrt{\sum_{j=1}^k \|x_j - \bar{x}\|_{\mathbb{R}^2}^2}}, \dots, \frac{x_k - \bar{x}}{\sqrt{\sum_{j=1}^k \|x_j - \bar{x}\|_{\mathbb{R}^2}^2}} \right)$$

où $\bar{x} = \frac{1}{k} \sum_{j=1}^k x_j \in \mathbb{R}^2$

Pre-shape space : $\tau(x_1, \dots, x_k) \in \mathbb{S}_*^{2k-3} = \mathbb{F}^{2k-2} \cap \mathbb{S}^{2k-1}$, où

$$\mathbb{F}^{2k-2} = \{(x_1, \dots, x_k) \in \mathbb{R}^{2k} : \sum_{j=1}^k x_j = 0\}$$

$$\mathbb{S}^{2k-2} = \{(x_1, \dots, x_k) \in \mathbb{R}^{2k} : \sum_{j=1}^k \|x_j\|_{\mathbb{R}^2}^2 = 1\}$$

Espace des formes de Kendall

Soit $\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ une rotation dans le plan autour de l'origine.

Equivalence de deux pré-formes $\tau_1, \tau_2 \in \mathbb{S}_*^{2k-3}$:

$$\tau_1 \sim \tau_2 \text{ s'il existe } \theta \text{ tel que } \theta(\tau_1) = \tau_2$$

où $\theta(\tau) = (\theta\tau^1, \dots, \theta\tau^k)$ pour $\tau = (\tau^1, \dots, \tau^k) \in \mathbb{R}^{2 \times k}$

Espace des formes de Kendall : ensemble de classes d'équivalence

$$\Sigma_2^k = \{[\tau] : \tau \in \mathbb{S}_*^{2k-3}\},$$

où $[\tau]$ est la classe d'équivalence de τ pour la relation \sim

Construction d'une distance sur Σ_2^k

Soit $\tau_1, \tau_2 \in \mathbb{S}_*^{2k-3}$, la distance (géodésique) entre τ_1 et τ_2 donnée par

$$d_g(\tau_1, \tau_2) = \cos^{-1}(\langle \tau_1, \tau_2 \rangle),$$

(avec $\langle \cdot, \cdot \rangle$ produit scalaire usuel sur \mathbb{C}^k) induit une distance sur Σ_2^k

$$\tilde{d}([\tau_1], [\tau_2]) = \inf_{0 \leq \theta < 2\pi} d_g(\tau_1, \theta(\tau_2))$$

On peut montrer que pour $[\tau_1]$ et $[\tau_2]$:

Proposition (Procrustean distance)

$$\tilde{d}([\tau_1], [\tau_2]) = \cos^{-1}(|\langle \tau_1, \tau_2 \rangle|) = \cos^{-1} \left(\left| \sum_{j=1}^k \tau_{1,j} \tau_{2,j}^* \right| \right)$$

Espace des formes = variété Riemannienne

Proposition (Variété Riemannienne)

(Σ_2^k, \tilde{d}) est une variété Riemannienne de dimension $2k - 4$

Différences avec le cas Euclidien $(\mathbb{C}^k, \|\cdot\|_{\mathbb{C}^k})$

- Σ_2^k n'est pas un espace linéaire car " $[\tau_1] + [\tau_2] \notin \Sigma_2^k$ "
- \tilde{d} n'est pas une distance associée à un produit scalaire euclidien

Problèmes :

- comment définir une moyenne sur Σ_2^k ?
- comment définir une notion d'ACP ?
- si $Y \in \mathcal{H}$ alors on définit pour $v \in \mathcal{H}$ l'opérateur de covariance

$$\mathbb{E}Y \langle Y, v \rangle_{\mathcal{H}} (= \mathbb{E}YY'v = \Sigma v \text{ si } \mathcal{H} = \mathbb{R}^p)$$

- si $[Y] \in \Sigma_2^k$ alors " $\mathbb{E}[Y] \langle [Y], [v] \rangle$ " n'a pas d'équivalent simple !

- 1 Motivations
- 2 Choix d'une distance
- 3 Moyenne de Fréchet**
- 4 Mesures de dissimilarité et ACP géométrique
 - Action de groupe
 - Mesure de dissimilarité
 - ACP géométrique
- 5 Statistique non-paramétrique et modèles déformables

Moyenne au sens de Fréchet

Definition

Soit X une variable aléatoire de loi \mathbb{P} à valeur dans un espace métrique (\mathcal{M}, d) . Une moyenne (pas nécessairement unique !) au sens de Fréchet de la distribution \mathbb{P} est un point x^* qui est un minimum **global** de la fonction

$$F(x) = \int_{\mathcal{M}} d^2(x, y) d\mathbb{P}(y) \text{ et } x^* \in \arg \min_{x \in \mathcal{M}} F(x)$$

La **moyenne empirique** au sens de Fréchet est donné par la distribution empirique associée à un échantillon X_1, \dots, X_n de loi \mathbb{P} i.e.

$$\bar{X}_n \in \arg \min_{x \in \mathcal{M}} \sum_{i=1}^n d^2(x, X_i)$$

Moyenne au sens de Fréchet - cas euclidien

Proposition

Si \mathcal{M} est un espace de Hilbert muni de la distance euclidienne usuelle alors la moyenne de Fréchet est unique et correspond à la notion de moyenne usuelle

$$x^* = \mathbb{E}X = \int_{\mathcal{M}} y d\mathbb{P}(y),$$

et

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

Moyenne au sens de Fréchet - propriétés statistiques

Question : convergence quand $n \rightarrow +\infty$ de

$$\bar{X}_n \in \arg \min_{x \in \mathcal{M}} \sum_{i=1}^n d^2(x, X_i)$$

vers

$$x^* \in \arg \min_{x \in \mathcal{M}} \int_{\mathcal{M}} d^2(x, y) d\mathbb{P}(y)$$

Références

- R. Bhattacharya and V. Patrangenaru. Large sample theory of intrinsic and extrinsic sample means on manifolds (i). *Annals of statistics*, 31(1) :1–29, 2003.
- R. Bhattacharya and V. Patrangenaru. Large sample theory of intrinsic and extrinsic sample means on manifolds (ii). *Annals of statistics*, 33 :1225–1259, 2005.

- 1 Motivations
- 2 Choix d'une distance
- 3 Moyenne de Fréchet
- 4 Mesures de dissimilarité et ACP géométrique**
 - Action de groupe
 - Mesure de dissimilarité
 - ACP géométrique
- 5 Statistique non-paramétrique et modèles déformables

Observations : collection de n “objets” $Y_i, i = 1, \dots, n$ qui peuvent être :

- des courbes ou des images i.e. $Y_i : \Omega \rightarrow \mathbb{R}$ avec $\Omega \subset \mathbb{R}^d$ pour $d = 1, 2, 3$
- des points dans le plan $Y_i \in \mathbb{R}^{2k}$ (ensemble de k landmarks dans \mathbb{R}^2)

Problème : comment choisir un espace \mathcal{M} et une distance (Riemannienne) d pour modéliser les données ?

Choix d'une action de groupe

Une possibilité (dans cette direction) : se donner un groupe G de transformations qui agit sur l'espace de Hilbert \mathcal{H}

Principe : soit $g \in G$ et $y \in \mathcal{H}$

- on note par \cdot l'action de G sur \mathcal{H}
- on suppose $g \cdot y \in \mathcal{H}$ pour tout $(g, y) \in G \times \mathcal{H}$

