UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA

"APROXIMACIÓN TIPO KOROVKIN Y SISTEMAS DE CHEBYSHEV"

T E S I S

PARA OBTENER EL GRADO DE:
MAESTRO EN MODELACIÓN MATEMÁTICA

PRESENTA:
JOSÉ LUIS CARRASCO PACHECO

DIRECTOR DE TESIS:
DR. JOSÉ MARGARITO HERNÁNDEZ MORALES

HUAJUAPAN DE LEÓN, OAXACA, MÉXICO, MARZO DE 2016

Agradecimientos

Agradezco a mi asesor de tesis Dr. José Margarito Hernández Morales por su comprensión, tiempo, paciencia y apoyo.

A mis sinodales: Luz del Carmen Álvarez Marín, Cuauhtemoc Héctor Castañeda Roldán, Tirso Miguel Ángel Ramírez Solano y Vulfrano Tochihuitl Bueno, por las observaciones que hicieron de este trabajo y por el tiempo que invirtieron en la revisión de esta tesis.

Gracias a todos mis compañeros de trabajo que me apoyaron en el transcurso de la maestría, pues sin ellos esto no hubiera sido posible.

También le agradezco a mis amigos por estar conmigo en las buenas y en las malas.

Por último y no menos importante le doy gracias a Oli, que durante la mayor parte del tiempo estuvo apoyándome y dándome aliento para no desistir.

Dedicatoria

A Josué, David, Oli y a mis padres, Gloria y Temo.

Por su infinita paciencia, amor y el apoyo incondicional que me han brindado, principalmente en aquellos momentos tan difíciles de mi vida.

Índice

Simbología

Notación	**Significado** (eventualmente puede tener otro uso)
$\mathbb{R}^+$	$\{t \in \mathbb{R} : t \geq 0\}$
$\widehat{\mathbb{R}^+}$	$\mathbb{R}^+ \cup \{\infty\}$
μ	Medida
$\mu(f)$	Número o constante de Lipschitz para la función f
$\theta(f)$	Abreviación de $\theta_\alpha(f)$
$C(X)$	Espacio de funciones reales continuas definidas en X
$L^p(X)$	Espacio de funciones medibles de p-ésima potencia integrable
$K(X)$	Espacio de funciones continuas de soporte compacto
ß(X)	Espacio de funciones reales acotadas
$C_{\mathrm{b}}(X)$	Espacio de funciones reales continuas acotadas
$C_0(X)$	Espacio de funciones continuas que se anulan en el infinito
$Sop(f)$	Soporte de la función f
$convB$	Casco convexo del conjunto B
$Sop(\mu)$	Soporte de la medida μ
$\mu(X)$	Espacio de medidas de Radon sobre $K(X)$
$\mu^+(X)$	Cono de medidas de Radon positivas
$\mu_b(X)$	Medidas de Radon acotadas

Simbología

$\mu_b^+(X)$	Medidas de Radon positivas y acotadas
K_n	Operadores lineales
$(E, \|\| \cdot \|\|_E)$	Espacio lineal normado
$g \cdot \mu$	Medida con densidad g con respecto a μ
$D_+(H,\mu)$	Subespacio de Korovkin para μ
$U_+(H,\mu)$	Subespacio igual a $D_+(H,\mu)$
$D_+^1(H,\mu)$	Igual a $D_+(H,\mu)$ para $\\|\mu\\| = 1$
$U_+^1(H,\mu)$	Igual a $U_+(H,\mu)$ para $\\|\mu\\| = 1$
$K_+(H,T)$	Subconjunto de Korovkin para el operador lineal T
$K_+^1(H,T)$	Igual a $K_+(H,T)$ con $\\|T\\| = 1$.
$\mathcal{L}_n(X,Y)$	Conjunto de operadores finitamente definidos
$\mathcal{L}_n^1(X,Y)$	Conjunto de contracciones finitamente definidos
$C_+(x_1,...,x_n)$	Medidas con soporte $x_1,...,x_n$
$C_+^1(x_1,...,x_n)$	Medidas contracción con soporte $x_1,...,x_n$
$\widetilde{C}_+^1(x_1,...,x_n)$	Medidas probabilidad
$r(\mu^*)$	Resultante de la medida μ
$\partial_e K$	Frontera de Choquet
H_u^*	Funciones f aproximables a la derecha por $h \in H$
H_l^*	Funciones f aproximables a la izquierda por $h \in H$
H^*	$H_u^* \cap H_l^*$
p	Norma asimétrica en $\mathbb{R}$
$\overline{p}$	Norma asimétrica conjugada de p
p^s	$\max\{p, \overline{p}\}$

Simbología

ω	Punto en el infinito
$w(\cdot)$	Función auxiliar
f	Función real continua en X
$\widetilde{f}$	Extensión de f a X_ω
ξ_x	Medida de Dirac
$\sum_{i=1}^{n} \alpha_i \xi_x$	Medida discreta
$\oplus$	Suma directa
$\Re e$	Parte real
$\mathbb{T}$	Espacio $[0, 2\pi)$
$C(\mathbb{T})$	Funciones continuas sobre $\mathbb{T}$

Introducción

A continuación explicamos a grandes rasgos el contenido de esta tesis y recopilamos algunos de los resultados que la motivan. A partir del Capítulo 1 se retoman los temas con más detalle.
Algunas veces podemos aproximar funciones complicadas mediante una sucesión de funciones más simples (con las cuales es más fácil de trabajar en la solución de un determinado problema) que dan una solución satisfactoria en ciertas aplicaciones. Hoy en día existe bastante teoría sobre tipos de aproximación de funciones, algunos de estos son: aproximación lineal, aproximación no lineal, aproximación uniforme, aproximación de Chebyshev, aproximación tipo Korovkin, aproximación de Lipschitz, aproximación con peso, aproximación con peso sensible al signo, aproximación lateral, etc.

Fue el matemático ruso P. L. Chebyshev, en 1854, quien desarrolló los conceptos que sientan las bases de la Teoría de Aproximación, mediante el problema: dada una función continua f encontrar un polinomio algebraico p de grado menor o igual a n, de tal forma que el máximo de su desviación con respecto a f sobre un intervalo dado, sea más pequeño que el de los otros polinomios con las mismas características. En el problema mencionado, los sistemas de Chebyshev juegan un papel muy importante, ya que tienen una caracterización interesante del polinomio de mejor aproximación. De hecho, un teorema clásico de Haar establece que toda función $f \in C(X)$ tiene un polinomio de mejor aproximación en el subespacio generado por $\{h_0, ...h_n\} \subseteq C(X)$ y, además éste es único si y sólo si $h_0, ..., h_n$ es un sistema de Chebyshev de orden $n + 1$, [5] .

Por otra parte, en 1953 Korovkin (otro matemático ruso) estableció un teorema que con el tiempo se haría muy célebre. Por su simplicidad y al mismo tiempo su poder fue que despertó el interés de muchos matemáticos. Se trata de un criterio que permite decidir cuándo una sucesión de operadores lineales positivos $K_n : C([0, 1]) \longrightarrow C([0, 1])$ converge uniformemente al operador identidad $Id : C([0,1]) \longrightarrow C([0,1])$. Korovkin estableció que basta con verificar la convergencia uniforme para $f \in \{1, x, x^2\}$, este último es llamado un subconjunto de Korovkin. Después se han encontrado otros subconjuntos con las mismas propiedades que $\{1, x, x^2\}$ y se les ha llamado también subconjuntos de Korovkin, mientras que a los espacios generados por estos conjuntos se les llama subespacios de Korovkin, [1] y [2]. En tiempos recientes se ha encontrado una relación muy interesante entre los subespacios de Chebyshev (espacios generados por un sistema de Chebyshev) y los subespacios de Korovkin.

El trabajo en esta tesis trata sobre la relación existente entre las dos teorías mencionadas. Específicamente la relación entre los subespacios de Chevyshev y los subespacios de Korovkin, así como el orden de un subespacio visto como un subespacio de Chebyshev y, el orden de ese mismo subespacio visto como un subespacio de Korovkin. Lo anterior se desarrolla en los cuatro capítulos siguientes, de los cuales damos aquí una breve descripción. En el capítulo inicial de preliminares damos algunos conceptos escenciales, como son: Espacios de funciones con los que trabajamos principalmente, medidas de Radon, así como otros relacionados con éstas. También en esta parte coleccionamos las principales características topológicas de los espacios resultantes, las cuales se utilizan en los capítulos restantes para estudiar las teorías de Chebyshev y de Korovkin.

En el Capítulo 2 mencionamos los conceptos básicos de la teoría de aproximación de Chebyshev, haciendo énfasis en los conceptos de mejor aproximación, polinomio de la mejor aproximación, resultados sobre la caracterización del mejor aproximante, sistemas y subespacios de Chebyshev, sistemas de Haar y caracterizaciones de sistemas de Chebyshev.

En el Capítulo 3 consideramos los D_+-subespacios ($D_+(H,\mu)$) para una medida de Radon, la igualdad de estos últimos con los U_+-subespacios ($U_+(H,\mu)$), con los cuales es más fácil de trabajar y llegar a la igualdad de los $K_+(H,T)$ con la intersección de los subespacios $D_+(H,\mu_y^T)$, esto es

$$K_+(H,T) = \bigcap_{y \in Y} D_+(H,\mu_y^T) = \bigcap_{y \in Y} U_+(H,\mu_y^T).$$

Relaciones similares son establecidas para Operadores contracción y medidas de Radon contractivas. Finalmente en el Capítulo 4 tratamos sobre los órdenes de los subespacios de Chebyshev y los órdenes de los subespacios de Korovkin, para despues establecer la relación entre los dos órdenes mencionados en espacios de funciones con un dominio particular.

Capítulo 1

Preliminares

1.1 Conceptos topológicos

En esta sección se presenta las definiciones y las principales propiedades de los espacios compactos, localmente compactos y σ-compactos, así como los principales espacios de funciones definidas sobre éstos, los cuales trabajamos principalmente en la teoría de Korovkin. En un espacio topológico X, una red (o sucesión generalizada) de elementos de X es una familia $(x_i)_{i\in I}$ tal que sobre el conjunto de índices I existe un orden parcial $\leq$ con respecto al cual $(I, \leq)$ es un conjunto dirigido. Se dice que la red $(x_i)_{i\in I}$ converge a x_0 si para cada vecindad V de x_0 existe un indice i_v, tal que $x_i \in V$ para cada $i \geq i_v$. En tales casos se dice que x_0 es un límite de la red $(x_i)_{i\in I}$.

Una familia filtrada de elementos de X es una familia $(x_i)_{i\in I,\mathcal{F}}$ tal que sobre el conjunto de índices I está definido un filtro fijo $\mathcal{F}$. la familia filtrada $(x_i)_{i\in I,\mathcal{F}}$ converge a x_0 si para cada vecindad V de x_0 existe un elemento F de $\mathcal{F}$, tal que $x_i \in V$ para cada $i \in F$. Similarmente en tales casos se dice que x_0 es un límite de la familia filtrada $(x_i)_{i\in I,\mathcal{F}}$.

En los dos casos anteriores, si X es Hausdorff, el límite si existe es único. Si X y Y son dos espacios topológicos, entonces una función $f : X \longrightarrow Y$ es continua en un punto x_0 si y sólo si para cada red (familia filtrada) $(x_i)_{i\in I}$ que converge a x_0 en X, se tiene que $(f(x_i))_{i\in I}$ converge a $f(x_0)$ en Y.

Dado el espacio topológico X una función numérica $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ ($f : X \longrightarrow \mathbb{R} \cup \{-\infty\}$, respectivamente) se dice que es semicontinua inferiormente (semicontinua superiormente, respectivamente) en el punto x_0, si para todo $\lambda \in \mathbb{R}$ satisfaciendo $f(x_0) > \lambda$ ($f(x_0) < \lambda$, respectivamente),

existe una vecindad V de x_0 tal que $f(x) > \lambda$ ($f(x) < \lambda$, respectivamente) para cada $x \in V$. La función f se dice que es semicontinua inferiormente (semicontinua superiormente, respectivamente), si lo es en cada punto de X. Una caracterización que es la que más se utiliza es que f se dice que es semicontinua inferiormente (semicontinua superiormente, respectivamente), si y sólo si el subconjunto $\{x \in X : f(x) > \lambda\}$ ($\{x \in X : f(x) < \lambda\}$, respectivamente) es abierto para cada $\lambda \in \mathbb{R}$, o de manera equivalente, el subconjunto $\{x \in X : f(x) \leq \lambda\}$ ($\{x \in X : f(x) \geq \lambda\}$, respectivamente) es cerrado para cada $\lambda \in \mathbb{R}$. Mediante las definiciones anteriores, se tiene que f es continua, si y sólo si f es ambas semicontinua inferior y superiormente.
El espacio X es llamado compacto si cada cubierta abierta de X tiene una subcubierta finita, o de manera equivalente, X es compacto si y sólo si para toda familia filtrada $(x_i)_{i\in I,\mathcal{F}}$ en X, existe un filtro $\mathcal{F}_1$ más fino que $\mathcal{F}$ sobre I, para el cual $(x_i)_{i\in I,\mathcal{F}_1}$ es convergente en X. Cuando X es metrizable, es decir, la topología es la generada por una métrica, entonces X es compacto si y sólo si toda sucesión de puntos en X admite una subsucesión convergente en X. Un espacio métrico compacto es separable.
Un espacio topológico X es localmente compacto si cada uno de sus puntos tiene una vecindad compacta. De hecho, cuando X es Hausdorff y localmente compacto, cada punto tiene un sistema fundamental de vecindades compactas.
Un espacio topológico es llamado σ-compacto si es la unión numerable de subconjuntos compactos. Todo espacio localmente compacto, Hausdorff y σ-compacto es normal.
La compactificación a un punto de Alexandrov X_ω de X, se define como X_ω=$X \cup \{\omega\}$, donde ω es un objeto que no pertenece a X (ω es llamado frecuentemente el punto en el infinito de X). La topología τ_ω en X_ω se define por

$$\tau_\omega = \tau \cup \{X_\omega\text{-}K | K \text{ es compacto en } X\},$$

donde τ denota la topología en X. Resulta que X_ω es un espacio compacto y Hausdorff, X es abierto en X_ω y si X no es compacto, entonces X es denso en X_ω.
Si una red $(x_i)_{i\in I}$ (familia filtrada $(x_i)_{i\in I,\mathcal{F}}$, respectivamente) de elementos de X converge a ω en X_ω, decimos que la red (familia filtrada, respectivamente) converge al punto en el infinito.
Análogamente, si $f : X \longrightarrow \mathbb{R}$ y $\alpha \in \mathbb{R}$, decimos que f converge al punto en el infinito, si para cada $\epsilon > 0$ existe un subconjunto compacto K de X tal

que $|f(x) - \alpha| < \epsilon$ para toda $x \in X - K$. Lo anterior se denota por

$$\lim_{x \to \omega} f(x) = \alpha.$$

1.2 Espacios de funciones

Dado un conjunto X, denotamos por $ß(X)$ al espacio de Banach de todas las funciones reales acotadas definidas sobre X, provisto con la norma de la convergencia uniforme (brevemente, la norma del supremo) definida mediante

$$\|f\| = \sup_{x \in X} |f(x)|, \text{ para cada } f \in ß(X).$$

Para cuando X es un espacio topológico, denotamos por $C(X)$ al espacio de todas las funciones reales continuas sobre X y por $C_b(X) = C(X) \cap ß(X)$ el espacio de funciones reales continuas y acotadas definidas en X. El espacio $C_b(X)$ provisto de la norma del supremo es un espacio de Banach. Si X es localmente compacto, denotamos por $C_0(X)$ al espacio de todas las funciones en $C(X)$ que se anulan en el infinito, lo que significa que la función $\widetilde{f} : X_\omega \longrightarrow \mathbb{R}$ definida por

$$\widetilde{f}(x) = \begin{cases} f(x) & \text{si } x \in X \\ 0 \text{ si} & x = \omega, \end{cases} \tag{1.1}$$

es continua en X_ω.

Claramente, si X es compacto, entonces $C_0(X) = C(X) = C_b(X)$. En general, $C_0(X)$ es un subespacio cerrado de $C_b(X)$ y por tanto él mismo es un espacio de Banach con la norma del supremo. Sobre $C_0(X)$ consideramos el orden natural inducido por el cono $C_0^+(X) := \{f \in C_0(X) : f(x) \geq 0,$ $x \in X\}$. De esta forma, si $f \in C_0(X)$, entonces f^+, f^- y $|f| \in C_0^+(X)$, más aún $\|f\| = \||f|\| = \max\{\|f^+\|, \|f^-\|\}$. Así, $C_0(X)$ y $C(X)$ son latices de Banach, siendo un latice normado E un espacio lineal con una norma que satisface

$$|f| \leq |g| \Longrightarrow \|f\| \leq \|g\| \text{ para cada } f, g \in E.$$

Una propiedad importante de los latices de Banach es la continuidad automática de operadores lineales sobre ellos. Más precisamente, si E es un latice de Banach y F es otro latice normado, entonces todo operador lineal positivo $T : E \longrightarrow F$ es continuo. Más aún, si $E = C(X)$, con X compacto, entonces $\|T\| = \|T(1)\|$ donde 1 denota la función constante 1. En particular, toda forma lineal positiva sobre un latice de Banach es continua. Otro espacio de funciones que juega un rol fundamental en este trabajo es $K(X)$ el espacio de funciones reales continuas definidas en X y que tienen soporte compacto. Siendo el soporte de f definido por

$$Sop(f) = \overline{\{x \in X : f(x) \neq 0\}}.$$

El espacio $K(X)$ es denso en $C_0(X)$. Si X es compacto, entonces éstos coinciden con $C(X)$, y por tanto $K(X) = C_0(X) = C(X) = C_b(X)$.

