UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA

"Teoremas de existencia y unicidad para la aproximación con restricciones mediante sistemas de Chebyshev"

Tesis

Para obtener el título de

LICENCIADO EN MATEMÁTICAS APLICADAS

Presenta

Miriam Cisneros Martínez

Director
M.C. Luz del Carmen Álvarez Marín

Co-director
Dr. Jorge Bustamante González

Huajuapan de León, Oaxaca, noviembre de 2011

Dedicatoria

Dedico esta tesis

A mis padres Rosa Elia y Rodrigo

*A mis hermanos Jerzi, Iracema,Yanel
Rodrigo, Alondra y Daniel.*

Agradecimientos

Agradezco de manera muy especial a mis padres por el apoyo y la confianza que me brindaron.

Doy gracias a mi director de tesis, la M.C. Luz del Carmen Álvarez Marín por todo el tiempo, la paciencia y los consejos brindados para la realización de la presente tesis.

Agradezco a mi co-director de tesis, el Dr. Jorge Bustamante González por su valiosa colaboración en la realización de esta tesis.

Quisiera agradecer a mis sinodales: M.C.Vulfrano Tochihuitl Bueno, Dr. Cuauhtémoc H. Castañeda Roldán, M.C. Juan Luis Hernández López, por el tiempo invertido en cada una de las revisiones, así como también los consejos que me dieron sobre éstas.

Quisiera agradecer a la Dra. Virginia Berrón Lara por el apoyo, sobre todo el tiempo, brindado en la carrera.

Agradezco también a mis profesores por sus enseñanzas durante este tiempo.

Asimismo, quisiera agradecer a todas las personas que estuvieron apoyándome desde el inicio hasta el fin de la carrera.

Contenido

Introducción

En 1854, el matemático ruso P. L Chebyshev, en su libro "Théorie des mécanismes connus sous le nom de parallélogrammes", sienta las bases de la Teoría de Aproximación para que ésta sea considerada como una rama de las matemáticas, ésto lo hace mediante la formulación del siguiente problema: Dada una función continua f, encontrar un polinomio de grado fijo, tal que el máximo de su desviación de f en un intervalo dado, sea más pequeño que el de los otros polinomios del mismo grado.

La idea fundamental tras la Teoría de Aproximación es la de reemplazar un objeto de estudio por uno más fácil de manipular. Por ejemplo, se conocen funciones con expresiones muy simples, como es el caso de $f(x) = e^{x^2}$, para las cuales no se tienen expresiones exactas para calcular sus integrales. La Teoría de Aproximación proporciona métodos para encontrar polinomios (u otros objetos) muy parecidos a las funciones dadas con los cuales se puede obtener, de forma aproximada, el valor de dicha integral.

El desarrollo posterior de la aproximación de funciones se vio influido por un importante descubrimiento matemático hecho a finales del siglo XIX por el matemático alemán Weierstrass, quien demostró con absoluto rigor la posibilidad de aproximar una función continua arbitraria, con cualquier grado de exactitud, mediante un polinomio algebraico [7].

Las ideas de Chebyshev y el teorema de Weierstrass jugaron un papel importante para sentar las bases de la actual Teoría de Aproximación, la cual considera problemas similares tanto en espacios normados como métricos.

En los últimos años se han desarrollado nuevas técnicas que permiten ampliar las investigaciones por un lado, y por otro obligan a replantearse problemas ya conocidos. Tal es el caso del siguiente resultado clásico, debido a Chebyshev: " Dada una función real continua no trivial sobre un intervalo real $[a, b]$, el polinomio algebraico de grado a lo sumo n, que mejor aproxima a la función, se caracteriza por alternar sus valores máximos y mínimos, en $n + 2$ puntos consecutivos del intervalo de definición ", el cual ha sido generalizado en varias direcciones como se muestra en [3], [8], [12], [14] y [15].

Para obtener tales tipos de afirmaciones se requiere contar previamente con resultados que aseguren la existencia y unicidad del polinomio que mejor aproxima.

A partir de los años 50 del siglo pasado, se hizo evidente que era necesario trabajar con clases de funciones más generales que los polinomios, lo que condujo al estudio de los llamados sistemas de Chebyshev.

En el presente trabajo, se analizan problemas relacionados con la aproximación con restricciones tanto para polinomios algebraicos como para sistemas de Chebyshev (polinomios generalizados), tales como lo son la existencia, caracterización y unicidad del mejor aproximante.

Para presentar los resultados obtenidos, hemos dividido esta tesis en tres capítulos distribuidos de la siguiente manera:

En el primer capítulo se desarrollan resultados referentes a la aproximación polinomial uniforme, considerando la clase de los polinomios de grado menor o igual a n.

El segundo capítulo, está destinado a la aproximación con restricciones para el caso en el que la clase de aproximaciones son los polinomios algebraicos de grado menor o igual a n, cabe mencionar que las condiciones adicionales que se le han pedido a estos polinomios es que su rango esté restringido.

En el tercer capítulo, se encuentran los resultados principales, es decir, hacemos el planteamiento del problema de aproximación con restricciones, usando ahora sistemas de Chebyshev, obteniendo resultados análogos a los expuestos en el capítulo anterior para polinomios.

Por último, se presentan las conclusiones derivadas de los capítulos anteriores.

Capítulo 1

Aproximación polinomial uniforme

1.1. Existencia del mejor aproximante

En este capítulo nos enfocaremos al desarrollo de resultados referentes a la mejor aproximación (o mejor aproximante) polinomial en la norma uniforme, tales como la existencia, unicidad y caracterización.

Definición 1.1.1. Sea V un espacio lineal sobre $\mathbb{R}$. Llamamos norma a una función

$$\| \cdot \| : V \to \mathbb{R}^+ \cup \{0\},$$

que cumple con las siguientes condiciones:

Dados $w, v \in V,\ \lambda \in \mathbb{R}$ arbitrarios.

$$\begin{array}{ll} (i) & \| v \| = 0 \text{ si y sólo si } v = 0. \\ (ii) & \| \lambda v \| = |\lambda| \, \| v \| \, . \\ (iii) & \| v + w \| \leq \| v \| + \| w \| \quad \text{(la desigualdad del triángulo).} \end{array} \tag{1.1}$$

La norma da una noción de distancia en V. Si $w, v \in V$, entonces la distancia de w a v es $\| v - w \|$.

Teorema 1.1.2. (Existencia del mejor aproximante) Si V es un espacio lineal normado y W un subespacio de V de dimensión finita, entonces dado $v \in V$, existe $w^* \in W$ tal que

$$\| v - w^* \| \leq \| v - w \|$$

para todo $w \in W$.

Demostración. Sea $v \in V$. Supongamos que W tiene dimensión k y que $\{w_1, w_2, \dots, w_k\}$ es una base para W. Probaremos que la función $f : \mathbb{R}^k \to \mathbb{R}$ definida por

$$f(\lambda_1, \lambda_2, \dots, \lambda_k) = \| v - (\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_k w_k) \| \tag{1.2}$$

tiene un mínimo sobre $\mathbb{R}^k$.

Veamos primero que la función f es continua.
Por equivalencia de normas, dados $\lambda,\ \lambda' \in \mathbb{R}^k$

$$\| \lambda - \lambda' \|_{máx} < c \| \lambda - \lambda' \|_2,$$

donde $c > 0$, $\| \cdot \|_2$ es la norma euclidiana y $\| \cdot \|_{máx}$ es la norma del máximo.

Sea $\epsilon > 0$, $\lambda \in \mathbb{R}^k$. Haciendo $\delta = \frac{\epsilon}{c \sum_{i=1}^{k} \|w_i\|}$, tenemos que para toda $\lambda' \in \mathbb{R}^k$,

si $\| \lambda' - \lambda \|_2 < \delta$, entonces $\| \lambda' - \lambda \|_{máx} < \frac{\epsilon}{\sum_{i=1}^{k} \|w_i\|}$,

es decir

$$|\lambda'_i - \lambda_i| < \frac{\epsilon}{\sum\limits_{i=1}^{k} \|w_i\|}, \qquad i = 1, 2, \dots, k;$$

luego

$$|\lambda'_1 - \lambda_1| \|w_1\| + |\lambda'_2 - \lambda_2| \|w_2\| + \cdots + |\lambda'_k - \lambda_k| \|w_k\| < \frac{\epsilon}{\sum\limits_{i=1}^{k} \|w_i\|} \left(\sum_{i=1}^{k} \|w_i\| \right) = \epsilon,$$

y como

$$\begin{aligned} |f(\lambda') - f(\lambda)| &= | \| v - (\lambda'_1 w_1 + \lambda'_2 w_2 + \cdots + \lambda'_k w_k) \| - \\ &- \| v - (\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_k w_k) \| | \\ &\leq |\lambda'_1 - \lambda_1| \|w_1\| + |\lambda'_2 - \lambda_2| \|w_2\| + \cdots + |\lambda'_k - \lambda_k| \|w_k\| \end{aligned}$$

entonces

$$|f(\lambda') - f(\lambda)| < \epsilon.$$

Con esto, queda demostrada la continuidad de f.

De igual forma la función $h : \mathbb{R}^k \to \mathbb{R}$ definida por

$$h(\lambda_1, \lambda_2, \ldots, \lambda_k) = \| \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_k w_k \|$$

es continua.

Consideremos la superficie esférica:

$$S = \left\{ (\lambda_1, \lambda_2, \ldots, \lambda_k) \in \mathbb{R}^k : \ \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_k^2 = 1 \right\}. \tag{1.3}$$

Tenemos que S es cerrada y acotada, por tanto h alcanza su mínimo m aquí. Este valor mínimo, m, es positivo. Para ver ésto, observemos que si $h(\lambda_1^*, \lambda_2^*, \ldots, \lambda_k^*) = m$, entonces $m \geq 0$. Supongamos que $m = 0$ entonces $\| \lambda_1^* w_1 + \lambda_2^* w_2 + \cdots + \lambda_k^* w_k \| = 0$, lo cual por (1.1), se sigue que $\lambda_1^* w_1 + \lambda_2^* w_2 + \cdots + \lambda_k^* w_k = 0$. Como $w_1, w_2, \ldots, w_k$ son linealmente independientes, concluimos que $\lambda_1^* = \lambda_2^* = \cdots = \lambda_k^* = 0$, contradiciendo (1.3). De donde $m > 0$.

Para $(\lambda_1, \lambda_2, \ldots, \lambda_k) \in \mathbb{R}^k - \{0\}$, consideremos $r = (\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_k^2)^{1/2}$, luego, se tiene

$$\frac{1}{r}(\lambda_1, \lambda_2, \ldots, \lambda_k) \in S \quad \text{y}$$

$$h(\lambda_1, \lambda_2, \ldots, \lambda_k) = r \, \| (\frac{\lambda_1}{r} w_1 + \frac{\lambda_2}{r} w_2 + \cdots + \frac{\lambda_k}{r} w_k) \|,$$

así que

$$h(\lambda_1, \lambda_2, \ldots, \lambda_k) \geq mr.$$

Más aún,

$$\begin{aligned} f(\lambda_1, \lambda_2, \ldots, \lambda_k) \quad &= \| v - (\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_k w_k) \| \\ &\geq \| \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_k w_k \| - \| v \| \\ &\geq mr - \| v \| . \end{aligned}$$

Hagamos

$$\rho = \inf_{\lambda \in \mathbb{R}^k} f(\lambda) \quad \text{y} \quad d = \frac{1 + \rho + \| v \|}{m}.$$

Si $\sum_{i=1}^{k} \lambda_i^2 > d^2$, entonces $f(\lambda_1, \lambda_2, \ldots, \lambda_k) > md - \| v \| = 1 + \rho > \rho$, por tanto el mínimo de f sobre $\mathbb{R}^k$ es igual al mínimo de f sobre

$$A = \left\{ (\lambda_1, \lambda_2, \ldots, \lambda_k) \in \mathbb{R}^k : \sum_{i=1}^{k} \lambda_i^2 \leq d^2 \right\}.$$

Como f es una función continua de $(\lambda_1, \lambda_2, \ldots, \lambda_k)$, entonces f alcanza su mínimo en A.

□

Definición 1.1.3. Un elemento w^* que cumple con las condiciones del teorema 1.1.2 es llamado mejor aproximación (o mejor aproximante) a v en W.

En lo que resta, denotaremos el espacio vectorial de las funciones continuas sobre un intervalo cerrado $[a, b]$ por $C[a, b]$. Asimismo, denotaremos por $\mathbb{R}_n[x]$ al espacio de polinomios de grado menor o igual a n con coeficientes reales.

Teniendo en consideración lo anterior, se tiene que el teorema 1.1.2 asegura que, dada $f \in C[a, b]$, existe un polinomio $p_n^* \in \mathbb{R}_n[x]$ tal que

$$\| f - p_n^* \| \leq \| f - p \|, \quad \text{para todo } p \in \mathbb{R}_n[x],$$

donde $\| \cdot \|$ es la norma uniforme sobre el intervalo $[a, b]$, ésto es,

$$\|g\| = \max_{a \leq x \leq b} |g(x)|, \quad g \in C[a, b] \tag{1.4}$$

Denotaremos como $E_n(f; [a, b]) = E_n(f) = \| f - p_n^* \|$ al error cometido al aproximar f por elementos de $\mathbb{R}_n[x]$.

1.2. Caracterización del mejor aproximante (Teorema de Chebyshev)

Entramos ahora al estudio de las propiedades de los polinomios $p_n^* \in \mathbb{R}_n[x]$ de mejor aproximación a una función continua dada f sobre el intervalo $[a,b]$.

Sea $e = f - p_n^*$; tenemos $\|e\| = E_n(f;[a,b])$.

Definición 1.2.1. Un conjunto de $k+1$ puntos distintos $x_0,\ldots,x_k$ que satisfacen $a \leq x_0 < x_1 < \cdots < x_{k-1} < x_k \leq b$ es llamado un conjunto alternante para la función de error $f - p_n$ si

$$\begin{aligned} |f(x_j) - p_n(x_j)| = \| f - p_n \|, \quad j = 0,1,\ldots,k \quad \text{y} \\ [f(x_j) - p_n(x_j)] = -[f(x_{j+1}) - p_n(x_{j+1})], \quad j = 0,1,\ldots,k-1. \end{aligned} \tag{1.5}$$

Teorema 1.2.2. Sea $f \in C[a,b]$; $p_n^* \in \mathbb{R}_n[x]$ es un mejor aproximante para f sobre $[a,b]$ si y sólo si existe un conjunto alternante para $e = f - p_n^*$ de $n+2$ puntos.

Demostración. Supongamos que $x_0, x_1, \ldots, x_{n+1}$ forman un conjunto alternante para $f - p_n^*$. Se mostrará que p_n^* es un mejor aproximante. Supongamos que no lo es, entonces existe $q_n \in \mathbb{R}_n[x]$ tal que

$$\| f - q_n \| < \| f - p_n^* \|, \tag{1.6}$$

en particular, como $x_0, x_1, \ldots, x_{n+1}$ forman un conjunto alternante, entonces

$$|f(x_j) - q_n(x_j)| < \| f - p_n^* \| = |f(x_j) - p_n^*(x_j)|, \quad j = 0,1,\ldots,n+1. \tag{1.7}$$

Se tiene que (1.7) y (1.5) implican que la diferencia

$$[f(x_j) - p_n^*(x_j)] - [f(x_j) - q_n(x_j)] \tag{1.8}$$

alterna de signo cuando j va de 0 a $n+1$. Veamos cómo sucede ésto, supongamos que para $j=0$, $f(x_j) - p_n^*(x_j) > 0$, entonces debido a (1.7) la diferencia (1.8) es mayor que cero, después para $j=1$, $f(x_j) - p_n^*(x_j) < 0$ por (1.5), y una vez más por (1.7) la diferencia (1.8) es menor que cero, esta alternancia se va presentando hasta $j = n+1$. La alternancia de signos se presenta de manera similar cuando $f(x_j) - p_n^*(x_j) < 0$ para $j = 0$. Ahora bien,

$$[f(x_j) - p_n^*(x_j)] - [f(x_j) - q_n(x_j)] = q_n(x_j) - p_n^*(x_j).$$

Así que $q_n - p_n^* \in \mathbb{R}_n[x]$ tiene un cero en cada intervalo de la forma $(x_j, x_{j+1}),\ j = 0, 1, \ldots, n$ para un total de $n+1$ ceros, luego $q_n - p_n^*$ debe ser el polinomio cero, lo cual implica que $q_n = p_n^*$. Esto contradice (1.6), por lo tanto p_n^* es un mejor aproximante.