Choix d'une action de groupe

Exemple d'action de groupes pour $\mathcal{H} = \mathbb{R}^{2 \times k}$

- **Espace des formes** - choix de

$$G = \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{R}_+$$

groupe des transformations “affines” (translation + rotation + scaling) du plan, et pour $g = (b, \theta, a) \in G$ et $y \in \mathbb{R}^{2 \times k}$ on définit

$$g \cdot y = aR_\theta y + \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} b'$$

où

$$R_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

Choix d'une action de groupe

Exemple d'action de groupes pour $\mathcal{H} = L^2([0, 1])$

- **Déformation rigide de courbes 1D** - choix de

$$G = \mathbb{R}$$

groupe des translations, et pour $g = b \in G$ et $y \in L^2([0, 1])$ on définit

$$g \cdot y(t) = y(t - b), \quad t \in [0, 1]$$

Choix d'une action de groupe

Exemple d'action de groupes pour $\mathcal{H} = L^2([0, 1]^2)$

- **Déformation rigide d'images 2D** - choix de

$$G = \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{R}_+$$

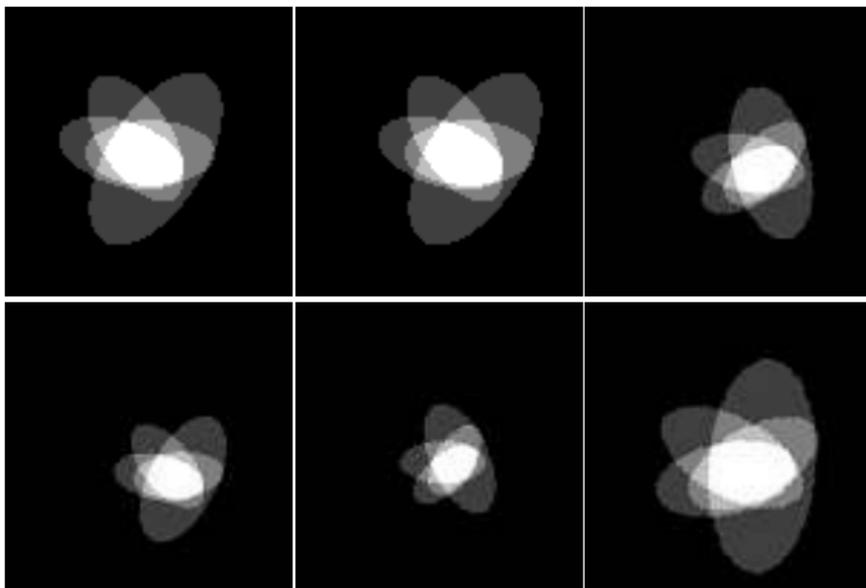
groupe des translations et rotations, et pour $g = (b, \theta, a) \in G$ et $y \in L^2([0, 1]^2)$ on définit

$$g \cdot y(u) = y(aR_\theta(u - b)), \quad u \in [0, 1]^2$$

avec

$$R_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

Déformation rigide d'images 2D



Choix d'une action de groupe

Exemple d'action de groupes pour $\mathcal{H} = L^2(\Omega)$ avec $\Omega \subset \mathbb{R}^d$

- **Déformation non-rigide de courbes ou d'images** - choix de

$$G = \{\phi : \Omega \rightarrow \Omega\}$$

groupe de difféomorphismes de Ω i.e. tel que

$$\phi(\Omega) = \Omega = \phi^{-1}(\Omega),$$

et pour $g = \phi \in G$ et $y \in L^2(\Omega)$ on définit

$$g \cdot y(u) = y(\phi(u)), \quad u \in \Omega$$

Déformation non-rigide d'images 2D



Choix de “distances” non-euclidiennes

Théorie des formes de Grenander (1993) : variabilité géométrique des images sous l'action d'un groupe de Lie

Espace des formes ou des images : $\mathcal{H} = \mathbb{R}^{2 \times k}$ ou $\mathcal{H} = L^2(\Omega)$

Espace des déformations : G groupe (de Lie) agissant sur \mathcal{H}

Mesure de dissimilarité induite par G : pour $y, y' \in L^2(\Omega)$

$$d_G^2(y, y') = \inf_{g \in G} \left\{ \|y - g^{-1} \cdot y'\|_{\mathcal{H}}^2 + \lambda D(g, e) \right\}, \text{ où } \lambda \geq 0$$

Interprétation de d_G^2 dans le cas $\lambda = 0$

$d_G^2(y, y') = 0$ s'il existe $g \in G$ tel que

$$y = g^{-1} \cdot y'$$

Estimation par moyenne de Fréchet

Observations : Y_1, \dots, Y_n variables aléatoires iid à valeur dans \mathcal{H}

Schéma général d'estimation (registration/warping/alignment)

$$\begin{aligned}\hat{y}_n &= \arg \min_{y \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n d_G^2(y, Y_i) \\ &= \arg \min_{y \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n \inf_{g_i \in G} \left\{ \|g_i^{-1} \cdot Y_i - y\|_{\mathcal{H}}^2 + \lambda D(g_i, e) \right\}\end{aligned}$$

Estimation par moyenne de Fréchet

Observations : Y_1, \dots, Y_n variables aléatoires iid à valeur dans \mathcal{H}

Schéma général d'estimation (regISTRATION/warping/alignment)

Etape 1 Estimation des paramètres de déformations

$$(\hat{g}_1, \dots, \hat{g}_n) = \arg \min_{(g_1, \dots, g_n) \in G^n} \left\{ \frac{1}{n} \sum_{i=1}^n \left\| g_i^{-1} \cdot Y_i - \frac{1}{n} \sum_{j=1}^n g_j^{-1} \cdot Y_j \right\|_{\mathcal{H}}^2 + \lambda D(g_i, e) \right\}$$

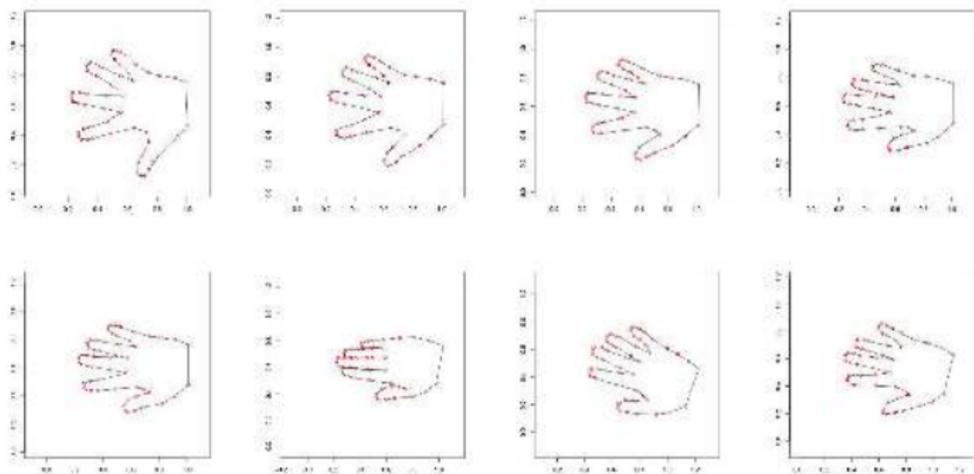
Etape 2 Alignement puis moyenne usuelle des données

$$\hat{y}_n = \frac{1}{n} \sum_{i=1}^n \hat{g}_i^{-1} \cdot Y_i$$

Exemple : moyenne de formes

G - groupe de translation + rotation + scaling du plan \mathbb{R}^2

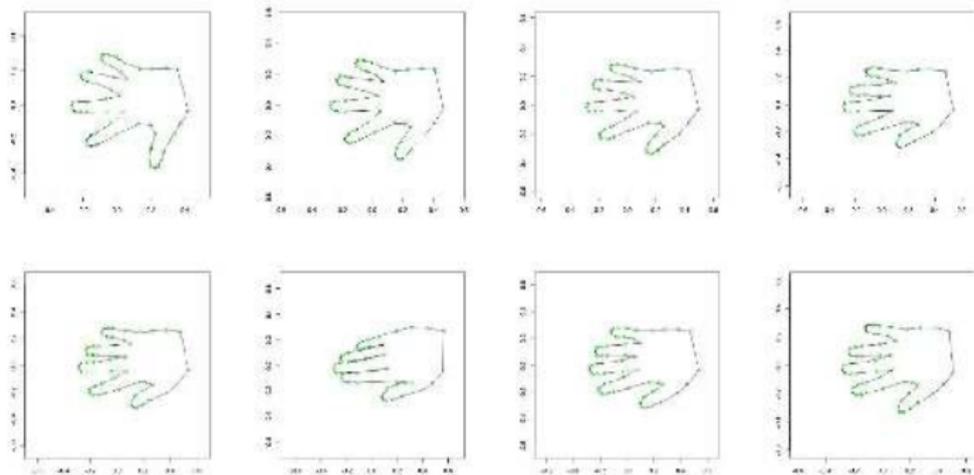
Données : formes dans le plan décrites par $k = 13$ landmarks (points rouges dans \mathbb{R}^2)