1.3 Medidas de Radon

Las medidas de Radon son utilizadas en muchas áreas del análisis, como es en la teoría de la probabilidad, la teoría del potencial o la teoría de la representación integral.
En está sección supondremos, al igual que antes, que X es un espacio localmente compacto y Hausdorff. La definición siguiente resulta ser una generalización natural cuando estamos tratando funcionales lineales positivos en espacios localmente compactos Hausdorff, pues la constante positiva M_K siempre existe para cada compacto K, utilizando el lema de Urysohn.

Definición 1 *Un funcional lineal μ: $K(X) \longrightarrow \mathbb{R}$ es llamado una medida de Radon si para cada K compacto, existe una constante $M_K > 0$ tal que $|\mu(f)| \leq M_K \|f\|_\infty$ para toda $f \in K(X)$ con $sop(f) \subset K$.*

El espacio de todas las medidas de Radon definidas sobre $K(X)$ se denota por $\mu(X)$. En nuestro caso nos centramos en el conjunto de las medidas de Radon acotadas $\mu_b(X)$, aquellas que son continuas con respecto a la norma suprema definida por

$$\|\mu\| = \sup_{||f|| \leq 1} |\mu(f)|. \tag{1.2}$$

Además, μ es positiva si $\mu(f) \geq 0$ para toda $f \geq 0$ con $f \in K(X)$, o equivalentemente si $f \leq g$, entonces $\mu(f) \leq \mu(g)$.
Observemos que el conjunto de las medidas de Radon positivas forma un cono lineal normado y es denotado por $\mu^+(X)$, mediante la norma definida por (1.2) en $\mu^+(X)$ por restricción de la norma suprema definida en $\mu(X)$. Al conjunto de medidas de Radon positivas y acotadas lo denotamos por $\mu_b^+(X)$. Una medida de Radon acotada es una contracción si $\|\mu\| \leq 1$, y se denota al conjunto de todas las medidas de Radon positivas de norma 1 por

$\mu_1^+(X)$ y se les llama probabilidades. Es fácil ver que toda medida $\mu \in \mu_b(X)$ puede ser extendida de manera única a una forma lineal sobre $C_0(X)$, a tal extensión se le continúa denotando por μ. Como ejemplos tenemos las más simples medidas, conocidas como medidas de Dirac sobre X. Para $x \in X$ éstas son definidas por:

$$\xi_x(f) = f(x), \text{ para cada } f \in K(X) \text{ (o, } f \in C_0(X)). \tag{1.3}$$

Una medida discreta sobre X es una combinación lineal de medidas de Dirac sobre X. Así, medidas de Radon discretas son medidas de la forma:

$$\mu = \sum_{i=1}^{n} \lambda_i \xi_{x_i}, \tag{1.4}$$

donde $n \geq 1$, $x_1, ..., x_n \in X$ y $\lambda_1, ..., \lambda_n \in \mathbb{R}$. En este caso μ es positiva si y sólo si todo λ_i es positivo. Más aún, se tiene

$$\|\mu\| = \sum_{i=1}^{n} |\lambda_i|.$$

Otro ejemplo es, dada una medida $\mu \in \mu(X)$ y $g \in C(X)$, entonces la funcional lineal $\upsilon : K(X) \longrightarrow \mathbb{R}$ definida por

$$\upsilon(f) = \mu(f \cdot g) \text{ para cada } f \in K(X), \tag{1.5}$$

donde $\cdot$ es el producto de funciones, es nuevamente una medida de Radon en X. Esta es llamada la medida con densidad g y relativa a μ, y es denotada por $g \cdot \mu$.

Decimos que μ es cero sobre un subconjunto abierto U de X si $\mu(f) = 0$ para cada función $f \in K(X)$ cuyo soporte está contenido en U. Se denota por $\Im(\mu)$ a la colección de todos los subconjuntos de X sobre los cuales μ es cero. De tal forma que el soporte de μ es denotado y definido por

$$Sop(\mu) = X \setminus \bigcup_{U \in \Im(\mu)} U.$$

Así, $Sop(\mu)$ es el complemento del mayor subconjunto abierto de X sobre el cual μ es cero. Claramente, un punto x_0 pertenece a $Sop(\mu)$ si para toda vecindad V de x_0 existe una función $f \in K(X)$ con $Sop(f) \subset V$ y $\mu(f) \neq 0$.

Por otra parte, un punto x_0 no pertenece a $Sop(\mu)$ si existe una vecindad V de x_0 sobre la cual μ es cero.

Por un método muy simple es posible extender toda medida $\mu \in \mu_b^+(X)$ a una medida de Radon $\widetilde{\mu}$ sobre la compactificación a un punto X_ω de X. La medida $\widetilde{\mu}$ es definida por

$$\widetilde{\mu}(g) = \mu(g|_X - g(\omega)) + g(\omega) \|\mu\|, \text{ para cada } g \in C(X_\omega). \tag{1.6}$$

Algunas referencia donde se prueban algunas propiedades importantes sobre medidas de Radon son [6],pag.151 y [3]. Algunas de las más importantes son mencionadas en el teorema siguiente y damos demostraciones constructivas, diferentes a las que se han mencionado.

Teorema 1 *Sea X un espacio Hausdorff localmente compacto y $\mu \in \mu(X)$. Entonces*
(1) $Sop(\mu + \upsilon) \subset Sop(\mu) \cup Sop(\upsilon)$ para cada $\upsilon \in \mu(X)$, y si ambas son positivas tenemos la igualdad en la inclusión anterior.
(2) Si $f \in K(X)$ y $f = 0$ sobre $sop(\mu)$, entonces $\mu(f) = 0$. Si $\mu \in \mu_b(X)$, entonces la misma propiedad es válida para cada $f \in C_0(X)$.
(3) Si $f, g \in K(X)$ (o $f, g \in C_0(X)$ siempre que $\mu \in \mu_b(X)$) y $f = g$ sobre $sop(\mu)$, entonces $\mu(f) = \mu(g)$.
(4) Si $\mu \in \mu^+(X)$ y $f \in K(X)$ (o $f \in C_0(X)$ siempre que $\mu \in \mu_b^+(X)$), entonces $\mu(f) \geq 0$ siempre que $f \geq 0$ sobre $Sop(\mu)$.
(5) Si $\mu \in \mu^+(X)$ y $f \in K(X)$, $f \geq 0$ (o $f \in C_0^+(X)$ siempre que $\mu \in \mu_b^+(X)$) y si $\mu(f) = 0$, entonces $f = 0$ sobre $Sop(\mu)$.
(6) Para cada $g \in C(X)$ se tiene $Sop(g \cdot \mu) = \overline{\{x \in Sop(\mu) : g(x) \neq 0\}} \subset Sop(\mu) \cap sop(g)$.
(7) Si $x_1, ..., x_n$ son puntos distintos en X, entonces $Sop(\mu) = \{ x_1, ..., x_n\}$ si y sólo si $\mu = \sum_{i=1}^n \lambda_i \xi_{x_i}$, para algunos $\lambda_1, ..., \lambda_n \in \mathbb{R} - \{0\}$.
(8) Si $Sop(\mu)$ es compacto, entonces μ es acotada.

Demostración. Probemos 1). Si $x \in Sop(\mu + \upsilon)$, entonces para toda vecindad V_x de x, existe $f \in K(X)$ tal que $sop(f) \subseteq V_x$ con $\mu(f) + \upsilon(f) \neq 0$, luego al menos uno de los términos $\mu(f)$ o $\upsilon(f)$ es diferente de cero, por tanto $x \in Sop(\mu)$ o $x \in Sop(\upsilon)$, así se tiene $Sop(\mu + \upsilon) \subseteq Sop(\mu) \cup Sop(\upsilon)$. Ahora supongamos que ambas μ y υ son positivas y notemos que si $f \in K(X)$, entonces $f^+, f^- \in K(X)$, más aún, como $sop(|f|) = sop(f^+ + f^-) \subseteq sop(f^+) \cup sop(f^-) \subseteq sop(f)$, entonces $sop(|f|) = sop(f)$.
Sea $x \in Sop(\mu)$, entonces para toda vecindad V_x de x, existe $f \in K(X)$

tal que $sop(f) \subseteq V_x$ y $\mu(f) \neq 0$, entonces $\mu(|f|) = \mu(f^+) + \mu(f^-) > 0$ con $sop(|f|) = sop(f) \subseteq V_x$, Así como $\mu(f) = \mu(f^+) - \mu(f^-)$, se debe tener $\mu(f^+) > 0$ o $\mu(f^-) > 0$, siguiendose que $\mu(|f|) > 0$, por tanto $(\mu + \upsilon)(|f|) = \mu(|f|) + \upsilon(|f|) > 0$, así dado V_x se tiene $|f| \in K(X)$ tal que $sop(|f|) \subseteq V_x$ con $(\mu + \upsilon)(|f|) = \mu(|f|) + \upsilon(|f|) > 0$, por tanto $x \in Sop(\mu + \upsilon)$, luego $Sop(\mu) \cup Sop(\upsilon) \subseteq Sop(\mu + \upsilon)$.
Probemos 2). Sea $f \in K(X)$ y $f|_{Sop(\mu)} = 0$. Supongamos que $f \geq 0$. sean los conjuntos cerrados $B_n = \{x \in X | f(x) \leq \frac{1}{n+1}\}$ y $A_n = \{x \in X | f(x) \geq \frac{1}{n}\}$, bajo la hipótesis que X es un espacio normal, entonces existe una función continua $g_n : X \longrightarrow [0, 1]$ tal que $g_n(B_n) = 0$ y $g_n(A_n) = 1$ para cada n, así, afirmamos que la sucesión de funciones continuas $h_n = g_n f$ converge a f uniformemente; en efecto tenemos los siguientes casos.

a) Si $x \in A_n$, entonces $|g_n(x)f(x) - f(x)| = 0$.

b) Si $x \in B_n$, entonces $|f(x)| \leq \frac{1}{n+1} < \frac{1}{n}$.

c) Si $x \in \{x \in X | \frac{1}{n+1} < f(x) < \frac{1}{n}\}$, entonces $0 \leq g_n(x)f(x) \leq f(x) < \frac{1}{n}$ y por tanto $|g_n(x)f(x) - f(x)| < \frac{1}{n}|$.

De a),b) y c) se sigue que $|g_n f(x) - f(x)| < \frac{1}{n}$ para toda $x \in X$, así, por la propiedad arquimediana se sigue $||g_n f - f|| \leq |g_n f(x) - f(x)| < \frac{1}{n} < \frac{1}{n_0}$ para toda $n_0 < n$ y toda $x \in X$, es decir, $g_n f \longrightarrow f$ uniformente; además observe que $sop(fg_n) \subseteq sop(f) \cap sop(g_n) \subseteq sop(f)$, entonces $fg_n \in K(X)$ para toda n, es decir, $\{g_n f\}$ es una sucesión de funciones continuas de soporte compacto que convergen a f. Finalmente probemos que $sop(g_n f) \subseteq X - Sop(\mu)$, notemos que $Sop(\mu) \subseteq \{x \in X | f(x) = 0\} \subset E_n := \{x \in X | f(x) \leq \frac{1}{n+2}\} \subseteq \{x \in X | f(x) \leq \frac{1}{n+1}\} \subseteq \{x \in X |\ g_n f(x) = 0\}$, tomando complementos en está serie de contenciones y como E_n es cerrado, se sigue $sop(g_n f) \subseteq E_n^c \subseteq X - Sop(\mu)$, por lo tanto como $X - Sop(\mu)$ es el más grande abierto para el cual μ es cero, entonces $sop(g_n f) \subseteq X - Sop(\mu)$ implica $\mu(g_n f) = 0$ para toda n.
Por otro lado como μ es una medida de Radon, si $K = sop(f)$, entonces existe una constante $M_K > 0$ tal que $||\mu(h)|| \leq M_K ||h||$ para todo $h \in K(X)$ con $sop(h) \subseteq K$, así como $sop(g_n f - f) \subseteq sop(g_n f) \cup sop(-f) \subseteq sop(f)$, entonces $||g_n f - f|| \leq M_K ||g_n f - f||$ para toda n, por tanto se sigue $\mu(g_n f) \longrightarrow \mu(f)$, pues $g_n f \longrightarrow f$ converge uniformemente, es decir, $\lim_{n \longrightarrow \infty} \mu(g_n f) = \mu(f)$, pero como $\mu(g_n f) = 0$ para toda n, entonces $0 = \lim_{n \longrightarrow \infty} \mu(g_n f) = \mu(f)$. En resumen se tiene la afirmación para funciones positivas.

Sea f una función que cumple las hipótesis del teorema, esto es $f|_{Sop(\mu)} = 0$, note que $f|_{Sop(\mu)} = 0$ si y sólo si $|f| = 0$, así como $|f| - f \leq 0$ y $(|f| - f)|_{Sop(\mu)} = 0$, y aplicando lo anterior para funciones positivas, se sigue $\mu(|f| - f) = 0$, por tanto $\mu(f) = \mu(|f|) = 0$.
(3) se sigue de (2), suponiendo $f - g = 0$ sobre $Sop(\mu)$ y la linealidad de μ.
(4) Sea $f = f^+ - f^-$, se tiene que $f^+, f^- \in C^+(X)$, luego si se define

$$g(x) = \begin{cases} f^+(x) & \text{si } x \in Sop(\mu) \\ 0 & \text{si } x \in Sop(\mu)^c \end{cases} \text{ y } g(x) = \begin{cases} f^+(x) & \text{si } x \in Sop(\mu) \\ 0 & \text{si } x \in Sop(\mu)^c, \end{cases}$$

entonces $g = f^+$ y $h = f^-$ sobre $Sop(\mu)$, luego por el resultado anterior $\mu(g) = \mu(f^+)$ y $\mu(h) = \mu(f^-)$. Ahora, de $f \geq 0$ sobre $Sop(\mu)$ se sigue $g \geq h$ (sobre todo X), luego por ser μ positiva se tiene $\mu(g) \geq \mu(h)$ y por tanto $\mu(f^+) \geq \mu(f^-)$. De manera equivalente $\mu(f) = \mu(f^+) - \mu(f^-) \geq 0$. Los restantes incisos son probados de manera similar. ■
Sobre el espacio $\mu(X)$ se considera a la topología vaga, que es por definición la topología más gruesa sobre $\mu(X)$ para la cual todos los funcionales φ_f son continuos, donde los φ_f son definidos mediante

$$\varphi_f(\mu) := \mu(f), \text{ para todo } f \in K(X) \text{ y } \mu \in \mu(X).$$

Con lo anterior, una red $(\mu_i)_{i\in I}$ en $\mu(X)$ converge a una medida μ con respecto a la topología vaga, si y sólo si $\lim_{i\in I} \mu_i(f)) = \mu(f)$ para cada $f \in K(X)$. Si es el caso, decimos que la red $(\mu_i)_{i\in I}$ converge vagamente a μ. Un subconjunto $\mathcal{F}$ de $\mu(X)$ es vagamente acotado si los conjuntos $\{\mu(f) : \mu \in \mathcal{F}\}$ son acotados para cada $f \in K(X)$. Más aún, $\mathcal{F}$ es fuertemente acotado si para todo subconjunto compacto K de X, existe una constante $M_K \geq 0$ tal que $|\mu(f)| \leq M_K \|f\|_\infty$ para toda $f \in K(X)$ con $sop(f) \subset K$ y cada $\mu \in \mathcal{F}$.

Teorema 2 *En $\mu(X)$ con la topología vaga, el concepto de conjunto relativamente compacto es equivalente a los de conjunto acotado y conjunto fuertemente acotado.*

1.4 Teoremas de extensión

En esta sección presentaremos teoremas de extensión que necesitaremos más adelante. Para este fin recordemos que un espacio vectorial ordenado es un espacio vectorial E con un orden parcial $\leq$ que satisface las propiedades

$$x + z \leq y + z, \text{ para todo } z \in E, \text{ siempre que } x \leq y,$$

y

$$\lambda x \leq \lambda y, \text{ para todo } \lambda \geq 0, \text{ siempre que } x \leq y.$$

Un latice vectorial es un espacio vectorial ordenado para el cual siempre existe $\sup\{x, y\}$ para cualquier par x, $y \in E$. Una funcional lineal $\varphi : E \longrightarrow \mathbb{R}$ se dice que es positiva si $\varphi(x) \geq 0$ para cada $x \in E$ con $x \geq 0$.

