

# Local Integrated Dual Depth

## Una Profundidad Local en Espacios de Banach

Lucas Fernández Piana<sup>1</sup>    Marcela Svarc<sup>2</sup>  
lucasfernandezpiana@gmail.com

<sup>1</sup>Departamento de Matemática, UBA and IMAS-CONICET

<sup>2</sup>Universidad de San Andrés, Departamento de Matemática and CONICET

September 20, 2016

# Profundidades

Dada una medida de probabilidad  $P$  en  $\mathbb{R}^d$ . Una **profundidad**,  $D(x, P)$ , es una función que provee una noción de orden para puntos  $x \in \mathbb{R}^d$  basada en  $P$  del centro hacia afuera.

Interpretar “orden del centro hacia afuera sugiere dos cosas”

- ▶ Una noción relevante de centro.
- ▶ Los puntos que se encuentren más cercanos al centro deberán ser más profundos que los que estén más alejados.

A partir de estas consideraciones se desprende que el “centro” debe ser un punto que maximice globalmente la función de profundidad.

Yijun Zuo y Robert Serfling en su artículo del *Annals of Statistics*, “General Notion of Statistical Depth Function” (2000) dan una lista de propiedades deseables que debe cumplir una función acotada y no negativa en  $\mathbb{R}^d$  para ser una profundidad.

### P.1. Invarianza Afín

La profundidad de un punto  $x \in \mathbb{R}^d$  no debería depender del sistema de coordenadas, en particular de la escala del sistema de medición.

### P.2. Maximalidad al centro

Para una distribución que tenga un único *centro* definido, por ejemplo, un punto de simetría, la profundidad se maximizaría en dicho punto.

### **P.3. Monotonía respecto del punto de máxima profundidad**

A medida que  $x \in \mathbb{R}^d$  se aleja del punto *más profundo* a lo largo de cualquier rayo fijo la profundidad de  $x$  debería decrecer.

### **P.4. Se anula en el infinito**

La profundidad de un punto  $x$  debería aproximarse a cero si  $\|x\|$  es demasiado grande.

### **P.5. Continuidad como función de $x$**

Es esperable pensar que si dos puntos se encuentran lo suficientemente cerca las profundidades no sean demasiado distintas

### **P.6. Continuidad como función de $P$ .**

Análogo a la propiedad anterior para distribuciones.

**Notación:** Llamamos  $\mathbb{P}$  a la clase de distribuciones sobre los conjuntos Borelianos de  $\mathbb{R}^d$  y llamamos  $P_\xi$  la distribución dada por el vector aleatorio  $\xi$ .

**Definición:** Sea  $D(\cdot, \cdot) : \mathbb{R}^d \times \mathbb{P} \rightarrow \mathbb{R}$  una función acotada y no negativa que satisfice:

1.  $D(Ax + b, P_{AX+b}) = D(x, P_X)$  para todo vector aleatorio  $X \in \mathbb{R}^d$  vector aleatorio, para todo matriz  $A \in R^{d \times d}$  no singular y para cualquier vector  $b \in \mathbb{R}^d$ .
2.  $D(\theta, P) = \sup_{x \in R^d} D(x, P)$ , para cualquier  $P \in \mathbb{P}$  con centro  $\theta$ .
3. Para toda  $P \in \mathbb{P}$  con punto más profundo  $\theta$   
 $D(x, P) \leq D(\theta + \alpha(x - \theta), P)$ , para cualquier  $\alpha \in [0, 1]$ .
4.  $\lim_{\|x\| \rightarrow +\infty} D(x, P) = 0$ , para todo  $P \in \mathbb{P}$ .

# Profundidad del Semi-espacio

La semilla de esta teoría fue plantada por Tukey (1975) con la *Profundidad del Semi-espacio* (Halfspace Depth, HD).

La profundidad del semi-espacio de un punto  $x \in \mathbb{R}^d$  con respecto a una medida de probabilidad  $P$  en  $\mathbb{R}^d$  está definida por la mínima probabilidad de que un semi-espacio cerrado que contenga a  $x$ , esto es,

$$HD(x, P) = \inf \{P(H) : H \text{ semi-espacio cerrado, } x \in H\}, \quad x \in \mathbb{R}^d.$$

# Profundidad Simplicial

Otra noción importante fue planteada por Liu (1990). La *Profundidad Simplicial (SD)* de un punto  $x \in \mathbb{R}^d$  con respecto a una medida de probabilidad  $P \in \mathbb{R}^d$  es la probabilidad de que  $x$  pertenezca a un simplex aleatorio en  $\mathbb{R}^d$ , i.e,

$$SD(x, P) = P(x \in S[X_1, \dots, X_{d+1}]), \quad x \in \mathbb{R}^d.$$

donde  $X_1, \dots, X_{d+1}$  es una muestra de vectores aleatorios con distribución  $P$  y  $S[X_1, \dots, X_{d+1}]$  es el simplex  $d$ -dimensional con vértices  $X_1, \dots, X_{d+1}$ .