Exemple : moyenne de formes

G - groupe de translation + rotation + scaling du plan \mathbb{R}^2

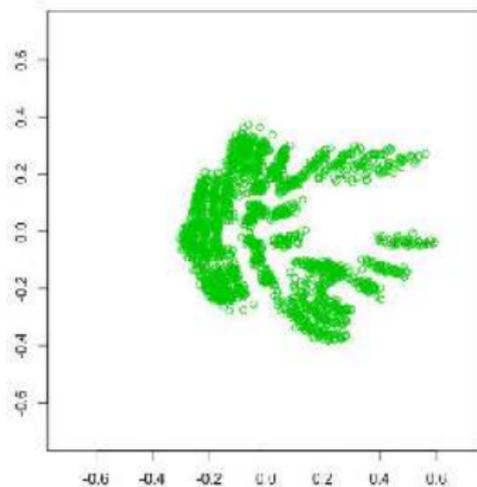
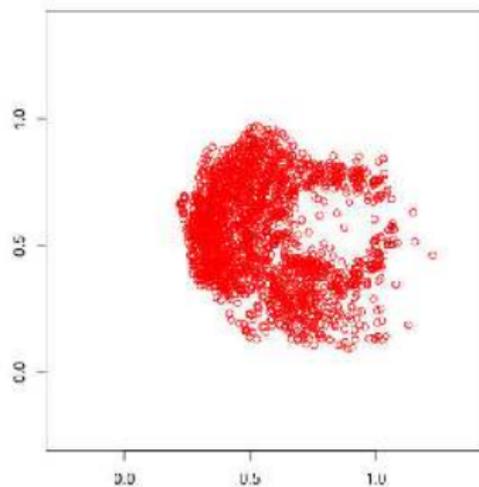
Données alignées : formes dans le plan décrites par $k = 13$ landmarks (points verts dans \mathbb{R}^2)



Exemple : moyenne de formes

G - groupe de translation + rotation + scaling du plan \mathbb{R}^2

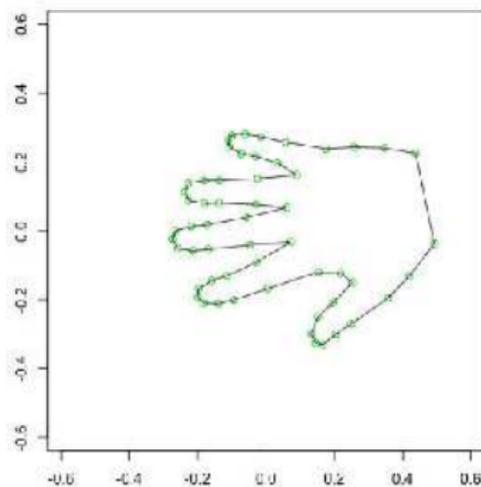
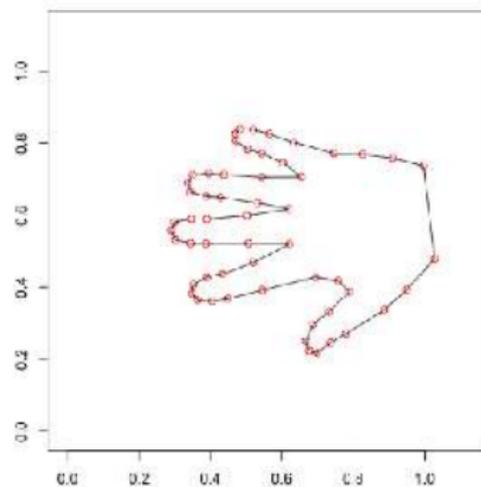
Données non-alignées / **Données alignées** ($n = 56$)



Exemple : moyenne de formes

G - groupe de translation + rotation + scaling du plan \mathbb{R}^2

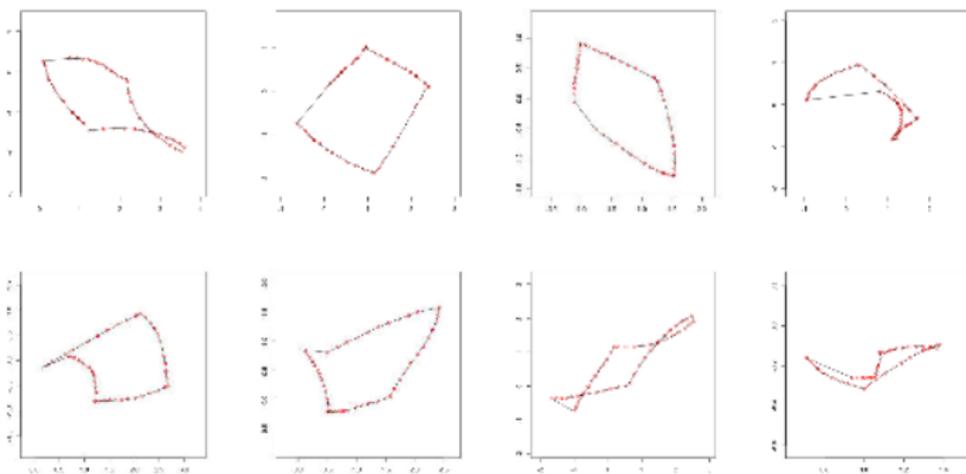
Moyenne euclidienne usuelle / **Moyenne de Fréchet** ($n = 56$)



Exemple : moyenne de formes

G - groupe de translation + rotation + scaling du plan \mathbb{R}^2

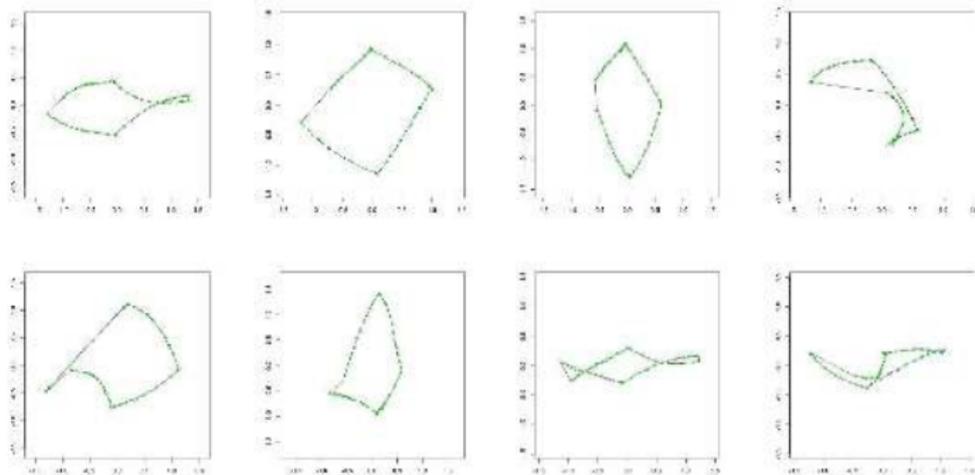
Données : formes dans le plan décrites par $k = 50$ landmarks (points rouges dans \mathbb{R}^2)



Exemple : moyenne de formes

G - groupe de translation + rotation + scaling du plan \mathbb{R}^2

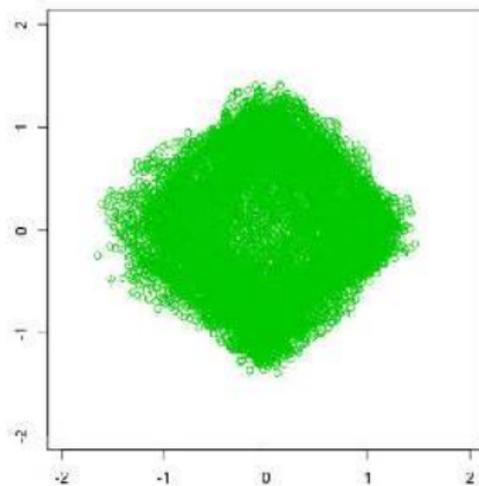
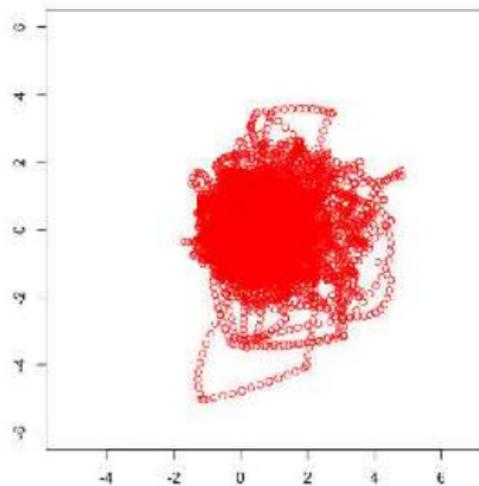
Données alignées : formes dans le plan décrites par $k = 50$ landmarks (points verts dans \mathbb{R}^2)



