UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA

"El teorema de espacios complementarios y análisis local de espacios normados"

T E S I S

que para obtener el título de

Licenciado en Matemáticas Aplicadas

presenta

Octavio Alberto Agustín Aquino

Directora de tesis:

M. en C. Luz del Carmen Álvarez Marín

Huajuapan de León, Oaxaca, a 9 de agosto de 2006

Dedico este trabajo con cariño:

a mis padres,
Zoila D. Aquino Pérez y
Gregorio Alberto Agustín Escamilla;

a mis hermanas,
Adriana y Valeria;

y a Angélica,
que tanto me ayudaron para no
desanimarme estando en Guanajuato.

Prefacio

Kanu ko ini.
(Que muy grande sea su alma).

Alocución mixteca

Hablando un tanto informalmente, podemos decir que la caracterización de un objeto matemático A es de la forma "A es el único objeto con la propiedad P", donde P generalmente es una hermosa propiedad que se puede verificar fácilmente. O, si no se puede verificar fácilmente, al menos justifica el estudio del objeto A, porque algunas manipulaciones que podemos hacer con él son útiles.

Los problemas de caracterización suelen ser muy interesantes y difíciles. En el caso del análisis funcional esto aplica particularmente bien. La exposición que sigue versa principalmente sobre una caracterización importante de los espacios de Hilbert y algunos preliminares necesarios para comprenderla. Me refiero al Teorema de Espacios Complementarios, que fue una conjetura durante gran parte del siglo XX hasta que Joram Lindenstrauss y Lior Tzafriri consiguieron demostrarlo en 1971.

La estructura del trabajo es la siguiente. Empezamos por fijar notación y definir algunas cuestiones básicas en el Capítulo 1. Ahí demostramos una caracterización de espacios de Banach que son de Hilbert para motivar el problema de los espacios complementarios. Luego haremos una observación geométrica que nos conduce (de forma un tanto inesperada) a la desigualdad de la media aritmética y geométrica (MA/MG).

En el Capítulo 2 se estudia con cierto detalle la medida de Haar. Sobre ella desarrollaremos algunos resultados de concentración respecto a funciones lipschitzianas sobre la esfera n-dimensional. Para todo eso se utilizan varias desigualdades que tienen el mismo sabor que el de la MA/MG. Al final también discutimos la concentración para almártagas.

Durante el Capítulo 3 hacemos un interludio e incursionamos en las aplicaciones de la concentración en la teoría de códigos.

Retomamos el tema original en el Capítulo 4, con los preámbulos para comprender la demostración del Teorema de Espacios Complementarios. Aquí destaca el teorema de Aryeh Dvoretzky, que examinamos detenidamente. La demostración de este resultado es toda una aventura, y sus consecuencias permean profundamente el análisis funcional, al grado de considerársele el primer resultado del llamado análisis local de espacios normados.

A continuación, en el Capítulo 5, se expone lo concerniente a las demostraciones de los resultados de Joichi y Davis, Dean y Singer que también se requieren para demostrar el Teorema de Espacios Complementarios.

Concluimos en el Capítulo 6 con la demostración del resultado principal que motiva esta tesis, con una previa discusión de la relación que tiene el problema de espacios complementarios con la mejor aproximación en espacios de Hilbert.

Se incluyen cuatro apéndices. Los tres primeros contienen los resultados y definiciones de teoría de la medida, probabilidad y análisis funcional que completan y facilitan la comprensión del resto de la obra. El último trata sobre un resultado combinatorio que se requiere en el Capítulo 2.

Aunque he tratado, tanto como me ha sido posible, hacer el texto autocontenido, es necesario pedirle al lector que tenga algunos conocimientos básicos sobre los espacios de Hilbert y de Banach. Como referencias en este sentido, recomiendo a [Jam74], [KF75], [FG97], [Day73] y [Yos80]. También se requieren conceptos de la teoría de la medida. Para estos temas, consúltense [GF02] o [Rud79].

Este escrito es, en gran medida, producto de mi estancia profesional en el Centro de Investigación en Matemáticas durante los meses de julio y agosto de 2004, bajo la supervisión de la M. en C. Helga Fetter Nathansky. Mis agradecimientos expresos van para ella por su incisivo énfasis en el rigor y sus ingeniosas sugerencias, el inmenso apoyo que me brindó durante mi estancia en el CIMAT, y la paciente revisión bajo su agudo criterio de lo que escribí.

Aprovecho aquí también para agradecer a la M. en C. Luz del Carmen Álvarez Marín por ser mi directora de tesis, sus valiosas sugerencias, su paciencia para ayudarme a pulir y completar esta tesis y señalarme una gran cantidad de errores e imprecisiones. Más aún, por enseñarme muchos hábitos que un buen matemático debe tener. Tampoco puedo dejar de mostrar mi gratitud al M. en C. Adolfo Maceda Méndez por sus magníficas clases y ejercicios (que me mantuvieron más de una noche en vela); sus acertadas observaciones sin duda han contribuído a mejorar enormemente este escrito. Agradezco asimismo al M. en C. José Margarito Hernández Morales y al Dr. Guillermo Arturo Lancho Romero sus numerosas correcciones y comentarios.

Octavio Alberto Agustín Aquino,
Oaxaca de Juárez, Oaxaca,
9 de agosto de 2006.

Índice general

Capítulo 1

Introducción

Un viaje de mil kilómetros empieza
por el primer paso.

Lao Tsé

En este capítulo presentaré algunos de los resultados y definiciones pertinentes para la comprensión del texto. Empezaremos por ver qué es lo que vamos a caracterizar. Entendemos por $\mathbb{K}$ ya sea al cuerpo de los reales o al de los complejos.

1.1. Algunas definiciones

Definición 1.1. Un espacio normado es un espacio vectorial X sobre $\mathbb{K}$ que posee una norma $\|\cdot\| : X \to \mathbb{K}$. Si X también es completo, entonces se dirá que es un espacio de Banach.

Un caso particular de los espacios de Banach es el que tiene mayor interés para nosotros.

Definición 1.2. Un espacio vectorial X es un espacio de producto interior si en él está definido un producto interior $(\cdot, \cdot) : X \times X \to \mathbb{K}$.

Escolio 1.1. Cualquier espacio X de producto interior tiene una norma inducida por

$$\|x\| = \sqrt{(x, x)}.$$

Definición 1.3. Si X es un espacio de producto interior que es completo respecto a la norma inducida por su producto interior, entonces se dice que es un espacio de Hilbert.

Precisamente aquí empieza lo interesante de nuestra discusión. Es obvio que todo espacio de Hilbert es también de Banach pero, ¿bajo qué condiciones podemos asegurar que un espacio de Banach es también de Hilbert? Dicho de otra forma, queremos saber los requerimientos que debe satisfacer un espacio de Banach para que podamos definir un producto interior en él, de modo que su norma esté inducida por dicho producto. Esto es precisamente de lo que trata este escrito: *caracterizar* a los espacios de Banach que también son de Hilbert.

1.2. Paralelogramo y polarización

Hay una caracterización muy sencilla que quiero incluir en esta introducción, por la belleza que posee. Necesitamos algunos resultados preliminares. En lo sucesivo, y a menos que se indique otra cosa, H denotará un espacio de Hilbert.

Lema 1.1 (Ley del paralelogramo)**.** *Para cualesquiera $x, y \in H$ se satisface*

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right). \tag{1.1}$$

Demostración. Por un lado,

$$\begin{aligned} \|x+y\|^2 &= (x+y, x+y) \qquad (1.2)\\ &= (x,x) + (x,y) + (y,x) + (y,y) \\ &= \|x\|^2 + 2\Re(x,y) + \|y\|^2, \end{aligned}$$

y por otro

$$\begin{aligned} \|x-y\|^2 &= (x-y, x-y) \qquad (1.3)\\ &= (x,x) - (x,y) - (y,x) + (y,y) \\ &= \|x\|^2 - 2\Re(x,y) + \|y\|^2. \end{aligned}$$

Sumando (1.2) y (1.3) obtenemos la identidad deseada. □

Se escribe la norma de H en términos de su producto interior. ¿Podremos hacerlo al revés, y escribir al producto interior en términos de la norma? Respondemos con el siguiente lema para el caso en el que $\mathbb{K} = \mathbb{C}$.

Lema 1.2 (Identidad de polarización)**.** *Si $x, y \in H$, entonces*

$$(x,y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right). \tag{1.4}$$

Demostración. Por una parte tenemos que,

$$\begin{aligned} \|x+y\|^2 - \|x-y\|^2 &= (x+y, x+y) - (x-y, x-y) \\ &= 2(x,y) - 2(y,x) \\ &= 2\left[(x,y) + \overline{(x,y)}\right] \\ &= 4\Re(x,y), \end{aligned}$$

y por otra,

$$\begin{aligned} \|x+iy\|^2 - \|x-iy\|^2 &= (x+iy, x+iy) - (x-iy, x-iy) \\ &= -2i(x,y) + 2i(y,x) \\ &= 2i\left[\overline{(x,y)} - (x,y)\right] \\ &= 4\Im(x,y). \end{aligned}$$

Luego

$$\Re(x,y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$$

e

$$\Im(x,y) = \frac{1}{4}\left(\|x+iy\|^2 - \|x-iy\|^2\right).$$

Puesto que $(x,y) = \Re(x,y) + i\Im(x,y)$, se sigue la ecuación (1.4). □

Veamos que si H es un espacio normado real, entonces a través de la parte real de la expresión (1.4) podemos definir un producto interior en H.

Lema 1.3. *Sea H un espacio normado real cuya norma satisface la ley del paralelogramo. Defínase, para cualquier $x, y \in H$,*

$$(x,y)_1 = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right).$$

Entonces se satisfacen las siguientes propiedades.

1. $(\alpha x, y)_1 = \alpha(x,y)_1$.
2. $(x+y, z)_1 = (x,z)_1 + (y,z)_1$.
3. $(x,y)_1 = (y,x)_1$.
4. $(x,x)_1 = \|x\|^2$.

Demostración. Las dos últimas son claras a partir de la definición. Consideremos que

$$(x,z)_1 + (y,z)_1 = \frac{1}{4}\left(\|x+z\|^2 - \|x-z\|^2 + \|y+z\|^2 - \|y-z\|^2\right),$$

y que, según la ley del paralelogramo, esto equivale a

$$\begin{aligned}(x,z)_1 + (y,z)_1 = {} & \frac{1}{4}\left[\frac{1}{2}\left(\|x+y+2z\|^2 + \|x-y\|^2\right)\right] \\ & - \frac{1}{4}\left[\frac{1}{2}\left(\|x+y-2z\|^2 + \|x-y\|^2\right)\right].\end{aligned}$$

Haciendo las manipulaciones pertinentes obtenemos

$$\begin{aligned}(x,z)_1 + (y,z)_1 &= \frac{1}{2}\left(\left\|\frac{x+y}{2}+z\right\|^2 - \left\|\frac{x+y}{2}-z\right\|^2\right) \\ &= 2\left(\frac{x+y}{2}, z\right)_1,\end{aligned}$$

y podemos escoger $y = 0$; esto nos conduciría a $(x,z)_1 = 2\left(\frac{x}{2}, z\right)_1$, lo cual implica que

$$(x,z)_1 + (y,z)_1 = 2\left(\frac{x+y}{2}, z\right)_1 = (x+y,z)_1,$$

y se satisface el segundo inciso. Por lo tanto

$$(2x,y)_1 = (x+x,y)_1 = (x,y)_1 + (x,y)_1 = 2(x,y)_1.$$

Suponiendo que se cumpla

$$((n-1)x,y)_1 = (n-1)(x,y)_1,$$

con $n \in \mathbb{N}$, entonces, usando el segundo inciso,

$$\begin{aligned}(nx,y)_1 &= ((n-1)x,y)_1 + (x,y)_1 \\ &= (n-1)(x,y)_1 + (x,y)_1 = n(x,y)_1,\end{aligned} \tag{1.5}$$

lo que concluye el argumento inductivo que prueba que la primera propiedad es válida para $\alpha \in \mathbb{N}$. Además,

$$(x,y)_1 = \left(\frac{n}{n}x, y\right)_1 = n\left(\frac{1}{n}x, y\right)_1,$$

luego la primera propiedad se cumple para todo racional por (1.5). De la desigualdad

$$|\|\alpha x + y\| - \|\beta x + y\|| \leq \|(\alpha - \beta)x\| = |\alpha - \beta| \, \|x\|$$

se tiene que $\|\alpha x + y\|$ es continua en α. La densidad de los racionales en $\mathbb{R}$ nos garantiza ahora que para todo $\alpha \in \mathbb{R}$ existe una sucesión de racionales $\{\alpha_n\}_{n=1}^{\infty}$ tales que $\lim_{n\to\infty} \alpha_n = \alpha$. Se sigue que

$$(\alpha x, y)_1 = \left(\lim_{n\to\infty} \alpha_n x, y\right)_1 = \lim_{n\to\infty} \alpha_n (x, y)_1 = \alpha(x, y)_1. \qquad \square$$

Teorema 1.1. *En un espacio de Banach H la norma satisface la ley del paralelogramo si, y sólo si, H es un espacio de Hilbert.*

Demostración. La suficiencia es obvia. Probemos la necesidad. Definamos en H para cualesquiera $x, y \in H$

$$(x, y)_2 = (x, y)_1 + i(x, iy)_1.$$

Por lo demostrado en el lema, para que tal operación sea un producto interior sólo falta demostrar que $(x, y)_2 = \overline{(y, x)_2}$ y que $(ix, y)_2 = i(x, y)_2$. Atendamos primero al hecho de que

$$\begin{aligned} (ix, iy)_1 &= \frac{1}{4}\left(\|ix + iy\|^2 - \|ix - iy\|^2\right) \\ &= \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2\right) \\ &= (x, y)_1. \end{aligned}$$

Ahora bien,

$$(x, iy)_1 = (-iix, iy)_1 = (-ix, y)_1 = -(ix, y)_1 = -(y, ix)_1,$$

y por lo tanto

$$\begin{aligned} (x, y)_2 &= (x, y)_1 + i(x, iy)_1 \\ &= (y, x)_1 - i(y, ix)_1 \\ &= \overline{(y, x)_2}. \end{aligned}$$

Finalmente,

$$\begin{aligned} (ix, y)_2 &= (ix, y)_1 + i(ix, iy)_1 \\ &= (ix, -iiy)_1 + i(x, y)_1 \\ &= (x, -iy)_1 + i(x, y)_1 \\ &= -(x, iy)_1 + i(x, y)_1 \\ &= i(x, y)_2, \end{aligned}$$

que, junto con lo ya demostrado, implica que H es un espacio con producto interior. $\square$

Existen espacios de Banach que no son de Hilbert. Consideremos al espacio ℓ_∞ de las sucesiones reales acotadas bajo la norma del supremo. Sean

$$x = (1, 0, 1, 0, \ldots), \quad y = (0, 1, 0, 1, \ldots),$$

entonces

$$x + y = (1, 1, 1, 1, \ldots), \quad x - y = (1, -1, 1, -1, \ldots)$$

y así

$$\|x\| = 1, \quad \|y\| = 1, \quad \|x + y\| = 1, \quad \|x - y\| = 1.$$

A través de un cálculo, nos damos cuenta de que no se cumple la ley del paralelogramo,

$$\|x + y\|^2 + \|x - y\|^2 = 2 \neq 2(\|x\|^2 + \|y\|^2) = 4.$$

Por lo tanto, según el Teorema 1.1, ℓ_∞ es de Banach pero no de Hilbert.

1.3. Una interpretación geométrica

Quiero concluir este breve capítulo con una observación sobre la identidad de polarización.

Definición 1.4. Sean V y W espacios vectoriales sobre $\mathbb{C}$, y $f : V \times V \to W$ una función tal que para cada $x_0, y_0 \in V$ los operadores $f_{x_0}(x) = f(x_0, x)$ y $f_{y_0}(x) = f(x, y_0)$ son lineal y lineal conjugado, respectivamente. Entonces se dice que f es una forma sesquilineal sobre V. Si $f(x, y) = \overline{f(y, x)}$, se dice que la forma sesquilineal es hermítica o autoadjunta.

Recuérdese que todo cuerpo es un espacio vectorial sobre cualquier subcampo de sí mismo. Pongamos que V es un espacio con producto interior. Si en la definición anterior tomamos $W = \mathbb{C}$, vemos precisamente que el producto interior es una forma sesquilineal; hermítica, además.

Definición 1.5. Con la notación de la definición anterior, un mapeo $Q : V \to W$ es una forma cuadrática si existe una forma sesquilineal hermítica f tal que

$$Qx = f(x, x).$$

Obsérvese que, en un espacio de Hilbert, el cuadrado de la norma es una forma cuadrática. Ahora bien, si nos dan un mapeo $Q : V \to W$ entre los espacios vectoriales V y W, y sabemos que es una forma cuadrática, podemos recuperar la forma sesquilineal hermítica que lo engendró, pues

$$\begin{aligned} Q(x+y) &= f(x+y,x+y) \\ &= f(x,x+y)+f(y,x+y) \\ &= f(x,x)+f(x,y)+f(y,x)+f(y,y) \\ &= Qx+2\Re f(x,y)+Qy, \end{aligned}$$

y

$$\begin{aligned} Q(x+iy) &= f(x+iy,x+iy) \\ &= f(x,x+iy)+f(y,x+iy) \\ &= f(x,x)+f(x,iy)+f(iy,x)+f(iy,iy) \\ &= Q(x)-if(x,y)+if(y,x)+f(y,y) \\ &= Qx+2\Im f(x,y)+Qy, \end{aligned}$$

de donde

$$Q(x+y)+Q(x+iy)-2Qx-2Qy=2f(x,y).$$

Es fácil ver también, cambiando y por $-y$ en los cálculos anteriores, que se satisface la siguiente identidad de polarización para formas cuadráticas,

$$4f(x,y)=Q(x+y)-Q(x-y)+i\left[Q(x+iy)-Q(x-iy)\right].$$

Ahora bien, la caracterización de espacios con producto interior que acabamos de ver nos indica cuándo podemos decir que el cuadrado de una norma en un espacio normado es una forma cuadrática, y esto ocurre precisamente cuando satisface la identidad de polarización. Veamos un ejemplo en $\mathbb{R}^2$. La función

$$f(x,y)=xy, \tag{1.6}$$

es claramente una forma bilineal hermítica. Observamos que el polinomio

$$Q(x)=f(x,x)=x^2$$

es una forma cuadrática. La identidad de polarización nos revela algo sumamente sorprendente,

$$xy=\frac{1}{4}\left[(x+y)^2-(x-y)^2\right]. \tag{1.7}$$

Consideremos a las funciones

$$g_1(x,y)=\left(\frac{x+y}{2}\right)^2 \tag{1.8}$$

y

$$g_2(x,y) = -\left(\frac{x-y}{2}\right)^2. \tag{1.9}$$

Las curvas definidas por $g_1(x,y) = 0$ y $g_2(x,y) = 0$ son las rectas

$$x + y = 0, \quad x - y = 0, \tag{1.10}$$

que son perpendiculares entre sí. Regresemos nuestra atención a $f(x,y) = xy$, con las curvas definidas por $f(x,y) = k$. Tenemos las hipérbolas

$$xy = k \tag{1.11}$$

y, curiosamente, las rectas (1.10) son los ejes de las hipérbolas (1.11). Entonces hemos expresado un paraboloide hiperbólico (1.6) como la suma de dos cilindros paraboloidales (1.8) y (1.9), cuyas curvas de nivel son perpendiculares entre sí, es decir, van sobre dos direcciones perpendiculares de polarización. De ahí el nombre de la identidad: nos permite conocer tal descomposición.

1.3.1. La desigualdad MA/MG

En lo que sigue haremos uso de un resultado muy socorrido por Cauchy: la desigualdad de la media aritmética y geométrica, que abreviaremos como MA/MG.

Recordemos que para un conjunto finito $A = \{a_k\}_{k=1}^n$ de números reales, su media aritmética es

$$M_\alpha(A) = \frac{1}{n}\sum_{k=1}^{n} a_k,$$

y su media geométrica (siempre que A consista solamente en números reales no negativos) es

$$M_\gamma(A) = \left(\prod_{k=1}^{n} a_k\right)^{\frac{1}{n}}.$$

Teorema 1.2 (Desigualdad MA/MG)**.** *Sea $A = \{a_k\}_{k=1}^n$ un conjunto de números reales no negativos. Entonces*

$$M_\gamma(A) \leq M_\alpha(A).$$

Demostración. A partir de (1.7) es fácil concluir que, de dos pares de reales con igual suma, el par que posee el mayor producto es aquél cuyos elementos presentan la menor diferencia. Si todos los elementos de A son iguales, el teorema es trivial, lo mismo que si $n = 1$. Supongamos, pues, que A tiene dos o más elementos no todos iguales. Necesariamente existen elementos $a_i, a_j \in A$ tales que $a_i > M_\alpha$

y $a_j < M_\alpha$. Sin pérdida de generalidad, podemos suponer que $i = 1$ y $j = 2$. Formamos el conjunto A' tomando $a'_1 = M_\alpha$, $a'_2 = a_1 + a_2 - M_\alpha$ y $a'_j = a_j$ para $j \geq 2$ (si aplica el caso). Entonces

$$a'_1 + a'_2 = a_1 + a_2$$

y

$$\begin{aligned} a'_1 - a'_2 &= 2M_\alpha - (a_1 + a_2) \\ &= a_1 - a_2 - 2(a_1 - M) \\ &\leq a_1 - a_2, \end{aligned}$$

de donde se concluye que

$$\prod_{k=1}^{n} a_k \leq \prod_{k=1}^{n} a'_k = M_\alpha \prod_{k=2}^{n} a'_k.$$

Como el conjunto A es finito, después de definir un conjunto A' continuando este proceso de manera análoga, debemos terminar en que

$$\prod_{k=1}^{n} a_k \leq (M_\alpha)^n .$$

Finalmente,

$$M_\gamma = \left(\prod_{k=1}^{n} a_k \right)^{\frac{1}{n}} \leq M_\alpha. \qquad \square$$

En particular, tomando $a \in \mathbb{R}$ positivo, m menor que n y definiendo

$$a_k = \begin{cases} a & \text{si } k \leq m, \\ 1 & \text{si } k > m, \end{cases},$$

se tiene que

$$a^{\frac{m}{n}} \leq \frac{ma + (n-m)}{n} = \frac{m}{n}a + \left(1 - \frac{m}{n}\right).$$

Puesto que las funciones $f(x) = a^x$ y $g(x) = xa + (1-x)$ son continuas, se sigue de la densidad de los racionales en los reales que para todo $\beta \in [0,1]$ se satisface

$$a^\beta \leq \beta a + 1 - \beta. \tag{1.12}$$

Escolio 1.2. Sean $\beta, \delta \in \mathbb{R}$ tales que $\beta + \delta = 1$. Para cualesquiera $x, y \in \mathbb{R}$ no negativos se satisface

$$x^\beta y^\delta \leq \beta x + \delta y.$$

Si x o y son nulos, el resultado es trivial. Por lo tanto supongamos que tanto y como x no son cero. Claramente $\frac{x}{y} > 0$, y según la desigualdad (1.12)

$$\left(\frac{x}{y}\right)^\beta \leq \beta\left(\frac{x}{y}\right) + 1 - \beta = \frac{\beta x}{y} + \delta,$$

y multiplicando por y obtenemos el resultado.

Capítulo 2

Concentrando la medida

> [El fenómeno de concentración] ... condujo a un nuevo, probabilístico, entendimiento de la estructura de los cuerpos convexos en altas dimensiones.
>
> *Keith Ball*, [Bal97]

El sorpresivo hecho de que la mayor parte de la "masa" de la esfera n-dimensional esté aglutinada alrededor de su ecuador es conocido como "fenómeno de la concentración de la medida". Fue bautizado así por Vitali Milman al descubrir una nueva demostración del Teorema de Dvoretzky que discutiremos más adelante. Esta singular propiedad tiene muchas aplicaciones y es de particular relevancia en la teoría local de espacios normados. El tratamiento aquí será siguiendo en gran medida a [Sch].

Notación 2.1. Denotaremos por S^{n-1} a la esfera unitaria de $\mathbb{R}^n$, es decir

$$S^{n-1} = \left\{x \in \mathbb{R}^n : \|x\|_2 = 1\right\}.$$

2.1. Medida de Haar

Recordemos lo que significa la acción de un grupo sobre un conjunto. Sea X un conjunto y G un grupo, con e el neutro del grupo. Una acción izquierda de G sobre X es un mapeo $* : G \times X \to X$ tal que $ex = x$ y $(g_1 g_2)(x) = g_1(g_2 x)$ para toda $x \in X$ y $g_1, g_2 \in G$. De manera totalmente análoga se define una acción derecha, utilizando en tal caso un mapeo de la forma $* : X \times G \to X$.

Consideremos un espacio métrico compacto (M, ρ), y el espacio vectorial de todas las funciones reales continuas en M, $C(M)$. Una acción isométrica sobre M

es una acción con la propiedad de que para todo $t, s \in M$ se tiene

$$\rho(gt, gs) = \rho(t, s).$$

Definición 2.1. Una medida de Haar sobre M es una medida $\mu : \Sigma \to [0, \infty)$ (donde Σ es una σ-álgebra que contiene a todos los subconjuntos de Borel de M) con las siguientes condiciones.

1. Está normalizada, es decir, $\mu(M) = 1$.
2. Es invariante bajo la acción de un grupo G, es decir, $\mu(gS) = \mu(S)$ para todo $g \in G$ y $S \in \Sigma$.

Para cada $f \in C(M)$ definimos el funcional

$$E(f) = \int_M f(x)\, d\mu(x).$$

El funcional E es lineal y acotado en $C(M)$, y de la forma en que lo hemos definido se desprende que, si χ_S es la función característica del conjunto $S \in \Sigma$

$$\begin{aligned} E(\chi_{gS}) &= \int_M \chi_{gS}(x)\, d\mu(x) \\ &= \int_M \chi_S(gx)\, d\mu(gx) \\ &= \int_M \chi_S(x)\, d\mu(x) = E(\chi_S), \end{aligned}$$

en particular, $E(\chi_M) = E(1) = 1$.

Recordemos que la acción de un grupo G sobre un conjunto M es transitiva si para cualesquiera $t, s \in M$ existe $g \in G$ tal que $gt = s$.

Para una $t \in M$ fija, el conjunto de los elementos de G para los cuales t es un punto fijo

$$G_t = \{g \in G : gt = t\},$$

es un subgrupo de G, y se dice que es un subgrupo de isotropía.

Definición 2.2. Si la acción (izquierda o derecha) isométrica de un grupo G sobre un espacio métrico compacto M es transitiva, entonces se dice que M es un espacio homogéneo de G.

Si M es un espacio homogéneo de G y G_t es el grupo de isotropía de algún $t \in M$, podemos establecer una correspondencia biyectiva entre M y G/G_t. En efecto, identificando a cada $s \in M$ con la clase de equivalencia gG_t de g de tal forma que $gt = s$ se tiene la relación requerida. Si $g_1 t = s$ y $g_2 t = s$, entonces $g_1 t = g_2 t$, lo que implica que $g_1 G_t = g_2 G_t$, y la sobreyectividad de la correspondencia se sigue del hecho de que G actúa transitivamente.

Ejemplo 2.1. Sea $(\cdot,\cdot)$ un producto interior en $\mathbb{R}^n$, y para $x \in \mathbb{R}^n$ hagamos $|x| = \sqrt{(x,x)}$. Sea $G = O(n)$, el grupo ortogonal en $(\mathbb{R}^n, (\cdot,\cdot))$, conformado por todos los operadores lineales en $\mathbb{R}^n$ que preservan el producto interior; esto es,

$$(x,y) = (Lx, Ly), \quad L \in O(n).$$

Podemos identificar a cada elemento de $L \in O(n)$ con una matriz de vectores ortogonales, a saber, su representación matricial

$$[L]_\beta = \begin{pmatrix} Le_1 & \cdots & Le_n \end{pmatrix},$$

donde $\beta = \{e_k\}_{k=1}^n$ es una base ortogonal de $\mathbb{R}^n$ respecto a $(\cdot,\cdot)$. Sea $M = S^{n-1}$ y $\varphi : O(n) \to S^{n-1}$ definida a través de

$$\varphi \begin{pmatrix} Le_1 & \cdots & Le_n \end{pmatrix} = Le_1.$$

Es claro que cada $t \in S^{n-1}$ puede identificarse con un elemento de $O(n-1)$, pues el subconjunto de $O(n)$ cuya representación matricial tiene su primer renglón y su primera columna fijas es un subgrupo de $O(n)$ isomorfo a $O(n-1)$. Precisamente este subgrupo es el grupo isotrópico de t. Por lo tanto,

$$S^{n-1} = O(n)/O(n-1).$$

Teorema 2.1. *Sea M un espacio homogéneo de G tanto por la acción izquierda como por la derecha. Existe una medida regular μ en los subconjuntos de Borel de M que es invariante bajo la acción de los miembros de G, es decir, se cumple la ecuación $\mu(A) = \mu(gA)$ para todo subconjunto boreliano $A \subseteq M$. Alternativamente, para todo $g \in G$ y todo $f \in C(M)$ se tiene*

$$\int f(t)\, d\mu(t) = \int f(gt)\, d\mu(t).$$

Demostración. Para cada $f \in C(M)$ y $g \in G$ sean

$$\begin{aligned} L_g(f)(x) &= f(gx), \\ R_g(f)(x) &= f(xg), \\ \Gamma(f) &= \max_{x\in M} f(x), \\ \gamma(f) &= \min_{x\in M} f(x), \\ v(f) &= \Gamma(f) - \gamma(f), \end{aligned}$$

y

$$\begin{aligned} \mathcal{L}(f) &= \left\{ \sum_{i=1}^n a_i L_{g_i} f : n \in \mathbb{N}, g_i \in G, a_i > 0, \sum_{i=1}^n a_i = 1 \right\} \subseteq C(M), \\ \mathcal{R}(f) &= \left\{ \sum_{i=1}^n a_i R_{g_i} f : n \in \mathbb{N}, g_i \in G, a_i > 0, \sum_{i=1}^n a_i = 1 \right\} \subseteq C(M). \end{aligned}$$

Denotamos por $\overline{\mathcal{L}(f)}$ a la cerradura de $\mathcal{L}(f)$ en $C(M)$. Observamos que para $h \in G$

$$\mathcal{L}(L_h f) = \mathcal{L}(f).$$

El operador L_h es continuo en virtud de la acción isométrica de G. Por lo tanto

$$\overline{\mathcal{L}(L_h f)} = \overline{\mathcal{L}(f)},$$

y además los elementos de $\overline{\mathcal{L}(f)}$ son funciones no negativas siempre que f es no negativa. Afirmamos que el conjunto $\overline{\mathcal{L}(f)}$ es compacto. Atendemos al hecho de que

$$\left\|\sum_{i=1}^{n} a_i L_{g_i} f\right\|_\infty \leq \sum_{i=1}^{n} \|a_i L_{g_i} f\|_\infty = \sum_{i=1}^{n} a_i \|f\|_\infty = \|f\|_\infty .$$

En otras palabras, $\overline{\mathcal{L}(f)}$ es un subconjunto acotado de $C(M)$. Veamos que la familia $\overline{\mathcal{L}(f)}$ es equicontinua. Sea $\epsilon > 0$. Como f tiene un dominio compacto, es uniformemente continua. Por lo tanto, existe $\delta > 0$ tal que, siempre que $\rho(x, y) \leq \delta$, entonces $|f(x) - f(y)| < \epsilon$. Ahora bien $\rho(gx, gy) \leq \delta$ para cualquier $g \in G$, de donde

$$\begin{aligned} \left|\sum_{i=1}^{n} a_i (L_{g_i} f)(x) - \sum_{i=1}^{n} a_i (L_{g_i} f)(y)\right| &\leq \sum_{i=1}^{n} a_i \left|(L_{g_i} f)(x) - (L_{g_i} f)(y)\right| \\ &\leq \sum_{i=1}^{n} a_i \left|f(g_i x) - f(g_i y)\right| \\ &\leq \sum_{i=1}^{n} a_i \epsilon = \epsilon. \end{aligned}$$

Por el teorema de Arzelà-Ascoli, se sigue que $\overline{\mathcal{L}(f)}$ es un subconjunto compacto de $C(M)$. Procedamos a probar que la función $\Gamma(f)$ es continua. Supongamos que $\|f - g\|_\infty < \epsilon$. Existe x_0 tal que

$$\Gamma(f) = f(x_0),$$

y además $f(x_0) - g(x_0) \leq \|f - g\|_\infty < \epsilon$, luego

$$\Gamma(f) = f(x_0) < g(x_0) + \epsilon \leq \Gamma(g) + \epsilon,$$

es decir

$$\Gamma(f) - \Gamma(g) \leq \epsilon,$$

y análogamente

$$\Gamma(g) - \Gamma(f) \leq \epsilon,$$

por lo que $|\Gamma(g) - \Gamma(f)| \leq \epsilon$, y Γ es continua. Eso basta para ver que v es continua, pues por la misma razón que Γ es continua también γ lo es, y la suma de funciones continuas es continua. Por lo tanto, v alcanza un mínimo en $f_* \in \overline{\mathcal{L}(f)}$. O bien $v(f_*) = 0$ y f_* es una función constante, o bien $v(f_*) \neq 0$ y $\Gamma(f_*) > \gamma(f_*)$. Lo último no puede ser. En tal caso, $\Gamma(f_*) > \gamma(f_*)$, y entonces el conjunto

$$F = \left\{ x \in M : f_* > \frac{\Gamma(f_*) + \gamma(f_*)}{2} \right\}$$

es no vacío (por la propiedad de los valores intermedios) y abierto. Sea $g \in G$ arbitrario. El conjunto

$$gF = \{gx : x \in F\}$$

también es abierto, por lo cual la familia de conjuntos abiertos $\{gF\}_{g \in G}$ es una cubierta de M, dada la transitividad de la acción de G. De la compacidad de M se deduce que existe una familia finita $\{g_i F\}_{i=1}^n$ que cubre a M. Hagamos

$$\widetilde{f}_*(x) = \frac{1}{n} \sum_{i=1}^{n} L_{g_i} f_*(x).$$

Esta función pertenece a $\overline{\mathcal{L}(f)}$, y además $\Gamma\left(\widetilde{f}_*\right) \leq \Gamma(f_*)$. Sea $x \in M$; existe $1 \leq k \leq n$ tal que $x \in g_k F$. Por lo tanto, $g_k^{-1} x \in F$, y por la definición de F se tiene que

$$f_*\left(g_k^{-1} x\right) > \frac{\Gamma(f_*) + \gamma(f_*)}{2},$$

y de aquí que

$$\begin{aligned} \widetilde{f}_*(x) &= \frac{1}{n} f_*(g_k x) + \frac{1}{n} \sum_{i \neq k} L_{g_i} f_*(x) \\ &> \frac{\Gamma(f_*) + \gamma(f_*)}{2n} + \frac{1}{n} \sum_{i \neq k} \gamma(f_*) \\ &> \frac{1}{2n} \Gamma(f_*) + \left(1 - \frac{1}{2n}\right) \gamma(f_*) \\ &> \gamma(f_*) + \frac{1}{2n} v(f_*). \end{aligned}$$