Recíprocamente, supongamos que $p_n^* \in \mathbb{R}_n[x]$ es un mejor aproximante para f.
Hagamos

$$\rho = \| f - p_n^* \|,$$

y seleccionemos $\epsilon > 0$ suficientemente pequeño de tal manera que

$$|x_1 - x_2| < \epsilon, \quad \text{implique}$$

$$|e(x_1) - e(x_2)| \leq \frac{1}{2}\rho$$

para $x_1, x_2 \in [a, b]$, esto es posible por la continuidad uniforme de la función de error e (e es una función continua sobre un intervalo cerrado).

Dividamos $[a, b]$ en intervalos cerrados consecutivos de longitud menor o igual a ϵ. Denotemos por $I_1, I_2, \ldots, I_m$, a los intervalos en los cuales $|e(x)|$ alcanza su máximo valor. Como e puede variar a lo más $\frac{1}{2}\rho$ en cualquiera de estos intervalos, entonces

$$e(x) > \frac{1}{2}\rho \quad \text{ó} \quad e(x) < -\frac{1}{2}\rho,$$

en $I_1, I_2, \ldots, I_m$.

Consideremos la sucesión $u_1, u_2, \ldots, u_m$ formada por valores 1 ó -1 de acuerdo al signo que presente e sobre los intervalos $I_1, I_2, \ldots, I_m$, por ejemplo, si e toma el valor $-\rho$ sobre I_j entonces $u_j = -1$.

Debemos demostrar que en esta sucesión se presentan al menos $n + 1$ cambios de signo, con lo cual podemos formar un conjunto alternante para $e = f - p_n^*$ de al menos $n+2$ puntos. Haremos ésto mostrando que si hubieran menos de $n + 1$ cambios, podríamos encontrar un polinomio para el cual el error es menor que ρ.

Si $u_1 = u_2 = \cdots = u_m$ sumando una constante apropiada a p_n^* obtendríamos una mejor aproximación, para ver ésto supongamos que $u_i = 1$, $i = 1, 2, \ldots, m$, es decir, $e(x) > -\rho$ sobre todo $[a, b]$; entonces

$$\min_{a \leq x \leq b} e(x) = M > -\rho,$$

hagamos

$$c = \frac{\rho + M}{2} > 0.$$

Tomando $q_n = p_n^* + c,\ f(x) - q_n(x) = f(x) - p_n^*(x) - c = e(x) - c$ se tiene

$$-(\rho - c) = c - \rho = M - c \leq e(x) - c \leq \rho - c,$$

por tanto $\| f - q_n \| = \rho - c$. Lo que significa que q_n es mejor que p_n^* al aproximar f. El caso $u_i = -1$ se establece de manera similar.

Agrupemos entonces los intervalos consecutivos $I_1, I_2, \ldots, I_m$ en grupos donde el signo de las $u_i,\ i = 1, 2, \ldots, m$ sea el mismo:

$$\begin{array}{ll} \text{Grupo 1 } \{I_1, I_2, \ldots, I_{j_1}\} & \text{primer signo} \\ \text{Grupo 2 } \{I_{j_1+1}, I_{j_1+2}, \ldots, I_{j_2}\} & \text{segundo signo} \\ \text{Grupo 3 } \{I_{j_2+1}, I_{j_2+2}, \ldots, I_{j_3}\} & \text{primer signo} \\ \quad \vdots & \quad \vdots \\ \text{Grupo k } \{I_{j_{k-1}+1}, I_{j_{k-1}+2}, \ldots, I_{j_k} = I_m\}. & \end{array} \tag{1.9}$$

Este esquema muestra $k - 1$ cambios de signo, supongamos entonces que se presentan menos de $n + 1$ cambios, es decir,

$$k - 1 < n + 1.$$

Consideremos los intervalos I_{j_1}, I_{j_1+1}. Estos intervalos son disjuntos pues para uno de ellos $e(x) > \frac{1}{2}\rho$ y para el otro $e(x) < -\frac{1}{2}\rho$. Por tanto, podemos encontrar un x_1 que cumpla:

$$s < x_1 < t \text{ para todo } s \in I_{j_1} \text{ y } t \in I_{j_1+1}. \tag{1.10}$$

De igual manera existen $x_2, \ldots, x_{k-1}$ tales que

$$\begin{array}{ll} s < x_2 < t & \text{para todo } s \in I_{j_2} \text{ y } t \in I_{j_2+1} \\ \quad \vdots & \\ s < x_{k-1} < t & \text{para todo } s \in I_{j_{k-1}} \text{ y } t \in I_{j_{k-1}+1}. \end{array} \tag{1.11}$$

Hagamos
$$q(x) = (x_1 - x)(x_2 - x) \cdots (x_{k-1} - x).$$
Como $k - 1 < n + 1,\ k - 1 \leq n$ luego $q \in \mathbb{R}_n[x]$. Las raíces del polinomio q son los valores x_i, los cuales no pertenecen a ninguno de los intervalos $I_1, I_2, \ldots, I_m$.

Supongamos que la función e toma el valor ρ sobre I_1, por construcción e tomará este valor sobre el grupo 1 de intervalos en (1.9), sobre el grupo 2 tomará entonces el valor $-\rho$ y así alternadamente. Ahora bien, sobre el primer grupo de intervalos en (1.9) el valor de $q(x) = (x_1 - x)(x_2 - x) \cdots (x_{k-1} - x)$ es positivo, pues en vista de (1.10) el valor de todos sus factores es positivo. Sobre el segundo grupo el valor de q es negativo pues únicamente su primer factor es negativo (debido a (1.10)) y los restantes positivos (por (1.11)), sobre el tercero el valor de q es positivo ya que el valor de sus primeros dos factores es negativo y del resto positivo, así sucesivamente se presenta la alternancia de signo en q.

Observamos entonces que las evaluaciones de e y q coinciden en signo sobre los intervalos $I_1, I_2, \ldots, I_m$. Cuando e toma el valor de $-\rho$ sobre I_1, multiplicando q por -1 se tiene el mismo resultado.

Sea $T = [a, b] - I_1 - I_2 - \cdots - I_m$;
$$\max_{x \in T} |e(x)| = \rho' < \rho,$$
supongamos que $\|q\| = M$ y que $\lambda > 0$ se escoge tan pequeño como se requiera para hacer válida la desigualdad
$$\lambda M < \min(\rho - \rho', \frac{\rho}{2}).$$

Mostraremos que $p = \lambda q + p_n^* \in \mathbb{R}_n[x]$ aparece como mejor aproximante para f, en lugar de p_n^*.

Si $x \in T$ entonces
$$|f(x) - p(x)| = |e(x) - \lambda q(x)| \leq |e(x)| + \lambda |q(x)| \leq \rho' + \lambda M < \rho. \quad (1.12)$$

Por otro lado, si $x \notin T$ entonces $x \in I_j$ para algún j entre 1 y m, aquí $e(x)$ y $\lambda q(x)$ tienen el mismo signo, y $|e(x)| \geq \frac{\rho}{2} > \lambda M \geq |\lambda q(x)|$, así tenemos

$$|f(x) - p(x)| = |e(x) - \lambda q(x)| = |e(x)| - \lambda|q(x)| \leq \rho - |\lambda q(x)| < \rho$$

que junto con (1.12) implican

$$\| f - p \| < \rho = \| f - p_n^* \|,$$

lo cual es una contradicción, por lo tanto $k-1 \geq n+1$, con lo que el teorema queda demostrado.

$\square$

1.3. Unicidad del mejor aproximante

Una vez que tenemos los teoremas previos, es decir, el de existencia y el de caracterización, lo que resta es contar con un resultado que nos asegure la unicidad del mejor aproximante.

Teorema 1.3.1. Si p_n^* es una mejor aproximación a $f \in C[a,b]$ en norma uniforme, $f \notin \mathbb{R}_n[x]$, entonces p_n^* es único; ésto es, si $p \in \mathbb{R}_n[x]$ y $p \neq p_n^*$, $\| f - p \| > \| f - p_n^* \|$.

Demostración. Supongamos que $p \in \mathbb{R}_n[x]$ es también un mejor aproximante para f, entonces $\| f - p \| = \| f - p_n^* \| = E_n(f)$ y

$$q = \frac{p + p_n^*}{2}$$

es también un mejor aproximante para f, pues

$$\begin{aligned} E_n(f) \quad & \leq \| f - q \| = \| f - \tfrac{p+p_n^*}{2} \| = \| \tfrac{1}{2}(f-p) + \tfrac{1}{2}(f - p_n^*) \| \\ & \leq \tfrac{1}{2}(\| f - p \| + \| f - p_n^* \|) = E_n(f). \end{aligned}$$

Sea $x_0, x_1, \ldots, x_{n+1}$ un conjunto alternante para $f - q$, entonces para algún entero l,

$$\begin{aligned} f(x_j) - q(x_j) \quad & = \tfrac{f(x_j) - p(x_j)}{2} + \tfrac{f(x_j) - p_n^*(x_j)}{2} \\ & = (-1)^{l+j} E_n(f), \quad j = 0, 1, \ldots, n+1, \end{aligned}$$

o bien,

$$f(x_j) - p(x_j) + f(x_j) - p_n^*(x_j) = 2(-1)^{l+j} E_n(f), \quad j = 0, 1, \ldots, n+1$$

lo cual implica que

$$f(x_j) - p(x_j) = f(x_j) - p_n^*(x_j) = (-1)^{l+j} E_n(f), \quad j = 0, 1, \ldots, n+1,$$

pues

$$|f(x_j) - p(x_j)| \leq E_n(f) \quad \text{y} \quad |f(x_j) - p_n^*(x_j)| \leq E_n(f),$$

así obtenemos

$$p(x_j) = p_n^*(x_j) \quad j = 0, 1, 2, \ldots, n+1,$$

luego

$$p = p_n^*.$$

Por lo tanto p_n^* es único. $\square$

Capítulo 2

Aproximación polinomial con restricciones

2.1. Conceptos preliminares

En este capítulo nos enfocaremos al estudio de la aproximación con restricciones, es decir, a la mejor aproximación se le pedirán condiciones adicionales. Para ser más precisos, la aproximación se hará mediante polinomios y la condición adicional que se le pedirá a la mejor aproximación es que su rango esté restringido.

Definición 2.1.1. Sea $X \subset \mathbb{R}$, donde $\mathbb{R}$ denota a los números reales y X es compacto. Se dice que $W : X \times \mathbb{R} \to \mathbb{R}$, es una función de peso generalizada si

(a) $sgn\, W(x, y) = sgn\ y$ (en particular, $W(x, 0) = 0$),

(b) W es continua,

(c) Para cada $x \in X$, W es una función monótona estrictamente creciente con respecto a y, con $\lim\limits_{|y| \to \infty} |W(x, y)| = \infty$.

$$\tag{2.1}$$

Algunos ejemplos de funciones de peso generalizadas que son de interés para problemas de aproximación son los siguientes:

(1) $W(x, y) = y$.

(2) $W(x,y) = \frac{y}{|f(x)|}$, donde f es una función de valores reales, con $f(x) \neq 0$ para toda $x \in X$.

(3) $W(x,y) = \begin{cases} y^2 & si \quad y > 0, \\ y & si \quad y \leq 0. \end{cases}$

En donde los errores correspondientes a las funciones de peso (1) y (2) se reducen a los casos del error uniforme (o de Chebyshev) y el error relativo, respectivamente.

En lo sucesivo, denotaremos con L (conjunto de aproximantes) a un subconjunto no vacío del conjunto de todas las funciones reales definidas sobre X.

Definición 2.1.2. Sea $P \in \mathbb{R}_n[x]$ y f una función de valores reales definida sobre X. Se dice que P es una aproximación a f con respecto a la función de peso generalizada W, si

$$\sup_{x \in X} |W[x, f(x) - P(x)]| < \infty.$$

Observación 2.1.3. Si en la definición, 2.1.2 consideramos que

- las funciones f, P y W son continuas,
- $X \subset \mathbb{R}$ compacto,

entonces se obtine $\sup\limits_{x \in X} |W[x, f(x) - P(x)]| < \infty$, es decir, simpre podemos asegurar la existencia de una aproximación a f.

Definición 2.1.4. Sea f una función de valores reales definida sobre X, se dice que $h \in L$ es una mejor aproximación (o mejor aproximante) generalizada a f con respecto a W y L si

$$\sup_{x \in X} |W[x, f(x) - h(x)]| \leq \inf_{g \in L} \sup_{x \in X} |W[x, f(x) - g(x)]| \, . \qquad (2.2)$$

En el caso en el que $L = \mathbb{R}_n[x]$ y f es una función continua, se muestra en [10] que la mejor aproximación existe y es única. Además, se da un algoritmo con el cual podemos encontrar la mejor aproximación generalizada con el apoyo de una computadora. Cuando $X = [a, b]$ y $W(x, y) = y$ se obtiene el teorema de Chebyshev, estudiado en el capítulo 1.

Como hemos mencionado anteriormente, en el presente trabajo, se abordará el tema referente a aproximaciones con rango restringido, esto es, para u y l dos funciones continuas de valores reales dadas y definidas sobre X, tales que $u(x) \geq l(x)$ para todo $x \in X$, definimos a $\mathbb{R}_n^*[x]$ de la siguiente manera:

$$\mathbb{R}_n^*[x] = \{P \in \mathbb{R}_n[x] : \ l(x) \leq P(x) \leq u(x) \text{ para todo } x \in X\}. \qquad (2.3)$$

Donde u y l se eligen de tal manera que se puede garantizar que $\mathbb{R}_n^*[x] \neq \varnothing$. Sea W una función de peso generalizada, consideraremos ahora (2.2) cuando $L = \mathbb{R}_n^*[x]$ y f es una función continua para la cual se cumple que $l(x) \leq f(x) \leq u(x)$.

Notemos que las funciones u y l se pueden manipular, con lo que podemos cambiar la clase $\mathbb{R}_n^*[x]$ de aproximaciones polinomiales. Por ejemplo, si $u(x) = M$, donde M es lo suficientemente grande y $l(x) = f(x)$, tendremos aproximaciones unilaterales a f.

También, cabe mencionar, que si X es un intervalo cerrado y acotado, y definimos las funciones u y l de la siguiente forma:

$$u(x) = \sup_{y \in X} f(y) \;\text{ y }\; l(x) = \inf_{y \in X} f(y).$$

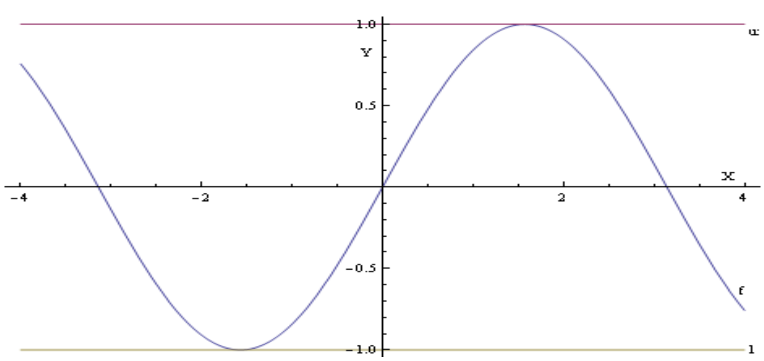

Figura 2.1: Ejemplo de gráfica de las funciones f, u y l.

entonces, podemos aproximarnos a f mediante polinomios cuyos rangos estén contenidos en el rango de f.

2.2. Existencia del mejor aproximante generalizado a f

Se abordará ahora el tema referente a la existencia de una mejor aproximación generalizada, por lo que en esta sección y en lo que resta del capítulo, asumiremos lo siguiente

(i) f es una función real continua definida sobre X,

(ii) $L = \mathbb{R}_n^*[x]$,

(iii) W^* es una función real de peso generalizada y

(iv) $X \subset \mathbb{R}$ compacto.

Antes de enunciar el teorema de existencia de una mejor aproximación a la función f con respecto a W^* y $\mathbb{R}_n^*[x]$, citaremos el siguiente teorema tomado de [9], asimismo, enunciaremos un lema, ya que en ellos nos apoyaremos para poder asegurar la existencia de una mejor aproximación.

Cabe mencionar también, que consideraremos que $l(x) \leq f(x) \leq u(x)$.