## Ejemplo univariado

Si consideramos una medida de probabilidad en  $\mathbb{R}$  y  $X$  una variable aleatoria con función de distribución acumulada  $F$  tenemos,

$$HD(x, P) = \min\{P(X \leq x), 1 - P(X < x)\} = \min\{F(x), 1 - F(x^-)\}.$$

$$SD(x, P) = 2P(X \leq x)(1 - P(X < x)) = 2F(x)(1 - F(x^-)).$$

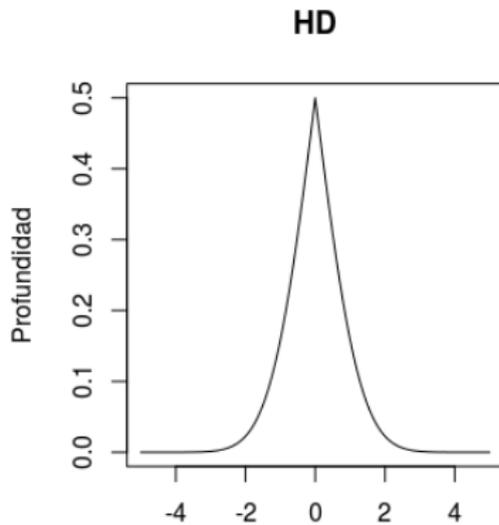
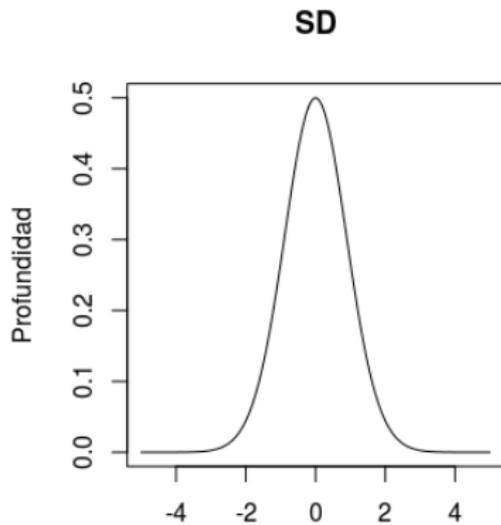


Figure : Profundidad para  $N(0,1)$

# Datos Funcionales

La noción de profundidad se puede extender al caso en que el espacio de las observaciones tenga dimensión infinita, por ejemplo un espacio de Hilbert o un espacio de Banach separable.

- ▶ Profundidad de Tukey Aleatorizada (Cuesta-Albertos y Nieto-Reyes, 2008)
- ▶ Profundidad Integrada (Fraiman y Muniz, 2001)
- ▶ Profundidad por Bandas (López-Pintado y Romo, 2009)
- ▶ Profundidad Dual Intregrada (Cuevas y Fraiman, 2009)

# Profundidad Dual Integrada (IDD)

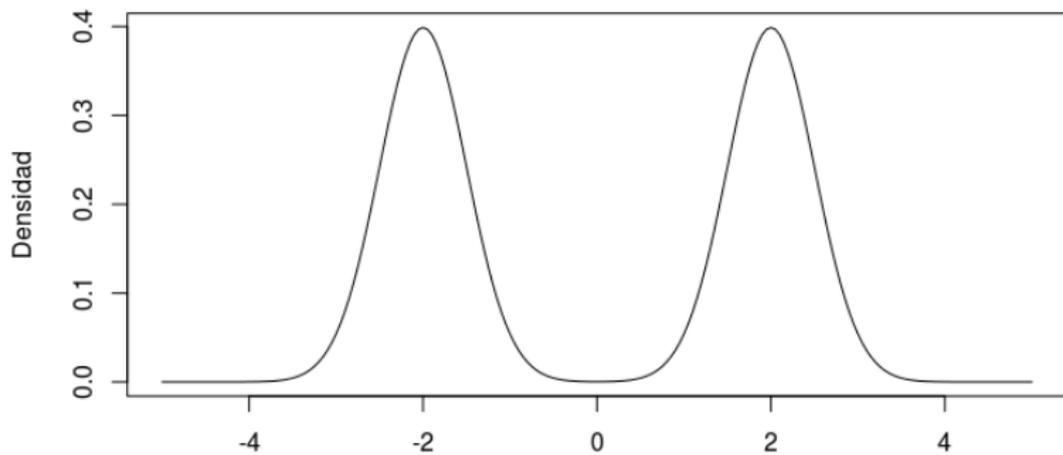
Sea  $\Omega$  un espacio de probabilidad y  $\mathbb{E}$  un espacio de Banach separable con dual  $\mathbb{E}'$  separable. Sea  $X : \Omega \rightarrow \mathbb{E}$  un elemento aleatorio en  $\mathbb{E}$  con distribución  $P$ .