Como esto es válido para todo $x \in M$, se sigue que $\gamma\left(\widetilde{f}_*\right) \geq \gamma(f_*) + \frac{1}{2n} v(f_*)$. Esto nos conduce a una contradicción, pues

$$\begin{aligned} v\left(\widetilde{f}_*\right) &= \Gamma\left(\widetilde{f}_*\right) - \gamma\left(\widetilde{f}_*\right) \\ &\leq \Gamma(f_*) - \gamma(f_*) - \frac{1}{2n} v(f_*) \\ &< \Gamma(f_*) - \gamma(f_*) = v(f_*), \end{aligned}$$

y así f_* no puede ser de variación mínima. Luego $v(f_*) = 0$ y f_* es constante. Reemplazando adecuadamente en el razonamiento anterior a $\mathcal{R}(f)$ y la acción derecha, obtenemos que existe una $f_* \in \overline{\mathcal{R}(f)}$ constante. Demostremos a continuación que $\mathcal{R}(f) \cap \mathcal{L}(f)$ contiene exactamente un elemento. Sean f_*, f'_* cualesquiera funciones constantes en $\mathcal{R}(f) \cap \mathcal{L}(f)$. Sea $\epsilon > 0$ y $u, w \in \mathcal{L}(f)$ de modo que

$$\begin{aligned} \|f_* - u\|_\infty &= \left\| f_* - \sum_{i=1}^{n} a_i L_{g_i} f \right\|_\infty \leq \frac{\epsilon}{2} \\ \|f'_* - w\|_\infty &= \left\| f'_* - \sum_{i=1}^{n} b_i L_{h_i} f \right\|_\infty \leq \frac{\epsilon}{2}. \end{aligned}$$

En particular, para cada $1 \leq j \leq n$,

$$\sup_{x \in M} \left| f_* - \sum_{i=1}^{n} a_i f(g_i h_j x) \right| \leq \frac{\epsilon}{2},$$

por lo cual

$$\begin{aligned} \left| f_* - \sum_{j=1}^{n} \sum_{i=1}^{n} a_i b_j f(g_i h_j x) \right| &= \left| \sum_{j=1}^{n} b_j f_* - \sum_{j=1}^{n} \sum_{i=1}^{n} a_i b_j f(g_i h_j x) \right| \\ &\leq \sum_{j=1}^{n} b_j \left| f_* - \sum_{i=1}^{n} a_i f(g_i h_j x) \right| \\ &\leq \sum_{j=1}^{n} b_j \frac{\epsilon}{2} = \frac{\epsilon}{2}. \end{aligned}$$

De forma completamente análoga para $\mathcal{R}(f)$ deducimos que

$$\left| f'_* - \sum_{j=1}^{n} \sum_{i=1}^{n} a_i b_j f(g_i h_j x) \right| \leq \frac{\epsilon}{2},$$

y por la desigualdad del triángulo se sigue que $|f_* - f'_*| \leq \epsilon$. Por lo tanto, $f_* = f'_*$. Sea $E(f) = f_*(x)$. Vemos que esta función satisface

$$\begin{aligned} |E(f)| &\leq \|f\|_\infty, \\ E(1) &= 1, \\ E(\lambda f) &= \lambda E(f), \\ E(L_g f) &= E(f), \\ E(R_g f) &= E(f). \end{aligned}$$

Resta probar que es lineal. Sean $f_1, f_2 \in C(M)$. Tomemos $u \in \mathcal{L}(f_1)$ de modo que

$$|E(f_1) - u| = \left| E(f_1) - \sum_{i=1}^{m} a_i L_{g_i} f_1(x) \right| \leq \frac{\epsilon}{2} \tag{2.1}$$

para todo $x \in M$. Denotemos por $\phi = \sum_{i=1}^{m} a_i L_{g_i} f_2$. Como $\phi \in \mathcal{L}(f_2)$, entonces $\mathcal{L}(\phi) \subset \mathcal{L}(f_2)$, así que $\overline{\mathcal{L}(\phi)} \subset \overline{\mathcal{L}(f_2)}$ y $E(\phi) = E(f_2)$. Escojamos $w \in \mathcal{L}(\phi)$ de modo que

$$\|E(f_2) - w\|_\infty = \left\| E(f_2) - \sum_{i=1}^{m} b_i L_{h_i} \phi \right\|_\infty \leq \frac{\epsilon}{2}.$$

Por la definición de ϕ,

$$\begin{aligned} &\left| E(f_2) - \sum_{j=1}^{m} \sum_{i=1}^{m} a_i b_i L_{g_i h_i} f_2(x) \right| \\ &\qquad = \left| E(f_2) - \sum_{j=1}^{m} b_j L_{h_j} \sum_{i=1}^{m} a_i L_{g_i} f_2(x) \right| \\ &\qquad\qquad = \left| E(f_2) - \sum_{j=1}^{m} b_j L_{h_j} \phi(x) \right| \leq \frac{\epsilon}{2}. \end{aligned}$$

Tomando $h_j x$ en lugar de x en (2.1), obtenemos que

$$\begin{aligned} &\left| E(f_1) - \sum_{j=1}^{m} \sum_{i=1}^{m} a_i b_i L_{g_i h_i} f_1(x) \right| \\ &\qquad = \left| \sum_{j=1}^{m} b_j E(f_1) - \sum_{j=1}^{m} b_j \sum_{i=1}^{m} a_i L_{g_i} f_1(h_j x) \right| \\ &\qquad \leq \sum_{j=1}^{m} b_j \left| E(f_1) - \sum_{i=1}^{m} a_i L_{g_i} f_1(h_j x) \right| \\ &\qquad\qquad \leq \sum_{j=1}^{m} b_j \frac{\epsilon}{2} = \frac{\epsilon}{2}. \end{aligned}$$

Por la desigualdad del triángulo

$$\left| E(f_1) - E(f_2) - \sum_{j=1}^{m} \sum_{i=1}^{m} a_i b_j (f_1 + f_2)(g_i h_j x) \right| \leq \epsilon.$$

y como $\sum_{j=1}^{m}\sum_{i=1}^{m} a_i b_j L_{g_i} L_{h_j}(f_1+f_2) \in \mathcal{L}(f_1+f_2)$, se sigue que $E(f_1)+E(f_2) \in \mathcal{L}(f_1+f_2)$. Por la unicidad de la función constante, podemos deducir que

$$E(f_1)+E(f_2)=E(f_1+f_2).$$

Este funcional E es único. Sea otro funcional $E' \in C(M)$ que satisfaga $E'(1)=1$ y $E'(f)=E'(L_g f)$ para todo $f \in C(M)$ y $g \in G$. Por la linealidad $E'(\phi)=E'(f)$ para todo $\phi \in \mathcal{L}(f)$. Por la continuidad, $E'(\phi)=E'(f)$ para todo $\phi \in \overline{\mathcal{L}(f)}$. En particular, como $\phi = E(f) \in \overline{\mathcal{L}(f)}$, se sigue que

$$E(f)=E'(E(f))=E'(f).$$

Usando este funcional y el Teorema A.2 obtenemos el resultado, pues M es compacto. □

La demostración anterior va por la vía larga. Podemos dar otra que es más corta y elegante, atribuída por Milman y Schechtman a W. Maak. El observador sagaz notará que ambas consisten en construir un funcional lineal que encaje en las hipótesis del Teorema A.2.

Otra demostración. Para cada $\epsilon > 0$ sea $N_\epsilon \subset M$ una ϵ-red minimal en M, es decir,

$$\bigcup_{t\in N_\epsilon} B(t,\epsilon)=M$$

y $n_\epsilon = |N_\epsilon|$ (la cardinalidad de N_ϵ) es mínima entre todos los conjuntos con dicha propiedad. Aquí,

$$B(t,\epsilon)=\{s\in M : \rho(t,s)\le \epsilon\}.$$

Definimos los funcionales

$$\overline{\Lambda}_\delta(f)=\frac{1}{n_\delta}\sum_{t\in N_\delta}\sup\{f(t): t\in B(t,\delta)\},$$

$$\underline{\Lambda}_\delta(f)=\frac{1}{n_\delta}\sum_{t\in N_\delta}\text{ínf}\{f(t): t\in B(t,\delta)\}.$$

Por construcción:

$$\underline{\Lambda}_\delta(f)\le \overline{\Lambda}_\delta(f), \quad |\overline{\Lambda}_\delta(f)-\underline{\Lambda}_\delta(f)|\le \epsilon. \tag{2.2}$$

Si $\epsilon_i \to 0$, existe una sucesión δ_i tal que para todo i se satisface (2.2). Dado que las sucesiones $\{\underline{\Lambda}_{\delta_i}(f)\}$ y $\{\overline{\Lambda}_{\delta_i}(f)\}$ son acotadas, existe una subsucesión i_k tal que $\{\overline{\Lambda}_{\delta_{i_k}}(f)\}$ converge a su límite inferior, y por (2.2) dicho límite es igual al límite superior de $\{\underline{\Lambda}_{\delta_i}(f)\}$.

Por lo tanto, dados los funcionales

$$\Lambda_\delta(f) = \frac{1}{n_\delta} \sum_{t \in N_\delta} f(t),$$

se tiene que $\Lambda_{\delta_{i_k}}(f)$ converge a un límite que denotaremos como $\Lambda(f)$. Queda así definido un funcional lineal positivo sobre $C(M)$ con $\Lambda(1) = 1$. Por el Teorema A.2, este funcional induce una única medida de probabilidad de Borel regular μ.

A continuación queremos mostrar que el funcional Λ está unívocamente determinado, esto es, si Λ'_ϵ está definido usando ϵ-redes minimales diferentes, entonces se tiene que

$$\Lambda'_\epsilon(f) \to \Lambda(f).$$

Si N'_ϵ es otra ϵ-red minimal en M, afirmamos que existe un mapeo biyectivo $\varphi : N_\epsilon \to N'_\epsilon$ con $\rho(t, \varphi(t)) \leq 2\epsilon$ para todo $t \in N_\epsilon$. Sea $K \subseteq N_\epsilon$ y para cada $t \in K$ definamos

$$S_t = \{s \in N'_\epsilon : B(t, \epsilon) \cap B(s, \epsilon) \neq \emptyset\}.$$

Haciendo

$$L = \bigcup_{t \in K} S_t = \left\{s \in N'_\epsilon : B\left(s, \epsilon\right) \cap \left(\cup_{t \in K} B(t, \epsilon)\right) \neq \emptyset\right\},$$

tenemos que $|L| \geq |K|$, pues de lo contrario $L \cup (N_\epsilon \setminus K)$ es una ϵ-red con menos de $|N_\epsilon|$ elementos. El Teorema D.1 nos dice que en tal situación existe un mapeo inyectivo

$$\varphi : N_\epsilon \to N'_\epsilon,$$

tal $\varphi(t) \in S_t$ para todo $t \in N_\epsilon$. Esto se traduce en que $B(t, \epsilon) \cap B(\varphi(t), \epsilon) \neq \emptyset$ o, lo que es lo mismo, $\rho(t, \varphi(t)) \leq 2\epsilon$, pues si s está en $B\left(t, \epsilon\right) \cap B(\varphi(t), \epsilon)$, entonces

$$\rho(t, \varphi(t)) \leq \rho(t, s) + \rho(s, \varphi(t)) \leq 2\epsilon.$$

Se sigue que, si Λ'_ϵ se define usando N'_ϵ en una manera análoga a la de Λ_ϵ, entonces

$$|\Lambda_\epsilon(f) - \Lambda'_\epsilon(f)| \leq \frac{1}{n_\epsilon} \sum_{t \in N_\epsilon} |f(t) - f\left(\varphi(t)\right)| \leq \omega\left(2\epsilon\right),$$

donde

$$\omega(\epsilon) = \sup_{\rho(s,t) \leq \epsilon} \{|f(t) - f\left(s\right)|\}$$

es el módulo de continuidad de f. Por lo tanto, $\lim_{\epsilon \to 0} \Lambda'_\epsilon(f)$ existe y es igual a $\Lambda(f)$. Resta probar que μ es invariante bajo G. Sea $g \in G$ y sea $N'_\epsilon = (gt)_{t \in N_\epsilon}$. Entonces N'_ϵ es una ϵ-red minimal (pues g actúa isométricamente sobre N_ϵ) y

$$\begin{aligned} \Lambda\left(fg\right) &= \lim_{\epsilon \to 0} \frac{1}{n_\epsilon} \sum_{t \in N_\epsilon} f(gt) = \lim_{\epsilon \to 0} \frac{1}{n_\epsilon} \sum_{gt \in N'_\epsilon} f(t) \\ &= \lim_{\epsilon \to 0} \Lambda'_\epsilon\left(f\right) = \Lambda(f). \end{aligned}$$

□

El siguiente teorema sería redundante, de no ser por la expresión que obtenemos al final que relaciona dos medidas regulares e invariantes sobre un mismo espacio.

Teorema 2.2. *Si (M, ρ) es un espacio métrico compacto y es espacio homogéneo del grupo G, entonces la medida del Teorema 2.1 es única, salvo por una constante.*

Demostración. Definimos una semimétrica en G por

$$d(g, h) = \sup_{t \in M} \rho(gt, ht). \tag{2.3}$$

En efecto, si $g = h$, entonces $\rho(gt, ht) = 0$ para todo $t \in M$, luego $d(g, g) = 0$. Que $d(g, h) = d(h, g)$ es claro a partir de las propiedades de ρ, y la desigualdad del triángulo es consecuencia de

$$\rho(gt, ht) \leq \rho(gt, kt) + \rho(kt, ht) \leq d(g, k) + d(k, h).$$

Identificando a los elementos cuya distancia entre sí es nula, tenemos un grupo H que todavía actúa isométricamente sobre M y sobre sí mismo, ya sea por la izquierda o por la derecha. En efecto, si tomamos $[u] = [v]$ y $[x] = [y]$, entonces $xt = yt$ para todo $t \in M$, y además

$$d(ux, vy) = \sup_{t \in M} \rho(uxt, vyt) = \sup_{s = xt \in M} \rho(us, vs) \leq \sup_{s \in M} \rho(us, vs) = 0,$$

lo que significa que $[ux] = [vy]$ y que la operación de grupo está bien definida.

Afirmamos que H es compacto. En efecto, considerando la cubierta

$$\mathcal{C} = \{B_d([g], r) : [g] \in H\}$$

definimos el conjunto de abiertos en M

$$\mathcal{D} = \left\{ \bigcup_{t \in M} B_\rho(gt, r) : B_d([g], r) \in \mathcal{C} \right\}.$$

Supongamos que $\mathcal{D}$ no es una cubierta de M y que por eso existe $s \in M$ tal que $\rho(s, gt) \geq r$ para todo $g \in G$ y todo $t \in M$. Consideremos a un t fijo. Por la transitividad de la acción de G en M existe $h \in G$ tal que $s = ht$, luego $\rho(ht, gt) \geq r$, lo que implica que $d(h, g) = \sup_{s \in M} \rho(hs, gs) \geq r$, y por lo tanto $[h]$ no está incluído en ninguno de los elementos de $\mathcal{C}$, lo cual es contradictorio. En consecuencia, $\mathcal{D}$ es una cubierta de M.

Dada la compacidad de M, existe una subcubierta finita

$$\mathcal{D}' = \left\{ \bigcup_{t \in M} B_\rho(g_1 t, r), \ldots, \bigcup_{t \in M} B_\rho(g_n t, r) \right\}$$

de $\mathcal{D}$. Supongamos que existe $[h] \in H$ tal que $d([h],[g_i]) \geq r$ para todo $i = 1,\ldots,n$. Entonces para todo $t \in M$ y $1 \leq i \leq n$ se cumple que $\rho(ht, g_i t) \geq r$. Pero esto contradice el hecho de que $\mathcal{D}'$ es una cubierta abierta de M. Por lo tanto $\mathcal{C}' = \{B_d([g_i], r)\}$ es una subcubierta finita de $\mathcal{C}$ y se se sigue que H es compacto.

Sean μ en M y ν en H medidas invariantes bajo la acción de G. Entonces, para todo $f \in C(M)$,

$$\nu(1)\mu(f) = \int_H \int_M f(gt)\, d\mu(t) d\nu(g) = \int_M \int_H f(gt)\, d\nu(g) d\mu(t).$$

Por la transitividad de H en M y la invariancia de ν, la integral interior de la derecha depende de f pero no de t. Llamémosla $\overline{\nu}(f)$. Entonces

$$\nu(1)\mu(f) = \overline{\nu}(f)\mu(1). \tag{2.4}$$

Ahora, si μ' es otra medida invariante en M entonces

$$\mu(f)\mu'(1) = \mu'(f)\mu(1). \qquad \square$$

Escolio 2.1. Si dotamos a $O(n)$ con la distancia (2.3), por el Teorema 2.2 se tiene que $M = (O(n), d)$ es un espacio métrico compacto. Sobre M actúa $O(n)$ isométrica y transitivamente. De acuerdo con el teorema de Haar, existe una única medida normalizada de probabilidad ν sobre $O\,(n)$ que es invariante bajo la multiplicación.

Sea $A \subseteq S^{n-1}$ un conjunto medible según la medida de Haar. Por la ecuación (2.4),

$$\mu(A) = \int_{S^{n-1}} \chi_A \, d\mu = \int_{O(n)} \chi_A(Ux)\, d\nu = \nu(\{U \in O(n) : Ux \in A\}).$$

Escolio 2.2. Introduciremos ahora otra medida de probabilidad sobre S^{n-1} como sigue. Sea $A \subseteq S^{n-1}$ y consideremos el conjunto

$$\overline{A} = \{\alpha x : x \in A, \alpha \in [0,1]\}.$$

que llamaremos cono levantado sobre A. Definamos

$$\nu A = \frac{\mu \overline{A}}{\mu B_n^2},$$

donde μ es la medida de Lebesgue. Evidentemente, siempre que A sea medible, $\overline{A}$ será también medible y νA estará bien definida. Adicionalmente, $\nu(A) = \nu(gA)$ para todo $g \in O(n)$. Por el Ejemplo 2.1 y el Teorema 2.1, tenemos que ν es de hecho la medida de Haar sobre S^{n-1} inducida por la acción de $O(n)$, pues

$$\nu S^{n-1} = \frac{\mu \overline{A}}{\mu B_n^2} = \frac{\mu B_n^2}{\mu B_n^2} = 1.$$

2.2. La desigualdad de Brunn-Minkowski

La desigualdad MA/MG tiene un paralelo muy interesante para los volúmenes en $\mathbb{R}^n$: la desigualdad de Brunn-Minkowski. Llegaremos a ella a través de la desigualdad de Prékopa-Leindler y en algunas partes utilizaremos, precisamente, la desigualdad MA/MG. Hagamos un poco de calentamiento con la versión unidimensional de la desigualdad de Brunn-Minkowski.

Lema 2.1. *Sean A, B dos conjuntos medibles no vacíos en $\mathbb{R}$, donde μ es la medida de Lebesgue. Entonces*

$$\mu\,(A+B) \geq \mu A + \mu B,$$

o, equivalentemente,

$$\mathrm{Vol}(A+B) \geq \mathrm{Vol}A + \mathrm{Vol}B.$$

Demostración. Si $\mu A = \infty$, en virtud de que $A + b \subseteq A + B$ para $b \in B$, se tiene que $\mu(A+B) = \infty$ y la desigualdad es trivial. Supongamos en consecuencia que $\mu A < \infty$ y $\mu B < \infty$. Sea $\epsilon > 0$. Como μ es una medida regular, entonces existen conjuntos A' y B' compactos tales que $\mu(A') + \epsilon/2 \leq \mu(A)$ y $\mu(B') + \epsilon/2 \leq \mu(B)$. Además, μ es invariante bajo traslaciones, así que es posible considerar que $A' \cap (-\infty, \epsilon] = A'$ y $B' \cap [-\epsilon, \infty) = B'$. Más aún, podemos suponer que $0 \in A \cap B$. Entonces, por ser $A' = A' + 0$ y $B' = B' + 0$,

$$A' \cup B' \subset A' + B' \subseteq A + B.$$

También $A' \cap B' \subseteq [-\epsilon, \epsilon]$, así que $\mu\,(A' \cap B') \leq 2\epsilon$. Por lo tanto:

$$\begin{aligned}
\mu(A+B) \geq \mu\,(A'+B') &\geq \mu\,(A' \cup B') \\
&= \mu(A') + \mu(B') - \mu(A' \cap B') \\
&= \mu A + \mu B - \mu(A' \cap B') - \epsilon \\
&\geq \mu A + \mu B - 3\epsilon,
\end{aligned}$$

y haciendo tender ϵ a cero obtenemos el resultado. □

Teorema 2.3 (Prékopa-Leindler)**.** *Sean $f, g, m : \mathbb{R}^n \to [0, \infty)$ tres funciones integrables no negativas que satisfacen, para alguna $0 < \lambda < 1$ y toda $x, y \in \mathbb{R}^n$ la desigualdad*

$$m\,(\lambda x + (1-\lambda)\,y) \geq f\,(x)^{\lambda} \cdot g\,(y)^{1-\lambda}. \tag{2.5}$$

Entonces

$$\int_{\mathbb{R}^n} m \geq \left(\int_{\mathbb{R}^n} f\right)^{\lambda} \left(\int_{\mathbb{R}^n} g\right)^{1-\lambda}.$$

Demostración. Probemos el resultado para $n = 1$. Sean $f, g, m : \mathbb{R} \to [0, \infty)$ que satisfacen (2.5). Supongamos inclusive que $f(x), g(x) > 0$ para todo $x \in \mathbb{R}$. Por el teorema de Fubini,

$$\begin{aligned}
\int_{\mathbb{R}} f &= \int_{\mathbb{R}^2} \chi_{\{(x,t):0\leq t\leq f(x)\}}\, d\mu \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \chi_{\{(x,t):0\leq t\leq f(x)\}}\, dx \right) dt \\
&= \int_0^{\infty} \mu\left(\{x : f(x) \geq t\}\right) dt. \qquad (2.6)
\end{aligned}$$

Los conjuntos

$$A = \{x \in \mathbb{R} : |f(x)| = \infty\} \quad \text{y} \quad B = \{x \in \mathbb{R} : |g(x)| = \infty\}$$

son de medida cero al ser f y g integrables, así que podemos redefinir f y g en A y B (respectivamente) de modo que valgan 1, sin alterar el valor las integrales. Entonces f y g están acotadas y por lo tanto podemos normalizarlas de modo que $\sup f(x) = \sup g(x) = 1$. De este modo los conjuntos

$$\{x : f(x) \geq t\} \quad \text{y} \quad \{y : g(y) \geq t\}$$

son no vacíos para cada $0 \leq t < 1$. Sean $x, y \in \mathbb{R}$ y $t \in [0, 1]$ tales que $f(x) \geq t$ y $g(y) \geq t$. Entonces se cumple que

$$m(\lambda x + (1-\lambda)\, y) \geq t.$$

De aquí que, por la desigualdad MA/MG,

$$\{z : m(z) \geq t\} \supseteq \lambda \{x : f(x) \geq t\} + (1-\lambda)\{y : g(y) \geq t\}.$$

Aplicando el Lema 2.1 obtenemos

$$\begin{aligned}
&\mu\left(\{z : m(z) \geq t\}\right) \\
&\qquad \geq \mu\left(\lambda \{x : f(x) \geq t\} + (1-\lambda)\{y : g(y) \geq t\}\right) \\
&\qquad\qquad \geq \lambda\mu\left(\{x : f(x) \geq t\}\right) + (1-\lambda)\,\mu\left(\{y : g(y) \geq t\}\right).
\end{aligned}$$

De aquí se desprende que la desigualdad es válida para $n = 1$, utilizando la expresión (2.6) y la desigualdad MA/MG.

Procedemos ahora por inducción. Supongamos que el resultado es válido para n. Consideremos las tres funciones integrables reales no negativas $f, g, m : \mathbb{R}^{n+1} \to [0, \infty)$ que satisfacen (2.5) para alguna $0 < \lambda < 1$. Fijemos $x_0, y_0 \in \mathbb{R}$ y hagamos

$$z_0 = \lambda x_0 + (1-\lambda)\, y_0. \qquad (2.7)$$

Defínanse nuevas funciones $f_{x_0}, g_{y_0}, m_{z_0} : \mathbb{R}^n \to [0,\infty)$ fijando la primera coordenada de las originales:

$$\begin{aligned} f_{x_0}(x) &= f(x_0, x), \\ g_{y_0}(y) &= g(y_0, y), \\ m_{z_0}(z) &= m(z_0, z). \end{aligned}$$

Como las nuevas funciones satisfacen las hipótesis del teorema, por la hipótesis de inducción se tiene que

$$\int_{\mathbb{R}^n} m_{z_0} \geq \left(\int_{\mathbb{R}^n} f_{x_0}\right)^{\lambda} \left(\int_{\mathbb{R}^n} g_{y_0}\right)^{1-\lambda}.$$

Nótese que la ecuación anterior se verifica para cada x_0, y_0, z_0 que satisfacen la relación (2.7). Definimos ahora las funciones

$$\begin{aligned} \widetilde{f}(u) &= \int_{\mathbb{R}^n} f_u(x)\,dx, \\ \widetilde{g}(u) &= \int_{\mathbb{R}^n} g_u(x)\,dx, \\ \widetilde{m}(u) &= \int_{\mathbb{R}^n} m_u(x)\,dx. \end{aligned}$$

Ya demostramos que cada una de estas funciones satisface la desigualdad del teorema para el caso unidimensional. Aplicando nuevamente el Lema 2.1, tenemos

$$\begin{aligned} \int_{\mathbb{R}^{n+1}} m &= \int_{\mathbb{R}} \widetilde{m} \\ &\geq \left(\int_{\mathbb{R}} \widetilde{f}\right)^{\lambda} \left(\int_{\mathbb{R}} \widetilde{g}\right)^{1-\lambda} = \left(\int_{\mathbb{R}^{n+1}} f\right)^{\lambda} \left(\int_{\mathbb{R}^{n+1}} g\right)^{1-\lambda}, \end{aligned}$$

lo que completa la inducción. □

Como consecuencia inmediata obtenemos el siguiente resultado.

Teorema 2.4 (Desigualdad de Brunn-Minkowski). *Sean A, B conjuntos medibles en $\mathbb{R}^n$. Entonces*

$$\mathrm{Vol}\,(A+B)^{\frac{1}{n}} \geq \mathrm{Vol}\,(A)^{\frac{1}{n}} + \mathrm{Vol}\,(B)^{\frac{1}{n}}$$

o, equivalentemente, para toda $\lambda \in (0,1)$

$$\mathrm{Vol}\,(\lambda A + (1-\lambda)\,B) \geq \mathrm{Vol}\,(A)^{\lambda}\,\mathrm{Vol}\,(B)^{1-\lambda}.$$

Demostración. Sean m, f y g las funciones características de $\lambda A + (1-\lambda) B$, A y B, respectivamente. Si x está en A y y está en B, entonces $\lambda x + (1-\lambda) y$ está en $\lambda A + (1-\lambda) B$ pero no necesariamente en A o B. De aquí que m, f y g satisfagan las condiciones del Teorema 2.3, y puesto que

$$\begin{aligned} \mathrm{Vol}\left(\lambda A + (1-\lambda) B\right) &= \int_{\mathbb{R}^n} m \\ &\geq \left(\int_{\mathbb{R}^n} f\right)^{\lambda} \left(\int_{\mathbb{R}^n} g\right)^{1-\lambda} = \mathrm{Vol}\left(A\right)^{\lambda} \mathrm{Vol}\left(B\right)^{1-\lambda} \end{aligned}$$

se sigue el resultado deseado. Sólo falta probar la equivalencia de las desigualdades propuestas. La primera implica a la segunda, pues

$$\begin{aligned} \mathrm{Vol}\left(\lambda A + (1-\lambda) B\right)^{\frac{1}{n}} &\geq \mathrm{Vol}\left(\lambda A\right)^{\frac{1}{n}} + \mathrm{Vol}\left((1-\lambda) B\right)^{\frac{1}{n}} \\ &\geq \lambda \mathrm{Vol}\left(A\right)^{\frac{1}{n}} + (1-\lambda) \mathrm{Vol}\left(B\right)^{\frac{1}{n}} \\ &\geq \mathrm{Vol}\left(A\right)^{\frac{\lambda}{n}} \mathrm{Vol}\left(B\right)^{\frac{1-\lambda}{n}}, \end{aligned}$$

esto último por la desigualdad MA/MG.

Finalmente, veamos que la segunda implica la primera. Si cualesquiera de los dos volúmenes de la desigualdad es nulo, no hay nada que demostrar. Por lo tanto, supongamos que $\mathrm{Vol}\left(A\right)$ y $\mathrm{Vol}\left(B\right)$ son positivos. Haciendo

$$\lambda = \frac{\mathrm{Vol}\left(A\right)^{\frac{1}{n}}}{\mathrm{Vol}\left(A\right)^{\frac{1}{n}} + \mathrm{Vol}\left(B\right)^{\frac{1}{n}}},$$

tenemos que

$$\begin{aligned} \frac{\mathrm{Vol}\left(A+B\right)}{\left(\mathrm{Vol}\left(A\right)^{\frac{1}{n}} + \mathrm{Vol}\left(B\right)^{\frac{1}{n}}\right)^{n}} &= \mathrm{Vol}\left(\frac{A+B}{\mathrm{Vol}\left(A\right)^{\frac{1}{n}} + \mathrm{Vol}\left(B\right)^{\frac{1}{n}}}\right) \\ &= \mathrm{Vol}\left(\lambda \frac{A}{\mathrm{Vol}\left(A\right)^{\frac{1}{n}}} + (1-\lambda) \frac{B}{\mathrm{Vol}\left(A\right)^{\frac{1}{n}}}\right) \\ &\geq \mathrm{Vol}\left(\frac{A}{\mathrm{Vol}\left(A\right)^{\frac{1}{n}}}\right)^{\lambda} + \mathrm{Vol}\left(\frac{B}{\mathrm{Vol}\left(A\right)^{\frac{1}{n}}}\right)^{(1-\lambda)} \\ &\geq 1, \end{aligned}$$

y resulta la equivalencia de las desigualdades. □

Definición 2.3. Sea $A \subseteq S^{n-1}$. Definimos el ϵ-engrosamiento de A como

$$A_{\epsilon} = \left\{x \in S^{n-1} : d\left(x, A\right) \leq \epsilon\right\},$$

donde $d\left(x,y\right) = \left\|x-y\right\|_2$ es la métrica euclidiana y $d(x,A) = \inf\{d(x,y) : y \in A\}$.

Obsérvese que el complemento del ϵ-engrosamiento de A es

$$A_\epsilon^c = \left\{x \in S^{n-1} : d\left(x, A\right) > \epsilon\right\}.$$

Antes de pasar al resultado siguiente, haremos una observación útil sobre los conjuntos convexos. Sean $\alpha, \beta \in \mathbb{R}$ positivos y A un conjunto convexo. Por un lado, es evidente que

$$(\alpha+\beta)A \subseteq \alpha A + \beta A.$$

Por otro lado, si $u \in \alpha A + \beta A$ entonces

$$u = \alpha a_1 + \beta a_2, \quad a_1, a_2 \in A,$$

por lo cual

$$u = (\alpha+\beta)\left(\frac{\alpha}{\alpha+\beta} a_1 + \frac{\beta}{\alpha+\beta} a_2\right).$$

Dada la convexidad de A, se hace manifiesto que $u \in (\alpha+\beta)A$ y que $\alpha A + \beta A \subseteq (\alpha+\beta)A$. Así,

$$(\alpha+\beta)A = \alpha A + \beta A.$$

Disponiendo de todo lo anterior, podemos ya obtener el resultado del que derivarán los principales hechos de concentración que emplearemos.