Teorema 2.2.1. Si $W(x, y)$ satisface las siguientes condiciones

(a) Para cada $x \in X$, $W(x, y)$ es una función monótona no decreciente de y,

(b) $sgn\ W(x, y) = sgn\ y$ para todo (x, y),

(c) $\lim\limits_{|y|\to\infty} |W(x, y)| = \infty$ para cada x,

(d) para cada $x \in X$ y $t \neq 0$, $\lim\limits_{y\to t,\ |y|<|t|} W(x, y) = W(x, t)$,

y existe una aproximación a f en $\mathbb{R}_n[x]$; entonces existe una mejor aproximación generalizada a f con respecto a W y $\mathbb{R}_n[x]$.

Lema 2.2.2. Si W^* es una función de peso generalizada y definimos a

$W : X \times \mathbb{R} \to [-\infty, \infty]$ de la siguiente manera

$$W(x, y) = \begin{cases} +\infty & \text{si} \quad y > f(x) - l(x), \\ W^*(x, y) & \text{si} \quad l(x) \leq f(x) - y \leq u(x), \\ -\infty & \text{si} \quad y < f(x) - u(x). \end{cases} \tag{2.4}$$

entonces W satisface las condiciones del teorema (2.2.1).

Demostración. Antes de iniciar con la prueba, se muestra una figura en la cual se puede apreciar las regiones de evaluación de W.

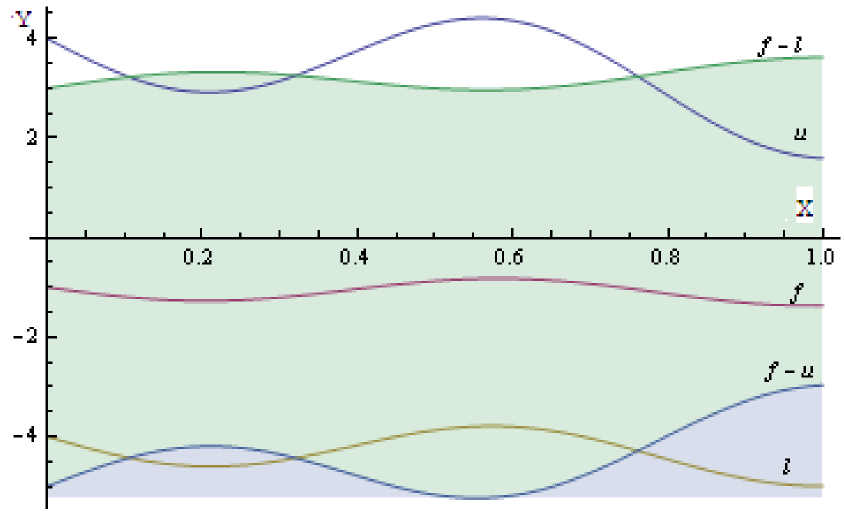

Figura 2.2: Gráfica de las regiones de evaluación de W.

Notemos, que para $y = 0$, $W(x, 0) = 0$ puesto que f cumple con

$$l(x) \leq f(x) \leq u(x), \quad \text{o lo que es lo mismo,}$$

$$l(x) \leq f(x) - 0 \leq u(x) \quad \text{para cada } x \in X,$$

con lo que se tiene que $W(x, 0) = W^*(x, 0)$, luego, como W^* es una función de peso generalizada, $W^*(x, 0) = 0$.

Probemos (a).
Sean $y_1,\ y_2 \in \mathbb{R}$ tales que $y_1 < y_2$, luego, se tienen los siguientes casos:

1. Para
$$y_2 > f(x) - l(x),$$
se cumple que
$$W(x, y_2) = +\infty,$$
para toda $x \in X$, de donde se sigue que
$$W(x, y_1) \leq W(x, y_2).$$

2. Para
$$l(x) \leq f(x) - y_2 \leq u(x)$$
se tiene que, $W(x, y_2) = W^*(x, y_2)$. Como $y_1 < y_2 \leq f(x) - l(x)$, entonces se tienen los siguientes dos casos referentes a y_1:

   - Si $l(x) \leq f(x) - y_1 \leq u(x)$, entonces $W(x, y_1) = W^*(x, y_1)$, además $W(x, y_2) = W^*(x, y_2)$ y como W^* es estrictamente monótona, entonces dado que $y_1 < y_2$, se sigue que $W^*(x, y_1) < W^*(x, y_2)$. Por lo tanto $W(x, y_1) < W(x, y_2)$.
   - Si $y_1 < f(x) - u(x)$, entonces $W(x, y_1) = -\infty$, de donde se sigue que $W(x, y_1) \leq W(x, y_2)$.

3. Por último, $y_2 < f(x) - u(x)$, implica $y_1 < f(x) - u(x)$. En consecuencia,
$$W(x, y_1) = -\infty = W(x, y_2).$$

Luego, de todo lo anterior, se sigue que $W(x, y_1) \leq W(x, y_2)$ con lo que $W(x, y)$ es una función monótona no decreciente de y.

Probemos (b).
Al igual que en (a), la prueba se hará por casos:

1. Supongamos que
$$y > f(x) - l(x).$$
Como $l(x) \leq f(x) \leq u(x)$, entonces $0 \leq f(x) - l(x)$. De donde
$$y > f(x) - l(x) \geq 0 \text{ y } sgn\, W(x, y) = sgn\, (+\infty) = sgn\, y.$$

2. Supongamos ahora que
$$l(x) \leq f(x) - y \leq u(x),$$
luego
$$W(x, y) = W^*(x, y)$$
y como W^* es una función de peso generalizada, obtenemos
$$sgn\, W(x, y) = sgn\, W^*(x, y) = sgn\, y.$$

3. Supongamos que
$$y < f(x) - u(x),$$
entonces $W(x,y) = -\infty$.

   Por otro lado, como
$$l(x) \leq f(x) \leq u(x),$$
entonces $y < f(x) - u(x) \leq 0$. Por lo tanto
$$sgn\ W(x,y) = sgn\ (-\infty) = sgn\ y.$$

Así,
$$sgn\ W(x,y) = sgn\ y; \ \text{ para todo } (x,y).$$

Probemos (c)

1. Si
$$y \to +\infty,$$
entonces existe $y_1 \in \mathbb{R}$, digamos, $y_1 = f(x) - l(x)$ tal que si $y > y_1$, entonces $y > f(x) - l(x)$, de donde se sigue que
$$\lim_{y\to\infty} W(x,y) = +\infty.$$

2. Por último, si
$$y \to -\infty,$$
entonces existe $y_2 \in \mathbb{R}$, digamos $y_2 = f(x) - u(x)$, tal que si $y < y_2$, entonces $y < f(x) - u(x)$, de donde se sigue que
$$\lim_{y\to-\infty} W(x,y) = -\infty.$$

Con lo que
$$\lim_{|y|\to\infty} |W(x,y)| = \infty.$$

Por último, probemos (d).

Para la demostración de este inciso, tomaremos en cuenta los siguientes casos

1. Supongamos que $t > f(x) - l(x)$. Como $l(x) \leq f(x)$ para cada $x \in X$, se sigue que $t > 0$.

   Tomemos $N_x = t - (f(x) - l(x))$ y $\delta = \frac{N_x}{2}$.

   Sea $y \in X$ tal que $|y - t| < \delta$ y $y < t$, con lo que $t - y < \delta$, o bien

   $$t - y < 2\delta - \delta$$

   luego,

   $$t - y < t - (f(x) - l(x)) - \frac{N_x}{2},$$

   $$y > (f(x) - l(x)) + \frac{N_x}{2} > f(x) - l(x)$$

   así, $y > f(x) - l(x)$ y por tanto, $W(x, y) = +\infty = W(x, t)$.

   De donde,

   $$\lim_{y \to t,\ y < t} W(x, y) = W(x, t).$$

2. Supongamos ahora que $l(x) \leq f(x) - t \leq u(x)$.

   Sea $A_x = t - (f(x) - u(x))$ y $B_x = (f(x) - l(x)) - t$. Definimos a C_x de la siguiente manera:

   $$C_x = \begin{cases} A_x & \text{si} \quad B_x = 0, \\ B_x & \text{si} \quad A_x = 0, \\ \min\{A_x, B_x\} & \text{si} \quad A_x, B_x > 0. \end{cases} \tag{2.5}$$

   Tomemos

   $$\delta = \frac{C_x}{2}.$$

   Sea $y \in X$ tal que $|y - t| < \delta$

   luego, $-\delta < t - y < \delta$.

   Sustituyendo el valor de δ, obtenemos

   $$-\frac{C_x}{2} < t - y < \frac{C_x}{2},$$

   por (2.5)

   $$-B_x < t - y < A_x,$$

luego

$$t-(f(x)-l(x))<t-y<t-(f(x)-u(x)),$$

$$-(f(x)-l(x))<-y<-(f(x)-u(x))$$

sumando $f(x)$ obtenemos

$$f(x)-(f(x)-l(x))<f(x)-y<f(x)-(f(x)-u(x))$$

luego,

$$l(x)<f(x)-y<u(x).$$

Por lo tanto, $W(x,y)=W^*(x,y)$ y por la continuidad de la función W^*, se sigue que $\lim\limits_{y\to t,\ |y|<|t|} W^*(x,y)=W^*(x,t)=W(x,t)$.

3. Por último, supongamos que $t<f(x)-u(x)$. Al ser $f(x)\leq u(x)$ para cada $x\in X$, se sigue que $t<0$. Sea $L_x=(f(x)-u(x))-t$, $\delta=\frac{L_x}{2}$ y $y\in X$ tal que $|y-t|<\delta$, entonces

$$-\delta<y-t<\delta.$$

Sustituyendo el valor de δ

$$-\frac{L_x}{2}<y-t<\frac{L_x}{2},$$

$$-L_x<y-t<L_x,$$

sustituyendo el valor de L_x,

$$t-(f(x)-u(x))<y-t<(f(x)-u(x))-t,$$

$$2t-(f(x)-u(x))<y<f(x)-u(x),$$

así,

$$y<f(x)-u(x).$$

Por lo tanto $W(x,y)=W(x,t)=-\infty$, de donde,

$$\lim_{y\to t,\ |y|<|t|} W(x,y)=W(x,t)$$

De los casos analizados, se concluye que

$$\lim_{y\to t,\ |y|<|t|} W(x,y)=W(x,t).$$

Por tanto, W cumple satisface las condiciones del teorema (2.2.1). $\square$

Observación 2.2.3. Cabe mencionar que una mejor aproximación a f en $\mathbb{R}_n[x]$ es equivalente a una mejor aproximación generalizada a f con respecto a W y $\mathbb{R}_n[x]$.

Teorema 2.2.4. Sea f una función continua sobre X. Si $W^*(x, y)$ es una función de peso generalizada y $\mathbb{R}_n^*[x]$ es la clase descrita en (2.3), entonces existe una mejor aproximación a f con respecto a W^* y $\mathbb{R}_n^*[x]$.

Demostración. Dada la función de peso generalizada W^*, definimos

$$W : X \times \mathbb{R} \to [-\infty, \infty] \text{ como sigue:}$$

$$W(x,y) = \begin{cases} +\infty & \text{si} \quad y > f(x) - l(x), \\ W^*(x,y) & \text{si} \quad l(x) \leq f(x) - y \leq u(x), \\ -\infty & \text{si} \quad y < f(x) - u(x). \end{cases} \tag{2.6}$$

Por el lema 2.2.2, se sigue que la función de peso generalizada W cumple con los requisitos que se piden en el teorema 2.2.1 y por la observación 2.1.3 existe una aproximación a f en $\mathbb{R}_n[x]$, entonces existe una mejor aproximación $P^* \in \mathbb{R}_n[x]$, es decir ,

$$\sup_{x\in X} |W[x, f(x) - P^*(x)]| \leq \inf_{R\in\mathbb{R}_n[x]} \sup_{x\in X} |W[x, f(x) - R(x)]|.$$

Sea $P_1 \in \mathbb{R}_n^*[x] \subset \mathbb{R}_n[x]$, luego,

$$\sup_{x\in X} |W[x, f(x) - P^*(x)]| \leq \sup_{x\in X} |W[x, f(x) - P_1(x)]| \tag{2.7}$$

ya que $P^* \in \mathbb{R}_n[x]$ es una mejor aproximación y $P_1 \in \mathbb{R}_n[x]$.

Como $P_1 \in \mathbb{R}_n^*[x]$, entonces

$$l(x) \leq P_1(x) \leq u(x),$$

luego

$$l(x) \leq f(x) - (f(x) - P_1(x)) \leq u(x).$$

Tomando $y = f(x) - P_1(x)$ en (2.6), se obtiene $W(x, y) = W^*(x, y) < \infty$, así

$$\sup_{x\in X} |W[x, f(x) - P_1(x)]| = \max_{x\in X} |W^*[x, f(x) - P_1(x)]| < \infty, \tag{2.8}$$

puesto que W^* es continua.

Por (2.7),

$$\sup_{x\in X} |W[x, f(x) - P^*(x)]| \leq \max_{x\in X} |W^*[x, f(x) - P_1(x)]| < \infty$$

luego,

$$\sup_{x\in X} |W[x, f(x) - P^*(x)]| < \infty \quad \text{y}$$

$$W[x, f(x) - P^*(x)] = W^*[x, f(x) - P^*(x)] \tag{2.9}$$

tomando

$$y_1 = f(x) - P^*(x)$$

se cumple que

$$l(x) \leq f(x) - (f(x) - P^*(x)) \leq u(x),$$

con lo que

$$l(x) \leq P^*(x) \leq u(x) \text{ para toda } x \in X.$$

Luego, por (2.9), (2.8) y la desigualdad (2.7), se tiene que

$$\sup_{x\in X} |W^*[x, f(x) - P^*(x)]| \leq \sup_{x\in X} |W^*[x, f(x) - P_1(x)]|.$$

Por lo tanto, $P^* \in \mathbb{R}_n^*[x]$ es la mejor aproximación generalizada para f con respecto a W^* y $\mathbb{R}_n^*[x]$. Con lo cual se prueba la existencia de una mejor aproximación a f con respecto a W^* y $\mathbb{R}_n^*[x]$.

$\square$

2.3. Caracterización del mejor aproximante

Para probar la unicidad debemos asumir que $l(x) < u(x)$ para toda $x \in X$. Entonces la unicidad de la mejor aproximación generalizada de f con respecto de W^* y $\mathbb{R}_n^*[x]$ se sigue de la caracterización de una mejor aproximación generalizada en términos de $n+2$ puntos como en el caso usual de Chebyshev.

Antes de proceder al enunciado y prueba del teorema referente a la caracterización del mejor aproximante, daremos algunas definiciones que nos serán de utilidad para tal efecto.

Definición 2.3.1. Sea $P \in \mathbb{R}_n^*[x]$. Definimos $sgn^*\left[f(x) - P(x)\right]$ por

$$sgn^*\left[f(x) - P(x)\right] = \begin{cases} -1 & \text{si} \quad P(x) = u(x), \\ +1 & \text{si} \quad P(x) = l(x), \\ sgn[f(x) - P(x)] & \text{en otro caso.} \end{cases} \tag{2.10}$$

Definición 2.3.2. Sea $P \in \mathbb{R}_n^*[x]$. Se dice que la curva de error $W^*[x, f(x) - P(x)]$ alterna $n+1$ veces sobre X si existen $n+2$ puntos, $x_1, x_2, \ldots, x_{n+2}$ en X, tales que $x_1 < x_2 < \cdots < x_{n+2}$, con

$$sgn^*[f(x_i) - P(x_i)] = -sgn^*\left[f(x_{i+1}) - P(x_{i+1})\right], \quad i = 1, 2, \ldots, n+1,$$

y cada x_i cumple al menos una de las siguientes condiciones:

a) $|W[x_i, f(x_i) - P(x_i)]| = \max\limits_{x \in X} |W[x, f(x) - P(x)]|$,
b) $P(x_i) = l(x_i)$,
c) $P(x_i) = u(x_i)$.

Nos referiremos a cada x_i como *punto extremo*.
Si $sgn^*[f(x_i) - P(x_i)] = +1$, entonces x_i es llamado *extremo positivo*, y si $sgn^*[f(x_i) - P(x_i)] = -1$, x_i es llamado un *extremo negativo*.

Antes de proceder, citaremos el siguiente lema adaptado de [12], el cual nos será de utilidad en lo sucesivo.