La demostración del teorema siguiente puede ser encontrada en Choquet [6].

Teorema 3 *Sea E un espacio vectorial ordenado y F un subespacio de E tal que, para cada $x \in E$ existe un $y \in F$ que satisface $x \leq y$. Se tiene que dada una forma lineal $\varphi : F \longrightarrow \mathbb{R}$, entonces para cada $x \in E$ y $\alpha \in \mathbb{R}$ satisfaciendo*

$$\sup_{\substack{y\in F \\ y\leq x}} \varphi(y) \leq \alpha \leq \inf_{\substack{z\in F \\ x\leq z}} \varphi(z),$$

existe una forma lineal positiva $\widetilde{\varphi} : E \longrightarrow \mathbb{R}$ que extiende a φ y $\widetilde{\varphi}(x) = \alpha$.

De este resultado, se deduce el teorema clásico de Hahn -Banach. Para establecer el resultado mencionado recordemos que dado un espacio vectorial E, un mapeo $p : E \longrightarrow \mathbb{R}$ se dice que es sublineal, si

$$p(\lambda x) = \lambda p(x) \text{ para } x \in E \text{ y } \lambda \geq 0,$$

y

$$p(x + y) \leq p(x) + p(y) \text{ para cada } x, y \in E.$$

Por ejemplo, una seminorma $p : E \longrightarrow \mathbb{R}$ es un mapeo sublineal tal que $p(-x) = p(x)$ para toda $x \in E$.

Teorema 4 *Sea E un espacio vectorial real y $p : E \longrightarrow \mathbb{R}$ un mapeo sublineal. Si F es un subespacio de E y $\varphi : F \longrightarrow \mathbb{R}$ es una forma lineal que satisface $\varphi \leq p_{|F}$, entonces existe una forma lineal $\widetilde{\varphi} : E \longrightarrow \mathbb{R}$ que extiende a φ y se mantiene la desigualdad $\widetilde{\varphi} \leq p$.*

Mencionamos otros dos resultados que son consecuencia directa del Teorema de Hahn-Banach.

Corolario 1 *Dado un espacio real localmente convexo E y una seminorma continua $p : E \longrightarrow \mathbb{R}$, entonces para todo $x_0 \in E$ existe $\varphi \in E'$ (E' es el dual de E) tal que $\varphi(x_0) = p(x_0)$ y $\varphi \leq p$. Además, si E es Hausdorff, para todo $x_0 \in E$ existe $\varphi \in E'$ tal que $\varphi(x_0) \neq 0$.*

Corolario 2 *Sea E un espacio real normado y F un subespacio de E. Si $\varphi \in F'$, entonces existe $\widetilde{\varphi} \in E'$ tal que $\widetilde{\varphi}_{|F} = \varphi$ y $\|\widetilde{\varphi}\| = \|\varphi\|$.*

En particular, de este último corolario se deduce que, si E es un espacio normado y $x_0 \in E$, $x_0 \neq 0$, entonces siempre existe un $\varphi \in E'$ tal que $\varphi(x_0) = \|x_0\|$ y $\|\varphi\| = 1$.

Capítulo 2

Subespacios de Chebyshev

2.1 Mejor aproximación

En esta sección supondremos que X es un espacio de Banach con escalares reales o complejos, a menos que se indique otra cosa. Sea Y un subespacio vectorial de X. Si $f \in X$, el error de aproximación $E(f)$ de f por elementos de Y se define por $E(f) := E(f, Y)_X := \inf_{P \in Y} ||f - P||$.

En particular si $\{x_1, \ldots, x_n\} \subset X$ se tiene el subespacio generado $Y = gen\{x_1, \ldots, x_n\}$ y a la combinación lineal $y = \sum_{i=1}^{n} \lambda_i x_i$ se le llama polinomio de grado n. El error de aproximación (o error de grado n) E_n de f por polinomios de Y está dado por $E_n(f) = \inf_{P \in Y} ||f - P||$.

Si existe un polinomio $P_0 \in Y$ para el cual $E_n(f) = ||f - P_0||$, entonces decimos que P_0 es un elemento de mejor aproximación de f en Y. La función error E es una función que se comporta muy bien, pues es una función continua, más aún es sublineal.

Teorma 5 *Si X es un espacio, Y un subespacio de X y $E : X \to \mathbb{R}$ es la función error de aproximación por elementos de Y, entonces E es continua.*

Demostración. Sean $f, g \in X$ fijos y $p \in Y$, entonces por la desigualdad triangular y la definición de ínfimo se sigue, $E(f) \leq ||f - p|| \leq ||f - g|| + ||g - p||$ para toda $p \in Y$, luego $E(f) - ||f - g|| \leq ||g - p||$, aplicando ínfimo sobre toda $p \in Y$, se sigue $E(f) - ||f - g|| \leq E(g)$, esta última desigualdad es válida indistintamente del orden de f y g. Así, se tiene

$$|E(f) - E(g)| \leq ||f - g||. \tag{2.1}$$

Dado $\epsilon > 0$, tomemos $0 < \delta < \epsilon$, para obtener: $||f - f_0|| < \delta$ implica $|E(f) - E(f_0)| \leq ||f - f_0|| < \delta < \epsilon$. Por tanto E es continua. ■
Ahora nos centramos en analizar las propiedades de los elementos de mejor aproximación, y como se muestra en [7], pag.59, la existencia de la mejor aproximación está garantizada cuando Y es un subespacio de X de dimensión finita, es decir, $B(f) = \{P \in Y | E(f) = ||f - P||\} \neq \emptyset$, más aún este conjunto tiene propiedades interesantes como se mostrará más adelante.

Proposición 2.1 *Si Y es un subespacio de dimensión finita de X, entonces E es una función sublineal.*

Demostración. Probemos que E cumple

i) $E(f + g) \leq E(f) + E(g)$.

ii) $E(\alpha f) = |\alpha| E(f)$.

Sean $P_f \in B(f)$ y $P_g \in B(g)$ (pues $B(f), B(g) \neq \emptyset$), luego $E(f + g) \leq ||(f + g) - (P_f + P_g)|| \leq ||f - P_f|| + ||g - P_g|| = E(f) + E(g)$.
Por otro lado, sea $\alpha \in \mathbb{R}$. Si $\alpha \neq 0$, entonces $E(\alpha f) = ||\alpha(f - \frac{1}{\alpha} P(\alpha f))|| = |\alpha| ||f - \frac{1}{\alpha} P(\alpha f))|| \geq E(f)$ por tanto $E(\alpha f) \geq |\alpha| E(f)$. Inversamente $|\alpha| E(f) = |\alpha| ||f - P(f)|| = ||\alpha f - \alpha P(f))|| \geq E(\alpha f)$. Las desigualdades anteriores son válidas ya que toda combinación lineal de elementos de mejor aproximación están en Y y para el caso $\alpha = 0$ el resultado es obvío. Así se tiene la igualdad y por tanto E es sublineal. ■
Notemos por *ii*) de la proposición anterior que $E(-f) = E(f)$, es decir, el error E es simétrico con respecto a el inverso aditivo de f en $C(X)$.

Proposición 2.2 *El conjunto $B(f)$ es cerrado, acotado y convexo.*

Demostración. Sea $f \in X - Y$. Primero probemos que

$$E(f) \leq ||f - p|| \leq E(f) + d(p, B(f)) \text{ para todo } f, p \in X,$$

donde $d(p, B(f)) = \inf\limits_{q \in B(f)} ||q - p||$. Sea $p \in X$, por la desigualdad triangular se sigue que $||f - p|| \leq ||f - q|| + ||q - p||$, para todo $q \in Y$ y como $\inf\limits_{q \in Y} ||q - p|| \leq \inf\limits_{q \in B(f)} ||q - p||$ pues $B(f) \subset Y$, entonces se tienen las siguientes desigualdades

$$\begin{aligned} E(f) \leq ||f - p|| &\leq \inf_{q \in Y} \{||f - q|| + ||q - p||\} = \inf_{q \in Y} ||f - q|| + \inf_{q \in Y} ||q - p|| \\ &\leq E(f) + \inf_{q \in B(f)} ||q - p|| = E(f) + d(p, B(f)). \end{aligned}$$

Por tanto

$$E(f) \leq ||f-p|| \leq E(f) + d(p, B(f)) \text{ para cada } p \in Y. \tag{2.2}$$

Por otro lado, recordemos que $p \in \overline{B(f)}$ si y sólo si $d(p, B(f)) = 0$. Así, afirmamos que $\overline{B(f)} \subset B(f)$. Si $p \in \overline{B(f)}$, entonces $d(p, B(f)) = 0$ y por (2.2) se sigue que $||f-p|| = E(f)$, es decir, $p \in B(f)$ y por tanto $B(f)$ es cerrado. Más aún, de $E(f) + ||f|| = ||f-p|| + ||f|| \geq ||p||$ para todo $p \in B(f)$, se sigue que $B(f)$ es acotado. Finalmente, $B(f)$ es convexo. Sean α y β no negativos con $\alpha + \beta = 1$ y $p, q \in B(f)$, se deduce que, $||\alpha f + \beta f - (\alpha p + \beta q)|| \leq \alpha||f-p|| + \beta||f-q|| = \alpha E(f) + \beta E(f)$ y factorizando se tiene $||f - (\alpha p + \beta q)|| \leq E(f)$. Además como Y es un subespacio vectorial, entonces $\alpha p + \beta q \in Y$, así $E(f) = ||f - (\alpha p + \beta q)||$, donde $\alpha + \beta = 1$, es decir, $B(f)$ es convexo. ■

Si Y es de dimensión finita, entonces Y es topológicamente isomorfo a $\mathbb{R}^n$, es decir, con $\{x_1, \ldots, x_n\}$ y $\{e_1, \ldots, e_n\}$ bases, respectivamente de Y y de $\mathbb{R}^n$. La función $\phi(y) = \phi(\sum_{i=1}^{n} \lambda_i x_i) := \sum_{i=1}^{n} \lambda_i e_i$, es un isomorfismo de espacios vectoriales y un homeomorfismo de espacios topológicos, luego se sigue que $\phi(B(f))$ es cerrado y acotado en $\mathbb{R}^n$ y por tanto compacto. Por otro lado como la compacidad es un invariante topológico se sigue que $B(f)$ es compacto en Y y más aún es compacto en X, pues Y es cerrado en X. Así se tiene el siguiente resultado.

Corolario 3 *$B(f)$ es compacto y convexo para cada $f \in X$.*

Una observación importante es que ϕ preserva la convexidad por ser isomorfismo de espacios vectoriales (note que la convexidad no es invariante topológico).

En muchos casos prácticos es importante saber que existe un único polinomio de mejor aproximación. Decimos que un espacio es estríctamente convexo, si dado $f_1 \neq f_2$ donde $||f_1|| = ||f_2|| = 1$ y $\alpha_1, \alpha_2 > 0$ con $\alpha_1 + \alpha_2 = 1$, entonces $||\alpha_1 f_1 + \alpha_2 f_2|| < 1$. Está es una condición suficiente para la unicidad del mejor aproximante y como ejemplo de espacio estríctamente convexo se tiene L_p con $1 < p < \infty$. En estos espacios el polinomio de mejor aproximación es único.

Definición 2 *Si Y es un subespacio de X tal que para cada $f \in X$ existe un único polinomio de aproximación en Y, entonces se dice que Y es un conjunto de unicidad de X.*

Proposición 2.3 *Sea Y un subespacio de unicidad de dimensión finita de X. Si $P : X \longrightarrow Y$ es el operador de mejor aproximación definido por $f \longrightarrow P(f)$, entonces P es continua.*

Demostración. Por la unicidad de $P(f)$, entonces el operador P está bien definido. Para la continudad basta con demostrar que P es acotado. De la desigualdad (2.1) y haciendo $g = 0$ se sigue que $E(f) \leq ||f||$ para toda $f \in X$, así $||P(f)|| \leq ||f - P(f)|| + ||f|| = E(f) + ||f|| \leq 2||f||$, por tanto $||P|| \leq 2$. ■

En algunos espacios especiales existen los subespacios de unicidad, uno de ellos es cuando el espacio es de Hilbert, es decir, cuando el espacio tiene definido un producto interno. En estos espacios una caracterización del mejor aproximante es el siguiente.

Teorma 6 *Sea $(H, \langle\rangle)$ un espacio de Hilbert y sea H_0 un subespacio de H, un elemento $g \in H_0$ es de mejor aproximación para $f \in H$ si y sólo si cumple la siguiente condición de ortogonalidad.*

$$\langle f - g, h\rangle = 0 \text{ para toda } h \in H_0. \tag{2.3}$$

Demostración. Procedemos por contradicción. Supongamos que g es un elemento de mejor aproximación a f y que existe un $h \in H_0$ para el cual no se cumple (2.3), es decir, $\langle f - g, h\rangle \neq 0$ para algún $h \in H_0$.

Sin perdida de generalidad se puede escoger h tal que $||h|| = 1$ pues el vector normalizado también es no ortogonal, de hecho todos los vectores sobre la recta generada por h. Además se puede escoger a h o $-h$ según sea el caso para suponer $\langle f - g, h\rangle = \delta > 0$. Luego sea el vector $k = g + \delta h \in H$, así se tiene

$$||f - (g + \delta h)||^2 = ||f - g||^2 - \langle f - g, \delta h\rangle - \langle \delta h, f - g\rangle + ||\delta h||^2 \tag{2.4}$$
$$= ||f - g||^2 - \langle f - g, \delta h\rangle - \langle \delta h, f - g\rangle + \delta^2. \tag{2.5}$$

Como $\langle f - g, h\rangle = \delta$, entonces $\delta\langle f - g, h\rangle = \langle f - g, \delta h\rangle = \delta^2$; de manera semejante se tiene $\langle \delta h, f - g\rangle = \delta^2$, de lo cual se sigue que $||f - k||^2 = ||f - (g + \delta h)||^2 = ||f - g||^2 - \delta^2 - \delta^2 + \delta^2 = ||f - g||^2 - \delta^2 < ||f - g||^2$, por tanto $||f - k|| < ||f - g||$, pero esto último no puede ser pues g es de mejor aproximación, entonces $\langle f - g, h\rangle = 0$ para toda $h \in H_0$.

Inversamente, como H_0 es un subespacio de H, entonces $H_0 = H_0 + g$. Se definen los conjuntos $A = \{||f-(g+h)|| : h \in H_0\}$ y $B = \{||f-k|| : k \in H_0\}$, respectivamente. Afirmamos que $A = B$; si $||f-k|| \in A$ y como $k = g+h$ para alguna $h \in H_0$, entonces $||f-k|| = ||f-(g+h)|| \in B$, luego $A \subset B$. La otra contención es clara, pues $g+h \in H_0$, por tanto se sigue que

$$E(f) = \inf_{k \in H_0} ||f-k|| = \inf_{b \in B} b = \inf_{a \in A} a = \inf_{h \in H_0} ||f-(g+h)|| \geq ||f-g||.$$

Utilizando la ecuación (2.4) con $\delta = 1$ y la hipótesis $\langle f-g, h\rangle = 0$, $E(f) = ||f-g||$. ■ Un producto interior define un funcional lineal real o complejo según sea el producto interior, así la ecuación (2.3) se puede interpretar como sigue: para el funcional $\lambda(h) := \langle f-g, h\rangle$ se tiene $\lambda(h) = 0$ para toda $h \in H_0$, luego podemos tomar lo anterior como motivación para la siguiente definición en espacios de Banach.

Sea X un espacio y Y un subespacio vectorial de X, decimos que un funcional lineal real o complejo λ es ortogonal a Y si $\lambda(h) = 0$ para toda $h \in Y$ y denotamos por $Y^{\perp}$ al conjunto de todos los funcionales lineales ortogonales a Y.

