

TESIS DE DOCTORADO

# Técnicas de teoría geométrica de la medida en estimación de conjuntos

Alejandro Cholaquidis

Orientadores: Antonio Cuevas, Ricardo Fraiman.

Doctorado en Matemática  
Facultad de Ciencias  
Universidad de la República  
Montevideo Uruguay

*Dedico esta tesis a los familiares y amigos que me acompañaron en estos años*

# Agradecimientos

Quiero agradecer, sabiéndome injusto con la enorme lista de personas que debería figurar aquí, a tres personas que han sido pilares fundamentales en mi formación como matemático. En diferentes aspectos y momentos de mi vida, las tres hicieron aportes que viajarán conmigo y dejarán su huella a lo largo de toda mi existencia. Espero que esa huella, que hoy dejo plasmada en esta tesis, quede en el futuro, (como dijo Benedetti) “como un ladrillo en otra casa”.<sup>0</sup>

En primer lugar agradecer a Ricardo, quien fue mi tutor desde la licenciatura hasta la actualidad. Ha sido un privilegio enorme poder trabajar tantos años con él y dejarse contagiar de su optimismo y buen humor. Con su capacidad de motivar y grandes ideas aprendí, desde el principio, que hacer matemática tiene que ser (y es) divertido y desafiante. Que hay que buscar problemas interesantes y abordarlos desde una perspectiva original. Que ser tutor es hacer sentir cómodo al estudiante de poder plantear sus ideas y valorar sus resultados, aún, como en mi caso, muy magros muchas veces.

En segundo lugar quiero agradecer a Antonio, quien tuve el gusto de conocer en el octavo ERPEN, en Buenos Aires. Además de haber sentido siempre que mis modestas ideas iniciales eran bien recibidas, logró generar en mi un interés y gusto por aspectos “artesanales” (a decir de él mismo) de la matemática. Recuerdo una frase que me quedó grabada y ha sido de motivación a la hora de enfrentarme a la ardua tarea de corregir erratas y revisar trabajos: “El trabajo del matemático es como el del artesano”. Me llevó cuatro años entender bien lo que aquella revelación significaba. Con el tiempo descubrí que Antonio era además una persona extremadamente amable y que se interesa muchísimo por sus estudiantes. Siempre, a pesar de la distancia, con buena disposición para hablar o leer los (a veces demasiados) mails que le envié.

En tercer lugar quiero recordar y agradecer a Mario Wschebor. Es difícil, si no imposible, decir de Mario algo que no haya sido dicho ya. Y no es sino el eco que aún resuena, de todo lo que ha hecho, de la gente que ha formado en diferentes aspectos, de sus trabajos como investigador, de sus aportes esenciales en temas universitarios, la imagen viva y más nítida de su personalidad y legado.

---

<sup>0</sup>Esta tesis contó con la financiación de la Agencia Nacional de Investigación e Innovación a través de una beca de Doctorado.

# Prólogo

En términos muy generales, puede decirse que esta tesis es un estudio de la incorporación a la metodología estadística de algunas ideas básicas de geometría euclídea. Más concretamente, con la excepción parcial del último capítulo, la tesis gira en torno a ciertos conceptos de geometría convexa examinados desde el punto de vista estadístico.

La noción de convexidad es, sin duda alguna, una de las ideas centrales en matemáticas, comparable por su importancia a conceptos como continuidad, diferenciabilidad u orden. La restricción de convexidad aplicada a conjuntos del espacio euclídeo  $\mathbb{R}^d$  es muy intuitiva y da lugar, en varias ramas de la matemática, a resultados clásicos de gran belleza y utilidad. Su único inconveniente es el hecho de que para muchas aplicaciones resulta demasiado restrictiva. Sin embargo, de manera análoga a lo que sucede con otros conceptos matemáticos centrales, la convexidad extiende sus dominios más allá de sí misma, en el sentido de que uno puede renunciar a la restricción estricta de convexidad manteniendo alguna de sus propiedades importantes y trabajando con ella. Este es exactamente el enfoque que se ha desarrollado en este trabajo. En las introducciones de los capítulos se pueden encontrar resúmenes más detallados del contenido y aportaciones de cada uno de ellos. A continuación esbozamos una breve panorámica general del trabajo desde el punto de vista de las propiedades geométricas que se han utilizado.

*La propiedad de volumen polinómico.*

Así, el capítulo 2 toma como punto de partida la propiedad de “volumen polinómico” de los conjuntos convexos. Esta propiedad, probada por Steiner a mediados del siglo XIX, establece que el volumen de la dilatación  $B(S, r)$  de un conjunto convexo compacto  $S \subset \mathbb{R}^d$ , considerado como una función del factor de dilatación  $r$ , es un polinomio de grado  $d$ , cuyos coeficientes tienen una interpretación muy directa en términos de las propiedades geométricas (medida de la frontera, curvatura,...) del conjunto. A lo largo del siglo XX se han obtenido algunas generalizaciones de este resultado (por ejemplo, Federer (1959)) que están estrechamente vinculadas con avances decisivos en la teoría geométrica de la medida. En el capítulo 2 “aislamos” esta propiedad de volumen polinómico, considerando la clase de conjuntos (mucho más amplia que la clase de los conjuntos convexos) que tienen volumen polinómico, al menos en un cierto intervalo  $[0, R)$ . Nuestra conclusión principal es que esta clase de conjuntos (que estudiamos en los casos  $d = 2$  y  $d = 3$ ) es lo bastante sencilla como para permitir una estimación paramétrica (con procedimientos estándar de máxima verosimilitud y método de momentos) de algunos funcionales importantes como, por ejemplo, la medida de su frontera.

*La propiedad de hiperplano soporte.*

Otra propiedad esencial de los conjuntos convexos cerrados (que de hecho los caracteriza cuando son de interior no vacío) es la existencia de un hiperplano soporte en todo punto de su frontera. El capítulo 3 es el resultado de “aislar” esta propiedad en una versión generalizada en la que el hiperplano (o más bien su correspondiente semiespacio exterior) es reemplazado por un cono finito de altura  $h$  y amplitud fija  $\rho$  (en el caso convexo tendríamos  $\rho = \pi$  y  $h = \infty$ ). Denominamos

---

conjuntos  $\rho, h$ -cono convexos a los que satisfacen esta condición de “cono exterior”, reminiscente de la propiedad convexa de hiperplano soporte. En realidad esta propiedad de cono-convexidad, en una versión algo más general, ya fue considerada por Poincaré y Zaremba, a principios del siglo XX, en relación con la solución del clásico problema de Dirichlet. Nuestro uso aquí es completamente diferente, mucho más próximo a la geometría estocástica, en la que se ha estudiado durante muchos años las propiedades del cierre convexo de una muestra finita de puntos en  $\mathbb{R}^d$ . En nuestro caso, hemos definido de manera natural el cierre cono-convexo y hemos analizado sus propiedades asintóticas y su implementación práctica como estimador del soporte  $S$  del cual se suponen extraídos los puntos muestrales. Como una interesante propiedad adicional de carácter probabilístico, la clase de los conjuntos  $\rho, h$ -cono convexos resulta ser una clase de Glivenko-Cantelli.

*La distancia de Hausdorff entre conjuntos considerada como distancia funcional.*

El capítulo 4 está algo más alejado de las nociones generalizadas de convexidad pero, aún así, incide en el uso estadístico de una herramienta geométrica, que tiene especial relieve en la geometría convexa: la llamada distancia de Hausdorff (o de Hausdorff-Pompeiu)  $d_H$  entre conjuntos compactos de  $\mathbb{R}^d$ . Esta distancia ha sido muy utilizada en teoría de fractales, análisis de imágenes, conjuntos aleatorios y teoría de la aproximación. Nuestra propuesta sobre el uso de  $d_H$  está más cerca a esta última aplicación. En el contexto de regresión o clasificación con datos funcionales, definimos una métrica “visual” entre funciones basada en la distancia de Hausdorff entre los correspondientes hipografos. Estudiamos las propiedades de este espacio métrico demostrando, en particular, que es completo, separable y localmente compacto. Discutimos las implicaciones del uso de este espacio en problemas de clasificación funcional supervisada. El estudio de algunos problemas con datos de espectrogramas (de masas o de resonancia magnética) muestra que la distancia propuesta es especialmente adecuada (cuando se compara con las distancias tradicionales entre funciones) en problemas de clasificación.

**Publicaciones basadas en el contenido de esta tesis:**

1. Berrendero, J.R., Cholaquidis, A., Cuevas, A. y Fraiman, R. (2013). A geometrically motivated parametric model in manifold estimation. Aparecerá en *Statistics*.  
DOI:10.1080/02331888.2013.800264. Published online: 30 May 2013.  
[Corresponde básicamente al contenido del capítulo 2 de la tesis].
2. Cholaquidis, A., Cuevas, A. y Fraiman, R. (2014). On Poincaré cone property. Aparecerá en *The Annals of Statistics*.  
[Corresponde básicamente al contenido del capítulo 3 de la tesis].

# Índice general

<b>1. Introducción</b>	<b>8</b>
1.1. Sobre la estimación de conjuntos . . . . .	8
1.1.1. Conceptos básicos y resultados previos de la estimación de conjuntos . . . . .	8
1.2. Aspectos computacionales . . . . .	15
1.3. Ideas geométricas en el estudio de los datos funcionales . . . . .	15
<b>2. Estimación paramétrica</b>	<b>17</b>
2.1. Conjuntos con volumen polinómico . . . . .	18
2.1.1. Alcance positivo y fórmula de Federer . . . . .	19
2.1.2. El caso general . . . . .	20
2.2. Estimación paramétrica: Caso tridimensional . . . . .	21
2.2.1. Introducción . . . . .	21
2.2.2. Método de los Momentos . . . . .	21
2.2.3. Aspectos prácticos de los estimadores. . . . .	24
2.3. Gráficos y Simulaciones en $\mathbb{R}^3$ . . . . .	26
<b>3. Aplicaciones de la propiedad de cono convexidad a la estimación de conjuntos</b>	<b>30</b>
3.1. Propiedad de Poincaré y conjuntos cono-convexos. Definiciones y propiedades . . . . .	31
3.1.1. Antecedentes históricos . . . . .	31
3.1.2. Notación y definiciones básicas . . . . .	32
3.1.3. Propiedades de convergencia . . . . .	36
3.2. Los conjuntos cono convexos son una clase de Glivenko-Cantelli . . . . .	37
3.3. Estimación de conjuntos cono convexos . . . . .	39
3.3.1. Tasas de convergencia para la distancia de Hausdorff . . . . .	39
3.3.2. Tasas de convergencia en la distancia en medida . . . . .	40
3.3.3. Tasas de convergencia para la envolvente por complemento . . . . .	42
3.4. Estimación no paramétrica de la medida de la frontera de conjuntos cono-convexos . . . . .	45
3.5. Inferencia sobre los parámetros $r$ y $\rho$ . . . . .	47
3.5.1. Espacios maximales en conjuntos cono convexos . . . . .	47
3.5.2. Contraste de hipótesis sobre $\rho$ . . . . .	50
3.6. Contraste de hipótesis de convexidad . . . . .	54
3.6.1. Algoritmo y simulaciones . . . . .	56
3.7. Un estimador consistente del parámetro . . . . .	59
3.8. Aspectos computacionales de la propiedad de Poincaré . . . . .	60
3.8.1. Algoritmo en $\mathbb{R}^d$ . . . . .	60
3.8.2. Aplicación a diferentes datos. . . . .	61

---

<b>4. Ideas geométricas en el análisis de datos funcionales y regresión no paramétrica</b>	<b>66</b>
4.1. Definición y estudio de las propiedades de un nuevo espacio métrico de funciones . . .	68
4.1.1. Distancias “visuales” entre funciones: Un poco de historia . . . . .	69
4.1.2. Una métrica funcional basada en la distancia de Hausdorff . . . . .	70
4.2. Un algoritmo para el cálculo de la distancia . . . . .	74
4.3. Aplicaciones prácticas a problemas de clasificación binaria con espectros (de masas o de resonancia magnética). . . . .	76
4.4. Análisis de la consistencia: condición de Besicovitch . . . . .	79
4.5. Aplicaciones a problemas de regresión no paramétrica (finito-dimensional) con restricciones geométricas de forma. . . . .	83
4.5.1. Algoritmo de construcción de funciones $\rho$ como convexas a partir de una muestra.	85
<b>A. Apéndice</b>	<b>89</b>
A.1. Cálculo del volumen de algunos conjuntos . . . . .	89
A.2. Pseudo código del algoritmo del Capítulo 4 . . . . .	90
<b>Bibliografía</b>	<b>93</b>

# Capítulo 1

## Introducción

En este capítulo introduciremos la notación que será utilizada en la tesis y presentaremos un breve repaso de las principales definiciones y resultados necesarios para la lectura de los siguientes capítulos.

### 1.1. Sobre la estimación de conjuntos

El término *estimación de conjuntos* refiere al problema estadístico de reconstrucción de un conjunto desconocido, que usualmente denotaremos  $S$ , a partir de una muestra de vectores aleatorios  $X_1, \dots, X_n$  cuya distribución está relacionada con  $S$  de algún modo. Dicho conjunto puede ser, por ejemplo, el soporte de una distribución, el borde de la misma, o un conjunto de nivel de una densidad con soporte  $S$ . Asimismo, como veremos en el capítulo dos de la presente tesis, la estimación de conjuntos aborda problemas como la estimación del volumen del borde del mismo, su característica de Euler, la integral de la curvatura media, etc. El estimador natural, en el caso más simple en el que  $S$  es convexo, es la envolvente convexa de los puntos. Dicho estimador fue el primero estudiado, para subconjuntos del plano, por Renyi y Sulanke, (ver Rényi y Sulanke (1963, 1964)). Resultados recientes referentes a la estimación del perímetro de un conjunto convexo pueden encontrarse en Bräker y Hsing (1998). En 1980 Devroye y Wise, interesados en la detección del comportamiento anormal de un sistema, propusieron un estimador de  $S$  que es consistente para cualquier conjunto (ver Devroye y Wise (1980)). Si bien dicho estimador es general no es posible obtener, por tal motivo, tasas de convergencia. En este capítulo veremos diferentes restricciones de forma que permiten debilitar la hipótesis de convexidad (es decir considerar familias más grandes) pero aún así obtener tasas de convergencia para los estimadores.

#### 1.1.1. Conceptos básicos y resultados previos de la estimación de conjuntos

A efectos de facilitar la lectura de los siguientes capítulos vamos a introducir en la presente sección las definiciones y resultados necesarios. Si bien las definiciones que daremos a continuación son para espacios métricos, en general consideraremos subconjuntos del espacio  $\mathbb{R}^d$ , dotado con el producto interno usual y norma euclídea, que denotaremos  $\langle \cdot, \cdot \rangle$  y  $\|\cdot\|$  respectivamente. Dado  $S \subset \mathbb{R}^d$ ,  $\bar{S}$  denota la clausura de  $S$  mientras que  $\text{int}(S)$ ,  $S^c$  y  $\partial S$  denotan el conjunto de puntos interiores, el complemento, y el borde respectivamente. En general denotaremos  $B(x, r) = \{y \in \mathbb{R}^d : \|y - x\| < r\}$  y  $S_1^d = \{x \in \mathbb{R}^{d+1} : \|x\| = 1\}$ .



### Distancia entre conjuntos

Para estimar un conjunto debemos tener alguna noción de distancia que nos permita evaluar la proximidad entre conjuntos. Dicha noción dependerá de las particularidades que estemos interesados en estudiar. Utilizaremos mayormente la distancia en medida, y la distancia de Hausdorff, ya que son las que frecuentemente se estudian en la teoría de la estimación de conjuntos y que se complementan en cuanto a la información que nos aportan respecto a la proximidad de los conjuntos.

*Distancia de Hausdorff.*

**Definición 1.1.** Sea  $(E, d)$  un espacio métrico, sean  $A$  y  $C$  subconjuntos de  $E$ , la “distancia” de Hausdorff entre  $A$  y  $C$  se define como:

$$d_H(A, C) = \max \left\{ \sup_{a \in A} d(a, C), \sup_{c \in C} d(c, A) \right\}, \quad (1.1)$$

siendo

$$d(x, C) = \inf \{ d(x, c) : c \in C \},$$

donde convenimos en que  $\inf \emptyset = \infty$ .

Observemos que la distancia de Hausdorff mide la proximidad entre los puntos de  $A$  y los de  $C$  por lo tanto basta modificar la posición de un punto para que la misma se vea afectada. Es claro que (1.1) no define una distancia en el sentido usual, ya que si alguno de los conjuntos  $A$  o  $C$  no son acotados, puede tomar el valor infinito. Más aún, puede valer 0 (por ejemplo:  $d_H(S, \bar{S})$ ) sin que ambos sean iguales. No obstante, si nos restringimos a subconjuntos compactos de el espacio métrico  $E$ ,  $d_H$  está bien definida y es una distancia en el sentido usual.

**Definición 1.2.** Sea  $A$  un subconjunto no vacío, de un espacio métrico  $(E, d)$ . El conjunto paralelo abierto de radio  $\varepsilon > 0$  de  $A$ , que denotaremos:  $\overset{\circ}{B}(A, \varepsilon)$ , se define como:

$$\overset{\circ}{B}(A, \varepsilon) = \{ x \in E : d(x, A) < \varepsilon \}.$$

Análogamente definimos:

**Definición 1.3.** Sea  $A$  un subconjunto no vacío de un espacio métrico  $(E, d)$ . El conjunto paralelo cerrado de radio  $\varepsilon > 0$  de  $A$ , que denotaremos:  $B(A, \varepsilon)$ , se define como:

$$B(A, \varepsilon) = \{ x \in E : d(x, A) \leq \varepsilon \}.$$

Si bien la notación empleada anteriormente es usual en la literatura, en ambas definiciones el conjunto  $A$  no es considerado como un punto de un espacio métrico, es conveniente aclarar que  $B(A, \varepsilon)$  no se corresponde con la notación usual de bola con centro en  $A$  y radio  $\varepsilon$ .

Teniendo en cuenta estas definiciones es sencillo ver que, para  $A$  y  $C$  no vacíos, la distancia de Hausdorff 1.1 es equivalente a:

$$d_H(A, C) = \inf \{ \varepsilon > 0 : A \subset \overset{\circ}{B}(C, \varepsilon) \text{ y } C \subset \overset{\circ}{B}(A, \varepsilon) \}.$$

**Definición 1.4.** Sean  $A$  y  $C$  subconjuntos de  $\mathbb{R}^d$ . La suma de Minkowski, que se denota usualmente con el símbolo  $\oplus$ , se define como:

$$A \oplus C = \{ a + c : a \in A, c \in C \},$$

y la resta, que se denota con el símbolo  $\ominus$ , como:

$$A \ominus C = \{ x : \{x\} \oplus C \subset A \}.$$

Para  $\lambda > 0$  definimos

$$\lambda C = \{ \lambda c : c \in C \},$$

por lo tanto, si denotamos  $B = B(0, 1)$  entonces

$$B(A, \varepsilon) = A \oplus \varepsilon \overline{B}.$$

Es inmediato verificar que, para  $A$  y  $C$  subconjuntos compactos, la distancia de Hausdorff es equivalente a:

$$d_H(A, C) = \inf \{ \varepsilon > 0 : A \subset C \oplus \varepsilon B \text{ y } C \subset A \oplus \varepsilon B \}.$$

*Distancia en medida.*

En general utilizaremos el espacio de medida  $(\mathbb{R}^d, \mathcal{B}, \mu)$  donde  $\mathcal{B}$  denota la sigma álgebra de Borel en  $\mathbb{R}^d$  y  $\mu$  la medida de Lebesgue. En algunos casos nos será de interés tomar otra medida, por ejemplo la distribución  $P_X$  de una variable aleatoria  $X$ , con soporte en cierto conjunto  $S \subset \mathbb{R}^d$ .

**Definición 1.5.** Consideremos el espacio de medida  $(\mathbb{R}^d, \mathcal{B}, \nu)$ . Sean  $A$  y  $C$  pertenecientes a  $\mathcal{B}$  con medida  $\nu$  finita. La distancia en medida entre  $A$  y  $C$  se define como

$$d_\nu(A, C) = \nu(A \Delta C),$$

donde  $A \Delta C$  denota la diferencia simétrica entre  $A$  y  $C$ , esto es:

$$A \Delta C = (A \setminus C) \cup (C \setminus A).$$

Es fácil ver que el hecho de que dos conjuntos estén próximos en distancia en medida no implica que lo estén en distancia de Hausdorff, y recíprocamente, que lo estén en distancia de Hausdorff tampoco implica que lo estén en medida. Aunque  $d_\nu$  no es obviamente una distancia en el espacio de los conjuntos compactos, puede convertirse en distancia si se identifican (al cocientar por la relación de equivalencia:  $A \sim B$  si  $\nu(A) = \nu(B)$ ) los compactos que difieren en un conjunto de medida cero.

Teniendo en cuenta las métricas antes definidas, podemos hablar de tasas de convergencia de un estimador  $S_n$  de un conjunto  $S$ . Se puede demostrar por ejemplo que si  $S$  es convexo  $d_H(S_n, S) = \mathcal{O}((\log(n)/n)^{1/d})$  (ver Dümbgen y Walther (1996)), donde  $S_n$  denota la envolvente convexa de los puntos. Dicho resultado se puede mejorar imponiendo restricciones de regularidad en el borde de  $S$ , y obtener así  $d_H(S_n, S) = \mathcal{O}((\log(n)/n)^{2/(d+1)})$ . Si queremos obtener velocidades de convergencia, es necesario imponer restricciones de forma sobre los conjuntos. Por otro lado, a los efectos prácticos la convexidad puede resultar demasiado restrictiva. En la siguiente sección veremos algunas restricciones de forma usuales en la literatura de estimación de conjuntos.

## Geometría de los conjuntos

En esta sección presentaremos algunas definiciones que, inspiradas en la idea de conjunto convexo, buscan generalizarlo. Se puede ver que un conjunto  $S$ , cerrado, con interior no vacío es convexo si para cada  $x \in \partial S$  podemos tomar un punto  $y$  (que depende de  $x$ ) tal que el producto interno  $\langle y - x, s \rangle \leq 0$  para todo  $s \in S$ . Es decir, el conjunto  $S$  se puede obtener como la intersección de los complementos de los semiespacios incluidos en  $S^c$ .

*Conjuntos  $r$ -convexos y condición de rodamiento libre.*

La definición de conjunto  $r$ -convexo está motivada por la idea antes expuesta, cambiando hiperplanos por bolas de un radio fijo  $r$ . Algunos resultados referentes a la estimación de dichos conjuntos pueden encontrarse en Rodríguez-Casal (2002, 2006), Pateiro-López (2008) y en Walther (1999, 1997).

**Definición 1.6.** Diremos que un conjunto compacto  $S \subset \mathbb{R}^d$  es  $r$ -convexo ( $r > 0$ ) si  $S = C_r(S)$ , siendo

$$C_r(S) = \bigcap_{\{B(x,r): B(x,r) \cap S = \emptyset\}} B(x,r)^c. \quad (1.2)$$

Se puede demostrar (ver Walther (1999)) que:

$$C_r(S) = \bigcap_{\{S \subset B, B \text{ es } r\text{-convexo}\}} B.$$

**Definición 1.7.** Sea  $S \subset \mathbb{R}^d$  compacto, diremos que una bola de radio  $r$  rueda libremente por  $S^c$  si para todo  $x \in \partial S$  existe una bola  $B_x$  de radio  $r$  tal que  $x \in \partial B_x$  y  $B_x \cap S = \emptyset$ .

De forma análoga se define la condición de rodamiento libre por  $S$ .

En Cuevas y otros (2012) se prueba que la clase de los subconjuntos de  $\mathbb{R}^d$  para los cuales rueda por su complemento una bola de radio  $r$  incluye a la clase de los subconjuntos  $r$ -convexos. Observemos que, si  $S$  tiene interior no vacío, en el caso límite en que  $r = \infty$  obtenemos conjuntos convexos. Se puede ver que, utilizando la notación anterior, un conjunto  $S \subset \mathbb{R}^d$  es  $r$ -convexo si  $S = (S \oplus rB) \ominus rB$ . Una propiedad interesante de los conjuntos  $r$ -convexos, (que no cumplirán los conjuntos como convexos que definiremos más adelante) es que la envolvente  $r$ -convexa de  $S$  (es decir el menor  $r$ -convexo que lo contiene) coincide con la intersección de los complementos de las bolas abiertas que no cortan al conjunto. Si además de ser  $S$   $r$ -convexo, lo es  $S^c$ , el orden de convergencia (en distancia en medida y Hausdorff) de la envolvente  $r$ -convexa de la muestra es  $\mathcal{O}((\log(n)/n)^{2/(d+1)})$ . Es decir, obtenemos el mismo orden que para conjuntos convexos cuyo borde cumple condiciones adicionales de regularidad. Lo llamativo de dicho resultado es que la clase de los conjuntos  $r$ -convexos es *mucho* más amplia que la de los conjuntos convexos.

**Definición 1.8.** Sea  $S \subset \mathbb{R}^d$  tal que  $S \in \mathcal{B}$ , diremos que  $S$  es estándar respecto de una medida de Borel  $\nu$  si existe  $\lambda > 0$  y  $\delta > 0$  tal que:

$$\nu(B(x, \varepsilon) \cap S) \geq \delta \mu(B(x, \varepsilon)) \quad \text{para todo } x \in S, 0 < \varepsilon \leq \lambda,$$

donde  $\mu$  denota la medida de Lebesgue en  $\mathbb{R}^d$ .

**Definición 1.9.** Definimos la función  $\lambda \mapsto \psi_\lambda(S)$ , para  $\lambda > 0$

$$\psi_\lambda(A) = (S \ominus \lambda \overline{B}) \oplus \lambda \overline{B},$$

$$\psi_{-\lambda}(A) = (A \oplus \lambda \overline{B}) \ominus \lambda \overline{B}.$$

Donde  $B = B(0, 1)$ . La función  $\psi_\lambda$  se llama granulometría de  $A$ .

**Definición 1.10.** Diremos que un conjunto compacto  $A \subset \mathbb{R}^d$  pertenece al modelo regular de Serra si:

$$A = (A \oplus \varepsilon \overline{B}) \ominus \varepsilon \overline{B} = (A \ominus \varepsilon \overline{B}) \oplus \varepsilon \overline{B} \quad \text{para algún } \varepsilon > 0.$$

Un estudio detallado de estas dos últimas definiciones y su aplicación al análisis de imágenes puede encontrarse en Serra (1982) y en Walther (1999). Si bien, como observamos anteriormente, la clase de conjuntos para los cuales rueda una bola de radio  $r$  es más grande, el Teorema siguiente (cuya demostración puede encontrarse en Walther (1999)) nos dice que si dicha condición la verifica además el complemento del conjunto, ambas clases coinciden.

**Teorema 1.11.** Sea  $S \neq \emptyset$  un subconjunto de  $\mathbb{R}^d$ , supongamos que  $S$  es compacto y conexo por caminos, sea  $r_0 > 0$ . Son equivalentes:

- (i)  $\psi_\lambda(S) = S$  para  $\lambda \in (-r_0, r_0]$ ;
- (ii)  $S$  y  $\overline{S^c}$  son  $r_0$ -convexos y  $\text{int}(S) \neq \emptyset$ ;
- (iii) una bola de radio  $r$  rueda libremente por  $S$  y  $\overline{S^c}$  para todo  $0 \leq r \leq r_0$ ;

(iv)  $\partial S$  es una  $d-1$  subvariedad de  $\mathbb{R}^d$  de clase  $C^1$  cuyo vector normal saliente  $n(s)$  en  $s \in \partial S$  satisface la siguiente condición de Lipschitz:

$$\|n(s) - n(t)\| \leq \frac{1}{r_0} \|s - t\| \text{ para todo } s, t \in \partial S.$$

Más aún, para algún  $r_0 > 0$ , las condiciones anteriores son equivalentes a

(v)  $S$  pertenece al modelo regular de Serra.

*Conjuntos con alcance positivo.*

El concepto de alcance de un conjunto fue definido en Federer (1959), aquí simplemente presentaremos su definición y algunas propiedades, puede encontrarse un estudio muy completo del mismo en Thäle (2008), en Colesanti y Manselli (2010) y en Federer (1959).

Siguiendo la notación de Federer (1959), dado un conjunto  $S \subset \mathbb{R}^d$  compacto, denotamos  $\text{Unp}(S)$  al conjunto:

$$\text{Unp}(S) = \{x \in \mathbb{R}^d : \text{existe una única proyección } \xi_S(x) \text{ sobre } S\},$$

es decir,  $\xi_S(x)$  es el punto en  $S$  que minimiza  $d(x, S)$ .

**Definición 1.12.** Sea  $S \subset \mathbb{R}^d$  definimos:

$$\begin{aligned} \text{reach}(S)(x) &= \sup \{r \geq 0 : B(x, r) \subset \text{Unp}(S)\}, \\ \text{reach}(S) &= \inf \{\text{reach}(S)(x) : x \in S\}. \end{aligned}$$

Se puede demostrar (véase Pateiro-López (2008)) que si  $S$  es cerrado no vacío, tal que una bola de radio  $r$  rueda libremente por  $S$  y  $S^c$  entonces el alcance de  $\partial S$  es mayor o igual que  $r$ . Asimismo en Cuevas y otros (2012) se demuestra que si un conjunto tiene alcance  $r$  es  $r$ -convexo; se presenta además un ejemplo de un conjunto  $r$ -convexo que no tiene alcance  $r$ . En dicho trabajo se demuestra que, con la hipótesis adicional de *conectividad local interior* un conjunto  $r$ -convexo tiene alcance positivo, aunque no necesariamente  $r$ . Una conjetura de K. Borsuk no resuelta aún establece que los conjuntos  $r$ -convexos son localmente contractibles (es decir homotópicamente equivalentes a un punto). Recordemos que (ver Observación 4.15 en Federer (1959)) un conjunto con alcance positivo es localmente contractible por lo tanto, si un conjunto  $r$ -convexo tuviese alcance positivo, la conjetura sería cierta.

Es interesante observar que, a diferencia de lo que sucede con los conjuntos  $r$ -convexos, no es posible definir, para cualquier conjunto, el menor conjunto con alcance mayor o igual que un cierto  $r$  (ver Colesanti y Manselli (2010)) que lo contiene, más aún, cuando tal conjunto existe, coincide con la envolvente  $r$ -convexa del mismo.

*Relación entre  $d_H$  y  $d_\mu$ .*

Como dijimos anteriormente no es cierto en general que  $S_n \xrightarrow{d_H} S \Rightarrow S_n \xrightarrow{d_\mu} S$  así como tampoco se verifica, en general, el recíproco. Sin embargo, como veremos a continuación, si agregamos algunas hipótesis adicionales se pueden probar resultados como los anteriores. En Cuevas y otros (2012) se demuestra el siguiente resultado:

**Teorema 1.13.** *Sea  $\{S_n\}$  una sucesión de subconjuntos de  $\mathbb{R}^d$ , compactos no vacíos. Sea  $\nu$  una medida de Borel tal que  $\nu(C) < \infty$  para todo conjunto compacto  $C$ , consideremos  $S$  compacto no vacío tal que  $\nu(\partial S) = 0$ . Si  $d_H(\partial S_n, \partial S) \rightarrow 0$  y  $d_H(S_n, S) \rightarrow 0$  entonces  $d_\nu(S_n, S) \rightarrow 0$ .*

**Teorema 1.14.** *Sea  $\{S_n\}$  una sucesión de subconjuntos de  $\mathbb{R}^d$ , compactos no vacíos. Supongamos que los conjuntos  $S_n$  son estándar respecto de una medida de Borel  $\nu$  (para el mismo  $\lambda$  y  $\delta$ , ver Definición 1.8). Supongamos que  $S$  es estándar respecto de  $\nu$  (para algún  $\lambda'$  y  $\delta'$ ).*

$$\text{Si } d_\nu(S_n, S) \rightarrow 0 \text{ entonces } d_H(S_n, S) \rightarrow 0$$

*Demostración.* Supongamos por reducción al absurdo, que no se cumple la tesis, tenemos dos casos:

1. existe  $\varepsilon > 0$  y  $y_{n_k} \in S_{n_k}$  tal que  $B(y_{n_k}, \varepsilon) \cap S = \emptyset$
2. existe  $\gamma > 0$  y  $x_{n_j} \in S$  tal que  $B(x_{n_j}, \gamma) \cap S_{n_j} = \emptyset$

Veamos primero el caso 1): sea  $\alpha < \min\{\varepsilon, \lambda\}$  como  $B(y_{n_k}, \alpha) \cap S = \emptyset$  tenemos que

$$\nu(S \Delta S_{n_k}) \geq \nu(B(y_{n_k}, \alpha) \cap S_{n_k}) \geq \delta \mu(B(y_{n_k}, \alpha)),$$

donde  $\mu$  denota la medida de Lebesgue, pero esto último contradice que  $S_n \xrightarrow{d_\mu} S$ . De forma análoga se llega a un absurdo suponiendo 2).  $\square$

**Corolario 1.15.** *Sea  $\{S_n\}$  una sucesión de subconjuntos compactos de  $\mathbb{R}^d$ , no vacíos, supongamos que en  $S_n$  rueda libremente una bola de radio  $\alpha_n$  con  $\alpha_n > \alpha > 0$ . Sea  $S$  tal que en  $S$  rueda libremente una bola de radio  $\beta > 0$ .*

$$\text{Si } d_\mu(S_n, S) \rightarrow 0 \text{ entonces } d_H(S_n, S) \rightarrow 0.$$

*Demostración.* En virtud del teorema anterior, basta demostrar que si en  $S_n$  rueda libremente una bola de radio  $\alpha_n$  con  $\alpha_n > \alpha > 0$  entonces dichos conjuntos son estándar respecto de la medida de Lebesgue, para el mismo  $\lambda$  y  $\delta$ . Veremos que esto se cumple para  $\lambda = \alpha$ . Lo mismo resultará para  $S$  (con  $\lambda' = \beta$ ). Sea  $x \in S_n$ , si  $x \in \partial S_n$  existe  $y \in S_n$  tal que  $x \in \partial B(y, \alpha)$  y  $B(y, \alpha) \subset S_n$ , sea  $\varepsilon < \alpha$  entonces

$$\mu(B(x, \varepsilon) \cap S_n) \geq \mu(B(x, \varepsilon) \cap B(y, \alpha)) \geq \delta_1 \mu(B(x, \varepsilon)),$$

y se cumple la estandaridad para  $x$ . Sea  $x \in \text{int}(S_n)$  y  $0 < \varepsilon < \alpha$ , si  $B(x, \varepsilon/2) \subset S_n$  se cumple la estandaridad para  $x$  y  $\delta_2 > 0$ . Si  $B(x, \varepsilon/2) \cap S_n^c \neq \emptyset$  existen  $z \in \partial S_n$  e  $y$  tal que  $z \in \partial B(y, \alpha)$  y  $z \in B(x, \varepsilon/2)$ , entonces

$$\mu(B(x, \varepsilon) \cap S_n) \geq \mu(B(x, \varepsilon) \cap B(y, \alpha)) \geq \delta_3 \mu(B(x, \varepsilon)),$$

tomando ahora  $\delta = \min\{\delta_1, \delta_2, \delta_3\}$  se sigue la tesis.  $\square$

*Conjuntos como convexos: una primera aproximación.*

Tanto la hipótesis de que el conjunto tenga alcance positivo, que sea  $r$ -convexo, o que ruede libremente por su complemento una bola de radio  $r$ , impiden que el conjunto tenga picos hacia adentro. No obstante, como se ve en Cuevas y otros (2012), dichas definiciones no son equivalentes en general. En el mencionado trabajo se demuestra que si un conjunto compacto tiene alcance mayor o igual que  $r$  es  $r$ -convexo y que si es  $r$ -convexo, entonces rueda libremente por su complemento una bola de radio  $r$ . Asimismo se presentan ejemplos en los cuales se ve que las inclusiones son estrictas. El concepto de cono convexidad exterior (que será dado en detalle en el capítulo 3) incluye el caso de conjuntos con picos hacia adentro, no obstante impone una restricción en la forma de dichas entrantes. Vamos a pedir que en cada punto del borde se pueda poner un cono con vértice en el punto, de ángulo  $\rho > 0$  totalmente incluido en el complemento. De manera intuitiva estamos pidiendo que desde cada punto del borde del conjunto se pueda *ver* un ángulo  $\rho$  del complemento del mismo. Es decir, excluimos conjuntos con picos hacia adentro *demasiado agudos* (no lineales). En el caso extremo en que  $\rho = \pi$  obtenemos la familia de los conjuntos convexos. Es fácil construir ejemplos de conjuntos que cumplan la condición de cono convexidad exterior para los cuales no ruede libremente una bola de radio  $r$  para ningún  $r$ , basta pensar por ejemplo en dos segmentos que se intersecan. Por otro lado es inmediato verificar que si una bola de radio  $r$  rueda libremente por el complemento del conjunto, este verifica la condición de cono convexidad exterior *finita* (en la variante *finita* si bien *vemos* el complemento del conjunto en un ángulo  $\rho$ , el alcance de la visión es acotado). De esto último se sigue que la clase de los conjuntos que cumplen la condición de cono convexidad exterior

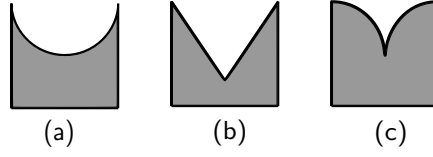


Figura 1.1: (a) Conjunto  $r$ -convexo, para algún  $r$  (b) Conjunto que cumple la condición de cono convexidad exterior pero no es  $r$ -convexo para ningún  $r$  (c) Conjunto que no es  $r$ -convexo ni cumple la condición de cono convexidad exterior.

es *más grande* que la clase para los cuales rueda libremente una bola por su complemento. En la Figura 1.1 se ve claramente lo antes expresado. El otro concepto que será introducido es el de cono convexidad por complementos, que es el análogo para conos del concepto de  $r$ -convexidad. Dado un ángulo  $\rho$  y una altura  $h$  diremos que un conjunto es cono convexo por complementos si es igual a la intersección de los complementos de los conos (de ángulo  $\rho$  y altura  $h$ ) que no cortan al conjunto. Más adelante veremos que si bien la definición es similar a la de conjunto  $r$ -convexo, no es cierto en general que un conjunto  $\rho$  cono convexo por complementos satisfaga (para el mismo  $\rho$  y  $h$ ) la condición de cono convexidad finita.

### Estimación de longitudes y superficies

Así como para estimar conjuntos fue necesario definir nociones de distancia, si lo que nos interesa es medir el volumen (longitud, área) del borde de  $S$  debemos tener una noción de longitud, área, etc. Con este fin vamos a introducir a continuación el concepto de contenido de Minkowski.

**Definición 1.16.** Sea  $S \subset \mathbb{R}^d$  no vacío y acotado. Definimos el contenido inferior de Minkowski, de dimensión  $m$  ( $m \leq d$ ), de  $S$  como:

$$\mathcal{M}_{*m}(S) = \liminf_{\varepsilon \rightarrow 0} \frac{\mu_d(B(S, \varepsilon))}{\varepsilon^{d-m} \omega_{d-m}},$$

y el superior como:

$$\mathcal{M}^{*m}(S) = \limsup_{\varepsilon \rightarrow 0} \frac{\mu_d(B(S, \varepsilon))}{\varepsilon^{d-m} \omega_{d-m}},$$

donde  $\mu_d$  denota la medida de Lebesgue en  $\mathbb{R}^d$  y  $\omega_{d-m}$  es la medida de Lebesgue de la bola de centro 0 y radio 1 en  $\mathbb{R}^{d-m}$ . Si ambos valores coinciden, dicho valor se denomina contenido de Minkowski de dimensión  $m$  de  $S$  y se denota  $\mathcal{M}^m(S)$ . Si  $S$  es un subconjunto compacto de  $\mathbb{R}^d$ , denotaremos  $L_0(\partial S)$  al contenido de Minkowski de dimensión  $d-1$  de  $\partial S$  en caso de que éste exista. Es decir:

$$L_0(\partial S) = \lim_{\varepsilon \rightarrow 0} \frac{\mu_d(B(\partial S, \varepsilon))}{2\varepsilon}.$$

Definiremos también el contenido exterior de Minkowski de un subconjunto compacto  $S$  de  $\mathbb{R}^d$  como el límite

$$\lim_{\varepsilon \rightarrow 0} \frac{\mu_d(B(S, \varepsilon) \setminus S)}{\varepsilon}, \quad (1.3)$$

en caso de que dicho límite exista.

No nos extenderemos en el análisis del contenido de Minkowski, un estudio más completo puede encontrarse en Matilla (1995) y en Federer (1969). Se puede demostrar (véase Ambrosio y otros (2008)) que si  $S \subset \mathbb{R}^d$  es compacto no vacío y tiene alcance positivo, existe el contenido de Minkowski  $d-1$  dimensional de  $\partial S$  y coincide con la medida de Hausdorff  $d-1$  dimensional de  $\partial S$ .

En lo referente a estimación del contenido de Minkowski del borde de un conjunto, además del ya mencionado trabajo de Bräker (ver Bräker y Hsing (1998)) sobre el problema de la estimación del perímetro de un subconjunto convexo de  $\mathbb{R}^2$ , existen numerosos trabajos. En Cuevas y otros (2007) se propone un estimador consistente para el caso en que el conjunto y el complemento cumplen la hipótesis de estandaridad, y se cuenta con dos muestras, una dentro y otra fuera del conjunto. Más aún la distribución asintótica de dicho estimador es normal (ver Armendáriz y otros (2009)). Para el caso en que el conjunto y el complemento son  $r$ -convexos, y se cuenta también con dos muestras, en Pateiro-López (2008) se propone un estimador del contenido de Minkowski del borde, y se demuestra que el orden de convergencia del mismo es  $\mathcal{O}((\log(n)/n)^{1/(d+1)})$ .

## 1.2. Aspectos computacionales

Si bien el principal énfasis en la tesis está puesto en los resultados teóricos que garantizan la consistencia de los estimadores que se proponen, hemos incorporado en cada capítulo aplicaciones y comentarios relacionados con la implementación computacional de los mismos. Dado el rápido crecimiento que ha tenido en la estadística en los últimos años, y sus ventajas en cuanto a facilidad de uso, hemos usado el software R (de licencia GNU) para programar los algoritmos. Dicho software cuenta con paquetes que permiten calcular ciertos objetos como por ejemplo la envolvente  $r$ -convexa de una muestra de puntos en  $\mathbb{R}^2$  (ver: <http://cran.r-project.org/web/packages/alphahull>). Este paquete permite además obtener y dibujar, mediante el comando `delvor` dos objetos que usaremos más adelante, que definiremos a continuación:

**Definición 1.17.** Dado  $\aleph_n = X_1, \dots, X_n$  un conjunto de puntos de  $\mathbb{R}^d$ , el diagrama de Voronoi de  $\aleph_n$  es el cubrimiento de  $\mathbb{R}^d$  por  $n$  regiones  $V_i$  donde:

$$V_i = \{x \in \mathbb{R}^d : \|x - X_i\| \leq \|x - X_j\| \quad \forall X_j \in \aleph_n\}.$$

**Observación 1.18.** *Se puede demostrar que las regiones  $V_i$  son cerradas, convexas, y que  $\text{int}(V_i) \cap \text{int}(V_j) = \emptyset \quad \forall i \neq j$ . Ver Aurenhammer y Klein (2000) para un estudio de los mismos. Una observación que nos será de utilidad más adelante es que, si denotamos  $v$  el número de vértices del diagrama de Voronoi, (es decir los puntos que están en la intresección de 3 regiones) entonces  $v \leq 2n - 2$ .*

**Definición 1.19.** Dado  $\aleph_n = X_1, \dots, X_n$  un conjunto de puntos de  $\mathbb{R}^2$ , la triangulación de Delaunay asociada a  $\aleph_n$  es el conjunto de triángulos con vértices en  $\aleph_n$  que cumplen que la circunferencia circunscrita a cada triángulo no contiene puntos de  $\aleph_n$ . Es fácil ver que dichos triángulos tienen como vértices los puntos de la muestra y que una arista  $e_i$  que une  $X_i$  con  $X_j$  está en un triángulo si y solo si las regiones  $V_i$  y  $V_j$  del diagrama de Voronoi, que contienen a  $X_i$  y  $X_j$  respectivamente, tienen un segmento del borde en común.

En la Figura 1.2 se muestran la triangulación de Delaunay y las celdas de Voronoi hechas con el paquete `alphahull`.

## 1.3. Ideas geométricas en el estudio de los datos funcionales

En el capítulo cuatro abordaremos el problema de regresión y clasificación (supervisada) en el contexto infinito dimensional mediante el uso de la distancia de Hausdorff. Dado un vector  $(X, Y) \in \mathcal{E} \times \mathbb{R}$ , tomaremos  $\mathcal{E}$  como el espacio de las funciones definidas en  $[0, 1]$  a valores reales, semicontinuas superiormente y positivas. Introduciremos en  $\mathcal{E}$  la métrica (que denotaremos  $\mathbb{H}$ ) dada por la distancia de Hausdorff entre los hipografos de las funciones (recordemos que dada  $f : [0, 1] \rightarrow \mathbb{R}$  positiva, el hipografo de  $f$  es  $H_f = \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq f(x), 0 \leq x \leq 1\}$ ). Los resultados y definiciones necesarios para la lectura del mismo se encuentran o bien en la sección anterior del presente capítulo o bien en el capítulo cuatro.

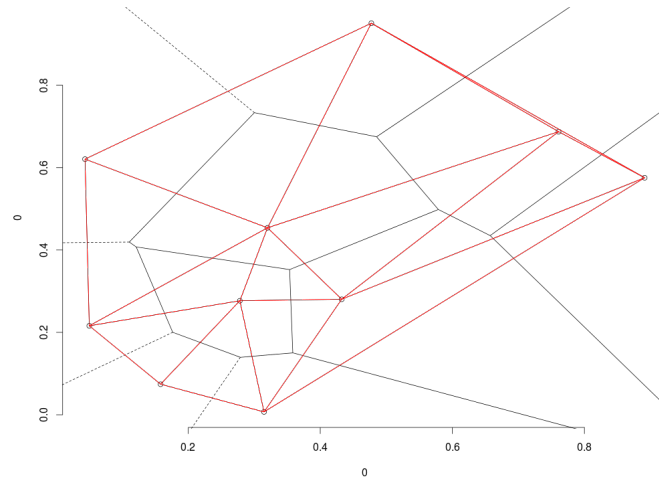


Figura 1.2: Para una muestra de  $n = 10$  puntos uniformemente distribuidos en el cuadrado unidad se representan en negro los bodes de las celdas de Voronoi y en rojo la triangulación de Delaunay.



## Capítulo 2

# Estimación paramétrica de algunas características geométricas

En este capítulo consideramos el problema de estimar algunas características geométricas de un conjunto compacto  $S$ , en particular la medida de su frontera y la curvatura integral media, a partir de una adecuada muestra aleatoria de puntos. Diferentes estimadores no paramétricos de la medida de la frontera fueron considerados por Cuevas y otros (2007, 2012, 2013), Pateiro-López (2008), Armendáriz y otros (2009) y Jiménez y Yukich (2011), utilizando muestras “dobles” de puntos dentro y fuera de  $S$ .

*La función de volumen y su interpretación estadística*

En general, para estimar (a partir de una muestra finita de puntos) ciertas características de un conjunto compacto  $S$ , como la medida de su frontera, es necesario imponer algunas condiciones de regularidad sobre dicho conjunto. Este es el caso en todas las referencias citadas en el párrafo anterior, en las que se imponen condiciones de carácter geométrico-topológico sobre  $S$ . Nuestro enfoque aquí es bastante distinto ya que nuestra condición básica sobre  $S$  se establece en términos de la llamada “función de volumen”  $V(r; S) = \mu(B(S, r))$  que se define como la medida de Lebesgue del conjunto “ $r$ -paralelo”.

