Universidad Nacional Autónoma de México

Facultad de Ciencias

PUNTOS EXTREMOS DE CONJUNTOS DE MEDIDAS

TESIS

que para obtener el título de:

MATEMÁTICO

presenta:

FEDERICO LASA GONSEBATT

Director de tesis: DRA. ANA MEDA GUARDIOLA

2008

1. Datos del Alumno Lasa Gonsebatt Federico 56 76 93 71 Universidad Nacional Autónoma de México Facultad de Ciencias Matemáticas 402067407
2. Datos del Tutor Dra. Ana Meda Guardiola
3. Datos del sinodal 1 Dra. Magali Louise Marie Folch Gabayet
4. Datos del sinodal 2 Dr. Francisco Marcos López García
5. Datos del sinodal 3 M. en C. Alejandro Bravo Mojica
6. Datos del sinodal 4 Dr. Javier Páez Cárdenas
7. Datos del trabajo escrito Puntos extremos de conjuntos de medidas 76 p 2008

Índice general

Introducción iii

1. Preliminares **1**
- 1.1. Conjuntos convexos 1
- 1.2. Puntos extremos 7

2. Medidas extremas y densidad de subespacios **11**
- 2.1. Teorema de Douglas 11
- 2.2. Teorema de Lindenstrauss. 13

3. Medidas con marginales fijas **19**
- 3.1. Soporte de medidas doblemente estocásticas 19
- 3.2. Soporte de medidas extremas discretas con marginales fijas 23

4. Medidas extremas con momentos fijos **31**
- 4.1. Problema de momentos y densidad de subespacios 31
- 4.2. Problema generalizado de momentos truncos 32
- 4.3. Soporte de medidas extremas discretas con momentos fijos 33
- 4.4. Generación por medidas extremas 36
- 4.5. Conexión entre los problemas de momentos y marginales fijas 40

5. Teorema de Birkhoff-von Neumann **43**
- 5.1. Caso finito 44
- 5.2. Caso infinito numerable 47
- 5.3. Caso no numerable 56

A. Apéndice **65**

Introducción

Los conjuntos convexos son por sí mismos objetos interesantes de estudio. En el estudio de conjuntos convexos, los puntos extremos son de importancia pues en general contienen la información suficiente para describir a todo el conjunto vía el Teorema de Krein-Milman.

En este trabajo se trata el tema de algunos conjuntos convexos de medidas y resultados sobre sus puntos extremos. Los resultados que se exponen fueron elegidos porque ilustran métodos de análisis y porque relacionan problemas clásicos, como es el caso del problema de los momentos y el problema de marginales fijas.

El primer capítulo esta dedicado a las nociones básicas sobre los conjuntos convexos y se limita a la exposición de resultados que son necesarios para el desarrollo del material subsecuente.

En el segundo capítulo se expone el Teorema de Douglas-Lindenstrauss, se presenta el teorema como lo publicó Douglas en [Dou64], y después la forma en que lo obtuvo Lindenstrauss en [Lin65], así como la implicación del primero al segundo. Este teorema representará la principal herramienta para el estudio de puntos extremos y sus implicaciones se muestran a lo largo del trabajo en cada capítulo. También se definen las medidas doblemente estocásticas. Estas medidas resultan de particular importancia por las diversas representaciones que tiene en diversas areas de estudio, un ejemplo de ésto es la identificación con operadores doblemente estocásticos del capítulo 5.

El tercer capítulo se dedica a caracterizaciones del soporte de conjuntos de medidas con marginales fijas. Se verá que los puntos extremos del conjunto de medidas doblemente estocásticas están soportadas en conjuntos que no contienen ciclos. También se expone una caracterización completa del soporte de puntos extremos del conjunto de medidas finitas discretas con marginales fijas. En ambos resultados, el teorema de Douglas-Lindenstrauss es escencial para las demostraciones. Las referencias principales de éste capítulo son [Den80] y [HW95].

El cuarto capítulo, sobre medidas con momentos fijos, esta basado principalmente en los resultados de [Gir95] y se dedica en mayor parte al problema generalizado de momentos truncos. Se exhiben resultados sobre el soporte de puntos extremos y se muestra como estos últimos generan al todo el conjunto solución por cerradura débil.

Por último, el quinto capítulo trata el Teorema de Birkhoff-von Neumann y el llamado problema 111 de Birkhoff en el marco de la teoría desarrollada principalmente en el capítulo 3. En el capítulo se estudian por separado tres casos: finito, infinito numerable e infinito no numerable. Para los casos finito e infinito numerable se consideran matrices finitas e infinitas doblemente estocásticas, se identifican los puntos extremos y se demuestra que estos generan al conjunto total, para el caso infinito numerable es necesario definir una topología adecuada, la que se presenta en éste trabajo fue propuesta en [Ken60]. En el caso no numerable se identifica la clase de medidas doblemente estocásticas con la de operadores lineales doblemente estocásticos (también conocidos como operadores de Markov), se encuentran los puntos extremos y se demuestra que son un conjunto generador, esta identificación se planteó en [Bro66]. Otra referencia importante para este capítulo es [KS04].

Capítulo 1

Preliminares

En este capítulo se presentan las definiciones básicas de conjuntos convexos, conjuntos cónicos, puntos extremos y resultados técnicos que se usarán mas adelante. Los resultados para conjuntos convexos en $\mathbb{R}^n$ serán utilizados en el capítulo 4, aunque estos conceptos pueden ser generalizados a espacios más generales.

1.1. Conjuntos convexos

Definición 1.1. Sea X un espacio vectorial sobre $\mathbb{R}$. Un conjunto $C \subset X$ es *convexo* si contiene todos los segmentos cuyos extremos son elementos de C. Esto es, si α_1, α_2 son números reales tales que $\alpha_1, \alpha_2 \geq 0$ y $\alpha_1 + \alpha_2 = 1$, entonces para todo $x_1, x_2 \in C$,

$$\alpha_1 x_1 + \alpha_2 x_2 \in C.$$

Definición 1.2. Una *combinación convexa* de los puntos $x_1, x_2, \ldots, x_n$, es una expresión de la forma

$$\sum_{i=1}^{n} \alpha_i x_i,$$

donde $\sum_{i=1}^{n} \alpha_i = 1$ y $\alpha_i \geq 0$, $i = 1, \ldots, n$.

Proposición 1.3. *Un conjunto es convexo si y sólo si contiene a todas las combinaciones convexas de sus elementos.*

Demostración. Probamos por inducción. Sea C convexo, $\{x_1, \cdots, x_n\} \subset C$ con $n \geq 2$. Sea $x = \sum_{i=1}^{n} \alpha_i x_i$ una combinacion convexa. Por definición, si $n = 2$ entonces $x \in C$. Supongamos ahora que $n = k$ y $\sum_{i=1}^{k} \alpha_i = 1$ implica que $\sum_{i=1}^{k} \alpha_i x_i \in C$. Sea

$\sum_{i=1}^{k+1} \alpha_i x_i$ una combinación convexa de $k+1$ puntos de C, con $\alpha_i \in (0,1)$ para toda i, entonces $\sum_{i=1}^{k} \alpha_i = (1-\alpha_{k+1}) \neq 0$. Hacemos

$$y = \sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_{k+1}} x_i,$$

entonces $y \in C$ por hipótesis de inducción. Luego

$$(1-\alpha_{k+1})y + \alpha_{k+1}x_{k+1} = \sum_{i=1}^{k+1} \alpha_i x_i \in C.$$

Por lo tanto C contiene a todas sus combinaciones convexas. La otra implicación es inmediata. □

Definición 1.4. Dado un conjunto no vacío, $S \subset X$, el *casco convexo* de S, denotado $\mathrm{cc}(S)$, es la intersección de todos los conjuntos convexos que contienen a S.

Observación 1.5. Es fácil ver de la definición de convexidad que la intersección de conjuntos convexos es convexa, entonces el casco convexo de un conjunto no vacío S es el conjunto convexo más pequeño que lo contiene.

Proposición 1.6. *El casco convexo de S es el conjunto de todas las combinaciones convexas de elementos de S. Es decir,*

$$cc(S) = \left\{ \sum_{i=1}^{n} \alpha_i s_i : \alpha_i \geq 0, \sum_{i=1}^{n} \alpha_i = 1, s_i \in S, n \geq 1 \right\}.$$

Demostración. Sea $D = \{\sum_{i=1}^{n} \alpha_i s_i : \alpha_i \geq 0, \sum_{i=1}^{n} \alpha_i = 1, s_i \in S, n \geq 1\}$.
Como $\mathrm{cc}(S)$ es convexo y $S \subset \mathrm{cc}(S)$, se sigue de la Proposición 1.3 que $D \subset \mathrm{cc}(S)$.

Ahora, sean $x, y \in D$. Entonces $x = \sum_{i=1}^{n} \alpha_i x_i$, $y = \sum_{j=1}^{m} \beta_j y_j$ para algunas α_i, $\beta_j \in [0,1]$, x_i, $y_j \in S$, $i = 1, ..., n$, $j = 1, ..., m$ tal que $\sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{m} \beta_i = 1$. Sea $t \in [0,1]$, se tiene que

$$\begin{aligned} tx + (1-t)y &= t \sum_{i=1}^{n} \alpha_i x_i + (1-t) \sum_{j=1}^{m} \beta_j y_j \\ &= \sum_{i=1}^{n} (t\alpha_i) x_i + \sum_{j=1}^{m} ((1-t)\beta_j) y_j \qquad (1.1) \end{aligned}$$

es una combinación lineal de puntos en S con $0 \leq (t\alpha_i)$, $(1-t)\beta_j \leq 1$ para toda $i = 1, ...n$, $j = 1, ...m$. Además

$$\begin{aligned}\sum_{i=1}^{n} t\alpha_i + \sum_{j=1}^{m} (1-t)\beta_j &= \sum_{i=1}^{n} t\alpha_i + \sum_{j=1}^{m} \beta_j - \sum_{j=1}^{m} t\beta_j \\ &= t\left(\sum_{i=1}^{n} \alpha_i - \sum_{j=1}^{m} \beta_j\right) + 1 = 1.\end{aligned}$$

Es decir, (1.1) es una combinación convexa de elementos de S. Entonces D es convexo y contiene a S. Se sigue que $\mathrm{cc}(S) \subset D$ y por lo tanto $\mathrm{cc}(S) = D$. □

Definición 1.7. Si X además es un espacio topológico se define el *casco cerrado convexo* de S, denotado $\overline{\mathrm{cc}}(S)$, como la intersección de todos los conjuntos convexos y cerrados que contienen a S.

Observación 1.8. $\overline{\mathrm{cc}}(S)$ es el convexo cerrado mas pequeño que contiene a S.

Denotaremos por $\overline{A}$ a la cerradura topológica de A y por ∂A a su frontera.

Definición 1.9. Decimos que un espacio vectorial X sobre $\mathbb{R}$ es un *espacio vectorial topológico*, si (X, τ) es un espacio topológico y las operaciones de producto escalar y suma

$$\begin{aligned}\cdot &: (\mathbb{R}, |\cdot|) \times (X, \tau) \to (X, \tau) \\ + &: (X, \tau) \times (X, \tau) \to (X, \tau)\end{aligned}$$

son funciones continuas.

Lema 1.10. *Sea X un espacio vectorial topológico y $\emptyset \neq C \subset X$ convexo, entonces $\overline{C}$ es convexo.*

Demostración. Sea $C \subset X$ convexo. Para cada $t \in [0, 1]$, definimos la función $s_t : X \times X \to X$ como $s_t(x, y) := tx + (1-t)y$. Por hipótesis, s_t es continua y además C es cerrado bajo s_t, es decir, si $x, y \in C$ entonces $s_t(x, y) \in C$. Luego, como el producto cartesiano finito de cerraduras es la cerradura del producto

$$s_t[\overline{C} \times \overline{C}] = s_t[\overline{C \times C}],$$

por continuidad de s_t

$$s_t[\overline{C \times C}] \subset \overline{s_t[C \times C]}$$

y como $s_t[C \times C] \subset C$ tenemos que

$$\overline{s_t[C \times C]} \subset \overline{C}$$

por lo tanto

$$s_t[\overline{C} \times \overline{C}] \subset \overline{C}.$$

Es decir, $\overline{C}$ es cerrado bajo combinaciones convexas, por lo tanto es convexo. □

Teorema 1.11. *Sea X un espacio vectorial topológico y $S \subset X$, entonces*

$$\overline{cc(S)} = \overline{cc}(S).$$

Demostración. Por un lado $\mathrm{cc}(C) \subset \overline{\mathrm{cc}}(C)$ pues $\overline{\mathrm{cc}}(C)$ es un convexo que contiene a C. Por lo que

$$\overline{\mathrm{cc}(C)} \subset \overline{\overline{\mathrm{cc}}(C)} = \overline{\mathrm{cc}}(C).$$

Además, por el Lema 1.10 $\overline{\mathrm{cc}(C)}$ es convexo, cerrado y contiene a C, entonces $\overline{\mathrm{cc}}(C) \subset \overline{\mathrm{cc}(C)}$ y por lo tanto $\overline{\mathrm{cc}(S)} = \overline{\mathrm{cc}}(S)$. □

A modo de ejemplo consideremos un conjunto de puntos $\{x_i\}$ en $\mathbb{R}^n$ y $S = cc(\{x_i\})$, entonces S es el conjunto de todas las combinaciones convexas de los puntos $\{x_i\}$. En este caso decimos que $\{x_i\}$ *genera* a S. El siguiente teorema nos dice aún más, que cada elemento de S se puede expresar como una combinación convexa de $n+1$ puntos en $\{x_i\}$. Este hecho cobrará importancia cuando consideremos los puntos extremos.

Teorema 1.12 (Carathéodory)**.** *Sea $C \subset \mathbb{R}^n$. Dado $x \in cc(C)$ entonces*

$$x = \sum_{i=1}^{m} \alpha_i x_i$$

para algunos $x_i \in C$, $\alpha_i > 0$, con $i = 1, \ldots, m$, $\sum_{i=1}^{m} \alpha_i = 1$ y $m \leq n+1$.

Demostración. Mostraremos que si $k > n$ y $x = \sum_{i=1}^{k+1} t_i x_i$ es una combinación convexa de $k+1$ puntos $x_i \in \mathbb{R}^n$, entonces x es una combinación convexa de k de éstos.

Supongamos que $t_i > 0$ para $i = 1, 2, ..., k+1$. Consideremos la función lineal

$$(a_1, ..., a_{k+1}) \to \left(\sum_{i=1}^{k+1} a_i x_i, \sum_{i=1}^{k+1} a_i \right)$$

de $\mathbb{R}^{k+1}$ en $\mathbb{R}^n \times \mathbb{R}$. Su espacio nulo debe ser de dimensión positiva pues $k > n$. Entonces existe $(a_1, ..., a_{k+1})$, con $a_i \neq 0$ para alguna i, tal que $\sum_{i=1}^{k+1} a_i x_i = 0$ y $\sum_{i=1}^{k+1} a_i = 0$. Como $t_i > 0$ para toda i, existe una constante α tal que $|\alpha a_i| \leq t_i$ para toda i y $\alpha a_j = t_j$ para al menos una j. Hacemos $c_i = t_i - \alpha a_i$, entonces $x = \sum_{i=1}^{k+1} c_i x_i$ donde al menos una de estas c_j es 0. Además $\sum_{i=1}^{k+1} c_i = \sum_{i=1}^{k+1} t_i = 1$ y $c_i \geq 0$ para toda i. □

Definición 1.13. Dado un vector no nulo $a = (a_1, ..., a_n) \in \mathbb{R}^n$ y $h \in \mathbb{R}$, un conjunto de la forma

$$\left\{ x = (x_1, ..., x_n) : \sum_1^n a_i x_i = h \right\},$$

se le llama *hiperplano* con normal a. Identificaremos a cada hiperplano con la ecuación que lo determina ya sea como $\sum_1^n a_i x_i = h$ o bien $a \cdot x = h$.

Observación 1.14. Un hiperplano $\mathcal{H}$ con ecuación $\sum_{i=1}^n a_i x_i = h$ o $a \cdot x = h$ genera en $\mathbb{R}^n$ dos semi-espacios abiertos

$$\mathcal{H}^+ := \left\{ x = (x_1, ..., x_n) : \sum_{i=1}^n a_i x_i > h \right\},$$

$$\mathcal{H}^- = \left\{ x = (x_1, ..., x_n) : \sum_{i=1}^n a_i x_i < h \right\}$$

y dos semi-espacios cerrados

$$\overline{\mathcal{H}^+} := \left\{ x = (x_1, ..., x_n) : \sum_{i=1}^n a_i x_i \geq h \right\},$$

$$\overline{\mathcal{H}^-} = \left\{ x = (x_1, ..., x_n) : \sum_{i=1}^n a_i x_i \leq h \right\}.$$

Definición 1.15. Decimos que un hiperplano $\mathcal{H}$ es un *hiperplano soporte* de un conjunto S, si $S \subset \overline{\mathcal{H}^+}$ o $S \subset \overline{\mathcal{H}^-}$ y $\mathcal{H}$ contiene al menos un punto frontera de S. A los semi-espacios que contienen al S los llamamos *hiperplanos soporte*.

Teorema 1.16 (Minkowsky). *En $\mathbb{R}^n$, cualquier punto p que es exterior a un conjunto convexo C puede ser separado de C por un hiperplano. Es decir, existe un hiperplano $\mathcal{H}$ tal que $C \subset \mathcal{H}^+$ y $p \in \mathcal{H}^-$ o viceversa.*

Demostración. Podemos suponer sin pérdida de generalidad que $p = 0$, pues podemos reemplazar C por $\{x - p : x \in C\}$. Ahora, como 0 es exterior a C tenemos que $\inf_{x\in C} \{d(x, c)\} > 0$, donde $d(x, y)$ es la métrica usual en $\mathbb{R}^n$. Entonces la cerradura $\overline{C}$ contiene un punto a que tal que la distancia de 0 a a es mínima con respecto a todas las distancias de 0 a x, para $x \in \overline{C}$. Entonces el hiperplano $a \cdot x = 0$, que pasa por el origen y es perpendicular a a no contiene puntos de $\overline{C}$.

Ahora sea x_0 un punto interior arbitrario del segmento que une a 0 y a. Este punto es exterior a $\overline{C}$ y el punto más cercano a este en $\overline{C}$ es a. Por el mismo razonamiento, el hiperplano $a \cdot x = a \cdot x_0$ no contiene puntos de $\overline{C}$ y además separa a 0 y a $\overline{C}$ □

Corolario 1.17. *Por cada punto frontera de un conjunto convexo $C \subset \mathbb{R}^n$ pasa un hiperplano soporte.*

Demostración. Sea $c \in \partial C$. Sea $\{c_n\}$ una sucesión de puntos exteriores a C que converge a c. Para cada c_i, consideremos un hiperplano $a_i \cdot x = a_i \cdot c_i$, que contiene a c_i y de un lado quede C, es decir, $a_i \cdot x \geq a_i \cdot c_i$ o $a_i \cdot x \leq a_i \cdot c_i$ para todo $x \in C$. Supongamos que $\|a_i\| = 1$ y que $a_i \cdot x \geq a_i \cdot c_i$ para todo $x \in C$. Tomamos una subsucesión de $\{a_i\}$ convergente a un punto a. Entonces $a \cdot x = a \cdot c$ es un hiperplano soporte. □

Corolario 1.18. *Un conjunto convexo cerado en $\mathbb{R}^n$ es la intersección de todos sus semi espacios soporte cerrados.*

Demostración. Por definición, todo punto del conjunto está en cada semi-espacio soporte. Por otro lado, cualquier punto común de los semi espacios soporte debe estar en el conjunto, de otro modo podría ser separado del conjunto por algún hiperplano soporte y entonces no pertenecería a todos sus semi-espacios soporte. □

Definición 1.19. Un conjunto K es *cónico* si $x \in K$ implica que $tx \in K$ para toda $t \geq 0$.

Teorema 1.20. *Sea $K \subsetneq \mathbb{R}^n$ un conjunto cónico y convexo. Entonces cualquier hiperplano soporte de K pasa por el origen.*

Demostración. Sea $a \cdot x = a \cdot c_0$ un hiperplano soporte de K que pasa por $c_0 \in \overline{K}$. Como $tc_0 \in \overline{K}$ para toda $t > 0$, se sigue que la diferencia $a \cdot tc_0 - a \cdot c_0 = (t - 1)a \cdot c_0$ no cambia de signo, lo cual no puede pasar si $a \cdot c_0 \neq 0$, por lo tanto el hiperplano pasa por el origen. □

Proposición 1.21. *Un conjunto cónico K es convexo si y sólo si x_1, $x_2 \in K$ implica que $x_1 + x_2 \in K$.*

Demostración. Sea K cónico y convexo y sea $t \in (0,1)$ entonces $t^{-1}x_1 \in K$, $(1-t)^{-1}x_2 \in K$. Por convexidad de K tenemos que

$$t(t^{-1}x_1) + (1-t)((1-t)^{-1}x_2) = x_1 + x_2 \in K.$$

Ahora sea K cónico y supongamos que para todo x_1, $x_2 \in K$ se cumple que $x_1+x_2 \in K$. Sea $t \in (0,1)$.Dado que K es cónico se tiene que $tx_1 \in K$ y $(1-t)x_2 \in K$, entonces

$$(tx_1) + ((1-t)x_2) \in K,$$

por lo tanto K es convexo. □

Definición 1.22. Dado un conjunto $S \subset \mathbb{R}^n$ el *casco cónico* de S es la intersección de todos los conjuntos convexos y cónicos que contienen a S. El *casco cerrado cónico* de S es la intersección de todos los conjuntos convexos, cónicos y cerrados que contienen a S.

Observación 1.23. El casco cónico de un conjunto convexo C es el conjunto de puntos de la forma tx, con $t \geq 0$ y $x \in C$. Equivalentemente, es el conjunto de combinaciones lineales positivas de elementos de C.