Exemple : moyenne de formes

G - groupe de translation + rotation + scaling du plan \mathbb{R}^2

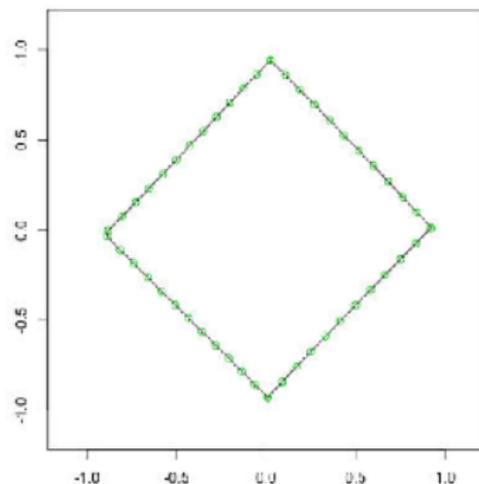
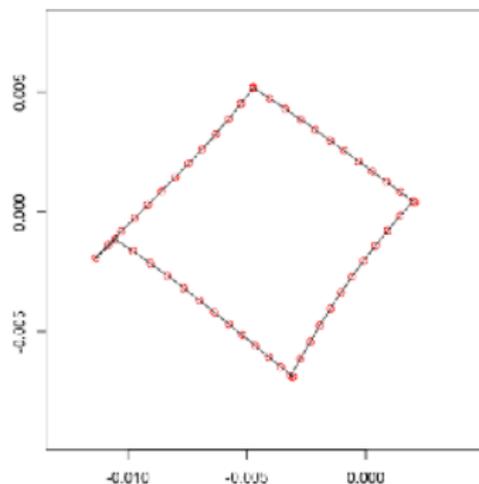
Données non-alignées / **Données alignées** ($n = 200$)



Exemple : moyenne de formes

G - groupe de translation + rotation + scaling du plan \mathbb{R}^2

Moyenne euclidienne usuelle / **Moyenne de Fréchet** ($n = 200$)



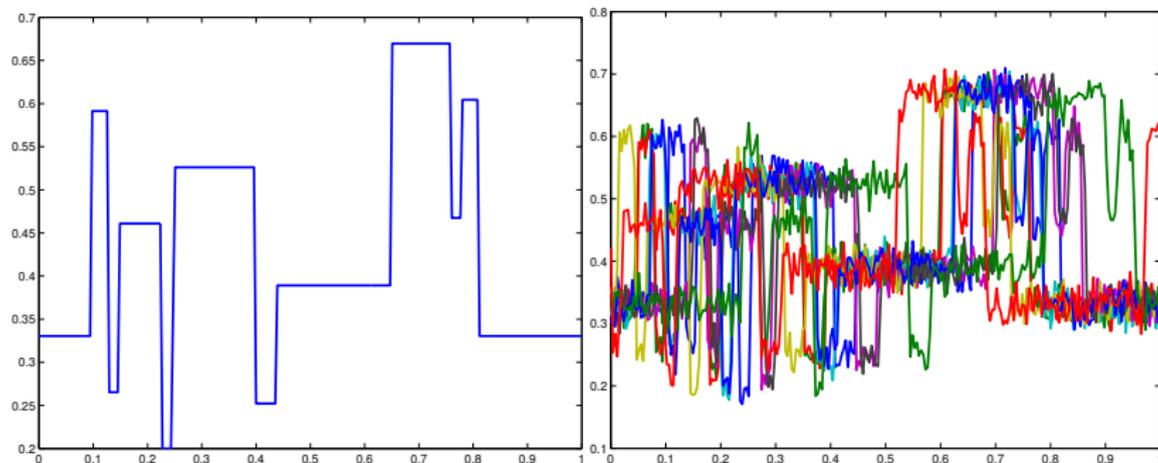
Exemple : moyenne de courbes

$G = \mathbb{S}^1$ - groupe de translation

Données : courbes $Y_i \in L^2([0, 1])$ shiftées aléatoirement + bruit

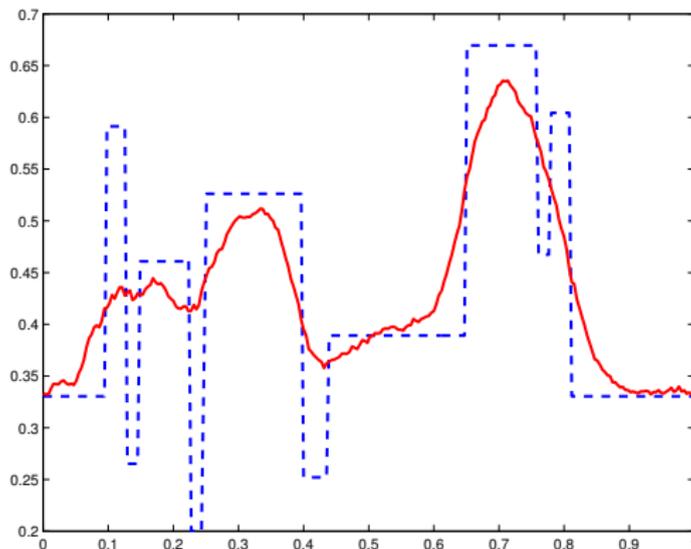
$Y_i(x) = f(x - g_i) + W_i(x)$, pour $x \in [0, 1]$, et g_i translations aléatoires iid

Courbe f / Sous-échantillon de 10 données ($n = 200$)



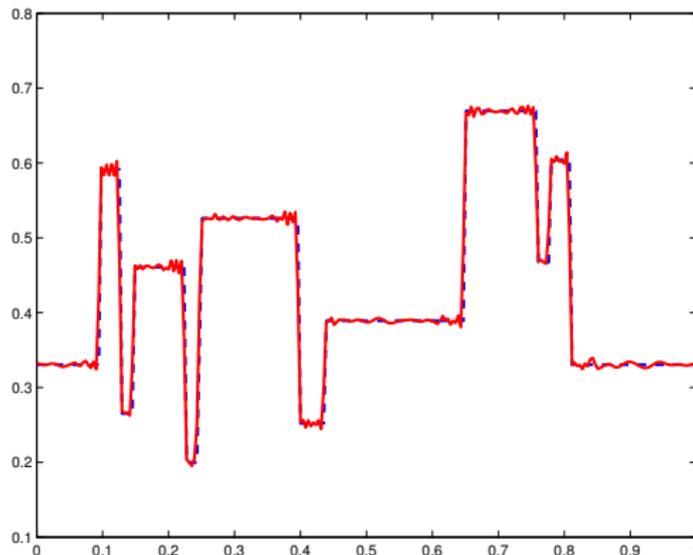
Exemple : moyenne de courbes

Moyenne euclidienne usuelle



Exemple : moyenne de courbes

Moyenne de Fréchet



Estimation par moyenne de Fréchet

Observations : Y_1, \dots, Y_n variables aléatoires iid à valeur dans \mathcal{H}

Schéma général d'estimation (registration/warping/alignment)

Estimation des paramètres de déformations

$$(\hat{g}_1, \dots, \hat{g}_n) = \arg \min_{(g_1, \dots, g_n) \in G^n} \left\{ \frac{1}{n} \sum_{i=1}^n \left\| g_i^{-1} \cdot Y_i - \frac{1}{n} \sum_{j=1}^n g_j^{-1} \cdot Y_j \right\|_{\mathcal{H}}^2 + \lambda D(g_i, e) \right\}$$

Source de variabilité géométrique

Inférence statistique et ACP à partir des $\hat{g}_1, \dots, \hat{g}_n$?

Moyenne de deux images

$$y \in L^2([0, 1]^2) \quad / \quad y' \in \mathcal{H} = L^2([0, 1]^2)$$



Moyenne de deux images



Moyenne euclidienne usuelle $(y + y')/2$

Moyenne de deux images

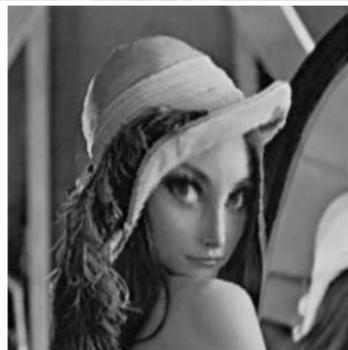
$$y \in L^2([0, 1]^2) \quad / \quad y' \in \mathcal{H} = L^2([0, 1]^2)$$



$$\hat{g}_\lambda = \arg \min_{g \in G} \left\{ \|y - g^{-1} \cdot y'\|_{\mathcal{H}}^2 + \lambda D(g, e) \right\},$$

où G ensemble de difféomorphismes de $[0, 1]^2$

Moyenne de deux images



Moyenne “géométrique” $([\hat{g}_\lambda/2] \cdot y + [\hat{g}_\lambda^{-1}/2] \cdot y') / 2$

Difféomorphismes générés par des petites déformations

Question : comment construire un difféomorphisme $\phi : \Omega \rightarrow \Omega$ avec $\Omega \subset \mathbb{R}^d$?