Se define la *Profundidad Dual Integrada* (IDD) como

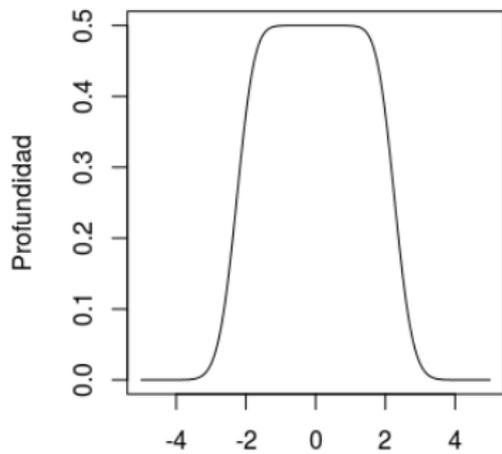
$$D(P, x) = \int D(P_f, f(x)) dQ(f), \quad (1)$$

donde  $f \in \mathbb{E}'$ ,  $x \in \mathbb{E}$  y  $P_f = P_{f(X)}$  es la distribución de una variable aleatoria  $f(X)$ ,  $Q$  es una medida de probabilidad en  $E'$  y  $D$  es la Profundidad del Semi-Espacio (HD) o la Profundidad Simplicial (SD) univariada.

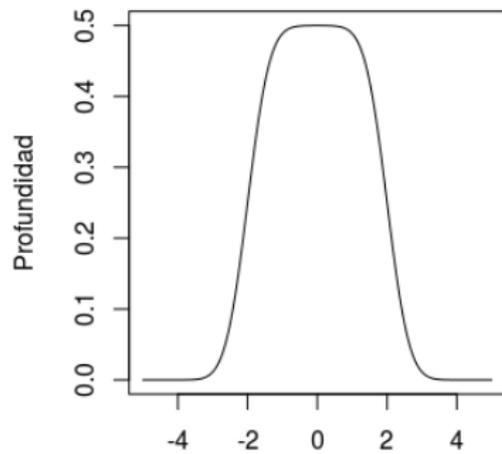
**Densidad Mixtura (N(-2,0.25),N(2,0.25))**



**SD**



**HD**



Las profundidades globales no detectan propiedades locales. No pueden lidiar con múltiples centros, encontrar su ubicación e importancia relativa. La profundidad local debería ser capaz de resolver estos problemas.

En caso de tener una distribución unimodal tendrá que comportarse como una profundidad global.

Algunas aplicaciones pueden ser:

- ▶ Distribuciones mixtas.
- ▶ Clusters.

# Propuestas

Algunas propuestas se han hecho en estos últimos años

- ▶ Local Depth (Agostinelli y Romanazzi, 2011)
- ▶ From Depth to Local Depth: A Focus on Centrality (Paindaveine y Van Bever, 2013)
- ▶ Local Half-Region Depth for Functional Data (Agostinelli, 2016)

## Agostinelli-Romanazzi (2011)

La propuesta para la profundidad local es restringir la cantidad de simplices a considerar dependiendo del volumen. Es decir, sea  $X$  en  $\mathbb{R}^d$  un vector aleatorio con distribución  $F$ ,  $x \in \mathbb{R}^d$  y  $\tau > 0$ . Definen la profundidad local simplicial como

$$LSD(x, F, \tau) = P_F(S_{p+1} : x \in S_{p+1} \cap \text{Vol}(S_{p+1}) \leq \tau),$$

donde  $S_{p+1}$  es un simplex aleatorio de vértices  $X_1, \dots, X_{p+1}$  con  $\{X_i\}_{i=1}^{p+1}$  independientes con la misma distribución que  $X$ .

# Paindaveine y Van Bever, 2013

Proponen un método para construir una profundidad local a partir de una global.

**Definición:** una región de profundidad de orden  $\alpha$  se define como el conjunto  $R_\alpha(P) = \{x \in \mathbb{R}^d : D(x, P) \geq \alpha\}$ .

**Definición:** sea  $\beta \in (0, 1]$  definen  $R^\beta(P) = \bigcap_{\alpha \in A(\beta)} R_\alpha(P)$  donde  $A(\beta) = \{\alpha > 0 : P(R_\alpha(P)) \geq \beta\}$ .

**Notación:** sean  $X$  un vector aleatorio en  $\mathbb{R}^d$ ,  $P$  la una medida de probabilidad asociada y  $x \in \mathbb{R}^d$ . Llamamos  $P^X = P$  y  $P_x = \frac{1}{2}P^X + \frac{1}{2}P^{2x-X}$ .

**Definición:** se define el entorno de profundidad de  $x$  de orden  $\alpha$  con respecto a  $P$  como  $R_{\alpha,x}(P) = R_{\alpha}(P_x)$ .

**Definición:** se define el entorno de probabilidad de  $x$  de nivel  $\beta$  con respecto a  $P$  como  $R_x^{\beta}(P) = R^{\beta}(P_x)$ .