Teorma 7 *Sea Y un subespacio cerrado de X y $f \in X-Y$. Entonces $f_0 \in Y$ es elemento de mejor aproximación a f en Y si y sólo si existe un funcional lineal acotado λ tal que $\lambda \in Y^{\perp}$, $||\lambda|| = 1$ y $||f-f_0|| = \lambda(f)$. Más aún*

$$E(f) = \sup_{\substack{\lambda \in Y^{\perp} \\ ||\lambda||=1}} \lambda(f). \tag{2.6}$$

Demostración. Sea $f \in X-Y$ y f_0 el mejor aproximante a f en Y, es decir, $||f-f_0|| = E(f)$. Definimos el funcional lineal $\lambda : H \longrightarrow \mathbb{R}$ por

$$\lambda(y+mf) = mE(f), \ \textit{donde } H = Y \oplus gen\{f\}. \tag{2.7}$$

Notemos que H es un subespacio vectorial de X y λ está bien definido, más aún es lineal ya que si $h = y+mf$, $h' = y' + m'f$ y $\alpha \in \mathbb{R}$, entonces $\lambda(h+\alpha h') = \lambda((y+\alpha y') + (m+\alpha m')f) = (m+\alpha m')E(f) = mE(f) + \alpha m' E(f) = \lambda(y+mf) + \alpha\lambda(y'+m') = \lambda(h) + \alpha\lambda(h')$. Con lo anterior demostraremos las propiedades para el λ del teorema. Primero notemos que $h \in Y$ si y sólo si $m = 0$, entonces $\lambda(h) = 0$ para toda $h \in Y$, es decir, $\lambda \in Y^{\perp}$. Además se tiene que

$$|\lambda(y+mf)| = |m|E(f) \leq ||mf+y|| \text{ para toda } y \in Y \text{ y } m \in \mathbb{R}.$$

Por tanto $||\lambda|| = 1$ sobre H y $\lambda(f) = \lambda(f-y) = E(f) = ||f - f_0||$. Por otro lado notemos que λ es un funcional lineal sobre H el cual está acotado por un funcional sublineal E continuo sobre H (por la proposición (2.1)), más aún λ coincide con E sobre H, por el teorema de Hahn-Banach (ver [1], pag.48) existe una extensión λ a todo X con $||\lambda|| = 1$ sobre X, además λ mantiene la ortogonalidad sobre Y.
Inversamente, como $\lambda(f-y) \leq ||\lambda||\;||f-y||$ y por hipótesis $||\lambda|| = 1$, $\lambda \in Y^{\perp}$ y $||f - f_0|| = \lambda(f)$, entonces $\lambda(f) = \lambda(f) - \lambda(y) = \lambda(f-y) \leq ||f-y||$ para todo $y \in Y$, así se sigue que

$$||f - f_0|| = \lambda(f) \leq \inf_{y\in Y} ||f-y|| = E(f) \text{ para cada } f \in X. \tag{2.8}$$

Por tanto $||f - f_0|| = E(f)$, es decir, f_0 es el mejor aproximante de f en Y. Más aún, de la desigualdad (2.8), si fijamos f, entonces para toda $\lambda \in Y^{\perp}$ y $||\lambda|| = 1$, $E(f)$ es una cota superior, luego $\sup\limits_{\substack{\lambda\in Y^{\perp}\\ ||\lambda||=1}} \lambda(f) \leq E(f)$.
Finalmente, como f_0 es el mejor aproximante de f sobre Y y por (2.7) existe un funcional ortogonal sobre Y tal que $||\lambda|| = 1$, se sigue la desigualdad contraria

$$E(f) = ||f - f_0|| = \lambda(f) \leq \sup_{\substack{\lambda\in Y^{\perp}\\ ||\lambda||=1}} \lambda(f).$$

■

2.2 Caracterización del mejor aproximante

Ahora nuestro interés es la caracterización del mejor aproximante en $C(X)$ donde X es un espacio compacto Hausdorff, pues estos espacios en la práctica son unos de los más usados por las propiedades que tienen. En el desarrollo de este capítulo se mostrará la fuerza del siguiente teorema ya que al final se caracterizará el polinomio de mejor aproximación de la manera más simple con el teorema de alternancia de Chebyshev.

Teorma 8 (***Kolmogorov***) *Si Y es un subespacio de dimención finita de $C(X)$, entonces P es un elemento de mejor aproximación de $f \in C(X)$ sobre Y si y sólo si para cada $Q \in Y$ se tiene*

$$\max_{x\in A_0} \Re e\{[f(x) - P(x)]\overline{Q(x)}\} \geq 0, \tag{2.9}$$

donde A_0 denota el conjunto (el cual depende de f y P) de todos los $x \in X$ para los cuales $|f(x) - P(x)| = ||f - P||$.

Demostración. Procedemos por contradicción, en efecto, supongamos que existe $Q \in Y$ tal que $\max\limits_{x \in A_0} \Re e\{[f(x) - P(x)]\overline{Q(x)}\} = -2\epsilon < 0$, para algún $\epsilon > 0$. Por la continuidad de la parte real $\Re e$,

$$|\Re e\{[f - P]\overline{Q}(x)\} - \Re e\{[f - P]\overline{Q}(x_0)\}| < \epsilon,$$

para toda $x \in B_\delta(x_0)$, para algún δ y $x_0 \in A_0$. Por lo tanto,

$\Re e\{[f - P]\overline{Q}(x)\} < \epsilon + \Re e\{[f - P]\overline{Q}(x_0)\}$ para toda $x \in B_\delta(x_0)$. Además como $\Re e\{[f - P]\overline{Q}(x_0)\} \leq \max\limits_{x \in A_0} \Re e\{[f(x) - P(x)]\overline{Q(x)}\} = -2\epsilon$, entonces $\Re e\{[f(x) - P(x)]\overline{Q(x)}\} < \epsilon - 2\epsilon = -\epsilon$ para toda $x \in B_\delta(x_0)$. Tomando ahora el conjunto abierto $G := \bigcup\limits_{x \in A_0} B_\delta(x)$, se tiene que G contiene a A_0, (notemos que $\delta(\epsilon, x)$ y $x \in A_0$), luego

$$\Re e\{[f(x) - P(x)]\overline{Q(x)}\} < -\epsilon \text{ para todo } x \in G \supseteq A_0. \tag{2.10}$$

Por otro lado, supongamos la función $P_1 = P - \lambda Q$ donde $\lambda > 0$ y $M = \max |Q(x)|$, entonces para cada $x \in G$ se tiene que

$$\begin{aligned}|f(x) - P_1(x)|^2 = |f(x) - P(x) + \lambda Q(x)|^2 = |f(x) - P(x)|^2 + \\ 2\lambda \Re e[(f - P)(x)\overline{Q(x)}] + \lambda^2 |Q(x)|^2 \leq E(f)^2 - 2\lambda\epsilon + \lambda^2 M^2.\end{aligned}$$

Ahora tomando $\lambda < M^{-2}\epsilon$, se sigue que $\lambda^2 M^2 < \lambda\epsilon$ y se tiene la desigualdad $|f(x) - P_1(x)|^2 < E(f)^2 - \lambda\epsilon < E(f)^2$, es decir,

$$|f(x) - P_1(x)| < E(f) \text{ para toda } x \in G. \tag{2.11}$$

Por otro lado

$$|f(x) - P(x)| \leq ||f - P|| = E(f) = \inf_{P \in Y} ||f - P|| \text{ para todo } x \in X,$$

así se obtiene la desigualdad estricta $|f(x) - P(x)| < E(f)$ para todo $x \in G^c$. Sea $\delta > 0$ tal que $|f(x) - P(x)| < E(f) - \delta$ para todo $x \in G^c$, entonces se tiene que, $|f(x) - P_1(x)| \leq |f(x) - P(x)| + \lambda |Q(x)| \leq E(f) - \delta + \lambda M$ y tomando λ de tal manera que $\lambda < (2M)^{-1}\delta$ se deduce que $|f(x) - P_1(x)| < E(f) - \delta + \frac{1}{2}\delta < E(f)$ para todo $x \in G^c$. De está última desigualdad y de (2.11) se sigue

$$|f(x) - P_1(x)| < E(f) \text{ para todo } x \in X \tag{2.12}$$

Por lo anterior, como $P_1 \in Y$ se obtiene que $||f - P_1|| < E(f)$, es decir, existe un menor error al que se tiene inicialmente, es una contradicción y por tanto.

$$\max_{x \in A_0} \Re e\{[f(x) - P(x)]\overline{Q(x)}\} \geq 0, \text{para todo } Q \in Y.$$

Inversamente, sea $P_1 \in Y$. Si $P - P_1 := Q$ y $x_0 \in X$, entonces $|f(x_0) - P_1(x_0)|^2 = |f(x_0) - P(x_0)|^2 + 2\Re e\{[f(x_0) - P(x_0)]\overline{Q(x_0)}\} + |Q(x_0)|^2$, donde $\Re e\{[f(x_0 - P(x_0)]\overline{Q(x_0)}\} \geq 0$. Luego $|f(x_0) - P_1(x_0)| \geq |f(x_0) - P(x_0)| = ||f - P||$, por tanto P_1 no es de mejor aproximación, pues si lo fuera se tendría $|f(x) - P_1(x)| \leq ||f - P_1|| = E(f)$ para todo $x \in X$, así P es el mejor aproximante. ■

En adelante A_0 siempre denotará el conjunto utilizado en la hipótesis del teorema anterior. Como consecuencia del Teorema de Kolmogorov y considerando el espacio de funciones continuas reales $C(X)$ se tiene.

Corolario 4 *Si Y es un subespacio de dimensión finita de $C(X)$ y $P \in Y$, entonces P es un elemento de mejor aproximación de $f \in C(X)$ si y sólo si para cada $Q \in Y$ se tiene*

$$\max_{x \in A_0} \{[f(x) - P(x)]Q(x)\} \geq 0.$$

Los teorema anteriores son caracterizaciones del elemento de mejor aproximación, utilizando funcionales lineales ortogonales. El Teorema Rivlin-Shapiro que demostraremos utiliza el Teorema de Kolmogorov, así como del Teorema de Caratheodory, se puede ver en ([1], pag.46). Además para el teorema de Hanh-Banach en su versión geométrica que también se utiliza en el desarrollo de la prueba del Teorema Rivlin-Shapiro, se puede ver en ([8],pag.204).

Teorma 9 *Sean G y G_0 conjuntos disjuntos convexos en un espacio lineal normado X. Si G_0 es abierto, entonces existe un funcional lineal λ en X y un $\gamma \in \mathbb{R}$ que cumplen*

$$\lambda(g) < \gamma \leq \lambda(f) \text{ para toda } f \in G,\ g \in G_0.$$

Teorma 10 (***Caratheodory***) *Si $B \subset \mathbb{R}^n$ y $r \leq n$, entonces el casco convexo conv B consta de todos los elementos de la forma*

$$x = \sum_{i=0}^{r} p_i x_i, \text{ con } x_i \in B, \; p_i \geq 0 \text{ y } \sum_{i=0}^{r} p_i = 1 \text{ para } r \in \{1, 2, \ldots, n\}.$$

Teorma 11 (***Rivlin-Shapiro***) *Sea Y un subespacio lineal de $C(X)$ de dimensión n. Una función $P \in Y$ es de mejor aproximación a $f \in C(X)$ si y sólo si existen puntos $x_i \in A_0$, $i = 0, \ldots, r$ y números $p_i \geq 0$ tal que $\sum_{i=0}^{r} p_i = 1$ y $\sum_{i=0}^{r} p_i[f(x_i) - P(x_i)]Q(x_i) = 0$, para todo $Q \in Y$ con $r \leq n$.*

Demostración. Note que todo funcional lineal $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}$, está definido en $\mathbb{R}^n$ si y sólo si está definido sólo en la base canónica de $\mathbb{R}^n$, es decir, λ es funcional lineal en $\mathbb{R}^n$ si y sólo si $\lambda(\mathbf{x}) = \sum_{i=1}^{n} a_i x_i$, donde $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ y $\mathbf{a} = (\lambda(e_1), \ldots, \lambda(e_n))$. Si $\{\phi_1, \ldots, \phi_n\}$ es una base para $Y \subset C(X)$, entonces para toda $Q \in Y$, se tiene $Q = \sum_{i=1}^{n} c_i \phi_i$, con $c_i \in \mathbb{R}$, para $i = 1, \ldots, n$. Por otro lado sea la función continua $\Phi = ((f-P)\phi_1, \ldots, (f-P)\phi_n) : X \longrightarrow \mathbb{R}^n$ definida por $\Phi(x) = ((f-P)(x)\phi_1(x), \ldots, (f-P)(x)\phi_n(x))$. Como A_0 es un subconjunto cerrado (A_0 es la imagen inversa del escalar $||f-p||$ bajo la función continua $|f-p|$) en un espacio X compacto y Hausdorff, entonces A_0 es compacto, y por tanto $\Phi(A_0)$ es un subespacio compacto de $\mathbb{R}^n$. Más aún, el casco convexo *cov* $\Phi(A_0)$ es compacto.
Por otro lado afirmamos:

i) P es de mejor aproximación a f en Y si y sólo si $\max\limits_{x \in \Phi(A_0)} \lambda(x) \geq 0$ para toda λ funcional lineal.

ii) Si $\max\limits_{x \in \Phi(A_0)} \lambda(x) \geq 0$ para toda λ funcional lineal, entonces $0 \in$ *cov* $\Phi(A_0)$.

Probemos i), supongamos que P es de mejor aproximación a f y sea λ un funcional lineal. Luego $(\lambda\Phi)(x) = \lambda(\Phi(x)) = \sum_{i=0}^{n} \lambda(e_i)(f-P)(x)\phi_i(x) = (f-P)(x)(\sum_{i=0}^{n} \lambda(e_i)\phi_i(x))$, es decir, $\lambda\Phi(x) = (f-P)(x)Q(x)$, donde $Q(x) = \sum_{i=0}^{n} \lambda(e_i)\phi_i(x) \in Y$. Por el teorema de Kolmogorov, se sigue que

$$0 \leq \max_{x \in A_0}(f-P)(x)Q(x) = \max_{x \in A_0} \lambda\Phi(x) = \max_{x \in \Phi(A_0)} \lambda(x). \tag{2.13}$$

Inversamente, supongamos que $0 \leq \max_{x \in \Phi(A_0)} \lambda(x)$. Sea $Q(x) = \sum_{i=0}^{n} c_i \phi_i(x) \in Y$ y el funcional lineal definido por $\lambda(\mathbf{x}) = \sum_{i=0}^{n} c_i x_i$, por (2.13) y el teorema de Kolmogorov se tiene el resultado. Para la prueba de ii), procedemos por contradicción. Supongamos que $0 \notin cov\ \Phi(A_0)$, como $cov\ \Phi(A_0)$ es compacto, éste es cerrado en $\mathbb{R}^n$, sea la bola abierta $B_\delta(0)$ en el complemento de $cov\ \Phi(A_0)$, si hacemos $G := cov\ \Phi(A_0)$ y $G_0 = B_\delta(0)$, entonces se cumple la hipótesis del teorema (9) y por tanto existe un funcional lineal λ en $\mathbb{R}^n$ y un $\gamma \in \mathbb{R}$ que cumplen

$$\lambda(g) < \gamma \leq \lambda(f) \textit{ para toda } f \in G,\ g \in G_0.$$

En particular, $0 = \lambda(0) < \gamma \leq \lambda(f)$ *para toda* $f \in \Phi(A_0)$, es decir, $-\lambda(f) < 0$ para toda $f \in \Phi(A_0) \subseteq cov\ \Phi(A_0)$, luego $\max_{f \in \Phi(A_0)} -\lambda(f) \leq 0$, lo cual contradice la hipótesis.

Finalmente, por i), ii) y aplicando el Teorema de Caratheodory, se sigue que $\sum_{i=0}^{r} p_i y_i = 0$, donde $y_i = \Phi(x_i)$ para algún $x_i \in A_0$ y los $p_i \geq 0$ son tales que $\sum_{i=0}^{r} p_i = 1$, con $r \leq n$. Ahora sea $Q = \sum_{j=0}^{r} c_j \phi_j \in Y$ y el correspondiente funcional lineal λ, donde $\lambda(e_j) = c_j$, $j = 1, \ldots, n$, entonces $0 = \lambda(0) = \sum_{i=0}^{r} p_i \lambda(y_i) = \sum_{i=0}^{r} p_i (\sum_{j=0}^{n} \lambda(e_j)(f-P)(x_i)\phi_j(x_i)) =$
$\sum_{i=0}^{r} p_i (\sum_{j=0}^{n} (f-P)(x_i) c_j \phi(x_i)) = \sum_{i=0}^{r} p_i (f-P)(x_i)(\sum_{j=0}^{n} c_j \phi(x_i)) =$
$\sum_{i=0}^{r} p_i (f-P)(x_i) Q(x_i)$, por tanto, existen puntos $x_i \in A_0$, $i = 0, \ldots, r$ y números $p_i \geq 0$ tal que $\sum_{i=0}^{r} p_i = 1$ y $\sum_{i=0}^{r} p_i [f(x_i) - p(x_i)] Q(x_i) = 0$ para todo $Q \in Y$ con $r \leq n$.

Para el recíproco del teorema, supongamos que $\sum_{i=0}^{r} p_i [f(x_i) - p(x_i)] Q(x_i) = 0$. Así $[f(x_i) - p(x_i)] Q(x_i) \geq 0$, para algún $x_i \in A_0$ con $0 \leq i \leq r$, pues $p_i \geq 0$, y por tanto

$\max_{x \in A_0} [f(x) - p(x)] Q(x) \geq [f(x_i) - p(x_i)] Q(x_i) \geq 0$, para todo $Q \in Y$.

Finalmente, por el teorema de Kolmogorov se sigue el resultado. ■

Corolario 5 *Si P es de mejor aproximación a f en X, entonces P es de mejor aproximación a f en algún subconjunto $\{x_0, \ldots, x_r\} \subseteq A_0$.*

Como motivación del teorema Rivlin-Shapiro surge la siguiente definición que es importante porque permite caracterizar de otra forma a los elementos de mejor aproximación en términos de funcionales lineales ortogonales como se vera más adelante.