Lema 2.3.3. Si $P \in \mathbb{R}_n[x]$ y $x_1, x_2, \ldots, x_k \in X$ con $k \le n$; entonces dado $\epsilon > 0$ existe $P^* \in \mathbb{R}_n[x]$ tal que

(i) $\max_{x \in X} |P(x) - P^*(x)| < \epsilon$,

(ii) $P^*(x_i) = P(x_i), \quad i = 1, \ldots, k$, y

(iii) $sgn[P^*(x) - P(x)] = (-1)^i$ para $x \in X \cap (x_i, x_{i+1}), \quad i = 0, 1, \ldots, k$, donde $x_0 = a = \text{ínf}\,\{x : x \in X\}$ y $x_{k+1} = b = \sup\{x \ : \ x \in X\}$.

Teorema 2.3.4. Sea f una función continua sobre $X \subset \mathbb{R}$, W^* una función de peso generalizada, $l(x) < u(x)$ para todo $x \in X$ y $\mathbb{R}_n^*[x]$ la clase descrita en (2.3). Entonces $P^* \in \mathbb{R}_n^*[x]$ es la mejor aproximación generalizada para f con respecto a W^* y $\mathbb{R}_n^*[x]$, si y sólo si la curva de error $W^*[x, f(x) - P^*(x)]$ alterna al menos $n+1$ veces sobre X.

Demostración. La prueba de la parte necesaria se hará por contradicción. Para lo cual, supongamos que $P^* \in \mathbb{R}_n^*[x]$ es la mejor aproximación generalizada para f con respecto a W^* y $\mathbb{R}_n^*[x]$, y que $W^*[x, f(x) - P^*(x)]$ alterna $k < n+1$ veces sobre X.

Como $W^*[x, f(x) - P^*(x)]$ alterna k veces, entonces por definición existen $k+1$ puntos extremos en X, digamos $y_1, y_2, \ldots, y_{k+1}$ con $y_1 < y_2 < \cdots < y_{k+1}$, para esta curva de error. Consideremos i, con $1 \leq i \leq k$ fijo y supongamos que y_i es un extremo positivo. Sean

$$y' = \sup\{y \ : \ y \in X \ \text{ con } \ y_i \leq y \leq y_{i+1}, \ y \ \text{ es un extremo positivo}\},$$

$$y'' = \text{ínf}\,\{y \ : \ y \in X \text{ con } y_i \leq y \leq y_{i+1}, \ y \ \text{ es un extremo negativo}\}.$$

Como X es compacto y todas las funciones involucradas son continuas, se sigue que y' es un extremo positivo, y'' es un extremo negativo. Además $y' < y''$, puesto que si $y' = y''$, tendríamos que y' es un extremo positivo y negativo al mismo tiempo, asimismo, si $y'' < y'$ entonces $W^*[x, f(x) - P^*(x)]$ alternaría más de k veces. Por otro lado, si $y \in (y', y'')$ entonces y no es un punto extremo por la manera en que hemos elegido a y' y y''.

Para el caso en el que y_i sea un extremo negativo, consideraremos

$$y' = \sup\{y \ : \ y \in X \ \text{ con } \ y_i \leq y \leq y_{i+1}, \ y \ \text{ es un extremo negativo}\},$$

$$y'' = \text{ínf}\,\{y \ : \ y \in X \text{ con } \ y_i \leq y \leq y_{i+1}, \ y \ \text{ es un extremo positivo}\}.$$

Una vez más, por la compacidad de X y la continuidad de cada una de las funciones que estamos considerando, se sigue que y' es un extremo negativo y y'' es un extremo positivo con $y' < y''$.

Ahora, si $X \cap (y', y'') \neq \varnothing$, elegimos $x_i \in X \cap (y', y'')$ arbitrario, x_i no es punto extremo, luego, en x_i tenemos:

$$\begin{aligned} &(i)\ l(x_i) < P^*(x_i) < u(x_i), \\ &(ii)\ \ |W^*[x_i, f(x_i) - P^*(x_i)]| < \rho^*, \end{aligned} \tag{2.11}$$

donde

$$\rho^* = \max_{x \in X} |W^*[x, f(x) - P^*(x)]| \,. \tag{2.12}$$

Por otro lado, si $X \cap (y', y'') = \varnothing$, entonces elegimos

$$x_i = \frac{y' + y''}{2}, \quad \text{para cada } i = 1, 2, \ldots, k.$$

Sea

$$x_0 = \inf\{y \ : \ y \in X\} \ \text{ y } \ x_{k+1} = \sup\{y \ : \ y \in X\}\,.$$

Consideremos ahora los siguientes puntos

$$x_i^* = \inf_{x \in X}\{x \in [x_i, x_{i+1}]\},$$

$$x_{i+1}^* = \sup_{x \in X}\{x \in [x_i, x_{i+1}]\},$$

con $i = 0, 1, 2, \ldots, k$, para construir el siguiente conjunto de intervalos

$$\left\{X \cap [x_i^*, x_{i+1}^*]\right\}, \ \ i = 0, 1, \ldots, k.$$

Con lo que obtenemos una familia de $k+1$ intervalos consecutivos en X, cada uno de estos intervalos, por la misma construcción, no contienen alternancias. Además, $W^*[x, f(x) - P^*(x)]$ alterna exactamente una vez sobre dos intervalos adyacentes cualquiera.

Notemos que $X \cap [x_i^*, x_{i+1}^*]$, con $i \in \{0, 1, \ldots, k\}$ fijo contiene al menos un extremo, digamos y_{i+1}; además, no contiene extremos con signo opuesto al de y_{i+1}, de lo contrario, $W^*[x, f(x) - P^*(x)]$ alternaría más de una vez sobre dos intervalos adyacentes.

Sea $i = 0$, sin pérdida de generalidad, podemos suponer que y_1 es un extremo positivo. Ahora, dado que X es compacto, que todas las funciones que estamos considerando son continuas y $u(x) > l(x)$, existe $\delta_1 > 0$ tal que:

- $W^*[x, f(x) - P^*(x)] + \rho^* > \delta_1$, o bien, $W^*[x, f(x) - P^*(x)] > -\rho^* + \delta_1$ para todo $x \in X \cap [x_0^*, x_1^*]$.
- $P^*(x) + \delta_1 < u(x)$, es decir, $P^*(x) < u(x) - \delta_1$ para todo $x \in X \cap [x_0^*, x_1^*]$.

Sea $\epsilon_1 > 0$ tal que

$$\epsilon_1 < \delta_1/2 \ \text{ y } \ W^*[x, f(x) - P^*(x) - \epsilon_1] \geq -\rho^* + \delta_1/2 \tag{2.13}$$

para todo $x \in X \cap [x_0, x_1]$.

Repitiendo este proceso sobre $X \cap [x_1^*, x_2^*]$, obtenemos $\epsilon_2 > 0$ y $\delta_2 > 0$ tal que

$$\epsilon_2 < \delta_2/2, \ P^*(x) > l(x) + \delta_2 \tag{2.14}$$

y

$$W^*[x, f(x) - P^*(x) + \epsilon_2] \leq \rho^* - \delta_2/2$$

para todo $x \in X \cap [x_1^*, x_2^*]$.

Continuando con este proceso para cada intervalo de la forma

$$X \cap [x_i^*, x_{i+1}^*], \quad i = 0, 1, \ldots, k,$$

obtenemos $\epsilon_1, \epsilon_2, \ldots, \epsilon_k$, donde $\epsilon_i > 0$ para todo $i = 1, 2, \ldots, k$. Sea

$$\epsilon = \text{mín}\{\epsilon_1, \ldots, \epsilon_k\}.$$

Así, se tiene lo siguiente:

- $P^* \in \mathbb{R}_n^*[x] \subset \mathbb{R}_n[x]$,
- $x_1^*, x_2^*, \ldots, x_k^* \in X$,
- $k \leq n$,
- $\epsilon > 0$.

Luego, por el lema 2.3.3, existe $P \in \mathbb{R}_n[x]$ tal que

1. $P(x_i^*) = P^*(x_i^*), \quad i = 1, \ldots, k,$
2. $\max\limits_{x \in X} |P(x) - P^*(x)| < \epsilon$
3. $sgn\,(P(x) - P^*(x)) = (-1)^i$ para $x \in X \cap (x_i^*, x_{i+1}^*), \quad i = 0, 1, \ldots, k.$

Probemos ahora que $P \in \mathbb{R}_n^*[x]$.

Se tiene que para $x \in X \cap [x_0^*, x_1^*]$:

- si $x = x_1^*$, entonces $P(x) = P^*(x)$,
- si $x \neq x_1^*$ entonces $sgn[P(x) - P^*(x)] = +1$, lo que implica que $P(x) > P^*(x)$.

Así, de los dos casos anteriores, se tiene que $P(x) \geq P^*(x)$ y como $l(x) \leq P^*(x)$, entonces $l(x) \leq P(x)$.

También, se tiene que $P(x) \leq u(x) - \delta_1/2 \leq u(x)$ y puesto que

$$\max_{x \in X} |P^*(x) - P(x)| < \epsilon < \delta_1/2,$$

$$P^*(x) < u(x) - \delta_1, \text{ para toda } x \in X \cap [x_0^*, x_1^*],$$

entonces,

$$l(x) \leq P(x) \leq u(x), \text{ para toda } x \in X \cap [x_0^*, x_1^*].$$

Una vez más, por la continuidad de las funciones y el hecho de que
$X \cap [x_0^*, x_1^*]$ es compacto, tenemos

$$\rho^* \geq \max_{x \in X \cap [x_0^*, x_1^*]} W^* [x, f(x) - P^*(x)] > \max_{z \in X \cap [x_0^*, x_1^*]} W^* [z, f(z) - P(z)]$$

de lo anterior junto con (2.13) se tiene que

$$\min_{z \in X \cap [x_0^*, x_1^*]} W^* [z, f(z) - P^*(z) - \epsilon] \geq -\rho^* + \frac{\delta_1}{2} > -\rho^*.$$

Luego,

$$-\rho^* < \max_{z \in X \cap [x_0^*, x_1^*]} W^* [z, f(z) - P(z)] < \rho^*.$$

Entonces $\max_{z \in X \cap [x_0^*, x_1^*]} |W^* [z, f(z) - P(z)]| < \rho^*$.

De manera similar para $x \in X \cap [x_1^*, x_2^*]$:

- si $x = x_1^*$ ó $x = x_2^*$, entonces $P(x) = P^*(x)$,

- si $x \neq x_1^*$ y $x \neq x_2^*$, entonces $sgn[P(x) - P^*(x)] = -1$, lo que implica que $P(x) < P^*(x)$.

De los casos anteriores, podemos concluir que

$$P(x) \leq P^*(x),$$

y como $P^*(x) \leq u(x)$, entonces $P(x) \leq u(x)$.
También, se tiene que $l(x) < l(x) + \delta_2/2 \leq P(x)$ y puesto que

$$\max_{x \in X} |P^*(x) - P(x)| < \epsilon < \delta_2/2,$$

$$l(x) + \delta_2 < P^*(x), \text{ para } \text{ todo } x \in X \cap [x_1^*, x_2^*].$$

Entonces,

$$l(x) \leq P(x) \leq u(x), \text{ para } \text{ todo } x \in X \cap [x_1^*, x_2^*].$$

Por la continuidad de las funciones consideradas y la compacidad de $X \cap [x_1^*, x_2^*]$ se sigue que

$$\begin{aligned} -\rho^* \leq \min_{x \in X \cap [x_1^*, x_2^*]} W^*[x, f(x) - P^*(x)] &\leq \max_{x \in X \cap [x_1^*, x_2^*]} W^*[x, f(x) - P^*(x)] \\ < \max_{z \in X \cap [x_1^*, x_2^*]} W^*[z, f(z) - P(z)] &\leq \max_{z \in X \cap [x_1^*, x_2^*]} W^*[z, f(z) - P^*(z) + \epsilon] \\ &\leq \rho^* - \delta_2/2 < \rho^*. \end{aligned}$$

Por lo tanto, como

$$-\rho^* < \max_{x \in X \cap [x_1^*, x_2^*]} W^*[x, f(x) - P(x)] < \rho^*,$$

entonces

$$\max_{x \in X \cap [x_1^*, x_2^*]} |W^*[x, f(x) - P(x)]| < \rho^*.$$

Repitiendo estos argumentos para cada uno de los intervalos $X \cap [x_i, x_{i+1}]$, con $i = 0, 1, \ldots, k$ obtenemos :

- $P \in \mathbb{R}_n^*[x]$, pues $l(x) \leq P(x) \leq u(x)$ para todo $x \in X$ y

- $\max_{x \in X} |W^*[x, f(x) - P(x)]| < \max_{x \in X} |W^*[x, f(x) - P^*(x)]| = \rho^*$, con lo que tenemos que $P \in \mathbb{R}_n^*[x]$ aproxima mejor a f que P^*, lo cual contradice el hecho de que P^* es la mejor aproximación generalizada de f con respecto a W^* y $\mathbb{R}_n^*[x]$.

Por lo tanto, la curva de error $W^*[x, f(x) - P^*(x)]$ alterna al menos $n+1$ veces sobre X.

Antes de continuar con la demostración, recordemos que, por el teorema del valor intermedio, si g es una función continua en $[x_1, x_{n+1}]$ y es tal que $g(x_1) \geq 0$, $g(x_2) \leq 0, \ldots, (-1)^n g(x_{n+1}) \geq 0$, entonces, existen $y_1, y_2, \ldots, y_n \in [x_1, x_{n+1}]$ tal que:

$$x_1 \leq y_1 \leq x_2 \leq y_2 \leq x_3 \leq \cdots \leq y_n \leq x_{n+1} \;\text{ y }\; g(y_i) = 0.$$

Con ayuda de este resultado, probaremos la parte suficiente.

Supongamos ahora que la curva de error $W^*[x, f(x) - P^*(x)]$ para $P^* \in \mathbb{R}_n^*[x]$ alterna al menos $n+1$ veces. Por simplicidad asumiremos que ésta alterna exactamente $n+1$ veces, supongamos también que existe $P \in \mathbb{R}_n^*[x]$ tal que

$$\max_{x \in X} |W^*[x, f(x) - P(x)]| < \max_{x \in X} |W^*[x, f(x) - P^*(x)]| \,.$$

Sea $\{x_1, x_2, \ldots, x_n, x_{n+1}, x_{n+2}\}$ un conjunto de $n+2$ puntos extremos en X, para la curva de error $W^*[x, f(x) - P^*(x)]$ con $x_i < x_{i+1},\ i = 1, 2, \ldots, n+1$. Sin pérdida de generalidad, podemos suponer que

$$sgn^* \left(f(x_i) - P^*(x_i) \right) = (-1)^i, \quad i = 1, \ldots, n+2.$$

Vamos a considerar a ρ^* como se definió en (2.12). Entonces para i un entero impar y de la definición de punto extremo,

$$W^*[x_i, f(x_i) - P^*(x_i)] = -\rho^* \text{ o } P^*(x_i) = u(x_i),$$

y para cuando i es un entero par,

$$W^*[x_i, f(x_i) - P^*(x_i)] = \rho^* \text{ o } P^*(x_i) = l(x_i).$$

Al ser W^* una función monótona se tiene que

$$(-1)^{i+1}(P^*(x_i) - P(x_i)) \geq 0, \quad i = 1, \ldots, n+2.$$

Luego para $Q_n(x) = P^*(x) - P(x)$ se cumple que

$$Q_n(x_1) \geq 0,$$

$$Q_n(x_2) \leq 0,$$

$$Q_n(x_3) \geq 0,$$

$$\vdots$$

$$(-1)^{n+1} Q_n(x_{n+2}) \geq 0,$$

por el teorema del valor intermedio, $Q_n(x)$ tiene ceros en $\{y_1, y_2, \ldots, y_{n+1}\}$, tales que

$$x_1 \leq y_1 \leq x_2 \leq y_2 \leq \cdots \leq y_{n+1} \leq x_{n+2}.$$

De donde $Q_n(x)$ tiene que ser el polinomio nulo y

$$P^*(x) \equiv P(x).$$

Por lo tanto, se debe de cumplir que

$$\max_{x \in X} |W^*[x, f(x) - P^*(x)]| \leq \max_{x \in X} |W^*[x, f(x) - P(x)]| \, .$$

Así, P^* es la mejor aproximación generalizada para f con respecto a W^* y $\mathbb{R}_n^*[x]$.