Teorema 2.5 (Gromov y Milman)**.** *Existen constantes* $0 < c, C < \infty$ *tales que para cualquier* n *entero positivo,* $A \subseteq S^{n-1}$ *medible y* μ *la medida de Haar en* S^{n-1} *se cumple*

$$\mu\left(A_\epsilon^c\right) \leq \frac{1}{c \cdot \mu\left(A\right)} \exp\left(-C\epsilon^2 n\right).$$

Demostración (Ball, Arias y Villa). Sean A_1 y A_2 un subconjunto de B_2^n, la bola euclidiana unitaria n- dimensional, tales que

$$d\left(A_1, A_2\right) = \inf_{x \in A_1, y \in A_2} d\left(x, y\right) \geq \epsilon.$$

Para cada $x \in A_1$ y $y \in A_2$ tenemos:

$$\left\|\frac{x+y}{2}\right\|_2^2 + \frac{\epsilon^2}{4} \leq \left\|\frac{x+y}{2}\right\|_2^2 + \left\|\frac{x-y}{2}\right\|_2^2 = \frac{\|x\|_2^2 + \|y\|_2^2}{2} \leq 1,$$

donde la igualdad es consecuencia del Lema 1.1. Por lo tanto,

$$\left\|\frac{x+y}{2}\right\|_2 \leq \left(1 - \frac{\epsilon^2}{4}\right)^{\frac{1}{2}},$$

es decir,

$$\frac{A_1 + A_2}{2} \subseteq (1-\delta)B_n^2$$

donde $\delta = 1-(1-\frac{\epsilon^2}{4})^{\frac{1}{2}} \geq \frac{\epsilon^2}{8}$. Tomemos $\overline{A}_1$ y $\overline{A}_2$, los conos levantados sobre A_1 y A_2, respectivamente. Afirmamos que si $\overline{x} \in \overline{A_1}$ y $\overline{y} \in \overline{A_2}$ entonces $(\overline{x}+\overline{y})/2 \in (1-\delta)B_n^2$. En efecto, $\overline{x} = \alpha x$ y $\overline{y} = \beta y$ para algunos $x \in A$, $y \in B$ y $0 \leq \alpha, \beta \leq 1$, y sin pérdida alguna de generalidad, podemos suponer que $\alpha \geq \beta > 0$ y $\beta/\alpha = \gamma \leq 1$. Entonces

$$\begin{aligned}\frac{\overline{x}+\overline{y}}{2} = \frac{\alpha x + \beta y}{2} &= \alpha\left(\frac{x+\gamma y}{2}\right) \\ &= \alpha\left(\gamma\frac{x+y}{2} + (1-\gamma)\frac{x}{2}\right) = \alpha\gamma\left(\frac{x+y}{2}\right) + \alpha(1-\gamma)\frac{x}{2}\end{aligned}$$

y por lo tanto

$$\frac{\overline{x}+\overline{y}}{2} \in \alpha\gamma(1-\delta)B_n^2 + \alpha(1-\gamma)(1-\delta)B_n^2 = \alpha(1-\delta)B_n^2 \subset (1-\delta)B_n^2$$

según las observaciones precedentes. De aquí que

$$\tfrac{1}{2}\overline{A_1} + \tfrac{1}{2}\overline{A_2} \subset (1-\delta)B_n^2$$

y por consiguiente

$$\mathrm{Vol}\left(\tfrac{1}{2}\overline{A}_1 + \tfrac{1}{2}\overline{A}_2\right) \leq (1-\delta)^n \mathrm{Vol}(B_n^2).$$

La aplicación de la desigualdad de Brunn-Minkowksi nos revela que

$$\mathrm{Vol}\left(\tfrac{1}{2}\overline{A}_1 + \tfrac{1}{2}\overline{A}_2\right) \geq \mathrm{Vol}^{\frac{1}{2}}(\overline{A}_1)\mathrm{Vol}^{\frac{1}{2}}(\overline{A}_2),$$

y así

$$\mu A_2 = \frac{\mathrm{Vol}\overline{A}_2}{\mathrm{Vol}(B_n^2)} \leq (1-\delta)^{2n}\frac{\mathrm{Vol}(B_n^2)}{\mathrm{Vol}\overline{A}_1} = \frac{(1-\delta)^{2n}}{\mu A_1} \leq \frac{e^{-2n\delta}}{\mu A_1} \leq \frac{e^{-n\epsilon^2/4}}{\mu A_1},$$

así que tomando $A_1 = A$, $A_2 = A_\epsilon^c$, $C = \frac{1}{4}$ y $c = 1$ tenemos el resultado. □

2.3. La desigualdad de Lévy

La siguiente desigualdad es precisamente la que nos da cuenta del fenómeno de concentración de la medida alrededor de una mediana. La noción de mediana puede consultarse en el Apéndice B.

Teorema 2.6 (Lévy). *Sea $f : S^{n-1} \to \mathbb{R}$ una función no negativa de Lipschitz con constante L, es decir, para toda x y y en S^{n-1} se cumple*

$$|f(x) - f(y)| \leq L \left\| x - y \right\|_2 .$$

Existen constantes $0 < c, C < \infty$ tales que para toda $f : S^{n-1} \to \mathbb{R}$ de Lipschitz y todo $t > 0$ se cumple

$$\mu \left(\{ x : |f(x) - M| > t \} \right) \leq c \exp \left(-\frac{Ct^2 n}{L^2} \right),$$

donde M es una mediana de f.

Demostración. Supongamos que $L = 1$. Denotemos $A = \{x : f(x) \leq M\}$. Sea x tal que $d(x, A) \leq t$. Entonces, para todo ϵ existe $y \in A$ tal que $d(x, y) \leq d(x, A) + \epsilon \leq t + \epsilon$. Como f es de Lipschitz,

$$f(x) - f(y) \leq t + \epsilon,$$

entonces

$$f(x) \leq t + \epsilon + f(y) \leq t + \epsilon + M$$

es decir,

$$f(x) \leq t + M.$$

Extraemos como conclusión que

$$A_t \subseteq \{x : f(x) - M \leq t\}$$

o, lo que es lo mismo

$$A_t^c \supseteq \{x : f(x) - M > t\} .$$

Por lo tanto, de acuerdo con el Teorema 2.5, tenemos

$$\begin{aligned} \mu \left(\{ x : f(x) - M > t \} \right) &\leq \mu \left(A_t^c \right) \\ &\leq \frac{1}{c' \cdot \mu(A)} \exp \left(-\frac{t^2 n}{16} \right) \\ &\leq \frac{2}{c'} \exp \left(-\frac{t^2 n}{4} \right). \end{aligned}$$

Si x es tal que $f(x) < M - t < M$, entonces $x \in A \subset A_t$. Por consiguiente

$$\begin{aligned} \mu \left(\{ x : f(x) - M < -t \} \right) &\leq \mu(A) \\ &\leq \frac{1}{c' \cdot \mu(A_t^c)} \exp \left(-\frac{t^2 n}{16} \right) \\ &\leq \frac{2}{c'} \exp \left(-\frac{t^2 n}{4} \right). \end{aligned}$$

Tomando $c = \frac{2}{c'}$ obtenemos

$$\mu\left\{|f(x) - M| > t\right\} \leq c \exp\left(-\frac{t^2 n}{4}\right).$$

En caso de que $L \neq 1$, tenemos

$$\mu\left\{\left|\frac{f(x) - M}{L}\right| > \frac{t}{L}\right\} \leq c \exp\left(-\frac{t^2 n}{4L^2}\right). \qquad \square$$

Para poder extender a la esperanza respecto a la medida μ el teorema anterior, requerimos de una desigualdad que implica (muy elegantemente) a la MA/MG.

Lema 2.2 (Jensen)**.** *Sea μ una medida positiva sobre una σ-álgebra $\mathcal{M}$ en un conjunto Ω, tal que $\mu(\Omega) = 1$. Si f es una función real en $L^1(\mu)$, $a < f(x) < b$ para todo $x \in \Omega$ y ϕ es convexa en (a, b), entonces*

$$\phi\left(\int_\Omega f \, d\mu\right) \leq \int_\Omega (\phi \circ f) \, d\mu.$$

Demostración. Hagamos $t = \int_\Omega f \, d\mu$, lo que asegura que $a < t < b$. Si β es el supremo de los cocientes del primer miembro de la desigualdad

$$\frac{\phi(t) - \phi(s)}{t - s} \leq \frac{\phi(u) - \phi(t)}{u - t}$$

con $a < s < t$, entonces β no supera a ninguno de los cocientes del segundo miembro de la desigualdad para cualquier $t \in (t, b)$. Deducimos que

$$\phi(s) \geq \phi(t) + \beta \cdot (s - t), \quad (a < s < b),$$

y, por lo tanto,

$$\phi(f(x)) - \phi(t) - \beta(f(x) - t) \geq 0 \tag{2.8}$$

para todo $x \in \Omega$. Como ϕ es continua, $\phi \circ f$ es medible. Si integramos ambos miembros de (2.8) respecto a μ, tenemos la conclusión del teorema, en virtud de la elección de t y la hipótesis $\mu(\Omega) = 1$. $\square$

Escolio 2.3. Si a $\Omega = \{t_i\}_{i=1}^n$ (n entero positivo) lo dotamos con la medida

$$\mu(A) = \frac{|A|}{n}, \quad A \subseteq \Omega$$

y siempre que

$$\phi(\alpha t) \geq |\alpha| \phi(t),$$

podemos concluir que

$$\sum_{i=1}^n \phi(f(t_i)) \geq \phi\left(\sum_{i=1}^n f(t_i)\right).$$

Teorema 2.7 (Lévy)**.** *Existen constantes* $0 < c,\, C < \infty$ *tales que para todo* $t > 0$ *y toda función de Lipschitz* $f : S^{n-1} \to R$ *no negativa con constante* L *se cumple que*

$$\mu\left(\{x : |f(x) - E(f)| > t\}\right) \leq c \exp\left(-\frac{Ct^2 n}{L^2}\right).$$

Demostración. Como en el teorema anterior, podemos suponer que $L = 1$. Denotemos por $\mu \times \mu$ la medida producto en $S^{n-1} \times S^{n-1}$. Sea

$$A = \{(x, y) \in S^{n-1} \times S^{n-1} : |f(x) - M| + |f(y) - M| > t\}$$

y

$$B = \left\{x \in S^{n-1} : |f(x) - M| > \frac{t}{2}\right\}.$$

Se satisface

$$A \subseteq (B \times S^{n-1}) \cup (S^{n-1} \times B)$$

y podemos deducir que

$$\begin{aligned}
\mu \times \mu & \left\{(x, y) \in S^{n-1} \times S^{n-1} : |f(x) - f(y)| > t\right\} \\
& \leq \mu \times \mu(A) \leq \mu \times \mu(B \times S^{n-1}) + \mu \times \mu(S^{n-1} \times B) \\
& = 2\mu \left\{x \in S^{n-1} : |f(x) - M| > \frac{t}{2}\right\} \leq 2c' \exp\left(-\frac{t^2 n}{16}\right),
\end{aligned}$$

donde M es una mediana de f. Con el teorema de Fubini estimamos la esperanza de $\exp(\lambda^2 |f(x) - E(f)|)$:

$$\begin{aligned}
E_x E_y \exp\left(\lambda^2 |f(x) - f(y)|^2\right) &= \int_0^\infty 2\lambda^2 t \exp\left(\lambda^2 t^2\right)\, \mu \times \mu\{|f(x) - f(y)| > t\}\, dt \\
&\leq 4c'\lambda^2 \int_0^\infty t \exp\left[t^2\left(\lambda^2 - \frac{n}{16}\right)\right] dt.
\end{aligned}$$

Eligiendo $\lambda = \sqrt{\frac{n}{32}}$ y observando que

$$\int_0^\infty t \exp\left(-\lambda t^2\right) dt = \frac{1}{2\lambda} \int_0^\infty \exp(-u)\, du = \frac{1}{2\lambda},$$

obtenemos

$$E_x E_y \exp\left(\lambda^2 |f(x) - f(y)|^2\right) \leq 2c'.$$

Como $\phi(t) = \exp(t^2)$ es una función convexa, la desigualdad de Jensen nos indica que

$$\begin{aligned}
2c' &\geq E_x E_y \exp\left(\lambda^2 |f(x) - f(y)|^2\right) \\
&\geq E_x \exp\left(\lambda^2 |f(x) - E(f)|^2\right) = E\left[\exp\left(\lambda^2 |f(x) - E(f)|^2\right)\right].
\end{aligned}$$

Combinando los resultados anteriores con la desigualdad de Chebyshev obtenemos

$$\begin{aligned}
\mu\left\{x : |f(x) - E(f)| > t\right\} &= \mu\left\{x : e^{\frac{n}{32}|f(x)-E(f)|^2} > e^{\frac{n}{32}t^2}\right\} \\
&\leq \frac{E\exp\left(\frac{n}{32}|f(x) - E(f)|^2\right)}{\exp\left(\frac{n}{32}t^2\right)} \\
&\leq 2c'\exp\left(-\frac{n}{32}t^2\right),
\end{aligned}$$

que es lo que queríamos probar. □

2.4. Almártagas

Por último, discutiremos el fenómeno de concentración para almártagas. Aplicaremos esto al análisis de los algoritmos por pase de mensajes, empleados con cierta clase de códigos lineales.

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, $\mathcal{G}$ una subσ-álgebra de $\mathcal{F}$ y $f \in L^1(\Omega, \mathcal{F}, P)$. Entonces

$$\mu(A) = \int_A f\,dP, \quad A \in \mathcal{G}$$

define una medida sobre $\mathcal{G}$ que es absolutamente continua con respecto a $P|_{\mathcal{G}}$. Según el Teorema A.3, existe una única $h \in L^1(\Omega, \mathcal{F}, P)$ tal que

$$\int_A h\,dP = \int_A f\,dP$$

para toda $A \in \mathcal{G}$. Llamamos a esta h la esperanza condicional de f respecto a $\mathcal{G}$ y lo denotaremos como $h = E(f|\mathcal{G})$.

Definición 2.4. Dada una sucesión

$$\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \ldots \subseteq \mathcal{F}$$

de σ-álgebras, una sucesión $f_1, f_2, \ldots$ de funciones $f_i \in L^1(\Omega, \mathcal{F}_i, P)$ es una almártaga respecto a $\{\mathcal{F}_i\}_{i=1}^{\infty}$ si

$$E(f_i|\mathcal{F}_{i-1}) = f_{i-1}, \quad i = 2, 3, \ldots.$$

Demostraremos ahora una de las desigualdades básicas de la teoría de almártagas.

Teorema 2.8 (Azuma y Hoeffding). *Sea $X_0, X_1, \ldots,$ una almártaga respecto a $\{\mathcal{F}_i\}_0^\infty$ tal que para cada $i \geq 1$*

$$|X_i - X_{i-1}| \leq \gamma_i,$$

donde γ_i puede depender de i. Entonces, para $n \geq 1$ y para $\alpha > 0$,

$$P(|X_n - X_0| \geq \alpha\sqrt{n}) \leq 2\exp\left(-\frac{\alpha^2 n}{2\sum_{i=1}^n \gamma_i^2}\right).$$

Más aún, si $E(X) = X_n$ para algún $n \geq 1$,

$$P(|E(X) - X_0| \geq \alpha\sqrt{n}) \leq 2\exp\left(-\frac{\alpha^2 n}{2\sum_{i=1}^n \gamma_i^2}\right).$$

Demostración. Normalizando si es necesario, podemos suponer que $\gamma_i \leq 1$. Consideremos la función

$$f(x) := \exp(\gamma x)$$

donde $\gamma \geq 0$ y la línea $c(x)$ intersecta a $f(x)$ en los dos puntos $(-1, e^{-\gamma})$ y $(1, e^{\gamma})$. Como $f(x)$ es convexa, tenemos

$$f(x) \leq c(x), \quad x \in [-1, 1].$$

También necesitamos la desigualdad

$$\cosh(x) \leq e^{x^2/2}, \quad x \in \mathbb{R},$$

que se deduce directamente de sus expansiones de Taylor. Sea $Y_i = X_i - X_{i-1}$ para $i \geq 1$. Observamos que la condición $|X_i - X_{i-1}|$ se transforma en $|Y_i| \leq 1$. Más aún

$$E(Y_i|\mathcal{F}_{i-1}) = E(X_i - X_{i-1}|\mathcal{F}_{i-1}) = E(X_i|\mathcal{F}_{i-1}) - X_{i-1} = 0,$$

y así

$$\begin{aligned} E(f(Y_i)|\mathcal{F}_{i-1}) &\leq E(c(Y_i)|\mathcal{F}_{i-1}) \\ &= c(E(Y_i|\mathcal{F}_{i-1})) = c(0) = \cosh(\gamma) \leq \exp\left(\frac{\gamma^2}{2}\right). \end{aligned}$$

Notemos que $X_n - X_0 = \sum_{i=1}^n Y_i$ y claramente todas esas variables son medibles respecto a $\mathcal{F}_0$. En consecuencia,

$$\begin{aligned}
E(f(X_n - X_0)) &= E\left(\exp\left(\gamma \sum_{i=1}^{n} Y_i\right)\right) = E\left(\prod_{i=1}^{n} f(Y_i)\right) \\
&= E\left(\prod_{i=1}^{n} f(Y_i) \middle| \mathcal{F}_0\right) = E\left(E\left(\prod_{i=1}^{n} f(Y_i) \middle| \mathcal{F}_{n-1}\right)\right) \\
&= E\left(\left(\prod_{i=1}^{n-1} f(Y_i)\right) E(f(Y_n)|\mathcal{F}_{n-1})\right) \\
&\leq E\left(\prod_{i=1}^{n-1} f(Y_i)\right) \exp\left(\frac{\gamma^2}{2}\right) \leq \exp\left(\frac{n\gamma^2}{2}\right).
\end{aligned}$$

Se sigue de la desigualdad de Chebyshev que

$$\begin{aligned}
P(X_n - X_0 \geq \alpha\sqrt{n}) &= P(f(X_n - X_0) \geq f(\alpha\sqrt{n})) \\
&\leq \frac{E(f(X_n - X_0))}{f(\alpha\sqrt{n})} \\
&\leq \exp\left(\frac{\gamma^2 n}{2}\right) \exp\left(-\alpha^2\right) = \exp\left(-\frac{\alpha^2}{2}\right),
\end{aligned}$$

donde elegimos $\gamma = \alpha/\sqrt{n}$. De manera totalmente análoga probamos que

$$P(X_n - X_0 \leq -\alpha\sqrt{n}) \leq \exp(-\alpha^2/2),$$

y juntando todo tenemos el resultado. □

Como fue observado por G. Pisier en [Pis83], sin mucha dificultad y con mucho ingenio podemos extender el Teorema 2.8 al resultado que sigue.

Teorema 2.9 (Pisier)**.** *Bajo las hipótesis del Teorema 2.8, y definiendo*

$$\|\{a_k\}_{k=1}^{n}\|_{p,\infty} = \max_{1\leq i\leq n} a_i^*$$

donde $\{a_i^*\}_{i=1}^{k}$ *es un reacomodo decreciente de* $\{\|a_i\|\}_{i=1}^{k}$*, tenemos*

$$P(|X_n - X_0| \geq \alpha\sqrt{n}) \leq 2\exp\left(-\exp\delta\frac{\alpha\sqrt{n}}{\|\gamma_i\|_{1,\infty}}\right).$$

donde δ *es una constante absoluta.*

Demostración. Normalizando si fuera necesario, podemos suponer que

$$\|\{\|\gamma_k\|_\infty\}_{k=1}^n\|_{1,\infty} = 1.$$

Ahora escojamos una permutación π de $\{1, \dots, n\}$ tal que

$$\|\gamma_{\pi(k)}\|_\infty = \|\gamma_k\|^*, \quad k = 1, 2, \dots, n.$$

Entonces, para $n \geq k \geq 1$

$$\|\gamma_{\pi(k)}\|_\infty \leq \frac{1}{k}.$$

Dado un entero $N \leq n$, tenemos:

$$P\left(\left|\sum_1^n d_k\right| \geq 2\ln N + 1\right) \leq P\left(\left|\sum_1^N d_k\right| \geq 2\ln N\right) + P\left(\left|\sum_{N+1}^n d_k\right| \geq 1\right), \quad (2.9)$$

pero

$$\left|\sum_1^N \gamma_{\pi(k)}\right| \leq \sum_1^N \|\gamma_{\pi(k)}\|_\infty \leq \sum_1^N \frac{1}{k} \leq 2\ln N,$$

y de aquí que el primer término del lado derecho de la desigualdad (2.9) es 0. Ahora, por el Teorema 2.8,

$$\begin{aligned} P\left(\left|\sum_1^n \gamma_k\right| \geq 2\ln N + 1\right) &\leq 2\exp\left(-\frac{1}{4\sum_{k=N+1}^n \|\gamma_{\pi(k)}\|_\infty^2}\right) \\ &\leq 2\exp\left(-\frac{N}{4}\right). \end{aligned}$$

Si $t \geq 1$, hagamos

$$N = \left\lfloor \exp\left(\frac{t-1}{2}\right)\right\rfloor,$$

de modo que $1 \leq N \leq \exp\left(\frac{t-1}{2}\right) \leq 2N$. En consecuencia,

$$\begin{aligned} P(|X_n - X_0| \geq t) &= P\left(\left|\sum_{k=1}^n \gamma_k\right| \geq t\right) \\ &\leq P\left(\left|\sum_{k=1}^n \gamma_k\right| \geq 2\ln N + 1\right) \\ &\leq 2\exp(-N/4) \\ &\leq 2\exp\left(-\frac{\exp\left(\frac{t-1}{2}\right)}{8}\right). \end{aligned}$$

Por otro lado, si $t < 1$,

$$2\exp\left(-\frac{\exp\left(\frac{t-1}{2}\right)}{8}\right) > 2\exp(-1/8) > 1,$$

y la desigualdad es trivialmente cierta en este caso. $\square$

Capítulo 3

Una aplicación a la teoría de códigos

Aseverar que un código tiene una propiedad específica es casi siempre una tarea difícil. Por otro lado, muchas veces es fácil mostrar que la mayoría de códigos, en un ensamble elegido apropiadamente, poseen una propiedad.

T. Richardson y R. Urbanke, [RU06]

Este capítulo es un interludio donde aplicaremos cuestiones de concentración a la teoría de códigos. Empezaremos por motivar e introducir brevemente las generalidades de la teoría y después nos enfocaremos en los códigos lineales de baja densidad. Finalmente, demostraremos un teorema que nos dice cómo se comporta "respecto al promedio" un algoritmo de decodificación sobre este tipo de códigos.

3.1. Cuestiones básicas

Supongamos que dos agentes, un *emisor* y un *receptor* desean comunicarse. Acuerdan utilizar un cierto conjunto finito de símbolos Σ con el cual codifican sus mensajes. Puesto que los mensajes pueden ser muy largos, también deciden descomponerlos en paquetes o *bloques* de longitud n.

En un mundo ideal, los dos agentes tendrán a su disposición un medio de transmisión perfecto, donde todo lo enviado por el emisor es recibido de forma fiel por el receptor. Sin embargo, en el mundo real, es poco probable que la transmisión sea perfecta y algunos símbolos del mensaje llegarán alterados, o inclusive serán reemplazados por ruido. Ante esta situación, el receptor deseará contar con algún mecanismo que le permita recuperar toda o parte de la información perdida a partir de lo que ha recibido.

Esta situación, cuyo estudio ha estado estimulado por las telecomunicaciones modernas, en lo concerniente a la naturaleza de los mensajes se captura matemáticamente a través de la teoría de códigos.

Definición 3.1. Sea Σ un alfabeto de cardinalidad q. Un código C es un subconjunto de Σ^n para algún entero positivo n. A cada elemento del código C le llamaremos palabra.

En Σ^n se puede introducir una distancia muy natural denominada métrica de Hamming, y que denotaremos por Δ. Definimos a Δ como

$$\Delta(x, y) = |\{x_i \neq y_i : 1 \leq i \leq n\}|,$$

es decir, el número de coordenadas en las cuales difieren x y y. Con esta métrica podemos definir la mínima distancia de un código a través de

$$\Delta(C) := \min_{x,y \in C, x \neq y} \Delta(x, y).$$

Hay cuatro parámetros asociados a un código C.

- El tamaño del bloque, n, donde $C \subseteq \Sigma^n$.
- Longitud del mensaje, $k = \log_q |C|$. Conviene observar que, a pesar de que se envíen n símbolos, sólo k contienen al mensaje.
- La distancia mínima $d = \Delta(C)$.
- El tamaño del alfabeto, $q = |\Sigma|$.

Es bastante bueno conseguir un código C con un valor grande de d. ¿Por qué? Veamos lo que definió Hamming al respecto.

Definición 3.2. El código C es e-detector de errores si, bajo la garantía de que no ocurren más de e errores durante las transmisión, es posible detectar si han ocurrido errores o no, y e es el mayor entero con esta propiedad. El código C es t-corrector de errores si, siempre que no ocurran más de t errores durante la transmisión, es siempre posible determinar cuáles son los errores y corregirlos, y t es el mayor entero con esta propiedad.

Teorema 3.1. *Un código C con distancia mínima d puede corregir hasta $\lfloor \frac{d-1}{2} \rfloor$ errores y detectar hasta $d-1$ errores.*

Demostración. Si el código es capaz de detectar e errores no puede ser que dos palabras disten en e o menos, pues entonces una palabra correcta y una con e errores serían indistinguibles. Por lo tanto $d > e$, así que e es a lo más $d-1$. Por otro lado, para corregir t errores es necesario que las bolas de radio t no se intersecten, así que $d > 2t$, y entonces $d - 1 \geq 2t$; luego $t \leq \lfloor \frac{d-1}{2} \rfloor$. $\square$

3.1.1. Códigos lineales

Los códigos lineales se obtienen cuando el alfabeto Σ es, de hecho, un campo finito $\mathbb{F}_q$ para algún q. Lo más interesante es que podemos elegir el código C como un subespacio de $\mathbb{F}_q^n$, y en este caso decimos que el código es lineal.

Notación 3.1. A un código lineal con longitud de bloque n y longitud de mensaje k lo denotaremos como $C[n,k]$.

Siendo así, podemos aprovechar la rica teoría algebraica disponible para espacios vectoriales y obtener propiedades útiles de los códigos lineales. Por simplicidad, consideraremos sólo códigos para los cuales $q = 2$, y que llamaremos binarios.

Digamos, de manera específica, que C es un subespacio k-dimensional de $\mathbb{F}_2^n$ sobre $\mathbb{F}_2$. Podemos encontrar una base $B = \{g_i\}_{i=0}^{k-1}$ para C. Tenemos que para todo $c \in C$ existen $u_0, \ldots, u_{k-1} \in \mathbb{F}_2$ tales que

$$c = \sum_{i=0}^{k-1} u_i g_i,$$

o, de forma más sucinta,

$$c = uG,$$

donde G es una matriz de tamaño $k \times n$ cuyas filas son los vectores g_i. A G se le denomina matriz generadora del código C. El espacio nulo $C^{\perp}$ de C es un subespacio $(n-k)$-dimensional que consiste en todos los vectores x tales que

$$xG^T = 0.$$

Consideremos ahora una base $\{h_i\}_{i=0}^{n-k-1}$ de $C^{\perp}$. Para $c \in C$ se cumple que $ch_i^T = 0$ para todo i, que se resume en la expresión

$$cH^T = 0,$$

donde H es una matriz es una matriz de tamaño $(n-k) \times n$ cuyas filas son los vectores $\{h_i\}$. A H se le llama matriz de verificación de paridad de C.

Definición 3.3. Un código de verificación de paridad de baja densidad es un código $C[n,k]$ cuya matriz de verificación de paridad tiene $O(n)$ entradas no nulas.

Abreviamos los códigos anterioremente definidos como códigos VPBD. Fueron introducidos por Gallager en 1961, y pasaron desapercibidos a pesar de sus excelentes propiedades.

3.2. El canal binario con pérdida

El canal binario con pérdida (CBP) es quizá el modelo de canal de comunicación más simple no trivial, y fue introducido como ejemplo excesivamente simplificado por Elias en 1954. No obstante, la aparición de la internet ha hecho retomar a este modelo como algo factible en la realidad.

En el canal binario se transmiten solamente dos símbolos: 0 y 1. Si el emisor transmite al tiempo t el símbolo x_t, con probabilidad $1-\epsilon$ el receptor recibe x_t y con probabilidad ϵ se pierde; esto se representa con el símbolo ?. Además, la probabilidad de pérdida es independiente del tiempo t. Dicho de otra forma, el canal no tiene memoria, lo cual se expresa como

$$P(x_{t'} = ?|x_t) = \epsilon, \quad t' > t.$$

A un canal binario con probabilidad de pérdida ϵ lo denotaremos como CBP(ϵ).

3.2.1. Transmisión usando códigos lineales

Consideremos un código binario lineal $C[n,k]$ definido en términos de su matriz de verificación de paridad H. Supongamos que el transmisor escoge de forma uniformemente aleatoria una palabra X de C y que la transmisión se realiza sobre el CBP(ϵ). Sea Y la palabra recibida con elementos en el alfabeto extendido $\{0,1,?\}$, $\mathcal{E} \subset [n] := \{1,\ldots,n\}$ el conjunto de índices de pérdidas y $\overline{\mathcal{E}} = [n] \setminus \mathcal{E}$. En otras palabras, $i \in \mathcal{E}$ si, y sólo si, $Y_i = ?$.

Consideremos ahora un decodificador de bloque con una regla de máxima probabilidad a posteriori (MAP), es decir

$$\hat{x}_B^{\mathrm{MAP}}(y) := \arg \max_{x \in C} P_{X|Y}(x|y).$$

Escribamos la ecuación $Hx^T = 0$ de la forma

$$H_{\mathcal{E}} x_{\mathcal{E}}^T + H_{\overline{\mathcal{E}}} x_{\overline{\mathcal{E}}}^T = 0$$

que, dado que trabajamos sobre $\mathbb{F}_2$, es equivalente a

$$H_{\mathcal{E}} x_{\mathcal{E}}^T = H_{\overline{\mathcal{E}}} x_{\overline{\mathcal{E}}}^T.$$

Desde luego, $s^T := H_{\overline{\mathcal{E}}} x_{\overline{\mathcal{E}}}^T$ es conocido por el receptor, lo que nos indica que debemos considerar la ecuación $H_{\overline{\mathcal{E}}} x_{\overline{\mathcal{E}}}^T = s^T$. Puesto que, por hipótesis, la palabra transmitida es una palabra válida, sabemos que esta ecuación tiene al menos una solución, lo cual quiere decir que el rango de $H_{\mathcal{E}}$ es a lo más $|\mathcal{E}|$. Cuando se cumple la igualdad, el decodificador MAP puede hallar exactamente una solución a la ecuación y la decodificación queda unívocamente determinada. En otro caso, hay múltiples soluciones igualmente probables en virtud de la elección de X. Sea

$$\mathcal{X}^{\mathrm{MAP}}(y) := \{x \in C : H_{\mathcal{E}} x_{\mathcal{E}}^T = H_{\overline{\mathcal{E}}} y_{\overline{\mathcal{E}}}^T, x_{\overline{\mathcal{E}}} = y_{\overline{\mathcal{E}}}^T\},$$

entonces

$$\hat{x}_B^{\mathrm{MAP}}(y) = \begin{cases} x \in \mathcal{X}^{\mathrm{MAP}}(y), & |\mathcal{X}^{\mathrm{MAP}}(y)| = 1, \\ \text{pérdida}, & \text{en otro caso.} \end{cases}$$

Esto quiere decir que cuando podemos decodificar de manera unívoca una palabra recibida, bastará resolver un sistema de ecuaciones. Usando la reducción gaussiana y una sustitución hacia atrás, podemos lograrlo en $O(n^3)$ iteraciones. Sin embargo, usando códigos VPBD, existen algoritmos de decodificación que corren casi siempre en tiempo $O(n)$, aprovechando precisamente el hecho de que las pocas entradas no nulas de la matriz de verificación de paridad muy probablemente eliminen en gran parte o totalmente la necesidad de hacer la reducción gaussiana.

3.3. Grafos de Tanner

A cada código binario C podemos asociarle un grafo bipartito denominado grafo de Tanner de la siguiente manera. Sea H la matriz de verificación de paridad de C, y supongamos que es de tamaño $m \times n$. Diremos entonces que el grafo de Tanner tendrá n vértices de *variable* y m vértices de *verificación*, y un vértice de verificación i estará conectado a un vértice j de variable si, y sólo si, $H_{i,j} = 1$, lo que significa que la variable j participa de la i-ésima restricción de verificación de paridad.

Cabe observar que para un código dado existe más de un grafo de Tanner asociado, de la misma manera que existen varias bases para $C^{\perp}$. A pesar de que todos estos grafos representan al mismo código, no son equivalentes desde el punto de vista del decodificador por paso de mensajes que veremos después.

Para el caso de los códigos VPBD, los grafos de Tanner asociados tienen pocas aristas. Esto se hace patente para ciertos casos particulares que definimos a continuación.

Definición 3.4. Un código VPBD (l, r)-regular es tal que, para su grafo de Tanner asociado, cada vértice de variable es de grado l y cada vértice de verificación es de grado r.

El número de aristas del grafo de Tanner asociado a un código VPDB (l, r)-regular es ln, donde n es la longitud del código. Conforme n crece, el número de aristas (o entradas no nulas de la matriz de verificación de paridad) crece linealmente respecto a n.

La situación se vuelve más ventajosa cuando examinamos el caso general. Consideremos a un código VPBD de longitud n tal que el número de vértices de variables de grado i es Λ_i, de tal suerte que $\sum_i \Lambda_i = n$. De la misma forma, denotando el número de vértices de verificación de grado i como P_i, tenemos que $\sum_i P_i = n\bar{r}$, donde $r = m/n$ es la razón de diseño del código y $\bar{r} = 1 - r$. Puesto que el conteo de aristas debe coincidir,

$$\sum_i i\Lambda_i = \sum_i iP_i.$$

Definimos también los polinomios

$$\Lambda(x) := \sum_{i=2}^{l_{\text{máx}}} \Lambda_i x^i, \quad P(x) := \sum_{i=1}^{r_{\text{máx}}} P_i x^i$$

donde $\Lambda_1 = 0$, pues los códigos con estas características tienen buenas propiedades.

De estas definiciones se siguen inmediatamente las siguientes relaciones

$$\Lambda(1) = n, \quad P(1) = n\bar{r}, \quad r(\Lambda, P) = 1 - \frac{P(1)}{\Lambda(1)}, \quad \Lambda'(1) = P'(1).$$

Llamamos a Λ y P las distribuciones de grados de variable y de verificación, respectivamente. También consideraremos las distribuciones normalizadas de grado

$$L(x) := \frac{\Lambda(x)}{\Lambda(1)}, \quad R(x) := \frac{P(x)}{P(1)}.$$

Definición 3.5. Dado un par de distribuciones de grado (Λ, P), definimos el ensamble de grafos bipartitos VPBD(Λ, P) de la siguiente manera. Cada grafo en VPBD(Λ, P) tiene $\Lambda(1)$ vértices de variable y $P(1)$ vértices de verificación, Λ_i vértices de variable y P_i vértices de verificación de grado i. Un vértice de grado i tiene i enchufes de los cuales i aristas emanan, de modo que hay $\Lambda'(i) = P'(i)$ en cada lado. Etiquetemos los enchufes de cada lado con el conjunto

$$[\Lambda'(i)] := \{1, \ldots, \Lambda'(i)\}$$

de alguna manera arbitraria pero fija. Sea σ una permutación de $[\Lambda'(i)]$. Asociemos a σ un grafo bipartito conectando el i-ésimo enchufe en el lado de las variables con el $\sigma(i)$-ésimo enchufe en el lado de la verificación. Dejando a σ correr sobre todo $S_{\Lambda'(1)}$ genera un conjunto de grafos bipartitos, que junto con una distribución de probabilidad uniforme sobre los mismos, constituye el ensamble VPBD(Λ, P).