- (a) *Funciones de volumen polinómicas.* En principio, las hipótesis sobre la estructura de  $V(r; S)$  no aparecen directamente como condiciones geométricas pero, como veremos, tienen interpretaciones geométricas muy profundas, relacionadas con la noción de convexidad. Un primer resultado clásico en este sentido es el Teorema de Steiner, probado en 1840, en el que se demuestra que si  $S \subset \mathbb{R}^d$  es convexo entonces  $V(r; S)$  resulta ser un polinomio en  $r$  de grado  $d$ . Más concretamente,

$$V(r; S) = \sum_{j=0}^d r^{d-j} \omega_{d-j} V_j(S), \quad (2.1)$$

donde  $\omega_k$  denota el volumen ( $k$ -dimensional) de la bola unidad en  $\mathbb{R}^k$ ,  $\omega_0 = 1$  y los coeficientes  $V_0(S), \dots, V_d(S)$  son los volúmenes intrínsecos de  $S$ . En particular,  $V_0(S) = 1$ ,  $V_d(S)$  es el volumen de  $S$ ,  $2V_{d-1}(S)$  es el área lateral y  $V_{d-2}(S)$  es la curvatura media integrada de  $\partial S$ .

En un famoso artículo, considerado como una referencia pionera en teoría geométrica de la medida, Federer (1959) obtuvo una importante generalización del resultado de Steiner para conjuntos de “alcance positivo”, definiendo el alcance de  $S$ ,  $\text{reach}(S)$ , como el supremo (posiblemente infinito) de los valores  $r > 0$  tales que, si  $\mathcal{D}(x, S) < r$  entonces  $S$  contiene un único punto más cercano a  $x$ . Federer probó que, si  $\text{reach}(S) = R$  entonces la fórmula de Steiner

(2.1) es válida para  $0 \leq r < R$  con la única diferencia de que el coeficiente de  $r^d$  no es necesariamente igual al volumen de la bola unidad  $\omega_d$ , pero coincide con la característica de Euler de  $S$  multiplicada por  $\omega_d$ .

La clase de conjuntos compactos  $S$  para los cuales  $\text{reach}(S) > 0$  es mucho más amplia que la clase de los convexos compactos (de hecho si  $S$  es cerrado la convexidad de  $S$  es equivalente a  $\text{reach}(S) = \infty$ ). Además, la propiedad de alcance positivo tiene una simple y atractiva interpretación “visual” (excluye la presencia de entrantes “puntiagudos” en el conjunto).

Nuestro desarrollo en este capítulo se basará precisamente en imponer la condición de que la función de volumen sea polinómica, con coeficiente de  $r^d$  igual a  $\omega_d$ , en algún entorno de 0. Como veremos, la clase de conjuntos que cumplen esto es aún más amplia que la clase de los conjuntos de alcance positivo y característica de Euler 1.

- (b) *Interpretación estadística de la condición de volumen polinómico.* Por otra parte, como consecuencia de resultados en Stachó (1976), si  $S$  es uno de nuestros conjuntos de volumen polinómico y  $X$  es una variable uniformemente distribuida en la banda  $B(S, R) \setminus S$  la distancia de  $X$  a  $S$ ,  $D = \mathcal{D}(X, S)$  resulta tener una distribución con densidad de expresión muy simple, asociada a la función de volumen, que en el caso  $d = 2$  tiene como único parámetro desconocido la longitud de la frontera de  $S$  y en el caso de dimensión  $d = 3$  depende de dos parámetros desconocidos: el área de la frontera y la curvatura integral media de  $S$ .

### Resultados obtenidos

Las anteriores consideraciones sugieren de manera natural estimar la longitud  $L_0$  de la frontera  $\partial S$  (en el caso  $d = 2$ ), el área lateral (denotada también  $L_0$ ) y la curvatura media integrada  $M$  (en el caso  $d = 3$ ) a partir de una muestra aleatoria  $D_1, \dots, D_n$  de las distancias a  $S$  de puntos uniformes en  $B(S, R) \setminus S$ . Este modelo muestral puede motivarse (con una cierta simplificación) en términos de problemas de “remote sensing” en los que se conocen las distancias a un objeto, por ejemplo un cuerpo sumergido, a partir de diferentes puntos exteriores. Existen en la literatura algunos trabajos donde se consideran modelos y problemas similares. En Genovese y otros (2012a) se busca estimar  $\Gamma_f = \{f(u) : 0 \leq u \leq 1\}$  siendo  $f : [0, 1] \rightarrow \mathbb{R}^d$ , a partir de observaciones  $Y_i$  que siguen un modelo aditivo del tipo  $Y_i = f(U_i) + \varepsilon_i$   $i : 1, \dots, n$ , donde  $U_i$  son i.i.d de una distribución  $H$  en  $[0, 1]$  y  $\varepsilon_i$  ruidos con distribución  $F$  con esperanza 0. En este caso la hipótesis de regularidad requerida se expresa en términos de las fibras  $L(u) = \{f(u) + tN(u), -\sigma \leq u \leq \sigma\}$  con  $u \in [0, 1]$ , siendo  $N(u)$  un vector normal. Se requiere que exista  $\sigma > 0$  que haga que las fibras  $L(u)$  y  $L(v)$  sean disjuntas, para  $u \neq v$ . El caso de variedades de mayor dimensión se estudia por ejemplo en Genovese y otros (2012b) con un modelo similar al que usaremos aquí y con hipótesis de regularidad en la variedad, expresada en términos de las fibras, análogas a las que se imponen en dimensión uno.

En definitiva, las principales contribuciones de este capítulo son:

1. Un estudio detallado de la condición de volumen polinómico, incluyendo varios resultados que interpretan esta condición en términos estadísticos.
2. Un análisis estadístico de los estimadores de momentos y de máxima verosimilitud para los parámetros  $L_0$  y  $M$  en los casos  $d = 2, 3$ . En concreto, se prueba la normalidad asintótica (calculando la correspondiente varianza asintótica) de los estimadores de momentos de estos parámetros (para  $d = 2, 3$ ) y del estimador de máxima verosimilitud de  $L_0$  (para  $d = 2$ ).
3. Un estudio de simulación para ilustrar el comportamiento de estos estimadores.

## 2.1. Conjuntos con volumen polinómico

En esta sección estudiaremos la función de volumen  $V(r; S)$  de un conjunto  $S \subset \mathbb{R}^d$  y nos centraremos en los conjuntos para los cuales  $V(r; S)$  es polinomial en  $r$ . Presentaremos brevemente

algunos resultados clásicos, para el caso en que  $S$  es un subconjunto compacto, con alcance positivo. Un resumen muy detallado con los resultados conocidos para conjuntos con alcance positivo puede encontrarse en Thäle (2008). Luego veremos que existen conjuntos con alcance 0 tal que la función de volumen es polinomial en  $r$ , para  $0 \leq r < R$  con  $R > 0$ .

### 2.1.1. Alcance positivo y fórmula de Federer

Consideremos un subconjunto compacto  $S \subset \mathbb{R}^d$  tal que  $\text{reach}(S) = R > 0$ , la fórmula de Federer, (ver Federer (1959)) nos dice que la función de volumen  $V(r; S)$  es un polinomio de grado a lo sumo  $d$ , si  $0 \leq r < R$ . El resultado demostrado es el siguiente:

**Teorema 2.1.** *Sea  $S$  compacto tal que  $\text{reach}(S) > 0$ . Sea  $K$  un subconjunto de Borel de  $\mathbb{R}^d$ . Existen y son únicas, medidas de Radon  $\Phi_0(S, \cdot), \dots, \Phi_d(S, \cdot)$  en  $\mathbb{R}^d$  tal que si  $0 \leq r < \text{reach}(S)$*

$$\mu\left(B(S, r) \cap \{x \in \mathbb{R}^d : \xi_S(x) \in K\}\right) = \sum_{i=0}^d r^{d-i} \omega_{d-i} \Phi_i(S, K),$$

donde  $\mu$  denota la medida de Lesbesgue en  $\mathbb{R}^d$ ,  $\omega_0 = 1$  y, para  $j \geq 1$ ,  $\omega_j$  denota el volumen de la bola de radio 1, en  $\mathbb{R}^j$ .

Dicha expresión generaliza la fórmula de Steiner, que establece que, para el caso en que  $S$  es convexo, con interior no vacío y su borde  $\partial S$  es una hipersuperficie de clase  $C^2$ , entonces:

$$V(r; S) = \mu(B(S, r)) = \sum_{k=0}^d \Phi_k(S) r^k,$$

donde

$$\Phi_0(S) = \mu(S), \text{ y si } k \geq 1, \Phi_k(S) = \frac{\binom{d}{k}}{d} M_{k-1}(\partial S),$$

siendo

$$M_k(\partial S) = \frac{1}{\binom{d-1}{k}} \int_{\partial S} \Xi_k(\xi) d\nu_{\partial S}.$$

Hemos usado la notación  $\Xi_k$  para indicar la  $k$ -ésima curvatura media. Observemos que si  $S \subset \mathbb{R}^3$  está en dichas hipótesis y  $\mathcal{X}(S) = 1$  (donde  $\mathcal{X}(S)$  denota la característica de Euler de  $S$ ), entonces:

$$V(r; S) = \mu(S) + L_0(\partial S)r + M_1 r^2 + \frac{4}{3} \pi r^3,$$

aquí  $L_0(\partial S)$  denota el contenido exterior de Minkowski de  $\partial S$ . Si  $S$  es una variedad de dimensión 2, sin borde, inmersa en  $\mathbb{R}^3$  se puede demostrar (ver Morvan (2008)) que:

$$V(r; S) = 2L_0(\partial S)r + \frac{4}{3} \pi \mathcal{X}(S)r^3.$$

En Ambrosio y otros (2008) puede encontrarse una demostración de que si  $S$  es compacto y tiene alcance positivo es posible definir su contenido de Minkowski.

En Santaló (1976) se calculan los coeficientes  $M_0$  y  $M_1$  para diferentes subconjuntos de  $\mathbb{R}^3$ . Aquí presentaremos los resultados para el caso en que el conjunto es un cono, ya que nos será de utilidad más adelante.

Si  $S$  es un *cono de revolución*, de altura  $h$  y radio  $t$  tenemos que

$$L_0(\partial S) = \pi t^2 + \pi t \sqrt{t^2 + h^2},$$

$$M_1 = \pi^2 t + \pi h - \pi t \arctan(h/t).$$

### 2.1.2. El caso general

Una pregunta que surge de manera natural a partir de la fórmula de Federer es si el hecho de que  $V(r; S)$  sea polinomial en  $r$ , para  $0 \leq r < R$ , implica que dicho conjunto tenga alcance positivo. Es fácil ver que esto no es cierto en general, basta pensar en dos segmentos en  $\mathbb{R}^2$  que se intersequen. El siguiente teorema relaciona, para subconjuntos compactos de  $\mathbb{R}^2$ , el intervalo de validez de la expresión polinomial, con la convexidad del conjunto:

**Teorema 2.2.** *Sea  $S \subset \mathbb{R}^2$  compacto, tal que*

$$V(r; S) = \mu(S) + c_1 r + c_2 r^2 \quad \forall r \geq 0,$$

*con  $c_1, c_2 \in \mathbb{R}$ , entonces  $S$  es convexo.*

Una demostración de este resultado puede encontrarse en Heveling y otros (2004). Más adelante veremos que el mismo (como se muestra en Heveling y otros (2004)) no es válido en  $\mathbb{R}^d$  para  $d \geq 3$ , es decir existen subconjuntos compactos, tal que  $V(r; S)$  es polinomial para todo  $r \geq 0$  pero que no son convexos.

**Definición 2.3.** Llamaremos  $VP_d(R)$  a la familia de subconjuntos de  $\mathbb{R}^d$  cuya función de volumen  $V(r; S)$  es polinomial en  $[0, R)$  para algún  $R > 0$  (no necesariamente el mismo para cada conjunto). Observemos por ejemplo que si  $S \subset \mathbb{R}^3$  pertenece a la familia  $VP_3(R)$  entonces su función de volumen será de la forma:

$$V(r; S) = \mu(S) + L_0 r + ar^2 + br^3 \quad r \in [0, R),$$

donde  $L_0 = L_0(\partial S)$  denota el contenido exterior de Minkowski de  $\partial S$ , para ciertos coeficientes  $a$  y  $b$ .

Ejemplos de conjuntos en  $VP_2(R)$  y  $VP_3(R)$  que no tienen alcance 0 se muestran en la Figura 2.1. Los cálculos que muestran que la función de volumen del conjunto en (b) (constituido por un círculo al que se le ha quitado un cono de ángulo  $\rho$ , con vértice en el centro) es polinomial se encuentran en la Proposición A.1 del apéndice. Como se ve en dicho cálculo la expresión polinomial depende del ángulo del cono que se quita de forma tal que cuanto más pequeño es el ángulo menor es el intervalo en el cual la función es polinomial. La función de volumen de los conjuntos que se muestran en (c) y (d) es también polinomial pero, a diferencia de lo que sucedía en (b) la expresión polinomial (ver los cálculos de la Proposición A.1 del apéndice) es válida para  $r \in [0, +\infty)$ .

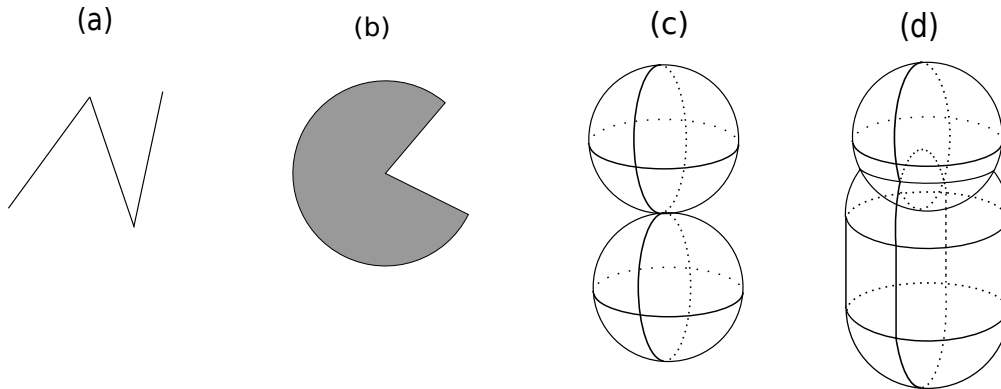


Figura 2.1: (a) (b) conjuntos en  $VP_2(R)$ . (c) y (d) conjuntos en  $VP_3(R)$  para todo  $R > 0$ .

**Observación 2.4.** Si consideramos  $S = \overline{B((0,1),2) \cup B((0,-1),2)}$  dicho conjunto es  $\rho$  cono convexo para cualquier  $\rho \leq 2\pi/3$ , es decir en cada punto del borde es posible poner un cono con vértice en el punto, y ángulo  $\rho$ , incluido en el complemento del conjunto, (en el siguiente capítulo daremos una definición formal de la  $\rho$  cono convexidad). El valor  $2\pi/3$  se debe a que éste es el ángulo que forman las rectas tangentes a las bolas, en el punto de intersección del borde de las mismas. La función de volumen del conjunto paralelo de  $S$  es (sumando el área de las dos bolas y restando la sección que se suma dos veces):

$$(2+r)^2 \left( 2\pi - 2 \arctan(\sqrt{r^2 + 2r + 3}) + 2 \frac{\sqrt{r^2 + 2r + 3}}{(2+r)^2} \right).$$

Por lo tanto la condición de cono convexidad no garantiza que la función de volumen del conjunto paralelo sea polinomial en algún intervalo. El recíproco tampoco es cierto para todo  $d$ , basta considerar en  $\mathbb{R}^d$ , para  $d \geq 3$  el conjunto formado por la unión de dos bolas tangentes, del mismo radio.

Consideremos  $S = \partial B((0,1),1) \cup \partial B((0,-1),1)$ , es claro que una bola de radio  $r \leq 1$  rueda libremente por  $S^c$ , sin embargo es fácil ver que el volumen del conjunto paralelo de  $S$  no es polinomial. Es decir, no es cierto en general que para cualquier conjunto  $S$  tal que en  $S^c$  rueda libremente una bola, el volumen del conjunto paralelo es polinomial.

## 2.2. Estimación paramétrica: Caso tridimensional

### 2.2.1. Introducción

Sea  $S \subset \mathbb{R}^3$  compacto perteneciente a la familia  $VP_3(R)$  introducida en la definición 2.3, con  $b = 4/3\pi$ . Dicha familia incluye, por ejemplo, los subconjuntos compactos, con alcance mayor o igual que  $R$  homeomorfos a una bola. Incluye además (ver Proposición A.1 del apéndice) los conjuntos c) y d) que se muestran en la Figura 2.1. Consideremos  $X$  una variable aleatoria con distribución uniforme en  $B(S,R) \setminus S$ . Sea  $D = d(X,S)$  la distancia euclidiana de  $X$  a  $S$ . Denotamos como  $F(r) = \mathbb{P}(D \leq r)$  la función de distribución de  $D$ . Como  $S \in VP_3(R)$  y estamos suponiendo que  $b = 4/3\pi$ , su función de volumen es:

$$V(r; S) = \mu(S) + L_0 r + ar^2 + \frac{4}{3}\pi r^3, \quad 0 \leq r < R,$$

de donde la densidad de  $D$  es:

$$F'(r) = \frac{V'(r; S)}{V(R; S) - \mu(S)} = \frac{L_0 + 2ar + 4\pi r^2}{L_0 R + aR^2 + \frac{4}{3}\pi R^3} \quad 0 \leq r < R.$$

A partir de una muestra  $D_1, \dots, D_n$  i.i.d de la variable  $D$  hallaremos los estimadores de  $L_0$  y  $a_0$  por el método de los momentos. Veremos que los estimadores que así se obtienen son asintóticamente normales y convergen c.s. a  $L_0$  y  $a$ . Dado que en dimensión dos (es decir para subconjuntos del plano) presentan la particularidad de tener esperanza infinita, estudiaremos una versión *truncada* que, si bien no es asintóticamente consistente, resuelve el problema de la esperanza. Al final del capítulo presentaremos los resultados de algunas simulaciones hechas para estos estimadores, para el caso en que el conjunto es un cono de altura 1 y ángulo  $\pi/3$ . Alternativamente podríamos haber planteado el estimador máximo verosímil de los parámetros; si bien eso es posible, (para todo  $L_0$  y  $a$ ) la condición de invertibilidad de la matriz de información no es claro que se cumpla en general.

### 2.2.2. Método de los Momentos

**Teorema 2.5.** Sea  $S \subset \mathbb{R}^3$  compacto no vacío,  $S \in VP_3(R)$  con  $b = 4/3\pi$ , los estimadores por momentos de  $a$  y de  $L_0$  son

$$\tilde{a} = -\frac{4\pi R}{5} \left[ \frac{3R^2 - 16\overline{D}R + 15\overline{D}^2}{R^2 - 6\overline{D}R + 6\overline{D}^2} \right], \quad (2.2)$$

y

$$\tilde{L}_0 = \frac{2\pi R^2}{5} \left[ \frac{3R^2 - 12\bar{D}R + 10\bar{D}^2}{R^2 - 6\bar{D}R + 6\bar{D}^2} \right], \quad (2.3)$$

respectivamente. Además, si denotamos

$$g_1(u, v) = \frac{2\pi R^2}{5} \left[ \frac{3R^2 - 12uR + 10v}{R^2 - 6uR + 6v} \right],$$

y

$$g_2(u, v) = -\frac{4\pi R}{5} \left[ \frac{3R^2 - 16uR + 15v}{R^2 - 6uR + 6v} \right],$$

entonces,

$$\sqrt{n}(\tilde{L}_0 - L_0) \xrightarrow{\omega} N(0, \sigma_{\tilde{L}_0}^2) \quad y \quad \sqrt{n}(\tilde{a} - a) \xrightarrow{\omega} N(0, \sigma_a^2),$$

siendo

$$\sigma_{\tilde{L}_0}^2 = \nabla g_1^t \Sigma_{D, D^2} \nabla g_1 \quad y \quad \sigma_a^2 = \nabla g_2^t \Sigma_{D, D^2} \nabla g_2.$$

*Demostración.* Haciendo algunos cálculos elementales obtenemos que

$$\mathbb{E}(D) = \int_0^R r \frac{L_0 + 2ar + 4\pi r^2}{L_0R + aR^2 + \frac{4}{3}\pi R^3} dr = \frac{3L_0R + 4aR^2 + 6\pi R^3}{6(L_0 + aR + \frac{4}{3}\pi R^2)}, \quad (2.4)$$

$$\mathbb{E}(D^2) = \int_0^R r^2 \frac{L_0 + 2ar + 4\pi r^2}{L_0R + aR^2 + \frac{4}{3}\pi R^3} dr = \frac{10L_0R^2 + 15aR^3 + 24\pi R^4}{30(L_0 + aR + \frac{4}{3}\pi R^2)}, \quad (2.5)$$

$$\text{Var}(D) = \frac{R^2(12\pi^2 R^4 + 24\pi aR^3 + 10a^2 R^2 + 44\pi L_0R^2 + 30L_0aR + 15L_0^2)}{20(4\pi R^2 + 3aR + 3L_0)^2},$$

$$\text{Var}(D^2) = \frac{R^4(768\pi^2 R^4 + 1360\pi aR^3 + 525a^2 R^2 + 1920\pi L_0R^2 + 1260L_0aR + 560L_0^2)}{700(4\pi R^2 + 3aR + 3L_0)^2},$$

$$\text{Cov}(D, D^2) = \frac{R^3(16\pi^2 R^4 + 30\pi aR^3 + 12a^2 R^2 + 48\pi L_0R^2 + 32L_0aR + 15L_0^2)}{20(4\pi R^2 + 3aR + 3L_0)^2},$$

los estimadores  $\tilde{L}_0$  y  $\tilde{a}$  se obtienen resolviendo el sistema formado por (2.4) y (2.5).

Definimos

$$g_1(u, v) = \frac{2\pi R^2}{5} \left[ \frac{3R^2 - 12uR + 10v}{R^2 - 6uR + 6v} \right],$$

entonces  $\tilde{L}_0 = g_1(\bar{D}, \bar{D}^2)$ .

Podemos escribir

$$\sqrt{n}(\tilde{L}_0 - L_0) = \sqrt{n} \left[ g_1(\bar{D}, \bar{D}^2) - g_1(\mathbb{E}(D), \mathbb{E}(D^2)) \right].$$

Si hacemos un desarrollo de Taylor de  $g_1$  en el punto  $(\mathbb{E}(D), \mathbb{E}(D^2))$  y denotamos  $v = ((\bar{D}, \bar{D}^2) - (\mathbb{E}(D), \mathbb{E}(D^2)))$  obtenemos:

$$g_1(\bar{D}, \bar{D}^2) - g_1(\mathbb{E}(D), \mathbb{E}(D^2)) = \nabla g_1(\mathbb{E}(D), \mathbb{E}(D^2))^t v + r(v).$$

Por lo tanto basta ver que  $\sqrt{nr}(v) \xrightarrow{\omega} 0$ , para eso vamos a usar el siguiente lema:

**Lema 2.6.** Sea  $[z, z + v] \subset U$  con  $U$  abierto de  $\mathbb{R}^2$  y  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  de clase  $C^1$ , 2 veces diferenciable en todo punto del abierto  $(z, z + v)$  con

$$\left| \frac{\partial^2 g}{\partial w^2}(x) \right| \leq M|w|^2,$$

para todo  $x \in (z, z + v)$  y todo  $w \in \mathbb{R}^2$ , entonces, el resto  $r$  del desarrollo de Taylor de orden 1 de  $g$  verifica:

$$|r(v)| \leq \frac{M}{(p+1)!} |v|^2.$$

En virtud del lema anterior basta ver que  $g_1$  es  $C^2$  en un entorno de  $(\mathbb{E}(D), \mathbb{E}(D^2))$ . Observemos que

$$\begin{aligned} \frac{\partial g_1}{\partial u} &= \frac{12\pi R^3}{5} \frac{R^2 - 2v}{(R^2 - 6uR + 6v)^2}, \\ \frac{\partial g_1}{\partial v} &= \frac{8\pi R^3}{5} \frac{3u - 2R}{(R^2 - 6uR + 6v)^2}, \\ \frac{\partial g_1}{\partial u}(\mathbb{E}(D), \mathbb{E}(D^2)) &= \frac{3(5L_0 - 4\pi R^2)(4\pi R^2 + 3aR + 3L_0)}{\pi R^3}, \\ \frac{\partial g_1}{\partial v}(\mathbb{E}(D), \mathbb{E}(D^2)) &= \frac{5(2\pi R^2 - 3L_0)(4\pi R^2 + 3aR + 3L_0)}{\pi R^4}. \end{aligned}$$

Y las derivadas de orden 2 son:

$$\begin{aligned} \frac{\partial^2 g_1}{\partial u^2} &= \frac{144\pi R^4}{5} \frac{R^2 - 2v}{(R^2 - 6Ru + 6v)^3}, \\ \frac{\partial^2 g_1}{\partial u \partial v} &= \frac{24\pi R^3}{5} \frac{-7R^2 + 6Ru + 6v}{(R^2 - 6Ru + 6v)^3}, \\ \frac{\partial^2 g_1}{\partial v^2} &= \frac{96\pi R^3}{5} \frac{2R - 3u}{(R^2 - 6Ru + 6v)^3}. \end{aligned}$$

Para estudiar la continuidad de estas funciones es suficiente ver que los denominadores no se anulan en  $(\mathbb{E}(D), \mathbb{E}(D^2))$ , si sustituimos  $u$  por  $\mathbb{E}(D)$  y  $v$  por  $\mathbb{E}(D^2)$  en  $R^2 - Ru + 6v$  obtenemos:

$$\frac{2\pi R^4}{5} \frac{1}{3L_0 + 3aR + 4\pi R^2},$$

que es no nulo siempre que  $R > 0$ .

Para obtener la distribución asintótica de  $\tilde{a}$  definimos

$$g_2(u, v) = -\frac{4\pi R}{5} \left[ \frac{3R^2 - 16uR + 15v}{R^2 - 6uR + 6v} \right],$$

de donde  $\tilde{a} = g_2(\overline{D}, \overline{D^2})$  y

$$\sqrt{n}(\tilde{a} - a) = \sqrt{n} \left[ g_2(\overline{D}, \overline{D^2}) - g_2(\mathbb{E}(D), \mathbb{E}(D^2)) \right].$$

Nuevamente hay que calcular las derivadas parciales de orden 1 y 2 y estudiar la continuidad en

$(\mathbb{E}(D), \mathbb{E}(D^2))$ .

$$\begin{aligned}\frac{\partial g_2}{\partial u} &= -\frac{8\pi R^2(R^2 - 3v)}{5(R^2 - 6uR + 6v)^2}, \\ \frac{\partial g_2}{\partial v} &= \frac{12\pi R^2(R - 2u)}{5(R^2 - 6uR + 6v)^2}, \\ \frac{\partial g_2}{\partial u}(\mathbb{E}(D), \mathbb{E}(D^2)) &= \frac{(32\pi R + 15a)(4\pi R^2 + 3aR + 3L_0)}{\pi R^3}, \\ \frac{\partial g_2}{\partial v}(\mathbb{E}(D), \mathbb{E}(D^2)) &= -\frac{15(2\pi R + a)(4\pi R^2 + 3aR + 3L_0)}{\pi R^4}.\end{aligned}$$

□

### 2.2.3. Aspectos prácticos de los estimadores.

Dado que los estimadores por momentos presentados en la sección anterior en dimensión 2 tienen esperanza infinita, lo mismo debería suceder en dimensión 3. Por lo tanto tomar el error cuadrático medio o la distancia  $L^1$  no es apropiado en este caso. Presentaremos dos alternativas a dicho problema, la primera consiste en cambiar la distancia por una distancia acotada. La segunda consiste en usar una versión de los estimadores con esperanza finita. Dichas alternativas serán ilustradas en simulaciones al final del capítulo.

#### Una distancia apropiada

La distancia que usaremos será la siguiente

$$d_{BE}(T_n; \theta) = \mathbb{E} \left( \frac{|T_n - \theta|}{|T_n - \theta| + 1} \right).$$

La motivación para definir esta distancia se sigue de que si  $Z_n$  (con  $n \in \mathbb{N}$ ) y  $Z$  son variables aleatorias se tiene que

$$Z_n \xrightarrow{P} Z \text{ si y solo si } \mathbb{E} \left( \frac{|Z_n - Z|}{|Z_n - Z| + 1} \right) \rightarrow 0.$$

#### Una versión truncada

Observemos que la densidad

$$f(r; L_0, a_0) = \frac{L_0 + 2ar + 4\pi r^2}{L_0R + aR^2 + \frac{4}{3}\pi R^3}, \quad 0 \leq r < R,$$

puede expresarse como

$$f(r) = \lambda_1 f_1(r) + \lambda_2 f_2(r) + \lambda_3 f_3(r), \quad 0 \leq r < R, \quad (2.6)$$

donde

$$\lambda_1 = \frac{L_0}{L_0 + aR + 4\pi R^2/3}, \quad \lambda_2 = \frac{aR}{L_0 + aR + 4\pi R^2/3}, \quad \lambda_3 = \frac{4\pi R^2/3}{L_0 + aR + 4\pi R^2/3},$$

y, para  $i = 1, 2, 3$ ,  $f_i$  es la densidad de una variable aleatoria  $RZ_i$ , donde  $Z_1$  tiene distribución uniforme en  $(0, 1)$ ,  $Z_2$  tiene distribución Beta(2, 1) y  $Z_3$  tiene distribución Beta(3, 1). Por (2.4) y



(2.5) podemos plantear el sistema

$$\begin{aligned}\mathbb{E}(D) &= \frac{R}{2}\lambda_1 + \frac{2R}{3}\lambda_2 + \frac{3R}{4}\lambda_3 \\ \mathbb{E}(D^2) &= \frac{R^2}{3}\lambda_1 + \frac{R^2}{2}\lambda_2 + \frac{3R^2}{5}\lambda_3 \\ \mathbb{E}(D^3) &= \frac{R^3}{4}\lambda_1 + \frac{2R^3}{5}\lambda_2 + \frac{R^3}{2}\lambda_3\end{aligned}$$

de donde obtenemos

$$\begin{aligned}\tilde{\lambda}_1 &= \frac{12(6R^2\bar{D} - 20R\bar{D}^2 + 15\bar{D}^3)}{R^3} \\ \tilde{\lambda}_2 &= -\frac{30(4R^2\bar{D} - 15R\bar{D}^2 + 12\bar{D}^3)}{R^3} \\ \tilde{\lambda}_3 &= \frac{20(3R^2\bar{D} - 12R\bar{D}^2 + 10\bar{D}^3)}{R^3}\end{aligned}$$

y despejando, los estimadores de  $L_0$  y  $a$  son:

$$\tilde{L}_0 = \frac{4\pi R^2}{3} \frac{\tilde{\lambda}_1}{1 - \tilde{\lambda}_1} \left( \frac{1}{1 - \tilde{\lambda}_2} \right) \left( 1 - \frac{\tilde{\lambda}_1}{1 - \tilde{\lambda}_1} \frac{\tilde{\lambda}_2}{1 - \tilde{\lambda}_2} \right)^{-1} = \frac{4\pi R^2}{3\tilde{\lambda}_2} \sum_{j=1}^{\infty} \left[ \frac{\tilde{\lambda}_1}{1 - \tilde{\lambda}_1} \frac{\tilde{\lambda}_2}{1 - \tilde{\lambda}_2} \right]^j, \quad (2.7)$$

y

$$\begin{aligned}\tilde{a} &= \frac{4\pi R}{3} \frac{\tilde{\lambda}_2}{1 - \tilde{\lambda}_2} \left[ \frac{1}{1 - \tilde{\lambda}_2} \frac{\tilde{\lambda}_1}{1 - \tilde{\lambda}_1} \left( 1 - \frac{\tilde{\lambda}_1}{1 - \tilde{\lambda}_1} \frac{\tilde{\lambda}_2}{1 - \tilde{\lambda}_2} \right)^{-1} + 1 \right] \\ &= \frac{4\pi R}{3} \frac{1}{1 - \tilde{\lambda}_2} \left[ \sum_{j=1}^{\infty} \left[ \frac{\tilde{\lambda}_1}{1 - \tilde{\lambda}_1} \frac{\tilde{\lambda}_2}{1 - \tilde{\lambda}_2} \right]^j + \tilde{\lambda}_2 \right]. \quad (2.8)\end{aligned}$$

Las versiones truncadas de los estimadores son

$$\tilde{L}_0^T = \frac{4\pi R^2}{3\tilde{\lambda}_2} \sum_{j=1}^K \left[ \frac{\tilde{\lambda}_1}{1 - \tilde{\lambda}_1} \frac{\tilde{\lambda}_2}{1 - \tilde{\lambda}_2} \right]^j,$$

y

$$\tilde{a}^T = \frac{4\pi R}{3} \frac{1}{(1 - \tilde{\lambda}_2)} \left[ \sum_{j=1}^K \left[ \frac{\tilde{\lambda}_1}{1 - \tilde{\lambda}_1} \frac{\tilde{\lambda}_2}{1 - \tilde{\lambda}_2} \right]^j + \tilde{\lambda}_2 \right].$$

respectivamente. A su vez, en la práctica vamos a expresar  $1/(1 - \tilde{\lambda}_1)$ ,  $\tilde{\lambda}_1/(1 - \tilde{\lambda}_1)$  y  $\tilde{\lambda}_2/(1 - \tilde{\lambda}_2)$  por sus respectivas series truncadas.

*Algunas consecuencias de la representación (2.6)*

Como una ventaja adicional de la expresión (2.6), el estimador máximo verosímil de  $\lambda_1$  y  $\lambda_2$  puede ser calculado mediante el algoritmo *E-M* (ver capítulo 1 en Krishnan y McLachlan (2008)):

1. Paso inicial:  $\hat{\lambda}_1^{(0)} = 0,5 = \hat{\lambda}_2^{(0)}$ .
2. Iteramos mientras que  $|\hat{\lambda}_1^{(k)} - \hat{\lambda}_1^{(k-1)}| > tol_1$  o  $|\hat{\lambda}_2^{(k)} - \hat{\lambda}_2^{(k-1)}| > tol_2$ .

- a) **Paso E.** Para  $i = 1, \dots, n$ , denotemos  $Y_i^1$  la variable aleatoria (no observable) que indica si  $D_i$  se obtuvo de  $f_1$  o no, mientras que  $Y_i^2$  denota la variable aleatoria (no observable) que indica si  $D_i$  se obtuvo de  $f_2$ . Usando la fórmula de Bayes

$$Y_{i,k}^1 = \mathbb{E}(Y_i^1 | D_i, \hat{\lambda}_1^{(k)}, \hat{\lambda}_2^{(k)}) = \frac{\hat{\lambda}_1^{(k)} f_1(D_i)}{\hat{\lambda}_1^{(k)} f_1(D_i) + \hat{\lambda}_2^{(k)} f_2(D_i) + (1 - \hat{\lambda}_1^{(k)} - \hat{\lambda}_2^{(k)}) f_3(D_i)},$$

y

$$Y_{i,k}^2 = \mathbb{E}(Y_i^2 | D_i, \hat{\lambda}_1^{(k)}, \hat{\lambda}_2^{(k)}) = \frac{\hat{\lambda}_2^{(k)} f_2(D_i)}{\hat{\lambda}_1^{(k)} f_1(D_i) + \hat{\lambda}_2^{(k)} f_2(D_i) + (1 - \hat{\lambda}_1^{(k)} - \hat{\lambda}_2^{(k)}) f_3(D_i)},$$

definimos:

$$Q(\lambda_1, \lambda_2; \hat{\lambda}_1^{(k)}, \hat{\lambda}_2^{(k)}) = \sum_{i=1}^n [Y_{i,k}^1 \log \lambda_1 + Y_{i,k}^2 \log \lambda_2 + (1 - Y_{i,k}^1 - Y_{i,k}^2) \log(1 - \lambda_1 - \lambda_2)].$$

- b) **Paso M.** Al maximizar  $Q$  en  $\lambda_1$  y  $\lambda_2$  obtenemos  $\hat{\lambda}_1^{(k+1)}$  y  $\hat{\lambda}_2^{(k+1)}$ . Se puede ver que  $\hat{\lambda}_1^{(k+1)} = n^{-1} \sum_{i=1}^n Y_{i,k}^1$  y  $\hat{\lambda}_2^{(k+1)} = n^{-1} \sum_{i=1}^n Y_{i,k}^2$ .

### 2.3. Gráficos y Simulaciones en $\mathbb{R}^3$

Las gráficas de la varianza asintótica de los estimadores (2.2) y (2.3) se muestra en la Figura 2.2.

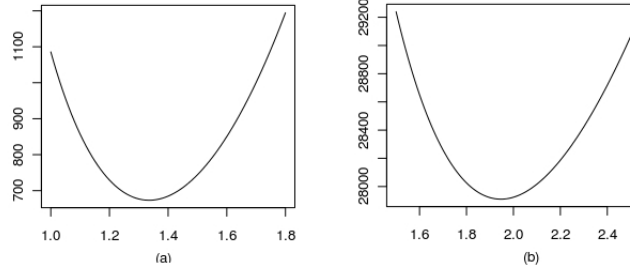


Figura 2.2: (a) Varianza asintótica del estimador  $\tilde{L}_0$  (2.3) para un cono de altura 1 y ángulo  $\pi/3$ . (b) Varianza asintótica del estimador  $\tilde{a}$  (2.2) para un cono de altura 1 y ángulo  $\pi/3$ .

Los resultados numéricos para el caso en que el conjunto es un cono de altura 1 y ángulo  $\pi/3$  (en cuyo caso  $L_0 = \pi$  y  $a = 6,9404$ ) se muestran en las tablas 2.1 hasta 2.6. Todas ellas están basadas en 2000 réplicas.

$n$	R=1.3 $\tilde{L}_0$	R=1.9 $\tilde{a}$	R=1.3 $\tilde{L}_0^T$	R=1.9 $\tilde{a}^T$
5000	3.0989	7.1621	3.0789	7.1567
10000	3.1449	7.0103	3.1338	7.0053
15000	3.1231	6.8696	3.1121	6.8656
20000	3.1245	6.9277	3.1383	6.8910
25000	3.1325	6.9341	3.1189	6.9300
30000	3.1530	6.9947	3.1331	6.9782
35000	3.1478	6.9510	3.1354	6.9470
40000	3.1398	6.9825	3.1299	6.9379

Tabla 2.1: Mediana sobre 2000 réplicas de los estimadores por momentos (2.7) y (2.8) y sus respectivas versiones truncadas  $\tilde{L}_0^T$  y  $\tilde{a}^T$  (tomando en ambos casos  $K = 8$ ).

$n$	R=1.3 $\tilde{L}_0$	R=1.9 $\tilde{a}$	R=1.3 $\tilde{L}_0^T$	R=1.9 $\tilde{a}^T$
5000	0.6557	4.2540	0.6810	4.2150
10000	0.4786	2.9131	0.4855	2.9008
15000	0.3703	2.3161	0.3840	2.3106
20000	0.3246	2.0613	0.3234	2.1194
25000	0.2883	1.8974	0.2989	1.8900
30000	0.2692	1.6455	0.2832	1.6742
35000	0.2490	1.5492	0.2544	1.5435
40000	0.2266	1.5175	0.2323	1.4426

Tabla 2.2: Desviación media absoluta (M.A.D. por sus siglas en inglés) sobre 2000 réplicas de los estimadores por el método de los momentos (2.7) y (2.8) y sus respectivas versiones truncadas  $\tilde{L}_0^T$  y  $\tilde{a}^T$  (tomando en ambos casos  $K = 8$ ).

**Observación 2.7.** *Observemos que generar datos de la densidad  $f(x; L_0)$  es simple ya que al depender únicamente de  $L_0$  y  $R$  podemos (suponiendo conocido el valor de  $L_0$ ) tomar un círculo de radio  $r$  de modo que su perímetro sea  $L_0$  y tomarlos uniformemente en el conjunto  $B(0, r + R) \setminus B(0, r)$ . Lo mismo sucede en  $\mathbb{R}^3$ , podemos generar los datos tomando conjuntos para los cuales es simple calcular la distancia hacia el borde, que tengan la misma curvatura media y superficie lateral.*

$n$	R=1.3 $\tilde{L}_0$	R=1.9 $\tilde{a}$	R=1.3 $\tilde{L}_0^T$	R=1.9 $\tilde{a}^T$
5000	0.3098	0.6715	0.3162	0.6700
10000	0.2463	0.6055	0.2509	0.6044
15000	0.2083	0.5611	0.2128	0.5602
20000	0.1924	0.5394	0.1934	0.5394
25000	0.1731	0.5139	0.1771	0.5129
30000	0.1637	0.4856	0.1702	0.4933
35000	0.1544	0.4789	0.1573	0.4780
40000	0.1437	0.4630	0.1469	0.4668

Tabla 2.3: Error ( $d_{EB}$ ) promedio sobre 2000 réplicas de los estimadores por el método de los momentos (2.7) y (2.8) y sus respectivas versiones truncadas  $\tilde{L}_0^T$  y  $\tilde{a}^T$  (tomando en ambos casos  $K = 8$ ).

$n$	R=1.3 $\hat{L}_0$	R=1.9 $\hat{a}$	R=1.3 $\hat{L}_0^T$	R=1.9 $\hat{a}^T$
5000	3.1521	7.1437	3.1393	7.3081
10000	3.1480	7.0499	3.1383	7.1004
15000	3.1433	7.0510	3.1324	7.0549
20000	3.1442	7.0419	3.1343	7.0573
25000	3.1404	7.0786	3.1374	7.0124
30000	3.1499	7.0610	3.1400	7.0741
35000	3.1395	7.0599	3.1290	7.0153
40000	3.1425	7.0361	3.1332	7.0123

Tabla 2.4: Mediana sobre 2000 réplicas de los estimadores por el método de máxima verosimilitud ( $\hat{L}_0$  y  $\hat{a}$ ) que se obtienen mediante el algoritmo  $E$ - $M$  y mediana de las respectivas versiones truncadas  $\hat{L}_0^T$  y  $\hat{a}^T$  (tomando en ambos casos  $K = 8$ ). En el algoritmo  $E$ - $M$  hemos tomado  $tol_1 = tol_2 = 0,0001$ .

$n$	$R = 1,3$ $\hat{L}_0$	$R = 1,9$ $\hat{a}$	$R = 1,3$ $\hat{L}_0^T$	$R = 1,9$ $\hat{a}^T$
5000	0.3842	1.8927	0.3814	2.1942
10000	0.2725	1.2942	0.2644	1.5391
15000	0.2268	1.0847	0.2248	1.2392
20000	0.1879	0.9594	0.1868	1.0846
25000	0.1772	0.8852	0.1579	0.9798
30000	0.1525	0.7773	0.1505	0.9321
35000	0.1380	0.7554	0.1368	0.8881
40000	0.1338	0.7083	0.1326	0.7721

Tabla 2.5: Desviación media absoluta, (M.A.D. por sus siglas en inglés) sobre 2000 réplicas de los estimadores por el método de máxima verosimilitud ( $\hat{L}_0$  y  $\hat{a}$ ) que se obtienen mediante el algoritmo  $E$ - $M$  y MAD de las respectivas versiones truncadas  $\hat{L}_0^T$  y  $\hat{a}^T$  (tomando en ambos casos  $K = 8$ ). En el algoritmo  $E$ - $M$  hemos tomado  $tol_1 = tol_2 = 0,0001$ .

$n$	$R = 1,3$ $\hat{L}_0$	$R = 1,9$ $\hat{a}$	$R = 1,3$ $\hat{L}_0^T$	$R = 1,9$ $\hat{a}_T$
5000	0.2172	0.5178	0.2135	0.5483
10000	0.1702	0.4425	0.1674	0.4794
15000	0.1468	0.4081	0.1444	0.4340
20000	0.1330	0.3845	0.1308	0.4082
25000	0.1255	0.3630	0.1178	0.3920
30000	0.1135	0.3448	0.1113	0.3794
35000	0.1057	0.3384	0.1078	0.3680
40000	0.0987	0.3183	0.0972	0.3460

Tabla 2.6: Error ( $d_{EB}$ ) promedio sobre 2000 réplicas de los estimadores por el método de máxima verosimilitud ( $\hat{L}_0$  y  $\hat{a}$ ) que se obtienen mediante el algoritmo  $E-M$  y sus respectivas versiones truncadas  $\hat{L}_0^T$  y  $\hat{a}^T$  (tomando en ambos casos  $K = 8$ ). En el algoritmo  $E-M$  hemos tomado  $tol_1 = tol_2 = 0,0001$ .

## Capítulo 3

# Aplicaciones de la propiedad de cono convexidad a la estimación de conjuntos

En este capítulo se introduce una nueva familia de conjuntos, llamados cono-convexos; estudiamos sus propiedades y abordamos el problema de estimar un conjunto  $S$  que cumpla esta condición de cono-convexidad usando un mecanismo de “cierre cono-convexo” similar a la noción usual de envoltura convexa. Una extensión análoga, considerada previamente en la literatura, es la llamada  $r$ -convexidad. La noción de cono-convexidad que aquí se introduce es notablemente menos restrictiva pero resulta aún manejable y mantiene un claro contenido intuitivo.

### *Cono-convexidad.*

Es bien sabido que un conjunto  $S$  convexo compacto de interior no vacío se caracteriza por la propiedad de hiperplano soporte: por cada punto de su frontera pasa un hiperplano que deja a todo el conjunto en uno de los dos semiespacios que define. Si cambiamos los semiespacios asociados a los hiperplanos por bolas de un radio fijo, obtenemos la condición de “bola exterior” (“rolling property”) que, según probó Henri Poincaré en 1899, desempeña un papel relevante en la teoría de potencial (problema de Dirichlet). En 1911, Stanislaw Zaremba debilitó la condición de Poincaré reemplazándola por la condición de “cono-exterior” (llamada a veces condición de Poincaré-Zaremba): en cada punto de la frontera del dominio  $S$  existe un cono finito con vértice en dicho punto cuyo interior no interseca a  $S$ . Diremos que el conjunto  $S$  es  $\rho, h$ -cono convexo cuando esta condición se verifica, para valores prefijados de  $\rho$  (la amplitud de los conos) y  $h$  (su altura) para todo  $x \in \partial S$ . De manera alternativa diremos que  $S$  es cono-convexo por complemento si  $S$  puede expresarse como la intersección de los complementarios de los  $\rho, h$ -conos (abiertos) que no intersecan a  $S$ . Probamos que, a diferencia de lo que sucede en el caso convexo, estas dos definiciones de cono-convexidad no coinciden.

Este capítulo está dedicado al estudio de las propiedades geométricas de los conjuntos cono-convexos y al problema estadístico de estimar un conjunto  $S$  cono-convexo (en cualquiera de las dos versiones) a través del correspondiente cierre cono-convexo de una muestra.

### *Principales aportaciones de este capítulo*

1. Se obtienen algunas propiedades geométricas topológicas de los conjuntos cono-convexos. En particular, se prueba que tienen frontera nula (Lebesgue) y que la convergencia Hausdorff de conjuntos cono-convexos implica la convergencia de sus fronteras.