□

Antes de proceder con el teorema de unicidad, debemos observar que utilizamos la suposición de que $l(x) < u(x)$ para demostrar la parte necesaria del teorema 2.3.4.

El siguiente ejemplo nos muestra que cuando $l(x) \leq u(x)$ la caracterización del teorema 2.3.4 no se cumple necesariamente.

Ejemplo 2.3.5. Sea

- $X = [0, 1]$,
- $W^*(x, y) = y$, la norma de Chebyshev,
- $u(x) = x^2$,
- $l(x) = -x$,

- $f(x) = x^4$,
- $\mathbb{R}_1[x]$.

En la siguiente figura, se pueden apreciar las funciones anteriores.

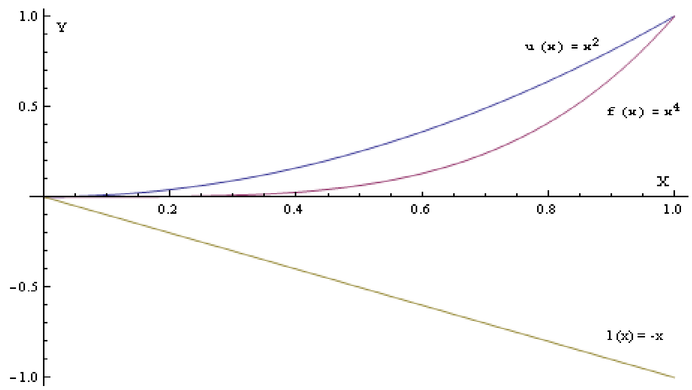

Figura 2.3: Gráfica del comportamiento de las funciones del ejemplo 2.3.5

En este caso,

$$\mathbb{R}_1^*[x] = \{P(x) = ax : \quad -1 \leq a \leq 0\}$$

pues, se tiene que

$$\begin{aligned}\mathbb{R}_1^*[x] &= \left\{P \in \mathbb{R}_1[x] : -x \leq P(x) \leq x^2 \text{ para } \text{ toda } x \in [0,1]\right\}\\ &= \left\{P \in \mathbb{R}_1[x] : -x \leq ax + b \leq x^2 \text{ para } \text{ toda } x \in [0,1]\right\}\end{aligned}$$

Analicemos cada caso.

(i) Si $x = 0$, entonces $0 \leq a \cdot 0 + b \leq 0$, por tanto $b = 0$.

(ii) Si $x \in (0,1]$, podemos dividir por x la siguiente desigualdad

$$-x \leq ax \leq x^2$$

con lo que obtenemos

$$-1 \leq a \leq x;$$

pero la desigualdad anterior ocurre para toda $x \in (0,1]$, entonces, el valor del coeficiente a debe ser menor a cualquier valor del intervalo $(0,1]$ y el ínfimo es el cero, luego $a < 0$. Por tanto $-1 \leq a < 0$.

Por lo que de (i) y (ii) se sigue que $-1 \leq a \leq 0$ y en consecuencia

$$\mathbb{R}_1^*[x] = \{P(x) = ax \text{ para } -1 \leq a \leq 0\},$$

y $P^*(x) = 0$ es la mejor aproximación generalizada de f con respecto a la norma de Chebyshev y $\mathbb{R}_1^*[x]$. Pero P^* no alterna al menos dos veces.

2.4. Unicidad del mejor aproximante

En esta parte se desarrolla un resultado referente a la unicidad del mejor aproximate generalizado, para este caso asumiremos que $u(x) > l(x)$ para toda $x \in X$.

Teorema 2.4.1. Sea f una función continua sobre X, W^* una función de peso generalizada, u, l dos funciones continuas, tales que $u(x) > l(x)$ para toda $x \in X$ y $\mathbb{R}_n^*[x]$ la clase descrita en (2.3). Entonces existe una y sólo una mejor aproximación generalizada para f con respecto a W^* y $\mathbb{R}_n^*[x]$.

Demostración. La existencia de la mejor aproximación generalizada se mostró en el teorema 2.2.4 en la sección 2.2.

Para probar la unicidad de la mejor aproximación, supongamos que la función f tiene dos mejores aproximaciones, digamos $P_1, P_2 \in \mathbb{R}_n^*[x]$.

Como P_1 es una mejor aproximación para f, entonces por el teorema 2.3.4, la curva de error

$$W^*[x, f(x) - P_1(x)] \tag{2.15}$$

alterna al menos $n+1$ veces sobre X, así, existen al menos $n+2$ puntos en X, digamos, $x_1, x_2, \ldots, x_n, x_{n+1}, x_{n+2}$, tales que:

$$x_1 < x_2 < \cdots < x_n < x_{n+1} < x_{n+2},$$

los cuales son puntos extremos de la función (2.15).

Sin pérdida de generalidad, podemos suponer que

$$sgn^*[f(x_i) - P_1(x_i)] = (-1)^i, \quad i = 1, 2, \ldots, n, n+1, n+2.$$

Luego, para $i \in \{1, 2, \ldots, n, n+1, n+2\}$ impar,

$$W^*[x_i, f(x_i) - P_1(x_i)] = -\rho^* \quad \text{o} \quad P_1(x_i) = u(x_i), \tag{2.16}$$

siendo ρ^* como en (2.12) con $P^* = P_1$.

para $i \in \{1, 2, \ldots, n, n+1, n+2\}$ par,

$$W^*[x_i, f(x_i) - P_1(x_i)] = \rho^* \quad \text{o} \quad P_1(x_i) = l(x_i). \tag{2.17}$$

Además, como P_1 es una mejor aproximación y $P_2 \in \mathbb{R}_n^*[x]$, entonces

$$\max_{x\in X} |W^*[x, f(x) - P_1(x)]| \leq \max_{x\in X} |W^*[x, f(x) - P_2(x)]| \tag{2.18}$$

Así, de (2.16), (2.17) y (2.18),

$$(-1)^{i+1}(P_1(x_i) - P_2(x_i)) \geq 0, \quad i = 1, 2, \ldots, n, n+1, n+2.$$

Luego, aplicando el teorema del valor intermedio, $Q_n(x) = P_1(x) - P_2(x)$ cuenta con al menos $n+1$ ceros, digamos $y_1, y_2, \ldots, y_n, y_{n+1}$, tales que

$$x_1 < y_1 < x_2 < y_2 < \cdots < x_n < y_n < x_{n+1} < y_{n+1} < x_{n+2},$$

y como $Q_n(x) \in \mathbb{R}_n[x]$, entonces $Q_n(x) \equiv 0$. Así, $P_1(x) \equiv P_2(x)$.
Por lo tanto, la mejor aproximación es única. $\square$

Capítulo 3

Aproximación con restricciones mediante sistemas de Chebyshev

3.1. Sistemas de Chebyshev

Consideremos a X, un espacio topológico compacto y de Hausdorff y $n \in \mathbb{N}$.
Sea $\{\varphi_1, \varphi_2, \dots, \varphi_n\}$ un conjunto de funciones, tales que $\varphi_i : X \to \mathbb{R}$ para $i = 1, 2, \dots, n$, a las expresiones

$$P(A, x) = a_1\varphi_1(x) + a_2\varphi_2(x) + \cdots + a_n\varphi_n(x), \tag{3.1}$$

las llamaremos *polinomios generalizados*, donde

$$A \in \mathbb{R}^n, \quad A = (a_1, a_2, \dots, a_n).$$

Definición 3.1.1. Sea $n \in \mathbb{N}$. Un conjunto de funciones continuas, reales o complejas, $\{\varphi_1, \dots, \varphi_n\}$, sobre X es un *sistema de Chebyshev* si cumple con las siguientes condiciones:

(a) X contiene al menos n puntos.

(b) Cada polinomio generalizado, $P(A, x)$ con $A \neq \mathbf{0}$, tiene a lo más $n - 1$ ceros distintos sobre X.

Notemos que en la definición 3.1.1, cada polinomio generalizado es una combinación lineal de funciones continuas, por tanto, cada uno de estos polinomios, es continuo sobre X.

Ejemplos 3.1.2. A continuación, se muestran algunos ejemplos de sistemas de Chebyshev, sobre su respectivo conjunto.

1. El conjunto de funciones $\{1, x, x^2, \ldots, x^n\}$ sobre $[a, b] \subset \mathbb{R}$.

2. Las funciones $1, z, z^2, \ldots, z^n$ sobre cada subconjunto W, con al menos $n+1$ elementos, del plano complejo, forman un sistema de Chebyshev.

3. El conjunto de funciones $\{1, \cos x, \cos 2x, \ldots, \cos nx\}$, sobre $[0, \pi]$.

4. Las funciones $1, \mathrm{sen}x, \mathrm{sen}2x, \ldots, \mathrm{sen}nx$, sobre el intervalo $[0, \pi]$, forman un sistema de Chebyshev.

De la definición 3.1.1, se sigue que las funciones que forman un sistema de Chebyshev, son linealmente independientes.

A continuación se muestran distintas formas en que la condición (b) de la definición 3.1.1 puede expresarse.

1. Si $x_1, \ldots, x_n$ son puntos distintos de X, entonces el siguiente sistema de n ecuaciones,

$$a_1\varphi_1(x_k) + a_2\varphi_2(x_k) + \cdots + a_n\varphi_n(x_k) = 0, \quad k = 1, \ldots, n, \tag{3.2}$$

   con n incógnitas $a_1, \ldots, a_n$, tiene sólo la solución:

$$a_1 = a_2 = \cdots = a_n = 0.$$

   Puesto que en caso contrario $P(A, x) = a_1\varphi_1(x_k) + a_2\varphi_2(x_k) + \cdots + a_n\varphi_n(x_k)$ tendría al menos n raíces distintas.

2. Si $x_1, \ldots, x_n$ son n puntos distintos de X, entonces por la afirmación 1 y la regla de Cramer, el determinante

$$\begin{vmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \varphi_1(x_2) & \varphi_2(x_2) & \cdots & \varphi_n(x_2) \\ \vdots & \vdots & & \vdots \\ \varphi_1(x_n) & \varphi_2(x_n) & \cdots & \varphi_n(x_n) \end{vmatrix} \tag{3.3}$$

es diferente de cero.

3. Si $x_1, \ldots, x_n$ son n puntos distintos de X y $c_1, \ldots, c_n$ son constantes arbitrarias, entonces el sistema de ecuaciones

$$a_1\varphi_1(x_k) + a_2\varphi_2(x_k) + \cdots + a_n\varphi_n(x_k) = c_k, \quad k = 1, \ldots, n, \tag{3.4}$$

tiene solución única para $a_1, \ldots, a_n$, lo cual se sigue de la afirmación 1 y 2

Como $P(A, x) = a_1\varphi_1(x) + \cdots + a_n\varphi_n(x)$, entonces la condición (3.4) la podemos escribir como $P(A, x_k) = c_k$, para $k = 1, \ldots, n$, de donde, podemos llamar a $P(A, x)$ *polinomio interpolante* con valores prescritos c_k en los puntos x_k.
Con lo anterior, podemos decir que un polinomio interpolante $P(A, x)$ existe y es único, sin embargo, si $k < n$, entonces el polinomio interpolante existe pero no es único.

4. La siguiente afirmación se cumple para un sistema de Chebyshev real, digamos $\varphi_1, \ldots, \varphi_n$, para cuando $X = [a, b]$.
Si $x_1, \ldots, x_{n-1}$ son $n - 1$ puntos distintos de X, entonces existe un polinomio $D(A, x)$ que se anula en los puntos x_k y cambia de signo en cada uno de éstos puntos, a excepción de cuando los puntos x_k coinciden con a o b.
Si las funciones son $\varphi_i = x^i$, para $i = 0, \ldots, n - 1$, sobre el intervalo $[a, b]$ entonces:

$$D(A, x) = (x - x_1)(x - x_2) \cdots (x - x_{n-1}).$$

Para el caso general, se tiene lo siguiente, si

$$R(A, x) = \begin{vmatrix} \varphi_1(x) & \varphi_2(x) & \cdots & \varphi_n(x) \\ \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \vdots & \vdots & & \vdots \\ \varphi_1(x_{n-1}) & \varphi_2(x_{n-1}) & \cdots & \varphi_n(x_{n-1}) \end{vmatrix} \tag{3.5}$$

se tiene que $R(A, x)$ se anula en los puntos $x_1, \ldots, x_{n-1}$. Si $x \neq x_i$ para $i = 1, \ldots, n - 1$, entonces se sigue de la afirmación 2, que $R(A, x) \neq 0$. Por tanto $R(A, x)$ mantiene un signo constante sobre cada intervalo de la forma $[x_k, x_{k+1}]$ con $k = 1, \ldots, n - 2$.

3.2. Existencia del mejor aproximante

En lo que resta de este capítulo, consideraremos lo siguiente:

- $X \subset \mathbb{R}$, un espacio topológico compacto y de Hausdorff,
- W una función de peso generalizada.
- La clase de funciones que consideraremos es la siguiente:

$$\mathbf{C_n} = \{P(A,x) : \{\varphi_1, \varphi_2, \dots, \varphi_n\} \text{ es un sistema de Chebyshev}\}. \tag{3.6}$$

Definición 3.2.1. Sea f una función continua sobre X, se dice que un polinomio de la forma $P(A,x)$, se aproxima a f si

$$\sup_{x \in X} |W[x, f(x) - P(A,x)]| < \infty.$$

Observación 3.2.2. Notemos que como las funciones que usaremos, a saber, f, P y W son continuas y el conjunto X sobre el que trabajaremos es compacto, entonces de la definición 3.2.1, podemos asegurar que siempre existe una aproximación a f.

Definición 3.2.3. Sea f una función continua sobre X, se dice que un polinomio $P(A,x)$ es la mejor aproximación (o mejor aproximante) a f si $P(A,x)$ es un aproximante a f y además cumple que

$$\sup_{x \in X} |W[x, f(x) - P(A,x)]| \leq \sup_{x \in X} |W[x, f(x) - P(B,x)]|$$

para todo $B \in \mathbb{R}^n$.

Definición 3.2.4. Sea $W(x,y)$ una función de peso generalizada. Llamaremos *error pesado* a $\mathbf{e}$, si

$$\mathbf{e} = \inf_{A \in \mathbb{R}^n} \sup_{x \in X} |W[x, f(x) - P(A,x)]|.$$

De la definición de $\mathbf{e}$ se sigue que existe una sucesión de puntos $\{A_i\}_{i=1}^{\infty}$ (no necesariamente distintos) tales que

$$\lim_{i\to\infty} \sup_{x\in X} |W[x, f(x) - P(A_i, x)]| = \mathbf{e}.$$

Sean u y l dos funciones continuas, definidas sobre X y consideremos la siguiente clase

$$\mathbf{C_n^*} = \{P(A, x) \in \mathbf{C_n} : l(x) \leq P(A, x) \leq u(x), \text{ para toda } x \in X\}, \tag{3.7}$$

cabe mencionar, que podemos elegir a las funciones u, l de tal manera que $\mathbf{C_n^*} \neq \varnothing$.

Asimismo, consideraremos que la función f, que queremos aproximar, cumple con la condición $l(x) \leq f(x) \leq u(x)$.

Antes de proceder con los resultados referentes a la existencia de la mejor aproximación, recordemos la siguiente equivalencia de compacidad, la cual nos será de utilidad en la demostración de los resultados posteriores.

Proposición 3.2.5. Sea $K \subset \mathbb{R}^n$, entonces K es compacto si, y sólo si de toda sucesión $\{A_k\}_{k=1}^{\infty}$ en K se puede obtener una subsucesión convergente, la cual converge a un punto $x \in K$.

Asimismo, se establece el siguiente lema, como un requisito previo para la demostración del teorema 3.2.7.

Lema 3.2.6. Sea $P(A, x) \in \mathbf{C_n}$, para $A \in \mathbb{R}^n$, $A = (a_1, a_2, \ldots, a_n)$ e $I = \{1, 2, \ldots, n\}$, entonces existe $\mu > 0$, tal que si

$$\max_{i\in I} |a_i| = 1, \tag{3.8}$$

entonces

$$\max_{x\in X} |P(A, x)| \geq \mu. \tag{3.9}$$

Demostración. La prueba se hará por contadicción.