Definición 3 *Sea Y un subespacio lineal de $C(X)$ de dimensión finita n, λ^* se llama funcional lineal de aproximación de f y P en Y, si dados $\{x_0, \ldots, x_r\} \subseteq A_0$ y $r \leq n$, se tiene $\lambda^*(g) = \sum_{i=0}^{r} \alpha_i g(x_i)$,*

donde $\alpha_i = \; p_i sign\ [f(x_i) - P(x_i)],\ p_i \geq 0,\ i = 0, \ldots, r,\ \sum_{i=0}^{r} p_i = 1.$

Corolario 6 *Sea Y un subespacio lineal de $C(X)$ de dimensión finita n. Si $f \notin Y$, entonces P es de mejor aproximación a f en Y si y sólo si existe un funcional lineal de aproximación de f y P, para algunos puntos $\{x_0, \ldots, x_r\} \subseteq A_0$ y $r \leq n$, el cual es ortogonal en Y.*

Demostración. Supongamos que P es de mejor aproximación a f en Y. Por el teorema de **Rivlin-Shapiro**, existen puntos $x_i \in A_0$, $i = 0, \ldots, r$ y números $p_i \geq 0$ ta que $\sum_{i=0}^{r} p_i = 1$ y $\sum_{i=0}^{r} p_i[f(x_i) - P(x_i)]Q(x_i) = 0$, para todo $Q \in Y$ con $r \leq n$, luego se define el funcional lineal ortogonal de f y P en Y por $\lambda^*(g) = \sum_{i=0}^{r} \alpha_i g(x_i)$ donde $\alpha_i = p_i sign\ [f(x_i) - P(x_i)],\ p_i \geq 0$ y $\sum_{i=0}^{r} p_i = 1$. Como $||f - P||\lambda^*(Q) = \lambda^*(||f - P||\ Q) = \sum_{i=0}^{r} \alpha_i ||f - P|| Q(x_i) = \sum_{i=0}^{r} p_i[f(x_i) - P(x_i)]Q(x_i) = 0$, se sigue que $\lambda^*(Q) = 0$ para todo $Q \in Y$, es decir, λ^* es ortogonal en Y. Inversamente, por el teorema (7), basta probar que el funcional lineal ortogonal λ^* cumple $||\lambda^*|| = 1$ y $||f - P|| = \lambda^*(f)$. Como $|\lambda^*(g)| = |\sum_{i=0}^{r} \alpha_i g(x_i)| \leq \sum_{i=0}^{r} p_i |g(x_i)| \leq \sum_{i=0}^{r} p_i ||g|| = ||g||$, luego $||\lambda^*|| = 1$. Por otro lado, como λ^* es ortogonal en Y se sigue que $\lambda^*(f) = \lambda^*(f - P) = \sum_{i=0}^{r} p_i sign\ [f(x_i) - P(x_i)][f(x_i) - P(x_i)] = \sum_{i=0}^{r} p_i (sign\ [f(x_i) - P(x_i)])^2 ||f - P|| = ||f - P||$. ∎

2.3 Sistemas de Haar

En la sección anterior todos los teoremas son caracterizaciones del mejor aproximante, finalizando con el teorema y corolario que caracteriza de la manera más sencilla tal polinomio. Ahora veremos como garantizar la unicidad del polinomio de mejor aproximación. Para ello definimos los subespacios de Haar y en particular los subespacios de Chebyshev en espacios de funciones definidas sobre espacios topológicos de Hausdorff.

Definición 12 *Sea X un espacio Hausdorff con al menos $n+1$ puntos y sean $h_0, h_1, \ldots, h_n$ funciones reales continuas definidas sobre X. Decimos que el conjunto de funciones $\Phi := \{h_0, h_1, \ldots, h_n\}$ es un sistema de Haar de orden $n+1$ en X, si cada combinación lineal $\lambda_0 h_0 + \lambda_1 h_1 + \cdots + \lambda_n h_n$ con coeficientes $\lambda_0, \lambda_1, \ldots, \lambda_n \in \mathbb{R}$ no simultáneamente todos ceros, no tienen más de n ceros distintos en X. El subespacio generado $H = gen\{h_0, \ldots, h_n\}$ de X por el sistema de Haar de orden $n+1$ en X es llamado subespacio de Haar de orden $n+1$ en $C(X)$. En particular si X es compacto Hausdorff decimos que Φ es un Sistema de Chebyshev, respectivamente H es llamado subespacio de Chebyshev.*

Si X tiene cardinalidad $n+1$ y Φ es un sistema de Haar, entonces Φ forma un conjunto linealmente independiente en $C(X)$. El conjunto $\{1, x, \ldots, x^n\}$ en $[a, b]$ es un sistema de Chebyshev.

Proposición 2.4 *Sean X un espacio Hausdorff de cardinalidad $n+1$ y $\{h_0, h_1, \ldots, h_n\} \subseteq C(X)$. $\Phi = \{h_0, h_1, \ldots, h_n\}$ forma un sistema de Haar de orden $n+1$ en $C(X)$ si y sólo si $\det(h_i(x_j)) \neq 0$ para cualesquiera puntos distintos $\{x_0, \ldots, x_n\} \in X$.*

Demostración. Sean $x_0, \ldots, x_n$ puntos distintos de X. Como $\sum_{i=0}^{n} \lambda_i h_i = 0$ implica $\lambda_i = 0$ con $i = 0, \ldots, n$, ya que $\{h_0, h_1, \ldots, h_n\}$ es un conjunto linealmente independientes, entonces evaluando la combinación lineal en cada x_i para $i = 0, \ldots, n$ se tiene el sistema de ecuaciones lineales homogéneo de orden $n+1$:

$$\begin{aligned} \lambda_0 h_0(x_0) + \lambda_1 h_1(x_0) + \cdots + \lambda_n h_n(x_0) &= 0 \\ \lambda_0 h_0(x_1) + \lambda_1 h_1(x_1) + \cdots + \lambda_n h_n(x_1) &= 0 \\ &\vdots \\ \lambda_0 h_0(x_n) + \lambda_1 h_1(x_n) + \cdots + \lambda_n h_n(x_n) &= 0 \end{aligned} \tag{2.14}$$

que tiene la solución trivial $\lambda_0 = \lambda_1 = \cdots = \lambda_n = 0$ si y sólo si $\det(h_i(x_j)) \neq 0$. Inversamente, supongamos que $\det(h_i(x_j)) = 0$, entonces existe una solución no trivial del sistema de ecuaciones lineales (2.14), luego la combinación lineal $\sum_{i=0}^{n} \lambda_i h_i$ tiene al menos $n+1$ ceros, por tanto Φ no es un sistema de Haar de orden $n+1$, así el regreso está demostrado por contra recíproca. ■

El teorema anterior es importante porque nos permite encontrar una única función en el subespacio de Haar, cuya gráfica pase por algunos puntos determinados, es decir, se tiene el siguiente corolario.

Corolario 7 *Sean $x_0, \ldots, x_n \in X$ y $\alpha_0, \ldots, \alpha_n \in \mathbb{R}$. Si Φ forma un sistema de Haar en X, entonces existe una única combinación lineal $h = \sum_{i=0}^{n} \lambda_i h_i$ tal que $h(x_i) = \alpha_i$ con $i = 1, \ldots, n$.*

Demostración. Por el teorema anterior se tiene $\det(h_i(x_j)) \neq 0$ y por tanto el sistema

$$\begin{aligned}
\lambda_0 h_0(x_0) + \lambda_1 h_1(x_0) + \cdots + \lambda_n h_n(x_0) &= \alpha_1 \\
\lambda_0 h_0(x_1) + \lambda_1 h_1(x_1) + \cdots + \lambda_n h_n(x_1) &= \alpha_2 \\
&\vdots \\
\lambda_0 h_0(x_n) + \lambda_1 h_1(x_n) + \cdots + \lambda_n h_n(x_n) &= \alpha_n.
\end{aligned}$$

tiene solución única, $\lambda_0, \ldots, \lambda_n$, así se tiene que $h = \sum_{i=0}^{n} \lambda_i h_i$ tal que $h(x_i) = \alpha_i$. ■

De está forma decimos que hay un polinomio de interpolación h de $h_0, \ldots, h_n$ con $h(x_i) = \alpha_i$ para $i = 0, \ldots, n$.

Proposición 2.5 *Si Φ es un sistema de Chebyshev de orden $n+1$ en $[a, b]$ y $x_1, \ldots, x_n \in (a, b)$, entonces la función definida por*

$$D(x, x_1, \ldots, x_n) := \begin{vmatrix} h_0(x) & h_0(x_1) & \ldots & h_0(x_n) \\ h_1(x) & h_1(x_1) & \ldots & h_1(x_n) \\ \vdots & \vdots & & \vdots \\ h_n(x) & h_n(x_1) & \ldots & h_n(x_n) \end{vmatrix} \tag{2.15}$$

tiene exactamente por ceros los puntos x_i, para $i = 1, \ldots, n$. Más aún, en cada x_i cambia de signo $D(x, x_1, \ldots, x_n)$.

Demostración. La primera parte es clara. Por las propiedades de los determinantes se sigue que $D(x_i, x_1, \ldots, x_n) = 0$ para $i = 1, \ldots, n$. Más aún, como $D(x, x_1, \ldots, x_n)$ es una combinación lineal del sistema de Chebyshev Φ, entonces $D(x, x_1, \ldots, x_n)$ tiene a lo más n ceros, y por tanto tiene exactamente n ceros.
Sea $0 < \delta$ lo suficientemente pequeño tal que $(x_1 - \delta, x_1 + \delta) \subset [a, b]$. Definimos la función continua $g_\delta : [0, \delta] \longrightarrow \mathbb{R}$ por

$$g_\delta(t) := \begin{vmatrix} h_0(x_1 - \delta + t) & h_0(x_1 + t) & \ldots & h_0(x_n) \\ h_1(x_1 - \delta + t) & h_1(x_1 + t) & \ldots & h_1(x_n) \\ \vdots & \vdots & & \vdots \\ h_n(x_1 - \delta + t) & h_n(x_1 + t) & \ldots & h_n(x_n) \end{vmatrix}$$

Afirmamos que $g_\delta(t) \neq 0$ para toda $t \in [0, \delta]$. En efecto, supongamos que existe $t_0 \in [0, \delta]$ tal que $g_\delta(t_0) = 0$, entonces $\lim\limits_{\delta \to t_0} g_\delta(t_0) = 0 = -D(x_1 + t_0, x_1, \ldots, x_n)$ contradiciendo la primera parte de la proposición, así $g_\delta(t) \neq 0$ para toda $t \in [0, \delta]$. Finalmente observemos que $g(0) = D(x_1 - \delta, x_1, \ldots, x_n)$ y $g(\delta) = -D(x_1 + \delta, x_1, \ldots, x_n)$, luego $D(x_1 - \delta, x_1, \ldots, x_n)$ y $D(x_1 + \delta, x_1, \ldots, x_n)$ tienen signos opuestos, pués $g \neq 0$ en todo el intervalo $[0, \delta]$, de está misma forma se prueba que también cambia de signo $D(x, x_1, \ldots, x_n)$, para $x_2, \ldots, x_n$.
■
Se tiene un resultado más general de la proposición anterior, de hecho la función $D(x, x_1, \ldots, x_n)$ genera a todas las funciones en un subespacio de Chebyshev que cambian de signo exactamente en n puntos de un intervalo cerrado $[a, b]$.

Proposición 2.6 *Sean Φ un sistema de Chebyshev de orden $n+1$ en $[a, b]$ y H el correspondiente subespacio de Chebyshev de orden $n+1$. Si $h \in H$ tiene exactamente n ceros distintos $x_1, \ldots, x_n$ en (a, b), entonces h necesariamente cambia de signo en cada x_i, para $i = 1, \ldots, n$.*

Demostración. $h \in H$ y $x_1, \ldots, x_n$ son los ceros de h, entonces $h = \sum\limits_{i=0}^{n} \lambda_i h_i$, pues H es un subespacio diferente del trivial y por tanto existe algún $\lambda_i \neq 0$. Sin pérdida de generalidad supongamos que $\lambda_0 \neq 0$, así usando la propiedad de los determinantes, multiplicando a los λ_i diferentes de cero con el renglón i-ésimo correspondiente y sumando al primer renglón en $D(x, x_1, \ldots, x_n)$, se sigue que

$$D(x,x_1,\ldots,x_n) = 1/\lambda_0 \begin{vmatrix} \sum\limits_{i=0}^{n} \lambda_i h_i & 0 & \ldots & 0 \\ h_1(x) & h_1(x_1) & \ldots & h_1(x_n) \\ \vdots & \vdots & & \vdots \\ h_n(x) & h_n(x_1) & \ldots & h_n(x_n) \end{vmatrix} = (D_1/\lambda_0)h(x),$$

donde $$D_1 = \begin{vmatrix} h_1(x_1) & h_1(x_2) & \ldots & h_1(x_n) \\ h_2(x_1) & h_2(x_2) & \ldots & h_2(x_n) \\ \vdots & \vdots & & \vdots \\ h_n(x_1) & h_n(x_2) & \ldots & h_n(x_n) \end{vmatrix}.$$

Entonces h tiene los mismos ceros que $D(x,x_1,\ldots,x_n)$, luego por la proposición (2.5), h cambia de signo en cada x_i si y sólo si lo hace $D(x,x_1,\ldots,x_n)$ en cada x_i. ■

Lema 1 *Sean X un espacio de Hausdorff que tiene al menos $n+2$ puntos, Φ un sistema de Haar de orden $n+1$ y P un polinomio de mejor aproximación a la función continua f, entonces el subconjunto A_0 para el cual $|f(x)-P(x)| = ||f-P|| = E(f)$ contiene al menos $n+2$ puntos.*

Demostración. Supongamos que $A_0 = \{x_0,\ldots,x_s\}$ con $s \leq n+1$, entonces existe un punto $x \in X$, tal que

$$0 \leq |f(x)-P(x)| < ||f-P|| = E(f),$$

luego entonces $0 < E(f)$. Por otro lado, por el sistema de ecuaciones lineales de $(s+1)\times(n+1)$

$$\begin{aligned} \lambda_0\phi_0(x_0) + \lambda_1\phi_1(x_0) + \cdots + \lambda_n\phi_n(x_0) &= -[f(x_0) - P(x_0)] \\ \lambda_0\phi_0(x_1) + \lambda_1\phi_1(x_1) + \cdots + \lambda_n\phi_n(x_1) &= -[f(x_1) - P(x_1)] \\ &\vdots \\ \lambda_0\phi_0(x_s) + \lambda_1\phi_1(x_s) + \cdots + \lambda_n\phi_n(x_s) &= -[f(x_s) - P(x_s)] \end{aligned}$$

se tiene un polinomio $Q = \sum\limits_{i=0}^{n} \lambda_i\phi_i$ (no necesariamente único) tal que $Q(x_k) = -[f(x_k) - P(x_k)]$ con $k = 0,\ldots,s$. Aplicando el teorema de Kolmogorov se

tiene $\max\limits_{x\in A_0}[f(x)-P(x)]Q(x) = \max\limits_{0\leq k\leq s}[f(x_k)-P(x_k)]Q(x_k) = \max\limits_{0\leq k\leq s} -[f(x_k)-P(x_k)]^2 = -E^2(f) < 0$, pero esto no puede ser ya que contradice la caracterización del polinomio de mejor aproximación. ■
De forma inmediata se prueba la unicidad del polinomio de mejor aproximación usando el teorema anterior.

Teorma 13 *Si Φ es un sistema de Haar de orden $n+1$ en X, entonces toda función continua f tiene un único polinomio de mejor aproximación P en X.*

Finalmente, si en X está definido un sistema de Haar se garantiza la unicidad del mejor aproximante. El teorema principal en está sección es el teorema de Alternancia de Chebyshev que se pueden consultar en ([7], pag.74).

Teorma 14 (***de Alternancia de Chebyshev***)
Sean Φ un sistema de Haar de orden n en $[a,b]$ y $A \subseteq [a,b]$ de al menos $n+1$ puntos. Si $f \in C(A)$, entonces P es el polinomio de mejor aproximación a f en A si y sólo si existen $n+1$ puntos $x_0 < \cdots < x_n$ en A_0 y un número $\epsilon = \pm 1$, para el cual $f(x_i) - P(x_i) = d(-1)^i\epsilon$ con $i = 0,\ldots,n$ y donde $d = ||f-P||$.

Notemos que el teorema de alternancia de Chebyshev sigue siendo un teorema de caracterización del polinomio de mejor aproximación y para calcular tal polinomio se puede utilizar el algoritmo de Remez.([7] pag.78)

Un sistema de Chebyshev completo de orden $n+1$ sobre X es un sistema de Chebyshev $\{h_0,...,h_n\}$ de orden $n+1$ sobre X, tal que para cada $k = 1,...,n$, el conjunto $\{h_0,...,h_k\}$ es también un sistema de Chebyshev de orden $k+1$ sobre X. Mientras que un subespacio de Chebyshev completo de orden $n+1$ sobre X es un subespacio de X generado por un sistema de Chebyshev completo de orden $n+1$ en X.