1.2. Puntos extremos

Definición 1.24. Sea C un subconjunto de un espacio vectorial X. Un subconjunto no vacío $E \subset C$ se llama *conjunto extremo* de C si siempre que $x_1 \in C$, $x_2 \in C$, $0 < \alpha < 1$, y

$$\alpha x_1 + (1-\alpha)x_2 \in E,$$

entonces $x_1 \in E$ y $x_2 \in E$.

Es decir, ningún punto de E es punto intermedio de algún segmento cuyos extremos son elementos de C, excepto cuando ambos extremos del segmento están en E. Por ejemplo si E es convexo entonces E es conjunto extremo de E.

Definición 1.25. Los *puntos extremos* de C son los conjuntos extremos de C que consisten de sólo un punto. Entonces $x \in C$ es punto extremo de C si siempre que $x_1 \in C$, $x_2 \in C$, $\alpha \in (0,1)$ y

$$x = \alpha x_1 + (1-\alpha)x_2$$

entonces $x_1 = x_2 = x$.

Observación 1.26. De la definición se sigue que $x \in C$ no es punto extremo de C si y sólo si existen $x_1, x_2 \in C$, $x_1 \neq x_2$ y $\alpha > 0$ tales que $x = \alpha x_1 + (1-\alpha)x_2$

Proposición 1.27. *Sea C convexo. Si $x \in C$ no es punto extremo, entonces existen $y_1, y_2 \in C$, $y_1 \neq y_2$ tales que $x = \frac{1}{2}(y_1 + y_2)$.*

Demostración. Si $x \in C$ no es punto extremo de C, entonces existen $x_1, x_2 \in C$, $x_1 \neq x_2$ y $0 < t_0 < 1$ tales que $x = t_0 x_1 + (1-t_0)x_2$. Sin perdida de generalidad supongamos que $t_0 > \frac{1}{2}$, si hacemos $t = 2t_0 - 1$ entonces $t \in (0,1)$. Sean $y_1 = x_1$, $y_2 = tx_1 + (1-t)x_2$. Tenemos que

$$\begin{aligned}
\frac{1}{2}y_1 + \frac{1}{2}y_2 &= \frac{1}{2}x_1 + \frac{1}{2}\left[tx_1 + (1-t)x_2\right] \\
&= \frac{1}{2}x_1 + \frac{1}{2}tx_1 + \frac{1-t}{2}x_2 \\
&= \frac{1+t}{2}x_1 + \frac{1-t}{2}x_2 \\
&= t_0x_1 + (1-t_0)x_2 = x.
\end{aligned}$$

□

Denotamos al conjunto de puntos extremos de C como $E(C)$. La relevancia de los puntos extremos se hace clara por el Teorema de Krein-Milman, el cual nos dice que bajo ciertas condiciones del espacio, los puntos extremos de un convexo compacto son un conjunto generador.

Lema 1.28. *Sea X un espacio vectorial topológico y $\emptyset \neq K \subset X$ un conjunto convexo y compacto. Sea $\mathcal{P}$ la colección de todos los conjuntos extremos y compactos de K, entonces*

(a) La intersección S de cualquier subcolección no vacía de $\mathcal{P}$ es un elemento de $\mathcal{P}$, a menos que $S = \emptyset$.

(b) Si $S \in \mathcal{P}$, $f : X \to R$ lineal y continua, $m = \max\{f(x) : x \in S\}$ y

$$S_f = \{x \in S : f(x) = m\},$$

Entonces $S_f \in \mathcal{P}$.

Demostración. Sea $S = \bigcap_\alpha E_\alpha \neq \emptyset$ con $E_\alpha \in \mathcal{P}$ y $x \in S$ Si $x = tx_1 + (1-t)x_2$ con $t \in [0,1]$ y x_1, $x_2 \in K$. $x \in S$ implica que $x \in E_\alpha$ para toda α, entonces x_1, $x_2 \in E_\alpha$ para toda α por lo tanto x_1, $x_2 \in S$ y $S \in \mathcal{P}$. Lo cual prueba (a).

Ahora sea $S \in \mathcal{P}$ y $z = tx + (1-t)y \in S_f$ con x, $y \in K$, $t \in (0,1)$. Como $z \in S$ y $S \in \mathcal{P}$ entonces x, $y \in S$. Por lo que $f(x) \leq m$ y $f(y) \leq m$. Como $z \in S_f$, $f(z) = m$. Luego $m = f(z) = tf(x) + (1-t)f(y)$ implica que $f(x) = f(y) = m$. Entonces x, $y \in S_f$. Por lo tanto $S_f \in \mathcal{P}$.

□

Teorema 1.29 (Krein-Milman)**.** *Sea X un espacio vectorial topológico localmente convexo. Si K es un conjunto convexo compacto no vacío en X, entonces K es el casco cerrado convexo de sus puntos extremos, es decir*

$$K = \overline{cc}(E(K)).$$

Demostración. Sea $\mathcal{P}$ la colección de todos los conjuntos extremos y compactos de E. Sea $S \in \mathcal{P}$ y $\mathcal{P}'$ la colección de todos los elementos de $\mathcal{P}$ que son subconjuntos de S. Como $S \in \mathcal{P}$, $\mathcal{P}' \neq \emptyset$. Ordenamos parcialmente a $\mathcal{P}'$ por contención, sea Ω una subcoleción de $\mathcal{P}'$ totalmente ordenada y maximal, y sea M a intersección de todos los miembros de Ω, entonces M es no vacío pues la colección Ω de conjuntos compactos tiene la propiedad de la intersección finita. Entonces por el inciso (a) del lema 1.28 $M \in \mathcal{P}$. Por maximalidad de Ω, ningún subconjunto propio de M pertenece a $\mathcal{P}$. Se sigue de (b) del lema 1.28 que todo funcional lineal continuo f en X es constante en M. Entonces M consta de un sólo punto y M es punto extremo de K. Tenemos entonces que

$$E(K) \cap S \neq \emptyset$$

para toda $S \in \mathcal{P}$. Es decir, todo conjunto extremo y compacto de K contiene un punto extremo de K.

Se sigue de la convexidad de K que

$$\overline{\mathrm{cc}}(E(K)) \subset \overline{\mathrm{cc}}(K) = K$$

y por lo tanto $\overline{\mathrm{cc}}(E(K))$ es compacto.

Supongamos ahora que existe $x_0 \in K - \overline{\mathrm{cc}}(E(K))$ entonces existe $f \in X^*$ (ver Teorema A.1 del apéndice) tal que $f(x) < f(x_0)$ para todo $x \in \overline{\mathrm{cc}}(E(K))$. Definimos K_f como en el lema 1.28, entonces $K_f \in \mathcal{P}$. La elección de f implica que $K_f \cap \overline{\mathrm{cc}}(E(K)) = \emptyset$, lo que es una contradicción. Por lo tanto $K = \overline{cc}(E(K))$. □

Capítulo 2

Medidas extremas y densidad de subespacios

Como se vió en el capítulo anterior, es importante conocer los puntos extremos de un conjunto convexo. Este capítulo se enfoca en dos teoremas: el Teorema de Douglas y el Teorema de Lindenstrauss. Ambos teoremas relacionan la propiedad de ser punto extremo de un conjunto de medidas convexo con la densidad en L_1 de un subespacio de funciones. Se muestra que el Teorema de Lindenstrauss es un caso particular del Teorema de Douglas, por esta razón, nos referimos a éste resultado como el Teorema de Douglas-Lindestrauss.

Se introducen también las medidas doblemente estocásticas y medidas con marginales fijas. Éstas serán el centro de atención en el siguiente capítulo y tienen un papel importante en el capítulo 5.

2.1. Teorema de Douglas

Sea X un espacio Hausdorff localmente compacto, sea $\mathcal{M}_+(X)$ el espacio de medidas de Borel definidas en X, regulares y no negativas. Sea $\mathcal{F}$ un espacio vectorial de funciones con valores reales definidas en X, Borel-medibles y que contiene a las funciones constantes. Sea $\mu \in \mathcal{M}_+(X)$ tal que

$$\int_X |f| d\mu < \infty, \qquad \text{para toda } f \in \mathcal{F},$$

en particular tenemos que $\mu(X) < \infty$. Definimos $E_\mu = E_\mu(\mathcal{F})$ como sigue

$$E_\mu := \left\{ \nu \in \mathcal{M}_+(X) : \int_X |f| d\nu < \infty \text{ y } \int_X f d\nu = \int_X f d\mu \quad \forall f \in \mathcal{F} \right\}. \tag{2.1}$$

Proposición 2.1. *E_μ es convexo*

Demostración. Sean ν_1, $\nu_2 \in E_\mu$, $f \in \mathcal{F}$ y $\alpha \in [0,1]$. Entonces

$$\begin{aligned} \int_X f d(\alpha\nu_1 + (1-\alpha)\nu_2) &= \int_X \alpha f d\nu_1 + \int_X (1-\alpha) f d\nu_2 \\ &= \alpha \int_X f d\mu + (1-\alpha) \int_X f d\mu \\ &= \int_X f d\mu. \end{aligned}$$

Es decir, $(\alpha\nu_1 + (1-\alpha)\nu_2) \in E_\mu$ y por lo tanto E_μ es convexo. □

El siguiente teorema caracteriza el conjunto de puntos extremos de $E_\mu(\mathcal{F})$.

Teorema 2.2 (Douglas)**.** *Sea $\mathcal{F}$ un subespacio vectorial de funciones que contiene a las funciones constantes. Entonces $\mathcal{F}$ es denso en $L_1(\mu)$ si y sólo si μ es un punto extremo de $E_\mu(\mathcal{F})$.*

Demostración. Supongamos que μ no es punto extremo de E_μ, entonces existen medidas μ_1, μ_2 en E_μ tales que $\mu = (\mu_1 + \mu_2)/2$ y $\mu_1 \neq \mu_2$. Esto implica que $2\mu \geq \mu_1 \geq 0$, entonces μ_1 es absolutamente continua con respecto a μ. Por el teorema de Radon-Nikodym existe una función $h \in L_\infty(\mu)$ tal que $d\mu_1 = hd\mu$ y $(1-h) \neq 0$. La función $(1-h) \in L_\infty(\mu)$ pues $\mathcal{F}$ contiene a las constantes, además cumple que:

$$\int_X f(1-h) d\mu = \int_X f d\mu - \int_X f h d\mu = \int_X f d\mu - \int_X f d\mu_1 = 0$$

para toda $f \in \mathcal{F}$. Por lo tanto, $\mathcal{F}$ no es denso en $L_1(\mu)$.

Ahora supongamos que $\mathcal{F}$ no es denso en $L_1(\mu)$, se sigue entonces del Teorema de Hahn-Banach y de la identificación $L_1^*(\mu) = L_\infty(\mu)$ que existe una función $h \in L_\infty$ distinta de cero tal que que $\langle f, h \rangle_{L_1(\mu) - L_\infty(\mu)} = \int f h d\mu = 0$ para toda $f \in \mathcal{F}$. Sea

$$\nu = \frac{1}{||h||_\infty} \int h d\mu, \qquad \mu_1 = \mu + \nu, \qquad \mu_2 = \mu - \nu.$$

Entonces las medidas μ_1 y μ_2 son positivas porque $1 \pm h/||h||_\infty \geq 0$. Además, μ_1 y μ_2 están en E_μ, pues:

$$\int_X fd(\mu \pm \nu) = \int_X fd\mu \pm \int_X fd\nu = \int_X fd\mu \pm \frac{1}{||h||_\infty} \int_X fhd\mu = \int_X fd\mu.$$

Por lo tanto, μ no es punto extremo de E_μ, pues $\mu = \frac{1}{2}(\mu_1 + \mu_2)$ y $\mu_1 \neq \mu_2$.

$\square$

En la siguiente sección veremos el Teorema de Lindestrauss, que no es más que un caso particular del Teorema de Douglas cuando se consideran medidas llamadas doblemente estocásticas. Al mismo tiempo generalizamos esta situacion a conjuntos de medidas con marginales fijas.

2.2. Teorema de Lindenstrauss.

Definición 2.3. Sean (X, Σ_1), (Y, Σ_2) dos espacios medibles y $(X \times Y, \Sigma_1 \times \Sigma_2)$ el correspondiente espacio producto medible. Si μ es una medida en $(X \times Y, \Sigma_1 \times \Sigma_2)$ se definen las marginales de μ en X y Y como

$$\mu_X(A) := \int_{A \times Y} 1 d\mu = \mu(A \times Y)$$

y

$$\mu_Y(B) := \int_{X \times B} 1 d\mu = \mu(X \times B)$$

con $A \in \Sigma_1$ y $B \in \Sigma_2$ de modo que μ_X y μ_Y son medidas sobre Σ_1 y Σ_2 respectivamente.

Si $X = Y = I$, donde I es el intervalo unitario $[0, 1] \subset \mathbb{R}$ y m es la medida de Lebesgue sobre los borelianos de I. Una medida positiva y regular μ en $I \times I$ se llama *doblemente estocástica* si para cada boreliano $A \subset I$ se tiene

$$\mu(A \times I) = \mu(I \times A) = m(A)$$

Es decir, la medida μ y la medida producto $m \times m$ coinciden en los cilindros medibles $A \times I$, $I \times A$ es la medida producto $(m \times m)$. O dicho de otra forma, las marginales de μ coinciden con m. Por ejemplo, la medida producto $m \times m$ es doblemente estocástica.

Podemos generalizar este concepto de la siguente manera. Sean (X, Σ_1, m_1) y (Y, Σ_2, m_2) dos espacios de medida finita. Consideremos ahora el conjunto de medidas en $(X \times Y, \Sigma_1 \times \Sigma_2)$ con marginales m_1 y m_2 en X y Y respectivamente. Es decir, medidas positivas μ sobre $X \times Y$ que cumplan $\mu(A \times Y) = m_1(A)$ y $\mu(X \times B) = m_2(B)$ para cualesquiera $A \subset X$, $B \subset Y$ medibles. Dicho de otra forma, medidas positivas tales que $\mu_X = m_1$ y $\mu_Y = m_2$.

Ahora, si μ_1, μ_2 son dos medidas sobre $X \times Y$ con marginales m_1 y m_2 en X y Y respectivamente y $\lambda \in [0,1]$, entonces si definimos $\mu := \lambda\mu_1 + (1-\lambda)\mu_2$ tenemos que:

$$\begin{aligned}\mu(A \times Y) =& \lambda\mu_1(A \times Y) + (1-\lambda)\mu_2(A \times Y)\\ =& \lambda m_1(A) + (1-\lambda)m_1(A) = m_1(A)\end{aligned}$$

y

$$\begin{aligned}\mu(X \times B) =& \lambda\mu_1(X \times B) + (1-\lambda)\mu_2(X \times B)\\ =& \lambda m_2(B) + (1-\lambda)m_2(B) = m_2(B).\end{aligned}$$

Por lo que el conjunto de medidas con marginales m_1 y m_2 es un conjunto convexo. Las medidas doblemente estocásticas son un caso particular cuando X y Y son el intervalo unitario y m_1 y m_2 son la medida de Lebesgue en éste.

Proposición 2.4. *Sea μ una medida definida en un espacio medible $(X \times Y, \Sigma_1 \times \Sigma_2)$, entonces para toda $f \in L_1(\mu_X)$ y $g \in L_1(\mu_Y)$ se cumple que*

$$\int_{X \times Y} f(x) + g(y) d\mu = \int_X f(x) d\mu_X + \int_Y g(y) d\mu_Y.$$

Demostración. Notemos que para toda $(x,y) \in X \times Y$

$$(\chi_A(x) + \chi_B(y)) = \chi_{A \times Y}(x,y) + \chi_{X \times B}(x,y).$$

Entonces

$$\begin{aligned}\int (\chi_A(x) + \chi_B(y)) d\mu(x,y) &= \int \chi_{A \times Y}(x,y) + \chi_{X \times B}(x,y) d\mu(x,y)\\ &= \int \chi_{A \times Y}(x,y) d\mu(x,y) + \int \chi_{X \times B}(x,y) d\mu(x,y)\\ &= \mu_X(A) + \mu_Y(B)\\ &= \int \chi_A(x) d\mu_X + \int \chi_B(x) d\mu_Y.\end{aligned}$$

Ahora sean

$$\phi(x) = \sum_{i=1}^{n} a_i \chi_{A_i}(x),$$
$$\psi(y) = \sum_{i=1}^{m} b_i \chi_{B_i}(y)$$

funciones simples. Entonces

$$\begin{aligned}
\int \phi(x) + \psi(y) d\mu &= \int \phi(x) d\mu + \int \psi(y) d\mu \\
&= \sum_{i=1}^{n} \alpha_i \int \chi_{A_i} d\mu + \sum_{i=1}^{m} \beta_i \int \chi_{B_i} d\mu \\
&= \sum_{i=1}^{n} \alpha_i \int \chi_{A_i} d\mu_X + \sum_{i=1}^{m} \beta_i \int \chi_{B_i} d\mu_Y \\
&= \int \phi(x) d\mu_X + \int \psi(y) d\mu_Y
\end{aligned}$$

Se sigue por los argumentos usuales de aproximación por funciones simples que si $f(x)$ y $g(y)$ son funciones medibles entonces

$$\int f(x) + g(y) d\mu = \int f d\mu_X + \int g d\mu_Y.$$

□

Proposición 2.5. *Sean m_1, m_2 medidas finitas en X y Y, μ una medida en $X \times Y$ con $\mu_X = m_1$ y $\mu_Y = m_2$. Sea*

$$M(\mu) := \{\nu : \nu_X = \mu_X = m_1, \quad \nu_Y = \mu_Y = m_2\}$$

el conjunto de todas las medidas en $X \times Y$ con marginales m_1 y m_2 en X y Y respectivamente. Sea

$$\mathcal{F} := \{f(x) + g(y) : f \in L_1(m_1), \quad g \in L_1(m_2)\},$$

entonces

$$E_\mu(\mathcal{F}) = M(\mu)$$

con $E\mu(\mathcal{F})$ definido en (2.1).

Lo que nos dice esta proposición es que el conjunto de medidas que integran igual a los elementos de $\mathcal{F}$ coincide con el conjunto de todas las medidas con marginales m_1 y m_2. De forma que podemos aplicar el teorema de Douglas al conjunto de todas las medidas con esas marginales y a al conjunto $\mathcal{F}$ definido arriba.

Demostración. Sea $\nu \in E_\mu(\mathcal{F})$, entonces si $(f(x) + g(y)) \in \mathcal{F}$:

$$\begin{aligned} \int_{X\times Y} (f(x)+g(y))d\nu &= \int_{X\times Y} (f(x)+g(y))d\mu \\ &= \int_X f(x)dm_1 + \int_Y g(y)dm_2 \end{aligned}$$

y por otro lado

$$\int_{X\times Y} (f(x)+g(y))d\nu = \int_X f(x)d\nu_X + \int_Y g(y)d\nu_Y$$

de donde

$$\int_X f(x)d\nu_X + \int_Y g(y)d\nu_Y = \int_X f(x)dm_1 + \int_Y g(y)dm_2$$

y la igualdad se da para toda $f \in L_1(m_1)$, $g \in L_1(m_2)$. Tomando $f \equiv 0$ y $g = \chi_B$, con $B \subset Y$ medible, tenemos que

$$\nu_Y(B) = \int \chi_B d\nu_Y = \int \chi_B dm_2 = m_2(B),$$

para todo $B \subset Y$ medible, por lo que $\nu_Y = m_2$. Similarmente,

$$\nu_X(A) = \int \chi_A d\nu_X = \int \chi_A dm_1 = m_1(A)$$

para todo $A \subset X$ medible, por lo que $v_X = m_1$. Entonces $\nu \in M(\mu)$.

Ahora, si ν es una medida con marginales m_1 y m_2 la igualdad tambien se cumple para toda $f \in L_1(m_1)$, $g \in L_1(m_2)$ de donde $\nu \in E_\mu(\mathcal{F})$ y $E_\mu(\mathcal{F})$ es justamente el conjunto de todas las medidas en $X \times Y$ con marginales m_1 y m_2 en X y Y, respectivamente. $\square$

Formulamos ahora el teorema de Lindenstrauss que no es más que el Teorema de Douglas en el caso particular de medidas en un espacio producto con marginales fijas:

Teorema 2.6 (Lindenstrauss)**.** *Sean X, Y espacios Hausdorff localmente compactos. Sean m_1, m_2 medidas regulares sobre X y Y respectivamente, con $m_1(X) < \infty$, $m_2(Y) < \infty$. Sea μ una medida en $X \times Y$ tal que $\mu_X = m_1$ y $\mu_Y = m_2$. Sea $\mathcal{F}$ el subespacio de $L_1(\mu)$ que consiste de todas las funciones de la forma $(f(x)+g(y))$ con $f \in L_1(m_1)$ y $g \in L_1(m_2)$. Entonces μ es punto extremo del conjunto de medidas con marginales m_1 y m_2 si y sólo si $\mathcal{F}$ es denso en $L_1(\mu)$*

Demostración. Se sigue del la teorema de Douglas y de la Proposición 2.5. □

Capítulo 3

Medidas con marginales fijas

Como una aplicación del Teorema de Douglas-Lindenstrauss caracterizaremos el soporte de medidas con marginales fijas en dos casos distintos. En el primer caso, veremos que los puntos extremos del conjunto de medidas doblemente estocásticas estan soportadas en conjuntos que no contienen ciclos. Una consecuencia interesante de este hecho, es que los puntos extremos resultan ser medidas singulares respecto a la medida de Lebesgue en el cuadrado unitario. El segundo resultado, el Teorema de Denny, es una caracterización completa del soporte de los puntos extremos del conjunto de medidas discretas con marginales fijas. Éste resulta ser la unión de las gráficas de una pareja aperiódica de funciones. Se muestra también que en éste caso, el soporte de medidas extremas no puede contener ciclos. Este último hecho se utilizará en el capítulo 5 para identificar los puntos extremos del conjunto matrices doblemente estocásticas y del de matrices doblemente estocásticas infinitas.