Escolio 3.1. Para cada elemento en el conjunto subyacente en VPBD(Λ, P) podemos asociarle un código considerando la matriz de verificación de paridad H

que tiene un 1 en la entrada i, j-ésima si, y sólo si, el i-ésimo vértice de verificación está conectado con la j-ésimo vértice de variable un número impar de vértices. Por esta razón, indistintamente consideraremos que VPBD(Λ, P) está conformado de códigos o grafos bipartitos, aún cuando la distribución de probabilidad resultante en el caso de los códigos no es uniforme.

Para terminar, esta sección, cambiaremos de perspectiva respecto al número de vértices por el número de aristas. Definimos los polinomios

$$\lambda(x) = \sum_i \lambda_i x^{i1} := \frac{\Lambda'(x)}{\Lambda'(1)} = \frac{L'(x)}{L'(1)},\ \rho(x) = \sum_i \rho_i x^{i-1} := \frac{P'(x)}{P'(1)} = \frac{R'(x)}{R'(1)}$$

y se verifica fácilmente que λ_i es la fracción de aristas conectadas a vértices de variables de grado i y ρ_i es la fracción de aristas conectadas a vértices de verificación de grado i. Evidentemente, tanto el par (Λ, P) como la tripla (n, λ, ρ) conllevan la misma información. Por eso, al conjunto VPBD(Λ, P) lo denotaremos también como VPBD(n, λ, ρ).

3.4. Decodificación por paso de mensajes

Como habíamos mencionado, cuando la matriz de verificación de paridad es de baja densidad, existe un algoritmo que casi siempre corre en tiempo lineal y decodifica correctamente un mesaje Y recibido en caso de presentar algunos símbolos perdidos al transmitir sobre el CBP(ϵ). Este algoritmo se conoce como decodificador por paso de mensajes.

El algoritmo procede por varias rondas de paso de mensajes. Primero, se envían los mensajes de los vértices de verificación a los vértices de variable a lo largo de todas las aristas. Estos mensajes se procesan a continuación en los vértices de variable y se reenvían a los vértices de verificación. Esto constituye una ronda de paso de mensajes. La propiedad característica de este algoritmo es que los mensajes emanados de un vértice particular a lo largo de diferentes aristas pueden ser diferentes.

Los mensajes se toman del conjunto $\{0, 1, ?\}$, donde ? significa pérdida, como hemos acordado. Asimismo indica una variable cuyo valor aún no ha sido determinado.

Dado un vértice de variable v, sea $E(v)$ el conjunto de las aristas que emanan de él. De la misma forma definimos $E(c)$ para un vértice de verificación c. Decimos que un mensaje o valor recibido que no sea un pérdida como *conocido*, y denotamos al mensaje de entrada enviado por la arista e como $m_i(e)$, y a los de salida como $m_s(e)$.

Definición 3.6. Decimos que un vértice de variable es conocido si su valor recibido es conocido o si tiene al menos mensaje de entrada que es conocido. Decimos que un vértice de verificación es conocido si a lo más uno de sus mensajes de entrada es conocido.

Algoritmo 3.1 (Decodificación por paso de mensajes)**.** Dada una palabra Y recibida por el CBP y el grafo de Tanner T asociado al código C, decodifica la palabra original X.

Hacer todos los mensajes de vértices de verificación a vértices de variable iguales a ?.
repetir
  para todo $c \in T$ de verificación **hacer**
    para todo $e \in E(c)$ **hacer**
      si $\exists f \in E(c) \setminus e, m_i(f) = ?$ **entonces**
        $m_o(e) \leftarrow ?$.
      si no
        $m_o(e) \leftarrow \sum_{f \in E(c) \setminus e} m(f)$.
      fin si
    fin para
  fin para
  para todo $v \in T$ de variable **hacer**
    para todo $e \in E(v)$ **hacer**
      si $Y_v = ? \wedge \forall f \in E(v) \setminus e, m_i(f) = ?$ **entonces**
        $m_o(e) \leftarrow ?$.
      si no
        $m_o(e) \leftarrow Y_v$.
      fin si
    fin para
  fin para
hasta que no haya más mensajes que enviar.

Si al término del algoritmo anterior todos los vértices de variable son conocidos, tenemos éxito y la palabra original es recuperada. Puede verse que el algoritmo visita cada arista a lo más dos veces, así que corre en tiempo lineal sobre la longitud del mensaje.

Teorema 3.2. *Sea T el grafo de Tanner que representa a un código lineal binario C. Supongamos que C se usa para transmitir sobre el* CBP(ϵ) *y que el decodificador realiza una decodificación por paso de mensajes en T. Denotemos por $P^{\mathrm{BP}}(T, \epsilon, \ell, x)$ la probabilidad condicional de pérdida después de la ℓ-ésima* *iteración, suponiendo que $x \in C$ fue enviado. Entonces*

$$P^{\mathrm{BP}}(T, \epsilon, \ell, x) = \frac{1}{|C|} \sum_{x' \in C} P^{\mathrm{BP}}(T, \epsilon, \ell, x') := P^{\mathrm{BP}}(T, \epsilon, \ell),$$

esto es, $P^{\mathrm{BP}}(T, \epsilon, \ell, x)$ es independiente de la palabra transmitida.

Demostración. Obsérvese que el mensaje a lo largo de la arista e es ya sea pérdida o es igual al valor de la entrada asociada. Consideremos el paso de mensajes suponiendo que dos palabras distintas, digamos x y x', fueron transmitidas pero que ha ocurrido el mismo patrón de pérdida ha ocurrido. Veamos el mensaje de verificación a variable enviado durante la ℓ-ésima iteración a lo largo del vértice e que está conectado con el i-ésimo vértice de variable. Procedemos por inducción. En la primera iteración todos los mensajes son de pérdida, así que es trivialmente cierto el aserto. Examinemos un mensaje subsecuente de verificación a variable a lo largo de la arista e que está conectado al vértice de verificación c. Este mensaje es pérdida si, y sólo si, cualquiera de los mensajes de variable a verificación a lo largo de los ejes $E(c) \setminus e$ son de pérdida. En otro caso, el mensaje es la suma de todos los mensajes entrantes a lo largo de las aristas $E(c) \setminus e$. Pero al ser x y x' palabras código validas y en virtud de las restricciones, esta suma es igual a x_i y x'_i, respectivamente. Un argumento análogo comprueba el paso inductivo para los mensajes a vértices de variable. □

Sea Z el número de mensajes de pérdida de variables a verificación entre las nl mensajes de pérdida, pasados durante la ℓ-ésima iteración para un $(T, R) \in \Omega$. Aquí G es un grafo del ensamble $\mathrm{VPBD}(n, x^{l-1}, x^{r-1})$, R es una entrada particular al decodificador y Ω es un espacio de probabilidad. Sea $=_i$ para $0 \leq i \leq (l+1)n$ una sucesión de relaciones de equivalencia en Ω ordenadas por refinamiento, esto es,

$$(G', R') =_i (G'', R'') \Rightarrow (G', R') =_{i-1} (G'', R'').$$

Estas clases de equivalencia están definidas por igualdades parciales. Para ser más precisos: supóngase que se exponen las ln aristas del grafo una a la vez, es decir, en el paso $i \in [nl]$ exponemos al enchufe $\pi(i)$ de un vértice de verificación que está conectado con el i-ésimo vértice de variable. De modo similar definimos en los siguientes n pasos se exponen n valores recibidos, uno a la vez. Tenemos que $(G', R') =_i (G'', R'')$ si, y sólo si, la información revelada durante los primeros i pasos es la misma para ambos pares. Denotemos por h el número de pasos en este procedimiento de exposición, notando que $h = (l+1)n$.

A continuación, definamos $Z_0, Z_1, \ldots, Z_n$ por

$$Z_i(G, R) := E(Z(G', R') | (G', R') =_i (G, R)).$$

Por construcción, $\{Z_i\}_{i=0}^h$ es una almártaga. Procederemos a encontrar las cotas γ_i de modo que

$$|Z_{i+1}(G,R) - Z_i(G,R)| \leq \gamma_i, \quad i = 0, \dots, h-1, \tag{3.1}$$

tratando de que dependan de l, r y ℓ, pero no de n.

Consigamos la expresión (3.1) para $i \in [nl]$, esto es, los pasos donde exponemos a las aristas. Recordemos que $\pi(i) = j$ significa que el enchufe del i-ésimo vértice de variable está conectado al enchufe del j-ésimo vértice de verificación. Sea $\mathcal{G}(G,i)$ el conjunto de grafos en el ensamble VPBD(n, x^{l-1}, x^{r-1}) tal que las primeras i aristas son iguales a las aristas de G. Tenemos

$$\mathcal{G}(G,i) = \{G' : (G', R) =_i (G,R)\}.$$

Sea $\mathcal{G}_j(G,i)$ el subconjunto de $\mathcal{G}(G,i)$ que consiste en aquellos grafos para los cuales $\pi(i+1) = j$. Entonces

$$\mathcal{G}(G,i) = \bigcup_j \mathcal{G}_j(G,i).$$

Según lo anterior,

$$\begin{aligned} Z_i(G,R) &= E(Z(G',R')|G \in (\mathcal{G}(G,i) \\ &= \sum_{j \in [nl]} E(Z(G',R')|G' \in \mathcal{G}_j(G,i))P(G' \in \mathcal{G}_j(G,i)|G' \in \mathcal{G}(G,i)) \end{aligned} \tag{3.2}$$

Afirmamos que si j y k son tales que

$$P(G' \in \mathcal{G}_j(G,i)|G' \in \mathcal{G}(G,i)) \neq 0 \quad \text{y} \quad P(G' \in \mathcal{G}_k(G,i)|G' \in \mathcal{G}(G,i)) \neq 0$$

entonces

$$|E(Z(G',R')|G' \in \mathcal{G}_j(G,i)) - E(Z(G',R')|G' \in \mathcal{G}_k(G,i))| \leq 8(lr)^\ell. \tag{3.3}$$

Para probar este aserto, definimos un mapeo $\phi_{j,k} : \mathcal{G}_j(G,i) \to \mathcal{G}_k(G,i)$ como sigue. Sea π la permutación que define la asignación de aristas para un grafo dado $H \in \mathcal{G}_j(G,i)$ y sea $i' = \pi^{-1}(k)$. Definimos la permutación π' como una idéntica a π, salvo por el hecho de que $\pi'(i+1) = k$ y $\pi'(i') = j$. Sea H' el grafo que resulta de esta operación. Observemos que $H' \in \mathcal{G}_k(G,i)$ y que $\phi_{k,j}$ es una biyección que conserva probabilidades. Veremos ahora que, para un H y R fijos

$$|Z(H,R) - Z(\phi_{j,k}(H),R)| \leq 8(lr)^\ell \tag{3.4}$$

y de aquí deduciremos nuestra afirmación inicial.

Para este fin, notemos primero que el mensaje enviado por una arista dada durante la ℓ-ésima iteración es sólo una función del grafo sobre el cual se trabaja y la información que se tiene hasta ese momento, que llamaremos C_ℓ. Por lo tanto, un mensaje se ve afectado por un intercambio de extremos de dos aristas si al menos una de estas aristas pertenece a C_ℓ. Tenemos C_ℓ tiene a lo sumo $2(lr)^\ell$ aristas distintas y, por simetría, una arista puede ser parte de a lo más $2(lr)^\ell$ grafos de esta naturaleza. Se sigue que a lo más $8(lr)^\ell$ grafos pueden ser afectados por el intercambio de extremos de aristas, lo que prueba la desigualdad (3.4).

Ahora cualquier par $Z(H, R)$ y $Z(\phi_{j,k}(H), R)$ tienen una diferencia acotada por $8(lr)^\ell$ y, puesto que para cualquier variable aleatoria $|E(W)| \leq E(|W|)$, se sigue el aserto (3.3).

Según la definición, $Z_{i+1}(G, R) = E(Z(G', R')|G' \in \mathcal{G}_j(G, i))$ para algún $j \in \Psi_i$, donde $\Psi_i \subseteq [nl]$ denota el conjunto de enchufes desocupados en el lado de los vértices de verificación después de revelar i aristas de G. Por lo tanto

$$\begin{aligned}|Z_{i+1}(G,R) - Z_i(G,R)| &\leq \max_{j\in\Psi_i} |E(Z(G',R')|G' \in \mathcal{G}_j(G,i)) - Z_i(G,R)| \\ &\leq \max_{j,k\in\Psi_i} |E(Z(G',R')|G' \in \mathcal{G}_j(G,i)) - E(Z(G',R')|G' \in \mathcal{G}_k(G,i))| \\ &\leq 8(lr)^\ell,\end{aligned}$$

donde usamos la expresión (3.2). Esto demuestra que podemos tomar $\gamma_i = 8(lr)^\ell$ para $i \in [nl]$ en (3.1).

Para la revelación de valores, el argumento es esencialmente el mismo, excepto que al no considerarse cuatro posibles combinaciones de intercambios de extremos de vértices basta considerar la cota $2(lr)^\ell$ dada por la simetría. Queda probado así que podemos hacer $\gamma_i = 2(lr)^\ell$ para $i \in \{nl + 1, \ldots, (n + 1)l\}$ en (3.1).

La desigualdad (3.1) conjuntamente con el Teorema 2.8, arroja inmediatamente este resultado.

Teorema 3.3. *Sea G elegido aleatoriamente de* VPBD(n, λ, ρ) *para ser usado en la transmisión por el* CBP(ϵ) *y supongamos que el decodificador realiza ℓ iteraciones del Algoritmo 3.1. Entonces, para $\delta > 0$, se satisface la desigualdad*

$$P(|Z - E(Z)| > \delta n) \leq 2 \exp\left(-\frac{\delta^2 n^2}{128[(l+1)n+1](lr)^{2\ell}}\right).$$

Este teorema nos habla del desempeño del decodificador por paso de mensajes usando como parámetro la media de los códigos en VPBD(n, λ, ρ), que es arquetípico del paradigma de decodificación iterativa. En la práctica da muy buenos resultados, y el Teorema 3.3 da "apoyo moral" a este paradigma, pues necesita valores muy grandes de n para caer por debajo de 1 y decirnos algo útil sobre la decodificación. Si se puede refinar lo suficiente la cota (3.1), entonces puede aplicarse el Teorema 2.9 para obtener una cota potencialmente más útil. Esto, no obstante, no es tan sencillo pues exige replantear profundamente el análisis anteriormente hecho a la decodificación por paso de mensajes.

Capítulo 4

El Teorema de Dvoretzky

Se tenía la idea de que los espacios de dimensión grande eran modelizados por los espacios de dimensión infinita. El teorema de Dvoretzky dió la confirmación concluyente de que este hecho era rotundamente falso.

Jesús Bastero Eleizalde, [BE00]

Una bola es parte esencial del juego.

Johan Cruijff

Es momento de enunciar (y demostrar) el teorema de Dvoretzky. Nos tomará todo el capítulo así que procederemos con calma.

4.1. Cuerpos convexos y seminormas

Veamos una generalización del concepto de norma, que nos ayudará a replantear el teorema de Dvoretzky para facilitar su demostración. Con X denotaremos a un espacio vectorial real.

Definición 4.1. Una función $\rho : X \to \mathbb{R}$ es una seminorma si:

1. Es subaditiva, es decir, para todo $x, y \in X$ se cumple

$$\rho(x+y) \leq \rho(x) + \rho(y). \tag{4.1}$$

2. Satisface, para $\alpha \in \mathbb{R}$,
$$\rho(\alpha x) = |\alpha|\, \rho(x). \tag{4.2}$$

Obsérvese que toda norma es una seminorma pero no recíprocamente. Por ejemplo, en cualquier espacio vectorial X podemos definir la seminorma trivial

$$\rho(x) = 0 \text{ para toda } x \in X,$$

que obviamente satisface (4.1) y (4.2), pero no cumple que $\rho(x) = 0$ si, y sólo si, $x = 0$.

Proposición 4.1. *Una seminorma ρ en X satisface*

$$\begin{aligned} \rho(0) &= 0, & (4.3) \\ \rho(x-y) &\geq |\rho(x) - \rho(y)|\,, & (4.4) \end{aligned}$$

en particular, $\rho(x) \geq 0$.

Demostración. Por (4.2)

$$\rho(0) = \rho(0x) = 0\rho(x) = 0.$$

Nótese que $\rho(-x) = |-1|\, \rho(x) = \rho(x)$. De la desigualdad (4.1) se tiene

$$\rho(x) = \rho(x - y + y) \leq \rho(x-y) + \rho(y),$$

e intercambiando x e y obtenemos

$$\rho(y) \leq \rho(y-x) + \rho(x) = \rho(x-y) + \rho(x),$$

y de aquí que
$$-\rho(x-y) \leq \rho(x) - \rho(y) \leq \rho(x-y),$$
lo que implica (4.4). □

Teorema 4.1. *Sea ρ una seminorma en X, y $c > 0$. Entonces el conjunto*

$$K = \{x \in X : \rho(x) \leq c\}$$

disfruta de las siguientes propiedades.

1. *El elemento 0 pertenece a K.*

2. *El conjunto K es convexo: si $x, y \in K$ y $\lambda \in [0,1]$ entonces $\lambda x + (1-\lambda) y \in K$.*

3. *El conjunto K es equilibrado: si $x \in K$ y $|\lambda| \leq 1$, entonces $\lambda x \in K$. En particular, K es simétrico.*

4. *El conjunto K es absorbente: para cualquier $x \in X$ existe $\lambda > 0$ tal que $\frac{x}{\lambda} \in K$.*

5. *Se cumple*
$$\rho(x) = \inf_{\lambda > 0} \left\{ \lambda c : \frac{x}{\lambda} \in K \right\}.$$

Demostración. La primera afirmación es clara a partir de (4.3). Supóngase que $|\lambda| \leq 1$. La tercera y cuarta son consecuencia de
$$\rho\left(\lambda x\right) = |\lambda|\, \rho(x) \leq |\lambda|\, c \leq c$$
y
$$\rho\left(\frac{c}{\rho(x)+1}x\right) < c.$$

El segundo aserto se debe a que
$$\begin{aligned} \rho\left(\lambda x + (1-\lambda)y\right) &\leq \rho\left(\lambda x\right) + \rho\left((1-\lambda)y\right) \\ &\leq \lambda c + (1-\lambda)\, c = c. \end{aligned}$$

Finalmente, tenemos que $\frac{x}{\lambda} \in K$ si, y sólo si, $\rho\left(\frac{x}{\lambda}\right) \leq c$, y esto último si, y sólo si, $\rho(x) \leq \lambda c$. Por lo tanto $\rho(x)$ es cota inferior del conjunto $\left\{\lambda c : \frac{x}{\lambda} \in K, \lambda > 0\right\}$. Pongamos
$$\beta := \inf_{\lambda > 0} \left\{ \lambda c : \frac{x}{\lambda} \in K \right\}.$$

Entonces $\rho(x) \leq \beta$. Sea $\epsilon > 0$. Escogiendo
$$\lambda_\epsilon = \frac{\rho\left(x\right) + \epsilon}{c},$$
obtenemos
$$\rho\left(\frac{x}{\lambda_\epsilon}\right) = \rho\left(\frac{c}{\rho\left(x\right)+\epsilon}x\right) = c\frac{\rho(x)}{\rho\left(x\right)+\epsilon} < c,$$
por lo tanto $\lambda_\epsilon c \in \left\{\lambda c : \frac{x}{\lambda} \in K\right\}$, y
$$\beta \leq \lambda_\epsilon c = \rho\left(x\right) + \epsilon$$
o, lo que es lo mismo, $\beta \leq \rho(x)$, y juntando las dos desigualdades obtenemos lo deseado. □

Definición 4.2. Sea $K \subseteq X$. El funcional
$$\rho_K(x) = \inf_{\lambda > 0} \left\{ \lambda : \frac{x}{\lambda} \in K \right\}$$
se llama el funcional de Minkowski del conjunto K.

El funcional de Minkowski es sumamente importante: si K es un conjunto apropiado, entonces su comportamiento es bastante familiar.

Teorema 4.2. *El funcional de Minkowski $\rho_K : X \to \mathbb{R}$ sobre un conjunto K convexo, absorbente y equilibrado de X es una seminorma.*

Demostración. Escribimos ρ en lugar de ρ_K, por simplicidad. Como K es absorbente, el conjunto $\left\{\lambda > 0 : \frac{x}{\lambda} \in K\right\}$ es no vacío, pues existe $\alpha > 0$ tal que

$$\frac{1}{\alpha}x \in K$$

para cualquier $x \in X$. El conjunto, además, está acotado inferiormente por 0. Por lo tanto $\rho(x)$ está bien definido. Sean $x, y \in X$ y $\epsilon > 0$. Por la definición del funcional de Minkowski deben existir $\lambda_1, \lambda_2 > 0$ tales que

$$\lambda_1 < \rho(x) + \epsilon$$

y

$$\lambda_2 < \rho(y) + \epsilon,$$

que satisfacen

$$\frac{1}{\lambda_1}x, \frac{1}{\lambda_2}y \in K.$$

Por la convexidad de K se tiene que

$$\left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)\frac{1}{\lambda_1}x + \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)\frac{1}{\lambda_2}y \in K.$$

Se cumple ahora que $\frac{x+y}{\lambda_1+\lambda_2} \in K$. Nuevamente, por la definición del funcional de Minkowski,

$$\rho(x + y) \leq \lambda_1 + \lambda_2 < \rho(x) + \rho(y) + 2\epsilon,$$

lo que arroja la desigualdad del triángulo al tender ϵ a cero. Sea $\alpha \in \mathbb{R}$. Como K es equilibrado entonces

$$\frac{-1}{\lambda_1}x = \frac{1}{\lambda_1}(-x) \in K,$$

lo que implica que $\rho(-x) \leq \lambda_1 < \rho(x) + \epsilon$, es decir, $\rho(-x) \leq \rho(x)$, lo cual es válido para toda $x \in K$. En particular, $\rho(-(-x)) = \rho(x) \leq \rho(-x)$, por lo que

$\rho(-x) = \rho(x)$. Ahora bien, para cada $\alpha > 0$, si $y = \alpha x$, entonces

$$\begin{aligned} \rho(\alpha x) &= \rho(y) = \operatorname*{ínf}_{\lambda>0}\left\{\lambda : \frac{y}{\lambda} \in K\right\} \\ &= \operatorname*{ínf}_{\lambda>0}\left\{\lambda : \frac{\alpha x}{\lambda} \in K\right\} \\ &= \operatorname*{ínf}_{\lambda=\alpha\lambda'>0}\left\{\lambda : \frac{\alpha x}{\lambda} \in K\right\} \\ &= \operatorname*{ínf}_{\alpha\lambda'>0}\left\{\alpha\lambda' : \frac{x}{\lambda'} \in K\right\} \\ &= \alpha \operatorname*{ínf}_{\lambda'>0}\left\{\lambda' : \frac{x}{\lambda'} \in K\right\} = \alpha\rho(x), \end{aligned}$$

ya que si $\alpha > 0$ entonces $|\alpha| = \alpha$. Si $\alpha < 0$ entonces $\alpha = -|\alpha|$ y

$$\rho(\alpha x) = \rho(-|\alpha|\, x) = \rho(|\alpha|\, x) = |\alpha|\, \rho(x).$$

Por último, como $0x = 0$ y para todo $\lambda > 0$ se cumple que $\lambda 0 = 0$, entonces

$$\rho(0x) = \operatorname*{ínf}_{\lambda>0}\{\lambda : \lambda \in \mathbb{R}\} = 0 = 0\rho(x),$$

por lo que se cumple la propiedad (4.2). □

Para una seminorma ρ en X, el conjunto $K = \{x : \rho(x) \leq 1\}$ es convexo, equilibrado y absorbente. Si ρ fuese además una norma, K tiene las mismas propiedades. ¿Tendrá otras? En este caso, si $x \neq 0$, se tiene que $\rho(x) \neq 0$, es decir

$$\rho_K(x) = \operatorname*{ínf}_{\lambda>0}\left\{\lambda : \frac{x}{\lambda} \in K\right\} > 0.$$

Traducción: para cada $x \in X$ no nulo existe un $0 < C_x < \infty$ tal que, si $kx \in K$, entonces $k \leq C_x$. Supongamos ahora que ρ es una seminorma y que, si $x \in X$ es distinto de cero, existe $0 < C_x < \infty$ tal que $kx \in K$ implica que $k \leq C$. Puesto que $\frac{1}{k} \geq \frac{1}{C_x} > 0$, entonces $\frac{1}{C_x}$ es una cota inferior para $\left\{\lambda : \frac{x}{\lambda} \in K\right\}$, y de aquí se desprende que

$$\rho_K(x) \geq \frac{1}{C_x} > 0.$$

Helo ahí: ahora ρ es una norma. Resumimos estas observaciones en el siguiente teorema.

Teorema 4.3. *Sea K un subconjunto de X. El funcional de Minkowski ρ_K es una norma en X si, y sólo si, K es convexo, equilibrado, absorbente y tiene la propiedad*

$$(\forall x \in X, x \neq 0)\,(\exists C_x, 0 < C_x < \infty)\,(k > 0, kx \in K \Rightarrow k \leq C_x)\,. \tag{4.5}$$

Definición 4.3. Sea $Y \subset X$. El núcleo de Y se define como el conjunto

$$N(Y) = \{y \in Y : (\forall x \in X)(\exists \epsilon > 0)(|t| < \epsilon \Rightarrow y + tx \in Y)\}.$$

Definición 4.4. Un subconjunto $K \subset \mathbb{R}^n$ es un cuerpo convexo si es convexo y de núcleo no vacío.

Proposición 4.2. *Un subconjunto $K \subseteq X$ es un cuerpo convexo cuyo núcleo contiene al cero si, y sólo si, K es convexo, equilibrado y absorbente. En particular, el funcional de Minkowski ρ_K es una seminorma si, y sólo si, K es un cuerpo convexo con $0 \in N(K)$.*

Demostración. Demostremos la parte directa. Ya tenemos que K es convexo, y la absorbencia es consecuencia inmediata de la pertenencia de 0 al núcleo de K. Si $x \in K$, se satisface que $tx \in K$ para todo t tal que $|t| < \epsilon$, pues $0 \in N(K)$. Dado que $x \in K$, $\epsilon > 1$, así que en particular $tx \in X$ siempre que $|t| < 1$.

Examinemos ahora la veracidad de la afirmación recíproca. Como K es absorbente, para todo $x \in X$ existe $\lambda > 0$ tal que $\frac{x}{\lambda} \in K$. En particular, $0 = \frac{0}{\lambda} \in K$. Además, dado que K es equilibrado, $-\frac{x}{\lambda} \in K$. La convexidad de K asegura que para cualquier t tal que $|t| < \frac{1}{\lambda}$ se cumple que $tx \in K$. Esto prueba que $0 \in N(K)$. □

Ejemplo 4.1. Consideremos a la familia de espacios vectoriales normados ℓ_p^n sobre $\mathbb{R}$, que son justamente $\mathbb{R}^n$ dotados con la norma

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$$

para $1 \leq p < \infty$ y

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| .$$

Entonces las bolas unitarias de ℓ_p^n:

$$B_p^n = \left\{ x \in \mathbb{R}^n : \|x\|_p \leq 1 \right\},$$

son ejemplos de cuerpos convexos simétricos. En general, la bola unitaria cerrada de una norma en $\mathbb{R}^n$ es un cuerpo convexo simétrico.

Proposición 4.3. *Sean K y L dos subconjuntos de X tales que ρ_K y ρ_L son sus respectivos funcionales de Minkowski y además son normas de X. Entonces $K \subseteq L$ si, y sólo si, para toda x se cumple $\rho_K(x) \geq \rho_L(x)$.*

Necesidad. Para todo $\epsilon > 0$ existe λ tal que

$$\frac{1}{\lambda}x \in K \subseteq L$$

y

$$\rho_L(x) \leq \lambda < \rho_K(x) + \epsilon.$$

Haciendo tender ϵ a cero obtenemos el resultado. □

Suficiencia. Sea $x \in X$ tal que $\rho_K(x) \leq 1$. Supongamos que $x \notin K$. Si $\lambda_0 > 0$ es tal que $\frac{1}{\lambda_0}x \in K$ entonces, claramente, $\lambda_0 \neq 1$. Si ocurriera que $\lambda_0 < 1$, como K es equilibrado, se tendría que $\frac{\lambda_0}{\lambda_0}x = x \in K$, lo cual es una contradicción. Entonces $\lambda_0 > 1$. Luego $\rho_K(x) \geq 1$, lo que implica que $\rho_K(x) = 1$. Se deduce que para cada $1 > \epsilon > 0$ existe $\lambda > 0$ tal que $\frac{1}{\lambda}x \in K$ y

$$\lambda < 1 + \epsilon,$$

y como $\frac{\lambda}{1+\epsilon} < 1$, al ser K equilibrado, se tiene que $\frac{1}{1+\epsilon}x = \frac{\lambda}{1+\epsilon}\frac{1}{\lambda}x \in K$ y $\frac{\epsilon}{1+\epsilon}x \in K$. Puesto que K es convexo,

$$x = \frac{1}{1+\epsilon}x + \frac{\epsilon}{1+\epsilon}x \in K,$$

lo cual nuevamente es contradictorio. Se concluye que $\rho_K(x) \leq 1$ implica que $x \in K$. Por otra parte, si $x \in K$, entonces

$$\rho_K(x) \leq 1,$$

luego $x \in K$ si, y sólo si, $\rho_K(x) \leq 1$. Es más, si $x \in K$, se satisface

$$\rho_L(x) \leq \rho_K(x) \leq 1,$$

y se sigue que $x \in L$. □

4.2. El enunciado y algunas observaciones

Teorema 4.4 (Dvoretzky, 1960). *Para cada $\epsilon > 0$ existe una constante $c(\epsilon) > 0$ tal que para cada $n \in \mathbb{N}$ y cada cuerpo convexo simétrico $K \subset \mathbb{R}^n$ existe un subespacio $V \subset \mathbb{R}^n$ tal que:*

1. *Se satisface* $\dim V = k$, *donde* $k \geq c \log n$.
2. *El conjunto $V \cap K$ es "ϵ-euclidiano", lo que significa que existe $r > 0$ tal que*
$$V \cap r(1+\epsilon)^{-\frac{1}{2}}B_2^n \subset V \cap K \subset V \cap r(1+\epsilon)^{\frac{1}{2}}B_2^n.$$

Antes de dar la demostración (que es bastante larga), veamos algunos ejemplos que ilustran el teorema.

Ejemplo 4.2. Consideremos como cuerpo simétrico convexo a la bola unitaria en ℓ_∞^n (el cubo n-dimensional), la cual es lejana a la bola euclidiana. Es fácil ver que la razón de radios de bolas acotadas y acotantes para tal cubo es $\sqrt{n}$, es decir,

$$B_2^n \subset B_\infty^n \subset \sqrt{n} B_2^n.$$

Escolio 4.1. De acuerdo al teorema de Dvoretzky, podemos encontrar un subespacio de $\mathbb{R}^n$ de dimensión proporcional a $\log n$ en el cual la razón de bolas acotadas y acotantes para un cuerpo simétrico y convexo sea exactamente $1+\epsilon$.

Escolio 4.2. La constante c de la formulación del Teorema 4.4 depende de la calidad de la aproximación ϵ. Se sabe que existen constantes c_1 y c_2 tales que

$$c_1\epsilon^2 \leq c \leq c_2 \frac{1}{\log \frac{1}{\epsilon}}.$$

Claramente, existe un gran espacio entre las cotas inferiores y superiores y la dependencia exacta es todavía un importante problema abierto.

La caracterización de las normas de un espacio vectorial nos permite plantear el Teorema 4.4 en la siguiente forma equivalente.

Teorema 4.5. *Para cada $\epsilon > 0$ existe una constante $c(\epsilon) > 0$ tal que para cada $n \in \mathbb{N}$ y cada norma $\|\cdot\|$ en $\mathbb{R}^n$, existe un subespacio $V \subset \mathbb{R}^n$ tal que:*

1. *Se satisface $\dim V = k$, donde $k \geq c \log n$.*

2. *Existe $0 < M < \infty$ tal que para todo $x \in V$ se cumple*

$$M(1+\epsilon)^{-\frac{1}{2}} \|x\|_2 \leq \|x\| \leq (1+\epsilon)^{\frac{1}{2}} M \|x\|_2,$$

es decir, las normas $\|\cdot\|_2$ y $\|\cdot\|$ son equivalentes en V salvo por ϵ, donde $\|\cdot\|_2$ es la norma euclideana.

4.3. Demostración del Teorema de Dvoretzky

Lema 4.1. *Para cada $0 < \epsilon < 1$ existe una ϵ-red N de S^{n-1} de cardinalidad no mayor que $\left(1+\frac{2}{\epsilon}\right)^n$.*

Demostración. Denotemos por B_2 a la bola unitaria euclidiana de $\mathbb{R}^n$. Sea $N = \{x_i\}_{i=1}^m$ un subconjunto maximal de S^{n-1} tal que para toda $x, y \in N$ se cumple que $\|x-y\|_2 \geq \epsilon$. La maximalidad de N implica que es una ϵ-red en S^{n-1}. Supongamos que N no fuera una ϵ-red. Entonces existe $u \in S^{n-1}$ tal que para todo $x \in N$, $\|x-u\|_2 \geq \epsilon$. Pero entonces $N_1 = \{x_i\}_{i=1}^m \cup \{u\}$ es tal que para todo $x, y \in N_1$ se cumple que $\|x-y\|_2 \geq \epsilon$, y además $N \subset N_1$, lo cual contradice la maximalidad de η.