Para el caso que nos interesa, que es el intervalo $[a,b]$ o el círculo unitario $\mathbb{T}$=$[0,2\pi]$ con la distancia $\rho(x,y) = \min\{|x-y|, 2\pi - |x-y|\}$, una función auxiliar que es de mucho apoyo es $w : [a,b] \longrightarrow \mathbb{R}$ dada por

$$w(x) := \begin{cases} 2 & \text{si } x \in (a,b), \\ 1 & \text{si } x = a \text{ o } x = b. \end{cases}$$

Para $X = \mathbb{T}$, hacemos $w(x) := 2$, para toda $x \in \mathbb{T}$.

Teorema 15 *Sea* X *un intervalo real compacto* $[a, b]$ *o el círculo unitario* $\mathbb{T}$. *Para* $n \in \mathbb{N}$ *sea* H *un subespacio de Chebyshev orden* $n+1$ *en* $C(X)$ *generado por las funciones* $h_0, ..., h_n$. *Si* $k = 1, ..., n$ *y si* $x_1, ..., x_n$ *son puntos diferentes en* X *que satisfacen* $\sum_{i=1}^{k} w(x_i) \leq n$, *entonces existe una función* $h \in H$, *tal que*

$$h \geq 0,\ \ h(x_1) = \cdots = h(x_k) = 0 \ y \ h > 0 \ sobre \ X \backslash \{x_1, ..., x_n\}. \tag{2.16}$$

Si X *es un intervalo real compacto* $[a, b]$, n *es par y uno de los puntos extremos* a *o* b *está en* $\{x_1, ..., x_n\}$, *entonces la función* h *puede anularse en ambos puntos* a *y* b. *Esto último no puede ocurrir cuando* H *es un subespacio de Chebyshev completo de orden* $n + 1$ *en* $C[a, b]$.

Capítulo 3

Subespacios de Korovkin

3.1 Subespacios de Korovkin para medidas de Radon

Definición 4 *Dada $\mu \in \mu_b^+(X)$, con X un espacio topológico Hausdorff y localmente compacto, se define para $H \subset C_0(X)$ el conjunto*

$$D_+(H,\mu) = \{f \in C_0(X) : \textit{para cada red } (\mu_i)_{i\in I} \textit{ de elementos de } \mu_b^+(X) \textit{ se satisface la propiedad } \lim_{i\in I} \mu_i(f) = \mu(f), \textit{ siempre y cuando } \sup_{i\in I} \|\mu_i\| < \infty \textit{ y } \lim_{i\in I} \mu_i(h) = \mu(h) \textit{ para cada } h \in H\}.$$

De la definición es claro que $H \subset D_+(H,\mu) \subset C_0(X)$. Lo interesante es saber en que casos se tiene $D_+(H,\mu) = C_0(X)$. En tales casos H es llamado un D_+-subconjunto de Korovkin con respecto a la medida de Radon positiva y acotada μ. Por lo anterior se desea conocer las propiedades que debe tener el subconjunto H para asegurar la igualdad. Como observación, es fácil ver que si $gen(H)$ es el subespacio de $C_0(X)$ generado por H, entonces se tiene $D_+(H,\mu) = D_+(gen(H),\mu)$, por tal razón asumiremos que H es ya un subespacio. Un subconjunto de $C_0(X)$ que está muy relacionado con $D_+(H,\mu)$ es el siguiente.

Definición 5 *Sea $\mu \in \mu_b^+(X)$, con X un espacio topológico Hausdorff localmente compacto, se define para $H \subset C_0(X)$ el conjunto*

$$U_+(H,\mu) = \{f \in C_0(X) : \upsilon(f) = \mu(f) \textit{ siempre que se cumple } \upsilon \in \mu_b^+(X), \textit{ y } \upsilon(h) = \mu(h) \textit{ para cada } h \in H\}.$$

El teorema siguiente es muy importante porque relaciona a $D_+(H,\mu)$ y $U_+(H,\mu)$.

Teorema 16 *Si H un subespacio de $C_0(X)$ y $\mu \in \mu_b^+(X)$, entonces $D_+(H,\mu) = U_+(H,\mu)$.*

Demostración. Es claro que $D_+(H,\mu) \subset U_+(H,\mu)$, ya que si $f \in D_+(H,\mu)$ y $\upsilon \in \mu_b^+(X)$ satisface $\upsilon(h) = \mu(h)$ para cada $h \in H$, al definir la red $(\mu_i)_{i\in I}$ con $\mu_i = \upsilon$, se ve fácilmente que $\sup_{i\in I} \|\mu_i\| = \|\upsilon\| < \infty$ y $\lim_{i\in I} \mu_i(h) = \mu(h)$ para cada $h \in H$. Esto último asegura que $\lim_{i\in I} \mu_i(f) = \mu(f) = \upsilon(f)$, y por tanto $f \in U_+(H,\mu)$. Inversamente, al suponer que $f \in U_+(H,\mu)$ y $(\mu_i)_{i\in I}$ es una red de medidas positivas y acotadas que cumplen $\sup_{i\in I} \|\mu_i\| < \infty$, $\lim_{i\in I} \mu_i(h) = \mu(h)$ para cada $h \in H$ y $\lim\limits_{i\in I} \mu_i(f) \neq \mu(f)$, existiría un $\epsilon > 0$, tal que para cada $i \in I$ hay un $\alpha(i) > i$ con la propiedad $|\mu_{\alpha(i)}(f) - \mu(f)| \geq \epsilon$. Por otra parte, la familia $A = \{\mu_{\alpha(i)}\}$ es fuertemente acotada, es decir, para todo compacto, existe una constante $M_K \geq 0$ tal que $|\mu_{\alpha(i)}(f)| \leq M_K \|f\|$ para toda $\alpha(i)$ y toda f cuyo soporte está contenido en K. Como es fuertemente acotada, entonces A es un conjunto relativamente compacto, por tanto en A existe una subfamilia $(\mu_r)_{r\in I}$ con un filtro $\mathcal{F}$ más fino que el filtro de secciones finales del conjunto de índices I tal que $\lim_{r\in I,\mathcal{F}} \mu_r(h) = \lim_{i\in I} \mu_i(h) = \mu(h)$ para cada $h \in H$. Por tanto se sigue que $\upsilon = \mu$ sobre H, luego, en particular $\mu(f) = \upsilon(f)$ lo cual contradice $|\mu_{\alpha(i)}(f) - \mu(f)| \geq \epsilon$. ■

Para la caracterización de los subconjuntos de Korovkin para medidas de Radon necesitamos del concepto siguiente.

Definición 6 *Un subespacio H de $C_0(X)$ es llamado cofinal si para cada $f \in C_0(X)$, existe $h \in H$ con $f \leq h$.*

La anterior definición es equivalente a que existe un $h_0 \in H$ con $h_0(x) > 0$ para toda $x \in X$.

Teorema 17 *Si X es compacto y H es cofinal en $C(X)$, entonces para $f \in C(X)$ los enunciados siguientes son equivalentes.*

(1) $f \in U_+(H,\mu)$.

(2) $\sup\limits_{\substack{h\in H\\ h\leq f}} \mu(h) = \mu(f) = \inf\limits_{\substack{k\in H\\ f\leq k}} \mu(k)$.

(3) *Para cada $\epsilon > 0$, existen $h, k \in H$ con $h \leq f \leq k$ y $\mu(k) - \mu(h) < \epsilon$.*

Demostración. La equivalencia entre (2) y (3) es clara. Veamos que (1) implica (2). Si al mapeo $p : C(X) \to \mathbb{R}$ definido por $p(g) = \inf_{\substack{k \in H \\ g \leq k}} \mu(k)$ para cada $g \in C(X)$, le aplicamos el Teorema de Hahn-Banach, entonces existe una forma lineal υ sobre $C(X)$ tal que $\upsilon \leq p$ y $\upsilon(f) = p(f)$. Ahora notemos que si $g \in C(X)$ y $g \leq 0$, entonces $\upsilon(g) \leq p(g) \leq 0$. Por tanto υ es positivo y entonces $\upsilon \in \mu^+(X)$. Más aún, $\upsilon = \mu$ sobre H puesto que $p = \mu$ sobre H. Así $\mu(f) = \upsilon(f) = p(f) = \inf_{\substack{k \in H \\ g \leq k}} \mu(k)$, por (1). Para completar esta parte de la demostración, veamos que $-f \in U_+(H, \mu)$, por lo que $\mu(-f) = p(-f)$, o bien, $\mu(f) = -p(-f) = \sup_{\substack{h \in H \\ h \leq f}} \mu(h)$. Para ver que (2) implica (1), si $\upsilon \in \mu^+(X)$ y $\upsilon = \mu$ sobre H, entonces de las desigualdades $\sup_{\substack{h \in H \\ h \leq f}} \upsilon(h) = \mu(f) = \inf_{\substack{k \in H \\ f \leq k}} \upsilon(k)$ se sigue que $\upsilon(f) = \mu(f)$. ■

Otro resultado similar al anterior (con una demostración también muy parecida) que caracteriza a los D_+-subespacios de Korovkin con respecto a una medida de Radon positiva y acotada, es el siguiente.

Teorema 18 *Si X es compacto y H un subespacio de $C(X)$, entonces para $f \in C(X)$ los enunciados siguientes son equivalentes.*
(1) $U_+(H, \mu) = C(X)$ para $\mu \in \mu_b^+(X)$.
(2) H es cofinal y $\sup_{\substack{h \in H \\ h \leq f}} \mu(h) \leq \mu(f) \leq \inf_{\substack{k \in H \\ f \leq k}} \mu(k)$ para cada $f \in C(X)$.
(3) H es cofinal y para cada $f \in C(X)$ y $\epsilon > 0$, existen $h, k \in H$ con $h \leq f \leq k$ y $\mu(k) - \mu(h) < \epsilon$.

Caracterizaremos ahora los subespacios determinantes para medidas de Radon discretas por medio de un criterio *casi puntual.* Supondremos ahora que X es un espacio Hausdorff localmente compacto. Preliminarmente el siguiente resultado dice que si queremos que una medida de Radon $\mu \in \mu_b^+(X)$ admita un subespacio determinante de dimensión finita, entonces necesariamente se debe tener que μ sea una medida de Radon discreta.

Teorema 19 *Sea $\mu \in \mu_b^+(X)$, $\mu \neq 0$ y H es un subespacio de dimensión n de $C_0(X)$. Supongase que H separa los puntos de X, y si X no es compacto para cada $x \in X$ existe un $h \in H$, tal que $h(x) \neq 0$. Entonces existe un número finito de puntos $x_1, ..., x_p \in X$, $p \leq n + 1$ y números reales positivos*

$\lambda_1, ..., \lambda_p$ *tales que* $\sum\limits_{i=i}^{p} \lambda_i \leq \|\mu\|$ *y* $\mu = \sum\limits_{i=1}^{p} \lambda_i \xi_{x_i}$ *sobre* H. *Consecuentemente, si* H *es un* D_+*-subespacio para* μ *o un* D_+^1*-subespacio para* μ, *si es que* $\|\mu\| \leq 1$, *entonces* $\mu = \sum\limits_{i=1}^{p} \lambda_i \xi_{x_i}$.

Demostración. Sin pérdida de generalidad podemos suponer que $\|\mu\| = 1$ y supongamos en principio que X es compacto. Fijemos una base algebraica $\{h_1, ..., h_n\}$ de H y consideremos el mapeo inyectivo $\Phi^* : X \longrightarrow \mathbb{R}^n$ definido por

$$\Phi^*(x) = (h_1(x), ..., h(x)) \text{ para cada } x \in X.$$

Entonces $\Phi^*(X)$ y su casco convexo $K = conv(\Phi^*(X))$ son subconjuntos compactos de $\mathbb{R}^n$. Por el Teorema de Krein-Milman se tiene $\partial_e K \subset \Phi^*(X)$. Consideremos ahora la medida de Radon probabilidad $\mu^* \in \mu_1^+(K)$ definida por $\mu^*(f) = \mu(f \circ \Phi^*)$ para cada $f \in C(K)$. Por otro lado, si denotamos por $r(\mu^*) \in K$ la resultante de μ^*, por el Teorema de Caratheodory existen a lo más $n+1$ puntos $x_1, ..., x_{n+1} \in X$ y $\lambda_1, ..., \lambda_{n+1}$ tal que $\sum\limits_{i=i}^{n+1} \lambda_i = 1$ y $r(\mu^*) = \sum\limits_{i=1}^{n+1} \lambda_i \Phi^*(x_i)$. De acuerdo a lo anterior, para cada $j = 1, ..., n$ obtenemos

$$\mu(h_j) = \mu^*(pr_j \Phi^*) = pr_j(r(\mu^*)) = \sum_{i=1}^{n+1} \lambda_i pr_j \Phi^*(x_i) = \sum_{i=1}^{n+1} \lambda_i h_j(x_i),$$

donde $pr_j : \mathbb{R}^n \longrightarrow \mathbb{R}$ denota la j-ésima proyección. Así, $\mu = \sum\limits_{i=1}^{n+1} \lambda_i \xi_{x_i}$.

Para el caso no compacto, considere la compactificación a un punto X_ω de X esto es para cada $f \in C_0(X)$, la función $\widetilde{f}$ definida por (1.1). Luego consideremos la medida de Radon $\widetilde{\mu}$ definida por (1.6). Entonces $\|\widetilde{\mu}\| = \|\mu\| = 1$ y $\mu(f) = \widetilde{\mu}(\widetilde{f})$ para toda $f \in C_0(X)$. Puesto que $\widetilde{H} = gen\{\widetilde{h}_1, ..., \widetilde{h}_n\}$ separa los puntos de X_ω, por el razonamiento del caso compacto, existen a los más $n+1$ puntos $x_1, ..., x_{n+1} \in X$ y $\lambda_1, ..., \lambda_{n+1}$ tal que $\sum\limits_{i=i}^{n+1} \lambda_i = 1$ y $\widetilde{\mu} = \sum\limits_{i=1}^{n+1} \lambda_i \xi_{x_i}$ sobre $\widetilde{H}$. Sea $I = \{i = 1, ..., n+1 : x_i \neq 0\}$. Entonces $I \neq \emptyset$, ya que de lo contrario $\mu = 0$; así $\mu = \sum\limits_{i \in I} \lambda_i \xi_{x_i}$ sobre H. ■

3.2 Subespacio de Korovkin para operadores

La aproximación tipo Korovkin se desarrolla en medio de los conceptos y los resultados obtenidos sobre la convergencia de las redes de operadores lineales positivos. La caracterización obtenida en relación con la convergencia del operador identidad tiene una subsecuente estimulación a la investigación sobre otra clase de operadores lineales. Esta sección sólo menciona los conceptos básicos y algunos resultados necesarios para comprender la relación entre los espacios de Chebyshev y los subespacios de Korovkin en el espacio de las funciones continuas.

Definición 7 *Un subconjunto H de $C_0(X)$ es llamado subconjunto de Korovkin con respecto a un operador lineal positivo $T : C_0(X) \longrightarrow C_0(Y)$, si se satisface la propiedad siguiente: para cada red $(L_i)_{i\in I}$ de operadores lineales positivos de $C_0(X)$ en $C_0(Y)$ tal que $\sup\limits_{i\in I} \|L_i\| < \infty$ y $\lim\limits_{i\in I} L_i(h) = T(h)$ para cada $h \in H$, entonces se tiene que $\lim\limits_{i\in I} L_i(f) = T(f)$, para cada $f \in C_0(X)$.*

Si H satisface la definición anterior decimos que H es un K_+-subconjunto de Korovkin para T.

Obsérvese que la equicontinuidad de la red de operadores lineales positivos está descrita por la condición $\sup_{i\in I} \|L_i\| < \infty$. Si T es un operador lineal positivo con $\|L\| \leq 1$, entonces T es llamado una contracción, y si H satisface la definición anterior sólo para redes $(L_i)_{i\in I}$ de operadores lineales positivos que son contracciones, entonces se dice que es un subconjunto de Korovkin con respecto a contracciones lineales positivas y se denota por K_+^1-subconjunto de Korovkin para T. Al igual que para medidas de Radon, un subconjunto H de $C_0(X)$ es un subconjunto de Korovkin para T con respecto a operadores lineales positivos si y sólo si el subespacio generado por H es un conjunto de Korovkin, así podemos suponer que en la definición (7), H es ya un subespacio y lo llamaremos K_+-subespacio para T, y de manera similar K_+^1-subespacio para T para contracciones lineales positivas. Notemos que si el espacio X es compacto y $1 \in H$, entonces la red automáticamente es equicontinua, ya que por ser X compacto, entonces $C_0(X) = C(X)$, luego, de ser necesario se reemplaza la red $(L_i)_{i\in I}$ por una subfamilia $(L_i)_{i\geq i_0}$ la condición $\sup_{i\in I} \|L_i\| < \infty$ es satisfecha siempre que $\lim_{i\in I} L_i(h) = T(h)$ para cada $h \in H$.