3.1. Soporte de medidas doblemente estocásticas

Definición 3.1. Un conjunto de puntos $\{(x_i, y_i)\}_{i=1}^{2n} \subset [0,1]^2$ es un *ciclo* si las (x_i, y_i) son disntintas, $y_{2i} = y_{2i-1}$, $x_{2i+1} = x_{2i}$ y $x_1 = x_{2n}$ para toda $1 \leq i \leq 2n$.

Definición 3.2. Sea $P_n = \{[0, 2^{-n}]\} \cup \{(i2^{-n}, (i+1)2^{-n}]\}_{i=1}^{2^n-1}$. Y $P_n \times P_n$ la partición de $[0,1]^2$ inducida por P_n. Un $P_n \times P_n$-*ciclo básico* es un subconjunto Γ de $P_n \times P_n$ tal que los centros de los cuadrados de Γ forman un ciclo.

Definición 3.3. Dado Γ un $P_n \times P_n$-ciclo básico, un ciclo $\{(x_i, y_i)\}_{i=1}^{\#(\Gamma)}$, donde $\#(\Gamma)$ denota la cardinalidad de Γ, es un Γ-*ciclo* si cada (x_i, y_i) está en un elemento distinto de Γ. En este caso, también decimos que $\{(x_i, y_i)\}_{i=1}^{\#(\Gamma)}$ es un $P_n \times P_n$-ciclo.

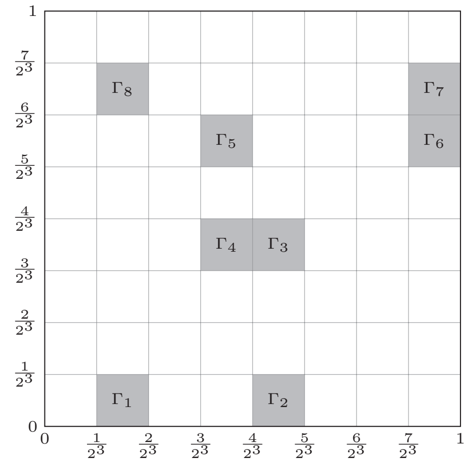

Figura 3.1: Ejemplo de un $P_3 \times P_3$-ciclo básico, $\Gamma = \bigcup_{i=1}^{8} \Gamma_i$.

Lema 3.4. *Si μ es una medida doblemente estocástica extrema entonces dado m un entero positivo y $\epsilon > 0$, existe $n \geq m$ y $\Pi \subset P_n \times P_n$ tal que $\bigcup_{O \in \Pi} O$ no contiene ningún $P_m \times P_m$-ciclo y $\mu\left(\bigcup_{O\in\Pi} O\right) > 1 - \epsilon$.*

Demostración. Sean m y ϵ fijas. Por el Teorema de Douglas-Lindenstrauss, dada $\delta > 0$, para cada $L \in P_m \times P_m$, existen funciones f_L y g_L en $L_1([0,1],\lambda)$ tales que

$$\int |\chi_L(x,y) - f_L(x) - g_L(y)| d\mu < \delta.$$

Como las funciones simples en $\bigcup_k P_k$ son densas en L_1, podemos tomar f_L y g_L funciones simples en P_n para alguna $n \geq m$. Sea $a(m)$ un entero finito tal que $\#(\Gamma) < a(m)$ para todo $P_m \times P_m$-ciclo Γ. Fijemos un $P_m \times P_m$-ciclo básico Γ, sea $\{(x_i, y_i)\}_{i=1}^{2n}$ su ciclo de centros y elijamos $L(\Gamma) \in \Gamma$ tal que $(x_1, y_1) \in L(\Gamma)$. Definimos

$$D(\Gamma) = \left\{ B \in P_n \times P_n : \chi_B |\chi_{L(\Gamma)} - f_{L(\Gamma)}(x) - g_{L(\Gamma)}(y)| \leq \frac{1}{a(m)} \right\}.$$

Sea $U(\Gamma) = \bigcup_{B \in D(\Gamma)} B$. Tenemos que

$$\frac{1}{a(m)} \mu\{(U(\Gamma))^c\} \leq \int_{U(\Gamma)^c} |\chi_{L(\Gamma)} - f_{L(\Gamma)} - g_{L(\Gamma)}(y)| d\mu \leq \delta,$$

y entonces

$$1 - \delta a(m) < \mu\left(U(\Gamma)\right).$$

Veremos ahora que $U(\Gamma)$ no tiene ningún Γ-ciclo. Supongamos que $\{(x_i, y_i)\}_{i=1}^{\#(\Gamma)}$ es un Γ-ciclo contenido en $U(\Gamma)$. Por definición de $L(\Gamma)$ podemos tomar los índices de forma que $(x_1, y_1) \in L(\Gamma)$. Entonces tenemos que

$$\begin{aligned} 1 = \chi_{L(\Gamma)}(x_1, y_1) &= \sum_{i=1}^{\#(\Gamma)} (-1)^{i+1} \left(\chi_{L(\Gamma)}(x_i, y_i) - f_{L(\Gamma)}(x_i) - g_{L(\Gamma)}(y_i)\right) \\ &\leq \sum_{i=1}^{\#(\Gamma)} |\chi_{L(\Gamma)}(x_i, y_i) - f_{L(\Gamma)}(x_i) - g_{L(\Gamma)}(y_i)| \leq \frac{\#(\Gamma)}{a(m)} < 1. \end{aligned}$$

Lo cual es un contradicción.

Ahora, sea $\{\Gamma_i\}_{i=1}^{M}$ el conjunto de todos los $P_m \times P_m$-ciclos. Para cada Γ_i sea $\delta = \delta_i$ del argumento anterior con $0 < \delta_i < \frac{\epsilon}{2^i a(m)}$. Construimos $U(\Gamma_i) = \bigcup_{B \in D(\Gamma_i)} B$. Entonces $U(\Gamma_i) \subset P_{n_i} \times P_{n_i}$ para algún n_i, $U(\Gamma_i)$ no tiene ningún Γ_i-ciclo, y

$$1 - \frac{\epsilon}{2^i} < 1 - \delta_i a(m) < \mu(U(\Gamma_i)).$$

Finalmente, sea $D = \bigcap_{i=1}^{M} U(\Gamma_i)$, D cumple que $\mu(D) > 1 - \epsilon$, no tiene ningún $P_m \times P_m$-ciclo, y $D = \bigcup_{O \in \Pi} O$ para $\Pi \subset P_n \times P_n$ con $n = \text{máx}_i\, n_i$.

□

Teorema 3.5. *Sea μ una medida doblemente estocástica extrema. Entonces existe un boreliano B sin ciclos tal que $\mu(B) = 1$*

Demostración. Por el Lema 3.4, para cada $m \geq 1$ existe un boreliano K_m tal que K_m no contiene $P_m \times P_m$-ciclos y $\mu(K_m) > 1 - 2^{-m}$. Entonces $B := \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} K_m$ no contiene ningún $P_n \times P_n$-ciclo para ninguna $n \geq 1$ y $\mu(B) = 1$. Por último notamos que cualquier ciclo debe ser un $P_n \times P_n$-ciclo para alguna n, por lo tanto B no tiene ciclos. □

Veremos como consecuencia de este resultado que las medidas doblemente estocásticas extremas son singulares respecto a la medida de Lebesgue en $[0, 1]^2$. Si un conjunto contiene las esquinas de un rectángulo, entonces contiene un ciclo. Probaremos que si un conjunto no contiene las esquinas de ningun rectangulo, entonces debe tener medida de Lebesgue cero.

Teorema 3.6. *Sea $A \subset [0, 1]^2$ un conjunto boreliano que no contenga las esquinas de ningún rectángulo, entonces A tiene medida de Lebesgue igual a cero.*

Demostración. Para la demostración usaremos el Teorema de Densidad de Lebesgue. (Apendice A)

Sea $\lambda \times \lambda$ la medida de Lebesgue en $[0,1]^2$ y sea $A \subset [0,1]^2$ con $(\lambda \times \lambda)(A) > 0$. Sea $(x,y) \in A$ un punto de densidad de Lebesgue de A. Entonces para cada $\delta > 0$ existe $\epsilon > 0$ tal que

$$(\lambda \times \lambda)(A \cap [x-\epsilon, x+\epsilon] \times [y-\epsilon, y+\epsilon]) > 4\epsilon^2(1-\delta).$$

Sea $B = A \cap ([x-\epsilon, x+\epsilon] \times [y-\epsilon, y+\epsilon])$. Dividamos B en 4 cuadrados de ancho ϵ: $B_1 = B \cap [x-\epsilon, x] \times [y-\epsilon, y]$, $B_2 = B \cap [x-\epsilon, x] \times [y, y+\epsilon]$, $B_3 = B \cap [x, x+\epsilon] \times [y-\epsilon, y]$ y $B_4 = B \cap [x, x+\epsilon] \times [y, y+\epsilon]$.

Entonces $(\lambda \times \lambda)(B_i) \geq \epsilon^2(1-4\delta)$, $i = 1,2,3,4$. Pues si $(\lambda \times \lambda)(B_j) < \epsilon^2(1-4\delta)$ para alguna j, entonces

$$(\lambda \times \lambda)(B) = (\lambda \times \lambda)(\bigcup_i B_i) \leq \sum_i (\lambda \times \lambda)(B_i) < \epsilon^2(1-4\delta) + 3\epsilon^2 = 4\epsilon^2(1-\delta)$$

Ahora, trasladamos cada B_i por $(\pm\epsilon/2, \pm\epsilon/2)$ de forma que estén centrados en (x,y). Sea C la intersección de los cuatro B_is trasladados. Entonces, como $\mu(B_i) \geq \epsilon^2(1-4\delta)$ y cada B_i esta contenido en un cuadrado de ancho ϵ, tenemos que $\mu(B_i^c) \leq \epsilon^2 4\delta$. Entonces $\mu(C^c) \leq \sum_{i=1}^4 \mu(B_i^c) \leq \epsilon^2 16\delta$, esto implica que

$$(\lambda \times \lambda)(C) \geq \epsilon^2(1-16\delta),$$

pues C esta contenido en un cuadrado de ancho ϵ.

Entonces $C \neq \emptyset$ para δ suficientemente pequeña. Elegimos $(x_0, y_0) \in C$ y sean $(x_i, y_i) = (x_0, y_0) + (\pm\epsilon/2, \pm\epsilon/2)$ para $i = 1,2,3,4$ los puntos que resultan al trasladar (x_0, y_0) a cada B_i. Entonces los cuatro puntos, (x_i, y_i), están en A y forman las esquinas de un cuadrado. $\square$

Corolario 3.7. *Si μ es punto extremo del conjunto de medidas doblemente estocásticas, entonces μ es singular con respecto a la medida de Lebesgue en $[0,1]^2$.*

Demostración. Sea μ una medida doblemente estocástica extrema, entonces existe un borelian B tal que $\mu(B) = 1$ y que no contiene ciclos. Entonces B, no contiene las esquinas de ningun rectangulo, por lo tanto $\lambda(B) = 0$ y μ es singular a λ. $\square$

3.2. Soporte de medidas extremas discretas con marginales fijas

Sean X y Y conjuntos numerables y μ una medida definida en los subconjuntos de $X \times Y$. Entonces la marginal de μ en X, μ_X, se puede escribir de la forma

$$\mu_X(x) := \mu(x \times Y) = \sum_{y \in Y} \mu(x, y)$$

y de la misma forma, la marginal de μ en Y

$$\mu_Y(y) := \mu(X \times y) = \sum_{x \in X} \mu(x, y)$$

Luego el conjunto

$$M(\mu) := \{\nu \mid \nu \text{ medida en } X \times Y, \nu_X = \mu_X, \nu_Y = \mu_Y\}$$

es convexo, pues si ν_1, $\nu_2 \in M(\mu)$ y $\nu = \lambda\nu_1 + (1 - \lambda)\nu_2$ entonces

$$\begin{aligned}
\nu_X(x) =& \sum_{y \in Y} (\alpha\nu_1 + (1 - \alpha)\nu_2)(x, y) \\
=& \sum_{y \in Y} \alpha\nu_1(x, y) + \sum_{y \in Y} (1 - \alpha)\nu_2(x, y) \\
=& \alpha \sum_{y \in Y} (\nu_1)_X(x) + (1 - \alpha) \sum_{y \in Y} (\nu_2)_X(x) \\
=& \alpha\mu_X(x) + (1 - \alpha)\mu_X(x) \\
=& \mu_X(x).
\end{aligned}$$

Análogamente $\nu_Y = \mu_Y$ por lo tanto $M(\mu)$ es convexo.

Ahora introducimos la noción de pareja aperiódica de funciones. Sean $A \subset X$, $B \subset Y$ con $A \neq \emptyset$ o $B \neq \emptyset$. Sean $f : A \to Y$, $g : B \to X$ funciones. Ponemos

$D_1 := f^{-1}[B]$. Si $D_1 \neq \emptyset$ sea $(g \circ f)(x) = g(f(x))$ para $x \in D_1$ y para $n \geq 2$ $D_n := D_{n-1} \cap ((g \circ f)^{n-1})^{-1}[D_1]$.

Observemos que para $x \in D_n$:

$$(g \circ f)^n(x) := (g \circ f)((g \circ f)^{n-1}(x)),$$

es decir, D_n es el dominio de la funcion $(g \circ f)^n$.

Definición 3.8. Una pareja de funciones (f, g), con $f : A \subset X \to Y$, $g : B \subset Y \to X$, es aperiódica si para toda $n \in \mathbb{N}$ y cada $x \in D_n$

$$(g \circ f)^n(x) \neq x.$$

Observación 3.9. Si A o B son vacíos entonces (f, g) es aperiódica.

Definición 3.10. Dadas $f : A \subset X \to Y$ y $g : B \subset Y \to X$ funciones, denotaremos por $G(f)$ y $G(g)$ a las gráficas de f y g como subconjuntos de $X \times Y$. Es decir,

$$G(f) := \{(x, f(x)) \in X \times Y : x \in A\}$$

y

$$G(g) := \{(g(y), y)) \in X \times Y : y \in B\}$$

Proposición 3.11. *Si (f, g) es aperiódica, entonces las gráficas de f y g son conjutos disjuntos de $X \times Y$.*

Demostración. Si $(x, f(x)) = (g(y), y)$ entonces $(g \circ f)(x) = g(f(x)) = g(y) = x$. □

Lema 3.12. *Sea (f, g) una pareja aperiódica con $A = Dom(f) \subset X$ no vacío. Sea $x_0 \in A$ y $A_0 := \{x_0\}$. Para cada entero $i \geq 1$ sean $B_i := g^{-1}(A_{i-1})$ y $A_i := f^{-1}(B_i)$. Entonces $A_m \cap A_k = \emptyset$ si $m > k$.*

Demostración. Si $x \in A_m \cap A_k$, con $m > k$ entonces

$$(g \circ f)^m(x) = x_0 = (g \circ f)^k(x).$$

Pero

$$(g \circ f)^{m-k}(x_0) = (g \circ f)^{m-k}((g \circ f)^k(x)) = (g \circ f)^m(x) = x_0,$$

lo cual contradice la aperiodicidad de (f, g). □

Lema 3.13. *Sean ν un punto extremo de $M(\mu)$, $U := \{(x, y) : \nu((x, y)) > 0\}$, $f : A \subset X \to Y$ y $g : B \subset Y \to X$ que satisfagan:*

(i) $G(f) \cup G(g) \subset U$

(ii) $G(f) \cap G(g) = \emptyset$.

Entonces para cada $n \geq 1$, $x \in D_n$ implica que

$$(g \circ f)^m(x) \neq (g \circ f)^p(x),$$

para $0 \leq p < m \leq n$. En particular, $f((g \circ f)^m(x)) \neq f((g \circ f)^p(x))$ siempre que $0 \leq p < m < n - 1$.

Demostración. Si $n = 1$ y suponemos que $(g \circ f)^0(x) = (g \circ f)(x)$ entonces $x = (g \circ f)(x)$ esto es, $g(f(x)) = x$ por lo que el punto $(x, f(x)) = (g(f(x)), f(x)) \in G(f) \cap G(g)$ lo que contradice la primera hipótesis. Se procede por inducción. Supongamos cierto el resultado para n y sea $x \in D_{n+1}$. Es suficiente probar que

$$(g \circ f)^{n+1} \neq (g \circ f)^p(x), \qquad 0 \leq p \leq n.$$

Pues $D_{n+1} \subset D_n$.

Ahora, si $(g \circ f)^{n+1}(x) = (g \circ f)^p(x)$ para $1 \leq p \leq n$ entonces como

$$(g \circ f)^p(x) \in D_{n+1-p}$$

$$(x \in D_{n+1} \Rightarrow (g \circ f)(x) \in D_n ... \Rightarrow (g \circ f)^r(x) \in D_{n+1-r})$$

tenemos que

$$(g \circ f)^{n+1-p}((g \circ f)^p(x)) = (g \circ f)^p(x),$$

lo que contradice la hipótesis de inducción. Falta el caso $p = 0$ para finalizar la prueba de inducción, es decir, mostrar que $(g \circ f)^{n+1}(x) = x$ contradice alguna hipótesis. Veremos que contradice la suposición de que ν es punto extremo de $M(\mu)$ (si ν es punto extremo, entonces $(g \circ f)^{n+1}(x) \neq x)$).

Supongamos que $(g \circ f)^{n+1}(x) = x$. Por hipótesis de inducción tenemos que

$$(g \circ f)^m(x) \neq (g \circ f)^p(x) \qquad 0 \leq p < m \leq n.$$

Definimos $z_k := (x_k, y_k)$ para $k = 0, 1, ..., 2n + 1$ como sigue

$$\begin{aligned}
&z_{2p} = ((g \circ f)^p(x), f((g \circ f)^p(x))), \\
&z_{2p+1} = ((g \circ f)^{p+1}(x), f((g \circ f)^p(x))), \\
&0 \leq p < m \leq n.
\end{aligned}$$

La hipótesis de inducción implica:

(i) Si $1 \leq k \leq 2n - 1$, $k < j \leq 2n$, entonces $x_k = x_j$ si y sólo si k es impar y $j = k + 1$.

(ii) Si $0 \leq k \leq 2n - 2$, $k < j \leq 2n$, entonces $y_k = y_j$ si y sólo si k es par y $j = k + 1$.

(iii) $x_0 \neq x_k$ si $1 \leq k \leq 2n$ y $y_{2n} \neq y_k$ si $0 \leq k \leq 2n - 1$.
Además, si $(g \circ f)^{n+1}(x) = x$ entonces

(iv) $x_{2n+1} = x_0$ y $y_{2n+1} = y_{2n}$.

Entonces $C := \{z_0, z_1, ..., z_{2n+1}\}$ es un ciclo. Sea

$$\alpha = \text{mín}\,\{\nu(z_k) : k = 0, ..., 2n+1\}.$$

Notamos que los z_k tienen la forma $(x, f(x))$ o $(g(y), y)$ por lo que $C \subset U$ y $\alpha > 0$. Definimos ahora una medida con signo m_1 en C:

$$m_1(z_k) = (-1)^k \alpha/2, \qquad k = 0, ..., 2n+1.$$

Sea $m_2 = -m_1$. Las marginales de las m_i son identicamente cero por (i)-(iv). Entonces $\nu_i = \nu + m_i$, $i = 1, 2$, son medidas de probabilidad con las mismas marginales que ν. Como $\nu = \frac{1}{2}(Q_1 + Q_2)$ tenemos la contradicción deseada y la prueba está completa. □

Observación 3.14. Nótese que la prueba del Lema 3.13 tambíen demuestra que el soporte de un punto extremo no contiene ciclos.

Lema 3.15. *Sea $\nu \in M(\mu)$ punto extremo de $M(\mu)$ y $U := \{(x, y) : \nu((x, y)) > 0\}$. Entonces para cada subconjunto finito $V \subset U$ existen $\tilde{f} : \tilde{A} \to Y$, $\tilde{g} : \tilde{B} \to X$ tal que*

$$V = G(\bar{f}) \cup G(\bar{g}).$$

Demostración. Si V contiene un sólo elemento, el resultado es claro. Continuamos la demostración por por inducción. Supongamos que todo subconjunto de U de n elementos se puede escribir como la union de gráficas de dos funciones $f : A \subset X \to Y$ y $g : B \subset Y \to X$. Entonces sea $V = \{(x_k, y_k) : k = 1, ..., n\} = G(f) \cup G(g)$ donde $f : A \to Y$ y $g : B \to X$. Queremos entonces una representación para $V \cup \{x_{n+1}, y_{n+1}\}$.

Si $x_{n+1} \in A^c$, sean $\tilde{A} := A \cup x_{n+1}$, $\tilde{f} = f$ en A y $\tilde{f}(x_{n+1}) := y_{n+1}$, $\tilde{B} := B$ y $\tilde{g} = g$. Del mismo modo, si $y_{n+1} \in B^c$ definimos $\tilde{B} := B \cup y_{n+1}$, $\tilde{g} = g$ en B, $\tilde{g}(y_{n+1}) := x_{n+1}$, $\tilde{A} := A$ y $\tilde{f} = f$.

Si no pasa lo anterior, es decir, si $x_{n+1} \in A$ y $y_{n+1} \in B$, sea $D_0 = A$ y D_m definido como antes (el dominio de la función $(g \circ f)^m$). Si $m \geq 0$ y $x_{n+1} \in D_m - D_{m+1}$ entonces $(g \circ f)^m(x_{n+1}) \in A^c$ ó $f((g \circ f)^m(x_{n+1})) \in B^c$.

Además si $x_{n+1} \in D_m$, $m \geq 1$, el Lema 3.13 implica que los puntos

$$((g \circ f)^p(x_{n+1}), f(g \circ f)^p(x_{n+1})), \qquad 0 \leq p \leq m-1$$

son elementos distintos de $G(f)$. Como $G(f)$ es finita, podemos suponer que para alguna $m \geq 0$, $x_{n+1} \in D_m - D_{m+1}$.

Si $x_{n+1} \in D_0 - D_1$ entonces $f((g \circ f)^0(x_{n+1})) \in B^c$ (Si $f((g \circ f)^0(x_{n+1})) = f(x_{n+1}) \in B$ entonces $g(f(x))$ está definida y $x_{n+1} \in D_1$, una contradicción).