Idée : composer un ensemble de déformations infinitésimales

Difféomorphismes et composition de petites déformations

- Soit $v : \Omega \rightarrow \mathbb{R}^d$ une fonction lisse, alors si ϵ est suffisamment petit $\phi(x) = x + \epsilon v(x)$ est une fonction bijective
- Soit $v : \Omega \rightarrow \mathbb{R}^d$ une fonction lisse telle et $\epsilon > 0$ tels que

$$\phi_{p+1} = (Id + \epsilon v) \circ \dots \circ (Id + \epsilon v) = (Id + \epsilon v) \circ \phi_p = \phi_p + \epsilon v \circ \phi_p$$

est une suite de fonctions bijectives qui peut s'écrire comme

$$\frac{\phi_{p+1} - \phi_p}{\epsilon} = v \circ \phi_p,$$

- Si $\epsilon \rightarrow 0$, on obtient (en introduisant une variable de temps t)

$$\frac{\partial \phi_t(x)}{\partial t} = v(\phi_t(x)), \quad x \in \Omega$$

avec la condition initiale $\phi_0(x) = x$ pour tout $x \in \Omega$.

Difféomorphismes et composition de petites déformations

- Soit $v : \Omega \rightarrow \mathbb{R}^d$ une fonction lisse, alors si ϵ est suffisamment petit $\phi(x) = x + \epsilon v(x)$ est une fonction bijective
- Soit $v : \Omega \rightarrow \mathbb{R}^d$ une fonction lisse telle et $\epsilon > 0$ tels que

$$\phi_{p+1} = (Id + \epsilon v) \circ \dots \circ (Id + \epsilon v) = (Id + \epsilon v) \circ \phi_p = \phi_p + \epsilon v \circ \phi_p$$

est une suite de fonctions bijectives qui peut s'écrire comme

$$\frac{\phi_{p+1} - \phi_p}{\epsilon} = v \circ \phi_p,$$

- Si $\epsilon \rightarrow 0$, on obtient (en introduisant une variable de temps t)

$$\frac{\partial \phi_t(x)}{\partial t} = v(\phi_t(x)), \quad x \in \Omega$$

avec la condition initiale $\phi_0(x) = x$ pour tout $x \in \Omega$.

Difféomorphismes et composition de petites déformations

- Soit $v : \Omega \rightarrow \mathbb{R}^d$ une fonction lisse, alors si ϵ est suffisamment petit $\phi(x) = x + \epsilon v(x)$ est une fonction bijective
- Soit $v : \Omega \rightarrow \mathbb{R}^d$ une fonction lisse telle et $\epsilon > 0$ tels que

$$\phi_{p+1} = (Id + \epsilon v) \circ \dots \circ (Id + \epsilon v) = (Id + \epsilon v) \circ \phi_p = \phi_p + \epsilon v \circ \phi_p$$

est une suite de fonctions bijectives qui peut s'écrire comme

$$\frac{\phi_{p+1} - \phi_p}{\epsilon} = v \circ \phi_p,$$

- Si $\epsilon \rightarrow 0$, on obtient (en introduisant une variable de temps t)

$$\frac{\partial \phi_t(x)}{\partial t} = v(\phi_t(x)), \quad x \in \Omega$$

avec la condition initiale $\phi_0(x) = x$ pour tout $x \in \Omega$.

Difféomorphismes et composition de petites déformations

Proposition (Trouvé, Younes)

Soit $\mathcal{V} = \{v : \Omega \rightarrow \mathbb{R}^d\}$ un espace de Hilbert de champ de vecteurs tels que

$$v(x) = 0 \text{ pour } x \in \partial\Omega$$

Alors

- il existe une unique solution de l'EDO $\frac{\partial \phi_t}{\partial t} = v(\phi_t)$ avec la condition initiale $\phi_0(x) = x$ i.e.

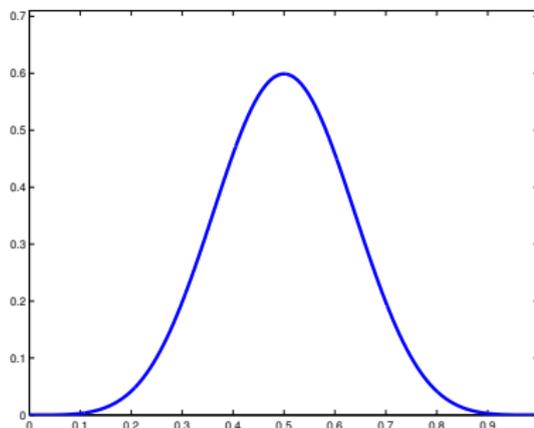
$$\phi_t(x) = x + \int_0^t v(\phi_s(x)) ds$$

- pour tout $t \in [0, 1]$, ϕ_t est un difféomorphisme de Ω dans Ω tel que

$$\frac{\partial \phi_t^{-1}}{\partial t} = -v(\phi_t^{-1})$$

Exemple numérique

Construction d'un difféomorphisme de $[0, 1]$

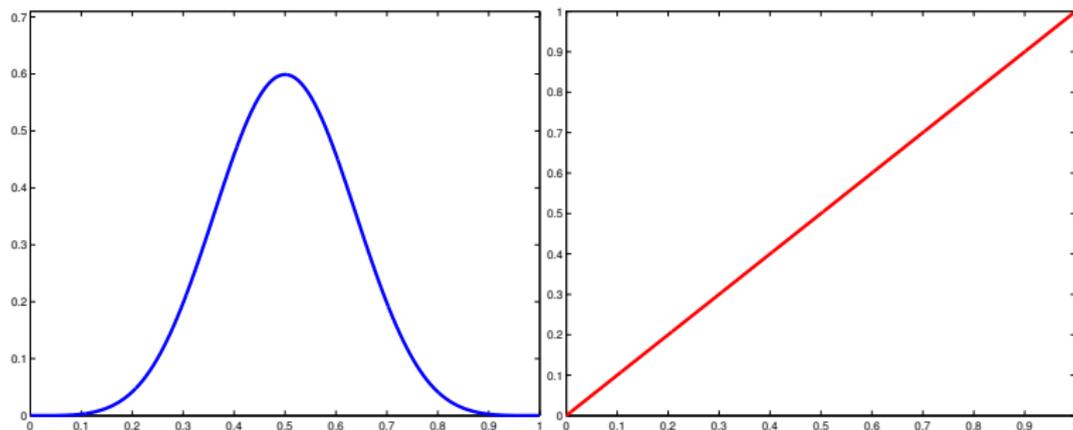


Champ de vecteur $v : [0, 1] \rightarrow \mathbb{R}$

$$v(0) = v(1) = 0$$

Exemple numérique

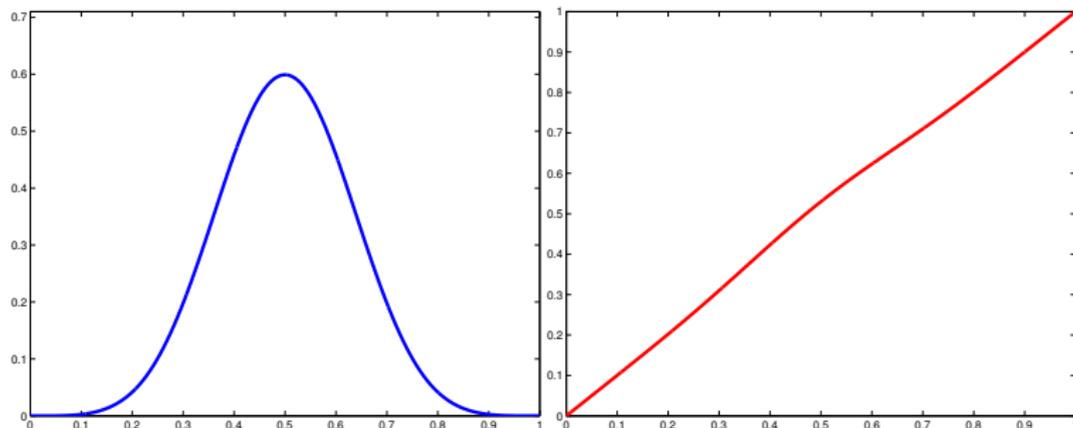
Construction d'un difféomorphisme de $[0, 1]$ (**fonction monotone**)