**Definición:** sea  $D(\cdot, P)$  una función de profundidad. Se define la correspondiente función de profundidad local de nivel de localización  $\beta$  como

$$\begin{aligned} LD^{\beta}(\cdot, P) &: \mathbb{R}^d \longrightarrow \mathbb{R} \\ LD^{\beta}(x, P) &= D(x, P_x^{\beta}) \end{aligned}$$

donde  $P_x^{\beta} = P(\cdot | R_x^{\beta}(P))$ .

## Ejemplo Univariado

Sea  $X$  variable aleatoria continua con función de distribución acumulada  $F$ , sea  $x \in \mathbb{R}$  y sea  $\beta \in (0, 1)$ .

Sea  $\lambda_x = \inf\{\lambda > 0 : F(x + \lambda) - F(x - \lambda) \geq \beta\}$ , entonces:

$$LHD^\beta(x, F) = \frac{1}{\beta} \min(F(x + \lambda_x) - F(x), F(x) - F(x - \lambda_x)),$$

$$LSD^\beta(x, F) = \frac{2}{\beta^2} (F(x + \lambda_x) - F(x))(F(x) - F(x - \lambda_x)).$$

# Profundidad Local Dual Integrada

Sea  $(\Omega, \Sigma, P)$  un espacio de probabilidad. Sean  $E$  un espacio de Banach separable,  $E'$  su dual y  $X : \Omega \rightarrow E$  un elemento aleatorio.

Para cada funcional,  $f \in E'$ , tenemos que  $f(X)$  es una variable aleatoria con una probabilidad asociada  $P_f$ . Donde

$$P_f((-\infty, t)) = P(f^{-1}(-\infty, t)) \quad \forall t \in \mathbb{R}.$$

Análogamente, tiene una función de distribución acumulada asociada  $F_f(t) = P(f(X) \leq t) \quad \forall t \in \mathbb{R}$ .

Dado  $\beta \in (0, 1)$  y  $x \in E$  definimos:

$$\lambda_{f(x)}^\beta = \inf \{ \lambda > 0 : F_f(f(x) + \lambda) - F_f(f(x) - \lambda) \geq \beta \}.$$

# Profundidad Local Dual Integrada

**Definición:** Sea  $(\Omega, \Sigma, P)$  un espacio de probabilidad. Sean  $E$  un espacio de Banach separable,  $E'$  su dual y  $X : \Omega \rightarrow E$  un elemento aleatorio continuo. Definimos la Profundidad Local Dual Integrada de nivel  $\beta$ :

$$\begin{aligned} LD^\beta : E &\longrightarrow \left[0, \frac{1}{2}\right] \\ LD^\beta(x, P) &= \int LSD^\beta(f(x), P_f) dQ(f) \end{aligned}$$

donde  $Q$  es una medida de probabilidad en  $E'$  independiente de  $P$ .

**Definición:**  $\theta \in E$  es un punto de  $\beta$ -simetría si  $F_{f(X)}(f(\theta) + \lambda\theta) - F_{f(X)} = \frac{\beta}{2}$ .

# Propiedades

- P.1.** Sólo se cumple para transformaciones ortonormales.
- P.2.** Se maximiza globalmente en cada punto de  $\beta$ -simetría.
- P.3.** Se cumple en un entorno de cada punto de  $\beta$ -simetría.
- P.4.** Requiere que la medida  $Q$  cumpla que  $\sup_{\|x\|=1} \{f : f(x) \leq (\epsilon)\} = O(\epsilon)$ .
- P.5.** Es continua respecto de  $x$  en norma.
- P.6.** Si  $P_n \xrightarrow[n \rightarrow +\infty]{w} P$ , entonces  $LD^\beta(x, P_n) \xrightarrow[n \rightarrow +\infty]{P} LD^\beta(x, P)$ .

# Versión Empírica

**Definición:** Sea  $\beta \in (0, 1)$ ,  $X : \Omega \rightarrow E$  un elemento aleatorio continuo y  $X_1, \dots, X_n$  una muestra aleatoria con la misma distribución que  $X$ . Para cada  $x \in E$ ,  $f \in E'$  y sea  $k = \lfloor n\beta \rfloor$  definimos

$$\lambda_{f(x),n}^\beta = \inf \left\{ \lambda > 0 : F_{f,n}(f(x) + \lambda) - F_{f,n}(f(x) - \lambda) \geq \frac{k}{n} \right\}$$

Sea  $\beta(k) = \frac{k}{n}$ . Definimos la versión empírica de la Profundidad Local Dual Integrada de nivel  $\beta(k)$ .

$$ELD^{\beta(k)}(x, P) = LD^{\beta(k)}(x, P_n)$$

# Consistencia

**Teorema:** Sea  $X$  un elemento aleatorio continuo en un espacio de Banach  $E$  separable con medida de probabilidad asociada  $P$ . Sea  $X_1, \dots, X_n$  una muestra aleatoria en  $E$  con la misma distribución que  $X$  y sea  $\beta \in (0, 1)$ . Entonces,

$$\text{a) } \sup_{P \in \mathbb{P}} E \left[ \sup_{x \in E} \left| ELD^{\beta(k)}(x, P) - LD^{\beta}(x, P) \right| \right] \xrightarrow{n \rightarrow +\infty} 0.$$

$$\text{b) } \sup_{x \in E} \left| ELD^{\beta(k)}(x, P) - LD^{\beta}(x, P) \right| \xrightarrow{n \rightarrow +\infty} 0 \text{ c.t.p.}$$

# LDC

Proponemos un método de clusters (LDC) basado en centroides que se calculan a partir de una profundidad local.