Definimos $\tilde{B} = B \cup \{f(x_{n+1})\}$, $\tilde{g} = g$ en B, $\tilde{g}(f(x_{n+1})) = x_{n+1}$, $\tilde{A} = A$, $\tilde{f} = f$ en $A - \{x_{n+1}\}$ y $\tilde{f}(x_{n+1}) = y_{n+1}$. Entonces $\tilde{f}$ y $\tilde{g}$ cumplen con la propiedad buscada.

Supongamos ahora que $x_{n+1} \in D_m - D_{m+1}$, $m \geq 1$. Consideraremos el caso $(g \circ f)^m(x_{n+1}) \in A^c$. El otro es análogo. Sean

$$A_1 := \{(g \circ f)^p(x_{n+1}) : p = 0, ..., m\} \subset A \cup \{(g \circ f)^m(x_{n+1})\}$$

y

$$B_1 := \{f((g \circ f)^p(x_{n+1})) : p = 0, ..., m - 1\} \subset B.$$

Definimos

$$\tilde{A} := A \cup \{(g \circ f)^m(x_{n+1})\},$$
$$\tilde{B} = B,$$

y $\tilde{f}$ en A_1 como

$$\begin{aligned}
&\tilde{f}(x_{n+1}) = y_{n+1}, \\
&\tilde{f}((g \circ f)^p(x_{n+1})) = f((g \circ f)^{p-1}(x_{n+1})), \quad 1 \leq p \leq m
\end{aligned}$$

y $\tilde{f} = f$ en $A - A_1$.

De manera similar definimos $\tilde{g}$ en B_1 como

$$\tilde{g}(f((g \circ f)^p(x_{n+1}))) = (g \circ f)^p(x_{n+1}), \quad 0 \leq p \leq m - 1$$

y $\tilde{g} = g$ en $B - B_1$. Por el Lema 3.13 $\tilde{f}$ y $\tilde{g}$ están bien definidas.

Veamos ahora que $V \cup \{(x_{n+1}, y_{n+1})\} = G(\tilde{f}) \cup G(\tilde{g})$. Supongamos que $(x, y) \in G(\tilde{f})$. Si $x = x_{n+1}$ o $x \in A - A_1$ es claro que $(x, y) \in V \cup \{(x_{n+1}, y_{n+1})\}$.

Ahora, si $x \in A_1$ y $x \neq x_{n+1}$, entonces para alguna $1 \leq p \leq m$

$$\begin{aligned}
(x, y) =& ((g \circ f)^p(x_{n+1}), \tilde{f}((g \circ f)^p(x_{n+1}))) \\
=& (g(f(g \circ f)^{p-1}(x_{n+1})), f((g \circ f)^{p-1}(x_{n+1}))) \in G(g) \subset V.
\end{aligned}$$

Análogamente, si $(x, y) \in G(\tilde{g})$ y $y \in B_1$ entonces

$$\begin{aligned}
(x, y) =& (\tilde{g}(f((g \circ f)^P(x_{n+1})), f((g \circ f)^p(x_{n+1}))) \\
=& ((g \circ f)^p(x_{n+1}), f((g \circ f)^p(x_n + 1))) \in G(f) \subset V.
\end{aligned}$$

Por lo que $G(\tilde{f}) \cup G(\tilde{g}) \subset V \cup (x_{n+}, y_{n+1})$.

Tomemos ahora $(x, y) \in G(f)$ Si $x \in A - A_1$ claramente $(x, y) \in G(\tilde{f})$, si $x \in A_1 \cap A$ entonces para alguna $1 \leq p \leq m$

$$\begin{aligned}(x, y) =&((g \circ f)^p(x_{n+1}), f((g \circ f)^p(x_{n+1}))) \\ =&(\tilde{g}(f(g \circ f)^p(x_{n+1})), f((g \circ f)^p(x_{n+1}))) \in G(\tilde{g}).\end{aligned}$$

De la misma forma, si $(x, y) \in G(g)$ y $y \in B_1 \cap B$ entonces para alguna $1 \leq p \leq m-1$

$$\begin{aligned}(x, y) =&(g(f((g \circ f)^p(x_{n+1}))), f((g \circ f)^p(x_{n+1}))) \\ =&((g \circ f)^{p+1}(x_{n+1}), \tilde{f}((g \circ f)^{p+1}(x_{n+1}))) \in G(\tilde{f}) \subset V.\end{aligned}$$

Por lo tanto $V \cup \{(x_{n+1}, y_{n+1}) \subset G(\tilde{f}) \cup G(\tilde{g})\}$ y la prueba está completa. □

Lema 3.16. *. Sean $G(f_i) \cup G(g_i) \subset G(f_{i+1}) \cup G(g_{i+1})$, $i = 1, 2, ...$ gráficas de funciones. Entonces existe $f : A \to Y$ y $g : B \to X$ tales que*

$$\bigcup_i \{G(f_i) \cup G(g_i)\} = G(f) \cup G(g).$$

Demostración. Supongamos falso el Lema, es decir, que no existe una pareja de funciones tal que la unión de sus gráficas sea igual a la unión de las gráficas de las f_i y g_i. Si entendemos una función como la correspondencia que define su gráfica, esto querría decir que para cualquier correspondencia de puntos en $\bigcup_i \{G(f_i) \cup G(g_i)\}$ a funciones f y g, existe al menos un punto (x_n, y_n) tal que los conjuntos $G(f) \cup \{(x_n, y_n)\}$ y $G(g) \cup \{(x_n, y_n)\}$ no pueden definir una función, es decir, ya habría un punto de la forma $(x_n, y) \in G(f)$ y un $(x, y_n) \in G(g)$. Pero $(x_n, y_n) \in G(f_i) \cup G(g_i)$ para alguna i, $(x_n, y) \in G(f_j) \cup G(g_j)$ para alguna j y $(x, y_n) \in G(f_k) \cup G(g_k)$ para alguna k. Tomando $l = max\{i, j, k\}$ tenemos por hipótesis que $(x_n, y_n), (x, y_n), (x_n, y) \in G(f_l) \cup G(g_l)$ lo cual es una contradicción pues supusimos que no podían existir tales funciones. □

Teorema 3.17 (Denny)**.** *Sea μ una medida de probabilidad en $X \times Y$ y sea $\nu \in M(P)$. Entonces son equivalentes*

(i) ν es punto extremo de $M(\mu)$.

(ii) Existe una pareja de funciones $f : A \subset X \to Y$, $g : B \subset Y \to X$ aperiódica (f, g), tal que $\nu(G) = 1$, donde $G = \{(x, f(x)) : x \in A)\} \cup \{(g(y), y) : y \in B\}$.

Demostración. Veamos que (i) implica (ii). Por el Lema 3.15 cualquier subconjunto finito del soporte de un punto extremo ν de $M(\mu)$ puede representarse como la unión de gráficas de dos funciones f y g con $f : A \subset X \to Y$ y $g : B \subset Y \to X$. Por el Lema 3.16 podemos garantizar la existencia de una pareja de funciones tal que la unión de sus graficas sea todo el soporte de ν y podemos suponer sin perdida de generalidad, que la intersección de las gráficas de estas funciones es vacía. Entonces por el Lema 3.13, éstas funciones forman una pareja aperiódica.

Para probar que (ii) implica (i) introducimos la siguiente colección de funciones

$$\mathcal{F} := \{f + g : f \in L_1(\mu_X), g \in L_1(\mu_Y)\}.$$

Por el Teorema de Douglas-Lindenstrauss, $\nu \in M(\mu)$ es punto extremo de $M(\mu)$ si y sólo si $\mathcal{F}$ es denso en $L_1(\nu)$.

Probaremos que

$$\int |h - (\phi + \psi)| d\nu$$

se puede hacer arbitrariamente pequeña para $h \in L_1(Q)$, $\phi + \psi \in \mathcal{F}$. Es suficiente probar esto para $h = \chi_V$, con V un subconjunto finito de $X \times Y$ pues podemos aproximar cualquier función en $L_1(\nu)$ por funciones con soporte finito. Además, como $\mathcal{F}$ es un espacio vectorial, es suficiente probarlo para $V = \{z\}$, $z \in X \times Y$. También podemos suponer que $z \in G$ pues G es el soporte de Q. Ahora construimos una sucesión de funciones $\phi_n - \psi_n$, con $\phi_n : X \to \mathbb{R}$, $\psi_n : Y \to \mathbb{R}$ que converge a h en $L_1(\nu)$.

Si A es vacío y $z = (g(y), y)$, definimos $\phi_0 = 0$ y $\psi = \chi_y$. Entonces $\int |h - (\phi + \psi)| dQ = 0$

Ahora supongamos que $z = (x_0, f(x_0))$, definimos $\psi_0 = 0$, $\phi_i = I_{A_i}$, $\psi_i = \chi_{B_i}$ con A_i, B_i definidos como en Lema 3.12. Entonces los conjuntos A_i, B_i y $B'_i := \{(g(y), y) : y \in B_i\}$ son disjuntos dos a dos. Ahora veamos que

$$\int |h - (\sum_{i=o}^{n} (\phi_i - \psi_i))| dQ = Q(B'_{n+1}).$$

Para ver esto analicemos para que subconjuntos de $X \times Y$ la integral arroja valores distintos de cero.

Observación 3.18. Si $x \in A_i$ entonces $\phi_j(x) = 0$ para toda $j \neq i$ y $y \in B_i$ implica $\psi_j(y) = 0$ para toda $j \neq i$.

Observación 3.19. El integrando se anula en los puntos (x, y) tales que $x \notin A_i$ para toda i y $y \notin B_j$ para toda j. Por lo que podemos ignorar los puntos de esta forma.

Observación 3.20. El integrando se anula en $A_i \times B_i$

Por las últimas dos observaciones sólo consideramos los puntos (x, y) tales que $x \in A_i$ para algún $i = 1, ..., n$ pero $y \notin B_i$ y viceversa.

Cuando $x \in A_i$, buscamos las $y \in Y$ tales que las parejas $(x, y) \in G$. Tenemos dos posibilidades:

(i) $y = f(x)$ es decir $y \in f[A_i] \subset B_i$, o

(ii) y es preimagen bajo g de x. Es decir $y \in g^{-1}[A_i] \subset B_{i+1}$.

En el primer caso $(x, y) \in A_i \times B_i$ y la integral se anula (Observación 3.20). En el segundo caso tendríamos que $(x, y) \in A_i \times B_{i+1}$ para alguna $i = 1, ..., n$. En este caso tenemos que la integral se reduce a

$$\int_{A_i \times B_{i+1}} |\phi_i - \psi_{i+1}| d\nu = 0$$

si $i < n$, o

$$\int_{A_i \times B_{i+1}} |\phi_n| d\nu = \nu(B'_{i+1})$$

si $i = n$, pues por lo anterior $x = g(y)$

Análogamente, si $y \in B_i$ y $(x, y) \in G$, entonces $x \in g[B_i] \subset A_{i-1}$ o $x \in A_i$. Aquí en ambos casos la integral se anula. Por lo que podemos concluir que

$$\int |h - (\sum_{i=1}^{n} (\phi_i - \psi_i))| dQ = Q(B'_{n+1}).$$

Ahora, como los B'_i son disjuntos dos a dos y ademas $B'_i \subset G$ para toda i, entonces

$$\sum_n Q(B'_n) \leq Q(G) = 1,$$

de donde $Q(B'_n)$ tiende a 0 cuando n tiende a infinito. Entonces $\mathcal{F}$ es denso en $L_1(Q)$ y por lo tanto (ii) implica (i). $\square$

Capítulo 4

Medidas extremas con momentos fijos

Sea $I \subset \mathbb{R}$ un intervalo. Dada una medida positiva μ en I se define su n-ésimo momento como $\int_I x^n d\mu(x)$ si la integral existe. Si $(c_n)_{n\geq 0}$ es una sucesion de numeros real el problema clásico de los momentos consiste en encontrar medidas en I con momentos (c_n), determinar cuando una medida esta únicamente determinada por sus momentos o describir la clase de soluciones.

Se puede generalizar esta situación reemplazando las funciones $\{x^n\}$ por una sucesión $\{\phi_n\}_{n\geq 0}$ de funciones linealmente independientes con valores reales o complejos, usualmente se toma $\phi_0 \equiv 1$.

En ambos casos, el conjunto de soluciones es convexo. En el cápitulo describe el problema generalizado de momentos truncos y se presentan algunos resultados sobre el soporte de los puntos extremos.

4.1. Problema de momentos y densidad de subespacios

Sean $\{\phi_n\}$ una sucesión de funciones reales linealmente independientes definidas en un intervalo, $\{c_n\}$ una sucesion de numeros reales y μ una medida tal que $\int \phi_n d\mu = c_n$. Si hacemos

$$\mathcal{F} = \langle\{\phi_n\}\rangle,$$

es decir, $\mathcal{F}$ es espacio lineal de funciones generado por $\{\phi_n\}$, entonces podemos ver que $E_\mu(\mathcal{F})$ del teorema de Douglas definido en (2.1) es exactamente el conjunto de soluciones del problema de momentos. En efecto, si $\nu \in E_\mu$ tenemos que para toda

$f \in \mathcal{F}$
$$\int f d\nu = \int f d\mu.$$

Tomando $f = \phi_n$, tenemos que $\int \phi_n d\nu = \int \phi_n d\mu = c_n$ para toda $n \geq 0$, entonces ν es solución. Por otro lado, sea ν una medida tal que para toda $n \geq 0$, $\int \phi_n d\nu = c_n$. Si $f \in \mathcal{F}$ entonces $f = \sum_{i=1}^{k} \alpha_i \phi_{n_i}$ con $\alpha_i \in \mathbb{R}$ y $\{\phi_{n_i}\} \subset \{\phi_j\}_{j\in\mathbb{N}}$. Se sigue que para toda $f \in \mathcal{F}$

$$\int f d\nu = \int \sum_{i=1}^{k} \alpha_i \phi_{n_i} d\nu = \sum_{i=1}^{k} \alpha_i c_{n_i} = \int \sum_{i=1}^{k} \alpha_i \phi_{n_i} d\mu = \int f d\mu$$

y $\nu \in E_\mu$.

Así que podemos aplicar el Teorema de Douglas (dado que se cumplan las hipótesis sobre el espacio) y obtenemos que una medida ν es punto extremo del conjunto de soluciones si y sólo si $\mathcal{F}$ es denso en la norma de $L_1(\nu)$.

4.2. Problema generalizado de momentos truncos

El problema de momentos truncos, como el nombre lo indica, es el problema de momentos para un número finito de momentos. A continuación describimos el problema generalizado de momentos truncos y establecemos las hipótesis y notación que se usaran para el resto del capítulo.

Sea I un subconjunto compacto de un espacio métrico completo y separable. Sea $\phi_K = (\phi_k)_{k\in K}$, $K = \{1, ..., n\}$, una familia linealmente independiente de funciones medibles definidas en I. Sea $\mathcal{M}^\phi$ el espacio de medidas en I tales que

$$\int_I \phi_k d\mu < +\infty \qquad k \in K$$

Sea Φ_K el funcional lineal definido en $\mathcal{M}^\phi$ por:

$$\Phi_K \mu = \left(\int_I \phi_k d\mu \right)_{k\in K}$$

Denotamos:
$\mathcal{M}_+^\phi$ al subconjunto de $\mathcal{M}^\phi$ de medidas no negativas en I,
$\mathcal{M}_+^d$ al subconjunto de $\mathcal{M}_+^\phi$ de medidas con soporte discreto,
$\mathcal{M}_+^f$ al subconjunto de $\mathcal{M}_+^d$ de medidas con soporte finito,
S_μ al soporte de una medida μ.

Dada una sucesión $c = (c_k)_{k\in K} \in \mathbb{R}^{|K|}$ el problema de los momentos consiste en determinar medidas en $\mathcal{M}_+^{\phi}$ tales que satisfagan la ecuacion

$$\Phi_K \mu = c$$

4.3. Soporte de medidas extremas discretas con momentos fijos

Estableceremos algunos resultados sobre el soporte de puntos extremos del conjuntos de soluciones a un problema de momentos truncos. Veremos que una si una medida discreta es punto extremo entonces su soporte debe ser finito, mas aún, se establece la cardinalidad del soporte. También se da el recíproco, es decir, si una medida discreta tiene soporte finito de cierta cardinalidad entonces esta medida debe ser punto extremo del conjunto de soluciones al problema de momentos. Antes algunas observaciones y definiciones preliminares.

Una medida μ con soporte discreto puede escribirse como sigue

$$\mu(x) = \sum_{\lambda_i \in S_\mu} \mu_i \chi_{\lambda_i}(x),$$

donde S_μ es el soporte de μ y $\mu_i = \mu(\lambda_i)$. De modo que el problema de los momentos para una medida discreta se reduce a encontrar sucesiones (μ_i) de números reales positivos que satisfagan la ecuacion de momentos

$$\sum_{i=1}^{\infty} \phi_k(\lambda_i)\mu_i = c_k.$$

Definamos la siguiente matriz para cualquier subconjunto finito S de I,

$$M_S^K = (\phi_k(\lambda))_{\lambda\in S}^{k\in K}$$

Teorema 4.1. *Sea $\mu \in \Phi_K^{-1}(c) \cap \mathcal{M}_+^d$, entonces μ es un punto extremo de $\Phi_K^{-1}(c) \cap \mathcal{M}_+^{\phi}$ si y sólo si μ tiene soporte finito S_μ, con $|S_\mu| = Rango[M_{S_\mu}^K]$*

Demostración. Supongamos que μ tiene soporte finito: si $\mu \in \Phi_K^{-1}(c) \cap \mathcal{M}_+^f$, entonces (μ_i), para $i = 1, ..., |S_\mu|$ es una solución del siguiente sistema con $|S_\mu|$ incógnitas y $|K|$ ecuaciones:

$$\sum_{\lambda_i \in S_\mu} \phi_k(\lambda_i)\mu_i = c_k, \qquad k \in K$$

El rango de este sistema es Rango$[M_{S_\mu}^K]$ y la dimensión del espacio afín de soluciones A es $|S_\mu|-$ Rango$[M_{S_\mu}^K]$. Entonces μ representa, una solución positiva $(\mu_i)_{i=1}^{|S_\mu|}$ de éste sistema de ecuaciones. Si $\mu = \alpha\nu_1 + (1-\alpha)\nu_2$ con $\nu_1, \nu_2 \in \Phi_K^{-1}(c) \cap \mathcal{M}_+^f$ y $\alpha \in (0,1)$, entonces los soportes de ν_1 y ν_2 están contenidos en S_μ por lo que μ_1 y μ_2 representan soluciones no negativas del sistema de ecuaciones. Si Rango$[M_{S_\mu}^K] = |S_\mu|$ entonces el sistema tiene solución única y por lo tanto $\mu_1 = \mu_2 = \mu$ y μ es punto extremo de $\Phi_K^{-1}(c) \cap \mathcal{M}_+^\phi$.

Por otro lado, si Rango$[M_{S_\mu}^K] < |S_\mu| < \infty$, entonces el sistema de ecuaciones tiene soluciones distintas a (μ_i). Entonces el vector $(\mu_i)_{i=1}^{|S_\mu|}$, pertence a una recta $\mathcal{L}$ contenida en A. Ya que $(\mu_i) \in \mathbb{R}_+^{|S_\mu|}$, el conjunto $\mathcal{L} \cap \mathbb{R}_+^{|S_\mu|}$ no es un solo punto. Así que (μ_h) puede ser escrito como una combinación convexa de vectores distintos (η_h) y (ν_h) que pertencecen a $\mathcal{L} \cap \mathbb{R}_+^{|S_\mu|}$, y por lo tanto μ se puede escribir como una combinación convexa de medidas η y ν que pertenecen a $\Phi_K^{-1}(c) \cap \mathcal{M}_+^f$. Entonces μ no es punto extremo de $\Phi_K^{-1}(c) \cap \mathcal{M}_+^\phi$.

Ahora supongamos que $\mu \in \Phi_K^{-1}(c) \cap (\mathcal{M}_+^d - \mathcal{M}_+^f)$. Para $S \subset S_\mu$, tal que $|K| < |S| < \infty$. Sea

$$\sum_{\lambda_h \in S} \phi_k(\lambda_h)\alpha_h = 0, k \in K.$$

Este sistema homogéneo tiene $|K|$ ecuaciones y $|S|$ incógnitas. Dado que

$$\text{Rango}[M_K^S] < |S|,$$

el sistema tiene soluciones distintas de cero. Sea $(\alpha_i)_{i=1,\ldots,|S|}$ una de ellas. Definimos una medida en I como sigue:

$$\tilde{\alpha}_{\lambda_i} = \begin{cases} M\alpha_i, & \text{si } \lambda_i \in S, \\ 0 & \text{en caso contrario,} \end{cases}$$

donde $M = \frac{\min(\mu_i | i=1,\ldots,|S|)}{\max(|\alpha_i| | i=1,\ldots,|S|)}$.

Nótese que $\tilde{\alpha}_{\lambda_i}$ sigue siendo solución al sistema homogéneo. Además, por la elección de M, $(\mu - \tilde{\alpha}) \geq 0$ por lo que $(\mu - \tilde{\alpha})$ y $(\mu + \tilde{\alpha})$ son medidas en I.