2. Se prueba que la familia de conjuntos compactos cono-convexos de  $\mathbb{R}^d$  (en cualquiera de las dos versiones) es cerrada y localmente compacta con la métrica de Hausdorff.
3. Utilizando los resultados anteriores y un teorema de Billingsley y Topsøe (1967), se establece que la familia de los conjuntos cono-convexos (en cualquiera de las dos versiones) es una clase de Glivenko-Cantelli.
4. Se estudia el comportamiento asintótico de los cierres cono-convexos (ordinario y por complemento) de una muestra aleatoria  $X_1, \dots, X_n$  extraída de una distribución absolutamente continua con soporte  $S \subset \mathbb{R}^d$  y se obtienen las tasas de convergencia respecto a la distancia de tipo  $L_1$  (distancia de la diferencia simétrica) y la distancia de Hausdorff.
5. Se propone un test para la hipótesis de convexidad basado en la noción de “multivariate spacing” de Janson (1987).
6. Se propone un algoritmo estocástico para el cálculo del cierre cono-convexo por complementos. La flexibilidad de forma de este estimador sugiere que puede ser especialmente adecuado en el problema de estimar el “home range” que ha sido objeto de gran atención en Zoología y Ecología.

### 3.1. Propiedad de Poincaré y conjuntos cono-convexos. Definiciones y propiedades

#### 3.1.1. Antecedentes históricos

La cono convexidad (en una versión más débil que la que daremos) fue relacionada por Stanislaw Zaremba con el problema de Dirichlet (ver Zaremba (1911)). Recordemos que, dada una función  $f$  continua definida en el borde de un dominio  $\Omega \subset \mathbb{R}^d$ , el problema de Dirichlet, propuesto por Gauss en 1840, consiste en encontrar una función armónica  $u : \Omega \cup \partial\Omega \rightarrow \mathbb{R}$  tal que  $f(x) = u(x)$  para todo  $x \in \partial\Omega$ . Perrón en 1923 demostró que, en caso de haber solución, coincide con el supremo de las funciones subarmónicas definidas en  $\Omega$ , acotadas superiormente por  $f$  en  $\partial\Omega$  (ver Teorema 2.12 en Gilbarg y Trudinger (2001)). Si bien Gauss creía que existía solución del mismo para todo  $U$ , en Zaremba (1909) y Lebesgue (1924) se presentan ejemplos de que esto no es cierto en general. El ejemplo de Zaremba consiste en considerar  $U$  como el disco de radio 1 en  $\mathbb{R}^2$ , quitarle el centro, y definir  $f$  como 1 en 0 y 0 en el borde del disco. El ejemplo de Lebesgue (conocido como espina de Lebesgue) es válido en  $\mathbb{R}^d$  para  $d \geq 3$ , el mismo consiste en tomar una bola y en un punto de la misma hacer una entrante suficientemente aguda. En Zaremba (1911) se demuestra que si para todo  $x \in \partial U$  es posible encontrar un cono incluido en el complemento del conjunto, con vértice en  $x$ , entonces existe solución al problema de Dirichlet en  $U$  (observemos que lo que estamos haciendo es impedir que el conjunto pueda tener una entrante muy aguda). No es necesario que el ángulo o la altura del cono sean el mismo para cada  $x$ . El resultado de Zaremba generaliza el trabajo Poincaré (1899), en este último Poincaré demostró que si en cada punto del borde es posible poner una bola tangente al conjunto, incluida en el complemento, entonces existe solución. Una referencia más nueva del resultado de Zaremba es Mörters y Peres (2010). En otros contextos, como por ejemplo en el estudio de la existencia de soluciones de las ecuaciones elípticas dicha condición es requerida (ver Che y otros (1999)). Permite además (junto con otra condición geométrica que se denomina *conjunto estrellado*) obtener cotas para la aproximación (en la norma  $H^0$  y  $H^1$ ) de una función armónica  $u$  (considerada como función en el espacio de Sobolev  $H^k(S)$  para algún  $k \geq 1$ ) por un polinomio armónico (ver Blowey y Craig (2004)). La variante *interior* de la condición (es decir el caso en que para cada punto del borde del conjunto es posible poner un cono de ángulo y altura fijos, con vértice en dicho punto, incluido en el conjunto) es también de interés y ha sido estudiada en relación a los espacios de Sobolev sobre el conjunto. Es necesaria por ejemplo, para la prueba de desigualdades como la *desigualdad de Sobolev-Poincaré* y otros resultados (ver Burenkov (1998) y

Blowey y Craig (2004)). En este capítulo nos centraremos en los conjuntos que satisfacen la condición de cono convexidad para conos de ángulo y altura (finita o no) fijos.

### 3.1.2. Notación y definiciones básicas

Antes de definir formalmente las dos clases de conjuntos cono convexos mencionadas vamos a introducir la notación que usaremos. Dados  $x \in \mathbb{R}^d$  y  $\rho \in (0, \pi]$ , llamaremos cono abierto con vértice en  $x$ , ángulo  $\rho$ , y eje el vector  $\xi$  (con  $\|\xi\| = 1$ ) al conjunto

$$C_{\rho,\xi}(x) = \left\{ z \in \mathbb{R}^d, z \neq x : \left\langle \xi, \frac{z-x}{\|z-x\|} \right\rangle > \cos(\rho/2) \right\}.$$

Dependiendo del contexto, denotaremos  $C_{\rho,\xi}(x)$  o simplemente  $C_\rho(x)$ . Consideraremos también los conos de *altura finita*:

$$C_{\rho,\xi,h}(x) = B(x, h) \cap C_{\rho,\xi}(x).$$

Por simplicidad, cuando no sea necesario especificar  $\xi$  usaremos la notación  $C_{\rho,h}(x)$ .

Vamos a introducir a continuación dos definiciones de cono convexidad, la primera es similar a la de rodamiento libre de una bola por el complemento del conjunto (ver Definición 1.7) donde en lugar de una bola pedimos que para cada punto  $x$  del borde del conjunto se pueda tomar un cono con vértice en  $x$ , ángulo  $\rho$  y altura  $h$ , incluido en el complemento del conjunto. La segunda es el análogo para conos de la condición 1.2 que define la  $r$ -convexidad sustituyendo los complementos de bolas por complementos de conos.

**Definición 3.1.** Decimos que un conjunto  $S \subset \mathbb{R}^d$  compacto es  $\rho$ -cono convexo en sentido amplio si existen  $\rho \in (0, \pi]$  y  $h > 0$ , tal que:

$$\forall x \in \partial S \exists \xi = \xi(x) \text{ tal que } C_{\rho,\xi,h}(x) \subset S^c. \quad (3.1)$$

Es claro que si se cumple la condición anterior para  $\rho \in (0, \pi]$ , también se cumple para  $0 < \rho' < \rho$ . Tomaremos el máximo  $\rho$  que verifica la propiedad anterior. Denotaremos  $\mathcal{C}_{\rho,h}$  a la clase de los conjuntos que satisfacen la condición (3.1).

**Definición 3.2.** Decimos que un conjunto  $S \subset \mathbb{R}^d$  compacto es  $\rho, h$ -cono convexo por complementos si existen  $\rho \in (0, \pi]$  y  $h > 0$ , tal que:

$$S = \bigcap_{\{x: C_{\rho,h}(x) \subset S^c\}} C_{\rho,h}(x)^c. \quad (3.2)$$

Denotaremos  $\tilde{\mathcal{C}}_{\rho,h}$  a la clase de los conjuntos que satisfacen la condición (3.2).

Es claro a partir de la definición de  $\mathcal{C}_{\rho,h}$  que dicha clase contiene a la clase de los conjuntos para los cuales rueda por su complemento una bola de radio  $r$ , tomando  $\rho$  y  $h$  adecuadamente. La Figura 3.1 muestra un conjunto que está en la clase  $\mathcal{C}_{\rho,h}$  para algún  $\rho$  y  $h$  pero en el cual no rueda libremente por su complemento una bola para ningún radio.

En la sección 3.5 vamos a considerar el caso en que los conos tienen altura infinita, introduciremos a continuación dichas definiciones:

**Definición 3.3.** Decimos que un conjunto  $S \subset \mathbb{R}^d$  compacto satisface la condición de cono convexidad exterior en sentido estricto si existe  $\rho \in (0, \pi]$  tal que

$$\forall x \in \partial S \exists \xi = \xi(x) \text{ tal que } C_{\rho,\xi}(x) \subset S^c. \quad (3.3)$$



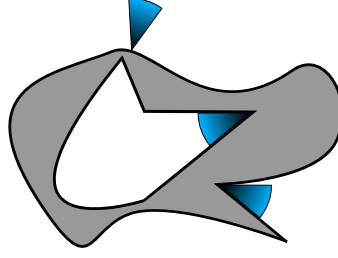


Figura 3.1: En gris un conjunto cono convexo en sentido amplio

Denotaremos  $\mathcal{C}_\rho$  a la familia de los conjuntos que satisfacen la condición (3.3). Dicha condición recuerda la propiedad clásica de hiperplano de apoyo de los conjuntos convexos. En este caso el semiespacio exterior correspondiente al hiperplano se reemplaza por un cono. En particular obtenemos que  $\mathcal{C}_\pi$  corresponde a la familia de los convexos.

Una vez definida una restricción de forma (por ejemplo convexidad,  $r$ -convexidad, rodamiento libre de una bola, etc) se define, para un conjunto cualquiera  $S$ , el *cierre*, esto es: la intersección de todos los conjuntos que incluyen a  $S$  y que cumplen dicha restricción. Claramente la restricción será útil para estimar el conjunto si el cierre también tiene la propiedad geométrica. En cuyo caso vamos a estimar el conjunto a partir de una muestra  $\aleph_n = X_1, \dots, X_n$  de puntos en  $S$  por medio del cierre de la muestra. A continuación introduciremos cuatro nociones de *cierre cono convexo* que generalizan en diferentes sentidos el clásico concepto de cierre convexo.

**Definición 3.4.** Definimos la *envolvente cono convexa en sentido amplio* de un conjunto  $S \subset \mathbb{R}^d$  acotado, que denotaremos  $\mathcal{C}_{\rho,h}(S)$ , como la intersección de todos los subconjuntos de  $\mathcal{C}_{\rho,h}$  que lo contienen. Definimos la *envolvente cono convexa en sentido estricto* de un conjunto  $S$  acotado, que denotaremos  $\mathcal{C}_\rho(S)$ , como la intersección de todos los subconjuntos de  $\mathcal{C}_\rho$  que lo contienen. Definimos la *envolvente por complemento en sentido amplio* de un conjunto  $S \subset \mathbb{R}^d$  acotado, que denotaremos:

$$\tilde{\mathcal{C}}_{\rho,h}(S) = \bigcap_{\{x: C_{\rho,h}(x) \subset S^c\}} C_{\rho,h}(x)^c,$$

como la intersección de los complementos de los conos de ángulo  $\rho$  y altura  $h$  con vértice en los puntos  $x$  tal que  $C_{\rho,h}(x) \subset S^c$ . Definimos la *envolvente por complemento en sentido estricto*, de un conjunto  $S \subset \mathbb{R}^d$  acotado, que denotaremos:

$$\tilde{\mathcal{C}}_\rho(S) = \bigcap_{\{x: C_\rho(x) \subset S^c\}} C_\rho(x)^c,$$

como la intersección de los complementos de los conos de de ángulo  $\rho$ , con vértice en los puntos  $x$  tal que  $C_\rho(x) \subset S^c$ .

**Observación 3.5.** Utilizaremos principalmente la *envolvente cono convexa en sentido amplio*. Observemos que en el caso en que  $S$  es convexo, las cuatro definiciones coinciden y son igual a la *envolvente convexa del conjunto*. En general la *envolvente por complemento* y la *envolvente cono convexa* no son iguales. El siguiente ejemplo ilustra esta situación: tomemos  $E = E_1 \cup E_2 \subset \mathbb{R}^2$  donde  $E_1$  es el gráfico de la función  $f(t) = t^2$  con  $t \in [0, 1]$  y  $E_2 = \{(t, 0) : 0 \leq t \leq 1\}$ . Dicho conjunto pertenece a  $\mathcal{C}_{\rho,h}$  para todo  $\rho \in (0, \pi/2]$ ,  $h > 0$  pero  $E \notin \tilde{\mathcal{C}}_{\rho,h}$  para ningún  $\rho$ . Otro ejemplo interesante se muestra en la figura 3.2, la misma está formada por el triángulo con vértices  $(1, 0)$ ,  $(1 + s, s)$  y  $(1 + s, -s)$  con  $s = (\sqrt{2} - 1)/2$  y los 7 triángulos congruentes a dicho triángulo, que se obtienen al aplicarle la rotación de centro en el origen y ángulo  $\pi/4$ . El conjunto es  $\pi/4$  cono convexo, es decir

coincide con su envolvente cono convexa en sentido amplio, mientras que  $(0,0) \notin \tilde{\mathcal{C}}_{\pi/4,h}(S)$  para  $h > 2$ . Esto se debe a que el ángulo entre los vértices es  $\pi/4$  y que los triángulos son tangentes al círculo de radio 1, por lo tanto si un cono de ángulo  $\pi/4$  y altura mayor que 2 contiene al origen corta a alguno de los vértices.

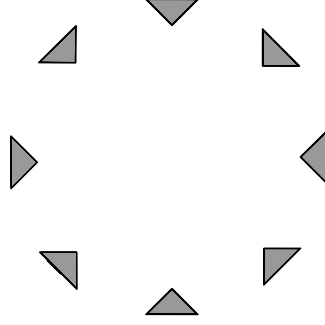


Figura 3.2: Conjunto  $\pi/4$  como convexo para el cual las envolventes cono convexas y por complemento no coinciden.

**Proposición 3.6.** Sea  $S \subset \mathbb{R}^d$  acotado, entonces

- $\mathcal{C}_{\rho,h}(S) \in \mathcal{C}_{\rho,h}$ .
- $\mathcal{C}_{\rho}(S) \in \mathcal{C}_{\rho}$ .
- $\tilde{\mathcal{C}}_{\rho,h}(S) \in \mathcal{C}_{\rho',h'}$  donde  $\rho' = (\pi - \rho)/2$  si  $\rho > \pi/3$  y  $\rho' = \rho$  si  $\rho \leq \pi/3$ ,  $h' = (h/2) \sin(\rho/2)$ .
- $\tilde{\mathcal{C}}_{\rho}(S) \in \mathcal{C}_{\rho}$ .
- $S \subset \mathcal{C}_{\rho',h'}(S) \subset \tilde{\mathcal{C}}_{\rho,h}(S) \subset \tilde{\mathcal{C}}_{\rho}(S) \subset \text{conv}(S)$ , donde  $\rho' = (\pi - \rho)/2$  si  $\rho > \pi/3$  y  $\rho' = \rho$  si  $\rho \leq \pi/3$ ,  $h' = (h/2) \sin(\rho/2)$ .

*Demostración.* a) Por definición tenemos que

$$\mathcal{C}_{\rho,h}(S) = \bigcap_{S \subset B_t, B_t \in \mathcal{C}_{\rho,h}} B_t.$$

Sea  $x \in \partial \mathcal{C}_{\rho,h}(S)$  queremos encontrar un cono que denotaremos  $C_{\rho,h}(x)$ , con vértice en  $x$  tal que  $C_{\rho,h}(x) \subset \mathcal{C}_{\rho,h}(S)^c$ . Tenemos que  $B(x, 1/n) \cap B_t \neq \emptyset$  para todo  $B_t$  tal que  $S \subset B_t$  y  $B_t \in \mathcal{C}_{\rho,h}$ , además  $B(x, 1/n) \cap B_n^c \neq \emptyset$  para algún  $B_n$  con  $S \subset B_n$  y  $B_n \in \mathcal{C}_{\rho,h}$ . Sea  $z_n \in B(x, 1/n) \cap \partial B_n$  tenemos que  $z_n \rightarrow x$ . Como  $z_n \in \partial B_n$  y  $B_n \in \mathcal{C}_{\rho,h}$  existe  $C_{\rho,\xi_n,h}(z_n) \subset B_n^c$ , un cono con vértice en  $z_n$  y altura  $h$ . Como  $\|\xi_n\| = 1$  existe  $\xi = \lim_n \xi_n$  (nuevamente pasando a subsucesiones) y  $z_n \rightarrow x$  tenemos que  $\overline{C_{\rho,\xi_n,h}(z_n)}$  converge en distancia de Hausdorff a  $\overline{C_{\rho,\xi,h}(x)}$ . Nos resta observar que  $C_{\rho,\xi,h}(x) \cap S = \emptyset$ . Supongamos que existe  $s \in S \cap C_{\rho,\xi,h}(x)$ , existe  $\delta > 0$  tal que

$$\left\langle \frac{s-x}{\|s-x\|}, \xi \right\rangle = \cos(\rho/2) + \delta,$$

entonces

$$\left\langle \frac{s-z_n}{\|s-z_n\|}, \xi_n \right\rangle > \cos(\rho/2),$$

para  $n > n_0$  para algún  $n_0$ , lo cual contradice que  $C_{\rho,\xi,h}(z_n) \cap S = \emptyset$ .

- Es análoga a a)

- c) Queremos encontrar  $\rho'$  y  $h'$  tal que para todo  $z \in \partial\tilde{\mathcal{C}}_{\rho,h}(S)$  existe un cono  $C_{\rho',h'}(z)$  con  $C_{\rho',h'}(z) \subset \tilde{\mathcal{C}}_{\rho,h}(S)^c$ . Tomemos  $z_n \rightarrow z$  con  $z_n \in \tilde{\mathcal{C}}_{\rho,h}(S)^c$ . Usando la definición de  $\tilde{\mathcal{C}}_{\rho,h}(S)$  existe una sucesión de conos  $C_{\rho,h}(u_n)$ , disjuntos con  $\tilde{\mathcal{C}}_{\rho,h}(S)$ , tal que  $z_n \in C_{\rho,h}(u_n)$ . Tomemos una subsucesión tal que  $u_n$  sea convergente, es decir,  $u_n \rightarrow u$ . Por construcción tenemos que  $z \in \partial C_{\rho,h}(u)$  y  $C_{\rho,h}(u) \cap \tilde{\mathcal{C}}_{\rho,h}(S) = \emptyset$ . Sólo resta probar que  $C_{\rho',h'}(z) \subset C_{\rho,h}(u)$ . Esto se sigue por los valores de  $\rho'$  y  $h'$  que hemos elegido (ver Figura 3.3).

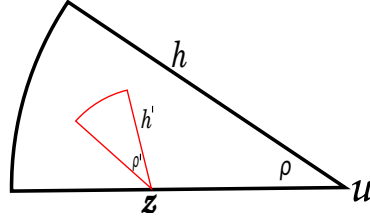


Figura 3.3: Para todo  $z \in \partial C_{\rho,h}(u)$  existe un cono  $C_{\rho',h'}(z) \subset C_{\rho,h}(u)$ .

- d) La demostración es similar a la demostración de la parte anterior. Tomemos  $s \in \partial\tilde{\mathcal{C}}_{\rho}(S)$ , queremos encontrar un cono  $C_{\rho}(s)$  con vértice en  $s$  y  $C_{\rho}(s) \subset \tilde{\mathcal{C}}_{\rho}(S)^c$ . Tenemos que  $B(s, 1/n) \cap C_{\rho}(x)^c \neq \emptyset$  para todo  $x$  tal que  $C_{\rho}(x) \subset S^c$ , además  $B(s, 1/n) \cap C_{\rho}(x_n) \neq \emptyset$  para algún  $x_n$  tal que  $S \subset C_{\rho}(x_n)^c$ . Existe por lo tanto  $z_n \in B(s, 1/n) \cap \partial C_{\rho}(x_n)$ . Tenemos que  $z_n \rightarrow_n s$ . Como  $z_n \in \partial C_{\rho}(x_n)$  existe  $\xi_n$  tal que el cono  $C_{\rho,\xi_n}(z_n)$  verifica que  $C_{\rho,\xi_n}(z_n) \subset C_{\rho}(x_n)$  (observemos que aquí hacemos uso del hecho de que los conos tienen *altura infinita*). El resto de la prueba es análogo a lo hecho en a).
- e) La primera, tercer y cuarta inclusión son inmediatas. La segunda es consecuencia inmediata de la parte c). □

**Ejemplo 3.7.** Como ejemplo sencillo en el cual la segunda inclusión de la Proposición 3.6 (e) no se cumple para el mismo  $\rho$  y  $h$  basta tomar  $\rho = \pi/4$ ,  $h = 1/4$  y el subconjunto de  $\mathbb{R}^2$ ,  $S = B((0,0), 1) \setminus C_{\pi/4, e_1, 1/4}((0,0))$ , siendo  $e_1 = (1,0)$ . En este caso  $\tilde{\mathcal{C}}_{\rho,h}(S) = S$  mientras que  $\mathcal{C}_{\rho,h}(S) = B(0,1)$ .

**Proposición 3.8.** Sea  $S \subset \mathbb{R}^d$  acotado y conexo, para todo  $\rho \in (0, \pi]$  y  $h > 0$   $\mathcal{C}_{\rho,h}(S)$  y  $\mathcal{C}_{\rho}(S)$  son conexos.

*Demostración.* Veamos que  $\mathcal{C}_{\rho,h}(S)$  es conexo. Supongamos por reducción al absurdo que existen  $U$  y  $V$  abiertos disjuntos, tal que  $\mathcal{C}_{\rho,h}(S) = (U \cap \mathcal{C}_{\rho,h}(S)) \cup (V \cap \mathcal{C}_{\rho,h}(S))$ . Como  $S$  es conexo, o bien  $S \subset U \cap \mathcal{C}_{\rho,h}(S)$  o, en caso contrario,  $S \subset V \cap \mathcal{C}_{\rho,h}(S)$ . Supongamos que  $S \subset U \cap \mathcal{C}_{\rho,h}(S)$ , como  $\mathcal{C}_{\rho,h}(S)$  es compacto y  $U$  y  $V$  son disjuntos, también lo es  $U \cap \mathcal{C}_{\rho,h}(S)$ . Veamos que  $U \cap \mathcal{C}_{\rho,h}(S) \in \mathcal{C}_{\rho,h}$ . Sea  $x \in \partial(U \cap \mathcal{C}_{\rho,h}(S))$  probemos primero que  $x \in \partial\mathcal{C}_{\rho,h}(S)$ : para todo  $\delta > 0$   $B(x, \delta) \cap (U \cap \mathcal{C}_{\rho,h}(S)) \neq \emptyset$  y  $B(x, \delta) \cap (U^c \cup \mathcal{C}_{\rho,h}(S)^c) \neq \emptyset$ . Es decir  $B(x, \delta) \cap \mathcal{C}_{\rho,h}(S) \neq \emptyset$ , si  $B(x, \delta) \cap \mathcal{C}_{\rho,h}(S)^c = \emptyset$  (y por lo tanto  $B(x, \delta) \cap U^c \neq \emptyset$ ) entonces  $B(x, \delta) \subset \mathcal{C}_{\rho,h}(S)$ , como  $B(x, \delta)$  es conexas tiene que ser  $B(x, \delta) \subset U$  que contradice que  $B(x, \delta) \cap U^c \neq \emptyset$ . Como  $x \in \partial\mathcal{C}_{\rho,h}(S)$  existe  $\xi$  de norma 1 tal que  $C_{\rho,\xi,h}(x) \subset \mathcal{C}_{\rho,h}(S)^c \subset (U \cap \mathcal{C}_{\rho,h}(S))^c$ . Como  $U \cap \mathcal{C}_{\rho,h}(S) \in \mathcal{C}_{\rho,h}$  tiene que ser  $\mathcal{C}_{\rho,h}(S) \subset U \cap \mathcal{C}_{\rho,h}(S) \subset \mathcal{C}_{\rho,h}(S)$ , de donde se sigue que  $V \cap \mathcal{C}_{\rho,h}(S) = \emptyset$  y por lo tanto  $\mathcal{C}_{\rho,h}(S)$  es conexo. □

**Proposición 3.9.** Sea  $S \subset \mathbb{R}^d$  acotado, tal que  $S = \overline{\text{int}(S)}$  entonces, para todo  $\rho \in (0, \pi]$  y  $h > 0$ ,

$$\mathcal{C}_{\rho,h}(S) = \overline{\text{int}(\mathcal{C}_{\rho,h}(S))}$$

y

$$\mathcal{C}_\rho(S) = \overline{\text{int}(\mathcal{C}_\rho(S))}.$$

*Demostración.* Vamos a demostrar la primer igualdad, la segunda es análoga. Como  $\mathcal{C}_{\rho,h}(S)$  es cerrado tenemos que  $\overline{\text{int}(\mathcal{C}_{\rho,h}(S))} \subset \mathcal{C}_{\rho,h}(S)$  por lo tanto nos resta demostrar que  $\mathcal{C}_{\rho,h}(S) \subset \overline{\text{int}(\mathcal{C}_{\rho,h}(S))}$ . Como  $S = \overline{\text{int}(S)} \subset \overline{\mathcal{C}_{\rho,h}(S)}$  y  $\overline{\text{int}(S)} \subset \overline{\text{int}(\mathcal{C}_{\rho,h}(S))}$ , para demostrar que  $\mathcal{C}_{\rho,h}(S) \subset \overline{\text{int}(\mathcal{C}_{\rho,h}(S))}$  es suficiente ver que  $\overline{\text{int}(\mathcal{C}_{\rho,h}(S))} \in \mathcal{C}_{\rho,h}$ .

Sea  $x \in \overline{\text{int}(\mathcal{C}_{\rho,h}(S))}$  tenemos que encontrar  $\xi$  de norma 1 tal que  $C_{\rho,\xi,h}(x) \subset \overline{\text{int}(\mathcal{C}_{\rho,h}(S))}^c$ . Veamos que  $x \in \partial\mathcal{C}_{\rho,h}(S)$ . Para todo  $\delta > 0$  tenemos que  $B(x,\delta) \cap \overline{\text{int}(\mathcal{C}_{\rho,h}(S))} \neq \emptyset$  y  $B(x,\delta) \cap \overline{\text{int}(\mathcal{C}_{\rho,h}(S))}^c \neq \emptyset$ . Sea  $z \in B(x,\delta) \cap \overline{\text{int}(\mathcal{C}_{\rho,h}(S))}$ .  $B(z,\delta - d(z,\partial B(x,\delta))) \cap \overline{\text{int}(\mathcal{C}_{\rho,h}(S))}$  entonces  $B(x,\delta) \cap \mathcal{C}_{\rho,h}(S) \neq \emptyset$ . Sea  $y \in B(x,\delta) \cap \overline{\text{int}(\mathcal{C}_{\rho,h}(S))}^c$ , tenemos que o bien  $y \in \mathcal{C}_{\rho,h}(S)^c$  en cuyo caso  $x \in \partial\mathcal{C}_{\rho,h}(S)$ , o bien  $y \in \partial\mathcal{C}_{\rho,h}(S)$  en cuyo caso  $B(y,\delta - d(y,\partial B(x,\delta))) \cap \mathcal{C}_{\rho,h}(S)^c \neq \emptyset$  de donde  $B(x,\delta) \cap \mathcal{C}_{\rho,h}(S)^c \neq \emptyset$ . Como  $x \in \partial\mathcal{C}_{\rho,h}(S)$  y, por la Proposición 3.6 a)  $\mathcal{C}_{\rho,h}(S) \in \mathcal{C}_{\rho,h}$  podemos tomar  $\xi$  de norma 1 tal que  $C_{\rho,\xi,h}(x) \subset \mathcal{C}_{\rho,h}(S)^c \subset \overline{\text{int}(\mathcal{C}_{\rho,h}(S))}^c$ .  $\square$

### 3.1.3. Propiedades de convergencia

**Teorema 3.10.** *Sea  $\{S_n\} \subset \mathcal{C}_{\rho,h}$  y  $S$  un subconjunto compacto, tal que  $d_H(S_n, S) \rightarrow 0$ . Entonces*

$$d_H(\partial S_n, \partial S) \rightarrow 0.$$

*Demostración.* Supongamos que el resultado no es cierto, entonces:

- (i) existe  $\varepsilon > 0$  tal que para infinitos  $n \in \mathbb{N}$  existe  $x_n \in \partial S$ , tal que  $d(x_n, \partial S_n) > \varepsilon$ , o
- (ii) existe  $\varepsilon > 0$  tal que para infinitos  $n \in \mathbb{N}$  existe  $x_n \in \partial S_n$  tal que  $d(x_n, \partial S) > \varepsilon$ .

Supongamos primero que se cumple (i). Como  $S$  es compacto existe una subsucesión de  $\{x_n\}$  que converge, la cual seguiremos denotando  $x_n$ , sea  $x \in \partial S$  el límite de dicha sucesión. Como  $S_n$  converge a  $S$  en distancia de Hausdorff tenemos que a partir de un cierto  $n$ ,  $d(x, S_n) \leq \varepsilon/2$ , además para todo  $y \in \partial S_n$ , usando la desigualdad triangular  $\|x_n - y\| \leq \|x_n - x\| + \|x - y\|$ , y tomando ínfimo en  $y \in \partial S_n$

$$d(x, \partial S_n) \geq d(x_n, S_n) - \|x - x_n\| > \varepsilon/2.$$

Dado que  $d(x, S_n) \leq \varepsilon/2$  y  $d(x, \partial S_n) > \varepsilon/2$  tenemos que  $x \in \text{int}(S_n)$  para infinitos  $n$ , y  $B(x, \varepsilon/2) \cap S_n^c = \emptyset$ . Como  $x \in \partial S$  y  $S$  es cerrado podemos tomar  $y \in S^c$  tal que  $\|x - y\| < \varepsilon/2$  y  $d(y, S) > 0$ . Sabemos que  $S_n$  converge a  $S$  en distancia de Hausdorff por lo tanto  $y \in S_n^c$  a partir de un cierto  $n$ , lo cual contradice que  $B(x, \varepsilon/2) \cap S_n^c = \emptyset$  para infinitos  $n$ .

Supongamos que se cumple (ii). La convergencia en distancia de Hausdorff de  $S_n$  a  $S$  implica que  $d(x_n, S) < \varepsilon$  a partir de un cierto  $n$ , como además  $d(x_n, \partial S) > \varepsilon$ , obtenemos (razonando de forma análoga que en (i)) que  $x_n \in \text{int}(S)$  para infinitos  $n$ . Tomemos una subsucesión convergente de  $x_n$  que seguiremos denotando  $x_n$ , tenemos que  $x_n \rightarrow x \in S$  y  $d(x, \partial S) > \varepsilon/2$ , es decir  $B(x, \varepsilon/2) \subset S$ . Observemos que, como  $x_n \in \partial S_n$  y  $S_n \in \mathcal{C}_{\rho,h}$  existen conos  $C_{\rho,h}(x_n)$  con vértice en  $x_n$  y  $C_{\rho,h}(x_n) \cap S_n = \emptyset$ . Tomemos  $c_n \in S_n^c$  puntos en el eje de los conos  $C_{\rho,h}(x_n)$  tal que  $0 < \|x_n - c_n\| = \min\{h/2, \varepsilon/4\} = k$  con  $k$  constante. Podemos suponer pasando nuevamente a subsucesiones que  $c_n \rightarrow c$  y  $\|c - x\| \leq \varepsilon/4$ . Por construcción  $d(c_n, S_n) \geq k \sin(\rho/2)$ , por lo tanto, a partir de un cierto  $n$ ,  $d(c, S_n) > k'$ , con  $k' > 0$  constante. Como  $S_n$  converge a  $S$  en distancia de Hausdorff tiene que ser  $c \in S^c$ , lo cual contradice que  $B(x, \varepsilon/2) \subset S$ .  $\square$

**Teorema 3.11.** *Sea  $\{S_n\} \subset \mathcal{C}_{\rho,h}$  y  $S$  compacto, tal que  $d_H(S_n, S) \rightarrow 0$ . Entonces  $S \in \mathcal{C}_{\rho,h}$ .*

*Demostración.* Sea  $x \in \partial S$ , queremos encontrar un cono  $C_{\rho,h}(x) \subset S^c$ . Sabemos que, por el teorema anterior  $d_H(\partial S_n, \partial S) \rightarrow 0$  por lo tanto podemos suponer que existe  $x_n \in \partial S_n$  tal que  $x_n \rightarrow x$ , ahora el razonamiento para encontrar el cono  $C_{\rho,h}(x)$  es análogo al que hicimos en la prueba de la Proposición 3.6 a). Consideramos nuevamente la sucesión acotada  $\{\overline{C_{\rho,\xi_n,h}(x_n)}\}$ , dicha sucesión converge en distancia de Hausdorff al cono cerrado  $\overline{C_{\rho,\xi,h}(x)}$  para algún  $\xi$  tal que  $\|\xi\| = 1$ . Veamos que  $C_{\rho,\xi,h}(x) \cap S = \emptyset$ . Supongamos por absurdo que existe  $z \in C_{\rho,\xi,h}(x) \cap S$ , como  $\xi_n \rightarrow \xi$  y  $x_n \rightarrow x$  podemos tomar  $\varepsilon > 0$  tal que para  $n$  suficientemente grande,  $B(z, \varepsilon) \subset C_{\rho,\xi,h}(x_n)$  de donde  $d(z, S_n) \geq \varepsilon$  para infinitos  $n$ , lo cual contradice que  $S_n$  tiende a  $S$  en distancia de Hausdorff.  $\square$

## 3.2. Los conjuntos cono convexos son una clase de Glivenko-Cantelli

Consideremos  $X_1, \dots, X_n, \dots$  una sucesión de variables aleatorias independientes e idénticamente distribuidas, definidas en el espacio de probabilidad  $(\Omega, \mathfrak{F}, \mathbb{P})$  a valores en un espacio de medida  $(E, \mathcal{B})$ . Denotemos como  $P$  la distribución de  $X_1$  en  $\mathcal{B}$  y  $\mathbb{P}_n$  a la distribución empírica de  $P$  asociada a  $X_1, \dots, X_n$ . Decimos que una clase  $\mathcal{A} \subset \mathcal{B}$  es una clase de Glivenko-Cantelli, si:

$$\sup_{A \in \mathcal{A}} |\mathbb{P}_n(A) - P(A)| \rightarrow 0, \quad c.s. \quad (3.4)$$

El estudio de las clases de Glivenko-Cantelli es un tema clásico en la teoría de los procesos empíricos. Un estudio detallado de la misma y sus aplicaciones en estadística puede encontrarse en Shorack y A (1986) y van der Vaart (1998) cap. 19. Por ejemplo, los Teoremas 19.4 y 19.13 en van der Vaart (1998) dan condiciones suficientes para que una clase sea de Glivenko-Cantelli. Dichas condiciones son expresadas en términos de restricciones en la entropía de la clase, la cual cuantifica el “tamaño” de la clase  $\mathcal{A}$ . Otra forma clásica de demostrar que una clase es de Glivenko-Cantelli es a través de la teoría de Vapnik-Cervonenkis. En esta teoría se obtienen cotas exponenciales para la probabilidad  $\mathbb{P}\{\sup_{A \in \mathcal{A}} |\mathbb{P}_n(A) - P(A)| > \epsilon\}$ . Si la serie (en  $n$ ) de dichas cotas es convergente, entonces, usando el Lema de Borel-Cantelli se obtiene (3.4). Las cotas requieren que la clase  $\mathcal{A}$  tenga dimensión de Vapnik-Cervonenkis finita (ver Definición 12.2 en Devroye y otros (1996)). Se puede ver, por ejemplo, que la clase de los subconjuntos cerrados, convexos, de  $\mathbb{R}^d$  no tienen dimensión de Vapnik-Chervonenkis finita, por lo tanto no es posible aplicar dicha teoría a esa clase. En este caso se usa el enfoque dado en Talagrand (1987)

En Cuevas y otros (2012) se demuestra que la familia de los subconjuntos no vacíos de un subconjunto compacto, con alcance mayor o igual que un cierto  $r$ , es una clase de Glivenko-Cantelli. En esta sección demostraremos ese resultado para la familia de los subconjuntos cerrados que verifican la condición (3.1). Es decir vamos a ampliar el resultado a una clase de conjuntos más grande (observemos que hemos quitado la restricción de que los conjuntos estén incluidos en un compacto) que incluye, como dijimos anteriormente, a la familia de subconjuntos compactos para los cuales rueda libremente una bola de radio  $r$  para algún  $r$ . Para eso vamos a introducir algunos conceptos.

**Definición 3.12.** Sean  $P_n$  y  $P$  medidas de probabilidad en  $\mathcal{B}$ , un conjunto  $A \in \mathcal{B}$  se dice  $P$ -continuo si  $P(\partial A) = 0$ . Una subclase  $\mathcal{A} \subset \mathcal{B}$  se dice que es una clase  $P$ -continua si todo conjunto  $A \in \mathcal{A}$  es  $P$ -continuo.

La siguiente Proposición prueba que la clase de los conjuntos cono convexos es una clase  $P$ -continua, para toda  $P$  absolutamente continua respecto de la medida de Lebesgue.

**Proposición 3.13.** Si  $S \in \mathcal{C}_{\rho,h}$  entonces  $\mu(\partial S) = 0$ , donde  $\mu$  denota la medida de Lebesgue en  $\mathbb{R}^d$ .

*Demostración.* Observemos primero que  $\partial S$  es medible por ser  $S$  cerrado. Para demostrar la proposición recordemos previamente el concepto de densidad métrica de un punto (véase por ejemplo Erdos (1945)): decimos que un punto  $x \in S$  tiene densidad métrica 1 si para todo  $\varepsilon > 0$  existe  $\delta > 0$

tal que  $\mu(S \cap B(x, r)) > (1 - \varepsilon)\mu(B(x, r))$  para todo  $r < \delta$ . En virtud del Corolario 2.9 en Morgan (2000) todo conjunto de medida positiva tiene al menos un punto de densidad métrica uno; pero esto último no es posible ya que para todo  $x \in \partial S$  existe  $C_{\rho, h}(x) \subset S^c$  un cono con vértice en  $x$  y por lo tanto

$$\mu(\partial S \cap B(x, r)) \leq \mu(S \cap B(x, r)) = \mu(S \cap C_{\rho, h}(x)^c \cap B(x, r)) \leq k_{\rho, h}\mu(B(x, r)),$$

para todo  $r < h$ , con  $k_{\rho, h} < 1$ . □

Recordemos que una sucesión  $P_n$  converge débilmente a  $P$  si

$$P_n(A) \rightarrow P(A), \quad c.s.$$

para todo conjunto  $A \in \mathcal{B}$ ,  $P$ -continuo. Ahora estamos en condiciones de definir el concepto fundamental de esta sección. El mismo es una generalización del de clase de Glivenko-Cantelli, donde sustituimos la distribución empírica por una distribución cualquiera que converja débilmente a  $P$ :

**Definición 3.14.** Decimos que una subclase  $\mathcal{A} \subset \mathcal{B}$  es una clase  $P$ -uniforme si

$$\sup_{A \in \mathcal{A}} |P_n(A) - P(A)| \rightarrow 0,$$

para toda sucesión  $P_n$  que converja débilmente a  $P$ .

El siguiente teorema, debido a Billingsley y Topsøe (ver Billingsley y Topsøe (1967)), establece condiciones sobre la subclase  $\mathcal{A}$  para que sea una clase  $P$ -uniforme. Antes de dar su enunciado vamos a introducir algo de notación: dado un espacio métrico  $(E, d)$ , denotamos  $\mathcal{M}(E)$  a la familia de subconjuntos compactos no vacíos de  $E$ , dotados con la métrica de Hausdorff. Observemos que  $\mathcal{M}(E)$  es localmente compacto.

**Teorema 3.15.** Si  $E$  es localmente conexo y  $\mathcal{A}$  es una clase  $P$ -continua de subconjuntos de  $E$ , cualquiera de las siguientes condiciones implica que  $\mathcal{A}$  es una clase  $P$ -uniforme:

- (i) La clase  $\partial\mathcal{A} = \{\partial A : A \in \mathcal{A}\}$  es un subconjunto compacto del espacio  $\mathcal{M}(E)$ .
- (ii) Existe una sucesión  $\{C_n\}$  de conjuntos acotados, con  $P(\text{int}(C_n)) \rightarrow 1$ , tal que para cada  $n$ , la clase  $\partial(C_n \cap \mathcal{A}) = \{\partial(C_n \cap A) : A \in \mathcal{A}\}$  es un subconjunto compacto de  $\mathcal{M}(E)$ .
- (iii) Existe una sucesión  $\{C_n\}$  de conjuntos cerrados, acotados, con  $P(\text{int}(C_n)) \rightarrow 1$ , tal que para cada  $n$ , la clase  $C_n \cap \partial\mathcal{A} = \{C_n \cap \partial A : A \in \mathcal{A}\}$  es un subconjunto compacto de  $\mathcal{M}(E)$ .

Podemos ahora, utilizando los Teoremas 3.10 y 3.11 demostrar que, fijado  $h$  y  $\rho$ , la familia de subconjuntos cerrados que verifican la condición de cono convexidad finita son una clase  $P$ -uniforme.

**Teorema 3.16.** Sea  $\rho \in (0, \pi]$  y  $h > 0$ . Sea  $P$  una medida de probabilidad absolutamente continua respecto de la medida de Lebesgue  $\mu$ . La clase  $\mathcal{C}_{\rho, h}^U$  de subconjuntos cerrados no vacíos que satisfacen la condición de  $\rho, h$  cono convexidad (3.1) es una clase  $P$ -uniforme (y en particular una clase de Glivenko-Cantelli)

*Demostración.* Veremos primero que el resultado es cierto si tomamos  $\mathcal{A} = \{A \subset K : A \neq \emptyset, A \in \mathcal{C}_{\rho, h}\}$  donde  $K$  es compacto y luego lo extenderemos a la familia  $\mathcal{C}_{\rho, h}^U$  del enunciado. Observemos que, por la Proposición 3.13 las familias  $\mathcal{A}$  y  $\mathcal{C}_{\rho, h}$  son  $P$ -continuas. Consideremos  $\{\partial A_n\}$  una sucesión en  $\partial\mathcal{A}$  donde  $\partial\mathcal{A} = \{\partial A : A \in \mathcal{A}\}$ , existe (ya que  $\mathcal{M}(E)$  es localmente compacto) una subsucesión  $A_{n_k}$  y  $A \in \mathcal{M}(E)$  tal que  $d_H(A_{n_k}, A) \rightarrow 0$ . Por el Teorema 3.11  $A \in \mathcal{C}_{\rho, h}$ , por lo tanto  $A \in \mathcal{A}$ . Por el Teorema 3.10 sabemos que  $d_H(\partial A_{n_k}, \partial A) \rightarrow 0$ . Tenemos entonces que  $\partial\mathcal{A}$  es compacta en  $\mathcal{M}(E)$ . De la parte (i) del Teorema 3.15 se sigue que es  $P$ -uniforme. Veamos ahora como extender el resultado

a la familia  $\mathcal{C}_{\rho,h}^U$ . Dado  $\varepsilon > 0$  sea  $R > 0$  tal que  $P(B(0, R)^c) < \varepsilon/8$ , denotemos  $K = \overline{B(0, R)}$ , como  $P_n$  converge débilmente a  $P$  existe  $n_0$  tal que  $P_n(K^c) < \varepsilon/4$  para  $n > n_0$ . Tenemos que

$$\begin{aligned} \sup_{A \in \mathcal{C}_{\rho,h}^U} |P_n(A) - P(A)| &\leq \sup_{A \in \mathcal{C}_{\rho,h}^U} |P_n(A \cap K) - P(A \cap K)| + \sup_{A \in \mathcal{C}_{\rho,h}^U} |P_n(A \cap K^c) - P(A \cap K^c)| \\ &\leq \sup_{A \in \mathcal{C}_{\rho,h}^U} |P_n(A \cap K) - P(A \cap K)| + P_n(K^c) + P(K^c) < \varepsilon, \end{aligned}$$

para  $n$  es suficientemente grande. En la última desigualdad hemos usado que si  $A \cap K \neq \emptyset$  entonces  $A \cap K \in \mathcal{A}$ .  $\square$

Observemos que en virtud de la Proposición 3.6 c), el mismo resultado es válido para la clase de conjuntos que satisfacen la condición 3.2, como establece el siguiente Teorema.

**Teorema 3.17.** *Sea  $\rho \in (0, \pi)$  y  $h > 0$ . Sea  $P$  una medida de probabilidad absolutamente continua respecto de la medida de Lebesgue  $\mu$ . Entonces la clase  $\tilde{\mathcal{C}}_{\rho,h}^U$  de subconjuntos cerrados no vacíos que satisfacen la condición de  $\rho, h$  como convexidad (3.2) es una clase  $P$ -uniforme (y en particular una clase de Glivenko-Cantelli).*

### 3.3. Estimación de conjuntos cono convexos

#### 3.3.1. Tasas de convergencia para la distancia de Hausdorff

En la sección anterior estudiamos algunas propiedades de los conjuntos que verifican la condición de cono convexidad exterior (finita e infinita); en la presente supondremos que tenemos una muestra  $X_1, \dots, X_n$  de variables i.i.d con distribución  $P_X$  con soporte  $S$  y veremos que las envolventes cono convexas de la muestra, definidas anteriormente, son consistentes en distancia de Hausdorff y en medida.

**Teorema 3.18.** *Sea  $S \subset \mathbb{R}^d$  y  $X_1, X_2, \dots, X_n, \dots$  i.i.d. de  $P_X$  con soporte  $S$ . Supongamos que  $S$  es estandar respecto a  $P_X$ . Si denotamos  $\aleph_n = \{X_1, X_2, \dots, X_n\}$ , entonces*

a) Si  $S \in \mathcal{C}_{\rho,h}$  entonces

$$d_H(\mathcal{C}_{\rho,h}(\aleph_n), S) = \mathcal{O} \left( \left( \frac{\log(n)}{n} \right)^{1/d} \right) \quad c.s..$$

b) Si  $S \in \tilde{\mathcal{C}}_{\rho,h}$  entonces

$$d_H(\tilde{\mathcal{C}}_{\rho,h}(\aleph_n), S) = \mathcal{O} \left( \left( \frac{\log(n)}{n} \right)^{1/d} \right) \quad c.s..$$

*Demostración.* Haremos la prueba de a) ya que b) es análogo. Para a) observemos que  $\aleph_n \subset \mathcal{C}_{\rho,h}(\aleph_n) \subset \mathcal{C}_{\rho,h}(S) = S$  por lo tanto el resultado es consecuencia del siguiente Teorema (cuya demostración se encuentra en Cuevas y Rodríguez-Casal (2004)).  $\square$

**Teorema 3.19.** *Sea  $X_1, X_2, \dots$  una sucesión de vectores aleatorios i.i.d. en  $\mathbb{R}^d$  con distribución común  $P_X$ . Sea  $S$  el soporte de  $P_X$ . Supongamos que  $S$  es compacto y estándar con respecto a  $P_X$ . Se verifica que*

$$\limsup_{n \rightarrow \infty} \left( \frac{n}{\log(n)} \right)^{\frac{1}{d}} d_H(\aleph_n, S) \leq \left( \frac{2}{\delta \omega_d} \right)^{\frac{1}{d}} \quad c.s.,$$

donde  $\omega_d$  es el volumen de la bola unidad  $d$ -dimensional y  $\delta$  es la constante de estandaridad.