El centro de cada cluster en vez de ser un sólo punto es definido por una región núcleo. Estas regiones se calculan a partir de las regiones de profundidad local. Es decir, que cada núcleo será una componente conexa de  $R_\alpha(P)$ . El valor de  $\alpha$  es un número entre 0 y  $\frac{1}{2}$  que debe ser determinado por el usuario.

Esperamos que LDC tenga la habilidad de detectar la verdadera estructura de los clusters para un rango importante de distribuciones debido a que las regiones núcleo representan mejor la estructura que un sólo punto. Las funciones de profundidad tiene una relación cercana con la robustez por lo que también esperamos que las regiones núcleo sean resistentes a la presencia de outliers.

## Algoritmo

Sean  $X_1, \dots, X_n$  observaciones en  $\mathcal{H}$ , un espacio de Hilbert separable agrupados en  $K$  clusters  $C_1, \dots, C_K$ . Sea  $ld_i = ELD^\beta(X_i)$  para  $i = 1, \dots, n$  la profundidad local correspondiente a la observación  $i$  –ésima.

LDC se define en dos pasos:

**Paso 1** Sea  $\mathcal{S} = \{X_{i_1}, \dots, X_{i_n}\}$ , el subconjunto de observaciones donde  $ld_i > \alpha$ ,  $i = 1, \dots, n$ . Asignar las observaciones de  $\mathcal{S}$  en  $K$  clusters,  $\tilde{C}_1, \dots, \tilde{C}_K$ .

**Paso 2** Sea  $\mathcal{S}^c = \{X_i, \dots, X_{i_n}\} \setminus \mathcal{S}$  el subconjunto de observaciones con  $ld_i \leq \alpha$ ,  $i = 1, \dots, n$ . Asignar la observación  $X_i \in \mathcal{S}^c$  al cluster más cercano definido en el paso anterior,  $\tilde{C}_j$ ,  $j = 1, \dots, K$ , i.e.,

$$\hat{k}_i = \arg \min_{k=1, \dots, K} d(X_i, \tilde{C}_k),$$

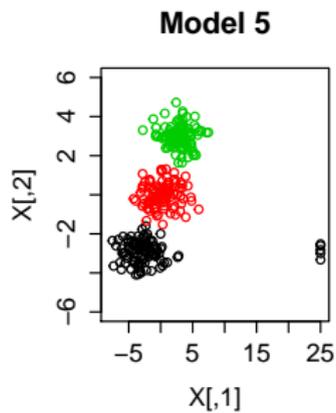
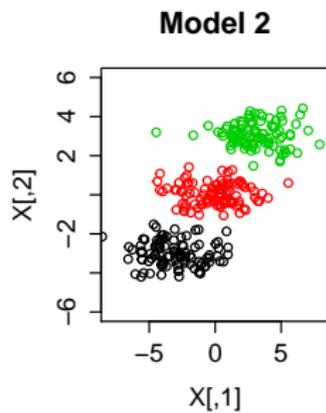
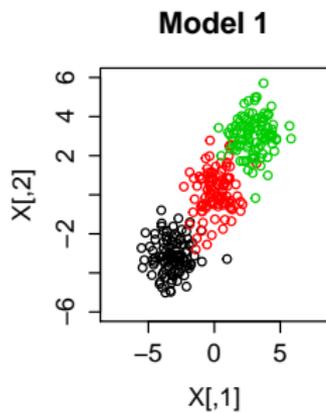
donde  $d(X_i, \tilde{C}_k) = \min_{X_j \in \tilde{C}_k} \|X_i - X_j\|$ .

# Simulaciones Datos Multivariados

Trabajamos sobre doce diferentes escenarios basados en una propuesta de Witten y Tibshirani (2010) y extendida por Kondo et al (2016). En todos los casos el modelo tiene tres grupos con 100 observaciones cada uno.

- ▶ **Modelo 1:** Los datos son generados a partir de una  $N(\mu_i, \Sigma)$ , for  $i = 1, 2, 3$ , con centros  $(-3, -3, 0)$ ,  $(0, 0, 0)$ ,  $(3, 3, 0)$ , y matriz de covarianza igual a la identidad.
- ▶ **Modelo 2:**  $N(\mu_i, \Sigma)$ , for  $i = 1, 2, 3$ , con centros  $(-3, -3, 0)$ ,  $(0, 0, 0)$ ,  $(3, 3, 0)$ , y matriz de covarianza  $\Sigma = \text{diag}(3, 0.25, 0)$ .
- ▶ **Modelo 3:** igual que el **Modelo 1** con dos variables más de ruido con distribution  $N(0, 1)$ .
- ▶ **Modelo 4:** igual que el **Modelo 2** con dos variables más de ruido con distribution  $N(0, 1)$ .