Luego

$$\sum_{\lambda_h \in S_\mu} \phi_k(\lambda_h)(\mu - \tilde{\alpha})_h = \sum_{\lambda_h \in S_\mu - S} \phi_k(\lambda_h)(\mu_h) + \sum_{\lambda_h \in S} \phi_k(\lambda_h)\left(\mu - \tilde{\alpha}\right)_h = I_1 + I_2$$

Dado que $\mu \in \Phi_K^{-1}(c)$ tenemos que

$$I_1 = \sum_{\lambda_h \in S_\mu - S} \phi_k(\lambda_h)(\mu_h) = c_k - \sum_{\lambda_h \in S} \phi_k(\lambda_h)\mu_h, \quad k \in K.$$

Por otro lado,

$$I_2 = \sum_{\lambda_h \in S} \phi_k(\lambda_h)\,(\mu - \tilde{\alpha})_h = \sum_{\lambda_h \in S} \phi_k(\lambda_h)\mu_h - \sum_{\lambda_h \in S} \phi_k(\lambda_h)\tilde{\alpha}_h$$

Pero $\sum_{\lambda_h \in S} \phi_k(\lambda_h)\tilde{\alpha}_h = 0$ por lo que

$$\sum_{\lambda_h \in S_\mu} \phi_k(\lambda_h)(\mu - \tilde{\alpha})_h = c_k$$

Lo mismo se da para $(\mu + \tilde{\alpha})$, por lo que $\mu - \tilde{\alpha}$ y $\mu + \tilde{\alpha}$ pertenecen a $\Phi_K^{-1}(c) \cap \mathcal{M}_+^\phi$, y entonces μ no es punto extremo de $\Phi_K^{-1}(c) \cap \mathcal{M}_+^\phi$. □

Teorema 4.2. *Sea $\mu \in \Phi_K^{-1} \cap \mathcal{M}_+^\phi$. Si $\phi_k \in L^2(\mu)$ para toda $k \in K$, entonces μ es punto extremo de $\Phi_K^{-1} \cap \mathcal{M}_+^\phi$ si y sólo si μ tiene soporte finito S_μ con $|S_\mu|$=Rango$[M_{S_\mu}^K]$.*

Demostración. Sea F_ϕ el subespacio de $L^2(\mu)$ generado por $(\phi_k)_{k\in K}$. Denotemos por S al espacio de funciones medibles y simples en I. Si $F_\phi \neq L^2(\mu)$, como F_ϕ tiene dimensión finita existe un subespacio E de S tal que

$$dim F_\phi < dim E < +\infty.$$

Sea $f \in E, 0 < f \leq 1$, ortogonal a F_ϕ y definamos $d\nu = (1-f)d\mu$ y $d\eta = (1+f)d\mu$. Luego

$$\int \phi_k d\nu = \int \phi_k (1+f) d\mu = \int \phi_k d\mu + \int \phi_k f d\mu = c_k$$

y

$$\int \phi_k d\eta = \int \phi_k (1-f) d\mu = \int \phi_k d\mu - \int \phi_k f d\mu = c_k$$

pues $\int \phi_k f d\mu = 0$, para toda $k \in K$. Entonces ν, η pertencen a $\Phi_K^{-1}(c) \cap \mathcal{M}_+^\phi$ y $\mu = \frac{1}{2}(\nu + \eta)$ por lo que μ no es punto extremo de $\Phi_K^{-1} \cap \mathcal{M}_+^\phi$.

Se sigue que si μ es punto extremo de $\Phi_K^{-1} \cap \mathcal{M}_+^\phi$ entonces $F_\phi = L^2(\mu)$, entonces existen a lo mas K conjuntos disjuntos con medida postiva, pues no puede haber mas de K funciones linealmente independientes en $L_2(\mu)$. Mas aún, dado un conjunto maximal de conjuntos disjuntos de medida positiva, estos deben ser átomos de μ por lo que μ es una medida discreta y $|S_\mu| \leq K < \infty$ y del Teorema 4.1 se sigue el resultado. □

4.4. Generación por medidas extremas

Hemos mostrado algunas caracterizaciones de los puntos extremos del conjunto de soluciones al problema de momentos. Ahora nos quisieramos generar el conjunto solución a partir de éstos. Restringiendonos a medidas con soporte finito veremos que el conjunto solución $\Phi_K^{-1}(c) \cap \mathcal{M}_+^f$ está generado por sus puntos extremos.

En el caso general, agregando mas hipótesis sobre la familia de funciones $(\phi_k)_{k\in K}$, veremos que las medidas extremas con soporte finito generan todo el conjunto de soluciones respecto a la topología débil.

En lo siguiente supondremos que $K = \{0, ..., n\}$, la hipótesis $\phi_0 = 1$ que se presenta en los teoremas sirve para garantizar que el conjunto solución de medidas con soporte finito no sea trivial.

Teorema 4.3. *Sea $\mu \in \Phi_k^{-1}(c) \cap \mathcal{M}_+^f$ Si $\phi_0 \equiv 1$ sobre I, entonces μ es combinacion convexa de $|S_\mu| + 2$ medidas extremas de $\mu \in \Phi_k^{-1}(c) \cap \mathcal{M}_+^f$, con soporte contenido en S_μ.*

Demostración. Sea $\mathcal{M}_+^{S_\mu}$ el subespacio de $\mathcal{M}_+^\phi$ de medidas cuyo soporte esta contenido en S_μ. Entonces $\Phi_K^{-1}(c) \cap S_\mu$ es un conjunto convexo contenido en $R^{|S_\mu|}$, aplicando el theorema de Carathéodory podemos concluir que μ es una combinación convexa de $|S_\mu| + 2$ medidas extremas de $\Phi_K^{-1}(c) \cap \mathcal{M}_+^{S_\mu}$ □

Lema 4.4. *Si $\phi_0 \equiv 1$ sobre I, entonces $\Phi_k(\mathcal{M}_+^\phi) = \Phi_k(\mathcal{M}_+^f)$.*

Demostración. Es claro que $\Phi_K(\mathcal{M}_+^f) \subset \Phi_K(\mathcal{M}_+^\phi)$. Probaremos que si $c \in \Phi_K(\mathcal{M}_+^\phi)$ entonces $c \in \Phi_K(\mathcal{M}_+^f)$ separadamente para puntos interiores y frontera, esto es, que $\Phi_K(\mathcal{M}_+^\phi) \subset \overline{\Phi_K(\mathcal{M}_+^f)}$ y $\Phi_K(\mathcal{M}_+^\phi) \cap \partial\Phi_K(\mathcal{M}_+^f) \subset \Phi_K(\mathcal{M}_+^f)$.

Tenemos que

$$\Phi_K(\mathcal{M}_+^f) = \left\{ \left(\sum_{\lambda\in S} \rho_\lambda \phi_k(\lambda) \right)_{k\in K} : \rho_\lambda \geq 0, S \subset I, |S| < \infty \right\} \tag{4.1}$$

es decir, $\Phi_K(\mathcal{M}_+^f)$ es el casco cónico del conjunto

$$\mathcal{U}_K^\phi = \left\{ (\phi_k(\lambda))_{k\in K} : \lambda \in I \right\}$$

Veamos que $\Phi_K(\mathcal{M}_+^\phi) \subset \overline{\Phi_K(\mathcal{M}_+^f)}$. Como podemos expresar al casco cónico de un conjunto como la interesección de todos los semiespacios soporte que pasan por el origen y que contienen al conjunto, tenemos que

$$\overline{\Phi_K(\mathcal{M}_+^f)} = \left\{ c \in \mathbb{R}^{|K|} : \forall (a_k)_{k\in K}, \sum_K a_k\phi_k \geq 0 \Rightarrow \sum_K a_k c_k \geq 0 \right\},$$

Luego si $c \in \Phi_K(\mathcal{M}_+^\phi)$ y $\mu \in \Phi_K^{-1}(c) \cap \mathcal{M}_+^\phi$, entonces

$$\int_I \sum_K a_k\phi_k(\lambda)d\mu(\lambda) = \sum_K a_k \int_I \phi_k(\lambda)d\mu(\lambda) = \sum_K a_k c_k.$$

De modo que si (a_k) es tal que $\sum_K a_k\phi_k(\lambda) \geq 0$ para toda $\lambda \in I$, entonces $\int_I \sum_K a_k\phi_k(\lambda)d\mu(\lambda) \geq 0$ de donde $\sum_K a_k c_k \geq 0$, por lo tanto $c \in \overline{\Phi_K(\mathcal{M}_+^f)}$ y $\Phi_K(\mathcal{M}_+^\phi) \subset \overline{\Phi_K(\mathcal{M}_+^f)}$.

Ahora veamos que $\Phi_K(\mathcal{M}_+^\phi) \cap \partial\Phi_K(\mathcal{M}_+^f) \subset \Phi_K(\mathcal{M}_+^f)$. Usamos inducción. Si $|K| = 1$ entonces $\partial\Phi_K(\mathcal{M}_+^f) = \Phi_K(\mathcal{M}_+^f)$ de donde se sigue el resultado. Además, si $\Phi_K(\mathcal{M}_+^f) = \mathbb{R}^{|K|}$ entonces $\partial\Phi_K(\mathcal{M}_+^f) = \emptyset$ y también se da la contención. Sea $c \in \Phi_K(\mathcal{M}_+^\phi) \cap \partial\Phi_K(\mathcal{M}_+^f)$ y supongamos que $\Phi_K(\mathcal{M}_+^f) \neq \mathbb{R}^{|K|}$. Entonces c pertenence a un hiperplano soporte $\mathcal{H}$ de $\Phi_K(\mathcal{M}_+^f)$.

Probaremos que

$$\Phi_K(\mathcal{M}_+^\phi) \cap \mathcal{H} \subset \Phi_K(\mathcal{M}_+^f) \cap \mathcal{H} \tag{4.2}$$

para cualquier hiperplano soporte de $\mathcal{M}_+^f$ $\mathcal{H}$.

Primero veamos que $\Phi_K(\mathcal{M}_+^\phi) \cap \mathcal{H} \subset \mathcal{C}_\mathcal{H}^\phi$, donde $\mathcal{C}_\mathcal{H}^\phi$ es el casco cónico de $\mathcal{U}_K^\phi \cap \mathcal{H}$.

La ecuación de $\mathcal{H}$ es

$$\sum_K h_k b_k = 0, \tag{4.3}$$

con $\sum_K h_k^2 = 1$ y $\sum_K h_k\phi_k(\lambda) \geq 0$ para todo λ en I.

Si $c \in \Phi_K(\mathcal{M}_+^\phi) \cap \mathcal{H}$ y $\mu \in \Phi_K^{-1}(c) \cap \mathcal{M}_+^\phi$, entonces

$$\int_I \sum_K h_k\phi_k(\lambda)d\mu(\lambda) = \sum_K h_k \int_I \phi_k(\lambda)d\mu(\lambda) = \sum_K h_k c_k = 0,$$

entonces $\sum_K h_k\phi_k(\lambda) = 0$, para toda $\lambda \in I - E$ para algun cojunto E con $\mu(E) = 0$. Definimos

$$\phi_k^* = \begin{cases} \phi_k & \text{en } I - E \\ 0 & \text{en } E \end{cases}$$

Entonces $c_k = \int_I \phi_k^*(\lambda)d\mu(\lambda)$ y $\sum_K h_k\phi_k^* \equiv 0$ en I. Sea $j \in K$ tal que $h_j \neq 0$, entonces

$$\phi_j^* = \sum_{k\neq j} \frac{h_k}{h_j}\phi_k^* \text{ y } c_j = \sum_{k\neq j} \frac{h_k}{h_j} c_k. \tag{4.4}$$

Entonces $(c_k)_{k\in K-\{j\}} \in \Phi^*_{K-\{j\}}(\mathcal{M}^\phi_+)$, donde $\Phi^*_{K-\{j\}}$ es el funcional definido como $\left(\int_I \phi^*_k d\mu\right)_{k\in K-\{j\}}$. Entonces por hipótesis de inducción,

$$(c_k)_{k\in K-\{j\}} \in \Phi^*_{K-\{j\}}(\mathcal{M}^f_+),$$

o de otra forma, como en la igualdad (4.1):

$$c_k = \sum_{\lambda\in S} \rho_\lambda \phi^*_k(\lambda), \text{ para } k \in K-\{j\}.$$

Ahora, por (4.4) tenemos que

$$c_j = \sum_{k\neq j} \frac{h_k}{h_j} \sum_{\lambda\in S} \rho_\lambda \phi^*_k(\lambda) = \sum_{\lambda\in S} \rho_\lambda \sum_{k\neq j} \frac{h_k}{h_j} \phi^*_k(\lambda) = \sum_{\lambda\in S} \rho_\lambda \phi^*_j(\lambda).$$

Es decir, $(c_k)_{k\in K} \in \Phi^*_K(\mathcal{M}^f_+) \cap \mathcal{H}$.

Ahora veamos que

$$\Phi_K(\mathcal{M}^f_+) \cap \mathcal{H} = \mathcal{C}^\phi_\mathcal{H}.$$

Por un lado es claro que $\mathcal{C}^\phi_\mathcal{H} \subset \Phi_K(\mathcal{M}^f_+) \cap \mathcal{H}$ pues $\mathcal{C}^\phi_\mathcal{H} \subset \Phi_K(\mathcal{M}^f_+)$ y $\mathcal{C}^\phi_\mathcal{H} \subset \mathcal{H}$.

Sea entonces $c \in \Phi_K(\mathcal{M}^f_+) \cap \mathcal{H}$ entonces por (4.3).

$$c_k = \sum_{\lambda\in S} \rho_\lambda \phi_k(\lambda) \text{ con } \rho_\lambda \geq 0 \text{ y } \sum_K h_k c_k = 0$$

sustituyendo c_k en la segunda ecuación tenemos que

$$\sum_{\lambda\in S} \rho_\lambda \sum_K h_k \phi_k(\lambda) = 0, \text{ para toda } \lambda \in S,$$

pero como $\sum_k h_k\phi_k \geq 0$ en I y $\rho_\lambda \geq 0$ para toda $\lambda \in S$, entonces

$$\sum_K h_k \phi_k(\lambda) = 0, \text{ para toda } \lambda \in S.$$

De forma que $(\phi_k(\lambda)_{k\in K} \in \mathcal{H}$ para toda $\lambda \in S$ y $c \in \mathcal{C}^\phi_\mathcal{H}$.

Ahora como $\mathcal{C}^{\phi^*}_\mathcal{H} \subset \mathcal{C}^\phi_\mathcal{H}$ tenemos que:

$$\Phi_K(\mathcal{M}^\phi_+) \cap \mathcal{H} \subset \Phi^*_K(\mathcal{M}^f_+) \cap \mathcal{H} = \mathcal{C}^{\phi^*}_\mathcal{H} \subset \mathcal{C}^\phi_\mathcal{H} = \Phi_K(\mathcal{M}^f_+) \cap \mathcal{H}$$

De donde $\Phi_K(\mathcal{M}^\phi_+ \cap \partial\Phi_K(\mathcal{M}^f_+) \subset \Phi_K(\mathcal{M}^f_+)$ y por lo tanto $\Phi_K(\mathcal{M}^\phi_+) = \Phi_K(\mathcal{M}^f_+)$.

□

Definición 4.5 (Topología debil de espacios de medida)**.** Sea (X, τ) un espacio topológico y $\{\mu_n\}_{n\in\mathbb{N}}$ una sucesión de medidas definidas en la σ-álgebra de Borel de τ. Entonces decimos que μ_n converge débilmente a μ si para toda función continua y acotada $f : X \to \mathbb{R}$ se tiene que $\lim_{n\to\infty} \int_X f d\mu_n = \int_X f d\mu$.

Teorema 4.6. *Sea $C(I)$ el espacio de funciones reales continuas definidas sobre I. Si el subespacio generado por $(\phi_k)_{k\in\mathbb{N}}$ es denso en $C(I)$, y $\phi_0 \equiv 1$ sobre I, entonces los puntos extremos de $\Phi_K^{-1}(c) \cap \mathcal{M}_+^f$, generan $\Phi_K^{-1}(c) \cap \mathcal{M}_+^\phi$ en la topología débil.*

Demostración. Sea $\mu \in \Phi_K^{-1}(c) \cap \mathcal{M}_+^\phi$. Si $f \in C(I)$, entonces f es el límite de una sucesión de combinaciones lineales finitas de elementos de $(\phi_k)_{k\in\mathbb{N}}$. Es decir es límite de una sucesión de funciones de la forma

$$F_H = \sum_{k\in H} a_k^H \phi_k$$

donde H es un subconjunto finito de $\mathbb{N}$.

Ahora, para cada $H \subset \mathbb{N}$ definimos

$$\overline{\mu} := (\overline{\mu_k})_{k\in K\cup H},$$

donde

$$\overline{\mu_k} = \int_I \phi_k(\lambda) d\mu(\lambda).$$

Es decir $\overline{\mu_k}$ es el k-ésimo momento de μ, entonces por definicón

$$\mu \in \Phi_{K\cup H}^{-1}(\overline{\mu}) \cap \mathcal{M}_+^\phi.$$

Luego por el Lema anterior $\Phi_k(\mathcal{M}_+^\phi) = \Phi_k(\mathcal{M}_+^f)$, entonces existe una medida $\mu^H \in \Phi_{K\cup H}^{-1}(\overline{\mu}) \cap \mathcal{M}_+^f$ tal que para $k \in K \cup H$

$$\sum_{\lambda\in S_{\mu^H}} \mu_\lambda^H \phi_k(\lambda) = \overline{\mu_k}$$

.

Luego

$$\begin{aligned}\int_I F_H(\lambda) d\mu(\lambda) &= \sum_{k\in K\cup H} a_k^H \int_I \phi_k(\lambda) d\mu \\ &= \sum_{k\in K\cup H} a_k^H \overline{\mu_k}\end{aligned}$$

y por otro lado:

$$\begin{aligned}\int_I F_H(\lambda)d\mu^H(\lambda) &= \sum_{k\in K\cup H} a_k^H \int_I \phi_k(\lambda)d\mu^H \\ &= \sum_{k\in K\cup H} a_k^H(\sum_{\lambda\in S_{\mu^H}} \mu_\lambda^H\phi_k(\lambda)) \\ &= \sum_{k\in K\cup H} a_k^H\overline{\mu_k}\end{aligned}$$

por lo que $\int_I F_H(\lambda)d\mu(\lambda) = \int_I F_H(\lambda)d\mu^H(\lambda)$ y

$$\int_I (f-F_H)(\lambda)d\mu(\lambda) - \int_I (f-F_H)(\lambda)d\mu(\lambda) = \int_I f(\lambda)d\mu(\lambda) - \int_I f(\lambda)d\mu^H(\lambda)$$

Entonces tenemos que:

$$\left|\int_I f(\lambda)d\mu(\lambda) - \int_I f(\lambda)d\mu^H(\lambda)\right| \leq \int_I |f-F_H|(\lambda)d\mu(\lambda) + \int_I |f-F_H|(\lambda)d\mu(\lambda).$$

Y como $F_H \to f$ entonces

$$\left|\int_I f(\lambda)d\mu(\lambda) - \int_I f(\lambda)d\mu^H(\lambda)\right| \to 0$$

para toda $f \in C(I)$, es decir, μ^H converge débilmente a μ. Tenemos entonces que $\Phi_K^{-1}(c)\cap\mathcal{M}_+^\phi$ esta contenido en la cerradura débil de $\Phi_K^{-1}(c)\cap\mathcal{M}_+^f$ entonces del teorema anterior se sigue el resultado. □

4.5. Conexión entre los problemas de momentos y marginales fijas

Consideremos $X = (x_i)_{i\in I}$, $Y = (y_j)_{j\in J}$ dos conjuntos finitos, $I = 0,...,n$, $J = 0,...,m$ y ρ, τ dos medidas en X y Y respectivamente. Denotamos como $\mathcal{M}_{\rho\tau}(X\times Y)$ al conjunto de medidas en $X\times Y$ con marginales ρ y τ en X y Y respectivamente. Para cada $k\in I$, $l\in J$ definimos las siguientes funciones en X y Y:

$$\phi_k := \chi_{x_k} \text{ y } \psi_l := \chi_{y_l}$$

Si $\mu \in \mathcal{M}_{\rho\tau}(X \times Y)$ entonces

$$\sum_{j \in J} \mu_{ij} = \rho_i = \int \phi_i d\rho$$

y de la misma forma

$$\sum_{i \in I} \mu_{ij} = \tau_j = \int \psi_j d\tau$$

Si $(c_{kl})_{(k,l) \in I \times J}$ es tal que $c_{k0} = \rho_k$ y $c_{0l} = \tau_l$ para toda $k \in I$, $l \in J$. Entonces las funciones ϕ_i, ψ_j transforman el problema de marginales fijas en un problema de momentos y viceversa. Y de hecho tienen las mismas soluciones.

El siguiente Teorema nos da condiciones para dar equivalencias entre problemas de momentos y de marginales fijas para funciones mas generales.

Teorema 4.7. *Sean* $X = (x_i)_{i \in I}$ *y* $Y = (y_j)_{j \in J}$*, donde* $I = \{0, ..., n\}$*,* $J = \{0, ..., m\}$*. Sean* $(\phi_k)_{k \in I}$ *y* $(\psi_l)_{l \in J}$ *familias de funciones definidas en* X *y* Y *respectivamente, tales que:*

- $\phi_0 \equiv \psi_0 \equiv 1$
- *Los rangos de las matrices* M_E^I *y* M_F^J *sean maximales para todo* $E \subset X$*,* $F \subset Y$*.*

Sea $(\Phi, \Psi)_{I,J}$ *el funcional lineal asociado a la familia de funciones*

$$((\phi_k)_{k \in I}, (\psi_l)_{l \in J}).$$

Sean ρ *y* τ *dos medidas en* X *y* Y *respectivamente.*
Sean $r_k := \int \phi_k d\rho$ *y* $t_l := \int \psi_l d\tau$*.*
Sea $c = (c_{kl})_{(k,l) \in I \times J}$ *tal que* $c_{k0} = r_k$ *para* $k \in I$ *y* $c_{0l} = t_l$ *para* $l \in J$*.*

Entonces $(\Phi, \Psi)_{I,J}^{-1}(c) = \mathcal{M}_{\rho\tau}(X \times Y)$*.*

Demostración. Para $k \in I$ y $l \in J$ tenemos que

$$r_k = \sum_{i \in I} \rho_i \phi_k(x_i) \text{ y } t_l = \sum_{j \in J} \tau_j \psi_l(y_l)$$

Entonces si $\mu \in \mathcal{M}_{\rho\tau}(X \times Y)$ es decir, $\sum_{j \in J} \mu_{ij} = \rho_i$ y $\sum_{i \in I} \mu_{ij} = \tau_j$, de modo que:

$$\int \phi_k d\mu = \sum_{i,j} \mu_{ij} \phi_k = \sum_{i \in I} \rho_i \phi_k = r_k$$

y

$$\int \psi_l d\mu = \sum_{i,j} \mu_{ij}\psi_l = \sum_{j\in J} \tau_j \psi_l = t_l$$

Por lo tanto $\mu \in (\Phi, \Psi)^{-1}_{I,J}(c) \cap \mathcal{M}^f_+$

Ahora, si $\mu \in (\Phi, \Psi)^{-1}_{I,J}(c)$ entonces para cada $k \in I$, $l \in J$ μ cumple

$$\sum_{i,j} \mu_{ij}\phi_k(x_i) = r_k \text{ y } \sum_{i,j} \mu_{ij}\psi_l(y_j) = t_l.$$

Sean

$$b_i = \sum_{j\in J} \mu_{ij} \text{ y } d_j = \sum_{i\in I} \mu_{ij}$$

los valores marginales de μ en X y Y respectivamente. Entonces $\{(b_i)_{i\in I}, (d_j)_{j\in J}\}$ es una solución al siguiente sistema con $n + m + 2$ incógnitas.

$$\begin{cases} \sum_i \phi_k(x_i) b_i & = r_k \\ \sum_j \psi_l(y_l) d_j & = t_l \end{cases}$$

Este sistema tiene una solución única, por lo tanto para toda $i \in I$, $b_i = \rho_i$ y para toda $j \in J$, $d_j = \tau_j$, y $\mu \in \mathcal{M}_{\rho\tau}(X \times Y)$.