$$\phi_t(x) = x \text{ au temps } t = 0$$

Exemple numérique

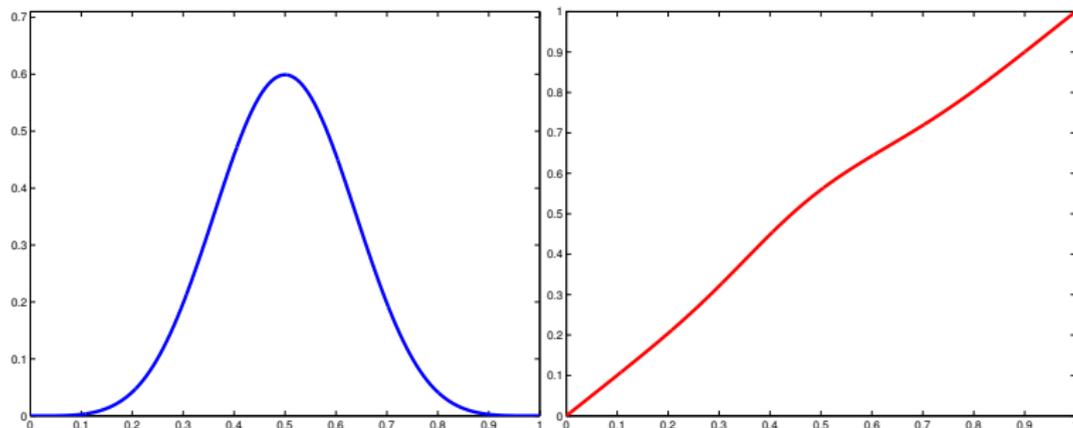
Construction d'un difféomorphisme de $[0, 1]$ (**fonction monotone**)



$$\phi_t(x) = x + \int_0^t v(\phi_s(x)) ds \text{ au temps } t = 0.05$$

Exemple numérique

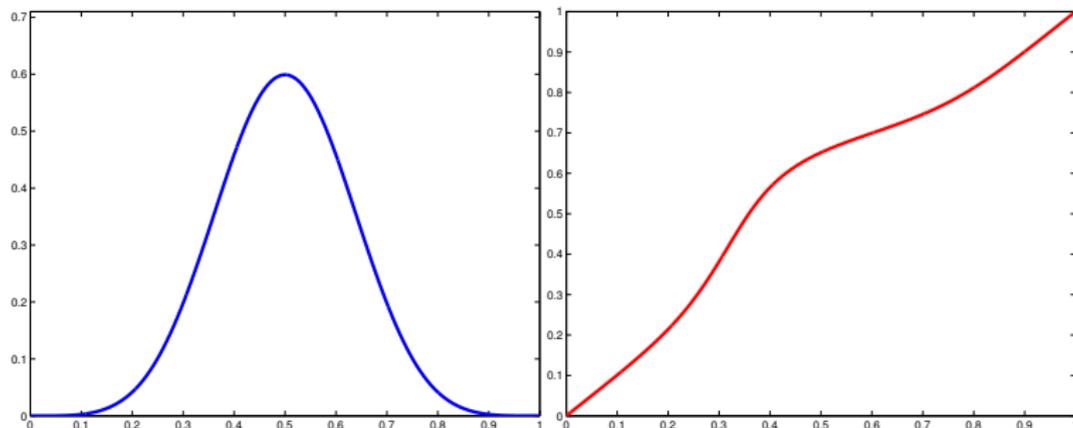
Construction d'un difféomorphisme de $[0, 1]$ (**fonction monotone**)



$$\phi_t(x) = x + \int_0^t v(\phi_s(x)) ds \text{ au temps } t = 0.1$$

Exemple numérique

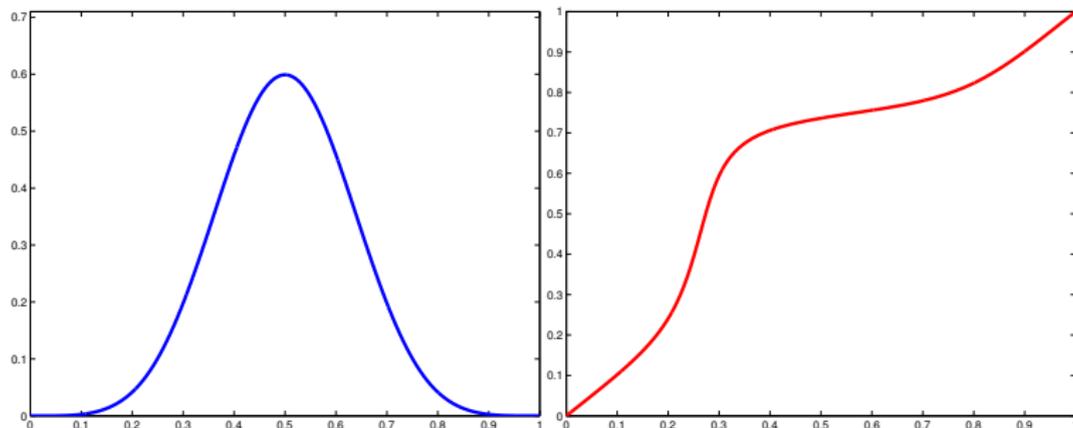
Construction d'un difféomorphisme de $[0, 1]$ (**fonction monotone**)



$$\phi_t(x) = x + \int_0^t v(\phi_s(x)) ds \text{ au temps } t = 0.3$$

Exemple numérique

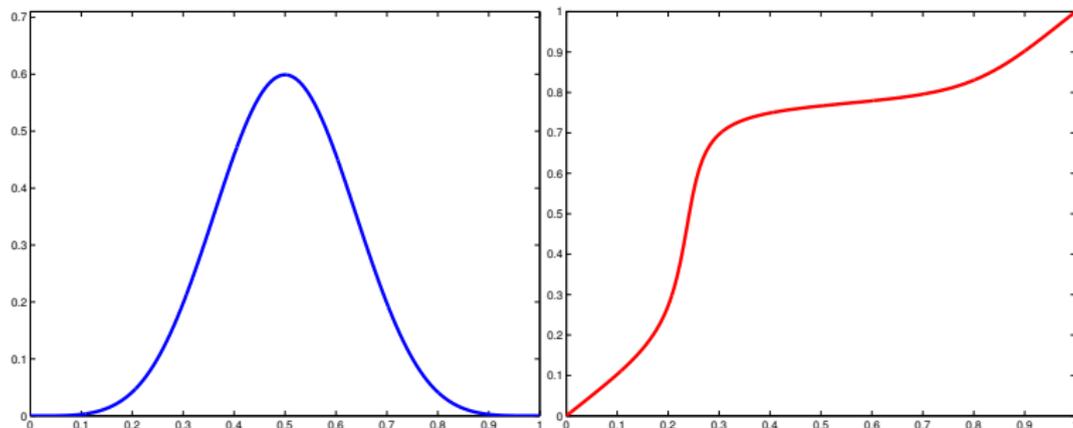
Construction d'un difféomorphisme de $[0, 1]$ (**fonction monotone**)



$$\phi_t(x) = x + \int_0^t v(\phi_s(x)) ds \text{ au temps } t = 0.7$$

Exemple numérique

Construction d'un difféomorphisme de $[0, 1]$ (**fonction monotone**)



$$\phi(x) := \phi_1(x) = x + \int_0^t v(\phi_s(x)) ds \text{ au temps } t = 1$$

Moyenne de deux images

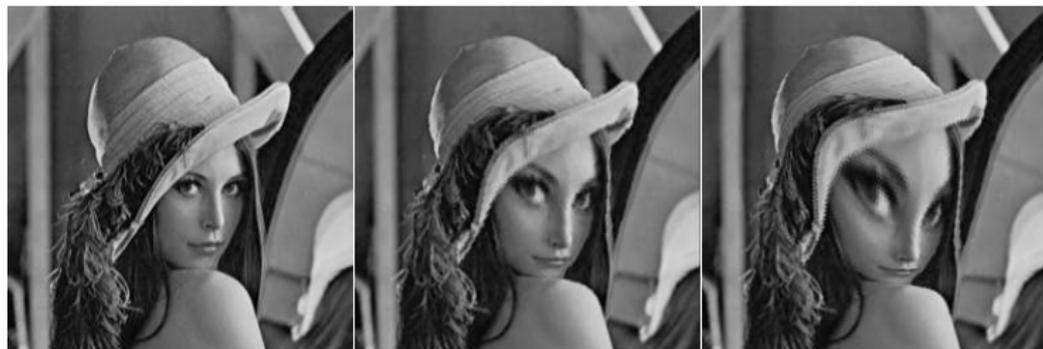
$$y \in L^2([0, 1]^2) \quad / \quad y' \in \mathcal{H} = L^2([0, 1]^2)$$



$$\hat{v}_\lambda = \arg \min_{v \in \mathcal{V}} \left\{ \int_{\Omega} |y(x) - y' \circ \phi^v(x)|^2 dx + \lambda \|v\|_{\mathcal{V}}^2 \right\},$$

$$\text{où } \phi^v(x) = \phi_1(x) = x - \int_0^1 v(\phi_t^v(x)) dt$$

Moyenne de deux images



Moyenne “géométrique” $([\hat{g}_\lambda/2] \cdot y + [\hat{g}_\lambda^{-1}/2] \cdot y') / 2$ avec

$$[\hat{g}_\lambda^{-1}/2] = \phi^{\hat{v}_\lambda/2}(x)$$

et

$$[\hat{g}_\lambda/2] = \phi^{-\hat{v}_\lambda/2}(x)$$

Estimation par moyenne de Fréchet

Observations : Y_1, \dots, Y_n variables aléatoires iid à valeur dans $L^2(\Omega)$