En dicho trabajo se demuestra también

**Teorema 3.20.** Para  $S = [0, 1]^d$ ,

$$\liminf_{n \rightarrow \infty} \left( \frac{n}{\log(n)} \right)^{\frac{1}{d}} d_H(\aleph_n, S) \geq \left( \frac{1}{2^d \omega_d} \right)^{\frac{1}{d}} \quad c.s.$$

### 3.3.2. Tasas de convergencia en la distancia en medida

En esta sección hallaremos el orden de convergencia de  $d_\mu(\mathcal{C}_{\rho,h}(\aleph_n), S)$ , para el caso en que  $S \in \mathcal{C}_{\rho,h}$  donde  $\mu$  es la medida de Lebesgue en  $\mathbb{R}^d$ . Recordemos la notación:  $S \ominus tB(0, 1) = \{x \in S : B(x, t) \subset S\}$ . El conjunto  $S \ominus tB(0, 1)$  se denomina conjunto paralelo interior de  $S$  a distancia  $t$ . Es fácil ver que  $A \ominus tB(0, 1) = (A^c \oplus tB(0, 1))^c$ ; un estudio detallado de las operaciones de Minkowski puede encontrarse en Serra (1982). Contrario a lo que sucede con el conjunto paralelo exterior, como se señala en el trabajo de Saorín (2008), aún en el caso en el que el conjunto original es convexo, no existe una fórmula precisa (por ejemplo la ya mencionada formula de Steiner) para el volumen del conjunto paralelo interior.

**Teorema 3.21.** Sea  $S \subset \mathbb{R}^d$  en las hipótesis del Teorema 3.18, supongamos que además se cumple

$$P_X(S \setminus S \ominus tB(0, 1)) = \mathcal{O}(t), \quad (3.5)$$

entonces

a) Si  $S \in \mathcal{C}_{\rho,h}(S)$

$$d_{P_X}(\mathcal{C}_{\rho,h}(\aleph_n), S) = \mathcal{O}\left(\left(\frac{\log(n)}{n}\right)^{1/d}\right) \quad c.s..$$

b) Si  $S \in \tilde{\mathcal{C}}_{\rho,h}(S)$

$$d_{P_X}(\tilde{\mathcal{C}}_{\rho,h}(\aleph_n), S) = \mathcal{O}\left(\left(\frac{\log(n)}{n}\right)^{1/d}\right) \quad c.s..$$

*Demostración.* a) Observemos que

$$d_{P_X}(\mathcal{C}_{\rho,h}(\aleph_n), S) = P_X(S \setminus \mathcal{C}_{\rho,h}(\aleph_n)),$$

de donde se sigue que es suficiente ver, denotando  $\varepsilon_n = d_H(\aleph_n, S)$ , que existe  $k \in \mathbb{R}$  tal que:

$$S \ominus k\varepsilon_n B(0, 1) \subset \mathcal{C}_{\rho,h}(\aleph_n) \quad \text{para } n \text{ suficientemente grande } c.s., \quad (3.6)$$

ya que en este caso:

$$P_X(S \setminus \mathcal{C}_{\rho,h}(\aleph_n)) \leq P_X(S \setminus (S \ominus k\varepsilon_n B(0, 1))) = \mathcal{O}(k\varepsilon_n) = \mathcal{O}\left(\left(\frac{\log(n)}{n}\right)^{1/d}\right) \quad c.s.,$$

donde en la última igualdad hemos usado el Teorema 3.19. Veamos que (3.6) se cumple con  $k = 3 + 2/\sin(\rho/2)$ . Tomemos  $n_0 = n_0(\omega)$  tal que si  $n > n_0$ ,  $2\varepsilon_n/\sin(\rho/2) < h/2$ , supongamos por reducción al absurdo que existe  $x_n$  una sucesión,  $x_n \in S \ominus k\varepsilon_n B(0, 1)$  tal que  $x_n \notin \mathcal{C}_{\rho,h}(\aleph_n)$ , es decir, podemos tomar  $B_n \in \mathcal{C}_{\rho,h}$  con  $\aleph_n \subset B_n$  y  $x_n \notin B_n$ . Por la forma en que fue tomado  $\varepsilon_n$  sabemos que existe  $X_i$  tal que  $X_i \in \overline{B(x_n, \varepsilon_n)}$ . Como  $\aleph_n \subset B_n$  tenemos que  $\overline{B(x_n, \varepsilon_n)} \cap B_n \neq \emptyset$  y  $\overline{B(x_n, \varepsilon_n)} \cap B_n^c \neq \emptyset$ , de donde se sigue que existe  $z_n \in \partial B_n$  y  $z_n \in \overline{B(x_n, \varepsilon_n)}$ . Dado que  $B_n \in \mathcal{C}_{\rho,h}$  podemos tomar  $\xi_n$  de norma 1 y  $C_{\rho,\xi_n,h}(z_n) \subset B_n^c$  un cono con vértice en  $z_n$ , por lo tanto  $C_{\rho,\xi_n,h}(z_n) \cap \aleph_n = \emptyset$ . Consideremos

$$y_n = z_n + \frac{2\varepsilon_n}{\sin(\rho/2)} \xi_n,$$



como se muestra en la Figura 3.4. Basta probar que  $B(y_n, 2\varepsilon_n) \subset S$ , lo cual contradice que  $d_H(S, \mathfrak{N}_n) = \varepsilon_n$  ya que  $B(y_n, 2\varepsilon_n) \cap \mathfrak{N}_n = \emptyset$ . Por la forma en que fue tomado  $y_n$  es fácil ver que  $B(y_n, 2\varepsilon_n) \subset C_{\rho, \xi_n, h}(z_n)$ . Sabemos que  $B(x_n, k\varepsilon_n) \subset S$ , por lo tanto basta ver que  $B(y_n, 2\varepsilon_n) \subset B(x_n, k\varepsilon_n)$  pero esto es inmediato dado que, si  $t \in B(y_n, 2\varepsilon_n)$ , entonces  $\|t - x_n\| \leq \|t - y_n\| + \|y_n - z_n\| + \|z_n - x_n\| \leq 2\varepsilon_n + 2\varepsilon_n/\text{sen}(\rho/2) + \varepsilon_n = \varepsilon_n(3 + 2/\text{sen}(\rho/2))$ .

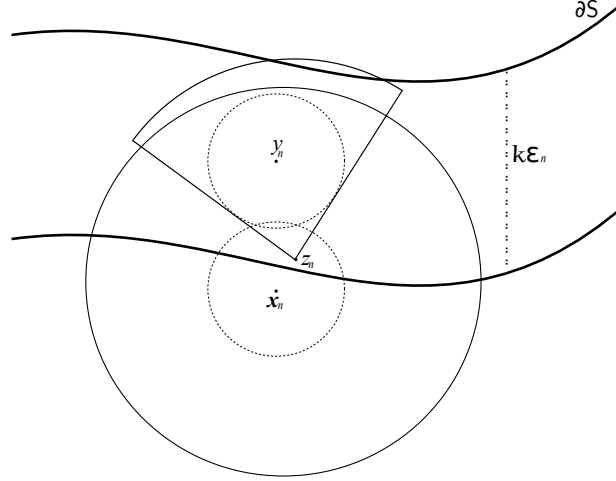


Figura 3.4:  $B(y_n, \varepsilon_n) \subset C_{\rho, \xi_n, h}(z_n)$

- b) El resultado se sigue de la parte a) y de la Proposición 3.6 d) ya que  $\tilde{\mathcal{C}}_{\rho, h}(\mathfrak{N}_n) \in \mathcal{C}_{\rho', h'}$ , y por lo tanto

$$\mathcal{C}_{\rho', h'}(\mathfrak{N}_n) \subset \tilde{\mathcal{C}}_{\rho, h}(\mathfrak{N}_n).$$

□

**Observación 3.22.** En el caso en que  $P_X$  es la distribución uniforme con soporte en  $S$  la hipótesis (3.5) se cumple por ejemplo cuando existe el contenido de Minkowski exterior de  $S^c$ . El orden que se obtiene en el teorema anterior es el mismo orden de convergencia que se obtiene para los conjuntos  $r$ -convexos aunque en este caso el orden se puede mejorar si se impone además que  $S^c$  también sea  $r$ -convexo (ver Rodríguez-Casal (2006)).

En virtud de (3.6) tenemos el siguiente resultado:

**Proposición 3.23.** Si  $S$  está en las hipótesis del Teorema 3.21 entonces

$$d_H(\partial S, \partial \mathcal{C}_{\rho, h}(\mathfrak{N}_n)) \leq kd_H(S, \mathfrak{N}_n) \quad \text{para } n \text{ suficientemente grande,}$$

donde

$$k = (3 + 2/\text{sen}(\rho/2)).$$

*Demostración.* La tesis se sigue de (3.6), de  $\partial \mathcal{C}_{\rho, h}(\mathfrak{N}_n) \subset S \setminus \text{int}(S \ominus k\varepsilon_n B(0, 1))$ , y de que

$$\partial(S \ominus k\varepsilon_n B(0, 1)) \subset \left\{ x \in \mathbb{R}^d : d(x, \partial \mathcal{C}_{\rho, h}(S)) = k\varepsilon_n \right\}.$$

□

Para construir el estimador  $\mathcal{C}_{\rho, h}(\mathfrak{N}_n)$  es necesario conocer los valores reales de  $\rho$  y  $h$ , si bien más adelante veremos un test para saber si estamos tomando un  $\rho$  mayor que el real (en el caso cono convexo en sentido estricto), el siguiente teorema nos permite tomar  $\rho_n \rightarrow 0$ . Recordemos que si  $S \in \mathcal{C}_{\rho, h}$  entonces  $S \in \mathcal{C}_{\rho_1, h_1}$  si  $0 < \rho_1 < \rho$  y  $h_1 < h$ .

**Teorema 3.24.** *Sea  $S \subset \mathbb{R}^d$ ,  $S \in \mathcal{C}_{\rho,h}$ , sea  $\aleph_n = X_1, \dots, X_n$  i.i.d con distribución  $P_X$  con soporte  $S$ , supongamos que  $P_X$  verifica la condición 3.5. Sean  $\rho_n, h_n$  sucesiones decrecientes de números reales tal que*

$$\rho_n, h_n \rightarrow 0, \quad \frac{d_H(S, \aleph_n)}{\rho_n} \rightarrow 0 \quad \text{y} \quad \frac{d_H(S, \aleph_n)}{h_n} \rightarrow 0,$$

entonces

$$d_{P_X}(\mathcal{C}_{\rho_n, h_n}(\aleph_n), S) = \mathcal{O}(d_H(S, \aleph_n)/\rho_n) \quad \text{c.s.}$$

*Demostración.* Como  $S \in \mathcal{C}_{\rho,h}$ ,  $\mathcal{C}_{\rho_n, h_n}(\aleph_n) \subset S$  para  $n > n_0$  con  $n_0$  tal que  $\rho_{n_0} \leq \rho$  y  $h_{n_0} < h$ . Observemos que

$$d_{P_X}(\mathcal{C}_{\rho_n, h_n}(\aleph_n), S) = P_X(S \setminus \mathcal{C}_{\rho_n, h_n}(\aleph_n)) \quad \text{c.s.},$$

para  $n > n_0$ . Es suficiente ver que si  $\varepsilon_n = d_H(\aleph_n, S)$ , existe  $k_n \in \mathbb{R}$  tal que

$$S \ominus k_n \varepsilon_n B(0, 1) \subset \mathcal{C}_{\rho_n, h_n}(\aleph_n), \quad (3.7)$$

y  $k_n \rho_n \rightarrow c$ , siendo  $c$  constante ya que en este caso

$$P_X(S \setminus \mathcal{C}_{\rho_n, h_n}(\aleph_n)) \leq P_X(S \setminus (S \ominus k_n \varepsilon_n B(0, 1))) = \mathcal{O}(k_n \varepsilon_n).$$

La demostración de (3.7) es igual a la de (3.6) tomando  $k_n = 3 + 2/\sin(\rho_n/2)$ .  $\square$

### 3.3.3. Tasas de convergencia para la envolvente por complemento

En las secciones anteriores obtuvimos, bajo la hipótesis de que el conjunto  $S$  es cono convexo en sentido amplio, tasas de convergencia para la distancia de Hausdorff y en medida. Claramente dichos resultados son ciertos para la envolvente cono convexa en sentido estricto. Siguiendo algunas ideas expuestas en Pateiro-López (2008), vamos a estudiar  $\mathbb{E}\left(d_{P_X}(S, \tilde{\mathcal{C}}_{\rho,h}(\aleph_n))\right)$ , siendo  $\aleph_n = X_1, \dots, X_n$  una muestra i.i.d. de  $X$ , cuya distribución denotaremos  $P_X$ , con soporte  $S$ . Supondremos que  $P_X$  es absolutamente continua, cuya densidad,  $f$ , satisface que existen  $k_1$  y  $k_2$  tal que  $0 < k_1 \leq f(x) \leq k_2 < \infty$  para todo  $x \in S$  c.s.. Supondremos además que

$$S = \tilde{\mathcal{C}}_{\rho,h}(S).$$

Dados  $a \in (0, \pi)$  y  $b > 0$ , denotamos

$$\mathcal{G}_{a,b} = \{C_{a,\xi,b}(x) : x \in \mathbb{R}^d, \xi \in S_1^{d-1}\}.$$

Definiremos, al igual que lo hecho en Pateiro-López (2008) para conjuntos  $r$ -convexos, *familias inevitables*:

**Definición 3.25.** Dado  $x \in \mathbb{R}^d$ ,  $\rho \in (0, \pi)$  y  $h > 0$ , denotemos  $\Lambda_{\rho,h}(x) = \{C \in \mathcal{G}_{\rho,h} : x \in C\}$ . La familia de conjuntos  $\mathfrak{U}_{x,\rho,h}$  es inevitable para  $\Lambda_{\rho,h}(x)$  si para todo  $C \in \Lambda_{\rho,h}(x)$  existe  $U \in \mathfrak{U}_{x,\rho,h}$  tal que  $U \subset C$ .

Supongamos que podemos definir, para cada  $x \in S$ , familias inevitables  $\mathfrak{U}_{x,\rho,h}$  con una cantidad finita de elementos. En tal caso

$$\mathbb{P}(x \in S \setminus \tilde{\mathcal{C}}_{\rho,h}(\aleph_n)) = \mathbb{P}(\exists C \in \Lambda_{\rho,h}(x) : C \cap \aleph_n = \emptyset) \leq \sum_{U \in \mathfrak{U}_{x,\rho,h}} \mathbb{P}(U \cap \aleph_n = \emptyset).$$

De donde se sigue que:

$$\begin{aligned} \mathbb{E}\left(d_{P_X}(S, \tilde{\mathcal{C}}_{\rho,h}(\aleph_n))\right) &= \mathbb{E} \int_S \mathbb{I}_{\{x \in S \setminus \tilde{\mathcal{C}}_{\rho,h}(\aleph_n)\}} f(x) dx \\ &\leq k_2 \int_S \mathbb{P}(x \in S \setminus \tilde{\mathcal{C}}_{\rho,h}(\aleph_n)) dx \\ &\leq k_2 \int_S \sum_{U \in \mathfrak{U}_{x,\rho,h}} \mathbb{P}(U \cap \aleph_n = \emptyset) dx \\ &\leq k_2 \int_S \sum_{U \in \mathfrak{U}_{x,\rho,h}} (1 - k_1 \mu(U \cap S))^n dx. \end{aligned}$$

Por lo tanto lo que haremos será definir, para cada  $x \in S$ , familias inevitables de modo tal que  $\frac{\mu(U \cap S)}{\mu(S)}$  sea *grande*.

**Proposición 3.26.** *Sea  $\gamma = \rho$  si  $\rho \leq \pi/3$  y  $\gamma = (\pi - \rho)/2$  en caso contrario. Sea  $h_1 = \frac{h}{2} \sin(\frac{\rho}{2})$ . Dado  $x \in S$ , consideremos un cubrimiento minimal  $\{\overline{C_{\gamma/2, \xi_j, h_1}(x)} : \|\xi_j\| = 1, j = 1, \dots, k\}$  de la bola  $\overline{B(x, h_1)}$  por conos de ángulo  $\gamma/2$ , eje  $\xi_j$  y altura  $h_1$ . La familia*

$$\mathfrak{U}_{x,\rho,h} = \{C_{\gamma/2, \xi_j, h_1}(x), \|\xi_j\| = 1, j = 1, \dots, k\},$$

*es inevitable para  $\Lambda_{\rho,h}(x)$ . Más aún el cardinal de  $\mathfrak{U}_{x,\rho,h}$  no depende de  $x$ .*

*Demostración.* La idea de la construcción que haremos consiste en encontrar, para cada  $C \in \Lambda_{\rho,h}(x)$  otro cono  $C' = C'_{\gamma, \xi_j, h_1}(x)$  con vértice en  $x$ , ángulo  $\gamma$ , altura  $h_1$  y eje el vector  $\xi_j$  con  $j = 1, \dots, k$ , tal que  $C' \subset C$ . En lugar de considerar todos los conos  $C \in \Lambda_{\rho,h}(x)$  podemos fijar  $C$  y variar  $x$ . Por ejemplo hacemos variar  $x$  en  $C = C_{\rho, e_1, h}(0)$  de ángulo  $\rho$ , altura  $h$ , eje el vector  $e_1 = (1, 0, \dots, 0)$  y vértice en el origen de coordenadas. Por definición de  $\Lambda_{\rho,h}(x)$  tenemos que  $x \in C_{\rho, e_1, h}(0)$ .

Vamos a separar en 3 casos:

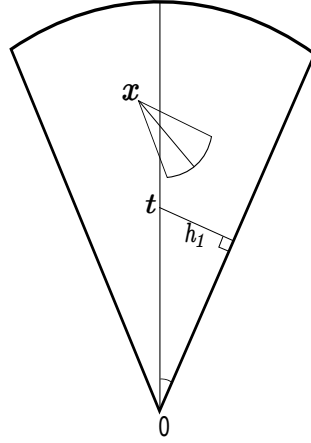
1. Si  $x \in C_{\rho, e_1, h/2}(0)$  tomamos  $C' = C_{\gamma, e_1, h/2}(x)$ . Es claro que en este caso  $C' \subset C_{\rho, e_1, h}(0)$ .
2. Si  $x = r e_1$  con  $h/2 \leq r < h$ , tomamos  $C' = C_{\gamma, -e_1, h_1}(x)$ ; es claro que en este caso también se cumple que  $C' \subset C_{\rho, e_1, h}(0)$  ya que  $h_1$  es la distancia del punto medio del eje del cono  $C_{\rho, e_1, h/2}(0)$  al borde del mismo.
3. Si  $x \in C_{\rho, e_1, h}(0)$  pero no está en los casos anteriores, definimos  $\nu = \frac{\lambda_1 t - x}{\|\lambda_1 t - x\|}$ , donde tomamos  $t = (h/2)e_1$ , y  $\lambda_1$  tal que  $\langle \frac{\lambda_1 t - x}{\|\lambda_1 t - x\|}, \frac{-x}{\|x\|} \rangle = \cos(\gamma/2)$ . Tomamos  $C' = C_{\gamma, \nu, h_1}(x)$ . En este caso es claro que  $C' \subset C_{\rho, e_1, h}(0)$ , ver Figura 3.5.

Para construir la familia inevitable consideremos un cubrimiento minimal de  $\overline{B(x, h_1)}$  de la forma  $\{\overline{C_{\gamma/2, \xi_j, h_1}(x)}\}_{j=1, \dots, k}$  para algunos  $\xi_j \in S_1^{d-1}$ . Veamos que  $\mathfrak{U}_{x,\rho,h} = \{C_{\gamma/2, \xi_j, h_1}(x)\}_{j=1, \dots, k}$  es una familia inevitable para  $\Lambda_{\rho,h}(x)$ . Sea  $C \in \Lambda_{\rho,h}(x)$ ,  $x$  está en alguno de los 3 casos y por lo tanto existe  $\xi$  de norma 1, tal que  $C_{\gamma, \xi, h_1}(x) \subset C$ , sea  $j_0$  tal que  $\langle \xi_{j_0}, \xi \rangle \geq \gamma/2$  (existe porque  $\{\overline{C_{\gamma/2, \xi_j, h_1}(x)}\}_{j=1, \dots, k}$  es un cubrimiento), entonces  $C_{\gamma/2, \xi_{j_0}, h_1}(x) \subset C_{\gamma, \xi, h_1}(x)$ . □

Ahora que hemos definido familias inevitables podemos demostrar el siguiente teorema:

**Teorema 3.27.** *Sea  $S \subset \mathbb{R}^d$ ,  $S \in \tilde{\mathcal{C}}_{\rho,h}$  y  $F(z) = \mu(\{x \in S : d(x, \partial S) \leq z\})$ . Supongamos que  $F'$  es acotada en un entorno de 0, y  $X_1, X_2, \dots$  son i.i.d. con distribución  $P_X$  con soporte  $S$ . Supongamos que  $P_X$  es absolutamente continua con densidad  $f$ , tal que  $0 < k_1 \leq f(x) \leq k_2 < \infty$  para casi todo  $x \in S$ , siendo  $k_1$  y  $k_2$  constantes. Entonces,*

$$\mathbb{E}\left(d_{P_X}(S, \tilde{\mathcal{C}}_{\rho,h}(\aleph_n))\right) = \mathcal{O}(n^{-1/d}).$$


 Figura 3.5: Caso 3,  $C_{\gamma, \nu, h_1}(x) \subset C_{\rho, e_1, h}(0)$ .

*Demostración.* En virtud del Lemma 2 en Stachó (1976)  $F'$  es absolutamente continua y existe excepto en un conjunto  $N$  numerable, por hipótesis existen  $s$  y  $q$  tal que  $F'(t) < q \forall t \in [0, s] \cap N^c$ . Sea  $h_2 = \min\{s, h_1\}$ , donde  $h_1 = \frac{h}{2} \sin(\rho/2)$ , entonces

$$\begin{aligned} \mathbb{E}\left(d_{P_X}(S, \tilde{\mathcal{C}}_{\rho, h}(\mathbb{N}_n))\right) &\leq k_2 \int_{\{x \in S: d(x, \partial S) \leq h_2\}} \sum_{U \in \mathfrak{U}_{x, \rho, h}} (1 - k_1 \mu(U \cap S))^n dx \\ &+ k_2 \int_{\{x \in S: d(x, \partial S) > h_2\}} \sum_{U \in \mathfrak{U}_{x, \rho, h}} (1 - k_1 \mu(U \cap S))^n dx. \end{aligned}$$

Respecto al último término observemos que si  $d(x, \partial S) > h_2$  entonces  $B(x, h_2) \subset S$  y, para todo  $U \in \mathfrak{U}_{x, \rho, h}$ , tenemos que:  $k_1 \mu(U \cap S) \geq k_1 \mu(U \cap B(x, h_2)) = c_0 h_2^d$  siendo  $c_0$  una constante positiva. Por lo tanto si  $k = k(\rho, h)$  denota el cardinal de la familia  $\mathfrak{U}_{x, \rho, h}$  entonces

$$\int_{\{x \in S: d(x, \partial S) > h_2\}} \sum_{U \in \mathfrak{U}_{x, \rho, h}} (1 - k_1 \mu(U \cap S))^n dx \leq k(1 - c_0 h_2^d)^n \mu(\{x \in S : d(x, \partial S) > h_2\}),$$

esta última expresión se puede acotar por  $k_3 e^{-n h_2^d c_0}$ , donde  $k_3$  es una constante positiva.

Para acotar la primer integral observemos que:  $U = C_{\gamma/2, \xi_j, h_1}(x)$  y  $t \leq h_2 \leq h_1$  entonces  $U \cap B(x, t) = C_{\gamma/2, \xi_j, t}(x)$  por lo tanto si  $d(x, \partial S) = t$

$$k_1 \mu(U \cap S) \geq k_1 \mu(U \cap B(x, t)) = c_0 t^d.$$

$$\begin{aligned} \int_{\{x \in S: d(x, \partial S) \leq h_2\}} \sum_{U \in \mathfrak{U}_{x, \rho, h}} (1 - k_1 \mu(U \cap S))^n dx &\leq \int_{\{x \in S: d(x, \partial S) \leq h_2\}} k(1 - c_0 d(x, \partial S)^d)^n dx \\ &\leq \int_{\{x \in S: d(x, \partial S) \leq h_2\}} k e^{-c_0 n d(x, \partial S)^d} dx. \end{aligned}$$

Sea  $g(z) = k e^{-c_0 n z^d}$ . Haciendo un cambio de variable obtenemos:

$$\begin{aligned} \int_{\{x \in S: d(x, \partial S) \leq h_2\}} k e^{-c_0 n d(x, \partial S)^d} dx &= \int_{\{x \in S: d(x, \partial S) \leq h_2\}} g(d(x, \partial S)) dx = \int_{[0, h_2]} g(z) dF(z) \\ &= \int_{[0, h_2]} g(z) F'(z) dz \leq \int_{[0, h_2]} k_4 e^{-c_0 n z^d} dz, \end{aligned}$$

donde  $k_4$  es una constante positiva (en la última desigualdad hemos usado que  $F'$  es acotado en  $[0, s]$ ). Finalmente, existe una constante positiva  $k_5 > 0$  tal que

$$\int_{[0, h_2]} k_4 e^{-c_0 n z^d} dz = n^{-1/d} \int_0^{c_0 n h_2^d} k_5 e^{-u} u^{(1-d)/d} du \leq n^{-1/d} \int_0^{+\infty} k_5 e^{-u} u^{(1-d)/d} du = \mathcal{O}(n^{-1/d}).$$

Juntando ambas cotas:  $\mathbb{E}(d_{P_X}(S, \tilde{\mathcal{C}}_{\rho, h}(\aleph_n))) = \mathcal{O}(e^{-nh_2^d c_0} + n^{-1/d}) = \mathcal{O}(n^{-1/d})$ .  $\square$

**Observación 3.28.**

1. La condición de que  $F'(0) < \infty$  es equivalente a la finitud del contenido de Minkowski exterior, según fue definido en 1.3. Un estudio detallado del contenido de Minkowski exterior puede encontrarse en Ambrosio y otros (2008).
2. El orden de convergencia que obtuvimos es más lento que el que se obtiene en Pateiro-López (2008) (ver Teorema 2.5.2) para los conjuntos  $r$ -convexos, en los que rueda una bola de radio  $\alpha > 0$  en  $S$  y  $\overline{S^c}$ , no obstante es importante señalar que la clase de conjuntos en la que estamos trabajando es mucho mayor.

### 3.4. Estimación no paramétrica de la medida de la frontera de conjuntos cono-convexos

En esta sección vamos a comentar brevemente el problema de la estimación de la longitud (o volumen de la superficie en dimensión mayor) del borde de un conjunto que verifica la condición de cono convexidad exterior. Veamos con un ejemplo sencillo que, aún en el caso en que la muestra está sobre el borde del conjunto, la longitud del borde de la envolvente cono convexa de un conjunto  $S$ , no converge a la longitud del borde de dicho conjunto, por lo tanto es razonable pensar que lo mismo sucederá cuando la muestra es uniforme en el conjunto. Consideremos el conjunto en  $\mathcal{C}_{\pi/2}$  de la Figura 3.6 formado por un cuadrado de lado uno cortado con el complemento de un disco de radio un medio, cuyo centro está en el punto medio de uno de los lados. Es fácil ver que si tomamos ángulo  $\rho = \pi/2$  y construimos la envolvente  $\rho$  cono convexa de una muestra de puntos uniforme en cualquiera de los 3 lados rectos, la suma de las longitudes de los arcos que constituyen el borde de dicha envolvente es  $\pi/2$ . Por otro lado si la muestra está en el arco de la circunferencia la envolvente  $\rho$  cono convexa de los puntos coincide con el arco. Es decir la longitud del borde de la envolvente es  $2\pi$  y la del conjunto  $3 + \pi/2$ .

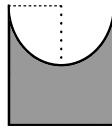


Figura 3.6: Conjunto  $S \in \mathcal{C}_{\pi/2}$  para el cual  $L_0(\partial S) \neq L_0(\partial \mathcal{C}_{\pi/2}(\aleph_n))$ .

**Ejemplo 3.29.** Veamos con un ejemplo que  $S_n \xrightarrow{d_H} S$  y  $\partial S_n \xrightarrow{d_H} \partial S$ , aún en el caso en que  $S_n \in \mathcal{C}_\rho$  y  $S \in \mathcal{C}_\rho$  no implica que el contenido de Minkowski de  $\partial S_n$  converge al de  $S$ . Consideremos  $S = [0, 1]^2 \subset \mathbb{R}^2$  y  $S_n$  de la siguiente manera:

$$S_0 = S \cup A_0^1,$$

siendo  $A_0^1$  el triángulo isósceles con base el segmento formado por los puntos  $(0, 1)$  y  $(1, 1)$ , como se muestra en la Figura 3.7 y lado  $1/\sqrt{2}$ .

$$S_n = A_n^1 \cup A_n^2 \cup \dots \cup A_n^{2^n},$$

siendo  $A_n^i$  con  $i = 1, \dots, 2^n$  el triángulo isósceles con base en el segmento formado por los puntos  $((i-1)/2^n, 1)$  y  $(i/2^n, 1)$ , de lado  $1/(2^n\sqrt{2})$ .

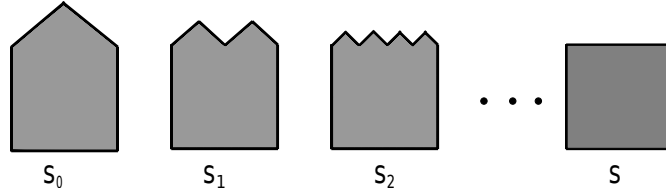


Figura 3.7:  $S_n \in \mathcal{C}_{\pi/2}$  para todo  $n$ .

Observemos que el contenido de Minkowski de  $\partial S_n$  es  $3 + \sqrt{2}$  para todo  $n$  y que  $S_n$  y  $S$  pertenecen a  $\mathcal{C}_{\pi/2, 1}$ .

**Lema 3.30.** Sean  $S_n$  y  $S$  subconjuntos compactos de  $\mathbb{R}^d$  tal que existe el contenido de Minkowski  $L_0(\partial S)$  y  $\varepsilon_n := d_H(\partial S_n, \partial S) \rightarrow 0$ , consideremos  $\varepsilon'_n \rightarrow 0$  tal que  $\varepsilon_n/\varepsilon'_n \rightarrow 0$  entonces

$$\lim_{n \rightarrow \infty} \frac{\mu((\partial S_n) \oplus \varepsilon'_n B(0, 1))}{2\varepsilon'_n} = L_0(\partial S).$$

*Demostración.* Demostraremos que

1.  $\forall \beta > 1$   $(\partial S_n) \oplus \varepsilon'_n B(0, 1) \subset (\partial S) \oplus \beta \varepsilon'_n B(0, 1)$  para  $n \geq n_0$ .
2.  $\forall \alpha < 1$   $(\partial S) \oplus \alpha \varepsilon'_n B(0, 1) \subset (\partial S_n) \oplus \varepsilon'_n B(0, 1)$  para  $n \geq n_1$ .

Veamos 1.: sea  $x_n \in (\partial S_n) \oplus \varepsilon'_n B(0, 1)$ , existe  $t_n \in \partial S_n$  tal que  $x_n \in B(t_n, \varepsilon'_n)$ . Como  $\varepsilon_n = d_H(\partial S_n, \partial S)$  existe  $y_n \in B(t_n, \varepsilon_n)$ ,  $y_n \in \partial S$ . Es suficiente ver que  $x_n \in B(y_n, \beta \varepsilon'_n)$  para  $n \geq n_0$ , donde  $n_0$  no depende de  $x_n$ , pero esto se sigue de que:

$$\|y_n - x_n\| \leq \|y_n - t_n\| + \|t_n, x_n\| \leq \varepsilon_n + \varepsilon'_n = \varepsilon'_n(1 + \varepsilon_n/\varepsilon'_n) < \beta \varepsilon'_n \quad n \geq n_0.$$

La demostración de 2. es análoga, tomemos  $x_n \in (\partial S) \oplus \alpha \varepsilon'_n B(0, 1)$ , existe  $t_n \in \partial S$  tal que  $x_n \in B(t_n, \alpha \varepsilon'_n)$ . Como  $\varepsilon_n = d_H(\partial S_n, \partial S)$  existe  $y_n \in B(t_n, \varepsilon_n)$ ,  $y_n \in \partial S_n$ . Es suficiente ver que  $x_n \in B(y_n, \varepsilon'_n)$  para  $n \geq n_1$  donde  $n_1$  no depende de  $x_n$ , pero esto se sigue de que:

$$\|y_n - x_n\| \leq \|x_n - t_n\| + \|t_n - y_n\| \leq \alpha \varepsilon'_n + \varepsilon_n = \varepsilon'_n(\alpha + \varepsilon_n/\varepsilon'_n) < \varepsilon'_n \quad n \geq n_1.$$

El resultado se sigue de forma inmediata a partir de 1 y 2:

$$\begin{aligned} \alpha L_0(S) &= \alpha \lim_{n \rightarrow \infty} \frac{\mu((\partial S) \oplus \alpha \varepsilon'_n B(0, 1))}{2\alpha \varepsilon'_n} \leq \liminf_{n \rightarrow \infty} \frac{\mu((\partial S_n) \oplus \varepsilon'_n B(0, 1))}{2\varepsilon'_n} \\ &\leq \limsup_{n \rightarrow \infty} \frac{\mu((\partial S_n) \oplus \varepsilon'_n B(0, 1))}{2\varepsilon'_n} \leq \beta \lim_{n \rightarrow \infty} \frac{\mu((\partial S) \oplus \beta \varepsilon'_n B(0, 1))}{2\beta \varepsilon'_n} = \beta L_0(S). \end{aligned}$$

□

**Observación 3.31.** Observemos que, si bien el ejemplo con el que comenzamos la sección nos sugiere que no es posible estimar  $L_0(\partial S)$  por  $L_n = L_0(\partial \mathcal{C}_{\rho, h}(\mathbb{N}_n))$  para el caso en que  $S \in \mathcal{C}_{\rho, h}$ , como conocemos la tasa de convergencia de  $\varepsilon_n = d_H(\partial S, \partial \mathcal{C}_{\rho, h}(\mathbb{N}_n))$  (Proposición 3.23), basta tomar en la definición de  $L_0$  un  $\varepsilon'_n$  de orden mayor.

## 3.5. Inferencia sobre los parámetros $r$ y $\rho$

### 3.5.1. Espacios maximales en conjuntos cono convexos

Una pregunta que surge de manera natural cuando suponemos restricciones de forma del tipo  $r$ -convexo o  $\rho$  cono convexo es cómo determinar el valor de  $r$  o de  $\rho$ . Si bien es cierto que en ambos casos se puede tomar un valor más pequeño que el verdadero valor, esto produce velocidades de convergencia más lentas. Cuando sobre estimamos el valor el estimador deja de ser consistente, (en distancia de Hausdorff o en medida) por lo tanto es de interés cuantificar el error que se comete en dicho caso, así como tener un procedimiento estadístico que nos permita aproximar el valor real. Para el estudio del valor real de  $r$  o de  $\rho$  para el caso de conjuntos  $r$ -convexos o  $\rho$  cono convexos vamos a introducir el concepto de espacio maximal (ver Janson (1987)):

**Definición 3.32.** Sea  $A \subset \mathbb{R}^d$  acotado tal que  $\text{int}(A) \neq \emptyset$ , definimos el *espacio maximal* de  $A$  (que denotaremos  $m(A)$ ) como

$$\sup \{m : \exists x \in A \text{ tal que } B(x, m) \subset A\}.$$

La idea intuitiva que nos permitirá construir luego un test para saber si estamos tomando el  $\rho$  óptimo se basa en el hecho de que si  $S$  es  $\rho$  cono convexo y tomamos un  $\rho'$  mayor que el mínimo de los valores que verifican  $S = \mathcal{C}_\rho(S)$ , el conjunto  $\mathcal{C}_{\rho'}(S) \setminus S$  tendrá un espacio maximal positivo que se puede estimar por medio de una muestra  $X_1, \dots, X_n$  de variables i.i.d con distribución uniforme, con soporte  $S$ . Para eso vamos a usar el siguiente teorema (ver Janson (1987)).

**Teorema 3.33.** Sea  $K \subset \mathbb{R}^d$  acotado tal que  $\mu_d(K) = 1$  y  $\mu_d(\partial K) = 0$  (donde  $\mu_d$  denota la medida de Lebesgue en  $\mathbb{R}^d$ ). Sean  $\aleph_n = \{X_1, \dots, X_n\}$  una muestra de variables i.i.d. con distribución uniforme en  $K$ , si denotamos

$$V_n = \mu_d\left(B(0, m(K \setminus \aleph_n))\right),$$

entonces

$$nV_n - \log(n) - (d-1) \log(\log(n)) - \log(\beta) \rightarrow U,$$

donde

$$\mathbb{P}(U \leq u) = e^{-e^{-u}},$$

y

$$\beta = \frac{1}{d!} \left( \frac{\sqrt{\pi} \Gamma\left(\frac{d}{2} + 1\right)}{\Gamma\left(\frac{d+1}{2}\right)} \right)^{d-1}.$$

Dado que en general no conocemos la medida del conjunto, suponer que la misma es uno es una restricción muy fuerte, por lo tanto en lugar del Teorema 3.33 vamos a usar una versión un poco más general.

**Teorema 3.34.** Sea  $K \subset \mathbb{R}^d$  acotado tal que  $\mu_d(\partial K) = 0$  y  $\mu_d(K) > 0$  (donde  $\mu_d$  denota la medida de Lebesgue en  $\mathbb{R}^d$ ). Sean  $\aleph_n = \{X_1, \dots, X_n\}$  una muestra de variables i.i.d. con distribución uniforme en  $K$ , si denotamos

$$V_n = \mu_d\left(B(0, m(K \setminus \aleph_n))\right),$$

entonces

$$\frac{n}{\mu_d(K)} V_n - \log(n) - (d-1) \log(\log(n)) - \log(\beta) \rightarrow U,$$

donde

$$\mathbb{P}(U \leq u) = e^{-e^{-u}},$$

y

$$\beta = \frac{1}{d!} \left( \frac{\sqrt{\pi} \Gamma\left(\frac{d}{2} + 1\right)}{\Gamma\left(\frac{d+1}{2}\right)} \right)^{d-1}. \quad (3.8)$$

*Demostración.* Consideremos la transformación lineal  $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$  tal que

$$T(x_1, \dots, x_d) = \frac{1}{(\mu_d(K))^{1/d}}(x_1, \dots, x_d),$$

es inmediato verificar a partir de un cambio de variable que, para todo boreliano  $B \subset \mathbb{R}^d$ , se tiene que

$$\mu_d(T(B)) = \frac{1}{\mu_d(K)}\mu_d(B). \quad (3.9)$$

De (3.9) se sigue que  $T(\aleph_n)$  es una muestra con distribución uniforme en  $T(K)$ . Podemos aplicar el Teorema 3.33 para  $T(K)$  y  $T(\aleph_n)$  y obtener:

$$nV_n^T - \log(n) - (d-1)\log(\log(n)) - \log(\beta) \rightarrow U,$$

donde  $V_n^T = \mu_d\left(B(0, m(T(K) \setminus T(\aleph_n)))\right) = \mu_d\left(T(B(0, K \setminus \aleph_n))\right)$ . Por (3.9) tenemos que

$$V_n^T = \frac{1}{\mu_d(K)}\mu_d\left(B(0, m(K \setminus \aleph_n))\right) = \frac{1}{\mu_d(K)}V_n$$

□

Es claro que el resultado anterior sigue siendo cierto si en lugar de  $\mu_d(K)$  tomamos una variable aleatoria  $Y_n$  tal que  $Y_n \rightarrow \mu_d(K)$  casi seguramente.

Veamos ahora un resultado que nos permite relacionar el espacio maximal de conjuntos que están próximos en distancia de Hausdorff y cuyos bordes también lo están.

**Proposición 3.35.** Sean  $A$  y  $B$  subconjuntos acotados, no vacíos de  $\mathbb{R}^d$  tal que  $d_H(A, B) \leq \varepsilon$  y  $d_H(\partial A, \partial B) \leq \varepsilon$  entonces

$$|m(A) - m(B)| \leq 2\varepsilon.$$

*Demostración.* Es suficiente ver:

1.  $\{x \in A : d(x, \partial A) > 2\varepsilon\} \subset B$

y

2.  $\{x \in B : d(x, \partial B) > 2\varepsilon\} \subset A$

ya que de 1 obtenemos que  $m(B) \geq m(A) - 2\varepsilon$ , mientras que de 2 se obtiene que  $m(A) \geq m(B) - 2\varepsilon$ , de donde se deduce que  $|m(A) - m(B)| \leq 2\varepsilon$ . Para ver 1 (2 es análogo) supongamos que existe  $x \in A$  tal que  $d(x, \partial A) > 2\varepsilon$  pero  $x \notin B$ . Como  $A \subset \overset{\circ}{B}(B, \varepsilon)$  tenemos que  $x \in \overset{\circ}{B}(B, \varepsilon) \setminus B$  por lo tanto  $d(x, \partial B) \leq \varepsilon$ ,

$$d(x, \partial A) = d_H(\{x\}, \partial A) \leq d_H(\{x\}, \partial B) + d_H(\partial B, \partial A) \leq 2\varepsilon.$$

□

**Observación 3.36.** Observemos que de la Proposición anterior se sigue de manera inmediata que  $m$  es continuo en  $(E, d)$  el espacio métrico de los subconjuntos compactos de  $\mathbb{R}^d$ , con la métrica  $d$  definida como:  $d(A, B) = d_H(A, B) + d_H(\partial A, \partial B)$ .

Dado que lo que queremos es usar la Proposición 3.35 para comparar el espacio maximal entre  $\tilde{\mathcal{C}}_\rho(\aleph_n) \setminus \aleph_n$  y el de  $\tilde{\mathcal{C}}_\rho(S) \setminus \aleph_n$  tenemos que demostrar que quitar  $\aleph_n$  en  $\tilde{\mathcal{C}}_\rho(\aleph_n)$  y en  $\tilde{\mathcal{C}}_\rho(S)$  no nos modifica la distancia de Hausdorff entre los conjuntos y entre los bordes, para eso:



**Corolario 3.37.** Sean  $A$ ,  $B$  y  $C$  subconjuntos de  $\mathbb{R}^d$ , acotados, no vacíos, tal que  $C \subset B \subset A$  y  $d_H(A, B) \leq \varepsilon$ ,  $d_H(\partial A, \partial B) \leq \varepsilon$  entonces

$$|m(A \setminus C) - m(B \setminus C)| \leq 2\varepsilon.$$

*Demostración.* Es consecuencia inmediata de la proposición anterior y de que

$$d_H(A \setminus C, B \setminus C) = d_H(A, B) \quad \text{y} \quad d_H(\partial(A \setminus C), \partial(B \setminus C)) = d_H(\partial A, \partial B).$$

□

**Teorema 3.38.** Sea  $S$  compacto,  $\rho \in (0, \pi]$  y  $\aleph_n = \{X_1, \dots, X_n\}$  una muestra de variables aleatorias i.i.d. con distribución uniforme con soporte  $S$ , entonces con probabilidad uno

$$d_H(\partial \tilde{\mathcal{C}}_\rho(S), \partial \tilde{\mathcal{C}}_\rho(\aleph_n)) \leq kd_H(S, \aleph_n) \text{ para } n \text{ suficientemente grande,} \quad (3.10)$$

y

$$d_H(\tilde{\mathcal{C}}_\rho(S), \tilde{\mathcal{C}}_\rho(\aleph_n)) \leq kd_H(S, \aleph_n) \text{ para } n \text{ suficientemente grande,} \quad (3.11)$$

siendo

$$k = \frac{2}{\text{sen}(\rho/2)}.$$

*Demostración.* Veamos que, para todo  $n$

$$\tilde{\mathcal{C}}_\rho(S) \ominus k\varepsilon_n B(0, 1) \subset \tilde{\mathcal{C}}_\rho(\aleph_n) \subset \tilde{\mathcal{C}}_\rho(S) \quad \text{c.s.,} \quad (3.12)$$

donde  $\varepsilon_n = d_H(S, \aleph_n)$ . La segunda inclusión es inmediata, veamos la primera: sea  $x \in \tilde{\mathcal{C}}_\rho(S) \ominus k\varepsilon_n B(0, 1)$ , supongamos que  $x \notin \tilde{\mathcal{C}}_\rho(\aleph_n)$ . Como  $x \notin \tilde{\mathcal{C}}_\rho(\aleph_n)$  existe  $\xi$  de norma uno y un cono  $C_{\rho, \xi}(y)$  con vértice en algún punto  $y$  tal que  $x \in C_{\rho, \xi}(y)$  y  $C_{\rho, \xi}(y) \cap \aleph_n = \emptyset$ . Como  $x \in \tilde{\mathcal{C}}_\rho(S) \ominus k\varepsilon_n B(0, 1)$  tenemos que  $B(x, k\varepsilon_n) \subset \tilde{\mathcal{C}}_\rho(S)$ . Por la forma en que fue tomado  $k$  existe  $z \in B(x, k\varepsilon_n) \cap C_{\rho, \xi}(y)$  tal que  $d(z, \partial C_{\rho, \xi}(y)) > 2\varepsilon_n$ . Consideremos el cono  $C_{\rho, \xi}(z)$ , es claro que  $C_{\rho, \xi}(z) \subset C_{\rho, \xi}(y)$ . Como  $z \in \tilde{\mathcal{C}}_\rho(S)$  y  $z \in B(x, k\varepsilon_n)$  tenemos que  $C_{\rho, \xi}(z) \cap S \neq \emptyset$ . Sea  $s \in C_{\rho, \xi}(z) \cap S$ , por la forma en que tomamos  $z$  es claro que  $B(s, 2\varepsilon_n) \subset C_{\rho, \xi}(y)$ , pero esto contradice que  $C_{\rho, \xi}(y) \cap \aleph_n = \emptyset$ . De (3.12) se sigue (3.11) de forma inmediata.

Para probar (3.10) hay que observar que si  $x \in \partial \tilde{\mathcal{C}}_\rho(\aleph_n)$  entonces  $x \notin \tilde{\mathcal{C}}_\rho(S) \ominus k\varepsilon_n B(0, 1)$ . Para ver esto último supongamos por reducción al absurdo que existe  $x \in \partial \tilde{\mathcal{C}}_\rho(\aleph_n)$  y  $x \in \tilde{\mathcal{C}}_\rho(S) \ominus k\varepsilon_n B(0, 1)$ . Dado que  $x \in \partial \tilde{\mathcal{C}}_\rho(\aleph_n)$  existe  $\xi$  de norma 1, y un cono  $C_{\rho, \xi}(x)$  tal que  $C_{\rho, \xi}(x) \cap \aleph_n = \emptyset$ , además sabemos que  $B(x, k\varepsilon_n) \subset \tilde{\mathcal{C}}_\rho(S)$ , es decir si consideramos los conos

$$C_\lambda = C_{\rho, \xi}(x + \lambda\xi) \quad 0 \leq \lambda < k\varepsilon_n,$$

tenemos que  $x + \lambda\xi \in B(x, k\varepsilon_n)$  y por lo tanto  $C_\lambda \cap S \neq \emptyset$  para todo  $0 \leq \lambda < k\varepsilon_n$ . Como  $S$  es compacto  $\overline{C_{k\varepsilon_n}} \cap S \neq \emptyset$ . Sea  $s \in \overline{C_{k\varepsilon_n}} \cap S \neq \emptyset$ , es claro que  $B(s, 2\varepsilon_n) \subset C_{\rho, \xi}(x)$  lo cual es un absurdo, por la forma en que fue tomado  $\varepsilon_n$ . □

El siguiente teorema es una consecuencia inmediata del teorema anterior y del Corolario 3.37

**Teorema 3.39.** Sea  $S$  en las hipótesis del teorema anterior, entonces con probabilidad uno

$$\left| m(\tilde{\mathcal{C}}_\rho(\aleph_n) \setminus \aleph_n) - m(\tilde{\mathcal{C}}_\rho(S) \setminus \aleph_n) \right| \leq 2kd_H(S, \aleph_n),$$

para  $n$  suficientemente grande c.s., siendo

$$k = \frac{2}{\text{sen}(\rho/2)}.$$

### 3.5.2. Contraste de hipótesis sobre $\rho$

#### Contraste para $\rho$ .

En esta sección supondremos que tenemos  $S \subset \mathbb{R}^d$  tal que  $S = \tilde{\mathcal{C}}_\rho(S)$  (siendo  $\rho$  el supremo de los valores para los cuales se verifica dicha propiedad), y que existe  $r > 0$  tal que  $S$  está en las hipótesis del Teorema 1.11 con  $r_0 = r$ . Como se dijo al comienzo de la sección anterior, nos interesa testear a partir de una muestra  $\aleph_n$ , si estamos sobrestimando  $\rho$ . Veamos primero una proposición que nos dice que si tomamos un  $\rho$  mayor que el verdadero valor eso nos produce un espacio maximal positivo.