Los elementos de $\left\{B(x_i, \frac{\epsilon}{2})\right\}_{i=1}^m$ son disjuntos entre sí y tal conjunto está completamente contenido en $\left(1+\frac{\epsilon}{2}\right)B_2$. Por un lado

$$\begin{aligned}
\mathrm{Vol}\left(\bigcup B\left(x_i, \frac{\epsilon}{2}\right)\right) &= \sum_{i=1}^m \mathrm{Vol}\left(B\left(x_i, \frac{\epsilon}{2}\right)\right) \\
&= m\mathrm{Vol}\left(B\left(x_i, \frac{\epsilon}{2}\right)\right) \\
&= m\mathrm{Vol}\left(\frac{\epsilon}{2}B\left(x_i, 1\right)\right) \\
&= m\left(\frac{\epsilon}{2}\right)^n \mathrm{Vol}\left(B_2\right),
\end{aligned}$$

pero

$$\mathrm{Vol}\left(\bigcup B\left(x_i, \frac{\epsilon}{2}\right)\right) \leq \mathrm{Vol}\left(\left(1+\frac{\epsilon}{2}\right)B_2\right),$$

de donde

$$m \leq \left(\frac{1+\epsilon/2}{\epsilon/2}\right)^n = \left(1+\frac{2}{\epsilon}\right)^n. \qquad \square$$

Lema 4.2. *Supóngase que V es un espacio de Banach dimensionalmente finito dotado con las normas $|\cdot|$ y $\|\cdot\|$ que satisface*

$$(1-\delta) \leq \|x\| \leq (1+\delta)$$

para cada $x \in N$, donde N es una δ-red en $B_{|\cdot|} = \{x : |x| = 1\}$ para algún $0 < \delta < 1$. Entonces, para cada $x \in V$

$$\frac{1-3\delta}{1-\delta}|x| \leq \|x\| \leq \frac{1+\delta}{1-\delta}|x|.$$

Demostración. Sea $x \in B_{|\cdot|}$ y tómese $x_1 \in N$ tal que $|x - x_1| \leq \delta$. Luego escójase $x_2 \in N$ tal que

$$\left|\frac{x-x_1}{|x-x_1|} - x_2\right| \leq \delta,$$

y obtenemos

$$|x - x_1 - |x-x_1|\,x_2| \leq \delta\,|x-x_1| \leq \delta^2.$$

Repitiendo este procedimiento de aproximación podemos obtener una sucesión infinita $\{x_i\}$ y números reales $0 < \delta_i \leq \delta^i$ tal que para cada n

$$\left| x - \sum_{i=1}^{n} \delta_i x_i \right| \leq \delta^n.$$

Por lo tanto, $\sum_{i=1}^{n} \delta_i x_i$ converge a x tanto en $(V, |\cdot|)$ como en $(V, \|\cdot\|)$, pues cualesquiera dos normas en un espacio de Banach dimensionalmente finito son equivalentes (ver el Teorema 4.9). Más aún

$$\begin{aligned} \|x\| &\leq \sup_n \left\| \sum_{i=1}^{n} \delta_i x_i \right\| \\ &\leq \sum_{i=1}^{\infty} \delta_i \|x_i\| \\ &\leq (1+\delta) \sum_{i=1}^{\infty} \delta_i \leq \frac{1+\delta}{1-\delta}. \end{aligned}$$

Para la cota superior, sea $x \in B_{|\cdot|}$ y escójase $y \in N$ tal que $|x - y| \leq \delta$. Entonces

$$\begin{aligned} \|x\| &= \|y + x - y\| \\ &\geq \|y\| - \|x - y\| \\ &\geq 1 - \delta - \frac{1+\delta}{1-\delta}|x - y| \\ &\geq 1 - \delta - \delta\frac{1+\delta}{1-\delta} \\ &\geq \frac{1 - \delta - \delta + \delta^2 - \delta - \delta^2}{1-\delta} = \frac{1-3\delta}{1-\delta}. \quad \square \end{aligned}$$

Escolio 4.3. Sea $\epsilon > 0$. Si

$$\delta \leq \frac{\sqrt{1+\epsilon} - 1}{3\sqrt{1+\epsilon} - 1},$$

entonces se cumple que

$$\frac{1}{\sqrt{1+\epsilon}} \leq \frac{1-3\delta}{1-\delta} \quad \text{y} \quad \frac{1+\delta}{1-\delta} \leq \sqrt{1+\epsilon}.$$

Teorema 4.6. *Para cada $\epsilon > 0$ existe una constante $c(\epsilon) > 0$ tal que para cada $n \in \mathbb{N}$ y cada norma $\|\cdot\|$ en $\mathbb{R}^n$, existe un subespacio $V \subset \mathbb{R}^n$ tal que:*

1. *Se satisface* $\dim V = k$, *donde* $k = \lfloor c \left(\frac{\eta}{b}\right)^2 n \rfloor$.

2. Para cada $x \in V$

$$(1+\epsilon)^{-\frac{1}{2}} \|x\|_2 \le \|x\| \le (1+\epsilon)^{\frac{1}{2}} \|x\|_2 , \tag{4.6}$$

donde $\eta = \int_{S^{n-1}} \|x\| \, d\mu$ y b es la constante más pequeña tal que $\|x\| \le b \|x\|_2$.

Demostración. Sea $V_0 \subseteq \mathbb{R}^n$ un subespacio de dimensión k. Escribimos

$$S^{k-1} = S^{n-1} \cap V_0$$

y consideramos a N una δ-red de S^{k-1} de cardinalidad a lo más $\left(\frac{3}{\delta}\right)^k$, con $\delta < 1$, la cual sabemos que existe por el Lema 4.1. Fijemos $x_0 \in N$ y consideremos las funciones

$$\|\cdot\| : S^{n-1} \to \mathbb{R}$$

y

$$F : O(n) \to \mathbb{R}$$

definida esta última a través de $F(U) = \|Ux_0\|$. Primero: nótese que b es una constante de Lipschitz de $\|\cdot\|$:

$$|\|x\| - \|y\|| \le \|x-y\| \le b\|x-y\|_2.$$

Segundo: por el Escolio 2.1

$$\nu \{U \in O(n) : |F(U) - E_\nu(F)| > t\} = \mu \left\{x \in S^{n-1} : |\|x\| - E_\mu(\|\cdot\|)| > t\right\},$$

y según la ecuación (2.4),

$$\begin{aligned} \eta &= E_\mu(\|\cdot\|) = \int_{S^{n-1}} \|x\| \, d\mu(x) = \int_{O(n)} F(U) \, d\nu(U) \\ &= E_\nu(F). \end{aligned}$$

Podemos ahora aplicar el Teorema 2.7,

$$\nu \{U \in O(n) : |\|Ux_0\| - \eta| > t\} \le c \exp\left(-\frac{Ct^2 n}{b^2}\right),$$

y haciendo $t = \delta\eta$,

$$\nu \{U \in O(n) : |\|Ux_0\| - \eta| > \delta\eta\} \le c \exp\left[-Cn\left(\frac{\delta\eta}{b}\right)^2\right].$$

Esto puede repetirse para cada $x \in N$, lo que implica que

$$\begin{aligned} \nu \{U \in O(n) : |\|Ux\| - \eta| > \delta\eta\} &\leq c \exp\left[-Cn\left(\frac{\delta\eta}{b}\right)^2\right] |N| \\ &\leq c \exp\left[-Cn\left(\frac{\delta\eta}{b}\right)^2 + k \ln \frac{3}{\delta}\right]. \end{aligned}$$

Existe una constante T tal que

$$\ln(c) + k \ln\left(\frac{3}{\delta}\right) \leq Tk \ln\left(\frac{3}{\delta}\right),$$

así que podemos elegir el entero k de modo que

$$\frac{C\delta^2}{T \ln \frac{3}{\delta}} \left(\frac{\eta}{b}\right)^2 n > k = \left\lfloor \frac{C\delta^2}{T \ln \frac{3}{\delta}} \left(\frac{\eta}{b}\right)^2 n \right\rfloor.$$

Con este k podemos asegurar con probabilidad positiva que existe $U \in O(n)$ que satisface para todo $x \in N$

$$(1 - \delta)\eta \leq \|Ux\| \leq (1 + \delta)\eta.$$

Por lo tanto, de acuerdo al Lema 4.2, para todo $x \in V_0$

$$\left(\frac{1 - 3\delta}{1 - \delta}\right) \|x\|_2 \leq \|Ux\| \leq \left(\frac{1 + \delta}{1 - \delta}\right) \|x\|_2,$$

y eligiendo δ según el Escolio 4.3, obtenemos que se satisface la desigualdad (4.6). El subespacio buscado es $V = UV_0$. □

Podemos usar este teorema para calcular el valor k para los espacios ℓ_p^n, con $1 \leq p < \infty$. Antes necesitamos el siguiente resultado.

Proposición 4.4. *Sea $f : S^{n-1} \to \mathbb{R}$ una función integrable y $\hat{f}(t) = \|t\|_2 f\left(\frac{t}{\|t\|_2}\right)$. Se satisface*

$$\int_{S^{n-1}} f(t)\, d\mu(t) = \frac{\int_{\mathbb{R}^n} \hat{f}(t)\, d\nu(t)}{\int_{\mathbb{R}^n} \|t\|_2\, d\nu(t)},$$

donde μ es la medida de Haar sobre S^{n-1} y ν es la medida sobre $\mathbb{R}^n$ con densidad $\exp(-\pi \sum_{k=1}^n t_k^2)$.

Demostración. Se tiene que $\int_{\mathbb{R}^n} \hat{f}(t)\, d\nu(t)$ es invariante bajo la acción de $O(n)$, y por la unicidad de la medida de Haar sobre S^{n-1} se sigue que

$$\int_{\mathbb{R}^n} \hat{f}(t)\, d\nu(t) = \int_{\mathbb{R}^n} \|t\|_2 \hat{1}\left(\frac{t}{\|t\|_2}\right) d\nu(t) \int_{S^{n-1}} f(t)\, d\mu(t),$$

que basta trasponer para obtener lo requerido. □

Por lo desigualdad de Jensen,

$$\left(\sum_{i=1}^{n}|x_i|\right)^p \leq \sum_{i=1}^{n}|x_i|^p,$$

lo que implica que $\|x\|_1 \leq \|x\|_p$ para $p > 1$; adicionalmente, tenemos que

$$\begin{aligned}
\|x\|_p^p &= \sum_{i=1}^{n}|x_i|^p \\
&\leq \sum_{i=1}^{n} 1 \cdot |x_i|^p \\
&\leq \left(\sum_{i=1}^{n} 1^{\frac{2}{2-p}}\right)^{\frac{2-p}{2}} \left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{p}{2}} \\
&= n^{\frac{2-p}{2}} \|x\|_2^p ,
\end{aligned}$$

para $p \geq 1$, luego

$$\|x\|_p \leq n^{\frac{1}{p}-\frac{1}{2}} \|x\|_2 .$$

Ahora tenemos que evaluar η. Sea $x = (g_1, \ldots, g_n) \in \mathbb{R}^n$ un vector de variables gaussianas independientes. Sea $\overline{x} = \frac{x}{\|x\|_2}$ y tomando $f(t) = \|t\|_p$ en la Proposición 4.4 nos devuelve la identidad

$$E_\mu \|\overline{x}\|_p = \frac{E\|x\|_p}{E\|x\|_2}.$$

Para acotar esta última cantidad, usamos primeramente la desigualdad

$$E\left(\sum|g_i|^p\right)^{\frac{1}{p}} \leq \left(\sum E|g_i|^p\right)^{\frac{1}{p}} \leq \left(\sum(E|g_i|)^p\right)^{\frac{1}{p}} = kn^{\frac{1}{p}}$$

donde k es una constante (lo cual sabemos por el Escolio B.1) y segundamente

$$\left(E\left(\sum|g_i|^p\right)^{\frac{1}{p}}\right)^p \geq E\left(\sum|g_i|^p\right) = K(p)n$$

donde $K(p) = E(|g_i|^p)$ es una constante. Ambas desigualdades resultan de aplicar de la desigualdad de Jensen. Por lo tanto

$$\eta = E_\mu \|\overline{x}\|_p \geq c_p n^{\frac{1}{p}-\frac{1}{2}}.$$

Finalmente tenemos

$$k \geq \begin{cases} c_p(\epsilon)\, n^{\frac{2}{p}} & \text{si } p > 2, \\ c_p(\epsilon)\, n & \text{si } p = 1. \end{cases}$$

pues

$$\left(\frac{\eta}{b}\right)^2 n \geq \begin{cases} \left(\frac{c_p n^{\frac{1}{p}-\frac{1}{2}}}{1}\right)^2 n = c_p^2 n^{\frac{2}{p}} & \text{si } p > 2, \\ \left(\frac{c_p n^{\frac{1}{2}}}{n^{\frac{1}{2}}}\right)^2 n = c_p^2 n & \text{si } p = 1. \end{cases}$$

4.3.1. Elipsoides y el lema de Dvoretzky y Rogers

Para completar la demostración del teorema de Dvoretzky, sólo nos falta estimar E y b. En lugar de probar el resultado para la bola euclidiana, lo haremos para un elipsoide. Veremos que será fácil intercambiarlos después.

Definición 4.5. Sea $\{x_i\}_{i=1}^n \subset \mathbb{R}^n$ un conjunto linealmente independiente. Un elipsoide $\mathcal{E} \in \mathbb{R}^n$ es el conjunto

$$\mathcal{E} = \left\{ \sum_{i=1}^n \alpha_i x_i : \sum_{i=1}^n \alpha_i^2 \leq 1 \right\}.$$

Sea $\{e_i\}_{i=1}^n$ la base canónica de $\mathbb{R}^n$. Definimos un automorfismo T de $\mathbb{R}^n$ a través de

$$Te_i = x_i,$$

que satisface

$$T\left(B_2^n\right) = \mathcal{E}.$$

Un elipsoide induce un producto interior

$$(x, y)_{\mathcal{E}} = \left(T^{-1}x, T^{-1}y\right).$$

Lema 4.3 (Dvoretzky y Rogers)**.** *Sea $\|\cdot\|$ alguna norma en $\mathbb{R}^n$ y denotemos su bola unitaria por K. Sea $\mathcal{E}$ el único elipsoide de máximo volumen inscrito en K (Lema C.2) y sea $|\cdot|$ la norma inducida por $\mathcal{E}$. Entonces existe $\{x_i\}_{i=1}^n \in \partial\mathcal{E}$ (la frontera de $\mathcal{E}$) ortonormal con respecto al producto interior $(\cdot,\cdot)_{\mathcal{E}}$ de modo tal que hay una constante universal δ que satisface*

$$\delta\left(1 - \frac{i-1}{n}\right) \leq \|x_i\| \leq 1$$

para toda $1 \leq i \leq n$.

Demostración. Seleccionemos primero un punto arbitrario $x_1 \in \partial\mathcal{E}$ de norma máxima. Ciertamente, $\|x_1\| = |x_1| = 1$. Supongamos que hemos escogido $\{x_k\}_{k=1}^{j-1}$

que son ortonormales con respecto a $\mathcal{E}$. Escójase x_j de norma máxima entre los $x \in \partial\mathcal{E}$ que son ortogonales a $\{x_k\}_{k=1}^{j-1}$. Nuestro elipsoide original es

$$\mathcal{E} = \left\{ \sum_{i=1}^{n} a_i x_i : \sum_{i=1}^{n} a_i^2 \leq 1 \right\}.$$

Definamos un nuevo elipsoide que está deformado apropiadamente

$$\overline{\mathcal{E}} = \left\{ \sum_{i=1}^{n} a_i x_i : \sum_{i=1}^{j-1} \frac{a_i^2}{a^2} + \sum_{i=j}^{n} \frac{a_i^2}{b^2} \leq 1 \right\}.$$

Supongamos que $\sum_{i=1}^{n} a_i x_i \in \overline{\mathcal{E}}$. Entonces, para cada $x \in \operatorname{gen}\{x_k\}_{k=j}^{n} \cap \partial\mathcal{E}$ tenemos que

$$\|x\| \leq \|x_j\|,$$

pues de lo contrario se contradice la elección de x_j, que es el vector de norma máxima perpendicular a los x_i con $1 < i < j$. Más aún, $\sum_{i=1}^{j-1} a_i x_i \in a\mathcal{E}$ y $\sum_{i=j}^{n} a_i x_i \in b\mathcal{E}$, por lo tanto

$$\begin{aligned} \left\| \sum_{i=1}^{j-1} a_i x_i \right\| &\leq a, \\ \left\| \sum_{i=j}^{n} a_i x_i \right\| &\leq \|x_j\| \, b, \end{aligned}$$

y de aquí que

$$\left\| \sum_{i=1}^{n} a_i x_i \right\| \leq \left\| \sum_{i=1}^{j-1} a_i x_i \right\| + \left\| \sum_{i=j}^{n} a_i x_i \right\| \leq a + \|x_j\| \, b.$$

Los volúmenes de $\mathcal{E}$ y $\overline{\mathcal{E}}$ satisfacen $\operatorname{Vol}\left(\overline{\mathcal{E}}\right) = a^{j-1} b^{n-j+1} \operatorname{Vol}(\mathcal{E})$. Si $a + \|x_j\| \, b \leq 1$, entonces $\overline{\mathcal{E}} \subseteq K$. En virtud de que $\mathcal{E}$ es el elipsoide de volumen máximo inscrito en K concluímos que para todo a, b, j tales que $a + \|x_j\| \, b \leq 1$, entonces $a^{j-1} b^{n-j+1} \leq 1$. Sustituyendo $b = \frac{1-a}{\|x_j\|}$ y $a = \frac{j-1}{n}$ se sigue que para cada $j \geq 2$

$$\|x_j\| \geq a^{\frac{j-1}{n-j+1}} (1-a) = \left(\frac{j-1}{n} \right)^{\frac{j-1}{n-j+1}} \left(1 - \frac{j-1}{n} \right) \geq e^{-1} \left(1 - \frac{j-1}{n} \right). \quad \square$$

Teorema 4.7. *Para cada $\epsilon > 0$ existe una constante absoluta $c(\epsilon)$ tal que si $\|\cdot\|$ es alguna norma de $\mathbb{R}^n$ entonces existe un subespacio $V \subseteq \mathbb{R}^n$ y un elipsoide $\overline{\mathcal{E}}$ que satisface lo siguiente.*

1. *Se cumple que* $\dim V = k$, *con* $k \geq c(\epsilon) \log n$.

2. *Para cada* $x \in V$:

$$(1+\epsilon)^{-\frac{1}{2}} \|x\|_{\mathcal{E}} \leq \|x\| \leq (1+\epsilon)^{\frac{1}{2}} \|x\|_{\mathcal{E}} .$$

Demostración. Sea $\mathcal{E}$ un elipsoide de máximo volumen inscrito en $K = B_{\|\cdot\|}$ y T un automorfismo lineal de $\mathbb{R}^n$ tal que $T(B_2^n) = \mathcal{E}$. Denotemos $H = T^{-1}(K)$. Entonces la bola euclidiana B_2^n es el elipsoide de máximo volumen inscrito en H. Por lo tanto basta probar el teorema para la norma $\|\cdot\|_H$.

Por el Lema 4.3 existe un conjunto ortogonal de vectores $\{x_i\}_{i=1}^n \subset S^{n-1}$, y además existe una constante δ' que para todo $1 \leq i \leq \lfloor \frac{n}{2} \rfloor$ satisface

$$\|x_i\|_H \geq \delta'.$$

Por lo tanto, por la ley del paralelogramo y la ortogonalidad de $\{x_i\}_{i=1}^n$ respecto a $\|\cdot\|_H$

$$\begin{aligned}
\int_{S^{n-1}} \|x\|_H \, d\mu &= \int_{S^{n-1}} \left\| \sum_{i=1}^n a_i x_i \right\|_H d\mu \\
&= \int_{S^{n-1}} \frac{1}{2} \left(\left\| \sum_{i=1}^{n-1} a_i x_i + a_n x_n \right\|_H + \left\| \sum_{i=1}^{n-1} a_i x_i - a_n x_n \right\|_H \right)_H d\mu \\
&\geq \int_{S^{n-1}} \text{máx} \left\{ \left\| \sum_{i=1}^{n-1} a_i x_i \right\|_H , \|a_n x_n\|_H \right\} d\mu \\
&\geq \int_{S^{n-1}} \text{máx} \left\{ \left\| \sum_{i=1}^{n-2} a_i x_i \right\|_H , \|a_{n-1} x_{n-1}\|_H , \|a_n x_n\|_H \right\} d\mu \\
&\vdots \\
&\geq \int_{S^{n-1}} \max_{1 \leq i \leq \lfloor \frac{n}{2} \rfloor} \{\|a_i x_i\|_H\} \, d\mu \geq \delta' \int_{S^{n-1}} \max_{1 \leq i \leq \lfloor \frac{n}{2} \rfloor} \{|a_i|\} \, d\mu.
\end{aligned}$$

Por lo tanto, la última integral es exactamente la esperanza del vector con las coordenadas distribuidas normal e independientemente, normalizada según la norma en ℓ_2:

$$\begin{aligned}
\int_{S^{n-1}} \max_{1 \leq i \leq n} \{|a_i|\} \, da &= E \frac{\text{máx}_{1 \leq i \leq \lfloor \frac{n}{2} \rfloor} \{|g_i|\}}{\left(\sum_{i=1}^n g_i^2\right)^{\frac{1}{2}}} \\
&= \frac{E \, \text{máx}_{1 \leq i \leq \lfloor \frac{n}{2} \rfloor} \{|g_i|\}}{E \left(\sum_{i=1}^n g_i^2\right)^{\frac{1}{2}}}.
\end{aligned}$$

Para acotar el denominador, obsérvese que por la concavidad de la función raíz y la desigualdad de Jensen

$$E\left(\sum_{i=1}^{n} g_i^2\right)^{\frac{1}{2}} \leq \left(E\sum_{i=1}^{n} g_i^2\right)^{\frac{1}{2}} = \sqrt{n}.$$

Ahora utilizamos las propiedades de la distribución normal para obtener

$$P\left(|g_i| > t\right) \geq c \exp\left(\frac{-t^2}{2}\right)$$

para alguna constante $c > 0$. Hagamos $m = \left\lfloor \frac{n}{2} \right\rfloor$ y

$$\begin{aligned} P\left(\max |g_i| > t\right) &= 1 - \prod_{i=1}^{m} P\left(|g_i| \leq t\right) \\ &= 1 - \prod_{i=1}^{m} \left(1 - P\left(|g_i| > t\right)\right) \\ &\geq 1 - \left(1 - c\exp\left(\frac{-t^2}{2}\right)\right)^m . \end{aligned}$$

Sustituyendo $t = \sqrt{2\log m}$ obtenemos

$$P\left(\max |g_i| > \sqrt{2\log m}\right) \geq 1 - \left(1 - \frac{c}{m}\right)^m \geq 1 - \exp\left(-c\right) \geq C,$$

para alguna constante absoluta C. Ahora aplicamos la desigualdad de Chebyshev para terminar la estimación

$$E \max_{1\leq i\leq \left\lfloor \frac{n}{2}\right\rfloor} \{|g_i|\} \geq \sqrt{2\log m} P\left(\max |g_i| > \sqrt{2\log m}\right) \geq C\sqrt{\log n}.$$

Ahora tenemos la desigualdad

$$\eta = \int_{S^{n-1}} \|x\|_H \, dx \geq C\sqrt{\frac{\log n}{n}}.$$

Para completar la demostración, haciendo $b = 1$ (porque $B_2^n \subseteq H$) en el Teorema 4.6 tenemos que existe un subespacio $V \subseteq \mathbb{R}^n$ de dimensión no menor que $C\log n$ que satisface para cada $x \in V$

$$(1+\epsilon)^{-\frac{1}{2}} \|x\|_2 \leq \|x\|_H \leq (1+\epsilon)^{\frac{1}{2}} \|x\|_2 .$$

El elipsoide buscado es simplemente $\mathcal{E} = T(B_2^n)$. $\square$

Usando la Proposición 4.3 podemos hacer el siguiente replanteamiento.

Teorema 4.8. *Para cada* $\epsilon > 0$ *existe una constante absoluta* $c(\epsilon)$ *tal que para cualquier cuerpo convexo simétrico de* $\mathbb{R}^n$ *existe un subespacio* $V \subseteq \mathbb{R}^n$ *y un elipsoide* $\mathcal{E}$ *que satisface lo siguiente.*

1. *Se cumple que* $\dim V = k$, *con* $k \geq c(\epsilon) \log n$.

2. *Una sección de* K *en* V *es "casi" un elipsoide*

$$(1+\epsilon)^{-\frac{1}{2}}\mathcal{E} \cap V \subseteq K \cap V \subseteq (1+\epsilon)^{\frac{1}{2}}\mathcal{E} \cap V.$$

La siguiente proposición nos conducirá directamente al Teorema 4.4, también por la Proposición 4.3.

Definición 4.6. Sea H un espacio con producto interior y M un subconjunto de H. Definimos

$$M^{\perp} = \{y \in H : (x,y) = 0, \forall x \in M\}.$$

Proposición 4.5. *Sea* $\mathcal{E}$ *un elipsoide en* $\mathbb{R}^k$. *Entonces existe un subespacio* $W \subseteq \mathbb{R}^k$ *de dimensión* $m = \left\lfloor \frac{k}{4} \right\rfloor$ *y* $0 < r < \infty$ *tal que*

$$r\left(B_2^k \cap W\right) = \mathcal{E} \cap W.$$

Demostración. Primero encontraremos $\left\lfloor \frac{k}{4} \right\rfloor$ vectores que son ortogonales con respecto al producto interno usual $(\cdot,\cdot)$ y también respecto al producto $(\cdot,\cdot)_{\mathcal{E}}$ inducido por $\mathcal{E}$. Sea E algún subespacio de $\mathbb{R}^n$, denotaremos por $E^{\perp}$ su complemento ortogonal con respecto a $(\cdot,\cdot)$ y por $E_{\mathcal{E}}^{\perp}$ su complemento ortogonal con respecto al producto $(\cdot,\cdot)_{\mathcal{E}}$.

Escojamos algún $v_1 \in \mathbb{R}^n$ tal que $\|v_1\|_2 = 1$. Nótese que la dimensión del subespacio

$$(\operatorname{gen}\{v_1\})^{\perp} \cap (\operatorname{gen}\{v_1\})_{\mathcal{E}}^{\perp}$$

es al menos $n-2$, pues cada $(\operatorname{gen}\{v_1\})^{\perp}$ tiene dimensión $n-1$. Por lo tanto, podemos elegir un vector v_2 que satisface $\|v_2\|_2 = 1$ y es ortogonal v_1 con respecto a ambos productos interiores. Podemos repetir este proceso $\left\lfloor \frac{k}{2} \right\rfloor$ veces para obtener un conjunto de vectores $\{v_i\}_{i=1}^{\left\lfloor \frac{k}{2} \right\rfloor}$ que es ortogonal respecto a ambos productos. A partir de esta familia vamos a construir otra de $\left\lfloor \frac{k}{4} \right\rfloor$ elementos todos de la misma longitud y ortogonal respecto a ambos productos.

Ordenamos los vectores respecto a las normas

$$\|v_1\|_{\mathcal{E}} \geq \|v_2\|_{\mathcal{E}} \geq \cdots > \left\|v_{\left\lfloor \frac{k}{2} \right\rfloor}\right\|_{\mathcal{E}} > 0,$$

y escogemos algún número a tal que

$$\left\|v_{\lfloor \frac{k}{4} \rfloor}\right\|_{\mathcal{E}} \geq a \geq \left\|v_{\lfloor \frac{k}{4} \rfloor + 1}\right\|_{\mathcal{E}}.$$

Ahora para cada $1 \leq i \leq \lfloor \frac{k}{4} \rfloor$ escogemos $0 < \lambda_i \leq 1$ de tal suerte que

$$\lambda_i \|v_i\|_{\mathcal{E}}^2 + (1 - \lambda_i) \left\|v_{\lfloor \frac{k}{2} \rfloor - i + 1}\right\|_{\mathcal{E}}^2 \leq a^2,$$

y construimos una sucesión de $\lfloor \frac{k}{4} \rfloor$ vectores

$$u_i = \sqrt{\lambda_i} v_i + \sqrt{1 - \lambda_i} v_{\lfloor \frac{k}{2} \rfloor - i + 1}.$$

De la construcción es fácil calcular que la nueva familia es ortogonal respecto a ambos productos. Más aún

$$\|u_i\|_2^2 = \lambda_i \|v_i\|_2^2 + (1 - \lambda_i) \left\|v_{\lfloor \frac{k}{2} \rfloor - i + 1}\right\|_2^2 = 1$$

y

$$\|u_i\|_{\mathcal{E}}^2 = \lambda_i \|v_i\|_{\mathcal{E}}^2 + (1 - \lambda_i) \left\|v_{\lfloor \frac{k}{2} \rfloor - i + 1}\right\|_{\mathcal{E}}^2 = a^2.$$

Haciendo $W = \operatorname{gen}\{u_i\}_{i=1}^{\lfloor \frac{k}{4} \rfloor}$ y

$$\begin{aligned} B_2^n \cap W &= \left\{ \sum_{i=1}^{\lfloor \frac{k}{4} \rfloor} a_i u_i : \left\| \sum_{i=1}^{\lfloor \frac{k}{4} \rfloor} a_i u_i \right\| = \left(\sum_{i=1}^{\lfloor \frac{k}{4} \rfloor} a_i^2 \right)^{\frac{1}{2}} \leq 1 \right\} \\ &= \left\{ \sum_{i=1}^{\lfloor \frac{k}{4} \rfloor} a_i u_i : \left\| \sum_{i=1}^{\lfloor \frac{k}{4} \rfloor} a_i u_i \right\|_{\mathcal{E}} = a \left(\sum_{i=1}^{\lfloor \frac{k}{4} \rfloor} a_i^2 \right)^{\frac{1}{2}} \leq a \right\} \\ &= a\mathcal{E} \cap W, \end{aligned}$$

y vemos que tomando $r = \frac{1}{a}$ se obtiene el resultado. □

4.4. Una observación final

Hemos hecho todo un argüende de la demostración del Teorema de Dvoretzky; sin embargo, todavía hay que enunciar una forma equivalente (ahora sí, es la última) que necesitaremos.

Definición 4.7. Sean E y F espacios de Banach. Si E y F son isomorfos, definimos su distancia de Banach-Mazur como

$$d(E,F) = \text{ínf}\left\{\|T\|\left\|T^{-1}\right\|\right\},$$

donde el ínfimo recorre todos los isomorfismos $T : E \to F$. Si E y F no son isomorfos, hacemos $d(E,F) = \infty$.

Escolio 4.4. La distancia de Banach-Mazur satisface una desigualdad del triángulo multiplicativa, es decir, si E es isomorfo a F, y F es isomorfo a G, tenemos

$$d(E,G) \leq d(E,F)\, d(F,G).$$

Nótese que

$$1 = \|I\| = \left\|T \circ T^{-1}\right\| \leq \|T\|\left\|T^{-1}\right\|, \tag{4.7}$$

por lo cual $d(E,F) \geq 1$.

Hemos hablado profusamente de los espacios de Banach. Sin embargo, resulta que todos los espacios de Banach de dimensión n son esencialmente idénticos.

Teorema 4.9. *Si X es un espacio vectorial de dimensión n, entonces cualesquiera dos normas en X son equivalentes.*

Demostración. Sea $|\cdot|$ cualquier norma en X y sea $\{\xi_i\}_{i=1}^n$ una base de Hamel para X. Sea $\{e_i\}_{i=1}^n$ la base canónica $\mathbb{K}^n$. Si para $x = \sum_{i=1}^n a_i e_i$ definimos

$$|x|_n = \left|\sum_{i=1}^n a_i \xi_i\right|,$$

es claro que $|\cdot|_n$ define una norma en $\mathbb{K}^n$, y por ello podemos reducir la demostración del teorema considerar solamente a $\mathbb{K}^n$.

Consideremos entonces a $\mathbb{K}^n$ con la norma usual $\|\cdot\|$ y sea $|\cdot|'$ cualquier otra norma en $\mathbb{K}^n$. Procederemos por inducción sobre n. Para $n = 1$, tomando

$$M_1 = \text{máx}\left\{|e_1|', \frac{1}{|e_1|'}\right\},$$

se cumple que $|e_1|' \leq M_1$, por lo que si $x = a_1 e_1$

$$\begin{aligned} |x|' &= |a_1 e_1|' \\ &= |a_i|\,|e_i|' \leq |a_i|\, M_1 = \|x\|\, M_1 \end{aligned}$$

e igualmente $|e_1|' \leq \frac{1}{M_1}$, y obtenemos $|x|' \geq \frac{1}{M_1}\|x\|$. Supongamos que la afirmación es verdadera para todo $m \leq n$. Sea

$$J = \{\chi = (j_1, \ldots, j_{m+1}) : j_i \in \mathbb{N}, 1 \leq j_1 \leq \cdots \leq j_{m+1} \leq n\},$$

entonces para cualquier $(j_1,\dots,j_m)$ el espacio generado por $\{e_{j_i}\}_{i=1}^m$ es cerrado en la norma $|\cdot|'$, por que es cerrado en la norma $\|\cdot\|$. Sea ahora

$$Y^{(\chi)} = \left\{ \sum_{i=1}^{m+1} a_i e_{j_i} \right\} \subset \left(\mathbb{K}^n, |\cdot|'\right),$$

y consideremos las $r = 1,\dots,m$ funcionales $P_{j_r}^{(\chi)} : Y^{(\chi)} \to \mathbb{K}^n$ dadas por

$$P_{j_r}^{(\chi)} \left(\sum_{i=1}^{m+1} a_i e_{j_i} \right) = a_r.$$

Es claro que

$$\text{núc}\, P_{j_r}^{(\chi)} = \text{gen} \{e_{j_i}\}_{i \neq r} \subset Y^{(\chi)},$$

y este subespacio es cerrado. Por lo tanto $P_{j_r}^{(\chi)}$ es continua (ver [FG97, pp. 81-82]) para $r = 1,\dots,m+1$. Por un lado, si $x = \sum_{i=1}^{m+1} a_i e_j$,

$$\|x\| = \left\| \sum_{i=1}^{m+1} \left(P_{j_r}^{(\chi)} x\right) e_{j_i} \right\| \leq \sum_{i=1}^{m+1} \left\| P_{j_r}^{(\chi)} \right\| |x|' \leq K_1 \, |x|',$$

donde

$$K_1 = \text{máx} \left\{ \sum_{i=1}^{m+1} \left\| P_{j_r}^{(\chi)} \right\| : \chi = (j_1,\dots,j_{m+1}) \in J \right\}.$$

Este número existe porque J es un conjunto finito. Por otro lado tenemos

$$\begin{aligned} |x|' &\leq \sum_{i=1}^{m+1} |a_i| \, |e_{j_i}|' \leq \left(\sum_{i=1}^{m+1} |a_i|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{m+1} \left(|e_{j_i}|'\right)^2 \right)^{\frac{1}{2}} \\ &\leq \left(\sum_{i=1}^{m+1} |a_i|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} \left(|e_{j_i}|'\right)^2 \right)^{\frac{1}{2}} \leq K_2 \, \|x\|, \end{aligned}$$

donde $K_2 = \left(\sum_{i=1}^n \left(|e_{j_i}|'\right)^2 \right)^{\frac{1}{2}}$, de manera que si $M_{n+1} = \text{máx}\{K_1, K_2\}$ entonces

$$\frac{1}{M_{n+1}} \|x\| \leq |x|' \leq M_{n+1} \|x\|. \qquad \square$$

Corolario 4.1. *Sea* $(X, \|\cdot\|_X)$ *un espacio normado de dimensión* n*, y sea* $\{\xi_i\}_{i=1}^n$ *una base de Hamel de* X*. Entonces, si* $\{e_i\}_{i=1}^n$ *es la base canónica de* $\mathbb{K}^n$*, el operador* T *dado por*

$$T\left(\sum_{i=1}^n a_i \xi_i \right) = \sum_{i=1}^n a_i e_i$$

es un isomorfismo. Es decir, todos los espacios de Banach de dimensión n son isomorfos entre sí, y a $\mathbb{K}^n$ en particular.