Si el subespacio H no es un K_+-subespacio para T o un K_+^1-subespacio para T es aún importante conocer al mayor subconjunto $\{f \in C_0(X) :$

$\lim_{i\in I} L_i(f) = T(f)$ siempre que $\sup_{i\in I} \|L_i\| < \infty$ y $\lim_{i\in I} L_i(h) = T(h)$ para cada $h \in H\}$.

Definición 8 *Se llama la clausura de Korovkin para H con respecto al operador lineal positivo T o K_+-clausura de Korovkin de H con respecto a T al conjunto:*

$$\begin{aligned} K_+(H,\mu) \quad &= \{f \in C_0(X) : \textit{para cada red } (L_i)_{i\in I} \textit{ de operadores lineales} \\ &\quad \textit{positivos } L_i : C_0(X) \longrightarrow C_0(Y) \textit{ se satisface } \lim_{i\in I} L_i(f) = T(f), \\ &\quad \textit{siempre que } \sup_{i\in I} \|L_i\| < \infty \textit{ y } \lim_{i\in I} L_i(h) = T(h) \\ &\quad \textit{para cada } h \in H \ \}. \end{aligned} \tag{3.1}$$

Sin dificultad alguna se puede verificar que $K_+(H,\mu)$ es un subespacio de $C_0(X)$ para cualquier subconjunto H.
Nuevamente hacemos notar que para clausuras de Korovkin, si H es un subconjunto de $C_0(X)$, la K_+-clausura de Korovkin con respecto a T es igual a la K_+-clausura de Korovkin con respecto a T del subespacio generado por H. También, según la definiciones de conjunto de Korovkin y de clausura de Korovkin, H es un subconjunto de Korovkin si y sólo si $K_+(H,\mu) = C_0(X)$.
De manera similar se define la K_+^1-clausura de Korovkin de H con respecto a la contracción T, al tomar en la definición 3.1 sólo operadores lineales que son contracciones.

3.3 Relación entre $\mathbf{K_+(H,T)}$ y $\mathbf{D_+(H,\mu)}$

En esta sección describimos la relación existente entre la K_+-clausura de Korovkin de H con respecto al operador lineal positivo T o K_+^1-clausura de Korovkin de H con respecto a la contracción T y los D_+-subconjunto de Korovkin con respecto a la medida de Radon positiva y acotada μ.

Para lo anterior, sea $T : C_0(X) \longrightarrow C_0(Y)$ un operador lineal positivo. Para cada $y \in Y$ fijo se tiene una medida de Radon $\mu_y^T : C_0(X) \longrightarrow \mathbb{R}$ definida por

$$\mu_y^T(f) = Tf(y) \text{ para cada } f \in C_0(X). \tag{3.2}$$

Como puede verse, se tiene $\mu_y^T = \xi_y \circ T$. Un conjunto que aparece en el siguiente teorema es el siguiente:

Definición 9 *Dado un subespacio H, el conjunto de todos las $f \in C_0(X)$, para los cuales se cumple que dado cada $\epsilon > 0$, existen $h, k \in H$ tales que h-$\epsilon \leq f \leq k + \epsilon$ se denota por H^*.*

Se puede verificar que $f \in H^*$ si y sólo si siempre que $\mu \in \mu_b^+(X)$ y $\mu = 0$ sobre H, entonces $\mu(f) = 0$. Para mostrar la relación de $K_+(H, \mu)$ y los $D_+(H, \mu_y^T)$, veamos primero la relación entre $K_+^1(H, \mu)$ y los $D_+^1(H, \mu_y^T)$.

Teorema 20 *Sea $T : C_0(X) \longrightarrow C_0(Y)$ una contracción lineal positiva y sea H un subespacio de $C_0(X)$. Para cada $f \in C_0(X)$ los enunciados siguientes son equivalentes:*

(1) $\quad f \in K_+^1(H, T)$.

(2) $\quad$ *i)* $f \in D_+^1(H, \mu_y^T)$ *para cada* $y \in Y$.

$\qquad$ *ii)* $f \in H^*$.

Demostración. (1) $\Rightarrow$ (2). Primero estableceremos la condición (i). Consideremos $y \in Y$ y $\upsilon \in \mu_b^+(X)$, tal que $\|\upsilon\| \leq 1$ y $\upsilon = \mu_y^T$ sobre H. Fijando una base $\wp$ de vecindades abiertas relativamente compactas para $y \in Y$. Para cada $V \in \wp$ existe una función continua $g_{_V} : Y \longrightarrow [0, 1]$, tal que $g_{_V}(y) = 1$ y $g_{_V} = 0$ sobre $Y \backslash V$. Con lo anterior podemos considerar el operador lineal $L_{_V} : C_0(X) \longrightarrow C_0(Y)$ definido por $L_{_V}(g) = \upsilon(g)g_{_V} + T(g)(1 - g_{_V})$ para cada $g \in C_0(X)$. Así, cada $L_{_V}$ es una contracción, ya que

$$\begin{aligned} |L_{_V}(g)(z)| &\leq |\upsilon(g)|g_{_V}(z) + \|T(g)\|\,(1 - g_{_V})(z) \\ &\leq ||g||g_{_V}(z) + \|g\|\,(1 - g_{_V})(z) \\ &\leq ||g|| \end{aligned}$$

para cada $g \in C_0(X)$ y $z \in Y$. Consideremos ahora el orden parcial $\leq$ sobre $\wp$ definido por $U \leq V$ si $V \subset U$. Si $h \in H$ y $\epsilon > 0$, existe un $U_0 \in \wp$ tal que $|Th(x) - Th(y)| \leq \epsilon$ siempre que $x \in U_0$. Así, si $V \in \wp$ y $U_0 \leq V$, entonces $\|L_{_V}(h) - T(h)\| \leq \epsilon$ puesto que $L_{_V}(h)(x) - T(h)(x) = (T(h)(y) - T(h)(x))g_{_V}(x)$ para cada $x \in X$. Con lo anterior hemos probado que $\lim\limits_{V \in \wp} L_{_V}(h) = T(h)$ para cada $h \in H$ y luego, por (1) la red $(L_{_V}(f))_{V \in \wp}$ converge a $T(f)$. Por tanto $\mu_y^T(f) = \lim\limits_{V \in \wp} L_{_V}(f) = \upsilon(f)$. Así, $f \in U_+^1(H, \mu)$ o de manera equivalente $f \in D_+^1(H, \mu)$. Para mostrar (ii) de (2), sea $\mu \in \mu_b^+(X)$ tal que $\mu = 0$ sobre H. Denotemos por Θ el conjunto de todos los subconjuntos compactos de Y ordenados por inclusión. Para cada K de Θ elijamos un punto $x_K \in Y \backslash K$ y una función $\varphi_K \in K(Y)$ tal que $0 \leq \varphi_K \leq 1$, $\varphi_K(x_K) = 1$ y $\varphi_K = 0$ sobre K. Con lo anterior, $\|\varphi_K\| = 1$ y es fácil verificar que $\lim\limits_{K \in \Theta} \varphi_K f = 0$ uniformemente sobre Y para toda $f \in C_0(Y)$. Si para cada $K \in \Theta$ consideramos la contracción lineal positiva $L_K : C_0(X) \longrightarrow C_0(Y)$ definida por $L_K(g) = \frac{\mu(g)}{\|\mu\|}\varphi_K + T(g)(1 - \varphi_K)$ para cada $g \in C_0(X)$, es claro que la red $(L_K(h))_{K \in \Theta}$ converge a $T(h)$ para cada $h \in H$; luego por (1) la red $(L_K(f))_{K \in \Theta}$ converge a $T(f)$. Consecuentemente $\frac{\mu(g)}{\|\mu\|}\varphi_K = L_K(f) - T(f) + T(f)\varphi_K \longrightarrow 0$. Esto último indica que $\mu(f) = 0$.

(2) $\implies$ (1) Sea $(L_i)_{i \in I}$ una red de contracciones lineales positivas de $C_0(X)$ en $C_0(Y)$, que convergen a $T(h)$ para cada $h \in H$. Si $(L_i(f))_{i \in I}$ no converge a $T(f)$, podemos hallar un $\epsilon > 0$ y para cada $i \in I$ un índice $\alpha(i) \in I$, con $\alpha(i) \geq i$ además de un $y_i \in Y$ tales que

$$|L_{\alpha(i)} f(y_i) - Tf(y_i)| \geq \epsilon. \tag{3.3}$$

Si la red $(y_i)_{i \in I}$ converge al punto en el infinito de Y, tenemos $\lim\limits_{i \in I} \psi(y_i) = 0$ para cada $\psi \in C_0(Y)$. Podemos considerar para cada $i \in I$ la medida de Radon contractiva positiva $\mu_i : C_0(X) \longrightarrow \mathbb{R}$ dada por $\mu_i(g) = L_{\alpha(i)}(g)(y_i)$ para cada $g \in C_0(X)$. Por la compacidad vaga de $\{\mu \in \mu_b^+(X) : \|\mu\| \leq 1\}$ existe un filtro $\mathcal{F}$ más fino que el filtro de secciones finales en I y una medida μ con $\|\mu\| \leq 1$, tal que $(\mu_i)_{i \in I}$ es vagamente convergente a $\mu \in \mu_b^+(X)$ con respecto al filtro $\mathcal{F}$. Con lo anterior, si $h \in H$ tenemos

$$|\mu_i(h)| \leq |L_{\alpha(i)} h(y_i) - Th(y_i)| + |Th(y_i)| \leq \left\|L_{\alpha(i)}(h) - T(h)\right\| + |Th(y_i)|$$

para cada $i \in I$ y entonces la red $(\mu_i(h))_{i \in I}$ converge vagamente a 0 (con respecto al filtro $\mathcal{F}$). Por (ii) de (2), se sigue que $\lim\limits_{i \in I} \mu_i(f) = \mu(f) = 0$, lo cual contradice (3.3).

Si la red $(\mu_i(h))_{i \in I}$ no converge vagamente al punto en el infinito de Y, podemos hallar un subconjunto compacto K de Y y, para cada $i \in I$ un índice $\beta(i) \in I$ con $\beta(i) \geq i$, tal que $y_{\beta(i)} \in K$. Luego podemos considerar la medida de Radon contractiva $\mu_i : C_0(X) \longrightarrow \mathbb{R}$ definida por $\mu_i(g) := L_{\alpha(i)} g(y_{\beta(i)})$ para cada $g \in C_0(X)$. Aplicando nuevamente la compacidad vaga de $\{\mu \in \mu_b^+(X) : \|\mu\| \leq 1\}$ para hallar un filtro $\mathcal{F}_1$ en I más fino que el filtro de secciones finales y una medida $\mu \in \mu_b^+(X)$ con $\|\mu\| \leq 1$, tal que $(\mu_i)_{i \in I}$ es vagamente convergente a μ con respecto al filtro $\mathcal{F}$, también por la compacidad de K podemos considerar otro filtro $\mathcal{F}_2$ sobre I más fino que

$\mathcal{F}_1$, tal que $(y_{\beta(i)})_{i\in I}$ converge a un $y \in Y$ con respecto al filtro $\mathcal{F}_2$. De tal manera que, si $h \in H$, para cada $i \in I$, tenemos

$$\mu_i(h) = L_{\alpha(i)}h(y_{\beta(i)}) - Th(y_{\beta(i)}) + Th(y_{\beta(i)})$$

y puesto que

$$|L_{\alpha(i)}h(y_{\beta(i)}) - Th(y_{\beta(i)})| \leq \left\|L_{\alpha(i)}(h) - T(h)\right\| \longrightarrow 0.$$

Se obtiene que $\mu(h) = \lim\limits_{i\in I} \mu_i(h) = Th(y) = \mu_y^T(h)$. Dado que $f \in U_+^1(H, \mu_y^T)$, tenemos también $\mu(f) = \mu_y^T(f) = Tf(y)$, que es una contradicción a (3.3). Ahora observemos que la condición (ii) de (2) significa que $f \in U_+(H,0) = H^*$. ■
Este resultado en breve dice que

$$K_+^1(H,T) = H^* \bigcap \left(\bigcap_{y\in Y} D_+^1(H, \mu_y^T) \right).$$

Más aún, si Y es compacto, entonces (1) es equivalente a la parte (i) de (2), y

$$K_+^1(H,T) = \bigcap_{y\in Y} D_+^1(H, \mu_y^T) = \bigcap_{y\in Y} U_+^1(H, \mu_y^T).$$

El caso en que $T : C_0(X) \longrightarrow C_0(Y)$ es un operador lineal positivo que no es una contracción, lo mencionamos enseguida y la demostración es muy similar a la anterior.

Teorema 21 *Sea $T : C_0(X) \longrightarrow C_0(Y)$ un operador lineal positivo y sea H un subespacio de $C_0(X)$. Para cada $f \in C_0(X)$ los enunciados siguientes son equivalentes:*
(1) $f \in K_+(H,T)$.
(2) $f \in D_+(H, \mu_y^T)$ *para cada* $y \in Y$.

En breve, se tiene

$$K_+(H,T) = \bigcap_{y\in Y} D_+(H, \mu_y^T) = \bigcap_{y\in Y} U_+(H, \mu_y^T).$$

Capítulo 4

Subespacios de Chebyshev y de Korovkin

4.1 Orden de un subespacio de Chebyshev

Hemos mencionado en el Capítulo 2 que, si X es un espacio de Hausdorff con al menos $n+1$ puntos y $h_0, h_1, ..., h_n$ funciones reales continuas, entonces decimos que el conjunto de funciones $\Phi = \{ h_0, h_1, ..., h_n\}$ es un sistema de Haar de orden $n+1$ en X, si cada combinacion lineal $\lambda_0 h_0 + \lambda_1 h_1 + \cdots + \lambda_n h_n$ con coeficientes λ_0, λ_1, ..., $\lambda_n \in \mathbb{R}$ no simultáneamente nulos, tienen a lo más de n ceros distintos en X. El subespacio $H = gen\{ h_0, h_1, ..., h_n\}$ es llamado subespacio de Haar de orden $n+1$ en $C(X)$. En particular, en el caso en que X es compacto decimos que Φ es un sistema de Chebyshev, y H es llamado subespacio de Chebyshev.

Una caracterización de los sistemas de Haar es que $\Phi = \{ h_0, h_1, ..., h_n\}$ forma un sistema de Haar de orden $n+1$ en $C(X)$ si y sólo si $\det(h_i(x_j)) \neq 0$ para cualesquiera puntos distintos $x_0, ..., x_n$ de X.

También, si Φ forma un sistema de Haar en X de orden $n+1$, $x_0, ..., x_n \in X$ y $\alpha_0, ..., \alpha_n \in \mathbb{R}$, entonces existe una única combinación lineal $h = \sum_{i=0}^{n} \lambda_i h_i$ tal que $h(x_i) = \alpha_i$ con $i = 0, ..., n$. Un resultado importante que relaciona el orden de un subespacio de Chebyshev con los D_+-subespacios es el siguiente.

Teorema 22 *Sea X el intervalo real $[a,b]$ o el círculo unitario $\mathbb{T}$. y sea H un subespacio de Chebyshev en $C(X)$ de dimensión $n+1$. Entonces para toda elección de puntos diferentes $x_1,...,x_p \in X$ satisfaciendo $\sum_{i=1}^{p} w(x_i) \leq n$ y para todo $\mu \in C_+(x_1,...,x_p)$ el subespacio H es un D_+-subespacio para μ.*

Demostración. Supóngase primero que n es impar y consideremos p puntos distintos $x_1,...,x_p \in X$ satisfaciendo $\sum_{i=1}^{p} w(x_i) \leq n$. Fijemos una medida $\mu = \sum_{i=1}^{p} \alpha_i \xi_{x_i} \in C_+(x_1,...,x_p)$. Si $\upsilon \in \mu^+(X)$ cumple que $\upsilon = \mu$ sobre H. Entonces por (2.16) del teorema (15), existe una $h \in H$ para la cual se tiene $h \geq 0$, $h(x_1) = \cdots = h(x_p) = 0$ y $h(x) > 0$ para $x \in X \backslash \{x_1,...,x_p\}$, luego se tiene $\upsilon(h) = 0$, por lo que $Sop(\upsilon)$ debe estar contenido en el conjunto $\{x_1,...,x_p\}$ y entonces $\upsilon = \sum_{i=1}^{p} \lambda_i \xi_{x_i}$ para algunos $\lambda_1,...,\lambda_p \in \mathbb{R}^+$. Consecuentemente, para toda $h \in H$ obtenemos $\sum_{i=1}^{p}(\alpha_i - \lambda_i)h(x_i) = 0$. Dado que $p \leq n+1$, podemos elegir de ser necesario $(n+1)-p$ puntos distintos $x_{p+1},...,x_{n+1}$ para los cuales $\sum_{i=1}^{n+1} \beta_i h(x_i) = 0$ para cada $h \in H$, siendo

$$\beta_i := \begin{cases} \alpha_i - \lambda_i & \text{si } 1 \leq i \leq p, \\ 0 & \text{si } p+1 \leq i \leq n+1. \end{cases}$$

Dado que H es un subespacio de Chebyshev de orden $n+1$ debemos tener $\beta_i = 0$ para $i = 1,...,n+1$ y entonces $\upsilon = \mu$. Esto muestra que H es un D_+-subespacio para μ.