**Proposición 3.40.** *Sea  $S$  tal que  $S = \tilde{\mathcal{C}}_{\rho_1}(S)$ , con  $\rho_1$  el supremo de los valores que cumplen dicha propiedad. Sea  $\rho > \rho_1$  entonces*

$$0 < m(\tilde{\mathcal{C}}_\rho(S) \setminus S) \leq m(\tilde{\mathcal{C}}_\rho(S) \setminus \aleph_n) \leq m(\tilde{\mathcal{C}}_\rho(S) \setminus S) + d_H(S, \aleph_n) \quad \forall \rho > \rho_1 \text{ y } n > 0.$$

*Demostración.* Veamos primero la demostración de la primer desigualdad (la segunda es inmediata). Como  $S$  no es  $\rho$  cono convexo existe  $x \in \partial S$  tal que  $x \notin \partial \tilde{\mathcal{C}}_\rho(S)$  es decir  $x \in \text{int}(\tilde{\mathcal{C}}_\rho(S))$ . Sea  $\delta > 0$  tal que  $B(x, \delta) \subset \tilde{\mathcal{C}}_\rho(S)$ . Como  $x \in \partial S$  y  $S$  es  $\rho_1$  cono convexo existe  $\xi$  de norma 1 y un cono  $C_{\rho_1, \xi}(x)$  tal que  $C_{\rho_1, \xi}(x) \cap S = \emptyset$ . Es decir  $B(x, \delta) \cap C_{\rho_1, \xi}(x) \subset \tilde{\mathcal{C}}_\rho(S) \setminus S$ . Para demostrar la tercer desigualdad tomemos  $\gamma > 0$  y  $m_\gamma$  tal que  $0 < m(\tilde{\mathcal{C}}_\rho(S) \setminus \aleph_n) - m_\gamma < \gamma$ , sea  $c_\gamma$  tal que  $B(c_\gamma, m_\gamma) \subset \tilde{\mathcal{C}}_\rho(S) \setminus \aleph_n$ , es claro que  $B(c_\gamma, m_\gamma - d_H(S, \aleph_n)) \subset \tilde{\mathcal{C}}_\rho(S) \setminus S$  de donde se sigue que  $m_\gamma - d_H(S, \aleph_n) \leq m(\tilde{\mathcal{C}}_\rho(S) \setminus S)$ .  $\square$

**Teorema 3.41.** *Consideremos la siguiente prueba de hipótesis:*

$$\begin{cases} H_0 : & \rho \geq \rho_1 \\ H_1 : & \rho < \rho_1 \end{cases} \quad (3.13)$$

el conjunto  $C_n = \tilde{\mathcal{C}}_{\rho_1}(C_r(\aleph_n))$ , el estadístico

$$\tilde{V}_n = \mu_d\left(B(0, m(C_n \setminus \aleph_n))\right),$$

y la región crítica

$$RC = \{\tilde{V}_n > c_{n, \alpha}\},$$

donde

$$c_{n, \alpha} = \frac{\mu_d(S)}{n} \left( -\log(-\log(1 - \alpha)) + \log(n) + (d - 1) \log(\log(n)) + \log(\beta) \right) = \mathcal{O}(\log(n)/n),$$

y  $\beta$  es la constante (3.8). Dado  $\alpha \in (0, 1)$  la prueba (3.13) tiene potencia 1, y es asintóticamente de nivel  $\alpha$ .

*Demostración.* Tenemos que demostrar que

$$P_{H_0}(\tilde{V}_n > c_{n, \alpha}) \rightarrow \alpha,$$

y, para  $n$  suficientemente grande

$$P_{H_1}(\tilde{V}_n > c_{n, \alpha}) = 1.$$

Observemos que bajo  $H_0$ ,  $\aleph_n \subset C_r(\aleph_n) \subset C_n \subset S$  para todo  $n$ , con probabilidad uno. Como estamos suponiendo que  $S$  es  $r$ -convexo sabemos, por el Teorema 4 en Rodríguez-Casal (2006) que:

$$d_H(C_r(\aleph_n), S) = \mathcal{O}\left(\left(\log(n)/n\right)^{2/(d+1)}\right),$$

y

$$d_H(\partial C_r(\aleph_n), \partial S) = \mathcal{O}\left(\left(\frac{\log(n)}{n}\right)^{2/(d+1)}\right),$$

por lo tanto si usamos el Corolario 3.37 tenemos que, bajo  $H_0$  :

$$\left| m(C_n \setminus \aleph_n) - m(S \setminus \aleph_n) \right| = \mathcal{O}\left(\left(\frac{\log(n)}{n}\right)^{2/(d+1)}\right).$$

Si denotamos  $V_n = \mu_d\left(B(0, m(S \setminus \aleph_n))\right)$ , tenemos que:

$$V_n - \tilde{V}_n = \mathcal{O}\left(\left(\frac{\log(n)}{n}\right)^{2d/(d+1)}\right),$$

$$\alpha_n = P_{H_0}(\tilde{V}_n > c_{n,\alpha}) = P_{H_0}((\tilde{V}_n - V_n) + V_n > c_{n,\alpha}) = P(V_n > c_{n,\alpha} + o(c_{n,\alpha})).$$

Observemos que conocemos (Teorema 3.34) la distribución asintótica de  $V_n$  bajo  $H_0$  por lo tanto

$$\alpha_n \rightarrow \alpha.$$

Bajo  $H_1$ , por el Teorema 3.39 y la Proposición 3.40,

$$0 < m(\tilde{\mathcal{C}}_{\rho_1}(S) \setminus S) \leq m(\tilde{\mathcal{C}}_{\rho_1}(S) \setminus \aleph_n) \leq m(\tilde{\mathcal{C}}_{\rho_1}(\aleph_n) \setminus \aleph_n) + 2kd_H(S, \aleph_n) \leq m(C_n \setminus \aleph_n) + 2kd_H(S, \aleph_n)$$

es decir, para  $n$  suficientemente grande

$$P_{H_1}(\tilde{V}_n > c_{n,\alpha}) = 1,$$

ya que  $c_{n,\alpha} \rightarrow 0$ . □

### Contraste para $r$

En esta sección, siguiendo las ideas de la sección anterior, vamos a proponer un test de hipótesis asintóticamente de nivel  $\alpha$ , para el parámetro  $r$  de los conjuntos  $r$ -convexos, que están en las hipótesis del Teorema 1.11 (con  $r = r_0$ ) y tal que  $S = \overline{\text{int}(S)}$ . Para eso, vamos a estudiar al igual que lo hecho para el contraste de  $\rho$ ,  $d_H(C_r(S), C_r(\aleph_n))$  y  $d_H(\partial C_r(S), \partial C_r(\aleph_n))$  para  $r > r_1$  (donde  $\aleph_n = \{X_1, \dots, X_n\}$  es una muestra de vectores aleatorios con distribución uniforme en  $S$ ). Veremos que se puede aproximar el espacio maximal  $m(C_r(S) \setminus S)$  por  $m(C_r(\aleph_n) \setminus \aleph_n)$ .

**Observación 3.42.** *Observemos que*

$$m(C_r(\aleph_n)) \geq m(\overline{\text{int}(C_r(\aleph_n))}) \geq m(\text{int}(C_r(\aleph_n))) = m(C_r(\aleph_n)),$$

por lo tanto  $m(C_r(\aleph_n)) = m(\overline{\text{int}(C_r(\aleph_n))})$ . En virtud de esto último es que trabajaremos con  $\overline{\text{int}(C_r(\aleph_n))}$  en lugar de  $C_r(\aleph_n)$ .

Veamos primero algunos lemas que serán de utilidad en esta sección

**Lema 3.43.** *Sea  $S \subset \mathbb{R}^d$  compacto tal que  $S = \overline{\text{int}(S)}$  entonces  $C_r(S) = \overline{\text{int}(C_r(S))}$ .*

*Demostración.* La demostración es análoga a la hecha en la Proposición 3.9 para las envolventes  $\rho$  cono convexas. Como  $C_r(S)$  es cerrado tenemos que  $\overline{\text{int}(C_r(S))} \subset C_r(S)$  por lo tanto nos resta demostrar que  $C_r(S) \subset \overline{\text{int}(C_r(S))}$ . Como  $S = \overline{\text{int}(S)} \subset C_r(S)$  y  $\overline{\text{int}(S)} \subset \overline{\text{int}(C_r(S))}$ , para demostrar que  $C_r(S) \subset \overline{\text{int}(C_r(S))}$  es suficiente ver que  $\overline{\text{int}(C_r(S))}$  es  $r$ -convexo.

Sea  $x \in \overline{\partial \text{int}(C_r(S))}$  tenemos que encontrar  $t$  tal que  $x \in \partial B(t, r) \subset \overline{\text{int}(C_r(S))}^c$ . Veamos que  $x \in \partial C_r(S)$ . Para todo  $\delta > 0$  tenemos que  $B(x, \delta) \cap \overline{\text{int}(C_r(S))} \neq \emptyset$  y  $B(x, \delta) \cap \overline{\text{int}(C_r(S))}^c \neq \emptyset$ . Sea  $z \in B(x, \delta) \cap \overline{\text{int}(C_r(S))}$ .  $B(z, \delta - d(z, \partial B(x, \delta))) \cap \text{int}(C_r(S))$  entonces  $B(x, \delta) \cap C_r(S) \neq \emptyset$ .

Sea  $y \in B(x, \delta) \cap \overline{\text{int}(C_r(S))}^c$ , tenemos que o bien  $y \in C_r(S)^c$  en cuyo caso  $x \in \partial C_r(S)$ , o bien  $y \in \partial C_r(S)$  en cuyo caso  $B(y, \delta - d(y, \partial B(x, \delta))) \cap C_r(S)^c \neq \emptyset$  de donde  $B(x, \delta) \cap C_r(S)^c \neq \emptyset$ . Como  $x \in \partial C_r(S)$  podemos tomar  $t$  tal que  $x \in \partial B(t, r) \subset C_r(S)^c \subset \overline{\text{int}(C_r(S))}^c$ .  $\square$

**Lema 3.44.** *Sea  $S \subset \mathbb{R}^d$  tal que  $S = C_{r_1}(S)$ , (con  $r_1$  el supremo de los  $r \in \mathbb{R}$  que cumplen dicha propiedad) entonces  $m(C_r(S) \setminus S) > 0$  para todo  $r > r_1$ .*

*Demostración.* Como  $S$  no es  $r$ -convexo existe  $x \in \partial S$  tal que  $x \notin \partial C_r(S)$ , además como  $S \subset C_r(S)$  tenemos que  $x \in \text{int}(C_r(S))$  por lo tanto existe  $\delta > 0$  tal que  $B(x, \delta) \subset C_r(S)$ . Como  $S$  es  $r_1$ -convexo y  $x \in \partial S$  existe  $y$  tal que  $x \in \partial B(y, r_1)$  y  $B(y, r_1) \subset S^c$ . Tenemos que  $B(y, r_1) \cap B(x, \delta) \subset C_r(S) \cap S^c$ , de donde se sigue la tesis.  $\square$

**Lema 3.45.** *Sea  $S \subset \mathbb{R}^d$  tal que  $S = C_{r_1}(S)$ , (con  $r_1$  el supremo de los  $r \in \mathbb{R}$  que cumplen esa propiedad) y sean  $\aleph_n = X_1, \dots, X_n$  variables i.i.d. con distribución uniforme en  $S$ . Sea  $\varepsilon_n = d_H(S, \aleph_n)$ , y  $r > r_1$  entonces*

$$C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B \subset C_r(\aleph_n) \subset C_r(S) \quad \text{para } n \text{ suficientemente grande.} \quad (3.14)$$

*Demostración.* La segunda inclusión es inmediata ya que  $\aleph_n \subset S$ . Veamos la primera: sea  $x \in C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B$ , supongamos por absurdo que  $x \notin C_r(\aleph_n)$  es decir, existe  $y$  tal que  $B(y, r) \cap \aleph_n = \emptyset$  y  $x \in B(y, r)$ . Es claro que  $B(x, 2\varepsilon_n) \subset C_{r-2\varepsilon_n}(S)$ . Es claro que  $B(y, r - 2\varepsilon_n) \cap B(x, 2\varepsilon_n) \neq \emptyset$ . Sea  $z \in B(y, r - 2\varepsilon_n) \cap B(x, 2\varepsilon_n)$  como  $z \in B(x, 2\varepsilon_n)$  tenemos que  $z \in C_{r-2\varepsilon_n}(S)$  de donde  $B(y, r - 2\varepsilon_n) \cap S \neq \emptyset$ , sea  $s \in B(y, r - 2\varepsilon_n) \cap S$  entonces  $B(s, 2\varepsilon_n) \subset B(y, r)$  pero esto es absurdo ya que  $B(y, r) \cap \aleph_n = \emptyset$ .  $\square$

Es inmediato, pero será de importancia más adelante, que las inclusiones en (3.14) siguen siendo válidas si tomamos interior y luego clausura, en cada uno de los conjuntos, es decir

$$\overline{\text{int}(C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B)} \subset \overline{\text{int}(C_r(\aleph_n))} \subset \overline{\text{int}(C_r(S))} \quad \text{para } n \text{ suficientemente grande.}$$

**Lema 3.46.** *Sea  $S \subset \mathbb{R}^d$  tal que  $S = \overline{\text{int}(S)}$  y  $S = C_{r_1}(S)$  (con  $r_1$  el supremo de los  $r \in \mathbb{R}$  que cumplen esa propiedad) y sean  $\aleph_n = \{X_1, \dots, X_n\}$  variables i.i.d. con distribución uniforme en  $S$ . Sea  $\varepsilon_n = d_H(S, \aleph_n)$  y  $r > r_1$ , con probabilidad uno*

1.  $\overline{\text{int}(C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B)} \xrightarrow{d_\mu} C_r(S)$
2.  $\overline{\text{int}(C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B)} \xrightarrow{d_H} C_r(S)$

*Demostración.* Veamos la demostración de 1. Como

$$\partial(C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B) \subset \{x \in \mathbb{R}^d : d(x, \partial C_{r-2\varepsilon_n}(S)) = 2\varepsilon_n\},$$

y  $\mu\{x \in \mathbb{R}^d : d(x, \partial C_{r-2\varepsilon_n}(S)) = 2\varepsilon_n\} = 0$  tenemos que

$$\mu\left(\overline{\text{int}(C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B)}\right) = \mu\left(C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B\right).$$

Vamos a demostrar 1 para  $C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B$ . Observemos que:

$$C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B = (S \oplus (r - 2\varepsilon_n)B) \ominus rB \subset (S \oplus (r - 2\varepsilon_{n+1})B) \ominus rB. \quad (3.15)$$

Para demostrar 1 veamos que como  $C_r(S)$  es compacto se tiene que

$$C_r(S) \setminus \bigcup_n \left( (S \oplus (r - 2\varepsilon_n)B) \ominus rB \right) \subset \partial C_r(S), \quad (3.16)$$

para eso sea  $x \in C_r(S)$  tal que  $\forall n \ x \notin (S \oplus (r - 2\varepsilon_n)B) \ominus rB$ . Existe  $t_n \in B(x, r)$  tal que  $B(t_n, r - 2\varepsilon_n) \cap S = \emptyset$ , sea  $t = \lim_n t_n$  (pasando eventualmente a subsucesiones), es claro que  $t \in \overline{B(x, r)}$ , veamos que

$$B(t, r) \cap S = \emptyset. \quad (3.17)$$

Supongamos que existe  $z \in B(t, r) \cap S$  y tomemos  $n$  suficientemente grande de modo que  $\|t_n - t\| < d(z, \partial B(t, r))/4$  y  $\varepsilon_n < d(z, \partial B(t, r))/4$ , entonces

$$\|t_n - z\| \leq \|t_n - t\| + \|t - z\| = \|t_n - t\| + (r - d(z, \partial B(t, r))) < r - 3\varepsilon_n < r - 2\varepsilon_n,$$

pero esto contradice que  $B(t_n, r - 2\varepsilon_n) \cap S = \emptyset$ . Veamos que  $t \in \partial B(x, r)$ : sabemos que  $t \in \overline{B(x, r)}$ , supongamos que  $t \in B(x, r)$  es decir  $x \in B(t, r)$ . Como  $x \in C_r(S)$ , tiene que ser  $B(t, r) \cap S \neq \emptyset$  lo cual contradice (3.17). Tenemos que  $x \in \partial B(t, r) \subset S^c$  entonces  $x \in \partial C_r(S)$  como queríamos demostrar. Como  $C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B = (S \oplus (r - 2\varepsilon_n)B) \ominus rB \subset C_r(S)$  de (3.15), (3.16) y  $\mu(\partial C_r(S)) = 0$  se sigue la tesis.

Veamos la demostración de 2, supongamos que no fuese cierto, como  $\overline{\text{int}(C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B)} = \text{int}\left(\overline{(S \oplus (r - 2\varepsilon_n)B) \ominus rB}\right) \subset C_r(S)$  podemos tomar  $x_n \in C_r(S)$  y

$$d\left(x_n, \overline{\partial \text{int}(C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B)}\right) = d_H\left(C_r(S), \overline{\text{int}(C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B)}\right) > \delta \text{ para todo } n.$$

Tomamos  $x = \lim x_n$  (pasando eventualmente a subsucesiones), entonces  $x \in C_r(S)$  ya que  $C_r(S)$  es compacto. Tenemos que  $B(x, \delta/2) \cap \text{int}(C_r(S)) \neq \emptyset$  lo cual contradice que  $\overline{\text{int}(C_{r-2\varepsilon_n}(S) \ominus 2\varepsilon_n B)} \xrightarrow{d_\mu} C_r(S)$   $\square$

**Lema 3.47.** *Sea  $S$  en las hipótesis anteriores, entonces, con probabilidad uno*

1.  $d_H\left(\overline{\partial \text{int}(C_r(\aleph_n))}, \partial C_r(S)\right) \rightarrow 0,$

y

2.  $d_H\left(\overline{\text{int}(C_r(\aleph_n))}, C_r(S)\right) \rightarrow 0.$

*Demostración.* Veamos la demostración de 1: si  $x \in \overline{\partial \text{int}(C_r(\aleph_n))}$  entonces  $x \in \partial C_r(\aleph_n)$ , como  $C_r(\aleph_n)$  es  $r$ -convexo existe  $y$  tal que  $x \in \partial B(y, r)$  y  $B(y, r) \cap \aleph_n = \emptyset$ . Veamos que para todo  $n \ x \notin (S \oplus (r - 2\varepsilon_n)B) \ominus rB$ : supongamos que existe  $n$  tal que  $x \in (S \oplus (r - 2\varepsilon_n)B) \ominus rB$  sea  $z \in B(y, \varepsilon) \cap B(x, r)$  entonces  $z \in S \oplus (r - 2\varepsilon_n)B$  de donde  $B(z, r - 2\varepsilon_n) \cap S \neq \emptyset$  lo cual contradice que  $B(y, r) \cap \aleph_n = \emptyset$ . La tesis se sigue de la parte 2 del lema anterior. La demostración de 2 es inmediata a partir de la parte 2 del lema anterior y de (3.14).  $\square$

Veamos ahora una adaptación del Teorema 3.39 para los conjuntos  $r$ -convexos.

**Teorema 3.48.** *Sea  $S \subset \mathbb{R}^d$  tal que  $S = \overline{\text{int}(S)}$  y  $S = C_{r_1}(S)$ , (con  $r_1$  el supremo de los  $r \in \mathbb{R}$  que cumplen esa propiedad) y sean  $\aleph_n = X_1, \dots, X_n$  variables i.i.d. con distribución uniforme en  $S$ . Sea  $\varepsilon_n = d_H(S, \aleph_n)$ , y  $r > r_1$  entonces*

$$\left| m(C_r(\aleph_n) \setminus \aleph_n) - m(C_r(S) \setminus \aleph_n) \right| \rightarrow 0.$$

*Demostración.* Como dijimos antes, podemos suponer que  $C_r(\aleph_n) = \overline{\text{int}(C_r(\aleph_n))}$ . El resultado es entonces una consecuencia del Corolario 3.37 y del lema anterior.  $\square$

Veamos ahora el teorema que nos asegura que la prueba de hipótesis antes mencionada tiene potencia 1 y es asintóticamente de nivel  $\alpha$ . La demostración es análoga a la hecha para conjuntos  $\rho$  como convexos.

**Teorema 3.49.** *Sea  $S \subset \mathbb{R}^d$  un conjunto compacto no vacío tal que  $\mu_d(S) > 0$ , sea  $r > 0$  el supremo de los valores que hacen que  $S$  esté en las hipótesis del Teorema 1.11 con  $r_0 = r$ . Consideremos la siguiente prueba de hipótesis:*

$$\begin{cases} H_0 : & r \geq r_1 \\ H_1 : & r < r_1 \end{cases} \quad (3.18)$$

el conjunto  $C_n = C_{r_1}(\aleph_n)$  el estadístico:

$$\tilde{V}_n = \mu_d\left(B(0, m(C_n \setminus \aleph_n))\right),$$

y la región crítica

$$RC = \{\tilde{V}_n > c_{n,\alpha}\},$$

donde

$$c_{n,\alpha} = \frac{\mu_d(S)}{n} \left( -\log(-\log(1-\alpha)) + \log(n) + (d-1)\log(\log(n)) + \log(\beta) \right) = \mathcal{O}(\log(n)/n),$$

y  $\beta$  es la constante 3.8). Dado  $\alpha \in (0, 1)$  la prueba (3.18) tiene potencia 1, y es asintóticamente de nivel  $\alpha$ .

### 3.6. Contraste de hipótesis de convexidad

En la sección anterior hicimos una prueba de hipótesis para conjuntos cono convexos, que nos permite testear si estamos tomando o no un  $\rho$  más grande que el máximo valor posible. Supongamos que queremos testear, a partir de una muestra  $X_1, \dots, X_n$  uniformemente distribuida en un conjunto  $S$  compacto tal que  $\mu(S) > 0$  siendo  $\mu$  la medida de Lebesgue, si dicho conjunto es convexo, (asumiremos que, bajo la hipótesis nula, el borde cumple ciertas restricciones de suavidad que veremos más adelante). Observemos que en este caso no estamos haciendo hipótesis sobre la forma de  $S$  es decir no estamos suponiendo que es cono convexo y queremos verificar si el  $\rho$  que estamos tomando es o no el adecuado. Lo que buscamos es saber si dicho conjunto pertenece o no a una determinada familia (la de los convexos cuyo borde cumple una condición de regularidad). Observemos que  $\tilde{\mathcal{C}}_\pi(\aleph_n)$  no es otra cosa que la envolvente convexa de la muestra, teniendo eso presente, sabemos que para cualquier conjunto convexo  $S \subset \mathbb{R}^d$  se verifica (ver Dümbgen y Walther (1996)) que:

$$d_H(S, \tilde{\mathcal{C}}_\pi(\aleph_n)) = \mathcal{O}\left((\log(n)/n)^{1/d}\right).$$

No obstante, bajo una hipótesis adicional de regularidad, en dicho trabajo se demuestra que

$$d_H(S, \tilde{\mathcal{C}}_\pi(S)) = \mathcal{O}\left((\log(n)/n)^{2/(d+1)}\right).$$

Esta condición es:

**Condición (P)** Para cada  $x \in \partial S$  existe un único vector  $\xi = \xi(x)$  de norma uno tal que  $\langle y, \xi \rangle \leq \langle x, \xi \rangle$  para todo  $y \in S$ , y para alguna constante  $l \in \mathbb{R}$  se cumple que:

$$\|\xi(x) - \xi(y)\| \leq l\|x - y\| \quad \forall x, y \in \partial S.$$

Veamos ahora una prueba de hipótesis que nos permite testear si un conjunto compacto  $S$  igual a la clausura de sus puntos interiores está en la clase de los conjuntos convexos que verifican la condición (P). Para eso usaremos el espacio maximal  $m(\tilde{\mathcal{C}}_\pi(S) \setminus \aleph_n)$ . Veamos primero que si  $S$  no es convexo, entonces  $m(\tilde{\mathcal{C}}_\pi(S) \setminus S) > 0$ .

**Lema 3.50.** Sea  $S \subset \mathbb{R}^d$  compacto, tal que  $S = \overline{\text{int}(S)}$ , supongamos que  $S$  no es convexo, entonces

$$0 < m(\tilde{\mathcal{C}}_\pi(S) \setminus S).$$

*Demostración.* Observemos primero que por la Proposición 3.9, si usamos que la intersección de los convexos que contienen a  $S$  es igual a la intersección de los semiespacios que no lo cortan obtenemos:

$$\tilde{\mathcal{C}}_\pi(S) = \overline{\text{int}(\tilde{\mathcal{C}}_\pi(S))}. \quad (3.19)$$

Como  $S$  no es convexo existen  $x, y \in S$  tal que el segmento que los une no está contenido en  $S$ , como  $\tilde{\mathcal{C}}_\pi(S)$  es convexo, dicho segmento tiene que estar incluido en  $\tilde{\mathcal{C}}_\pi(S)$ . De la compacidad de  $S$  se sigue que podemos tomar un punto  $t$  en dicho segmento tal que  $d(t, S) > \delta > 0$  para algún  $\delta$ . Por 3.19  $B(t, \delta) \cap \text{int}(\tilde{\mathcal{C}}_\pi(S)) \subset \tilde{\mathcal{C}}_\pi(S) \setminus S$ , de donde se sigue la tesis de forma inmediata.  $\square$

**Observación 3.51.** Si bien el teorema siguiente presenta una prueba de hipótesis fácil de implementar y que arroja buenos resultados en las simulaciones que veremos más adelante, es importante aclarar que la condición **(P)** es una restricción en la forma del borde del conjunto que deja fuera conjuntos convexos muy simples como lo son por ejemplo un cuadrado. Si bien creemos que se puede generalizar al caso en que el borde cumple la condición **(P)** salvo en un conjunto finito de puntos, no tenemos una prueba de consistencia para este caso.

**Teorema 3.52.** Denotemos  $\mathcal{A}_P$  a la clase de los subconjuntos compactos convexos que verifican la condición **(P)**, sea  $S$  compacto igual a la clausura de su interior, consideremos la prueba de hipótesis

$$\begin{cases} H_0 : S \in \mathcal{A}_P \\ H_1 : S \notin \mathcal{A}_P \end{cases} \quad (3.20)$$

el conjunto  $C_n = \tilde{\mathcal{C}}_\pi(\aleph_n)$ , el estadístico

$$\tilde{V}_n = \mu_d\left(B(0, m(C_n \setminus \aleph_n))\right),$$

y la región crítica

$$RC = \{\tilde{V}_n > c_{n,\alpha}\},$$

donde

$$c_{n,\alpha} = \frac{\mu_d(S)}{n} \left( -\log(-\log(1-\alpha)) + \log(n) + (d-1)\log(\log(n)) + \log(\beta) \right) = \mathcal{O}(\log(n)/n),$$

y  $\beta$  es la constante (3.8). Entonces dado  $\alpha \in (0, 1)$  la prueba (3.20) es asintóticamente de nivel  $\alpha$ .

*Demostración.* La prueba es casi idéntica a la hecha en 3.41 no obstante la incluiremos a efectos de completitud. Tenemos que demostrar que

$$P_{H_0}(\tilde{V}_n > c_{n,\alpha}) \rightarrow \alpha,$$

y, para  $n$  suficientemente grande

$$P_{H_1}(\tilde{V}_n > c_{n,\alpha}) = 1.$$

Observemos que bajo  $H_0$ ,  $\aleph_n \subset C_n \subset S$  para todo  $n$ . Como estamos suponiendo que se cumple la condición **(P)** sabemos (ver Dümbgen y Walther (1996)) que:

$$d_H(C_n, S) = \mathcal{O}\left((\log(n)/n)^{2/(d+1)}\right),$$

por lo tanto

$$d_H(\partial C_n, \partial S) = \mathcal{O}\left((\log(n)/n)^{2/(d+1)}\right).$$

Si usamos el Corolario 3.37 tenemos que, bajo  $H_0$

$$\left| m(C_n \setminus \aleph_n) - m(S \setminus \aleph_n) \right| = \mathcal{O}\left( (\log(n)/n)^{2/(d+1)} \right).$$

Bajo  $H_0$ , si denotamos  $V_n = \text{Vol}\left(B(0, m(S \setminus \aleph_n))\right)$ , tenemos que,

$$V_n - \tilde{V}_n = \mathcal{O}\left( (\log(n)/n)^{2d/(d+1)} \right).$$

$$\alpha_n = P_{H_0}(\tilde{V}_n > c_{n,\alpha}) = P_{H_0}((\tilde{V}_n - V_n) + V_n > c_{n,\alpha}) = P(V_n > c_{n,\alpha} + o(c_{n,\alpha})).$$

Observemos que conocemos (Teorema 3.34) la distribución asintótica de  $V_n$  bajo  $H_0$  por lo tanto  $\alpha_n \rightarrow \alpha$ .  $\square$

Si bien la hipótesis alternativa no significa que el conjunto no sea convexo, si no lo es tenemos, por el Teorema 3.39 y el lema anterior, con probabilidad uno, para  $n$  suficientemente grande

$$0 < m(\tilde{\mathcal{C}}_\pi(S) \setminus S) - d_H(\aleph_n, S) \leq m(C_n \setminus \aleph_n),$$

es decir, para  $n$  suficientemente grande

$$P_{\tilde{H}_1}(\tilde{V}_n > c_{n,\alpha}) = 1,$$

ya que  $c_{n,\alpha} \rightarrow 0$ , donde  $\tilde{H}_1$  significa que  $S$  no es convexo,

### 3.6.1. Algoritmo y simulaciones

#### Algoritmo

Para definir un algoritmo que calcule  $\tilde{V}_n = \mu_d\left(B(0, m(C_n \setminus \aleph_n))\right)$  vamos a usar el diagrama de Voronoi definido en 1.17 para calcular  $m_n := m(C_n \setminus \aleph_n)$ . Primero presentaremos un algoritmo exacto que si bien se puede calcular es computacionalmente más costoso que otro aproximado que usaremos en las simulaciones. Observemos primero que si  $B(x, m_n) \subset C_n \setminus \aleph_n$  tenemos dos posibilidades:

- 1)  $B(x, m_n) \cap \partial C_n = \emptyset$ , en cuyo caso, como  $m$  es maximal, existen al menos tres puntos  $X_i, X_j, X_k \in \partial B(x, m)$ , es decir  $x$  está en la intersección de las regiones  $V_i, V_j, V_k$  del diagrama de Voronoi definido por  $\aleph_n$ .
- 2)  $B(x, m_n) \cap \partial C_n \neq \emptyset$ , en este caso, como  $\partial C_n$  es una unión finita de segmentos, tenemos nuevamente 2 casos
  - a)  $B(x, m_n)$  es tangente a uno y sólo uno de dichos segmentos, en cuyo caso existen, como  $m_n$  es maximal, al menos dos puntos de la muestra  $X_i, X_j \in \partial B(x, m_n)$ .
  - b)  $B(x, m_n)$  es tangente a dos o más segmentos de  $\partial C_n$ , en cuyo caso  $x$  pertenece a la bisectriz del ángulo interior que forman los segmentos tangentes.

Lo que haremos ahora es considerar los centros y radios que se encuentren en los casos 1 y 2, y tomar el máximo de dichos radios.

En el caso 1) si  $y_1, \dots, y_k$  denotan los puntos de intersección de tres regiones del diagrama de Voronoi, definimos  $m_1 = \max\{d(y_i, \aleph_n) : i = 1, \dots, k\}$  donde  $d(y_i, \aleph_n) = \min\{d(y_i, X_j) : j = 1, \dots, n\}$ .

En el caso 2 a) tenemos que considerar  $\binom{n}{2}$  parejas de puntos  $X_i, X_j, i, j = 1, \dots, n, i \neq j$ . Para cada una de dichas parejas y cada segmento del borde existe un único  $x$  tal que  $X_i, X_j \in \partial B(x, d)$  siendo  $d = d(x, X_j)$ ,  $B(x, d)$  es tangente a la recta que determina el segmento. Nos interesan los



pares  $x, d$  tal que  $B(x, d) \subset C_n \setminus \mathcal{N}_n$  (observemos que esta condición es fácil de verificar). Definimos  $m_2$  como el máximo de los radios  $d$  que verifican dicha condición.

En el caso 2 b) la situación es análoga a la 2 a), si es tangente a más de dos segmentos de  $\partial C_n$  entonces el centro  $x$  está en la intersección de dos o más bisectrices. Si es tangente sólo a dos segmentos de  $\partial C_n$  el centro  $x$  pertenece a la bisectriz de un ángulo interior, y, como  $m_n$  es maximal, es tangente a uno o más puntos de la muestra, dichos centros y radios son una cantidad finita, que depende de  $n$ . Definimos  $m_3$  como el máximo de dichos radios.

Finalmente, el espacio maximal  $m(C_n \setminus \mathcal{N}_n) = \max\{m_1, m_2, m_3\}$ . Así definido el algoritmo, es de orden a lo sumo  $n^3$ . En las simulaciones que presentaremos a continuación consideraremos solamente las bolas que se encuentran en el punto uno, esto nos da un espacio maximal  $\hat{m}_n$  menor o igual que  $m_n$ , pero que es computacionalmente más fácil de calcular que el anterior. En este caso la prueba es asintóticamente de potencia uno y de nivel menor o igual que  $\alpha$ .

### Simulaciones

Veamos algunas simulaciones hechas para algunos subconjuntos de  $\mathbb{R}^2$ , vamos a tomar en ambos casos  $\alpha = 0,05$  con lo cual

$$\beta = \frac{1}{2} \frac{\sqrt{\pi} \Gamma(2)}{\Gamma(\frac{3}{2})} = 1,$$

y

$$c_{n,\alpha} = \mu_2(K) \frac{2,97 + \log(n) + \log(\log(n))}{n},$$

siendo  $K$  es el conjunto sobre el cual haremos la prueba de hipótesis y  $\mu$  la medida de Lebesgue en  $\mathbb{R}^2$

- El primero conjunto que consideraremos es el disco de centro en el origen y radio  $1/\sqrt{\pi}$ . Hemos repetido el test 1000 veces, para diferentes valores de  $n$  desde 100 a 2000, en la Tabla 3.1 se muestra  $\hat{\alpha}$  la estimación de  $\alpha$ .

$n$	$\hat{\alpha}$
100	0.035
200	0.037
500	0.044
1000	0.042
1500	0.045
2000	0.05

Tabla 3.1: En la tabla se muestran la estimación del nivel, en base a 1000 repeticiones, para diferentes tamaños muestrales.

- El segundo conjunto que hemos considerado es el cuadrado  $[0, 2]^2$ , dicho conjunto no verifica la condición **(P)** no obstante los resultados que se obtienen sugieren que el Teorema 3.52 es válido en este contexto. Hemos repetido el test 1000 veces, para diferentes valores de  $n$ . En la Tabla 3.2 se muestra  $\hat{\alpha}$  la estimación de  $\alpha$ .
- El tercer conjunto es el cuadrado  $[0, 1]^2$  al que le hemos quitado un triángulo isósceles con ángulo  $\varphi$ , y altura  $1/2$  como se muestra en la Figura 3.8 (b). En las simulaciones que se presentan en la Tabla 3.3 se muestra la potencia estimada  $\hat{\beta}$  de la prueba, en base a 1000 repeticiones de la misma, para diferentes valores de  $\varphi$  y diferentes tamaños muestrales.

$n$	$\hat{\alpha}$
500	0.06
800	0.052
1000	0.057
1500	0.061
2000	0.059

Tabla 3.2: En la tabla se muestran la estimación del nivel, en base a 1000 repeticiones, para diferentes tamaños muestrales.

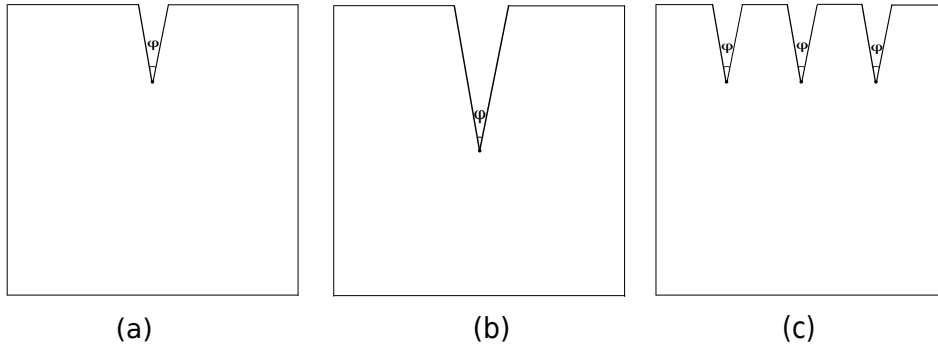


Figura 3.8: (b)  $[0, 1]^2 \setminus T_b$  siendo  $T_b$  el triángulo isósceles de altura  $1/2$  y ángulo  $\varphi$ . (a)  $[0, 1]^2 \setminus T_a$  siendo  $T_a$  el triángulo isósceles de ángulo  $\varphi$  y altura  $h$ , donde  $h$  lo tomamos de modo tal que su área sea  $1/3$  de la del triángulo  $T_a$ . (c)  $[0, 1]^2 \setminus S$  donde  $S$  son tres triángulos iguales a  $T_a$

- En la cuarta simulación hemos comparado el efecto de poner una entrante, como en la Figura 3.8 (a), contra el de poner tres de la misma área. Para el caso de tres entrantes tomamos  $\varphi = \pi/6$  y realizamos, para diferentes valores de  $N$ , 1000 repeticiones. Como se ve en la primera de las tablas en 3.4, es más fácil detectar la no convexidad a través de una entrante grande, que de varias entrantes pequeñas, aún en el caso en el que se mantiene el ángulo de entrante. La segunda tabla calcula la potencia estimada, para una entrante de ángulo  $\varphi = \pi/6$  Figura 3.8 (a), cuya área es un tercio del área de la entrante en 3.8 (b).

$\varphi = \pi/4$	
$n$	$\hat{\beta}$
100	0.40
130	0.636
160	0.835
200	0.946
300	0.997

$\varphi = \pi/6$	
$n$	$\hat{\beta}$
200	0.565
250	0.787
300	0.926
400	0.996
500	1

$\varphi = \pi/8$	
$n$	$\hat{\beta}$
300	0.543
350	0.679
400	0.846
500	0.976
600	0.997

Tabla 3.3: En las tablas se muestra  $\hat{\beta}$  la estimación de la potencia, en base a 1000 repeticiones, para diferentes valores de  $\varphi$ .

Tres entrantes $\varphi = \pi/6$	
$n$	$\hat{\beta}$
600	0.741
700	0.889
800	0.967
900	0.985
1000	1

Una entrante $\varphi = \pi/6$	
$n$	$\hat{\beta}$
800	0.642
900	0.776
1000	0.882
1250	0.977
1500	1

Tabla 3.4: Potencia estimada en base a 1000 repeticiones

### 3.7. Un estimador consistente del parámetro

Hasta ahora hemos abordado el problema de sobre estimar el verdadero valor del ángulo para subconjuntos que verifican la condición de cono convexidad en sentido estricto mediante una prueba de hipótesis que permite evaluar si estamos tomando un valor mayor que el real. En esta sección veremos como, siguiendo las ideas de la sección anterior, podemos construir un estimador consistente del verdadero ángulo, que denotaremos como  $\rho_1$ . Recordemos que  $\rho_1$  es el supremo de los  $\rho$  para los cuales  $\tilde{\mathcal{C}}_\rho(S) = S$  por lo tanto en virtud de la Proposición 3.40 tenemos que

$$\rho_1 = \inf \left\{ \rho > 0 : m(\tilde{\mathcal{C}}_\rho(S) \setminus S) > 0 \right\},$$

esto último nos sugiere un estimador para  $\rho_1$ :

$$\hat{\rho}_n = \inf \left\{ \rho > 0 : m(\tilde{\mathcal{C}}_\rho(\mathfrak{N}_n) \setminus \mathfrak{N}_n) > \beta_n \right\},$$

siendo  $\beta_n$  una sucesión de números reales positivos que converge a 0. Lo que haremos será encontrar el orden que debe tener  $\beta_n$  de modo que el estimador antes propuesto sea consistente casi seguramente. Observemos que por el Teorema 3.34 sabemos que

$$m(\tilde{\mathcal{C}}_{\rho_1}(S) \setminus \mathfrak{N}_n) = \mathcal{O} \left( \left( \frac{\log(n)}{n} \right)^{1/d} \right). \quad (3.21)$$

**Teorema 3.53.** *Sea  $S$  tal que  $S = \tilde{\mathcal{C}}_{\rho_1}(S)$ , con  $\rho_1$  el supremo de los valores que cumplen dicha propiedad, supongamos además que  $S$  es estándar. Sea  $\rho > \rho_1$  y  $\mathfrak{N}_n = \{X_1, \dots, X_n\}$  una muestra de variables aleatorias i.i.d, con distribución uniforme con soporte  $S$ , entonces con probabilidad uno*

$$\hat{\rho}_n = \inf \left\{ \rho > 0 : m(\tilde{\mathcal{C}}_\rho(\mathfrak{N}_n) \setminus \mathfrak{N}_n) > \beta_n \right\} \rightarrow \rho_1, \quad (3.22)$$

donde

$$\beta_n = \mathcal{O} \left( \left( \frac{\log(n)}{n} \right)^{1/d} \right).$$

*Demostración.* Definamos

$$\beta_n = m(\tilde{\mathcal{C}}_{\rho_1}(S) \setminus \aleph_n) + \varepsilon_n,$$

donde  $\varepsilon_n = d_H(S, \aleph_n)$ . Observemos que por (3.21) y el Teorema 3.18,  $\beta_n$  tiene orden  $(\log(n)/n)^{1/d}$ . Si tomamos  $\rho > \hat{\rho}_n$  tenemos que

$$m(\tilde{\mathcal{C}}_{\rho_1}(S) \setminus \aleph_n) + \varepsilon_n < m(\tilde{\mathcal{C}}_{\rho}(\aleph_n) \setminus \aleph_n).$$

Como  $\tilde{\mathcal{C}}_{\rho}(\aleph_n) \subset \tilde{\mathcal{C}}_{\rho}(S)$  tenemos que, para todo  $\rho > \hat{\rho}_n$ :

$$m(\tilde{\mathcal{C}}_{\rho}(\aleph_n) \setminus \aleph_n) \leq m(\tilde{\mathcal{C}}_{\rho}(S) \setminus \aleph_n).$$

Por la Proposición 3.40 se verifica, para todo  $n$ :

$$m(\tilde{\mathcal{C}}_{\rho}(S) \setminus \aleph_n) \leq m(\tilde{\mathcal{C}}_{\rho}(S) \setminus S) + \varepsilon_n.$$

Es importante observar que la desigualdad anterior se cumple para todo  $n$ , y todo  $\rho > \rho_1$ . Hemos demostrado entonces que para todo  $\rho > \hat{\rho}_n$

$$0 \leq m(\tilde{\mathcal{C}}_{\rho_1}(S) \setminus \aleph_n) < m(\tilde{\mathcal{C}}_{\rho}(S) \setminus S),$$

de donde  $0 < m(\tilde{\mathcal{C}}_{\rho}(S) \setminus S)$  para todo  $\rho > \hat{\rho}_n$  y por lo tanto  $\hat{\rho}_n \geq \rho_1$ . Supongamos por reducción al absurdo que existe  $\rho_t$  tal que  $\hat{\rho}_n > \rho_t > \rho_1$  para todo  $n$ . Por el Teorema 3.39 y la Proposición 3.40 tenemos que, para  $n > n_0(\rho_t)$

$$m(\tilde{\mathcal{C}}_{\rho_t}(\aleph_n) \setminus \aleph_n) \geq m(\tilde{\mathcal{C}}_{\rho_t}(S) \setminus S) - \varepsilon_n > 0, \quad (3.23)$$

de donde se sigue que

$$m(\tilde{\mathcal{C}}_{\rho_t}(\aleph_n) \setminus \aleph_n) > \beta_n,$$

para algún  $n > n_1$ , y por lo tanto  $\hat{\rho}_n < \rho_t$ . □

## 3.8. Aspectos computacionales de la propiedad de Poincaré

### 3.8.1. Algoritmo en $\mathbb{R}^d$

Hasta ahora hemos presentado resultados asintóticos y características geométricas de las diferentes envolventes de una muestra  $\aleph_n = X_1, \dots, X_n$  de vectores i.i.d. con distribución  $P_X$  en un conjunto cono convexo  $S$  (y cono convexo por complementos). Una pregunta que surge de manera natural es cómo calcular dichas envolventes. En esta sección vamos a presentar un algoritmo estocástico para aproximar la envolvente  $\rho, h$  cono convexa por complementos  $\tilde{\mathcal{C}}_{\rho, h}(\aleph_n)$ . La idea del algoritmo está basada en el hecho de que si  $x \notin \tilde{\mathcal{C}}_{\rho, h}(S)$  entonces existe un cono  $C_{\rho, h}(y)$  que no contiene puntos de la muestra y tal que  $x \in C_{\rho, h}(y)$ . Siguiendo esta idea lo que haremos es partir de un rectángulo  $E = [a, b] \times [c, d]$  que contenga a la muestra, y sortear al azar un vértice  $y$  y un eje  $\xi$ , de modo tal que si el cono de altura  $h$  que determinan  $y$  y  $\xi$  no corta a la muestra “borramos” de  $E$  dicho cono (para optimizar el procedimiento vamos a borrar el cono con vértice en  $y$ , eje  $\xi$ , con el mayor ángulo posible, de modo tal que no corte a la muestra). Luego repetimos el sorteo de un vértice y un eje pero esta vez en  $E := E \setminus C_{\rho, h}(y)$ . Si el cono inicial cortaba a la muestra, volvemos a sortear. Este algoritmo de sorteo de vértices, de ejes, y de borrado, se repite una cantidad  $N$  de veces y se obtiene así un conjunto  $E$  que es  $\rho, h$  cono convexo por complementos.

---

INPUT: Una muestra  $\aleph_n = \{X_1, \dots, X_n\} \subset \mathbb{R}^d$ , los parámetros  $\rho \in (0, \pi]$  y  $h > 0$ , el rectángulo  $E = [a, b] \times [c, d]$  con  $\aleph_n \subset E$ ,  $N$  un entero positivo indicando el número de iteraciones completas de los pasos 1 a 3 del algoritmo.

PASO 1. **Generar conos aleatoriamente:** Elegir al azar un vértice  $x \in E$  y un eje  $\xi$  con  $\|\xi\| = 1$ , considerar el cono  $C_{\rho, h, \xi}(x)$ .

PASO 2. **Verificar que el cono no contiene puntos de la muestra:** Si  $C_{\rho, h, \xi}(x) \cap \aleph_n \neq \emptyset$  volvemos al Paso 1.