Escolio 4.5. Haremos una manipulación de las desigualdades establecidas por el Teorema 4.5: para cada $x \in V$ y cada isomorfismo $T : \ell_2^k \to V$,

$$\|T\| \leq (1+\epsilon)^{\frac{1}{2}} M$$

y usando (4.7)

$$\|T^{-1}\|_2 \leq (1+\epsilon)^{\frac{1}{2}} \frac{1}{M},$$

luego $\|T\|\|T^{-1}\|_2 \leq (1+\epsilon)$.

Teorema 4.10. *Para cada $\epsilon > 0$ y cada natural k, hay un número $n(\epsilon, k)$ con la siguiente propiedad: existe un espacio de Banach Y de dimensión por lo menos $n(\epsilon, k)$ que contiene un subespacio X de dimensión k que satisface $d(X, \ell_2^k) \leq 1+\epsilon$.*

Demostración. Sea $\epsilon > 0$ y $c(\epsilon)$ la constante del Teorema 4.5. Hagamos

$$n(\epsilon, k) = \lfloor \exp(k - c(\epsilon)) \rfloor,$$

y sea $Y = \mathbb{R}^n$ con una norma arbitraria $\| \cdot \|$. Ahora existe un subespacio X de Y de dimensión k tal que

$$M(1+\epsilon)^{-\frac{1}{2}} \|x\|_2 \leq \|x\| \leq (1+\epsilon)^{\frac{1}{2}} M \|x\|_2 ,$$

para alguna constante $0 < M < \infty$. Que $d(X, \ell_2^k) \leq 1 + \epsilon$ es consecuencia del Escolio 4.5. $\square$

A la luz de estos hechos y teniendo en cuenta los Teoremas 4.9 y 4.10, podemos dar otra equivalencia (¡la última!) del teorema estrella de este capítulo.

Teorema 4.11 (Dvoretzky, 1961)**.** *Sea X un espacio de Banach dimensionalmente infinito. Entonces, para cada entero n y cada $\epsilon > 0$ existe un subespacio n-dimensional C de X tal que*

$$d\left(C, \ell_2^n\right) \leq 1 + \epsilon.$$

Capítulo 5

Los Teoremas de Joichi y Davis-Dean-Singer

> La estructura del espacio de Hilbert es bien conocida, en parte porque cada subespacio está complementado.
>
> *W. Davis, D. Dean y I. Singer*, [DDS68]

Pienso que con el capítulo anterior ya llevamos la mitad del camino. Los siguientes resultados son un poco más sencillos, y es inmediata su conexión con el problema de los espacios complementarios. A decir verdad, no he dicho todavía de qué se trata ese problema. Pronto lo veremos.

5.1. El Teorema de Joichi

En esta sección seguiremos al artículo original [Joi66].

5.1.1. Redes y semirretículos

Definición 5.1. Un conjunto A parcialmente ordenado (es decir, que tiene una relación $\prec$ reflexiva, antisimétrica y transitiva) se denomina conjunto dirigido si para cualesquiera $\alpha, \beta \in A$ existe un $\gamma \in A$ tal que $\alpha \prec \gamma$ y $\beta \prec \gamma$. Tal elemento γ se denomina cota superior.

Ejemplo 5.1. Para cualquier conjunto X, su conjunto potencia está parcialmente ordenado bajo la relación de inclusión. Además, es un conjunto dirigido, pues para cualquier $A, B \subseteq X$, el conjunto $A \cup B$ pertenece al conjunto potencia y $A \cup B \supseteq A$ y $A \cup B \supseteq B$.

Sea A un conjunto dirigido y $\alpha, \beta, \gamma \in A$. Sea γ una cota superior de α y β. Si para cualquier otra cota superior γ' de α y β se satisface $\gamma \prec \gamma'$, entonces se dice que γ es el supremo de α y β, y se denota $\gamma = \alpha \vee \beta$.

Definición 5.2. Un conjunto dirigido donde todo par de elementos tiene un supremo es un semirretículo.

Ejemplo 5.2. El conjunto dirigido del ejemplo anterior también es un semirretículo, pues si C es otro conjunto tal que $A \subseteq C$ y $B \subseteq C$, entonces $A \cup B \subseteq C$.

Definición 5.3. Sea X un conjunto. Una red en X es un subconjunto $\{x_\alpha\}_{\alpha \in A} \subseteq X$ indizado por un conjunto parcialmente ordenado A. Si (X, d) es un espacio métrico se dice que la red converge a $a \in X$ si para toda $\epsilon > 0$ existe un $\alpha_0 \in A$ tal que $\alpha_0 \prec \alpha$ implica que

$$d(x_\alpha, a) \leq \epsilon,$$

esto lo denotamos como $\lim_A x_\alpha = a$. Diremos que una red es de Cauchy si para cada $\epsilon > 0$ existe un $\alpha_0 \in A$ tal que $\alpha_0 \succ \alpha, \beta$ trae como consecuencia que

$$d(x_\alpha, x_\beta) \leq \epsilon.$$

5.1.2. El Teorema de Hahn-Banach

Es difícil exagerar la importancia de cierto resultado del análisis funcional: el teorema de extensión de Hahn-Banach. Sobre todo cuando resulta de utilidad para nuestro objetivo.

Teorema 5.1 (Hahn-Banach)**.** *Sea X un espacio vectorial real y $p(x)$ una función de valores reales definida en X y que satisface las siguientes condiciones*

$$p(x+y) \leq p(x) + p(y) \tag{5.1}$$
$$p(\alpha x) = \alpha p(x) \tag{5.2}$$

para $x, y \in X$ y α real no negativa. Sea M un subespacio vectorial de X y f_0 un funcional lineal de valores reales definido en M. Supongamos que f_0 satisface $f_0(x) \leq p(x)$ para $x \in M$. Entonces existe un funcional lineal de valores reales F definido en X que satisface lo siguiente.

1. *Extiende a f, es decir, $F(x) = f_0(x)$ para todo $x \in M$.*

2. *Cumple la desigualdad $F(x) \leq p(x)$ sobre X.*

Demostración. Supongamos primero que X está generado por M y otro elemento $x_0 \notin M$, esto es,

$$X = \{x = m + \alpha x_0 : m \in M, \alpha \in \mathbb{R}\}.$$

La representación de cada $x \in X$ como $x = m + \alpha x_0$ es única. De no ser así, y

$$x = m_1 + \alpha_1 x_0 = m_2 + \alpha_2 x_0$$

entonces $x_0 = \frac{m_2 - m_1}{\alpha_1 - \alpha_2} \in M$, lo cual es contradictorio. Se sigue que, para todo número real c, si

$$F(x) = F(m + \alpha x_0) = f_0(m) + \alpha c,$$

entonces F es un funcional sobre X que extiende a f_0, pues si $x \in M$, entonces $\alpha = 0$. Tenemos que escoger c de tal modo que $F(x) \leq p(x)$, es decir

$$f_0(m) + \alpha c \leq p(m + \alpha x_0).$$

Tal condición es equivalente a otras dos:

$$\begin{aligned} f_0\left(\frac{m}{\alpha}\right) + c &\leq p\left(x_0 + \frac{m}{\alpha}\right), \alpha > 0, \\ f_0\left(\frac{m}{-\alpha}\right) - c &\leq p\left(-x_0 + \frac{m}{-\alpha}\right), \alpha < 0. \end{aligned}$$

Para satisfacer dichas condiciones, debemos escoger c tal que

$$f_0(m') - p(m' - x_0) \leq c \leq p(m'' + x_0) - f_0(m'')$$

para toda $m', m'' \in M$. Tal elección de c es posible, pues

$$\begin{aligned} f_0(m') + f_0(m'') &= f_0(m' + m'') \leq p(m' + m'') \\ &\leq p(m' - x_0 + m'' + x_0) \\ &\leq p(m' - x_0) + p(m'' + x_0), \end{aligned}$$

lo que quiere decir que

$$f_0(m') - p(m' - x_0) \leq p(m'' + x_0) - f_0(m'').$$

El número c debe elegirse entonces en el intervalo

$$\left[\sup_{m' \in M} \{f_0(m') - p(m' - x_0)\}, \inf_{m'' \in M} \{p(m'' + x_0) - f_0(m'')\}\right].$$

Consideremos ahora a la familia de todas las extensiones reales g de f_0, para las cuales se cumple $g(x) \leq p(x)$ en el dominio de g. Podemos dotar a esta familia de un orden parcial definiendo $h \succ g$ si, y sólo si, h extiende a g. Por el lema de Zorn existe una extensión lineal maximal F para la cual la desigualdad $F(x) \leq p(x)$ se satisface en todo el dominio de F. De hecho, el dominio de F es X. De lo contrario existiría, por lo demostrado previamente, una extensión H de F propia que satisface la desigualdad $H(x) \leq p(x)$ en el todo el dominio de H, lo que contradice la elección de F. □

Corolario 5.1. *Dado un funcional $p(x)$ definido sobre un espacio vectorial real X que satisface* (5.1) *y* (5.2), *existe un funcional lineal f definido sobre X tal que*

$$-p(-x) \leq f(x) \leq p(x).$$

Demostración. Tómese cualquier punto $x_0 \in X$ y defínase

$$M = \{x : x = \alpha x_0, \alpha \in \mathbb{R}\}.$$

El conjunto M es un subespacio, pues si x y y pertenecen a M, entonces

$$\begin{aligned} \gamma x + \beta y &= \gamma \alpha x_0 + \beta \alpha x_0 \\ &= (\gamma \alpha + \beta \alpha) x_0 \in M. \end{aligned}$$

Hágase $f_0(\alpha x_0) = \alpha p(x_0)$. Entonces f_0 es un funcional real lineal definido sobre M, pues

$$\begin{aligned} f_0(\beta_1 x + \beta_2 y) &= f_0(\beta_1 \alpha_1 x_0 + \beta_2 \alpha_2 x_0) \\ &= f_0((\beta_1 \alpha_1 + \beta_2 \alpha_2) x_0) \\ &= (\beta_1 \alpha_1 + \beta_2 \alpha_2) p(x_0) \\ &= \beta_1 \alpha_1 p(x_0) + \beta_2 \alpha_2 p(x_0) \\ &= \beta_1 f_0(x) + \beta_2 f_0(x). \end{aligned}$$

Se tiene así que $\alpha p(x_0) = p(\alpha x_0)$ si $\alpha > 0$ y

$$\alpha p(x_0) \leq -\alpha p(-x_0) = p(\alpha x_0)$$

para $\alpha < 0$, pues

$$0 = p(0) \leq p(x_0) + p(-x_0).$$

Por lo tanto, existe un funcional f definido sobre X tal que $f(x) = f_0(x)$ sobre M y $f(x) \leq p(x)$ sobre X. Puesto que

$$-f(x) = f(-x) \leq p(-x),$$

obtenemos el funcional deseado. □

Recordemos que el dual de un espacio vectorial X sobre un campo $\mathbb{K}$ es el conjunto de todas las transformaciones lineales de X en $\mathbb{K}$, y se denota por X^*. Si X es normado, sólo se consideran en el dual a las transformaciones continuas. El siguiente teorema, cuya demostración recuerda a la del teorema de existencia de la medida de Haar, aparece en [Roy63].

Teorema 5.2 (Woodbury). *Bajo las hipótesis del Teorema 5.1, sea G un semigrupo conmutativo de operadores lineales sobre X tales que para cada $g \in G$ y para todo $x \in X$ se tiene $p(gx) \leq p(x)$. Supóngase también que f_0 es invariante bajo la acción de G, es decir, $f_0(gx) = f_0(x)$. Entonces existe una extensión F de f_0 que satisface la conclusión del Teorema 5.1 y que además también es invariante bajo la acción del grupo.*

Demostración. Sea q un funcional sobre X definido a través de

$$q(x) = \inf_{n \in \mathbb{N}, \{g_i\}_{i=1}^n \in G} \left\{ \frac{1}{n} p\left(\sum_{i=1}^n g_i x \right) \right\}.$$

Claramente, $q(x) \leq p(x)$ y $q(\alpha x) = \alpha q(x)$ para $\alpha \geq 0$. Para cualquier $x, y \in X$ y $\epsilon > 0$, escojamos g_i y h_j de modo que

$$\begin{aligned} \frac{1}{n} p\left(\sum_{i=1}^n g_i x \right) &< q(x) + \frac{\epsilon}{2}, \\ \frac{1}{m} p\left(\sum_{i=1}^m h_i x \right) &< q(y) + \frac{\epsilon}{2}. \end{aligned}$$

Entonces

$$\begin{aligned} q(x+y) &\leq \frac{1}{nm} p\left(\sum_{i=1}^n \sum_{j=1}^m g_i h_j (x+y) \right) \\ &\leq \frac{1}{nm} \left[p\left(\sum_{i=1}^n g_i \sum_{j=1}^m h_j x \right) + p\left(\sum_{i=1}^n g_i \sum_{j=1}^m h_j y \right) \right] \\ &\leq \frac{1}{n} p\left(\sum_{i=1}^n g_i x \right) + \frac{1}{m} p\left(\sum_{j=1}^m h_j y \right) \\ &\leq q(x) + q(y) + \epsilon, \end{aligned}$$

y la arbitrariedad de ϵ permite concluir que $q(x+y) \leq q(x) + q(y)$. Para $x \in M$, se cumple

$$\begin{aligned} f_0(x) &= \frac{1}{n} f_0\left(\sum_{i=1}^{n} g_i x\right) \\ &\leq \frac{1}{n} p\left(\sum_{i=1}^{n} g_i x\right) \\ &< q(x) + \frac{\epsilon}{2}, \end{aligned}$$

de donde $f_0(x) \leq q(x)$. Podemos ahora aplicar el Teorema 5.1 para obtener una extensión F de f_0 de modo que $F(x) \leq q(x) \leq p(x)$. Resta probar que F es invariante ante la acción de G. Puesto que

$$\begin{aligned} q(x - gx) &\leq \frac{1}{n} p\left(\sum_{i=0}^{n-1} g^i (x - gx)\right) \\ &\leq \frac{1}{n} p\left(x - g^{n+1} x\right) \\ &\leq \frac{1}{n}\left[p(x) + p(-x)\right], \end{aligned}$$

haciendo tender $n \to \infty$, vemos que $q(x - gx) \leq 0$. De

$$F(x) - F(gx) = F(x - gx) \leq q(x - gx) \leq 0,$$

se sigue que $F(x) \leq F(gx)$. Aplicando esto a $-x$, obtenemos que $F(x) = F(gx)$. $\square$

Teorema 5.3 (Banach). *Sea S un semirretículo, y $B(S)$ el espacio de las funciones reales acotadas definidas sobre S con la norma supremo. Entonces $B(S)$ es un espacio de Banach con*

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) \\ (\lambda f)(x) &= \lambda f(x), \end{aligned}$$

para todo escalar λ y $x \in S$. Además, podemos definir un funcional lineal acotado Φ en $B(S)$ que satisface la condición

$$\sup_{\alpha \in S} \inf_{\beta \succ \alpha} f(\beta) \leq \Phi(f(\alpha)) \leq \inf_{\alpha \in S} \sup_{\beta \succ \alpha} f(\beta),$$

y $\Phi(f)(\alpha) = \lim_{\alpha \in S} f(\alpha)$ si este último límite existe.

Demostración. Es claro que $B(S)$ es un espacio vectorial. El espacio $B(S)$ es un espacio de Banach con la norma supremo. Para verlo, sea $\{f_i\}_{i=1}^{\infty}$ una sucesión de Cauchy en $B(S)$. Esto significa que para cada $\epsilon > 0$ existe N tal que si $m > n > N$ entonces

$$|f_m(x) - f_n(x)| \leq \|f_m - f_n\|_\infty \leq \epsilon \tag{5.3}$$

para todo $x \in S$. Luego las sucesiones $\{f_i(x)\}_{i=1}^{\infty}$ son sucesiones reales de Cauchy para cada $x \in S$, y por lo tanto convergen a $f(x)$.

Cada f_i está acotada. Es decir, existe $M_i > 0$ tal que $|f_i(x)| \leq M_i$. Según la continuidad del valor absoluto y la desigualdad (5.3), para todo $x \in S$ vale decir

$$|f(x)| \leq \epsilon + M_i,$$

luego f es acotada. Como S es un conjunto dirigido, cada $f \in B(S)$ es una red en $\mathbb{R}$.

Sea $\lambda = \operatorname{ínf}_{\alpha \in S} \sup\{f(\beta) : \beta \succ \alpha\}$. Para cada $\epsilon > 0$ existe α_0 tal que

$$\lambda - \epsilon < \sup\{f(\beta) : \beta \succ \alpha_0\} < \lambda + \epsilon.$$

Más aún, si $\alpha \succ \alpha_0$,

$$\lambda - \epsilon < \sup\{f(\beta) : \beta \succ \alpha\} \leq \sup\{f(\beta) : \beta \succ \alpha_0\} < \lambda + \epsilon,$$

que significa precisamente que

$$\lim_{s \in S} \sup\{f(\beta) : \beta \succ \alpha\} = \lambda = \inf_{\alpha \in S} \sup_{\beta \succ \alpha} f(\beta).$$

Sabiendo esto podemos definir

$$p(f) = \inf_{\alpha \in S} \sup_{\beta \succ \alpha} f(\beta). \tag{5.4}$$

Claramente p así definido satisface (5.1) y (5.2). Por lo tanto, por el Corolario 5.1, existe un funcional F definido sobre $B(S)$ tal que

$$-p(-f) \leq \Phi(f) \leq p(f)$$

en $B(S)$, pues es evidente que $-p(-f) = \sup_{\alpha \in S} \operatorname{ínf}_{\beta \succ \alpha} f(\beta)$. □

El subconjunto $K(S)$ de $B(S)$ que consta de las funciones que son redes convergentes es cerrado en $B(S)$. Sea $\{f_i\}_{i=1}^{\infty}$ una sucesión convergente de elementos de $K(S)$. Probemos que $f = \operatorname{lím}_{i \to \infty} f_i$ es una red convergente.

Sea $a_i = \operatorname{lím}_{\alpha \in S} f_i(\alpha)$. Para $\epsilon > 0$ existen $N \in \mathbb{N}$ y $\alpha_0 \in S$ tal que si $m > n \geq N$ se cumple

$$|f_m(x) - f_n(x)| \leq \frac{\epsilon}{3}$$

para todo $x \in S$. Escojamos, pues, $m > n \geq N$ y sea $\beta \succ \alpha_0$ tal que

$$|f_i(\beta) - a_i| \leq \frac{\epsilon}{3}, \quad i = m, n,$$

entonces

$$|a_n - a_m| \leq |a_n - f_n(\beta)| + |f_n(\beta) - f_m(\beta)| + |f_m(\beta) - a_m| \leq \epsilon.$$

Habiendo comprobado que $\{a_i\}_{i=1}^{\infty}$ es una sucesión de Cauchy, tenemos que converge a un número a. Para $\epsilon > 0$ arbitrario, existe N de modo que simultáneamente

$$\|f - f_N\|_\infty \leq \frac{\epsilon}{3} \quad \text{y} \quad |a_N - a| \leq \frac{\epsilon}{3}.$$

Elijamos α_0 de tal suerte que para todo $\beta \succ \alpha_0$ se satisface

$$|f_N(\beta) - a_N| \leq \frac{\epsilon}{3},$$

como consecuencia obtenemos la desigualdad

$$|f(\beta) - a| \leq |f(\beta) - f_N(\beta)| + |f_N(\beta) - a_N| + |a_N - a| \leq \epsilon,$$

que es equivalente a

$$\lim_{\alpha \in S} f(\alpha) = a.$$

Para $f \in K(S)$, sea

$$\phi(f) = \lim_S f;$$

así, ϕ es un funcional lineal acotado en $K(S)$.

Para cada $s \in S$, sea L_s el operador lineal en $B(S)$ definido a través de

$$(L_s f)(t) = f(s \vee t).$$

Vemos lo siguiente.

1. El conjunto $\{L_s : s \in S\}$ es un semigrupo conmutativo. En efecto, si $s, u \in S$ entonces

$$\begin{aligned}
(L_s L_u f)(t) &= f(s \vee (u \vee t)) \\
&= f((s \vee u) \vee t) \\
&= f((u \vee s) \vee t) \\
&= f(u \vee (s \vee t)) \\
&= (L_u L_s f)(t).
\end{aligned}$$

2. La imagen bajo L_s de un elemento en $K(S)$ es de nuevo un elemento de $K(S)$. Supongamos que $f \in K(S)$ es una red convergente. Entonces para todo $\epsilon > 0$ existe $x_0 \in S$ tal que si $x \succ x_0$, entonces
$$|f(x) - \phi(f)| \leq \epsilon.$$
En particular, $s \vee x \succ x \succ x_0$, luego
$$|(L_s f)(x) - \phi(f)| \leq \epsilon,$$
lo que implica que $L_s f$ converge también.

3. Se satisface $\phi(L_s f) = \phi(f)$ para cada $s \in S$ y $f \in K(S)$, lo cual es consecuencia del punto anterior.

4. Se cumple $p(L_u f) = p(f)$. Para verlo, definamos
$$F(f, s) = \sup\{f(t) : t \succ s\};$$
claramente
$$F(L_u f, s \vee u) = F(f, s \vee u).$$
Sea $\psi = \lim_S F(f, s)$. Dado $\epsilon > 0$ existe s_0 tal que si $s \succ s_0$ entonces $|\psi - F(f, s)| < \epsilon$. En particular $s \vee u \succ s$, así que $|\psi - F(f, s \vee u)| < \epsilon$, de donde se sigue el aserto.

Por los Teoremas 5.2 y 5.3, existe una extensión Φ de ϕ a $B(S)$ con las siguientes propiedades.

1. El funcional Φ es invariante bajo la acción de L_g, es decir, $\Phi(L_g f) = \Phi(f)$.

2. Además, considerando el funcional p definido por (5.4) y haciendo $q(f) = -p(-f) \leq \Phi(f)$, entonces $q(f) \leq \Phi(f) \leq p(f)$ para cada $f \in B(S)$.

Teorema 5.4 (Joichi, 1965). *Un espacio de Banach X es un espacio de Hilbert si, y solamente si, existe una constante $k \geq 1$ tal que para cada subespacio dimensionalmente finito M de X existe un mapeo lineal T_M de M en un espacio de Hilbert H tal que $\left(\frac{1}{k}\right)|x| \leq |T_M x| \leq k|x|$ para cada $x \in M$.*

Demostración. La necesidad es consecuencia del hecho de que M también es un espacio de Hilbert y la función identidad sobre M satisfaría la conclusión del teorema. Probemos entonces la suficiencia. Sea S el conjunto de los subespacios dimensionalmente finitos de X ordenados por inclusión como subespacio vectorial. Entonces S es un semirretículo donde
$$M \vee N = M + N,$$

en el sentido de la suma vectorial, para $M, N \in S$. Por lo tanto, según la discusión anterior, existe un funcional Φ en $B(S)^*$ que satisface las tres condiciones antes mencionadas.

Para $x \in X$, defínase f_x en $B(S)$ a través de

$$f_x(M) = \begin{cases} |T_M x| & \text{si } x \in M, \\ 0 & \text{si } x \notin M. \end{cases}$$

Ahora, hagamos

$$n(x) = \left|\Phi\left(f_x^2\right)\right|^{\frac{1}{2}},$$

vamos a probar que n es una norma procedente de un producto interior en X equivalente a la norma original.

Es claro que n es no negativa. Además

$$\begin{aligned} n(\alpha x) &= \left|\Phi\left(f_{\alpha x}^2\right)\right|^{\frac{1}{2}} \\ &= \left||\alpha|^2 \Phi\left(f_x^2\right)\right|^{\frac{1}{2}} \\ &= |\alpha| \left|\Phi\left(f_x^2\right)\right|^{\frac{1}{2}} = |\alpha|\, n(x), \end{aligned}$$

para todo escalar α. También

$$n(x)^2 = \Phi\left(f_x^2\right) \begin{cases} \leq p\left(f_x^2\right) = \left(\operatorname{lím\,sup}_S f_x\right)^2 \leq (k|x|)^2, \\ \geq q\left(f_x^2\right) = \left(\operatorname{lím\,inf}_S f_x\right)^2 \geq \left(\frac{1}{k}|x|\right)^2, \end{cases}$$

por lo que

$$\frac{1}{k}|x| \leq n(x) \leq k|x|. \tag{5.5}$$

De aquí se deduce fácilmente que $n(x) = 0$ si, y sólo si, $x = 0$. Esto comprueba que si n es una norma, es equivalente a la original. Falta probar que n satisface la desigualdad del triángulo. Para tal efecto, demostraremos primero que n satisface (1.1), la ley del paraleogramo. Eso, de paso, mostrará que n debe venir de un producto interior. Para $x, y \in X$, sea $N = \operatorname{gen}\{x, y\}$. Entonces, para cualquier subespacio tal que $M \succ N$, tenemos

$$\begin{aligned} 2\left[f_x^2(M) + f_y^2(M)\right] &= 2\left[|T_M x|^2 + |T_M y|^2\right] \\ &= |T_M x - T_M y|^2 + |T_M x + T_M y|^2 \\ &= f_{x-y}^2(M) + f_{x+y}^2(M), \end{aligned}$$

donde la segunda igualdad es consecuencia de que $T_M(M)$ es un subespacio de un espacio de Hilbert, donde la norma satisface (1.1). Aplicando Φ a la identidad obtenida, resulta

$$2\left[n(x)^2 + n(y)^2\right] = n(x-y)^2 + n(x+y)^2,$$

que no es otra cosa que decir que n satisface (1.1).

Para concluir la prueba, observamos que n es el funcional de Minkowski de $B = \{x : n(x) \leq 1\}$. Como $n(-x) = n(-1x) = |-1|\, n(x) = n(x)$, B es equilibrada. Es absorbente, porque para cualquier $x \in X$ distinto de 0 (el 0 está contenido en B por $n(0) = 0n(x) = 0 \leq 1$),

$$n\left(\frac{1}{k\,|x|}x\right) \leq \frac{1}{k\,|x|}k\,|x| = 1,$$

lo que significa que $\frac{1}{k|x|}x \in B$. Sólo falta probar que B es convexo para comprobar que n es una norma. Tenemos que n es continua, según la desigualdad (5.5). Supongamos que $x, y \in X$ son tales que $n(x) = 1$ y $n(y) = 1$ pero existe un escalar $\alpha \in (0, 1)$ tal que

$$n(\alpha x + (1 - \alpha) x) > 1.$$

Entonces, por la continuidad de n, existen escalares α_1 y α_2 tal que

$$0 \leq \alpha_1 < \alpha < \alpha_2 \leq 1,$$

se cumple $n(\alpha_i x + (1 - \alpha_i) x) > 1$ para $i = 1, 2$ y siempre que $\beta \in (\alpha_1, \alpha_2)$ entonces

$$n(\beta x + (1 - \beta)) > 1.$$

Sea

$$x_i = \alpha_i x + (1 - \alpha_i) x$$

para $i = 1, 2$. Por la ley del paralelogramo,

$$\begin{aligned} 4 &= 2\left[n(x_1)^2 + n(x_2)^2\right] \\ &= n(x_1 + x_2)^2 + n(x_1 - x_2)^2 \\ &\geq n(x_1 + x_2)^2 = 4n\left(\left(\frac{1}{2}\right)(x_1 + x_2)\right)^2 \\ &> 4, \end{aligned}$$

lo que es imposible, luego B es convexa y n es una norma. □

5.2. El Teorema de Davis-Dean-Singer

En esta sección sabremos que significa que dos subespacios de un espacio de Banach sean complementarios.

5.2.1. Subespacios de codimensión finita

Lema 5.1. *Cada subespacio unidimensional X de un espacio de Banach Y es cerrado en Y.*

Demostración. Sea $\{x_\alpha\}_{\alpha\in A}$ una red en X tal que existe $x \in Y$ para el cual $\lim_A x_\alpha = x$. Entonces $\{x_\alpha\}$ es una red de Cauchy en X. Si $z \neq 0$ está en X, el mapeo uniformemente continuo $t \mapsto tz$ entre X y $\mathbb{R}$ lleva a $\{x_\alpha\}$ a la red de Cauchy $\{t_\alpha\} \subset \mathbb{R}$. Como $\mathbb{R}$ es completo, $\{t_\alpha\}$ converge a algún límite t. Entonces

$$tz = \lim_A t_\alpha z = \lim_A x_\alpha = x,$$

por lo que $x \in X$. □

Corolario 5.2. *Sean Y un espacio de Banach y X un subespacio cerrado del mismo. Sea Z un subespacio unidimensional tal que $Z\cap X = \{0\}$ y sea $W = X+Z$. Entonces*

1. *El espacio W es cerrado en Y.*
2. *Si $0 \neq z \in Z$, la correspondencia natural $(x, r) \leftrightarrow x + rz$ entre $X \times \mathbb{R}$ y W es un homeomorfismo.*

Demostración. En W definamos f a través de

$$f(w) = f(x + rz) = r$$

donde $r \in \mathbb{R}$ y $y \in X$. Entonces f es lineal y $f^{-1}(0) = X$, el cual es cerrado en W pues es cerrado en Y. Tenemos que f es continua, luego para $w = y + rx$ se tiene que y y r son funciones continuas de z. Por lo tanto, la función F definida por $F(y+rx) = (y, r)$ es continua de W sobre $X\times\mathbb{R}$. La continuidad de las operaciones vectoriales asegura que F^{-1} es continua. Por lo tanto F es un homeomorfismo.

Si $\{w_\alpha\}_{\alpha\in A}$ es una red en W y $\lim_A w_\alpha = w \in Y$, entonces $w_\alpha = y_\alpha + r_\alpha x$, $y_\alpha \in X$, así que

$$w_\alpha - w_\beta = (y_\alpha - y_\beta) + (r_\alpha - r_\beta)x \to 0.$$

Por la continuidad de f, $r_\alpha - r_\beta \to 0$, y al ser $\mathbb{R}$ completo se sigue que existe r tal que $\lim_A r_\alpha = r$. Entonces $\lim_A r_\alpha x = rx$, y $\lim_\alpha y_\alpha = w - rx$. Pero X es cerrado, así que $y = w - rx \in X$, luego $w = y + rx \in W$, por lo que W es cerrado en Y. □

Teorema 5.5. *Si X y Z son subespacios de un espacio de Banach Y tal que X es cerrado y Z es de dimensión finita, entonces $X + Z$ es cerrado en Y. Si también $X \cap Z = \{0\}$, entonces $X + Z$ tiene la topología de $X \times Z$; es decir, cualquier isomorfismo algebraico entre estos espacios es bicontinuo.*

Demostración. El teorema para el caso en que $\dim Z = 1$ es el corolario anterior. Para $\dim Z = n+1$, sabemos que Z tiene un subespacio n dimensional Z'. Por la hipótesis de inducción resultan las conclusiones del teorema para X y Z'. Aplicamos el corolario anterior sobre $X + Z'$ y $\text{gen}\{z_0\}$ (donde $z_0 \notin Z'$) para terminar la demostración por inducción. □

Proposición 5.1. *Supongamos que X es un subespacio de un espacio vectorial Y. Son equivalentes lo siguientes enunciados.*

1. *Se cumple* $\dim Y/X = n < \infty$.
2. *Se satisface $Y = X \oplus Z$ donde Z es un subespacio de Y de dimensión n.*

Demostración. Supongamos que $\dim Y/X = n$ y $U = Y/X$. Entonces existe un base para U de la forma

$$\mathcal{B} = \{z_1 + X, \dots, z_n + X\}.$$

Sean $Z = \text{gen}\{z_i\}_{i=1}^n$, $y \in Y$ arbitrario y $u + X$ su clase de equivalencia en Y/X. Claramente, $y = u + x$ para algún $x \in X$ y $u \in Z$. Al tomar $w \in Z \cap X$, podemos escoger escalares $\{\alpha_i\}_{i=1}^n$ tales que $w = \sum_{i=1}^n \alpha_i z_i = x$ para algún $x \in X$. Es válido escribir

$$\sum_{i=1}^n \alpha_i (z_i + X) = \left(\sum_{i=1}^n \alpha_i z_i\right) + X = x + X = 0 + X,$$

de donde deducimos que $\alpha_i = 0$ para todo $1 \leq i \leq n$ (por ser $\mathcal{B}$ una base), y con ello que $w = 0$. En resumen, $Y = X \oplus Z$ y Z es un espacio vectorial n dimensional.