Estimation des paramètres de déformations

$$(\hat{v}_1, \dots, \hat{v}_n) = \arg \min_{(v_1, \dots, v_n) \in \mathcal{V}^n} \left\{ \frac{1}{n} \sum_{i=1}^n \left\| Y_i \circ \phi^{v_i} - \frac{1}{n} \sum_{j=1}^n Y_j \circ \phi^{v_j} \right\|_{\mathcal{H}}^2 + \lambda \|v_i\|_{L^2(\Omega)}^2 \right\}$$

où $\phi^{v_i}(x) := \phi_1(x) = x - \int_0^1 v_i(\phi_t(x)) dt$

Source de variabilité géométrique

Inférence statistique et ACP sur les variables aléatoires $\hat{v}_1, \dots, \hat{v}_n$ à valeur dans \mathcal{V} espace de Hilbert

Estimation par moyenne de Fréchet

Observations : Y_1, \dots, Y_n variables aléatoires iid à valeur dans $L^2(\Omega)$

Estimation des paramètres de déformations

$$(\hat{v}_1, \dots, \hat{v}_n) = \arg \min_{(v_1, \dots, v_n) \in \mathcal{V}^n} \left\{ \frac{1}{n} \sum_{i=1}^n \left\| Y_i \circ \phi^{v_i} - \frac{1}{n} \sum_{j=1}^n Y_j \circ \phi^{v_j} \right\|_{\mathcal{H}}^2 + \lambda \|v_i\|_{L^2(\Omega)}^2 \right\}$$

où $\phi^{v_i}(x) := \phi_1(x) = x - \int_0^1 v_i(\phi_t(x)) dt$

Source de variabilité géométrique

On note $\hat{u}_1, \hat{u}_2, \dots$ les vecteurs propres associés aux plus grandes valeurs propres $e_1 \geq e_2 \geq \dots \geq 0$ de l'opérateur de covariance des v.a. $\hat{v}_1, \dots, \hat{v}_n$ (à valeur dans \mathcal{V} espace de Hilbert)

Estimation par moyenne de Fréchet

Observations : Y_1, \dots, Y_n variables aléatoires iid à valeur dans $L^2(\Omega)$

Estimation des paramètres de déformations

$$(\hat{v}_1, \dots, \hat{v}_n) = \arg \min_{(v_1, \dots, v_n) \in \mathcal{V}^n} \left\{ \frac{1}{n} \sum_{i=1}^n \left\| Y_i \circ \phi^{v_i} - \frac{1}{n} \sum_{j=1}^n Y_j \circ \phi^{v_j} \right\|_{\mathcal{H}}^2 + \lambda \|v_i\|_{L^2(\Omega)}^2 \right\}$$

où $\phi^{v_i}(x) := \phi_1(x) = x - \int_0^1 v_i(\phi_t(x)) dt$

Source de variabilité géométrique

On note $\hat{\psi}_1, \hat{\psi}_2$ les principaux modes de déformations qui sont des difféomorphismes de Ω définis par

$$\hat{\psi}_1(x) := \phi_1(x) = x - \int_0^1 \sqrt{e_1} \hat{u}_1(\phi_t(x)) dt$$

Estimation par moyenne de Fréchet

Observations : Y_1, \dots, Y_n variables aléatoires iid à valeur dans $L^2(\Omega)$

Estimation des paramètres de déformations

$$(\hat{v}_1, \dots, \hat{v}_n) = \arg \min_{(v_1, \dots, v_n) \in \mathcal{V}^n} \left\{ \frac{1}{n} \sum_{i=1}^n \left\| Y_i \circ \phi^{v_i} - \frac{1}{n} \sum_{j=1}^n Y_j \circ \phi^{v_j} \right\|_{\mathcal{H}}^2 + \lambda \|v_i\|_{L^2(\Omega)}^2 \right\}$$

où $\phi^{v_i}(x) := \phi_1(x) = x - \int_0^1 v_i(\phi_t(x)) dt$

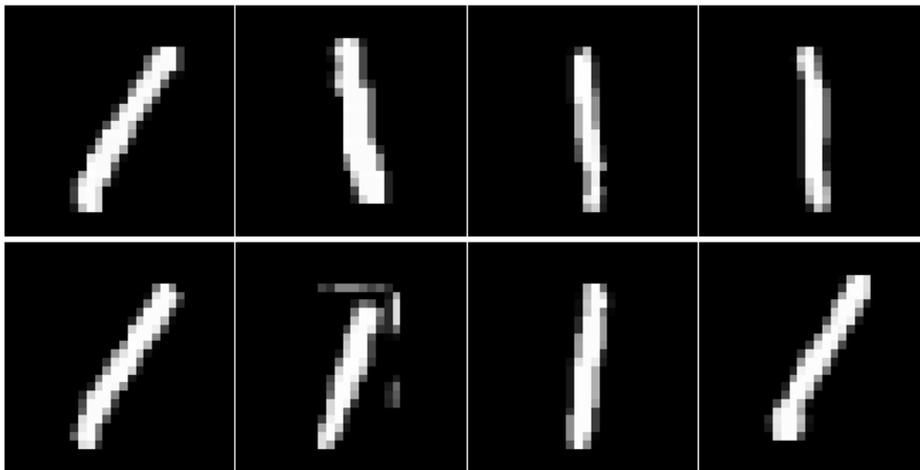
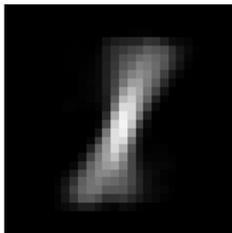
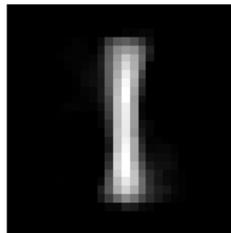
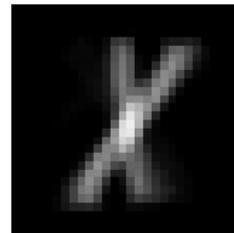
Source de variabilité géométrique

Premier et deuxième modes de variabilité géométrique

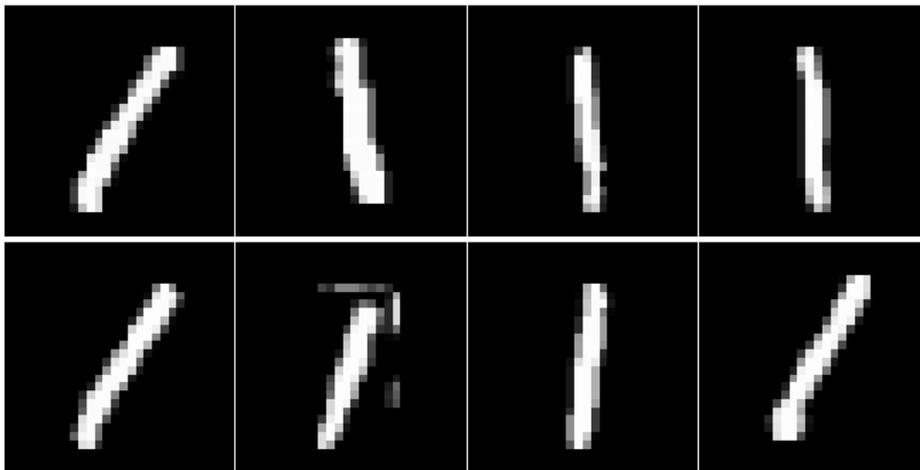
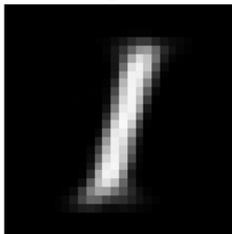
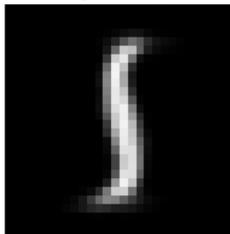
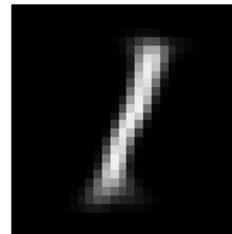
$$\hat{y}_n \circ \hat{\psi}_1 \text{ et } \hat{y}_n \circ \hat{\psi}_2$$