PASO 3. **Borrar un cono maximal:** Si  $C_{\rho, h, \xi}(x) \cap \aleph_n = \emptyset$  borramos el cono maximal con vértice en  $x$  que no contiene puntos de la muestra. Es decir, encontrar  $X_l, X_r \in \aleph_n$  que satisfagan:

$$\frac{\langle X_l - x, \xi \rangle}{\|X_l - x\|} = \max \left\{ \frac{\langle X_i - x, \xi \rangle}{\|X_i - x\|} : \langle X_i - x, \xi \rangle \geq 0, \langle X_i - x, \xi^\perp \rangle > 0, \|X_i - x\| \leq h \right\},$$

$$\frac{\langle X_r - x, \xi \rangle}{\|X_r - x\|} = \max \left\{ \frac{\langle X_i - x, \xi \rangle}{\|X_i - x\|} : \langle X_i - x, \xi \rangle \geq 0, \langle X_i - x, \xi^\perp \rangle \leq 0, \|X_i - x\| \leq h \right\},$$

donde  $\xi^\perp$  denota un vector de norma uno, ortogonal a  $\xi$  (ver Figura 3.9).

Si  $\left\{ \frac{\langle X_i - x, \xi \rangle}{\|X_i - x\|} : \langle X_i - x, \xi \rangle \geq 0, \langle X_i - x, \xi^\perp \rangle > 0, \|X_i - x\| \leq h \right\} = \emptyset$  elegir  $X_l = x + \xi^\perp$ .

Si  $\left\{ \frac{\langle X_i - x, \xi \rangle}{\|X_i - x\|} : \langle X_i - x, \xi \rangle \geq 0, \langle X_i - x, \xi^\perp \rangle \leq 0, \|X_i - x\| \leq h \right\} = \emptyset$  elegir  $X_r = x - \xi^\perp$ .

Borramos el cono (que vamos a denotar  $C$ ) de altura  $h$ , con vértice en  $x$  y lados de longitud  $h$ , determinados por las direcciones  $X_l - x$  y  $X_r - x$ , es decir reemplazamos  $E$  con  $E \setminus C$  y volvemos al Paso 1.

OUTPUT: El conjunto  $E$  resultante de aplicar el Paso 3  $N$  veces.  $N$  es por lo tanto el número de conos borrados durante el proceso de iteración.

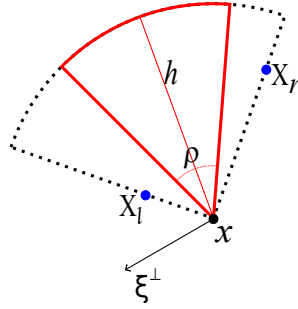


Figura 3.9: Paso 3:  $\langle X_l, \xi^\perp \rangle > 0$  y  $\langle X_r, \xi^\perp \rangle < 0$

**Observación 3.54.** El algoritmo dado anteriormente se puede mejorar con algunos cambios simples, por ejemplo podemos sortear puntos en la envolvente convexa de la muestra o, eligiendo adecuadamente  $r$ , en la envolvente  $r$ -convexa de la misma. Encontrar un algoritmo exacto (no estocástico) para calcular  $\mathcal{C}_{\rho, h}(\aleph_n)$  o  $\mathcal{C}_{\rho, h}(\aleph_n)$  es un problema abierto.

### 3.8.2. Aplicación a diferentes datos.

Hemos realizado un estudio computacional para varios subconjuntos de  $\mathbb{R}^2$  y comparado los resultados que se obtienen con el algoritmo anterior y con la envolvente  $r$ -convexa de la muestra.

Un algoritmo exacto para calcular el cierre  $r$ -convexo para muestras en  $\mathbb{R}^2$  puede encontrarse en el paquete `alphahu11` de R (ver Pateiro-López y Rodríguez-Casal (2010)). Hemos aplicado el algoritmo a diferentes datos, en algunos casos dichos datos se obtuvieron por medio de simulaciones, en otros tomamos datos reales (tales como la posición de plantas en un área determinada). Mientras que en los primeros, como conocemos el conjunto a estimar, podemos aproximar (por medio del Método de Montecarlo) la distancia en medida entre el conjunto y la envolvente, para diferentes valores de los parámetros, viendo así el efecto y estabilidad frente a cambios de los mismos; en los últimos nos limitamos a describir los datos y mostrar algunas imágenes que comparan el resultado de nuestro estimador y el que se obtiene mediante la envolvente  $r$ -convexa. En todas las figuras que presentaremos hemos representado en rojo el borde de la envolvente  $r$ -convexa de la muestra. En los gráficos  $N$  indica el número de iteraciones completas del algoritmo, hemos usado la notación  $d_1 = d_\mu(S, \tilde{\mathcal{C}}_{\rho,h}(\mathfrak{N}_n))$ ,  $d_2 = d_\mu(S, C_r(\mathfrak{N}_n))$ . Las distancias en medida son estimadas por el método de montecarlo, en base a 2000 uniformes para el ejemplo 1 y 4, y 4000 para el 3.

1. En el primer ejemplo tomamos (a la izquierda en la Figura 3.10) el conjunto  $\pi/4$  cono convexo  $S = [0, 1] \times [0, t + 1/2] \setminus T$  donde  $t = \frac{1}{2} \tan(3\pi/8)$ ,  $T$  es el triángulo isósceles con vértices  $(0, t + 1/2)$ ,  $(1, t + 1/2)$   $(1/2, 1/2)$ . El tamaño muestral es  $n = 500$ . Para la envolvente cono convexa hemos usado  $\rho = \pi/4$  y  $h = 1/2$  con  $N = 200$  iteraciones. Para la envolvente  $r$ -convexa (en rojo) tomamos  $r = 1/2$  a izquierda y  $r = 1/4$  en el centro. La elección de estos parámetros esta motivada por el hecho de que, aún en conjuntos muy simples, la presencia de una entrante conduce a situaciones en las cuales la envolvente  $r$ -convexa no es adecuada, ya que produce un “sobresuavizado” del conjunto. En las tablas 3.5 y 3.6 se muestra el resultado de algunas simulaciones que hemos realizado, en las mismas hemos tomado el conjunto  $S = [0, 1] \times [0, t + 1/2] \setminus T$  donde  $t = \frac{1}{2} \tan((\pi - \rho_0)/2)$  y  $T$  es el triángulo con vértices  $(0, t + 1/2)$   $(1/2, 1/2)$  y  $(1, t + 1/2)$  para diferentes valores de los parámetros.

Respecto al tiempo computacional del algoritmo, hemos obtenido un tiempo promedio de ejecución (para el conjunto antes mencionado, con  $\rho = \pi/4$ ,  $h = r = 1/4$  y el procesador Intel i7-2620M) de 36,453 segundos. La desviación estándar del tiempo fue: 3,362 segundos.

2. En el segundo ejemplo (a la derecha en la Figura 3.10) tomamos como datos una muestra de puntos que representan la posición de cañas de moras en un terreno de 9 metros cuadrados. Para nuestro estimador hemos tomado  $\rho = \pi/3$  y  $h = 1/8$ , y para la envolvente  $r$ -convexa  $r = 1/10$ . Estos datos fueron analizados por Hutchings (ver Hutchings (1979)) y luego por Diggle (ver Diggle (1983)) entre otros (*Hutchings' Bramble Canes* en la librería *spatstat* de R). En este caso la envolvente cono convexa detecta mejor en general la ausencia de cañas en algunas áreas.
3. En el tercer ejemplo vamos a considerar el conjunto  $S_1$  de la Figura 3.11,  $S_1 = [0, 1]^2 \setminus (T_1 \cup T_2 \cup T_3 \cup T_4)$  donde:  $T_1$  es el triángulo con vértices  $(0, 1)$   $(1/2, 1/2)$  y  $(1, 1)$ .  $T_2$  es el triángulo con vértices  $(0, 0)$   $(1/2, 1/2)$  y  $(1, 0)$ .  $T_3$  es el triángulo con vértices  $(0, 1/3)$   $(1/2, 1/2)$  y  $(0, 2/3)$ .  $T_4$  es el triángulo con vértices  $(1, 1/3)$   $(1/2, 1/2)$  y  $(1, 2/3)$ . Es fácil ver que dicho conjunto es  $\rho_0 = 2 \arctan(1/3)$  cono convexo. Como se ve en la tabla 3.7, se obtienen mejores resultados para el cierre cono convexo por complementos que en el ejemplo 1. Esto se debe a que hemos agregado dos entrantes lineales y reducido así la medida de Lebesgue del conjunto  $S$ , con lo cual el tamaño muestral necesario para obtener mejores resultados que con el cierre  $r$ -convexo es mas pequeño.
4. Vamos a mostrar el resultado de la aplicación del algoritmo para el caso en que el conjunto es muy irregular en el borde, y, al mismo tiempo, ilustrar cómo en cierta familia de conjuntos es posible incorporar información de la misma para mejorar las estimaciones. Hemos tomamos como conjunto el hipografo de una poligonal construida a partir del proceso browniano (recordemos que si una función  $f : [0, 1] \rightarrow \mathbb{R}$  es positiva, su hipografo se define como  $S = \{(x, y) : 0 \leq y \leq f(x), x \in [0, 1]\}$ ). Para generar dicha poligonal tomamos una partición de

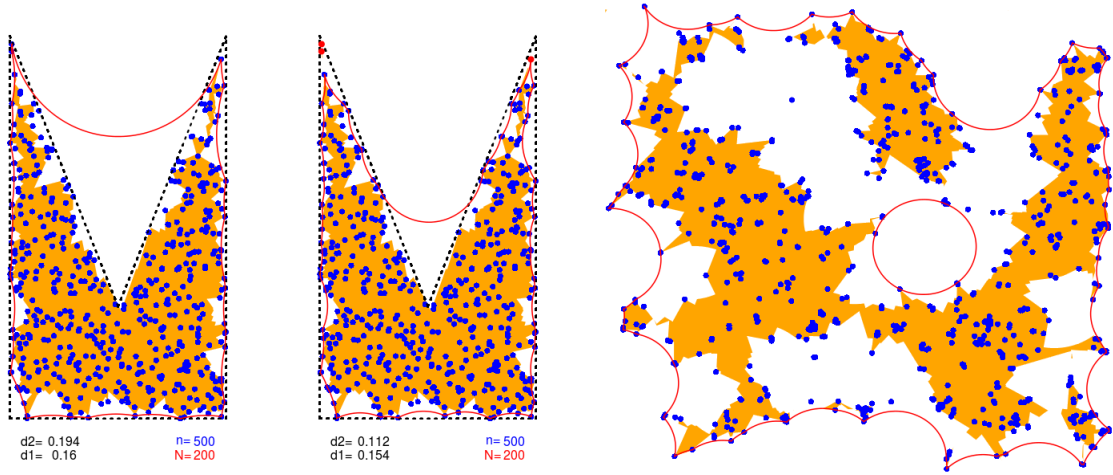


Figura 3.10: Izquierda: El cierre  $\pi/4, 1/4$ - como convexo por complementos, y en rojo el borde de la envolvente  $1/2$ -convexa. Centro: Idem, con  $r = 1/4$ . Derecha: El cierre  $\pi/3, 1/8$ - como convexo por complementos, y en rojo el borde de la envolvente  $1/10$ -convexa para los datos *Bramblecanes* de la librería *spatstat*.

$[0, 1]$  en  $m$  puntos:  $0 = t_0 < t_1 < \dots < t_{m-1} = 1$  con  $t_i - t_{i-1} = 1/m \forall i = 1, \dots, m-1$  y generamos  $m$  variables i.i.d. con distribución normal con media 0 y varianza  $1/m$ . A efectos de que sea más fácil la visualización hemos reescalado el proceso, para que el máximo sea 1, y el mínimo 0. Observemos que nuestro algoritmo permite incorporar la información de que el conjunto es un hipografo, basta tomar en el Paso 1 las direcciones del eje entre  $[\pi/2 - \rho/2, \pi/2 + \rho/2]$  siendo  $\rho$  el ángulo para la envolvente  $\rho$  como convexa. En la Figura 3.12 se muestra el resultado de aplicar el algoritmo con  $\rho = \pi/6, h = 1$  para la envolvente como convexa por complementos, y  $r = 1/8$  para la  $r$ -convexa. Hemos tomado  $m = 200$  para la partición y  $n = 500$  puntos distribuidos uniformemente en el hipografo. Para el algoritmo tomamos  $N = 300$  conos. Como se ve en la figura, la estimación por medio de la envolvente convexa por complementos no sólo es visualmente mejor sino que  $d_1$  es más chico que  $d_2$ . En este caso lo que sucede es que si achicamos  $r$  la envolvente  $r$ -convexa da como resultado un conjunto que no es un hipografo. Como se ve comparando  $d_1$  con  $d_2$ , no se reduce sustancialmente la distancia en medida.





$$\rho_0 = \pi/4$$

$n$	$d_1$	$d_2$	$d_1$	$d_2$
	$\rho = \pi/4, h = 1/2$	$r = 1/6$	$\rho = \pi/4, h = 2$	$r = 1/4$
200	0.293 (0.021)	0.239 (0.018)	0.269 (0.019)	0.163 (0.016)
300	0.234 (0.017)	0.224 (0.015)	0.216 (0.016)	0.136 (0.013)
400	0.198 (0.014)	0.217 (0.013)	0.183 (0.014)	0.119 (0.011)
500	0.173 (0.012)	0.213 (0.012)	0.162 (0.013)	0.110 (0.010)
1000	0.113 (0.010)	0.201 (0.010)	0.107 (0.009)	0.086 (0.007)
1500	0.087 (0.008)	0.197 (0.009)	0.084 (0.008)	0.076 (0.006)
2000	0.072 (0.007)	0.194 (0.009)	0.070 (0.007)	0.071 (0.006)
$n$	$d_1$	$d_2$	$d_1$	$d_2$
	$\rho = \pi/6, h = 1$	$r = 1/3$	$\rho = \pi/10, h = 2$	$r = 1/6$
200	0.327 (0.021)	0.174 (0.015)	0.389 (0.022)	0.174 (0.017)
300	0.260 (0.017)	0.151 (0.012)	0.313 (0.018)	0.139 (0.014)
400	0.219 (0.015)	0.139 (0.011)	0.262 (0.017)	0.117 (0.011)
500	0.189 (0.013)	0.132 (0.010)	0.227 (0.015)	0.104 (0.010)
1000	0.120 (0.010)	0.112 (0.008)	0.143 (0.011)	0.073 (0.008)
1500	0.092 (0.008)	0.105 (0.008)	0.107 (0.009)	0.061 (0.006)
2000	0.075 (0.007)	0.101 (0.007)	0.087 (0.008)	0.053 (0.006)

Tabla 3.5: Media y Desviación estándar en base a 500 repeticiones

$$\rho_0 = \pi/6$$

$n$	$d_1$	$d_2$	$d_1$	$d_2$
	$\rho = \pi/6, h = 1/2$	$r = 1/2$	$\rho = \pi/6, h = 1/3$	$r = 1/6$
200	0.352 (0.020)	0.307 (0.023)	0.417 (0.023)	0.204 (0.018)
300	0.279 (0.018)	0.295 (0.018)	0.306 (0.019)	0.161 (0.014)
400	0.236 (0.015)	0.288 (0.016)	0.245 (0.017)	0.137 (0.011)
500	0.207 (0.014)	0.284 (0.016)	0.210 (0.013)	0.123 (0.011)
1000	0.132 (0.011)	0.274 (0.012)	0.134 (0.010)	0.088 (0.008)
1500	0.102 (0.009)	0.270 (0.011)	0.101 (0.009)	0.074 (0.007)
2000	0.085 (0.008)	0.267 (0.011)	0.083 (0.008)	0.066 (0.006)

Tabla 3.6: Media y Desviación estándar en base a 500 repeticiones

$n$	$d_1$	$d_2$	$d_1$	$d_2$
	$\rho = \rho_0, h = 1/3$	$r = 1/4$	$\rho = \pi/5, h = 1/2$	$r = 1/6$
200	0.204 (0.011)	0.191 (0.009)	0.197 (0.011)	0.161 (0.010)
400	0.138 (0.009)	0.180 (0.008)	0.134 (0.010)	0.140 (0.008)
600	0.107 (0.008)	0.174 (0.007)	0.105 (0.008)	0.132 (0.007)
800	0.090 (0.007)	0.172 (0.007)	0.089 (0.007)	0.127 (0.007)
1000	0.080 (0.007)	0.170 (0.007)	0.078 (0.007)	0.124 (0.006)
1200	0.070 (0.006)	0.169 (0.006)	0.070 (0.006)	0.122 (0.006)
10000	0.022 (0.004)	0.165 (0.006)	0.024 (0.004)	0.112 (0.005)

Tabla 3.7: Media y Desviación estándar en base a 500 repeticiones

## Capítulo 4

# Ideas geométricas en el análisis de datos funcionales y regresión no paramétrica

En este capítulo se consideran diferentes problemas con dos rasgos en común: primero, en todos ellos se desea estimar la función de regresión  $\eta(x) = \mathbb{E}(Y|X = x)$ , donde  $Y$  es una v.a real y  $X$  un elemento aleatorio que toma valores en un espacio métrico  $(\mathcal{E}, m)$  separable y completo. Segundo, la distancia  $m$  definida en este espacio está inspirada en criterios geométricos de cercanía “visual” entre conjuntos (en particular, los hipografos de funciones). En todos los casos la información muestral consiste en una muestra i.i.d.  $(X_1, Y_1), \dots, (X_n, Y_n)$  de elementos aleatorios con la misma distribución de  $(X, Y)$ .

*El problema de clasificación.*

Se dedicará especial atención al problema de clasificación binaria en el cual  $Y$  toma valores en  $\{0, 1\}$  (típicamente  $Y$  indica la “clase” a la que pertenece el elemento  $X$ ) y, por tanto,  $\eta(x) = \mathbb{P}(Y = 1|X = x)$ . El objetivo aquí es predecir la clase  $Y$  de una nueva observación  $x$  mediante un “clasificador”, o “regla de clasificación”  $g(x)$ , donde  $g$  es una función con valores en  $\{0, 1\}$ . En la práctica, esta regla de clasificación se construye a partir del conocimiento de una “muestra de entrenamiento”  $(X_1, Y_1), \dots, (X_n, Y_n)$  (“training sample”) de observaciones en las que la clase  $Y_i$  es conocida sin error. Es sabido [ver Devroye et al. (1996)] que el clasificador óptimo (la llamada “regla Bayes”), respecto al criterio natural de minimizar el error de clasificación  $\mathbb{P}(g(X) \neq Y)$ , viene dado por

$$g(x) = \mathbb{1}_{\{\eta(x) > 1/2\}}.$$

Dado que  $\eta(x)$  es usualmente desconocida, un procedimiento natural (enfoque “plug-in”) que consiste en reemplazar  $\eta(x)$  por un estimador  $\eta_n(x)$  basado en los datos muestrales. Un estudio muy completo del problema de clasificación para datos funcionales puede encontrarse en el trabajo de Baíllo y otros (2010), incluido en el libro colectivo editado por Ferraty y Romain, (2011).

*El caso de la espectrometría: la utilidad de una distancia “visual”.*

Un situación de gran importancia práctica que involucra problemas de clasificación binaria con datos funcionales surge en el diagnóstico médico basado en la observación de espectros (de resonancia magnética, de masas, etc.) obtenidos a partir de el plasma sanguíneo o los tejidos de los pacientes. En estos casos, la variable  $Y$  sería el diagnóstico (sano/enfermo) y el elemento aleatorio  $X$  es el correspondiente espectro. En el caso de los espectros de masas y también en los de resonancia

magnética, las correspondientes funciones son típicamente muy abruptas con picos muy agudos en los que se concentra una buena parte de la información. Para cada una de las clases, (por ejemplo pacientes sanos o enfermos) los picos están ubicados en distintas posiciones. Sin embargo, para elementos de una misma clase las abscisas de los picos están cerca entre los individuos pero no exactamente en la misma posición: necesitamos, por tanto, una métrica en la que la distancia entre dos aproximaciones a “deltas de Dirac” en puntos cercanos sea pequeña. Las métricas usuales en  $L^1, L^2, L^\infty$  no son adecuadas para este propósito. Necesitamos de una métrica “visual” que tenga en cuenta no sólo la distancia “integral” entre los grafos o su distancia en vertical, sino que mida la distancia entre los grafos “en todas las direcciones”. Intentaremos mostrar aquí que una métrica basada en la distancia de Hausdorff entre los hipografos cumple con este objetivo.

*Regresión funcional no paramétrica: algunas ideas básicas sobre el comportamiento asintótico*

Como se deduce de la discusión anterior, el objetivo básico (incluso en el problema de clasificación) es estimar la función de regresión  $\eta(x) = \mathbb{E}(Y|X = x)$ . Es natural preguntarse qué tipo de estimadores de  $\eta$  pueden usarse cuando la “variable auxiliar”  $x$  es una función. Una clase muy general de estimadores de  $\eta$ , considerada (para el caso  $\mathcal{E} = \mathbb{R}^d$ ) por Stone (1977), tienen la forma

$$\eta_n(X) = \sum_{i=1}^n W_{ni}(X)Y_i. \quad (4.1)$$

Bajo hipótesis débiles sobre las funciones de pesos  $W_{ni}$  (que describiremos más adelante) Stone obtuvo la convergencia en media cuadrática,  $\lim_{n \rightarrow \infty} \mathbb{E}((\eta_n(X) - \eta(X))^2) = 0$ , sin restricción alguna sobre la distribución de  $(X, Y)$  (excepto  $\mathbb{E}(\eta^2(X)) < \infty$ ).

Para el caso de clasificación, de dicho resultado se deduce que la regla de clasificación “plug-in”  $g_n(x) = \mathbb{I}_{\{\eta_n(x) > 1/2\}}$ , es consistente, en el sentido de que su error de clasificación  $P(g_n(X) \neq Y)$  tiende, en probabilidad, al error de Bayes  $P(g(X) \neq Y)$ .

El caso en que  $X$  toma valores en un espacio general (no finito-dimensional) es mucho más delicado en lo que respecta a las condiciones para asegurar la consistencia.

Si bien es posible, cuando  $X$  toma valores en un espacio métrico general, probar el mismo resultado de consistencia en media cuadrática de estimadores de la forma (4.1), con la hipótesis adicional de que el espacio sea separable y localmente compacto (ver Teorema 3.1 en Forzani y otros (2012)), una de las hipótesis que en dimensión finita es cierta para los estimadores clásicos basados en núcleos y vecinos más cercanos no es fácil de verificar en dimensión infinita. Sin embargo, si el espacio es separable, y además la función  $\eta$  es acotada y satisface una condición de diferenciación similar a la conclusión del clásico teorema de diferenciación de Lebesgue, es posible probar la consistencia. Esta condición, denominada “Condición de Besicovich” (relacionada al clásico teorema de cubrimiento de Besicovich), será objeto de estudio en profundidad más adelante. Dicha condición se verifica por ejemplo si  $\eta$  es continua.

En el problema de clasificación binaria, la hipótesis de separabilidad junto con la “Condición de Besicovich” garantizan la consistencia (débil) del clasificador por el método de vecinos más cercanos (ver Cérou y Guyader (2006)).

*Principales aportaciones de este capítulo*

### 1. Definición y estudio de las propiedades de un nuevo espacio métrico de funciones

En la primera sección veremos que en el espacio de las funciones semicontinuas superiormente, a valores reales, definidas en un intervalo finito (que denotaremos  $\mathcal{E}$ ), puede definirse una métrica  $\mathbb{H}$  (definida como la distancia de Hausdorff entre los hipografos) que proporciona un espacio  $(\mathcal{E}, \mathbb{H})$  en el que los conjuntos cerrados y acotados son compactos y, por tanto, es localmente compacto. Como es sabido, esta propiedad no se verifica en ninguno de los espacios

funcionales de uso común ( $L^p$ ,  $L^\infty$ , ...). Resultará, además, que el espacio  $(\mathcal{E}, \mathbb{H})$  es separable y completo.

2. Un algoritmo para el cálculo de la distancia

En la Sección 2 daremos algunas caracterizaciones y propiedades de la distancia  $\mathbb{H}$  y propondremos un algoritmo para su cálculo en el caso de funciones continuas.

3. Aplicaciones prácticas a problemas de clasificación binaria con espectros (de masas o de resonancia magnética).

La sección 3 está dedicada a estudiar el funcionamiento práctico de la nueva distancia  $\mathbb{H}$  cuando se utiliza (comparada con la distancias clásicas  $L^2$  y  $L^\infty$ ) para definir clasificadores  $k$ -NN en tres problemas reales de clasificación binaria con datos funcionales. Dos de ellos corresponden a problemas en Medicina (cardiología y oncología) y el tercero corresponde a la clasificación de dos variedades de café. En los dos ejemplos médicos, el funcionamiento de la distancia  $\mathbb{H}$  es claramente superior. En el problema de las variedades de café, la distancia del supremo  $L^\infty$  muestra una ligera superioridad.

4. Análisis de la consistencia: condición de Besicovitch.

Como hemos mencionado más arriba, el principal problema teórico asociado con el uso de estimadores de la regresión (o clasificadores) de tipo  $k$ -NN en espacios abstractos es la posible falta de consistencia. La condición de Besicovitch (que es válida automáticamente en dimensión finita) es un requerimiento natural (y no trivial) en espacios abstractos. En la sección 4 demostramos que nuestro espacio  $(\mathcal{E}, \mathbb{H})$  no es  $\sigma$ -finito-dimensional en el sentido de Preiss (1983) y, como consecuencia, la condición de Besicovitch no se verifica automáticamente para toda función del espacio. La conclusión práctica es que debemos imponer la condición de Besicovitch para garantizar la consistencia.

5. Aplicaciones a problemas de regresión no paramétrica (finito-dimensional) con restricciones geométricas de forma.

Finalmente, en la última sección abordaremos un problema diferente. Consideramos el modelo de regresión no paramétrica clásico

$$Y = \eta(X) + e,$$

con  $Y, X \in \mathbb{R}$ . El objetivo es obtener estimadores de la función de regresión  $\eta$  con restricciones geométricas de forma. En lugar de poner restricciones de regularidad analítica, como por ejemplo en la teoría de splines, buscamos imponer restricciones geométricas de forma sobre el gráfico de la función. Más precisamente, definiremos el espacio de las funciones  $\rho$ -cono convexas  $\mathcal{E}_\rho \subset \mathcal{E}$  como el de aquellas funciones semicontinuas superiormente y no negativas cuyo hipografo es un conjunto  $\rho$ -cono convexo. Probaremos que  $\mathcal{E}_\rho$  es un subconjunto localmente compacto de  $\mathcal{E}$ , y daremos un algoritmo que basado en  $n$  observaciones

$$(t_1, f(t_1)), \dots, (t_n, f(t_n)),$$

de una función  $f : [0, 1] \rightarrow \mathbb{R}$  arbitraria, obtener una aproximación  $f_\rho^n$  de  $f$ .

## 4.1. Definición y estudio de las propiedades de un nuevo espacio métrico de funciones

Al comienzo de esta sección veremos algunas métricas clásicas definidas en espacios de funciones. Luego introduciremos una nueva distancia, basada en la distancia de Hausdorff entre conjunto, que será la utilizada en el resto del capítulo. Para dicha distancia probaremos algunas propiedades que permitirán en la siguiente sección obtener un algoritmo para calcularla.

### 4.1.1. Distancias “visuales” entre funciones: Un poco de historia

Como señala A.N. Shyriaev en el libro “Kolmogorov in perspective” (ver Shyriaev (2007)) ya en el año 1948, en un encuentro de la Sociedad Matemática de Moscú, Kolmogorov propone la idea de considerar las distribuciones de los procesos estocásticos como medidas en un álgebra de Borel en algún espacio funcional. En este contexto, señala, es natural considerar la convergencia en probabilidad de procesos aleatorios como la convergencia débil de las medidas en el espacio funcional, asociadas a los mismos. En 1955 Skorohod define una métrica en el espacio  $D$  de las funciones càdlàg (continuas por derecha y con límite por izquierda) que lo hace un espacio separable. En 1956, en un trabajo titulado “On Skorohod convergence” Kolmogorov muestra que existe una métrica en  $D$ , equivalente a la definida por Skorohod, (construida de forma explícita por su estudiante Yu. V. Prokhorov poco después) que hace que dicho espacio sea además completo. Tales métricas, así como las distancias  $L^1$  y  $L^2$  no son fácilmente interpretables desde un punto de vista visual, como se señala en Marron y Tsybakov (1995) y en Cuevas y Fraiman (1998). No obstante, a diferencia de la métrica del supremo, la métrica de Prokhorov tiene la importante propiedad de ser una metrización de la convergencia débil (ver Billingsley (1999)). Dadas las aplicaciones en el tratamiento de imágenes y en la teoría de la aproximación, vamos a definir una métrica (la distancia de Hausdorff entre los hipografos) que haga que el espacio de las funciones semicontinuas superiormente sea localmente compacto.

#### Métrica de Skorohod

Como se indica en Billingsley (1999) dos funciones  $f$  y  $g$  continuas, definidas en  $[0, 1]$  están próximas en la métrica uniforme ( $\|f - g\|_\infty$ ) si el gráfico de  $f$  puede *llevarse* uniformemente al gráfico de  $g$  a través de una pequeña perturbación uniforme de las ordenadas, dejando fija la abscisa. La métrica de Skorohod que definiremos a continuación contempla además, pequeñas perturbaciones uniformes en las ordenadas.

**Definición 4.1.** Una función  $f : [0, 1] \rightarrow \mathbb{R}$  se dice que es una función càdlàg si para todo  $t \in [0, 1]$

1.  $\exists f(t^-) = \lim_{x \rightarrow t^-} f(x)$
2.  $\exists f(t^+) = \lim_{x \rightarrow t^+} f(x) = f(t)$

Es decir, existe el límite por izquierda y es continua por derecha. Denotaremos  $D$  al espacio de las funciones càdlàg.

**Definición 4.2.** Denotemos  $\Lambda$  al conjunto de las biyecciones continuas, crecientes, de  $[0, 1]$  en  $[0, 1]$ . Sean  $f, g \in D$  dos funciones càdlàg, la distancia de Skorohod entre  $f$  y  $g$  es:

$$d_S(f, g) = \inf_{\lambda \in \Lambda} \max \{ \|\lambda - Id\|_\infty, \|f - g \circ \lambda\|_\infty \}.$$

Es fácil ver que la métrica  $d_S$  restringida al espacio de las funciones continuas  $C[0, 1]$  coincide con la métrica uniforme. Asimismo, se puede demostrar que  $(D, d_S)$  es separable (ver Billingsley (1999)). Una observación interesante, que muestra la utilidad de la distancia  $d_S$  es la siguiente: si consideramos  $U_1, \dots, U_n$  i.i.d uniformes en  $[0, 1]$  el proceso empírico  $X_n : ([0, 1], \mathcal{B}, \mu)^n \rightarrow (D, d_S)$  es medible, donde  $\mathcal{B}$  denota la sigma álgebra de Borel en  $[0, 1]^n$ ,  $\mu$  es la medida de Lebesgue y

$$X_n(t) = \sqrt{n}(F_n(t) - t) \quad 0 \leq t \leq 1,$$

siendo

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{[0,t]}(U_i) \quad 0 \leq t \leq 1.$$

Este resultado de medibilidad no es cierto si tomamos en  $D$  la métrica dada por la distancia del supremo.

### Métrica de Prokhorov-Lévy

Si bien la métrica de Skorohod tiene propiedades que la hacen interesante y útil desde el punto de vista de la medida, no es fácil su interpretación visual. Si nos restringimos a las funciones de distribución, una alternativa *más visual* (definida por Levy en 1937) es la siguiente:

**Definición 4.3.** Sean  $F$  y  $G$  dos distribuciones de probabilidad en  $\mathbb{R}$ , la distancia de Lévy entre  $F$  y  $G$  se define como

$$d_L(F, G) = \inf \{ \varepsilon > 0 : F(x - \varepsilon) - \varepsilon \leq G(x) \leq F(x + \varepsilon) + \varepsilon \quad \forall x \in \mathbb{R} \}.$$

En 1956 Prokhorov generaliza dicha métrica al conjunto de las medidas de probabilidad de modo tal que  $d_L(F_n, F) \rightarrow 0$  para  $F_n$  y  $F$  funciones de distribución, si y sólo si las medidas asociadas a dichas distribuciones convergen con la métrica de Prokhorov. Se puede demostrar que  $d_L$  es una metrización de la topología débil en el conjunto de las medidas de probabilidad en  $\mathbb{R}$ . La métrica de Lévy tiene una interpretación visual simple:  $\sqrt{2}d_L(F, G)$  coincide con el máximo de las distancias entre  $F$  y  $G$  medidas en la dirección de  $(-1, 1)$ .

#### 4.1.2. Una métrica funcional basada en la distancia de Hausdorff

Antes de definir la métrica que será la que usaremos a lo largo de este capítulo es necesario introducir algunas definiciones y notación.

**Definición 4.4.** Sea  $T$  un espacio topológico y  $f : D \rightarrow \mathbb{R}$  con  $D \subset T$ . Se dice que  $f$  es semicontinua superiormente en  $x_0 \in T$  si para todo  $\varepsilon > 0$  existe  $U$  entorno de  $x_0$  tal que  $f(x) \leq f(x_0) + \varepsilon$  para todo  $x \in U$ . Se dice simplemente que  $f$  es semicontinua superiormente (que abreviaremos USC por sus siglas en inglés) si lo es en cada punto de su dominio. En el caso en que  $T$  es un espacio métrico, es inmediato observar que  $f$  es semicontinua superiormente en  $x_0$  si y solo si  $\limsup_{x \rightarrow x_0} f(x) \leq f(x_0)$ .

**Definición 4.5.** Sea  $f : [0, 1] \rightarrow \mathbb{R}$  USC, no negativa, el hipografo de  $f$ , que denotaremos  $H_f$  se define como

$$H_f = \{ (x, y) \in \mathbb{R}^2 : x \in [0, 1], 0 \leq y \leq f(x) \}.$$

La siguiente proposición establece dos resultados conocidos de las funciones USC y será de utilidad más adelante, una prueba de la misma puede encontrarse por ejemplo en Natanson (1960).

**Proposición 4.6.** Sea  $f : [0, 1] \rightarrow \mathbb{R}$  no negativa y USC entonces

- 1) Para todo  $K \subset [0, 1]$  compacto, existe  $z \in K$  tal que  $\sup_{x \in K} f(x) = f(z)$ .
- 2)  $H_f$  es compacto.

**Definición 4.7.** Sean  $g : [0, 1] \rightarrow \mathbb{R}$  y  $f : [0, 1] \rightarrow \mathbb{R}$  USC, no negativas, definimos  $\mathbb{H}(f, g) = d_H(H_f, H_g)$ . Denotamos  $(\mathcal{E}, \mathbb{H})$  el espacio de las funciones no negativas, USC dotado con esta métrica.

Observemos que la distancia definida en 4.7 es invariante por traslaciones, no obstante nos restringimos al caso donde las funciones son positivas. Si queremos que el espacio sea localmente compacto es necesario que consideremos funciones o bien positivas, o, con mayor generalidad, uniformemente acotadas inferiormente (en este caso la distancia 4.7 se define como antes, trasladando ambas para que sean positivas). Para ver que es necesario que sean uniformemente acotadas inferiormente para obtener la compacidad local veamos el siguiente ejemplo: consideremos la sucesión de funciones  $f_n(x) = -nx$  si  $x \in [0, 1)$  y  $f_n(1) = 0$  para  $n = 1, 2, \dots$ . Es claro que  $f_n \subset B(f_1, 1)$  para  $n = 1, 2, \dots$ , pero no existe subsucesión convergente de  $f_n$ . Si bien así definida es una distancia entre funciones, veremos luego que hereda propiedades de la distancia de Hausdorff entre subconjuntos del plano. En lo que respecta al uso de esta distancia en estadística cabe destacar el trabajo de Cuevas y Fraiman (1998) en el cual se halla la tasa de convergencia para la distancia  $\mathbb{H}$  de estimadores basados en

núcleos de una densidad desconocida  $f$ . Variantes de la misma se encuentran en la literatura en trabajos que la relacionan con la teoría de la aproximación. Dadas las propiedades que tienen las funciones convexas en cuanto a optimización (ver Roberts y Varberg (1973) capítulo 5) usualmente se asume que la función a optimizar es convexa, y se toma como distancia entre funciones la distancia de Hausdorff entre los epígrafos (i.e:  $\{y : f(x) \leq y\}$ ). El interés de esta definición en el campo de la optimización está, entre otras cosas, en el hecho de que tanto los epígrafos como los conjuntos de nivel de una función convexa, son subconjuntos convexos del plano. Esto permite aproximar los mínimos de las funciones a través de los mínimos de las aproximaciones (ver Rockafellar y Wets (2009) Teorema 7.69). Un estudio muy completo del surgimiento y evolución del uso de la epiconvergencia en diferentes áreas de la matemática puede encontrarse en las páginas 292 a 297 en Rockafellar y Wets (2009).

En cuanto a hipógrafos, en Sendov (1990) se definen versiones análogas a la dada en 4.7 para *funciones segmento*. Veremos brevemente la construcción que allí se realiza y las diferencias con la definición 4.7. Consideremos  $S(\overline{\mathbb{R}})$  el conjunto de todos los intervalos de la forma  $[a, b]$  con  $a, b \in \overline{\mathbb{R}}$ . Denotemos  $\mathcal{A}_\Omega = \{f : \Omega \rightarrow S(\overline{\mathbb{R}})\}$  con  $\Omega \subset \overline{\mathbb{R}}$  cerrado. Observemos que este conjunto incluye como caso particular las funciones que toman en cada  $x$  un valor  $f(x) \in \mathbb{R}$ , a esta clase la denotaremos  $\mathcal{U}_\Omega$ . Sea  $\delta > 0$ , definimos:

$$I(\delta, f, x) = \inf \{y \in f(t) : t \in [x - \delta, x + \delta] \cap \Omega\},$$

$$S(\delta, f, x) = \sup \{y \in f(t) : t \in [x - \delta, x + \delta] \cap \Omega\},$$

y, tomando límite, obtenemos las funciones:  $I(f, x) = \lim_{\delta \rightarrow 0} I(\delta, f, x)$  y  $S(f, x) = \lim_{\delta \rightarrow 0} S(\delta, f, x)$  respectivamente. El *gráfico completo* de una función  $f \in \mathcal{A}_\Omega$  es la función de segmento

$$F(f, x) = [I(f, x), S(f, x)].$$

Denotemos  $\mathcal{F}_\Omega$  al conjunto de las funciones  $f \in \mathcal{A}_\Omega$  tal que  $F(f, x) = f(x)$  para todo  $x \in \Omega$ . Observemos que este conjunto contiene las funciones en  $\mathcal{U}_\Omega$  que son continuas. Teniendo en cuenta estas definiciones, se define la distancia de Hausdorff para funciones en  $\mathcal{F}_\Omega$ . Para eso consideraremos las funciones  $f \in \mathcal{F}_\Omega$  como subconjuntos de  $\mathbb{R}^2$  con la métrica usual (observemos que son cerrados), y tomar la distancia de Hausdorff entre dichos conjuntos, esto es:

$$r(f, g) = \max \left\{ \sup_{s \in f} \inf_{t \in g} \|s - t\|; \sup_{s \in g} \inf_{t \in f} \|s - t\| \right\}. \quad (4.2)$$

Esto define una métrica en  $\mathcal{F}_\Omega$ . Observemos que, si ambas funciones están en  $\mathcal{U}_\Omega$ ,  $r(f, g)$  no es igual a  $\mathbb{H}(f, g)$  siendo  $\mathbb{H}$  la distancia definida en 1.1. Esto último es fácil de ver, basta tomar  $\Omega = [0, 1]$   $g(x) = 1$  para todo  $x \in \Omega$  y  $f(x) = x$ . En este caso es fácil ver que  $r(f, g) = 1$  mientras que  $\mathbb{H}(f, g) = 1/\sqrt{2}$ . No obstante si consideramos ambas funciones como funciones de segmento (por ejemplo  $f$  sería la función en  $\mathcal{F}_\Omega$  que manda  $x$  en  $[0, f(x)]$ )  $r(f, g) = \mathbb{H}(f, g)$ .

En Sendov (1990) se toma una generalización de la métrica anterior, cambiando la métrica dada por la distancia euclidiana  $\|\cdot\|$  por  $\rho_\alpha(x, y) = \max\{x/\alpha, y\}$  con  $x, y \in \mathbb{R}^2$  (métrica que se denota  $r(\alpha, f, g)$ ). Esta nueva métrica es equivalente a la métrica uniforme, si nos restringimos al espacio de las funciones de  $\mathcal{U}_\Omega$  (lo cual no es cierto para la métrica  $\mathbb{H}$ ).

En el trabajo de Holá (1992) se presenta una generalización de las funciones de segmento. Se consideran dos espacios métricos  $(X, d_X)$  e  $(Y, d_Y)$  y se introduce el concepto de *multifunción*, esto es: una función cuyo valor en  $x \in X$  es un elemento de las partes de  $Y$ . El gráfico de una multifunción  $H$  se define como  $\{(x, y) : x \in X, y \in H(x)\}$ . En este contexto una multifunción es semicontinua superiormente en  $z \in X$  si para todo abierto  $V$  que contiene a  $H(z)$  el conjunto  $\{x : H(x) \subset V\}$  contiene un abierto que contiene a  $z$ . El conjunto de las multifunciones semicontinuas superiormente de  $X$  en  $Y$  se denota  $U(X, Y)$ . La métrica en  $X \times Y$  definida como:

$$\rho((x_1, x_2), (y_1, y_2)) = \max \left\{ d_X(x_1, x_2), d_Y(y_1, y_2) \right\},$$

induce una métrica  $h_\rho$  en  $U(X, Y)$  si identificamos los elementos de  $U(X, Y)$  con sus gráficos:

$$h_\rho(f, g) = \max \left\{ \sup_{x \in X} \sup_{z \in f(x)} \inf_{y \in Y} \inf_{t \in g(y)} \rho(z, t); \sup_{x \in X} \sup_{z \in g(x)} \inf_{y \in Y} \inf_{t \in f(y)} \rho(z, t) \right\}.$$

En el citado trabajo se demuestra que si  $Y$  es completo, la compacidad local de  $X$  es equivalente a que  $(U(X, Y), h_\rho)$  sea un subespacio cerrado del espacio métrico  $(CL(X \times Y), h_\rho)$  de los subconjuntos cerrados de  $X \times Y$ , con la métrica  $h_\rho$ .

Veamos la relación entre la distancia antes definida y las métricas usuales entre funciones.

**Observación 4.8.** Consideremos la función  $f : [0, 1] \rightarrow \mathbb{R}$  constante e igual a 1 y la sucesión  $\{f_n\}_n$ ,  $f_n : [0, 1] \rightarrow \mathbb{R}$ , donde  $f_n(x) = nx$  si  $x < 1/n$  y  $f_n(x) = 1$  si  $x \in [1/n, 1]$ , en este caso  $\mathbb{H}(f_n, f) \rightarrow 0$  pero no se verifica que  $f_n(0) \rightarrow f(0)$ , es decir la convergencia en la métrica  $\mathbb{H}$  no implica la convergencia puntual. El recíproco tampoco es cierto, basta considerar la función  $f$  constante e igual a 0, y las funciones  $f_n(x) = x^n(1 - x^n)$ , en este caso  $f_n(1/\sqrt[n]{2}) = 1/4$  para todo  $n$ . Observemos que  $\arg \max f_n(x) = 1/\sqrt[n]{2}$  mientras que  $\arg \max f(x) = [0, 1]$ . Por otro lado, si definimos  $f(x) = 1/4$  si  $x = 1$  y 0 en  $[0, 1)$  tenemos que  $\mathbb{H}(f_n, f) \rightarrow 0$ .

Es fácil ver que la convergencia en  $L^p$  no implica la convergencia en la métrica  $\mathbb{H}$ , basta tomar  $f_n(x) = \mathbb{I}_{[0, 1/n]}$  y  $f = 0$ . El recíproco tampoco es cierto, consideremos la sucesión de funciones

$$I_i^n(x) = \mathbb{I}_{[\frac{i-1}{2^n}, \frac{i}{2^n}]}(x) \quad i = 1, \dots, 2^n,$$

y las funciones

$$f_n(x) = \sum_{k=1}^{2^n-2} I_{2k}^n(x) \quad y \quad f(x) = \mathbb{I}_{[0, 1]}(x),$$

en este caso  $\mathbb{H}(f, f_n) = 1/2^{n+1}$  mientras que, para todo  $p$ ,  $\int_0^1 |f_n(x) - f(x)|^p dx = 1/2$ .

**Proposición 4.9.** Sean  $f_n$  y  $f$  USC, no negativas, tal que  $\mathbb{H}(f_n, f) \rightarrow 0$  entonces

$$\max_{x \in [0, 1]} f(x) = \lim_{n \rightarrow +\infty} \max_{x \in [0, 1]} f_n(x).$$

*Demostración.* Por la proposición anterior sabemos que existe  $z \in [0, 1]$  tal que  $f(z) = \max_{x \in [0, 1]} f(x)$ . Como  $\mathbb{H}(f_n, f) \rightarrow 0$  existe  $x^n = (x_1^n, x_2^n) \in H_{f_n}$  tal que  $x_n \rightarrow (z, f(z))$ , entonces

$$x_2^n \leq f_n(x_1^n) \leq \max_{x \in [0, 1]} f_n(x),$$

y si tomamos límite inferior tenemos que

$$\max_{x \in [0, 1]} f(x) = f(z) \leq \liminf_{n \rightarrow +\infty} \max_{x \in [0, 1]} f_n(x) \leq \limsup_{n \rightarrow +\infty} \max_{x \in [0, 1]} f_n(x),$$

veamos que  $\limsup_n \max_{x \in [0, 1]} f_n(x) \leq \max_{x \in [0, 1]} f(x)$ . Sea  $z_0 = \limsup_{n \rightarrow +\infty} \max_{x \in [0, 1]} f_n(x)$ . Existe  $x_n \in [0, 1]$  tal que  $f_n(x_n) \rightarrow z_0$  y  $f_n(x_n) = \max_{x \in [0, 1]} f_n(x)$ . Podemos, pasando a subsucesiones, suponer que  $x_n \rightarrow x_0 \in [0, 1]$ . Como  $(x_0, z_0) \in H_f$  tenemos que  $f(x) \geq z_0$  y por lo tanto  $\max_{x \in [0, 1]} f(x) \geq z_0$  que es lo que queríamos.  $\square$

**Proposición 4.10.** El espacio  $(\mathcal{E}, \mathbb{H})$  definido en 4.7 es localmente compacto.

*Demostración.* Veremos que un conjunto cerrado y acotado es compacto, y por lo tanto las bolas cerradas (que son una base local) son compactas. Tomemos  $\{f_n\} \subset \mathcal{E}$  una sucesión acotada con la métrica  $\mathbb{H}$ . Observemos que los subconjuntos compactos  $H_{f_n}$  forman una sucesión acotada, por lo tanto existe el límite en distancia de Hausdorff de una subsucesión de la misma, el cual llamaremos  $C$ . Tenemos que demostrar que si  $\{H_{f_n}\}_n$  es una sucesión tal que  $H_{f_n} \rightarrow C$  en distancia de Hausdorff, para algún conjunto  $C$  compacto,  $C = H_f$  para alguna  $f : [0, 1] \rightarrow \mathbb{R}$  USC.