# Simulaciones Datos Multivariados



# Simulaciones Datos Multivariados

## Modelos con contaminación

- ▶ Contaminación Tipo 1: cinco outliers reemplazando la primera coordenada por una variable  $U[25, 25.01]$ , contaminación en la variable informativa.
- ▶ Contaminación Tipo 2: cinco outliers reemplazando la última coordenada por una variable  $U[25, 25.01]$ , contaminación en la variable de ruido.

# Simulaciones Datos Multivariados

## Modelos con contaminación

- ▶ Contaminación Tipo 1: cinco outliers reemplazando la primera coordenada por una variable  $U[25, 25.01]$ , contaminación en la variable informativa.
- ▶ Contaminación Tipo 2: cinco outliers reemplazando la última coordenada por una variable  $U[25, 25.01]$ , contaminación en la variable de ruido.
- ▶ **Modelos 5-8:** Igual que **Modelos 1-4**, con contaminación Tipo 1.

# Simulaciones Datos Multivariados

## Modelos con contaminación

- ▶ Contaminación Tipo 1: cinco outliers reemplazando la primera coordenada por una variable  $U[25, 25.01]$ , contaminación en la variable informativa.
- ▶ Contaminación Tipo 2: cinco outliers reemplazando la última coordenada por una variable  $U[25, 25.01]$ , contaminación en la variable de ruido.
- ▶ **Modelos 5-8:** Igual que **Modelos 1-4**, con contaminación Tipo 1.
- ▶ **Modelos 9-12:** Igual que **Modelos 1-4**, con contaminación Tipo 2.

# Simulaciones Datos Multivariados

Métodos usados en la comparación:

- ▶  $k$ -means.
- ▶ Sparse  $k$ -means (SKM).
- ▶ Sparse  $k$ -means Robusto (RSKM).
- ▶ MCLUST.

Realizamos  $M = 100$  replicas para cada modelo. Para aplicar LDC, fijamos los parámetros de la profundidad local  $\beta = 0.2$ ,  $\alpha = 0.4$  y  $N = 50$  proyecciones al azar.

# Simulaciones Datos Multivariados

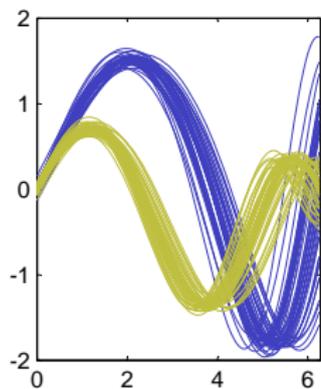
## Resultados

Proporción media de bien clasificados.

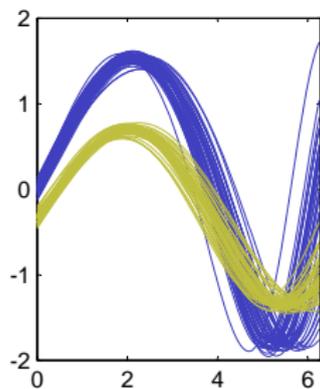
Modelo	<i>k</i> -means	LDC	SKM	RSKM	MCLUST
1	0.98	0.96	0.98	0.98	0.98
2	0.87	0.91	0.80	0.86	0.99
3	0.98	0.96	0.98	0.98	0.98
4	0.87	0.89	0.81	0.86	0.99
5	0.66	0.95	0.71	0.96	0.65
6	0.65	0.89	0.62	0.84	0.66
7	0.67	0.94	0.70	0.96	0.65
8	0.65	0.88	0.62	0.84	0.66
9	0.65	0.95	0.70	0.98	0.65
10	0.65	0.91	0.64	0.86	0.65
11	0.65	0.94	0.67	0.97	0.65
12	0.65	0.91	0.64	0.86	0.65

# Simulaciones Datos Funcionales

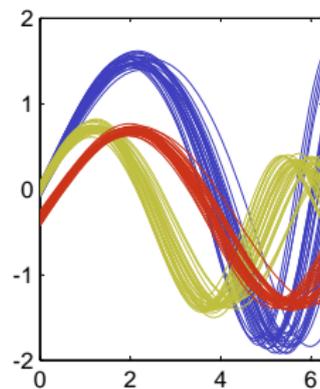
## Caso curvas con warping



Case A



Case B



Case C

# Simulaciones Datos Funcionales

*Modelo A:* Dos clusters con  $n/2$  funciones generadas de la siguiente forma,

$$\begin{aligned} X_i(t) &= (1 + \epsilon_{1i}) \sin(\epsilon_{3i} + \epsilon_{4i}t) + \\ &\quad (1 + \epsilon_{2i}) \sin\left(\frac{(\epsilon_{3i} + \epsilon_{4i}t)^2}{2\pi}\right) \end{aligned} \quad (2)$$

for  $i = 1, \dots, n/2$ ,

$$\begin{aligned} X_i(t) &= (1 + \epsilon_{1i}) \sin(\epsilon_{3i} + \epsilon_{4i}t) - \\ &\quad (1 + \epsilon_{2i}) \sin\left(\frac{(\epsilon_{3i} + \epsilon_{4i}t)^2}{2\pi}\right) \end{aligned} \quad (3)$$

for  $i = n/2 + 1, \dots, n$ .