Supongase ahora que n es par y que $X = [a,b]$. También fijemos $\mu = \sum_{i=1}^{p} \alpha_i \xi_{x_i} \in C_+(x_1,...,x_p)$, donde $a \leq x_1 < \cdots < x_p \leq b$ y $\sum_{i=1}^{p} w(x_i) \leq n$. Al considerar a $\upsilon = \mu$ sobre H, nuevamente por (2.16) del teorema (15) podemos encontrar una función $h \in H$ para la cual se tiene $h \geq 0$, $h(x_1) = \cdots = h(x_p) = 0$ y $h(x) > 0$ para $x \in X \backslash \{x_1,...,x_p\}$. No podemos proceder como en el caso anterior cuando sólo uno de los puntos extremos a o b está en $\{x_1,...,x_p\}$, puesto que en este caso h puede anularse en a y en b. Supongase que $a < x_1 < \cdots < x_p < b$. Si $h(b) > 0$, entonces la conclusión se sigue como antes. Si $h(b) = 0$, entonces consideremos la medida de Radon $\upsilon_1 = \upsilon + \xi_b$. Puesto que $\upsilon_1(h) = \upsilon(h) + h(b) = 0$, también se tiene $Sop(\upsilon_1) \subset \{x \in [a,b] : h(x) = 0\} = \{x_1, ..., x_p, b\}$. Por lo anterior, existen $\lambda_1, ..., \lambda_p, \lambda_{p+1} \in \mathbb{R}^+$ tal que $\upsilon_1 = \sum\limits_{i=1}^{p} \lambda_i \xi_{x_i} + \lambda_{p+1}\xi_b$. esto implica que para toda $h \in H$, $\sum\limits_{i=1}^{p} \alpha_i u(x_i) = \mu(u) = \upsilon(u) = \upsilon_1(u) - u(b) = \sum\limits_{i=1}^{p} \lambda_i u(x_i) + \lambda_{p+1}u(b) - u(b)$, es decir $\sum\limits_{i=1}^{p} \alpha_i u(x_i) + u(b) = \sum\limits_{i=1}^{p} \lambda_i u(x_i) \lambda_{p+1} u(b)$. Así $\upsilon = \mu$. ■

4.2 Orden de un subespacio de Korovkin

En esta parte consideraremos la convergencia de redes de operadores lineales positivos (o contracciones lineales positivas) hacia un operador que corresponde de forma natural a medidas de Radon discretas. Para lo anterior, primero damos la definición de los operadores en los que nos centramos.

Definición 10 *Sean X y Y espacios de Hausdorff localmente compactos y sea $T : C_0(X) \longrightarrow C_0(Y)$ un operador lineal y positivo. Decimos que T es un operador finitamente definido de orden n si existen n mapeos $\varphi_1, ..., \varphi_n : Y \longrightarrow X$ y n funciones reales $\psi_1, ..., \psi_n : Y \longrightarrow \mathbb{R}$ tal que*

$$T(f) = \sum_{i=1}^{n} \psi_i (f \circ \varphi_i) \text{ para cada } f \in C_0(X). \tag{4.1}$$

Se asume que el lado derecho de (4.1) pertenece $C_0(Y)$ para cada $f \in C_0(X)$. Lo anterior es válido siempre que Y sea compacto y cada una de las funciones $\varphi_1, ..., \varphi_n$, $\psi_1, ..., \psi_n$ sean continuas, o en el caso en que Y no es compacto, que las $\psi_1, ..., \psi_n$ estén en $C_0(Y)$ y las $\varphi_1, ..., \varphi_n$ sean continuas.

Denotaremos por $\mathcal{L}_n(X,Y)$ al conjunto de todos los operadores finitamente definidos de orden n de $C_0(X)$ en $C_0(Y)$ y por $\mathcal{L}_n^1(X,Y)$ al conjunto de todos los operadores finitamente definidos de orden n de $C_0(X)$ en $C_0(Y)$ que admiten una representación (4.1) con $\sum_{i=1}^{n} \psi_i = 1$. Se hace notar que $\mathcal{L}_n(X,Y) \subset \mathcal{L}_{n+1}(X,Y)$ y $\mathcal{L}_n^1(X,Y) \subset \mathcal{L}_{n+1}^1(X,Y)$ para cada $n \in \mathbb{N}$. Un operador finitamente definido $T : C_0(X) \longrightarrow C_0(Y)$ de orden 1 es de la forma $T(f) = \psi \cdot (f \circ \varphi)$ para cada $f \in C_0(X)$, donde $\varphi : Y \longrightarrow X$ y $\psi : Y \longrightarrow \mathbb{R}$. Los operadores fintamente definidos están estríctamente relacionados con medidas de Radon discretas positivas; de hecho, si $T : C_0(X) \longrightarrow C_0(Y)$ es un operador finitamente definido de orden n con una representación (4.1), tenemos entonces que

$$\mu_y^T(f) = Tf(y) = \sum_{i=1}^{n} \psi_i(y) \cdot f(\varphi_i(y)) = \sum_{i=1}^{n} \psi_i(y) \cdot \xi_{\varphi_i(y)}(f),$$

para $y \in Y$ y $f \in C_0(X)$. Consecuentemente, se tiene

$$\mu_y^T = \sum_{i=1}^{n} \psi_i(y) \cdot \xi_{\varphi_i(y)} \tag{4.2}$$

es una medida de Radon discreta positiva para cada $y \in Y$. Así

$$\mu_y^T \in C_+(\varphi_1(y), ..., \varphi_n(y)) \tag{4.3}$$

para cada $y \in Y$. Más aún, si $T \in \mathcal{L}_n^1(X, Y)$, entonces μ_y^T es también una medida contracción y

$$\mu_y^T \in C_+^1(\varphi_1(y), ..., \varphi_n(y)) \tag{4.4}$$

para cada $y \in Y$.

Las relaciones previas permiten establecer el teorema siguiente, que puntualiza la importancia de los operadores finitamente definidos en teoría de aproximación tipo Korovkin. De hecho, estos son los únicos operadores lineales positivos que pueden tener subespacios de Korovkin de dimensión finita.

Teorema 23 *Sea $T : C_0(X) \longrightarrow C_0(Y)$ un operador lineal y positivo y considere un subespacio n-dimensional H de $C_0(X)$. Supóngase que H separa los puntos de X y si X no es compacto, adicionalmente que para cada $x \in X$ exista $h \in H$ tal que $h(x) \neq 0$. Entonces existen $n+1$ mapeos $\varphi_1, ..., \varphi_{n+1} : Y \longrightarrow X$ y $n+1$ funciones reales $\psi_1, ..., \psi_{n+1} : Y \longrightarrow \mathbb{R}$ tales que*

$$T(h) = \sum_{i=1}^{n+1} \psi_i (h \circ \varphi_i) \text{ para cada } h \in H. \tag{4.5}$$

Consecuentemente, si H es un K_+-subespacio para T de dimensión n (o un K_+^1-subespacio para T de dimensión n, con $\|T\| \leq 1$), entonces T es finitamente definido de orden $n+1$. Finalmente, si X y Y son compactos, $T(1) = 1$. y si H es un K_+^1-subespacio para T , se tiene que $T \in \mathcal{L}_{n+1}^1(X, Y)$.

Hasta ahora, hemos tratado los K_+-subespacios con un operador fijo T. Ahora fijemos a n y veamos las K_+-clausuras y K_+^1-clausuras para todo operador finitamente definido de orden n.

Definición 11 *Sea X un espacio de Hausdorff localmente compacto. Un subconjunto H de $C_0(X)$ es llamado un K_+-subconjunto (o K_+^1-subconjunto, respectivamente) de orden n en $C_0(X)$, si H es un K_+-subconjunto para T (o un K_+^1-subconjunto para T, con $\|T\| \leq 1$, respectivamente) para todo espacio Hausdorff localmente compacto Y y todo operador finitamente definido $T \in \mathcal{L}_n(X,Y)$ ($T \in \mathcal{L}_n^1(X,Y)$, respectivamente). Cuando H es un subespacio de $C_0(X)$, se llama a éste un K_+-subespacio (o K_+^1-subespacio, respectivamente) de orden n.*

Claramente todo K_+-subespacio de orden n en $C_0(X)$ es un K_+-subespacio de orden p, para $1 \leq p \leq n$.

4.3 Relación entre órdenes de subespacios

En esta sección supondremos que X es un espacio Hausdorff, localmente compacto y con más de n puntos.

Teorema 24 *Sea H un subespacio de $C_0(X)$. Entonces los enunciados siguientes son equivalentes.*
(1) *H es un K_+-subespacio de orden n en $C_0(X)$.*
(2) *Para toda elección de n puntos diferentes $x_1, ..., x_n \in X$ y para toda $\mu \in C_+(x_1, ..., x_n)$ el subespacio H es un D_+-subespacio para μ.*
(3) *Para todo conjunto de n puntos distintos $x_1, ..., x_n \in X$, para todo compacto K de X satisfaciendo $K \cap \{ x_1, ..., x_n \} = \emptyset$ y para todo $\epsilon > 0$ existe un $h \in H$ y $u \in C_0^+(X)$ tal que $\|u\| \leq \epsilon$, $0 \leq h+u$ sobre H y $h(x_i)+u(x_i) < \epsilon$ para cada $i = 1, ..., n$.*

Demostración. (1) $\Longrightarrow$ (2) Sean $x_1, ..., x_n \in X$ y fijemos $\mu \in C_+(x_1, ..., x_n)$. Entonces $\mu = \sum\limits_{i=1}^{n} \alpha_i \xi_{x_i}$ con $\alpha_1, ..., \alpha_n \in \mathbb{R}^+$. Sea $Y = \{1\}$ y para cada $i = 1, ..., n$, consideremos los mapeos constantes $\varphi_i : Y \longrightarrow X$ y $\psi_i : Y \longrightarrow \mathbb{R}$ de valores x_i y α_i, respectivamente. Entonces el operador $T : C_0(X) \longrightarrow C_0(Y)$ definido por $T(f) = \sum_{i=1}^{n+1} \psi_i(f \circ \varphi_i)$ para cada $f \in C_0(X)$ es finitamente definido de orden n y consecuentemente por virtud del teorema (21), H es un D_+-subespacio para cada μ_y^T, $y \in Y$. Pero $\mu_y^T = \mu$ y entonces (2) es válido. (2) $\Longrightarrow$ (3) es una consecuencia directa del teorema 15. ■

Corolario 8 *Sea H un subespacio de $C_0(X)$ que satisface lacondición siguiente: si para toda elección de puntos distintos $x_1, ..., x_n \in X$ y para todo $x \in X \backslash \{ x_1, ..., x_n\}$ existe un $h \in H$, $h \geq 0$, tal que $h(x_1) = \cdots = h(x_n) = 0$ y $h(x) > 0$. Entonces H es un K_+-subespacio de orden n en $C_0(X)$.*

Teorema 25 *Sea X el intervalo real $[a, b]$ o el círculo unitario $\mathbb{T}$. Si H es un subespacio de $C(X)$ de dimensión $n+1$, entones los enunciados siguientes son válidos:*
(1) *Si n es par, H es un subespacio de Chebyshev de orden $n+1$ en $C(X)$ si y sólo si H es un K_+-subespacio de orden $\frac{n}{2}$ en $C(X)$.*
(2) *Si n es impar y H es un subespacio de Chebyshev de orden $n+1$ en $C(X)$, entonces H es un K_+-subespacio de orden $\frac{n-1}{2}$ en $C(X)$.*

Demostración. Hagamos $q_n = \frac{n}{2}$ si n es par y $q_n = \frac{n-1}{2}$ si n es impar. Si H es un subespacio de Chebyshev de orden $n+1$, entonces para cada elección de n puntos diferentes $x_1, ..., x_{q_n} \in X$ siempre se tiene $\sum\limits_{i=1}^{q_n} w(x_i) \leq n$. Por el teorema (22), H es un D_+-subespacio para μ para cada $\mu \in C_+(x_1, ..., x_{q_n})$. Consecuentemente ahora por el teorema (24), H es un K_+-subespacio de orden q_n en $C(X)$.
Para concluir la demostración de esta parte (1), sólo tenemos que mostrar que si n es par y si H es un K_+-subespacio de orden $\frac{n}{2}$ en $C(X)$, entonces H es un sub-espacio de Chebyshev de orden $n+1$ en $C(X)$. Para lo anterior, supóngase que H es algebráicamente generado por las funciones linealmente independientes $h_0, ..., h_n$. Si H no es un subespacio de Chebyshev de orden $n+1$ en $C(X)$, existen $n+1$ puntos diferentes $x_0, ..., x_n \in X$ y $n+1$ números reales $\alpha_0, ..., \alpha_n$ no todos cero, tal que $\sum\limits_{i=0}^{n} \alpha_i h_j(x_i) = 0$ para cada $j = 0, ..., n$. Si $J = \{i = 0, ..., n : \alpha_i > 0\}$, podemos asumir que $\mathrm{Card}(J) \leq \frac{n}{2}$, de otra forma multiplicamos cada α_i por -1. Al considerar la medida de Radon positiva $\mu = \sum\limits_{i \in J} \alpha_i \xi_{x_i}$ y $\upsilon = -\sum\limits_{i \notin J} \alpha_i \xi_{x_i}$ entonces $\mu = \upsilon$ sobre H y entonces por (2) del teorema 24, se debe tener $\mu = \upsilon$. Ésto último contradice que los puntos $x_0, ..., x_n$ son distintos. ■

Conclusiones

El objetivo de esta tesis ha sido el estudio de la relación entre los subespacios de Chebyshev y los subespacios de Korovkin en el espacio de las funciones reales continuas que se anulan en el infinito, definidas sobre un espacio localmente compacto y para algunos resultados especiales se ha requerido de la compacidad de estos espacios. Principalmente lo que se ha procurado en este trabajo es mostrar la relación entre las dos teorías mencionadas y los correspondientes órdenes de los subespacios.
El problema tratado tiene suma complejidad, ya que para establecer la igualdad de los órdenes mencionados se ha utilizado una serie de resultados que involucra trabajar con conjuntos de medidas de Radon asociadas a un operador lineal positivo, esto último para los espacios de Korovkin, mientras que para los espacios de Chebyshev se ha requerido de resultados sobre caracterización de los conjuntos que son sistemas de Chebyshev.
Aunque hemos establecido relaciones entre órdenes de subespacios de Chebyshev y de Korovkin para el espacio de funciones reales continuas definidas en los espacios topológicos $[a, b]$ y $\mathbb{T}$ el grupo unitario, los resultados son más generales, pero las demostraciones son aún más compleja.
Muchos de los resultados aquí presentados podrían extenderse al caso no simétrico, esto es, en lugar de utilizar la norma del supremo, utilizar las normas asimétricas definidas por

$$p(x) = \max(0, x),$$

y su conjugada que viene dada por

$$\overline{p}(x) = \max(0, -x).$$

Cabe mencionar que el caso no simétrico ha sido ya tratado para la teoría de Chevyshev en la tesis “Aproximación polinomial mediante bandas de amplitud variantes”, por la Dra. Ivonne Lilian Martínez Cortes ; sin embargo

aún no hay trabajos similares en la dirección de la teoría de Korovkin. Lo anterior da la pauta para un posible estudio con ese enfoque en un trabajo doctoral.
Creo que la forma en que se dan las demostraciones de los resultados y su presentación son un aporte importante.

Bibliografía

[1] Altomare F. and Campiti M., *Korovkin type Approximation Theory and its Applications,* Walter de Gruyter Studies in Mathematics 17 (1994).

[2] Altomare F. *Korovkin Type Theorems and Approximation by Positive Linear Operators,* Topology Appl., 155 (2008) 527-539.

[3] Bernd Anger and Claude Portenier, *Radon Integrals Theory,* Progress in Mathematics Volume 6 (2010).

[4] Bauer H. *Theorems of Korovkin type for adapted spaces*, French Summary, Ann. Inst. Fourier (Grenoble) 23 (1973).

[5] Cheney E. W. *Introduction to Approximation Theory,* MgGraw-Hill, USA (1993).

[6] Gustave Choquet. *Lectures on Analysis*, Aportaciones Matemáticas W. A. Benjamin, INC. Advanced Book Program, volume 1 (1996).

[7] DeVore R. A. and G. G. Lorentz, *Constructive Approximation*, Springer-Verlag, New York, (1993).

[8] Royden H. L. *Real Analysis*, 2nd McMillan, New York, (1998).