Sea  $(x, y) \in C$  y  $(x_n, y_n) \in H_{f_n}$  que converge a  $(x, y)$  (existe al menos una  $y$  que  $H_{f_n}$  converge en distancia de Hausdorff a  $C$ ), como los  $H_{f_n}$  son hipografos y convergen a  $C$  tenemos que:

$$\left[ (x, 0), (x, \limsup f_n(x_n)) \right] \in C. \quad (4.3)$$

Observemos que como  $f_n(x_n) \geq y_n$ , tomando límite superior en 4.3, tenemos que:  $y \leq \limsup f_n(x_n)$ . Por lo tanto

$$\{(x, z) : 0 \leq z \leq \limsup f(x_n)\} \subset C.$$

Sea  $f : [0, 1] \rightarrow \mathbb{R}^+$  definida como

$$x \mapsto \sup_{\{x_n\}:x_n \rightarrow x} \limsup f_n(x_n).$$

$f$  está bien definida ya que  $\{f_n\}$  es acotada. Veamos que  $C = H_f$ . Como  $C$  es cerrado, por 4.3, tenemos que  $H_f \subset C$ . Además, si  $(x, y) \in C$ , y  $(x_n, y_n) \rightarrow (x, y)$ , con  $(x_n, y_n) \in H_{f_n}$ ,  $f(x) \geq \limsup f_n(x_n) \geq y$ , por lo tanto  $(x, y) \in H_f$ .

Resta ver que  $f$  es USC. Supongamos por reducción al absurdo que existe  $a$  tal que  $\limsup_{x \rightarrow a} f(x) > f(a)$ . Existe  $\delta > 0$  y  $x_n \rightarrow a$ ,  $x_n \neq a$  para todo  $n$ , tal que  $f(x_n) > f(a) + \delta$  si  $n > n_0$ . Por definición de  $f$ ,

$$f(x_n) = \sup_{\{z_k\}:z_k \rightarrow x_n} \limsup f_k(z_k).$$

Podemos tomar  $z_k \rightarrow_k x_n$  una sucesión (que depende de  $n$ ), tal que  $f(x_n) = \lim_{z_k \rightarrow_k x_n} \limsup f_k(z_k)$ . Sea  $\varepsilon > 0$ , para todo  $n > n_0$  sea  $k(n)$  el mínimo  $k$  tal que si  $k > k(n)$  entonces

$$|f_{k(n)}(z_{k(n)}) - \limsup f_k(z_k)| < \varepsilon.$$

es decir, si  $n > n_0$

$$|f_{k(n)}(z_{k(n)}) - f(x_n)| < \varepsilon.$$

Pero como  $z_{k(n)} \rightarrow_n a$  esto contradice que  $f(x_n) > f(a) + \delta$  si  $n > n_0$ . □

**Observación 4.11.** *Observemos que  $(\mathcal{E}, \|\cdot\|_\infty)$  no es localmente compacto.*

**Proposición 4.12.** *El espacio  $(\mathcal{E}, \mathbb{H})$  definido en 4.7 es separable y completo.*

*Demostración.* La completitud se sigue de forma inmediata de la compacidad local. Para ver que es separable consideremos  $P_n$  el conjunto de todas las particiones de  $[0, 1]$  formadas con  $n$  elementos  $0 = x_0 < x_1 < \dots < x_{n-1} = 1$  donde los  $x_i$  son números racionales, y  $\mathcal{P} = \cup_n P_n$ . Observemos que  $\mathcal{P}$  es numerable. Dada una partición  $P \in P_n$  y un conjunto  $q_0, \dots, q_{n-1}$  de números racionales, vamos a definir una función USC constante e igual a  $q_i$  en cada intervalo de la partición  $P$ :

$$f_P(x) = \begin{cases} q_0 & \text{si } x \in [0, x_1) \\ q_i & \text{si } x \in (x_i, x_{i+1}) \quad 1 \leq i \leq n-3 \\ q_{n-1} & \text{si } x \in (x_{n-2}, 1] \\ \max\{q_i, q_{i+1}\} & \text{si } x = x_i \quad 1 \leq i \leq n-2 \end{cases} \quad (4.4)$$

Así definida  $f_P$  es USC y acotada inferiormente. Veamos que el conjunto (numerable) de todas las funciones definidas de esta forma, al variar la partición  $P$  y los valores  $q_i$  es denso en  $\mathcal{E}$  con la métrica  $\mathbb{H}$ . Sea  $f$  USC y acotada inferiormente, y  $\varepsilon > 0$ , podemos suponer sin pérdida de generalidad que  $f$  es positiva. Tomemos  $P \in P_n$  una partición formada por los puntos:  $0 = x_0 < x_1 < \dots < x_{n-1} = 1$  con  $x_i$  racionales y tal que  $\max_{i=0, \dots, n-1} |x_{i+1} - x_i| < \varepsilon/2$ . Por la Proposición 4.6 existe  $f_i = \max_{x \in [x_i, x_{i+1}]} f(x)$ . Tomemos  $q_0, \dots, q_{n-1}$  números racionales tal que  $q_i > f_i$  y  $q_i - f_i < \varepsilon/2$  para todo  $i$ . Para esta partición y este conjunto de números racionales consideremos la función  $f_P$

definida en 4.4. Veamos que  $\mathbb{H}(f_P, f) < \varepsilon$ . Es claro que  $f_P(x) > f(x)$  para todo  $x \in [0, 1]$  con lo cual  $H_f \subset H_{f_P}$ , de donde  $\mathbb{H}(f_P, f) = \sup_{z \in H_{f_P}} d(z, H_f)$ . Sea  $z = (z_1, z_2) \in H_{f_P}$  un punto cualquiera, existe  $0 \leq i_0 \leq n-1$  tal que  $x_{i_0} \leq z_1 \leq x_{i_0+1}$ . Sea  $t$  tal que  $x_{i_0} \leq t \leq x_{i_0+1}$  y  $f_{i_0} = f(t)$ . Observemos que  $z_2 < f(t) + \varepsilon/2$  y por lo tanto  $d(z, H_f) < \varepsilon$ . Como  $z$  es un punto cualquiera en  $H_{f_P}$  tenemos que  $\sup_{z \in H_{f_P}} d(z, H_f) < \varepsilon$ .  $\square$

Veremos a continuación un lema que nos será de utilidad más adelante. El mismo establece que las funciones continuas son densas, con la métrica de Hausdorff entre hipografos, en el espacio  $\mathbb{H}$ .

**Lema 4.13.** *Sea  $f \in (\mathcal{E}, \mathbb{H})$ , para todo  $\varepsilon > 0$  existe  $g$  positiva, continua, tal que  $\mathbb{H}(f, g) < \varepsilon$ .*

*Demostración.* Dado  $\varepsilon > 0$  tomemos  $\mathcal{P}_\varepsilon = \{0 = x_0 < x_1 < \dots < x_n = 1\}$  una partición de  $[0, 1]$  tal que  $\max_{i=1, \dots, n} x_i - x_{i-1} < \varepsilon$ . Definimos para  $i = 0, \dots, n$ ,  $g(x_i) = \limsup_{x \rightarrow x_i} f(x)$ . Sea  $t_i = \max_{x_i \leq x \leq x_{i+1}} f(x)$ , sea  $z_i \in [x_i, x_{i+1}]$  el mayor de los valores donde se realiza dicho máximo (podría coincidir con  $x_i$  o  $x_{i+1}$ ). Definimos  $g$  en los intervalos  $[x_i, x_{i+1}]$  como la poligonal que une los puntos  $(x_i, g(x_i))$ ,  $(z_i, t_i)$  y  $(x_{i+1}, g(x_{i+1}))$ , en caso de coincidir  $(z_i, t_i)$  con alguno de los otros dos,  $g$  se define como el segmento de recta que une  $(x_i, g(x_i))$  con  $(x_{i+1}, g(x_{i+1}))$ . Por la forma en que fue definida  $g$  es claro que es continua y que  $\mathbb{H}(f, g) < \varepsilon$ .  $\square$

## 4.2. Un algoritmo para el cálculo de la distancia

El cálculo mediante algoritmos de la distancia de Hausdorff entre conjuntos tiene aplicaciones en el procesamiento de imagenes digitalizadas y es por lo tanto objeto de estudios recientes. En el software `Matlab` la función `HausdorffDist` calcula la distancia de Hausdorff entre dos subconjuntos de puntos del plano. En Scharf (2003) se estudia el caso particular en que los conjuntos son curvas del plano, parametrizables por funciones racionales. Otra referencia reciente es el trabajo de Nutanong y otros (2011) y el de Alt y otros (2003). En este último los conjuntos son símlices de dimensión  $k$  en  $\mathbb{R}^d$ . Veremos primero algunos resultados que nos permitirán luego proponer un algoritmo para el cálculo de la distancia  $\mathbb{H}$ .

**Observación 4.14.** *Observemos que si  $f$  y  $g$  son funciones USC, no negativas, como  $H_f$  y  $H_g$  son compactos,  $\mathbb{H}(f, g) \leq d_H(\partial H_f, \partial H_g)$  no obstante, aún en el caso en que  $f$  y  $g$  son continuas, la igualdad no es cierta en general, ver Figura 4.1.*

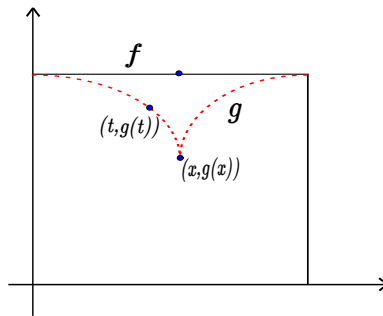


Figura 4.1:  $\mathbb{H}(f, g) = \|(t, g(t)) - (x, f(x))\|$  mientras que  $d_H(\partial H_f, \partial H_g) = \|(x, g(x)) - (x, f(x))\|$

Veamos, no obstante, que en el caso de funciones USC, la distancia de Hausdorff se realiza en puntos del borde del hipografo. Más aún si ambas funciones son continuas dichos punto son de la forma  $(t, f(t))$  y  $(u, g(u))$  con  $t, u \in [0, 1]$ . Esto nos permitirá plantear un algoritmo sencillo de implementar, de orden  $n^2$  para aproximar la distancia de Hausdorff entre funciones continuas.

**Lema 4.15.** Sean  $f, g : [0, 1] \rightarrow \mathbb{R}$  no negativas, USC, existen  $u \in \partial H_f$  y  $v \in \partial H_g$  tal que  $\mathbb{H}(f, g) = \|u - v\|$ .

*Demostración.* Como  $f$  y  $g$  son no negativas, tenemos que, por definición de  $\mathbb{H}$

$$\mathbb{H}(f, g) = d_H(H_f, H_g) = \max \left\{ \sup_{a \in H_f} d(a, H_g), \sup_{b \in H_g} d(b, H_f) \right\}.$$

Supongamos por absurdo que no existen  $u \in \partial H_f$  y  $v \in \partial H_g$  tal que  $\mathbb{H}(f, g) = \|u - v\|$ . Sean  $x = (x_1, x_2)$  e  $y = (y_1, y_2)$  que realizan la distancia  $\mathbb{H}$ , es decir  $\mathbb{H}(f, g) = \|x - y\|$ . Como estamos suponiendo que no se realiza en puntos del borde,  $x$  o  $y$  tienen que pertenecer al interior del hipografo. Por ejemplo  $x \in \text{int}(H_f)$ , en tal caso  $0 < x_1 < 1$ ; veremos que  $d((x_1, f(x_1)), H_g) \geq \mathbb{H}(f, g)$ . Para cada  $t \in [0, 1]$  tal que  $|t - x_1| < \mathbb{H}(f, g)$  denotemos  $u^t = (u_1^t, u_2^t)$  y  $v^t = (v_1^t, v_2^t)$  los puntos de corte de  $\partial B(x, \mathbb{H}(f, g))$  y la recta  $x_1 = t$ , con  $u_2^t < v_2^t$ . Como  $B(x, \mathbb{H}(f, g)) \cap H_g = \emptyset$ , y  $H_g$  es un hipografo (es decir si  $(a, b) \in H_g$  entonces el segmento que une  $(a, b)$  y  $(0, b)$  esta contenido en  $H_g$ ) es inmediato que  $g(t) \leq u_2^t$  para todo  $t \in [0, 1]$  tal que  $|t - x_1| < \mathbb{H}(f, g)$ . Por lo tanto  $B((x_1, f(x_1)), \mathbb{H}(f, g)) \cap H_g = \emptyset$  de donde  $d((x_1, f(x_1)), H_g) \geq \mathbb{H}(f, g)$ . Como  $(x_1, f(x_1)) \in \partial H_f$  tenemos que  $d((x_1, f(x_1)), H_g) = \mathbb{H}(f, g)$  conduce a un absurdo ya que estamos suponiendo que  $x \in \text{int}(H_f)$  para todo  $x$  tal que  $\mathbb{H}(f, g) = d(x, H_g)$ .  $d((x_1, f(x_1)), H_g) > \mathbb{H}(f, g)$  también conduce a una contradicción ya que  $\mathbb{H}(f, g) = d_H(H_f, H_g)$ . Como  $x \notin \text{int}(H_f)$  tiene que ser  $x \in \partial H_f$   $\square$

**Observación 4.16.** No es cierto en general que si  $A$  y  $B$  son subconjuntos compactos de  $\mathbb{R}^2$  existen  $a \in \partial A$  y  $b \in \partial B$  tal que  $\|a - b\| = d_H(A, B)$  basta tomar  $B = B(0, 1)$  y  $A = B(0, 1) \setminus B(0, 3/4)$ .

Por el lema anterior sabemos que existen  $x \in \partial H_f$  e  $y \in \partial H_g$  tal que  $\|x - y\| = \mathbb{H}(f, g)$ , veamos que los puntos  $x$  e  $y$  donde se realiza la distancia  $\mathbb{H}(f, g)$ , para el caso en que ambas funciones son continuas, son de la forma  $(t, f(t))$  y  $(s, g(s))$ .

**Proposición 4.17.** Sean  $f, g : [0, 1] \rightarrow \mathbb{R}$  no negativas, continuas, sean  $x$  e  $y$  tal que  $\mathbb{H}(f, g) = \|x - y\|$  entonces existen  $t \in [0, 1]$  y  $s \in [0, 1]$  tal que  $x = (t, f(t))$  e  $y = (s, g(s))$ .

*Demostración.* Por el lema anterior sabemos que existen  $x \in \partial H_f$  e  $y \in \partial H_g$  tal que  $\mathbb{H}(f, g) = \|x - y\|$ . Por lo tanto basta demostrar que  $x$  e  $y$  son de la forma  $(t, f(t))$  y  $(s, g(s))$  respectivamente. Para eso observemos que como  $f$  es continua si  $x \in \partial H_f$  hay 4 casos posibles (y lo mismo para  $y \in \partial H_g$ ):

1.  $x = (0, x_2)$  con  $x_2 < f(0)$
2.  $x = (1, x_2)$  con  $x_2 < f(1)$
3.  $x = (x_1, 0)$  con  $0 \leq x_1 \leq 1$
4.  $0 \leq x_1 \leq 1$  y  $x_2 = f(x_1)$

Veamos que  $x$  puede estar únicamente en el caso cuatro. Es claro que  $x$  no pueden estar en el tercer caso ya que ambas funciones son no negativas. Los casos 1 y 2 se excluyen razonando de la misma forma que se hizo en el lema anterior para ver que  $x \notin \text{int}(H_f)$ : supongamos que estamos en el caso 1 (ver Figura 4.2). Observemos primero que  $B((x_1, f(x_1)), \mathbb{H}(f, g)) \cap H_g = \emptyset$  ya que si  $(t_1, t_2) \in B((x_1, f(x_1)), \mathbb{H}(f, g)) \cap H_g$  entonces el segmento que une los puntos  $(t_1, 0)$  y  $(t_1, t_2)$  (el cual está contenido en  $H_g$  por ser  $H_g$  hipografo) intersecta a  $B(x, \mathbb{H}(f, g))$  lo cual contradice que  $\mathbb{H}(f, g) = d_H(H_f, H_g)$ . Tenemos entonces que  $d((x_1, f(x_1)), H_g) \geq \mathbb{H}(f, g)$ . No obstante como ahora estamos suponiendo que  $f$  es continua no puede ser  $d(x, H_g) = d((x_1, f(x_1)), H_g)$ , y por lo tanto  $d((x_1, f(x_1)), H_g) > \mathbb{H}(f, g)$ , lo cual nuevamente conduce a una contradicción. Es decir  $x$  solo puede estar en el cuarto caso. Razonando de forma análoga se llega a que  $y$  solamente puede estar en el cuarto caso, de donde se sigue la tesis.  $\square$

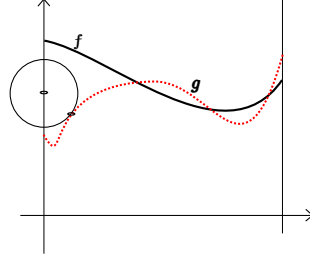


Figura 4.2: Si  $x = (0, x_2)$  entonces  $d(x, H_g) < \mathbb{H}(f, g)$  si  $x_2 < f(0)$  (o  $x_2 < f(1)$  si  $x = (1, x_2)$ ).

**Proposición 4.18.** Sean  $f, g : [0, 1] \rightarrow \mathbb{R}$  no negativas USC, entonces

$$\mathbb{H}(f, g) = \max \left\{ \sup_{\substack{x=(x_1, x_2) \in \partial H_g \\ g(x_1) \geq f(x_1)}} d(x, \partial H_f), \sup_{\substack{y=(y_1, y_2) \in \partial H_f \\ f(y_1) \geq g(y_1)}} d(y, \partial H_g) \right\}.$$

*Demostración.* Denotemos  $d = \max \left\{ \sup_{\substack{x=(x_1, x_2) \in \partial H_g \\ g(x_1) \geq f(x_1)}} d(x, \partial H_f), \sup_{\substack{y=(y_1, y_2) \in \partial H_f \\ f(y_1) \geq g(y_1)}} d(y, \partial H_g) \right\}$ .

Si  $\mathbb{H}(f, g) = 0$  el resultado es inmediato, supongamos  $\mathbb{H}(f, g) > 0$ . Veremos que  $\mathbb{H}(f, g) \leq d$ .

Como  $H_f$  y  $H_g$  son compactos tenemos dos casos

- 1) existe  $x \in H_g$  tal que  $\mathbb{H}(f, g) = d(x, \partial H_f)$ , o
- 2) existe  $y \in H_f$  tal que  $\mathbb{H}(f, g) = d(y, \partial H_g)$ .

Supongamos que estamos en el primero. Por el Lema 4.15 podemos suponer que  $x = (x_1, x_2) \in \partial H_g$ , en este caso, como  $\mathbb{H}(f, g) > 0$  y  $H_f$  es un hipografo  $g(x_1) > f(x_1)$ , por lo tanto  $x \in \{x = (x_1, x_2) \in \partial H_g : g(x_1) \geq f(x_1)\}$  de donde  $\mathbb{H}(f, g) \leq d$ . Si estamos en el caso dos, nuevamente por el Lema 4.15 podemos suponer que  $y = (y_1, y_2) \in \partial H_f$ , como  $\mathbb{H}(f, g) > 0$  y  $H_g$  es un hipografo tiene que ser  $f(y_1) > g(y_1)$ , por lo tanto  $y \in \{y = (y_1, y_2) \in \partial H_f : f(y_1) \geq g(y_1)\}$ , de donde  $\mathbb{H}(f, g) \leq d$ . La otra desigualdad es inmediata a partir de la definición de  $\mathbb{H}$ .  $\square$

**Observación 4.19.** Observemos que si bien en las hipótesis del Lema 4.15 y de la Proposición 4.18 pedimos que las funciones sean no negativas, es solo a los efectos de simplificar la notación, basta que estén acotadas inferiormente.

Si tenemos  $\aleph_n^f = (t_1, f(t_1)), \dots, (t_n, f(t_n))$  y  $\aleph_n^g = (t_1, g(t_1)), \dots, (t_n, g(t_n))$  y asumimos que las funciones  $f$  y  $g$  son continuas, la Proposición 4.18 nos da un algoritmo sencillo, de orden  $n^2$  que permite aproximar  $\mathbb{H}(f, g)$ , simplemente calculamos

$$\tilde{\mathbb{H}}(f, g) = \max \left\{ \max_{\{i: g(t_i) \geq f(t_i)\}} d((t_i, g(t_i)), \aleph_n^f), \max_{\{i: f(t_i) \geq g(t_i)\}} d((t_i, f(t_i)), \aleph_n^g) \right\}.$$

A partir del Lema 4.15, si pedimos que  $\max_i |t_{i+1} - t_i| \rightarrow 0$  es claro que  $\tilde{\mathbb{H}}(f, g) \rightarrow \mathbb{H}(f, g)$ .

### 4.3. Aplicaciones prácticas a problemas de clasificación binaria con espectros (de masas o de resonancia magnética).

El espectrómetro de masas es un instrumento que permite analizar con gran precisión la composición de diferentes compuestos químicos, separando los núcleos atómicos en función de su relación

masa/carga. Una espectrometría es una función que asigna a cada cociente masa/carga el número de moléculas encontradas correspondiente a dicho cociente. Trabajaremos con tres conjuntos de datos, el primero consiste de 216 espectrometrías hechas a un grupo de 216 mujeres, con el fin de determinar la presencia o no de cáncer de ovario. En el segundo tenemos una muestra de 56 espectrometrías de café. Tomaremos una muestra de entrenamiento de 28 datos. En dicha muestra 14 corresponden a espectrometrías de la cepa “Arabica” y 14 a la cepa “Robusta”. En la muestra de testeo tenemos 15 de la cepa “Arabica” y 13 de la “Robusta”. Buscamos clasificar en función de la cepa a la que pertenece. El tercer conjunto de datos consiste de 24 espectros RMN-H de resonancia magnética, hechas en extractos de tejidos de corazón de ratones a los cuales se les ha suministrado a la mitad una dieta rica en contenido graso (clase que denotaremos HFD por sus siglas en inglés) y a los restantes una dieta normal (que constituyen el grupo de control). Dentro de cada grupo 5 son hembras y 7 son machos. Hemos clasificado dicha muestra con el fin de determinar a partir de una espectrometría el género del ratón. En una segunda clasificación buscamos determinar se le ha suministrado una dieta con alto contenido graso o no.

- 1) Los valores de masa/carga de la muestra varían entre 700 y 12.000, y tenemos, para dicho rango, entre 320.000 y 360.000 valores de la función, dependiendo del paciente estudiado. En la primera gráfica de la Figura 4.3 se representa en negro la espectrometría de una mujer sin la enfermedad, mientras que en la segunda, también en negro, la espectrometría de una mujer con la enfermedad. Dado que el método de clasificación requiere el cálculo de la distancia de Hausdorff entre los hipografos y que esto es computacionalmente muy costoso hemos realizado algunas modificaciones de los datos. Dado que el algoritmo que definimos asume que las funciones están definidas en la misma grilla hemos construido un estimador de Nadaraya-Watson basado en la muestra, (usando núcleos gaussianos) y lo evaluamos en una grilla equiespaciada de  $N$  valores. Estos  $N$  valores fueron tomados en el intervalo  $[7.000,9.500]$  ya que es donde están los picos más altos de la función. Hemos asignado el valor 0 a los cocientes masa/carga cuyo valor funcional es menor que 5. El valor 5 se obtuvo luego de probar con diferentes umbrales para la eliminación del ruido presente en las espectrometrías originales. El resultado de dicho procedimiento se muestra en rojo en ambas gráficas de la Figura 4.3. Finalmente, reescalamos cada función dividiendo entre su máximo, para llevarlas a funciones a valores en  $[0,1]$ . Esto se debe a que nos interesa más medir la posición de los picos altos que su la magnitud. Cada uno de los 216 datos los hemos clasificado por el método de vecinos más cercanos, tomando como muestra de entrenamiento toda la muestra excepto el dato a clasificar. En las tablas 4.1 se muestran los resultados de la clasificación para el caso en que  $N = 20001$ ,  $k = 5$  y  $k = 7$ . Comparando con los resultados obtenidos con las distancias  $\mathbb{H}$ ,  $\|\cdot\|_\infty$  y  $L^2$ .

k=5, $\mathbb{H}$			k=5, $\ \cdot\ _\infty$			k=5, $L^2$		
	$\hat{C}$	$\hat{N}$		$\hat{C}$	$\hat{N}$		$\hat{C}$	$\hat{N}$
C	113	8	C	110	11	C	111	10
N	9	86	N	14	81	N	10	85

k=7, $\mathbb{H}$			k=7, $\ \cdot\ _\infty$			k=7, $L^2$		
	$\hat{C}$	$\hat{N}$		$\hat{C}$	$\hat{N}$		$\hat{C}$	$\hat{N}$
C	110	11	C	107	14	C	110	11
N	7	88	N	13	82	N	8	87

Tabla 4.1: Cada tabla muestra en la entrada 1,1 el número de pacientes con cáncer que fueron bien clasificados, en la 1,2 las pacientes con cáncer mal clasificadas. En la entrada 2,1 el número de pacientes sin cáncer que fueron mal clasificados, y finalmente en la 2,2 las pacientes sin cáncer, bien clasificadas.

- 2) Cada espectrometría de café está compuesta por 286 valores de masa/carga que fueron ajustados

tados a una grilla equiespaciada en el intervalo  $[0, 1]$ . Los valores funcionales en dicha grilla varían entre 6 y 23 aproximadamente. Los mismos fueron reescalados dividiendo entre el máximo de cada espectrometría. No fue necesario eliminar ruido ni considerar un conjunto menor que los 286 valores ya que el tiempo computacional de cálculo de la distancia de Hausdorff es bajo. Los resultados de la clasificación se muestran en la Tabla 4.2.

k=5, $\mathbb{H}$		
	$\hat{A}$	$\hat{R}$
A	15	0
R	1	12

k=5, $\ \cdot\ _\infty$		
	$\hat{A}$	$\hat{R}$
A	15	0
R	0	13

k=5, $L^2$		
	$\hat{A}$	$\hat{R}$
A	12	3
R	2	11

k=3, $\mathbb{H}$		
	$\hat{A}$	$\hat{R}$
A	15	0
R	2	11

k=3, $\ \cdot\ _\infty$		
	$\hat{A}$	$\hat{R}$
A	15	0
R	1	12

k=3, $L^2$		
	$\hat{A}$	$\hat{R}$
A	15	0
R	2	11

Tabla 4.2: Cada tabla muestra en la entrada 1,1 el número de espectrometrías de la cepa “Arabica” que fueron bien clasificadas. En la 1,2 las de la misma cepa, mal clasificadas. En 2,2 el número espectrometrías de la cepa “Robusta” bien clasificadas y en la 2,1 las de la misma cepa, mal clasificadas.

- 3) En este caso los espectros están compuestas cada uno por 438 valores, la grilla donde se toman varía entre 0,5 y 4,5, de a incrementos de aproximadamente 0,01. No hemos reescalado los datos en ninguna coordenada. En las tablas 4.3 se muestra el resultado de la clasificación por género, es decir sin distinguir el tipo de dieta que se ha proporcionado. En la 4.4 se muestra el resultado de la clasificación por dieta. Los valores de  $k$  tomados son los que arrojaron en cada caso mejores resultados.

k=3, $\mathbb{H}$		
	$\hat{M}$	$\hat{H}$
M	12	1
H	1	9

k=3, $\ \cdot\ _\infty$		
	$\hat{M}$	$\hat{H}$
M	9	4
H	2	8

k=3, $L^2$		
	$\hat{M}$	$\hat{H}$
M	9	4
H	1	9

Tabla 4.3: Cada tabla muestra en la entrada 1,1 el número de espectros de RMN bien clasificados, hechos a machos, en la 1,2 el número de mal clasificados, hechos a hembras. En la entrada 2,2 se muestra el número de espectros de RMN bien clasificados, hechos a hembras, finalmente en la 2,1 los mal clasificados, hechos a hembras.

k=3, $\mathbb{H}$		
	HFD	GC
HFD	7	5
GC	4	7

k=3, $\ \cdot\ _\infty$		
	HFD	GC
HFD	5	7
GC	9	2

k=3, $L^2$		
	HFD	GC
HFD	5	7
GC	5	6

Tabla 4.4: Cada tabla muestra en la entrada 1,1 el número de espectros de RMN bien clasificados, hechos a ratones de la clase HFD (dieta rica en grasas), en la 1,2 el número de espectros mal clasificados, del grupo de control. En la entrada 2,2 se muestra el número de espectros de RMN bien clasificados, hechas al grupo de control, finalmente en la 2,1 las del grupo HFD mal clasificados.

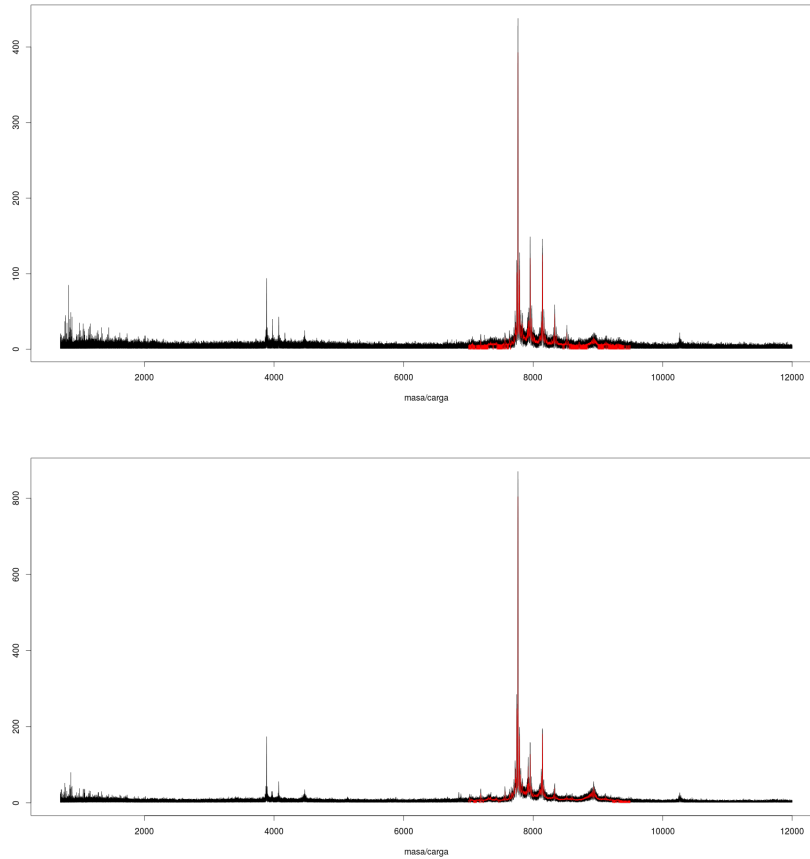


Figura 4.3: El primer gráfico muestra, en negro, la espectrometría de una paciente sin la enfermedad y en rojo la función que aproxima sin reescalar. El segundo gráfico muestra en negro, la de una paciente con cáncer de ovario y en rojo la función aproximante.

#### 4.4. Análisis de la consistencia: condición de Besicovitch

En esta sección enunciaremos un resultado de consistencia en media cuadrática para el estimador de la función de regresión  $\eta$  basado en el método de vecinos más cercanos, cuando la variable  $X$  toma valores en el espacio  $(\mathcal{E}, \mathbb{H})$ . Como dijimos al comienzo del capítulo, si bien es posible probar un resultado general de consistencia para espacios métricos localmente compactos y separables, en el mismo se pide que los pesos satisfagan una hipótesis (ver hipótesis (iii) del Teorema 3.1 en Forzani y

otros (2012)) difícil de verificar para los estimadores basados en el método de vecinos más cercanos y kernel. En estos casos, si el espacio es separable, se puede probar la consistencia, bajo la hipótesis de que  $\eta$  sea acotada y cumpla la condición de Besicovich, que estudiaremos en profundidad.

Dado un espacio métrico  $(\mathcal{E}, m)$  y una medida  $\mu$  de Borel, una función  $f$  medible satisface la condición de Besicovich si:

$$\lim_{\delta \rightarrow 0} \mu \left\{ x \in \mathcal{E} : \frac{1}{\mu(B(x, \delta))} \int_{B(x, \delta)} |f(y) - f(x)| d\mu(y) > \varepsilon \right\} = 0. \quad (4.5)$$

Esta condición, que recuerda en su forma al clásico Teorema de Diferenciación de Lebesgue, significa intuitivamente que el valor de la función  $f$  en un punto  $x$  se puede obtener promediando valores *próximos* a  $x$ . Observemos que si  $f$  es continua se verifica de forma inmediata. En el caso del problema de regresión lo que pediremos es que la misma se verifique cuando  $f = \eta$  y  $\mu$  es la distribución de la variable  $X$ . La siguiente condición, (más débil en general), ha sido objeto de estudio en numerosos trabajos: para todo  $\varepsilon > 0$

$$\lim_{\delta \rightarrow 0} \mu \left\{ x \in \mathcal{E} : \left| \frac{1}{\mu(B(x, \delta))} \int_{B(x, \delta)} f(y) d\mu(y) - f(x) \right| > \varepsilon \right\} = 0. \quad (4.6)$$

Se puede demostrar fácilmente, usando el Teorema 7.9 en Folland (1984), que si el espacio  $(\mathcal{E}, m)$  es separable, localmente compacto y vale la condición (4.6) para toda función acotada entonces vale la condición (4.5) para toda función acotada. Mencionaremos brevemente algunos resultados donde se dan condiciones para la validez de (4.6) para toda función  $f$  integrable:

- 1) Una condición suficiente para que se cumpla (4.6) casi seguramente (es decir cambiando el límite en medida por el límite casi seguro según  $\mu$ ) es que la medida  $\mu$  sea doblante (ver Heinonen (2001)), esto es: toda bola tiene medida finita y existe una constante  $C \geq 1$  tal que  $\mu(B(x, 2r)) \leq C\mu(B(x, r))$  para todo  $x$  y  $r$ . Nuevamente la condición de medida doblante es difícil de verificar en la práctica. Observemos que dicha condición es impuesta sobre la medida y no sobre el espacio. Es una condición muy restrictiva, se puede ver por ejemplo que en espacios de Banach no existen medidas doblantes finitas (ver Ruzhansky (2001)).
- 2) Otra condición que garantiza la validez de (4.6) es que el espacio sea de Vitali. Recordemos que un espacio métrico es de Vitali si, dada una familia  $\mathcal{F}$  de bolas cerradas con la propiedad de que, si para todo  $a$  en el espacio  $\inf\{r > 0 : B(a, r) \in \mathcal{F}\} = 0$  entonces existe un cubrimiento numerable, disjunto, de todo el espacio (a menos de un conjunto de medida nula) con bolas de la familia. Los espacios de Vitali (y generalizaciones) han sido objeto de estudio de numerosos trabajos, entre ellos: Larman (1967), Edgar (2001), Das (2001).
- 3) Un resultado aún más fino que los anteriores, cuya demostración puede encontrarse Preiss (1983) da una condición equivalente a la validez casi segura de la condición (4.6). Veamos antes de dar dicha condición el concepto de dimensión finita de un subconjunto de un espacio métrico:

**Definición 4.20.** Un espacio métrico  $(X, d)$  es finito dimensional en un conjunto  $Y \subset X$  si existe un número natural  $n$  y un número real  $s \in (0, +\infty)$  tal que

$$\bigcap_{x \in \mathfrak{F}} B(x, r(x)) = \emptyset, \quad (4.7)$$

donde  $\mathfrak{F} \subset Y$  es un conjunto finito que verifica:

- 1)  $\#\mathfrak{F} > n$
- 2)  $r(x) \in (0, s)$
- 3)  $x \notin B(y, r(y)) \quad \forall x \neq y \in \mathfrak{F}$



Un espacio métrico  $(X, d)$  es  $\sigma$  finito dimensional si es unión numerable de conjuntos respecto de los cuales  $X$  es de dimensión finita.

El estudio del concepto de finitud de un espacio métrico (y más en general, de dimensión de un espacio topológico) escapa al interés del presente trabajo, no obstante dos referencias clásicas que abordan en profundidad la temática son Hurewicz y Wallman (1948) y Kuratowski (1966). En los espacios separables y completos (espacios Polacos), es equivalente que sean  $\sigma$  finito dimensionales a la validez casi segura de (4.6) (ver Preiss (1983)). Si el espacio  $(\mathcal{E}, \mathbb{H})$  introducido en 4.7 fuese  $\sigma$  finito dimensional, obtendríamos (dado que es un espacio Polaco) que vale la condición (débil) (4.6) y por lo tanto, como es localmente compacto, la condición (4.5). Pero esto no es cierto, como se demuestra en la siguiente proposición:

**Proposición 4.21.** *El espacio  $(\mathcal{E}, \mathbb{H})$  definido en 4.7 no es  $\sigma$  finito dimensional.*

*Demostración.* Supongamos por absurdo que existen conjuntos  $\mathcal{E}_i$ ,  $i \in \mathbb{N}$ , tal que  $\mathcal{E} = \cup_i \mathcal{E}_i$  y  $\mathcal{E}$  es finito dimensional en  $\mathcal{E}_i$  para todo  $i$ . Como  $(\mathcal{E}, \mathbb{H})$  es completo es un espacio de Baire, por lo tanto existe  $i_0$  tal que  $\text{int}(\mathcal{E}_{i_0}) \neq \emptyset$ . Veamos que  $\mathcal{E}$  no puede ser finito dimensional en  $\mathcal{E}_{i_0}$ . Tomemos  $n$  un número natural cualquiera y  $s$  un real positivo, arbitrarios. Sea  $f \in \text{int}(\mathcal{E}_{i_0})$ , existe  $r > 0$  tal que  $B_{\mathbb{H}}(f, r) \subset \mathcal{E}_{i_0}$ , donde  $B_{\mathbb{H}}$  denota la bola con la métrica  $\mathbb{H}$ . Por el lema 4.13 podemos tomar  $g$  continua y positiva, tal que  $\mathbb{H}(f, g) < r/4$ . Es decir  $B_{\mathbb{H}}(g, r/2) \subset B_{\mathbb{H}}(f, r)$ . Tomemos  $n_0$  tal que  $1/n_0 < \min\{r/4, s\}$  y  $n_0 > n$ . Observemos que  $B_{\|\cdot\|_{\infty}}(g, r/2) \subset B_{\mathbb{H}}(g, r/2)$ , donde  $B_{\|\cdot\|_{\infty}}$  denota la bola con la métrica del supremo. Vamos a definir en  $B_{\|\cdot\|_{\infty}}(g, r/2)$  un conjunto  $\mathfrak{F}$  de  $n_0$  funciones que verifiquen 1) 2) y 3) de la Definición 4.20 pero que no se cumpla (4.7). Supondremos sin pérdida de generalidad que  $g(x) - r/2 > 0$  para todo  $x \in [0, 1]$ . Denotemos  $H_g$  el hipografo de  $g$  y su gráfico como  $H_g^+$  es decir:

$$H_g^+ = \{(x, g(x)) : 0 \leq x \leq 1\}.$$

Consideremos los puntos de  $H_g$  que distan  $1/n_0$  de  $H_g^+$ :

$$I = \{(x, y) \in H_g : d((x, y), H_g^+) = 1/n_0\}$$

Es fácil ver que dicho conjunto es el gráfico de una cierta función continua, es decir existe  $h$  continua tal que  $H_h^+ = I$ . Por la forma en que tomamos  $I$  se tiene que  $h \in B_{\mathbb{H}}(g, r/2)$ . Consideremos los  $n_0$  puntos  $x_i = i/n_0$  con  $i = 0, \dots, n_0 - 1$ .  $\mathfrak{F}$  será el conjunto de las  $n_0$  funciones definidas como:

$$f_i(x) = \begin{cases} g(x_i) & x = x_i \\ h(x) & x \in [0, 1] \setminus x_i. \end{cases}$$

Como  $g$  es continua y  $g(x) - r/2 > 0$  tenemos que  $f_i \in \mathcal{E}$  para todo  $i = 0, \dots, n_0 - 1$ . Observemos que  $f_j \notin B_{\mathbb{H}}(f_i, 2/(3n_0))$  para todo  $j \neq i$ . (Ya que  $\mathbb{H}(f_i, f_j) = 1/n_0$  para todo  $i \neq j$ ). Definimos  $r(f_i) = 2/(3n_0)$  que es menor que  $s$  por la forma en que tomamos  $n_0$ .

Veamos que no se cumple 4.7: para

$$t \in \left( \frac{1}{3n_0}, \frac{2}{3n_0} \right)$$

definimos

$$I_t = \{(x, y) \in H_g : d((x, y), H_g^+) = t\},$$

esto define una función  $g_t$  que cumple que  $g_t \in B_{\mathbb{H}}(f_i, r(f_i))$  para todo  $i$ . Con lo cual

$$\forall t \in \left( \frac{1}{3n_0}, \frac{2}{3n_0} \right) \quad g_t \in \bigcap_{f \in \mathfrak{F}} B_{\mathbb{H}}(f, r(f))$$

□

### Resultados de consistencia

Enunciaremos ahora el resultado de consistencia para el espacio métrico  $(\mathcal{E}, \mathbb{H})$  dotado de la sigma álgebra de Borel  $\mathcal{B}$  asociada a  $\mathbb{H}$ , siguiendo la notación e ideas expuestas en Forzani y otros (2012). En este contexto tenemos datos  $(X_1, Y_1), \dots, (X_n, Y_n)$  independientes, idénticamente distribuidos, con la misma distribución de un vector  $(X, Y)$  a valores en  $\mathcal{E} \times \mathbb{R}$  que satisfacen el modelo

$$Y = \eta(X) + \varepsilon,$$

donde el error  $\varepsilon$  satisface que  $\mathbb{E}(\varepsilon|X) = 0$  y  $\mathbb{E}(\varepsilon^2|X) = \sigma^2 < \infty$ . Al igual que como se hace en dimensión finita (ver Stone (1977)) la función de regresión  $\eta(X) = \mathbb{E}(Y|X)$  se aproxima por

$$\eta_n(X) = \sum_{i=1}^n W_{ni}(X)Y_i,$$

donde los pesos  $W_{ni}(X) = W_{ni}(X, X_1, \dots, X_n)$  son no negativos y  $\sum_{i=1}^n W_{ni}(X) = 1$ . En el caso particular en que

$$W_{ni}(X) = \frac{1}{k} \mathbb{1}_{\{X_i \in k(X)\}}, \quad (4.8)$$

donde  $X_i \in k(X)$  indica que  $X_i$  es uno de los  $k$  datos más próximos (respecto de  $\mathbb{H}$ ) de  $X$ , se obtiene el estimador por vecinos más cercanos. El siguiente teorema (ver Forzani y otros (2012)) garantiza la convergencia en media cuadrática de  $\eta_n(X)$  a  $\eta(X)$  cuando los pesos  $W_{ni}$  son de la forma (4.8):

**Teorema 4.22.** *Denotemos  $\mu_X$  la distribución de probabilidad de  $X$ , supongamos que  $\mu_X$  es de Borel. Supongamos que  $\eta$  es acotada y satisface la condición de Besicovich (4.5) con  $\mu = \mu_X$ . Sean  $\{W_{ni}(X)\}_{i=1}^n$  la sucesión de pesos definidos en (4.8), tomemos  $k_n \rightarrow \infty$  una sucesión de números reales positivos tal que  $k_n/n \rightarrow 0$ , entonces  $\eta_n$  es consistente en media cuadrática, i.e.:*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( (\eta(X) - \eta_n(X))^2 \right) = 0.$$

En el caso del problema de clasificación binaria, es decir cuando la variable  $Y$  puede tomar únicamente los valores 0 o 1 tenemos que  $\eta(x) = P(Y = 1|X = x)$ . Si denotamos

$$L^* = \inf_{g: \mathcal{E} \rightarrow \{0,1\}} P(g(X) \neq Y),$$

se puede demostrar (ver Devroye y otros (1996)) que:

$$L^* = P(g^*(X) \neq Y),$$

donde

$$g^*(x) = \begin{cases} 1 & \text{si } \eta(x) > 1/2 \\ 0 & \text{en otro caso.} \end{cases}$$

Dado  $\eta_n$ , el estimador *plug-in* de  $g^*$  es:

$$g_n(x) = \begin{cases} 1 & \text{si } \eta_n(x) > 1/2 \\ 0 & \text{en otro caso} \end{cases}$$

Por otro lado, estimar  $\eta(x)$  en media cuadrática nos garantiza la consistencia de la regla de decisión anterior ya que (ver Devroye y otros (1996)):

$$P(g_n(X) \neq Y) - L^* \leq 2 \left( \mathbb{E} |\eta(X) - \eta_n(X)|^2 \right)^{1/2}.$$

Es decir tenemos la consistencia para el espacio  $(\mathcal{E}, \mathbb{H})$ , de la regla de clasificación basada en el método de vecinos más cercanos si se verifica la condición 4.5 (observemos que aquí  $\eta$  es claramente acotada). Dicho resultado se puede obtener también como consecuencia del Teorema 2 en Cérou y Guyader (2006).

## 4.5. Aplicaciones a problemas de regresión no paramétrica (finito-dimensional) con restricciones geométricas de forma.

Vamos a introducir a continuación una definición que es el análogo para funciones del concepto de cono convexidad introducido en el capítulo tres para subconjuntos compactos de  $\mathbb{R}^d$ . Nuevamente vamos a considerar funciones no negativas. Los resultados que presentaremos se pueden generalizar para el caso de funciones uniformemente acotadas inferiormente.

**Definición 4.23.** Sea  $f : [0, 1] \rightarrow \mathbb{R}$  USC, no negativa, decimos que  $f$  es  $\rho$  cono convexa en sentido estricto si  $H_f \in \mathcal{C}_\rho$ .

**Definición 4.24.** Sea  $f : [0, 1] \rightarrow \mathbb{R}$  USC, no negativa, decimos que  $f$  es  $\rho$  cono convexa en sentido amplio si existe  $h > 0$  tal que  $H_f \in \mathcal{C}_{\rho,h}$ .

Así definidas, es inmediato verificar que las propiedades anteriores son invariantes por traslaciones de  $f$ , es decir si  $f$  es  $\rho$  cono convexa también lo será  $f + k$  para toda constante  $k > 0$ .

**Proposición 4.25.** Sean  $\rho \in (0, \pi)$ ,  $h > 0$  si

$$\mathcal{E}_{\rho,h} = \{f : [0, 1] \rightarrow \mathbb{R}, 0 \leq f, \text{ USC}, \rho \text{ cono convexa en sentido amplio}\},$$

denota el espacio de las funciones  $\rho$  cono convexas en sentido amplio, USC, no negativas, entonces el espacio métrico  $(\mathcal{E}_{\rho,h}, \mathbb{H})$  es localmente compacto, donde  $\mathbb{H}$  es la métrica 4.7.

**Proposición 4.26.** Sean  $\rho \in (0, \pi)$ , si

$$\mathcal{E}_\rho = \{f : [0, 1] \rightarrow \mathbb{R}, 0 \leq f, \text{ USC}, \rho \text{ cono convexa en sentido estricto}\},$$

denota el espacio de las funciones  $\rho$  cono convexas en sentido estricto, USC, no negativas, entonces el espacio métrico  $(\mathcal{E}_\rho, \mathbb{H})$  es localmente compacto, donde  $\mathbb{H}$  es la distancia definida en 4.7.

**Observación 4.27.** Si  $f : S \subset \mathbb{R}^d \rightarrow \mathbb{R}$  es Lipschitz con constante  $K$  entonces  $f$  es  $\rho$  cono convexa en sentido estricto con  $\rho \geq \pi/2 - \arctg(K)$ . El recíproco no es cierto, basta tomar  $f(x) = \sqrt{x}$ .