## Simulaciones Datos Funcionales

*Modelo B:* Dos clusters con  $n/2$  funciones generadas en el primer grupo siguiendo (2) y en el segundo grupo,

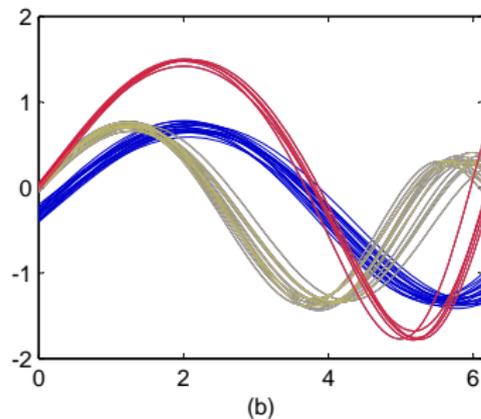
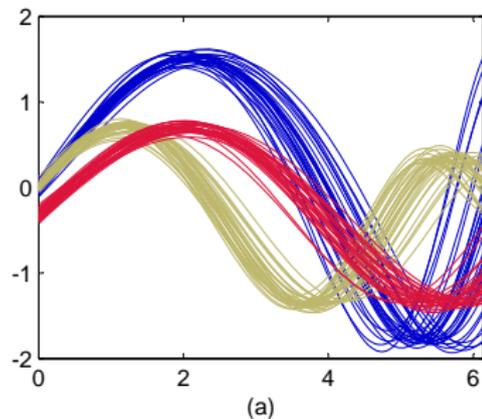
$$X_i(t) = (1 + \epsilon_{1i}) \sin \left( \epsilon_{3i} + \epsilon_{4i} \left( -\frac{1}{3} + \frac{3}{4}t \right) \right) - (1 + \epsilon_{2i}) \sin \left( \frac{\left( \epsilon_{3i} + \epsilon_{4i} \left( -\frac{1}{3} + \frac{3}{4}t \right) \right)^2}{2\pi} \right) \quad (4)$$

for  $i = n/2 + 1, \dots, n$ .

*Modelo C:* Tres clusters de  $n/3$  funciones, las del primer grupo siguiendo la distribución dada (2), el segundo siguiendo la distribución dada en (3) y el tercero siguiendo la distribución dada (4).

Simulamos 200 data sets de tamaño  $n = 90$  para cada modelo. Los errores  $\epsilon_{1i}, \dots, \epsilon_{4i}$  son independientes con distribución  $N(0, 0.025)$ .

# Simulaciones Datos Funcionales



(a) Observaciones del Modelo C. (b) Región núcleo de cada grupo

# Simulaciones Datos Funcionales

Proporción media de bien clasificados.

<b>Método</b>	<b>Model A</b>	<b>Model B</b>	<b>Model C</b>
funclust	75.62	73.57	48.48
kmeans.fd	67	76.77	68.45
funHDDC	85.81	47.39	37.86
LDC	<b>99.98</b>	<b>99.94</b>	<b>99.07</b>

# Simulaciones Datos Funcionales

## Caso Instant wise sampling errors

Simulamos 200 data sets con cuatro clusters con 150 funciones cada uno. Generados con un modelo similar al introducido por Serban y Wasserman (2005).

$$X_{ij}(t) = f_j(t) + \epsilon_i(t), \quad (5)$$

for  $t \in [0, 1], i = 1, \dots, 150$  and  $j = 1, \dots, 4,$

donde

$$f_1(t) = \min \left( \frac{2 - 5t}{2}, \left( \frac{2 - 5t^2}{2} \sin \left( \frac{5\pi t}{2} \right) \right) \right),$$

$$f_2(t) = -f_1(t), \quad f_3(t) = \cos(2\pi t) \quad \text{and} \quad f_4(t) = -f_4(t).$$

# Simulaciones Datos Funcionales

En (5) los autores consideran errores independientes, mientras que nosotros consideramos errores correlacionados provenientes de un proceso Gaussiano.