$\square$

Capítulo 5

Teorema de Birkhoff-von Neumann

El Teorema de Birkhoff-von Neumann dice lo siguiente

Teorema (Birkhoff)**.** *Sea (t_{ij}) cualquier matriz de términos no negativos de un campo ordenado tal que las sumas de los términos en cada fila y columna es uno. Entonces (t_{ij}) es un promedio pesado de matrices de permutación.*

El término *promedio pesado* se refiere a una combinación lineal de la forma

$$\sum_{i=1}^{n} \alpha_i P_i,$$

con $\alpha_i \geq 0$, $i = 1, ..., n$ y $\sum_{i=1}^{n} \alpha_i = 1$. También se conoce como combinación convexa de los puntos P_i, $i = 1, ..., n$. En la terminología de conjuntos convexos, el Teorema de Birkhoff nos dice que toda matriz cuyas entradas son no negativas y la suma de éstas en filas o columnas es igual a uno, es una combinación convexa de matrices de permutación. Esto es, la clase de matrices de $n \times n$ con entradas no negativas y con suma de filas y columnas igual a uno, es el casco convexo del conjunto de matrices de permutación de $n \times n$.

El llamado problema 111 de Birkhoff, es generalizar este resultado, bajo hipótesis apropiadas, a espacios de dimensión infinita.

En este capítulo se plantea el Teorema de Birkhoff en el contexto de conjuntos convexos y se demuestra mediante las herramientas desarrolladas principalmente en el capítulo 2. También se discute el problema 111 de Birkhoff, por lo que el capítulo está dividido en tres partes: los casos finito, infinito numerable e infinito no numerable.

5.1. Caso finito

Definición 5.1. Una matriz de $n \times n$, (p_{ij}) se llama matriz estocástica si y sólo si cada fila es un vector de probabilidad n-dimensional. Esto es, si y sólo si, $p_{ij} \geq 0$ y $\sum_{j=1}^{n} p_{ij} = 1$ para toda $i = 1, ..., n$.

Denotaremos por P^t a la matriz transpuesta de P. Esto es, si $P = (p_{ij})$ entonces $P^t = (p_{ji})$. Notemos que $(P^t)^t = P$.

Definición 5.2. Una matriz de $n \times n$, $P = (p_{ij})$, se llama doblemente estocástica si y sólo si P y P^t son matrices estocásticas. Esto es, P es doblemente estocástica si y sólo si para todo $i, j = 1, ..., n$, se cumple que

$$\sum_{j=1}^{n} p_{ij} = \sum_{i=1}^{n} p_{ij} = 1.$$

Definición 5.3. Una matriz P de $n \times n$ es una matriz de permutación si y sólo si P es doblemente estocástica y sus entradas sólo constan de ceros y unos. Entonces P es de permutación si cada fila y columna tienen exactamente un uno y las demas entradas son cero.

Denotaremos por $\mathcal{D}_n$ al conjunto de matrices doblemente estocásticas de $n \times n$ entradas y por $\mathcal{P}_n$ al conjunto de matrices de permutación de $n \times n$.

Observación 5.4. $\mathcal{D}_n$ es convexo para toda $n \geq 1$.

Veremos ahora que las matrices de permutación son los puntos extremos del conjunto de matrices doblemente estocásticas. Para esto identificaremos el conjunto de matrices doblemente estocásticas con un conjunto de medidas con marginales fijas; el soporte de éstas medidas se identificará con las entradas positivas de la matriz. Aplicando el Teorema 3.17 se verá que los puntos extremos del conjunto de matrices doblemente estocásticas deben ser matrices de permutación.

Teorema 5.5. *$\mathcal{P}_n$ es el conjunto de puntos extremos de $\mathcal{D}_n$.*

Demostración. Sea $P \in \mathcal{P}_n$ una matriz de permutación. Si $P = \alpha A + (1-\alpha)B$ con $0 < \alpha < 1$, $A = (a_{ij}) \in \mathcal{D}_n$ y $B = (b_{ij}) \in \mathcal{D}_n$. Entonces para todo (i, j) se tiene que $1 = \alpha a_{ij} + (1-\alpha)b_{ij} \leq 1$ o $0 = \alpha a_{ij} + (1-\alpha)b_{ij} \geq 0$. En el primer caso $a_{ij} = b_{ij} = 1$, y en el segundo $a_{ij} = b_{ij} = 0$. Entonces $A = B = P$, y por lo tanto P es punto extremo.

Ahora sean $X = Y = \{1, 2, ..., n\}$. Identificamos a cada matriz doblemente estocástica $D = (p_{ij}) \in \mathcal{D}_n$ con una medida μ en $X \times Y$ definida como $\mu((i,j)) = p_{ij}$. Para cada medida de esta forma sus marginales son

$$\mu_X(i) = \sum_{j=1}^{n} p_{ij} = 1$$

$$\mu_Y(j) = \sum_{i=1}^{n} p_{ij} = 1$$

para todo $(i,j) \in X \times Y$.

De la misma forma toda medida definida sobre $X \times Y$ con estas marginales se puede identificar con una matriz doblemente estocástica. Entonces para μ definida en $X \times Y$ que cumpla que $\mu_X(i) = 1$, $\mu(j)_Y = 1$ para toda $(i,j) \in X \times Y$, $M(\mu)$ del Teorema 3.17 se identifica con $\mathcal{D}_n$.

Entonces si $P \in \mathcal{D}_n$ es punto extremo y U es el soporte de la medida inducida por P, es decir, U son las parejas de índices de las entradas distintas de cero. Por el Teorema de Denny (Teorema 3.17) sobre soporte de medidas discretas con marginales fijas, $U = G(f) \cup G(g)$ donde (f,g) es una pareja aperiódica de funciones, $f : A \subset X \to Y$ y $g : B \subset Y \to X$ con A, $B \subset \{1, ..., n\}$

Sea entonces P un punto extremo de $\mathcal{D}_n$ y $U = G(f) \cup G(g)$ las parejas de índices de entradas positivas de P. Si $G(f) = \emptyset$, entonces $U = G(g) = \{(g(y), y) : y = 1, ..n\}$. El dominio de g debe ser $Y = \{1, .., n\}$ pues cada columna debe tener al menos una entrada distinta de cero, y como g es función tampoco puede haber dos entradas positivas en la misma columna, entonces P debe de ser de permutación pues $P \in \mathcal{D}_n$. Análogamente, si $G(g) = \emptyset$ entonces P es de permutación.

Afirmamos ahora que no hay pérdida de generalidad en suponer que el dominio de f sea igual a X.

Si $x \in X - A$ entonces $g^{-1}[\{x\}] \neq \emptyset$ ya que no puede haber filas totalmente cero pues $P \in \mathcal{D}_n$. Sea $y \in g^{-1}[\{x\}]$, definimos $f' = f$ en A y $f'(x) = y$ y g' igual a g restringida a $B - y$.

La pareja (f', g') es aperiódica, pues para $x' \in A$ entonces $(g' \circ f')^n(x') = (g \circ f)^n(x')$ para los valores de n que la expresion esté definida y $f'(x)$ no está en el dominio de g'. Procediendo de esta manera es posible construir a partir de (f, g), una nueva pareja aperiódica (f', g') tal que el dominio de f' sea igual a X y $U = G(f') \cup G(g')$.

Ahora mostramos que el rango de f' debe ser igual a Y. Si $y \in Y - f[X]$, entonces $Q(g'(y), y) = 1$, por ser doblemente estocástica y por definición de función. A su vez, $Q(g'(y), y) = 1$ implica que $Q(g'(y), f'(g'(y))) = 0$ pues el dominio de f' es X.

Esta igualdad contradice la definición de U, por lo tanto el rango de f es igual a Y. Notemos ahora que el dominio B' de g' no puede ser igual a Y pues (f', g') es aperiódica y X es finito.

Para concluir el resultado mostraremos ahora que B es vacío. Si $y \in B'$ entonces $y = f'(x)$, la aperiodicidad implica que existe un mínimo entero positivo k tal que $f'(g \circ f)^k(x) \in Y - B$. Como f' es sobre, $(x, y) \in G(f')$ $(y = f'(x))$ y por lo tanto $Q(g'(y), y) = Q(g'(f'(x)), f'(x)) < 1$; aplicando f' a $g'(y)$ tenemos que que $Q((g' \circ f')(x), f'(g' \circ f'(x))) < 1$. Siguiendo este procedimiento, aplicando f' y g' a las coordenadas izquierda y derecha respectivamente, obtenemos que $Q((g' \circ f')^k(x), f'(g' \circ f')^k(x)) < 1$. Esto último contradice la doble estocasticidad, pues $f'(g \circ f)^k(x) \in Y - B$ implica que la columna de la matriz correspondiente a $f'(g \circ f)^k(x)$ sólo tiene un elemento distinto de cero $((g' \circ f')^k(x), f'(g' \circ f')^k(x))$. □

Teorema 5.6 (Birkhoff-von Neuman)**.** *Toda matriz doblemente estocástica se puede expresar como combinación convexa de matrices de permutación.*

Demostración. Veamos que el conjunto de matrices doblemente estocásticas es cerrado en $\mathbb{R}^{n^2}$. Sea $D_n = (p_{ij}^{(n)})$ una sucesión de matrices doblemente estocásticas que convergen a una matriz $D = (p_{ij})$. Entonces para toda i,j se tiene que $\lim_{n\to\infty} p_{ij}^{(n)} = p_{ij}$. Entonces para toda j fija, se tiene que

$$\sum_{i=1}^{n} p_{ij} = \sum_{i=1}^{n} \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \sum_{i=1}^{n} p_{ij}^{(n)} = 1$$

y para toda i

$$\sum_{j=1}^{n} p_{ij} = \sum_{j=1}^{n} \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \sum_{j=1}^{n} p_{ij}^{(n)} = 1$$

Entonces D es doblemente estocástica. Dado que toda matriz doblemente estocástica (p_{ij}) cumple que $0 \leq p_{ij} \leq 1$, también tenemos que el conjunto es acotado. Se sigue entonces del teorema de Krein-Milman (1.29) que

$$\overline{\mathrm{cc}}(E(D_n)) = \overline{\mathrm{cc}}(P_n) = D_n$$

y en este caso, como P_n es un conjunto finito, $\mathrm{cc}(P_n) = \overline{\mathrm{cc}}(P_n)$. Se sigue del teorema de Caratheódory (1.12) que cada matriz doblemente estocástica se puede escribir como combinación convexa de $(n-1)^2 + 1$ matrices de permutación. □

5.2. Caso infinito numerable

Podemos extender naturalmente los conceptos de la sección anterior al caso de matrices infinitas.

Definición 5.7. Una sucesión $(p_n)_{n\geq 1}$ con $p_n \geq 0$, es un vector de probabilidad si y solo si $\sum_{n=1}^{\infty} p_n = 1$.

Definición 5.8. $P = (p_{ij})$ es una matriz infinita si y sólo si cada fila y cada columna es una sucesión infinita de números reales.

Definición 5.9. Una matriz infinita $P = (p_{ij})$, es estocástica si y sólo si $0 \leq p \leq 1$ y $\sum_j p_{ij} = 1$, para toda $i \in \mathbb{N}$.

Definición 5.10. Una matriz estocástica infinita P, es doblemente estocástica si y sólo si P y P^t son ambas matrices estocásticas.

Definición 5.11. Una matriz de permutación infinita es una matriz doblemente estocástica infinita cuyas entradas son sólo ceros y unos.

Como en la sección anterior, identificaremos los puntos extremos del conjunto de matrices doblemente estocásticas infinitas. Al igual que en el caso finito, éstas resultan ser las matrices de permutación.

Lema 5.12. *Sea (p_n) un vector de probabilidad tal que $0 < p_k < \frac{1}{2}$ para alguna $k \in \mathbb{N}$. Sean*

$$\Pi_m = \frac{p_m}{1-p_m}\frac{p_{m-1}}{1-p_{m-1}}\cdots\frac{p_1}{1-p_1}$$

y

$$\Pi_{m\sim k_0} = \Pi_m\left(\frac{1-p_{k_0}}{p_{k_0}}\right) \quad p_{k_0} > 0$$

Caso 1: Si $0 \leq p_n \leq \frac{1}{2}$ para toda n. Sea $m > 1$ fija y definamos para toda $n \neq m$:

$$\begin{aligned} a_m &= p_m\left(1-\Pi_{m-1}\right), \\ b_m &= p_m\left(1+\Pi_{m-1}\right), \\ a_n &= p_n(1+\Pi_m) \quad y \\ b_n &= p_n(1-\Pi_m). \end{aligned}$$

Entonces (a_n) y (b_n) son vectores de probabilidad y para toda n y $p_n = \frac{1}{2}(a_n + b_n)$.

Caso 2: Si $0 \leq p_n \leq \frac{1}{2}$ *para toda* $n \neq n_0$*,* $\frac{1}{2} < p_{n_0} < 1$*, sea* $m > 1$*,* $m \neq n_0$ *fija y sean para toda* $n \neq n_0$*:*

$$\begin{aligned} a_{n_0} &= p_{n_0} - (1 - p_{n_0})\Pi_{m\sim n_0}, \\ b_{m_0} &= p_{n_0} + (1 - p_{n_0})\Pi_{m\sim n_0}, \\ a_n &= p_n(1 + \Pi_{m\sim n_0}) \ \ y \\ b_n &= p_n(1 - \Pi_{m\sim n_0}) \end{aligned}$$

Entonces (a_n) *y* (b_n) *son vectores de probabilidad y para toda* n*,* $p_n = \frac{1}{2}(a_n + b_n)$*.*

Demostración. Caso 1. Si $0 \leq p_n \leq \frac{1}{2}$ para toda n, entonces $0 \leq \frac{p_n}{1-p_n} \leq 1$. De modo que $0 \leq \Pi_m \leq 1$ para toda m, y $0 \leq a_n \leq 1$, $0 \leq b_n \leq 1$ para toda n. Luego

$$\begin{aligned} \sum_{n=1}^{\infty} a_n &= p_m(1 - \Pi_{m-1}) + \sum_{n\neq m} a_n \\ &= p_m(1 - \Pi_{m-1}) + \sum_{n\neq m} p_n(1 + \Pi_m) \\ &= p_m(1 - \Pi_{m-1}) + \sum_{n\neq m} p_n + \sum_{n\neq m} p_n\Pi_m \\ &= p_m(1 - \Pi_{m-1}) + (1 - p_m) + \Pi_m \sum_{n\neq m} p_n \\ &= p_m - p_m\Pi_{m-1} + 1 - p_m + (1 - p_m)\Pi_m \\ &= p_m - p_m\Pi_{m-1} + 1 - p_m + p_m\Pi_{m-1} \\ &= 1 \end{aligned}$$

y para (b_n)

$$\begin{aligned} \sum_{n=1}^{\infty} b_n &= p_m(1 + \Pi_{m-1}) + \sum_{n\neq m} b_n \\ &= p_m(1 + \Pi_{m-1}) + \sum_{n\neq m} p_n(1 - \Pi_m) \\ &= p_m(1 + \Pi_{m-1}) + \sum_{n\neq m} p_n - \sum_{n\neq m} p_n\Pi_m \\ &= p_m(1 + \Pi_{m-1}) + (1 - p_m) - \Pi_m \sum_{n\neq m} p_n \\ &= p_m(1 + \Pi_{m-1}) + (1 - p_m) - \Pi_{m-1}p_m \\ &= 1 \end{aligned}$$

Entonces (a_n) y (b_n) son vectores de probabilidad, y por construcción es claro que $p_n = \frac{1}{2}(a_n + b_n)$.

Caso 2: $0 \leq p_n \leq \frac{1}{2}$ para toda $n \neq n_0$ y $\frac{1}{2} < p_{n_0} < 1$. Como $\Pi_{m\sim n_0} \leq 1$, tenemos que

$$a_{n_0} = p_{n_0} - (1 - p_{n_0})\Pi_{m\sim n_0} \leq p_{n_0} < 1$$

y

$$b_{n_0} = p_{n_0} + (1 - p_{n_0})\Pi_{m\sim n_0} \leq p_{n_0} + (1 - p_{n_0}) = 1.$$

Por lo que, $0 \leq a_n \leq 1$ y $0 \leq b_n \leq 1$ para toda n. Entonces

$$\begin{aligned}\sum_n a_n &= p_{n_0} - (1 - p_{n_0})\Pi_{m\sim n_0} + \sum_{n\neq n_0} p_n(1 + \Pi_{m\sim n_0}) \\ &= p_{n_0} - (1 - p_{n_0})\Pi_{m\sim n_0} + \sum_{n\neq n_0} p_n + \sum_{n\neq n_0} p_n\Pi_{m\sim n_0} \\ &= p_{n_0} - (1 - p_{n_0})\Pi_{m\sim n_0} + (1 - p_{n_0}) + \Pi_{m\sim n_0}(1 - p_{n_0}) \\ &= 1\end{aligned}$$

y

$$\begin{aligned}\sum_n b_n &= p_{n_0} + (1 - p_{n_0})\Pi_{m\sim n_0} + \sum_{n\neq n_0} p_n(1 - \Pi_{m\sim n_0}) \\ &= p_{n_0} + (1 - p_{n_0})\Pi_{m\sim n_0} + \sum_{n\neq n_0} p_n - \sum_{n\neq n_0} p_n\Pi_{m\sim n_0} \\ &= p_{n_0} + (1 - p_{n_0})\Pi_{m\sim n_0} + (1 - p_{n_0}) - \Pi_{m\sim n_0}(1 - p_{n_0}) \\ &= 1\end{aligned}$$

□

Definición 5.13. Una *camino* en una matriz con entradas no negativas $M = (m_{ij})$ es un conjunto finito $\{m_{r_1c_1}, m_{r_2,c_2}, ..., m_{r_uc_u}\}$ de entradas positivas de M, tal que $r_1 = r_2$ y $c_2 \neq 1$ o $c_1 = c_2$ y $r_2 \neq c_1$ y tal que si $r_i = r_{i-1}$ entonces $c_i = c_{i+1} \neq c_{i-1}$ y si $c_i = c_{i-1}$ entonces $r_{i+1} = r_i \neq r_{i-1}$. Esto es, una secuencia de entradas positvas de la matriz, en la cual los sucesores estan en la misma columna o fila, alternadamente, del antecesor. Un camino termina si regresa a una fila o columna ya visitada, en este caso el camino contiene un ciclo.

Denotaremos por $\mathcal{D}$ al conjunto de matrices infinitas doblemente estocásticas y por $\mathcal{P}$ al conjunto de matrices infinitas de permutación.

Teorema 5.14. *$P \in \mathcal{D}$ es punto extremo de $\mathcal{D}$ si y sólo si P es de permutación ($P \in \mathcal{P}$).*

Demostración. Usaremos la siguiente notación. Para una matriz doblemente estocástica P con una entrada $p_{ij} \in (0,1)$, sea P_{ij} el conjunto de parejas ordenadas de enteros positivos, (u,v) tal que existe un camino en P de p_{ij} a p_{uv}. Para $(u,v) \in P_{ij}$ si p_{rs} está en el camino que va desde p_{ij} a p_{uv} denotamos por $p_{\overrightarrow{rs}}$ al sucesor inmediato de p_{rs} en el camino.

Notemos que si $p_{ij} \in (0,1)$ y $(m,n) \notin P_{ij}$, entonces $(m,c) \notin P_{ij}$ para toda c y $(r,n) \notin P_{ij}$ para toda r. Esto es, si p_{ij} es una entrada positiva de la matriz, y no hay caminos de p_{ij} a p_{nm}, entonces no hay caminos de p_{ij} a ninguna otra entrada en la misma fila o columna que p_{nm}.

Sea $P \in \mathcal{P}$. Si $P = \alpha A + (1-\alpha)B$ con $0 < \alpha < 1$ y A,$B \in \mathcal{D}$, entonces para todo i,j pasa que $1 = \alpha a_{ij} + (1-\alpha)b_{ij}$ y $a_{ij} = b_{ij} = 0$, o que $0 = \alpha a_{ij} + (1-\alpha)b_{ij}$ y $a_{ij} = b_{ij} = 1$. Por lo tanto $A = B = P$ y P es punto extremo.