**Observación 4.28.** Observemos que la convergencia en la métrica  $\mathbb{H}$  de funciones  $\rho$  cono convexas, no implica la convergencia puntual y tampoco es cierto el recíproco. Veamos que esto no pasa con la convergencia en  $L^1$ :

**Proposición 4.29.** Sea  $f_n \in \mathcal{E}_\rho$  una sucesión tal que  $\mathbb{H}(f_n, f) \rightarrow 0$  para alguna  $f \in \mathcal{E}$ , entonces:  $\int_0^1 |f_n(x) - f(x)| dx \rightarrow 0$

*Demostración.* Basta probar que  $d_\mu(H_{f_n}, H_f) \rightarrow 0$  siendo  $\mu$  la medida de Lebesgue en  $\mathbb{R}^2$ . Observemos primero que por el Teorema 3.11  $H_f$  es un subconjunto  $\rho$  cono convexo de  $\mathbb{R}^2$ , de donde, por la Proposición 3.13  $\mu(\partial H_f) = 0$ . Por el Teorema 3.10  $d_H(\partial H_{f_n}, \partial H_f) \rightarrow 0$ . Por lo tanto que  $d_\mu(H_{f_n}, H_f) \rightarrow 0$  es consecuencia inmediata del Teorema 2 de Cuevas y otros (2012).  $\square$

Veamos ahora algunas proposiciones que relacionan las funciones USC con las funciones  $\rho$  cono convexas. Las primeras dos nos dan una idea de cómo son los puntos en los cuales no es posible poner un cono de ángulo  $\rho$  en el complemento del hipografo.

**Proposición 4.30.** Sea  $f : [0, 1] \rightarrow \mathbb{R}$  no negativa, USC, sea  $\rho \in (0, \pi)$ , si

$$T_\rho = \left\{x \in [0, 1] : \forall \xi \in S_1^1 \quad C_{\rho,\xi}((x, f(x)) \cap H_f \neq \emptyset)\right\}, \quad (4.9)$$

denota el conjunto de los puntos que no verifican la condición de  $\rho$  cono convexidad en sentido amplio. Entonces  $T_\rho$  es abierto.

*Demostración.* Supongamos que  $\rho \in (0, \pi)$  es tal que  $T_\rho \neq \emptyset$ . Veamos que  $\partial T_\rho \cap T_\rho = \emptyset$ . Sea  $x \in \partial T_\rho$ , existe  $x_n \rightarrow x$  tal que  $x_n \in T_\rho^c$ . Como  $x_n \in T_\rho^c$  existe  $\xi_n \in S_1^1$  tal que  $C_{\rho, \xi_n}((x_n, f(x_n))) \subset H_f^c$  para todo  $n$ . Podemos suponer, pasando eventualmente a subsucesiones, que  $\xi_n \rightarrow \xi$  con  $\xi$  de norma 1. Consideremos el cono  $C_{\rho, \xi}((x, \limsup f(x_n)))$ . Dado que  $C_{\rho, \xi}((x, \limsup f(x_n)))$  es abierto se sigue que  $C_{\rho, \xi}((x, \limsup f(x_n))) \cap H_f = \emptyset$ . Observemos que, como  $f$  es USC,  $\limsup f(x_n) \leq f(x)$ , si  $\limsup f(x_n) = f(x)$  entonces  $x \in T_\rho$  lo cual es absurdo. En caso contrario, si  $\limsup f(x_n) < f(x)$  se tiene que, si  $(z, t) \in H_f$  entonces el segmento  $[(z, 0), (z, t)] \in H_f$  por lo tanto,  $C_{\rho, \xi}((x, f(x))) \cap H_f = \emptyset$ , de donde  $x \in T_\rho^c$ .  $\square$

**Proposición 4.31.** *Sea  $f : [0, 1] \rightarrow \mathbb{R}$ , no negativa y USC. Entonces,*

$$\bigcap_{0 < \rho < \pi} T_\rho,$$

*es a lo sumo una cantidad numerable de puntos, donde  $T_\rho$  es el conjunto definido en (4.9).*

*Demostración.* Por la proposición anterior sabemos que los conjuntos  $T_\rho$  son abiertos, es decir son una unión numerable de intervalos abiertos. Observemos que si  $\rho_1 < \rho_2$  entonces  $T_{\rho_1} \subset T_{\rho_2}$ , por lo tanto tenemos dos casos: o bien  $\bigcap_{0 < \rho < \pi} T_\rho$  es una cantidad a lo sumo numerable de puntos, o bien existe  $(a, b) \subset \bigcap_{0 < \rho < \pi} T_\rho$ . Veamos que esto último no es posible. Supongamos por reducción al absurdo que existe  $(a, b) \subset \bigcap_{0 < \rho < \pi} T_\rho$ , tomemos  $a_1$  y  $b_1$  tal que  $a < a_1 < b_1 < b$  y consideremos el intervalo cerrado  $[a_1, b_1] \subset (a, b)$ . Observemos que como  $f$  es USC existen  $z \in [a_1, b_1]$  tal que  $f(z) \geq f(x)$  para todo  $x \in [a_1, b_1]$ , y  $t \in [0, 1]$  tal que  $f(t) \geq f(x)$  para todo  $x \in [0, 1]$ . Veamos que  $z \in (\bigcap_{0 < \rho < \pi} T_\rho)^c$ , para eso vamos a demostrar que existe  $\xi_z$  de norma uno y  $\rho_z \in (0, \pi)$ , tal que  $C_{\rho_z, \xi_z}(z, f(z)) \subset H_f^c$ . Observemos primero que si  $f(z) = f(t)$ , basta tomar  $\rho_z = \pi$  y  $\xi_z = (0, 1)$ . Supongamos entonces que  $f(z) < f(t)$ . En este caso, por la forma en que tomamos  $z$  y  $t$  es inmediato verificar que el cono que determinan los puntos  $(a_1, f(t))$ ,  $(z, f(z))$  y  $(b_1, f(t))$  cumple las condiciones requeridas de donde se sigue que  $z \in (\bigcap_{0 < \rho < \pi} T_\rho)^c$ , pero esto es absurdo ya que  $z \in [a_1, b_1] \subset (a, b) \subset \bigcap_{0 < \rho < \pi} T_\rho$ .  $\square$

**Teorema 4.32.** *Siguiendo la notación introducida anteriormente,*

$$\overline{\bigcup_{0 < \rho < \pi} \mathcal{E}_\rho} = \mathcal{E},$$

*donde la clausura es tomada respecto de la métrica  $\mathbb{H}$  definida en 4.7 y  $\mathcal{E}$  es el espacio definido en 4.7.*

*Demostración.* Observemos primero que, por la Proposición 4.10, como  $\mathcal{E}$  es cerrado

$$\overline{\bigcup_{0 < \rho < \pi} \mathcal{E}_\rho} \subset \mathcal{E}.$$

Para ver la otra inclusión sea  $f \in \mathcal{E}$ , vamos a demostrar que para todo  $\varepsilon > 0$  existe un ángulo  $\rho_\varepsilon$ ,  $0 < \rho_\varepsilon \leq \pi$ , y una función  $f_\varepsilon \in \mathcal{E}_{\rho_\varepsilon}$  tal que  $\mathbb{H}(f_\varepsilon, f) < \varepsilon$ . La construcción de  $f_\varepsilon$  es idéntica a la hecha para definir  $f_P$  en 4.4, asociada a una partición  $P$  de la forma  $0 = x_0 < x_1 < \dots < x_{n-1} = 1$ . Veamos que  $f_\varepsilon$  es  $\rho$  cono convexa para algún  $\rho$ . Observemos que como  $f$  es USC existe, nuevamente aplicando la Proposición 4.6 1),  $z \in [0, 1]$  tal que  $\sup_{x \in [0, 1]} f(x) = f(z)$ . Sea  $m = \min_{i=1, \dots, n-1} |x_{i+1} - x_i|$ , por la forma en que fue definida la partición tenemos que  $m > 0$ . Tomemos  $\rho < \arctan(\frac{m}{2f(z)})$ . Sea  $t = (t_1, t_2) \in \partial H_{f_\varepsilon}$ , sea  $i$  tal que  $x_i \leq t_1 \leq x_{i+1}$ . Supongamos primero que  $t_2 = f_\varepsilon(t_1)$ . Es claro que el ángulo que forman los puntos  $(x_i, f(z))$ ,  $(t_1, t_2)$  y  $(x_{i+1}, f(z))$  es mayor o igual que  $\rho$  y el cono que determinan no interseca a  $H_{f_\varepsilon}$ . Para los restantes puntos de  $\partial H_{f_\varepsilon}$  el cono se construye con el mismo  $\rho$ , razonando de manera análoga.  $\square$

#### 4.5.1. Algoritmo de construcción de funciones $\rho$ como convexas a partir de una muestra.

Supongamos que tenemos  $f(t_1), \dots, f(t_n)$  valores de una función  $f : [0, 1] \rightarrow \mathbb{R}$  no negativa, y USC, tomados en una grilla  $t_1 < \dots < t_n$ , vamos a construir un algoritmo para obtener una función  $f_\rho^n : [0, 1] \rightarrow \mathbb{R}$  no negativa, continua, y  $\rho$  como convexa en sentido estricto. Lo que haremos en esta sección es explicar dicho algoritmo de manera intuitiva, y demostrar que en el caso en que la función es  $\rho$  como convexa, coincide con la poligonal que une los puntos  $(t_1, f(t_1)) \dots (t_n, f(t_n))$ . El pseudocódigo del mismo puede encontrarse en el Apéndice A.

##### Idea intuitiva del algoritmo

El algoritmo será definido de manera inductiva. Partiendo de una función inicial  $f^0$  que será constante e igual al máximo de los valores  $f(t_i)$ , iremos obteniendo aproximaciones de  $f$  más finas, ver 4.4. En el paso  $k$  la función que obtenemos (que denotaremos  $f^k$ ) será definida como la poligonal  $P(t_1, t_2^k, \dots, t_{m(k)}^k, t_n)$  que une ciertos puntos  $(t_1, f(t_1)), (t_2^k, f(t_2^k)), \dots, (t_{m(k)}^k, f(t_{m(k)}^k)), (t_n, f(t_n))$  donde  $t_1 \leq t_2^k \dots < t_{m(k)}^k \leq t_n$  es un subconjunto de la grilla original. Veamos como se construye dicha poligonal.

**INPUT:**  $0 < \rho < \pi$  y  $f(t_1), f(t_2) \dots, f(t_n)$  siendo  $0 \leq t_1 < t_2 < \dots < t_n \leq 1$ .

**OUTPUT:** Una poligonal  $P(t_1, t_2^k, \dots, t_{m(k)}^k, t_n) \in \mathcal{E}_\rho$  que une los puntos

$$(t_1, f(t_1)), (t_2^k, f(t_2^k)), \dots, (t_{m(k)}^k, f(t_{m(k)}^k)), (t_n, f(t_n)),$$

siendo  $t_1 \leq t_2^k < \dots < t_{m(k)}^k \leq t_n$  un subconjunto de la partición original. El valor de  $k$  indica el número de la iteración del algoritmo donde se detuvo y  $1 \leq m(k) \leq n$

**IDEA GENERAL:** Partimos de  $f^0$  la función constante e igual al máximo de los valores  $f(t_i)$ . Si dicho máximo se da en cierto  $t_{i_0}$  tenemos una grilla inicial  $t_1 \leq t_{i_0} \leq t_n$  que iremos refinando, agregando puntos  $t_j$  de la partición original, siempre y cuando al agregarlo se cumpla la condición de cono convexidad. Cuando no es posible refinar más la grilla el algoritmo se detiene dando como resultado una función  $\rho$  como convexa, construida como la poligonal que une los valores de  $f$  en los puntos de la grilla que se obtuvieron.

**PASO INICIAL:** Tomamos la función  $f^0$  constante e igual al máximo de los valores  $f(t_i)$ . Supongamos que dicho máximo se da en un punto  $t_i$  tal que  $t_1 < t_i < t_n$ , definimos  $t_2^1 = t_i$ . Tenemos entonces la grilla  $t_1, t_2^1, t_n$  e intervalos  $I_1^1 = (t_1, t_2^1)$  y  $I_2^1 = (t_2^1, t_n)$ .

**ITERACIÓN:** En el paso  $k \geq 1$  tenemos  $m(k)$  intervalos no vacíos  $I_1^k, \dots, I_{m(k)}^k$  y  $f^k$ . Definiremos  $f^{k+1}$  primero en  $I_1^k = (t_1, t_2^k)$  luego en  $I_2^k$ , etc, hasta  $I_{m(k)}^k = (t_{m(k)}^k, t_n)$ . Denotaremos  $t_1^k = t_1$  y  $t_{m(k)+1}^k = t_n$ .

Mientras que  $1 \leq r \leq m(k)$

En  $I_r^k = (t_r^k, t_{r+1}^k)$ , sea  $l_1 = \arg \max_{t_r^k < t_j < t_{r+1}^k} f(t_j)$  (si hay más de uno tomamos el menor).

Consideraremos dos conos, con ángulo  $\rho$  y vértice en  $(l_1, f(l_1))$  contenidos en el cono determinado por los puntos  $(t_r^k, f(t_r^k)), (l_1, f(l_1))$  y  $(t_{r+1}^k, f(t_{r+1}^k))$ , como se muestra en la Figura 4.5. Para construir uno de los conos denotemos  $u = (t_r^k, f(t_r^k)) - (l_1, f(l_1))$ , tomamos  $t_s \leq t_r^k$  tal que el ángulo que forma el vector  $(t_s, f(t_s)) - (l_1, f(l_1))$  con  $u$  es el

mayor de los ángulos que forma  $u$  con vectores de la forma  $(t_i, f(t_i)) - (l_1, f(l_1))$  siempre que  $t_i \leq t_r^k$  y  $f(t_i) \geq f(t_r^k)$ . Es decir, de los puntos menores que  $t_r^k$  cuyas observaciones superan  $f(t_r^k)$ , el de mayor pendiente respecto de  $u$  (puede ser  $t_s = t_r^k$ ). Consideremos el cono con vértice en  $(l_1, f(l_1))$ , ángulo  $\rho$  y eje el vector que se obtiene como rotación de ángulo  $\rho/2$ , en sentido horario, del vector

$$\frac{(t_s, f(t_s)) - (l_1, f(l_1))}{\|(t_s, f(t_s)) - (l_1, f(l_1))\|}.$$

- Si dicho cono no corta a la muestra, definimos  $t_r^{k+1} = t_r^k$ ,  $t_{r+1}^{k+1} = l_1$ ,  $t_{r+2}^{k+1} = t_{r+1}^k$  y  $f^k$  en  $I_r^k$  como la poligonal que une los puntos, y vamos a  $I_{r+1}^k$ .
- En caso contrario, consideramos el cono, construido de forma análoga al anterior, de ángulo  $\rho$  y vértice  $(l_1, f(l_1))$ , cuyo eje se obtiene rotando en sentido antihorario del vector

$$\frac{(t_w, f(t_w)) - (l_1, f(l_1))}{\|(t_w, f(t_w)) - (l_1, f(l_1))\|}.$$

Donde  $t_w$  es un punto mayor o igual que  $t_{r+1}^k$  tal que  $f(t_w) \geq f(t_{r+1}^k)$  y de mayor pendiente respecto del vector  $(t_{r+1}^k, f(t_{r+1}^k)) - (l_1, f(l_1))$ .

- ◊ Si dicho cono no corta a la muestra definimos  $t_r^{k+1} = t_r^k$ ,  $t_{r+1}^{k+1} = l_1$ ,  $t_{r+2}^{k+1} = t_{r+1}^k$  y  $f^k$  en  $I_r^k$  como la poligonal que une los puntos, y vamos a  $I_{r+1}^k$ .
- ◊ Si ambos conos contienen puntos de la muestra, buscamos  $l_2$  el máximo de  $f$ , menor o igual que  $f(l_1)$ , en  $I_r^k$ . Para  $l_2$  construimos los conos de forma análoga a lo hecho para  $l_1$ . Si ambos conos cortan a la muestra, seguimos con el mismo procedimiento, con  $l_3$  etc. La búsqueda de puntos intermedios acaba o bien porque encontramos un máximo que cumple la condición de que al menos uno de los conos no corta a la muestra (en cuyo caso vamos a  $I_{r+1}^k$ ), o bien ningún punto cumple dicha condición y  $f^k$  se define en  $I_r^k$  como la poligonal que une los extremos del intervalo y vamos a  $I_{r+1}^k$ .

**CRITERIO DE PARADA:** Si luego de aplicar la iteración en el intervalo  $I_{m(k)}^k$  la función  $f^{k+1}$  que obtenemos es igual a la que teníamos antes, es decir no agregamos puntos nuevos a la grilla, el algoritmo se detiene, dando como resultado la función  $f^k$ .

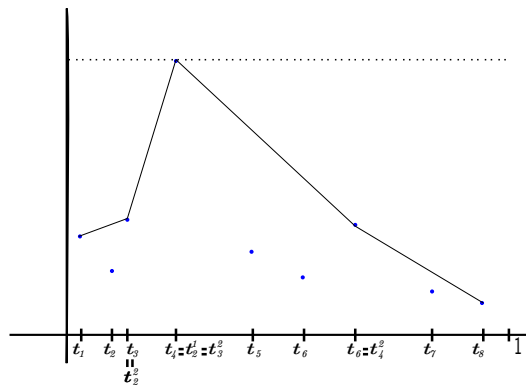


Figura 4.4: Gráfico de las funciones  $f^0$  en punteado, y  $f^1$  en línea sólida.

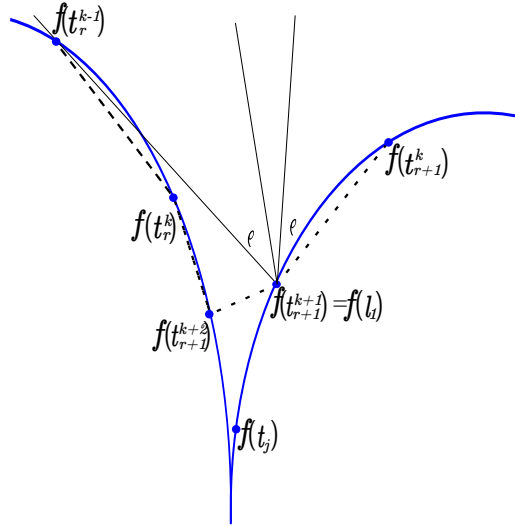


Figura 4.5: Gráfico de una función y el procedimiento de selección de puntos. El punto  $t_{r+1}^{k+1}$  se agrega tras observar que en los conos de la figura no hay más puntos de la muestra. En línea punteada la poligonal que construye el algoritmo.

**Proposición 4.33.** Sea  $f : [0, 1] \rightarrow \mathbb{R}$  no negativa, USC, y  $\rho$  como convexa en sentido estricto, sean  $\aleph_n = (t_1, f(t_1)), \dots, (t_n, f(t_n))$  entonces  $P = P(t_1, \dots, t_n)$  es  $\rho$  como convexa en sentido estricto.

*Demostración.* Consideremos los puntos  $(t_1, f(t_1))$  y  $(t_2, f(t_2))$  y la poligonal  $P(t_1, t_2)$  como se muestra en la Figura 4.6, vamos a suponer sin pérdida de generalidad que  $t_1 < t_2$  y  $f(t_2) \leq f(t_1)$ . Como  $(t_1, f(t_1)) \in \partial H_f$  existe  $\xi_1$  de normal 1 y  $C_{\rho, \xi_1}(u_{t_1})$  un cono de ángulo  $\rho$  y vértice en el punto  $u_{t_1}$  tal que  $C_{\rho, \xi_1}(u_{t_1}) \subset H_f^c$ . Observemos que como  $H_f$  es un hipografo, podemos suponer que  $(t_1, 2f(t_1)) \in C_{\rho, \xi_1}(u_{t_1})$  y  $C_{\rho, \xi_1}(u_{t_1}) \subset H_P^c$ . Análogamente, existe  $\xi_2$  de norma uno y  $C_{\rho, \xi_2}(u_{t_2})$  tal que  $(t_2, 2f(t_2)) \in C_{\rho, \xi_2}(u_{t_2})$  y  $C_{\rho, \xi_2}(u_{t_2}) \subset H_P^c$ . El borde de  $C_{\rho, \xi_1}(u_{t_1})$  determina dos rectas que

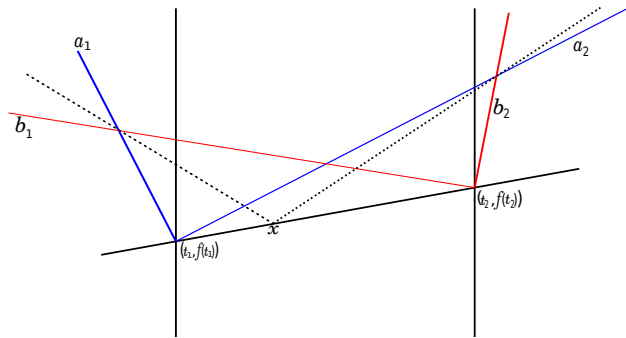


Figura 4.6:  $P(t_1, t_2)$

llamaremos  $a_1$  y  $a_2$  mientras que el de  $C_{\rho, \xi_2}(u_{t_2})$  determina dos rectas que llamaremos  $b_1$  y  $b_2$ , ver Figura 4.6. Si el ángulo que forma  $a_1$  con la recta  $x = t_1$  es mayor o igual que el ángulo que forma  $b_1$  con la recta  $x = t_2$  (observar que como  $\rho \leq \pi$  y estamos suponiendo que  $(t_1, 2f(t_1)) \in C_{\rho, \xi_1}(u_{t_1})$  y  $(t_2, 2f(t_2)) \in C_{\rho, \xi_2}(u_{t_2})$  dichos ángulos son menores o iguales que  $\pi/2$ ), para todo  $x \in P(t_1, t_2)$  el cono  $C_{\rho, \xi_1}(x)$  paralelo a  $C_{\rho, \xi_1}(u_{t_1})$ , con vértice en  $x$  y ángulo  $\rho$  verifica que  $C_{\rho, \xi_1}(x) \subset H_P^c$ . Si

denotamos  $y_1 = a_1 \cap b_1$ ,  $y_2 = a_2 \cap b_2$ , es claro que, para todo  $x \in P(t_1, t_2)$  se tiene que

$$\left\langle \frac{y_1 - x}{\|y_1 - x\|}, \frac{y_2 - x}{\|y_2 - x\|} \right\rangle \leq \cos(\rho),$$

es fácil ver que el cono determinado por los puntos  $y_1, x, y_2$  no contiene puntos de  $H_P$ .  $\square$

**Proposición 4.34.** *Sea  $f : [0, 1] \rightarrow \mathbb{R}$  USC, y  $\rho$  como convexa en sentido estricto. Sean  $(t_1, f(t_1)), \dots, (t_n, f(t_n))$  observaciones de  $f$ . Entonces la poligonal que une los puntos es igual a la función  $f_\rho^n$  que se obtiene luego de aplicar el algoritmo.*

*Demostración.* Dado  $t_1 < t_i < t_n$  supongamos que  $t_i$  no fue seleccionado en el algoritmo, es decir,  $t_r^k < t_i < t_{r+1}^k$  para algún  $r$ . Como  $(t_i, f(t_i)) \in \partial H_f$  existe  $\xi$  de norma 1 y  $C_{\rho, \xi}((t_i, f(t_i))) \subset H_f^c$ . Es inmediato verificar que en este caso, ambos conos que construye el algoritmo a partir de  $t_r^k, t_i$  y  $t_{r+1}^k$  no cortan a la muestra, y por lo tanto el punto se agrega.  $\square$



# Apéndice A

## Apéndice

### A.1. Cálculo del volumen de algunos conjuntos

#### Proposición A.1.

a) La función de volumen del subconjunto  $S$  de  $\mathbb{R}^2$  definido como  $S = \overline{B(0, T)} \setminus C_\rho((0, 0))$  es

$$(T+r)^2\pi - (T+r)^2\frac{\rho}{2} + \frac{r^2}{2}\pi + 2\left(Tr - \frac{r^2}{2\tan(\rho/2)}\right) = \\ \left(\frac{3\pi}{2} - \frac{\rho}{2} - \frac{1}{\tan(\rho/2)}\right)r^2 + (2T\pi - r\rho + 2T)r + \pi T^2 - T^2\frac{\rho}{2} \quad \text{si } 0 \leq r \leq T \tan(\rho/2).$$

b) La función de volumen del subconjunto  $S$  de  $\mathbb{R}^3$  definido como  $S = B((0, 0, 1), 1) \cup B((0, 0, -1), 1)$  es

$$\frac{4}{3}\pi r^3 + 6\pi r^2 + 8\pi r + \frac{8}{3}\pi \quad r \in [0, +\infty).$$

c) Consideremos  $S = [(0, 0, -1/2), (0, 0, -1)] \cup (0, 0, 1)$  es decir el segmento que une los puntos  $(0, 0, -1/2)$  y  $(0, 0, -1)$  unión el punto  $(0, 0, 1)$  observemos que  $B(S, 1)$  es el conjunto (d) de la Figura 2.1. La función de volumen de  $B(S, 1)$  es

$$\frac{4}{3}\pi(1+r)^3 + \frac{3}{2}\pi(1+r)^2 \quad r \in [0, +\infty).$$

*Demostración.* Haremos las cuentas de b) y c). Veamos b)

$$B(S, r) = B((0, 0, 1), 1+r) \cup B((0, 0, -1), 1+r)$$

El volumen de  $B(S, r)$  se puede calcular como  $2\mu_3(B((0, 0, 1), 1+r)) - \mu_3(A)$  donde  $\mu_3$  denota la medida de Lebesgue en  $\mathbb{R}^3$  y  $A$  es el conjunto  $B((0, 0, 1), 1+r) \cap B((0, 0, -1), 1+r)$ . Vamos a calcular  $\mu_3(S)$ . Tenemos que calcular una integral triple, la cual haremos mediante un cambio de variable a coordenadas cilíndricas. Tenemos que  $t = \sqrt{(1+r)^2 - 1}$ . Por lo tanto si  $x = \rho \cos(\phi)$ ,  $y = \rho \sin(\phi)$  y  $z = z$  tenemos que

$$\mu_3(A) = 8 \int_0^{\pi/2} d\phi \int_0^t d\rho \int_0^{\sqrt{(1+r)^2 - \rho^2} - 1} \rho dz = 4\pi \int_0^t (\sqrt{(1+r)^2 - \rho^2} - 1) \rho d\rho$$

haciendo  $u = (1+r)^2 - \rho^2$  tenemos que  $-du/2 = \rho d\rho$  por lo tanto

$$\begin{aligned} \mu_3(A) &= -2\pi \left( \frac{2}{3} ((1+r)^2 - \rho^2)^{3/2} + \rho^2 - (1+r)^2 \right) \Big|_0^t = \\ &= -2\pi \left( \frac{-1}{3} - \frac{2}{3}(1+r)^3 + (1+r)^2 \right) = \frac{4}{3}\pi(1+r)^3 - 2\pi(1+r)^2 + \frac{2\pi}{3} \end{aligned}$$

Por otro lado  $2\mu_3(B((0,0,1), 1+r)) = 2\frac{4}{3}\pi(1+r)^3$  es decir

$$\begin{aligned} \mu_3(B(S, r)) &= \frac{8}{3}\pi(1+r)^3 - \left( \frac{4}{3}\pi(1+r)^3 - 2\pi(1+r)^2 + \frac{2\pi}{3} \right) = \\ &= \frac{4}{3}\pi(1+r)^3 + 2\pi(1+r)^2 - \frac{2\pi}{3} = \frac{4}{3}\pi r^3 + 6\pi r^2 + 8\pi r + \frac{8}{3}\pi \end{aligned}$$

Veamos la demostración de c). Vamos a calcular  $\mu_3(B(S, 1+r))$ , para eso observemos que

$$\mu_3(B(S, 1+r)) = 2\mu_3(B((0,0,1), 1+r)) - \mu_3(C) + (1/2)\pi(1+r)^2$$

donde  $(1/2)\pi(1+r)^2$  es el volumen de la *parte cilíndrica* de  $B(S, 1+r)$  y

$$C = B((0,0,1), 1+r) \cap B((0,0,-1/2), 1+r).$$

Vamos a calcular  $\mu_3(C)$ . Para eso haremos un cálculo análogo al hecho para calcular  $\mu_3(A)$  en b). Pasando a coordenadas cilíndricas, (observemos que en este caso  $t = \sqrt{(1+r)^2 - (3/4)^2}$ ) tenemos que si  $x = \rho \cos(\phi)$   $y = \rho \sin(\phi)$  y  $z = z$  se obtiene

$$\mu_3(C) = 8 \int_0^{\pi/2} d\phi \int_0^t d\rho \int_0^{\sqrt{(1+r)^2 - \rho^2} - 1/2} \rho dz = 4\pi \int_0^t (\sqrt{(1+r)^2 - \rho^2} - 1/2) \rho d\rho$$

haciendo el mismo cambio de variable que en b), es decir  $u = (1+r)^2 - \rho^2$  obtenemos que

$$\mu_3(C) = -2\pi \left( \frac{2}{3} ((1+r)^2 - \rho^2)^{3/2} + \frac{1}{2}(\rho^2 - (1+r)^2) \right) \Big|_0^t = \frac{4}{3}\pi(1+r)^3 - \pi(1+r)^2.$$

Finalmente

$$\mu_3(B(S, 1+r)) = \frac{4}{3}\pi(1+r)^3 + \frac{3}{2}\pi(1+r)^2 \quad r \in [0, +\infty).$$

□

## A.2. Pseudo código del algoritmo del Capítulo 4

### Notación

- $u_{t_i} = (t_i, f(t_i))$  con  $1 \leq i \leq n$ .
- Dados  $t_{i_1} \leq t_{i_2} \leq \dots \leq t_{i_k}$   $P(t_{i_1}, t_{i_2}, \dots, t_{i_k})$  denota la poligonal formada por los puntos  $t_{i_1}, t_{i_2}, \dots, t_{i_k}$ .

### Función auxiliar:

Vamos a definir  $F(t_i, t_j, t_k) \rightarrow \{0, 1\}$  donde  $t_i < t_j < t_k \subset \{t_1, \dots, t_n\}$ .

Denotemos  $\theta_{t_j}^1 = \max\{\theta_1, \dots, \theta_k\}$  donde los  $\theta_s$  son los ángulos que forma el vector  $u_{t_k} - u_{t_j}$  con

vectores de la forma  $u_{t_s} - u_{t_j}$  y  $t_s$  son observaciones tal que  $t_k \leq t_s$  y  $f(t_k) \leq f(t_s)$ . Denotemos  $t_v$  el valor en el cual se da dicho máximo, (puede ser  $t_v = t_k$ ). Denotemos  $\xi_1$  el vector que resulta de aplicar la rotación en sentido anti horario, de ángulo  $\rho/2$  del vector  $\frac{u_{t_v} - u_{t_j}}{\|u_{t_v} - u_{t_j}\|}$  y el cono  $C_{\rho, \xi_1}(u_{t_j})$  con vértice en  $u_{t_j}$  y eje el vector  $\xi_1$ .

Denotemos  $\theta_{t_j}^2 = \max\{\theta_1, \dots, \theta_p\}$  donde los  $\theta_s$  son los ángulos que forma el vector  $u_{t_i} - u_{t_j}$  con vectores de la forma  $u_{t_s} - u_{t_j}$  y  $t_s$  son observaciones tal que  $t_s \leq t_i$  y  $f(t_i) \leq f(t_s)$ . Denotemos  $t_u$  el valor en el cual se da dicho máximo, (puede ser  $t_u = t_i$ ). Denotemos  $\xi_2$  el vector que resulta de aplicar la rotación en sentido horario, de ángulo  $\rho/2$  del vector  $\frac{u_{t_u} - u_{t_j}}{\|u_{t_u} - u_{t_j}\|}$  y el cono  $C_{\rho, \xi_2}(u_{t_j})$  con vértice en  $u_{t_j}$  y eje el vector  $\xi_2$ .

Definimos

- $F(t_i, t_j, t_k) = 1$  si  $C_{\rho, \xi_1}(u_{t_j}) \cap \{t_1, \dots, t_n\} = \emptyset$  o  $C_{\rho, \xi_2}(u_{t_j}) \cap \{t_1, \dots, t_n\} = \emptyset$ .
- $F(t_i, t_j, t_k) = 0$  en caso contrario.

**Paso base:** Definimos  $f^0 = \max_{t_1 \leq t_i \leq t_n} f(t_i)$ . Definimos  $t_2^1 = \arg \max_{1 \leq i \leq n} f(t_i)$ . Definimos  $I_1^1 = (t_1, t_2^1)$  y  $I_2^1 = (t_2^1, t_n)$ . Puede pasar que uno de dichos intervalos sea vacío, en cuyo caso  $t_2^1 = t_1$  o  $t_2^1 = t_n$ .

**Paso inductivo:** Supongamos que tenemos, para  $k \geq 1$ ,  $m(k)$  intervalos  $I_1^k, \dots, I_{m(k)}^k$ , donde  $I_1^k = (t_1, t_2^k)$  y  $I_{m(k)}^k = (t_{m(k)}, t_n)$ . Denotamos  $t_1^k = t_1$  y  $t_{m(k)+1}^k = t_n$  para todo  $k$ .

Mientras que  $m(k) < m(k+1)$

Mientras que  $r \leq m(k)$ .

Sea  $I_r^k = (t_r^k, t_{r+1}^k)$ , definimos  $l_1 = \arg \max_{j: t_r^k < t_j < t_{r+1}^k} f(t_j)$ .

- Si  $F(t_r^k, l_1, t_{r+1}^k) = 1$ , definimos  $t_r^{k+1} = t_r^k$ ,  $t_{r+1}^{k+1} = l_1$ ,  $t_{r+2}^{k+1} = t_{r+1}^k$  y  $r = r + 1$ .

- Si  $F(t_r^k, l_1, t_{r+1}^k) = 0$ , sea  $l_2 = \arg \max_{j: t_r^k < t_j < t_{r+1}^k, t_j \neq l_1} f(t_j)$ .

- Si  $F(t_r^k, l_2, t_{r+1}^k) = 1$ , definimos  $t_r^{k+1} = t_r^k$ ,  $t_{r+1}^{k+1} = l_2$ ,  $t_{r+2}^{k+1} = t_{r+1}^k$  y  $r = r + 1$ .

⋮

La iteración sigue hasta que  $F(t_r^k, l_i, t_{r+1}^k) = 1$  para algún  $l_i$ , en cuyo caso tomamos  $r = r + 1$  o  $F(t_r^k, l_i, t_{r+1}^k) = 0$  para todo  $t_r^k \leq l_i \leq t_{r+1}^k$  donde

$l_i = \arg \max_{j: t_r^k < t_j < t_{r+1}^k, t_j \notin \{l_1, \dots, l_{i-1}\}} f(t_j)$  en cuyo caso  $f^k(x) = P(t_r^k, t_{r+1}^k)$  en  $I_1^k$ , definimos  $t_{r+1}^{k+1} = t_{r+1}^k$  y  $r = r + 1$ .

$k = k + 1$

# Índice alfabético

- condición de Besicovich
  - débil, 78
  - fuerte, 78
- conjunto
  - cono convexo
    - estimación de la medida de la frontera, 43
    - tasas de convergencia en medida, 40
    - consistencia en Hausdorff, 37
    - consistencia en Medida, 38
    - definición, 30
  - cono convexo por complementos
    - definición, 30
  - estándar, 10
  - con alcance positivo, 11
  - de volumen polinómico, 18
  - P-continuo, 35
  - paralelo, 8
  - r-convexo, 10
- contenido de Minkowski, 13
  - exterior, 13
- diagrama de Voronoi, 14
- dimensión de un espacio métrico, 78
- distancia
  - en medida, 9
  - de Hausdorff, 8
- distancias entre funciones
  - de Hausdorff entre los hipografos, 68
  - Lévy, 68
  - Skorohod, 67
- envolvente cono convexa, 31
- espacio maximal, 45
- fórmula de Federer, 18
- familia inevitable, 40
- función
  - càdlàg, 67
  - hipografo de, 68
  - semicontinua superiormente, 68
- granulometría, 10
- P-Uniformidad, 36
- rodamiento de una bola, 10
- suma de Minkowski, 8
- triangulación de Delaunay, 14

# Bibliografía

- ALT, H.; BRASS, P.; GODAU, M.; KNAUER, C. y WENK, C.: *Computing the Hausdorff Distance of Geometric Patterns and Shapes*. Springer-Verlag Berlin Heidelberg, 2003.
- AMBROSIO, L.; COLESANTI, A. y VILLA, E.: «Outer Minkowski content for some classes of closed sets and applications to stochastic geometry». *Math. Ann.*, 2008, **342**, pp. 727–748.
- ARMENDÁRIZ, I.; CUEVAS, A. y FRAIMAN, R.: «Nonparametric estimation of boundary measures and related functionals: asymptotic results». *Adv. in Appl. Probab.*, 2009, **41**, pp. 311–322.
- AURENHAMMER, F. y KLEIN, R.: *Voronoi diagrams*. Elsevier Science Publishing, 2000.
- BAÍLLO, A.; CUEVAS, A. y FRAIMAN, R.: *Classification methods for functional data*. Oxford Handbooks in Mathematics, 2010.
- BILLINGSLEY, P.: *Convergence of probability measures, s.e.* Wiley, 1999.
- BILLINGSLEY, P. y TOPSØE, F.: «Uniformity in weak convergence.» *Wahrscheinlichkeitstheorie Verw. Geb.*, 1967, **7**, pp. 1–16.
- BLOWEY, J.F. y CRAIG, A.W.: *Frontiers of numerical analysis*. Durham, 2004.
- BRÄKER, H. y HSING, T.: «On the area and perimeter of a random convex hull in a bounded convex set». *Probability theory and related fields*, 1998, **11**, pp. 517–550.
- BURENKOV, V.I.: *Sobolev spaces on domains*. Teubner-TextezurMathematik, 1998.
- CÉROU, F. y GUYADER, A.: «Nearest neighbor classification in infinite dimension». *ESAIM: Probability and Statistics*, 2006, **10**, pp. 340–355.
- CHE, Z.Q.; WILLIAMS, R.J. y ZHAO, Z.: «On the existence of positive solutions for semilinear elliptic equations with singular lower order coefficients and Dirichlet boundary conditions». *Mathematische Annalen*, 1999, **315**, pp. 735–769.
- COLESANTI, A. y MANSELLI, P.: «Geometric and isoperimetric properties of sets of positive reach in  $E^d$ ». *Atti Semin. Mat. Fis. Univ. Modena Reggio Emilia*, 2010, **57**, pp. 97–113.
- CUEVAS, A. y FRAIMAN, R.: «On visual distances in density estimation: The Hausdorff choice», 1998, **40**, pp. 333–341.
- CUEVAS, A.; FRAIMAN, R. y GYÖRFI, L.: «Towards a universally consistent estimator of the Minkowski content». *ESAIM: Probability and Statistics.*, 2013, **17**, pp. 359–369.
- CUEVAS, A.; FRAIMAN, R. y PATEIRO-LÓPEZ, B.: «On statistical properties of sets fulfilling rolling-type conditions». *Adv. in Appl. Probab.*, 2012, **44**, pp. 311–329.
- CUEVAS, A.; FRAIMAN, R. y RODRIGUEZ-CASAL, A.: «A nonparametric approach to the estimation of lengths and surface areas». *Ann. Statist.*, 2007, **0**, pp. 1–23.

- 
- CUEVAS, A. y RODRIGUEZ-CASAL, A.: «On boundary estimation.» *Adv. in Appl. Probab.*, 2004, **36**, pp. 340–354.
- DAS, M.: «Strong and weak vitali properties». *Real Anal. Exchange*, 2001, **27**, pp. 7–16.
- DEVROYE, L.; GYÖRFI, L. y LUGOSI, G.: *A probabilistic theory of pattern recognition*. Springer-Verlag, New York., 1996.
- DEVROYE, L. y WISE, G.L.: «Detection of abnormal behavior via nonparametric estimation of the support». *Appl. Math*, 1980, **38**, pp. 480–488.
- DIGGLE, P.J.: «Statistical analysis of spatial point patterns.» *Academic Press.*, 1983.
- DÜMBGEN, L. y WALTHER, G.: «Rates of convergence for random approximations of convex sets». *Adv. Appl. Probab.*, 1996, **28**, pp. 384–393.
- EDGAR, G.: «Packing measures in general metric space». *Real Anal. Exchange*, 2001, **26**, pp. 831–852.
- ERDOS, P.: «Some remarks on the measurability of certain sets.» *Bull. Amer. Math. Soc.*, 1945, **51**, pp. 728–731.
- FEDERER, H.: «Curvature measures». *Trans. Amer. Math. Soc.*, 1959, **93**, pp. 418–491.
- : *Geometric measure theory*. Springer Verlag, 1969.
- FOLLAND, G.: *Real Analysis, Modern Techniques and Their Applications*. Jhon Wileyger & Sons, 1984.
- FORZANI, L.; FRAIMAN, R. y LLOP, P.: «Consistent nonparametric regression for functional data under the Stone-Besicovitch conditions». *Information Theory, IEEE Transactions on*, 2012, **58**, pp. 6697–6708.
- GENOVESE, C.; PERONE-PACIFICO, M.; VERDINELLI, I. y WASSERMAN, L.: «The Geometry of Nonparametric Filament Estimation». *Journal of the American Statistical Association*, 2012a, **107**, pp. 788–799.
- : «Minimax Manifold Estimation». *Journal of Machine Learning Research*, 2012b, **13**, pp. 1263–1291.
- GILBARG, D. y TRUDINGER, N.S.: *Elliptic partial differencial equations of second order*. Springer, 2001.
- HEINONEN, J.: *Lectures on analysis on metric spaces*. Springer, 2001.
- HEVELING, M.; HUG, D. y LAST, G.: «Does polynomial parallel volume imply convexisty?» *Math. Ann.*, 2004, **328**, pp. 469–479.
- HOLÁ, L.: «Hausdorff metric on the space of upper semicontinuous multifunctions», 1992, **22**, pp. 601–610.
- HUREWICZ, W. y WALLMAN, H.: *Dimension theory*, 1948.
- HUTCHINGS, M. J.: «Standing crop and pattern in pure stands of *Mercurialis perennis* and *Rubus fruticosus* in mixed deciduous woodland.» *Oikos.*, 1979, **31**, pp. 351–357.
- JANSON, S.: «Maximal spacings in several dimensions.» *Ann. Prob.*, 1987, **15**, pp. 274–280.
- JIMÉNEZ, R. y YUKICH, J.E.: «Nonparametric estimation of surface integrals.» *Annals of Statistics*, 2011, **39**, pp. 232–260.

- 
- KRISHNAN, T. y McLACHLAN, G.: *The EM Algorithm and Extensions*, s.e. Wiley, 2008.
- KURATOWSKI, K.: *Topology*. Academic Press, 1966.
- LARMAN, D.G.: «A new theory of dimension». *Proc. London Math. Soc.*, 1967, **17**, pp. 178–192.
- LEBESGUE, H.: «Conditions de régularité, conditions d'irrégularité, conditions d'impossibilité dans le probleme de Dirichlet». *C. R. Acad. Sci. Paris*, 1924, **178**, pp. 349–354.
- MARRON, J.S. y TSYBAKOV, A.B.: «Visual error criteria for qualitative smoothing», 1995, **90**, pp. 499–507.
- MATILLA, P.: *Geometry of sets and measures in Euclidean spaces*. Cambridge Studies in Advanced Mathematics, 1995.
- MORGAN, F.: *Geometric measure theory, a beginner's guide. Third Edition*. Academic Press, 2000.
- MÖRTERS, P. y PERES, Y.: *Brownian Motion*. Cambridge University Press, 2010.
- MORVAN, J.: *Generalized Curvatures*. Springer, 2008.
- NATANSON, I.: *The theory of functions of a real variable, vol. 2*. Frederick Ungar., 1960.
- NUTANONG, S.; JACOX, E. y SAMET, H.: «An incremental Hausdorff distance calculation algorithm», 2011, **4**, pp. 506–517.
- PATEIRO-LÓPEZ, B.: *Set estimation under convexity type restrictions*. Tesis doctoral, Santiago de compostela, 2008.
- PATEIRO-LÓPEZ, B. y RODRÍGUEZ-CASAL, A.: «Generalizing the convex hull of a sample: The R package alphahull.» *J. Stat. Softw.*, 2010, **5**, pp. 1–28.
- POINCARÉ, H.: «Théorie du potentiel Newtonien». *Leçons professées á la Sorbonne, Gauthier - Villars*, 1899.
- PREISS, D.: «Dimension of metrics and differentiation of measures». *General topology and its relations to modern analysis and algebra V*, 1983, pp. 565–568.
- RÉNYI, A. y SULANKE, R.: «Über die konvexe Hülle von n zufällig gewählten Punkten». *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 1963, **2**, pp. 75–84.
- : «Über die konvexe Hülle von n zufällig gewählten Punkten (II)». *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 1964, **3**, pp. 138–147.
- ROBERTS, A.W. y VARBERG, D.E.: *Convex Functions*. Academic Press, 1973.
- ROCKAFELLAR, R.T y WETS, R.: *Variational Analysis*. Springer, 2009.
- RODRÍGUEZ-CASAL, A.: *Estimación de conjuntos y sus fronteras, un enfoque geométrico*. Tesis doctoral, Santiago de Compostela, 2002.
- : «Set estimation under convexity type assumptions.» *Annales de l'I.H.P.- Probabilites & Statistiques*, 2006, **43**, pp. 763–774.
- RUZHANSKY, M.: «On uniform properties of doubling measures». *Proc. of Amer. Math. Soc.*, 2001, **129**, pp. 3413–3416.
- SANTALÓ, L.: *Integral Geometry and Geometric Probability*. Cambridge University Press, 1976.
-

- 
- SAORÍN, E.: *On inner parallel bodies. From the steiner polynomial to Poincaré inequality*. Tesis doctoral, Universidad de Murcia, 2008.
- SCHARF, L.: *Computing the Hausdorff distance between sets of curves*. Tesina o Proyecto, Freire Universität Berlin, 2003.
- SENDOV, B.: *Hausdorff approximations*. Kluwer, Dodrecht, 1990.
- SERRA, J.: *Image analysis and mathematical morphology*. Academic press, inc, 1982.
- SHORACK, G. R. y A, WELLNER J.: *Empirical Processes with Applications to Statistics*. Wiley, 1986.
- SHYRIAIEV, A.N.: *Andre Nikolaevich Kolmogorov. A Biographical Sketch of His Life and Creative Paths*. American Mathematical Society, 2007.
- STACHÓ, L.L.: «On the volume function of parallel sets». *Acta Sci. Math.*, 1976, **38**, pp. 365–374.
- STONE, C.: «Consistent Nonparametric Regression». *The Annals of Statistics*, 1977, **5**, pp. 595–620.
- TALAGRAND, M.: «The Glivenko-Cantelli problem.» *Ann. Probab.*, 1987, **15**, pp. 837–870..
- THÄLE, C.: «50 years of positive reach, a survey». *Surveys in Mathematics and its Applications*, 2008, **3**, pp. 123–165.
- VAN DER VAART, A. W.: *Asymptotic statistics*. Cambridge University Press, 1998.
- WALTHER, G.: «Granulometric smoothing». *Ann. Statist*, 1997, **25**, pp. 2273–2299.
- : «On a generalization of Blaschke’s rolling theorem and the smoothing of surfaces». *Math. Methods Appl. Sci.*, 1999, **22**, pp. 301–316.
- ZAREMBA, S.: «Sur le principe du minimum». *Bull. Int. Acad. Sci. Cracovie*, 1909, **7**, pp. 197–264.
- : «Sur le principe de Dirichlet». *Acta Mathematica*, 1911, **34**, pp. 293–316.