Las funciones  $\epsilon(t)$  tienen distribución normal con media 0.4, desvío estándar 0.9 y estructura de covarianza,

$$\rho(s, t) = 0.3 \exp\left(-\frac{(s-t)^2}{0.3}\right), \quad \text{for } s, t \in [0, 1].$$

# Simulaciones Datos Funcionales

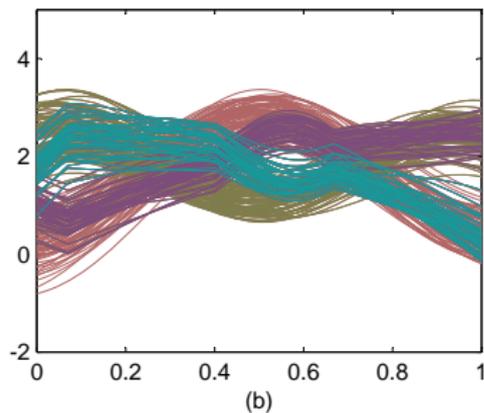
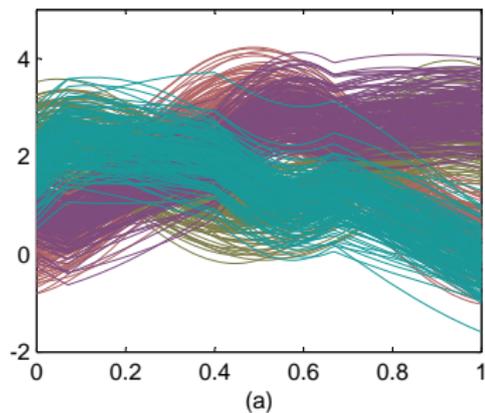


Figure : (a) Un ejemplo generado a partir de (5). (b) Región núcleo del mismo ejemplo.

# Simulaciones Datos Funcionales

## Resultados

Método	CCR media
funclust	40.05
kmeans.fd	60.12
funHDDC	63.21
LDC	99.39

**Ventaja:** LDC con respecto a otros métodos tiene una performance computacional muy buena debido al bajo costo que tiene hacer proyecciones unidimensionales.

-  Agostinelli, C. (2016). "Local half-region depth for functional data." *Preprint* <http://arxiv.org/pdf/1512.04395.pdf>
-  Agostinelli, C. and Romanazzi, M. (2011). "Local Depth." *Journal of Statistical Planning and Inference*, **141**, 817-830.
-  C. Bouveyron and J. Jacques, "Model-based clustering of time series in group-specific functional subspaces". *Advances in Data Analysis and Classification*, 5, 281–300, 2011.  
DOI:10.1007/s11634-011-0095-6.
-  Cuevas, A. and Fraiman, R. (2009). "On depth measures and dual statistics. A methodology for dealing with general data." *Journal of Multivariate Analysis*, **100**(4), 753-766.
-  M. Febrero-Bande, M. Oviedo de la Fuente, "Statistical Computing in Functional Data Analysis: The R Package *fda.usc*". *Journal of Statistical Software*, 51(4), 1-28 ,2012.  
DOI: 10.18637/jss.v051.i04

-  Fraley C. and Raftery, A.E. (2002). “Model-based clustering, discriminant analysis, and density estimation.” *Journal of the American Statistical Association*, **97**, 6116631.
-  Fraley C. and Raftery, A.E. (2009). “MCLUST Version 3 for R: Normal Mixture Modeling and Model-based Clustering”, Technical Report No. 504, Department of Statistics, University of Washington.
-  J. Jacques and C. Preda, “Model-based clustering of functional data.”, *Computational Statistics and Data Analysis*, 71, 92–106, 2014. DOI:10.1016/j.csda.2012.12.004
-  J. Jacques and C. Preda, “Functional Data Clustering: A Survey.”, *Advances in Data Analysis and Classification*, 8 (3), 231–255, 2014. DOI : 10.1007/s11634-013-0158-y.
-  Kondo, Y., Salibian-Barrera, M. and Zamar, R. H. (2016). “A robust and sparse K-means clustering algorithm.” *Journal of Statistical Software*, 72(5).

-  Paindavaine, D. and Van Bever (2013). “From depth to local depth: A focus in centrality.” *Journal of the American Statistical Association*, **108**(503), 1105-1119.
-  L.M. Sangalli, P. Secchi, S. Vantini and V. Vitelli, “*k*-means alignment for curve clustering.” *Computational Statistics and Data Analysis*, 54, 1219–1233, 2010.  
DOI:10.1016/j.csda.2009.12.008
-  R. Tibshirani, G. Walther and T. Hastie, “Estimating the number of data clusters via the gap statistic”. *Journal of the Royal Statistical Society, B*, 63, 411–423, 2001. DOI: 10.1111/1467-9868.00293.
-  N. Serban and L. Wasserman, “CATS: Cluster Analysis by Transformation and Smoothing”. *Journal of the American Statistical Association*, 100, 990–999, 2005. DOI 10.1198/016214504000001574



Witten, D. and Tibshirani, R. (2010). "A framework for feature selection in clustering." *Journal of the American Statistical Association*, **105**(490), 713-726.

MUCHAS GRACIAS!