Para demostrar el recíproco indentificamos cada matriz doblemente estocástica infinita $P = (p_{ij})$ con una medida μ en $\mathbb{N} \times \mathbb{N}$, definida como $\mu((i,j)) = p_{ij}$. Entonces cada medida de este tipo tiene las mismas medidas marginales. Usando la caracterización del Capítulo 3 del soporte de medidas discretas con marginales fijas y por la Observación 3.14, el soporte de un punto extremo del conjunto de medidas definidas de esta manera no contiene ciclos, esto es, el conjunto de entradas positivas de la matriz no contiene ciclos. Construiremos a partir de una matriz sin ciclos que no es de permutación dos matrices doblemente estocásticas cuya combinación convexa es P. Esto se hará modificando las filas y columnas de las entradas de un camino de forma que sigan siendo vectores de probabilidad según la construcción del Lema 5.12. El hecho de que la matriz no contenga ciclos garantiza que no se modifique alguna entrada modificada en algun paso anterior.

Sea P punto extremo de $\mathcal{D}$ y supongamos que no es de permutación. Entonces P no contiene ciclos. Elijamos una entrada positiva de P que sea menor o igual a $\frac{1}{2}$. Por conveniencia llamemos a esta entrada p_{11}. Sea P_{11} definido y $(u,v) \in P_{11}$. Consideremos la fila o columna que contiene a p_{11} y $p_{\overrightarrow{11}}$. Sin perdida de generalidad supongamos que es una fila. Es decir $p_{\overrightarrow{11}} = p_{12}$.

Si $0 \leq p_{1c} \leq \frac{1}{2}$ para toda c, sean

$$\begin{aligned}
a_{12} &= \left(1 - \frac{p_{11}}{1-p_{11}}\right)\\
b_{12} &= p_{12}\left(1 + \frac{p_{11}}{1-p_{11}}\right)\\
a_{1c} &= p_{1c}\left((1+\Pi_2)\right) \text{ y}\\
b_{1c} &= p_{1c}\left(1-\Pi_2\right)
\end{aligned}$$

Por 5.12, (a_{1m}) y (b_{1m}) son vectores de probabilidad con $(p_{1m}) = \frac{1}{2}((a_{1m})+(b_{1m}))$.

Si hay exactamente un $p_{1c_0} > \frac{1}{2}$, definimos

$$\begin{aligned}
a_{1c_0} &= p_{1c_0} + (1-p_{1c_0})\frac{p_{11}}{1-p_{11}},\\
b_{1c_0} &= p_{1c_0} - (1-p_{1c_0})\frac{p_{11}}{1-p_{11}}\frac{p_{11}}{1-p_{11}},\\
a_{1c} &= p_{1c}\left(1 - \frac{p_{11}}{1-p_{11}}\right) y\\
b_{1c} &= p_{1c}\left(1 + \frac{p_{11}}{1-p_{11}}\right) \text{ para todo } c \neq c_0.
\end{aligned}$$

Nuevamente por 5.12 (a_{1m}) y (b_{1m}) son vectores de probabilidad con $(p_{1m}) = \frac{1}{2}((a_{1m})+(b_{1m}))$.

Ahora sea

$$p_{22} = p_{\overrightarrow{12}}.$$

Si $p_{12} \leq \frac{1}{2}$, por el paso anterior $a_{12} = p_{12}\left(1 - \frac{p_{11}}{1-p_{11}}\right) \leq \frac{1}{2}$. La construcción va como sigue:

Para el caso 1 (todas las entras de la fila menores a $\frac{1}{2}$) hacemos

$$a_{r2} = p_{r2}\left(1-\Pi_{12}\right) = p_{r2}\left(1 - \frac{p_{12}}{1-p_{12}}\frac{p_{11}}{1-p11}\right) \quad r \neq 1$$

$$b_{r2} = p_{r2}\left(1+\Pi_{12}\right) = p_{r2}\left(1 + \frac{p_{12}}{1-p_{12}}\frac{p_{11}}{1-p11}\right) \quad r \neq 1$$

Caso 2. Si $p_{r_02} > \frac{1}{2}$, entonces $a_{r_02} = p_{r_02}+(1-p_{r_02})\frac{p_{11}}{1-p_{11}}$, $b_{r_02} = p_{r_02}-(1-p_{r_02})\frac{p_{11}}{1-p_{11}}$

y

$$a_{r2} = p_{r2}\left(1 - \frac{p_{11}}{1-p_{11}}\right)$$
$$b_{r2} = p_{r2}\left(1 + \frac{p_{11}}{1-p_{11}}\right)$$

para $r \neq r_0$.

Si $p_{12} > \frac{1}{2}$ entonces por la construcción del paso anterior, $a_{12} = p_{12} + (1-p_{12})\frac{p_{11}}{1-p_{11}}$ entonces hacemos

$$a_{r2} = p_{r2}\left(1 - \frac{p_{11}}{1-p_{11}}\right)$$
$$b_{r2} = p_{r2}\left(1 + \frac{p_{11}}{1-p_{11}}\right)$$

para toda $r \neq 1$.

La construcción se sigue inductivamente para de modo que a_{ij}, b_{ij} están definidas para toda $i, j \in P_{11}$. Definimos $a_{ij} = b_{ij} = p_{ij}$ para toda $i,j \notin P_{11}$, y entonces $A = (a_{ij})$ y $B = (b_{ij})$ son matrices doblemente estocásticas con $A \neq B$ y $\frac{1}{2}(A+B) = P$. $\square$

Ya identificados los puntos extremos del conjunto de matrices doblemente estocásticas infinitas, es de interés obtener un resultado análogo al teorema de Birkhoff-von Neumann para el caso de matrices infinitas. Por un lado es claro que no se puede obtener un análogo exacto, pues una matriz con un numero infinto de entradas positivas no puede ser una combinación lineal finita de matrices de permutación. Recurriendo al Teorema de Krein-Milman, una generalización adecuada estaría en terminas de cascos convexos, es decir, que

$$\overline{\text{cc}}(P) = D.$$

Bastaría entonces probar que el conjunto de matrices doblemente estocásticas es compacto en alguna topología. Sin embargo en las topologías inducidas por las normas $\|\cdot\|_p$ el conjunto de matrices doblemente estocásticas no es cerrado. Para ver esto usaremos los siguientes resultados.

Teorema 5.15. *Una matriz infinita P, de terminos no negativos es estocástica, si y sólo si $\sum_{i=1}^{\infty}(fP)_i = \sum_{i=1}^{\infty} f_i$ para toda $f = (f_1, f_2, ...)$ con $\sum i = 1^{\infty}|f_i| < \infty$.*

Demostración. Sea P una matriz estocástica infinita y sea $f \in \ell_1$ entonces por convergencia absoluta

$$\sum_{i=1}^{\infty} (fP)_i = \sum_{j=1}^{\infty} \left(\sum_{i=1}^{\infty} f_i p_{ij} \right) = \sum_{i=1}^{\infty} f_i \left(\sum_{j=1}^{\infty} p_{ij} \right) = \sum_{i=1}^{\infty} f_i.$$

Ahora supongamos que P es una matriz infinita positiva que cumple

$$\sum_{i=1}^{\infty} (fP)_i = \sum_{i=1}^{\infty} f_i$$

para toda f con $\sum i = 1^{\infty} |f_i| < \infty$. Tenemos que $e_i P = (p_{ij})_{ij}$ y $P = (p_{ij})$. Como P es no negativa entonces $p_{ij} \geq 0$ y ya que $\sum_{j=1}^{\infty} e_i P = \sum_{j=1}^{\infty} p_{ij}$ y $\sum_{j=1}^{\infty} (e_i P)_j = \sum_{j=1}^{\infty} (e_i)_j = 1$ tenemos que P es estocástica. $\square$

Lema 5.16. *Sean P_1 ,P_2 matrices doblemente estocásticos. Entonces $P_1 P_2$ y P_1^n son doblemente estocásticas para toda $n \in \{1, 2, 3, ...\}$.*

Demostración. Sean $P_1 = (a_{ij})$ y $P_2 = (b_{ij})$ matrices infinitas doblemente estocásticas. Por el teorema 5.15 $P_1 P_2 = P$ es una matriz infinita de términos no negativos. Además, $P = (c_{ij})$ donde $c_{ij} = \sum_{k=1}^{\infty} a_{ik} b_{kj}$. Por el Teorema 5.15 tenemos que

$$\sum_{i=1}^{\infty} c_{ij} = \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} a_{ik} b_{kj} = \sum_{k=1}^{\infty} b_{kj} = 1$$

y

$$\sum_{j=1}^{\infty} c_{ij} = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} a_{ik} b_{kj} = \sum_{j=1}^{\infty} b_{kj} = 1$$

Por lo que P es doblemente estocástica. Se sigue por inducción que P_1^n es doblemente estocástica para toda $n \in \mathbb{N}$. $\square$

Teorema 5.17. *$\mathcal{D}$ no es cerrado para la convergencia puntual.*

Demostración. Sea P la siguiente matriz infinita doblemente estocástica.

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & \dots \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & \dots \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \dots \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & \dots \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

P es doblemente estocástica, $P = P^t$, y por el Lemma 5.16, (P^{2n}) es una sucesión de matrices doblemente estocásticas con $P^{2n} = (P^{2n})^t$ para toda n. Notemos que cualquier entrada distinta de cero en P^{2n} es de la forma $\frac{1}{2^{2n}}\binom{2n}{k}$ para toda $0 \leq k \leq 2n$. Cada entrada $p_{ij}^{(2n)}$ de P^{2n} es el producto punto de la i-ésima fila con la j-ésima columna de la matrix $P^{2(n-1)}$. Por lo que, dada P^{2n} cada fila y columna tienen exactamente los mismos elementos, tal vez en distinto orden. Para toda P^{2n}, la entrada más grande de cada fila o columna siempre estará en la diagonal, pues para toda fila (r_n) y toda columna (c_m) se tiene que $\|(r_n)\|^2 = \|c_m\|^2$ ya que el número de entradas que no son cero es finito. Entonces, cada entrada $p_{ij}^{(2n)}$ es el producto $(x_i) \cdot (x_j)$, y un elemento en la diagonal $p_{ii}^{(2n)}$ está dado por $(x_i) \cdot (x_i) = \|(x_i)\|_2^2$. Entonces para cualquier entrada $p_{ij}^{(2n)}$ en P^{2n} tenemos por la desigualdad de Hölder, que

$$\begin{aligned} p_{ij}^{(2n)} &= (x_i) \cdot (x_j) = \sum x_i x_j \leq \left(\sum x_i^2\right)^{\frac{1}{2}} \left(\sum x_i^2\right)^{\frac{1}{2}} = \|(x_i)\|_2 \|(x_j)\|_2 \\ &= \|(x_i)\|_2^2 = p_{ii}^{(2n)} = \frac{1}{2^{2n}}\binom{2n}{n}. \end{aligned}$$

Estas entradas convergen a cero, por lo que la sucesion (P^{2n}) converge a una matriz cuyas entradas son todas cero. Esto también muestra que P^{2n} converge en norma a la matriz cero, por lo que $\mathcal{D}$ no es cerrado en norma. $\square$

Sin embargo, si se impone una topología correcta se puede ver que las matrices infinitas doblemente estocásticas son el casco cerrado convexo del conjunto de matrices de permutación.

Sea V el espacio de matrices con suma de filas y columnas finitas. Es decir, $M \in V$ si y sólo si

$$\sup_i \sum_j |m_{ij}| < \infty \quad \text{y} \quad \sup_i \sum |m_{ij}| < \infty$$

Definimos una topología τ en V como la topología más débil que hace a las funciones

$$r_k(M) = \sum_{j=1}^{\infty} m_{kj}, \quad c_k(M) = \sum_{i=1}^{\infty} m_{ik}, \quad s_{ij}(M) = m_i j$$

continuas.

(V, τ) es un espacio vectorial topológico, localmente convexo y Hausdorff, ya que si $M_1 \neq M_2$, existe algun $i_0 j_0$ tal que $s_{i_0 j_0}(M_1) \neq s_{i_0 j_0}(M_2)$.

Veamos que en (V,τ) el conjunto de matrices doblemente estocásticas es el casco cerrado convexo de las matrices de permutación.

Teorema 5.18. *En* (V,τ), $\mathcal{D} = \overline{cc}(\mathcal{P})$.

Demostración. Supongamos que $M \in D$ pero $M \notin \overline{\mathrm{cc}}(P)$. Entonces existe un funcional lineal continuo f (ver Teorema A.1 en el Apéndice) para el cual se cumple que

$$f(M) < \inf_{E\in\overline{\mathrm{cc}}(P)} f(E).$$

Pero f debe ser combinación lineal de los funcionales r_k, c_k y s_k (ver apédince, Teorema A.8), entonces podemos escribir a f como

$$f(M) = \sum_{m=1}^{n_1} a_m r_m(M) + \sum_{m=1}^{n_2} b_m c_m(M) + \sum_{m=1}^{n_3} d_m s_{i_m j_m}(M)$$

para toda $M \in V$. Además notemos que para toda matriz doblemente estocástica M

$$r_m(M) = c_m(M) = 1, \text{ y } 0 \leq s_{i_m j_m}(M) \leq 1.$$

Sea $n = \max\{n_1, n_2, i_{n_3}, j_{n_3}\}$. Notemos que ningún m_{ij} con $i,j \geq n$ contribuye al valor de f.

Sea S_n la matriz de $n\times n$, $S_n = (m_{ij})$ para $1 \leq i,j \leq n$. Sea U_n la matriz diagonal donde la entrada (i,i) sea $1 - \sum_{j=1}^{n} m_{ij}$ y las demás entradas son ceros. Sea V_n una matriz similar donde las entradas (j,j) sean igual a $1 - \sum_{i=1}^{n} m_{ij}$. Si S_n^t es la matriz transpuesta de S_n, entonces la matriz de $2n \times 2n$

$$W_{2n} = \begin{pmatrix} S_n & U_n \\ V_n & S_n^t \end{pmatrix}$$

es doblemente estocástica. Si expandimos W_{2n} a una matriz infinita, con ceros en las colas de las filas y columnas y con un matriz identidad infinita en la esquina inferior derecha, obtenemos la siguiente matriz infinita doblemente estocástica

$$W = \begin{pmatrix} S_n & U_n & 0 \\ V_n & S_n^t & 0 \\ 0 & 0 & I \end{pmatrix},$$

con $f(W) = f(M)$. Además, usando el teorema de Birkhoff para el caso finito, tenemos que $W_{2n} = \sum_{k=1}^{2n} \lambda_k P_k$, donde P_k son matrices de permutación de $2n \times 2n$ y $0 \leq \lambda_k \leq 1$ con $\sum_k \lambda_k = 1$. Entonces $f(W) = f(\sum_{k=1}^{2n} \lambda_k P_k)$ done P_k es expandida de la misma forma que W_{2n}. Tenemos entonces que $f(M) \geq \inf_{E\in\overline{\mathrm{cc}}(P)} f(E)$ lo cual es una contradicción. Lo cual prueba que $\mathcal{D} = \overline{\mathrm{cc}}(\mathcal{P})$. $\square$

5.3. Caso no numerable

Siguiendo con la idea de generalizar el teorema de Birkhoff, dirigimos nuestra atención al conjunto de medidas doblemente estocásticas. En el capítulo 3 se mostró una caracterización del soporte de los puntos extremos de este conjunto, ahora identificaremos algunos de estos puntos extremos y veremos que generan al resto del conjunto imponiendo una topología apropiada. Para esto identificaremos al conjunto de medidas doblemente estocásticas con una clase de operadores lineales en L_p. Esta clase de operadores es la generalización de los operadores lineales inducidos por matrices doblemente estocásticas para los casos finito e infinito numerable, esto es, si P es una matriz estocástica finito o infinita el operador inducido por P es el operador lineal definido por $Tf := fP$, donde f pertenece a $\mathbb{R}^n$ o a ℓ_1 según sea el caso. Definiremos los operadores estocásticos y doblemente estocásticos en L_p y examinaremos algunas propiedades de éstos. Estas propiedades también las cumplen los operadores inducidos por matrices doblemente estocásticas finitas o infinitas. Veremos además que todo operador con estas propiedades se puede identificar de manera única con una medida doblemente estocástica. Dotamos entonces al conjunto de operadores doblemente estocásticos con las topologías fuerte, y débil de operadores para demostrar la versión infinita no numerable del teoremoa de Birkhoff-von Nuemann.

Trabajaremos en los espacios $L_p[0,1]$ con la medida de Lebesgue, y con operadores $T : L_p \to L_p$.

Definición 5.19. Un operador lineal T es *positivo* si y sólo si $f \geq 0$ implica que $Tf \geq 0$.

Definición 5.20. Un operador lineal acotado $T : L_1 \to L_1$ es *estocástico* si T es positvo y $\int Tf = \int f$ para toda $f \in L_1$

Esta definición es consistente con la de matrices estocásticas. Reemplazando la integral por sumas de entradas, una matriz es estocástica si y sólo si cumple esta propiedad (ver Teorema 5.15, el caso finito es igual).

Si $T : L_p \to L_p$, $1 \leq p \leq \infty$ es un operador lineal acotado podemos definir $T^t : L_q \to L_q$ con $\frac{1}{p} + \frac{1}{q} = 1$, por la relación $\langle g, Tf\rangle_{L_q - L_p} = \langle T^t g, f\rangle_{L_q - L_p}$, donde $\langle f, g\rangle_{L_p - L_q} = \int fg$ con $f \in L_p$, $g \in L_q$ y lo llamamos operador transpuesto de T. El operador T^t es análogo al operador inducido por la matriz transpuesta de los casos finito y numerable. En lo que sigue denotaremos por 1 a la función constante 1.

Teorema 5.21. *Sea $T : L_1 \to L_1$ un operador lineal acotado y positivo. T es estocástico si y solo si $T^t 1 = 1$.*

Demostración. Supongamos que $T \geq 0$ y que $\int Tf = \int f$ para toda $f \in L_1$. Entonces $\langle T^t 1, f\rangle = \langle 1, Tf\rangle = \int Tf = \int f = \langle 1, f\rangle$ para toda $f \in L_1$ por lo tanto $T^t 1 = 1$. Ahora, si $T \geq 0$ y $T^t 1 = 1$, entonces $\int Tf = \langle 1, Tf\rangle = \langle T^t 1, f\rangle = \langle 1, f\rangle = \int f$ para toda $f \in L_1$ y por lo tanto T es estocástico. $\square$

Teorema 5.22. *Sea T_p un operador lineal acotado en L_p, $1 \leq p < \infty$. Si $T_p \geq 0$, entonces $T_p^t \geq 0$ y $|T_p^t g| \leq \|g\|_\infty$ para toda $g \in L_\infty$ y por ende $T^t : L_\infty \to L_\infty$.*

Demostración. Si $T_p \geq 0$ entonces $T_p \chi_A \geq 0$ para todo A medible. Sea $g \in L_\infty \subset L_q$, $1 = \frac{1}{p} + \frac{1}{q}$ y $g \geq 0$. Entonces $\langle g, T_p \chi_A\rangle = \int g T_p \chi_A \geq 0$, y $0 \leq \langle T_p^t g, \chi_A\rangle = \int_A T_p^t g$ para todo A medible. Esto implica que $T_p^t g \geq 0$. Luego, como $|g| \leq \|g\|_\infty$ entonces $|T_p^t g| \leq T_p^t |g| \leq T_p^t \|g\|_\infty = \|g\|_\infty$ ya que $T_p^t g$ manda constantes en constantes. $\square$

Teorema 5.23. *Si T es un operador estocástico entonces T^t esta definido en L_∞.*

Demostración. Por el Teorema anterior $|T^t g| \leq \|g\|_\infty$ para toda $g \in L_\infty$. $\square$

Definición 5.24. Un operador estocástico T es doblemente estocástico si existe E un operador estocástico tal que $E|_{L_\infty} = T^t$.

Teorema 5.25. *Si T es un operador doblemente estocástico, entonces $T : L_\infty \to L_\infty$.*

Demostración. Sean T doblemente estocástico y E estocástico tal que $E = T^t$ en L_∞. Entonces $E^t : L_\infty \to L_\infty$ y $E^t 1 = 1$. Sea $g \in L_\infty$. Tenemos que $\langle E^t g, f\rangle = \langle g, Ef\rangle = \langle g, T^t f\rangle = \langle Tg, f\rangle$ para toda $f \in L_1$. Por lo tanto, $T = E^t : L_\infty \to L_\infty$. E^t es continuo en un subconjunto denso de L_1 y por lo tanto se extiende a un operador acotado en en L_1. Esta extensión es igual a T en L_∞, por lo que esta extensión debe ser T. $\square$

Teorema 5.26. *T es un operador estocástico con $T1 = 1$ si y sólo si T es doblemente estocástico.*

Demostración. Por cada operador lineal positivo y acotado, T, definido en L_1, T^t se extiende a un operador lineal positivo y acotado E definido en L_1. Supongamos que $T1 = 1$, entonces $\int Ef = \int T^t f = \int (T1) f = \int f$ para toda $f \in L_1$. Entonces E es estocástico y por lo tanto T es doblemente estocástico. Si T es doblemente estocástico entonces es estocástico y $\int T^t f = \int f$ para toda $f \in L_1$. Entonces $\int T1 f = \int f$ para toda $f \in L_1$ y por lo tanto $T1 = 1$. $\square$

En los casos finito e infinito numerable es inmediato identificar los operadores estocásticos y doblemente estocásticos con medidas en $\{1, ..., n\} \times \{1, ..., n\}$ y $\mathbb{N} \times \mathbb{N}$ respectivamente, si (p_{ij}) es la matriz que induce a un operador entonces $\mu(i, j) := p_{ij}$

define una medida en estos espacios. Notemos que para el caso de los operadores doblemente estocásticos las medidas marginales son uniformes. Además cualquier medida con marginales uniformes en estos espacios se puede asociar con una matriz doblemente estocástica y por lo tanto con un operador doblemente estocástico.

En el caso infinito no numerable tambien es posible esta identificación, las medidas con las que identificamos los operadores doblemente estocásticos son las medidas doblemente estocásticas definidas en el Capitulo 2. Es decir, medidas de probabilidad definidas en $I \times I = [0,1] \times [0,1]$ tales que $\mu(A \times I) = \mu(I \times A) = m(A)$. El siguiente teorema nos dice cómo es esa correspondencia.