Recíprocamente, supongamos que $Y = X \oplus Z$ con Z un subespacio de Y de dimensión n. Sea $\mathcal{C} = \{z_i\}_{i=1}^n$ una base para Z. Demostraremos que $\{z_i + X\}_{i=1}^n$ es una base para Y/X. Para cualquier $y \in Y$, existen $x \in X$ y $z \in Z$ de modo que $y = x + z$. Entonces $y \in z + X$, y además $z = \sum_{i=1}^n \alpha_i z_i$, por lo que

$$z + X = \left(\sum_{i=1}^n \alpha_i z_i\right) + X = \sum_{i=1}^n \alpha_i \left(z_i + X\right).$$

El conjunto $\mathcal{B} = \{z_i + X\}_{i=1}^n$ es una base, pues si

$$\sum_{i=1}^n \alpha_i (z_i + X) = \left(\sum_{i=1}^n \alpha_i z_i\right) + X = 0,$$

entonces $\sum_{i=1}^n \alpha_i z_i = x$ para algún $x \in X$. Dada la hipótesis de suma directa, $x = 0$. Por ser $\mathcal{C}$ una base, $\alpha_i = 0$ para $1 \leq i \leq n$. Siendo $\mathcal{B}$ un subconjunto generador y linealmente independiente de Y/X, es una base, y por lo tanto $\dim Y/X = n$. □

Definición 5.4. Sea Y un espacio vectorial. Si X es un subespacio de Y que satisface cualquiera de las dos equivalencias de la Proposición 5.1 se dice que es de codimensión finita n.

5.2.2. Constantes de expansión y de proyección

Definición 5.5. Un operador lineal P sobre un espacio vectorial V es una proyección si $P^2 = P$.

Proposición 5.2. *Sea Y un subespacio cerrado de un espacio de Banach X. Existe un subespacio cerrado Z tal que $X = Y + Z$ si, y sólo si, existe una proyección acotada P de X sobre Y.*

Demostración. Demostremos la necesidad. Puesto que cada $x \in X$ se puede escribir de modo único como $x = y + z$ para $y \in Y$ y $z \in Z$, podemos definir $Px = P(y+z) = y$. Claramente $P^2 = P$ y además $Pz = 0$ para todo $z \in Z$. Sea $(x_n, Px_n) \subset X \times X$ una sucesión convergente a (x, y). Como Y es cerrado, $y \in Y$. Ya que P es una proyección, $Py = y$. Más aún, $(I - P)x_n = x_n - Px_n \in Z$ para cada n, pues $P(I-P)x_n = Px_n - P^2x_n = 0$. Pero $x_n - Px_n \to x - y$, y dado que Z es cerrado, $x - y \in Z$. Por lo tanto $P(x-y) = 0$, y de aquí que $Px = Py = y$, luego la gráfica de P es cerrada. Por el Teorema C.4, se sigue que P es continua, y por el Escolio C.1, es acotada.

Recíprocamente, si existe una proyección acotada de P sobre X, entonces es continua según el Escolio C.1. También $I - P$ es una proyección continua, pues

$$(I-P)^2 = I - 2P + P^2 = I - 2P + P = I - P.$$

Puesto que $Py = y$ para todo $y \in Y$, entonces $Y \subseteq \text{núc}(I-P)$. Sea $Z = \text{núc}(I-P)$. Cualquier $y \in X \cap Z$ satisface $y = Py = 0$, luego $X \cap Y = \{0\}$. Además, para cualquier $x \in X$

$$x = Px + (I-P)x.$$

Por estas razones $X = Y + Z$. Dado que cualquier punto en un espacio de Banach es cerrado y P es continua, $Z = P^{-1}(0)$ es cerrado. □

Definición 5.6. Para cualquier espacio métrico X definimos su constante de expansión $E(X)$ como el ínfimo de los números μ con la siguiente propiedad: para cualquier familia $\{B(x_\alpha, \rho_\alpha)\}_{\alpha \in A} \subset X$ de bolas cerradas que se intersectan a pares, se cumple que

$$\bigcap_{\alpha \in A} B(x_a, \mu\rho_\alpha) \neq \emptyset.$$

La constante de expansión para el caso de espacios de Minkowski fue estudiada primeramente por Grünbaum en su tesis doctoral.

Definición 5.7. Para cualquier espacio normado X la n-constante de proyección $p_n(X)$ es el ínfimo de los números μ con la siguiente propiedad: si Y es un espacio normado que contiene a X como subespacio de codimensión n (esto es,

$$Y = X \oplus \text{gen}\{z_k\}_{k=1}^n$$

con $z_k \notin X$ para cada $1 \leq k \leq n$ y $\{z_k\}_{k=1}^n$ es linealmente independiente) entonces existe una proyección $P : Y \to X$ tal que $\|P\| \leq \mu$.

Escolio 5.1. Sean X y Y son espacios de Banach. Si X es cerrado y de codimensión finita n en Y, por el Teorema 5.2 existe una proyección de $P : Y \to X$ que es acotada. En tal caso, $p_n(X) < \infty$.

Teorema 5.6 (Grünbaum, 1959)**.** *Para cualquier espacio de Banach X, se cumple que $E(X) = p_1(X)$.*

Demostración. Sea

$$Y = X \oplus \mathrm{gen}\{y_0\}$$

con $y_0 \notin X$. Dada cualquier familia $\{B(x_\alpha, \rho_\alpha)\}$ de bolas cerradas que se intersectan a pares, definiremos una norma tal que X sea un subespacio cerrado de Y y que para cualquier proyección continua $P : Y \to X$ tengamos

$$\bigcap_{\alpha \in A} B(x_\alpha, \|P\|\rho_\alpha) \neq \emptyset. \tag{5.6}$$

Si $\inf \rho_\alpha = 0$, cualquier norma sobre Y establece (5.6). En efecto, si $\rho_\beta = 0$ para algún $\beta \in A$, entonces $x_\beta \in B(x_\alpha, \rho_\alpha)$ para todo $\alpha \in A$. Por otro lado, si para alguna sucesión de índices α_n tenemos que $\lim \rho_{\alpha_n} = 0$, entonces $\{x_{\alpha_n}\}$ es una sucesión de Cauchy. Como X es completo existe $x_0 = \lim_{n\to\infty} x_{\alpha_n}$; ahora para cualesquiera $\alpha \in A$ y $\epsilon > 0$ existe n tal que $\rho_{\alpha_n} < \epsilon/2$ y $\|x_{\alpha_n} - x_0\| < \epsilon/2$, lo que implica que

$$\|x_\alpha - x_0\| \leq \|x_\alpha - x_{\alpha_n}\| + \|x_{\alpha_n} - x_0\| < \rho_\alpha + \epsilon$$

que junto con la arbitrariedad del ϵ demuestran que $x_0 \in B(x_\alpha, \rho_\alpha)$ para toda $\alpha \in A$.

Si $\inf \rho_\alpha > 0$, entonces definamos

$$y_\alpha = \frac{x_\alpha + y_0}{\rho_\alpha}$$

y $K = \{y_\alpha\}_{\alpha \in A}$. Sea S la bola unitaria en X y T la cerradura (en la topología producto $Y = X \times \mathbb{R}$) de la envolvente convexa del conjunto $S \cup K \cup (-K) \subset Y$. El conjunto T es equilibrado, pues $0 = 1 \cdot 0 + 0 \cdot y_\alpha \in T$ y, por ser T convexo y $-K \subseteq T$, si $y \in T$ y $|t| \leq 1$, $ty = (1-t)0 + ty \in T$.

A fin de demostrar que T es absorbente, sea en particular $y = ry_0$. Como $(1/r)y = y_0$, basta considerar el caso para y_0. Tomemos $\alpha \in A$ arbitrario. Por un lado, si $-x_\alpha \in S$,

$$\frac{1}{1+\rho_\alpha} y_0 = \frac{\rho_\alpha}{1+\rho_\alpha} y_\alpha + \frac{1}{1+\rho_\alpha}(-x_\alpha) \in T,$$

y por otro, si $-x_\alpha \notin S$, $\|x_\alpha\|_S > 1$, y

$$\frac{1}{\|x_\alpha\|_S(1+\rho_\alpha)}y_0 = \frac{\rho_\alpha}{\|x_\alpha\|_S(1+\rho_\alpha)}y_\alpha + \frac{1}{\|x_\alpha\|_S(1+\rho_\alpha)}(-x_\alpha) \in T.$$

En todos los casos, existe λ tal que $\frac{1}{r\lambda}y \in T$. En general, sea $y = x + ry_0$. Por ser T equilibrado, sin perder generalidad podemos suponer que $r > 0$. Existen λ_x y λ_r tales que $(1/\lambda_x)x \in S$ y $(1/\lambda_r)(ry_0) \in T$. Haciendo $\lambda = \text{máx}(\lambda_x, \lambda_r)$, resulta que

$$\frac{1}{\lambda(r+1)}y = \frac{1}{\lambda(r+1)}x + \frac{r}{\lambda(r+1)}y_0 \in T.$$

Siendo T un cuerpo convexo, el funcional de Minkowski asociado al mismo es una norma. Mostraremos que la intersección de un segmento que conecta a un punto de K con otro de $-K$ y X está en S, lo que implicará que X es subespacio de Y. Un simple cálculo revela que

$$[y_\alpha, -y_\beta] \cap X = \frac{x_\alpha - x_\beta}{\rho_\alpha + \rho_\beta} = u,$$

y además

$$\|u\|_S = \left\| \frac{x_\alpha - x_\beta}{\rho_\alpha + \rho_\beta} \right\|_S = \left\| \frac{x_\alpha - x_\gamma + x_\gamma - x_\beta}{\rho_\alpha + \rho_\beta} \right\|_S \leq \frac{\rho_\alpha + \rho_\beta}{\rho_\alpha + \rho_\beta} = 1,$$

pues hemos supuesto que las bolas se intersectan a pares, lo que implica que $u \in S$. Ahora sea $P : Y \to X$ una proyección continua y $x_0 = -P(y_0)$. Entonces

$$P(y_\alpha) = P\left(\frac{x_\alpha + y_0}{\rho_\alpha}\right) = \frac{x_\alpha - x_0}{\rho_\alpha}$$

y en consecuencia

$$\left\| \frac{x_\alpha - x_0}{\rho_\alpha} \right\|_T = \|P(y_\alpha)\|_T \leq \|P\|_T \|y_\alpha\|_T \leq \|P\|_T$$

para todo $\alpha \in A$. Entonces $x_0 \in B(x_\alpha, \|P\|_T \rho_\alpha)$ para cada α, lo que implica (5.6) y establece que $E(X) \leq p_1(X)$.

Para concluir la prueba, sea Y cualquier espacio de Banach que tenga a X como mínimo subespacio, y sea $y_0 \in Y \setminus X$. La desigualdad del triángulo implica que

$$B_X(x, \|x - y_0\|) \cap B_X(x', \|x' - y_0\|) \neq \emptyset$$

para cualesquiera $x, x' \in X$. Sea $\mu = E(X)$. Se cumple que $\bigcap_{x \in X} B_X(x, \mu\|x - y_0\|) \neq \emptyset$, así que tomaremos cualquier punto x_0 de dicha intersección. Entonces

$\|x - x_0\| \leq \mu\|x - y_0\|$ para cualquier $x \in X$. Definimos la proyección $P : Y \to X$ a través de

$$P(x + \lambda y_0) = x + \lambda x_0.$$

Naturalmente, podemos suponer que $\lambda \neq 0$ y entonces, por la definición de x_0

$$\|P(x + \lambda y_0)\| = \|x + \lambda x_0\| = \lambda\|\tfrac{-x}{\lambda} - x_0\| \leq \lambda\mu\|\tfrac{-x}{\lambda} - y_0\| = \mu\|x + \lambda y_0\|,$$

o lo que es lo mismo, $\|Px\| \leq \mu\|x\|$, lo que implica que $\|P\| \leq \mu$ y que $p_1(X) \leq E(X)$. □

Proposición 5.3. *Se cumple que $E(X) \leq 2$ para cualquier espacio de Banach X.*

Demostración. Sea $B(x_\alpha, \rho_\alpha)_{\alpha \in A}$ una familia de bolas cuya intersección a pares no es vacía. Mientras demostrábamos el Teorema 5.6 vimos que si $\inf \rho_\alpha = 0$ entonces $\bigcap_{\alpha \in A} B(x_\alpha, \rho_\alpha) \neq \emptyset$.

Supongamos entonces que $\inf \rho_\alpha = \rho_0 > 0$. Dado $\epsilon > 0$, sea $\beta \in A$ tal que $\rho_\beta \leq (1 + \epsilon)\rho_0$. En virtud de la secuencia de desigualdades

$$\|x_\alpha - x_\beta\| \leq \rho_\alpha + \rho_\beta \leq \rho_\alpha + (1 + \epsilon)\rho_0 \leq (2 + \epsilon)\rho_\alpha$$

para cualquier $\alpha \in A$ se cumple que

$$x_\beta \in \bigcap_{\alpha \in A} B(x_\alpha, (2 + \epsilon)\rho_\alpha)$$

de donde se sigue la proposición. □

Corolario 5.3. *Para todo espacio de Banach Y, se satisface que $p_n(Y) \leq 2^n$*

Demostración. El resultado es válido para $n = 1$ según la Proposición 5.3 y el Teorema 5.6. Supongamos que sigue siendo verdadero para $n = k$. Sea X un subespacio de codimensión $k + 1$ en Y. Entonces $X' = X \oplus z$ con $z \notin X$ es un espacio de codimensión k en Y y a su vez X es de codimensión 1 en X'. Por la hipótesis de inducción, $p_k(Y) \leq 2^k$, así que existen proyecciones $P : Y \to X'$ y $P' : X' \to X$ continuas. Se sigue que $P'P : Y \to X$ es una proyección continua y

$$\|P'P\| \leq \|P'\|\|P\| \leq 2^k 2 = 2^{k+1},$$

lo que completa la inducción. □

Definición 5.8. Dos subespacios M y N cerrados de un espacio de Banach B tales que $M + N$ es denso en B y $M \cap N = \{0\}$ se dicen complementarios si $B = M \oplus N$. Si, dado un subespacio cerrado M de B, existe otro subespacio cerrado N tal que M y N son complementarios, decimos que M está complementado.

Escolio 5.2. Por el Teorema 5.2, M y N son complementarios si, y sólo si, M es cerrado y existe una proyección acotada $P : B \to M$.

Definición 5.9. Sea E un espacio de Banach y X un subespacio cerrado del mismo. Definimos la constante de proyección $\lambda(X)$

$$\lambda(X) = \text{ínf}\{\|P\| : P \text{ es una proyección de } E \text{ sobre } X\}$$

y $\lambda(X) = \infty$ si X no está complementado en E.

Definimos también

$$\lambda_f(X) = \sup\{\lambda(F) : F \text{ es un subespacio dimensionalmente finito de } X\}.$$

Lema 5.2. *Si E es suma directa de dos subespacios G y Y, G es dimensionalmente finito y $\lambda_f(Y)$ es finito, entonces $\lambda_f(E)$ es finito.*

Demostración. Si Z es un subespacio cerrado de E, $Z \subset X$ y se cumple que

$$\dim X/Z = N < \infty,$$

entonces existe una proyección Q_ϵ de X sobre Z con norma no mayor a $2^N + \epsilon$ para todo $\epsilon > 0$. Por lo tanto, si $P : E \to X$ es una proyección cualquiera, entonces

$$Q_\epsilon \circ P : E \to Z$$

y

$$\lambda(Z) \leq 2^N \lambda(X),$$

pues

$$\lambda(Z) \leq \|Q_\epsilon P\| \leq \|Q_\epsilon\| \|P\| \leq \left(2^N + \epsilon\right) \|P\|,$$

es decir

$$\frac{\lambda(Z)}{2^N + \epsilon} \leq \|P\|,$$

y de aquí que $\lambda(X) \geq \frac{\lambda(Z)}{2^N+\epsilon}$. Entonces $\lambda(Z) \leq \left(2^N + \epsilon\right) \lambda(X)$, y la arbitrariedad de ϵ implica lo afirmado. Por estas razones existe una proyección $P_\epsilon : E \to Y$ tal que $\|P_\epsilon\| < 2^{\dim G} + \epsilon$.

Sea F un subespacio de E dimensionalmente finito arbitrario tal que

$$F \subset F' := (I - P_\epsilon)(E) \oplus P_\epsilon(F).$$

Como $\dim P_\epsilon(F) \leq \dim F < \infty$, existe una proyección Q de Y sobre $P_\epsilon(F)$ con

$$\|Q\| \leq \lambda_f(Y) + \epsilon,$$

por lo cual

$$(I - P_\epsilon) + Q \circ P_\epsilon$$

es una proyección de E sobre F'. De aquí que

$$\begin{aligned} \lambda(F') &\leq \|(I - P_\epsilon) + Q \circ P_\epsilon\| \\ &\leq \|I\| + \|P_\epsilon\| + \|Q \circ P_\epsilon\| \\ &\leq 1 + 2^{\dim G}(1 + \lambda_f(Y)), \end{aligned}$$

y como $\dim F'/F \leq \dim G$, entonces

$$\lambda(F) \leq 2^{\dim G}\left[1 + 2^{\dim G}(1 + \lambda_f(Y))\right] < \infty,$$

luego $\lambda_f(E) < \infty$. □

Teorema 5.7 (Davis, Dean y Singer, 1968)**.** *Sea E un espacio de Banach tal que $\lambda_f(F) = \infty$. Entonces E tiene un subespacio no complementado, F. El subespacio F tiene descomposición de Schauder en subespacios dimensionalmente finitos.*

Demostración. Sea E_1 un subespacio de E dimensionalmente finito con $\lambda(E_1) \geq 1$. Escójanse

$$\left\{f_1^1, \cdots, f_{n_1}^1\right\} \subset E_1^*$$

todos unitarios (es decir, $\left\|f_j^1\right\| = 1$ para cada j), y sean $\{g_i\}_{i=1}^{n_1} \subset E^*$ extensiones de Hahn-Banach de los $\left\{f_1^1, \cdots, f_{n_1}^1\right\}$, escogidas de modo que

$$\left[\bigcap_{j=1}^{n_1} g_j^{-1}([-1,1])\right] \cap E_1$$

esté completamente contenido en la bola de radio 2 en E_1. Sea

$$Y_1 = \left\{\bigcap_{j=1}^{n_1} g_j^{-1}(0)\right\},$$

entonces $E_1 \cap Y_1 = \{0\}$ y la proyección natural de $E_1 \oplus Y_1$ sobre E_1 tiene norma no mayor que 2. Por el Lema 5.2 podemos escoger el subespacio $E_2 \subset Y_1$ tal que $\dim E_2 < \infty$ y $\lambda(E_2) < \infty$. Como antes, hallamos $\{g_i\}_{i=n_1+1}^{n_2}$ en E^* tal que $\|g_j\| = 1$ y

$$\left[\bigcap_{j=1}^{n_2} g_j^{-1}([-1,1])\right] \cap E_1 \oplus E_2$$

está completamente contenido en la bola de radio 2 en $E_1 \oplus E_2$. Hacemos

$$Y_2 = \left\{\bigcap_{j=1}^{n_2} g_j^{-1}(0)\right\},$$

y se observa que $Y_2 \subset Y_1$, y la dimensión del complemento de Y_2 en Y_1 es no mayor que n_2. Continuando este proceso, obtenemos $\{E_n\}$ y $\{Y_n\}$ tales $\lambda(E_n) \geq n$, $Y_{n+1} \subset Y_n$ y la proyección natural de $E_1 \oplus \cdots \oplus E_n \oplus Y_n$ sobre $E_1 \oplus \cdots \oplus E_n$ tiene norma no mayor que 2.

Ahora definimos

$$F = \bigoplus_{i=1}^{\infty} E_n = \left\{ \sum e_n : e_n \in E_n, \sum e_n \text{ converge en } E \right\}.$$

Puede verse que F es un subespacio de E y tiene descomposición de Schauder $\{E_n\}$. Sea P_n la proyección natural de F sobre $E_1 \oplus \cdots \oplus E_n$. Entonces si P es una proyección de E sobre F tenemos que

$$(I - P_{n-1}) P_n P$$

es una proyección de E sobre E_n y como

$$\begin{aligned} \|(I - P_{n-1}) P_n P\| &\leq \|P_n P\| + \|P_{n-1} P_n P\| \\ &\leq 2\|P\| + 2 \cdot 2\|P\| = 6\|P\|, \end{aligned}$$

se sigue que $\lambda(E_n) \leq 6\|P\|$ lo cual no puede ser, a menos que F no sea complementado. $\square$

Con este intrincado teorema terminamos este capítulo y nos vamos directo a nuestro destino.

Capítulo 6

El Teorema de Espacios Complementarios

[El Teorema de Espacios Complementarios] ... ha sido conjeturado por mucho tiempo por varios autores. La prueba que damos aquí es sorprendentemente simple.

J. Lindenstrauss y L. Tzafriri, [LT71]

Ha tomado tanto trabajo arribar a este punto que el lector debe estar suponiendo que lo que se va a demostrar a continuación debe valer mucho la pena. Quizá si, quizá no. La verdad es que uno piensa que un enunciado tan sencillo y fácil de entender debiera ser igualmente fácil de demostrar. Pues no, y las páginas anteriores son testimonio de ello.

Aún no podemos llegar. Primero presentamos una propiedad muy importante de los espacios con producto interior.

6.1. Mejor aproximación en espacios de Hilbert

Teorema 6.1. *Todo subconjunto E no vacío, convexo y cerrado de un espacio con producto interior H contiene un único elemento de norma mínima.*

Demostración. Sea $\delta = \inf\{\|x\| : x \in E\}$. Aplicando la ley del paralelogramo a $\frac{1}{2}x$ y $\frac{1}{2}y$ obtenemos

$$\frac{1}{4}\|x-y\|^2 + \left\|\frac{x+y}{2}\right\|^2 = \frac{1}{2}\|x\|^2 + \frac{1}{2}\|y\|^2 .$$

Como E es convexo, $\frac{x+y}{2} \in E$, se tiene que $\left\|\frac{x+y}{2}\right\|^2 \geq \delta^2$ y entonces

$$\|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 - 4\left\|\frac{x+y}{2}\right\|^2 \leq 2\|x\|^2 + 2\|y\|^2 - 4\delta^2. \tag{6.1}$$

Por la definición de δ existe una sucesión $\{y_n\}_{n=1}^{\infty} \subset E$ tal que

$$\lim_{n\to\infty} \|y_n\| = \delta,$$

y de la desigualdad (6.1) inferimos que

$$\|y_n - y_m\|^2 \leq 2\|y_n\|^2 + 2\|y_m\|^2 - 4\delta^2.$$

Al pasar al límite vemos que $\{y_n\}_{n=1}^{\infty}$ es de Cauchy. Como E es cerrado, existe $x_0 \in E$ tal que $\lim_{n\to\infty} y_n = x_0$, y por la continuidad de la norma concluimos que

$$\delta = \lim_{n\to\infty} \|y_n\| = \|x_0\|.$$

Si $y \in E$ es tal que $\|y\| = \delta$, por (6.1) tenemos

$$\|x_0 - y\|^2 \leq 0,$$

y de aquí que $x_0 = y_0$. □

Lema 6.1. *Sea H un espacio de Hilbert, y sea M un subconjunto de H. Entonces $M^{\perp}$ es cerrado.*

Demostración. Definimos para $x \in H$

$$x^{\perp} = \{y \in H : (x,y) = 0\}.$$

Sean $y, z \in x^{\perp}$, y λ un escalar. Entonces

$$(x, y + \lambda z) = (x,y) + \overline{\lambda}\,(x,z) = 0,$$

luego $x^{\perp}$ es un subespacio de H. Sea $\{y_n\}_{n=1}^{\infty}$ una sucesión convergente en $x^{\perp}$ y sea $y = \lim_{n\to\infty} y_n$. Puesto que el producto interior es continuo, se sigue que

$$0 = \lim_{n\to\infty} (x, y_n) = (x,y),$$

y de aquí que $x^{\perp}$ es cerrado. Finalmente, como

$$M^{\perp} = \bigcap_{x\in M} x^{\perp},$$

y la intersección arbitraria de cerrados es cerrada, $M^{\perp}$ es cerrado. □

Teorema 6.2. *Sea M un subespacio cerrado de un espacio de Hilbert H. Entonces existen proyecciones P de H sobre M y Q de H sobre $M^{\perp}$ tales que para toda $x \in H$,*

$$\begin{aligned} x &= Px + Qx, \\ \|x\|^2 &= \|Px\|^2 + \|Qx\|^2, \\ \|x - Px\| &= \text{ínf}\{\|x - y\| : y \in M\}. \end{aligned}$$

Además, si

$$z = x_1 + x_2$$

con $x_1 \in M$ y $x_2 \in M^{\perp}$, entonces $x_1 = Px$ y $x_2 = Qx$.

Demostración. Sea $x \in H$. Entonces $x + M = \{x + y : y \in M\}$ es un subconjunto cerrado, convexo de H. Sea $z = Qx$ el único elemento de norma mínima en este conjunto, el cual sabemos que existe por el Teorema 6.1. Definamos

$$Px = x - Qx = x - z.$$

Como $Qx \in x + M$, se tiene que $Px \in M$, y por definición

$$\begin{aligned} \|x\| &= \|Qx\| = \|x - Px\| \\ &= \text{ínf}\{\|x - y\| : y \in M\}. \end{aligned}$$

Sea $y \in M$ con $\|y\| = 1$ y α un escalar. Entonces

$$z - \alpha y = x - (Pz - \alpha y) \in x + M$$

y por esto

$$(z, z) = \|z\|^2 \leq \|z - \alpha y\|^2 = (z - \alpha y, z - \alpha y).$$

Por lo tanto

$$0 \leq -\alpha (y, z) - \overline{\alpha} (z, y) + |\alpha|^2,$$

y eligiendo $\alpha = (z, y)$ obtenemos

$$0 \leq -|(z, y)|^2 - |(z, y)|^2 + |(z, y)|^2 = -|(z, y)|^2,$$

de donde se sigue que $(z, y) = 0$. Sea ahora $y \in M$ arbitrario no nulo. Entonces

$$0 = \left(z, \frac{y}{\|y\|}\right) = \frac{1}{\|y\|}(z, y),$$

lo que implica que $(z, y) = 0$ y que

$$Qx = z \in M^{\perp}.$$

Como $Px \perp Qx$, es claro que

$$\|x\|^2 = \|Px\|^2 + \|Qx\|^2 .$$

Supongamos ahora que $x = x_1 + x_2$ con $x_1 \in M$ y $x_2 \in M^\perp$. Entonces

$$Px - x_1 = x_2 - Qx \in M \cap M^\perp = \{0\} ,$$

es decir, $x_1 = Px$ y $x_2 = Qx$. Probaremos ahora que P y Q son transformaciones lineales. Sean $x, y \in H$ y α, β escalares. Tenemos

$$\begin{aligned} x &= Px + Qx \\ y &= Py + Qy \\ \alpha x + \beta y &= P(\alpha x + \beta y) + Q(\alpha x + \beta y) , \end{aligned}$$

lo que nos permite escribir

$$\begin{aligned} \alpha x + \beta y &= \alpha Px + \alpha Qx + \beta Py + \beta Qy \\ &= P(\alpha x + \beta y) + Q(\alpha x + \beta y) , \end{aligned}$$

de donde

$$P(\alpha x + \beta y) - \alpha Px - \beta Py = Q(\alpha x + \beta y) - \alpha Qx - \beta Qy,$$

y como este elemento nuevamente pertenece tanto a M como a $M^\perp$, se sigue que

$$\begin{aligned} P(\alpha x + \beta y) &= \alpha Px + \beta Py, \\ Q(\alpha x + \beta y) &= \alpha Qx + \beta Qy, \end{aligned}$$

lo que significa que P y Q son lineales. $\square$

6.2. Espacios complementarios

Escolio 6.1. En todo espacio de Hilbert H cada subespacio cerrado M está complementado. Esto es, $M^\perp$ es tal que $M + M^\perp$ es denso en H, $M \cap M^\perp = \{0\}$, y la conclusión del Teorema 6.2 indica que $H = M \oplus M^\perp$.

Redoble de tambores: ha llegado el momento de enunciar y probar el Teorema de Espacios Complementarios.

Teorema 6.3 (De los espacios complementarios)**.** *Todo espacio de Banach X en el cual cada subespacio cerrado está complementado es isomorfo a un espacio de Hilbert.*

Demostración (Lindenstrauss y Tzafriri, 1970). Supondremos que X es un espacio dimensionalmente infinito, pues el caso finito es consecuencia del Corolario 4.1 y el Teorema 5.5. Sea $k = \lambda_f(X)$, la constante cuya existencia está garantizada por el Teorema 5.7. Considérese un subespacio dimensionalmente finito B de X y sea

$$\alpha = d\left(B, \ell_2^n\right),$$

donde $\dim B = n$. Sea Q una proyección de X sobre B con $\|Q\| \leq k$. Por el Teorema 4.11 existe un subespacio C dimensionalmente finito de $(I-Q)\,X$ tal que (tomando $\epsilon = 1$),

$$d\left(C, \ell_2^n\right) \leq 2. \tag{6.2}$$

Se sigue que

$$d\left(B, C\right) \leq 2\alpha,$$

y por lo tanto existe un T operador de B sobre C tal que

$$\frac{\|x\|}{2\alpha} \leq \|Tx\| \leq \|x\|$$

para todo $x \in B$. Sea P una proyección tal que $\|P\| \leq k$ de X sobre el subespacio

$$D = \left\{x + \mu Tx : x \in B\right\},$$

donde $\mu = 2^6 k^2$. Por construcción, para todo $x \in B$,

$$PTx = x_0 + \mu T x_0, \quad x_0 \in B.$$

Si $x_1 \in B$ satisface también $PTx = x_1 + \mu T x_1$, entonces

$$x_2 = QPTx = x_1,$$

y por lo tanto, el operador $V = QPT$ está bien definido y cumple con

$$PTx = Vx + \mu TVx.$$

Ahora bien,

$$\begin{aligned} \|Vx\| &\leq \|QPTx\| \\ &\leq \|Q\|\,\|P\|\,\|Tx\| \\ &\leq k^2\,\|Tx\|. \end{aligned}$$

Más aún, nótese que para cada $x \in B$ tenemos

$$\begin{aligned} Px &= P\left(x + \mu Tx\right) - \mu PTx \\ &= x + \mu Tx - \mu\left(Vx + \mu TVx\right) \\ &= \left(x - \mu Vx\right) + \mu\left(Tx - \mu TVx\right), \end{aligned}$$

y por lo tanto

$$(I-Q)\,Px = \mu\,(Tx - \mu TVx)\,.$$

Combinando todo lo anterior tenemos

$$\begin{aligned}
\|(I-Q)\,Px\| &= \mu\,\|T\,(x-\mu Vx)\| \\
&\geq \frac{\mu}{2\alpha}\,\|x-\mu Vx\| \\
&\geq \frac{\mu}{2\alpha}\,(\|x\| - \mu\,\|Vx\|) \\
&\geq \frac{\mu}{2\alpha}\left(\|x\| - \mu k^2\,\|Tx\|\right).
\end{aligned}$$

Usando (6.2) podemos hallar un operador U de C sobre ℓ_2^n tal que

$$\frac{\|y\|}{2} \leq \|Uy\| \leq \|y\|$$

para todo $y \in C$. Sea $\widehat{T}$ el operador de B sobre el espacio de Hilbert $2n$-dimensional $\ell_2^n \oplus \ell_2^n$ definido a través de

$$\widehat{T}x = \left(\frac{UTx}{2}, \frac{U\,(I-Q)\,Px}{4k^2}\right).$$

Entonces, por un lado

$$\begin{aligned}
\left\|\widehat{T}x\right\| &\leq \frac{\|U\|\,\|T\|\,\|x\|}{2} + \frac{\|U\|\,\|I-Q\|\,\|P\|\,\|x\|}{4k^2} \\
&\leq \frac{\|x\|}{2} + \frac{k\,(k+1)}{4k^2}\,\|x\| \leq \|x\|\,,
\end{aligned}$$

y por otro

$$\begin{aligned}
\left\|\widehat{T}x\right\| &\geq \text{máx}\left\{\frac{\|UTx\|}{2}, \frac{\|U\,(I-Q)\,Px\|}{4k^2}\right\} \\
&\geq \text{máx}\left\{\frac{\|Tx\|}{4}, \frac{\|(I-Q)\,Px\|}{8k^2}\right\} \\
&\geq \text{máx}\left\{\frac{\|Tx\|}{4}, \frac{\mu}{16\alpha k^2}\left(\|x\| - \mu k^2\,\|Tx\|\right)\right\}.
\end{aligned}$$

Distinguimos entre dos casos. Si $k^4 2^7\,\|Tx\| \geq \|x\|$ entonces

$$\left\|\widehat{T}x\right\| \geq \frac{\|Tx\|}{4} \geq \frac{\|x\|}{k^4 2^9}.$$

En el caso contrario de que $k^4 2^7 \|Tx\| < \|x\|$, dado que $\mu = 2^6 k^2$ tenemos

$$\begin{aligned} \left\|\widehat{T}x\right\| &\geq \frac{\mu}{16\alpha k^2}\left(\|x\| - k^4 2^6 \|Tx\|\right) \\ &\geq \frac{\mu}{2^5 \alpha k^2}\|x\| = \frac{2}{\alpha}\|x\|. \end{aligned}$$

En ambos casos

$$\left\|\widehat{T}x\right\| \geq \|x\| \operatorname{mín}\left(\frac{2}{\alpha}, \frac{1}{k^4 2^9}\right),$$

esto es, $\widehat{T}$ es un isomorfismo de B sobre ℓ_2^n. Como α es mínima, se sigue inmediatamente que $\alpha \leq k^4 2^9$. Éstas son las hipótesis del Teorema 5.4, el cual garantiza que X es isomorfo a un espacio de Hilbert. $\square$

Obtenemos como corolario la maravillosa caracterización que quería probar con este trabajo.

Corolario 6.1. *En todo espacio de Banach X cada subespacio cerrado está complementado si, y sólo si, X es isomorfo a un espacio de Hilbert.*

Ejemplo 6.1. Sea c_0 el subespacio de ℓ_∞ que consta de todas las sucesiones convergentes a 0. Como puede verse en [Jam74], c_0 es un subespacio cerrado de ℓ_∞ que no está complementado. Luego ℓ_∞ no sólo no es un espacio de Hilbert: no es isomorfo a *ningún* espacio de Hilbert.