Supongamos que para $\mu = \frac{1}{k}$, con $k \in \mathbb{N}$, existen

$A_1 \in \mathbb{R}^n$, con $\max\limits_{i \in I} |a_{i1}| = 1$ y $\max\limits_{x \in X} |P(A_1, x)| < 1$
$A_2 \in \mathbb{R}^n$, con $\max\limits_{i \in I} |a_{i2}| = 1$ y $\max\limits_{x \in X} |P(A_2, x)| < \frac{1}{2}$
$A_3 \in \mathbb{R}^n$, con $\max\limits_{i \in I} |a_{i3}| = 1$ y $\max\limits_{x \in X} |P(A_3, x)| < \frac{1}{3}$
$\vdots$
$A_k \in \mathbb{R}^n$, con $\max\limits_{i \in I} |a_{ik}| = 1$ y $\max\limits_{x \in X} |P(A_k, x)| < \frac{1}{k}$

De esta manera se construye una sucesión $\{A_k\}_{k=1}^{\infty}$ tal que

$$\max_{i \in I} |a_{ik}| = 1, \quad \max_{x \in X} |P(A_k, x)| < \frac{1}{k}. \tag{3.10}$$

Sea D el conjunto de parámetros que cumplen con (3.8). Luego, el conjunto D, conformado por vectores en $\mathbb{R}^n$, es un conjunto compacto y $\{A_k\}_{k=1}^{\infty}$ es una sucesión en D, entonces por la proposición 3.2.5, existe una subsucesión convergente a un conjunto de parámetros $A' \in D$.
Por otra parte, $\max\limits_{x \in X} |P(A, x)|$ es una función continua y por (3.10), se sigue que

$$\max_{x \in X} |P(A', x)| = \lim_{k \to \infty} \max_{x \in X} |P(A_k, x)| = 0$$

lo cual implica que

$$\max_{x \in X} |P(A', x)| = 0,$$

sustituyendo $P(A', x) = \sum\limits_{i=1}^{n} a'_i \varphi_i(x)$,

$$\max_{x \in X} \left| \sum_{i=1}^{n} a'_i \varphi_i(x) \right| = 0,$$

de donde se obtiene $\sum\limits_{i=1}^{n} a'_i \varphi_i(x) = 0$.

Por otro lado, como A' satisface (3.8), se sigue que $\max\limits_{i \in I} |a'_i| = 1$.
Sea $a_j = \max\limits_{i \in I} |a'_i|$ y consideremos los siguientes casos:

Si $|a'_i| = 0$ para todo $i \neq j \in I$, entonces $\sum\limits_{i=1}^{n} |a'_i| = 1$.

Si al menos $|a'_i| \neq 0$ para algún $i \neq j \in I$, entonces $\sum\limits_{i=1}^{n} |a'_i| > 1$.

Por lo tanto, $\sum_{i=1}^{n} |a'_i| \geq 1$ lo cual contradice el hecho de que el conjunto $\{\varphi_i\}_{i=1}^{n}$ es linealmente independiente y en consecuencia, que sea un sistema de Chebyshev, con lo cual concluimos la prueba.

$\square$

Teorema 3.2.7. Sea f una función continua sobre X y $P(A,x) \in \mathbf{C_n}$, entonces, existe $A^* \in \mathbb{R}^n$, tal que

$$\max_{x\in X} |P(A^*,x) - f(x)| \leq \max_{x\in X} |P(A,x) - f(x)| \tag{3.11}$$

para todo $A \in \mathbb{R}^n$.

Demostración. Sea $\{A_k\}_{k=1}^{\infty}$ una sucesión de parámetros en $\mathbb{R}^n$, tal que

$$\lim_{k\to\infty} \max_{x\in X} |P(A_k,x) - f(x)| = \inf_{A\in\mathbb{R}^n} \max_{x\in X} |P(A,x) - f(x)| = \rho. \tag{3.12}$$

Así, para k_0 lo suficientemente grande y para $k > k_0$, se sigue de (3.12) que

$$\max_{x\in X} |P(A_k,x) - f(x)| \leq \rho + 1.$$

Luego, si tomamos $M = \|f(x)\|$, entonces

$$\begin{aligned}
\max_{x\in X} |P(A_k,x)| &= \max_{x\in X} |P(A_k,x) - f(x) + f(x)| \\
&\leq \max_{x\in X} |P(A_k,x) - f(x)| + \max_{x\in X} |f(x)| \\
&\leq \max_{x\in X} |P(A_k,x) - f(x)| + \|f(x)\| \\
&\leq \rho + M + 1
\end{aligned}$$

por tanto

$$\max_{x\in X} |P(A_k,x)| \leq \rho + M + 1. \tag{3.13}$$

Veamos ahora que (3.13) implica que los parámetros, es decir, que los componentes de A_k están acotados.
En particular, supongamos lo siguiente:

$$\max_{i\in I} |a_{ik}| \geq \frac{\rho + M + 1}{\mu'}, \qquad I = \{1, 2, \ldots, n\},$$

para algún $\mu' > 0$. Con ésto, obtenemos:

$$\begin{aligned}
\max_{x\in X}|P(A_k,x)| &= \max_{x\in X}\left[\left(\max_{i\in I}|a_{ik}|\right)\left|\frac{\sum\limits_{i=1}^{n} a_{ik}\varphi_i(x)}{\max\limits_{i\in I}|a_{ik}|}\right|\right] \\
&= \max_{i\in I}|a_{ik}|\cdot\max_{x\in X}\left|\frac{\sum\limits_{i=1}^{n} a_{ik}\varphi_i(x)}{\max\limits_{i\in I}|a_{ik}|}\right| \qquad (3.14)\\
&\geq \tfrac{\rho+M+1}{\mu'}\cdot\max_{x\in X}\left|\frac{\sum\limits_{i=1}^{n} a_{ik}\varphi_i(x)}{\max\limits_{i\in I}|a_{ik}|}\right|.
\end{aligned}$$

La función de aproximación del lado derecho de (3.14) satisface la condición (3.8). En efecto, se tiene que

$$\frac{\sum\limits_{i=1}^{n} a_{ik}\varphi_i(x)}{\max\limits_{i\in I}|a_{ik}|} = \sum_{i=1}^{n}\frac{a_{ik}}{\max\limits_{i\in I}|a_{ik}|}\varphi_i(x).$$

En general, se tiene que $|a_{ik}| \leq \max\limits_{j\in I}|a_{jk}|$, luego,

Si $|a_{ik}| = \max\limits_{j\in I}|a_{jk}|$, entonces $\left|\frac{a_{ik}}{\max\limits_{j\in I}|a_{jk}|}\right| = 1.$

Si $|a_{ik}| < \max\limits_{j\in I}|a_{jk}|$, entonces $\left|\frac{a_{ik}}{\max\limits_{j\in I}|a_{jk}|}\right| < 1.$
Por tanto,

$$\left|\frac{a_{ik}}{\max\limits_{j\in I}|a_{jk}|}\right| \leq 1.$$

Luego, por el lema 3.2.6, se sigue que existe $\mu > 0$, tal que

$$\max_{x\in X}\left|\frac{\sum\limits_{i=1}^{n} a_{ik}\varphi_i(x)}{\max\limits_{i\in I}|a_i|}\right| \geq \mu,$$

entonces, de (3.14), se tiene que

$$\max_{x\in X} |P(A_k, x)| \geq \frac{\rho + M + 1}{\mu'} \cdot \mu$$

ahora, eligiendo $0 < \mu' < \mu$, se cumple que $\frac{\mu}{\mu'} > 1$.
Así,

$$\max_{x\in X} |P(A_k, x)| > \rho + M + 1,$$

lo cual contradice (3.13).

Por lo tanto,

$$|a_{ik}| < \frac{\rho + M + 1}{\mu'}, \quad \text{para} \;\; k \geq k_0. \tag{3.15}$$

Además, el conjunto de parámetros D' que satisfacen (3.15) es compacto, entonces por la proposición 3.2.5, se tiene que $\{A_k\}_{k=1}^{n}$ tiene una subsucesión convergente, denotemos el límite de esta subsucesión con $A^* \in D'$, de donde, junto con (3.12) se tiene la existencia de una mejor aproximación. □

Teorema 3.2.8. Si $\lim\limits_{i\to\infty} \sup\limits_{x\in X} |W[x, f(x) - P(A_i, x)]| = \mathbf{e}$, entonces la sucesión $\{A_i\}_{i=1}^{\infty}$ está acotada.

Demostración. Supongamos que $\lim\limits_{i\to\infty} \sup\limits_{x\in X} |W[x, f(x) - P(A_i, x)]| = \mathbf{e}$.
Para $A = (a_1, a_2, \ldots, a_n)$ sea $\parallel A \parallel = \max\limits_{1\leq i\leq n} |a_i|$. Si la sucesión $\{A_i\}_{i=1}^{\infty}$ no estuviese acotada, entonces existiría una subsucesión $\{B_i\}_{i=1}^{\infty}$ de $\{A_i\}_{i=1}^{\infty}$, tal que

(a) $\lim\limits_{i\to\infty} \parallel B_i \parallel = \infty$,

(b) $\parallel B_i \parallel \neq 0$, para $i = 1, 2, \ldots$

(c) $\lim\limits_{i\to\infty} \frac{B_i}{\|B_i\|} = B_0$, para algún B_0.

(d) $\sup\limits_{x\in X} |W[x, f(x) - P(B_i, x)]| < \mathbf{e} + 1$, $i = 1, 2, \ldots$

Como las funciones φ_i son linealmente independientes y $\parallel B_0 \parallel = 1$, existe un punto $t \in X$ tal que $P(B_0, t) \neq 0$.

Luego, por (d) tenemos

$$|W[t, f(t) - P(B_i, t)]| < \mathbf{e} + 1, \quad \text{para } \; i = 1, 2, \ldots$$

de donde

$$-\mathbf{e} - 1 < W[t, f(t) - P(B_i, t)] < \mathbf{e} + 1, \quad \text{para } \; i = 1, 2, \ldots \tag{3.16}$$

Podemos reescribir lo anterior como

$$-\mathbf{e} - 1 < W\left[t, \| B_i \| \left\{ \frac{f(t)}{\| B_i \|} - P\left(\frac{B_i}{\| B_i \|}, t \right) \right\} \right] < \mathbf{e} + 1, \quad \text{para } \; i = 1, 2, \ldots$$

por otro lado, se tiene que

$$\begin{aligned} &\lim_{i \to \infty} P\left(\tfrac{B_i}{\|B_i\|}, t \right) = P(B_0, t) \neq 0, \\ &\lim_{i \to \infty} \tfrac{f(t)}{\|B_i\|} = f(t) \cdot \lim_{i \to \infty} \tfrac{1}{\|B_i\|} = 0. \end{aligned} \tag{3.17}$$

Luego, de (3.17)

$$\lim_{i \to \infty} \left| \| B_i \| \left\{ \frac{f(t)}{\| B_i \|} - P\left(\frac{B_i}{\| B_i \|}, t \right) \right\} \right| = \infty. \tag{3.18}$$

Por lo tanto, tomando $y = \| B_i \| \left\{ \frac{f(t)}{\|B_i\|} - P\left(\frac{B_i}{\|B_i\|}, t \right) \right\}$ en la propiedad (c) de la función de peso $W(x, y)$ y de (3.18)

$$\lim_{\substack{|y| \to \infty \\ i \to \infty}} |W(t, y)| = \infty,$$

lo cual contradice (3.16).

Luego, la sucesión $\{A_i\}_{i=1}^{n}$ está acotada. $\square$

El siguiente teorema ha sido adaptado para el caso en el cual se trabaja con la clase $\mathbf{C_n}$, esta adaptación surge de la necesidad de contar con ciertas condiciones para poder obtener el resultado principal de esta sección, referente a la existencia de la mejor aproximación en esta clase.

Teorema 3.2.9. Si $W(x, y)$ satisface las siguientes condiciones

(a) Para cada $x \in X$, $W(x, y)$ es una función monótona no decreciente de y,

(b) $sgn\ W(x, y) = sgn\ y$ para todo (x, y),

(c) $\lim\limits_{|y|\to\infty} |W(x, y)| = \infty$ para cada x,

(d) Para cada $x \in X$ y $t \neq 0$ $\lim\limits_{y\to t,\ |y|<|t|} W(x, y) = W(x, t)$,

y existe una aproximación a f en $\mathbf{C_n}$, entonces existe una mejor aproximación a f en $\mathbf{C_n}$.

Demostración. Sea $\{A_i\}_{i=1}^{\infty}$ una sucesión en $\mathbb{R}^n$, tal que

$$\begin{array}{l} 1.\ \lim\limits_{i\to\infty} |W[x, f(x) - P(A_i, x)]| = \mathbf{e}, \\ 2.\ \lim\limits_{i\to\infty} A_i = A_0 \in \mathbb{R}^n. \end{array} \tag{3.19}$$

La existencia de tal sucesión se sigue de la definición 3.2.4 y del teorema 3.2.8.

Lo que demostraremos es que $P(A_0, x)$ es una mejor aproximación.

Como $A_0 \in \mathbb{R}^n$, entonces, se cumple que

$$\sup_{x\in X} |W[x, f(x) - P(A_0, x)]| \geq \mathbf{e}.$$

Si $\sup\limits_{x\in X} |W[x, f(x) - P(A_0, x)]| > \mathbf{e}$, entonces existe $\epsilon > 0$ y un $t \in X$ tal que

$$|W[t, f(t) - P(A_0, t)]| \geq \mathbf{e} + \epsilon \geq \epsilon > 0, \tag{3.20}$$

lo cual implica que

$$|W[t, f(t) - P(A_0, t)]| > 0.$$

Notemos que

$$f(t) - P(A_0, t) \neq 0$$

de lo contrario, $|W[t, f(t) - P(A_0, t)]| = 0$.

Por otro lado, como (3.19) se satisface, entonces para i lo suficientemente grande, digamos $i > N$, se tiene que

$$\sup_{x\in X} |W[x, f(x) - P(A_i, x)]| \leq \mathbf{e} + \epsilon/2. \tag{3.21}$$

Además, también se cumple que

$$sgn[f(t) - P(A_i, t)] = sgn[f(t) - P(A_0, t)], \tag{3.22}$$

pues en caso contrario, si $f(t) - P(A_i, t) < 0$ y $f(t) - P(A_0, t) > 0$, entonces

$$f(t) - P(A_i, t) < f(t) - P(A_0, t),$$

calculando el límite

$$\lim_{i\to\infty} f(t) - P(A_i, t) \leq 0 < f(t) - P(A_0, t),$$

$$-\lim_{i\to\infty} P(A_i, t) < -P(A_0, t),$$

luego, por (3.19) y el hecho de que $P(A_i, t)$ y $P(A_0, t)$ sean continuos, se sigue que $-P(A_0, t) < -P(A_0, t)$, lo cual es una contradicción.

Por otro lado, de (3.20) y (3.21), se tiene

$$\sup_{x\in X} |W[x, f(x) - P(A_i, x)]| < |W[t, f(t) - P(A_0, t)]| .$$

En particular,

$$|W[t, f(t) - P(A_i, t)]| < |W[t, f(t) - P(A_0, t)]| \tag{3.23}$$

Analicemos los siguientes casos

- Si $W[t, f(t) - P(A_i, t)] > 0$, entonces por la propiedad (b) de W, se sigue que $f(t) - P(A_i, t) > 0$ y por (3.22) $f(t) - P(A_0, t) > 0$, así, $W[t, f(t) - P(A_0, t)] > 0$, luego, como W es monótona en y y

  $$W[t, f(t) - P(A_i, t)] < W[t, f(t) - P(A_0, t)]$$

  se obtiene que $f(t) - P(A_i, t) < f(t) - P(A_0, t)$, por lo tanto

  $$|f(t) - P(A_i, t)| < |f(t) - P(A_0, t)| .$$

- Si $W[t, f(t) - P(A_i, t)] < 0$, entonces por la propiedad (b) de W, se sigue que $f(t) - P(A_i, t) < 0$ y por (3.22) $f(t) - P(A_0, t) < 0$, así, $W[t, f(t) - P(A_0, t)] < 0$, luego, dado que W es monótona en y y

$$W[t, f(t) - P(A_i, t)] > W[t, f(t) - P(A_0, t)]$$

se tiene que $f(t) - P(A_i, t) > f(t) - P(A_0, t)$ por lo tanto

$$|f(t) - P(A_i, t)| < |f(t) - P(A_0, t)| .$$

Notemos que los signos de $W[t, f(t) - P(A_i, t)]$ y $W[t, f(t) - P(A_0, t)]$ no pueden ser distintos puesto que si ésto ocurriera se tendría que los signos de $f(t) - P(A_i, t)$ y de $f(t) - P(A_0, t)$ serían distintos, lo cual contradice (3.22).