Teorema 5.27. *Sea T un operador doblemente estocástico y μ una mediada doblemente estocástica. La relación $\mu(A \times B) = \int_I \chi_A T_{\chi_B}$ para todos borelianos A y B, define una correspondencia biyectiva entre el conjunto de operadores doblemente estocásticos y el de medidas doblemente estocásticas.*

Demostración. Sea μ una medida doblemente estocástica. Sea $g \in L_\infty(m)$ y f una función simple en $L_1(m)$. Como

$$\begin{aligned}\int_{I\times I} \chi_A(x) d\mu(x,y) &= \int_{I\times I} \chi_{A\times I}(x,y) d\mu(x,y) = \mu(A \times I) = m(A) \\ &= \int_I \chi_A(x) dm(x)\end{aligned}$$

entonces

$$\int_{I\times I} f(x) d\mu(x,y) = \int_I f(x) dm(x)$$

para toda funcion simple. Por argumentos de aproximación por funciones simples usuales tenemos que

$$\int_{I\times I} f(x) d\mu(x,y) = \int_{I\times I} f(y) d\mu(x,y) = \int_I f(x) dm(x)$$

para toda $f \in L_1(m)$.

Sea $g \in L_\infty(m)$, $g > 0$, definimos

$$G(f) := \int_{I\times I} f(x) g(y) d\mu(x,y).$$

Entonces

$$|G(y)| \leq \int_{I\times I} |f(x)||g(y)| d\mu(x,y) \leq \|g\|_\infty \int_{I\times I} |f(x)| d\mu(x,y) = \|g\|_\infty \|f\|_1.$$

Así que G es un funcional lineal acotado y por lo tanto continuo en L_1, entonces por el Teorema de representación de Riesz (Teorema A.4 del apéndice) existe $h \in L_\infty$ tal que $G(f) = \langle f, h\rangle$ para toda $f \in L_1$. Definimos $T^t : L_\infty(m) \to L_\infty(m)$ por $T^t(g) = h$, de modo que

$$G(f) = \int_{I\times I} f(x)g(y)d\mu(x,y) = \int_I f(x)h(x)dm(x) = \int_I f(x)T^t g(x)dm(x).$$

Como $G(f) \geq 0$ siempre que $f \geq 0$ y como $g > 0$, tenemos que $T^t g \geq 0$. Entonces T^t es positivo. Además

$$G(1) = \int_{I\times I} g(y)d\mu(x,y) = \int_I g(y)dm(y) = \langle 1, T^t g\rangle = \int_I T^t g(y)dm(y)$$

para toda $g \in L_\infty$, por lo que $T1 = 1$. Luego si $g = 1$ entonces $\int_{I\times I} f(x)1d\mu(x,y) = \int_I f(x)dm(x) = \langle f, T^t 1\rangle = \int_I fT^t dm(x)$ para toda $f \in L_1$ por lo que $T^t 1 = 1$. Tenemos entonces que $T^t : L_\infty \to L_\infty$, es tal que $T \geq 0$ y $T1 = T^t 1 = 1$, y por lo tanto T es doblemente estocástico.

Por otra parte supongamos que T es doblemente estocástico y definamos para borelianos A y B la función $\lambda(A \times B) = \int_I \chi_A T \chi_B dm$. Entonces λ es finitamente aditiva. Además como T es doblemente estocástico tenemos que

$$\lambda(A \times I) = \int \chi_A T \chi_I = \int \chi_A = m(A)$$

y

$$\lambda(I \times A) = \int \chi_I T \chi_A = \int T\chi_A = \int \chi_A = m(A),$$

donde m es la medida de Lebesgue. Como la medida de Lebesgue es regular, para cada boreliano A y $\epsilon > 0$ existen compactos $A_1 \subset A$ y $B_1 \subset B$ tales que $m(A - A_1) < \epsilon$ y $m(B - B_1) < \epsilon$. De modo que para el compacto $A_1 \times B_1 \subset A \times B$,

$$\begin{aligned} \lambda((A \times B) - (A_1 \times B_1)) &\leq \lambda(A \times (B - B_1)) + \lambda((A - A_1) \times B) \\ &\leq \lambda(I \times (B - B_1)) + \lambda((A - A_1) \times I) \\ &= m(B - B_1) + m(A - A_1) < 2\epsilon. \end{aligned}$$

Por lo tanto λ es regular en los rectángulos medibles de en $I \times I$, por el teorema de Alexandroff (A.5 en el Apéndice) λ es contablemente aditiva, y por el Teorema de extensión de Hahn (A.6 en el Apéndice) tiene una única extension a los borelianos de $I \times I$. $\square$

Observación 5.28. El teorema de Douglas-Lindenstrauss tambien caracteriza los puntos extremos del conjunto de operadores doblemente estocásticos.

Definición 5.29. Una función $\phi : [0,1] \to [0,1]$ es *preservadora de medida* si $m(\phi^{-1}(A)) = m(A)$ para todo conjunto medible A.

Proposición 5.30. *Si $\phi : I \to I$ es una función preservadora de medida, si para toda $f \in L_1$ definimos $T_\phi(f) := f \circ \phi$, entonces T_ϕ es un operador doblemente estocástico.*

Demostración. Si $f \geq 0$ entonces $f \circ \phi \geq 0$ por lo que T_ϕ es positivo. Luego para todo $A \subset I$ medible se tiene que $\int T_\phi \chi_A(x)dm = \int \chi_A \circ \phi(x)dm = \int \chi_{\phi^{-1}(A)}(x)dm = m(\phi^{-1}(A)) = m(A)$. Se sigue que que para toda función simple f, $\int T_\phi f dm = \int f dm$. Y por argumentos de aproximación por funciones simplies que $\int T_\phi f dm = \int f dm$ para toda $f \in L_1$. Ademas, $T_\phi 1 = 1$ por lo tanto T_ϕ es doblemente estocástico. □

Denotaremos por $\mathbb{D}$ al conjunto de operadores doblemente estocásticos.

Proposición 5.31. *Sea $\phi : I \to I$ una función preservadora de medida. Entonces T_ϕ es punto extremo de $\mathbb{D}$.*

Demostración. Supongamos que $T_\phi = \frac{1}{2}(T_1 + T_2)$ donde T_1 y T_2 son doblemente estocásticos. Entonces $T_\phi \chi_A = \chi_{\phi^{-1}A} = \frac{1}{2}T_1\chi_A + \frac{1}{2}T_2\chi_A$. Entonces si $\phi(x) \in A$, $2 = T_1\chi_A(x) + T_2\chi_A(x)$ y $T_1\chi_A(x) = T_2\chi_A(x) = 1$. De manera similar, si $\phi(x) \notin A$ entonces $T_1\chi_A(x) = T_2\chi_A(x) = 0$. Entonces $T_1\chi_A = T_2\chi_A = \chi_{\phi^{-1}A} = T_\phi A$ y T_ϕ es extremo. □

Definición 5.32. Un operador doblemente estocástico T es un *operador de permutación* si y sólo si existe una función preservadora de medida invertible $\phi : I \to I$ para la cual $T = f \circ \phi$.

Denotaremos por $\mathbb{P}$ al conjunto de operadores de permutación.

Para el caso continuo hay operadores extremos que no son operadores de permutación. Por ejemplo, operadores inducidos por transformaciones preservadoras de medida no invertibles. Sin embargo, mediante la elección de una topología adecuada el conjunto de operadores doblemente estocásticos es el casco cerrado convexo del conjunto de operadores de permutación, lo que sería la version no numerable del Teorema de Birkhoff.

Definición 5.33. Sean X y Y espacios de Banach (métricos y completos), y sea $B(X,Y)$ la clase de todos los operadores lineales y continuos $B : X \to Y$. La *top*

ología fuerte de operadores en $B(X,Y)$ es la topología definida por la base de vecindades
$$N(T;A,\epsilon) := \{S \in B(X,Y) : |(T-S)x| < \epsilon, x \in A\}$$
donde A es un subconjunto finito arbitrario de X, y $\epsilon > 0$ arbitraria.
La *topología débil de operadores* en $B(X,Y)$ es la topología definida por la base de vecindades

$$N(T;A,B,\epsilon) := \{S \in B(X,Y) : |y^*(T-S)x| < \epsilon, x \in A, y^* \in B\}$$

donde A y B son subconjuntos finitos arbitrarios de X y Y^* respectivamente, y $\epsilon > 0$ arbitraria.
Entonces una sucesión $\{T_n\}$ converge a T en la topología fuerte de operadores si y sólo si la sucesión $\{T_n x\}$ converge a Tx para todo $x \in X$. Y converge a T en la topología débil de operadores si y sólo si $\{y^* T_n x\}$ converge a $y^* T x$ para todo $y^* \in Y^*$ y para todo $x \in X$.

Definición 5.34. Sea B el espacio de Banach de operadores lineales $T : L_p \to L_p$, $1 \leq p < \infty$. La *topología fuerte de operadores* en B es la topología para la cual T_n converge a T si y sólo si $\|T_n f - Tf\|$ converge a cero para toda $f \in L_p$. La *topología débil de operadores* en B es la topología para la cual T_n converge a T si y sólo si $|\langle g, T_n f\rangle_{L_q - L_p} - \langle g, Tf\rangle_{L_q - L_p}|$ converge a cero para toda $f \in L_p$ y para toda $g \in L_q$, $\frac{1}{q} + \frac{1}{p} = 1$.

Veremos primero que $\mathbb{P}$ es denso en $\mathbb{D}$ en la topología débil de operadores. Se seguirá del hecho que cerradura de convexos en las topologías fuerte y débil de operadores coinciden ([DS58] p. 477) que $\mathbb{D}$ es el casco cerrado convexo de $\mathbb{P}$.

Lema 5.35. *$\mathbb{P}$ es denso en $\mathbb{D}$ en la topología debil de operadores.*

Demostración. Una base para la topología debil de operadores en $\mathbb{D}$ está dada por conjuntos básicos de la forma

$$\{T : |\langle f_k, Tg_k\rangle - \langle f_k, Sg_k\rangle| < \epsilon, k = 1, ..., n\}$$

donde f_k, g_k varían sobre un conjunto denso de L_2 y $\epsilon > 0$, y $S \in \mathbb{D}$. Tomaremos f_k y g_k continuas. En este caso son acotadas y dado que ϵ es arbitrario las tomaremos acotadas por 1. Sea $S \in \mathbb{D}$. Mostraremos que existe T_ϕ en un básico. Sean λ y λ_ϕ las medidas doblemente estocásticas asociadas con S y T_ϕ respectivamente. Definamos $h_k(x,y) = f_k(x) g_k(y)$, $k = 1, ..., n$. Entonces h_k es uniformemente continua en $I \times I$. Entonces
$$\langle f_k, Sg_k\rangle = \int_{I\times I} h_k(x,y) d\lambda$$

y

$$\langle f_k, T_\phi g_k \rangle = \int_{I\times I} h_k(x,y) d\lambda_\phi.$$

Dado que cada h_k es uniformemente continua en $I \times I$, podemos encontrar $I_1, I_2, ...I_s$ intervalos disjuntos tales que $I = \bigcup_{r=1}^{s} I_r$ y tales que h_k no varía mas de ϵ en cada rectángulo $I_i \times I_j$, $i, j = 1, ..., s$.

Como λ es doblemente estocástica, podemos escoger para cada $i = 1, ..., s$, conjuntos medibles disjuntos $X_{ij} \subset I_i$ tales que $m(X_{ij}) = \lambda(I_i \times I_j)$. De manera similar, para cada $j = 1, ..., s$ podemos encontrar conjuntos medibles disjuntos $Y_{ij} \subset I_j$ tales que $m(Y_{ij}) = \lambda(I_i \times I_j)$. Entonces para cada i y j existe una función preservadora de medida invertible (ver Teorema A.10 del apéndice) $\phi_{ij} : X_{ij} \to Y_{ij}$ pues $m(X_{ij}) = m(Y_{ij})$ para toda i, j. Definimos ϕ como

$$\phi(x) = \phi_{ij}(x) \quad x \in X_{ij}.$$

Entonces $\phi : I \to I$ es invertible y preservadora de medida, mostraremos que T_ϕ está en un básico. Dado que $\phi_{ij} : X_{ij} \to Y_{ij} \subset I_j$ y dado que $I_i = \bigcup_{r=1}^{s} X_{ir}$ se sigue que

$$\lambda_\phi(I_i \times I_j) = m(I_i \cap \phi^{-1} I_j) = m(X_{ij}) = \lambda(I_i \times I_j).$$

Entonces

$$\begin{aligned} |\langle f_k, T_\phi g_k \rangle - \langle f_k, S g_k \rangle| &= \left| \int_{I\times I} h_k d\lambda_\phi - \int_{I\times I} h_k d\lambda \right| \\ &\leq \sum_{i,j=1}^{s} \left| \int_{I_i \times I_j} h_k d\lambda_\phi - \int_{I_i\times I_j} h_k d\lambda \right| \\ &\leq \sum_{i,j=1}^{s} \epsilon[\lambda_\phi(I_i \times I_j) + \lambda(I_i \times I_j)] + |\lambda_\phi(I_i \times I_j) - \lambda(I_i \times I_j)| \\ &= 2\epsilon \sum_{i,j=1}^{s} \lambda(I_i \times I_j) = 2\epsilon. \end{aligned}$$

□

Teorema 5.36. $\mathbb{D} = \overline{cc}(\mathbb{P})$ *en la topología fuerte de operadores.*

Demostración. Probaremos que $\mathbb{D}$ es compacto en la topología débil de operadores en el espacio de Banach de operadores lineales acotados en L_p, $1 < p < \infty$. Dado que los operadores doblemente estocásticos pertenecen a este espacio para toda $1 \leq p \leq \infty$, basta mostrar que $\mathbb{D}$ es cerrado en la esfera unitaria. Como los conjuntos convexos tienen la misma cerradura en la topología débil y topología fuerte de operadores, se sesguirá el resultado del Teorema de Krein-Milman.

Supongamos que $T_n \in \mathbb{D}$ y $\langle g, T_n f\rangle$ converge a $\langle g, Tf\rangle$ para toda $f \in L_p$ y para toda $g \in L_q$ con $\frac{1}{p}+\frac{1}{q}=1$. Si $f \geq 0$ y $g \geq 0$ entonces $\langle g, T_n f\rangle \geq 0$ pues T_n es positivo. Entonces $\langle g, Tf\rangle = \int gTf \geq 0$ para toda $f \in L_p$, $g \in L_q$. Tomando $g = \chi_A$ entonces para todo conjunto medible A se tiene que $Tf \geq 0$ cuando $f \geq 0$. Entonces T es positivo. Tomando $g = 1$, entonces $\int T_n f$ converge a $\int Tf = \int f$ pues $\int T_n f = \int f$ para toda n. Además $T1 = 1$ pues $T_n 1 = 1$ para toda n. Entonces $T \in \mathbb{D}$ y $\mathbb{D}$ es cerrado en la topología débil. Entonces $\mathbb{D}$ es compacto y por el Teorema de Krein-Milman es el casco cerrado convexo de sus puntos extremos en la topología débil. Entonces es el casco cerrado convexo en la topología fuerte. Como $\mathbb{P}$ es denso en $\mathbb{D}$ en la topología fuerte se sigue que $\mathbb{D} = \overline{\mathrm{cc}}(\mathbb{P})$. $\square$

Apéndice A

Apéndice

A continuación se enuncian los teoremas que son usados en varias de las demostraciones de eśte trabajo y las referencias de donde se pueden consultar las demostraciones.

Teorema A.1. *Sean B un conjunto convexo y cerrado en un espacio localmente convexo X y $x_0 \in X - B$. Entonces existe $f \in X^*$ tal que $|f(x)| \leq 1$ para todo $x \in B$, pero $f(x_0) > 1$.*

Demostración. En [Rud91] p. 61. $\square$

Teorema A.2 (Radon - Nikodym)**.** *Sea μ una medida σ-finita en en espacio medible (X, Σ). Sea ν una medida finita, absolutamente continua respecto a μ. Entonces para alguna $h \in L_1(\mu)$,*

$$\nu(E) = \int_E h d\mu$$

para todo $E \in \Sigma$.

Demostración. En [Dud89] p. 134. $\square$

Teorema A.3 (Hahn-Banach)**.** *Sean M es un subespacio de un espacio vectorial topológico localmente convexo X y $x_0 \in X$. Si x_0 no esta en la cerradura de M, entonces existe $f \in X^*$ tal que $f(x_0) = 1$ pero $f(x) = 0$ para toda $x \in M$.*

Demostración. En [Rud91] p. 60 $\square$

Teorema A.4 (Representación de Riesz)**.** *Para cualquier espacio de medida σ-finito (X, Σ, μ), $1 \leq p < \infty$ con $\frac{1}{p} + \frac{1}{q} = 1$, la funcion T definida por*

$$T_g(f) := \int f g d\mu = \langle f, g \rangle_{L_q - L_p}$$

es una isometría lineal de $(L_q, \|.\|)$ *en* $(L_p^*, \|.\|_p^*)$

Demostración. En [Dud89] p. 162. □

Teorema A.5 (Alexandroff)**.** *Sea* μ *una funcion con valores reales, regular, acotada y finitamente aditiva, definida en un álgebra* Σ *de subjuconjuntos de un espacio topológico compacto* S*. Entonces* μ *es contablemente aditiva.*

Demostración. [DS58] p. 138. □

Teorema A.6 (Extensión de Hahn)**.** *Toda función* μ *contablemente aditiva definida en un álgebra de subconjuntos* Σ *con valores en los reales extendidos tiene una extensión al* σ*-álgebra generado por* Σ*. Si además* μ *es* σ*-finita en* Σ *entonces la extensión es única.*

Demostración. [DS58] p. 136. □

Teorema A.7 (Teorema de Densidad de Lebesgue)**.** *Seas* m *la medida de Lebesgue,* $A \subset [0,1]^2$ *un conjunto Lebesgue medible y* $B(x,r)$ *la bola con centro en* x *y radio* r*, entonces para casi todo* x *en* A*.*

$$\lim_{\epsilon \to 0} \frac{m(A \cap B(x,\epsilon))}{m(B(x,\epsilon))} = 1$$

A los elementos de A que cumplen esta propiedad se les llama *puntos de densidad de Lebesgue* de A.

Demostración. En [Rud86] p. 141. □

Teorema A.8. *Sea* X *un espacio vectorial topologógico y* $\mathcal{F}$ *un espacio vectorial de funcionales lineales que separa puntos de* X*. Sea* τ *es la topología más débil que hace continua a toda* $f \in \mathcal{F}$*. Entonces* (X, τ) *es un espacio localmente convexo cuyo dual* $X^* = \mathcal{F}*$*.*

Demostración. [Rud91] pp. 62-65. □

Teorema A.9. *Sean* X*,* Y *dos espacios de Banach. Sea* $B(X,Y)$ *el conjunto de operadores lineales* $T : X \to Y$*. Entonces un conjunto convexo en* $B(X,Y)$ *tiene la misma cerradura en la topología débil de operadores que en la topología fuerte de operadores.*

Demostración. En [DS58] p. 477. □

Teorema A.10 (Sirkoski)**.** *Sea* (X, A, μ) *un espacio de medida con conjuntos nulos, y sea* Φ *un homeomorfismo de* σ*-álgebras de medidas de la familia* $\mathcal{B}$ *de borelianos en* $[0, 1]$ *a la* σ*-álgebra* A*, con* $\Phi([0, 1]) = X$*. Entonces existe una transformación medible de* X *en* $[0, 1]$ *tal que para todo* $B \in \mathcal{B}$ $\phi^{-1}[B]$ *está en la clase de equivalencia de* $\Phi[B]$*.*

Demostración. En [Roy88] p. 397. $\square$

Bibliografía

[Bro66] James R. Brown, Aproximation theorems for Markov Operators,en *Pacific J. Math.* Vol. 16 Número 1, 1966, pp 13-23.

[Den80] J. L. Denny, The support of discrete extremal measures with given marginals. *Michigan Math. J.* 27, pp 59-64, 1980.

[DS58] Dunford N., Schwartz J., *Linear Operators Part 1*, New York, Wiley-Interscience, 1958.

[Dou64] R.G. Douglas, On extremal measures and subspace density, *Michigan Math. J.* 11, pp. 243-246. 1964.

[Dud89] Richard M. Dudley, *Real Analysis and Probability Theory*, Wadsworth & Brooks/Cole Mathematics Series, 1989.

[Gir95] Valerie Girardin, Extreme measures with given moments or marginals, *Can. J. Math.* Vol. 47 1995 pp 946-958.

[HW95] Kevin Hestir and Stanley Williams, Supports of doubly stochastic measures, *Bernoulli 1* No. 3, pp. 217-243, 1995.

[Ken60] David G. Kendall, On Infinite Doubly-Stochastic Matrices nad Birkhoff's Problem 111, *Journal London Math. Soc.* 35, 1960, pp. 81-84.

[KN77] M. Krein y A. Nudelman, The Markov Moment Problem and extremal problems, *Amer. Math. Soc.*, 1977.

[KS04] Sheila King y Ray shiflett, Doubly Stochastic Operators and the History of Birkhoff's Problem 111, en *Stochastic Processes and Functional Analysis: A Volume of Recent Advances in Honor of M.M. Rao*, A Dekker series in lecture notes in Pure and Applied Mathematics vol. 238, CRC Press, pp 411-440, 2004.

[Lin65] Joram Lindenstrauss, A Remark on extreme doubly stochastic measures, *Amer. Math. Monthly* 72 pp 279-382, 1965.

[Roy88] Royden, H.L. *Real Analysis*, Third Edition. McMillan Publishing Company. 1988.

[Rud86] Walter Rudin. *Real and Complex Analysis*, Third Edition. McGraw-Hill. 1986.

[Rud91] Walter Rudin. *Functional Analisis*, International Series on Pure and Applied Mathemathics, Mc. Graw-Hill. 1991.