Cerremos con broche de oro dando un ejemplo de subespacios de un espacio de Banach B cuya intersección es $\{0\}$, su suma es densa en B, pero no suman directamente a B.

Ejemplo 6.2. Sean H un espacio de Hilbert con una base numerable y $\{\phi_n\}_{n=0}^\infty$ una base ortonormal en H. Si para $n = 0, 1, 2, \ldots$, $a_n = \phi_{2n}$, $b_n = \phi_{2n} + \frac{1}{n+1}\phi_{2n+1}$, X es la cerradura del espacio vectorial $\operatorname{gen}\{a_n\}_{n=0}^\infty$ generado por $\{a_n\}_{n=0}^\infty$ y Y es la cerradura de $\operatorname{gen}\{b_n\}_{n=0}^\infty$.

1. Se cumple que $X \cap Y = \{0\}$.

2. Se satisface que $X + Y$ es denso en H pero no es un conjunto cerrado.

Demostración. Sea $u \in X \cap Y$. Entonces, por ser $\{\phi_n\}_{n=0}^\infty$ una base ortonormal y $u \in \overline{\operatorname{gen}\{a_n\}_{n=0}^\infty} \cap \overline{\operatorname{gen}\{b_n\}_{n=0}^\infty}$

$$\begin{aligned} u &= \sum_{i=0}^\infty \alpha_i \phi_{2i} \\ &= \sum_{i=0}^\infty \beta_i \left(\frac{i+1}{i+2}\phi_{2i} + \frac{1}{i+2}\phi_{2i+1}\right) \\ &= \sum_{i=0}^\infty \frac{i+1}{i+2}\beta_i \phi_{2i} + \sum_{i=0}^\infty \frac{\beta_i}{i+2}\phi_{2i+1}, \end{aligned}$$

lo que implica que

$$0 = \sum_{i=0}^{\infty} \left(\alpha_i - \frac{i+1}{i+2} \beta_i \right) \phi_{2i} + \sum_{i=0}^{\infty} \frac{\beta_i}{i+2} \phi_{2i+1},$$

de donde $\beta_i = 0$, dada la representación única del 0 en la base ortonormal, y de aquí que $\alpha_i = 0$. Luego $u = 0$.

Para cualquier $v \in H$ y $\epsilon > 0$, existe M tal que

$$\left| v - \sum_{n=0}^{M} \hat{v}(n) \phi_n \right| \leq \epsilon,$$

pero

$$\sum_{n=0}^{M} \hat{v}\,(n)\,\phi_n = \sum_{n=0}^{M} \left(\hat{v}(2n) - (n+1)\hat{v}(2n+1)\right) \phi_{2n} + \sum_{n=0}^{M} (n+1)\hat{v}\,(2n+1) \left(\phi_{2n} + \frac{1}{n+1} \phi_{2n+1} \right) \in X + Y,$$

así que definiendo

$$x_M = \sum_{n=0}^{M} \left(\hat{v}(2n) - (n+1)\hat{v}(2n+1)\right) \phi_{2n},$$

$$y_M = \sum_{n=0}^{M} (n+1)\hat{v}\,(2n+1) \left(\phi_{2n} + \frac{1}{n+1} \phi_{2n+1} \right)$$

se tiene que

$$|v - (x_M + y_M)| \leq \epsilon,$$

luego $X + Y$ es denso en H. Sin embargo, la serie

$$\sum_{n=0}^{\infty} (n+1)\hat{v}(2n+1)\phi_{2n}$$

no convergerá para el elemento $v = \sum_{k=0}^{\infty} \frac{1}{(k+1)^2} \phi_{2k+1} \in H$. Luego $X + Y$ no es cerrado. $\square$

Conclusiones

...Las matemáticas puras son, en su totalidad, distintivamente más útiles que las aplicadas... pues lo más útil por sobre todas las cosas es la *técnica*, y la técnica matemática se enseña, principalmente, a través de las matemáticas puras.

G. H. Hardy

Con este trabajo se dió una exposición lo más clara y completa posible de la teoría local de espacios de Banach requerida para demostrar el Teorema de Espacios Complementarios. El principal éxito es la unificación de la notación y un desarrollo coherente y pleno de los resultados.

Un aspecto fundamental de esta teoría es la concentración de la medida, que en este caso se presentó para la medida de Haar. En este sentido, como puede verse en [Pis89], se han desarrollado resultados de concentración para otro tipo de medidas. De hecho, usando medidas gaussianas es posible obtener una demostración alternativa del Teorema de Espacios Complementarios, aunque preferí hacerlo empleando la medida de Haar por tener una sabor geométrico más claro, y que por ende resulta más intuitivo.

La necesidad de este tipo de exposiciones queda patente con la aplicación realizada a la teoría de codigos. Los Teoremas 2.8 y 2.9 ya estaba en la literatura bastante tiempo antes de que se empezaran a aplicar los resultados de concentración a la teoría de códigos y de algoritmos (véase [RU01]). Sin embargo, una falta de comunicación entre los especialistas dedicados al fenómeno de concentración y aquéllos dedicados a la teoría de códigos impiden que resultados, considerados puramente abstractos, encuentren aplicaciones de gran valor.

Como trabajo futuro, se puede revisar el análisis del algoritmo de paso de mensajes para poder aplicar el Teorema 2.9. También que es deseable que las cotas involucradas en los Teoremas 2.8 y 2.9 no dependan exponencialmente del número de iteraciones, pues tal comportamiento compensa en demasía la concentración exponencial observada.

Apéndice A

Teoría de la medida

En este apéndice se definen conceptos y se enuncian (sin demostración) teoremas concernientes a la teoría de la medida necesarios para completar el texto.

Definición A.1. Una colección $\mathcal{M}$ de subconjuntos de un conjunto X se denomina σ-álgebra en X si tiene las siguientes propiedades.

1. El conjunto X está en $\mathcal{M}$.
2. Si $A \in \mathcal{M}$, entonces $\complement_X A \in \mathcal{M}$.
3. Si $A = \bigcup_{n=1}^{\infty} A_n$, y $A_n \in \mathcal{M}$ para $n = 1, 2, \ldots$, entonces $A \in \mathcal{M}$.

Al par $(X, \mathcal{M})$ se le denomina espacio medible, y cualquier $A \in \mathcal{M}$ se dice que es un conjunto medible. Una función $f : X \to Y$ donde (Y, τ) es un espacio topológico se dice medible si

$$f^{-1}(B) \in \mathcal{M}$$

para todo $B \in \tau$.

Teorema A.1. *Si $\mathcal{F}$ es una colección cualquiera de subconjuntos de X, existe una σ-álgebra mínima $\mathcal{M}^*$ en X tal que $\mathcal{F} \subseteq \mathcal{M}^*$. Esta $\mathcal{M}^*$ se denomina σ-álgebra generada por $\mathcal{F}$.*

Sea X un espacio topológico. Por el Teorema A.1, existe una σ-álgebra mínima $\mathcal{B}$ en X tal que todo conjunto abierto de X pertenece a $\mathcal{B}$. Los elementos de $\mathcal{B}$ se denominan conjuntos de Borel de X.

Definición A.2. Sea $(X, \mathcal{M})$ un espacio medible. Una medida positiva μ es una función

$$\mu : \mathcal{M} \to [0, \infty]$$

que satisface lo siguiente.

1. Para algún $A \in \mathcal{M}$, $\mu(A) < \infty$.

2. Para todo conjunto de índices I numerable y $\{A_\iota\}_{\iota\in I}$ una familia de conjuntos disjunta, se tiene la identidad

$$\mu\left(\bigcup_{\iota\in I} A_\iota\right) = \sum_{\iota\in I} \mu(A_\iota).$$

A la tripleta $(X, \mathcal{M}, \mu)$ se le denomina espacio de medida. Si $\mu(X) = 1$, entonces la medida se dice que es de probabilidad, y lo mismo se dice del espacio medible subyacente.

El siguiente resultado es de gran importancia para el análisis funcional y es uno de los llamados "de representación" debidos a Riesz. La demostración se encuentra en [Rud79].

Teorema A.2 (Riesz, 1909)**.** *Sea X un espacio de Hausdorff localmente compacto y sea Λ un funcional lineal positivo sobre el conjunto de funciones continuas sobre X con soporte compacto, $C_c(X)$. Entonces existe una σ-álgebra $\mathcal{M}$ en X que contiene a todos los conjuntos de Borel de X, y existe una única medida positiva μ sobre $\mathcal{M}$ que satisface lo siguiente.*

1. *Para toda $f \in C_c(X)$ tenemos que $\Lambda f = \int_X f\, d\mu$.*

2. *Para todo compacto $K \subset X$ se tiene que $\mu(K) < \infty$.*

3. *Para todo $E \in \mathcal{M}$ tenemos*

$$\mu(E) = \text{ínf}\{\mu(V) : E \subset V, V \textit{ abierto}\}.$$

4. *La relación*

$$\mu(E) = \sup\{\mu(K) : K \subset V, K \textit{ compacto}\}.$$

   se verifica para todo conjunto abierto E, y para todo $E \in \mathcal{M}$ con $\mu(E) < \infty$.

5. *Si $E \in \mathcal{M}$, $A \subset E$ y $\mu(E) = 0$, entonces $A \in \mathcal{M}$.*

Definición A.3. Una medida de Borel sobre X (donde X es un espacio de Hausdorff localmente compacto) es una medida μ definida sobre la σ-álgebra de todos los conjunto de Borel. Un conjunto de Borel se dice regular exterior o interior, respectivamente, si E posee la propiedad 3 o 4 del Teorema A.2. Si todos los conjuntos de Borel de X son a la vez regulares interiores y exteriores, entonces μ se dice regular.

Para las cuestiones de almártagas necesitamos la noción de medida compleja, continuidad absoluta de una medida y el Teorema de Radon y Nikodym. Este último aparece demostrado en [Rud79].

Definición A.4. Sea $\mathcal{M}$ una σ-álgebra en un conjunto X. Una medida compleja μ sobre $\mathcal{M}$ es una función compleja $\mu : M \to \mathbb{C}$ tal que

$$\mu(E) = \sum_{i=0}^{\infty} \mu(E_i), \quad E \in \mathcal{M}$$

para toda partición numerable $\{E_i\}_{i=0}^{\infty}$ de E.

Definición A.5. Sea μ una medida positiva sobre una σ-álgebra $\mathcal{M}$, y sea λ una medida compleja o positiva sobre $\mathcal{M}$. Decimos que λ es absolutamente continua respecto a μ si $\mu(E) = 0$ implica que $\lambda(E) = 0$ para todo $E \in \mathcal{M}$.

Teorema A.3 (Radon y Nikodym)**.** *Sean μ y λ medidas positivas acotadas sobre una σ-álgebra $\mathcal{M}$ sobre el conjunto X. Si λ es absolutamente continua respecto a μ, existe una única $h \in L^1(\mu)$ tal que*

$$\lambda = \int_E h \, d\mu, \quad E \in \mathcal{M}.$$

Finalmente, enunciamos un resultado que requeriremos en el siguiente apéndice. Su demostración se encuentra en [GF02].

Teorema A.4 (De convergencia dominada)**.** *Sea X un espacio de medida, y $\Phi, f_1, f_2, \ldots$ funciones medibles tales que $\int_X \Phi < \infty$ y $|f_n| \leq \Phi$ para cada n. Si $f_n \to f$ en casi todas partes, entonces f es integrable y*

$$\lim_{n\to\infty} \int_X f_n = \int_X f.$$

Apéndice B

Probabilidad

Introducimos algunas desigualdades y definiciones relativas a la probabilidad, en términos de la teoría de la medida.

Definición B.1. Cualquier función medible $X : \Omega \to \mathbb{R}$ sobre un espacio de probabilidad $(\Omega, \mathcal{M}, P)$ se denomina variable aleatoria.

Muchas veces consideraremos a Ω como un subconjunto de $\mathbb{R}$. De hecho, así lo haremos en lo sucesivo.

Definición B.2. Para cualquier variable aleatoria no negativa X definimos su esperanza o valor esperado como

$$E(X) = \int_0^\infty P(X \geq s)\, ds.$$

Escolio B.1. En particular, si X tiene una distribución normal con media 0 y varianza 1, esto es

$$X(t) = \frac{\exp\left(-\frac{t^2}{2}\right)}{\sqrt{2\pi}},$$

entonces

$$\begin{aligned} E(X) &= \int_0^\infty P(X \geq s)\, ds \\ &= \frac{1}{\sqrt{2\pi}} \int_0^\infty \int_s^\infty \exp\left(-\frac{t^2}{2}\right) dt\, ds \\ &\leq \frac{1}{\sqrt{2\pi}} \int_0^\infty \int_s^\infty \exp\left(-\frac{t}{2}\right) dt\, ds \\ &= \sqrt{\frac{2}{\pi}} \int_0^\infty \exp\left(-\frac{s}{2}\right) ds = 2\sqrt{\frac{2}{\pi}}, \end{aligned}$$

es decir, $E(X) \leq \infty$.

Es fácil obtener el siguiente hecho de la definición de esperanza y las propiedades de la probabilidad.

Teorema B.1 (Markov). *Para cualquier variable aleatoria X no negativa y cualquier $t > 0$, se satisface*

$$P(X \geq t) \leq \frac{E(X)}{t}.$$

Demostración. Sea $A = \{\omega : X(\omega) < t\}$ y $B = \{\omega : X(\omega) \geq t\}$. Claramente $A \cap B = \emptyset$ y $A \cup B = \Omega$. Entonces

$$\begin{aligned} E(X) &= E(\chi_A \cdot X + \chi_B \cdot X) \\ &= E(\chi_A \cdot X) + E(\chi_B \cdot X) \\ &\geq E(\chi_B \cdot X) \geq tE(\chi_B) = tP(X \geq t), \end{aligned}$$

lo cual es equivalente al aserto. □

Corolario B.1 (Chebyshev). *Sea X cualquier variable aleatoria no negativa y ϕ una función estrictamente creciente no negativa. Entonces, si $t > 0$,*

$$P(X \geq t) = P(\phi(X) \geq \phi(t)) \leq \frac{E(\phi(X))}{\phi(t)}.$$

En particular, para todo $t > 0$,

$$P(|X - E(X)| \geq t) \leq \frac{E(|X - E(X)|^2)}{t^2} = \frac{\mathrm{Var}(X)}{t^2}$$

Demostración. Como la variable aleatoria $Y = |X - E(X)|$ es no negativa y la transformación $\phi(x) = x^2$ es estrictamente creciente y positiva para $x \geq 0$, el segundo aserto se sigue inmediatamente. □

Para las siguientes proposiciones, μ es la medida de Haar sobre S^{n-1}, $f : S^{n-1} \to \mathbb{R}$ es una función continua y definimos la función $\psi : \mathbb{R} \to \mathbb{R}$ a través de

$$\psi(t) = \mu(\{x : f(x) \leq t\}).$$

Proposición B.1. *La función ψ es no decreciente, continua por la derecha, y tiene límites por la izquierda.*

Demostración. Si $u < t$, entonces $\{x : f(x) \leq u\} \subset \{x : f(x) \leq t\}$. En consecuencia,

$$\psi(u) = \mu(\{x : f(x) \leq u\}) \leq \mu(\{x : f(x) \leq t\}) = \psi(t).$$

.

Sea $\{t_i\}_{i=0}^{\infty}$ una sucesión que converge a t, tal que $t_i \geq t$ para todo $i \geq 0$. Tomemos $\epsilon > 0$ y $x \in S^{n-1}$. Si $f(x) > t$, sea $\delta < f(x) - t$; existe N tal que $n > N$ implica que $t_n - t < f(x) - t$, y esto a su vez que $t_n < f(x)$. Si $f(x) \leq t$, para todo $n \geq 0$ se cumple que $f(x) \leq t_n$. En suma, $|f_n(x) - f(x)| \leq \epsilon$. Resulta así que la sucesión de funciones

$$g_n(x) = \chi_{\{u:f(u)\leq t_n\}}(x)$$

converge casi en todas partes a $g(x) = \chi_{\{u:f(u)\leq t\}}(x)$. Por el Teorema A.4,

$$\psi(t) = \int_{S^{n-1}} \lim_{n\to\infty} g_n(x)\, d\mu = \lim_{n\to\infty} \int_{S^{n-1}} g_n(x)\, d\mu = \lim_{n\to\infty} \psi(t_n),$$

lo que significa que ψ es continua por la derecha.

Ahora sea $\{t_i\}_{i=0}^{\infty}$ una sucesión que converge a t, pero con $t_i < t$ para todo $i \geq 0$. Por un lado, si $f(x) \geq t$ entonces $f(x) \geq t > t_i$ para todo $i \geq 0$. Por otro, si $f(x) < t$, sea $\delta < t - f(x)$. Existe N tal que $n > N$ implica que $t - t_n < t - f(x)$, lo que es precisamente $f(x) < t_n$. Nuevamente las funciones g_n convergen en casi todas partes a la función $h(x) = \chi_{\{u:f(u)<t\}}$ y

$$\begin{aligned}\psi(t-) := \mu(\{u : f(u) < t\}) &= \int_{S^{n-1}} \lim_{n\to\infty} g_n(x)\, d\mu = \int_{S^{n-1}} h(x)\, d\mu \\ &= \lim_{n\to\infty} \int_{S^{n-1}} g_n(x)\, d\mu = \lim_{n\to\infty} \psi(t_n).\end{aligned}$$

Se cumple así que los límites por la izquierda existen. □

Proposición B.2. *Existe un número M tal que*

$$\mu\{x : f(x) \leq M\} \leq \frac{1}{2} \quad y \quad \mu\{x : f(x) \geq M\} \geq \frac{1}{2},$$

y que diremos que es una mediana de f.

Demostración. Sean $A = \{x : \psi(x) \geq 1/2\}$ y $B = \{x : \psi(x-) \leq 1/2\}$. Afirmamos que los números $m_\ell = \inf A$ y $m_u = \inf B$ son finitos. En efecto, S^{n-1} es compacto bajo la métrica euclídea y en consecuencia la función f es acotada. Sea $0 < C < \infty$ una cota para f; la desigualdad $-C < f(x) < C$ se satisface para todo x, de modo que $\psi(C) = 1$ y $\psi(-C) = 0$. Dada la monotonía de ψ, $-C \leq x$ para todo $x \in A$ y $x \leq C$ para todo $x \in B$.

Para cualquier $y \leq m_\ell$, tenemos que $\psi(y-) \leq \psi(y) < 1/2$, lo que implica que $m_u \geq m_\ell$. Por lo tanto el intervalo $[m_\ell, m_u]$ es no vacío. Cualquier $M \in (m_\ell, m_u)$ es una mediana de f, pues

$$1 - \psi(M-) = \mu(\{x : f(x) \geq M\}) \geq \frac{1}{2} \quad \text{y} \quad \psi(M) = \mu(\{x : f(x) \leq M\}) \geq \frac{1}{2}.$$

Por la continuidad por la derecha de ψ, incluso m_ℓ es una mediana. En todo caso, f siempre tiene una mediana. □

Apéndice C

Suplemento de análisis funcional

Conviene definir a los espacios L^p y ℓ_p.

Definición C.1. Sea $(X, \mathcal{M}, \mu)$ un espacio de medida. Si $0 < p < \infty$ y f una función compleja medible sobre X, hacemos

$$\|f\|_p = \left(\int_X |f|^p \, d\mu \right)^{1/p}$$

y definimos a $L^p(X, \mathcal{M}, \mu)$ como el conjunto de todas las f para las que

$$\|f\|_p < \infty.$$

A $\|f\|_p$ se le llama la norma en L^p de f.

Definición C.2. Definimos a ℓ_p como el espacio vectorial de todas las sucesiones $\{x_\alpha\}_{\alpha\in A} \subset \mathbb{K}$ tales que $\sum_{\alpha\in A} |x_\alpha|^p \leq \infty$, dotado con la norma $\|x\| = \left(\sum_{\alpha\in A} |x_\alpha|^p\right)^{1/p}$. Si $|A| = n \in \mathbb{N}$, denotamos a ℓ_p como ℓ_p^n

En $\mathbb{R}^n$ cualquier conjunto cerrado y acotado es compacto. Un criterio análogo en análisis funcional es el teorema de Arzelà y Ascoli. Una demostración se puede encontrar en [Yos80].

Teorema C.1 (Arzelà y Ascoli)**.** *Sea S un espacio métrico compacto, y $C(S)$ el espacio de Banach de todas las funciones reales o complejas $x(s)$ normadas por*

$$\|x\| = \sup_{s\in S} |x(s)|.$$

Entonces una sucesión $\{x_n(s)\}_{n=1}^{\infty} \subset C(S)$ es relativamente compacta en $C(S)$ si se satisfacen las siguentes condiciones.

1. *La sucesión $\{x_n(s)\}_{n=1}^{\infty} \subset C(S)$ es equiacotada, esto es,*

$$\sup_{n\geq 1} \sup_{s\in S} |x_n(s)| \leq \infty$$

2. La sucesión $\{x_n(s)\}_{n=1}^{\infty} \subset C(S)$ *es equicontinua (en* n*),*

$$\lim_{\delta \to 0} \sup_{n \geq 1, d(s', s'') \leq \delta} |x_n(s') - x_n(s'')| = 0.$$

Otro resultado que apareció sutilmente es la existencia de cierto elipsoide. Lo que sigue está basado en las notas de R. Howard, [How97]. Veamos un lema previo para la demostración.

Lema C.1. *Sea* $\mathcal{E}$ *un elipsoide en* $\mathbb{R}^n$ *de modo que* $\mathcal{E} = AB_2^n + b$ *con* $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ *y* $b \in \mathbb{R}^n$*. Entonces puede elegirse* A *como definida positiva.*

Demostración. Cualquier matriz no singular A puede escribirse como $A = PU$ donde P es definida positiva y U es ortogonal (usando la descomposición en valores singulares). Para toda matriz ortogonal U tenemos que $UB_2^n = B_2^n$, por lo que podemos reemplazar a A por P. □

Recordemos que si A es invertible

$$\mathrm{Vol}(AK) = |\det(A)| \cdot \mathrm{Vol}(K),$$

lo cual se obtiene de la fórmula de integración bajo un cambio de base.

Teorema C.2 (John). *Cada cuerpo convexo* $K \subset \mathbb{R}^n$ *contiene un único elipsoide de máximo volumen.*

Demostración. Sea $N = n^2 + n$. El conjunto de pares (A, b) donde A es una matriz de $n \times n$ y $b \in \mathbb{R}^n$ es $\mathbb{R}^N$. Sea

$$\mathcal{L} = \{(A, b) \in \mathbb{R}^N : AB_2^n + b \subset K\}.$$

El conjunto $\mathcal{L}$ es cerrado bajo la norma euclidiana. Por lo tanto es compacto en $\mathbb{R}^N$. Además, la función $f(A, b) = |\det A|$ es continua (dada la continuidad de las operaciones aritméticas en $\mathbb{R}$), por lo que existe (A_0, b_0) que maximiza el valor de f. Este valor no puede ser 0, pues $(kI, 0) \in \mathcal{L}$ para un número positivo k suficientemente pequeño, al ser K de interior no vacío. Por lo tanto A_0 es invertible, y de aquí que $A_0 B_2^n + b_0$ sea un elipsoide del máximo volumen posible.

Para probar la unicidad, consideremos ahora el conjunto de todas las matrices definidas positivas de $n \times n$, $\mathcal{P}_n$, y sea $\mathcal{K} = \mathcal{P}_n \times \mathbb{R}^n$. Tomemos

$$\mathcal{L} = \{(A, b) \in \mathcal{K} : AB_2^n + b \subset K\}.$$

Por el Lema C.1 cada elipsoide contenido en K es de la forma $AB_2^n + b$ para algunos $(A, b) \in \mathcal{K}$. Como K es convexo y una combinación lineal de matrices definidas positivas sigue siendo definida positiva, entonces $\mathcal{L}$ es convexo también.

Sean $\mathcal{E}_1 = A_1 B_2^n + b_1$ y $\mathcal{E}_2 = A_2 B_2^n + b_2$ de máximo volumen contenidos en K. Se cumple que $\det A_1 = \det A_2$. Además tenemos que

$$\mathcal{E}_3 = \frac{1}{2}(\mathcal{E}_1 + \mathcal{E}_3) \subset K.$$

por ser $\mathcal{L}$ convexo, y $\mathcal{E}_3 = \frac{1}{2}(A_1 + A_2)B_2^n + \frac{1}{2}(b_1 + b_2)$.

Denotemos por V el volumen de la bola B_2^n. Por el Teorema 2.4,

$$\begin{aligned}\left(|\det(\tfrac{1}{2}(A_1 + A_2))|V\right)^{\frac{1}{n}} &= \mathrm{Vol}(\mathcal{E}_3)^{\frac{1}{n}} \\ &\geq \mathrm{Vol}(\tfrac{1}{2}\mathcal{E}_1)^{\frac{1}{n}} + \mathrm{Vol}(\tfrac{1}{2}\mathcal{E}_2)^{\frac{1}{n}} \\ &= \left(|\det(\tfrac{1}{2}A_1)|V\right)^{\frac{1}{n}} + \left(|\det(\tfrac{1}{2}A_2)|V\right)^{\frac{1}{n}} \\ &= (|\det(A_1)|V)^{\frac{1}{n}},\end{aligned}$$

pero la maximalidad de $\mathcal{E}_1$ y $\mathcal{E}_2$ implican que $|\det A_3| = |\det A_1| = |\det A_3|$ donde $A_3 = \frac{1}{2}(A_1 + A_2)$. De aquí que $A_3 = A_1 = A_2$. Ahora, si $b_1 \neq b_2$, entonces la envolvente convexa de $\mathcal{E}_1$ y $\mathcal{E}_2$ contendría al elipsoide $\mathcal{E}_3$, pero tendría mayor volumen, lo cual es contradictorio. Esto concluye la demostración. $\square$

Los siguientes resultados y definiciones hablan sobre propiedades de continuidad de transformaciones lineales entre espacios normados.

Teorema C.3. *Sean X y Y espacios normados y $T : X \to Y$ una transformación lineal. Los siguientes asertos son equivalentes.*

1. *La transformación T es continua.*
2. *Existe una constante $C > 0$ tal que $\|Tx\| \leq C\|x\|$ para toda $x \in X$.*

Definición C.3. Sean X y Y espacios normados y $T : X \to Y$ una transformación lineal. Definimos la norma de T como

$$\|T\| = \text{ínf}\{C : \|Tx\| \leq C\|x\|, x \in C\} = \sup_{|x\| \leq 1} \|Tx\|.$$

Si $\|Tx\|$ no está acotado por arriba para $\|x\| \leq 1$, definimos $\|T\| = \infty$. En caso contrario, decimos que T está acotada.

Escolio C.1. Por el Teorema C.3, una transformación lineal es acotada si, y sólo si, es continua.

Teorema C.4 (De la gráfica cerrada)**.** *Si X y Y son espacios de Banach y $T : X \to Y$ es una función lineal cuya gráfica*

$$\{(x, Tx) \in X \times Y : x \in X\}$$

es cerrada en $X \times Y$ (con la topología producto), entonces T es continua.

Apéndice D

El Teorema de Hall

El siguiente hecho combinatorio tiene diversas aplicaciones en la teoría de grafos y optimización. El enunciado y la demostración que damos está basada en la que aparece en [Bal95].

Teorema D.1 (P. Hall). *Sea M un conjunto y $S = \{S_i\}_{i=1}^n$ un conjunto de subconjuntos finitos de M. Existe una función inyectiva $\phi : \{1, \ldots, n\} \to M$ con $\phi(i) \in S_i$ si, y sólo si, se cumple*

$$\left| \bigcup_{k \in J} S_k \right| \geq |J| \tag{D.1}$$

para todo $J \subseteq \{1, \ldots, n\}$.

Demostración. Si existe ϕ con las condiciones requeridas, entonces

$$\phi(J) \subseteq \bigcup_{k \in J} S_k,$$

lo que implica que $|J| = |\phi(J)| \leq \left|\bigcup_{k \in J} S_k\right|$. Para probar el recíproco, eliminemos sendos elementos de los miembros de S hasta quedar con $U = \{U_k\}_{k=1}^n$, de manera tal que si eliminamos un elemento más, para algún $J \subseteq \{1, \ldots, n\}$ ya no se satisface (D.1). Ningún U_i puede quedar vacío, pues en tal caso para $J = \{i\}$ se tiene $|U_i| = 0 \geq |J| = 1$, lo cual es imposible.

Afirmamos que, de hecho, $U_k = \{u_k\}$ para $1 \leq k \leq n$, quedando definida una función biyectiva $\phi(S_k) = u_k$. Supongamos, por el contrario, que U_1 tiene por lo menos dos elementos x y y. Dada la construcción de U, si removemos de U_1 a x o a y ya no se satisfará (D.1). Por lo tanto, existen dos conjuntos de índices P y Q tales que los conjuntos

$$X = (U_1 \setminus \{x\}) \cup \left(\bigcup_{i \in P} U_i \right) \quad \text{y} \quad Y = (U_1 \setminus \{y\}) \cup \left(\bigcup_{i \in Q} U_i \right),$$

cumplen con $|X| < |P|$ y $|Y| < |Q|$. Por un lado, usando el principio de inclusión y exclusión,

$$|X \cup Y| + |X \cap Y| = |X| + |Y| < |P| + |Q|,$$

mientras que, por otro,

$$X \cup Y = U_1 \cup \left(\bigcup_{i \in P \cup Q} U_i \right) \quad \text{y} \quad X \cap Y = \bigcup_{i \in P \cap Q} U_i.$$

La ecuación (D.1) aplicada a $T' = \{U_i : i \in \{1\} \cup P \cup Q\}$ y a $T'' = \{U_i : i \in P \cap Q\}$ nos indica que

$$| \cup T'| = |X \cup Y| \geq 1 + |P \cup Q| \quad \text{y} \quad | \cup T''| = |X \cap Y| \geq |P \cap Q|,$$

lo cual, conjuntamente con el principio de inclusión y exclusión, conduce a la desigualdad

$$|X \cup Y| + |X \cap Y| \geq 1 + |P \cup Q| + |P \cap Q| = 1 + |P| + |Q| > |P| + |Q|,$$

lo cual es contradictorio y establece el teorema. □

Bibliografía

[Bal95] Balakrishnan, V. K.: *Theory and problems of combinatorics.* Schaum's Outlines. McGraw-Hill, 1995.

[Bal97] Ball, Keith: *An elementary introduction to modern convex geometry.* En *Flavors of geometry.* MSRI Publications, 1997.

[BE00] Bastero Eleizalde, Jesús: *De la geometría de los espacios de Banach al análisis convexo y geométrico.* Academia de Ciencias Exactas, Físicas, Químicas y Naturales de Zaragoza, Zaragoza, 2000.

[Day73] Day, Mahlon Marsh: *Normed linear spaces.* Springer-Verlag, 1973.

[DDS68] Davis, W., D. Dean y I. Singer: *Complemented subspaces and Λ systems in Banach spaces.* Israel Journal of Mathematics, 6:303–309, 1968.

[Dör65] Dörrie, Heinrich: *100 great problems of elementary mathematics.* Dover, New York, 1965.

[FG97] Fetter Nathansky, Helga y Berta Gamboa de Buen: *Introducción al análisis funcional y a la geometría de espacios de Banach.* Grupo Editorial Iberoamérica, 1997.

[GF02] Galaz Fontes, Fernando: *Medida e integral de Lebesgue en $\mathbb{R}^N$.* Oxford University Press, México, D. F., 2002.

[Grü60] Grünbaum, B.: *Some applications of expansion constants.* Pacific Journal of Mathematics, 10:193–201, 1960.

[How97] Howard, Ralph: *The John ellipsoid theorem.* Notas del seminario de análisis funcional en la Universidad de Carolina del Sur, 1997.

[Jam74] Jameson, G. J.: *Topology and normed spaces.* Chapman and Hall, 1974.

[Joi66] Joichi, J. T.: *Normed linear spaces equivalent to inner product spaces.* Proceedings of the American Mathematical Society, 17:423–426, 1966.

[KF75] Kolmogorov, A. N. y S. V. Fomin: *Elementos de la teoría de las funciones y del análisis funcional*. Editorial Mir, Moscú, 1975.

[LP97] Lidl, Rudolf y Günter Pilz: *Applied abstract algebra*. Undergraduate Texts in Mathematics. Springer-Verlag, 1997.

[LT71] Lindentrauss, Joram y Lior Tzafriri: *On the complemented subspaces problem*. Israel Journal of Mathematics, 9:263–269, 1971.

[MS86] Milman, Vitali y Gideon Schechtman: *Asymptotic theory of finite dimensional normed spaces*, volumen 1200 de *Lecture Notes in Mathematics*. Springer-Verlag, 1986.

[Pis83] Pisier, Gilles: *On the dimension of the ℓ_p^n-subspaces of Banach spaces, for* $1 < p < 2$. Transactions of American Mathematical Society, 276:201–211, 1983.

[Pis89] Pisier, Gilles: *The volume of convex bodies and Banach space geometry*, volumen 94 de *Cambridge Tracts in Mathematics*. Cambridge University Press, 1989.

[Roy63] Royden, H. L.: *Real analysis*. Macmillan, New York, 1963.

[RU01] Richardson, Thomas y Ruediger Urbanke: *The capacity of low-density parity-check codes under message passing decoding*. IEEE Transactions on Information Theory, 47:599–618, 2001.

[RU06] Richardson, Thomas y Ruediger Urbanke: *Modern Coding Theory*. Por publicarse, 2006.

[Rud79] Rudin, Walter: *Análisis real y complejo*. Alhambra, Madrid, 1979.

[Rya04] Ryan, W. E.: *An introduction to LDPC codes*. En Vasic, B. (editor): *CRC Handbook for Coding and Signal Processing for Recording Systems*. CRC Press, 2004.

[Sch] Schechtman, Gideon: *Topics in the local theory of normed spaces*. En internet: `http://www.wisdom.weizmann.ac.il/~borisl/`.

[Yos80] Yosida, Kôsaku: *Functional analysis*. Grundlehren der Mathematischen Weissenschaften. Springer-Verlag, 6a. edición, 1980.