Por tanto, si $i > N$, se sigue que

$$|f(t) - P(A_i, t)| < |f(t) - P(A_0, t)| . \tag{3.24}$$

De todo lo anterior, tenemos

- $t \in X$,
- Se cumple (3.24) y
- $f(t) - P(A_0, t) \neq 0$

entonces, por la propiedad d) de W, se sigue que

$$\lim_{i\to\infty} |W[t, f(t) - P(A_i, t)]| = |W[t, f(t) - P(A_0, t)]| \geq \mathbf{e} + \epsilon$$

luego,

$$\lim_{i\to\infty} |W[t, f(t) - P(A_i, t)]| \geq \mathbf{e} + \epsilon$$

lo cual contradice (3.19).
Por lo tanto

$$\sup_{x\in X} |W[x, f(t) - P(A_0, x)]| = \mathbf{e}.$$

De donde, $P(A_0, x)$ es una mejor aproximación a f con respecto a W y $\mathbf{C_n}$.

□

Una vez que hemos analizado los resultados anteriores, tenemos las herramientas suficientes para poder demostrar la existencia de la mejor aproximación en la clase $\mathbf{C_n^*}$ descrita en (3.7).

Teorema 3.2.10. Sea f una función continua sobre X. Si $W^*(x,y)$ es una función de peso generalizada y consideramos la clase $\mathbf{C_n^*}$, entonces existe una mejor aproximación a f con respecto a W^* y $\mathbf{C_n^*}$.

Demostración. Definamos la siguiente función

$$W(x,y) = \begin{cases} +\infty & \text{si} \quad y > f(x) - l(x), \\ W^*(x,y) & \text{si} \quad l(x) \le f(x) - y \le u(x), \\ -\infty & \text{si} \quad y < f(x) - u(x). \end{cases} \tag{3.25}$$

Por el lema 2.2.2 se tiene que $W(x,y)$ es una función de peso generalizada que cumple con las condiciones que se piden en el teorema 3.2.9, entonces, por la observación 3.2.2 y por el teorema 3.2.9, existe una mejor aproximación para f en $\mathbf{C_n}$ con respecto a W, denotemos a esta mejor aproximación como $P(A^*,x)$, entonces

$$\sup_{x\in X} |W[x, f(x) - P(A^*,x)]| \le \inf_{P(A,x)\in \mathbf{C_n^*}} \sup_{x\in X} |W[x, f(x) - P(A,x)]|$$

Como se tiene que $\mathbf{C_n^*} \neq \varnothing$, sea $P(A',x) \in \mathbf{C_n^*} \subset \mathbf{C_n}$ y al ser $P(A^*,x)$ una mejor aproximación se cumple lo siguiente:

$$\sup_{x\in X} |W[x, f(x) - P(A^*,x)]| \le \sup_{x\in X} |W[x, f(x) - P(A',x)]|\,. \tag{3.26}$$

Como $P(A',x) \in \mathbf{C_n^*}$, entonces

$$l(x) \le P(A',x) \le u(x)$$

de donde, se sigue que

$$l(x) \le f(x) - (f(x) - P(A',x)) \le u(x),$$

tomando $y = f(x) - P(A',x)$ en (3.25), obtenemos

$$W(x,y) = W^*(x,y)$$

o bien,

$$W[x, f(x) - P(A',x)] = W^*[x, f(x) - P(A',x)] < \infty$$

de donde,

$$\sup_{x\in X} |W[x, f(x) - P(A',x)]| = \max_{x\in X} |W^*[x, f(x) - P(A',x)]| < \infty \tag{3.27}$$

pues W^* es continua y X es compacto.

Por lo tanto por (3.26) y (3.27),

$$\sup_{x\in X}|W[x,f(x)-P(A^*,x)]| \leq \max_{x\in X}|W^*[x,f(x)-P(A',x)]| < \infty.$$

Luego, $\sup_{x\in X}|W[x,f(x)-P(A^*,x)]| < \infty$, de donde se sigue que

$$W[x,f(x)-P(A^*,x)] = W^*[x,f(x)-P(A^*,x)] \tag{3.28}$$

tomando $y_1 = f(x) - P(A^*,x)$ en (3.25) se cumple que

$$l(x) \leq f(x) - (f(x) - P(A^*,x)) \leq u(x)$$

Así, $l(x) \leq P(A^*,x) \leq u(x)$ para toda $x \in X$.
Luego, por (3.26) junto con (3.28)

$$\sup_{x\in X}|W^*[x,f(x)-P(A^*,x)]| \leq \sup_{x\in X}|W^*[x,f(x)-P(A',x)]|$$

Por lo tanto, $P(A^*,x) \in \mathbf{C}^*_{\mathbf{n}}$ es una mejor aproximación generalizada con respecto a W^* y $\mathbf{C}^*_{\mathbf{n}}$.

□

3.3. Caracterización del mejor aproximante

Cabe mencionar que el siguiente lema únicamente se encuentra en la literatura para el caso de funciones racionales. Debido a la necesidad de contar con un resultado análogo se consideró la adaptación del mismo, para el caso en el que los elementos que consideramos pertenecen a la clase $\mathbf{C_n}$, descrita en (3.6), obteniendo de esta manera el siguiente lema:

Lema 3.3.1. Sea $P(A,x) \in \mathbf{C_n}$ y $x_1, x_2, \dots, x_k \in X$ con $k \leq n-1$ y $x_1 < x_2 < x_3 < \cdots < x_k$; entonces dado $\epsilon > 0$ existe $P(A^*,x) \in \mathbf{C_n}$ tal que

(i) $\max\limits_{x \in X} |P(A,x) - P(A^*,x)| < \epsilon$,

(ii) $P(A^*,x_i) = P(A,x_i), \;\; i = 1, \dots, k$, y

(iii) $sgn[P(A^*,x) - P(A,x)] = (-1)^i$ para $x \in X \cap (x_i, x_{i+1}), \;\; i = 0, 1, \dots, k$, donde $x_0 = a = \inf\{x : x \in X\}$ y $x_{k+1} = b = \sup\{x \; : \; x \in X\}$.

Demostración. Sea $P(A,x) = \sum\limits_{i=1}^{n} a_i \varphi_i(x) \in \mathbf{C_n}, \; x_1, x_2, \dots, x_k \in X$ y $\epsilon > 0$. Consideremos la siguiente función

$$R(C,x) = \begin{vmatrix} \varphi_1(x) & \varphi_2(x) & \cdots & \varphi_n(x) \\ \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \vdots & \vdots & & \vdots \\ \varphi_1(x_k) & \varphi_2(x_k) & \cdots & \varphi_n(x_k) \end{vmatrix} \tag{3.29}$$

cuya existencia la podemos asegurar de la equivalencia al inciso b) de la definición de sistemas de Chebyshev (afirmación 4).

Además, la función $R(C,x)$ tiene las siguientes propiedades:

- Se anula en los puntos $x_1, x_2, \dots, x_k$.
- Si consideramos que $x \neq x_i$, para $i = 1, 2, \dots, k$, entonces $R(C,x) \neq 0$.

Por tanto, $R(C,x)$ mantiene un signo constante sobre cada intervalo de la forma $X \cap (x_i, x_{i+1}), \; i = 0, 1, \dots, k$, donde $x_0 = a = \inf\{x : x \in X\}$ y $x_{k+1} = b = \sup\{x \; : \; x \in X\}$.

Definamos ahora a $P(A^*, x)$ de la siguiente forma:

$$P(A^*, x) = P(A, x) - tR(C, x)$$

para algún $t > 0$.

Sustituyendo, obtenemos

$$P(A^*, x) = \sum_{i=1}^{n} (a_i - tc_i)\varphi_i(x)$$

Por la manera en la que hemos definido $P(A^*, x)$, se sigue que

$$P(A^*, x) \in \mathbf{C_n}.$$

Para $P(A^*, x)$, así definido, se cumple lo siguiente:

- Se tiene que $|P(A, x) - P(A^*, x)| = \left|\sum_{i=1}^{n} (tc_i)\varphi_i(x)\right|$, tomando
  - $c_j = \max |c_i|$,
  - $M' = \max_{i \in I}\{\max_{x \in X} |\varphi_i(x)|\}$, con $I = 1, 2, \ldots, n$, lo cual se puede hacer puesto que cada φ_i es continua y X es compacto.
  - $M = \frac{\epsilon}{nc_j M'}$.

  Tomando $t < M$, se tiene

$$\begin{aligned} \left|\sum_{i=1}^{n} tc_i\varphi_i(x)\right| &\leq \sum_{i=1}^{n} |tc_i\varphi_i(x)| \\ &< \sum_{i=1}^{n} |Mc_i\varphi_i(x)| \\ &< Mc_jM' \sum_{i=1}^{n} 1 \\ &= Mc_jM'n \\ &= \epsilon, \end{aligned}$$

  entonces se tiene que

$$\max_{x \in X} |P(A, x) - P(A^*, x)| < \epsilon.$$

- También, se tiene que como $P(A,x) - P(A^*,x) = tR(C,x)$, entonces en $x_i,\ i = 1, 2, \ldots, k$, se cumple que

  $$P(A, x_i) - P(A^*, x_i) = 0$$

  con lo que $P(A^*, x_i) = P(A, x_i)$, para $i = 1, \ldots, k$.

- Por último, se tiene que

  $$sgn[P(A^*,x) - P(A,x)] = (-1)^i \quad \text{para } x \in X \cap (x_i, x_{i+1}), \quad i = 0, 1, \ldots, k,$$

  donde

  $$x_0 = a = \text{ínf}\{x : x \in X\}, \quad x_{k+1} = b = \sup\{x : x \in X\}.$$

  Lo cual se sigue del hecho de que $sgn[P(A^*,x) - P(A,x)] = sgn[tR(C,x)]$.

$\square$

Una vez que contamos con el teorema 3.2.10 junto con el lema 3.3.1, lo que resta, antes de proceder con la demostración de la unicidad de la mejor aproximación es caracterizar a ésta en términos de $n+1$ puntos.

Teorema 3.3.2. Sea f una función continua sobre $X \subset \mathbb{R}$, W^* una función de peso generalizada, $l(x) < u(x)$ para toda $x \in X$ y $\mathbf{C_n^*}$ la clase descrita en (3.7). Entonces $P(A^*,x) \in \mathbf{C_n^*}$ es la mejor aproximación generalizada para f con respecto a W^* y $\mathbf{C_n^*}$ si y sólo si la curva de error $W^*[x, f(x) - P(A^*,x)]$ alterna al menos n veces.

Demostración. La prueba de la parte necesaria se hará por contradicción, para lo cual, supongamos que $P(A^*,x) \in \mathbf{C_n^*}$ es la mejor aproximación generalizada para f con respecto a W^* y $\mathbf{C_n^*}$, y que $W^*[x, f(x) - P(A^*,x)]$ alterna $k < n$ veces sobre X.

Como $W^*[x, f(x) - P(A^*,x)]$ alterna k veces, entonces, por definición existen $k+1$ puntos, digamos $y_1, y_2, \ldots, y_{k+1}$, tales que $y_1 < y_2 < \cdots < y_{k+1}$, puntos extremos para esta curva de error . Consideremos i, con $1 \leq i \leq k$ fijo y supongamos que y_i es un extremo positivo. Sea

$$y' = \sup\{y : y \in X \text{ con } y_i \leq y \leq y_{i+1} \text{ y } y \text{ es un extremo positivo}\},$$

$$y'' = \text{ínf}\{y : y \in X \text{ con } y_i \leq y \leq y_{i+1} \text{ y } y \text{ es un extremo negativo}\}.$$

Como X es compacto y todas las funciones involucradas son continuas, se sigue que y' es un extremo positivo y y'' es un extremo negativo. Además $y' < y''$, puesto que si $y' = y''$, tendríamos que y' es un extremo positivo y negativo al mismo tiempo y si $y'' < y'$, se tendría que la curva de error $W^*[x, f(x) - P(A^*, x)]$ alternaría más de k veces. Por otro lado, si $y \in (y', y'')$ entonces y no es un punto extremo por la manera en que hemos elegido a y' y y''.

Para el caso en el que y_i sea un extremo negativo, consideraremos

$$y' = \sup\{y \ : \ y \in X \ \text{ con } \ y_i \leq y \leq y_{i+1} \text{ y } y \ \text{ es } \ \text{ un } \ \text{ extremo } \ \text{ negativo}\},$$

$$y'' = \inf\{y \ : \ y \in X \ \text{ con } \ y_i \leq y \leq y_{i+1} \text{ y } y \ \text{ es } \ \text{ un } \ \text{ extremo } \ \text{ positivo}\}.$$

Una vez más, por la compacidad de X y la continuidad de cada una de las funciones que estamos considerando, se sigue que y' es un extremo negativo y y'' es un extremo positivo con $y' < y''$.

Ahora, si $X \cap (y', y'') \neq \varnothing$, elegimos $x_i \in X \cap (y', y'')$ arbitrario, luego, en x_i tenemos:

$$(i) \ l(x_i) < P(A^*, x_i) < u(x_i),$$

$$(ii) \ |W^*[x_i, f(x_i) - P(A^*, x_i)]| < \max_{x \in X} |W^*[x, f(x) - P(A^*, x)]|. \tag{3.30}$$

Denotemos a

$$\rho^* = \max_{x \in X} |W^*[x, f(x) - P(A^*, x)]|. \tag{3.31}$$

Por otro lado, si $X \cap (y', y'') = \varnothing$, entonces elegimos

$$x_i = \frac{y' + y''}{2}, \quad \text{para cada } i = 1, 2, \ldots, k.$$

Sea

$$x_0 = \inf\{y \ : \ y \in X\} \text{ y } x_{k+1} = \sup\{y \ : \ y \in X\}.$$

Consideremos ahora los siguientes puntos

$$x_i^* = \inf_{x \in X}\{x \in [x_i, x_{i+1}]\}$$

$$x_{i+1}^* = \sup_{x \in X}\{x \in [x_i, x_{i+1}]\}$$

con $i = 0, 1, 2, \ldots, k$, para construir los siguientes intervalos

$$\left\{X \cap [x_i^*, x_{i+1}^*]\right\}, \quad i = 0, 1, \ldots, k.$$

Con lo que obtenemos una familia de $k+1$ intervalos consecutivos en X, cada uno de estos intervalos, por la misma construcción, no contienen alternancias. Además, $W^*[x, f(x) - P(A^*, x)]$ alterna exactamente una vez sobre dos intervalos adyacentes cualquiera.

Notemos que $X \cap [x_i^*, x_{i+1}^*]$, con $i \in \{0, 1, \ldots, k\}$ fijo contiene al menos un extremo, digamos y_{i+1}; además, no contiene extremos con signo opuesto al de y_{i+1}, de lo contrario, $W^*[x, f(x) - P(A^*, x)]$ alternaría más de una vez sobre dos intervalos adyacentes.

Sea $i = 0$, sin pérdida de generalidad, podemos suponer que y_1 es un extremo positivo y dado que X es compacto, que todas las funciones que estamos considerando son continuas y $u(x) > l(x)$, entonces existe $\delta_1 > 0$ tal que:

- $W^*[x, f(x) - P(A^*, x)] + \rho^* > \delta_1$, o bien, $W^*[x, f(x) - P(A^*, x)] > -\rho^* + \delta_1$ para todo $x \in X \cap [x_0^*, x_1^*]$.
- $P(A^*, x) + \delta_1 < u(x)$, es decir, $P(A^*, x) < u(x) - \delta_1$ para todo $x \in X \cap [x_0^*, x_1^*]$.

Sea $\epsilon_1 > 0$ tal que

$$\epsilon_1 < \delta_1/2 \;\text{ y }\; W^*[x, f(x) - P(A^*, x) - \epsilon_1] \geq -\rho^* + \delta_1/2 \tag{3.32}$$

para todo $x \in X \cap [x_0, x_1]$. Repitiendo este proceso sobre $X \cap [x_1^*, x_2^*]$, obtenemos $\epsilon_2 > 0$ y $\delta_2 > 0$ tal que

$$\epsilon_2 < \delta_2/2, \; P(A^*, x) > l(x) + \delta_2 \tag{3.33}$$

y

$$W^*[x, f(x) - P(A^*, x) + \epsilon_2] \leq \rho^* - \delta_2/2$$

para todo $x \in X \cap [x_1^*, x_2^*]$. Continuando con este proceso para cada intervalo de la forma $X \cap [x_i^*, x_{i+1}^*]$, $\quad i = 0, 1, \ldots, k$, obtenemos $\epsilon_1, \epsilon_2, \ldots, \epsilon_k$, donde

$\epsilon_i > 0$ para todo $i = 1, 2, \ldots, k$.