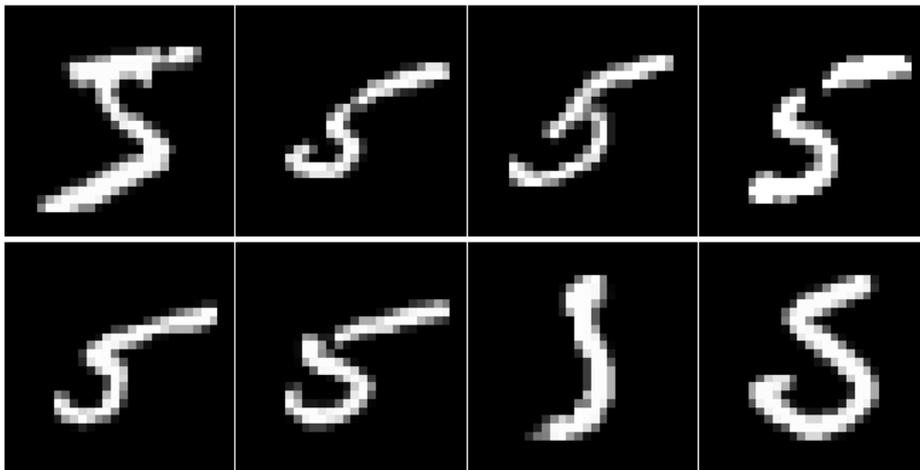
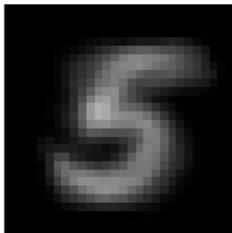
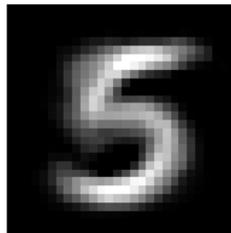
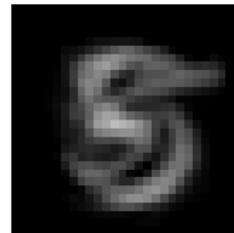
où \hat{y}_n est une moyenne de Fréchet (associée à un groupe de difféomorphismes)

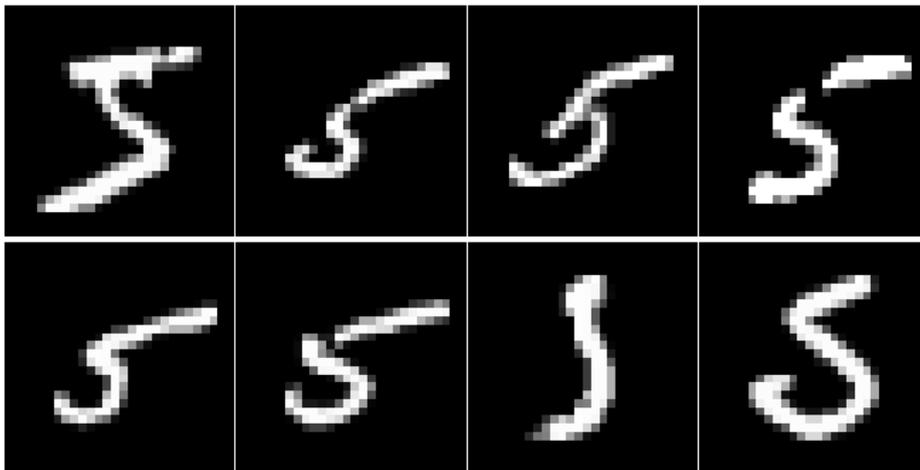
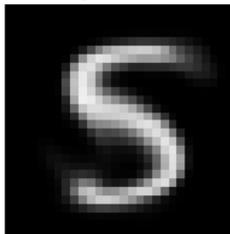
ACP Standard (Mnist database $n = 100$)

 \bar{Y}_n  $\bar{Y}_n + \sqrt{\lambda_1} \hat{w}_1$  $\bar{Y}_n + \sqrt{\lambda_2} \hat{w}_2$ 

ACP Géométrique (Mnist database $n = 100$)

 \hat{y}_n  $\hat{y}_n \circ \hat{\psi}_1$  $\hat{y}_n \circ \hat{\psi}_2$ 

ACP Standard (Mnist database $n = 100$) \bar{Y}_n  $\bar{Y}_n + \sqrt{\lambda_1} \hat{w}_1$  $\bar{Y}_n + \sqrt{\lambda_2} \hat{w}_2$ 

ACP Géométrique (Mnist database $n = 100$) \hat{y}_n  $\hat{y}_n \circ \hat{\psi}_1$  $\hat{y}_n \circ \hat{\psi}_2$ 

- 1 Motivations
- 2 Choix d'une distance
- 3 Moyenne de Fréchet
- 4 Mesures de dissimilarité et ACP géométrique
 - Action de groupe
 - Mesure de dissimilarité
 - ACP géométrique
- 5 Statistique non-paramétrique et modèles déformables

Lien avec les problèmes inverses B. & Gadat (2010)

Modèle le plus simple : courbes shiftées aléatoirement

$$dY_i(x) = f(x - \tau_i)dx + \epsilon dW_i(x), \quad x \in [0, 1], \quad i = 1, \dots, n, \quad \text{avec } \tau_i \sim_{iid} g$$

Moyenne empirique usuelle :

$$\bar{f}_n(x) = \frac{1}{n} \sum_{i=1}^n Y_i(x) \xrightarrow{n \rightarrow +\infty} \mathbb{E}Y_1(x) = \mathbb{E}f(x - \tau_1) = \int f(x - \tau)g(\tau)d\tau$$

Problème de déconvolution ?

Courbes shiftées aléatoirement

Contributions : étude de l'estimation de $f \in \mathcal{F} \subset L^2([0, 1])$ au sens du risque minimax

$$\mathcal{R}_n(\mathcal{F}) = \inf_{\hat{f}_n} \sup_{f \in \mathcal{F}} \mathbb{E} \|\hat{f}_n - f\|_{L^2}^2$$

Hypothèse : les coefficients de Fourier de g décroissent à la vitesse polynomiale $\nu > 0$ i.e. $\int g(x)e^{-i2\pi\ell x} dx \sim |\ell|^{-\nu}$ pour $\ell \in \mathbb{Z}$.

Résultats : $\mathcal{R}_n(\mathcal{F}) \sim n^{-\frac{2s}{2s+2\nu+1}}$ pour $\mathcal{F} = W^s(A)$ (boule de Sobolev) ou $\mathcal{F} = B_{p,q}^s(A)$ (boule de Besov)

Courbes shiftées aléatoirement

Contributions : étude du problème de la reconstruction des shifts $\tau_i, i = 1, \dots, n$ (variabilité géométrique)

Résultats : à partir de la notion de moyenne de Fréchet construction d'estimateurs $\hat{\tau}_i$ qui converge (en probabilité) vers τ_i avec **filtrage passe-bas** des courbes bruitées + **asymptotique** $\epsilon \rightarrow 0$ (niveau du bruit)

Borne inférieure pour l'estimation des shifts

$$\mathbb{E} \left(\frac{1}{n} \sum_{m=1}^n (\hat{\tau}_m - \tau_m)^2 \right) \geq \frac{\epsilon^2}{\|f'\|^2 + \epsilon^2 \int \left(\frac{\partial}{\partial \tau} \log g(\tau) \right)^2 g(\tau) d\tau} \xrightarrow{n \rightarrow +\infty} 0$$

Choix de l'asymptotique ?

- $n \rightarrow +\infty$ et ϵ fixé
- n fixé et $\epsilon \rightarrow 0$
- $n \rightarrow +\infty$ et $\epsilon \rightarrow 0$