Sea

$$\epsilon = \text{mín}\{\epsilon_1, \ldots, \epsilon_k\}.$$

Así, del análisis previo, se obtiene lo siguiente:

- $P(A^*, x) \in \mathbf{C_n}^* \subset \mathbf{C_n}$,
- $x_1^*, x_2^*, \ldots, x_k^* \in X$,
- $k < n$,
- $\epsilon > 0$.

Luego, por el lema 3.3.1, existe $P(A, x) \in \mathbf{C_n}$ tal que

$$P(A, x_i^*) = P(A^*, x_i^*), \;\; i = 1, \ldots, k,$$

$$\text{máx}_{x\in X} |P(A, x) - P(A^*, x)| < \epsilon$$

$$sgn\,(P(A, x) - P(A^*, x)) = (-1)^i \;\text{ para }\; x \in X \cap (x_i^*, x_{i+1}^*), \;\; i = 0, 1, \ldots, k.$$

Probemos ahora que $P(A, x) \in \mathbf{C_n}^*$.

Se tiene que para $x \in X \cap [x_0^*, x_1^*]$

- si $x = x_1^*$, entonces $P(A, x) = P(A^*, x)$,
- si $x \neq x_1^*$ entonces $sgn[P(A, x) - P(A^*, x)] = +1$, lo que implica que: $P(A, x) > P(A^*, x)$.

Así, de los dos casos anteriores, se tiene que $P(A, x) \geq P(A^*, x)$ y como $l(x) \leq P(A^*, x)$, entonces $l(x) \leq P(A, x)$.

También, se tiene que $P(A, x) \leq u(x) - \delta_1/2 \leq u(x)$ y puesto que

$$\text{máx}_{x\in X} |P(A^*, x) - P(A, x)| < \epsilon < \delta_1/2,$$

$$P(A^*, x) < u(x) - \delta_1, \text{ para toda } x \in X \cap [x_0^*, x_1^*],$$

entonces,

$$l(x) \leq P(A, x) \leq u(x), \text{ para toda } x \in X \cap [x_0^*, x_1^*].$$

Una vez más, por la continuidad de las funciones y el hecho de que $X \cap [x_0^*, x_1^*]$ es compacto, tenemos

$$\rho^* \geq \max_{x \in X \cap [x_0^*, x_1^*]} W^* [x, f(x) - P(A^*, x)] > \max_{z \in X \cap [x_0^*, x_1^*]} W^* [z, f(z) - P(A, z)]$$

de lo anterior junto con (3.32) se tiene que

$$\min_{z \in X \cap [x_0^*, x_1^*]} W^* [z, f(z) - P(A^*, z) - \epsilon] \geq -\rho^* + \frac{\delta_1}{2} > -\rho^*.$$

Luego,

$$-\rho^* < \max_{z \in X \cap [x_0^*, x_1^*]} W^* [z, f(z) - P(A, z)] < \rho^*,$$

entonces

$$\max_{z \in X \cap [x_0^*, x_1^*]} |W^* [z, f(z) - P(A, z)]| < \rho^*.$$

De manera similar para el intervalo $X \cap [x_1^*, x_2^*]$ tenemos que para $x \in X \cap [x_1^*, x_2^*]$,

- si $x = x_1^*$ ó $x = x_2^*$, entonces $P(A, x) = P(A^*, x)$,
- si $x \neq x_1^*$ y $x \neq x_2^*$, entonces $sgn[P(A, x) - P(A^*, x)] = -1$, lo que implica que $P(A, x) < P(A^*, x)$.

De los casos anteriores, podemos concluir que

$$P(A, x) \leq P(A^*, x),$$

asimismo, como $P(A^*, x) \leq u(x)$, entonces $P(A, x) \leq u(x)$.

También, se tiene que $l(x) < l(x) + \delta_2/2 \leq P(A, x)$ y puesto que

$$\max_{x \in X} |P(A^*, x) - P(A, x)| < \epsilon < \delta_2/2,$$

se tiene

$$l(x) + \delta_2 < P(A^*, x), \text{ para todo } x \in X \cap [x_1^*, x_2^*],$$

con lo que,

$$l(x) \leq P(A, x) \leq u(x), \text{ para todo } x \in X \cap [x_1^*, x_2^*].$$

Por la continuidad de las funciones consideradas y la compacidad de $X \cap [x_1^*, x_2^*]$ se sigue que

$$\begin{aligned} -\rho^* &\leq \min_{x \in X \cap [x_1^*, x_2^*]} W^*[x, f(x) - P(A^*, x)] \\ &\leq \max_{x \in X \cap [x_1^*, x_2^*]} W^*[x, f(x) - P(A^*, x)] \\ &< \max_{z \in X \cap [x_1^*, x_2^*]} W^*[z, f(z) - P(A, z)] \\ &\leq \max_{z \in X \cap [x_1^*, x_2^*]} W^*[z, f(z) - P(A^*, z) + \epsilon] \\ &\leq \rho^* - \delta_2/2 < \rho^*. \end{aligned}$$

Por lo tanto,

$$-\rho^* < \max_{x \in X \cap [x_1^*, x_2^*]} W^*[x, f(x) - P(A, x)] < \rho^*,$$

entonces

$$\max_{x \in X \cap [x_1^*, x_2^*]} |W^*[x, f(x) - P(A, x)]| < \rho^*.$$

Repitiendo estos argumentos para cada uno de los intervalos $X \cap [x_i, x_{i+1}]$, con $i = 0, 1, \ldots, k$ obtenemos :

- $P(A, x) \in \mathbf{C_n}^*$, pues $l(x) \leq P(A, x) \leq u(x)$ para todo $x \in X$ y
- $\max_{x \in X} |W^*[x, f(x) - P(A, x)]| < \max_{x \in X} |W^*[x, f(x) - P(A^*, x)]| = \rho^*$,

con lo que tenemos que $P(A, x) \in \mathbf{C}_\mathbf{n}^*$ aproxima mejor a f que $P(A^*, x)$, lo cual contradice el hecho de que $P(A^*, x)$ es la mejor aproximación generalizada de f con respecto a W^* y $\mathbf{C}_\mathbf{n}^*$. Por lo tanto, la curva de error $W^*[x, f(x) - P(A^*, x)]$ alterna al menos n veces sobre X.

Para probar la parte suficiente del teorema, supongamos que la curva de error $W^*[x, f(x) - P(A^*, x)]$ para $P(A^*, x) \in \mathbf{C}_\mathbf{n}^*$ alterna al menos n veces. Por simplicidad asumiremos que ésta alterna exactamente n veces, supongamos también que existe $P(A, x) \in \mathbf{C}_\mathbf{n}^*$ tal que

$$\max_{x \in X} |W^*[x, f(x) - P(A, x)]| < \max_{x \in X} |W^*[x, f(x) - P(A^*, x)]|.$$

Sea $\{x_1, x_2, \ldots, x_n, x_{n+1}\}$ con $x_1 < x_2 < \ldots < x_n < x_{n+1}$ un conjunto de $n + 1$ puntos extremos en X, para la curva de error $W^*[x, f(x) - P(A^*, x)]$. Sin pérdida de generalidad, podemos suponer que

$$sgn^* \left(f(x_i) - P(A^*, x_i) \right) = (-1)^i, \quad i = 1, \ldots, n + 1.$$

Vamos a considerar a ρ^* como se definió en (3.31). Entonces para i un entero impar y de la definición de punto extremo,

$$W^*[x_i, f(x_i) - P(A^*, x_i)] = -\rho^* \text{ o } P(A^*, x_i) = u(x_i),$$

para i un entero par,

$$W^*[x_i, f(x_i) - P(A^*, x_i)] = \rho^* \text{ o } P(A^*, x_i) = l(x_i).$$

Al ser W^* una función monótona se tiene que

$$(-1)^{i+1}(P(A^*, x_i) - P(A, x_i)) \geq 0, \ \ i = 1, \ldots, n+1.$$

Por el teorema del valor intermedio $P(A^*, x) - P(A, x)$ cuenta con ceros $y_1, y_2, \ldots, y_n$, tales que

$$x_1 \leq y_1 \leq x_2 \leq y_2 \leq \cdots \leq y_n \leq x_{n+1}.$$

Contando cualquier cero doble como dos ceros, luego, si definimos a $Q_n(C, x) = P(A^*, x) - P(A, x)$, obtenemos que $Q_n(C, y_i) = 0$, para todo $i = 1, 2, \ldots, n$.

Entonces, tenemos que $Q_n(C, x)$ tiene más de $n-1$ ceros distintos. Lo cual es una contradicción, puesto que $Q_n(C, x) \in \mathbf{C_n}$. De donde $Q_n(C, x) \equiv \mathbf{0}$, con lo que

$$P(A^*, x) \equiv P(A, x)$$

Así, $P(A^*, x)$ es la mejor aproximación generalizada para f con respecto a W^* y $\mathbf{C_n}^*$.

$\square$

3.4. Unicidad del mejor aproximante

Una vez que hemos caracterizado al mejor aproximante, lo que resta es ver bajo que condiciones ésta mejor aproximación es única, lo cual se muestra en el siguiente teorema.

Teorema 3.4.1. Sea f una función continua sobre X, W^* una función de peso generalizada, u, l dos funciones continuas, tales que $u(x) > l(x)$ para toda $x \in X$ y $\mathbf{C_n^*}$ la clase descrita en (3.7). Entonces existe una y sólo una mejor aproximación generalizada para f con respecto a W^* *y* $\mathbf{C_n^*}$.

Demostración. La existencia de la mejor aproximación generalizada se mostró anteriormente, por lo que únicamente se probará la unicidad.

Para probar la unicidad del mejor aproximante, supongamos que la función f tiene dos mejores aproximantes, digamos, $P(A_1, x),\ P(A_2, x) \in \mathbf{C_n^*}$. Como $P(A_1, x)$ es un mejor aproximante para f, entonces, por el teorema 3.3.2, la curva de error

$$W^*[x, f(x) - P(A_1, x)] \tag{3.34}$$

alterna al menos n veces sobre X, así, existen al menos $n+1$ puntos en X, digamos, $x_1, x_2, \dots, x_n, x_{n+1}$, tales que:

$$x_1 < x_2 < \cdots < x_n < x_{n+1},$$

los cuales son puntos extremos de la función (3.34).

Definamos a $I_{n+1} = \{1, 2, \dots, n, n+1\}$ y consideremos a ρ^* como se definió en (3.31).

Sin pérdida de generalidad, podemos suponer que

$$sgn^*[f(x_i) - P(A_1, x_i)] = (-1)^i, \quad i \in I_{n+1}.$$

Luego, para $i \in I_{n+1}$ impar,

$$W^*[x_i, f(x_i) - P(A_1, x_i)] = -\rho^* \ \text{ o } \ P(A_1, x_i) = u(x_i), \tag{3.35}$$

para $i \in I_{n+1}$ par,

$$W^*[x_i, f(x_i) - P(A_1, x_i)] = \rho^* \ \text{ o } \ P(A_1, x_i) = l(x_i). \tag{3.36}$$

Además, como $P(A_1, x)$ es una mejor aproximación y $P(A_2, x) \in \mathbf{C_n^*}$, entonces

$$\max_{x \in X} |W^*[x, f(x) - P(A_1, x)]| \leq \max_{x \in X} |W^*[x, f(x) - P(A_2, x)]| . \tag{3.37}$$

Así, de (3.35), (3.36) y (3.37), se sigue que

$$(-1)^{i+1}(P(A_1, x_i) - P(A_2, x_i)) \geq 0, \quad i \in I_{n+1}.$$

Luego, por el teorema del valor intermedio, se sigue que

$$Q_n(C, x) = P(A_1, x) - P(A_2, x)$$

cuenta con al menos n ceros, digamos $\{y_1, y_2, \ldots, y_n\}$, tales que

$$x_1 < y_1 < x_2 < y_2 < \cdots < x_n < y_n < x_{n+1}.$$

Luego, se sigue que $Q_n(x) \equiv \mathbf{0}$. Así, $P(A_1, x) \equiv P(A_2, x)$.
Por lo tanto, el mejor aproximante es único. $\square$

Conclusiones

El objetivo principal de esta tesis fue estudiar la aproximación con restricciones, mediante sistemas de Chebyshev, analizando la problemática referente a la mejor aproximación, considerando:

1. Existencia
2. Caracterización
3. Unicidad

Para poder desarrollar este trabajo, se necesitaron resultados que fueron tomados (en su mayoría) de la bibliografía utilizada, sin embargo, se detallaron un poco más las demostraciones, asimismo, se aportaron algunos nuevos resultados, como se especifica a continuación.

Para obtener un teorema que nos asegurara la existencia de la mejor aproximación, se tuvieron que analizar algunos resultados previos y aunque la mayoría fueron tomados de los textos (citados en la bibliografía), en esta parte se dió una adaptación de uno de ellos (teorema 3.2.7) para el caso en el que el conjunto X es compacto, la cual nos ayudaría posteriormente para la demostración del teorema principal de la existencia.

Cabe mencionar que para poder caracterizar a la mejor aproximación, se requería de un lema previo; el cual únicamente se encontraba en los textos para el caso de aproximación racional [12], así, debido a la necesidad de contar con un resultado análogo se realizó una adaptación de él, para sistemas de Chebyshev, obteniendo de esta manera el resultado fundamental para la demostración del teorema de caracterización.

Por último, una vez que se tenían los resultados referente a la existencia y caracterización de la mejor aproximación, con ayuda de éstos, se demostró la unicidad, concluyendo satisfactoriamente el objetivo principal de esta tesis.

Cabe mencionar también que al igual que se han generalizado los resultados acerca de existencia, caracterización y unicidad del mejor aproximante, un trabajo futuro podría ser la obtención de un algoritmo para encontrar el mejor aproximante, así como el estudio de algunas de sus aplicaciones.

Referencias

[1] B.L. Chalmers y G.D. Taylor, *Uniform approximation with constraints*, Jber. d. Dt. Math.- Verein, P. 49-86.

[2] E.W.Cheney, *Introduction to Approximation Theory*, Mc Graw Hill, 1966.

[3] R. A. DeVore y G G. Lorentz, *Cosntructive Approximation*, Springer-Verlag, erlin Heidelberg, 1993.

[4] J.W. Kammerer, *Optimal approximations of functions: one sided approximations and extrema preserving approximations*, Tesis doctoral, Universidad de Wisconsin, Madison, 1959.

[5] S. Karlin, W. J. Studden, *Tchebycheff systems: with applications to analysis and statistics*, Interscience, New York, 1996.

[6] N.P. Korneichuk, A. A. Ligun y V.G. Doronin, *Approximation with constraints*, Naukova Dumka, Kiev, 1982.

[7] E. Kreyszig, *Introductory Functional Anlaysis with Applicatiosn*, Wiley ans Sons, Inc., 1978.

[8] G.G Lorentz, *Approximation of Functions*, New York Holt, Renehart and Winston, 1996.

[9] D.G Moursund, *Chebyshev approximation using a generalized weight function*, SIAM J.Numer. Anal., 3, 1966, pp 435-450.

[10] D.G Moursund, *Computational aspects of Chebyshev approximation using a generalized weight function*, SIAM J.Numer. Anal. 5, 1968, pp. 126-137.

[11] D.G Moursundand y G.D.Taylor, *Optimal starting values for the Newton-Raphson calculation of inverses of certain functions*, SIAM J.Numer. Anal. 5, 1968, pp. 138-150.

[12] J. R. Rice, *The Approximation of Functions I, II*, Addison- Wesley, Reading,MA. Vol. I (1964), vol. II (1969).

[13] Raúl Nivón Santiago, *Métodos de la aproximación lineal ordinaria y semi-infinita para la aproximación uniforme de funciones y una aplicación a la óptica*, Tesis de licenciatura, UTM, 2003.

[14] G. D. Taylor, *On approximation by polynomials having restricted ranges I*, SIAM J. Numer. Analysis 5, 1968, pp. 258-268.

[15] A. F. Timan, *Theory of Approximation of Functions of Real Variable*, Pergamon Press, 1963.